JOURNAL OF BACTERIOLOGY, Mar. 2007, p. 2086–2100 Vol. 189, No. 5
0021-9193/07/$08.00+0 doi:10.1128/JB.01637-06

First Complete Genome Sequence of Two *Staphylococcus epidermidis* Bacteriophages▽†

Anu Daniel,[1]* Penelope E. Bonnen,[2] and Vincent A. Fischetti[1]

Laboratory of Bacterial Pathogenesis and Immunology[1] and Laboratory of Molecular Genetics,[2] The Rockefeller University, New York, New York 10021

Received 20 October 2006/Accepted 4 December 2006

***Staphylococcus epidermidis* is an important opportunistic pathogen causing nosocomial infections and is often associated with infections in patients with implanted prosthetic devices. A number of virulence determinants have been identified in *S. epidermidis*, which are typically acquired through horizontal gene transfer. Due to the high recombination potential, bacteriophages play an important role in these transfer events. Knowledge of phage genome sequences provides insights into phage-host biology and evolution. We present the complete genome sequence and a molecular characterization of two *S. epidermidis* phages, ϕPH15 (PH15) and ϕCNPH82 (CNPH82). Both phages belonged to the *Siphoviridae* family and produced stable lysogens. The PH15 and CNPH82 genomes displayed high sequence homology; however, our analyses also revealed important functional differences. The PH15 genome contained two introns, and in vivo splicing of phage mRNAs was demonstrated for both introns. Secondary structures for both introns were also predicted and showed high similarity to those of *Streptococcus thermophilus* phage 2972 introns. An additional finding was differential superinfection inhibition between the two phages that corresponded with differences in nucleotide sequence and overall gene content within the lysogeny module. We conducted phylogenetic analyses on all known *Siphoviridae*, which showed PH15 and CNPH82 clustering with *Staphylococcus aureus*, creating a novel clade within the *S. aureus* group and providing a higher overall resolution of the siphophage branch of the phage proteomic tree than previous studies. Until now, no *S. epidermidis* phage genome sequences have been reported in the literature, and thus this study represents the first complete genomic and molecular description of two *S. epidermidis* phages.**

Staphylococcus epidermidis, a member of the novobiocin-susceptible coagulase-negative staphylococci, is an important opportunistic pathogen and a major cause of nosocomial infections (42, 55). *S. epidermidis* predominantly colonizes the mucous membranes, groin, and axillar areas, as well as the cutaneous system of human body, with bacterial counts of up to 10 to 10^3 CFU/cm^2 (26). *S. epidermidis* is usually considered a harmless commensal microorganism; however, infections can occur in immunocompromised individuals and in patients with indwelling or implanted medical devices such as prosthetic heart valves and joint prostheses, where the staphylococci penetrate cutaneous and mucosal barriers (79). With the increasing use of such devices in medical practice, several million people are affected by complications arising from *S. epidermidis* infections (49). The widespread use of various antimicrobial agents, including penicillins, macrolides, aminoglycosides, and semisynthetic penicillins such as methicillin, has now led to the emergence of multiple-drug-resistant *S. epidermidis* strains (21, 72).

Two *S. epidermidis* genomes have been sequenced: those of the non-biofilm-forming, non-infection-associated strain ATCC 12228 and the infectious biofilm-forming strain RP62A (ATCC 35984) (21, 80). The difference between commensalism and pathogenicity was attributed to various factors: (i) single-nucleotide polymorphisms in pathogenicity-related genes (75), (ii) differential expression of genes contributing to *S. epidermidis* virulence (78), and (iii) acquisition of novel virulence factors and a high recombination and lateral gene transfer potential. These factors contributed toward the evolution of *S. epidermidis* from a commensal pathogen to a more aggressive opportunistic pathogen (21). Comparative genomic studies between staphylococcal species indicated that the majority of novel and unique genes could be accounted for by the presence or absence of prophages and genomic islands (21).

Bacteriophages are among the most abundant inhabitants of the biosphere, considering that an environmental sample contains nearly 10-fold more phage particles than prokaryotes (9). The bacteriophage population is estimated to be on the order of $\geq 10^{31}$, comprising approximately 10^8 distinct species (65, 77). The phage population is very dynamic, with rapid population turnover occurring within a relatively short period of time (7). They represent a rich and unique source of genetic and protein diversity, since less than 0.0002% of the global phage metagenome has been sampled (65) and a majority of genes have no assigned functions or matches in GenBank (6, 62). Bacteriophages also confer novel biological and physiological traits allowing host strains to adapt to new environments or obtain virulence determinants, thereby driving bacterial speciation and adaptation. The role of bacteriophages in bacterial biology, evolution, and diversity is now being truly appreciated due to concerted efforts toward characterizing bacteriophage genomes and analyzing their overall impact on global genetic and proteomic diversity, ecology, and recycling of organic matter (76).

Complete genomic sequences and molecular characteriza-

* Corresponding author. Mailing address: Laboratory of Bacterial Pathogenesis and Immunology, The Rockefeller University, New York, NY 10021. Phone: (212) 327-8170. Fax: (212) 327-7584. E-mail: adaniel@rockefeller.edu.

† Supplemental material for this article may be found at http://jb.asm.org/.

▽ Published ahead of print on 15 December 2006.

tion of several bacteriophages infecting important pathogenic bacteria, such as *Bacillus anthracis* (69), *Pseudomonas aeruginosa* (34), *Mycobacterium tuberculosis* (62), and *Staphylococcus aureus* (35), have been reported in recent years. This wealth of information has provided researchers a better understanding of the dynamics of phage-host interactions, mechanisms of virus evolution, acquisition of bacterial virulence determinants, and ecological and evolutionary changes of their hosts. Bacteriophages infecting *Staphylococcus epidermidis* have also been isolated (63) and are typically used for typing *S. epidermidis* strains (24, 73). However, to our knowledge, no *S. epidermidis* phage sequences have been reported in the literature.

In this study, we determined the ultrastructures and the complete genome sequences of two *S. epidermidis* phages, ϕPH15 (PH15) and ϕCNPH82 (CNPH82). The two genomes are highly similar and display typical modular organization seen in phages. Differential superinfection inhibition corresponded with nucleotide sequence differences and overall gene content within the lysogeny module. Detailed sequence analysis suggested the presence of two introns in PH15 which are spliced in vivo. We also conducted comparative phylogenetic analyses of the two phages with other sequenced phages of the family *Siphoviridae* reported in GenBank.

MATERIALS AND METHODS

Bacterial strains and phages. *Staphylococcus epidermidis* 414 (Felix d'Herelle Reference Center for Bacterial Viruses), referred as HER 1292 in this study, as well as phages PH15 and CNPH82 were donated by the Health Protection Agency (United Kingdom) and obtained from the Centers for Disease Control and Prevention, Atlanta, GA. HER 1292 was grown in Trypticase soy broth (TSB) (Difco) at 37°C in a shaking incubator at 250 rpm. Phages were propagated on strain HER 1292 in TSB supplemented with 4 mM $CaCl_2$ and 3 mM $MgCl_2$.

Preparation of phage lysates. PH15 and CNPH82 phage lysates were prepared by incubating phage stock with early-log-phase HER 1292 cells until complete lysis occurred (approximately 50 min). The lysates were cleared by centrifugation at 10,000 × *g* for 10 min, sterile filtered, and stored at 4°C. Phage titers were determined as described previously, using soft-agar overlay assays (68). The typical phage titers obtained by this method were in the range of 10^8 to 10^9 PFU/ml.

Electron microscopy of phages. Phages were pelleted by centrifugation at 100,000 × *g* for 1 h, washed twice with 100 mM ammonium acetate, and resuspended in 4 mM $CaCl_2$. The sample was adsorbed onto a carbon grid, which was floated on a small drop of 2% Na-phosphotungstate, pH 7.0, for 1 min and then blotted off. The sample was air dried and examined in a JEOL 100CX2 transmission electron microscope.

Production of HER 1292/PH15 and HER 1292/CNPH82 lysogens. Serial dilutions of PH15 and CNPH82 were spotted on HER 1292 lawns and incubated overnight at 37°C. Bacterial colonies growing within the plaques were harvested and passaged for several generations on TSB agar plates. After each passage, 25 PH15 lysogens and 20 CNPH82 lysogens were tested by colony PCR using the PH15-specific helicase primer set and the CNPH82-specific portal protein primer set for the presence of phage (primer sequences are available on request). Additionally, induction of PH15 and CNPH82 was determined by mitomycin C (5 μg/ml) treatment of the lysogenized HER 1292 cells followed by soft-agar overlays.

Preparation of phage DNA. Phage DNA was prepared from the lysates after treatment with RNase and DNase (10 μg each) at 37°C for 30 min. Phage were precipitated by adding NaCl (0.5 M final concentration) and polyethylene glycol 8000 (10%, wt/vol) and incubating on ice for 30 min. Phage were pelleted by ultracentrifugation at 100,000 × *g* for 1 h, resuspended in 10 mM Tris-Cl (pH 7.5) plus 50 mM $MgCl_2$, and lysed by adding sodium dodecyl sulfate (0.5% final concentration) and 20 mM EDTA. DNA was purified by phenol-chloroform extraction as previously described (68).

Sequencing of phage DNA. Complete genome sequencing of both phages was done by Macrogen Inc., Seoul, Korea. Briefly, phage genomic DNA was sheared using a nebulizer (Invitrogen, CA), and blunt-end repaired and dephosphorylated. DNA fragments of desired size (1 to 6 kb) were blunt-end ligated into pCR4Blunt-TOPO vector (Invitrogen, CA) and electroporated into *Escherichia coli* DH10B cells according to the manufacturer's instructions. Clones were sequenced until >10-fold redundancy was obtained. The complete genomes were assembled using the SeqMan II (DNASTAR, Inc.) sequence analysis software.

DNA sequence and bioinformatics analyses. Open reading frames (ORFs) in the final genome sequences were predicted using GeneMark.hmm for prokaryotes (http://opal.biology.gatech.edu/GeneMark) (47) and ORF Finder (http://ncbi.nlm.nih.gov/gorf/gorf.html) software. In our analyses, putative ORFs contained either AUG (methionine), UUG (leucine), or GUG (valine) as the starting codon, preceded by a Shine-Dalgarno sequence optimally placed (3 to 12 bp) upstream of the start codon and containing sequence encoding at least 50 amino acids before the termination codon. ORF numbers are preceded by either *ph* or *cn* designation for phage PH15 or CNPH82, respectively (Tables 1 and 2). The predicted proteins were searched against the protein database by using BLAST (1) and PSI-BLAST (2) algorithms. Structural predictions and motif searches of proteins were performed with the proteomic tools at ExPASy (http://us.expasy.org) and Pfam (http://www.sanger.ac.uk/Software/Pfam) as well as the software COILS (48), and PSORT-B (48). Transmembrane domains were predicted with the TMHMM version 2.0 program (http://www.cbs.dtu.dk/services/TMHMM-2.0) (71).

Reverse transcription-PCR (RT-PCR). HER 1292 cells were infected at a multiplicity of infection of 5, and the corresponding phage and RNA were isolated at 15 and 25 min postinfection (p.i.). RNA was isolated using the RNeasy kit (QIAGEN Sciences, MD) according to the manufacturer's instructions. Contaminating DNA was removed with the TURBO DNA-free kit (Ambion Inc., TX). First-strand cDNA synthesis was performed using the Superscript III first-strand Synthesis kit (Invitrogen, CA) with random hexamers according to the manufacturer's instructions. PCR was conducted with 2 μl each of cDNA and PH15 and CNPH82 large terminase subunit- and lysin-specific primers (sequences are available on request). The PCR conditions were as follows: 94°C for 3 min, 94°C for 30 s, 55 to 57°C for 30 s, and 72°C for 1 min 30 s, for 35 cycles. PCR products were separated on 1% agarose gels.

Phylogenetic analysis. The DOTTER program (70) was used to compare the nucleotide sequences of PH15 and CNPH82 with 35 *S. aureus* phage sequences present in the GenBank database, with a sliding window of 25 bp.

Proteomic distances were calculated using an approach similar to that described by Rohwer and Edwards (66). Predicted protein sequences for PH15 and CNPH82 were compared to all known siphophage protein sequences in GenBank (see Table S1A in the supplemental material for accession numbers). Protein distance scores were estimated using the same methods and parameters as described by Rohwer and Edwards (66). Like in that study, we considered protein distance scores (S_{PD}) significant if $-1 < S_{PD} < 5$, and we applied a penalty of 5 for pairwise protein comparisons that did not have a significant score. Proteomic distance was calculated using the formula {[Σ(significant S_{PD}) × 100] + (penalty × number of nonhits)}/[(number of hits × 100) + number of nonhits], where a "hit" is a pairwise comparison that had a significant S_{PD} and a "nonhit" is a pairwise comparison that did not have a significant S_{PD}. We did not use a length correction, as this caused a small number of phages to cluster in inappropriate places. For example, when using length, two members of *Mycoplasma* were drawn into the outer edge of the *Staphylococcus* branch. These phages clustered completely separately from *Staphylococcus* when length was not included in the proteomic distance formula. Because most S_{PD} ranged from 0.1 to 5, we applied a 100× amplification in the calculation of the proteomic distance score. This amplification of scores facilitated tree generation and resulted in a tree that was less dense. It did not affect relationships between genomes.

An unrooted phylogenetic tree was generated using FITCH (22) with the default parameters plus the additional options of randomizing the input order of sequences and global rearrangement optimization. The representative tree was produced using the MEGA 3.1 software package (33).

Nucleotide sequence accession numbers. The genome sequences of PH15 and CNPH82 were submitted to GenBank and were assigned accession numbers DQ834250 and DQ831957, respectively.

RESULTS AND DISCUSSION

We report the complete annotated genome sequence of two temperate *S. epidermidis* phages, PH15 and CNPH82. These phages, along with several other phages such as ϕPH48, ϕPH456, and ϕPH459, were obtained as a gift from the CDC (17) and have been previously used for *S. epidermidis* phage

TABLE 1. Genome organization of PH15

ph[a]	bp		Size (kDa) (pI)[b]	Predicted function	Most significant database match (E value)[c]	Structural features[d]	% Amino acid identity (% amino acid similarity)	Accession no.
	Start	End						
1	299	733	16.5 (5.4)	Terminase, small subunit	Orf35, phage 37 (9e−48); phage terminase small subunit (4e−26)	c-c domain	66 (80)	AAX91296.1
2a	720	1601	34 (9.2)	Terminase, large subunit	Orf9, phage EW (5e−132); phage terminase large subunit (3e−40)	c-c domain, P loop	90 (95)	AAX91349.1
3	1671	2300	24.5 (9.8)	Homing endonuclease	Enterobacteria, phage T4 (9e−08); GIYX(10-11YIG family of class I homing endonucleases (4e−04)		28 (46)	AAD42655.2
2b	2491	2952	17.9 (5.9)	Terminase, large subunit	Orf9, phage 37 (2e−82); phage terminase large subunit (2e−17)		94 (98)	AAX91349.1
*4**	2958	4394	56 (4.5)	Portal	Orf7, phage 37 (0.0); phage portal protein (5e−78)	c-c domains	77 (89)	AAX91270.1
5	4351	5304	36.6 (8.8)	Head morphogenesis	Orf12, phage 71 (1e−131); phage Mu protein F-like protein (1e−14)	c-c domains	74 (87)	AAX91582.1
6	5408	6004	22.9 (4.6)	Minor capsid protein	Orf20, phage 71 (5e−49); minor capsid protein (3e−21)	c-c domain	52 (74)	AAX91590.1
7	6022	6852	29.6 (4.8)	Major capsid protein	Orf13, phage 37 (6e−124); major capsid protein (9e−54)		82 (90)	AAX91276.1
8	6869	7159	11.2 (4.8)	Transcription termination?	Orf50, phage 71 (7e−25); Rho termination factor (N terminal) (1e−05)		55 (67)	AAX91616.1
9	7159	7473	12 (7.0)	Head morphogenesis	Orf50, phage X2 (7e−28)		65 (82)	AAX92052.1
10	7466	7795	12.7 (5.0)	Head-tail joining	Orf43, phage 52A (3e−46); gp16 of phage SPP1 (3e−23)		77 (85)	AAX91833.1
*11**	7788	8201	15.5 (8.9)	Tail component?	Orf46, phage 71 (1e−65)		81 (89)	AAX91613.1
12	8214	8651	16.8 (8.3)		Orf28, phage 29 (8e−74)		92 (97)	AAX91747.1
13	8638	9177	19.9 (4.5)		Orf26, phage 92 (2e−80)		80 (87)	AAX91957.1
14	9239	9733	18.8 (4.7)	Tail component	Orf99, phage EW (2e−63)	c-c domains	70 (82)	AAX91402.1
15	9796	10098	11.7 (10.8)		Orf48, phage 71 (9e−31)	P loop	62 (78)	AAX91614.1
16	10101	13205	112.4 (10.3)	Tape measure protein	Orf1, phage EW (0.0); TMP repeat (1e−03)	c-c domain (N terminal), 10 TM domains	62 (76)	AAX91341.1
*17**	13221	14159	36.4 (6.2)		Orf38, phage EW (4e−111); glycosyl hydrolase family 5 signature (7e−07)		62 (79)	AAX91376.1
*18**	14173	16029	69.5 (8.8)	Endopeptidase	Orf6, phage EW (0.0); SGNH_hydrolase subfamily (3e−39)	c-c domain	68 (84)	AAX91346.1
*19**	16044	18710	96.7 (6.6)	Preneck appendage protein	Orf3, phage 37 (0.0)	c-c domain (N terminal), beta-helix repeat	53 (71)	AAX91266.1
20	18710	20242	58 (4.7)		Orf5, phage 37 (3e−82)	c-c domain	35 (52)	AAX91268.1
21	20226	20585	13.7 (4.2)		Orf29, phage 37 (6e−14)	c-c domain	39 (57)	AAX91290.1
22	20587	20727	5.6 (8.9)		Orf143, phage 37 (9e−07)		62 (77)	AAX91338.1
23	20764	21063	11.6 (8.0)		Orf51, phage EW (4e−28)	c-c domains, 1 TM domain	58 (80)	AAX91384.1
24a	21202	21318	4.1 (9.9)	Endolysin	Orf3, phage EW (6e−13); CHAP domain (2e−15)		84 (92)	AAX91343.1
24b	21562	23349	65.7 (9.8)	Endolysin	Orf3, phage EW (0.0); amidase_4 (8e−17); CHAP domain (2e−15)	c-c domain	72 (82)	AAX91343.2
25	23405	23914	19.2 (4.5)		Phage 10750.2 (2e−07)		30 (43)	NP_795654.1
26	23914	24444	20.5 (4.9)		Orf70, phage G1 (1e−23)		37 (55)	AAX92151.1
27	24493	24765	9.8 (9.2)	Holin	Orf56, phage 71 (4e−32); phage holin (4e−09)	2 TM domains	67 (88)	AAX91620.1
28	24765	26147	52.6 (9.5)	Endolysin	Orf7, phage 71 (6e−114); amidase_3 (3e−08); CHAP domain (2e−07)		50 (64)	NP_510959.1
29	26247	27152	35.6 (9.5)		Orf114, phage LP65 (1e−14)		34 (56)	YP_164749.1
30	27582	27797	8.4 (9.3)					
31	29856	28480	54.2 (8.9)	Integrase	Orf7, phage 85 (7e−134); resolvase (N terminal) (7e−29)	c-c domains	52 (70)	AAX90917.1
32	30392	29910	18.6 (4.7)			c-c domains		
33	30870	30394	18.4 (9.0)		Orf25, phage ROSA (5e−22)	c-c domain (C-term)	43 (66)	AAX91520.1
34	31351	30890	18.3 (5.6)	Unknown (DUF955)	Orf26, phage ROSA (5e−63); DUF955 (3e−10)		73 (88)	AAX91521.1
35	31686	31363	12.3 (8.3)	cI-like repressor	Orf46, phage ROSA (2e−50); HTH-XRE type 3 family (4e−10)	HTH motif	91 (94)	AAX91538.1
36	31850	32098	9.4 (8.6)	Cro-like repressor	Orf63, phage ROSA (8e−35); HTH-XRE type 3 family (5e−07)	HTH motif	85 (96)	AAX91550.1
37	32272	32111	6.4 (5.6)		Orf114, phage ROSA (9e−15)		88 (97)	AAX91561.1
38	32403	32233	6.5 (7.2)					
39	32452	33219	29.7 (8.4)	Antirepressor	Orf17, phage 53 (2e−24); COG3645 (2e−14)	c-c domains	31 (47)	AAX90848.1
40	33222	33386	6.2 (9.9)					
41	33372	33578	8 (4.5)			c-c domain		
42	33616	33738	5 (4.4)					
43	33735	33911	6.7 (9.1)			1 TM domain		
44	33971	34246	11 (4.7)					
*45**	34218	34442	8.7 (5.0)		Orf72, phage ROSA (4e−11)		55 (76)	AAX91552.1
46	34435	35061	23.7 (5.6)	Topoisomerase	Orf23, phage 85 (6e−100); DUF1071 (8e−14)	c-c domain	84 (91)	AAX90932.1
47	35054	35470	15.5 (7.0)	Single-stranded-DNA-binding protein	Orf34, phage EW (2e−51); Single-stranded-DNA-binding protein (1e−20)		75 (84)	AAX91372.1

48	35484	36158	26.3 (7.4)	Unknown (DUF968)	phiSLTp19, phage phiSLT (4e−98)		75 (85)	NP_075482.1
49	36151	36945	30.7 (5.3)	DNA replication	Orf15, phage 71 (1e−74); phage replisome organizer (N terminal) (3e−45)	c-c domains	61 (76)	AAX91585.1
50	36945	37304	14.2 (8.7)		Orf42, phage 96 (3e−36)		75 (114)	AAX91459.1
51	37297	38532	47.2 (5.1)	Helicase	*S. aureus* phage phi 11 (1e−161); DnaB helicase C domain (C terminal) (5e−65)	c-c domain, P-loop	70 (84)	AAL82244.1
52	38529	38750	8.8 (6.6)		Orf78, phage 71 (5e−16)		50 (75)	AAX91630.1
53	38728	38973	9.6 (8.3)		Orf69, phage 187 (2e−02)	c-c domain (C-term)	30 (54)	AAX90732.1
54	38982	39389	16.0 (9.2)	Replication, recombination, and repair	Orf37, phage 85 (4e−49); Rus, Holliday junction resolvase (9e−23)		67 (78)	AAX91525.1
55	39390	39587	7.7 (4.3)	Transcriptional regulator	*L. innocua* CLIP 11262 (1e−03)		34 (55)	CAC97067.1
56	39588	39947	14.2 (9.9)		Orf50, phage phi PVL (1e−17)	c-c domain	40 (65)	NP_075488.1
57	39944	40378	16.8 (5.6)		Orf60, phage 96 (5e−24)		67 (79)	AAX91470.1
58	40381	41061	26.7 (9.5)	HNH homing endonuclease	Orf85, phage G1 (3e−06); intron-encoded nuclease repeat motif (C terminal) (9e−03)		37 (46)	AAX92166.1
59	41242	41568	13.1 (4.5)		Orf30, phage 96 (2e−03)	c-c domain	29 (56)	AAX91447.1
60	41558	41806	9.9 (5.5)					
61	41796	41975	6.7 (4.7)		Orf57, phage 55 (3e−06)		45 (81)	AAX91694.1
62	41968	42393	15.4 (5.3)	dUTPase	Orf36, phage 37 (3e−36); dUTPase (7e−22)		81 (90)	AAX91297.1
63	42472	42783	12.2 (8.7)					
*64**	42780	42932	5.7 (5.5)	Unknown (DUF1381)	Orf118, phage 71 (1e−09); DUF1381 (3e−06)	c-c domain	55 (71)	AAX91638.1
65	42925	43095	6.7 (5.4)	Transcriptional regulator	Orf144, phage 71 (6e−11); transcriptional activator RinB (3e−11)		63 (85)	AAX91640.1
66	43098	43238	5.3 (4.9)			1 TM domain		
*67**	43239	43460	8.6 (9.2)					
68	43478	43894	16.5 (5.4)	Transcriptional regulator	Orf26, phage 77 (2e−60); transcriptional activator RinA (1e−18)	c-c domain (N terminal)	76 (91)	NP_958669.1

[a] An asterisk indicates an ORF with an unusual start codon (TTG or GTG).
[b] Based on computer prediction.
[c] Only the most significant homology is listed.
[d] Based on computational predictions. c-c, coiled coil; TM, transmembrane; HTH, helix-turn-helix.

typing. To our knowledge, our study is the first report of genomic sequences of *S. epidermidis* phages with characterization at the molecular level.

Phage ultrastructure and host range studies. PH15 and CNPH82 were selected for further analyses because EcoRI and BamHI restriction profiles showed substantive sequence variation between these two phages (data not shown). Purified PH15 and CNPH82 particles observed by electron microscopy revealed that the two phages had similarly sized small icosahedral heads (diameter of approximately 55 nm) and long flexible noncontractile tails (length of approximately 160 nm) with dual disc baseplates (Fig. 1B and E), suggesting that they both belong in the *Siphoviridae* family within the order *Caudovirales*. A small collar-like structure between the head and tail was visible in phage particles, particularly where the tail had separated from the head (Fig. 1B and E). Similar collar-like structures have been observed in *Lactococcus lactis* phage TP901-1 (27). The phage particles appeared to aggregate and form a "bouquet-like" arrangement, possibly due to tail fiber adherence to wall fragments (Fig. 1A and D), similar to those observed in *Bacillus anthracis* γ phage (69).

Host range analysis of PH15 and CNPH82 revealed that both phages formed plaques within infected bacterial lawns of HER 1292, while no plaques were observed on lawns of *S. epidermidis* strains ATCC 12228 or RP62A. Additionally, no plaques were observed on 10 *S. aureus* strains tested, including NCTC 8325 (58) and RN4220 (31).

Both PH15 and CNPH82 were able to form stable lysogens in HER 1292, suggesting the presence of a functional lysogeny module in both phages. Phages could be induced from the respective lysogens by mitomycin C induction, and the induced phages formed plaques on HER 1292 lawns. The lysogens did not show any differences in colony morphology or growth curve characteristics in liquid media compared to the host (data not shown). The lysogens were tested for superinfection by infecting PH15-lysogenized HER 1292 and CNPH82-lysogenized HER 1292 with PH15 and CNPH82. PH15 formed plaques on CNPH82-lysogenized HER 1292 lawns. The efficiency was similar to that on HER 1292 lawns. CNPH82 was unable to form plaques on CNPH82-lysogenized HER 1292 (data not shown). Neither PH15 nor CNPH82 was able to form plaques on a PH15-lysogenized HER 1292 lawn.

The results of superinfection studies suggest that the PH15-lysogenized host exhibits superinfection immunity whereby the integrated PH15 prophage prevents infection by a homoimmune phage (PH15 itself and CNPH82). However, the ability of PH15 to infect CNPH82-lysogenized strain could be attributed to differences in the genomic organization within the lysogeny control module (see below). Such differences have been shown to be important in *Streptococcus thermophilus* phage Sfi19 for eliminating Sfi21 prophage control of superinfection immunity (46). We conclude that the genomic differences observed in the lysogeny module of PH15 may eliminate superinfection immunity by CNPH82 prophage.

Genome features of PH15 and CNPH82. The complete genome sequences of PH15 and CNPH82 were determined. The PH15 genome comprised 44,047 bp containing 68 putative ORFs, while the CNPH82 genome comprised 43,420 bp containing 65 putative ORFs (Fig. 2). The G+C contents of PH15

TABLE 2. Genome organization of CNPH82

cn[a]	bp		Size (kDa) (pI)[b]	Predicted function	Most significant database match (E value)[c]	Structural features[d]	% Amino acid identity (% amino acid similarity)	Accession no.
	Start	End						
1	245	679	16.5 (5.4)	Terminase, small subunit	Orf35, phage 37 (2e−46); terminase small subunit (4e−25)	c-c domains	65 (81)	AAX91296.1
2	663	1928	49.2 (7.8)	Terminase, large subunit	Orf9, phage X2 (0.0); terminase large subunit (9e−63)	c-c domain, P loop	92 (96)	AAX92014.1
*3**	1934	3370	56.2 (5.7)	Portal	Orf7, phage 37 (0.0); phage portal protein, SPP1 Gp6-like (5e−78)	c-c domains	77 (88)	AAX91270.1
4	3327	4280	36.6 (9.0)	Head morphogenesis	Orf12, phage 71 (1e−139); phage Mu protein F-like protein (1e−15)	c-c domains	76 (87)	AAX91582.1
5	4283	4387	4.1 (5.6)		Orf147, phage EW (6e−07)		72 (90)	AAX91409.1
6	4388	4594	7.9 (5.3)		Orf66, phage 96 (2e−14)	c-c domain	54 (75)	AAX91475.1
7	4709	5233	20.0 (4.8)	Minor head protein	Orf20, phage 71 (5e−44); minor capsid protein (6e−10)	c-c domain	51 (74)	AAX91590.1
8	5215	6204	36.0 (5.5)		Orf11, phage 52A (1e−152)	c-c domain	85 (93)	AAX91805.1
9	6221	6511	11.2 (4.8)	Transcription termination?	Orf50, phage 71 (3e−24); Rho-termination factor (N terminal) (8e−05)		56 (67)	AAX91616.1
10	6511	6825	12.0 (6.6)	Head morphogenesis	Orf50, phage X2 (3e−33)		65 (82)	AAX92052.1
11	6818	7147	12.7 (5.7)	Head-tail joining	Orf43, phage 52A (2e−45); gp16 of phage SPP1 (1e−23)		77 (86)	AAX91833.1
*12**	7140	7553	15.4 (8.6)	Tail component?	Orf46, phage 71 (4e−64)		81 (89)	AAX91613.1
13	7566	8003	16.9 (8.6)		Orf28, phage 29 (4e−74)		93 (97)	AAX91747.1
14	7990	8529	20.0 (4.7)		Orf26, phage 92 (2e−79)	c-c domain	80 (87)	AAX91957.1
15	8592	9086	18.8 (4.7)	Tail component	Orf99, phage EW (1e−63)	c-c domain	72 (82)	AAX91402.1
16	9149	9451	11.8 (10.7)		Orf48, phage 71 (2e−29)	c-c domain, P loop	62 (78)	AAX91614.1
17	9454	12558	112.3 (10.4)	Tape measure protein	Orf1, phage EW (0.0); TMP repeat (5e−13)	c-c domain (N terminal), 12 TM domains	62 (76)	AAX91341.1
*18**	12574	13512	25.5 (5.7)		Orf38, phage EW (1e−73); glycosyl hydrolase family 5 signature (7e−07)		63 (77)	AAX91376.1
*19**	13526	15382	69.8 (8.9)	Endopeptidase	Orf6, phage EW (0.0); SGNH_hydrolase subfamily (1e−34)	c-c domain	69 (83)	AAX91346.1
*20**	15397	18063	97.0 (6.9)	Preneck appendage	Orf3, phage 37 (0.0)	c-c domain	53 (71)	AAX91266.1
21	18063	19595	58.0 (5.0)		Orf5, phage 37 (2e−84)	c-c domain	35 (52)	AAX91268.1
22	19579	19938	13.7 (4.3)		Orf29, phage 37 (7e−13)	c-c domain (C terminal)	39 (57)	AAX91290.1
23	19940	20080	5.6 (9.0)		Orf143, phage 37 (1e−05)		62 (77)	AAX91338.1
*24**	20084	20416	12.8 (6.8)		Orf51, phage EW (5e−27); cell attachment sequence (1e−04)	c-c domains, 1 TM domain	58 (80)	AAX91384.1
25	20555	22456	72.5 (9.8)	Endolysin	Orf3, phage EW (0.0); amidase_4 (8e−17); CHAP domain (2e−15)	c-c domain	71 (81)	AAX91343.1
*26**	22646	23494	31.4 (4.8)	Tail component	Orf10, phage 71 (7e−34)	c-c domain	35 (52)	AAX91580.1
27	23506	23859	13.5 (4.4)		Orf29, phage 37 (7e−13)	c-c domain (C terminal)	36 (59)	AAX91290.1
28	23852	23998	5.7 (4.8)		Orf143, phage 37 (3e−05)		57 (76)	AAX91338.1
29	24053	24325	9.9 (9.2)	Holin	Orf56, phage 71 (6e−31); phage holin (4e−09)	2 TM domains	65 (88)	AAX91620.1
30	24325	25707	52.5 (9.5)	Endolysin	Orf7, phage 71 (1e = 125); amidase_3 (3e−08); CHAP domain (2e−07)		50 (65)	AAX91577.1
31	25774	26934	46 (9.5)		Orf114, *L. plantarum* phage LP65	c-c domain (C terminal)	28 (43)	YP_164749.1
32	28631	27255	54.2 (8.8)	Integrase	Orf26, phage 52A (8e−114); resolvase (N-term) (4e−20); PinR recombinase (1e−24)	c-c domains	51 (70)	AAX90917.1
33	29252	28686	20.9 (9.1)		Orf7, phage X2 (2e−03); quinonprotein alcohol dehydrogenase-like (2e−13)	c-c domain		
*34**	29740	29270	18.5 (7.2)	Unknown function (DUF955)	Orf28, phage 55 (2e−61); DUF955 (3e−8)		72 (86)	AAX91670.1

35	30078	29746	2.7 (5.5)	cI-like repressor	Orf44, phage 55 (1e−40); HTH-XRE type 3 family (7e−10)	HTH motif	75 (89)	AAX91685.1
36	30267	30509	9.4 (8.8)	Cro-like repressor	Orf54, phage 55 (3e−32); HTH-XRE type 3 family (3e−12)	HTH motif	84 (91)	AAX91692.1
*37**	30928	30674	9.8 (0.8)		Orf74, phage 85 (3e−20)		67 (83)	AAX90968.1
38	30930	31703	29.8 (9.1)	Antirepressor	Orf16, phage EW (1e−98); COG3645 (9e−20)	c-c domains	68 (82)	AAX91356.1
39	31741	31902	6 (9.8)					
40	32275	32550	10.9 (5.2)			c-c domain		
*41**	32522	32746	8.7 (5.2)		Orf59, phage 37 (2e−18)		60 (75)	AAX91316.1
42	32739	33365	23.7 (5.7)	Topoisomerase	Orf20, phage X2 (1e−99); DUF1071 (9e−14)	c-c domain	85 (92)	AAX92025.1
43	33358	33768	15.0 (7.8)	DNA replication	Orf34, phage EW (7e−53); Single-stranded-DNA-binding protein (2e−20)		79 (86)	AAX91372.1
44	33782	34447	25.8 (6.4)	Unknown function (DUF968)	Orf17, phage 37 (1e−100); DUF968 (3e−79)		77 (88)	AAX91280.1
45	34447	35175	28.0 (8.9)	DNA replication	Orf22, phi ETA (1e−60); phage replisome organizer (6e−04)	c-c domain, P loop, winged helix	78 (87)	AAL82345.1
46	35181	35534	14.6 (8.7)		Orf41, phage EW (4e−17)		34 (67)	AAX91379.1
47	35521	36771	47.5 (4.9)	DNA replication	Orf10, phage EW (1e−129); DnaB helicase C domain (C terminal) (8e−65)	c-c domain, P loop	54 (75)	AAX91350.1
48	36768	36989	8.8 (5.1)		Orf78, phage X2 (3e−13)		47 (70)	AAX92064.1
49	36967	37212	9.6 (7.6)		Orf69, phage 187 (6e−02)	c-c domain	31 (55)	AAX90732.1
*50**	37222	37638	16.7 (9.5)	Unknown function (DUF1064)	Orf36, phage X2 (5e−54); DUF1064 (9e−29)	c-c domain	73 (87)	AAX92040.1
51	37625	38089	18.0 (5.0)		Orf32, phage 55 (3e−10)		30 (52)	AAX91674.1
52	38041	38232	7.5 (4.8)					
53	38233	38592	14.2 (9.9)		Orf44, phage 71 (7e−17)	c-c domain	40 (62)	AAX91612.1
54	38589	39023	16.8 (5.7)		Orf48, phage 47 (5e−23)		71 (82)	AAX91237.1
55	39026	39706	26.8 (9.5)		Orf85, phage G1 (4e−06); intron-associated endonuclease (3e−13)		38 (48)	AAX92166.1
56	39703	39888	7.1 (10.1)		Orf74, phage 2638A (5e−15)		60 (80)	AAX91033.1
*57**	39876	40232	13.7 (5.5)		Orf13, phage Twort (3e−27); SNase homolog (1e−6)	c-c domain	60 (73)	AAX92396.1
58	40235	40774	21.3 (4.7)		Orf30, phage 96 (3e−07)	c-c domains	27 (42)	AAX91447.1
59	40764	40943	6.7 (4.9)		Orf57, phage 55 (3e−06)	c-c domain	48 (81)	AAX91694.1
60	40936	41361	15.6 (5.3)	dUTPase	Orf36, phage 37 (8e−58); dUTPase (7e−22)		82 (88)	AAX91297.1
61	41641	41892	9.7 (5.3)		Orf56, phage 37 (1e−12)		46 (60)	AAX91314.1
62	41889	42059	6.5 (5.4)	Transcriptional regulator	Orf85, phage 29 (1e−08); transcriptional activator RinB (6e−12)		60 (80)	AAX91783.1
*63**	42265	42663	15.2 (8.3)		Orf44, phage 187 (2e−14)	c-c domain, 2 TM domains, leucine zipper	38 (60)	AAX90719.1
64	42651	42791	5.3 (4.6)		Orf118, phage 29 (7e−06)	c-c domain, 1 TM domain	54 (77)	AAX91790.1
65	42794	43210	16.5 (5.6)	Transcriptional regulator	Orf26, phage 77 (1e−60); transcriptional activator RinA (1e−57)	c-c domain (N terminal)	76 (91)	NP_958669.1

[a] An asterisk indicates an ORF with an unusual start codon (TTG or GTG).
[b] Based on computer prediction.
[c] Only the most significant homology is listed.
[d] Based on computational predictions. c-c, coiled coil; TM, transmembrane; HTH, helix-turn-helix.

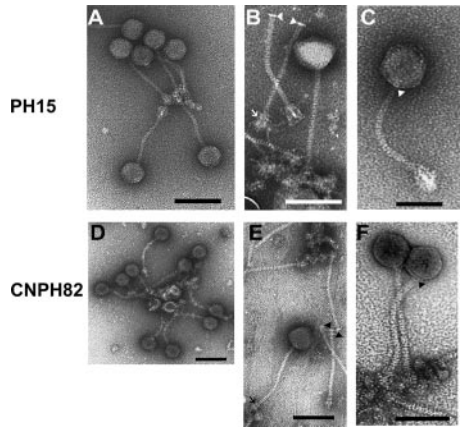

FIG. 1. Ultrastructures of phages PH15 (A to C) and CNPH82 (D to F). (A and D) Transmission electron micrographs of phages PH15 and CNPH82, showing their isometric heads and noncontractile tails. (B and E) Tails of PH15 and CNPH82 detached from the virion head, showing the attached collar (arrowheads) and dual disc plates (arrows). (C and F) Single intact phage virions of PH15 and CNPH82. The collar is marked by arrowheads. Bars, 100 nm.

and CNPH82 were 34.91% and 34.67%, respectively, which are comparable to that of the *S. epidermidis* genome (80).

Restriction analysis with endonucleases that have one cut site within the phage genomes (XhoI and SacI for PH15 and ApaLI and MscI for CNPH82) produced a single band on agarose gels, while double digestion produced two fragments (data not shown). These results suggested that both the PH15 and CNPH82 genomes were circularly permuted, since circularly permuted (and terminally redundant) linear phage chromosomes behave as do circular chromosomes during restriction analysis (5). We also observed submolar fragments of both PH15 and CNPH82 DNAs after heating the restriction endonuclease digestion mixture (data not shown), suggesting that both phages belonged to the *pac*-type group. In this group of phages, which typically have circularly permuted linear chromosomes, packaging of concatemeric phage DNA is initiated processively in a headful manner after the terminase cuts the DNA at a *pac* site (5).

Bioinformatic analyses revealed that the gene coding potentials for PH15 and CNPH82 were 91% and 94%, respectively, indicating tight packing and few intergenic regions, with approximately 1.5 genes/kbp of nucleotide sequence. The majority of ORFs initiated translation from an AUG start codon in both genomes, except for five ORFs in PH15 (*ph15*, *ph49*, *ph55*, *ph56*, and *ph57*) and eight ORFs in CNPH82 (*cn3*, *cn6*, *cn10*, *cn19*, *cn46*, *cn52*, *cn54*, and *cn55*) that initiated with UUG start codon and three ORFs (*ph34*, *ph37*, and *ph42*) in PH15 and five ORFs (*cn26*, *cn32*, *cn37*, *cn59*, and *cn61*) in CNPH82 that initiated with GUG start codon (Tables 1 and 2). Both genomes were annotated using comparisons to current databases. However, biological function could be assigned to only 43% of the PH15 proteome and only 41% of the CNPH82 proteome. Both genomes displayed organizations similar to those of the phages of the *Siphoviridae* family (8). The genomes were modularly organized and consisted of DNA packaging, head-and-tail morphogenesis, host cell lysis, lysogeny, and DNA replication and modification modules (Fig. 2).

PH15 and CNPH82 genome comparison. Database searches with the putative ORFs of PH15 and CNPH82 revealed that the gene orders were similar for the two genomes, and the proteins encoded by the two genomes showed one-to-one correspondence (Fig. 2). Pairwise nucleotide comparison revealed that the two genomes displayed similarity at the nucleotide level, with ~60% identity genome-wide and ~85% identity over the regions corresponding to the head-and-tail morphogenesis module and between the tail morphogenesis and lysis modules (Fig. 3). The interruption of the straight line seen within the head morphogenesis module corresponded to the swapping of ORF *ph7* (Table 2) with ORF *cn8*. A straight line was also observed in the DNA packaging region, except for a lateral shift marked by the presence of an intron (*terL-I*) (see below) within the putative large terminase subunit (Fig. 3). Another lateral shift of the straight line in the otherwise highly similar lysis module corresponded to the *lys-I* intron (see below). A gap corresponding to ORFs *ph25*/*26* and ORFs

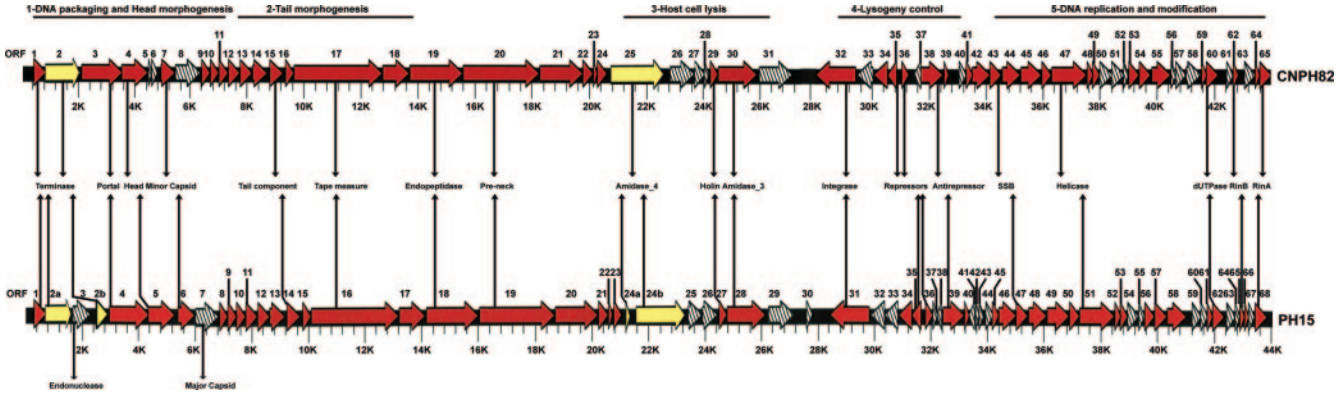

FIG. 2. Schematic representation of CNPH82 and PH15 genomes, showing the genome organization, predicted ORFs and some putative functions. The ORFs are depicted by arrows or arrowheads pointing in the direction of transcription and are numbered consecutively (see Tables 1 and 2, respectively). ORFs identical in both the PH15 and CNPH82 genomes are shown in red, while ORFs unique to either genome are shown as hatched arrows. The *terL* and *lys* ORFs of both phages are shown as yellow arrows. The phage modules determined by database matches and genome organization are labeled. The ruler marks the relative positions of the ORFs within the 44,047-bp PH15 genome and the 43,420-bp CNPH82 genome.

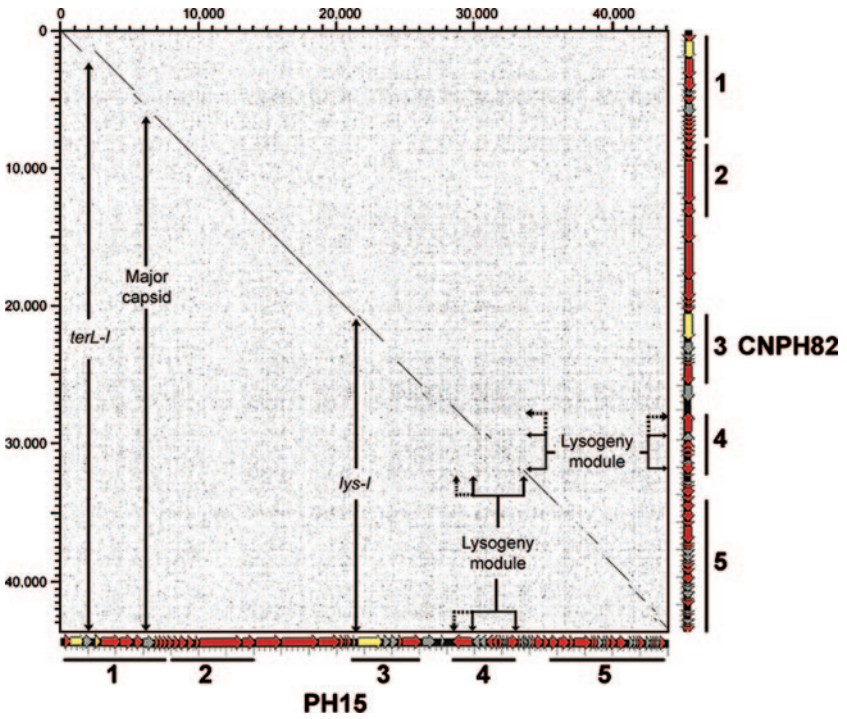

FIG. 3. Dot plot comparison of PH15 and CNPH82 genomes. Dot plot analysis was conducted for the genomic DNA sequences of PH15 (*x* axis) and CNPH82 (*y* axis) by using the DOTTER program (70) with a sliding window of 25 bp. The color-coded schematic genome maps of PH15 and CNPH82 along with the numbers of the corresponding modules (see Fig. 2) are presented at the respective axes for easy orientation. Specific regions of differences between the two genomes are marked and annotated. The *int* gene within the lysogeny module is identical in both genomes and is marked by broken arrow.

cn26/*27* was also observed. The lysogeny module did not show nucleotide similarity, which corresponded with the presence of numerous unique ORFs in the PH15 genome (Tables 1 and 2; Fig. 3). Likewise, there was a lack of similarity in the DNA replication module due to the presence of numerous unique ORFs in the two genomes (Fig. 3).

DNA-packaging and morphogenesis modules. BLAST and Pfam searches suggested that ORFs *ph1* and *cn1* encoded putative small terminase subunits. Sequence analysis indicated that the putative large terminase subunit of PH15 (and not that of CNPH82) was interrupted by an intron. An ORF (*ph3*) was also predicted within this intron. The DNA segment encoding the remainder of a large terminase subunit was downstream of the intron, which was spliced in vivo (see below). Thus, ORF *ph2* was subdivided into *ph2a* and *ph2b*, with *ph3* between them (Table 2; Fig. 2). ORF *cn2* also encoded a putative large terminase subunit but was not interrupted by an intron. Terminases are enzymes involved in phage DNA packaging into an empty capsid (13). These enzymes function as heteromultimers, with the small subunit involved in DNA binding and the large subunit, with its associated ATPase activity, involved in DNA packaging (19, 67). The highly conserved GKT/S Walker box (P loop) in the N termini of phage terminases (54) was predicted in both Ph2 and Cn2.

The overall organization of genes encoding the structural module in PH15 displayed similarity to that of *S. aureus* phage 37, while the CNPH82 genome showed similarity to that of *S. aureus* phage 52A (35). ORFs *ph4* and *cn3* were 99% identical and showed homology to *Bacillus subtilis* phage SPP1 Gp6-like putative portal protein (18). Portal proteins enable DNA passage into phage heads during packaging by forming a 12-fold-symmetrical ring (40, 74). ORFs *ph5* and *cn4* showed homology to the phage Mu protein F-like putative minor head protein.

The organization of genes in the CNPH82 genome downstream of *cn4* differed from that in the PH15 genome in that two CNPH82 ORFs, *cn5* and *cn6*, were unique to the CNPH82 genome (Fig. 2). These ORFs encoded proteins of unknown function. Similarity searches indicated that whereas Cn6 displayed 38% to 54% identity (over 27 to 37 amino acids) with ORF products of *S. aureus* phages of clade IIC, such as phages 96, 55, 29, and 52A, Cn5 displayed 72% identity (24/33) only with ORF147 of *S. aureus* phage EW (35). ORFs *ph6* and *cn7* encoded a putative phage minor capsid protein. ORF *ph7* was predicted to encode a putative major capsid protein (MCP). No MCP was predicted from the CNPH82 genome. The PH15 MCP, which shares high sequence identity with the *S. aureus* phage phiETA MCP, falls within the group of HK97-like MCPs due to the presence of nine strictly conserved amino acid residues (25). Neither a prohead protease nor a scaffolding protein was predicted in the genome, suggesting that the assembly of capsid differs from that in lambda-like phages. In lambda-like phages, a viral protease is involved in the processing of the prohead protein (25).

ORFs *ph16* and *cn17* were identified as putative tape measure proteins (TMPs). The TMP is one of the largest proteins in the phage genome and is found in almost all *Siphoviridae* phages (62). Phylogenetic analyses have revealed that phage

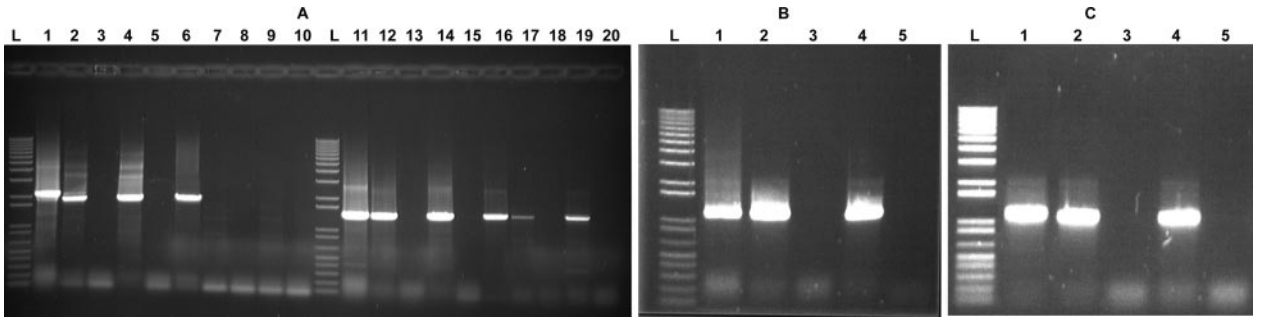

FIG. 4. In vivo splicing of intron DNA from the PH15 *lys* gene. (A) RT-PCR was conducted on RNA isolated from PH15- or CNPH82-infected host strain HER 1292 with primer pairs Lys-F/Ph24b-Lys-R (lanes 1 to 5) and Lys-F/Cn25-Lys-R (lanes 6 to 10) (Table 1). RT-PCR was also done using primer pairs Ph28-Lys-F/Ph28-Lys-R (lanes 11 to 15) and Cn30-Lys-F/Cn30-Lys-R (lanes 16 to 20) (Table 1). The template used for the PCR was as follows: lanes 1 and 11, PH15 genomic DNA; lanes 2 and 12, RNA isolated at 15 min p.i. from the PH15-infected host (cDNA); lanes 3 and 13, Similar to lanes 2 and 12 but with no RT; lanes 4 and 14, RNA isolated at 25 min p.i. from the PH15-infected host (cDNA); lanes 5 and 15, similar to lanes 4 and 14 but with no RT; lanes 6 and 16, CNPH82 genomic DNA; lanes 7 and 17, RNA isolated at 15 min p.i. from the CNPH82-infected host (cDNA); lanes 8 and 18, similar to lanes 7 and 17 but with no RT; lanes 9 and 19, RNA isolated at 25 min p.i. from the CNPH82-infected host; lanes 10 and 20, similar to lanes 9 and 19 but with no RT. (B and C) RT-PCR on RNA isolated from the PH15-infected host with helicase-specific primers (B) and from the CNPH82-infected host with portal protein-specific primers (C). Lanes: 1, genomic DNA; 2, RNA isolated at 15 min p.i. (cDNA); 3, similar to lane 2 but with no RT; 4, RNA isolated at 25 min p.i. (cDNA); 5, similar to lane 4 but with no RT; L, 1-kb plus DNA ladder (Invitrogen).

TMPs frequently contain a soluble lytic transglycosylase domain (66), and lytic transglycosylases have been proposed to be involved in phage DNA entry during early stages of infection (30, 39). However, the soluble lytic transglycosylase domain was not predicted in Ph16 and Cn17, suggesting that these phages utilize a different mechanism to deliver phage DNA during the early infection stages.

In phage λ, the length of the phage tail is directly proportional to the size of the TMP, where one amino acid of the TMP equals 0.15 nm of tail length (28, 61). Using this equation, the length of the tails in PH15 and CNPH82 corresponds to 155 nm (TMP, 1,034 amino acids), which correlates well with our estimation from the electron micrographs. Similar calculations of tail length based on TMP size have also been made for phages that are unrelated to λ (11, 44, 62, 81, 82).

Lysis module. The lysis module was located after the structural module and included ORFs *ph24a* and *ph24b*, *ph27*, and *ph28* in PH15 and *cn25*, *cn29*, and *cn30* in CNPH82 (Fig. 2). The holin-endolysin dual-lysis system responsible for cell lysis and phage progeny release in double-stranded DNA phages was present in both phages. The putative holins of both phages (Ph27 and Cn29) were composed of 90 amino acids, displayed 100% sequence identity with each other, and contained two transmembrane regions. The stop codons of ORFs *ph27* and *cn29* overlapped the corresponding downstream ORFs by 1 bp in a different reading frame. A similar organization is seen in the *S. aureus* phage K (60).

The endolysins encoded by ORFs *ph28* and *cn30* shared 99% sequence identity with each other and contained a CHAP domain (residues 14 through 144) (3) as well as an *N*-acetylmuramoyl-L-alanine amidase domain (residues 170 through 350). Both PH15 and CNPH82 encoded an additional peptidoglycan hydrolase (amidase) (Ph24a and -24b and Cn25, respectively) which contained an N-terminal CHAP domain as well as a C-terminal endo-β-*N*-acetyl-glucosaminidase domain. This amidase exhibited 73% identity with the amidase from *S. aureus* phage EW (35) and 65% identity with Ply187 from *S. aureus* phage 187 (43). ORF *ph24* (product of ORFs *ph24a* and *ph24b*) contained an intron which is spliced in vivo (see below). The arrangement of genes in the lysis module of CNPH82 resembled that of the clade IIC of *S. aureus* phages, with a putative tail fiber protein (Cn26) immediately downstream of amidase Cn25 (35). RT-PCR analysis revealed that while both endolysins were expressed during the PH15 infection cycle, only the endolysin encoded by ORF *cn30* was expressed during the CNPH82 infection cycle (Fig. 4).

Lysogeny module. The overall genetic organization of the lysogeny modules of both phages was similar to that of other *Siphoviridae* infecting low-GC-content gram-positive bacteria. In these phages, the lysogeny-related genes are compactly organized, unlike the case for *Siphoviridae* infecting high-GC-content gram-positive bacteria, where the lysogeny-related genes are more loosely organized and interspersed by many ORFs (45). The PH15 lysogeny module had more unique ORFs (*orf32*, *orf33*, *orf37*, and *orf38*) than CNPH82, while CNPH82 had only two unique ORFs (*orf33* and *orf37*). These differences could be accounted for by gene transfer among *S. aureus* phages (35) or DNA rearrangements/recombinational events similar to those observed in *S. thermophilus* phages, where the lysogeny module is shown to be a recombination hot-spot (45).

The nucleotide sequences of ORFs *ph31* and *cn32* were nearly identical and were predicted to encode site-specific phage recombinases (*int*). Preliminary experiments suggested that unlike those in other phages (23, 41, 82), the *attPP′* site might not be located close to the *int* gene (unpublished data). An excisionase-encoding gene (*xis*) was not identified in either the PH15 or CNPH82 genome, similar to the case for other phages where the *xis* gene is absent but an *int* gene is predicted (23, 32, 82).

ORFs *ph35* and *cn35* did not share sequence similarity; however, both ORFs showed homology to putative cI-like repressor protein involved in suppression of the phage lytic cycle. Likewise, ORFs *ph36* and *cn36* did not share sequence simi-

larity, but were homologous to the putative Cro-like repressor protein of the HTH-XRE family, which is required for the lytic cycle. However, ORFs *ph39* and *cn38* shared sequence similarity and were predicted to encode putative antirepressors.

Two adjacent and outward-facing putative promoters for the putative *cI*- and *cro*-like repressor genes were identified in both phages. Additionally, a 19-bp overlapping direct repeat between the two putative promoters of *cn35 and cn36* was also identified. Sequence repeats within the promoter region have been reported to be present in *L. lactis* phage r1t (56) as well as *Listeria monocytogenes* phage A118 (44). These repeats have been shown to regulate lysogeny module gene expression in phage r1t (56). The order and orientation of the *cI*- and *cro*-like repressor genes with corresponding outward facing promoters are seen in other phages, including *Lactobacillus* phage ϕg1e (29), *S. thermophilus* phages TPJ34 and Sfi21 (10, 57), *L. monocytogenes* phage A118 (44), and few *S. aureus* phages (35). The Cro-like repressor could play a role during the lytic cycle in lysogenized hosts, as suggested by Lucchini et al. (45).

An interesting biochemical property of the PH15 cI-like repressor is that the predicted isoelectric point (pI) is 8.3 (basic protein), unlike the CNPH82 cI-like repressor (pI 5.5) and other λ-like cI repressors that are acidic. Therefore, assignment of a possible function on the basis of non-sequence-alignment-based homology parameters as suggested by Chandry et al. (15) cannot be applied for the PH15 cI-like repressor.

DNA replication and metabolism module. Both PH15 and CNPH82 contained ORFs encoding proteins involved in DNA replication and metabolism (Tables 1 and 2). ORF *ph54* displayed homology to a Holliday junction resolvase (RusA) homolog, while ORF *cn57* displayed homology to a staphylococcal nuclease homolog. A 13-bp direct repeat was present in *cn45.* ORFs *ph51* and *cn47* showed homology to the DnaB_C family of helicases and also contained the ATP/GTP-binding P loop. A putative dUTPase gene (ORFs *ph62* and *cn60*) was predicted in both genomes. The ORFs had 99% sequence identity and displayed high homology with dUTPase genes of several staphylococcal and lactococcal phages. Several regulatory proteins were also identified in both genomes. ORFs *ph65* and *cn62* displayed similarity to the RinB family of transcriptional regulators, while *ph68* and *cn65* displayed similarity to the RinA family of transcriptional regulators.

PH15 has two group I introns. Two introns interrupting genes with crucial enzymatic functions were identified in the PH15 genome. The first intron, referred to as *terL-I*, was present within the sequence for the putative large terminase subunit (TerL)-encoding ORFs *ph2a* and *ph2b*. The *terL-I* intron harbored a putative endonuclease-encoding ORF *ph3*, which belonged to the GIY-YIG family of class I homing endonucleases typically found in introns (4). The second intron, referred to as *lys-I*, was present between the putative endolysin (Lys)-encoding ORFs *ph24a* and *ph24b*. A similar example of introns interrupting *terL* and *lys* genes within the same genome has been reported for *Streptococcus thermophilus* phage 2972 (41). Additionally, interrupted *lys* genes have been reported to occur in *S. thermophilus* and the *S. aureus* phage K genomes (20, 59). Also, an intron interrupting the *terL* gene has been characterized for *Lactobacillus delbrueckii* phage LL-H (52), *L. delbrueckii* phage JCL1032 (64), and *Lactobacillus plantarum* phage LP65 (16).

The presence of introns in crucial genes of PH15 confirms a common theme seen in many phage genomes, i.e., that phage introns target essential function genes, unlike eubacterial introns that are present in tRNA genes. The introns target conserved regions within these genes (*lys-I* targets the CHAP domain, and *terL-I* targets the nuclease motif) (59). This is particularly true for lysin genes, which contain a defined homing sequence (20). However, unlike the streptococcal intronless lysin genes that have a conserved 5′ATTT3′ sequence immediately upstream of the intron insertion site (20), both the intronless CNPH82 lysin gene (*cn25*) and the intron-containing PH15 lysin gene (*ph24*) have the sequence 5′GTGT3′, suggesting that intron homing in staphylococcal lysin genes may depend on a different homing sequence(s).

In vivo splicing in the *lys* RNA transcripts was tested by RT-PCR using cDNA prepared from HER 1292 RNA. Primers were designed based on identical sequences in *lys*-containing ORFs *ph24a*/*ph24b* and *cn25* (primer sequences are available on request). Primers specific for the PH15 helicase gene (*ph51*) and the CNPH82 portal protein gene (*cn3*) were designed as positive controls. PH15 and CNPH82 genomic DNAs were also included in the analysis. A 2.1-kb PCR product was amplified from PH15 genomic DNA, while a ~1.9-kb RT-PCR product was amplified from PH15-infected HER 1292 cDNA (Fig. 4A), indicating that a 246-bp intron was excised from the RNA precursor. In contrast, both CNPH82 genomic DNA and CNPH82-infected HER 1292 cDNA produced a 1.9-kb product. The PCR and RT-PCR products of the PH15 helicase and CNPH82 portal protein, included as positive controls, showed identical 1.2-kb bands (Fig. 4B and C).

RT-PCR analysis was also conducted to test in vivo splicing in the *terL* RNA transcripts by using cDNA prepared as described above. Primers recognizing identical sequences in both the PH15 and CNPH82 genomes that were located in *terL*-carrying ORFs *ph2a*/*ph2b* and *cn2* were designed. As a control, PCR was conducted using the same primer set with PH15 and CNPH82 genomic DNAs. A 750-bp RT-PCR product was amplified for PH15-infected HER 1292 cDNA, in contrast to the 1,720-bp PCR product amplified from the PH15 genomic DNA (Fig. 5), suggesting that a 971-bp intron was excised from the RNA precursor. The spliced *terL* mRNA was 1,263 bp long. In contrast, a 750-bp product was obtained when PCR was conducted with CNPH82 genomic DNA and CNPH82-infected HER 1292 cDNA (Fig. 5).

Secondary structure of the introns. The RT-PCR product was sequenced to determine the exact exon-intron boundary. As is the case with other group I introns, splicing of *lys-I* occurred after uridine (coordinate 21315, within *ph24a*) and guanosine (coordinate 21561, upstream of *ph24b*) residues. No new codon was generated, while the reading frame was maintained with a tryptophan residue following the conserved cysteine residue of the CHAP domain (Fig. 6, left panel). Similar to the case for *lys-I*, splicing of *terL-I* occurred after uridine (coordinate 1518, within *ph2a*) and guanosine (coordinate 2488, 2 bp upstream of *ph2b*) residues. A new codon, UAC rather than UAU, with the conserved coding potential for tyrosine was created at the splice junction (Fig. 6, right panel). The absence of an intron in CNPH82 could be due to sequence variation between PH15 and CNPH82 near the integration site, similar to that seen in *S. thermophilus* phages (20, 41).

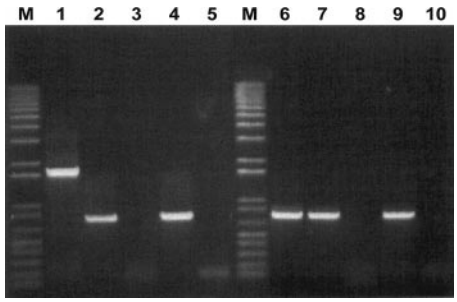

FIG. 5. In vivo splicing of intron DNA from the PH15 *terL* gene. RT-PCR was conducted on RNA isolated from PH15- or CNPH82-infected host strain HER 1292 with a *terL*-specific primer pair. The template used for the PCR was as follows: lane 1, PH15 genomic DNA; lane 2, RNA isolated at 15 min p.i. from the PH15-infected host (cDNA); lane 3, similar to lane 2 but with no RT; lane 4, RNA isolated at 25 min p.i. from the PH15-infected host (cDNA); lane 5, similar to lane 4 but with no RT; lane 6, CNPH82 genomic DNA; lane 7, RNA isolated at 15 min p.i. from the CNPH82-infected host (cDNA); lane 8, similar to lane 7 but with no RT; lane 9, RNA isolated at 25 min p.i. from the CNPH82-infected host; lane 10, similar to lane 9 but with no RT.

The putative secondary structures of *lys-I* and *terL-I* were determined by using comparative sequence analysis based on secondary structure predictions (51) as well as the mFOLD software (http://www.bioinfo.rpi.edu/applications/mfold) (83). Both introns could be folded into a secondary structure containing the entire canonical group I structural features essential for self-splicing and formation of the catalytic core (Fig. 6). The overall structure of corresponding introns were similar to that of the *S. thermophilus* phage 2972 introns, although closer structural comparisons revealed that some elements were swapped between the two PH15 introns (41). For example, in PH15 *terL-I*, P3.1 and P3.2 stems were predicted, similar to those in the phage 2972 *lys-I* intron, while P9 and P9.1 stems predicted in PH15 *lys-I* were predicted in the phage 2972 *ter-I* intron (Fig. 6). Also, ORF *ph3* was present in the *terL-I* intron, similar to the phage 2972 *lys-I* intron, which contained a putative endonuclease-encoding ORF (*orf28*) (41). A poly(U) loop was predicted in P6a of *terL-I* (Fig. 6, right panel). A putative tertiary interaction (P12) and the internal guide sequence, which facilitates the splicing process by bringing P1 and P10 together, were identified in the *lys-I* intron (Fig. 6, left panel) (14).

A BLASTN search with the *lys-I* intron sequence revealed significant matches to phage Twort *orf142* introns I2 and I3 (E value = 2.4e−04; identity, 39/43 bases). Since Twort *orf142* introns I2 and I3 are closely related to the phage T4 *nrdB* intron (37), pairwise comparisons of the *orf142* I2 and I3 introns and the T4 *nrdB* intron to *lys-I* intron revealed 72 to 75% and 60% overall identity, respectively (Table 3).

The structural differences between *lys-I* and *terL-I* indicated that the two introns were from different sources. The predicted secondary structures of both *terL-I* and *lys-I* contain all group I canonical structures. However, *terL-I* belongs to the IA1 subgroup while *lys-I* belongs to the IA2 subgroup, because *terL-I* contains the P7.1-P7.1a stems while *lys-I* has the decorative P7.1-P7.2 stems separated by a G-U-A sequence. The introns of subgroup IA2 occur more commonly in phage genomes than those of subgroup IA1 (51), although introns belonging to subgroup IA1 recently have been reported to occur in *Bacillus* (36, 38), *Lactobacillus* (64), *Streptococcus* (41), and

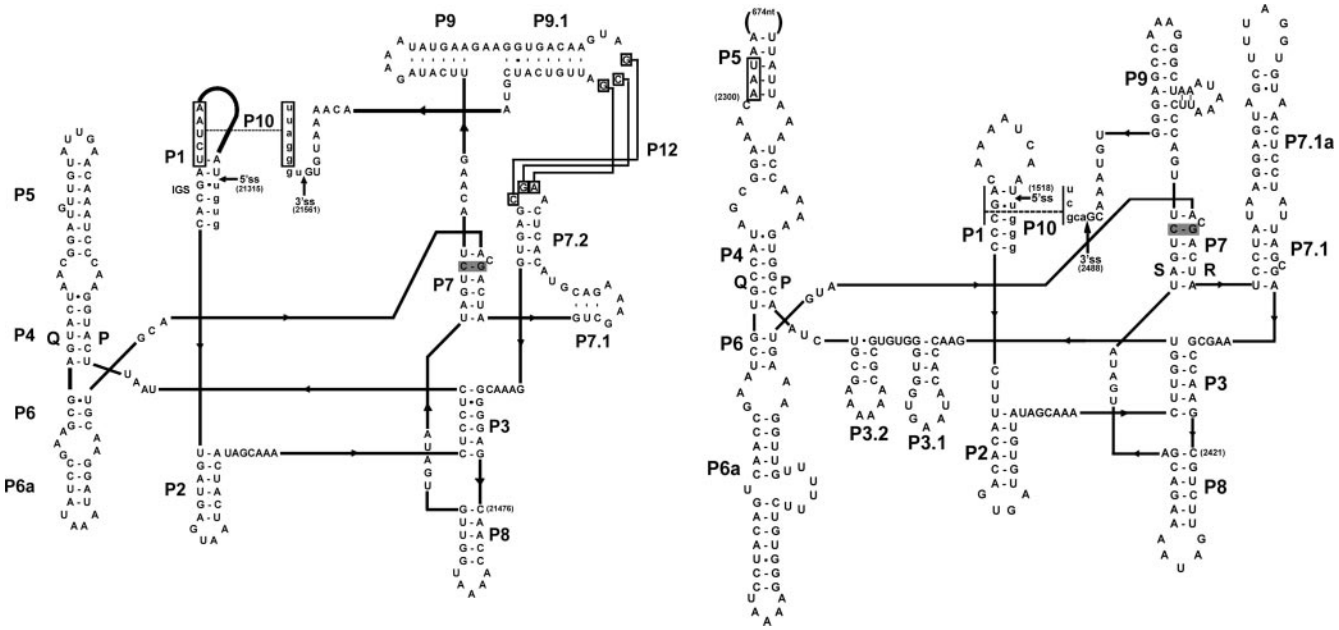

FIG. 6. Secondary structure predictions for *lys-I* (left) and *terL-I* (right) introns in the PH15 genome. The secondary structures are represented according to the structural convention of Burke et al. (12). Lower- and uppercase letters denote exon and intron sequences, respectively. The 5′ and 3′ splice sites (ss) are indicated by arrows. The conserved structural elements P1 through P10 and sequences P, Q, R, and S are indicated, and a putative tertiary interaction, P12, within the *lys-I* intron is shown (50). The putative guanosine binding site in P7 is shaded. The structural elements are connected with bold lines with pointers indicating the 5′-to-3′ direction. The stop codon of ORF *ph3* in *terL-I* is boxed. The predicted internal guide sequence (IGS) within the *lys-I* intron is marked. The nucleotide position in PH15 genome is shown in parentheses.

TABLE 3. Sequence identity between the Ph15 *lys-I* intron, Twort *orf142* introns, and T4 *nrdB* intron

Intron	No. of identical bases/total (% identity) in:		
	orf142 I2	*orf142* I3	T4 *nrdB*
lys-I	176/245 (72)	183/245 (75)	144/245 (59)
orf142 I2		210/245 (86)	156/245 (64)
orf142 I3			156/245 (64)

Synechococcus (53) phages. PH15 *terL-I* is similar to most subgroupIA1 introns because they contain an endonuclease-encoding ORF (*ph3* in PH15), unlike the *Streptococcus thermophilus* phage 2972 (41) and *Lactobacillus delbruekii* phage JCL1032 (64) introns, which do not contain an ORF.

Comparative genomics and phylogenetic analysis. Pairwise nucleotide sequence comparison of all *S. aureus* phage genomes reported in the GenBank database was conducted using dot plot analysis (Fig. 7). PH15 and CNPH82, along with *S. aureus* phage phiETA, show high sequence similarity with class II clade C phages and therefore belong in this clade. We also identified a novel clade within class II, clade D, consisting of *S. aureus* phages phi13, PV83, PVL, and N315. This fourth clade generated by our analyses also included phage 77, which previously could not be classified due to its similarity to members of multiple clades (35).

A phylogenetic tree based on proteomic distances between all known *Siphoviridae* phages in GenBank was generated using the approach detailed by Rohwer and Edwards (66). All staphylococcal phages, including PH15 and CNPH82, were present on one major branch of the tree (Fig. 8). The secondary grouping of phages within the staphylococcal branch correlated well with the results of our pairwise nucleotide analysis. Every phage of the corresponding clades clustered within a branch in the proteomic tree (Fig. 8). PH15 and CNPH82

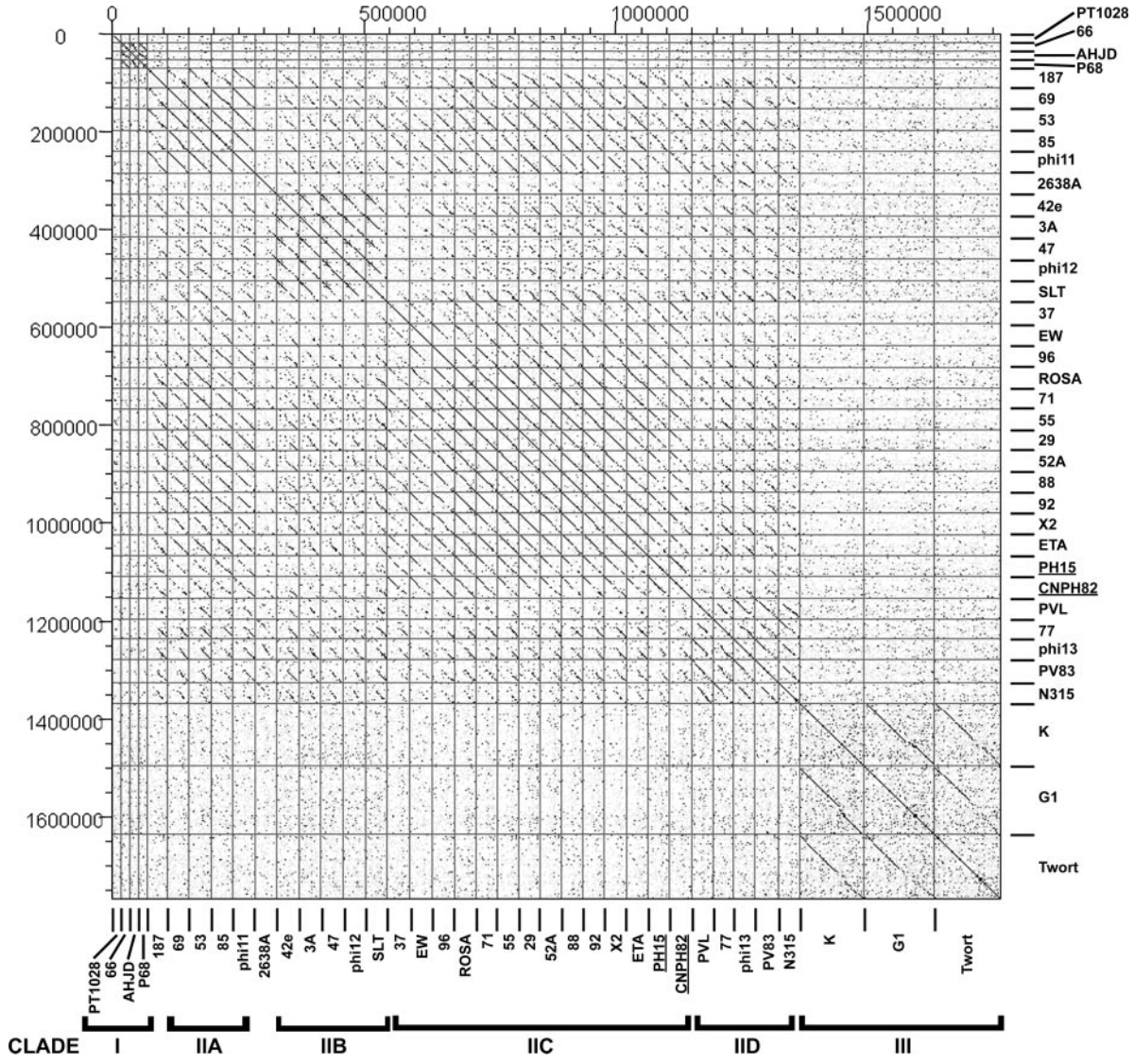

FIG. 7. Dot plot matrix of PH15 and CNPH82 with *S. aureus* phages. The nucleotide sequences of PH15 and CNPH82 along with those of all *S. aureus* phages reported in the database to date were compared using the DOTTER program (70). The sliding window was set at 25 bp.

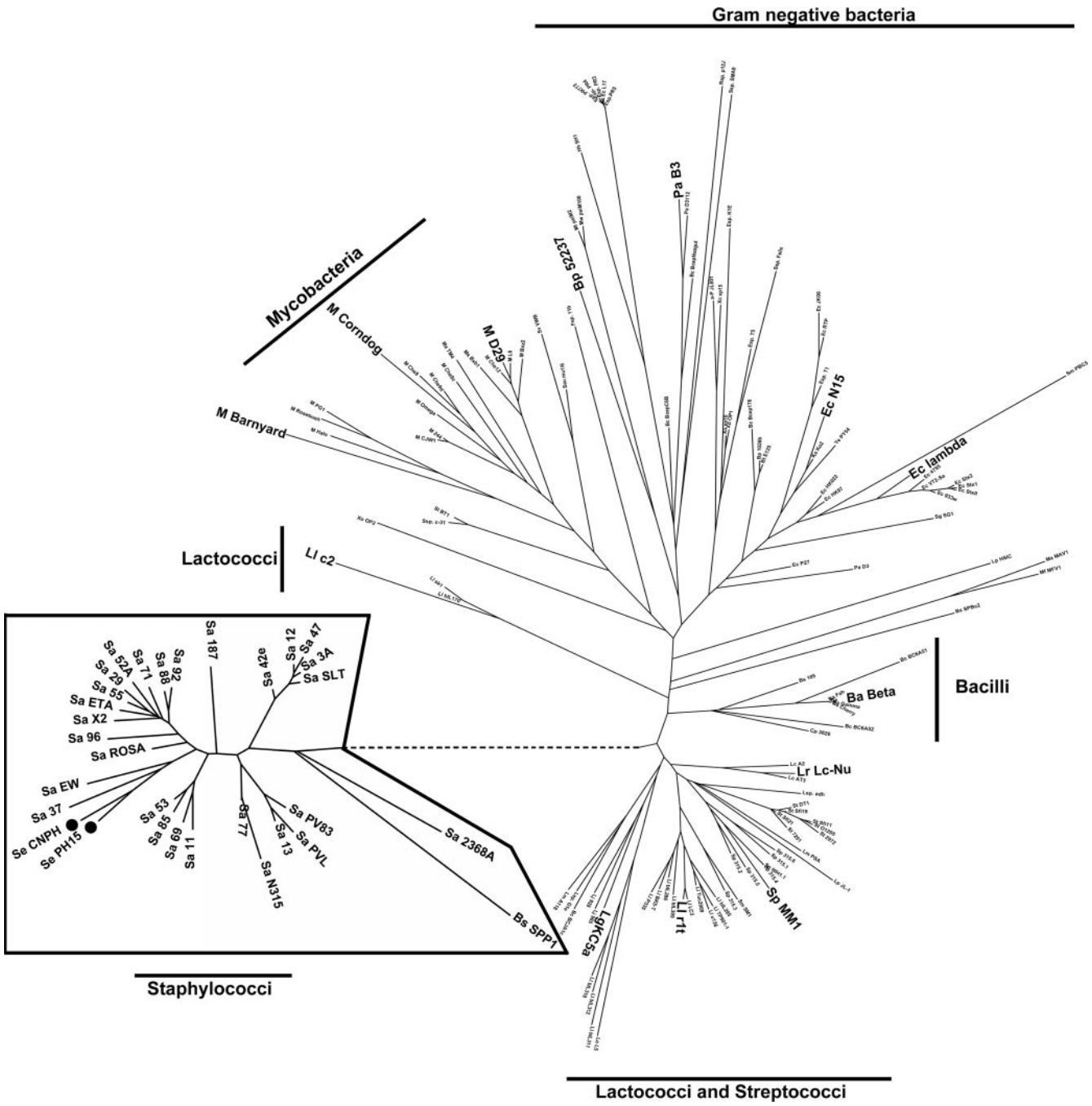

FIG. 8. The *Siphoviridae* section within the phage proteomic tree. The relationship between PH15, CNPH82, and all phages belonging to *Siphoviridae* available in the database is presented. The tree was constructed using 153 phage genomes within *Siphoviridae* collated from the GenBank database as well as the website http://phage.sdsu.edu/~rob/phage. The staphylococcal subbranch was expanded to show the phages clearly. PH15 and CNPH82 are highlighted by filled circles. Representative phages of different groups are highlighted in larger font. A high-resolution version of the tree is presented in Fig. S1A in the supplemental material. The phage abbreviations are listed in Table S1A in the supplemental material.

clustered with other phages of class II clade C, and the members of our newly described clade D clustered together, further confirming the close relationship of these phages.

The number of complete genome sequences of *Siphoviridae* phages in the GenBank database has increased exponentially since the first phage proteomic tree was published (66). Rohwer and Edwards (66) observed that the siphophage branch within the proteomic tree was most problematic due to rampant lateral gene transfer among siphophages. However, due to the large increase in the number of siphophage genomes available (153 genomes), our tree shows much higher resolution, which confirms the general findings and further

refines the siphophage branch of the phage proteomic tree. Rohwer and Edwards (66) identified five separate groups of siphophages, in which lambda-like, D29-like, SK-1-like, and TP901-like phages formed monophyletic taxa and Sfi21-like phages formed a polyphyletic group. We confirm this structure and observe that the mycobacterial and gram-negative groups are the most diffuse and therefore the most diverged of the five groups. Lactococci and streptococci group together in the broader context of the tree but show distinct clusters within the group. Bacilli and staphylococci show distinct and more intragroup clustering, indicating younger lineages. In fact, we were able to define a novel clade within the *Staphylococcus* grouping which was also confirmed by our pairwise sequence comparison. One intriguing finding in our phylogenetic analyses was the presence of *Bacillus subtilis* phage SPP1 clustering at the edge of the *Staphylococcus* branch of the tree. This could be due to high similarity between the head and tail morphogenesis proteins. This result may stem from a common ancestry for these proteins or a more recent gene transfer event.

Conclusion. In this study, we have presented the first detailed genomic and molecular characterization of two *S. epidermidis* phages. We have demonstrated that the two phages are highly similar in gene content, although differences exist within the lysogeny and DNA replication modules. We have identified two introns within essential genes in the PH15 genome, one of which belongs to subgroup 1A1. Phylogenetic analysis revealed that PH15 and CNPH82 are very similar to the *S. aureus* phages, and based on comparative sequence analysis, we propose a new clade D within class II to classify staphylococcal phages. Sequence information for more *S. epidermidis* phages will provide better insight into the role of phages in *S. epidermidis* pathogenesis and evolution.

ACKNOWLEDGMENTS

We thank John Curtin, CDC, Atlanta, GA, for his gifts of *S. epidermidis* strain HER 1292 and phages PH15 and CNPH82; Barry Kreiswirth for *S. epidermidis* strain ATCC 12228; Xuning Wang at the Department of Information Technology at The Rockefeller University for bioinformatic analyses; Elena Sphicas at the Bio-Imaging Resource Center at The Rockefeller University for electron microscopy; and Chad Euler, Daniel Nelson, and Ray Schuch for helpful discussions.

This research was supported by USPHS grants AI-057472 and AI-056510 to V.A.F.

REFERENCES

1. **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.
2. **Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389–3402.
3. **Bateman, A., and N. D. Rawlings.** 2003. The CHAP domain: a large family of amidases including GSP amidase and peptidoglycan hydrolases. Trends Biochem. Sci. **28:**234–237.
4. **Belfort, M., and R. J. Roberts.** 1997. Homing endonucleases: keeping the house in order. Nucleic Acids Res. **25:**3379–3388.
5. **Black, L. W.** 1989. DNA packaging in dsDNA bacteriophages. Annu. Rev. Microbiol. **43:**267–292.
6. **Breitbart, M., P. Salamon, B. Andresen, J. M. Mahaffy, A. M. Segall, D. Mead, F. Azam, and F. Rohwer.** 2002. Genomic analysis of uncultured marine viral communities. Proc. Natl. Acad. Sci. USA **99:**14250–14255.
7. **Breitbart, M., L. Wegley, S. Leeds, T. Schoenfeld, and F. Rohwer.** 2004. Phage community dynamics in hot springs. Appl. Environ Microbiol. **70:** 1633–1640.
8. **Brussow, H., and F. Desiere.** 2001. Comparative phage genomics and the evolution of *Siphoviridae*: insights from dairy phages. Mol. Microbiol. **39:** 213–222.
9. **Brussow, H., and R. W. Hendrix.** 2002. Phage genomics: small is beautiful. Cell **108:**13–16.
10. **Bruttin, A., F. Desiere, S. Lucchini, S. Foley, and H. Brussow.** 1997. Characterization of the lysogeny DNA module from the temperate *Streptococcus thermophilus* bacteriophage phi Sfi21. Virology **233:**136–148.
11. **Bukovska, G., L. Klucar, C. Vlcek, J. Adamovic, J. Turna, and J. Timko.** 2006. Complete nucleotide sequence and genome analysis of bacteriophage BFK20—a lytic phage of the industrial producer *Brevibacterium flavum*. Virology **348:**57–71.
12. **Burke, J. M., M. Belfort, T. R. Cech, R. W. Davies, R. J. Schweyen, D. A. Shub, J. W. Szostak, and H. F. Tabak.** 1987. Structural conventions for group I introns. Nucleic Acids Res. **15:**7217–7221.
13. **Catalano, C. E.** 2000. The terminase enzyme from bacteriophage lambda: a DNA-packaging machine. Cell Mol. Life Sci. **57:**128–148.
14. **Cech, T. R.** 1988. Conserved sequences and structures of group I introns: building an active site for RNA catalysis—a review. Gene **73:**259–271.
15. **Chandry, P. S., S. C. Moore, J. D. Boyce, B. E. Davidson, and A. J. Hillier.** 1997. Analysis of the DNA sequence, gene expression, origin of replication and modular structure of the *Lactococcus lactis* lytic bacteriophage sk1. Mol. Microbiol. **26:**49–64.
16. **Chibani-Chennoufi, S., M. L. Dillmann, L. Marvin-Guy, S. Rami-Shojaei, and H. Brussow.** 2004. *Lactobacillus plantarum* bacteriophage LP65: a new member of the SPO1-like genus of the family *Myoviridae*. J. Bacteriol. **186:** 7069–7083.
17. **Curtin, J. J., and R. M. Donlan.** 2006. Using bacteriophages to reduce formation of catheter-associated biofilms by *Staphylococcus epidermidis*. Antimicrob. Agents Chemother. **50:**1268–1275.
18. **Dube, P., P. Tavares, R. Lurz, and M. van Heel.** 1993. The portal protein of bacteriophage SPP1: a DNA pump with 13-fold symmetry. EMBO J. **12:** 1303–1309.
19. **Duffy, C., and M. Feiss.** 2002. The large subunit of bacteriophage lambda's terminase plays a role in DNA translocation and packaging termination. J. Mol. Biol. **316:**547–561.
20. **Foley, S., A. Bruttin, and H. Brussow.** 2000. Widespread distribution of a group I intron and its three deletion derivatives in the lysin gene of *Streptococcus thermophilus* bacteriophages. J. Virol. **74:**611–618.
21. **Gill, S. R., D. E. Fouts, G. L. Archer, E. F. Mongodin, R. T. Deboy, J. Ravel, I. T. Paulsen, J. F. Kolonay, L. Brinkac, M. Beanan, R. J. Dodson, S. C. Daugherty, R. Madupu, S. V. Angiuoli, A. S. Durkin, D. H. Haft, J. Vamathevan, H. Khouri, T. Utterback, C. Lee, G. Dimitrov, L. Jiang, H. Qin, J. Weidman, K. Tran, K. Kang, I. R. Hance, K. E. Nelson, and C. M. Fraser.** 2005. Insights on evolution of virulence and resistance from the complete genome analysis of an early methicillin-resistant *Staphylococcus aureus* strain and a biofilm-producing methicillin-resistant *Staphylococcus epidermidis* strain. J. Bacteriol. **187:**2426–2438.
22. **Golding, B., and J. Felsenstein.** 1990. A maximum likelihood approach to the detection of selection from a phylogeny. J. Mol. Evol. **31:**511–523.
23. **Govind, R., J. A. Fralick, and R. D. Rolfe.** 2006. Genomic organization and molecular characterization of *Clostridium difficile* bacteriophage PhiCD119. J. Bacteriol. **188:**2568–2577.
24. **Heczko, P. B., G. Pulverer, A. Kasprowicz, and A. Klein.** 1977. Evaluation of a new bacteriophage set for typing of *Staphylococcus epidermidis* strains. J. Clin. Microbiol. **5:**573–577.
25. **Helgstrand, C., W. R. Wikoff, R. L. Duda, R. W. Hendrix, J. E. Johnson, and L. Liljas.** 2003. The refined structure of a protein catenane: the HK97 bacteriophage capsid at 3.44 A resolution. J. Mol. Biol. **334:**885–899.
26. **Huebner, J., and D. A. Goldmann.** 1999. Coagulase-negative staphylococci: role as pathogens. Annu. Rev. Med. **50:**223–236.
27. **Johnsen, M. G., H. Neve, F. K. Vogensen, and K. Hammer.** 1995. Virion positions and relationships of lactococcal temperate bacteriophage TP901-1 proteins. Virology **212:**595–606.
28. **Katsura, I.** 1987. Determination of bacteriophage lambda tail length by a protein ruler. Nature **327:**73–75.
29. **Kodaira, K. I., M. Oki, M. Kakikawa, N. Watanabe, M. Hirakawa, K. Yamada, and A. Taketo.** 1997. Genome structure of the *Lactobacillus* temperate phage phi g1e: the whole genome sequence and the putative promoter/repressor system. Gene **187:**45–53.
30. **Koraimann, G.** 2003. Lytic transglycosylases in macromolecular transport systems of Gram-negative bacteria. Cell Mol. Life Sci. **60:**2371–2388.
31. **Kreiswirth, B. N., S. Lofdahl, M. J. Betley, M. O'Reilly, P. M. Schlievert, M. S. Bergdoll, and R. P. Novick.** 1983. The toxic shock syndrome exotoxin structural gene is not detectably transmitted by a prophage. Nature **305:**709–712.
32. **Kropinski, A. M.** 2000. Sequence of the genome of the temperate, serotype-converting, *Pseudomonas aeruginosa* bacteriophage D3. J. Bacteriol. **182:** 6066–6074.
33. **Kumar, S., K. Tamura, and M. Nei.** 2004. MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform. **5:**150–163.
34. **Kwan, T., J. Liu, M. Dubow, P. Gros, and J. Pelletier.** 2006. Comparative genomic analysis of 18 *Pseudomonas aeruginosa* bacteriophages. J. Bacteriol. **188:**1184–1187.
35. **Kwan, T., J. Liu, M. DuBow, P. Gros, and J. Pelletier.** 2005. The complete

genomes and proteomes of 27 *Staphylococcus aureus* bacteriophages. Proc. Natl. Acad. Sci. USA **102:**5174–5179.
36. **Landthaler, M., and D. A. Shub.** 2003. The nicking homing endonuclease I-BasI is encoded by a group I intron in the DNA polymerase gene of the *Bacillus thuringiensis* phage Bastille. Nucleic Acids Res. **31:**3071–3077.
37. **Landthaler, M., and D. A. Shub.** 1999. Unexpected abundance of self-splicing introns in the genome of bacteriophage Twort: introns in multiple genes, a single gene with three introns, and exon skipping by group I ribozymes. Proc. Natl. Acad. Sci. USA **96:**7005–7010.
38. **Lazarevic, V.** 2001. Ribonucleotide reductase genes of *Bacillus* prophages: a refuge to introns and intein coding sequences. Nucleic Acids Res. **29:**3212–3218.
39. **Lehnherr, H., A. M. Hansen, and T. Ilyina.** 1998. Penetration of the bacterial cell wall: a family of lytic transglycosylases in bacteriophages and conjugative plasmids. Mol. Microbiol. **30:**454–457.
40. **Leiman, P. G., S. Kanamaru, V. V. Mesyanzhinov, F. Arisaka, and M. G. Rossmann.** 2003. Structure and morphogenesis of bacteriophage T4. Cell Mol. Life Sci. **60:**2356–2370.
41. **Levesque, C., M. Duplessis, J. Labonte, S. Labrie, C. Fremaux, D. Tremblay, and S. Moineau.** 2005. Genomic organization and molecular analysis of virulent bacteriophage 2972 infecting an exopolysaccharide-producing *Streptococcus thermophilus* strain. Appl. Environ. Microbiol. **71:**4057–4068.
42. **Lim, S. M., and S. A. Webb.** 2005. Nosocomial bacterial infections in intensive care units. I. Organisms and mechanisms of antibiotic resistance. Anaesthesia **60:**887–902.
43. **Loessner, M. J., S. Gaeng, and S. Scherer.** 1999. Evidence for a holin-like protein gene fully embedded out of frame in the endolysin gene of *Staphylococcus aureus* bacteriophage 187. J. Bacteriol. **181:**4452–4460.
44. **Loessner, M. J., R. B. Inman, P. Lauer, and R. Calendar.** 2000. Complete nucleotide sequence, molecular analysis and genome structure of bacteriophage A118 of *Listeria monocytogenes*: implications for phage evolution. Mol. Microbiol. **35:**324–340.
45. **Lucchini, S., F. Desiere, and H. Brussow.** 1999. Comparative genomics of *Streptococcus thermophilus* phage species supports a modular evolution theory. J. Virol. **73:**8647–8656.
46. **Lucchini, S., F. Desiere, and H. Brussow.** 1999. Similarly organized lysogeny modules in temperate *Siphoviridae* from low GC content gram-positive bacteria. Virology **263:**427–435.
47. **Lukashin, A. V., and M. Borodovsky.** 1998. GeneMark.hmm: new solutions for gene finding. Nucleic Acids Res. **26:**1107–1115.
48. **Lupas, A., M. Van Dyke, and J. Stock.** 1991. Predicting coiled coils from protein sequences. Science **252:**1162–1164.
49. **Mack, D., P. Becker, I. Chatterjee, S. Dobinsky, J. K. Knobloch, G. Peters, H. Rohde, and M. Herrmann.** 2004. Mechanisms of biofilm formation in *Staphylococcus epidermidis* and *Staphylococcus aureus*: functional molecules, regulatory circuits, and adaptive responses. Int. J. Med. Microbiol. **294:**203–212.
50. **Michel, F., L. Jaeger, E. Westhof, R. Kuras, F. Tihy, M. Q. Xu, and D. A. Shub.** 1992. Activation of the catalytic core of a group I intron by a remote 3′ splice junction. Genes Dev. **6:**1373–1385.
51. **Michel, F., and E. Westhof.** 1990. Modelling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. J. Mol. Biol. **216:**585–610.
52. **Mikkonen, M., and T. Alatossava.** 1995. A group I intron in the terminase gene of *Lactobacillus delbrueckii* subsp. lactis phage LL-H. Microbiology **141:**2183–2190.
53. **Millard, A., M. R. Clokie, D. A. Shub, and N. H. Mann.** 2004. Genetic organization of the psbAD region in phages infecting marine *Synechococcus* strains. Proc. Natl. Acad. Sci. USA **101:**11007–11012.
54. **Mitchell, M. S., and V. B. Rao.** 2004. Novel and deviant Walker A ATP-binding motifs in bacteriophage large terminase-DNA packaging proteins. Virology **321:**217–221.
55. **National Nosocomial Infections Surveillance System.** 1999. System report data summary from January 1990-May 1999, issued June 1999. Am. J. Infect. Control **27:**520–532.
56. **Nauta, A., D. van Sinderen, H. Karsens, E. Smit, G. Venema, and J. Kok.** 1996. Inducible gene expression mediated by a repressor-operator system isolated from *Lactococcus lactis* bacteriophage r1t. Mol. Microbiol. **19:**1331–1341.
57. **Neve, H., K. I. Zenz, F. Desiere, A. Koch, K. J. Heller, and H. Brussow.** 1998. Comparison of the lysogeny modules from the temperate *Streptococcus thermophilus* bacteriophages TP-J34 and Sfi21: implications for the modular theory of phage evolution. Virology **241:**61–72.
58. **Novick, R.** 1967. Properties of a cryptic high-frequency transducing phage in *Staphylococcus aureus*. Virology **33:**155–166.
59. **O'Flaherty, S., A. Coffey, R. Edwards, W. Meaney, G. F. Fitzgerald, and R. P. Ross.** 2004. Genome of staphylococcal phage K: a new lineage of *Myoviridae* infecting gram-positive bacteria with a low G+C content. J. Bacteriol. **186:**2862–2871.
60. **O'Flaherty, S., R. P. Ross, W. Meaney, G. F. Fitzgerald, M. F. Elbreki, and A. Coffey.** 2005. Potential of the polyvalent anti-*Staphylococcus* bacteriophage K for control of antibiotic-resistant staphylococci from hospitals. Appl. Environ. Microbiol. **71:**1836–1842.
61. **Pedersen, M., S. Ostergaard, J. Bresciani, and F. K. Vogensen.** 2000. Mutational analysis of two structural genes of the temperate lactococcal bacteriophage TP901-1 involved in tail length determination and baseplate assembly. Virology **276:**315–328.
62. **Pedulla, M. L., M. E. Ford, J. M. Houtz, T. Karthikeyan, C. Wadsworth, J. A. Lewis, D. Jacobs-Sera, J. Falbo, J. Gross, N. R. Pannunzio, W. Brucker, V. Kumar, J. Kandasamy, L. Keenan, S. Bardarov, J. Kriakov, J. G. Lawrence, W. R. Jacobs, Jr., R. W. Hendrix, and G. F. Hatfull.** 2003. Origins of highly mosaic mycobacteriophage genomes. Cell **113:**171–182.
63. **Pulverer, G., J. Pillich, and A. Klein.** 1975. New bacteriophages of *Staphylococcus epidermidis*. J. Infect. Dis. **132:**524–531.
64. **Riipinen, K. A., and T. Alatossava.** 2004. Two self-splicing group I introns interrupt two late transcribed genes of prolate-headed *Lactobacillus delbrueckii* phage JCL1032. Arch. Virol. **149:**2013–2024.
65. **Rohwer, F.** 2003. Global phage diversity. Cell **113:**141.
66. **Rohwer, F., and R. Edwards.** 2002. The phage proteomic tree: a genome-based taxonomy for phage. J. Bacteriol. **184:**4529–4535.
67. **Rubinchik, S., W. Parris, and M. Gold.** 1994. The in vitro ATPases of bacteriophage lambda terminase and its large subunit, gene product A. The relationship with their DNA helicase and packaging activities. J. Biol. Chem. **269:**13586–13593.
68. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
69. **Schuch, R., and V. A. Fischetti.** 2006. Detailed genomic analysis of the Wbeta and gamma phages infecting *Bacillus anthracis*: implications for evolution of environmental fitness and antibiotic resistance. J. Bacteriol. **188:**3037–3051.
70. **Sonnhammer, E. L., and R. Durbin.** 1995. A dot-matrix program with dynamic threshold control suited for genomic DNA and protein sequence analysis. Gene **167:**GC1–GC10.
71. **Sonnhammer, E. L., G. von Heijne, and A. Krogh.** 1998. A hidden Markov model for predicting transmembrane helices in protein sequences. Proc. Int. Conf. Intell. Syst. Mol. Biol. **6:**175–182.
72. **Stevens, D. L.** 2003. Community-acquired *Staphylococcus aureus* infections: increasing virulence and emerging methicillin resistance in the new millennium. Curr. Opin. Infect. Dis. **16:**189–191.
73. **Talbot, H. W., Jr., and J. T. Parisi.** 1976. Phage typing of *Staphylococcus epidermidis*. J. Clin. Microbiol. **3:**519–523.
74. **Valpuesta, J. M., and J. L. Carrascosa.** 1994. Structure of viral connectors and their function in bacteriophage assembly and DNA packaging. Q. Rev. Biophys. **27:**107–155.
75. **Wei, W., Z. Cao, Y. L. Zhu, X. Wang, G. Ding, H. Xu, P. Jia, D. Qu, A. Danchin, and Y. Li.** 2006. Conserved genes in a path from commensalism to pathogenicity: comparative phylogenetic profiles of *Staphylococcus epidermidis* RP62A and ATCC12228. BMC Genomics **7:**112.
76. **Weinbauer, M. G., and F. Rassoulzadegan.** 2004. Are viruses driving microbial diversification and diversity? Environ. Microbiol. **6:**1–11.
77. **Wommack, K. E., and R. R. Colwell.** 2000. Virioplankton: viruses in aquatic ecosystems. Microbiol. Mol. Biol. Rev. **64:**69–114.
78. **Yang, X. M., N. Li, J. M. Chen, Y. Z. Ou, H. Jin, H. J. Lu, Y. L. Zhu, Z. Q. Qin, D. Qu, and P. Y. Yang.** 2006. Comparative proteomic analysis between the invasive and commensal strains of Staphylococcus epidermidis. FEMS Microbiol. Lett. **261:**32–40.
79. **Yao, Y., D. E. Sturdevant, A. Villaruz, L. Xu, Q. Gao, and M. Otto.** 2005. Factors characterizing *Staphylococcus epidermidis* invasiveness determined by comparative genomics. Infect. Immun. **73:**1856–1860.
80. **Zhang, Y. Q., S. X. Ren, H. L. Li, Y. X. Wang, G. Fu, J. Yang, Z. Q. Qin, Y. G. Miao, W. Y. Wang, R. S. Chen, Y. Shen, Z. Chen, Z. H. Yuan, G. P. Zhao, D. Qu, A. Danchin, and Y. M. Wen.** 2003. Genome-based analysis of virulence genes in a non-biofilm-forming *Staphylococcus epidermidis* strain (ATCC 12228). Mol. Microbiol. **49:**1577–1593.
81. **Zimmer, M., E. Sattelberger, R. B. Inman, R. Calendar, and M. J. Loessner.** 2003. Genome and proteome of *Listeria monocytogenes* phage PSA: an unusual case for programmed + 1 translational frameshifting in structural protein synthesis. Mol. Microbiol. **50:**303–317.
82. **Zimmer, M., S. Scherer, and M. J. Loessner.** 2002. Genomic analysis of *Clostridium perfringens* bacteriophage phi3626, which integrates into GuaA and possibly affects sporulation. J. Bacteriol. **184:**4359–4368.
83. **Zuker, M.** 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. **31:**3406–3415.