Extremophiles (2007) 11:95–104
DOI 10.1007/s00792-006-0014-5

ORIGINAL PAPER

**Michael Borriss · Thierry Lombardot
Frank Oliver Glöckner · Dörte Becher
Dirk Albrecht · Thomas Schweder**

Genome and proteome characterization of the psychrophilic *Flavobacterium* bacteriophage 11b

Received: 2 June 2006 / Accepted: 30 June 2006 / Published online: 25 August 2006

Abstract Virion DNA of bacteriophage 11b (*Φ*11b), which infects a psychrophilic *Flavobacterium* isolate from Arctic sea-ice, was determined to consist of 36,012 bp. With 30.6% its GC content corresponds to that of host-genus species and is the lowest of all phages of Gram-negative bacteria sequenced so far. Similarities of several of 65 predicted ORFs, genome organization and phylogeny suggest an affiliation to 'mesophilic' nonmarine siphoviruses, e.g. to bacteriophages SPP1 and HK97. Early genes presumably encode an essential recombination factor (ERF), a single strand binding (SSB) protein, an endonuclease, and a DNA methylase. The late gene segment is likely to contain a terminase, portal, minor head, protease and a major capsid gene. Five ORFs exhibited similarities to *Bacteroidetes* species and seem to reflect the host specificity of the phage. Among PAGE-separated virion proteins that were identified by MALDI-ToF mass spectrometry are the portal, the major capsid, and a putative conserved tail protein. The *Φ*11b genome is the first to be described of a cultivated virus infecting a psychrophilic host as well as a *Bacteroidetes* bacterium.

Keywords Psychrophilic Flavobacterium bacteriophage *Φ*11b · Genomics · Proteome · Low GC

Electronic Supplementary Material Supplementary material is available to authorised users in the online version of this article at http://dx.doi.org/10.1007/s00792-006-0014-5.

Communicated by K. Horikoshi

M. Borriss · T. Schweder (✉)
Institute of Marine Biotechnology,
W.-Rathenau-Str. 49a, 17489 Greifswald, Germany
E-mail: schweder@uni-greifswald.de
Tel.: +49-3834-864212
Fax: +49-3834-864238

T. Lombardot · F. O. Glöckner
Max-Planck-Institute for Marine Microbiology,
Celsius-Str. 1, 28359 Bremen, Germany

D. Becher · D. Albrecht
Institute of Microbiology, Ernst-Moritz-Arndt-University,
F.-L.-Jahn-Str. 15, 17487 Greifswald, Germany

T. Schweder
Institute of Pharmacy, Ernst-Moritz-Arndt-University,
F.-L.-Jahn-Str. 17, 17487 Greifswald, Germany

Introduction

The isolation and cultivation of only a few bacteriophages from the polar region has been reported to day. *Φ*11b was one out of a set of phages that were recently isolated together with their psychrophilic host bacteria from Arctic sea-ice and cultured under laboratory conditions (Borriss et al. 2003). The phage has a chromosome of ds DNA and morphologically resembles *Escherichia coli* phage *λ* (Siphoviridae, morphotype B1). While 16S rDNA sequencing indicated that the host bacterium of *Φ*11b, *Bacteroidetes* bacterium 11B, may represent a novel *Flavobacterium* species, no further characterization of the isolate has been conducted.

Flavobacteriophages are viruses that infect Flavobacteria (low GC Gram-negative bacteria of the *Bacteroidetes* phylum). According to data of one of the few published articles about flavobacteriophage-host systems (Jiang et al. 1998), *Φ*11b seems to share similarities in morphology and dimensions with a temperate phage isolate from Hawaii. The online name directory of phage names (Ackermann and Abedon 2001) lists four different virus species for the B1 morphotype of flavobacteria-infecting siphophages, whereas most phages infecting *Bacteroidetes* bacteria have been described for *Bacteroides* species (17 phage names). Bacteria of the *Cytophaga–Flavobacterium* cluster have been reported from diverse habitats, where they influence the carbon flow as numerically important aerobic heterotrophic consumers (Kirchman 2002). Strain 11B was initially described to be closely related (99.6% 16S rDNA identity) to an uncertainly classified 'Cytophaga' isolate from the Baltic

Sea (Borriss et al. 2003). In a BLASTN similarity search update, 11B's 16S rDNA turned out to be similar to several new entries. 97–99% identity was found to a few sequences from Arctic and Antarctic isolates [99% to ARK10270 (Brinkmeyer et al. 2003), 97–98% to *Flavobacterium frigoris* and *Flavobacterium degerlachei*]. Further *Flavobacterium* species with strong rDNA sequence similarities to strain 11B were isolated from genus-typical aquatic (estuarine, freshwater) or soil habitats (Kisand et al. 2002, Bradford et al. 1996). Notably, no 16S rDNA sequence of a closely related species from an offshore marine environment, which is not sea-ice associated, has been deposited in the databases so far.

While virus genome structure depends on the host species, in the past, single phage genomes that were sequenced and deposited in databases were mainly 'mesophilic' viruses of a restricted taxonomical spectrum of bacteria associated with human health or of industrial importance. The sequencing of a cultivated virus from polar sea-ice or from other extremely cold habitats, as well as the sequencing of a cultivated phage infecting Flavobacteria or other bacteria of the *Bacteroidetes* phylum, has not been reported yet. The following proteomics-supported analysis of the *Φ*11b genome aims at new insights into the genomics of a 'psychrophilic' phage including viral genome structure, properties of genes and phylogeny.

Materials and methods

DNA isolation and sequencing

Phage purification and DNA isolation were performed with DEAE-cellulose column chromatography and Phenol–Chloroform extraction as described previously (Borriss et al. 2003). Sequencing of the *Φ*11b genome was initiated by shot gun cloning and completed by primer-walking. For cloning, phage DNA was digested with endonuclease *Dra*I (Roche Diagnostics), ligated into pUC18 and electroporated into *E. coli* DH5α. Plasmid DNA was purified for sequencing by mini columns (High Pure Plasmid Isolation Kit, Roche). Plasmid inserts and virion DNA were sequenced applying Cycle Sequencing (Engelke et al.1988) and ABI PRISM Big Dye technology by Agowa GmbH, Berlin, FRG. Capillary electrophoresis with the polymer Pop 6 was carried out employing an ABI Prism 3700 Genetic Analyzer, sequencing data was obtained from pherogram analysis with ABI Prism Collection and ABI Prism Sequencing Analysis3.0 software (all materials PE Biosystems). M13 primers were used to sequence pUC18 inserted phage DNA. Specific primers were employed for walking on large plasmid inserts or directly on the phage DNA. Sequence assembly was carried out by Lasergene SeqMan software (DnaStar, Inc.).

Bioinformatic tools

Annotation was performed manually with the GenDB system (Meyer et al. 2003), which integrates Open Reading Frame (ORF) identification by Glimmer (Delcher et al. 1999) and Critica (Badger and Olsen 1999) and the standard tools for similarity (BLAST, McGinnis and Madden 2004), pattern and profile (InterPro) based search against the corresponding databases (last update September 2005). Selected online prediction tools used are given in S1 (Gasteiger et al. 2005; Krogh et al. 2001; Cserzo et al. 1997; Tusnady and Simon 2001; Hirokowa et al. 1998; Arai et al. 2004; Lupas et al. 1991; Berger et al. 1995; Wolf et al. 1997; Combet et al. 2000; Bendtsen et al. 2004; Juncker et al. 2003; Gruber et al. 2005; Jones 1999; Heger and Holm 2000; Andrade et al. 2000; Gardy et al. 2005). For illustrations BugView (Leader 2004) outputs were extracted and finished with Open Office Draw (http://www.openoffice.org/) or Incscape (http://www.inkscape.org/). Trees for glycosidase and ERF proteins were calculated by means of clustal and "muscle" parameters; respective amino acid (aa) sequences were extracted from the NCBI protein database (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Protein). Further calculations were performed with Phylip (3.6 package).

Separation and mass spectrometry (MS) of virion proteins

Purified virus particles were diluted in TE (10 mM Tris–HCl pH 7.4; 1 mM EDTA, pH 8) and denatured by adding 6 M Urea lysis buffer (stock: 2 ml 10% SDS, 3.6 g Urea, 699 µl mercaptoethanol, 3 ml glycerol). 0.1 µg protein was filled in a gel slot. Proteins were separated in a sodium dodecyl sulfate (SDS) 15% discontinuous polyacrylamide mini gel at 190 V for 45 min in a Miniprotean Cell (Bio-Rad). Gels were stained with either Coomassie Blue or silver nitrate according to Blum et al. (1987). For mass spectrometry analyses (MS), protein spots were excised from stained gels. In-gel trypsin digestion of the proteins was performed as described earlier (Büttner et al. 2001). Peptide masses were measured either in a Voyager-DE STR or in a Proteomics Analyzer 4700 (both PE Biosystems), mass fingerprints were analyzed using the gpmaw software. Gel free protein extracts were identified with NanoLC Tandem MS. After digestion the peptide extract was separated with high pressure liquid chromatography on an UltimateTM system (LC Packings). Peptides were first loaded and desalted on a reversed phase pre-column (*μ*-Precolumn, PepMapTM, C18, 300 µm i.d. × 5 mm, LC Packings) with a flow rate of 30 µl/min using 0.1% (v/v) formic acid as solvent. The separation was performed via a reversed phase nano column (PepMapTM, C18, 75 µm i.d. × 15 cm, LC Packings). As solvents 0.05% (v/v) formic acid (solvent A) and 90% (v/v) acetonitril in 0.05% (v/v) formic acid (solvent B) were used

with a linear gradient from 5 to 50% of solvent B over 70 min. The chromatography system was coupled via a nanoLC inlet (New Objective) to the Q-TOF mass spectrometer (Q-Star Pulsar, PE Biosystems) equipped with a nano-electrospray source (Protana). The eluted peptides were analyzed by tandem MS. The resulting MS/MS data were analyzed with the BioanalystTM Software (PE Biosystems) and the integrated Mascot (Matrix Science Ltd) script.

Nucleotide sequence accession number

The complete nucleotide sequence of the Φ11b virion DNA was deposited at EMBL under AJ 842011.

Results and discussion

DNA primary structure

After sequence assembly the chromosome of Φ11b turned out to be circular ds DNA, consisting of 36,012 base pairs. No physical ends of the supposed linear virion DNA molecule could be identified during sequencing and attempts to identify cohesive end (cos) sites by heat denaturation of restriction enzyme digested phage DNA failed (data not shown). A BlastN search of the whole sequence revealed similarity (expect (*E*) value 0.002 and 87% identities) of the region at 9,232–9,288 bp (PHG11b_19) to a 57 bp spanning coding sequence (X56064.1, 2571–2627) of ORF-5, a gene of the genomic packaging module of bacteriophage SPP1. No tRNA genes were found with tRNAscan (Lowe and Eddy 1997). Several potential transcription termination sites that fall in noncoding regions between potential genes were identified (S4). Statistical data is provided in S2.

The guanine-plus-cytosine (GC) content of the overall nucleotide sequence is 30.6 mol%. The GC percentage for genes ranges from 22 (PH11b_9, PHG11b_42) to 39 (PHG11b_32) and is on average 31.2. No genome segment of significantly higher or lower GC content indicating functional clusters and lateral gene transfer during evolution (Smith et al. 1999) could be identified for Φ11b (Fig. 1, Table 1). Its DNA appears to be in accordance with the assumed low GC content of host DNA since flavobacterial isolates are reported to have a GC content of 32–37 mol% (Bernardet et al. 1996). A correspondence of virus and host GC content is regarded to be the usual case in phage host systems and is

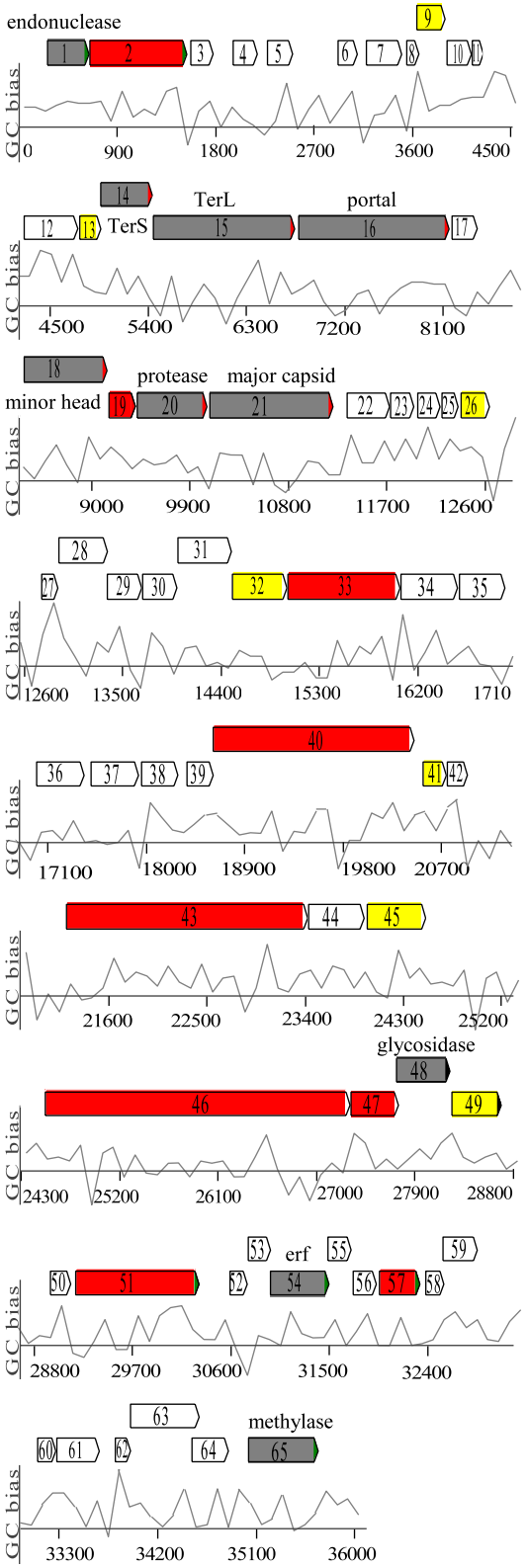

Fig. 1 Genome map of Φ11b shown in a randomly linearized form with ORFs underlined by the corresponding GC content diagram. The actual ends of the phage chromosome have not been identified. *ORF arrow color* indicates the respective annotation group (*grey* probable function resolved; *red* conserved hypothetical protein; *yellow* predicted membrane association). *Colored arrow tips* indicate the affiliation to the respective genomic functional cluster. A putative genomic organization of the phage might be as follows: PHG11b_50–_13, early genes (*green tips* with so far unresolved modules of recombination, modification, replication, transcriptional regulation); PHG11b_14–_21, DNA packaging and head morphogenesis (*red tips*); PHG11b_22–_47, head-tail joining and tail morphogenesis; PHG11b_48 and _49, destruction of the host cell envelope (*black tips*)

Table 1 Annotated ORFs of Φ11b: Bioinformatic prediction and similarity search results

ORF[a] (aa)	pI; mol. mass; %GC[a]	Annotation as[b]; sequence similarity (*) to ORFs (protein accession no.), *E* value; motif prediction profile (#)
1 (31)	8.9; 1494; 24	Endonuclease; *phage T4 MobB (NP_049674), −2e-28[G]; –
2 (297)	6.3; 33753; 30	CHP;**Bacteroides thetaiotaomicron* ORF(AAO79651.1),1e-33;–
3 (69)	6.2; 8271; 19	HP; –; –
4 (77)	9.2; 9003; 23	HP; *LysB of bacteriophage P2 (AAD03278.1), 0.62[G]
5 (83)	4.3; 9476; 31	HP; *phage HK022 putative endonuclease(NP_597900.1), 0.04[G]
6 (61)	4.9; 7083; 33	HP; –; cytoplasmic locus (Psort)
7 (109)	9.4; 12755; 27	HP; *RusA endonuclease (YP_240211.1) of phage EW, 0.35; –
8 (39)	10.0; 4536; 24	HP; –; –
9 (90)	9.4; 10852; 22	HP; –; #1 C terminal TM helix(ConPred-all), cytoplasmic(Psort)
10 (76)	9.6; 8333; 27	HP; *unknown protein of *Clostridium difficile* phage phiC2 (AAS02086.1), 0.16; #HTH (threshold 25%)
11 (38)	9.8; 4556; 27	HP; –; #coiled coil and cytoplasmic localization (Psort)
12 (165)	6.0; 19500; 29	HP; *unknown protein of phage sk1 (AAB70047.1), 0.2[G]; –
13 (67)	9.6; 7660; 29	HP; -; #2 (ConPred-all) TM helices
14 (161)	6.0; 18261; 36	Terminase small subunit (TerS); *phage HK620 TerS (NP_112075.1), 5e-18[G]; #1 (ConPred-elite-split) TM helix
15 (433)	5.8; 49237; 34	TerL; **H. influenzae* protein HI1410 (P44184), 1e-143[G]; –
16 (461)	5.47; 52062; 33	Portal protein, **Ehrlichia canis* phage-related protein (ZP_00211249.1), 5e-04; #signal peptide (SignalP3-NN)
17 (81)	6.3; 9312; 28	HP; –; –
18 (256)	8.8; 29028; 32	Minor head protein; *phage ϕCh1 gpC (NP_665925.1), 1e-12; –
19 (86)	5.1; 9994; 30	CHP; * SPP1 orf5 (NP_690660.1), 5e-16; #cytoplasmic (Psort)
20 (215)	5.0; 24898; 30	Prohead protease; * HK97 prohead proteinase (S54391), 5e-68[G]
21 (380)	5.0; 40958; 36	Major capsid protein precursor; *gp5 (P49861) of phage HK97, 1e-151[G]; #coiled coil (Paircoil)
22(133)	8.3; 15860; 23	HP; –; –
23(75)	5.8; 8627; 29	HP; *enterobacteria phage T7 RNA polymerase (NP_041960.1), 0.63; #cytoplasmic localization (Psort)
24 (74)	5.0; 8340; 31	HP; *PHG11b_23, 1e-07; #cytoplasmic localization (Psort)
25 (54)	5.2; 6467; 27	HP; –; –
26 (69)	6.2; 9648; 28	HP; *Sfi21 DNA package protein (NP_049972.1), 0.033[G]; #1 (ConPred-elite) TM helix, signal peptide (LipoP 1.0)
27(54)	6.2; 6009; 27	HP; –; #cytoplasmic localization (Psort)
28 (150)	4.7; 16958; 32	HP; *phage HK022(tail cluster) protein(NP_597887.1), 0.66[G]; –
29 (107)	6.2; 12308; 27	HP; –; –
30 (108)	5.0; 11990; 28	HP; **Ehrlichia ruminantium* protein (CAI28325.1), 0.16; –
31 (165)	4.6; 17465; 28	HP; –; –
32 (169)	4.6; 16593; 39	HP; **Gallus* (pore-, glyco-?) protein (XP_420179.1), 0.008; –
33 (343)	5.4; 35982; 35	CHP; **Rhodopseudomonas pallustris* (NP_947258.1), 1e-11;–
34 (174)	9.4; 19012; 34	HP; *phage TP901-1 head-tail protein(NP_112703.1), 0.005[G]; –
35 (143)	5. 5; 16191; 28	HP; –; –
36 (147)	4.4; 15541; 35	HP; –; –
37 (148)	7.9; 15516; 35	HP; *phage TP901-1 head-tail protein (NP_112702.1), 0.23[G]; –
38 (113)	5.7; 12712; 29	HP; –; #signalpeptide, cytoplasmic (membrane?) locus (Psort)
39 (75)	9.7; 9933; 33	HP; *phage P22 gp51 (NP_059610.1, Replication), 0.47[G]; –
40 (613)	9.2; 63111; 36	CHP; **Bacteroides fragilis* protein (CAH07988.1), 7e-05; #2 (ConPred-all) TM helices, coiled coils (Paircoil), OMP (Psort)
41 (74)	8.0; 8683; 26	HP; *bacteriophage sk1 unknown protein (AAB70080.1), 0.13[G]; #1 (DAS) TM helix or signal peptide (ConPred)
42 (66)	9.4; 7959; 22	HP; –; –
43 (740)	5.0; 82909; 29	CHP; **B. fragilis* hypothetical protein (YP_099731.1), 1e-05; #OMP (Psort) with signal peptide
44 (171)	4.4; 18667; 31	HP; *bacteriophage HK620 DNA transfer protein (NP_112087.1), 0.62 [G]; #predicted OMP (Psort)
45 (184)	5.1; 19371; 34	HP, *bacteriophage sk1 unknown protein (NP_044964.1), 0.52 [G]; #1 (ConPred-all) TM helix
46 (931)	5.1; 98599; 34	CHP; *gp134 (NP_818207.1) of mycophage Bxz1, 4e-19; #OMP or extracellular location (Psort)
47 (146)	6. 1; 17213; 30	CHP; **Clostridium thermocellum* conserved protein (ZP_00312461.1), 1e-04; #1 (SOSUI) TM helix or signal peptide (ConPred) and two coiled coils (Paircoil)
48 (167)	9.6; 18966; 31	Glycoside hydrolase, family 19; **Pseudomonas syringae* glycoside hydrolase 19 (YP_235844.1), 4e-29; –
49 (155)	9.7; 17680; 25	HP; **Saccharomyces cerevisiae* membrane glycoprotein (P53752), 0.079 [G]; #2 (ConPred-all) TM helices
50 (56)	4.7; 7916; 23	HP, –; #cytoplasmic localization (Psort)

Table 1 (contd.)

ORF[a] (aa)	pI; mol. mass; %GC[a]	Annotation as[b]; sequence similarity (*) to ORFs (protein accession no.), *E* value; motif prediction profile (#)
51 (377)	8.6; 42905; 30	CHP; * phage K1-5 protein 22 (AAR90064.1) of the "metabolism" gene cluster, 1e-24; #1(ConPred-all) TM helix
52 (57)	6.6; 6565; 31	HP, –; #coiled coil (Paircoil)
53 (71)	4.8; 8389; 28	HP; –; –
54 (180)	7.7; 19212; 33	erf; **Nostoc punctiforme* protein (ZP_00106022.1), 2e-17; –
55 (72)	7.9; 8712; 29	HP; *sfi21 ATPase similar protein (NP_050003.1), 0.8[G]; –
56 (73)	7.9; 8753; 23	HP; *phage T7 DNA ligase (NP_041963.1), 0.27[G]; –
57 (126)	4.8; 14508; 33	CHP; **Cytophaga hutchinsonii* DNA-binding protein (ZP_00308272.1), 2e-08; –
58 (59)	9.6; 6760; 32	HP; –; –
59 (108)	9.5; 12793; 29	HP; –; #cytoplasmic localization (Psort)
60 (59)	9.8; 6682; 31	HP; –; #cytoplasmic localization (Psort)
61 (132)	6.2; 15789; 28	HP; *phage HK620 DNA stabilization (NP_112082.1), 0.55[G]; –
62 (49)	6.7; 5692; 26	HP; –, #signal peptide (Conpred)
63 (213)	9.2; 25036; 31	HP; *phage P22 Tin protein (NP_046797.1, lysogenic conversion), 0.38 [G]; #cytoplasmic localization (Psort)
64 (116)	7. 7; 13454; 30	HP; *phage P22 gp52 (NP_059611.1), 0.33 [G]; #coiled coil (Paircoil), HTH(25%), cytoplasmic localization (Psort)
65 (217)	7.8; 25059; 32	Modification methylase; **Neisseria meningitidis* N-6 adenine-DNA methylase (AAF41140.1), 4e-50; –

[a]Cds (PHG11b_1-PHG11b_65) are specified by their locus tag number (ORF), protein length in base amino acids (aa), predicted isoelectric point and molecular mass (mol. mass) values in Da (pI/mass) as well as GC content in mol% (GC cont.)
[b]The column comprises annotation results by similarity search (underlined) and for some deduced proteins a predicted motif and subcellular localization profile including potential transmembrane (TM) helices, helix–turn–helices (HTH), outer membrane protein (OMP) prediction, with the respective tool used in parentheses. If not indicated otherwise ([G] indicates GenDB data), similarity data were obtained from BLASTP search.
HP hypothetical; *CHP* conserved hypothetical protein

explained by a coevolution of viral and host genes (Smith et al. 1999, Kwan et al. 2005). While no other complete low-GC siphophage genomes from the marine environment have been reported yet, Φ11b showed one of the lowest GC content values reported for sequenced siphoviral genomes.

ORF analysis and genome organization

Sixty-five Open Reading Frames (ORFs) were identified by GenDB that exclusively read from left to right (Table 1, Fig. 1). Φ11b sequences had strong similarity to coding sequences (cds) of phages and of eubacterial species of the phyla *Bacteroidetes*/*Chlorobi*, *Firmicutes*, as well as α-, β- and γ-proteobacteria and are likely to contain closely related prophage genes. Significant similarity to cds of closely host-related bacteria of the *Bacteroidetes* phylum was found for five ORFs (PHG11b_2, _40, _43, _57, _65) that might provide a link to the host specificity of Φ11b. Similarities to cds of eukaryotic viruses were found for PHG11b_32 (*E* value 0.16), _46 (*E* value 6e-04), _47 (*E* value 4e-04). For some ORFs of Φ11b a distinct motif prediction pattern was identified (Table 1, S3).

The putative genomic organization of the virus (Fig. 1) appears to follow that of other siphoviruses in which the ORFs are arranged in a manner reflecting the intracellular phage-development (Van Sinderen et al. 1996). With PHG11b_54 (erf), _65 (methylase), and _1 (endonuclease), the genome of Φ11b features three putative early genes that have significantly similar cds in databases and could serve as further potential markers for phage similarity analyses. Weak similarity hits (see Table 1), that might have a functional early gene analog in Φ11b, were observed for PHG11b_51 (DNA metabolism), _52 (ATPase), _56 (ligase), _57 (SSB protein), _61 (DNA stabilization or integrase), _63 (lysogenic conversion/growth inhibition), _64 (helicase). The endonuclease encoding ORF is located in front of a possible replication (primase) or recombination (transposase) protein gene (PHG11b_2) and 13 genes in front of the start of the head morphogenesis cluster. Upstream, the adjacent ORF encodes a putative DNA methyltransferase (MTase) with a proposed function in contributing to phage recombination or replication processes. Most prokaryotic MTases are components of a restriction-modification (RM) system in which they are associated with a restriction endonuclease (Bujnicki 2002). Thus, the PHG11b_1 encoded enzyme might be a component of an RM system or functions as a homing endonuclease, which seems to be common among phages (Crutz-LeCog et al. 2002). RM systems, one has been reported from the enterobacterial myophage P1 (Rao et al. 1989), can be associated with mobile genetic elements (Kobayashi et al. 2001). The other likely function of the MTase protein, however, is to provide protection of phage DNA against host restriction endonucleases by self-methylation, as it has been described from the temperate *Myxococcus xanthus* phage Mx8

(Magrini et al. 1997). In addition, a methylase gene adjacent to an endonuclease encoding ORF has also been reported for the lytic siphoviral *Lactobacillus casei* bacteriophage A2 (Proux et al. 2002).

*Φ*11b's head morphogenesis cluster appears to be the most conserved region of the phage genome. For six component late gene products with significant similarities to analogous proteins of other phages, a putative function could be assigned. The cluster which should be mainly responsible for packaging and phage head assembly consists of genes coding for the terminase subunits, portal, minor head, protease, and the major capsid protein after which transcription might be terminated by a predicted intergenic stem-loop (S4). A scaffolding protein, however, could not be identified for *Φ*11b. While for the arrangement of this segment of the genome (protease in front of one major capsid) *Φ*11b would be sorted into the cos-site sfi21-like phage genus, the presence of a putative minor head gene and an ORF involved in packaging similar to bacteriophage SPP1 suggests that a pac-site mechanism might be used, which is characteristic for phages of the sfi11-like genus (Desiere et al. 2002).

Both, the putative small and the large terminase subunit of *Φ*11b display at least significant similarity to the terminases of the related enterophages HK620 and sf6. For the highly similar (*E* value 4e-65) large subunit of SPP1 (G2P, NP_690654.1), an endonuclease, ATPase, and ds DNA binding activity was postulated (Gual et al. 2000). The protein is also reported to initiate the headful packaging series by transient interactions with two other gene products (Camacho et al. 2003). Comparisons of terminase aa sequences were found to give clues about the respective phage packaging mechanism (Casjens et al. 2005). Sequence similarities to terminases of pac-site phages (P22-like) suggest that *Φ*11b uses headful packaging and has circular permuted and terminal redundant virion DNA. Similar to some P22-like packaging phages the viral packaging initiation would not occur at a precise location but in a region of variable length in which the initiation cuts are distributed. This hypothesis is in agreement with the earlier finding that during electrophoresis digested ethidium bromide stained virion DNA samples separated with a smear, the intensity of which varied for the respective restriction enzyme used (Borriss et al. 2003) and which might contain the diagnostic submolar pac-site fragments.

Over its whole aa sequence and for a 57 bp spanning sequence PHG11b_19 displays strong similarity (*E* value 5e-16) to orf5 (CAA66579.1) of *Bacillus* phage SPP1, one of five ORFs that specify gene products required for pac-site cleavage (Chai et al. 1992). Apart from the observed significant similarity of PHG11b_20 and_21 to the protease or major capsid of phage HK97, the gene arrangement of the putative morphogenesis segment, too, would suggest an HK97-like head assembly mechanism for *Φ*11b. In HK97 the endoprotease is involved in head morphogenesis by catalyzing the cleavage of the major coat precursor; the major capsid protein guides a scaffold-lacking assembly from penta- and hexamers and the maturation of the icosahedral head (Duda et al. 1995; Wikoff et al. 2000).

In siphophages, packaging and head genes are usually followed by a head-to-tail and a tail protein encoding sector. For *Φ*11b, no single tail gene with significant similarity to analogous cds of phage origin could be identified. However, several phage tail proteins of weak similarity were identified for the following ORFs: PHG11b_28, _32, _34, _35, _37, _40, and _47. The large sized ORFs PHG11b_40-_46 may be structural tail genes. Transmembrane (TM) helices are reported to be typical for the tape measure protein (tmp) (Zimmer et al. 2003) and, indeed, a putative membrane association was predicted for PHG11b_40, an ORF with weak similarity to a tmp (CAD43903.1) of bacteriophage A2. Tmps are also suggested to be encoded by the first large ORF downstream of the major capsid encoding gene. With 613 aa corresponding to 75–90 nm of measured tail length (Borriss et al. 2003) the size of the potential tmp (PHG11b_40), however, does not appear to be well in line with the length of the tail as measured by electron microscopy. In comparison a tmp size of 853 aa corresponds to 150 nm tail length in phage lambda (Zimmer et al. 2003). PHG11b_46 was highly similar to the cds of gp134 of mycophage Bxz1, it seems that it specifies one of the latest, if not the last gene of the genomic tail morphogenesis segment of *Φ*11b. With predicted 98 kDa the largest putative protein of the phage might be a tail fiber protein as several viral tail or tail fiber BLASTP hits of weak similarity suggest.

At the very right end of the putative late-gene segment the putative endolysin gene of *Φ*11b PHG11b_48 might form a lysis cassette with an adjacent ORF with TM domain prediction, e.g. PHG11b_49, encoding the putative holin. As an unusual arrangement for a siphoviral genome, the holin would be preceded by the endolysin in the putative module; a characteristic observed for bacteriophage SPP1, too. Also PHG11b_47, a conserved hypothetical protein with similarity to a Canarypox virus protein (NP_955191.1, *E* value 6e-04) and three low similarity hits for phage tail proteins, was predicted to contain one TM domain. Notably, the endolysin gene of enterophage T1 is flanked by a gene for a deduced 1- and a 2-TM protein (Roberts et al. 2004).

Attacking glycosidic linkages of host cell murein, glycosidases were reported to be one of four muralytic classes of phage endolysins (Young et al. 2000). Focusing on the sequenced genomes of bacteria taxonomically related to the *Φ*11b host, a glycosidase of family 19 has neither been reported from the three completed *Bacteroides* genomes nor for those of *Chlorobium tepidum* and *P. gingivalis*. For viruses, the enzyme family is reported from sipho-, podo- and myophage representatives (12 Interpro entries) and seems to be conserved in mycobacteriophages. Both, phage and bacterial sequences of the glycosidase 19 are suggested to be either endolysins/lysozymes or chitinases. The deduced protein of

PHG11b_49, one of 11 potential membrane interacting proteins of Φ11b, was predicted to have two TM domains, which is typical for class-2 holins. However, its size of 155 residues is reported to be covered by holins of class-1 (Young et al. 2000). Moreover, no holin/antiholin-indicating dual start motif could be found for PHG11b_49 and it is possible that also one of the other ORFs with more than one predicted TM domain, i.e.

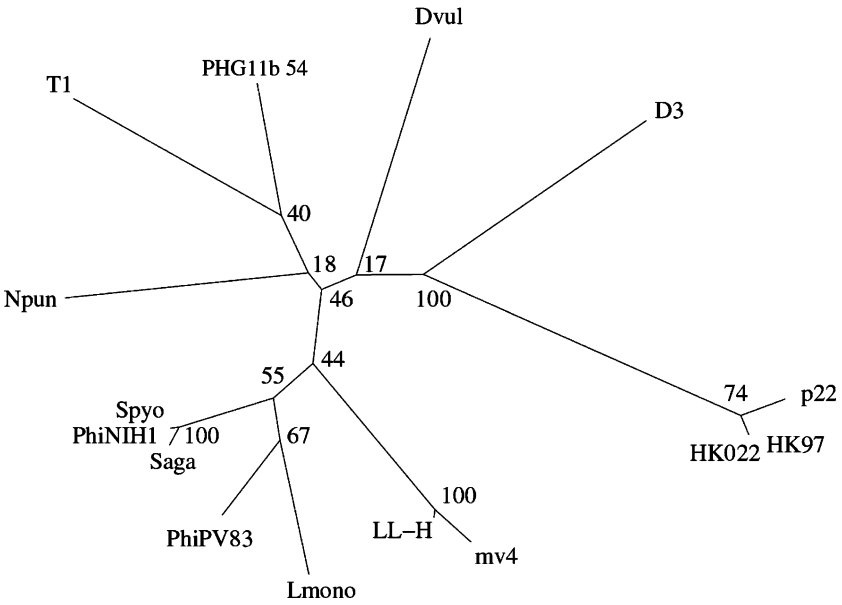

Fig. 2 Phylogenetic tree for the ERF proteins analogous to PHG11b_54. Numbers are bootstraps. The sequences are denoted by host species abbreviations, gene identification numbers are: *Desulvovibrio vulgaris* (*Dvu1*), gi|4658043; enterobacteria phage P22 (*p22*), gi|9911064; enterobacteria phage HK022 (*HK022*), gi|9634188; enterobacteria phage HK97 (*HK97*), gi|9635505; enterobacteria phage T1 (*T1*), gi|4568633; *Lactobacillus* phage LL-H (*LL-H*), gi|1865420; *Lactobacillus* phage mv4 (*mv4*), gi|1113833; *Listeria monocytogenes* (*Lmono*), gi|1680435; *Nostoc punctiforme* (*Npun*), gi|2312400; *Pseudomonas* phage D3 (*D3*), gi|9635643; *S. aureus* phage ϕPV83 (*PhiPV83*), gi|9635694; *Streptococcus agalactiae* (*Saga*), gi|2253674; *Streptococcus* phage NIH1/100 (*PhiNIH1*), gi|1627178; *Streptococcus pyogenes* (*Spyo*), gi|1974668

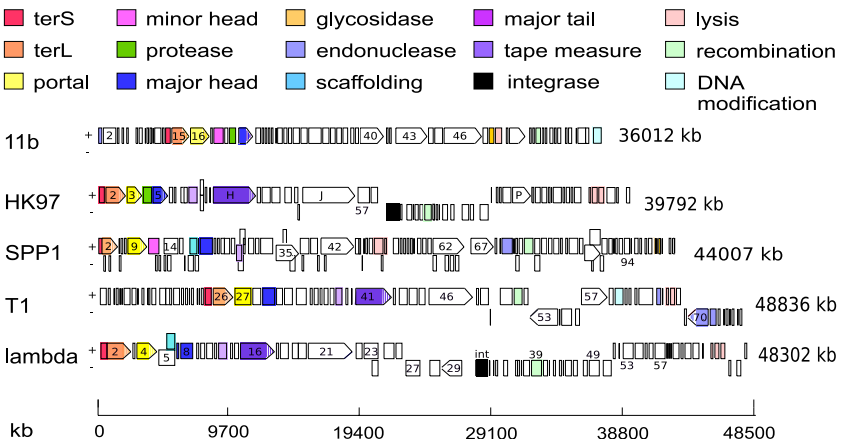

Fig. 3 Alignment of the Φ11b (11b) ORF map with the genomes of bacteriophage HK97, SPP1, lambda and enterobacteria phage T1. Functionally similar ORFs share the same color. Phage genome length is given at the right end of the gene map. All five phages share conserved head package/morphogenesis genes. Notably, the Φ11b gene map features a minor head and protease encoding ORF. The potential lysis module of Φ11b contains a glycosidase, whereas the non-essential glycosidases of SPP1 are situated aloof from the lysis cassette. The glycosidases of SPP1 apparently do not share synteny in function, position, number, gene size and similarity with the putative enzyme of Φ11b and are not required for phage growth. Any significant sequence similarity among the endonucleases (one in SPP1 and three in T1) of the three phages could not be found. PHG11b_54 was annotated as putative essential recombination factor (*Erf*). HK97 and lambda have their corresponding recombination protein encoding genes located close to the lysogeny module indicating an integrase gene. However, no ORF which clearly codes for a potential integrase was identified in Φ11b. Similar to Φ11b, in SPP1 and T1, the recombination protein genes are located shortly (about five ORFs) downstream of the "long orf" tail cluster. DNA modification methylases are found in Φ11b and T1. Their position in both phages appears to be similar: the modification enzyme encoding genes are located a few ORFs downstream of the *erf* gene

PHG11b_32, which also exhibits similarity to membrane glycoproteins, could function as a holin. In addition, PHG11b_49 was cloned in an *E. coli* T7-overexpression vector, where it turned out to be lethal after transformation (data not shown).

Phylogenetic and synteny analysis

Three potential proteins of *Φ*11b show strong similarity to aa sequences of bacteriophage SPP1 and HK97. According to a phylogenetic whole genome tree we calculated from a global BLASTP analysis (Clarke et al. 2002; Teeling et al. 2004) of selected siphoviruses including HK97 (data not shown), *Φ*11b is currently most closely related to the unclassified pac-site siphovirus SPP1 (Alonso et al. 1997). While the two phages differ clearly in their host specificities (Bacteroidetes vs. Firmicutes) the fact that numerous species of both groups of bacteria are known to occur in terrestrial habitats would give a basis for an interviral gene transfer during evolution.

We performed a phylogenetic analysis for two deduced proteins of *Φ*11b. The glycosidase (PHG11b_48) and erf (PHG11b_54) enzymes were among the few putative proteins of *Φ*11b suitable for phylogenetic tree calculation, because they showed extended conservation in their aa sequence to cds from independent species. In contrast to PHG11b_54 that branched monophyletically with the T1 sequence (Fig. 2), PHG11b_48 grouped solitarily in the glycosidase tree (data not shown) and seemed most closely related to Pseudomonade glycosidases and a cds of the beta-proteobacterium *Nitrosomonas europaea*. Notably, a Pfam phylogeny (http://www.sanger.ac.uk/cgi-bin/Pfam/getacc?PF00182) positions PHG11b_48 monophyletically with those of mycobacteriophage Rosebush and *N. europea*. Both the glycosidase and the erf tree appear to confirm BLASTP similarity data which suggest that—like in other phages—the genome of *Φ*11b is a mosaic of single evolutionarily acquired sequences, a string of linking hot spots for an interviral kinship which is even mirrored throughout fairly different environments.

In an attempt to illustrate the phylogenetic results, four potentially closely related phage genomes were aligned and analyzed for synteny (Fig. 3). Similarity search and tree calculation indicated that selected phage genomes share global (phage SPP1) or similarity for single genes (erf of bacteriophage T1, major capsid of phage HK97) with *Φ*11b. Bacteriophage *λ* was included as the prototype of a lambdoid phage.

*Φ*11b proteome

With silver staining, at least 11 prominent single bands were observed (Fig. 4). The largest protein band (> 200 kDa) might be caused by complex virion protein species, e.g. multimers of the major capsid that did not enter the gel (Van Sinderen et al. 1996; Wikoff et al. 2000). The gel separation band pattern of lactococcal phage r1t showed a profile partially similar to *Φ*11b, were also at least 11 protein bands could be identified (Van Sinderen et al. 1996). The published protein profile resembles *Φ*11b gel separations for six bands (190/190, 160/160, 50/50, 43/40, 16/16, 14/14 kDa) and for the total number of obtained bands (12 vs. 11 bands).

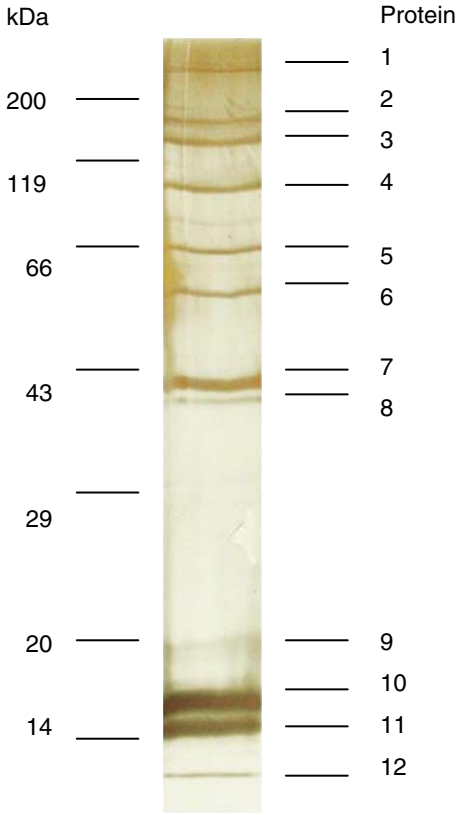

Fig. 4 Silver stained protein band pattern of denatured *Φ*11b virions, separated by SDS Polyacrylamide Gel Electrophoresis (*PAGE*). The observed approximate molecular mass (in kDa) of 12 visible bands is: > 200 (*1*), 190 (*2*), 160 (*3*), 90 (*4*), 70 (*5*), 50 (*6*), 40 (*7*), 37 (*8*), 20 (*9*), 16 (*10*), 14 (*11*), 0–5 (*12*, bottom band or electrophoresis front). The results of the MALDI-ToF analysis of the *Φ*11b protein bands are summarized in Table 2

At least five different phage proteins could be identified through MS (Table 2) during two independent experiments. Only one short peptide with a relatively low score (data not shown) could be identified from the PHG11b_15 encoded TerL subunit. The corresponding protein has a predicted mass of 49.2 kDa. However, it was identified from a migration band of about 15 kDa. Thus, one might speculate that this protein was cleaved or degraded into a smaller molecule, migrating in a band of about 15 kDa. The terminase is an enzyme involved in the virion packaging process, it might still be detectable in the bacteriophage virion. Twelve different peptides (data not shown) of the potential tail gene PHG11b_46 were identified from the visible band separated at 90 kDa (Fig. 4, Table 2). For phage PSA, Zimmer et al. (2003) reported the identification of an

Table 2 Proteome analysis of Φ11b: mass spectrometry (MS) identification data

No.[a]	Protein band (in kDa)	Corresponding ORF	Putative function obtained by similarity search	Calculated mass (Da)
2	190	PHG11b_21	Major capsid	40,958
3	160	PHG11b_21	Major capsid	40,958
4	90	PHG11b_46	Conserved hypothetical protein	98,599
6	55	PHG11b_40[b]	Conserved hypothetical protein	63,111
7	40	PHG11b_16	Portal protein	52,063
9-11	14–20	PHG11b_18	Minor head protein	29,028
9–11	14–20	PHG11b_34[b]	Hypothetical protein	19,012
9–11	14–20	PHG11b_37	Hypothetical protein	15,516
9–11	14–20	PHG11b_45	Hypothetical protein	19,371
10/11	14/16	PHG11b_36	Hypothetical protein	15,541
10/11	14/16	PHG11b_15[b]	TerL	49,237

[a]Protein band corresponding to the numbers shown in Fig. 4
[b]Low score in MS analysis

apparent tail-and-base-plate protein from peptides separated at ~68 kDa as well as a probable tmp at 80 kDa. The identification of sequences of PHG11b_21 (major capsid) in migration bands of 160 and 190 kDa might be interpreted as the result of two different multiprotein aggregation states (Van Sinderen et al.1996; Wikoff et al. 2000). MS analysis of virion proteins of bacteriophage r1t revealed four suggested structural proteins encoded by ORFs of a probable head tail morphogenesis gene cluster including the major capsid protein (Van Sinderen et al. 1996). Also, major capsid peptides were identified for this phage in the 160 and 190 kDa band; this separation characteristic was suggested to be caused by the formation of tetra- and pentameric capsid protein species (Van Sinderen et al. 1996). The most abundant proteins of Φ11b might be contained in the prominent bands that separated during PAGE at 14 and 16 kDa and at 40 kDa; accordingly. The major capsid protein, which is found in two migration bands, also occurs in abundance.

Bacteriophage 11b delivers the first described genome of a cultivated virus infecting a psychrophilic host and a *Bacteroidetes* bacterium. Bioinformatic analysis of the nucleotide sequence and deduced genes as well as a proteomic approach indicated an affiliation to 'mesophilic' siphoviruses of nonmarine habitats. It is tempting to speculate that the strikingly low general GC content and significant similarities for some conserved ORFs of the phage provide a link to its cold-adapted flavobacterial host.

Acknowledgments We thank Hansjörg Lehnherr for critical reading of the manuscript and Dr. Steffen Krüger for helpful comments during the sequencing and sequence assembly procedure. We are grateful for the support of Michael Hecker in the analysis of the phage proteome.

References

Ackermann H-W, Abedon ST (2001) Bacteriophage names. The Bacteriophage Ecology Group. (Online.) http://www.phage.org/names.htm

Alonso JC, Luder G, Stiege AC, Chai S, Weise F, Trautner TA (1997) The complete nucleotide sequence and functional organization of *Bacillus subtilis* bacteriophage SPP1. Gene 204:201–212

Andrade MA, Ponting CP, Gibson TJ, Bork P (2000) Homology-based method for identification of protein repeats using statistical significance estimates. J Mol Biol 298:521–537

Arai M, Mitsuke H, Ikeda M, Xia J-X, Kikuchi T, Satake M, Shimizu T (2004) ConPred II: a consensus prediction method for obtaining transmembrane topology models with high reliability. Nucleic Acids Res 32:390–393

Badger JH, Olsen GJ (1999) CRITICA: coding region identification tool invoking comparative analysis. Mol Biol Evol 16:512–524

Bendtsen JD, Nielsen H, von Heijne G, Brunak S (2004) Improved prediction of signal peptides: SignalP 3.0. J Mol Biol 340:783–795

Berger B, Wilson DB, Wolf E, Tonchev T, Milla M, Kim PS (1995) Predicting coiled coils by use of pairwise residue correlations. Proc Natl Acad Sci USA 92:8259–8263

Bernardet J-F, Segers P, Vancanneyt M, Berthe F, Kersters K, Vandamme P (1996) Cutting a Gordian knot: emended classification and description of the genus *Flavobacterium*, emended description of the family *Flavobacteriaceae*, and proposal of *Flavobacterium hydatis* nom. nov. (basonym, *Cytophaga aquatilis* Strohl and Tait 1978). Int J Syst Bacteriol 46:128–148

Blum H, Beier H, Gross HJ (1987) Improved silver staining of plant proteins, RNA and DNA in polyacrylamide gels. Electrophoresis 8:93–99

Borriss M, Helmke E, Hanschke R, Schweder T (2003) Isolation and characterisation of marine psychrophillic phage-host systems from Arctic sea ice. Extremophiles 7:377–384

Bujnicki JM (2002) Sequence permutations in the molecular evolution of DNA methyltransferases. BMC Evol Biol 2:3

Büttner K, Bernhardt J, Scharf C, Schmid R, Mäder U, Eymann C, Antelmann H, Völker A, Völker U, Hecker M (2001) A comprehensive two-dimensional map of cytosolic proteins of *Bacillus subtilis*. Electrophoresis 22:2908–2935

Bradford D, Hugenholtz P, Seviour EM, Cunningham MA, Stratton H, Seviour RJ, Blackall LL (1996) 16S rRNA analysis of isolates obtained from gram-negative filamentous bacteria micromanipulated from activated sludge. Syst Appl Microbiol 19:334–343

Brinkmeyer R, Knittel K, Jurgens J, Weyland H, Amann R, Helmke E (2003) Diversity and structure of bacterial communities in Arctic versus Antarctic pack ice. Appl Environ Microbiol 69: 6610–6599

Camacho AG, A. Gual A, Lurz R, Tavares P, Alonso JC (2003) *Bacillus subtilis* bacteriophage SPP1 DNA packaging motor requires terminase and portal proteins. J Biol Chem 278:23251–23239

Casjens SR, Gilcrease EB, Winn-Stapley DA, Schicklmaier P, Schmieger H, Pedulla ML, Ford ME, Houtz JM, Hatfull GF, Hendrix RW (2005) The generalized transducing *Salmonella* bacteriophage ES18: complete genome sequence and DNA packaging strategy. J Bacteriol 187:1091–1104

Chai S, Bravo A, Luder G, Nedlin A, Trautner TA, Alonso JC (1992) Molecular analysis of the *Bacillus subtilis* bacteriophage SPP1 region encompassing genes 1 to 6. The products of gene 1 and gene 2 are required for pac cleavage. J Mol Biol 224:87–102

Clarke GD, Beiko RG, Ragan MA, Charlebois RL (2002) Inferring genome trees by using a filter to eliminate phylogenetically discordant sequences and a distance matrix based on mean normalized BLASTP scores. J Bacteriol 184:2072–2080

Combet C, Blanchet C, Geourjon C, Deléage G (2000) NPS@: Network Protein Sequence Analysis TIBS 25:147–150

Crutz-Le Coq AM, Cesselin B, Commissaire J, Anba J (2002) Sequence analysis of the lactococcal bacteriophage bIL170: insights into structural proteins and HNH endonucleases in dairy phages. Microbiology 148:985–1001

Cserzo M, Wallin E, Simon I, von Heijne G, Elofsson A (1997) Prediction of transmembrane alpha-helices in prokaryotic membrane proteins: the Dense Alignment Surface method. Prot Eng 10:673–676

Delcher AL, Harmon D, Kasif S, White O, Salzberg SL (1999) Improved microbial gene identification with GLIMMER. Nucleic Acids Res 27:4636–4641

Desiere F, Lucchini S, Canchaya C, Ventura M, Brussow H (2002) Comparative genomics of phages and prophages in lactic acid bacteria. Antonie Van Leeuwenhoek 82:73–91

Duda RL, Martincic K, Hendrix RW (1995) Genetic basis of bacteriophage HK97 prohead assembly. J Mol Biol 247:636–647

Engelke DR, Hoener PA, Collins FS (1988) Direct sequencing of enzymatically amplified human genomic DNA. Proc Natl Acad Sci USA 85:544–548

Gardy JL, Laird MR, Chen F, Rey S, Walsh CJ, Ester M, Brinkman FSL (2005) PSORTb v.2.0: expanded prediction of bacterial protein subcellular localization and insights gained from comparative proteome analysis. Bioinformatics 21:617–623

Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A (2005) Protein identification and analysis tools on the ExPASy server. In Walker JM (eds) The proteomics protocols handbook. Humana Press, pp 571–607

Gruber M, Soding J, Lupas AN (2005) REPPER–repeats and their periodicities in fibrous proteins. Nucleic Acids Res 33:239–243

Gual A, Camacho AG Alonso JC (2000) Functional analysis of the terminase large subunit, G2P, of *Bacillus subtilis* bacteriophage SPP1. J Biol Chem 275:35311–35299

Heger A, Holm L (2000) Rapid automatic detection and alignment of repeats in protein sequences. Proteins 41:224–237

Hirokawa T, Seah BC, Mitaku S (1998) SOSUI: classification and secondary structure prediction system for membrane proteins. Bioinformatics 14:378–379

Jiang SC, Kellogg CA, Paul JH (1998) Characterization of marine temperate phage-host systems isolated from Mamala Bay, Oahu, Hawaii. Appl Environ Microbiol 64:535–542

Jones DT (1999) Protein secondary structure prediction based on position-specific scoring matrices. J Mol Biol 292:195–202

Juncker AS, Willenbrock H, von Heijne G, Nielsen H, Brunak S, Krogh A (2003) Prediction of lipoprotein signal peptides in Gram-negative bacteria. Protein Sci 12:1652–1662

Kirchman DL (2002) The ecology of cytophaga–flavobacteria in aquatic environments. FEMS Microbiol Ecol 39:91–100

Kisand V, Cuadros R, Wikner J (2002) Phylogeny of culturable estuarine bacteria catabolizing riverine organic matter in the northern Baltic Sea. Appl Environ Microbiol 68:379–388

Kobayashi I (2001) Behavior of restriction-modification systems as selfish mobile elements and their impact on genome evolution. Nucleic Acids Res 29:3742–3756

Krogh A, Larsson B, von Heijne G, Sonnhammer ELL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 305:567–580

Kwan T, Liu J, Dubow M, Gros P Pelletier J (2005) The complete genomes and proteomes of 27 *Staphylococcus aureus* bacteriophages. Proc Natl Acad Sci USA 102:5174–5179

Leader DP (2004) BugView: a browser for comparing genomes. Bioinformatics 20:129–130

Lowe TM, Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 25:955–964

Lupas A, Van Dyke M, Stock J (1991) Predicting coiled coils from protein sequences. Science 252:1162–1164

Magrini V, Salmi D, Thomas D, Herbert S, Hartzell P, Youderian P (1997) Temperate *Myxococcus xanthus* phage Mx8 encodes a DNA adenine methylase, Mox. J Bacteriol 179:4254–4263

McGinnis S, Madden TL (2004) BLAST: at the core of a powerful and diverse set of sequence analysis tools. Nucleic Acids Res 32:20–25

Meyer F, Goesmann A, McHardy AC, Bartels D, Bekel T, Clausen J, Kalinowski J, Linke B, Rupp O, Giegerich R, Puhler A (2003) GenDB-an open source genome annotation system for prokaryote genomes. Nucleic Acids Res 31:2187–2195

Proux C, van Sinderen D, Suarez J, Garcia P, Ladero V, Fitzgerald GF, Desiere F, Brussow H (2002) The dilemma of phage taxonomy illustrated by comparative genomics of Sfi21-like Siphoviridae in lactic acid bacteria. J Bacteriol 184:6026–6036

Rao DN, Eberle H, Bickle TA (1989) Characterization and mutations of the bacteriophage P1 mod gene encoding the recognition subunit of the EcoP1 restriction and modification system. J Bacteriol 171:2347–2352

Roberts MD, Martin NL, Kropinski AM (2004) The genome and proteome of coliphage T1. Virology 318:245–266

Smith MC, Burns RN, Wilson SE, Gregory MA (1999) The complete genome sequence of the *Streptomyces* temperate phage phiC31: evolutionary relationships to other viruses. Nucleic Acids Res 27:2145–2155

Teeling H., Lombardot T, Bauer M, Ludwig W, Glöckner FO (2004) Evaluation of the phylogenetic position of the planctomycete '*Rhodopirellula baltica*' SH 1 by means of concatenated ribosomal protein sequences, DNA-directed RNA polymerase subunit sequences and whole genome trees. Int J Syst Evol Microbiol 54:791–801

Tusnady GE, Simon I (2001) The HMMTOP transmembrane topology prediction server. Bioinformatics. 17:849–850

Van Sinderen D, Karsens H, Kok J, Terpstra P, Ruiters MH, Venema G, Nauta A (1996) Sequence analysis and molecular characterization of the temperate lactococcal bacteriophage r1t. Mol Microbiol 19:1343–1355

Wikoff WR, Liljas L, Duda RL, Tsuruta H, Hendrix RW, Johnson JE (2000) Topologically linked protein rings in the bacteriophage HK97 capsid. Science 289:2129–2133

Wolf E, Kim PS, Berger B (1997) MultiCoil: a program for predicting two- and three-stranded coiled coils. Protein Sci 6:1179–1189

Young R, Wang I-N, Roof WD (2000) Phages will out: strategies of host cell lysis. Trends Microbiol 8:120–128

Zimmer M, Sattelberger E, Inman RB, Calendar R, Loessner MJ (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