Genomics 97 (2011) 249–255

Contents lists available at ScienceDirect

Genomics

journal homepage: www.elsevier.com/locate/ygeno

Genomic analysis of bacteriophage ϕAB1, a ϕKMV-like virus infecting multidrug-resistant *Acinetobacter baumannii*

Kai-Chih Chang [a,1], Nien-Tsung Lin [b,1], Anren Hu [a], Yu-Shan Lin [a], Li-Kuang Chen [c,d], Meng-Jiun Lai [a,*]

[a] Department of Laboratory Medicine and Biotechnology, Tzu Chi University, Hualien 970, Taiwan
[b] Institute of Microbiology, Immunology and Molecular Medicine, Tzu Chi University, Hualien 970, Taiwan
[c] College of Medicine, Tzu Chi University, Hualien 970, Taiwan
[d] Department of Emergency Medicine, Buddhist Tzu Chi General Hospital, Hualien 970, Taiwan

ARTICLE INFO

Article history:
Received 14 September 2010
Accepted 4 January 2011
Available online 15 January 2011

Keywords:
Bacteriophage
phiKMV-like phages
Multidrug-resistant *Acinetobacter baumannii*
Genome sequence
HNH endonuclease
Genes-in-pieces
Holin
Endolysin

ABSTRACT

We present the complete genomic sequence of a lytic bacteriophage ϕAB1 which can infect many clinical isolates of multidrug-resistant *Acinetobacter baumannii*. The recently isolated bacteriophage displays morphology resembling *Podoviridae* family. The ϕAB1 genome is a linear double-stranded DNA of 41,526 bp containing 46 possible open reading frames (ORFs). The majority of the predicted structural proteins were identified as part of the phage particle by mass spectrometry analysis. According to the virion morphology, overall genomic structure, and the phylogenetic tree of RNA polymerase, we propose that ϕAB1 is a new member of the ϕKMV-like phages. Additionally, we identified four ORFs encoding putative HNH endonucleases, one of which is presumed to integrate and create a genes-in-pieces DNA polymerase. Also, a potential lysis cassette was identified in the late genome. The lytic power of this bacteriophage combined with its specificity for *A. baumannii* makes ϕAB1 an attractive agent for therapeutic or disinfection applications.

1. Introduction

Acinetobacter baumannii is an aerobic Gram-negative bacterium commonly found in soil and water. The prevalence of *A. baumannii* as a significant nosocomial pathogen is rising worldwide due to its ability to survive on various surfaces in the hospital environment for prolonged periods [1]. In addition to the traditional threat to patients in intensive care units, *A. baumannii* infection is also becoming a growing threat to healthy communities [2,3]. In recent years, *A. baumannii* has evolved increasing resistance to nearly all conventional antimicrobial drugs [4]. Many hospitals worldwide have reported outbreaks caused by multidrug-resistant strains of *A. baumannii* (MDRAB) [5,6]. The use of antimicrobials to effectively treat MDRAB infection remains possible such as using the new drug tigecycline, the old antibiotic colistin, or combination therapy with different agents [7]. However, the treatment is regarded as a difficult task for clinicians since the therapeutic options are highly dependent on clinical experience and the susceptibility patterns of the multidrug-resistant isolates [8]. Emerging resistance has been recently reported even against the most currently promising antimicrobial options [9]. These developments highlight the need to find alternative tools to combat *A. baumannii* infection for either prophylactic or therapeutic purposes. Phage therapy may be a useful strategy in fighting this bacterial infection.

Bacteriophages (phages) are viruses that specifically infect bacteria and are estimated to be the most widely distributed and diverse entities in the biosphere. The genes and activities of phages are suggested to be a driving force in maintaining genetic diversity amongst the bacterial community [10].

So far, more than 500 complete phage genomes have been deposited in GenBank [11], only one of which specifically infects *A. baumannii* (phage AB1, HM36820). The AB1 genome is a 45,159 bp double-stranded DNA (dsDNA) molecule whose result of the detailed genome analysis has not yet been published. The lytic phage AB1, capable of infecting the highly drug-resistant clinical *A. baumannii* strain KD311, was recently characterized. Phage morphology indicated that this phage might belong to the *Siphoviridae* family [12].

In this study, we present the analysis of the complete genomic sequence of the lytic phage ϕAB1, one of the ten newly isolated phages infecting MDRAB [13]. In contrast to phage AB1 particles, electron micrographs of ϕAB1 phage particles showed that this phage possesses an isometric head and a short tail, matching the typical morphological features of the *Podoviridae* virus family.

Based on morphology and our genomic analysis, ϕAB1 is suggested to be a new member of the ϕKMV-like group. The ϕKMV-like virus

* Corresponding author at: Department of Laboratory Medicine and Biotechnology, Tzu Chi University, 701, Section 3, Zhong-Yang Road, Hualien 97004, Taiwan. Fax: +886 3 8571917.
E-mail address: monjou@mail.tcu.edu.tw (M.-J. Lai).
[1] Equal contribution.

0888-7543/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.ygeno.2011.01.002

group was initially proposed following the analysis of three phages: ϕKMV, LKD16, and LKA1 [14,15]. These three phages may not all show significant DNA similarity, but they all possess typical *Podoviridae* morphology and common genome structures, generally corresponding to those of the T7-like phages. The key feature that distinguishes ϕKMV-like phages is that their genes encoding RNA polymerases (RNAPs) are located adjacent to the structural protein region, instead of in the early region as in other members of the T7-like phage group. At the time the analysis was done, there were ten fully-sequenced phages that have been previously designated as the ϕKMV-like viruses [16]. Knowledge of these ϕKMV-like phages provides information to initiate a comparative study of the biodiversity and evolution of the ϕAB1 genome.

2. Result and discussion

2.1. Overview of the ϕAB1 genome

The ϕAB1 genome sequence was determined by a combination of shotgun sequencing and primer walking. The linear 41,526 bp dsDNA genome has a GC content of 39.08%, which is similar to the 38% of the *A. baumannii* ATCC 17978 strain. The direct terminal repeats are 410 bp long. Through the combined use of gene prediction, BLAST, protein signature search, and whole-phage shotgun analyses, we have assigned 46 open reading frames (ORFs) (see Fig. 1 and Supplementary Table 1), which are all transcribed on the same strand of the DNA molecule. All predicted ORFs have an AUG as the initiation codon and are preceded by recognizable Shine–Dalgarno sequences. The genome is densely packed; 91.0% of the genome is occupied by coding sequences.

Only two of the 46 ORFs found statistically significant ($E<0.001$) matches when queried using nucleotide–nucleotide BLAST (BLASTn) searches in the *nr* (non-redundant) database. A downstream fragment within ORF41 matched to the genomic region of *A. baumannii* which was annotated as a gene encoding a putative tail fiber protein. This indicates the possibility of a phage remnant or prophage DNA existing in the bacterial genome. The BLASTn search for ORF45 showed three discontinuous regions with E values under 10^{-6}, matching to the ORF38 of Vibrio phage Vp93, a member of the ϕKMV-like group. A protein BLAST (BLASTp) search of ORF45 suggests that its product is DNA maturase B.

In a similarity search of potential ϕAB1 gene products (gp), 22 gp showed protein similarity to predicted products from the members of the ϕKMV-like group (17 gp to LKA1 as shown in Fig. 1). The majority of the ORFs encoding these 22 gp were also present in the same genomic gene order as observed in the genomes of the three phages ϕKMV, LKD16, and LKA1 [14,15], with the exceptions of ORF17 and ORF24. These two ORFs were predicted to encode HNH endonucleases and matched the product of ORF13 in Vibrio phage VP93. Another eight predicted products matched proteins from various microorganisms, and two (gp26 and gp35) of which have unknown functions. Sixteen ORFs were predicted to have products unique to ϕAB1. Among the first 11 gp of the ϕAB1 ORFs, BLASTp revealed similar proteins in the databases to only gp1 and gp10. Gp1 matched hypothetical proteins found in many phages, most of which belong to the ϕKMV-like group (gp2 PT2, gp3 PT5, gp2 phiKF77, gp3 LKD16, gp3 ϕKMV, and gp4 LKA1). This is consistent with the previous observation that these homologous proteins are less divergent relative to other products encoded within the early regions of the ϕKMV-like genomes [14].

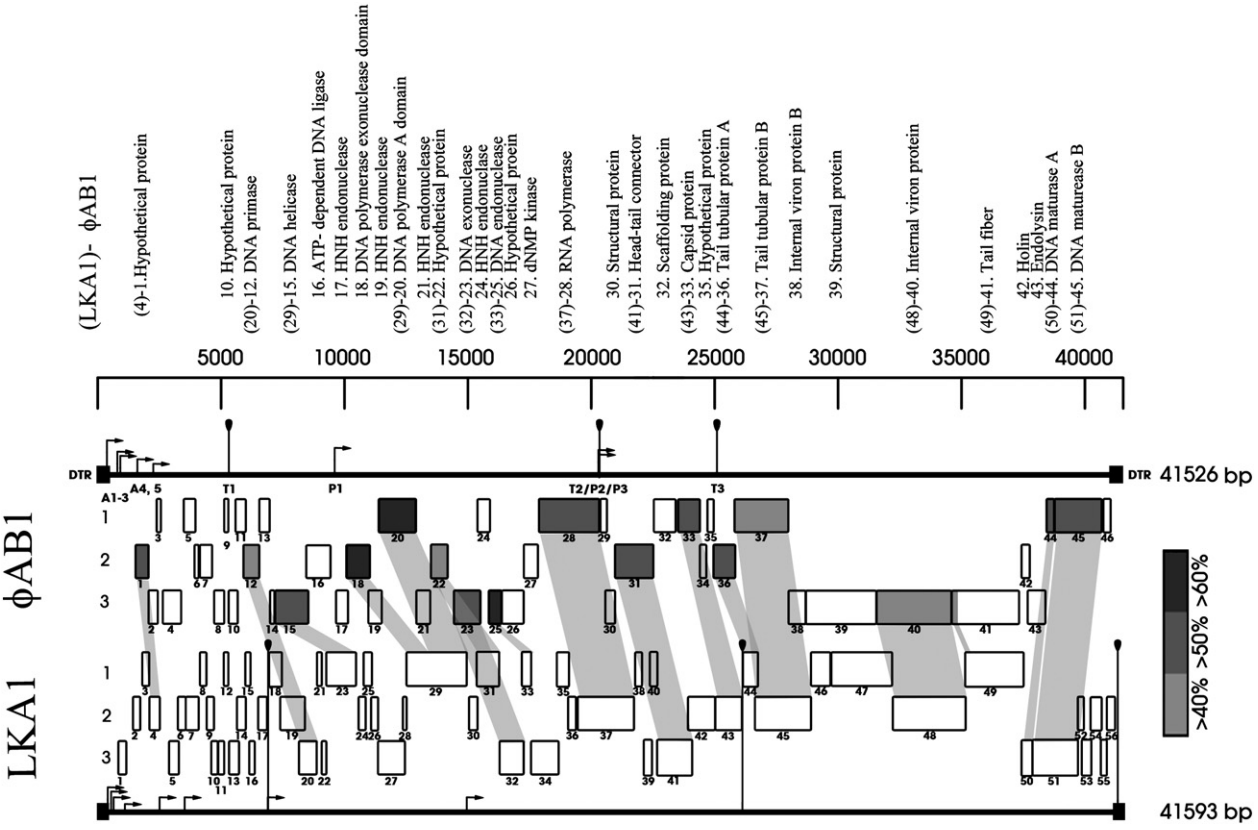

Fig. 1. Comparison of the genomic organization of phages LKA1 (NC_009936) and ϕAB1. Predicted open reading frames (ORFs) are presented numbered and arranged according to their reading frame. The percentages of amino acid similarities of similar ORFs in LKA1 and ϕAB1 are depicted in a grey scale. Predicted functions of the possible gene products encoded by ϕAB1 are shown in the right column. Putative host promoters (A) and phage promoters (P) are indicated. The vertical rods topped with closed circles indicate the presence of possible intrinsic transcription terminators.

The ϕAB1 genome can be divided roughly into two regions, analogous to the previously annotated ϕKMV-like phages [14,15]. The first region contains the early genes relating to host conversion and DNA replication, and the second region encompasses the late genes coding for structural and lysis proteins. The location of the gene for RNAP (ORF28) is at the end of the DNA metabolism genes.

In summary, the phage particles of ϕAB1 and the ϕKMV-like group display a similar morphology, characteristic of the *Podoviridae* group. In total, 22 of 46 gp (including RNAP) from regions distributed along the whole ϕAB1 genome show significant sequence similarity to proteins from the ϕKMV-like phages. The functional annotations of the predicted gp confirm the common genome structure between ϕAB1 and the ϕKMV-like phages. Most importantly, the ORF encoding the putative RNAP in ϕAB1 is situated adjacent to the structural protein region, a hallmark feature of the known genome organizations of the ϕKMV-like phages [14,17]. Therefore, we suggest that ϕAB1 is a new member of the ϕKMV-like phages.

We have also compared the genomic sequence of ϕAB1 to the available partial genomic sequence of phage ϕAB2. ϕAB2 was the first lytic phage infecting *A. baumannii* to be characterized [13]. The two phages resemble one another in morphology and genome size but vary in their spectrum of activity against clinical isolates of MDRAB. Surprisingly, DNA analysis (including the putative ORFs encoding RNA polymerases) suggested strong homology between these two phages, with the DNA identity ranging from 91 to 99% of the matched area. Therefore, we believe that ϕAB1 and ϕAB2 are similar phages. The difference in host specificity between these phages is possibly caused by an alteration of the tail fiber proteins (see below).

2.2. Structural protein analysis by whole-phage shotgun mass spectrometry

Whole-phage shotgun analysis (WSA) was used to identify ϕAB1 structural proteins by mass spectrometry. These data allowed us to assign nine structural proteins to annotated ORFs (Table 1), two of which (gp30 and gp39) showed no significant similarity in BLASTp searches and could be identified as structural proteins. Similar to results obtained in a WSA analysis of ϕKMV [18], the number of peptides identified within individual proteins was very limited in our study. For all WSA-confirmed structural proteins in ϕKMV, their homologs (as defined by BLASTp) were found in our WSA of ϕAB1, except for the putative scaffold protein gp32. This probably results from the absence of scaffold proteins in mature phage particles [18].

2.3. Putative regulatory elements

The potential prokaryotic promoters in the ϕAB1 genome were initially predicted using the BDGP program [19]. Compared to the three reported ϕKMV-like phages, which carry four to five putative host promoters in the upstream end of their early regions [14,15], five equivalent possible host σ^{70} promoters (A1–A5) were identified in ϕAB1 (Fig. 1). By identifying genome blocks displaying sequence similarity corresponding to conserved regulatory elements, three phage promoters in the ϕAB1 genome were predicted (Fig. 1 and Supplementary Table 2). The first putative phage promoter P1 lies in front of ORF17 while the other two (P2 and P3) are located downstream of the terminator T2 and preceding ORF29. The location of these two promoters conforms to our presumption that ORF29 marks the beginning of the late genomic region. The consensus sequences of the three possible phage promoters are dramatically different from other predicted ϕKMV-like promoters [14,20], suggesting a variation in transcriptional regulation between these phages. Based on a similar approach used in the analysis of the LKA1 and LKD16 genomes [14], three possible intrinsic transcription terminators were identified, located after ORF9, ORF28, and ORF35 (Fig. 1 and Supplementary Table 2). The second terminator-like sequence (T2) sits upstream of the predicted phage promoters P2 and P3, all at the junction between the presumed early and late genomic regions. However, in contrast to the three reported ϕKMV-like phages which possess a consistent intrinsic transcription terminator at the end of their genome, there is no intrinsic terminator found at the end of the genome in ϕAB1.

Table 1
Mass spectrometry data for whole-phage shotgun analysis of ϕAB1. Whole-phage particles were digested and analyzed by a 2D-nano LC–MS/MS system. The generated peptide MS/MS spectra were identified by comparison to all potential peptide spectra deduced from the ϕAB1 genome sequence. The number of identified peptides for each inferred parent gene product (gp), the corresponding function of each gp predicted by BLAST, and the protein sequence coverage are indicated.

ϕAB1 protein	Predicted function	Number of peptides	Sequence coverage (%)
gp30	Hypothetical protein	1	9.64
gp31	Head–tail connector protein	1	1.74
gp33	Capsid protein	40	59.48
gp36	Tail tubular protein A	2	12.02
gp37	Tail tubular protein B	1	0.92
gp39	Hypothetical protein	5	4.68
gp40	Internal virion (core) protein	1	1.84
gp41	Tail fiber	2	3.17
gp45	DNA maturase B	1	3.70

2.4. Phylogenetic analysis of RNA polymerase

Consistent with other reported ϕKMV-like phages, ORF28 (which possibly encodes RNAP) of ϕAB1 is located at the end of the DNA replication region. The majority of proteins displaying high similarity to gp28 in the BLASTp analysis were RNAPs from the ϕKMV-like group, with the exception being a protein from the *Klebsiella* phage KP34. According to the new classification standards from the International Committee on the Taxonomy of Viruses, the ϕKMV-like phages together with the T7 *sensu strictu* group and the SP6-like phages form the *Autographivirinae* subfamily [17]. Therefore, we constructed a phylogenetic tree (Fig. 2) based on the putative RNAP protein sequences from the phages ϕAB1, the ϕKMV-like group, T7 (T7-like family), and SP6 (SP6-like family). Ten RNAPs from the ϕKMV-like group were filtered, keeping five protein sequences with mutual sequence identities under 90%. While the topology of the tree supports the assignment of ϕAB1 as a new member of the ϕKMV-like group, ϕAB1 is still relatively distant from the other members in the group.

2.5. HNH endonucleases and genes-in-pieces DNA polymerase

Comparing to other reported ϕKMV-like viruses, an unusual feature of ϕAB1 is the existence of four ORFs encoding putative HNH endonucleases (gp17, gp19, gp21, and gp24). The HNH endonuclease is a member of the homing endonuclease family which confer the mobility or duplication of their coding and flanking sequences by a recombination-based process known as homing [21,22]. The BLASTp searches clearly demonstrated the similarity of the four proteins to various HNH endonucleases found in phages and bacteria, including matches between both gp17 and gp24 and gp13 from Vibrio phage VP93. Further queries against the HNH domain database (HNHDb) [23] revealed that the HNH domains in the N-terminals of the four putative HNH endonucleases all belong to subset 10 of the HNH class. This subset is one of the ten classes in the HNHDb characterized as containing an H–N–N motif (Fig. 3). An H–N–N motif is the conserved sequence pattern commonly found in HNH endonucleases encoded within self-splicing introns and inteins or free standing ORFs in bacteriophages, viruses, and bacteria. In addition to containing the HNH domain, HNH endonucleases are also commonly observed to carry DNA-binding domains such as AP2 and NUMOD1/NUMOD4 [21–24]. For the four putative HNH endonucleases in ϕAB1, InterProScan [25] suggested an AP2 domain located

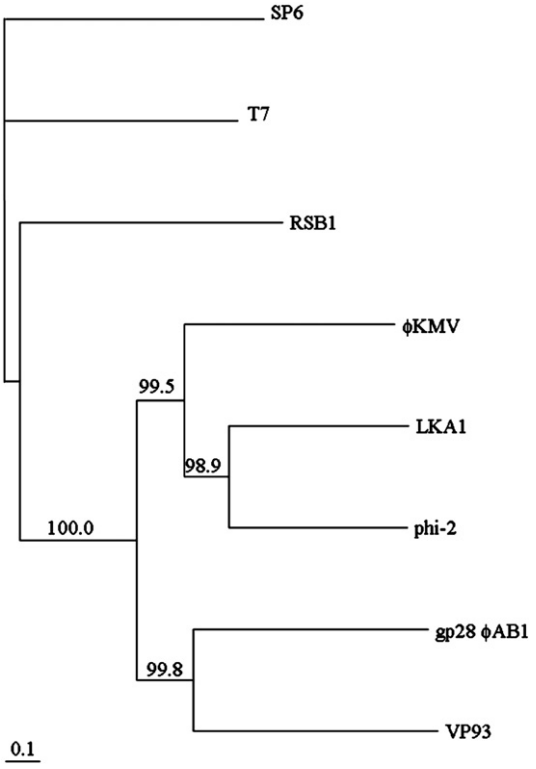

Fig. 2. Phylogenetic neighbor-joining tree constructed using Phylip-3.69 [55] based on the sequence alignment of RNA polymerases (RNAPs) between ϕAB1, ϕKMV-like phages (RSB1, VP93, ϕKMV, LKA1 and phi-2), SP6, and T7. Five redundant RNAPs from the ϕKMV-like phages (>90% sequence identity) were removed before the phylogenetic analysis. Multiple sequence alignments were viewed and managed using Jalview 2.4.0 [54]. Bootstrap support was estimated with 1000 replicates and the support percentages above 70% are shown at the nodes.

in the C-terminal region of gp19. Further sequence comparison between these four putative HNH endonucleases and the AP2 motif derived from the AP2 domain in the Conserved Domain Database (CDD) [26] indicated that adjacent to the H–N–N motif of gp21 there is a YRG motif [27], implying that this region is an incomplete AP2 domain. No DNA-binding domains have been identified thus far in the C-terminal regions of gp17 and gp24.

Based on these analyses, we divided the four putative HNH endonucleases of ϕAB1 into two groups. One consists of gp17 and gp24 and the other includes gp19 and gp21. The sequence identities based on the corresponding HNH domain in each group are 63.6% and 48.8%, respectively. The sequence alignment and domain organizations of the four putative HNH endonucleases from ϕAB1 are presented in Fig. 3. The possible involvement of HNH endonucleases in DNA duplication and mobility may suggest that gp 17, 19, 21, and 24 in ϕAB1 originated from early self-duplications followed by domain degeneration. Similar observations of degenerated H–N–N–AP2 endonucleases and a similar model of gain-and-loss have been emphasized in the study of the 21 putative HNH endonucleases in *Xanthomonas* phages [28].

Additionally, gp18 and gp20 showed significant similarity to two separate domains within the same putative DNA polymerase proteins in the ϕKMV-like phages (gp29 in LKA1 and gp19 in ϕKMV and LKD16): gp18 mapped to the N-terminal domain belonging to the DnaQ-like 3′–5′ exonuclease superfamily, and gp20 corresponded to the C-terminal DNA polymerase A domain. The gene encoding the putative HNH-AP2 endonuclease lies between ORF18 and ORF20 in ϕAB1. This leads us to hypothesize that a mobility-promoting HNH endonuclease gene integrated into a DNA polymerase gene in ϕAB1. This insertion may have split an intact ancestral gene into two independent genes, ORF18 and ORF20, which encode 3′–5′ exonuclease and DNA polymerase domains and resulted in the genes-in-pieces organization. Many examples of naturally occurring genes-in-pieces arrangements of DNA polymerases with highly diverse intergenic sequences have been described [29–33]. In one particular example in the gp43 B-type DNA polymerase of T4-like phage, the intervening region is a GIY–YIG homing endonuclease [33]. In *Aeromonas hydrophila* phage Aeh1, two functional genes were created after the insertion of an HNH endonuclease gene into the original *nrd*A gene [34].

2.6. Tail fiber

As previously mentioned, the downstream fragment of ORF41 was similar to genomic regions of *A. baumannii* in a BLASTn analysis. BLASTp searches further revealed strong similarity at the corresponding region (amino acids 751 to 882) of gp41 to the C-terminal region of many prophage tail fiber proteins, the most similar of which were found in various strains of *Acinetobacter* (Supplementary Fig. 1). These proteins are generally half the length of the complete gp41. In addition, the 150 N-terminal residues of gp41 displayed weak similarity to the N-terminal regions of gp49 from LKA1 and to gp38 from ϕKMV, a conserved region found among tail fibers of the T7-like group [14]. This may suggest that the 150 N-terminal residues of gp41 in ϕAB1 are responsible for attaching the tail fibers to the phage virion.

Given that the C-terminal regions of the tail fiber proteins in LKA1 are responsible for binding to the receptor on the host-surface [14] and that we observed strong conservation between the C-terminal region of gp41 and the prophage tail fibers in *Acinetobacter*, we speculate that the distal end of gp41 is important for ϕAB1 to

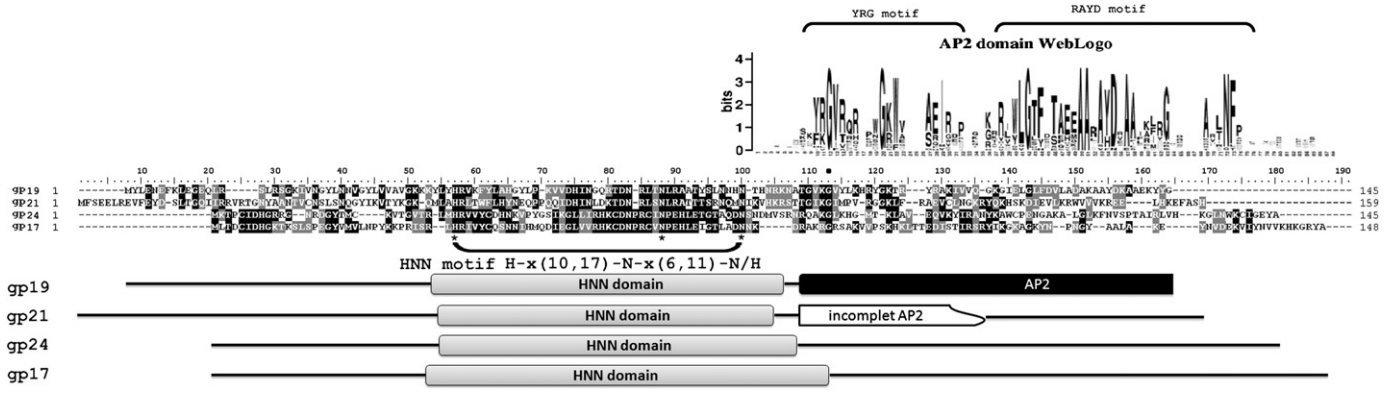

Fig. 3. Sequence alignment of the four putative HNH endonucleases (gp 17, 19, 21, and 24) of ϕAB1. The H–N–N motif is represented in the PROSITE pattern (H–x(10,17)–N–x(6,11)–N/H). AP2 logos derived from the AP2 domain in the Conserved Domain Database [26] are shown above the sequence alignment.

recognize *A. baumannii*. Among the five genomic fragments of ϕAB2 [13], sequence comparisons showed that the entire fragments were highly identical to their counterparts in ϕAB1, except for the DNA fragment GU979517. Only the first 250 bp of a total of 1980 bp of GU979517 matched to the 5′ end of ORF41 (2646 bp) in ϕAB1. This may support the presumption that the C-terminal region of gp41 determines the ϕAB1 host specificity, resulting in the diverse infectivity patterns for *A. baumannii* strains between ϕAB1 and ϕAB2 [13]. Interestingly, sequence similarity between gp41 and the tail fiber (gp76) of phage AB1 was detected in the region (amino acids 166 to 320) immediately downstream of the 150 N-terminal residues of gp41. Whether the corresponding regions contribute to the specificity of the two phages for *A. baumannii* remains to be defined.

2.7. Holin and endolysin

Many dsDNA bacteriophages use a holin–endolysin (or holin–lysozyme) system to lyse host bacterial cells to release progeny virions [35–37]. Holins are small transmembrane proteins that function to perforate the cytoplasmic membrane of the host bacteria. Once the lesion in the cytoplasmic membrane forms, other proteins, including endolysins, are allowed to pass. Endolysins are a type of muralytic enzymes that are responsible for the degradation of the peptidoglycan cell wall. Some phages also encode Rz and Rz-like proteins to facilitate the lysis process by destabilizing the outer membrane of Gram-negative bacteria [38].

The genetic elements encoding the products involved in the holin–endolysin system are commonly found adjacent to one another in the phage genomes [39]. In the three reported ϕKMV-like phages (ϕKMV, LKD16, and LKA1), the ORFs of the lysis cassette composed of holin, endolysin, and putative Rz/Rz-like proteins are located near the downstream end of the late region [14,15]. However, in contrast to the reported ϕKMV-like phages, the putative lysis cassette in ϕAB1 is upstream of the ORFs encoding DNA maturases. Furthermore, no protein sequence similarity was detected to the gene products from the ϕKMV-like phages. A conserved lysozyme-like domain was detected in gp43 using the BLASTp algorithm. Functional analysis of the product encoded by the identical gene in ϕAB2 confirmed the endolysin activity of gp43 (manuscript in preparation). It is notable that BLAST results indicated that gp43 in ϕAB1 is homologous to the gp73 of phage AB1, as it is one of the two gp (the other being the tail fiber protein) sharing detectable sequence similarity between the two phages ϕAB1 and AB1. A BLASTp search for gp42 resulted in only one protein with a very low homology score, which was shown to be a holin protein (ZP_04591001, BLASTp *E*-value 6.1) belonging to the phage holin 3 family (pfam05106) [40]. This family is a collection of mutually dissimilar phage holins and contains the representative phage λ S protein. Most holins can be grouped into class I or class II based on the possible presence of either two or three transmembrane domains (TMDs) [17,39,41]. Based on these clues, we chose 11 sequences from the 25 seed sequences of pfam05106 that have the same number (three) of predicted TMDs as gp42 and generated a sequence alignment along with gp42 and phage λ S protein (Fig. 4).

This sequence alignment demonstrates that the sequence conservation is mainly concentrated in two regions: TM2 and the loop connecting TM1 and TM2, reported to be essential for the lysis

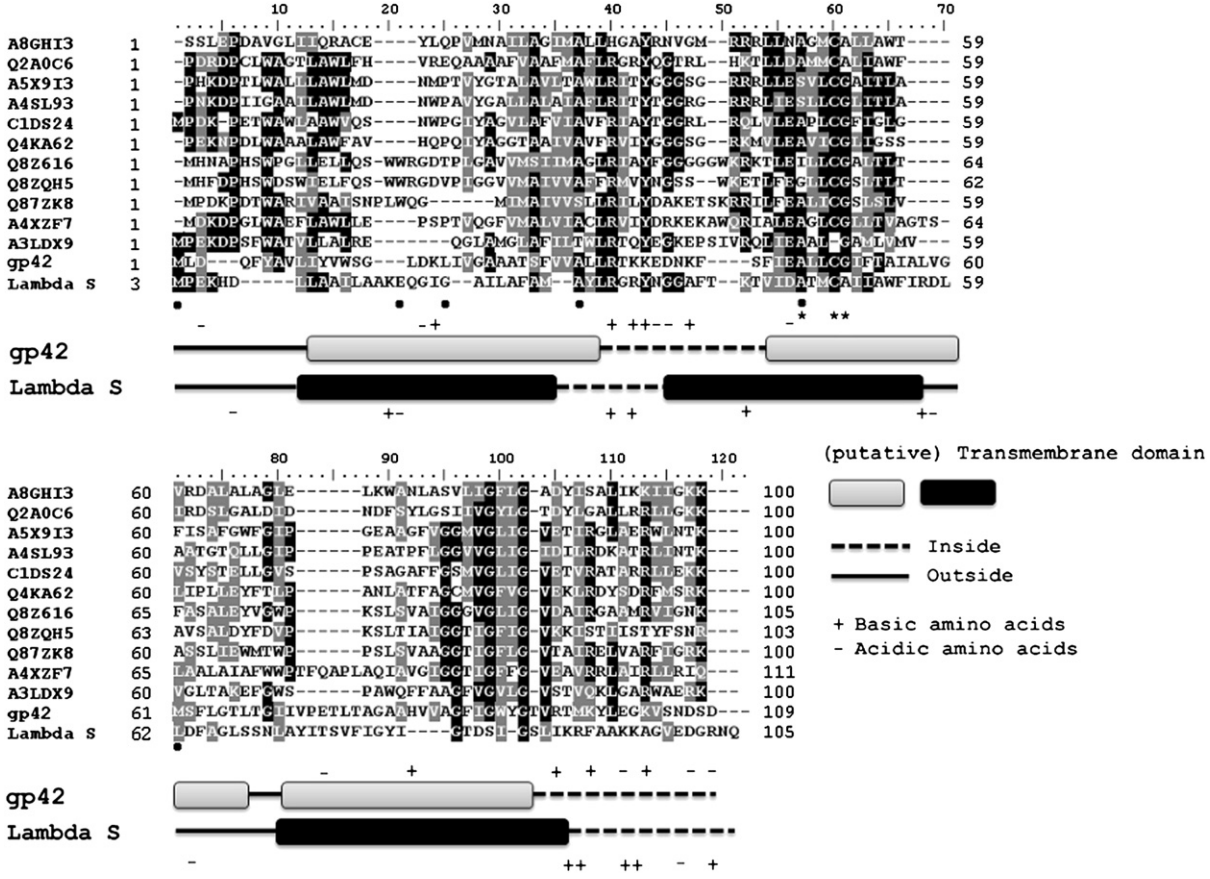

Fig. 4. Comparison of gp42 of ϕAB1, phage λ S, and selected holin proteins containing three putative transmembrane regions (TM1, TM2, and TM3). The conserved sequences are primarily in the TM2 domain and the loop connecting the TM1 and TM2 region, reported to be essential for the lysis function of the phage λ S protein [42]. +/−: basic and acidic residues observed in the gp42 and phage λ S proteins; .: M3, A30, and A48 in phage λ S are three of the residues identified to be essential for the host lethality of phage λ S [43] and are also found in gp42; *: C51 in phage λ S has been suggested to participate in dimerization [17] and is conserved in the alignment.

function of the phage λ S protein [42]. The rich distribution of charged residues in the loop between TM1 and TM2 and the C-terminal region of gp42 also match the common arrangement of charged residues observed in many holin proteins [36]. Some of the residues in phage λ S holin which were identified to be important for host lethality [43] were found in gp42 (M3, A30, and A48 in λ S numbering). Furthermore, residue C51, which was suggested to participate in dimerization of λ S protein [17], is also present in gp42. In contrast to phage λ S protein [44], no potential dual translational start motifs were predicted in the coding sequence of gp42.

In summary, despite lacking sequence similarity to any well-characterized holins, the sequence features displayed in the predicted gp of ORF42, such as the discernible TMDs and a hydrophilic, highly charged-C-terminus, suggest that ORF42 is a potential holin gene and together with ORF43 forms a lysis cassette in the genome of ϕAB1.

Downstream of the predicted lysis cassette of ϕAB1 are ORF44 and ORF45, which encode homologs of DNA maturases A and B in other ϕKMV-like phages. The intergenic region between ORF43 and ORF44 is only 118 bases, leaving little space to accommodate ORFs encoding Rz or Rz-like proteins.

3. Conclusion

In this study, we present the genomic analysis of ϕAB1, a lytic phage infecting clinical isolates of MDRAB. According to the morphological features presented in a previous study [13] and the overall genomic organization and sequence similarities revealed here, ϕAB1 is suggested to be a new member of the ϕKMV-like group. Phylogenetic analysis of the RNAPs also supported this assignment, although ϕAB1 was relatively divergent from other group members. The existence of four ORFs encoding putative HNH endonucleases is a key distinguishing feature that sets ϕAB1 apart from other ϕKMV-like phages. One of the HNH endonuclease genes was postulated to have integrated into and split the DNA polymerase gene. Furthermore, in the downstream region of the ϕAB1 genome, we have located a possible lysis cassette, which showed no detectable sequence similarity to lysis cassettes found in the ϕKMV-like phages. Facing the emerging threat from MDRAB, the lytic power of ϕAB1 combined with its specificity for *A. baumannii* makes ϕAB1 an appealing agent for therapeutic or disinfection applications.

4. Materials and methods

4.1. Bacteriophage amplification and purification

Phage ϕAB1 was isolated from wastewater as mentioned in [13]. Phages were amplified using *A. baumannii* strain M68316 (DSM 23587) as a host, and purified by CsCl density gradient centrifugation.

4.2. DNA sequencing

Multiple independent isolations of phage DNA were carried out a described previously [13]. After random shearing, gel-recovered fragments approximately 3 kb in length were cloned into the SmaI site of the pUC19 vector and transformed into *Escherichia coli* DH10B cells (Invitrogen). Plasmid clones prepared from multiple transformations were sequenced on an ABI 3730xl DNA analyzer (Applied Biosystems). A total of 180 fragments were sequenced on both strands. Base calling and assembly were performed and viewed with Phred/Phrap/Consed package [45–47] using default parameters. One contig with a total length of 21,885 bp was obtained, resulting in an approximate fifteen-fold redundancy. The contig was then extended by 28 primer-walking steps using purified ϕAB1 genomic DNA as the template. Every oligonucleotide primer with the length ranged from 20 to 27 was designed to extend the contig. The primering sites were chosen to ensure that there was a reliable overlapping region of at least 50 bases between the new and the old sequence data. All regions have been sequenced at least twice, resulting in an approximate 2.5-fold redundancy by primer walking. The complete genome of ϕAB1 is 41,526-bp in length.

The genomic sequence of ϕAB1 was deposited in GenBank (accession number HQ186308).

4.3. In silico analysis

The genome was scanned for potential open reading frames (ORFs) by the GeneMark.hmm program for prokaryotes [48]. Annotation was performed using a combination of NCBI-BLAST [49] searches on predicted gene products (BLASTp) as well as on the entire genome of ϕAB1 (BLASTx and BLASTn) against the non-redundant nucleotide and protein database. Scanning for the existence of known domains was performed using InterProScan [25]. The prediction of transmembrane regions was performed using TMHMM 2.0 [50]. Host promoters were predicted using the BDGP program [19]. The possible phage promoters were identified by the PHIRE program [51]. Putative intrinsic transcription terminators were located by the same approach mentioned in [14]. In brief, we searched for a potential stem-loop forming sequence followed by a uracil-rich stretch and with a stable secondary structure (ΔG less than −10 kcal/mol), as calculated by the MFold program [52].

4.4. Phylogenetic analysis

Ten RNA polymerase protein sequences from each phage of the ϕKMV-like group (RSB1, accession number YP_002213715; ϕKMV, NP_877465; LKA1, YP_001522878; phi-2, YP_003345489; VP93, YP_002875649; phiKF77, YP_002727849; LUZ19, YP_001671971; PT2, YP_002117810; LKD16, YP_001522818; and PT5, YP_002117751) were filtered to have fewer than 90% mutual sequence identities. The remaining five RNAPs (RSB1, VP93, ϕKMV, LKA1 and phi-2) together with the RNAPs from phages SP6 (NP_853568) and T7 (NP_041960) and the putative RNAP gp28 (ADQ12732) of ϕAB1 were used to generate a multiple sequence alignment by Expresso (3DCoffee) [53]. The sequences were viewed and managed using Jalview 2.4.0 [54]. The neighbor-joining tree was constructed using Phylip-3.69 [55] assignment of SP6 as the outgroup. One thousand replicates were run to obtain bootstrap support percentages at each node. The maximum likelihood method using PHYML 3.0 [56] was also employed to confirm the result.

4.5. Whole-phage shotgun mass spectrometry

Phage particles purified from CsCl gradient centrifugation were disrupted by sonication in 30% acetonitrile solution for 20 min. After sonication, the solution was filtered through a 0.45 μm membrane filter and concentrated using a YM-3 (Millipore) membrane. After trypsin digestion, the peptide products were analyzed by a 2D-nano LC–MS/MS system. The peptide mixture was fractionated using a strong-cation-exchange column followed by a reverse-phase C18 column and then analyzed by LCQ Fleet ion trap mass spectrometry (Thermo Fisher Scientific Inc.). The data generated were searched using Biowork/SEQUEST [57] against a local database of all possible peptide spectra deduced from the ϕAB1 genome sequence.

Acknowledgments

We thank Dr. J. C. Fowler for reading the manuscript. We also thank Dr. The-Yuan Chow, Genomics BioSci & Tech for technical assistance. This work was supported by grant TCSP-03-06 from the Buddhist Tzu Chi General Hospital.

Appendix A. Supplementary data

Supplementary data to this article can be found online at doi:10.1016/j.ygeno.2011.01.002.

References

[1] A.Y. Peleg, H. Seifert, D.L. Paterson, *Acinetobacter baumannii*: emergence of a successful pathogen, Clin. Microbiol. Rev. 21 (2008) 538–582.
[2] C.W. Ong, et al., Severe community-acquired *Acinetobacter baumannii* pneumonia: an emerging highly lethal infectious disease in the Asia-Pacific, Respirology 14 (2009) 1200–1205.
[3] T. Ozaki, et al., Community-acquired *Acinetobacter baumannii* meningitis in a previously healthy 14-month-old boy, J. Infect. Chemother. 15 (2009) 322–324.
[4] F. Perez, et al., Global challenge of multidrug-resistant *Acinetobacter baumannii*, Antimicrob. Agents Chemother. 51 (2007) 3471–3484.
[5] M. Meric, et al., Emergence and spread of carbapenem-resistant *Acinetobacter baumannii* in a tertiary care hospital in Turkey, FEMS Microbiol. Lett. 282 (2008) 214–218.
[6] J. Garnacho-Montero, R. Amaya-Villar, Multiresistant *Acinetobacter baumannii* infections: epidemiology and management, Curr. Opin. Infect. Dis. 23 (2010) 332–339.
[7] J. Garnacho-Montero, R. Amaya-Villar, Multiresistant *Acinetobacter baumannii* infections: epidemiology and management, Curr. Opin. Infect. Dis. 23 (2010) 332–339.
[8] J. Fishbain, A.Y. Peleg, Treatment of *Acinetobacter* infections, Clin. Infect. Dis. 51 (2010) 79–84.
[9] V. Manchanda, S. Sanchaita, N. Singh, Multidrug resistant acinetobacter, J. Glob. Infect. Dis. 2 (2010) 291–304.
[10] N.H. Mann, The third age of phage, PLoS Biol. 3 (2005) e182.
[11] Phage list in Entrenz Genome, http://www.ncbi.nlm.nih.gov/genomes/genlist.cgi?taxid=10239&type=6&name=PhagesJuly 2010.
[12] H. Yang, L. Liang, S. Lin, S. Jia, Isolation and characterization of a virulent bacteriophage AB1 of *Acinetobacter baumannii*, BMC Microbiol. 10 (2010) 131.
[13] N.T. Lin, P.Y. Chiou, K.C. Chang, L.K. Chen, M.J. Lai, Isolation and characterization of phiAB2, a novel bacteriophage of *Acinetobacter baumannii*, Res. Microbiol. 161 (4) (2010) 308–314.
[14] P.J. Ceyssens, et al., Genomic analysis of *Pseudomonas aeruginosa* phages LKD16 and LKA1: establishment of the phiKMV subgroup within the T7 supergroup, J. Bacteriol. 188 (2006) 6924–6931.
[15] R. Lavigne, et al., The genome of bacteriophage phiKMV, a T7-like virus infecting *Pseudomonas aeruginosa*, Virology 312 (2003) 49–59.
[16] phiKMV-like viruses (Taxonomy ID: 477967) in NCBI taxonomy database, http://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?mode=Undef&id=477967&lvl=3&keep=1&srchmode=1&unlockJuly 2010.
[17] A. Grundling, U. Blasi, R. Young, Biochemical and genetic evidence for three transmembrane domains in the class I holin, lambda S, J. Biol. Chem. 275 (2000) 769–776.
[18] R. Lavigne, et al., The structural proteome of *Pseudomonas aeruginosa* bacteriophage phiKMV, Microbiology 152 (2006) 529–534.
[19] M.G. Reese, Application of a time-delay neural network to promoter annotation in the *Drosophila melanogaster* genome, Comput. Chem. 26 (2001) 51–56.
[20] R. Lavigne, B. Roucourt, K. Hertveldt, G. Volckaert, Characterization of the bacteriophage PhiKMV DNA ligase, Protein Pept. Lett. 12 (2005) 645–648.
[21] B.L. Stoddard, Homing endonuclease structure and function, Q. Rev. Biophys. 38 (2005) 49–95.
[22] M.S. Jurica, B.L. Stoddard, Homing endonucleases: structure, function and evolution, Cell. Mol. Life Sci. 55 (1999) 1304–1326.
[23] A. Veluchamy, et al., HNHDb: a database on pattern based classification of HNH domains reveals functional relevance of sequence patterns and domain associations, Bioinformation 4 (2009) 80–83.
[24] E. Sitbon, S. Pietrokovski, New types of conserved sequence domains in DNA-binding regions of homing endonucleases, Trends Biochem. Sci. 28 (2003) 473–477.
[25] E.M. Zdobnov, R. Apweiler, InterProScan—an integration platform for the signature-recognition methods in InterPro, Bioinformatics 17 (2001) 847–848.
[26] A. Marchler-Bauer, et al., CDD: a Conserved Domain Database for protein classification, Nucleic Acids Res. 33 (2005) D192–D196.
[27] J.K. Okamuro, B. Caster, R. Villarroel, M. Van Montagu, K.D. Jofuku, The AP2 domain of APETALA2 defines a large new family of DNA binding proteins in *Arabidopsis*, Proc. Natl Acad. Sci. USA 94 (1997) 7076–7081.
[28] C.N. Lee, et al., Comparison of genomes of three *Xanthomonas oryzae* bacteriophages, BMC Genomics 8 (2007) 442.
[29] L. Huang, et al., The enzymological basis for resistance of herpesvirus DNA polymerase mutants to acyclovir: relationship to the structure of alpha-like DNA polymerases, Proc. Natl Acad. Sci. USA 96 (1999) 447–452.
[30] Z. Kelman, S. Pietrokovski, J. Hurwitz, Isolation and characterization of a split B-type DNA polymerase from the archaeon *Methanobacterium thermoautotrophicum* deltaH, J. Biol. Chem. 274 (1999) 28751–28761.
[31] D.R. Smith, et al., Complete genome sequence of *Methanobacterium thermoautotrophicum* deltaH: functional analysis and comparative genomics, J. Bacteriol. 179 (1997) 7135–7155.
[32] H. Wu, Z. Hu, X.Q. Liu, Protein trans-splicing by a split intein encoded in a split DnaE gene of Synechocystis sp. PCC6803, Proc. Natl Acad. Sci. USA 95 (1998) 9226–9231.
[33] V.M. Petrov, J.D. Karam, Diversity of structure and function of DNA polymerase (gp43) of T4-related bacteriophages, Biochemistry (Mosc.) 69 (2004) 1213–1218.
[34] N.C. Friedrich, et al., Insertion of a homing endonuclease creates a genes-in-pieces ribonucleotide reductase that retains function, Proc. Natl Acad. Sci. USA 104 (2007) 6176–6181.
[35] M. Takac, A. Witte, U. Blasi, Functional analysis of the lysis genes of *Staphylococcus aureus* phage P68 in *Escherichia coli*, Microbiology 151 (2005) 2331–2342.
[36] R. Young, Bacteriophage lysis: mechanism and regulation, Microbiol. Rev. 56 (1992) 430–481.
[37] G. Ziedaite, R. Daugelavicius, J.K. Bamford, D.H. Bamford, The Holin protein of bacteriophage PRD1 forms a pore for small-molecule and endolysin translocation, J. Bacteriol. 187 (2005) 5397–5405.
[38] E.J. Summer, et al., Rz/Rz1 lysis gene equivalents in phages of Gram-negative hosts, J. Mol. Biol. 373 (2007) 1098–1112.
[39] I.N. Wang, D.L. Smith, R. Young, Holins: the protein clocks of bacteriophage infections, Annu. Rev. Microbiol. 54 (2000) 799–825.
[40] R.D. Finn, et al., The Pfam protein families database, Nucleic Acids Res. 38 (2010) D211–D222.
[41] M.J. Loessner, Bacteriophage endolysins—current state of research and applications, Curr. Opin. Microbiol. 8 (2005) 480–487.
[42] R. Raab, G. Neal, C. Sohaskey, J. Smith, R. Young, Dominance in lambda S mutations and evidence for translational control, J. Mol. Biol. 199 (1988) 95–105.
[43] R. Raab, et al., Mutational analysis of bacteriophage lambda lysis gene S, J. Bacteriol. 167 (1986) 1035–1042.
[44] U. Blasi, C.Y. Chang, M.T. Zagotta, K.B. Nam, R. Young, The lethal lambda S gene encodes its own inhibitor, EMBO J. 9 (1990) 981–989.
[45] B. Ewing, P. Green, Base-calling of automated sequencer traces using phred. II. Error probabilities, Genome Res. 8 (1998) 186–194.
[46] B. Ewing, L. Hillier, M.C. Wendl, P. Green, Base-calling of automated sequencer traces using phred. I. Accuracy assessment, Genome Res. 8 (1998) 175–185.
[47] D. Gordon, C. Abajian, P. Green, Consed: a graphical tool for sequence finishing, Genome Res. 8 (1998) 195–202.
[48] A.V. Lukashin, M. Borodovsky, GeneMark.hmm: new solutions for gene finding, Nucleic Acids Res. 26 (1998) 1107–1115.
[49] M. Johnson, et al., NCBI BLAST: a better web interface, Nucleic Acids Res. 36 (2008) W5–W9.
[50] A. Krogh, B. Larsson, G. von Heijne, E.L. Sonnhammer, Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes, J. Mol. Biol. 305 (2001) 567–580.
[51] R. Lavigne, W.D. Sun, G. Volckaert, PHIRE, a deterministic approach to reveal regulatory elements in bacteriophage genomes, Bioinformatics 20 (2004) 629–635.
[52] M. Zuker, On finding all suboptimal foldings of an RNA molecule, Science 244 (1989) 48–52.
[53] F. Armougom, et al., Expresso: automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee, Nucleic Acids Res. 34 (2006) W604–W608.
[54] A.M. Waterhouse, J.B. Procter, D.M. Martin, M. Clamp, G.J. Barton, Jalview Version 2—a multiple sequence alignment editor and analysis workbench, Bioinformatics 25 (2009) 1189–1191.
[55] J. Felsenstein, PHYLIP (Phylogeny Inference Package) Version 3.69, in, Distributed by the author, Department of Genome Sciences, University of Washington, Seattle, 2005.
[56] S. Guindon, F. Lethiec, P. Duroux, O. Gascuel, PHYML Online—a web server for fast maximum likelihood-based phylogenetic inference, Nucleic Acids Res. 33 (2005) W557–W559.
[57] J.K. Eng, A.L. McCormack, J.R. Yates III, An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database, J. Am. Soc. Mass Spectrom. 5 (1994) 976–989.