RESEARCH ARTICLE
doi: 10.2306/scienceasia1513-1874.2009.35.125

ScienceAsia 35 (2009): **125–130**

Intriguing phylogenetic arrangement of tailed bacteriophages based on putative DNA polymerase sequences

Tirasak Pasharawipas[a,b], **Surayut Thaikua**[b], **Timothy W. Flegel**[b]

[a] Department of Microbiology, Faculty of Science, Rangsit University, Pathumthani 12000, Thailand
[b] Centex Shrimp, Faculty of Science, Mahidol University, Bangkok 10400, Thailand

*Corresponding author, e-mail: tirasak4124@yahoo.com

Received 14 Aug 2008
Accepted 5 Jun 2009

ABSTRACT: VHS1 (*Vibrio harveyi* siphoviridae like) is a newly described bacteriophage that has been tentatively placed in the family Siphoviridae based on its morphology by transmission electron microscopy. However, there are examples among emergent viruses where morphological characteristics do not correspond with genomic properties. Thus we attempted to verify the phylogenetic relationship of VHS1 to other tailed phages on the basis of its putative DNA polymerase. Although we found no clear support for placement of VHS1 in any tailed-phage family, we did discover that a general phylogenetic tree based on all putative tailed-phage sequences was not compatible with relationships based on phage morphology. However, construction of individual putative DNA polymerase trees for traditional morphological phage families revealed a clear tendency for grouping of putative DNA polymerase A and B types into either Gram negative or Gram positive bacteria hosts in a manner that varied with phage family. Based on these results, we propose that early branching of an ancestral DNA polymerase into polymerase A and B types may have accompanied phage specialization for Gram negative and Gram positive hosts. Whether this proposal is correct or not, the study suggests that comparisons based on major genes such as polymerases can constitute additional criteria for evaluation of DNA phage relationships and that genome comparisons may complement morphology in devising a more natural taxonomy.

KEYWORDS: tailed phages, phylogeny

INTRODUCTION

Bacteriophages are abundant in the marine environment including fish and shrimp aquaculture ponds. In shrimp ponds, *Vibrio harveyi* (VH) is a common bacterial species that sometimes causes luminescent disease, although luminescence is not a requirement for shrimp pathogenicity. On the other hand, it has been shown that bacteriophages can sometimes lysogenize VH and alter their pathogenicity for shrimp[1–3]. Increased virulence due to lysogeny is well known for human pathogens[4], including *Salmonella enteritica*[5], *Vibrio cholerae*,[6] and *Staphylococcus aureus*[7]. At the same time, it has been proposed that lytic bacteriophages may be important in limiting the population of marine bacteria[8,9].

DNA bacteriophages are classified as non-head-tailed and head-tailed phages. No non-head-tailed DNA phages have yet been reported for VH but DNA head-tailed phages have[1,2]. Based on the seventh report of the international committee on taxonomy of viruses (ICTV)[10], there are three families of head-tailed DNA bacteriophages. These are Myoviridae, Siphoviridae, and Podoviridae. Based on morphology, Myoviridae contains phages with icosahedral heads and long, rigid tails (80–455 $\times$ 16–20 nm) with a sheath, while Siphoviridae contains phages with icosahedral heads and long, non-contractile and flexible tails (65–570 nm $\times$ 7–10 nm), and the Podoviridae contains phages with icosahedral heads and very short, non-contractile tails (approximately 20 $\times$ 8 nm). VHS-1 is a Siphoviridae-like phage of *Vibrio harveyi* that was isolated from a black tiger shrimp pond in Southern Thailand. Its head is approximately 60–67 nm in diameter and its tail is approximately 100–120 $\times$ 9–10 nm[11].

DNA polymerase genes are crucial in genomic replication and mutagenic repair. They are classified into 6 families (i.e., A, B, C, D, X, and Y), the first four of which (A–D) contain replicative polymerases[12]. The latter two (X and Y) are newer families that contain repair polymerases[13,14]. The morphological criteria currently used to classify bacteriophage families by ICTV may be considered coarse in that phage

particles have very simple structures when compared to other living organisms. It is doubtful that the few morphological features available would be sufficient to reveal the relationships amongst the members of what may possibly be a very large group. The situation may be complicated further if one considers the potential opportunity for dynamic genetic exchange in marine and other aquatic environments. Here we describe the use of molecular genetic information from the VHS1 putative polymerase gene sequence to determine whether it supports the placement of VHS1 in the family Siphoviridae. This led to a more general comparison of tailed-phage DNA polymerase genes and the revelation of unexpected relationships.

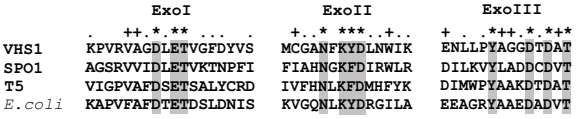

Fig. 1 Amino acid alignments of highly conserved segments of 3′-5′ exonuclease of VHS1 phage (Exo I: position 140–157, Exo II: 217–230, and Exo III: 291–304), SPO1 phage (Exo I: 205–222, Exo II: 290–303, and Exo III: 369–382), T5 phage (Exo I: 131–148, Exo II: 190–203, and Exo III: 281–294), and *E. coli* 0157:H7 (Exo I: 348–365, Exo II: 416–429, and Exo III: 492–505). Star (*): conserved residues; plus (+): strongly similar residues; dot (.): weakly similar residues.

MATERIALS AND METHODS

VHS-1

VHS-1 phage was derived from plaques that arose on lawns of *Vibrio harveyi* VH1114 after exposure to membrane filtered water from a shrimp pond. It was stored in 3% NaCl or phosphate buffered saline (PBS) pH 7.0 at room temperature and amplified when required as previously described[11]. Approximately 20% of the VHS-1 genome has been sequenced and deposited at GenBank[11]. One of the deposited sequences PH102-1 (4668 bp, AF465603) contains an ORF called PH102-768 with significant sequence homology to DNA polymerase sequences at GenBank. That sequence was used together with previous records at GenBank to carry out the analysis described herein.

Sequence analysis and phylogenetic tree prediction

The nucleic acid and deduced amino acid sequences of the selective reported putative phage polymerase genes (including VHS1) were used to identify their classification based on conserved domains by the RPS-blast program of GenBank. The conserved domains of the reported tailed-phage polymerases included the polymerase A family (pfam00136), the polymerase B family (pfam00476), the polymerase C family (cog0587, DnaE pol), and the poxvirus pol family (pfam03288) which is also a member of polymerase family B. The results are shown in Table 1. The nucleic acid sequences were converted to deduced amino acid sequences to carry out phylogenetic analysis using the program Phylogenetic Tree Prediction of the GeneBee service (www.genebee.msu.su/services/phtree_reduced.html). The amino acid sequences of all the phages were put into the query window in FASTA format and BLOSUM62 was the selected matrix used to prepare the unrooted trees.

RESULTS AND DISCUSSION

Polymerase of VHS-1

ORF102-786 in clone PH102-1 (AF465603) from nucleotide positions 2109 to 4169 showed homology to polymerase genes of both bacterial species and bacteriophages. The highest expectation value (5×10^{-36}) was for a polymerase of *Vibrio cholerae* (NC_002505). It also showed significant homology to polymerases of bacteriophages SP01 (family Myoviridae) of *Bacillus subtilis*[15,16], T5 phage (family Siphoviridae)[17], Mycobacterial phage D29 (family Siphoviridae)[18] and Mycobacterial phage bxb1 (family Siphoviridae)[19] with expectation values ranging from 2×10^{-29} to 4×10^{-7}.

A search for conserved domains (RPS-blast of GenBank) revealed that ORF102-786, contained the 3′-5′ exonuclease domain of polymerase family A. According to the report of Bernard et al[20] there are three highly conserved segments of 3′-5′ exonuclease in *E. coli* (i.e., Exo I, Exo II and Exo III) (AP-003944). The Exo I segment of *E. coli* contains two conserved residues Asp^{355} and Glu^{357} while Exo II possesses the conserved residue Asp^{424} and Exo III contains the conserved residues Tyr^{497} and Asp^{501}. These conserved residues are aligned in Fig. 1 with sequences from VHS1, SPO1, and T5. In addition, the region 515–728 of ORF 102–786 from VHS1 also showed significant similarity to the conserved DNA polymerase A domain, with a significant expectation value (6×10^{-32}).

The amino acid sequences from the polymerase genes of the DNA head-tailed phages were retrieved from the GenBank database and then searched for conserved domains to identify their replicative families (Table 1). A comparison of conserved domains in the 3′-5′ exonuclease region, a repair polymerase, was

Table 1 Details of phage polymerase gene sequences used for phylogenetic analysis. These are listed under phage family (as given in the relevant database) together with host bacterium, accession number and similarity (Blast expectation values or E values) relative to various representative polymerase family sequences. The code pfam00476 represents DNA polymerase family A while the codes pfam 0136 and pfam03175 both represent DNA polymerase family B. The pfam03288 and COG0587 represent poxvirus DNA polymerase and DnaE polymerase which are representatives of polymerase families B and C respectively.

Phage family	Phages	Hosts	Accession number	Amino acids	Replicative polymerase family with expectation value
Myoviridae	T4	*Escherichia coli*	AAD42468	898	Pfam00136: 2×10^{-49}
	RB69	*Escherichia coli*	AAA93077	903	Pfam00136: 8×10^{-49}
	KPV40	*Vibrio parahemolyticus*	AAQ64153	850	Pfam00136: 6×10^{-40}
	S-PM2	*Marine Synechococcus*	CAF34198	830	pfam00136: 4×10^{-36}
	Aeh1	*Aeromonas hydrophila*	NP_943895	919	pfam00136: 1×10^{-25}
	SPO1	*Bacillus substillis*	P30314	924	pfam00476: 5×10^{-117}
	LP65	*Lactobacillus plantarum*	AAV35879	988	pfam00476: 6×10^{-48}
	Phi Adh1	*Lactobacillus gasseri*	NP-050131	771	pfam03288 (Poxvirus pol): 9×10^{-57}
Siphoviridae	L5	*Mycobacterium tuberculosis*	CAA79420	595	pfam00476: 1×10^{-100}
	D29	*Mycobacterium tuberculosis*	O64235	607	pfam00476: 4×10^{-76}
	U2	*Mycobacterium smegmatis*	AAR89682	608	pfam00476: 7×10^{-63}
	Bxb1	*Mycobacterium smegmatis*	NP 075308	608	pfam00476: 9×10^{-60}
	T5	*Escherichia Coli*	P19822	829	pfam00476: 6×10^{-50}
	BT1	*Streptomyces lividans*	CAD80134	624	pfam00476: 6×10^{-36}
	SPO2	*Bacillus substillis*	P06225	648	pfam00476: 4×10^{-13}
	Phi 12	*Staphylococcus aureus*	AAL82293	650	pfam00476: 1×10^{-8}
	SPBc2	*Bacillus subtilis str. 168*	NP 046685	1305	Cog0587 (DnaE pol): 0.0
Podoviridae	T7	*Escherichia Coli*	P00581	704	pfam00476: 6×10^{-104}
	T3	*Escherichia Coli*	NP 523320	704	pfam00476: 4×10^{-101}
	Phi 1122	*Yersinia Pestis*	NP 848283	704	pfam00476: 3×10^{-102}
	GH-1	*Pseudomonas putida*	NP 813764	709	pfam00476: 2×10^{-46}
	SIO1	*Rosebacter* SIO67	NP_064753	580	pfam00476: 5×10^{-37}
	P60	*Synechococcus sp.* WH7803	NP 570330	587	pfam00476: 3×10^{-35}
	VP5	*Vibrio cholerae*	AAR92073	633	pfam00476: 2×10^{-33}
	PaP3	*Pseudomonas aeruginosa*	NP_775225	545	pfam00476: 8×10^{-30}
	VPV262	*Vibrio parahemolyticus*	NP640280	661	pfam00476: 6×10^{-20}
	Phi KMV	*Pseudomonas aeruginosa*	NP 877458	807	pfam00476: 3×10^{-10}
	M2	*Bacillus substillis*	AAA32368	572	pfam03175: 7×10^{-62}
	PZA	*Bacillus substillis*	ERBP2Z	572	pfam03175: 8×10^{-64}
	Phi 29	*Bacillus spp.*	P03680	575	pfam03175: 1×10^{-62}
	B103	*Bacillus subtilis* 9/3	NP 690635	572	pfam03175: 1×10^{-60}
	Cp-1	*Streptococcus pneumoniae*	NP 044817	568	pfam03175: 9×10^{-46}
	P82	*Mycoplasma pulmonis*	NO064636	694	pfam03175: 3×10^{-6}
	P68	*Staphylococcus aureus*	NP 817326	755	pfam03175: 9×10^{-3}
	44AHJD	*Staphylococcus aureus*	NP 817305	755	pfam03175: 3×10^{-3}
Unclassified	HF1	*Halorubrum coriense*	NP_861642	854	pfam03175: 2×10^{-27}
	HF2	*Halorubrum coriense*	AAL54976	854	pfam03175: 2×10^{-27}
	VHS1	*Vibrio harveyi*	AAL85287	786	pfam00476: 6×10^{-37}

also carried out (Table 1). The COG0749 domain is a conserved domain of replicative polymerase family A and its 3′-5′ exonuclease while the COG0417 domain is conserved in polymerase family B. The deduced amino acid sequences of all the phage polymerases in Table 1, including those from the unclassified phages HF1 and HF2 of the Archaebacterial genus *Halorubrum*[21] were included in the construction of phylogenetic trees. Since most of the polymerases fell either into the polymerase A or B category, the following discussion relates predominantly to the A and B polymerases.

Phylogenetic tree based on all reported polymerases

The first attempt at tree construction was made using all the reported putative tailed-phage polymerases with matching domains that could be found in the GenBank database (Table 1). The resulting tree showed three major clades, each with mixed A and B putative polymerases and no particular pattern that would suggest clear phylogenetic trends with respect to either phage family or type of bacterial host.

Separate phylogenetic trees for A and B putative polymerases

The second attempt at tree construction focused on separate trees for the A and B putative polymerase types. The tree for putative polymerase A phages showed some trends, in that two major clades and one smaller one contained phages of a single morphological type in a common type of bacterium (i.e., Gram + or −). However, 3 phages, including VHS1 fell outside the 3 clades. With the putative polymerase B tree, similar trends were shown as with the putative polymerase A tree in that two major clades and one minor clade were formed also containing phages of a single morphological type in a common type of bacterium (i.e, Gram + or −). In this case, only 1 phage fell outside the three clades.

Separate phylogenetic trees for tailed-phage families

After seeing the relational patterns in the phylogenetic trees for the putative polymerase types, a third attempt was made by focusing on phylogenetic trees for each of the individual phage families (Figs. 2–4). In this case, much clearer relational patterns were seen. In the family Podoviridae (Fig. 2) there were two distinct clades, one for putative polymerase A in Gram− bacteria and one for putative polymerase B in Gram+ bacteria. Similarly, in the Myoviridae (Fig. 3) and Siphoviridae (Fig. 4), there were two major clades grouped according to putative polymerase type in a particular bacterial type. Curiously, the A & B putative polymerases did not always group with the same bacterial type in every phage family.

Integrated phylogenetic tree

Based on the relational patterns seen in the previous phylogenetic trees and considering that the phage types in the Archaebacteria may represent types near the root of the phage putative polymerase tree, we constructed a hypothetical tree (Fig. 5). The hypothetical tree suggests that an ancestral phage DNA

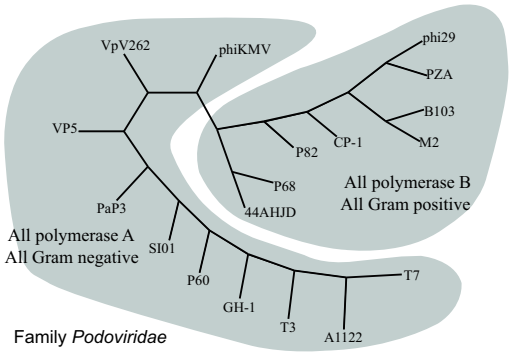

Fig. 2 Phylogenetic tree of phages in the family Podoviridae with reported putative polymerases showing a clean separation of the polymerase A types into a clade of Gram− bacteria and the polymerase B types into Gram+ bacteria.

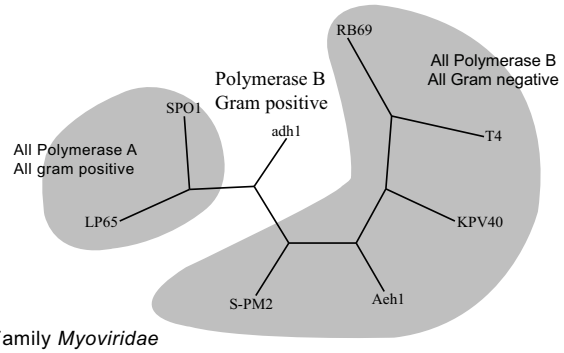

Fig. 3 Phylogenetic tree of phages in the family Myoviridae with reported putative polymerases showing two major clades with a clean separation of polymerase A types into a clade of Gram+ bacteria and polymerase B types into Gram− bacteria. There is one intermediate type adh1 with polymerase B in a Gram+ bacterium.

polymerase branched into A and B types early in the evolution of the 3 tailed-phage families in a more or less random fashion with respect to their Gram+ and Gram− hosts, such that subsequent evolution led to a single polymerase type for each phage in each bacterial type. This hypothesis would accommodate the variable patterns seen in the individual phage family trees.

There are some ramifications and predictions arising from this hypothetical tree. For example, it predicts that phages in the family Podoviridae with Gram+ bacterial hosts will use a DNA polymerase B, if indeed they still harbour their own DNA polymerase. Similar arguments would apply to the other

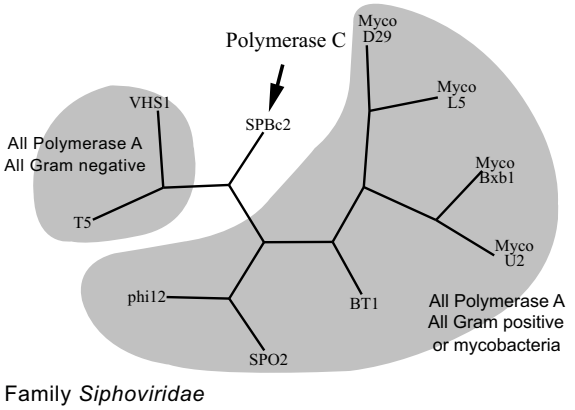

Fig. 4 Phylogenetic tree of phages in the family Siphoviridae with reported putative polymerases showing two major clades with a clean separation of polymerase A types into either Gram+ or Gram− bacteria, except for one intermediate phage with polymerase C in a Gram+ bacterium.

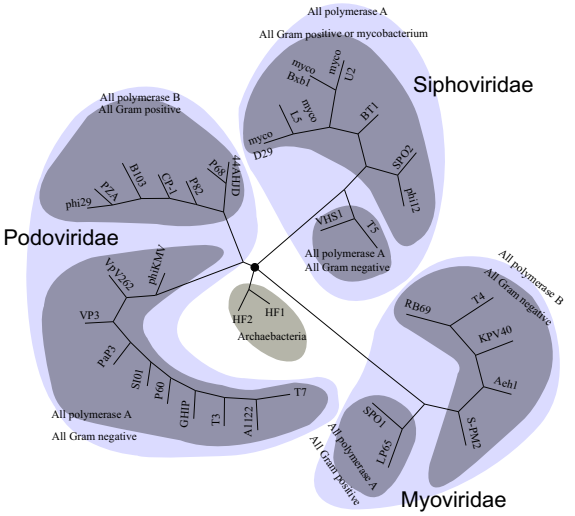

Fig. 5 A hypothetical phylogenetic tree interpolating from the information in previous figures and prepared by assembly of the trees for the individual phage families (Figs. 2–4). The tree is rooted (central black dot) for an ancestral phage type(s) that gave rise to separate branches for various phage families united by a common morphology. It proposes that these families then evolved separate sub-branches into the various bacterial groups. Thus, polymerase A types may be found in Gram− bacteria in one phage family but in Gram+ bacteria in another phage family. The same holds for polymerase A types. The branch lengths from the ancestral type(s) do not represent evolutionary distance but are intended to show theoretical evolutionary links. The phage Adh1 in the family Myoviridae is included in the diagram on the branch line between the ancestral type(s), before the sub-clades in the family because of its intermediate characteristics.

phage families. If the predictions hold, the principle revealed may apply also to other phage genes. The possibility could be tested by comparing related genes from a particular phage family to see whether they also fall into two distinct groups according to bacterial type. We made a preliminary attempt at this but could not find a suitable target gene.

In summary, our DNA polymerase comparisons provide no strong support one way or the other for inclusion of VHS1 in the family Siphoviridae. Its closest neighbours in the overall DNA polymerase A tree were myophages and a podophage. On the other hand, it did cluster with the siphophage T5 when included in the Siphoviridae tree (Fig. 4). It is unfortunate that more sequences were not available to make this tree more robust. Unfortunately, the only other VH phage associated with shrimp (*Vibrio harveyi* Myoviridae-like or VHML phage) described from Australia by Oakey and Owens[2] and Oakey et al[22] does not have a polymerase gene and could not be included in our analysis. Nor does it have other significant sequence homology with VHS1[11].

In the separate phage family trees, there were two 'odd men out', SPBc2 in the siphophage tree and adh1 in the myophage tree. SPBc2 is classified as a siphovirus but is unique in possessing a dnaE polymerase from the polymerase C family[23,24]. Polymerase C is not a common polymerase type and is usually reported only from bacteria[25]. Indeed, the whole polymerase sequence of SPBc2 shows very high homology (E value 0) to the DNA polymerase of its host bacterial species *Bacillus subtilis*. This suggests that one acquired its polymerase from the other and our siphophage tree would suggest that SPBc2 acquired its polymerase from its host.

With respect to adh1, its DNA polymerase B has also been called a Poxvirus polymerase[26,27]. In our overall polymerase B tree, it was most closely associated with polymerases of P68 and 44AHJD that are both members of the Podoviridae. Somewhat similar to SPBc2 above, the polymerase of adh1 shares reasonable homology with its Gram positive host *Lactobacillus gasseri* (E value 10^{-43}) suggesting that it might be another example of a phage that derived its polymerase gene from its host – in this case an ancestral host.

By removing these 2 'odd men out' from the 3 phage family trees, there appears to be a trend towards a characteristic single polymerase type (A or B) for the Gram positive and Gram negative bacteria in each family. As phage genome sequences accumulate in the database, it will be interesting to see whether this pattern holds and whether it will eventually include

other phage genes.

Acknowledgements: This work is supported by a research grant of the National Centre of Genetic Engineering and Biotechnology, Thailand.

REFERENCES

1. Ruangpan L, Danayadol Y, Direkbusarakom S, Sriurairatana S, Flegel TW (1999) Lethal toxicity of *Vibrio harveyi* to cultivated *Penaeus monodon* induced by a bacteriophage. *Dis Aquat Org* **35**, 195–201.
2. Oakey HJ, Owens L (2000) A new bacteriophage, VHML, isolated from a toxin-producing strain of *Vibrio harveyi* in tropical Australia. *J Appl Microbiol* **89**, 702–9.
3. Khemayan K, Pasharawipas T, Puiprom O, Sriurairatana S, Suthienkul O, Flegel TW (2006) Unstable lysogeny and pseudolysogeny in *Vibrio harveyi* siphovirus-like phage 1. *Appl Environ Microbiol* **72**, 1355–63.
4. Cheetham BF, Katz ME (1995) A role for bacteriophages in the evolution and transfer of bacterial virulence determinants. *Mol Microbiol* **18**, 201–8.
5. Rankin S, Platt DJ (1995) Phage conversion in *Salmonella entritica* serotype Enteritidis: implications for epidemiology. *Epidemiol Infect* **114**, 227–36.
6. Faruque SM, Asadulghani Saha MN, Alim AR, Albert MJ, Islam KM, Mekalanos JJ (1998) Analysis of clinical and environmental strains of nontoxigenic *Vibrio cholerae* for susceptibility to CTXΦ: Molecular basis for origination of new strains with epidemic potential. *Infect Immun* **66**, 5819–25.
7. Yoshizawa Y, Sakurada J, Sakurai S, Machida K, Kondo I, Masuda S (2000) An exfoliative toxin A-converting phage isolated from *Staphylococcus aureus* strain ZM. *Microbiol Immunol* **44**, 189–91.
8. Proctor LM, Fuhrman JA (1990) Viral mortality of bacteria and cyanobacteria. *Nature* **343**, 60–2.
9. Braithwaite DK, Ito J (1993) Compilation, alignment, and phylogenetic relationships of DNA polymerases. *Nucleic Acids Res* **21**, 787–802.
10. Van Regenmortel MHV, Fauquet CM, Bishop DHl Carstens EB, Estes MK, Lemon SM, Maniloff J, Mayo MA, McGeoch DJ, Pringle CR, Wickner RB (2000) *Virus Taxonomy: Seventh Report of the International Committee on Taxonomy of Viruses*, Academic Press, California.
11. Pasharawipas T, Thaikua S, Sriurairatana S, Ruangpan L, Direkbusarakum S, Manopvisetcharean J, Flegel TW (2005) Partial characterization of a novel bacteriophage of *Vibrio harveyi* isolated from shrimp culture ponds in Thailand. *Virus Res* **114**, 63–9.
12. Ito J, Braithwaite DK (1991) Compilation and alignment of DNA polymerase sequences. *Nucleic Acids Res* **19**, 4045–57.
13. Filee J, Forterre P, Sen-Lin T, Laurent J (2002) Evolution of DNA polymerase families: Evidences for multiple gene exchange between cellular and viral proteins. *J Mol Evol* **54**, 763–73.
14. Yang W (2005) Portraits of Y-family DNA polymerase. *FEBS Lett* **579**, 868–72.
15. Scarlato V, Gargano S (1992) The DNA polymerase-encoding gene of *Bacillus subtilis* bacteriophage SPO1. *Gene* **118**, 109–13.
16. Stewart CR, Gaslightwala I, Hinata K, Krolikowski KA, Needleman DS, Peng AS, Peterman MA, Tobias A, Wei P (1998) Genes and regulatory sites of the "host-takeover module" in the terminal redundancy of *Bacillus subtilis* bacteriophage SPO1. *Virology* **246**, 329–40.
17. Leavitt MC, Ito J (1989) T5 DNA polymerase: structural-functional relationships to other DNA polymerases. *Proc Natl Acad Sci USA* **86**, 4465–9.
18. Ford ME, Sarkis GJ, Belanger AE, Hendrix RW, Hatfull GF (1998) Genome structure of mycobacteriophage D29: Implications for phage evolution. *J Mol Biol* **279**, 143–64.
19. Mediavilla J, Jain S, Kriakov J, Ford ME, Duda RL, Jacobs WR, Hendrix RW, Hatfull GF (2000) Genome organization and characterization of mycobacteriophage Bxb1. *Mol Microbiol* **38**, 955–70.
20. Bernard A, Blanco L, Lazaro JM, Martin G, Salas M (1989) A conserved 3′-5′ exonuclease active site in prokaryotic and eukaryotic DNA polymerases. *Cell* **59**, 219–28.
21. Tang S, Nuttall S, Dyall-Smith M (2004) Haloviruses HF1 and HF2: evidence for a recent and large recombination event. *J Bacteriol* **186**, 2810–7.
22. Oakey HJ, Cullen WR, Owens L (2002) The complete nucleotide sequence of the *Vibrio harveyi* bacteriophage VHML. *J Appl Microbiol* **93**, 1089–98.
23. Brucks I, Goodman MF, O'Donnell M (2003) The essential C family DnaE polymerase is error-prone and efficient at lesion bypass. *J Biol Chem* **278**, 44361–8.
24. Ito J, Huang Y, Parekh T (1998) Studies on the cyanobacterial family C DNA polymerase. *FEMS Microbiol Lett* **158**, 39–43.
25. Kornberg A, Baker T (1992) *DNA Replication*, Freeman, New York.
26. Argos P (1988) A sequence motif in many polymerases. *Nucleic Acids Res* **16**, 9909–16.
27. Takemura M (2001) Poxviruses and the origin of the eukaryotic nucleus. *J Mol Evol* **52**, 419–25.