Available online at www.sciencedirect.com
ScienceDirect
Virology 364 (2007) 237–243
VIROLOGY
www.elsevier.com/locate/yviro

Genome of the *Acidianus* bottle-shaped virus and insights into the replication and packaging mechanisms ☆

Xu Peng [a,*], Tamara Basta [b], Monika Häring [b], Roger A. Garrett [a], David Prangishvili [b]

[a] *Centre for Comparative Genomics, Department of Molecular Biology, Biocenter, Copenhagen University, Ole Maaløes vej 5, DK-2200 Copenhagen N, Denmark*
[b] *Molecular Biology of the Gene in Extremophiles Unit, Institute Pasteur, rue Dr. Roux 25, 75724 Paris Cedex 15, France*

Received 9 February 2007; accepted 5 March 2007
Available online 6 April 2007

Abstract

The *Acidianus bottle-shaped virus*, ABV, infects strains of the hyperthermophilic archaeal genus *Acidianus* and is morphologically distinct from all other known viruses. Its genome consists of linear double-stranded DNA, containing 23,814 bp with a G+C content of 35%, and it exhibits a 590-bp inverted terminal repeat. Of the 57 predicted ORFs, only three produced significant matches in public sequence databases with genes encoding a glycosyltransferase, a thymidylate kinase and a protein-primed DNA polymerase. Moreover, only one homologous gene is shared with other sequenced crenarchaeal viruses. The results confirm the unique nature of the ABV virus, and support its assignment to the newly proposed viral family the *Ampullaviridae*. Exceptionally, one region at the end of the linear genome of ABV is similar in both gene content and organization to corresponding regions in the genomes of the bacteriophage ϕ29 and the human adenovirus. The region contains the genes for a putative protein-primed DNA polymerase, and a small putative RNA with a predicted secondary structure closely similar to that of the prohead RNA of bacteriophage ϕ29. The apparent similarities in the putative mechanisms of DNA replication and packaging of ABV to those of bacterial and eukaryal viruses are most consistent with the concept of a primordial gene pool as a source of viral genes.

Keywords: Viral genome; Protein-primed DNA polymerse; Terminal protein; Prohead RNA

Introduction

Studies of viral diversity in geothermally heated aquatic environments have resulted in the discovery of a plethora of double-stranded DNA viruses infecting hyperthermophilic members of the Crenarchaeota, a kingdom in the domain Archaea. The morphological diversity of these viruses is extraordinary, and includes forms that have never been observed in any other habitat. They have been assigned to six new viral families including the spindle-shaped *Fuselloviridae*, filamentous *Lipothrixviridae*, rod-shaped *Rudiviridae*, droplet-shaped *Guttaviridae*, spherical *Globuloviridae* and two-tailed *Bicaudaviridae* (reviewed by Prangishvili and Garrett, 2005; Prangishvili et al., 2006a). One of the most unusual morphotypes is represented by virions of the *Acidianus* bottle-shaped virus, ABV, which has been proposed as the first member of a new family, the *Ampullaviridae*. This virus was isolated from a hot acidic spring (87 to 93 °C, pH 1.5) at Pozzuoli, Italy and was propagated in "*Acidianus convivator*", a hyperthermophilic strain isolated from the same environment (Häring et al., 2005a).

The bottle-shaped virion of ABV has an overall length of about 230 nm and a width varying from 75 nm at the broad end to about 4 nm at the pointed end (Häring et al., 2005a). It carries an envelope encasing a funnel-shaped core composed of a tip resembling a "bottle-neck" and a torroidally supercoiled nucleoprotein (see Fig. S1). The base of the "bottle-neck" appears to be connected to the nucleoprotein filament. Virions were observed attached to membrane vesicles of host cells via their tips and the latter are inferred, therefore, to be involved in viral adsorption and channeling of viral DNA into host cells. About 20 short fine filaments, of unknown function, are inserted into a disk or ring at the broad end of the virion (Häring et al., 2005a).

☆ Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under Accession No. EF432053.

* Corresponding author. Fax: +45 35322040.
E-mail address: peng@mermaid.molbio.ku.dk (X. Peng).

0042-6822/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.virol.2007.03.005

In order to gain more insight into the biology of this exceptional virus, we have sequenced and analyzed the genome and provide some insights into possible replication and packaging mechanisms.

Results and discussion

Genome sequencing and genome structure

The genome was earlier shown to constitute double-stranded DNA (Häring et al., 2005a). However, since viral yields were very low, and only about 200 ng purified DNA was available for sequencing, we adopted a two-step genome sequencing strategy described in Materials and methods which yielded a single linear contig of 22 kb, 1.8 kb smaller than the size estimated from different restriction enzyme digests (Häring et al., 2005a, data not shown). We then sequenced further out from the contig ends by primer walking on amplified DNA which yielded about 1.8 kb of additional sequence beyond which sequence reads invariably terminated. Subsequently, the termini of total viral DNA extracted from ABV-infected host cells were tailed with poly dA (see Materials and methods) and the terminal regions were amplified by PCR and sequenced. This yielded a further 10 bp at each end. The total sequence obtained was 23,814 bp, consistent with the size estimated from restriction enzyme digests. The linear genome, with a 35% G+C content, carries a 590 bp inverted terminal repeat (ITR).

In order to confirm the linearity of the genome and to test for its possible circularization during replication, PCR experiments were performed on viral DNA, and on total DNA extracted from ABV-infected host cells, with different primer pairs annealing near to each end of the genome. However, none of these produced an amplified product (data not shown) and we inferred, therefore, that the ABV genome is linear in mature virions and that a circularization step is unlikely to occur during replication.

Gene content

Fifty seven ORFs ranging in size from 37 to 653 aa were predicted. Start codons were predominantly AUG and stop codons were mainly UAA. TATA-like promoter motifs and/or Shine-Dalgarno motifs were assigned as described earlier (Torarinsson et al., 2005). A genome map illustrating the sizes and directions of the genes is presented in Fig. 1. The predicted gene distribution shows a clear break around position 8.5 kb. All ORFs except three, ORF247, ORF53a and ORF156, are located on one strand in the left one third while all except one (ORF61b) are located on the opposite strand in the right two thirds (Fig. 1). The genome is compactly organized with little space between the ORFs, and 15 pairs of ORFs show small overlaps. ORF103, ORF257 and ORF653 are exceptional in that they each contain a putative internal start codon preceded by a Shine-Dalgarno motif and these internal ORFs were annotated together with the remainder of the genome (Table 1). In addition to the EMBL/GenBank database, the annotated genome is also available online in The Sulfolobus Database (Brügger, 2007).

About 20 ORFs are preceded by putative TATA-like motifs although for a very A+T-rich genome (65%) these can be difficult to identify unambiguously. Forty three ORFs are preceded by putative Shine-Dalgarno motifs and 13 ORFs are followed by putative T-rich terminators (Table 1). Thirty eight ORFs are arranged in putative operons, three of which carry 7 to 9 genes and, moreover, 27 genes located between positions 10 kb to 21 kb appear to constitute a superoperon (Table 1, Fig. 1). Fifteen genes are predicted to generate transcripts that are either leaderless or carry very short leaders. This correlates with the results of previous analyses showing that single genes and first genes of putative operons in crenarchaeal chromosomes and extrachromosomal elements, frequently produce leaderless transcripts (Torarinsson et al., 2005).

Only three ORFs could be assigned unambiguous functions based on homolog searches in public sequence databases (Table 2). ORF315 and ORF156 yielded significant matches with a glycosyltransferase and thymidylate kinase, respectively, while ORF653 showed significant sequence similarity to the family B DNA polymerases, with the best matches to protein-primed polymerases (see below). Several ORFs show recognizable structural motifs. ORF61c exhibits a periodic repetition of leucine residues from position 33 to 54, consistent with a leucine zipper motif and could, therefore, be involved in transcriptional regulation. The same pattern also occurs in ORF53a, ORF92 and ORF150a (Table 2). Moreover, three adjacent ORFs, ORF112, ORF166 and ORF346, carry predicted transmembrane helical segments and are likely to constitute membrane-binding proteins (Table 2). ORF346 also contains an EGF-like domain which commonly occurs in extracellular domains of membrane-bound proteins or in secreted proteins (Davis, 1990). Possibly, ORF346 constitutes a virion protein that either interacts with a host membrane protein or is a transmembrane protein which facilitates cellular infectivity and/or extrusion, possibly facilitated

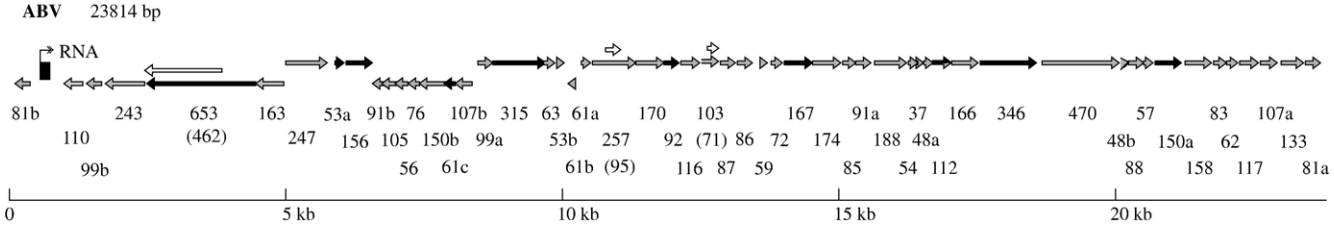

Fig. 1. Genome map of ABV showing the location, size and direction of the putative genes. Dark arrows indicate ORFs with assigned functions or recognizable structural motifs, while genes of unknown structure and function are denoted by gray arrows. Three internal ORFs are denoted by empty arrows and their amino acid sizes are given in brackets.

Table 1
Summary of predicted ORFs and operons in the ABV genome

ORF	TATA-like box	S-D motif	Distance to upstream ORF	Terminator	Operon
Forward strand					
247	AATTAAAA(−32)		–	CTTTTTT	
53a	ATTATA(−44)	GGGGA(−6)	−23	CTTTTTTCTT	
156	TTAAATA(−23)	–	33		
99a	ATTTTT(−26)	–	–		Operon
315		GGGGT(−5)	−4		
63	–	GGCG(−4)	−14		
53b	–	GGAGA(−8)	3	TCTTCTTTT	
61a	TTATTA(−35)	GTGA(−6)	–		Operon
257	–	GTGA(−6)	31		
95 (in)	–	GATGA(−5)	–		
170	–	AGATGA(−5)	5		
92	TAAATA(−38)	AGGTGTA(−5)	1		Operon
116	–	GGTGA(−5)	10	TCTCTTCT	
103	–	AGGT(−7)	51		
71 (in)	–	AGGTGA(−5)	–		
87	–	–	13		
86	ATTTAT(−32)		44		Operon
59	–	AGGTGA(−8)	15		
72	–	AGGTGA(−6)	42		
167	–	AGGTGA(−6)	32		
174	–	GGGTGA(−5)	5		
85		AGGTGA(−6)	22		
91a		AGGTGA(−6)	1	TCTTTTTTC	
188	–	GATGA(−6)	58	TTTTTT	
54	–	–	33		
37	–	GTGA(−6)	−4		Operon
48a	–	GGCGA(−6)	11		
112	–	AGAAGA(−7)	−4		
166	–	GAAGA(−9)	6		
346	–	AGGTGA(−5)	10		
470	–	GGTAA(−4)	83		
48b	–	AGGTGAG(−5)	21		
88	–	GTGA(−6)	−4		
57	–	GAGGT(−7)	−13		
150a	TTAAATA(−30)	–	52	TTTTTTCTTTCTT	
158	TTTAAATA(−32)	–	79	TTTTTTCTTTTT	
83	AAATTT(−35)	GGTGA(−6)	40		Operon
62	–	GGAGT(−4)	−1		
117	–	AGGTGA(−7)	20		
107a	–	AGGTGA(−5)	17	TCTTTTTCT	
133	TTTTAA(−30)	–	56	TTTTTT	
81a	TTTAAA(−31)	–	61		
Reverse strand					
61b	AATAAT(−40)	–	–		
107b	ATTTAT(−26)	–	–		Operon
61c	–	GGTGA(−5)	−8		
150b	–	GGAGA(−6)	−8		
76	–	GGAGA(−6)	−4		
56	–	AGGTGG(−5)	0		
105	–	GTTGA(−5)	−32		
91b	–	GGTGA(−6)	−92	TCTTTTT	
163	TTTAATT(−23)	–	–		
653	TTTAAA(−22)	–	13		Operon
462 (in)	–	GATG (−6)			
243	–	GGTGA(−5)	19	TCTTTTTTCCTT	
99b	AAATTTA(−27)	–	54		Operon
110	TTAAATA(−32)	GGTGA(−4)	57	CTTTTC	
81b	TTTTAAA(−31)	–	–		

Table 2
Functions assigned to ABV ORFs

ORF	Predicted function	Database/Search tool	Expectation value
156	Thymidylate kinase	pfam02223	6.1e−5
315	Glycosyl transferase	BlastP/Pfam00534	3e−10
653	DNA polymerase	BlastP/Pfam03175	9e−10
53a	Leucine zipper	Pdoc00029	–
61c	Leucine zipper	Pdoc00029	–
92	Leucine zipper	Pdoc00029	–
150a	Leucine zipper	Pdoc00029	–
112	3 TM helices	TMHMM	–
167	3 TM helices	TMHMM	–
346	5 TM helices	TMHMM	–
	EGF-like domain	Pdoc00021	–

Best BlastP match to the glycosyltransferase was with the crenarchaeal filamentous virus AFV2 and match with the DNA polymerase was with the mitochondrion of *A. aegerita*. – No expectation values available.

by ORF112 and ORF166. The C-terminal regions of ORF346 and the neighboring ORF470 exhibit low complexity sequences as have also been observed in a few large ORFs in other crenarchaeal viral genomes and especially that of the bicaudavirus ATV (Häring et al., 2005b; Prangishvili et al., 2006b).

Putative protein-primed DNA polymerase

ABV is the only member of known crenarchaeal viruses for which a DNA polymerase gene (ORF653) has been identified. Moreover, the putative ABV polymerase shows significant sequence similarity to protein-primed DNA polymerases which constitute a subfamily of the family B DNA polymerases (reviewed by Salas, 1991). These enzymes are exclusively linked with linear ds DNA genomes carrying ITRs with covalently linked terminal proteins and are encoded in some mitochondria, plasmids and viruses. BlastP analysis against NCBI revealed the mitochondrial polymerase of *Agrocybe aegeri* as the best match ($E=10^{-5}$) with sequence similarity highest in the central region of the gene. In contrast, pairwise sequence comparisons using Blast 2 showed that the putative polymerases of the euryarchaeal *Salterproviruses* His1 and His2 (Bath et al., 2006) aligned along the entire sequence of ORF653 albeit at a lower similarity level. A partial sequence alignment of ORF653 and the genes of the two haloarchaeal viruses is given in Fig. 2. The polymerase sequences of phage ϕ29 and human adenovirus, which are the best characterized of this subfamily, are included in the alignment (Fig. 2).

The alignment clearly illustrates that ORF653 belongs to the protein-primed subfamily of the B DNA polymerases. It carries the characteristic exonuclease domains (Exo I, II and III) in the N-terminal region and the conserved polymerizing domains (Pol I, IIa, IIb and III) towards the C-terminus (Bois et al., 1999). Moreover, the two insertion regions TPR-1 and TPR-2 which are characteristic of these protein-primed polymerases are also present (Fig. 2) (Dufour et al., 2000; Rodriguez et al., 2005). TPR-1 is similar in size (about 50 aa) to the corresponding region of all known protein-primed DNA polymerases, including that of the human adenovirus (Dufour et al., 2000) while TPR-2, located between motifs pol IIa and IIb, is smaller than other observed inserts which vary in size from 28 aa (phage ϕ29) to 120 aa (adenovirus type 2) (Bois et al., 1999). TPR-1 of the phage ϕ29 polymerase interacts with the terminal priming protein (Dufour et al., 2000) while TPR-2 is required for the strand displacement and high processivity

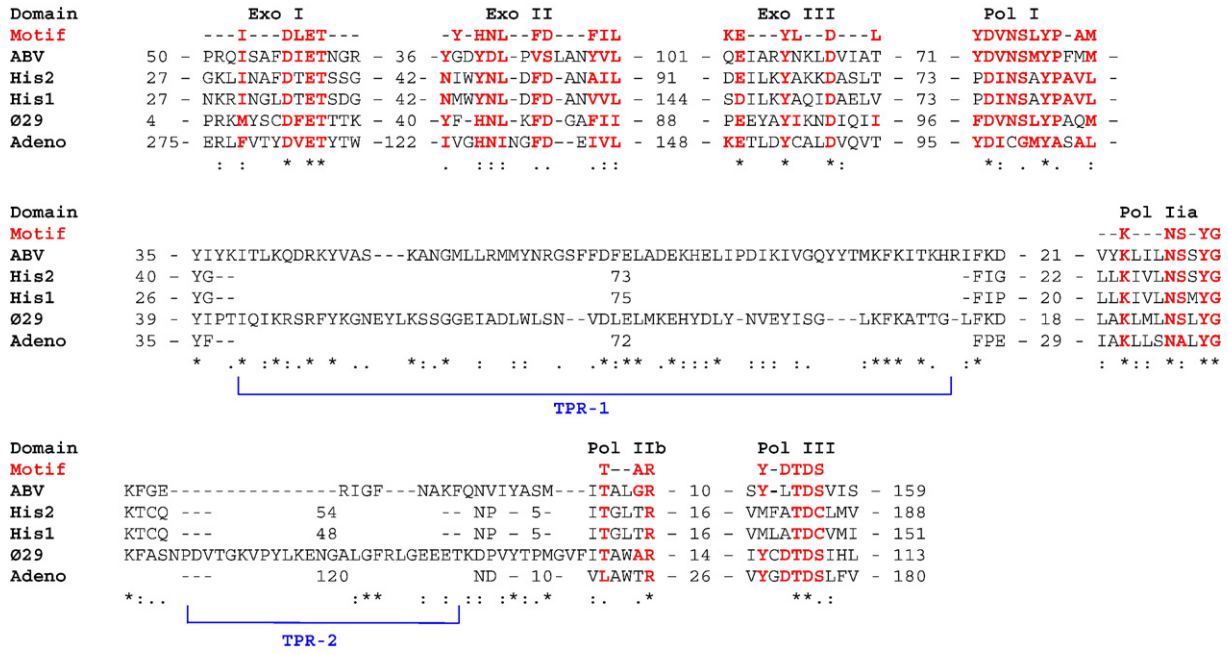

Fig. 2. Sequence alignment between ORF653 of ABV virus and the DNA polymerases encoded by the euryarchaeal viruses His1 and His2, phage ϕ29 and human adenovirus 5, showing two regions (TPR I and TPR II) which are specific for all known protein-primed DNA polymerases. TPR I and TPR II sequences are aligned only for ABV and phage ϕ29 due to low sequence conservation. The conserved motifs implicated in exonuclease activity (Exo I, II, III) and polymerization (Pol I, IIa, IIb, III and IV) are shown and numbers denote the numbers of amino acids between the motifs.

activities of the polymerase (Rodriguez et al., 2005). Although the function of the more conserved TPR-1 is likely to be conserved for all protein-primed DNA polymerases, it remains unclear whether this applies to the more variable TPR-2.

It is likely that ABV encodes a terminal protein and this inference is supported by the finding that treatment of viral DNA by proteinase K results in a significant decrease of material at the phenol–water interface during extraction, consistent with the presence of tightly bound protein (data not shown). However, in contrast to the protein-primed DNA polymerases, the terminal proteins show little conservation of either sequence or size. Nevertheless, the gene context of the terminal protein is conserved. Thus, in bacteriophages ϕ29, PRD1, GA-1 and CP-1 (Accession numbers P03681, P09009, X96987 and Z47794), as well as in human adenoviruses, the gene is always located immediately upstream of the DNA polymerase gene and the two genes are cotranscribed in all studied viruses (Fig. 4) (Meijer et al., 2001). Moreover, the two genes are sometimes fused (e.g., Takeda et al., 1996). Thus, ORF163 which is located 13 bp upstream from ORF653 is a strong candidate for encoding a terminal protein.

Putative prohead RNA

Analysis of the ABV genome revealed a candidate gene which is likely to produce an RNA molecule that is involved in DNA packaging, by analogy to the packaging system of the bacteriophage phi29. In the latter, a virus-encoded 174-nt RNA, termed prohead RNA, is essential for packaging of linear ds DNA into a preformed capsid (reviewed by Guo, 2005). Prohead RNA actively participates in DNA translocation by binding to the procapsid and ATP and interacting with a packaging protein (Guo, 2005). The prohead RNA is encoded adjacent to one ITR in the phage ϕ29 genome and it exhibits several secondary structural regions which are conserved in the RNAs from all known ϕ29-related bacteriophages (reviewed by Meijer et al., 2001).

Examination of the corresponding region in the ABV genome, revealed a 600-bp section, lacking ORFs which is relatively G+C-rich at the center and is located close to the left ITR (Fig. 3). The predicted secondary structure for the putative 200-nt transcript is highly similar to that of the pRNAs of phages ϕ29 and CP-1 (Fig. 3) (Meijer et al., 2001). It carries six, co-structural, helices labeled A to F and the only clear structural difference occurs to the left of helix F where an extra hairpin-loop is present in the putative ABV RNA (Fig. 3). Transcription of the ABV RNA could be initiated at the promoter-like sequence ATTTAAT located 20 bp upstream. Thus, the conserved genomic position, the presence of a putative promoter motif, the high G+C content and the similar secondary structure of the transcript, all support the presence of a gene encoding a prohead RNA. Further support derives from a genomic region adjacent to the ITR of the terminal protein-encoding adenoviral DNA which was also shown to be important for packaging although no RNA molecule has yet been identified (Fig. 4) (Grable and Hearing, 1990). The similar gene location adjacent to ITRs of representative viruses assigned to different families deriving from the three domains of life is illustrated in Fig. 4.

Whether the putative RNA is indeed involved in packaging of the ABV genome remains to be demonstrated experimentally.

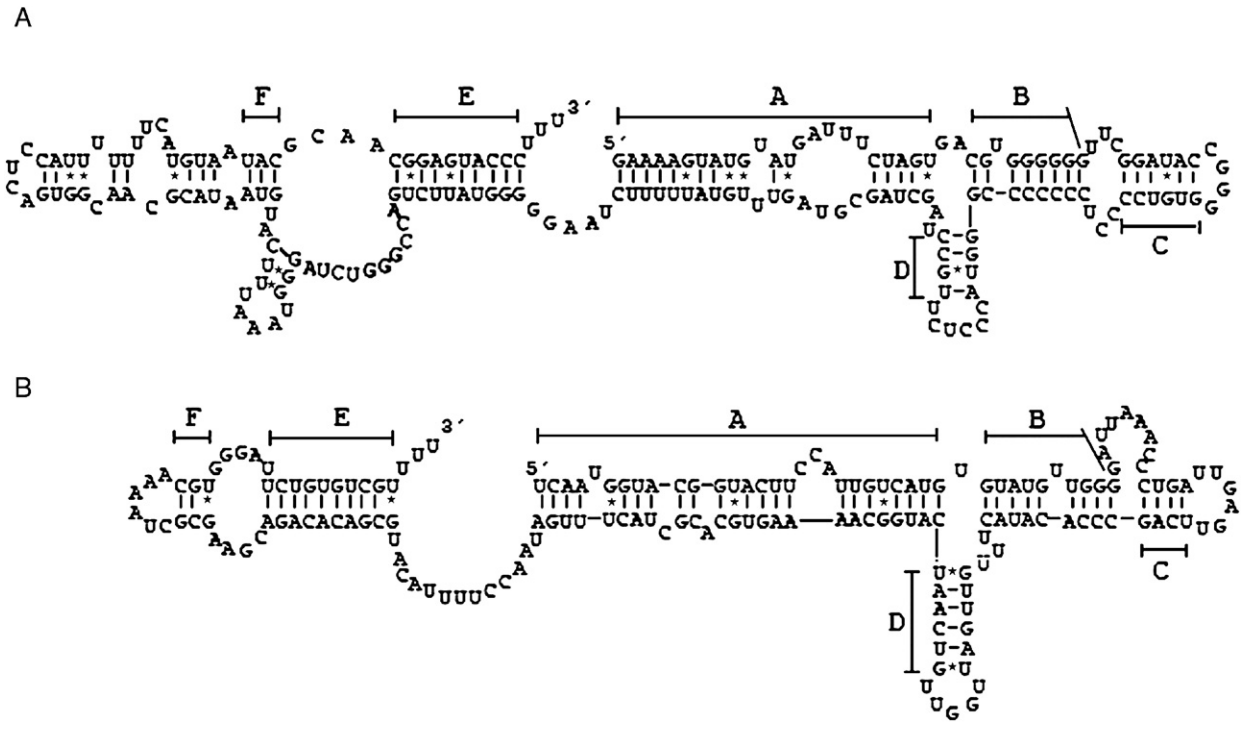

Fig. 3. (A) Predicted secondary structure of the putative prohead RNA involved in ABV packaging. (B) Secondary structure of the prohead RNA of phage ϕ29. The six helices conserved in the phage prohead RNAs are labeled A to F from each 5′ terminus (Bailey et al., 1990).

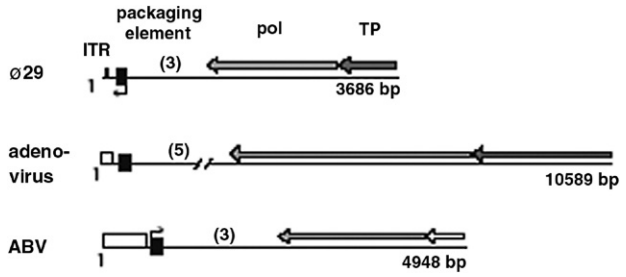

Fig. 4. Depiction of the similar gene organizations at the left ends of the linear genomes of phage ϕ29 (Accession No. V01121), adenovirus type 5 (Accession No. AC000008) and ABV where the ITRs of lengths 6, 103 and 590 bp, respectively, are represented by empty boxes. Genes encoding the polymerase (pol) and terminal protein (TP) are denoted by light and dark shaded gray arrows, respectively, while ORF163 upstream of the putative polymerase in ABV is shown by an empty arrow. The dark box denotes the genomic region indispensable for packaging in phage ϕ29 and adenovirus. Transcription directions (predicted for ABV) are shown by small arrows. The numbers of ORFs between the packaging elements and the protein-primed polymerase genes are given in brackets.

Nevertheless, if so this would represent a most unexpected similarity in packaging machineries in viruses with radically different morphologies.

Concluding remarks

Complex evolutionary issues are raised by the discovery of viruses which encode homologous protein-primed DNA polymerases but derive from different taxa and exhibit diverse morphotypes, including icosahedra, with ("*ϕ29-like viruses*") or without (*Adenoviridae*, *Tectiviridae*) short tails, spindles with one short tail (*Salterprovirus*) and now also the bottle-shaped ABV (*Ampullaviridae*). The sharing of these homologous genes is not consistent with close phylogenetic relationships between the viruses because they are extremely diverse in both morphotype and genomic properties. Moreover, horizontal gene transfer is an improbable explanation given the large evolutionary distances between hosts from the three domains of life and the low probability of cross-domain spreading. Thus, the incidence of genes encoding protein-primed DNA polymerases among viral taxa associated with the three domains of life is most consistent with the genes being acquired directly from a primordial gene pool, predating the divergence of the three domains of life. Furthermore, the prohead RNA of ϕ29 and its structural homolog in ABV could be sharing a primordial ancestor.

Materials and methods

Purification of viruses and preparation of viral DNA

Aerobic enrichment cultures were prepared from samples taken from a water reservoir in the crater of the Solfatara volcano at Pozzuoli, Italy, at 87–93 °C and pH 1.5–2, as described earlier (Häring et al., 2005a). They were grown at 75 °C, pH 3. Virions were purified by centrifugation in a CsCl buoyant density gradient and disrupted with 1% (w/v) SDS for 1 h at room temperature prior to extracting and precipitating DNA as described earlier (Häring et al., 2005a).

Sequencing of genomic DNA

Given that only 200 ng DNA was available for sequencing, first about 1 ng DNA was amplified *in vitro* to yield about 2 μg DNA using the GenomiPhi amplification kit (Amersham Biotech, Amersham, UK). A shot-gun library was then constructed from sonicated fragments in the size range 1.5 to 4 kbp, cloned into the *Sma*I site of pUC18. Sequencing produced a highly biased genome coverage (Peng et al., 2004) and many chimeric clones. However, on average a 20-fold sequence coverage was obtained for the genome and this enabled us to eliminate chimeric clones from the library on the basis of their lower frequency in contigs. Next, the validity of the sequence assembly obtained from the amplified DNA library was tested by mixing 50 ng of the original ABV DNA with 1 μg DNA extracted from *Acidianus* lipothrixviruses (Vestergaard G. et al., unpublished) and a mixed shotgun library was prepared which included larger clone inserts (2 to 6.5 kbp). The sequenced ABV clones were then assembled and the contigs generated were compared with those obtained from the first library. The results demonstrated that the genome assemblies obtained from the two libraries matched exactly. Moreover, sequence regions not covered by the second library were obtained after PCR amplification of these regions from original viral DNA and/or by primer walking on the large insert clones. In addition, sequences of a few clones at the left terminus were extended by primer walking which yielded a final contig of about 22 kb.

The 3′-OH ends of ABV were tailed with poly dA (Forstemann et al., 2000) using the terminal deoxynucleotidyl transferase and total DNA of ABV-infected host cells as a template using the protocol provided by the manufacturer (Amersham Biosciences, Amersham, UK). A 45-cycle PCR step was then performed to amplify the poly dA tailed ends using a 30-mer oligo dT and a primer annealing to the terminal region of ABV. This produced a 700-bp PCR product that was cloned with Zero Blunt TOPO PCR cloning kit (Invitrogen, Paisley, UK) and sequenced.

Sequence analyses

Open reading frames were located and adjusted using putative Shine-Dalgarno motifs and TATA-like promoter motifs (Torarinsson et al., 2005) in the Artemis V8 program (Sanger Institute, Hinxton, UK). BlastP searches were performed against the EMBL/NCBI database (Altschul et al., 1997). SMART and MotifScan were used to detect conserved domains or motifs. Coiled coils, secondary structures and transmembrane helices were detected by programs in ExPASy Proteomics Tools (http://www.expasy.org/tools/).

Acknowledgments

We thank Simonetta Gribaldo, Jonathan Filée and Patrick Forterre for their helpful discussions. The research was in part

supported by the Danish Natural Science Research Council, the European Union (grant QLK3-2000-00649), the Grundforskning Centre for Comparative Genomics and the Agence Nationale de Recherche, France (grant NT05-2_41674). X.P. received a Steno Stipend from the Danish Research Council for Natural Science, T.B. received Docteur Roux postdoctoral fellowship from the Institut Pasteur.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.virol.2007.03.005.

References

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Bailey, S., Wichitwechkarn, J., Johnson, D., Reilly, B.E., Anderson, D.L., Bodley, J.W., 1990. Phylogenetic analysis and secondary structure of the *Bacillus subtilis* bacteriophage RNA required for DNA packaging. J. Biol. Chem. 265, 22365–22370.

Bath, C., Cukalac, T., Porter, K., Dyall-Smith, M.L., 2006. His1 and His2 are distantly related, spindle-shaped haloviruses belonging to the novel virus group, *Salterprovirus*. Virology 350, 228–239.

Bois, F., Barroso, G., Gonzalez, P., Labarere, J., 1999. Molecular cloning, sequence and expression of *Aa-polB*, a mitochondrial gene encoding a family B DNA polymerase from the edible basidiomycete *Agrocybe aegerita*. Mol. Gen. Genet. 261, 508–513.

Brügger, K., 2007. The Sulfolobus database. Nucl. Acids Res. 35, D413–D415.

Davis, C.G., 1990. The many faces of epidermal growth factor repeats. New Biol. 2, 410–419.

Dufour, E.J., Mendez, J., Lazaro, M., de Vega, M., Blanco, L., Salas, M., 2000. An aspartic acid residue in TPR-1, a specific region of protein-priming DNA polymerases, is required for the functional interaction with primer terminal protein. J. Mol. Biol. 304, 289–300.

Forstemann, K., Hoss, M., Lingner, J., 2000. Telomerase-dependent repeat divergence at the 3 ends of yeast telomeres. Nucleic Acids Res. 28, 2690–2694.

Grable, M., Hearing, P., 1990. Adenovirus type 5 packaging domain is composed of a repeated element that is functionally redundant. J. Virol. 64, 2047–2056.

Guo, P., 2005. Bacterial virus phi29 DNA-packaging motor and its potential applications in gene therapy and nanotechnology. Methods Mol. Biol. 300, 285–324.

Häring, M., Rachel, R., Peng, X., Garrett, R.A., Prangishvili, D., 2005a. Diverse viruses in hot springs of Pozzuoli, Italy, including a unique bottle-shaped archaeal virus ABV from a new family, the *Ampullaviridae*. J. Virol. 79, 9904–9911.

Häring, M., Vestergaard, G., Rachel, R., Chen, L., Garrett, R.A., Prangishvili, D., 2005b. Independent virus development outside a host. Nature 436, 1101–1102.

Meijer, W.J., Horcajadas, J.A., Salas, M., 2001. Phi29 family of phages. Microbiol. Mol. Biol. Rev. 65, 261–287.

Peng, X., Kessler, A., Phan, H., Garrett, R.A., Prangishvili, D., 2004. Multiple variants of the archaeal DNA rudivirus SIRV1 in a single host and a novel mechanism of genome variation. Mol. Microbiol. 54, 366–375.

Prangishvili, D., Garrett, R.A., 2005. Viruses of hyperthermophilic Crenarchaea. Trends Microbiol. 13, 535–542.

Prangishvili, D., Forterre, P., Garrett, R.A., 2006a. Viruses of the Archaea: a unifying view. Nat. Rev. Microbiol. 4, 837–848.

Prangishvili, D., Vestergaard, G., Häring, M., Aramayo, R., Basta, T., Rachel, R., Garrett, R.A., 2006b. Structural and genomic properties of the hyperthermophilic archaeal virus ATV with an extracellular stage of the reproductive cycle. J. Molec. Biol. 359, 1203–1216.

Rodriguez, I., Lazaro, J.M., Blanco, L., Kamtekar, S., Berman, A.J., Wang, J., Steitz, T.A., Salas, M., de Vega, M., 2005. A specific subdomain in phi29 DNA polymerase confers both processivity and strand-displacement capacity. Proc. Natl. Acad. Sci. U. S. A. 102, 6407–6412.

Salas, M., 1991. Protein-priming of DNA replication. Annu. Rev. Biochem. 60, 39–71.

Takeda, M., Hiraishi, H., Takesako, T., Tanase, S., Gunge, N., 1996. The terminal protein of the linear DNA plasmid pGKL2 shares an N-terminal domain of the plasmid-encoded DNA polymerase. Yeast 12, 241–246.

Torarinsson, E., Klenk, H.P., Garrett, R.A., 2005. Divergent transcriptional and translational signals in Archaea. Environ Microbiol. 7, 47–54.