

## Genome of low-temperature T4-related bacteriophage vB\_EcoM-VR7

Laura Kaliniene · Vytautas Klausas ·  
Aurelija Zajančkauskaitė · Rimas Nivinskas ·  
Lidija Truncaite

Received: 24 May 2011 / Accepted: 26 July 2011 / Published online: 10 August 2011  
© Springer-Verlag 2011

**Abstract** The complete genome sequence of the T4-related low-temperature *Escherichia coli* bacteriophage vB\_EcoM-VR7 was determined. The genome sequence is 169,285 bp long, with a G+C content of 40.3%. Overall, 95% of the phage genome is coding. It encodes 293 putative protein-encoding open reading frames (ORFs) and tRNA<sup>Met</sup>. More than half (59%) of the genomic DNA lacks significant homology with the DNA of T4, but once translated, 72% of the vB\_EcoM-VR7 genome (211 ORFs) encodes protein homologues of T4 genes. Overall, 46 vB\_EcoM-VR7 ORFs have no homologues in T4 but are derived from other T4-related phages, nine ORFs show similarities to bacterial or non-T4-related phage genes, and 27 ORFs are unique to vB\_EcoM-VR7. This phage lacks several T4 enzymes involved in host DNA degradation; however, there is extensive representation of the DNA replication, recombination and repair enzymes as well as the viral capsid and tail structural genes.

Phages that share an evolutionary history with bacteriophage T4 comprise one of the most diverse and common groups of bacteriophages found in nature [1, 2]. The genomes of about 40 T4-related phages have been sequenced and annotated over the last several years. Recently, using the ORF composition of the T4 genome as a criterion, these

phages were grouped into 23 different types, each one representing a pool of closely interrelated phages with low levels of sequence divergence between pool members [3].

About 90% of T4-related bacteriophages grow on *E. coli* or other enterobacteria [4]. The vast majority of such phages are from municipal wastewater or sewage, indicating that their natural habitat is the human or animal intestine [5]. Consequently, the optimum temperature for coliphage development is about 37 to 40°C, and it correlates closely with the optimum growth temperature of the host [6]. Based on the effect of temperature on the efficiency of plating, three physiological types of bacteriophages have been recognized—high-temperature (HT) phages plating at or above 25°C, low-temperature (LT) phages at or below 30°C, and mid-temperature (MT) phages in the range of 15 to 42°C [7]. *E. coli* phage T4, the archetype of the T4-type superfamily, is the typical representative of MT bacteriophages. It is noteworthy that only three *E. coli*-specific bacteriophages, vB\_EcoM-VR5, vB\_EcoM-VR7 and vB\_EcoM-VR20, which were described in our previous paper [8], have been assigned to the LT group to date.

In this paper, we present an analysis of the complete genome sequence of low-temperature T4-related bacteriophage vB\_EcoM-VR7 (subsequently referred to by its shorter common laboratory name VR7).

Bacteriophage VR7 was originally isolated from sewage as described in refs. [8, 9]. *E. coli* strain B<sup>E</sup> (*sup*<sup>0</sup>), a gift from Dr. L. W. Black, was used for phage propagation. All phage experiments were carried out as described in ref. 9. For the isolation of phage DNA, aliquots of phage suspension (10<sup>11</sup>–10<sup>12</sup> pfu/ml) were subjected to phenol/chloroform extraction and ethanol-precipitation as described in ref. [8]. Isolated phage DNA was subsequently digested with the restriction enzyme *ApoI* (*XapI*). The

**Electronic supplementary material** The online version of this article (doi:10.1007/s00705-011-1084-y) contains supplementary material, which is available to authorized users.

L. Kaliniene (✉) · V. Klausas · A. Zajančkauskaitė ·  
R. Nivinskas · L. Truncaite  
Vilnius University Institute of Biochemistry,  
Mokslininkų 12, 08662 Vilnius, Lithuania  
e-mail: laura.kaliniene@bchi.vu.lt

resulting 1–5 kb fragments were purified, cloned into the plasmid vector pCC1BAC<sup>TM</sup>, and introduced into *E. coli* by electroporation. Plasmids were purified, and the cloned inserts were sequenced at the Sequencing Centre (Institute of Biotechnology, Lithuania) using a BigDye<sup>®</sup> Terminator v3.1. Cycle Sequencing Kit (Applied Biosystems). The DNA was sequenced with six- to eightfold coverage, and the sequences were assembled, resulting in 64 separate contigs (~110 kb of genomic DNA), which were projected on to T4 and RB69 maps. The sequencing of PCR amplicons obtained using VR7 DNA-specific primers closed sequence gaps.

Analysis of the genome sequence of phage VR7 was performed using “Glimmer 2.02.RBS finder & TransTerm” (<http://nbc11.biologie.uni-kl.de>), Fasta-Protein, Fasta-Nucleotide, Fasta-Genome, BLAST2, PSI-Search, Transeq and ClustalW2 programs (available on <http://www.ebi.ac.uk>), Sequence editor (available on <http://www.fr33.net/seqedit.php>) and Geneious v5.0 (available from <http://www.geneious.com>). The genome alignment was performed using MegaBLAST (<http://phage.bioc.tulane.edu/>), and Artemis (<http://www.webact.org>). tRNAscan-SE 1.21 (<http://lowelab.ucsc.edu/tRNAscan-SE/>) was used to search for tRNAs. In order to annotate the genome of VR7, genes with best hit E-values ( $<10^{-4}$ ) to known genes of T4 were designated by the T4 gene name. Putative genes without T4 orthologs (with the exception of genes for IP) were designated by their ORF numbers, starting with *rIIA* as *ORF001*. The strand of each ORF is designated “w” for clockwise-transcribed genes and “c” for counter clockwise. The terminus of the genome was defined as the start of translation of the gene *rIIB*.

The complete genome sequence of bacteriophage VR7 is 169,285 bp long, with an overall G+C content of 40.3%, in comparison with 35.3% for T4. Overall, 95.4% of the VR7 genome is coding. It encodes 293 putative protein-encoding open reading frames (ORFs) and tRNA<sup>Met</sup>. In total, 281 VR7 ORFs were found to initiate with AUG, 7 with UUG, and 5 with GUG. Interestingly, no ORFs of T4 have been found to initiate from UUG [10].

Only 41% of the VR7 genomic DNA shared detectable nucleotide sequence homology with the DNA of T4, but once translated, 72% of the VR7 genome (211 ORFs) appeared to encode protein homologues of T4 genes, with levels of amino acid (aa) identity ranging between 27% and 97% (Supplementary Fig S1). Thus, based on their similarity to biologically defined T4 proteins, 111 ORFs of VR7 were given a functional annotation (Fig. 1). With no homologues in T4, 46 of VR7 ORFs derived from other T4-type phages and 9 ORFs showed similarities to bacterial (Supplementary Table S1) or non-T4-type phage genes (Supplementary Table S2). Meanwhile 27 ORFs of VR7 lacked any database matches. Homologues to the T4  $\alpha$ -gt,

$\beta$ -gt, SegA, SegB, SegC, SegD, SegE, I-TevI, I-TevII, I-TevIII, gp42, Ac, NrdG, NrdD, Arn, IPI, IPII, IPIII, Mrh as well as the T4-specific tRNAs appeared to be absent from VR7.

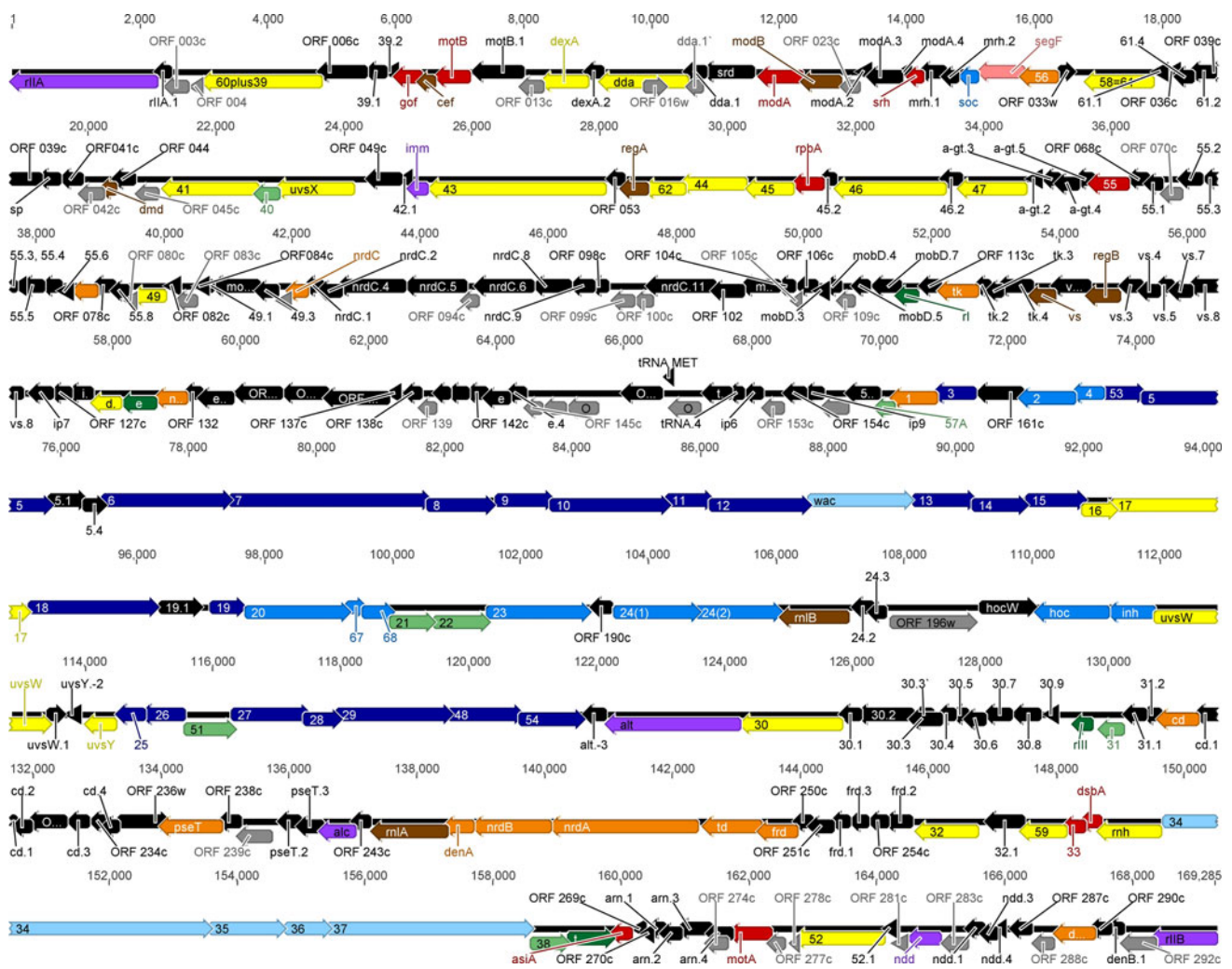
Restriction analysis of VR7 suggested that the DNA of VR7 and T4 might contain similar modifications [8]. This suggestion was also supported by the presence of homologues to the T4 Alc and DenB in the genome of VR7. However, as mentioned above, the gene for the dCMP hydroxymethylase (*g42*) as well as both glucosyltransferases ( $\alpha$ -gt and  $\beta$ -gt) were found to be absent from VR7. Therefore, it can be hypothesized that in the case of VR7, ORFs with yet unknown function may act as the functional analogues of these DNA-modifying enzymes.

Phage T4 encodes 15 homing endonucleases, accounting for ~11% of the coding potential of the genome [11, 12]. Five genes for putative homing endonucleases were present in VR7: two of the VR7 ORFs encoded proteins that shared 65% and 54% amino acid sequence identity with SegF and SegG of T4, while the remaining three putative mobile endonucleases of VR7 were homologous to MobE of RB43, RB32 and T4.

In spite of the differences mentioned, there was a distributed synteny between the genomes of VR7 and T4, with the two largest regions including the DNA replication-recombination-repair and virion structural genes.

The structural genes of VR7 were found to be organized into two separate clusters—a large, 51-kb cluster encoding most of the structural components of the virion (Supplementary Fig. S2) and a small, 10-kb cluster represented by five tail fiber genes (Supplementary Fig. S3). Both structural gene clusters were separated by a 27-kb segment of DNA transaction genes, as was also observed in T4 [13, 14]. The large virion structural module of VR7 contains no homologues of non-essential T4 genes *segC*, *segD*, *segE*, *ipI*, *ipII*, *ipIII* or 5.3. Gene 24 (the gene for a head vertex protein) of VR7 was found to be duplicated, and two *hoc*-like ORFs were detected on opposite strands of the genomic DNA. The duplicated *g24* has also been observed in JS98, although in the case of this phage, one *g24* copy was split into two open reading frames [15].

The structural proteins of VR7 shared amino acid sequence identity with T4, ranging from a minimum of 34% for gp11 (base plate wedge protein) to a maximum of 89% for gp21 (prohead core protein/protease). Overall, the head and tail proteins of VR7 and T4 showed the highest degree of conservation. In contrast, most of the tail fiber and base plate wedge proteins of VR7 showed a relatively low degree of sequence conservation with T4 (e.g., gp12 [45%], gpwac [36%] or gp9 [48%]). Moreover, several aa insertions similar to those in TuIb, which were found within the C-terminal region of gp37, suggested that VR7 and T4 might differ in their host cell recognition [16].



**Fig. 1 Functional genome map of bacteriophage VR7.** The coding capacity of the VR7 genome is shown. Functions are assigned according to the characterized ORFs in T4. The colour code for the online version of the article is as follows: yellow –DNA replication, recombination, repair and packaging; red – transcription; deep

brown – translation; orange –nucleotide metabolism; blue – head and neck proteins; dark blue – tail proteins; light blue – tail fibers; light green – chaperones/assembly; dark green – lysis; purple – host or phage interactions; dark grey – ORFs of unknown function; light grey – VR7 specific ORFs of unknown function

All of the essential DNA replication, recombination and repair (RRR) enzymes of T4 were found in VR7 as well. Eight of the VR7 RRR genes were detected within a single, contiguous gene cluster (>13 kb in length) (Supplementary Fig. S4). Meanwhile the genes encoding other RRR enzymes were distributed throughout the VR7 genome. The RRR gene order was relatively conserved between T4 [17] and VR7, with major differences mainly corresponding to non-essential genes that are not involved in DNA replication or recombination.

With the amino acid sequence identity ranging from a minimum of 55% for gp62 (sliding clamp loading subunit) to a maximum of 79% for UvsX (RecA-like protein), the RRR proteins of VR7 and T4 shared a high degree of sequence conservation within the sites that determine their catalytic properties. Meanwhile, the regions for protein-

protein interactions differed significantly. The same observations were made after comparison of the genomic DNA of T4 with the genomes of 11 different T4-related bacteriophages [17, 18].

Most of the proteins involved in the nucleotide metabolism in T4 [19] were found to have homologues in VR7. With the amino acid sequence identity ranging from 65% (Td) to 97% (NrdA), they showed a rather high degree of sequence conservation, suggesting that the VR7-directed DNA metabolism is quite similar to that of T4. However, the absence of genes for NrdD and NrdG implies that bacteriophage VR7 is less adapted to anaerobic growth conditions in comparison to T4 [10].

In summary, the comparison of the complete genome sequence of bacteriophage VR7 with the genome of the prototype phage T4 revealed substantial structural

differences. Many or some of these, however, are most probably related to the physiological properties of VR7 [8], which are rather distinct from those documented for T4. Unexpectedly, the analysis of the complete genome sequence of VR7 did not provide us with a straightforward answer as to why this phage belongs to the low-temperature class of T4-related bacteriophages. We believe that the comparison of the genome sequence of VR7 with the genome of another low-temperature phage, VR5 (which is being sequenced presently), will help to elucidate this matter.

As mentioned previously, the genomes of T4-related phages were grouped into 23 different types. The uniqueness of both the physiological properties and the genetic content of VR7 prevents us from classifying this phage into any of the established groups, suggesting that VR7 forms a distinct 24th type of T4-related genome.

### Nucleotide sequence accession number

The complete genome sequence of low-temperature bacteriophage VR7 was deposited in the NCBI database under the accession number HM563683.

**Acknowledgments** We thank Dr. L. W. Black (University of Maryland, Baltimore, US) for *E. coli* strain B<sup>E</sup>. This work was funded by the Research Council of Lithuania (grant no. MIP-10240).

### References

- Balter M (2000) Evolution on life's fringes. *Science* 289:1866–1867
- Karam JD, Miller ES (2010) Bacteriophage T4 and its relatives. *Virology* 7:293. doi:10.1186/1743-422X-7-293
- Petrov VM, Ratnayaka S, Nolan JM, Miller ES, Karam JD (2010) Genomes of the T4-related bacteriophages as windows on microbial genome evolution. *Virology* 7:292. doi:10.1186/1743-422X-7-292
- Comeau AM, Bertrand C, Letarov A, Tétart F, Krisch HM (2007) Modular architecture of the T4 phage superfamily: a conserved core genome and a plastic periphery. *Virology* 362:384–396. doi:10.1016/j.virol.2006.12.031
- Ackermann H-W, Krisch HM (1997) A catalogue of T4-type bacteriophages. *Arch Virol* 142:2329–2345
- Leclerc H, Mossel DAA, Edberg SC, Struijk CB (2001) Advances in the bacteriology of the coliform group: their suitability as markers of microbial water safety. *Annu Rev Microbiol* 55:201–234
- Seeley ND, Primrose SB (1980) The effect of temperature on the ecology of aquatic bacteriophages. *J Gen Virol* 46:87–95
- Kaliniene L, Klauska V, Truncaite L (2010) Low-temperature T4-like coliphages vB\_EcoM-VR5, vB\_EcoM-VR7 and vB\_EcoM-VR20. *Arch Virol* 155:871–880
- Klauska V, Piešiniene L, Staniulis J, Nivinskas R (2003) Abundance of T4-type bacteriophages in municipal wastewater and sewage. *Ekologija* 1:47–50
- Miller ES, Kutter E, Mosig G, Arisaka F, Kunisawa T, Rüger W (2003) Bacteriophage T4 genome. *Microbiol Mol Biol Rev* 67:86–156
- Wilson GW, Edgell DR (2009) Phage T4 mobE promoters trans homing of the defunct homing endonuclease I-TevIII. *Nucl Acids Res* 21:7110–7123
- Edgell DR, Gibb EA and Belfort M (2010) Mobile DNA elements in T4 and related phages: a review in the series on bacteriophage T4 and its relatives. *Virology* 7:290. doi:10.1186/1743-422X-7-290
- Comeau AM, Krish HM (2008) The capsid of the T4 phage superfamily: the evolution, diversity, and structure of some of the most prevalent proteins in the biosphere. *Mol Biol Evol* 25:1321–1332
- Comeau AM, Bertrand C, Letarov A, Tétart F, Krish HM (2007) Modular architecture of the T4 phage superfamily: a conserved core genome and a plastic periphery. *Virology* 362:384–396
- Chibani-Chennoufi S, Canchaya C, Bruttin A, Brüßow H (2004) Comparative genomics of the T4-like *Escherichia coli* phage JS98: implications for the evolution of T4 phages. *J Bacteriol* 186(24):8276–8286. doi:10.1128/JB.186.24.8276-8286.2004
- Black LW, Showe MK, Steven AC (1994) Morphogenesis of the T4 head. In: Karam JD (ed) *Molecular biology of bacteriophage T4*. American Society for Microbiology, Washington, DC, pp 369–382
- Petrov VM, Nolan JM, Bertrand C, Levy D, Desplats C, Krish HM, Karam JD (2006) Plasticity of the gene functions for DNA replication in the T4-like phages. *J Mol Biol* 361:46–68
- Filée J, Baptiste E, Susko E, Krisch HM (2006) A selective barrier to horizontal gene transfer in the T4-type bacteriophages that has preserved a core genome with the viral replication and structural genes. *Mol Biol Evol* 23:1688–1696
- Nolan JM, Petrov V, Bertrand C, Krish HM, Karam JD (2006) Genetic diversity among five T4-like bacteriophages. *Virology* 33:30