VILNIUS UNIVERSITY
INSTITUTE OF BIOCHEMISTRY

Laura Kalinienė

INVESTIGATION OF GENOME SEQUENCE AND GENE EXPRESSION REGULATION IN T4 RELATED BACTERIOPHAGES

Summary of doctoral thesis
Physical Sciences, Biochemistry (04 P)

Vilnius, 2010

This study was carried out at the Department of Gene Engineering of the Institute of Biochemistry during period of 2005-2009.

Scientific supervisor:

Dr. Habil. Rimantas Nivinskas (Institute of Biochemistry, physical sciences, biochemistry - 04 P).

Scientific consultant:

Dr. Lidija Truncaitė (Institute of Biochemistry, physical sciences, biochemistry - 04 P)

The thesis is defended at the Council of Biochemistry science direction of Vilnius University and Institute of Biochemistry:

Chairman:

Dr. Rolandas Meškys (Institute of Biochemistry, physical sciences, biochemistry - 04 P).

Members:

Prof. Dr. Habil. Kęstutis Sasnauskas (Institute of Biotechnology, physical sciences, biochemistry - 04 P);

Prof. Dr. Virginijus Šikšnys (Institute of Biotechnology, physical sciences, biochemistry - 04 P);

Prof. Dr. Donaldas Čitavičius (Vilnius University, biomedical sciences, biology – 01 B);

Dr. Habil. Juozas Benediktas Staniulis (Institute of Botany, biomedical sciences, biology – 01 B).

Opponents:

Prof. Dr. Edita Sužiedėlienė (Vilnius University, physical sciences, biochemistry - 04 P);

Dr. Daumantas Matulis (Institute of Biotechnology, physical sciences, biochemistry - 04 P).

The dissertation will be defended at the public meeting of the Council of Biochemistry at 11 a.m. on June 29, 2010 in the hall of the Institute of Biochemistry.

Address: Mokslininkų 12, LT – 08662, Vilnius, Lithuania

The abstract of the dissertation was sent on May 28, 2010.

The dissertation is available at the library of the Institute of Biochemistry and at the library of Vilnius University.

VILNIAUS UNIVERSITETAS
BIOCHEMIJOS INSTITUTAS

Laura Kalinienė

T4 GIMININGŲ BAKTERIOFAGŲ GENOMŲ SEKŲ NUSTATYMAS IR GENŲ RAIŠKOS TYRIMAS

Daktaro disertacijos santrauka
Fiziniai mokslai, biochemija (04P)

Vilnius, 2010

Disertacija rengta 2005-2009 metais Biochemijos institute

Mokslinis vadovas:

habil.dr. Rimantas Steponas Nivinskas (Biochemijos institutas, fiziniai mokslai, biochemija – 04P);

Mokslinė konsultantė:

dr. Lidija Truncaitė (Biochemijos institutas, fiziniai mokslai, biochemija – 04P).

Disertacija ginama Vilniaus universiteto ir Biochemijos instituto biochemijos mokslo krypties taryboje:

Pirmininkas:
dr. Rolandas Meškys (Biochemijos institutas, fiziniai mokslai, biochemija – 04P)

Nariai:
prof. habil. dr. Kęstutis Sasnauskas (Biotechnologijos institutas, fiziniai mokslai, biochemija – 04P);
prof. dr. Virginijus Šikšnys (Biotechnologijos institutas, fiziniai mokslai, biochemija – 04P);
prof. dr. Donaldas Čitavičius (Vilniaus universitetas, biomedicinos mokslai, biologija – 01 B);
habil. dr. Juozas Benediktas Staniulis (Botanikos institutas, biomedicinos mokslai, biologija – 01 B).

Oponentai:
prof. dr. Edita Sužiedėlienė (Vilniaus universitetas, fiziniai mokslai, biochemija – 04P);
dr. Daumantas Matulis (Biotechnologijos institutas, fiziniai mokslai, biochemija – 04P).

Disertacija bus ginama viešame Biochemijos mokslo krypties tarybos posėdyje 2010 m. birželio mėn. 29d. 11 val. Biochemijos instituto aktų salėje
Adresas: Mokslininkų 12, LT – 08662, Vilnius, Lietuva

Disertacijos santrauka išsiuntinėta 2010 m. gegužės mėn. 28 d.
Disertaciją galima peržiūrėti Biochemijos instituto ir Vilniaus universiteto bibliotekose

ACTUALITY OF THE PROBLEM

Bacteriophages are the most abundant entities in the biosphere. They are found wherever bacteria thrive. T4 like viruses comprise one of the most diverse and common group of bacteriophages found in nature. Despite their abundance, only ~25 fully sequenced T4 type bacteriophages can be found in various databases to date. The attempt to explore the molecular mechanisms of genetic adaptation as well as the evolutionary relationships between various groups of phages is limited mainly by the lack of information content. Based on the sequence comparison of the three major structural proteins, the T4-type phages are classified into four subgroups with an increasing divergence from T4: the T-evens, pseudo-T-evens, shizo-T-evens and exo-T-evens. The archetype of the T4-type superfamily bacteriophage T4 represents the subgroup of T-evens. Only 10% of a typical T-even genome diverges significantly from the DNA content of T4. Meanwhile the nucleotide sequence of genomic DNA of the typical pseudo T-even (e.g. Phage RB49) differs substantially and only ~ 45% of the genome encodes protein homologues of T4 with the levels of aa identity that range between 20% and 70%. The shizo-T-evens and exo-T-evens are even more distant – they differ from T4 in size, host range and the genomic content.

Despite the fact that bacteriophage T4 is being investigated for more than 50 years, properties of nearly half of the T4 genes are still unknown. More over, an accurate understanding of the contribution in the development of the phage, as well as the molecular principles of action are yet to be discovered in case of many identified proteins of T4. The lack of the structural variants encoded by other phages of the T4-type superfamily eminently limits the investigation of such proteins of T4.

The regulation of bacteriophage T4 gene expression as well as the activity of this process-related proteins (such as ModB, ModA, MotA and RegB) have been investigated by the scientists of the Department of Gene Engineering of the Institute of Biochemistry for many years. However, the genetic content disconcordance between the representatives of the distinct

subgroups of T4-type bacteriophages prevents the researchers from more detailed analysis of the processes mentioned. A set of the genes encoding regulatory proteins and the DNA signals that are important in the gene expression regulation of T-evens, are generally absent from the genomes of pseudo-, shizo- and exo-T-evens. Thus, the identification and subsequent investigation of the intermediate subgroups not only will help to explore the evolutionary relationships among the T4-related bacteriophages, but may also promote the clarification of the overall gene regulation process in T-evens.

This work is intended to identification of phages, occupying an intermediate phylogenetic subgroup between T-evens and pseudo T-evens.

AIM OF THE DISSERTATION WORK

The purpose of the research presented in this dissertation was to identify and subsequently investigate the whole genome sequence and gene expression regulation of such a T4 related bacteriophage, which might occupy the distinct intermediate subgroup between T-evens and pseudo-T-evens.

Towards this goal, the following tasks had been formulated:

1. To determine the structural organization of the genomic region between genes *30* and *31* in the group of uncharacterized T4-related bacteriophages. Based on the data obtained, to select the potential representative of the intermediate subgroup of the T4 type bacteriophages.
2. To determine the whole genome sequence of the bacteriophage selected.
3. To investigate the gene expression regulation principles of the bacteriophage selected.
4. To investigate the physiological properties of the selected phage and to perform its classification.

NOVELTY AND SIGNIFICANCE

The results of the first genomic project carried out in Lithuania were presented in this dissertation. The complete genome sequence of bacteriophage

vB_EcoM-VR7 (subsequently referred to by its shorter common laboratory name VR7) had been determined. 293 ORF`s as well as sequences for 43 early, 43 middle and 44 late putative promoters were identified. In total, the activities of 24 promoter sequences, which contained deviations from the consensus of those of T4, were examined and the consensus sequences for each class of VR7 promoters were established.

Three low temperature bacteriophages were characterized and classified. The results obtained indicated that these phages possessed physiological properties that were never observed among the characterized T4-related *E. coli* phages before. The analysis and the study of the defined sequences demonstrated the existence of the new phylogenetic subgroup within the superfamily of T4-related bacteriophages

THE DEFENSIVE STATEMENTS:

- The genomic region between genes *30* and *31* of bacteriophages LZ, LZ1 and LZ9, despite various divergences both in size and nucleotide sequence, exhibits properties that have already been observed within the T-evens.
- The genomic region between genes *30.4* and *30.2* of bacteriophage VR7 contains several structural features that are not common within the T-evens.
- The 169.285 bp genomic sequence of bacteriophage VR7 encodes 211 proteins homologous to those in T4 in the range 27%-97%.
- The temporal control of transcription in VR7 is based upon the use of the three major classes of promoters – early, middle and late.
- The plasmid encoded MotA of VR7 complements the T4*motA*Δ mutant for growth in *E. coli.*
- Bacteriophages VR5, VR7 and VR20 possess physiological properties that have never been observed among the characterized T4-related *E. coli* phages.

- Bacteriophages VR5, VR7 and VR20 together with JS98 fall into the distinct subgroup of T4-type bacteriophages.

DISSERTATION CONTENTS

The dissertation is written in Lithuanian and contains the following sections: Introduction, Review of the Literature, Materials and Methods, Results and Discussion, Conclusions, List of References (249 positions), Additional Material (2), Tables (2) and Figures (46). Total 155 pages.

MATERIALS AND METHODS

Enzymes and isotope. T4 polynucleotide kinase and *Taq* DNA polymerase were obtained from Fermentas AB. Avian myeloblastosis virus reverse transcriptase was obtained from Promega. [γ^{32}P]ATP was supplied by Amersham Biosciences.

Phages and bacterial strains. T4-like phages VR5, VR7 and VR20 were originally isolated from Lithuanian municipal wastewater (VR20) and sewage (VR5 and VR7). Phage T4D^{+} wild-type was kindly provided by Dr W. B. Wood. Phages T4*mot*A^{-} (33amN134-55amBL292-45amE10-amG1) and T4 K10*motA*Δ were kindly provided by Dr. N. Guild and Dr K. N. Kreuzer respectively. T4-related phages LZ, LZ1 and LZ9 were kindly provided by Dr. K. Carlson. *E. coli* strain B^{E} (*sup*^{0}), a gift from Dr L. W. Black, was used for phage propagation and phage growth experiments. *E. coli* strains: B40 (*supF strr* (Sup+)) (Dr L. W. Black), MH1 (*araD139 ΔlacX74 galU galK*$^{-}$ *hsr hsm rpsL*) (Dr K. N. Kreuzer), CR63 (*supD*, ser) (Dr K. N. Kreuzer), K-12 (AN180 (F^{-}, *argE3, thi, mtl, xyl, sir704*)) (Dr V. N. Krylov), BL21 (DE3) (F^{-} *dcm ompT hsdS*($r_B^{-}m_B^{-}$) *gal* λ(DE3)) (Novagene), Nova Blue (DE-3) (*endA1 hsdR17*(r_{K12}^{-} m_{K12}^{+}) *supE44 thi-1 recA1 gyrA96 relA1lac* [F' *proA*$^{+}$*B*$^{+}$ *lacI*^{q}*ZΔM15*::Tn*10* (Tc^{r})]) (Novagene) and GM 2163 (*dam-13::*Tn *9 dcm-6 hsdR2 leuB6 his-4 thi-1 ara-14 lacY1 galK2 galT22 xyl-5 mtl-1 rpsL136 tonA31 tsx-78 supE44* McrA^{-} McrB) (Fermentas) were used for host range determination. *E. coli* strains MCS1 (*supD*) and MCS1/pRS31 (*supD,* bearing a plasmid with gene *motA*), which were kindly provided by Dr. K. N. Kreuzer, were used in the functional complementation experiments of VR7 MotA.

For all phage experiments, bacteria were cultivated in Luria-Bertani broth (LB) or LB agar. Bacterial growth was monitored turbidimetrically by reading OD_{600}. An OD_{600} of 0.8 corresponded to 2.0×10^{8} cells/ml.

Phage techniques. Phage isolation, plating and titering were carried out as described by Klausa (2003), while, adsorption test and one-step growth experiments were carried out as described by Khusainov (1992) and Carlson (1994), respectively. Determination of the efficiency of plating (e.o.p.) was performed as described by N.D. Seeley and S.B. Primrose (1980). High titer phage stocks were diluted and plated in duplicate. Plates incubated at 17, 24, 30, 35, 37, 39 and 40ºC were read after 18-24 h, those at 12 ºC after 48h, and those at 7ºC after five days. The temperature at which the largest number of plaques was formed was taken as the standard for the e.o.p. calculation.

Viral DNA isolation and restriction analysis. Aliquots of phage suspension (10^{11}-10^{12} pfu/ml) were subjected to phenol/chloroform extraction and ethanol-precipitation as described by Carlson (1994). Isolated phage DNA was subsequently used in restriction analysis. Restriction digestion was performed with *EcoRV*, *DraI*, *Vsp*I, *NdeI*, *SmiI*, *Bsp1107I*, *AluI, Bst1170I*, *Bsp143I*, *HinfI*, *BglII* and *MunI* restriction endonucleases (Fermentas), according to the suppliers` recommendations. DNA fragments were separated by electrophoresis in 0.8% agarose gel containing ethidium bromide. Restriction digestion was repeated at least thrice to ascertain the results.

RNA isolation and primer extension analysis of phage mRNAs. Total RNA from phage-infected *E. coli* cells was phenol extracted and analyzed by primer extension under conditions of primer excess, using avian myeloblastosis virus reverse transcriptase as described by Uzan (1988). Phage-specific oligonucleotides were used to prime reverse transcriptase. The oligonucleotides were 5'–labeled with [γ32P]ATP using T4 polynucleotide kinase. In the case of VR7 P_E, mRNA was isolated from *E. coli* B^E cells at 3 min post-infection with VR7, in the case of P_L - at 3, 6, 9 and 15 min post-infection. The activity of middle promoters of VR7 was analyzed using mRNA isolated from chloramphenicole treated (+Cm) and untreated (-Cm) *E. coli* B_E cells at 6 minute post-infection with VR7. The ability of the plasmid encoded VR7 MotA to activate the middle-mode transcription of T4 was investigated using mRNA isolated from *E. coli* B_E cells at 6 minute post-infection with $T4D^+$ and *E. coli* B_E cells, harbouring either pET-*motA*-VR7 or pET-*motA*-T4, infected with T4*motA*$^-$ independently. The RNA isolated from *E. coli* B_E cells at 6 minute post-infection with T4*motA*$^-$ was used as control. All of the experiments mentioned were performed at 30ºC.

PCR and sequencing procedures. For the investigation of the genomic region *30-31*, the DNA fragments were amplified by PCR using the T4-specific, RB69-specific and RB49-specific primers. Later, primers based on the DNA sequences obtained were synthesized. The DNA templates for the PCR were denatured phage particles. The PCR was carried out as described by Saiki and co-authors (1988). The sequencing reactions were carried out using a CycleReader™ DNA sequencing kit (Fermentas AB). The oligonucleotide primers for the sequencing reactions were 5'–end labeled by T4 polynucleotide kinase (Fermentas AB) with [γ^{32}P]ATP (Amersham Biosciences).

In the case of the phylogenetic analysis of gp23, PCR with Pfu DNA polymerase (Fermentas), denaturated phage particles, as a template, g*23* targeting degenerate primers MZIA1bis (5` - GAT ATT TGI GGI GTT CAG CCI ATG A-3`) and MZIA6 (5` -CGC GGT TGA TTT CCA GCA TGA TTT C-3`) was performed following the method described by Filée et al. (2005). PCR products were cloned using CloneJET™ PCR Cloning Kit (Fermentas) and sequenced. DNA sequences were deposited in the EMBL database under the accession numbers: FN641800, FN641801, and FN641802.

Sequencing strategy of the genomic DNA of VR7. VR7 phage DNA was digested with the restriction enzyme *ApoI* (*XapI*). The resulting 1-3(5) kb fragments were purified and cloned into the plasmid vector pCC1BAC™ and transformed into Escherichia coli by electroporation. Plasmids were purified and the cloned inserts were sequenced at the Sequencing Center (Institute of Biotechnology, Lithuania) using BigDye® Terminator v3.1. Cycle Sequencing Kit (Applied Biosystems). The DNA was sequenced to a six-, eightfold coverage and the sequences were assembled, resulting in 64 separate contigs (~110 kb of genomic DNA) which were projected on

T4 and RB69 maps. Sequence gaps were closed by the sequencing of PCR amplicons obtained using VR7 DNA specific primers.

Genome annotation. Genes with best hit E-values $<10^{-4}$ to known genes of T4were designated by the T4 gene name. Putative genes without T4 orthologs (with the exeption 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 counterclockwise. The terminus of the genome was defined as the start of translation of the gene rIIB.

Sequence analysis. DNA sequence analysis 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. The rho-independent transcription terminators were detected with TransTerm and ARNold (http://rna.igmors.u-psud.fr/). Hypothetical proteins of VR7 were analyzed using YASPIN (http://www.ibi.vu.nl) and FeatureMap3D (http://www.cbs.dtu.dk/). Neighbor-joining tree was constructed using MEGA 4.0 (www.megasoftware.net). DNA sequences of diagnostic g*23* fragments were deposited in the EMBL database under the accession numbers: FN641800, FN641801, and FN641802.

RESULTS AND DISCUSSION

Investigation of the structural organization of the genomic region between genes *30* and *31* in T4-related bacteriophages LZ, LZ1, LZ9 and VR7. The genomic region between genes *30* and *31* of four T4-related bacteriophages LZ, LZ1, LZ9 and VR7 has been investigated in order to single out those, occupying the intermediate subgroup between the T-evens and the pseudo T-evens. In the case of T4, this region has been shown to contain eleven coding sequences and several regulatory elements, including two early, two middle and one late promoters. The comparative analysis of the genomic region *30-31* of 26 T4 related bacteriophages (unpublished data) indicate that *rIII*, *g30.9* and *g30.3,* in the case of all T-evens tested, encode proteins >92 % homologous to their corresponding counterparts in T4.

The sequence analysis of LZ, LZ1 and LZ9 reveales that the most diverged part of the genomic region *30-31* in these phages lies within early gene cluster *30.9-30.3*. The new orthologue of T4 *30.8* has been identified in LZ1. Two individual open reading frames *30.4A* and *30.4B* replace the counterpart of

T4 gene *30.4* in LZ, while any counterparts of T4 genes *30.8* and *30.5* are absent at all. No analogues of T4 genes *30.7* and *30.8* have been found in phage LZ9. One of two early promoters is absent in LZ and LZ9. However, the genomic region between genes *30* and *31* of all three LZ bacteriophages tested, despite various divergences both in size and nucleotide sequence, employs almost the same regulatory pattern of gene regulation as bacteriophage T4. Moreover, all three LZ phages tested preserve ~100% identity of genes *rIII* and *30.9* as well as >92% bp identity of g*30.3* with those of T4.

From these considerations, it has been concluded that bacteriophages LZ, LZ1 and LZ9 are typical representatives of the T-even subgroup of T4 related bacteriophages and have been excluded from the subsequent analysis of the genomic DNA.

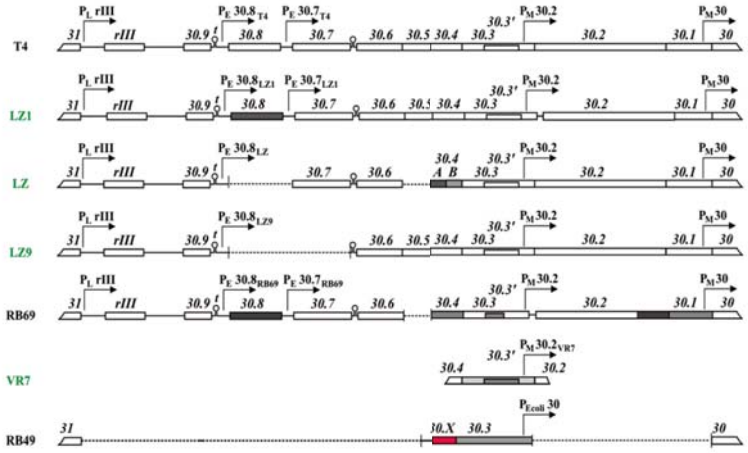

Fig. 1. **Schematic outline of the genomic region *30-31* of bacteriophages T4, LZ, LZ1, LZ9, VR7 and RB49.** Shown are positions of genes, as well as the positions of late (P_L), early (P_E), middle (P_M) promoters, rho-independent terminators (*t*) and putative stem-loop structures. Grayed boxes represent genes with low homology, red boxes – no homology to their T4 counterparts.

In the case of VR7, following the strategy described in Materials and Methods, only the sequence of gene *30.3* (together with completely embedded gene 30.3`) has been obtained. The deduced primary structure of gp30.3 and gp30.3` of VR7 share 85% and 58% sequence identity with those of T4.

Moreover, the sequence for MotA box of promoter $P_M30.2$ which have been identified in VR7 (Fig. 2) is completely new for the middle promoters of T4.

Fig. 2. The alignment of the sequences of VR7 $P_M30.2$ and T4 $P_M30.2$. Different nucleotides are given in black. The -30 and -10 consensus elements of middle promoters are boxed. The arrows indicate the initiating nucleotides.

When the activity of VR7 $P_M30.2$ was confirmed by mRNA sequencing (Fig. 3), the search for *motA* gene of VR7 has been performed. Interestingly, the sequence of VR7 *mot*A has been obtained by PCR using T4 *mod*B and *mod*A specific primers.

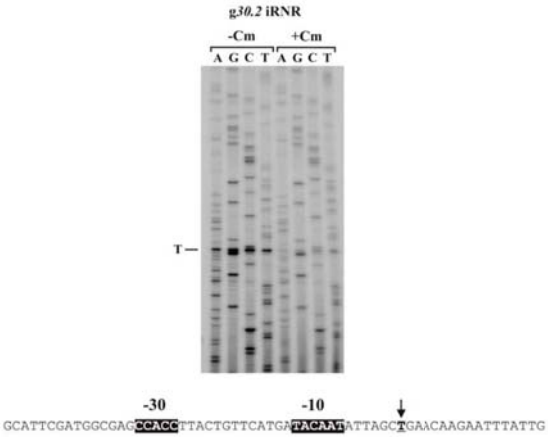

Fig. 3. Primer extension sequencing of the transcript for VR7 gene *30.2*. Primer extension sequencing reactions were performed using total mRNA (isolated from chloramphenicole treated (+Cm) and untreated (-Cm) *E. coli* B_E cells at 6 minute post infection with VR7), AMV and dNTP`s as indicated. The initiating nucleotides for the MotA dependent transcripts were noted.

The deduced primary structure of VR7 MotA share only 34% aa identity with corresponding counterpart in T4, indicating that phage VR7 encodes the most diverged middle transcription activator within the group of identified MotA proteins.

Based on the data obtained, bacteriophage VR7, as the potential representative of an intermediate subgroup of the T4 type bacteriophages, has been selected for the subsequent experiments.

Overview of the genome sequence of bacteriophage VR7. The complete genome sequence of bacteriophage VR7 has been determined. The genome sequence is 169,285 bp, with an overall G+C content of 40,3% in comparison with 35.3 % of T4. Overall, 95,4% of the VR7 genome is coding. It encodes 293 putative protein-encoding open reading frames (ORFs), 1 $tRNA^{MET}$, 43 P_E, 43 P_M, 50 P_L and 40 potential rho-independent terminators. In total, 281 of VR7 ORFs are initiated with AUG, while 7 with UUG and 5 with GUG. Interestingly, no ORFs of T4 has been found to initiate from UUG (Miller et al., 2003).

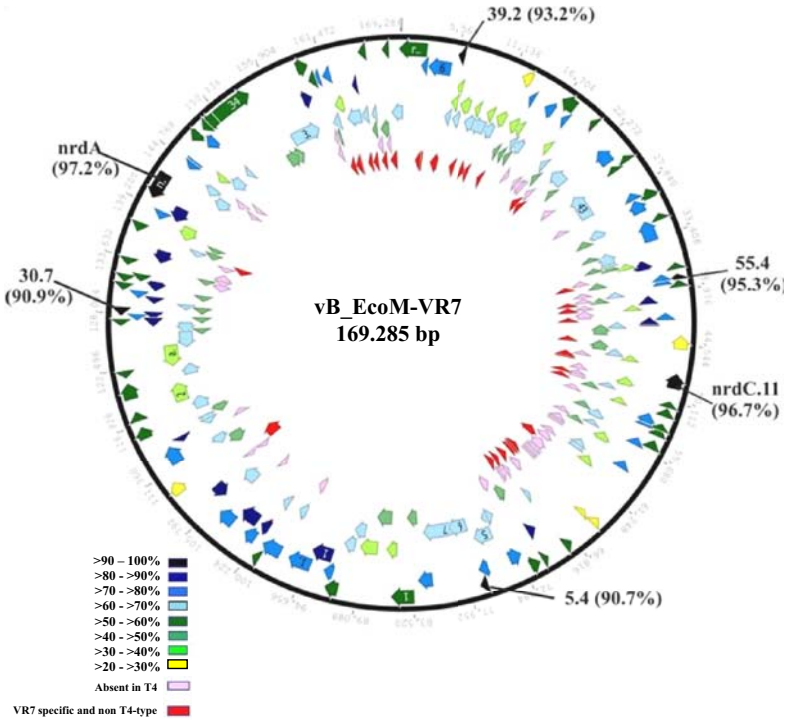

Fig. 4. The genomic map of VR7. The meanings of the colours are indicated.

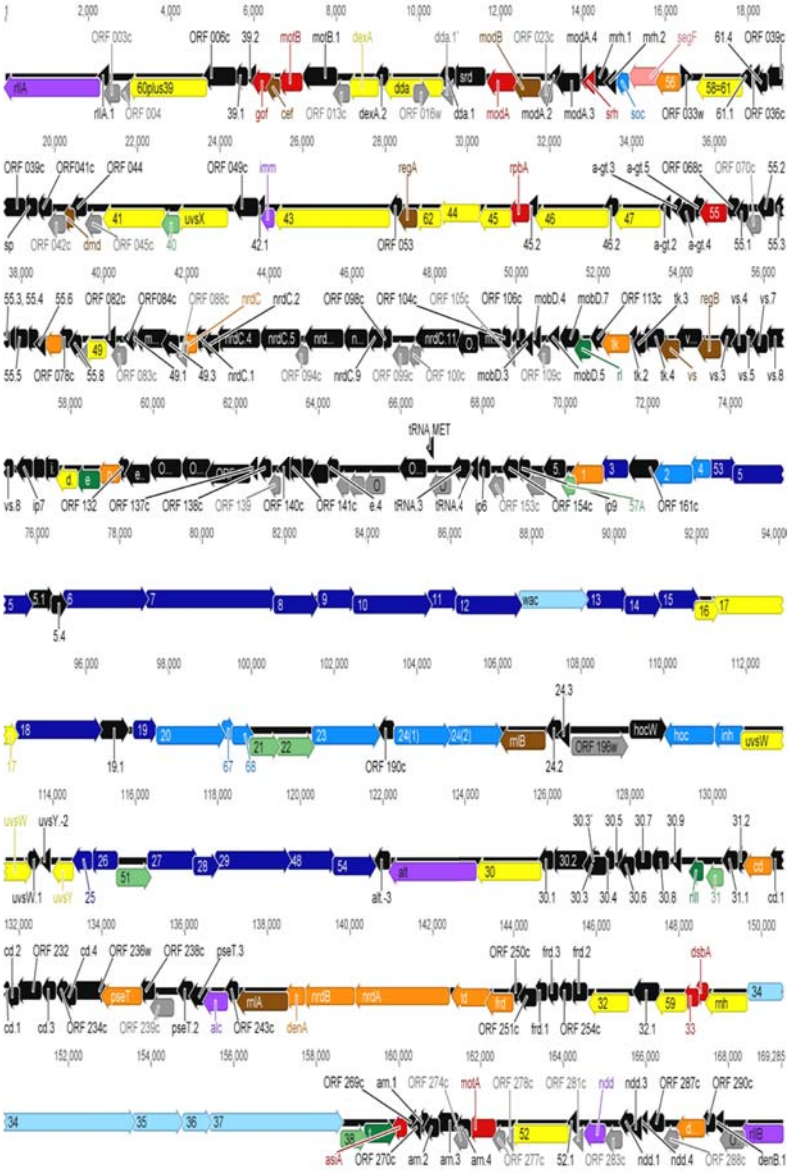

Fig. 5. Functional genome map of bacteriophage VR7. The coding capacity of the VR7 genome is shown. The colour scheme (by gene function) is as defined below in the text. Functions are assigned according to the characterized ORFs in T4.

Only 41% of VR7 genomic DNA share 73% base paire identity with DNA of T4, but once translated, 88% of VR7 genome (211 ORFs) encodes protein homologs of T4 genes (Fig. 4) with levels of aa identity that range between 27% and 97%. Thus, based on their similarity with biologically defined T4 proteins, 111 ORFs of VR7 may be given a functional annotation. The functional genome map of bacteriophage VR7 is presented in the figure 5. Genes are coloured 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.

VR7 differs from T4 by a duplicated gene *24* as well as an additional gene for an outer capsid protein Hoc. Four of VR7 putative ORFs *30.3*`, *dda.1*`, *ORF016w* and *ORF274c* are completely embeded within the coding sequence of *30.3*, *dda.1*, *dda* and *arn.4* respectively. Homologs to the T4 α-gt, β-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 are all absent in VR7.

Phage T4 encodes 15 homing endonucleases, accounting for ~11% of the coding potential of the genome (Wilson and Edgell, 2009). There are only five genes for putative homing endonucleases in VR7: two of VR7 ORFs encode proteins that share 65% and 54% amino acid sequence identity with SegF and SegG of T4, meanwhile the remaining 3 putative mobile endonucleases of VR7 are homologous to MobE of RB43, RB32 and T4.

With no homologues in T4, 46 VR7 ORFs derive from other T4-type phages and 9 ORFs show similarities to bacterial (Table 2) or non-T4-type phage genes (Table 1). Meanwhile, 27 ORFs of VR7 lack any database matches.

Table 1. VR7 ORFs with homologs in non-T4-type phages

VR7 ORF (aa)	Homologous ORF (aa)	Origin of the best match	Function	No. of identical aa/similar aa	E-value
ORF006c (231)	Gp13.5 (133)	Enterobacteria phage EcoDS1	Hypothetical protein	45/66 (131*)	3.5e-11
ORF134c (216)	ORF78 (210)	Pseudomonas phage D3	Hypothetical protein	54/87 (147*)	2.1e-22
ORF136c (55)	ORF (55)	Salmonella phage 5	Hypothetical protein	27/33(57*)	0.25
ORF290c (163)	ORF058 (211)	Staphylococcus phage G1	HNH nuclease	55/81 (150*)	3.0e-21

*- represents the length of overlapping segment

Table 2. VR7 ORFs with homologs in cellular organisms

VR7 ORF (aa)	Homologous ORF (aa)	Origin of the best match	Function	Number of identical aa/similar aa	E-value
ORF003c (99)	Avi_1970 (425)	*Agrobacterium vitis*	Major capsid-like protein	25/38 (77*)	0.14
ORF023c (77)	rpoH (299)	*Orientia tsusugamushi*	RNA polymerase sigma factor 32	27/41(79*)	0.43
ORF108c (72)	Cpin_4278 (213)	*Chitinophaga pinensis*	Two component transcriptional regulator, LuxR family	13/18(34*)	0.021
ORF146c (141)	RPC_1770 (522)	*Rhodopseudomonas palustris*	Cytidine/deoxycytidylate deaminase	33/57 (117*)	2.0e-05
ORF195w (427)	(842)	*Bacillus cereus subsp. cytotoxis*	Collagen triple helix repeat protein	(320*)	4.9e-44

*- represents the length of overlapping segment

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

Structural genes of VR7. The structural genes of VR7 are organized into two separate clusters – a large 51 kb cluster which encodes most of the structural components of the virion (Fig. 6) and a small 10 kb cluster which is represented by five tail fiber genes (Fig. 7). Both structural gene clusters are separated by a 27 kb segment of DNA transaction genes, as it is also observed in T4. Although the virion gene order is nearly identical between VR7 and T4, there are a few distinctions. Homologs of non-essential T4 genes *segC*, *segD*, *segE,ipI, ipII, ipIII* and *5.3* are all absent in VR7. *ORF161c*, showing 60% aa sequence identity with putative mobile endonuclease of RB32, has been found between VR7 structural genes *3* and *2*. Gene *24* (the gene for a head vertex protein) is duplicated in VR7, and there are two *hoc*-like ORFs lying on the opposite strands of genomic DNA. The duplicated *g24* has been also observed in JS98, although in the case of this phage, one *g24* copy is split into two open reading frames (Chibani-Chennoufi et al., 2004). The analysis of VR7 DNA sequences surrounding *hoc*(w) indicates that this gene together with ORF196w have been acquired by horizontal gene transfer (most likely from the host genome harbouring DNA of prophage or deffective prophage).

The structural proteins of VR7 share the 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 show the highest degree of conservation, however, the amino acid sequences of the three major structural proteins gp23, gp18 and gp19 of VR7 show greater relatedness with JS98 (97%, 90% and 91% aa identity respectively) than with gp23, gp18 and gp19 of T4 (85%, 71% and 70% aa identity respectively).

Most of the tail fiber and base plate wedge proteins of VR7 show low degree of sequence conservation with T4 (e.g. gp12 (45%), gpwac (36%) or gp9 (48%)). Thus, the fact that gp37 and gp38 of VR7 share 70% and 88% amino acid sequence identity with corresponding proteins of T4 is surprising, indeed. However, several aa insertions, similar to those in TuIb, have been

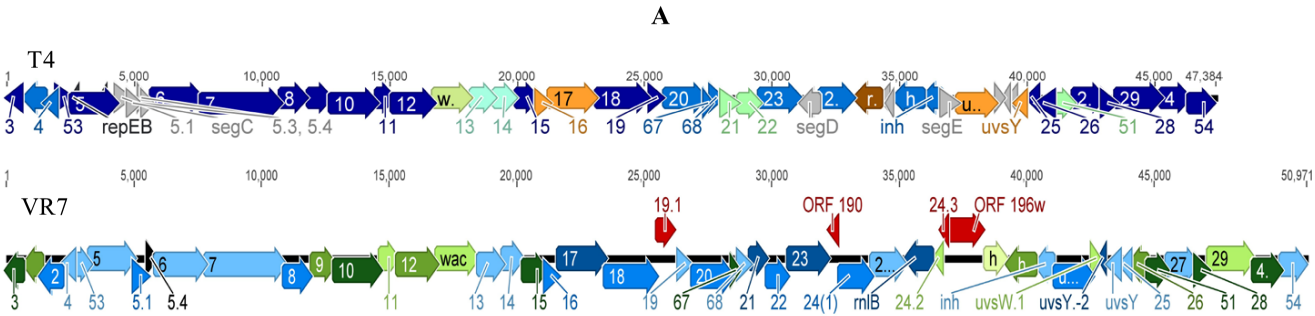

Fig 6. Comparison of conserved virion structural modules of VR7 and T4. In the case of T4 genes are coloured as follows: dark blue - tail; blue - head; light blue - neck; green – fibers; light green – chaperonins; orange – replication/recombination/reparation and packaging; brown - translation; grey – mobile nucleases and unknown function ORFs. In the case of VR7 colours represent aa identity as follows: black – 90%-100%; dark blue - 80%-90%; blue – 70%-80%; light blue – 60%-70%; dark green – 50%-60%; green – 40%-50%; light green – 30%-40%; very light green – 20%-30%; red – no homology with T4.

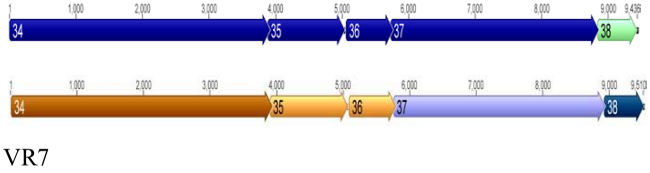

Fig 7. Tail fiber gene cluster of VR7 and T4. In the case of T4 genes are coloured as follows: dark blue – tail fibers; light green – assembly catalyst. In the case of VR7 colours represent amino acid identity with corresponding proteins in T4 (dark blue –80%-90%; light blue – 60% - 70%; brown – 50% - 60%; light brown – 40%-50%).

found within the C-terminal region of gp37, suggesting that VR7 and T4 may differ in the host cell recognition.

DNA replication, recombination, repair (RRR) and nucleotide metabolism. All of the essential DNA replication, recombination and repaire enzymes of T4 are encoded in VR7. Eight of the VR7 RRR genes are located in a single, contiguous gene cluster that is >13 kb in length (Fig. 8). Meanwhile the genes encoding other RRR enzymes are distributed throughout the VR7 genome. The RRR gene order is relatively conserved between T4 and VR7 with major differences mainly corresponding to the non-essential genes, not involved in the DNA replication and recombination. The gene for the larger topoisomerase subunit is the only exeption to the rule. Unlike in T4, genes *39* and *52* of VR7 encode two complete subunits of the DNA topoisomerase.

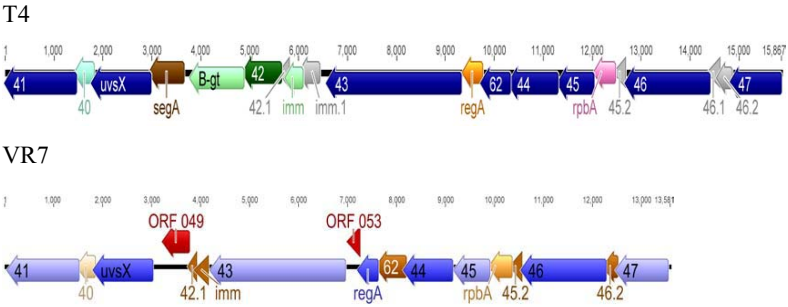

Fig. 8. Comparison of RRR gene cluster of VR7 and T4. In the case of T4 colours represent: dark blue – RRR; dark brown – mobile endonucleases; light green – host or phage interactions; dark green – nucleotide metabolism; pink – transcription. In the case of VR7 colours represent amino acid identity with corresponding proteins in T4 (blue – 70% - 80%; light blue – 60% - 70%; brown – 50% - 60%; orange – 40%-50%; peach – 30% - 40%; red – no homology to T4).

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 share a high degree of sequence conservation within the sites that determine their catalytic properties. Meanwhile the regions for protein-protein interaction have diverged significantly. The same observations have been made after the comparision of

the genomic DNA of T4 with the genomes of 11 different T4 related bacteriophages (Petrov et al., 2006).

Most of the proteins involved in the nucleotide metabolism in T4 have homologues in VR7. The gene for dCTPase/dUTPase *56*, dCMP deaminase *cd*, dTMP synthase *td*, dihidrofolate reductase *frd*, dNMP kinase *1*, thymidine kinase *tk* as well as aerobic ribonucleotide reductase complex *nrdABC* are all present in VR7. With the amino acid sequence identity ranging from 65% (Td) to 97% (NrdA) they show 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 may imply that bacteriophage VR7 is less adapted to anaerobic growth conditions in comparison to T4.

As it will be reviewed bellow, the restriction analysis suggests that the DNA of VR7 and T4 may contain similar modifications. This suggestion is also supported by the presence of homologs to the T4 Alc and DenB in the genome of VR7. However, the gene for the dCMP hydroxymethylase (g*42*) as well as both *α-gt* and *β-gt* genes are absent from VR7. Therefore, it can be hypothesized that in the case of VR7 some ORF with unknown function may act as the functional analog of gp42, as it has been observed in the phage RB69.

Promoters, transcription proteins and terminators. The temporal control of gene transcription in T4 derives specificity from modifications to the host RNA polymerase, phage encoded transcription factors, and unique promoter sequences. The analysis of the DNA sequence of VR7 reveales that T4 transcriptional regulatory circuits are all valid for the phage investigated. Putative promoter elements resembling T4 early, middle and late promoters have been identified in VR7. Moreover, the homologs to the T4 transcription proteins are all present in VR7 as well.

In total, 43 putative VR7 P_E have been identified, and the activities of P_E24.3, $P_E$motA.1 and $P_E$ip6 have been confirmed by primer extension analysis

(data not shown). The deduced consensus sequence (Fig. 9) implies that most of putative VR7 P_E show a definite preference for G rather than T at position -34 and/or C rather than T at position -12.

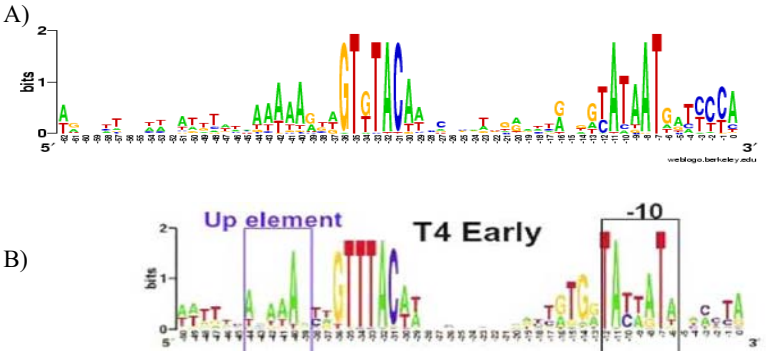

Fig. 9. The comparison of an early promoter consensus sequences of VR7 (A) and T4 (B). The conserved promoter elements are indicated. The logo of VR7 P_E was obtained using WebLogo program (http://weblogo.berkley.edu), the logo of T4 P_E is from Nolan et al (2006). In total 43 VR7 P_E were aligned.

In addition, an A-rich upstream promoter element at position -42 is highly conserved in the case of VR7 P_E. In T4, this element is important for acceleration of transcription with the ADP-ribosilated RNA polymerase. Interestingly, the ADP-ribosilating proteins Alt, ModA and ModB are all encoded in VR7 although the amino acid sequence identity (35%, 32% and 31% respectively) with their corresponding counterparts in T4 is rather low.

In total, 43 VR7 P_M have been identified, and the activities of 17 VR7 P_M have been confirmed by primer extension analysis (data not shown). Although the consensus sequence of VR7 P_M (Fig. 10) clearly resembles that of T4, there are some differences. First, the residues at positions -34, -33, -29, -28, -27, -10, -9 and -8 are more conserved in the case of VR7 P_M. In contrast, those at positions -25 and -12 exhibit low degree of conservation in comparison to P_M of T4. Nearly 10% of T4 P_M contain A at position -31 (Truncaite et al., 2002; Stoškiene et al., 2007), meanwhile in the case of VR7, there is a definite preference for T or C. Only 2 putative P_M with A within

MotA box sequence position -31 have been identified in VR7. However, no apparent promoter activity have been found in the case of both of these putative P_M-like sequences.

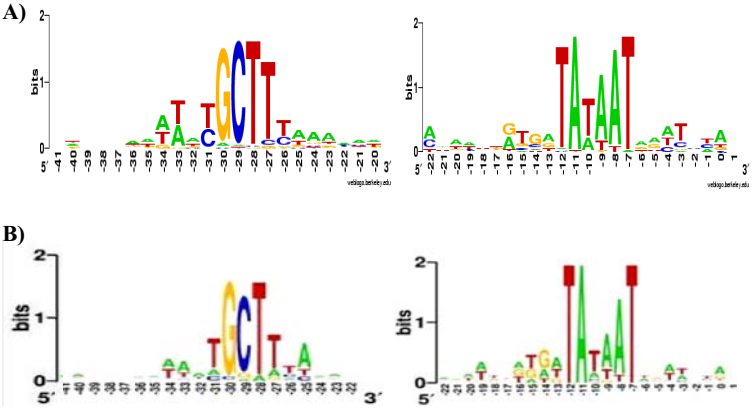

Fig. 10. The comparison of a middle promoter consensus sequences of VR7 (A) and T4 (B). The logo of VR7 P_M was obtained using WebLogo program (http://weblogo.berkley.edu), the logo of T4 P_M is from Stoškiene et al (2007). In total 43 VR7 P_M were aligned.

AsiA of VR7 exhibit 51% amino acid sequence identity with the middle transcription coactivator of T4. Meanwhile, as it has been mentioned above, the gene for the middle-mode transcription activator MotA in VR7 encodes the protein that exhibits only 34% sequence identity with its corresponding counterpart in T4. In order to test the ability of VR7 MotA to activate the middle-mode transcription of T4 *in vivo*, the functional complementation experiments have been performed.

It has been demonstrated, that the presence of plasmid encoded VR7 MotA complements the T4*motA*Δ mutant for growth in *E. coli,* and activates middle-mode transcription during the growth of T4*motA*$^-$ (e.g. P_M31, Fig. 11(I)).

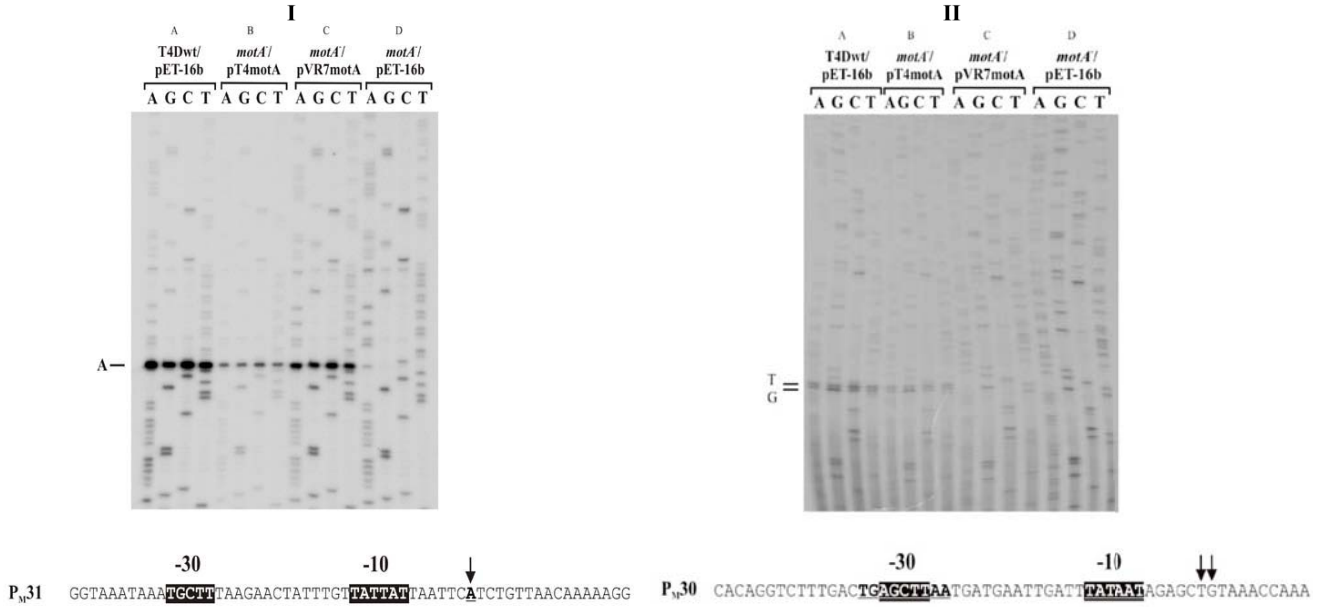

Fig. 11. Functional complementation of plasmid encoded VR7 *mot*A during the growth of T4motA$^-$ *in vivo*. Primer extension sequence analysis was performed using RNA isolated from *E. coli* B_E cells at 6 minute post infection, at 30ºC, with wild type phage T4 (A) and mutant phage T4*mot*A$^-$ (B, C, D). The sequencing lanes are labeled with the dNTP`s used in the sequencing reactions. The initiating nucleotides for the MotA dependent transcripts are noted.

yet fails to activate several T4 P_M containing A at position -31 (e.g. P_M30, Fig. 11 (II)).

As it has been observed in other T4 related bacteriophages, the characteristic T4 late-mode transcription also functions in VR7. In total, 50 VR7 P_L have been identified and the activities of 2 (P_Lg37 and P_LORF129w) have been confirmed by primer extension analysis (data not shown). The deduced consensus sequence of VR7 P_L (Fig. 12) implies that the motif TAAAT between nucleotide positions -11 and -7 is more conserved in the case of late promoters of VR7. In contrast, the T at position -13 is less conserved, and the A at position -12 of several VR7 P_L may be substituted with C.

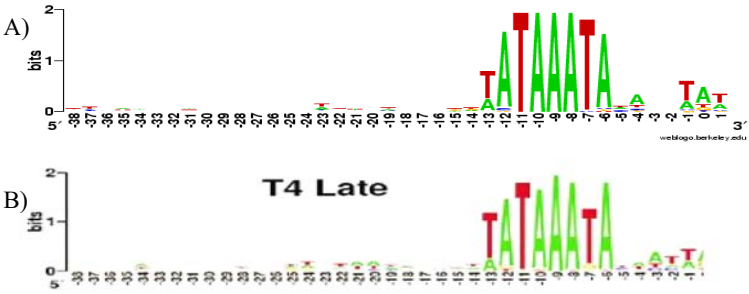

Fig. 12. The comparison of a late promoter consensus sequences of VR7 (A) and T4 (B). The logo of VR7 P_L was obtained using WebLogo program (http://weblogo.berkley.edu), the logo of T4 P_L is from Nolan et al., (2006).

All important proteins for the late transcription activation have been identified in the phage tested. VR7 gp55, gp33 and gp45 share 80%, 64% and 60% amino acid sequence identity with their corresponding counterparts in T4.

In total, 40 potential rho-independent terminators have been identified in VR7. The tetraloop sequence UUCG has been found in 18 of these terminators, while 7 have CGAA. The remaining terminators have noncanonical 4-nucleotide loop sequences (e.g. UUUC) or 3-, 5- and 6-base loop regions. However, the sequence characteristics of terminators mentioned allow us to predict that in most cases these terminators function on both strands.

Therefore, the analysis of the complete genome sequence of bacteriophage VR7 reveales substantial differences both within coding and non coding regions compared with the DNA of T4. Most of these differences, however, are most probably related with physiological properties of VR7 which are rather distinct from those documented for T4.

Investigation of the basic physiological properties of bacteriophages VR5, VR7 and VR20. T4-like phages VR5, VR7 and VR20 were originally isolated from Lithuanian municipal wastewater (VR20) and sewage (VR5 and VR7) as described by Klausa et al., 2003. Electron microscopy of VR5, VR7 and VR20 phage particles showed that all three phages tested resembled the virion morphology of the family *Myoviridae*, A2 morphotype (indistinguishable from phage T4) (Klausa et al., 2003). SDS-PAGE of VR5, VR7 and VR20 virion proteins confirmed their close phylogenetic relatedness to the T4-type bacteriophages.

The host range determination. In order to investigate the host range of VR5, VR7 and VR20 they were tested for the ability to infect different laboratory strains of *E. coli*. In the plaque assay test (Table 3), only the infectivity of bacteriophage VR5 was shown to be similar to that of T4 against all *E. coli* strains used. Meanwhile, phages VR7 and VR20 showed a clear preference for *E. coli* B^E and BL21(DE3) over B40, K12 (or K12 derivatives).

These results suggest that VR7 and VR20 have narrower host range specificity than VR5 or T4.

Table 3. Susceptibility of *E. coli* laboratory strains to VR5, VR7, VR20 and T4 (control).

	B_E	B40	K12	MH1	Nova Blue (DE3)	GM 2163	CR63	BL21 (DE3)
VR5	+	+	+	+	+	+	+	+
VR7	+	+	+*	+*	-	+**	+**	+
VR20	+	+*	+*	+	+*	+	+	+
T4	+	+	+	+	+	+	+	+

+* - very small plaques observed;
+** - diffused plaques observed.

VR5, VR7 and VR20 were also tested on several laboratory isolates of *Klebsiella*, *Aeromonas*, *Pseudomonas* and *Salmonella* (data not shown). None of these bacteria were susceptible to VR5, VR7 or VR20, indicating that these phages have known host range limited only to *E. coli*.

The effect of temperature on the efficiency of plating (e.o.p.) and adsorption. The e.o.p. of VR5, VR7, VR20 and T4 was examined in the temperature range of 7-48ºC. Unlike phage T4, all three VR phages tested had an optimum temperature for plating around 24ºC, but the e.o.p. dropped sharply to 0 between 30 and 37ºC in the case of VR7 and VR20 or between 30 and 40ºC in the case of VR5 (Fig 13).

These results indicate that VR5, VR7 and VR20 are adapted to grow at the lower temperatures of extracolonic environments rather than in the range of mammalian gut temperatures.

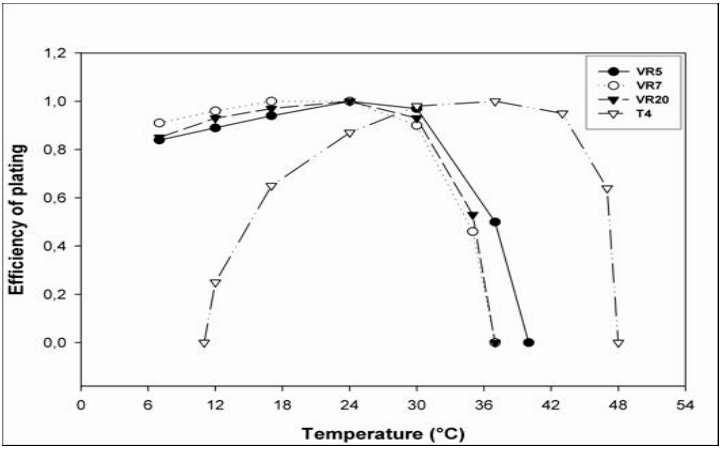

Fig. 13. Effect of temperature on the efficiency of plating of phages VR5, VR7, VR20 and T4 (control). Each point represents the mean of four individual experiments.

There are at least two possible reasons for the phage to fail to form plaques at a certain temperature: failure to adsorb to the host cell, or failure to multiply (Seeley and Primrose, 1980). In order to test the effect of temperature on the ability of phages VR5, VR7 and VR20 to adsorb the cells of *E. coli* B^E,

the adsorption tests were performed at 24, 30 and 35ºC (in the case of VR7 and VR20) or 24, 30 and 37ºC (in the case of VR5). Figure 14a shows that the decrease of the temperature from 37ºC to 24ºC negatively affected both adsorption and multiplication rates of phage T4.

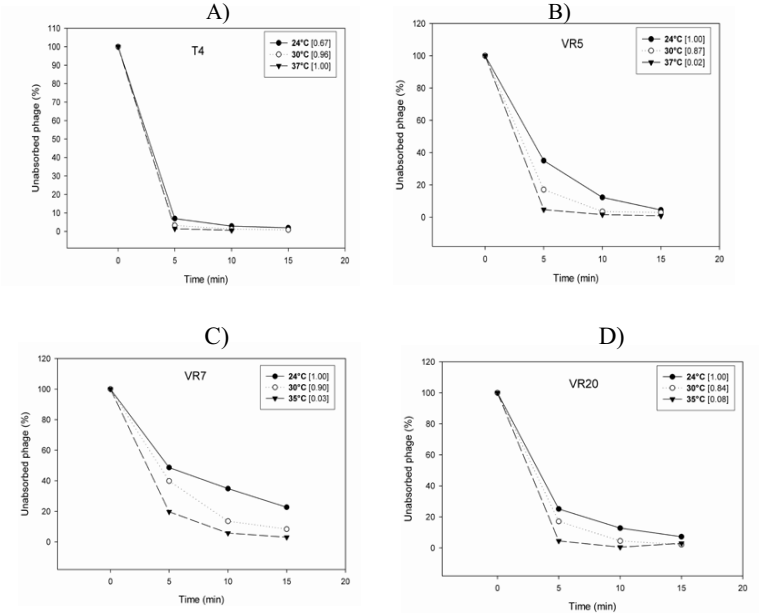

Fig. 14. Effect of temperature on the ability of T4 (a) (control), VR5 (b), VR7 (c) and VR20 (d) to adsorb *E. coli* B^E cells. Each point represents the mean of three individual experiments. *Numbers in square brackets* represent values of viable phage particles formed calculated after plating chloroform treated samples at the end of the rise period. The temperature at which the highest number of viable phage formed was taken as the standard [1.00].

The adsorption rates of VR5, VR7 and VR20 (Fig 14b, 14c, 14d respectively) showed that there was also a certain increase in the adsorption rates of VR5, VR7 and VR20 as the temperature was raised from 24 to 35ºC (or 37ºC). Meanwhile the burst sizes were reduced by 97-98% (in the case of VR7 and VR5) and by 92% (in the case of VR20).

These results suggest that multiplication (and/or injection step) of VR5, VR7 and VR20 has been affected at temperatures above 30ºC. On the other

hand, a slight abortive adsorption profile displayed by VR20 at 35ºC (Fig 14d) implies that the adsorption step at restrictive temperatures in the case of this phage may be also disturbed.

Plaque formation. T-evens generally display small, rough, cloudy bordered plaques (Abedon et al., 2003; Dressman and Drake, 1999). Bacteriophage T4 (as well as T-evens in general) can not produce bursts on stationary-phase cells, and that is why T4 plaques are limited in size, rather than continuing to grow indefinitely as do those of phiX174 and T7 (Kutter et al., 1994; Fort and Méndez, 2002).

Meanwhile, in the range of permissive temperatures, bacteriophages VR5, VR7 and VR20 produced constantly growing plaques, e.g. after 24 h of incubation at 24ºC, bacteriophages VR5, VR7 and VR20 formed plaques of 1.6 (±0.3), 1.5 (±0.5), 0.9 (±0.3) mm in diameter, respectively; after 48h - 1.8 (±0.4), 2.1 (±0.6), 1.75 (±0.2) mm, after 72h - 2.3 (±0.3), 2.5 (±0.5), 2.3 (±0.3) mm and the plaques reached 4.3 (±0.7), 4.8 (±0.7), 3.6 (±0.4) mm in diameter within a period of 6 days. In the case of T4, the growth of plaques at 24ºC could only be observed for 42-48 h reaching the size limit of 1.15 (±0.35) mm in diameter.

These results suggest that bacteriophages VR5, VR7 and VR20 are adapted to grow on stationary-phase cells. This observation has been also supported by the phage growth on stationary-phase cells

Phylogenetic relatedness. Degenerate PCR primers that had been reported to amplify a homologous segment of the g23 sequence in all of the subgroups of T4-type phages were used. PCR products of predicted size (~0, 6 kb) were detected for all three bacteriophages tested. The PCR-generated VR5, VR7 and VR20 DNA fragments were subsequently cloned, sequenced and the corresponding amino acid sequences were deduced. As was expected, all of the nucleotide sequences obtained were similar to an internal region of g23 of T4. The ClustalW2 sequence alignment revealed that PCR-generated fragments of VR5, VR7 and VR20 g23 were closely related to each other (97-98% nucleotide and amino acid sequence identity).

The FASTA-Protein search within NCBI database revealed that the deduced amino acid sequences of the internal part of gp23 of VR5, VR7 and VR20 were most closely related to corresponding part of gp23 in JS98 (95%, 94%, and 94% identical aa respectively). Meanwhile, the amino acid sequence identities shared with the corresponding part of gp23 of T4 were only 84%, 84% and 83% respectively. To determine their phylogenetic relatedness, the deduced sequences of gp23 fragments of VR5, VR7 and VR20 were aligned with the corresponding fragments of the major capsid proteins from 14 T4-type phages available in databases and a neighbor-joining tree was constructed (Fig 15).

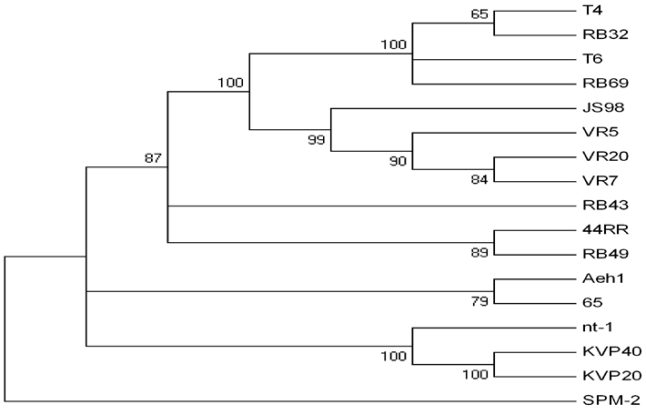

Fig. 15. Neighbor-joining tree based on the alignment of the central part of gp23 (corresponding to 110-303 aa of gp 23 of T4) of VR5, VR7, VR20 and fourteen T4-type bacteriophages available on http://phage.bioc.tulane.edu or NCBI-EBI databases. The evolutionary history was inferred using the Neighbor-Joining method (Saitou and Nei, 1987). The bootstrap consensus tree inferred from 1000 replicates (Felsenstein, 1985) is taken to represent the evolutionary history of the taxa analyzed (Felsenstein, 1985). Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches (Felsenstein, 1985). The evolutionary distances were computed using the Poisson correction method (Zuckerkandl and Pauling, 1965) and are in the units of the number of amino acid substitutions per site. All positions containing gaps and missing data were eliminated from the dataset (Complete deletion option). There were a total of 151 positions in the final dataset. Phylogenetic analyses were conducted in MEGA4 (Tamura et al., 2007).

Based on phylogenetic analysis as well as following all considerations given, bacteriophage VR7 together with phages VR5, VR20 and JS98 form

distinct subgroup of T4-related bacteriophages. Obviously, the members of this subgroup vary both in the DNA content and physiological properties.

CONCLUSIONS

1. The genomic region between genes *30* and *31* of bacteriophages LZ, LZ1 and LZ9 exhibits properties that have already been observed within the T-evens. Meanwhile, the genomic region between genes *30.4* and *30.2* of bacteriophage VR7 contains several structural features that are not common within the T-evens.
2. The complete genome sequence of the T4-like bacteriophage VR7 is 169.285 bp. It encodes 293 open reading frames amnong which 211 encode putative proteins that share amino acid sequence identity with corresponding counterparts in T4. 46 ORFs resemble genes from other T4-like phages, 9 have no T4-type homologues and 27 ORFs lack any database matches.
3. In the course of transcription regulation of VR7, three classes of promoter sequences are activated. Phage VR7 encodes homologues of all transcription-associated proteins of T4. The presence of plasmid encoded VR7 MotA complements the T4*motA*Δ mutant for growth in *E. coli.*
4. Bacteriophages VR5, VR7 and VR20 exhibit physiological properties that have never been observed among the characterized T4-related *E. coli* phages before.
5. Based on phylogenetic analysis, bacteriophages VR5, VR7 and VR20, together with JS98, fall into the distinct subgroup of T4 type bacteriophages

LIST OF PUBLICATIONS

1) Tiemann, B., Depping, R., Gineikiene, E., **Kaliniene, L.**, Nivinskas, R. and Rüger, W. (2004). ModA and ModB, Two ADP-Ribosyltransferases Encoded by Bacteriophage T4: Catalytic Properties and Mutation Analysis. *J. Bacteriology*, 186:7262-7272.

2) **Kalinienė, L**., Truncaitė, L., Arlauskas, A., Stoškienė, G., Zajančkauskaitė, A., Nivinskas, R. (2006). Structural organization of the genomic region between genes *30* and *31* in T4-related bacteriophages LZ, LZ1 and LZ9. *Biologija*, 4:5-9.

3) Truncaitė, L., Stoškienė, G., **Kalinienė, L.**, Strazdaitė, Ž., Nivinskas, R. (2006). *In vivo* and *in vitro* activities of the bacteriophage T4 early promoters. *Biologija,* 3:8-11.

4) **Kalinienė, L.**, Klausa, V., Truncaitė, L. (2010). Low-temperature T4-like coliphages vB_EcoM-VR5, vB_EcoM-VR7 and vB_EcoM-VR20. *Archives of Virology*, (accepted); doi:10.1007/s00705-010-0656-6.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my scientific supervisor Dr. Habil. Rimantas Nivinskas for giving me the opportunity to be involved in research work, for his scientific guidance, valuable advices, patience and support during my doctoral studies.

I am particularly grateful to Dr. Lidija Truncaitė for valuable consultations and assistance while summarizing the results for the doctoral thesis.

I would also like to thank my work colleagues of the Laboratory of Gene Engineering, especially Dr. Aurelija Zajančkauskaitė, Dr. Vytautas Klausa and Nijolė Uždavinienė for fruitful discussions, advices and help in carrying out the experiments.

I am very grateful to Dr. Rolandas Meškys for his valuable remarks and huge help in preparing my doctoral thesis. I would also like to thank the personell of the Department of Molecular Microbiology and Biotechnology for their assistance, support and wonderful neighbourship.

I would like to thank Dr. Edita Sužiedėlienė for her valuable advices and remarks.

I would also like to express my sincere gratitude to Dr. Virginija Bukelskienė for the invaluable help in preparing all the necessary completion documentation.

This work was granted by the stipend (assigned 2006 – 2009) from Lithuanian State Science and Studies Foundation.

REZIUMĖ

Nustačius 169,285 b.p. bakteriofago VR7 genomo nukleotidų seką aptikta viena $tRNR^{MET}$ ir 293 hipotetiniai ASR. Du šimtai vienuolika šio fago genų koduoja baltymus, kurie yra 30% – 97% homologiški atitinkamiems fago T4 baltymams. Keturiasdešimt šeši fago VR7 baltymai neturi analogų T4, bet yra homologiški įvairių kitų T4 gimininių fagų baltymams, 9 baltymai nėra artimi T4-gimininių fagų koduojamiems baltymams, o 27 bakteriofago VR7 ASR koduoja baltymus, kuriems homologų NCBI duomenų bazėje nėra. Fago VR7 genome nėra genų, koduojančių bakteriofago T4 α ir β gliukoziltransferazes (*α-gt* , *β-gt*) , DNR endonukleazes SegA, SegB, SegC, SegD, DNR metilazę Dam, dCMP hidroksimetilazę gp42, atsparumą akriflavinui sąlygojantį baltymą Ac bei ląstelės šeimininkės ^{32} fosforilinime dalyvaujančio *mrh* geno produkto. Nustatyta, kad GC sudaro 40,3% fago VR7 genominės DNR, kai tuo tarpu fago T4 - 35%. Taipogi nustatyta, kad VR7 gp18, gp19 ir gp23 yra tik 71.3% , 69.9% ir 84.9% homologiški atitinkamiems fago T4 baltymams.

Tiriant bakteriofago VR7 transkripcijos reguliaciją buvo aptikti 43 ankstyvieji, 43 vidurinieji bei 44 vėlyvieji promotoriai. Šio fago genominėje DNR taip pat buvo identifikuoti visų fago T4 transkripcijos reguliacijoje dalyvaujančių baltymų homologai.

Klonavus bakteriofago VR7 geną *mot*A buvo atliktas funkcinės komplementacijos tyrimas fago T4motA$^-$ sistemoje *in vivo*. Nustatyta, kad plazmidėje koduojamas fago VR7 viduriniosios transkripcijos aktyvatorius MotA, kurio homologija T4 MotA tėra 34%, atstato fago T4*motA*Δ gyvybingumą.

Ištyrus bakteriofagų VR7, VR5 bei VR20 fiziologines savybes nustatyta, kad šių fagų fiziologinės savybės yra nebūdingos jokiam žinomam T-lyginiam bakteriofagui. Palyginamoji diagnostinio gp23 fragmento analizė parodė, kad VR grupės fagai, kartu su bakteriofagu JS98, formuoja atskirą T4 gimininių bakteriofagų pogrupį.

CURRICULUM VITAE

Laura Kalinienė
Junior research associate, Institute of Biochemistry

Work Address:
Department of Gene Engineering
Institute of Biochemistry
Mokslininkų 12
LT-08662 Vilnius
Fax: (8-5) 272-91-96
E-mail: laura.kaliniene@bchi.lt

Home Address:
Baltupio 75-8
LT-08307 Vilnius
Tel.: (8-5) 271-16-36

Date and Place of Birth: February 11, 1974; Kaunas, Lithuania.

Education:
2001 – BSc, Biology (Molecular biology), Vilnius University, Lithuania.
2003 – MSc Biology (Genetics), Vilnius University, Lithuania.

Professional Experience:
2000-2002 – BSc diploma student/Technician, Institute of Biochemistry;
2002-2003 – MSc diploma student/Head Technician, Institute of Biochemistry;
2003-2005 – Biologist, Institute of Biochemistry;
2005-2009 – Ph.D. student, Institute of Biochemistry, Department of Gene Engineering;
2009- to date – Junior research associate, Institute of Biochemistry.

Research Interests: Microorganism genetics, phage genetics, T4 type bacteriophages, phage genome structure and control of gene expression.

Publications: 5 published scientific articles, including two in ISI-rated journals.

Honours/Scholarships:
2006-2009 – Ph.D. students research scholarship of the Lithuanian State Science and Studies Foundation.

The list of publications (2004-2010):

1. Tiemann, B., Depping, R., Gineikiene, E., **Kaliniene, L**., Nivinskas, R., Rüger, W. (2004). ModA and ModB, two ADP-ribosyltransferases encoded by bacteriophage T4: catalytic properties and mutation analysis. *Journal of Bacteriology*, 186:7262-7272.
2. Truncaitė, L., Stoškienė, G., **Kalinienė, L**., Strazdaitė, Ž., Nivinskas, R. (2006). *In vivo* and *in vitro* activities of the bacteriophage T4 early promoters. *Biologija,* 3:8-11.
3. **Kalinienė, L**., Truncaitė, L., Arlauskas, A., Stoškienė, G., Zajančkauskaitė, A., Nivinskas, R. (2006). Structural organization of the genomic region between genes *30* and *31* in T4-related bacteriophages LZ, LZ1 and LZ9. *Biologija,* 4:5-9.
4. Rubikas, J., **Kalinienė L**. (2006) Gyvybės evoliucijos veiksniai: sąveika, dauginimasis, kitimas, atranka. *Laboratorinė medicina*, 4(32):20-27.
5. **Kalinienė, L.**, Klausa, V., Truncaitė, L. (2010). Low-temperature T4-like coliphages vB_EcoM-VR5, vB_EcoM-VR7 and vB_EcoM-VR20. *Archives of Virology*, (accepted); doi:10.1007/s00705-010-0656-6.