Applied and Environmental Microbiology

Complete Genomic Sequence of Bacteriophage ϕEcoM-GJ1, a Novel Phage That Has Myovirus Morphology and a Podovirus-Like RNA Polymerase

Nidham Jamalludeen, Andrew M. Kropinski, Roger P. Johnson, Erika Lingohr, Josée Harel and Carlton L. Gyles
Appl. Environ. Microbiol. **2008, 74(2):516. DOI: 10.1128/AEM.00990-07.**
Published Ahead of Print 26 November 2007.

APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Jan. 2008, p. 516–525 Vol. 74, No. 2
0099-2240/08/$08.00+0 doi:10.1128/AEM.00990-07

Complete Genomic Sequence of Bacteriophage ϕEcoM-GJ1, a Novel Phage That Has Myovirus Morphology and a Podovirus-Like RNA Polymerase▽

Nidham Jamalludeen,[1] Andrew M. Kropinski,[2,3] Roger P. Johnson,[2] Erika Lingohr,[2] Josée Harel,[4] and Carlton L. Gyles[1]*

Department of Pathobiology, Ontario Veterinary College, University of Guelph, Guelph, Ontario N1G 2W1, Canada[1]; Laboratory for Foodborne Zoonoses, Public Health Agency of Canada, Guelph, Ontario N1G 3W4, Canada[2]; Department of Microbiology and Immunology, Queen's University, Kingston, Ontario K7L 3N6, Canada[3]; and Faculté de Médecine Vétérinaire, Departement de Pathologie et Immunologie, Uníversité de Montréal, Saint-Hyacinthe, Quebec J2S 7C6, Canada[4]

Received 2 May 2007/Accepted 12 November 2007

The complete genome of ϕEcoM-GJ1, a lytic phage that attacks porcine enterotoxigenic *Escherichia coli* of serotype O149:H10:F4, was sequenced and analyzed. The morphology of the phage and the identity of the structural proteins were also determined. The genome consisted of 52,975 bp with a G+C content of 44% and was terminally redundant and circularly permuted. Seventy-five potential open reading frames (ORFs) were identified and annotated, but only 29 possessed homologs. The proteins of five ORFs showed homology with proteins of phages of the family *Myoviridae*, nine with proteins of phages of the family *Podoviridae*, and six with proteins of phages of the family *Siphoviridae*. ORF 1 encoded a T7-like single-subunit RNA polymerase and was preceded by a putative *E. coli* σ^{70}-like promoter. Nine putative phage promoters were detected throughout the genome. The genome included a tRNA gene of 95 bp that had a putative 18-bp intron. The phage morphology was typical of phages of the family *Myoviridae*, with an icosahedral head, a neck, and a long contractile tail with tail fibers. The analysis shows that ϕEcoM-GJ1 is unique, having the morphology of the *Myoviridae*, a gene for RNA polymerase, which is characteristic of phages of the T7 group of the *Podoviridae*, and several genes that encode proteins with homology to proteins of phages of the family *Siphoviridae*.

Postweaning diarrhea (PWD) due to enterotoxigenic *Escherichia coli* (ETEC) is a major problem for the swine industry (2, 17, 36). The ETEC strains that cause this disease typically produce F4 (K88) or F18 fimbriae and belong to a small number of O serogroups, with O149 being dominant worldwide (2, 17, 19). Strains of this serogroup were responsible for recent exceptionally severe outbreaks of disease in pigs in Ontario (2, 26, 38). The outbreak strains were almost all O149:H10:F4, whereas O149 strains from earlier periods were all O149:H43:F4 (38). Since multidrug resistance is common among these ETEC strains (2, 32, 38), and the problem persists despite traditional approaches such as vaccination, antibiotic treatment, feed additives, and management strategies, it was decided to investigate the use of bacteriophages for prophylaxis and therapy.

Phage therapy has been effective in the prevention and treatment of experimentally induced diarrhea due to ETEC in neonatal pigs (46) but has never been investigated for PWD. We therefore used a mixture of 10 strains of O149:H10 ETEC from PWD as host strains in order to isolate phages by standard procedures from pig sewage (9). Six phages (ϕEcoM-GJ1 to ϕEcoM-GJ6) were isolated and characterized on the bases of morphology and spectrum of activity against O149:H10:F4 and O149:H43:F4 ETEC strains, strains of other serotypes of ETEC, and the 72 strains of the ECOR collection. The phages were also compared with respect to DNA fragment patterns resulting from digestion with three restriction enzymes. The six phages were similar morphologically and in their restriction enzyme digestion patterns as well as in their ability to lyse all or almost all O149:H10:F4 ETEC but not O149:H43:F4 ETEC strains. They differed in the percentages of strains of the ECOR collection and of non-O149 ETEC that they lysed (24). One phage (ϕEcoM-GJ1) was selected for further characterization by determination of its genome sequence.

Although over 400 phage genomes have been sequenced and knowledge of phage genomes and gene control mechanisms has accumulated remarkably in recent years, much is unknown about phage genes and their products, because both species numbers and diversity are enormous for phages (7). It is estimated that there are approximately 100 million phage species, with approximately 2 billion different phage open reading frames (ORFs) yet to be discovered. Most of the available sequencing data represent the tailed phages of the order *Caudovirales*. The sequence data, along with proteomic and functional studies of several phages, provide information about phage replication as well as their DNA packaging, morphogenesis, and lysis of host cell walls. The data also provide a good resource for the study of relationships of newly sequenced phages to those whose sequences are in databases.

The genome sequence data for phage ϕEcoM-GJ1 are expected to provide information on the relationship of this phage to other phages whose genomes have been sequenced. Homol-

* Corresponding author. Mailing address: Department of Pathobiology, Ontario Veterinary College, University of Guelph, Guelph, Ontario N1G 2W1, Canada. Phone: (519) 824-4120, ext. 54715. Fax: (519) 824-5930. E-mail: cgyles@uoguelph.ca.

▽ Published ahead of print on 26 November 2007.

ogy data may assist in predicting the likely behavior of the phage. Analysis of the genome sequence will also determine whether the phage is capable of establishing lysogeny, an important consideration for a phage intended for use in prophylaxis and therapy.

The objectives of this study were to obtain the genomic sequence of ϕEcoM-GJ1, analyze the sequence data, and assess the relationship of this phage to five other recently isolated lytic O149:H10:F4 phages (ϕEcoM-GJ2 to ϕEcoM-GJ6) and to previously sequenced phages.

MATERIALS AND METHODS

Bacterial strains. O149:H10:F4 ETEC strain JG280, a hemolytic *E. coli* strains that has genes for LT, STa, STb, and EAST1 enterotoxins, was used for propagation of phages. This strain was isolated in 1999 from an Ontario pig with PWD and was obtained from Gallant Custom Laboratories, Cambridge, Ontario, Canada.

Phages. Six phages that were lytic for a selection of 10 O149:H10:F4 ETEC strains from Ontario pigs were isolated from sewage collected from pig farms in Ontario. The phages were identified as ϕEcoM-GJ1 to ϕEcoM-GJ6. Phage ϕEcoM-GJ1 was examined by electron microscopy of negatively stained preparations. Forty milliliters of a suspension of approximately 10^9 PFU/ml of SM buffer (5.8 g NaCl, 2 g $MgSO_4 \cdot 7H_2O$, 50 ml 1 M Tris [pH 7.5], and 5 ml 2% gelatin per liter of distilled water) were centrifuged in a Beckman XL-90 ultracentrifuge (SW 28 rotor, 25,000 rpm) for 1 h, and the pellet was resuspended in 0.5 ml sterile 0.1 M *N*-(2-hydroxyethyl)-piperazine-*N*′-2-ethanesulfonic acid (HEPES) buffer (Roche Pharmaceuticals, Laval, PQ, Canada). A drop of sample was applied to the surface of a Formvar-coated grid (200 mesh copper grids), negatively stained with 2% uranyl acetate, and then examined in a Leo 912AB Energy Filter transmission electron microscope operated at 100 kV (Guelph Regional STEM Facility, University of Guelph).

Phage DNA extraction. DNA was extracted from the six phages as described by Sambrook and Russell (42). Phages were allowed to completely lyse ETEC strain JG280 in 10 ml of soft agar overlay in each of 10 large petri plates. The agar with the phage was added to an equal volume of SM buffer and held at 4°C for 3 to 4 h with gentle shaking. Following centrifugation at 4,000 × *g* for 15 min, DNase I and RNase I were added to a final concentration of 1 μg/ml of the supernatant, and the mixture was incubated at 37°C for 30 min. Solid NaCl was added to a final concentration of 1 M and dissolved by swirling. The suspension was incubated on ice for 1 h and then centrifuged at 11,000 × *g* for 10 min at 4°C. The supernatant was collected, solid polyethylene glycol (PEG 8000) was added to a final concentration of 10% (wt/vol), and the mixture was stirred slowly at room temperature. After cooling on ice and allowing to stand for 1 h on ice, the phage was pelleted at 14,000 × *g* for 10 min at 4°C and then resuspended in 1 ml TM buffer (50 mM Tris-Cl [pH 7.8], 10 mM $MgSO_4$). Following addition of an equal volume of chloroform and centrifugation, phage particles in the aqueous phase were purified by ultracentrifugation through a glycerol gradient. The phage pellet was resuspended in 1 ml TM buffer. After treatment with DNase, RNase, and EDTA, the phage was lysed with proteinase K and sodium dodecyl sulfate (SDS) at 56°C for 2 h. An equal volume of phenol-chloroform-isoamyl alcohol (25:24:1 [vol/vol]) was added to the sample and mixed. The DNA was collected in the aqueous phase following two rounds of treatment with phenol and chloroform. The DNA was precipitated with ethanol and then dissolved in TE buffer (10 mM Tris, 1 mM EDTA [pH 8.0]).

Sequencing. The DNA of ϕEcoM-GJ1 was sent to Fidelity Systems, Inc, Gaithersburg, MD, for sequencing. This company conducts direct sequencing of genomic DNA, a method that avoids multistep processes, such as PCR and subcloning, and minimizes errors due to DNA amplification. PCR fragments generated from ϕEcoM-GJ1, -GJ2, -GJ3, -GJ4, -GJ5, and -GJ6 were submitted to the University of Guelph Laboratory Services sequencing facility.

PCR evaluation of relationships among ϕEcoM-GJ1 to ϕEcoM-GJ6. Based on the nucleotide sequence of ϕEcoM-GJ1, two primers were designed to amplify a 992-bp EcoRI fragment of this phage, between base pairs 28526 and 29517. The primer sequences were CCTTCATGTCACGTCTTGCC (forward) and GGAGTAGTTGGCACTTGGGC (reverse). The amplification mixture consisted of template DNA (1 μl), primers (1 μl of each at 20 pmol/μl), and Platinum Supermix HiFi (22 μl) (Invitrogen). The conditions that successfully amplified the targeted region in ϕEcoM-GJ1 were 94°C for 3 min, followed by 35 cycles of denaturation at 94°C for 30 s, annealing at 55°C for 30 s, and extension at 68°C for 1 min. There was a final extension at 68°C for 5 min. These conditions were subsequently applied to template DNA from ϕEcoM-GJ2 to ϕEcoM-GJ6.

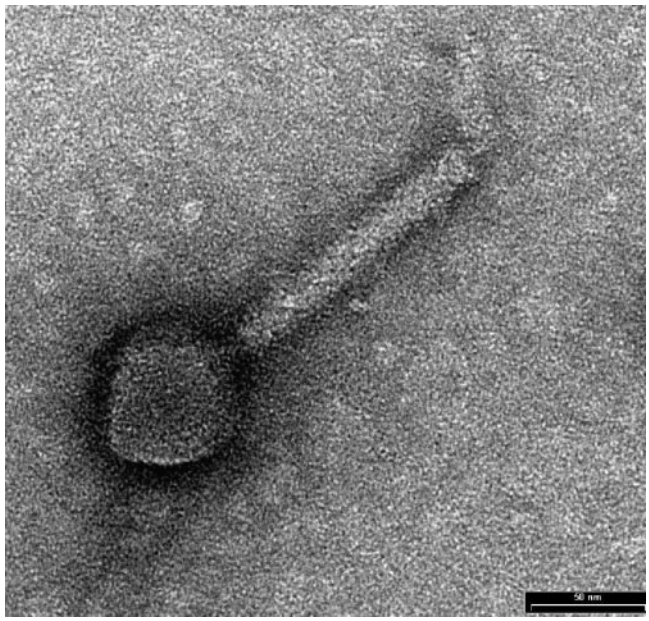

FIG. 1. Electron microscopic appearance of phage ϕEcoM-GJ1 isolated from pig sewage. The phage has an icosahedral head, a neck, and a long tail. Bar, 50 nm.

Structural proteins of ϕEcoM-GJ1. ϕEcoM-GJ1 was purified by a glycerol gradient procedure (42), concentrated by precipitation with 10% polyethylene glycol and 1 M NaCl, and then resuspended in TE buffer. The phage suspension was boiled in sample buffer and subjected to SDS-polyacrylamide gel electrophoresis on a 12.5% gel. Proteins were stained with Simply Blue safe stain (Invitrogen Canada, Burlington, ON). Six intense bands were removed and subjected to in-gel digestion with trypsin (http://www.uoguelph.ca/~bmsf/day1.shtml) at the University of Guelph Biological Mass Spectrometry Facility (Guelph, ON, Canada). The peptides were resolved and sized using a Reflex III matrix-assisted laser desorption ionization–time of flight instrument (Bruker Daltonics Inc., Billerica, MA). The masses of the tryptic peptides were analyzed using Protein Prospector's MS-FiT (12) at http://131.104.190.14/ucsfhtml4.0/msfit.htm against a database of phage proteins. Purification of the phage by the standard cesium chloride procedure was attempted three times, but in each case no phage was recovered.

Bioinformatic analysis. ORFs were identified using Kodon (Applied Maths, Inc., Austin, TX), which allows analysis of all possible start codons. The proteins were characterized by number of amino acids and molecular weight (Lasergene Suite; DNASTAR, Inc., Madison, WI), and screened for homologs using Psi-BLASTP analysis at NCBI (http://www.ncbi.nlm.nih.gov/) (1). FastRNA (bioweb.pasteur.fr/seqanal/interfaces/fastrna.html) (15) and tRNAscan-SE (http://lowelab.ucsc.edu/tRNAscan-SE/) (31) were used to search for tRNA genes. PHIRE (www.agr.kuleuven.ac.be/logt/PHIRE.htm) was used to analyze both strands of the genome in order to identify promoter sequences (30). DNAMAN (Lynnon Corporation, Vaudreuil-Dorion, QC, Canada) was used to identify codon distribution and restriction enzyme sites in the phage genome.

Nucleotide sequence accession number. The sequence of the phage genome has been deposited in GenBank as accession number EF460875.

RESULTS

Morphology of ϕEcoM-GJ1. ϕEcoM-GJ1 has an icosahedral head, a neck, and contractile tail, with tail fibers (Fig. 1). Based on morphology, the phage belongs to the family *Myoviridae*. Six images of the phage were measured, and the mean values were as follows: the head was 63 nm long by 55 nm wide, while the tail was 120 nm long and 23 nm wide.

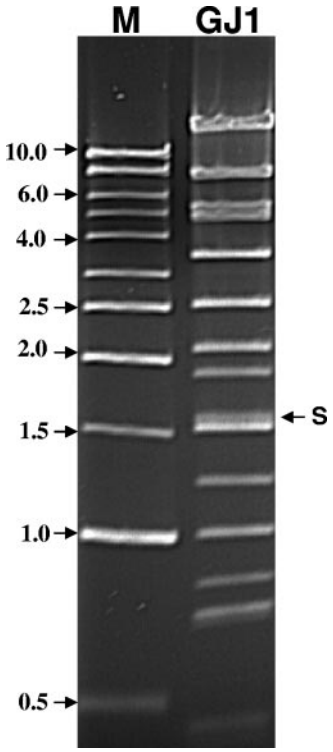

FIG. 2. Digestion of ϕEcoM-GJ1DNA with EcoRI reveals a fuzzy submolar band (S) probably due to the *pac* fragment.

Sequence of ϕEcoM-GJ1. The sequencing of ϕEcoM-GJ1 was completed with the probability of error being less than 1 nucleotide, as determined by using Phred/Phrap/Consed software (www.phrap.org) (16, 21). Fragments of ϕEcoM-GJ1 DNA generated by using three restriction endonucleases (24) match those obtained by in silico digestion. The genome of this phage is circularly permuted, as indicated by the sequence assembly and the presence of a submolar fuzzy band in the EcoRI digestion pattern (Fig. 2). The DNA was opened immediately downstream of ORF 75 (NrdF homolog) prior to annotation. This sequence is 52,975 bp and possesses an average G+C content of 44 mol%. Interestingly, the average G+C content of ORFs 2 to 19 is significantly lower than that of the remainder of the ORFs.

We have reported that ϕEcoM-GJ1 DNA is resistant to digestion by many common restriction enzymes (24), and analysis of the DNA sequence reveals that this is due not to modification but to a lack of many restriction sites. This is reminiscent of the T7-like phages, which commonly lack restriction sites. The DNA is sensitive to the type 1 *E. coli* restriction endonucleases EcoBI (nine sites) and EcoKI (four sites). Interestingly, the product of ϕEcoM-GJ1 ORF 22 possesses low level of homology to an antirestriction protein of *Yersinia pestis* strain KIM5 murine toxin plasmid pMT-1. This may indicate that, like phage T7, ϕEcoM-GJ1 produces an antirestriction protein that blocks the activity of type 1 *E. coli* restriction enzymes (3).

Genome organization. A total of 75 putative ORFs were identified in the genome (Fig. 3; Table 1). A total of 46,461 nucleotides (87.7% of the genome) were involved in coding for putative proteins. The longest stretch of noncoding sequence was 2,067 bp. Most adjacent genes had either a short overlap or 0 to 9 nucleotides between them. Sixty-seven of the proposed genes began with ATG, five started with GTG, and three started with TTG. All of the genes were read from the same DNA strand. One unusual characteristic of the genome of this phage was the high percentage of small ORFs, which were particularly prevalent at the beginning of the genome (Fig. 3). Twenty-nine ORFs were associated with some function and were assigned as putative genes with some specified function based on translated protein sequences and bioinformatic analysis. In several cases, protein homologies were with proteins of phages of the *Podoviridae* or *Siphoviridae* (Table 1). Forty-six ORFs showed no homology to sequences in databases.

The genes are organized into functional clusters (Fig. 3), of which one cluster is involved in the transition from host to phage machinery, another in replication of phage DNA, and a third in phage particle and DNA maturation and packaging (22). The precise delineation of the groups of genes is not possible because of the presence of many genes with no homology or genes that encode proteins of unknown function. However, it is clear that, following gene 1, there is a cluster of small genes reminiscent of those on coliphages T4 involved in

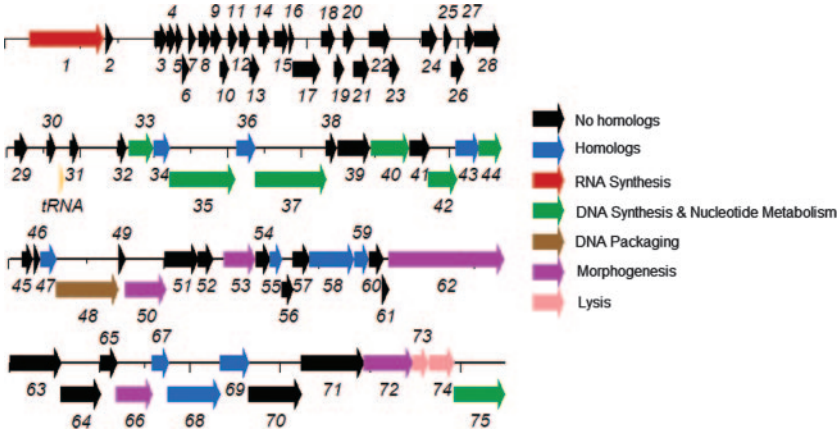

FIG. 3. Gene map for ϕEcoM-GJ1. Genes are color coded with respect to functionality.

TABLE 1. General features of putative ORFs of phage ϕEcoM-GJ1 and homology to proteins in databases[a]

ORF	Coordinates (bp)	Size (aa)	GC content (%)	Mol mass (kDa)	SD sequence (RBS and initiation codon)[b]	Predicted function	Best matches		
							Gene product	No. of identical aa (total)	E value
1	655–2601	648	42.6	72.4	ttggagacttacaaATG	<RNA polymerase	Putative DNA-directed RNA polymerase from *Xanthomonas oryzae* phage OP1 (YP_453591.1)	110 (512)	5e-08
							DNA-dependent RNA polymerase; phage monosubunit type (*Xanthomonas oryzae* phage Xp10) (NP_858979.1)	113 (525)	5e-08
							Putative RNA polymerase from *Pseudomonas aeruginosa* phage gh-1 (NP_813747.1)	103 (466)	2e-04
2	2681–2848	55	40.0	6.3	attggagactttataATG	Hypothetical phage protein	No homologs		
3	4046–4294	82	36.1	9.3	tttggagaataaaATG	Hypothetical phage protein	No homologs		
4	4296–4526	76	39.5	8.8	aaaggattctgaaATG	Hypothetical phage protein	No homologs		
5	4529–4717	62	38.7	7.4	cgaggaaataatcATG	Hypothetical phage protein	No homologs		
6	4714–4875	53	39.0	6.0	ttctggaggcaaaATG	Hypothetical phage protein	No homologs		
7	4872–5045	57	36.2	6.7	atctggagtataaaATG	Hypothetical phage protein	No homologs		
8	5141–5440	99	42.1	11.1	ctaaggtaatcaaATG	Hypothetical phage protein	No homologs		
9	5455–5718	87	42.2	9.8	acttaggagttattATG	Hypothetical phage protein	No homologs		
10	5708–5935	64	43.4	7.5	taggagttttgtaccATG	Hypothetical phage protein	No homologs		
11	5935–6141	68	38.7	7.7	ggtgtacttcaataATG	Hypothetical phage protein	No homologs		
12	6205–6495	96	41.7	10.8	aacttggagcatttATG	Hypothetical phage protein	No homologs		
13	6488–6721	77	39.8	8.8	gagaagattaagcGTG	Hypothetical phage protein	No homologs		
14	6718–7005	95	39.7	10.8	gaaactaaagcaATG	Hypothetical phage protein	No homologs		
15	7136–7516	126	47.9	14.1	agagaatcttaatcATG	Single-stranded DNA binding protein	Single-strand DNA-binding protein SSB (vibriophage VP2) (YP_024430.1)	21 (59)	6e-05
16	7530–7649	39	35.0	4.8	attggagatgccctATG	Hypothetical phage protein	No homologs		
17	7639–8328	229	42.1	25.5	cgtgaggaatatcATG	Hypothetical phage protein	No homologs		
18	8384–8734	116	39.4	13.1	tcactggagcatttATG	Hypothetical phage protein	No homologs		
19	8731–8964	77	39.4	8.8	tgtaaggaagttaaATG	Hypothetical phage protein	No homologs		
20	8961–9236	91	40.3	9.9	tggaaataattttcaATG	Hypothetical phage protein	No homologs		
21	9236–9631	131	44.5	14.9	ttaaagggtaaataATG	Conserved hypothetical phage protein	Hypothetical protein NTHI1519 (*Haemophilus influenzae* 86–028NP) (YP_248988.1)	60 (126)	1e-25
22	9652–10197	181	40.4	21.4	aatggagcatccgATG	Putative antirestriction protein	Antirestriction protein (*Yersinia pestis* KIM) (NP_857871.1)	26 (64)	0.014
23	10190–10429	79	41.4	9.2	agtattgttccgcaATG	Hypothetical phage protein	No homologs		
24	11029–11430	133	44.6	15.8	caaggagcataaATG	Hypothetical phage protein	No homologs		
25	11612–11821	69	37.7	7.7	attggagagttcctATG	Hypothetical phage protein	No homologs		

Continued on following page

TABLE 1—*Continued*

ORF	Coordinates (bp)	Size (aa)	GC content (%)	Mol mass (kDa)	SD sequence (RBS and initiation codon)[b]	Predicted function	Best matches		
							Gene product	No. of identical aa (total)	E value
26	11811–12131	106	43.4	12.2	aaggagttgcaaaATG	Hypothetical phage protein	No homologs		
27	12125–12418	97	42.9	10.7	tccagaaagagcATG	Hypothetical phage protein	No homologs		
28	12420–13094	224	42.6	25.8	caaggccactaaaATG	Hypothetical phage protein	No homologs		
29	13219–13557	112	50.3	12.3	taaggtatatacaaATG	Hypothetical phage protein	No homologs		
30	14095–14337	80	48.7	8.7	aagctatgccaatGTG	Hypothetical phage protein	No homologs		
31	14725–14961	78	42.3	8.6	catctaaggtgatttATG	Hypothetical phage protein	No homologs		
32	16028–16258	76	41.7	8.5	ctggagagaaacATG	Hypothetical phage protein	No homologs		
33	16315–16965	216	45.8	24.8	accggagagcaaATG	Putative thymidylate synthase	Thymidylate synthase (roseophage SIO1) (NP_064749.1)	98 (227)	6e-42
34	16986–17429	147	41.5	16.8	aggaggtctggcgATG	HNH endonuclease	HNH endonuclease F-TflVI (bacteriophage T5) (YP_006874.1)	48 (112)	1e-14
35	17429–19192	587	44.1	65.8	aaggtaaaatctaATG	Helicase/primase	Primase/helicase (*Pseudomonas aeruginosa* phage PaP3) (NP_775217.1)	177 (570)	2e-57
36	19254–19739	161	43.7	18.3	acatggattaagtaATG	Putative HNH endonuclease	Hypothetical protein RB49p040 (enterobacterial phage RB49) (NP_891611.1)	55 (177)	2e-10
37	19739–21673	644	44.2	74.5	ataggaacatataATG	DNA polymerase	DNA polymerase (podovirus GOM) (AAP82219.1)	104 (429)	9e-12
38	21673–21945	90	41.8	10.1	cgcagattcaactaATG	Hypothetical phage protein	No homologs		
39	21976–22842	288	48.6	31.0	aaaggtactcaaaATG	Hypothetical phage protein	No homologs		
40	20876–21913	345	43.4	39.3	cgggagcctttaattTTG	Probable 5′-3′ exonuclease	Exonuclease (*Pseudomonas aeruginosa* phage PaP3) (NP_775229.1) (SP6) (NP_853581.1)	78 (311)	1e-12
41	23925–24449	174	41.3	20.1	taattggagttggaATG	Hypothetical phage protein	No homologs		
42	24439–25194	251	45.3	28.6	agaaagaatcttaATG	Putative DNA ligase	DNA ligase (vibriophage VP4) (YP_249578.1)	45 (164)	3e-04
43	25187–25813	208	45.0	23.5	agtgaggaaagttTTG	Conserved hypothetical phage protein	Hypothetical protein T1p62 (enterobacterial phage T1) (YP_003883.1)	56 (169)	4e-15
44	25816–26415	199	48.6	21.0	tagagaaataatcATG	Putative deoxyuridine 5′-triphosphate nucleotidohydrolase	COG0756: dUTPase (*Rubrivivax gelatinosus* PM1) (ZP_00242767.1)	72 (132)	2e-28
45	26433–26762	109	43.1	12.2	gaacctatctgaaATG	Hypothetical phage protein	No homologs		
46	26818–27003	61	42.1	6.9	tattcggagtcgcttATG	Hypothetical phage protein	No homolog		
47	27023–27514	163	43.1	18.2	gagaacgaatcaATG	Putative HNH endonuclease	RB16 HNH(AP2) 3 (enterobacterial phage RB16) (AAY44388.1)	67 (158)	5e-28
48	27514–29529	671	43.8	75.7	caaggagtctgtgATG	Putative large subunit terminase	Putative terminase large subunit (bacteriophage RTP) (YP_398965.1)	146 (508)	6e-48
49	29532–29744	70	42.4	7.8	gcatggatgtaaatATG	Hypothetical phage protein	No homologs		
50	29744–31060	438	44.9	49.0	taaggaagaaataATG	Putative portal protein	Putative portal protein (enterobacterial phage T1) (YP_003893.1)	73 (396)	9e-07

Continued on following page

TABLE 1—*Continued*

ORF	Coordinates (bp)	Size (aa)	GC content (%)	Mol mass (kDa)	SD sequence (RBS and initiation codon)[b]	Predicted function	Best matches		
							Gene product	No. of identical aa (total)	E value
51	31029–32093	354	44.9	38.9	agggtaacgcaaGTG	Conserved hypothetical phage protein	Hypothetical protein Aaphi23p33 (bacteriophage Aaphi23) (NP_852758.1)	68 (210)	4e-19
52	32103–32576	157	50.5	16.2	ataaggtaagacaATG	Hypothetical phage protein	No homologs		
53	32925–33932	335	49.9	36.7	atggattaaactacATG	Conserved hypothetical phage protein	Hypothetical protein T1p47 (enterobacterial phage T1) (YP_003898.1)	84 (302)	4e-18
54	33976–34416	146	44.3	16.2	ttaagagaaatgtaATG	Hypothetical phage protein	No homologs		
55	34417–34806	129	46.7	14.5	gataagttgggtaaATG	Conserved hypothetical phage protein	Hypothetical protein XF1684 (*Xylella fastidiosa* 9a5c) (NP_298973.1)	42 (124)	4e-09
56	34803–35165	120	43.6	13.8	cctgggtcacagttTTG	Hypothetical phage protein	No homologs		
57	35162–35674	170	42.7	19.2	aggagttagagaaATG	Hypothetical phage protein	No homologs		
58	35675–37123	482	49.0	50.8	ttaagggaatctaaATG	Conserved hypothetical phage protein	Hypothetical protein Aaphi23p39 (bacteriophage Aaphi23) (NP_852765.1)	116 (472)	7e-19
59	37136–37591	152	48.1	16.4	aggggtgcgataaGTG	Conserved hypothetical phage protein	Unknown (bacteriophage Felix O1) (NP_944897.1)	39 (138)	2e-06
60	37604–38062	152	43.9	17.2	taatttggagtaagtATG	Hypothetical phage protein	No homologs		
61	38017–38238	73	42.9	8.4	ggcggaggaagcATC	Hypothetical phage protein	No homologs		
62	38231–41950	1239	47.2	133.2	gtagatgaagacaATG	Internal phage protein	Hypothetical protein predicted by GeneMark (*Bordetella* phage BIP-1) (NP_996621.1)	42 (118)	5e-10
63	42023–43132	369	43.3	40.7	ataggtatattgcaATG	Hypothetical phage protein	No homologs		
64	43132–44019	295	44.0	31.1	aatggagtcattttaATG	Hypothetical phage protein	No homologs		
65	44016–44378	120	40.6	13.7	agggacgtatcctATG	Hypothetical phage protein	No homologs		
66	44371–45167	264	46.8	28.1	gagtgtacttgaacGTG	Putative baseplate assembly protein	Putative bacteriophage protein (*Burkholderia vietnamiensis* G4) (ZP_00426837.1)	79 (255)	1e-18
67	45166–45537	123	45.6	13.4	ggccgatgaaataATG	Conserved hypothetical phage protein	Hypothetical protein NTHI1549 (*Haemophilus influenzae* 86–028NP) (YP_249016.1)	30 (90)	2e-07
68	45513–46670	385	45.2	41.2	acggaattcttaacATG	Conserved hypothetical phage protein	Putative bacteriophage protein (*Salmonella enterica* subsp. *enterica* serovar Typhi strain T18 (NP_456376.1)	113 (383)	5e-34
69	46671–47312	213	43.2	23.6	ataacggtgtaacATG	Hypothetical phage protein	No homologs		
70	47312–48460	382	43.6	42.2	ctccggagaaataATG	Hypothetical phage protein	No homologs		
71	48460–49857	465	46.1	50.0	ttagtgaggccataATG	Hypothetical phage protein	No homologs		
72	49866–50936	356	50.2	37.3	tttataggattcaaaATG	Putative tail fiber protein	COG5301: phage-related tail fiber protein (*Escherichia coli* 101–1) (ZP_00923911.1)	91 (273)	3e-13

Continued on following page

TABLE 1—*Continued*

ORF	Coordinates (bp)	Size (aa)	GC content (%)	Mol mass (kDa)	SD sequence (RBS and initiation codon)[b]	Predicted function	Best matches		
							Gene product	No. of identical aa (total)	E value
73	50944–51285	113	47.7	12.3	gcttaaggaaatcATG	Putative holin	Probable holin gp13 (*Acyrthosiphon pisum* bacteriophage APSE-2) (ABA29371.1)	28 (82)	0.002
74	51303–51857	184	40.4	20.6	ataaggaactcgaATG	Probable endolysin	Hypothetical protein T5.083 (bacteriophage T5) (YP_006911.1)	45 (136)	1e-10
75	51857–52975	372	42.9	42.5	ctaggaaatctgtaATG	Ribonucleotide reductase, beta subunit	COG0208: ribonucleotide reductase, beta subunit (*Escherichia coli* E24377A) (ZP_00700291.1)	301 (376)	2e-174

[a] Abbreviations: aa, amino acid; SD, Shine-Dalgarno; RBS, ribosome binding site.
[b] Start codons are in capital letters.

host takeover. Further downstream we find genes that are clearly involved in DNA replication and repair. These include proteins involved in nucleotide metabolism (Gp33, Gp40, and Gp44) along with primase/helicase (Gp35), DNA polymerase (Gp37), single-stranded DNA binding protein (Gp15), and DNA ligase (Gp42). Interestingly, the significantly downstream gene 75 encodes a ribonucleotide reductase, which is usually associated with early transcriptional events. This phage genome also contains two genes (ORFs 34 and 47) with homology to HNH endonucleases. While these are sometimes found within introns, there is no evidence for this in the case of the thymidylate synthase gene (ORF 33), nor does ORF 47 obscure identification of the gene encoding a small-subunit terminase. The lack of the latter has also been observed in the P2-like phages, mycobacteriophage D29 (18), and *Lactococcus* phage r1t (48). This cluster is followed by genes associated with DNA packaging (Gp50) and phage morphogenesis (genes 50, 53, 66, and 72). Finally, the adjacent genes 73 and 74 constitutes a lysis cassette. ORF 73 encodes a 113-amino-acid protein with three potential transmembrane domains and has homology to the class 1 holins. The putative lysin homolog, gp74, showed low-level sequence similarity to bacterial murein hydrolases and minimal homology to other phage lysins. It contains a pfam07486 (Hydrolase_2, cell wall hydrolase) domain and is likely an endolysin. ORF 75 encodes a homolog of the *E. coli* beta subunit of ribonucleotide reductase, but none of the genes in this phage encode a product with homology to the alpha subunit of this enzyme.

Putative tRNA gene with intron. A tRNA gene of 95 bp with a putative 18-bp intron was identified between positions 14457 and 14551. The sequence of this gene is GGACGAGTAGCT CAGAAGGTTTGCATTGTGCTCAGTCAGGTAGAGCA ATCTGT**TCG**ACAGATGTGTCGTTGGTTCGAATCCAA CCTTGTCCGCCA, where the anticodon is in bold letters. The Cove score of the anticodon arm is 32.9. When the suspected intron (underlined) is removed, the Cove score increases to 60.9. Analysis of the distribution of codons in the appropriate frame of the ϕEcoM-GJ1 sequence showed that the codons that specify arginine were used as follows: CGA, 22.1%; CGG, 15.3%; AGA, 41.4%; and AGG, 21.2%. This contrasts with the distribution in *E. coli* (CGA, 22%; CGG, 36.4%; AGA, 26.2%; and AGG, 15.4% [37]) but does not explain the physiological significance to the phage of possessing an $tRNA^{Arg}_{CGA}$.

Transcription. The unique feature of this phage is the presence of a gene encoding a single-subunit RNA polymerase (RNAP) phylogenetically related to the RNAP of *Xanthomonas* phage XP10 and more distantly to those of the T7-like phages (Table 1). A putative *E. coli* σ^{70}-like promoter $TTGACGN_{18}TATTAT$ displaying strong conservation to the consensus promoter ($TTGACAN_{15-17}TATAAT$) was detected upstream of ORF 1. The DNA sequence was scanned for potential phage promoters using the software program PHIRE (30). Nine potential phage promoter sequences with the consensus sequence aArCATTAGTTGGATRTGRr (lowercase indicates preferred, but not conserved, nucleotides) were recognized, all of which were located upstream of ϕEcoM-GJ1 genes (Table 2).

Mass spectrometry data. Unambiguous data were obtained for identification of two bands. The mass spectrometry data are consistent with the 133-kDa and 35-kDa proteins (Fig. 4) being products of gene 62 and gene 53, respectively (Table 1). The theoretical size of the product of gene 53 is 37 kDa (Table 1). Interestingly, Gp62 contains a cd00254 (LT_GEWL) lytic transglycosidase domain and homology to Gp16 of the T7-like phages and Gp15 of *Salmonella* phage ε15. The former protein, also called internal virion protein D, has been shown to be

TABLE 2. Promoter sequences recognized in the genome of phage ϕEcoM-GJ1

Gene	Sequence	Location	
		Beginning	End
3	CGACATTAGTTGGATGTGAA	02897	02916
3	TATCATTAGTTGCATGTGAA	03833	03852
15	CAACATTAGTTGGATGTGAG	07044	07063
24	AAACATTAGTGGGATTTGAT	10496	10515
24	AAACATTAGTTGGAATTGAA	10637	10656
29	AAGAATTAGTTGGAAGAGGA	12791	12810
29	AAACATTAGTTGGATGTGGA	13123	13142
39	AAGCATTAGTTGGATATGAC	21897	21916
63	TAACATTAGTTGAATATGTA	41981	42000

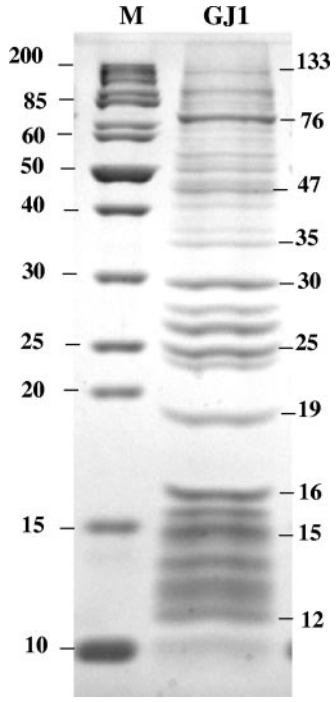

FIG. 4. SDS-polyacrylamide gel electrophoresis of ϕEcoM-GJ1. The phage was purified by a glycerol gradient procedure, as repeated attempts at purification with cesium chloride were unsuccessful. The 133-kDa and 35-kDa bands were identified by mass spectrometry as being consistent with the products of genes 62 and 53, respectively. Lane M, molecular weight markers.

part of the virion injection machinery (35). In phage ε15, Gp15 has also been shown to be an internal protein (25). The homology of the protein product of ORF 53 with coliphage T1 Gp47 (41) confirms that this is the major capsid protein. Surprisingly, the 35-kDa band (Fig. 4) is a weak band, unlike the prominent bands typical of the major capsid protein of other phages. It is possible that this protein breaks down into lower-molecular-weight structures, many of which are very prominent.

Sequencing data from PCR fragments of ϕEcoM-GJ2 to -GJ6. Fragment sequence alignment data (not shown) supported previous pulsed-field gel electrophoretic data (24) indicating that all six phages were highly related, since a 1-kb fragment selected at random was 93 to 94% identical among all six phages.

DISCUSSION

ϕEcoM-GJ1 is unique, since it is the first member of the *Myoviridae* which apparently possesses a coliphage T7-like transcriptional system. Experimentally, its genome appears to be circularly permuted and terminally redundant. This was verified by restriction analysis with EcoRI, which revealed a submolar fragment indicative of the *pac* fragment (Fig. 2). In addition, Casjens and Gilcrease have shown that the phylogeny of the *Caudovirales* large-subunit terminase proteins is correlated with the virus packaging strategy (10). Since the sequence of ϕEcoM-GJ1 Gp48 is related to that of the terminases from T1-like phages, which are known to package DNA by a headful mechanism, we assumed that phage ϕEcoM-GJ1 behaves similarly (20).

The gene for a phage-specific single-subunit RNAP, considered a hallmark of phages of the T7 family of phages (13, 29, 44, 45), is present in ϕEcoM-GJ1. The ϕEcoM-GJ1 RNAP has highest sequence similarity to the RNAPs of *Xanthomonas oryzae* phages OP1 and Xp10, which are members of the *Siphoviridae* (23, 49). There is also significant homology with the RNAP of the T7-like virus *Pseudomonas* phage gh-1 (Table 1).

The gene for the RNAP in ϕEcoM-GJ1 is relatively similar in its genomic location to the gene in phage T7. This location appears to be important for optimal function of the genome, since Springman et al. (47) found that when they moved it to an ectopic location in phage T7, the phage was severely debilitated, as indicated by a marked decline in doublings per hour. However, the location does not appear to be necessary for a high burst size, as Lavigne et al. (29) noted that the *Pseudomonas aeruginosa* phage ϕKMV has the T7-like RNAP gene located among the DNA metabolism genes but exhibits a burst size similar to that of phage T7.

Promoter sequences that are recognized by the phage-encoded RNAP are found in all the members of the T7 group for which genome data are available (11). The genome of ϕEcoM-GJ1 has nine genes with upstream motifs reminiscent of T7-like promoters (11, 29). ϕEcoM-GJ1 is also similar to members of the T7 supergroup of phages (T7, SP6 and ϕKMV) in that all the predicted genes are transcribed from one strand (45).

A putative tRNA gene specific for the codon CGA was detected. The Cove score of 32.9 is well above the cutoff of 20 and suggests that this gene encodes an authentic tRNA (14). tRNA genes have been detected in several members of the *Myoviridae*, including the enterobacterial phages T4 (34) and T5 (43) and *Pseudomonas* phage ϕKZ (33). However, there are no potential tRNA genes in members of the T7-like viruses. A recent analysis identified a significant relationship between tRNA distribution and codon usage and provided evidence for a role in synthesis of phage proteins (4). The tRNA gene in ϕEcoM-GJ1 is specific for arginine (anticodon TCG), which corresponds to a codon (CGA) that is used 22% of the time in the ϕEcoM-GJ1 genome. The genome of *E. coli* O157:H7 strain Sakai contains four tRNAs using this anticodon; hence, the presence of this specific tRNA gene in ϕEcoM-GJ1 cannot readily be explained.

The presence of an intron in the putative tRNA gene is to our knowledge the first such example in a phage genome. Group 1 introns that are self-splicing are found in a number of tRNA genes of alphaproteobacteria and cyanobacteria (6, 39, 40), and similar introns have been found in protein-encoding genes but not tRNA genes of phages (5, 27, 28). The significance of the intron in the tRNA gene is unknown.

Yet another unique feature of phage ϕEcoM-GJ1 is that it encodes the beta but not the alpha subunit of ribonucleotide reductase. It is possible that the phage uses the alpha subunit of its *E. coli* host, but there is no explanation of why it would not use both subunits of the host enzyme.

No genes with homology to bacterial toxins were present in the genome of ϕEcoM-GJ1. The presence of phage-encoded toxin genes, such as genes for cholera, diphtheria, Shiga, and botulinum C1 and D toxins (8), would have disqualified ϕEcoM-GJ1 as a candidate for phage therapy. In addition, no

integrases was detected. This is an important finding, since therapeutic and decontaminating phages should not be capable of integrating into the host chromosome to form lysogens.

The classification of phage ϕEcoM-GJ1 presents a conundrum. Chen and Schneider (11) classified T7-like phages based on their having an RNAP plus or minus T7-like promoters. ϕEcoM-GJ1 is thus clearly a member of the T7 group, since it possesses the gene for a single-subunit RNAP (Gp1) associated with a T7-like promoter. Among the phages that possess a single-subunit RNAP but lack T7-like promoters, the *Xanthomonas oryzae* phage XP10 has been described as "an odd T-odd phage" because it is morphologically a member of the *Siphoviridae*, rather than the *Podoviridae*, to which phage T7 and other T7-like phages belong (49). ϕEcoM-GJ1 is unique in that its genome organization and transcription superficially resemble those of the T7 superfamily of phages but its morphology clearly places it in the *Myoviridae*, strongly suggesting that it is the first member of a new genus of bacteriophages.

ACKNOWLEDGMENTS

This work was supported by the NSERC Canadian Research Network on Bacterial Pathogens of Swine and the Ontario Ministry of Agriculture, Food and Rural Affairs (OMAFRA) to C.L.G. A.M.K. was supported by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada.

Bob Harris assisted with electron microscopy, and Dyanne Brewer, in the University of Guelph Biological Mass Spectrometry Facility, carried out the mass spectroscopy.

REFERENCES

1. **Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389–4022.
2. **Amezcua, R., R. M. Friendship, C. E. Dewey, C. Gyles, and J. M. Fairbrother.** 2002. Presentation of postweaning *Escherichia coli* diarrhea in southern Ontario, prevalence of hemolytic *E. coli* serogroups involved, and their antimicrobial resistance patterns. Can. J. Vet. Res. **66:**73–78.
3. **Atanasiu, C., T. J. Su, S. S. Sturrock, and D. T. Dryden.** 2002. Interaction of the ocr gene 0.3 protein of bacteriophage T7 with EcoKI restriction/modification enzyme. Nucleic Acids Res. **30:**3936–3944.
4. **Bailly-Bechet, M., M. Vergassola, and E. Rocha.** 2007. Causes for the intriguing presence of tRNAs in phages. Genome Res. **17:**1486–1495.
5. **Belle, A., M. Landthaler, and D. A. Shub.** 2002. Intronless homing: site-specific endonuclease SegF of bacteriophage T4 mediates localized marker exclusion analogous to homing endonucleases of group I introns. Genes Dev. **16:**351–362.
6. **Bonocora, R. P., and D. A. Shub.** 2001. A novel group I intron-encoded endonuclease specific for the anticodon region of tRNA(fMet) genes. Mol. Microbiol. **39:**1299–1306.
7. **Breitbart, M., and F. Rohwer.** 2005. Here a virus, there a virus, everywhere the same virus? Trends Microbiol. **13:**278–284.
8. **Brüssow, H., C. Canchaya, and W. D. Hardt.** 2004. Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. Microbiol. Mol. Biol. Rev. **68:**560–602.
9. **Cappuccino, J. G., and N. Sherman.** 1992. Microbiology: a laboratory manual. The Benjamin/Cummings Publishing Company, Inc., San Francisco, CA.
10. **Casjens, S. R., and E. B. Gilcrease.** Determining DNA packaging strategy by analysis of the termini of the n chromosomes in tailed-bacteriophage virions. *In* M. R. J. Clokie and A. M. Kropinski (ed.), Bacteriophages: methods and protocols, in press. Humana Press, Totowa, NJ.
11. **Chen, Z., and T. D. Schneider.** 2005. Information theory based T7-like promoter models: classification of bacteriophages and differential evolution of promoters and their polymerases. Nucleic Acids Res. **33:**6172–6187.
12. **Clauser, K. R., P. R. Baker, and A. L. Burlingame.** 1999. Role of accurate mass measurement (±10 ppm) in protein identification strategies employing MS or MS/MS and database searching. Anal. Chem. **71:**2871–2882.
13. **Dunn, J. J., and F. W. Studier.** 1983. Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. J. Mol. Biol. **166:**477–535.
14. **Eddy, S. R., and R. Durbin.** 1994. RNA sequence analysis using covariance models. Nucleic Acids Res. **22:**2079–2088.
15. **El-Mabrouk, N., and F. Lisacek.** 1996. Very fast identification of RNA motifs in genomic DNA. Application to tRNA search in the yeast genome. J. Mol. Biol. **264:**46–55.
16. **Ewing, B., L. Hillier, M. C. Wendl, and P. Green.** 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res. **8:**175–185.
17. **Fairbrother, J. M., E. Nadeau, and C. L. Gyles.** 2005. *Escherichia coli* in postweaning diarrhea in pigs: an update on bacterial types, pathogenesis, and prevention strategies. Anim. Health Res. Rev. **6:**17–39.
18. **Ford, M. E., G. J. Sardis, A. E. Belanger, R. W. Hendrix, and G. F. Hatfull.** 1998. Genome structure of mycobacteriophage D29: implications for phage evolution. J. Mol. Biol. **279:**143–164.
19. **Frydendahl, K., J. T. Kare, A. J. Strodl, M. Fredholm, and G. Evans.** 2003. Association between the porcine *Escherichia coli* F18 receptor genotype and phenotype and susceptibility to colonisation and postweaning diarrhoea caused by *E. coli* O138:F18. Vet. Microbiol. **93:**39–51.
20. **German, G. J., R. Misra, and A. M. Kropinski.** 2006. The T1-like bacteriophages, p. 211–224. *In* R. L. Calendar (ed.), The bacteriophages. Oxford University Press, New York, NY.
21. **Gordon, D., C. Abajian, and P. Green.** 1998. Consed: a graphical tool for sequence finishing. Genome Res. **8:**195–202.
22. **Hausmann, R.** 1976. Bacteriophage T7 genetics. Curr. Top. Microbiol. Immunol. **75:**77–110.
23. **Inoue, Y., T. Matsuura, T. Ohara, and K. Azegami.** 2006. Bacteriophage OP_1, lytic for *Xanthomonas oryzae* pv. *oryzae*, changes its host range by duplication and deletion of the small domain in the deduced tail fiber gene. J. Gen. Plant Pathol. **72:**111–118.
24. **Jamalludeen, N., R. P. Johnson, R. Friendship, A. M. Kropinski, E. Lingohr, and C. L. Gyles.** 2007. Isolation and characterization of nine bacteriophages that lyse O149 enterotoxigenic *Escherichia coli*. Vet. Microbiol. **124:**47–57.
25. **Jiang, W., J. Chang, J. Jakana, P. Weigele, J. King, W. Chiu, W. Jiang, J. Chang, J. Jakana, P. Weigele, J. King, and W. Chiu.** 2006. Structure of epsilon15 bacteriophage reveals genome organization and DNA packaging/injection apparatus. Nature **439:**612–616.
26. **Josephson, G., M. Archambeault, and C. L. Gyles.** 2000. Colibacillosis in pigs in 1999. AHL Newsl. **4:**8–9.
27. **Landthaler, M., U. Begley, N. C. Lau, and D. A. Shub.** 2002. Two self-splicing group I introns in the ribonucleotide reductase large subunit gene of *Staphylococcus aureus* phage Twort. Nucleic Acids Res. **30:**1935–1943.
28. **Landthaler, M., and D. A. Shub.** 2003. The nicking homing endonuclease I-BasI is encoded by a group I intron in the DNA polymerase gene of the *Bacillus thuringiensis* phage Bastille. Nucleic Acids Res. **31:**3071–3077.
29. **Lavigne, R., M. V. Burkal'tseva, J. Robben, N. N. Sykilinda, L. P. Kurochkina, B. Grymonprez, B. Jonckx, V. N. Krylov, V. V. Mesyanzhinov, and G. Volckaert.** 2003. The genome of bacteriophage fKMV, a T7-like virus infecting *Pseudomonas aeruginosa*. Virology **312:**49–59.
30. **Lavigne, R., W. D. Sun, and G. Volckaert.** 2004. PHIRE, a deterministic approach to reveal regulatory elements in bacteriophage genomes. Bioinformatics **20:**629–6135.
31. **Lowe, T. M., and S. R. Eddy.** 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. **25:**955–964.
32. **Maynard, C., J. M. Fairbrother, S. Bekal, F. Sanschagrin, R. C. Levesque, R. Brousseau, L. Masson, S. Lariviere, and J. Harel.** 2003. Antimicrobial resistance genes in enterotoxigenic *Escherichia coli* O149:K91 isolates obtained over a 23-year period from pigs. Antimicrob. Agents Chemother. **47:**3214–3221.
33. **Mesyanzhinov, V. V., J. Robben, B. Grymonprez, V. A. Kostyuchenko, M. V. Bourkaltseva, N. N. Sykilinda, V. N. Krylov, and G. Volckaert.** 2002. The genome of bacteriophage ϕKZ of *Pseudomonas aeruginosa*. J. Mol. Biol. **317:**1–19.
34. **Miller, E. C., E. Kutter, G. Mosig, F. Arisaka, T. Kunisawa, and W. Rüger.** 2003. Bacteriophage T4 genome. Microbiol. Mol. Biol. Rev. **67:**86–156.
35. **Molineux, I. J.** 2001. No syringes please, ejection of phage T7 DNA from the virion is enzyme driven. Mol. Microbiol. **40:**1–8.
36. **Nagy, B., and P. Z. Fekete.** 2005. Enterotoxigenic *Escherichia coli* in veterinary medicine. Int. J. Med. Microbiol. **295:**443–454.
37. **Nakamura, Y., T. Gojobori, and T. Ikemura.** 2000. Codon usage tabulated from the international DNA sequence databases: status for the year 2000. Nucleic Acids Res. **28:**292.
38. **Noamani, B. N., J. M. Fairbrother, and C. L. Gyles.** 2003. Virulence genes of O149 enterotoxigenic *Escherichia coli* from outbreaks of postweaning diarrhea in pigs. Vet. Microbiol. **97:**87–101.
39. **Paquin, B., A. Heinfling, and D. A. Shub.** 1999. Sporadic distribution of $tRNA_{CCU}^{Arg}$ introns among α-purple bacteria: evidence for horizontal transmission and transposition of a group I intron. J. Bacteriol. **181:**1049–1053.
40. **Paquin, B., S. D. Kathe, S. A. Nierzwicki-Bauer, and D. A. Shub.** 1997. Origin and evolution of group I introns in cyanobacterial tRNA genes. J. Bacteriol. **179:**6798–6806.
41. **Roberts, M. D., N. L. Martin, and A. M. Kropinski.** 2004. The genome and proteome of coliphage T1. Virology **318:**245–266.
42. **Sambrook, J., and D. W. Russell.** 2001. Molecular cloning: a laboratory manual. Cold Spring Harbor Press, Cold Spring Harbor, NY.

43. **Sayers, J. R.** 2006. Bacteriophage T5, p. 268–276. *In* R. Calendar (ed.), The bacteriophages. Oxford University Press, New York, NY.
44. **Scholl, D., S. Adhya, and C. R. Merril.** 2002. Bacteriophage SP6 is closely related to phages K1-5, K5, and K1E but encodes a tail protein very similar to that of the distantly related P22. J. Bacteriol. **184:**2833–2836.
45. **Scholl, D., J. Kieleczawa, P. Kemp, J. Rush, C. C. Richardson, C. Merril, S. Adhya, and I. J. Molineux.** 2004. Genomic analysis of bacteriophages SP6 and K1-5, an estranged subgroup of the T7 supergroup. J. Mol. Biol. **335:** 1151–1171.
46. **Smith, H. W., and M. B. Huggins.** 1983. Effectiveness of phages in treating experimental *Escherichia coli* diarrhoea in calves, piglets and lambs. J. Gen. Microbiol. **129:**2659–2675.
47. **Springman, R., M. R. Badgett, I. J. Molineux, J. J. Bull., R. Springman, M. R. Badgett, I. J. Molineux, and J. J. Bull.** 2005. Gene order constrains adaptation in bacteriophage T7. Virology **341:**141–152.
48. **van Sinderen, D., H. Karsens, J. Kok, P. Terpstra, M. H. Ruiters, G. Venema, and A. Nauta.** 1996. Sequence analysis and molecular characterization of the temperate lactococcal bacteriophage r1t. Mol. Microbiol. **19:**1343–1355.
49. **Yuzenkova, J., S. Nechaev, J. Berlin, D. Rogulja, K. Kuznedelov, R. Inman, A. Mushegian, K. Severinov, J. Yuzenkova, S. Nechaev, J. Berlin, D. Rogulja, K. Kuznedelov, R. Inman, A. Mushegian, and K. Severinov.** 2003. Genome of *Xanthomonas oryzae* bacteriophage Xp10: an odd T-odd phage. J. Mol. Biol. **330:**735–748.