Complete genome sequence of bacteriophage vB_YenP_AP5 which infects Yersinia enterocolitica of serotype O:3

Carlos G Leon-Velarde1,5*, Andrew M Kropinski2,3, Shu Chen1, Arash Abbasifar4, Mansel W Griffiths4,5 and Joseph A Odumeru5

Abstract

Background: Bacteriophage vB_YenP_AP5 is a lytic bacteriophage capable of infecting Yersinia enterocolitica strains of serotype O:3, an epidemiologically significant serotype within this bacterial species that causes yersiniosis in humans. This work describes the complete genome sequence of this phage.

Results: The genome consists of linear double-stranded DNA of 38,646 bp, with direct terminal repeats of 235 bp in length, and a GC content of 50.7%. There are 45 open reading frames which occupy 89.9% of the genome. Most of the proteins encoded by this virus exhibit sequence similarity to Yersinia phage φYeO3-12 and Salmonella phage φSG-JL2 proteins.

Conclusions: Genomic and morphological analyses place the bacteriophage vB_YenP_AP5 in the T7likevirus genus of the subfamily Autographivirinae within the family Podoviridae.

Background

Yersinia enterocolitica, a facultatively anaerobic, Gram-negative, non-sporulating, short bacillus, is an important zoonotic pathogen leading to human and animal enteric infection [1]. Among the species of the genus Yersinia, Y. enterocolitica is considered highly heterogeneous and is grouped into a biochemical scheme composed of six biotypes divided into three lineages: avirulent strains belonging to biotype 1A, highly pathogenic strains of biotype 1B, and weakly pathogenic strains of biotypes 2–5 that do not kill mice [2,3]. Most strains associated with yersiniosis belong to bioserotypes 1B/O:8, 2/O:5,27, 2/O:9, 3/O:3, and 4/O:3, with the latter being the most common in Europe, Japan, Canada, and the USA [1,4]. Although several yersiniohpages have been described for typing Y. enterocolitica [5-8], few have been studied in detail via whole genome sequencing. To date, phage φYeO3-12 displaying specificity for Y. enterocolitica O:3 [9], phage PY54 exhibiting a host range restricted to Y. enterocolitica O:5 and O:5,27 [10], phage φR1-37 with a broader host range within Y. enterocolitica [11], and PY-100 exhibiting a broad host range restricted to the genus Yersinia [12], have been described. Given the considerable interest in bacteriophages because of their potential use as typing, diagnostic, therapeutic, decontaminating, and bio-control agents, our research is aimed at isolating and characterizing novel yersiniohpages in order to expand the repertoire of phages available for targeting clinically significant Y. enterocolitica bioserotypes. In this manuscript we report the morphology, genome sequence, and transcriptomic analysis of phage vB_YenP_AP5 (hereafter referred to as AP5).

Results and discussion

Isolation and host range

Analysis of preliminary treated sewage resulted in the initial isolation of 12 phages infecting Y. enterocolitica strains. From these, AP5 was chosen for detailed study because of its ability to infect Y. enterocolitica strains of serotype O:3. The host range of AP5 was determined using 60 strains belonging to ten Yersinia species at 25°C and at 37°C. The results (Table 1) show that AP5...
| Yersinia strains     | Description/Source            | Biotype | Serotype | Degree of lysis |
|---------------------|--------------------------------|---------|----------|-----------------|
| Y. aldovae Z1       | Unknown, Canada                |         |          | N               |
| Y. aldovae Z2       | Unknown, Canada                |         |          | N               |
| Y. bercovieri Z3    | Unknown, Canada                |         |          | N               |
| Y. bercovieri Z4    | Unknown, Canada                |         |          | N               |
| Y. enterocolitica gk132 | Unknown                       | 2       | O:1      | 2+              |
| Y. enterocolitica JDE029 | Human, Finland            |         | O:1      | 3+              |
| Y. enterocolitica gk2943 | Unknown, Finland       | 2       | O:2      | 3+              |
| Y. enterocolitica gk1142 | Hare, Finland               |         | O:2      | N               |
| Y. enterocolitica K3 | Patient isolate, Canada       | 3       | O:3      | 2+              |
| Y. enterocolitica K9 | Patient isolate, Canada       | 3       | O:3      | 3+              |
| Y. enterocolitica K10 | Patient isolate, Canada      | 3       | O:3      | 3+              |
| Y. enterocolitica K11 | Patient isolate, Canada      | 3       | O:3      | 3+              |
| Y. enterocolitica K12 | Patient isolate, Canada      | 3       | O:3      | 2+              |
| Y. enterocolitica K2 | Patient isolate, Canada       | 4       | O:3      | 3+              |
| Y. enterocolitica K6 | Patient isolate, Canada       | 4       | O:3      | 3+              |
| Y. enterocolitica A | Patient isolate, Canada       | 4       | O:3      | 3+              |
| Y. enterocolitica B | Patient isolate, Canada       | 4       | O:3      | 1+              |
| Y. enterocolitica K1 | Patient isolate, Canada       | 4       | O:3      | 2+              |
| Y. enterocolitica YeO3-6471/76 (YeO3) | Patient isolate, Finland | 4       | O:3      | 3+              |
| Y. enterocolitica YeO3-6471/76-c (YeO3-c) plasmid (pYV) cured derivative of YeO3 [15] | 4       | O:3      | 2+              |
| Y. enterocolitica YeO3-R1 | Spontaneous rough derivative of YeO3-c [13] | 4       | O:3      | N               |
| Y. enterocolitica YeO3-R2 | Spontaneous rough derivative of YeO3 [13] | 4       | O:3      | N               |
| Y. enterocolitica YeO3-OC $\Delta{(\text{wzx-wbcQ})}$ derivative of YeO3, mutant missing outer core operon [14] | 4       | O:3      | 3+              |
| Y. enterocolitica YeO3-OCR $\Delta{(\text{wzx-wbcQ})}$ spontaneous rough derivative of YeO3-OC [14] | 4       | O:3      | N               |
| Y. enterocolitica E | Patient isolate, Canada       | 1A      | O:5      | N               |
| Y. enterocolitica F | Patient isolate, Canada       | 1A      | O:5      | N               |
| Y. enterocolitica C | Patient isolate, Canada       | 2       | O:5,27   | N               |
| Y. enterocolitica D | Patient isolate, Canada       | 2       | O:5,27   | N               |
| Y. enterocolitica K5 | Patient isolate, Canada       | 2       | O:5,27   | N               |
| Y. enterocolitica K7 | Patient isolate, Canada       | 2       | O:5,27   | N               |
| Y. enterocolitica K8 | Patient isolate, Canada       | 2       | O:5,27   | N               |
| Y. enterocolitica ATCC 9610 | Patient isolate, USA      | 1       | O:8      | N               |
| Y. enterocolitica ATCC 23715 | Patient isolate, USA      | 1       | O:8      | N               |
| Y. enterocolitica ATCC 27729 | Patient isolate, Belgium     | 1       | O:8      | N               |
| Y. enterocolitica I | Patient isolate, Canada       | 2       | O:8      | N               |
| Y. enterocolitica J | Patient isolate, Canada       | 2       | O:8      | N               |
| Y. enterocolitica K4 | Patient isolate, Canada       | 2       | O:8      | N               |
| Y. enterocolitica K20 | Patient isolate, Canada      | 2       | O:8      | N               |
| Y. enterocolitica K21 | Patient isolate, Canada      | 2       | O:8      | N               |
| Y. enterocolitica G | Patient isolate, Canada       | 2       | O:9      | N               |
| Y. enterocolitica H | Patient isolate, Canada       | 2       | O:9      | N               |
| Y. enterocolitica K13 | Patient isolate, Canada      | 2       | O:9      | N               |
| Y. enterocolitica K14 | Patient isolate, Canada      | 2       | O:9      | N               |
can form plaques only on *Y. enterocolitica* serotypes O:3, O:2, and O:1 (serotypes with an O antigen known to contain 6-deoxy-L-altropyranose). Other serotypes of *Y. enterocolitica* as well as other species within the genus *Yersinia* were unaffected by the presence of phage AP5. Additionally none of the *Escherichia coli*, *Salmonella*, or *Listeria* species strains tested were infected by this bacteriophage (data not shown). Analysis of *Y. enterocolitica* O:3 rough mutants YeO3-R1 and YeO3-R2 [13] (which are missing the O antigen), were not sensitive to AP5. YeO3-OC, a *Y. enterocolitica* O:3 deletion mutant (Δwzx-wbcQ) which is missing the entire core operon yet produces O antigen [14], was sensitive to AP5. In contrast, its derivative YeO3-OCR, a rough mutant which is also missing the entire core operon yet is unable to produce O antigen [14] was not sensitive to phage AP5. These results indicate the host receptor for phage AP5 lies within the O antigen of the lipopolysaccharide of *Yersinia enterocolitica* O:3 strains, and suggests the O side chain of this serotype (6-deoxy-L-altropyranose) is involved.

### Morphology

AP5 was negatively stained and examined by transmission electron microscopy (Figure 1). The head is icosahedral in shape exhibiting T7 symmetry of approximately 55.0 nm in diameter. The phage particles are each decorated with a short non-contractile tail of approximately 12.0 nm in length and 8 nm in width. Collectively, these morphological features indicate that this virus belongs to the family *Podoviridae*.

### General features of the AP5 genome

The DNA sequence of the phage AP5 consists of linear double stranded DNA of 38,646 bp in length. The size of this phage correlates well with other T7-like phage members, which range from 37.4 kb (*Pseudomonas* phage gh-1) to 45.9 kb (*Erwinia* phage Era103 [17]). The genomes of T7-like phages typically contain direct terminal repeats (DTRs) that are used during genome replication and packaging [18]. The lengths of the DTRs of AP5 (235 bp) are in agreement with the reported lengths for members of the T7 group, for example phage *Salmonella* phage φSG-JL2 and *Yersinia* phage φYeO3-12 have DTRs of 230 bp and 232 bp, respectively [9], whereas Enterobacteria phage T7 has DTRs of 160 bp [19]. Moreover, an alignment of the DTR sequences of phage AP5 and representative members of the T7likevirus genus show a high degree of conservation (Figure 2). Phage AP5 has also an overall genomic guanine plus cytosine (GC) content of 50.7%, compared to 48.5 ± 1.5 mol% for its host [20]. The GC contents of the common representatives of the T7 group, T7 (accession no. V01146.1 [complete sequence of 39,937 bp]) and T3 (accession no. NC_003298.1 [complete sequence of 38,208 bp]), are 48.4% and 49.9%, respectively.

---

**Table 1 Bacterial strains used in this study; and host range of AP5 on 60 Yersinia strains at 25°C and at 37°C (Continued)**

| Strain                          | Source                     | Temperature | Results |
|---------------------------------|----------------------------|-------------|---------|
| Y. enterocolitica K15           | Patient isolate, Canada    | 2           | O:9     | N |
| Y. frederiksenii Q              | Unknown, Canada            |             |         | N |
| Y. frederiksenii S              | Unknown, Canada            |             |         | N |
| Y. intermedia M                 | Unknown, Canada            |             |         | N |
| Y. intermedia N                 | Unknown, Canada            |             |         | N |
| Y. kristensenii Y               | Unknown, Canada            |             |         | N |
| Y. kristensenii X               | Unknown, Canada            |             |         | N |
| Y. kristensenii ATCC 33639      | Hare, unknown              |             |         | N |
| Y. mollaretii T                 | Unknown, Canada            |             |         | N |
| Y. mollaretii U                 | Unknown, Canada            |             |         | N |
| Y. pseudotuberculosis K         | Patient isolate, Canada    | I           |         | N |
| Y. pseudotuberculosis L         | Patient isolate, Canada    | I           |         | N |
| Y. rohdei V                     | Unknown, Canada            |             |         | N |
| Y. rohdei W                     | Unknown, Canada            |             |         | N |
| Y. ruckeri O                    | Unknown, Canada            |             |         | N |
| Y. ruckeri P                    | Unknown, Canada            |             |         | N |
| Y. ruckeri ATCC 29473           | Rainbow trout, USA         |             |         | N |

1Wauters et al. [2].
2Degree of lysis: 4+, complete lysis; 3+ clearing throughout but with faint hazy background, 2+ substantial turbidity throughout cleared zone, 1+ a few individual plaques; N: No effect of phage on bacterial growth as described by Kutter [16].
The GC content of phage AP5 is in agreement with other T7-like phages which range from 46.2 - 62.3% [21].

**Open reading frames and comparative genomics**

The genome of AP5 was scanned for open reading frames (ORF) using computational software. A total of 34,743 nucleotides were involved in the coding of 45 ORFs with sizes ranging from 113 to 3,962 nucleotides (Table 2). The temporal and functional distributions of genes are tightly organized and packed close to each other so that they occupy 89.9% of the genome (Figure 3). The initiation codon ATG is present in 93.3% of the protein-coding genes. Only two other initiation codons occur, TTG and GTG at a frequency of 0.5%, and 0.2%, respectively. All predicted protein-coding genes were screened using BLASTP and Psi-BLAST algorithms against the non-redundant protein database at NCBI. From the 45 coding sequences (CDSs) of AP5, 30 (66.6%) have assigned function, and 15 (33.3%) are similar to proteins of unknown function. While the great majority of the homologs are to

---

**Figure 1** Electron micrographs of phage AP5. The phage has been negatively stained with 2% potassium phosphotungstic acid. AP5 is shown at 50,000× magnification or 150,000× magnification. Scale bar indicates size in nm.

**Figure 2** Multiple sequence alignment of the direct terminal repeats of phage AP5 and selected members of the T7-like virus genus. Multiple sequence alignment of direct terminal repeat sequences was performed using Clustal Omega [22]. Positions which have a single, fully conserved base pair are indicated by an asterisk (*).
| Gene name | ORF | Product | Start (bp) | Stop (bp) | pI | Protein mass (Da) | Function or Similarity | Evidence or Organism | Homolog (Accession number) | Id (%) | BlastP E-value |
|-----------|-----|---------|------------|-----------|----|------------------|------------------------|------------------------|------------------------|--------|---------------|
| -         | -   | regulatory element | 1          | 235       | -  | -                | terminal repeat         | -                      | -                      | -      | -             |
| -         | -   | regulatory element | 550        | 580       | -  | -                | promoter (sequence similarity to host promoter) | -                      | -                      | -      | -             |
| 0.3       | ORF1| S-adenosyl-L-methionine hydrolase | 1,017      | 1,475     | 5.94 | 17171.56        | DNA replication, repair and nucleotide metabolism | Yersinia phage phiYeO3-12 | NP_052065 | 92.1 | 1.94e-101   |
| 0.6       | ORF2| hypothetical protein | 1,748      | 1,951     | 10.5 | 7878.43         | unknown                 | Yersinia phage phiYeO3-12 | NP_052068 | 98.5 | 2.31e-30    |
| 0.65      | ORF3| hypothetical protein | 1,938      | 2,108     | 11.4 | 6546.78         | unknown                 | Enterobacteria phage T3 | NP_523299 | 98.0 | 1.98e-27    |
| 0.7       | ORF4| protein kinase | 2,123      | 2,323     | 7.19 | 42424.34        | host transcription shutoff and Col Iib exclusion | Salmonella phage phiSG-JL2 | YP_001949749 | 93.2 | 0            |
| 1         | ORF5| RNA polymerase | 3,303      | 3,957     | 7.10 | 98841.61        | RNA replication, transcription and modification | Enterobacteria phage T3 | NP_523301 | 99.3 | 0            |
| 1.05      | ORF6| hypothetical protein | 6,044      | 6,316     | 9.22 | 10529.28        | unknown                 | Enterobacteria phage T3 | NP_523302 | 92.2 | 2.75e-53    |
| 1.1       | ORF7| hypothetical protein | 6,409      | 6,549     | 10.93 | 5887.54         | unknown                 | Yersinia phage phiYeO3-12 | NP_052073 | 95.7 | 1.26e-22    |
| 1.2       | ORF8| deoxyguanosine triphospho-hydrolase inhibitor | 6,549      | 6,824     | 7.01 | 10433.64        | hydrolyzes dGTP and may affect cellular pool of dGTP | Salmonella phage phiSG-JL2 | YP_001949753 | 97.8 | 5.40e-50    |
| 1.3       | ORF9| DNA ligase | 6,919      | 7,965     | 5.16 | 39618.14        | DNA replication, recombination, and repair | Salmonella phage phiSG-JL2 | YP_001949754 | 92.5 | 0            |
| 1.6       | ORF10| hypothetical protein | 8,135      | 8,392     | 11.20 | 9893.68         | unknown                 | Yersinia phage phiYeO3-12 | NP_052078 | 100 | 1.26e-53    |
| 1.7       | ORF11| hypothetical protein | 8,392      | 8,871     | 9.38 | 17869.64        | unknown                 | Yersinia phage phiYeO3-12 | NP_052079 | 84.9 | 7.83e-95    |
| 1.8       | ORF12| hypothetical protein | 8,858      | 8,995     | 5.14 | 5267.92         | unknown                 | Yersinia phage phiYeO3-12 | NP_052080 | 100 | 2.81e-23    |
| 2         | ORF13| Host RNA polymerase inhibitor | 8,992      | 9,228     | 4.85 | 8839.87         | inhibition of host RNA Polymerase | Salmonella phage phiSG-JL2 | YP_001949759 | 100 | 2.40e-27    |
| 2.5       | ORF14| ssDNA binding protein | 9,281      | 9,979     | 4.80 | 25965.80        | helix-destabilizing protein | Yersinia phage phiYeO3-12 | NP_052082 | 98.7 | 2.40e-167   |
| 3         | ORF15| endonuclease | 9,979      | 10,440    | 9.58 | 17725.54        | endonuclease I           | Enterobacteria phage T7M | AFQ97046 | 98.0 | 1.14e-103   |
| 3.5       | ORF16| endolysin | 10,433     | 10,888    | 9.03 | 16900.28        | N-acetylmuramoyl-L-alanine amidase | Yersinia phage phiYeO3-12 | NP_052084 | 100 | 9.21e-109   |
| 4         | ORF17| primase/helicase | 11,255     | 12,769    | 5.11 | 55882.14        | DNA replication          | Yersinia phage phiYeO3-12 | NP_052088 | 99.8 | 0            |
| ORF  | Gene Description | Start Position | End Position | Length (bp) | E-value | E. coli Similarity (%) | Identity (%) | Accession # | Function | E. coli Similarity Protein |
|------|------------------|----------------|--------------|------------|---------|------------------------|-------------|-------------|----------|--------------------------|
| 4.3  | hypothetical protein | 12865          | 13077        | 1000       | 7762.16 | unknown                | Enterobacteria phage T3 | NP_523318  | 98.6     | 3.66e-20                   |
| 4.5  | hypothetical protein | 13090          | 13374        | 989        | 10749.39 | unknown                | Yersinia phage phiYeO3-12 | NP_052092  | 100     | 1.26e-62                   |
| 5    | DNA polymerase    | 13442          | 15556        | 624        | 79875.05 | DNA replication        | Salmonella phage phiSG-JL2 | YP_001949769 | 99.0 | 0                          |
| 5.5  | hypothetical protein | 15573          | 15869        | 555        | 11022.78 | unknown                | Klebsiella phage KP32 | YP_003347541 | 58.3 | 7.72e-25                   |
| 5.7  | hypothetical protein | 15866          | 16075        | 981        | 7260.42  | unknown                | Yersinia phage phiYeO3-12 | YP_052098  | 100     | 5.77e-42                   |
| 5.9  | host recBCD nuclease inhibitor | 16072          | 16254        | 31         | 6742.48  | Inhibits host recBCD nuclease | Yersinia phage phiYeO3-12 | YP_072071  | 98.3 | 1.08e-33                   |
| 6    | exonuclease       | 16251          | 17162        | 498        | 34799.67 | DNA Replication, repair, and recombination | Yersinia phage phiYeO3-12 | YP_052100  | 99.3 | 0                          |
| 6.3  | hypothetical protein | 17144          | 17257        | 969        | 4111.08  | unknown                | Yersinia phage phiYeO3-12 | YP_052102  | 97.3 | 8.79e-15                   |
| 6.5  | hypothetical protein | 17401          | 17595        | 594        | 7458.47  | unknown                | Yersinia phage phiYeO3-12 | YP_052103  | 100     | 6.28e-38                   |
| 6.7  | hypothetical protein | 17600          | 17851        | 931        | 8833.96  | unknown                | Enterobacteria phage T3 | NP_523330  | 98.8 | 1.76e-40                   |
| 7.3  | tail assembly protein | 17879          | 18199        | 320        | 11003.70 | scaffolding protein required for the assembly of tail fibers on capsids | Salmonella phage phiSG-JL2 | YP_001949779 | 95.3 | 2.72e-39                   |
| 8    | head to tail joining protein | 18210          | 19817        | 454        | 58649.37 | bacteriophage head to tail connecting protein | Yersinia phage phiYeO3-12 | NP_052106  | 100     | 0                          |
| 9    | capsid assembly protein | 19919          | 20851        | 432        | 33787.57 | scaffolding protein required for the formation of pro-capsids | Salmonella phage phiSG-JL2 | YP_001949781 | 98.4 | 0                          |
| 10A  | major capsid protein | 21008          | 22054        | 1046       | 36954.20 | scaffolding protein    | Yersinia phage phiYeO3-12 | NP_052108  | 99.1 | 0                          |
| 10B  | minor capsid protein | 22066          | 22188        | 62         | 4295.84  | scaffolding protein    | Salmonella phage phiSG-JL2 | YP_001949782 | 95.0 | 2.23e-15                   |
| 11   | tail tubular protein A | 22270          | 22860        | 490        | 22233.69 | required for assembly of tails of T7-like phages | Yersinia phage phiYeO3-12 | NP_052110  | 100     | 1.00e-141                 |
| 12   | tail tubular protein B | 22876          | 25281        | 645        | 89771.51 | required for assembly of tails of T7-like phages | Yersinia phage phiYeO3-12 | NP_052111  | 98.1 | 0                          |
| 13   | internal virion protein A | 25354          | 25779        | 425        | 16473.96 | forms internal core of virion | Salmonella phage phiSG-JL2 | YP_001949786 | 96.3 | 5.14e-33                   |
| 13.5 | hypothetical protein | 25766          | 26155        | 390        | 14582.99 | unknown                | Yersinia phage phiYeO3-12 | YP_072072  | 73.6 | 2.87e-50                   |
| 14   | internal virion protein B | 26158          | 26751        | 603        | 21308.02 | forms internal core of virion | Yersinia phage phiYeO3-12 | NP_052114  | 96.4 | 1.54e-133                 |
Table 2 Yersiniophage AP5 gene annotations (Continued)

|   | Gene ID | Gene Name | Start Position | Stop Position | Length | Identity | Accession | E-value  |
|---|---------|-----------|----------------|---------------|--------|----------|-----------|----------|
| 15 | ORF38   | internal virion protein C | 26,754 | 28,997 | 2,243 | 5.47 | Yersinia phage phiYeO3-12 | 99.3 | 0 |
| 16 | ORF39   | internal virion protein D | 29,016 | 32,978 | 3,962 | 8.41 | Yersinia phage phiYeO3-12 | 98.9 | 0 |
| 17 | ORF40   | tail fiber protein | 33,050 | 34,996 | 1,946 | 6.45 | Yersinia phage phiYeO3-12 | 89.3 | 0 |
| 17.5 | ORF41  | holin | 35,008 | 35,211 | 203 | 6.08 | Yersinia phage phiYeO3-12 | 94.0 | 2.79 \* 37 |
| 18 | ORF42   | DNA packaging protein A | 35,215 | 35,481 | 266 | 4.70 | Salmonella phage phiSG-JL2 | 100 | 1.01 \* 55 |
| 18.5 | ORF43  | Phage λ Rz-like lysis protein (Rz/Rz1 equivalent) | 35,570 | 36,022 | 452 | 9.41 | Salmonella phage phiSG-JL2 | 98.7 | 4.69 \* 105 |
| 19 | ORF44   | DNA packaging protein B | 35,997 | 37,760 | 763 | 5.32 | Yersinia phage phiYeO3-12 | 99.8 | 0 |
| 19.5 | ORF45  | hypothetical protein | 38,005 | 38,154 | 149 | 7.87 | Yersinia phage phiYeO3-12 | 98.0 | 1.78 \* 24 |
| - | regulatory element | 38,412 | 38,646 | - | - | terminal repeat | - | - |

Genes are listed by number, along with their predicted function, if known, followed by the nature of the evidence that supports the functional classification. Genes with no functional prediction, but with significant (E < 10\(^{-5}\)) sequence similarity to genes in the NCBI database as determined by BLASTP are listed, including the name of the organism in which the similar gene was found.
proteins of *Yersinia* phage φYeO3-12 (26), examples of primary sequence similarity to *Salmonella* phage φSG-JL2 (12), *Enterobacteria* phages T3 and T7 (6,1) and *Klebsiella* phage KP32 (1), exist. All of these phages are members of the *T7likevirus* genus. No function can be speculated about the hypothetical proteins of AP5 without further study. Based upon overall protein homology determined using CoreGenes [23,24], AP5 shares 42 (76.4%), similar proteins with *Enterobacteria* phage T7 and *Enterobacteria* phage T3, 43 (78.2%) similar proteins with *Salmonella* phage φSG-JL2, and 43 (72.9%) similar proteins with *Yersinia* phage φYeO3-12. Collectively these results indicate that AP5 is a member of the *Autographivirinae*, specifically a member of the *T7likevirus* genus [25]. Thus, the T7 gene nomenclature was adopted for naming the genes of AP5. Since at the protein level phage AP5 showed the greatest sequence identity with *Yersinia* phage φYeO3-12 proteins, the genomes of the two phages were compared using progressive Mauve [26] (Figure 4). The gene arrangement of essential genes is collinear, highly conserved, and only some genes coding for hypothetical proteins present in φYeO3-12 are dissimilar or absent in AP5. The pairwise % identity of the phage AP5 genome to *Yersinia* phage φYeO3-12 genome was estimated at 89.6%.

**Nucleotide metabolism, DNA replication and recombination**

In the AP5 genome, at least eleven genes were identified that play a role in nucleotide metabolism, DNA replication, and recombination. The transcribed genes function to overcome host restriction and to convert the metabolism of the host cell to the production of phage proteins. The product of gene 0.3 is a small protein, which mimics B-form DNA, and binds to and inhibits type I restriction endonucleases [28-30], as well as possessing S-adenosyl-L-methionine hydrolase (SAMase) activity acting to degrade the methyl group donor and the methylation activities present in the host [31]. Dam (DNA adenine methyltransferase) methylase modifies GATC, and Dcm (DNA cytosine methyltransferase) methylase modifies CC (A/T)GG sequences [9]. As in *Yersinia* phage φYeO3-12, the sequences corresponding to restriction enzyme
recognition sites GATC and CC(A/T)GG, are underrepresented in phage AP5 DNA, occurring only 4 and 3 times, respectively. Gene 0.7 codes for a protein kinase involved in host transcription shutoff [32] and phosphorylates host elongation factors G and P and ribosomal protein S6 [33]. Other genes include an RNA polymerase (gene 1), a deoxyguanosine triphospho-hydrolase inhibitor (gene 1.2), a DNA ligase (gene 1.3), a host RNA polymerase inhibitor (gene 2), a ssDNA binding protein/helix destabilizing protein (gene 2.5), an endonuclease (gene 3), a primase/helicase (gene 4), a DNA polymerase (gene 5), a host recBCD nuclease inhibitor (gene 5.9), and an exonuclease (gene 6).

DNA packaging and morphogenesis

Several genes were identified that play a role in morphogenesis and DNA packaging. We identified two CDSs which display sequence similarity to the capsid proteins of phages belonging to T7-like viruses. The upstream gene 10A displays homology to Yersinia phage φYeO3-12 major capsid protein 10A [NP_052108], while the downstream gene 10B is similar to the minor capsid protein 10B in Salmonella phage φSG-JL2 [YP_001949782]. Some T7-like phages display two “versions” of the major capsid protein, which are designated as 10A and 10B [34]. The sequences of the amino termini of these proteins are identical, but during translation a −1 ribosomal frameshift allows for alternative reading frames within one mRNA, permitting the elongation of the protein product. The features of this system are a slippery site in the DNA/RNA and a downstream stem-loop structure capable of forming a pseudoknot. Gene 9 was identified as the capsid assembly protein required for the formation of procapsids. The structure of this phage is therefore made up of gene 10A and gene 10B (capsid), the head to tail joining protein (gene 9), and an internal core formed by the products of gene 13 (internal virion protein A), gene 14 (internal virion protein B), gene 15 (internal virion protein C), and gene 16 (internal virion protein D). These proteins are homologous to those that form the internal core of the T7 virion. In T7, along with internal virion proteins B and C, the internal virion protein D, is ejected from the phage head and forms part of a putative channel that spans the entire host cell envelope and allows entry of DNA. The N-terminus of this protein has similarity to a lytic transglycosylase and may help form a channel for phage DNA translocation through the peptidoglycan layer of the host envelope [18]. BLASTN analysis of gene 16 (internal virion protein D) confirms the presence of a peptidoglycan hydrolase motif at the N-terminus. Gene 7.3 was identified as the tail assembly protein required for assembly of tail fibers on capsids. Genes 11 and 12 correspond to tail tubular proteins A and B respectively required for assembly of tails. Gene 17, codes for the tail fiber protein or host recognition binding protein and shares 89.3% identity with gp17, the tail fiber protein of Yersinia phage φYeO3-12 [NP_052117], and only 67% identity with gp17 of Salmonella phage φSG-JL2 [YP_001949790]. As with other gp17 homologs, sequence similarity is only found at the N-terminus, the part of the protein that is associated with the tail structure. The C-terminus is involved in ligand interactions and exhibits considerable differences, despite that phage AP5 shares a similar host range with Yersinia...
phage φYeO3-12 [9]. The large and small terminase subunit homologs were determined to be the products of gene 18 (DNA Packaging Protein A) and gene 19 (DNA Packaging Protein B).

**Host cell lysis**
The final stage of the phage lytic cycle is degradation of the bacterial cell wall and release of progeny phages. The lysis of the cell wall is typically induced by two phage encoded proteins, a holin and an endolysin [38]. Endolysins are murelytic enzymes produced by dsDNA phages, which hydrolyze the peptidoglycan layer of bacterial cell walls. As in other T7 phages, gene 3.5 of phage AP5 is proposed to be the endolysin protein since it possesses N-acetylmuramoyl-L-alanine amidase activity. Access of endolysins to the cell wall occurs through the presence of a secondary lysis factor, known as a holin. Holins are usually small proteins characterized by the presence of transmembrane domains (TMD) [39]. The predicted proteins of AP5 were scanned for TMDs using TMHMM [40]. TMDs were identified in gene 0.6, gene 6.3, gene 17.5, and gene 19.5, which code for small proteins of 67, 37, 67, and 49 amino acids, respectively. The derived protein from gene 17.5 of AP5 is proposed as a holin since it is a small protein containing an N-terminal TMD and shares sequence similarity to Yersinia phage φYeO3-12 lysis protein [NP_052118]. Phage AP5 has also one more lysis gene (gene 18.5) coding for a phage λ Rz-like lysis protein (PHA000276), an i-spanin of 150 amino acids which presents 98.7% sequence identity to λ Rz-like protein [YP_00194793] in Salmonella phage φSG-JL2. Further inspection of the gene 18.5 sequence, confirms the presence of a nested ORF of 255 bp (in the +1 reading frame) embedded entirely within the sequence coding for an o-spanin with homology to Rz1 (18.7) of bacteriophage T7. Based on these observations, gene 18.5 is proposed as an Rz/Rz1 equivalent lysis gene coding for transmembrane spanins involved in the disruption of the outer membrane of the host [41].

**Transcriptional and regulatory sequences**
Phage AP5 was not found to contain tRNA genes, which is not an unexpected observation since no T7-like phages have been found to harbour them. A promoter was identified at position 550–580 bp of the genome with sequence similarity to host promoter consensus TTGACA(N15-18)TATAAT with a 2 bp miss-match suggesting the early genes of this type of virus are transcribed by the host RNA polymerase. This is a major dissimilarity between phage AP5 and T3/T7 phages where the latter possess multiple strong promoters recognized by the host RNA polymerase. As with all T7 group phages, the AP5 phage encoded RNA polymerase (RNAP), is responsible for the recognition of phage specific promoters. In phage AP5, we identified 14 phage-specific promoters using PHIRE [42], which are named according to the downstream gene (Table 3). The promoter sequences lie within intergenic regions and show the greatest similarity to those of Yersinia phage φYeO3-12 and bacteriophage T3.

**Conclusions**
In this manuscript we have reported on the morphology and genome of the phage vB_YenP_AP5. Due to its lytic nature and marked specificity to Y. enterocolitica strains of serotype O:3, this phage is a potential biotechnological tool for diagnostic, therapeutic, and/or biocontrol uses, given that O:3 is the most predominant serotype involved in human food-borne infections [4]. Additionally, the genome of this phage does not contain any undesirable laterally transferable genes that are related to bacterial toxins, pathogenicity, antibiotic resistance and/or lysogeny on the basis of homologies with known virulence and resistance genes available in GenBank.

**Methods**

**Bacterial strains and growth media**
Tryptic Soy Broth (TSB), Tryptic Soy Agar (TSA), and Tryptic Soft Agar (TSB +0.6% agar) (Difco Laboratories, Detroit, MI) were used to grow the host bacteria and to propagate the phage. In procedures involving phage infection, media were supplemented with filter-sterilized CaCl₂2H₂O to a final concentration of 5 mM. Y. enterocolitica strains of serotype O:3 were used as indicator strains for phage isolation (Table 1). Y. enterocolitica 6471/76-c of bioserotype 4/O:3 [43] obtained from the Félix d’Hérelle Reference Center for bacterial viruses (Université Laval, QC, Canada) was used for phage propagation. Y. enterocolitica 6471/76, Y. enterocolitica strains of serotype O:1 and O:2, and Y. enterocolitica O:3 mutants, were acquired from the Haartman Institute, University of Helsinki, Finland. Other strains from the genus Yersinia were obtained from the Ontario Agency for Health Protection and Promotion (OAHPP) (Ontario, Canada), and the American Type Culture Collection (ATCC) (Manassas, Virginia, USA).

**Isolation and propagation of phage**
A 1 L sample of preliminary treated sewage from a local water treatment plant (Guelph, Ontario) was centrifuged twice at 10,000 g for 20 minutes at 4°C using a Beckman high-speed centrifuge and a JA-10 fixed-angle rotor (Beckman, Palo Alto, CA, USA) and the supernatant passed through a sterile filter membrane of 0.45 μm pore size (Fisher Scientific, Mississauga, ON, Canada). Equal 9 ml volumes of the filtered supernatant and TSB
were inoculated with 200 μL of an overnight mixed culture of selected *Y. enterocolitica* strains of serotype O:3 and incubated for 18–24 h at 30°C with gentle shaking. After incubation, the enrichments were centrifuged at 10,000 g for 20 minutes at 4°C and the supernatant filtered through a sterile disposable filter of 0.45 μm pore size, and the filtrates stored at 4°C. Phages were detected by spot tests [16] on indicator strains incubating for 16–20 h at 25°C. Complete or partial lysis zones were then removed by cutting the soft layer from the plates using a sterile pipette tip and placing them separately in 1 mL of SM buffer (5.8 g of NaCl per liter, 2.0 g of MgSO₄·7H₂O per liter, 50 mM Tris–HCl [pH 7.5]), and used in standard double agar overlay plaque assays [43] to identify plaques showing different size and plaque morphology. Three rounds of repeated single plaque isolation were then performed to ensure unique phages were obtained. Purified phages were named following the naming convention of Kropinski et al. [44].

**Transmission electron microscopy**

The phage was pelleted at 25,000 × g for 1 hour at 4°C, using a Beckman high-speed centrifuge and a JA-18.1 fixed-angle rotor (Beckman, Palo Alto, CA, USA). The phage pellet was washed twice under the same conditions in neutral 0.1 M ammonium acetate [46]. The final phage sediment was re-suspended in 150 μL of SM-buffer supplemented with 5 mM CaCl₂. Samples were then deposited onto carbon-coated Formvar films on copper grids, and stained with 2% uranyl acetate (pH 4) or 2% potassium phosphotungstate (PT, pH 7.2), air dried, and examined under a Tecnai G₂ F20 transmission electron microscope (FEI, Hillsboro, OR, USA), operating at 120 keV. Images were collected and analyzed using Digital Micrograph™ Software (Gatan, Pleasanton, CA, USA).

**Isolation of phage DNA**

To separate phage from bacterial debris, a crude phage lysate was centrifuged at 10,000 × g for 15 min at 4°C.
and the supernatant filtered through 0.22 μm low protein binding filter (Millipore, USA). Contaminating nucleic acids in the supernatant were digested with pancreatic DNase 1, and RNase A, each added to obtain a final concentration of 10 μg/mL. (Sigma-Aldrich Canada Ltd., Oakville, ON), for 15 min at room temperature. DNA isolation was then performed with a commercial Phage DNA Isolation Kit (Norgen BioTek Corp., Thorold, ON, Canada), as per the manufacturer’s instructions. The DNA was characterized spectrophotometrically.

**Genome sequencing and assembly**

Phage genomic DNA was fragmented using Ion Xpress™ Plus gDNA Fragment Library kit following the manufacturer’s protocol (Life Technologies, Foster City, CA). The fragmented DNA was collected using Pippin Prep DNA Size Selection System (Sage Science, Beverly, MA) and assessed for concentration and size distribution using a Bioanalyzer 2100 (Agilent Technologies, Mississauga, ON). The DNA fragments were then attached to the surface of Ion Sphere particles (ISPs) using an Ion Xpress Template kit (Life Technologies) according to the manufacturer’s instructions. Template-ISPs were sequenced using 316 micro-chips using an Ion Torrent Personal Genome Machine (PGM) with an Ion PGM Sequencing 400 kit (Life Technologies). The sequence reads were filtered using PGM software to remove low quality sequences, trimmed to remove adaptor sequences and the filtered sequences were assembled. The assembled genome had a coverage of 33.4x. Gaps were identified using the Lasergene® Genomics Suite of DNASTar software (DNASTar Inc., Madison, WI). The gaps were closed by PCR using primers flanking regions adjacent to the gaps and sequencing using a 3730 Genetic Analyzer (Life Technologies). The final assembled genome was manually curated for errors.

**Bioinformatics analysis**

The phage genome was analyzed for coding sequences using Kodon version 2.0 (Applied Maths Inc., Austin, TX, USA). Genes were identified from among the predicted coding sequences based on the presence of ATG, GTG, CTG or TTG start codons, followed by at least 30 additional codons, and an upstream sequence resembling the following ribosome-binding site, GGAGGT [47,48]. A search for phage-encoded tRNA genes was performed with tRNAscan-SE and Aragorn, using default parameters [49,50]. Preliminary annotation of genes was performed using myRAST [51]. Additional manual functional annotation was performed using the Geneious software version 7.1.5 (Biomatters) [52,53]. Phage-specific promoters were discovered using PHIRE [42] using a length (L) of 22 bp and a degeneracy (D) of 4 bp. Determination of theoretical molecular weight and isoelectric point employed ExPASy via http://web.expasy.org/compute_pi/ [54-56]. BLASTP and Psi-BLAST algorithms were used to determine the similarity to described proteins in the National Center for Biotechnology Information [NCBI] database (http://www.ncbi.nlm.nih.gov). Whole genome comparisons were carried out using Mauve [26], and CoreGenes [24].

**Genome sequence**

The annotated genome sequence for the phage vB_YenP_AP5 was deposited in the NCBI nucleotide database under the accession number KM253764.

**Competing interests**

The authors declare that they have no competing interest.

**Acknowledgements**

This research was supported by the Ontario Ministry of Agriculture Food and Rural Affairs (OMAFRA) Food Safety Research Program (research grant SF607S). The authors would like to thank Dr. Jiping Lee, Agriculture and Food Laboratory, Laboratory Services Division, University of Guelph for performing the genome sequencing. The authors would like to acknowledge the technical assistance with TEM imaging of Mr. Robert Harris, Molecular and Cellular Imaging Facility, at the University of Guelph, and the assistance of Dr. Mikail Skurnik, Haartman Institute, University of Helsinki, Finland for providing selected bacterial strains.

**Author details**

1. Laboratory Services Division, University of Guelph, Guelph, ON N1H 8J7, Canada.
2. Department of Pathobiology, University of Guelph, Guelph, ON N1G 2W1, Canada.
3. Department of Molecular and Cellular Biology, University of Guelph, Guelph, ON N1G 2W1, Canada.
4. Canadian Research Institute for Food Safety, University of Guelph, Guelph, ON N1G 2W1, Canada.
5. Department of Food Science, University of Guelph, Guelph, ON N1G 2W1, Canada.

**Received: 15 August 2014 Accepted: 19 October 2014 Published: 28 October 2014**

**References**

1. Bottone EJ: *Yersinia enterocolitica: Overview and epidemiologic correlates*. Microbes Infect 1999, 1(6):323–333.
2. Wauters G, Kandolo K, Janssens M: Revised biogrouping scheme of *Yersinia enterocolitica*. Clin Microbiol Immunol 1987, 3:21–29.
3. Wren BW: *The yersiniae—a model genus to study the rapid evolution of bacterial pathogens*. Nat Rev Microbiol 2003, 1(1):55–64.
4. Fredriksson-Ahomaa M, Stolle A, Korkeala H: Molecular epidemiology of *yersinia enterocolitica* infections. FEMS Immunol Med Microbiol 2006, 47(3):315–329.
5. Baker PM, Farmer JJ: New bacteriophage typing system for *Yersinia enterocolitica*, *Yersinia kristensenii*, *Yersinia frederiksenii*, and *Yersinia intermedia*: Correlation with serotyping, biotyping, and antibiotic susceptibility. J Clin Microbiol 1982, 15(3):491–502.
6. Bergan T, Norris JR: Bacteriophage typing of *Yersinia enterocolitica*. Methods Microbiol 1978, 12:225–236.
7. Kawaoka Y, Mitani T, Otsuki K, Tsubokura M: Isolation and use of eight phages for typing *Yersinia enterocolitica* O3. J Med Microbiol 1987, 23(4):349–352.
8. Nicolle P, Mollaret H, Hamon Y, Vieu JF, Brault J, Brault G: Lysogenic bacteriocinogenic and phage-typing study of species *Yersinia enterocolitica*. Ann Inst Pasteur (Paris) 1967, 112(1):86–92.
The complete genomic sequence of lytic bacteriophage phiYeO3-12 of Yersinia enterocolitica serotype O3. J. Bacteriol. 2001, 183(1):1928–1937.

Hertwig S, Klein I, Schmidt V, Beck S, Hammer J, Appel B: Sequence analysis of the genome of the temperate Yersinia enterocolitica phage PTV4. J. Mol Biol. 2003, 15:2505–2522.

Klijnjan S, Hakala K, Pinta E, Huttunen S, Puita P, Gador A, Lonnberg H, Skurnik M: Yersinia entercolitica phiR1–37 is a tailed bacteriophage having a 270 kb DNA genome with thymidine replaced by deoxyuridine. Microbiology 2005, 151(12):4093–4102.

Schwudke D, Egerin A, Michael K, Vollmar S, Appel B, Krubender D, Konietzny A, Strauch E: Broad-host-range yersinia phage PiY100: Genome sequence, proteome analysis of virions, and DNA packaging strategy. J. Bacteriol. 2008, 190(3):332–342.

Al-Hendy A, Toivanen P, Skurnik M: Lipopolysaccharide O side chain of Yersinia enterocolitica O3 is an essential virulence factor in an orally infected murine model. Infect Immun 1992, 60:870–875.

Biedzka-Sarek M, Vehno R, Skurnik M: Role of YadA, Ali, and Lipopolysaccharide in serum resistance of Yersinia enterocolitica Serotype O3. Infect Immun 2005, 73:2232–2244.

Skurnik M: Lack of correlation between the presence of plasmid and fimbriae in Yersinia enterocolitica and Yersinia pseudotuberculosis. J. Appl Bacteriol. 1984, 56:355–363.

Kutter E: Phage Host Range and Efficiency of Plating. In: Bacteriophages. Methods and protocols, Volume 1: isolation, characterization, and interactions. 1st edition. Edited by Clokie MRJ, Kropinski AM. New York: Humana Press; 2009:141–149.

Kovalyova IV, Kropinski AM: The complete genomic sequence of lytic bacteriophage gh-1 infecting Pseudomonas putida-evidence for close relationship to the T7 group. Virology 2003, 315:305–315.

Molineux I: The T7 Group. In: The bacteriophages. Edited by Calendar RL. Oxford: Oxford University Press; 2005.

Dunn UJ, Studier FW: Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. J. Mol Biol. 1983, 166:477–535.

Brenner DI, Uring J, Bercovici H, Steinweil AG, Fanning GR, Alonso JM, Bercovier H, Steigerwalt AG, Fanning GR, Alonso JM: T7 DNA and the locations of T7 genetic elements. J. Mol. Biol. 1983, 151:1–62.

Marchand I, Nicholson AW, Dreyfus M: The genome of bacteriophage ghMV, a T7-like virus infecting Pseudomonas aegginiosa. Virology 2008, 312:49–59.

Sievers F, Wilm A, Dineen G, Gibson TJ, Karplus K, Li W, Löwe R, McLachlan H, Remmert M, Soding J, Thompson JD, Higgins DG: Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 2011, 7:539.

Kropinski AM, Borovodsky M, Carver TJ, Cerdeño-Tárraga AM, Darling A, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 2011, 7:539.

Kropinski AM, Borovodsky M, Carver TJ, Cerdeño-Tárraga AM, Darling A, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 2011, 7:539.

Marchand I, Nicholson AW, Dreyfus M: The genome of bacteriophage ghMV, a T7-like virus infecting Pseudomonas aegginiosa. Virology 2008, 312:49–59.

Sievers F, Wilm A, Dineen G, Gibson TJ, Karplus K, Li W, Löwe R, McLachlan H, Remmert M, Soding J, Thompson JD, Higgins DG: Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 2011, 7:539.

Kropinski AM, Borovodsky M, Carver TJ, Cerdeño-Tárraga AM, Darling A, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 2011, 7:539.
55. Bjellqvist B, Basse B, Olsen E, Celis JE: Reference points for comparisons of two-dimensional maps of proteins from different human cell types defined in a pH scale where isoelectric points correlate with polypeptide compositions. *Electrophoresis* 1994, 15:529–539.

56. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A: Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook*. Edited by Walker JM. Totowa, New Jersey: Humana Press; 2005.

doi:10.1186/1743-422X-11-188

*Cite this article as:* Leon-Velarde et al.: Complete genome sequence of bacteriophage vB_YenP_AP5 which infects *Yersinia enterocolitica* of serotype O:3. *Virology Journal* 2014 11:188.