* **All RBS scores taken from PECAAN**

**Gene 1**

**1. Final Gene Information:**

**(a) FINAL Gene Co-ordinates: Start: 1, Stop: 567**

**(b) Is it a protein-coding gene? Yes**

**(c) What is its function? Adenylate kinase**

**(d) Final Summary: START:** Longest ORF, same call in Glimmer and Genemark, no overlap or gap with previous gene (final score: -6.721, z-value: 1.274, spacer distance: 6). Strong coding potential in +3 frame from ∼1–530 bp, 1:1 target:query for top 20 genes, most manual called start in Starterator (called 95.8% of the time when present), conserved synteny. **FUNCTION:** Function called by HHPred and CDD, adenylate kinase called by 4 of top 10 genes in BLAST, called by 81% (320 out of 395) of phages in same pham (84713), called by three closest whole-genome relatives.

**Gene 2**

**1. Final Gene Information:**

**(a) FINAL Gene Co-ordinates: Start: 564, Stop: 2351**

**(b) Is it a protein-coding gene? Yes**

**(c) What is its function? Terminase**

**(d) Final Summary: START:** Not longest ORF—longest ORF overlaps previous gene over 200 bp—called ORF is the first ORF without significant overlap (4 bp overlap), same call in Glimmer and Genemark (final score: -3.723, z-value: 2.975, spacer distance: 15). Strong coding potential in +1 frame from ∼600–2300 bp with small dips, 1:1 target:query for top 40 genes, most manual called start in Starterator (called 100% of the time when present), conserved synteny. **FUNCTION:** Function called by HHPred and CDD, function called by 10 of 10 top genes in BLAST, called by 86% (498 out of 582) of phages in same pham (220995), called by three closest whole-genome relatives.

**DELETED THIS AUTO-CALLED GENE**

**1. Final Gene Information:**

**(a) FINAL Gene Co-ordinates: Start: 3003, Stop: 2398 (Reverse) (NOT INCLUDING)**

**(b) Is it a protein-coding gene? Yes (no non-draft calls)**

**(c) What is its function? Hypothetical protein (no non-draft calls)**

**(d) Final Summary: START:** Not longest ORF—longest ORF has too small gap (1 base) for switching to reverse frame—called ORF is first ORF with reasonable gap (76 bases) but would overlap significantly with called genes at 2435–2650 and 2647–3012. Start called in Glimmer at 3003, uncalled in Genemark (final score: -4.035, z-value: 2.771, spacer distance: 18). Extremely weak coding potential in -3 frame from ∼3000–2600, +3 frame has more coding potential from ∼2650–2950, no BLAST alignment data, no manual calls for any start in pham 6989, no synteny. **FUNCTION:** Function not supported by HHPred score (low coverage, e-value of 400), no BLAST data for function call, no non-draft function calls in pham 6989, gene not called by three closest whole-genome relatives.

**Gene 3**

**1. Final Gene Information:**

**(a) FINAL Gene Co-ordinates: Start: 2435, Stop: 2650**

**(b) Is it a protein-coding gene? Yes**

**(c) What is its function? Hypothetical protein**

**(d) Final Summary: START:** Longest ORF, call supported by Genemark, 83 bp gap with previous gene—ORF present in gap, however there is no coding potential and none of three closest relatives call a gene in the gap. Chosen start has strongest RBS values (final score: -2.666, z-value: 3.224, spacer distance: 13). Strong coding potential in +2 frame from ∼2450–2600 bp, 1:1 target:query for top 9 genes, most manual called start in Starterator (called 100% of the time when present), conserved synteny. Start not called by Glimmer. **FUNCTION:** Hypothetical protein called by 10 of top 10 genes in BLAST, called by 100% (296 out of 296) of phages in same pham (216), called by three closest whole-genome relatives. Function not called in HHPred.

**Gene 4**

**1. Final Gene Information:**

**(a) FINAL Gene Co-ordinates: Start: 2647, Stop: 3012**

**(b) Is it a protein-coding gene? Yes**

**(c) What is its function? Hypothetical protein**

**(d) Final Summary: START:** Not longest ORF—longest ORF overlaps previous gene by 19 bp—called start overlaps by 4 bp, supported by Genemark (final score: -4.140, z-value: 2.491, spacer: 13), called start is most manual called start in Starterator (called 96.3% of the time when present). Some coding potential in +3 frame from ∼2650–2950—many rises and falls, 1:1 target:query for top 3 genes, conserved synteny. Start not called by Glimmer. **FUNCTION:** Hypothetical protein called by 9 of top 10 genes in BLAST (called by over 50 total), called by 100% (301 of 301) of phages in same pham (84759), called by three closest whole-genome relatives.Function not called by HHPred score.

**Gene 5**

**1. Final Gene Information:**

**(a) FINAL Gene Co-ordinates: Start: 3080, Stop: 3217**

**(b) Is it a protein-coding gene? Yes**

**(c) What is its function? Hypothetical protein**

**(d) Final Summary: START:** Longest ORF, same call in Glimmer and Genemark, 76 bp gap with previous gene—ORF present in gap, however no coding potential or BLAST results and none of closest relatives call a gene in the gap. Chosen start has strongest RBS scores (final score: -4.991, z-value: 1.893, spacer distance: 10). Strong coding potential in +2 frame from ∼3080–3180, 1:1 target:query for top genes #1 and #3 but not #2, most manual called start in Starterator (called 100% of the time when present), conserved synteny. **FUNCTION:** Hypothetical protein called by 8 of top 8 genes in BLAST, called by 100% (289 of 289) of phages in same pham (84776), called by three closest whole-genome relatives, function not called by HHPred.

**Gene 6**

**1. Final Gene Information:**

**(a) FINAL Gene Co-ordinates: Start: 3918, Stop: 3364 (Reverse)**

**(b) Is it a protein-coding gene? Yes**

**(c) What is its function? RuvC-like resolvase**

**(d) Final Summary: START:** Longest ORF, same call in Glimmer and Genemark (final score: -5.989, z-value: 1.721, spacer distance: 14), 46 bp gap with previous gene—ORF present in gap, however no coding potential and no close relatives call a gene in the gap. Strong coding potential in -3 frame from ∼3900–3450, 1:1 target:query for top 25 related genes, most manual called start in Starterator (called 90.2% of the time when present), conserved synteny. **FUNCTION:** Function called by HHPred, RuvC-like resolvase called by top 25 similar genes in BLAST, called by 83% (707 of 847) of phages in same pham (228150), called by three closest whole-genome relatives.

**Gene 7**

**1. Final Gene Information:**

**(a) FINAL Gene Co-ordinates: Start: 3965, Stop: 4411**

**(b) Is it a protein-coding gene? Yes**

**(c) What is its function? Hypothetical protein**

**(d) Final Summary: START:** Not longest ORF–longest ORF overlaps by 167 bp—called ORF is first ORF without significant overlap (46 bp gap), same call in Glimmer and Genemark (final score: -5.157, z-value: 2.135, spacer distance: 14). Strong coding potential in +2 frame from ∼3950–4400 save for dip at ∼4080, 1:1 target:query for top 3 genes, most manual called start in Starterator (called 95.6% of the time when present), mostly conserved synteny. **FUNCTION:** Function called by HHPred, hypothetical protein called by 8 of top 10 similar genes in BLAST, called by 100% (4988 out of 488) of phages in same pham (228172), called by three closest whole-genome relatives.

**Gene 8**

**1. Final Gene Information:**

**(a) FINAL Gene Co-ordinates: Start: 4441, Stop: 6360**

**(b) Is it a protein-coding gene? Yes**

**(c) What is its function? Portal protein**

**(d) Final Summary: START:** Longest ORF, strong coding potential in +3 frame save for dip at ∼ 5050, 1:1 target:query for top 20 genes, conserved synteny, 29 bp gap with previous gene—no ORF within the gap, no coding potential, none of closest relatives call a gene in the gap. Not called in Glimmer and GeneMark, however chosen start has stronger RBS scores (final score: -5.138, z-value: 2.368, spacer distance: 16) and is the most called start in Starterator (called 90.4% of the time when present). **FUNCTION:** Function called by HHPred, portal protein called by 30 plus top genes in BLAST, called by 82% (593 out of 726) of phages in same pham (228153), called by three closest whole-genome relatives.

**Gene 9**

**1. Final Gene Information:**

**(a) FINAL Gene Co-ordinates: Start: 6360, Stop: 8933**

**(b) Is it a protein-coding gene? Yes**

**(c) What is its function? Capsid maturation protease**

**(d) Final Summary: START:** Not longest ORF—longest ORF overlaps by 130 bp, called start overlaps by 1 bp and is most manual called start in Starterator (called 95.1% of the time when present), supported by Glimmer (final score: -5.111, z-value: 2.103, spacer distance: 15), strong coding potential in +1 frame from ∼6700–8900 save for rises and falls and other frame with minimal coding potential from ∼7300–7600. 1:1 target:query for top 25 genes, conserved synteny. Genemark start creates a gap >300 bp and has zero calls in Starterator. **FUNCTION:** Capsid maturation protease called by 5 of top 10 genes in BLAST, variations of this call called by 77% (384 out of 497) of phages in same pham (90), called by 3 most-related whole-genome relatives, function not called in HHpred.

**Gene 10**

**1. Final Gene Information:**

**(a) FINAL Gene Co-ordinates: Start: 9213, Stop: 9350**

**(b) Is it a protein-coding gene? Yes**

**(c) What is its function? Hypothetical protein**

**(d) Final Summary: START:** Longest ORF, same call in Glimmer and Genemark (final score: -6.259, z-value: 1.81, spacer distance: 16), strong coding potential in +1 frame from ∼9150– 9300, 1:1 target:query for top 2 genes, most manual called start in Starterator (called 100% of the time when present), conserved synteny, 279 base gap with previous gene—multiple ORFs in gap however there is no coding potential and no close relatives call a gene in the gap. **FUNCTION:** Hypothetical protein called by all top 5 genes in BLAST, called by 100% (262 out of 262) of phages in same pham (84814), called by three closest whole-genome relatives, function not called by HHPred.

**Gene 11**

**1. Final Gene Information:**

**(a) FINAL Gene Co-ordinates: Start: 9464, Stop: 11209**

**(b) Is it a protein-coding gene? Yes**

**(c) What is its function? Major capsid hexamer protein**

**(d) Final Summary: START:** Longest ORF, same call in Glimmer and Genemark (final score: -2.794, z-value: 3.16, spacer distance: 13), strong coding potential in +2 frame from ∼9550–11200, 1:1 target:query for top 3 genes, most manual called start in Starterator (called 97.7% of the time when present), conserved synteny. 113 bp gap with previous gene—ORF present in gap however no coding potential and none of closest relatives call a gene in the gap. **FUNCTION:** Major capsid hexamer protein called by 6 of top 10 genes in BLAST, called by 85% (431 out of 507) of phages in same pham (228170), called by three closest whole-genome relatives, function not called by HHPred.

**Gene 12**

**1. Final Gene Information:**

**(a) FINAL Gene Co-ordinates: Start: 11309, Stop: 12106**

**(b) Is it a protein-coding gene? Yes**

**(c) What is its function? Major capsid pentamer protein.**

**(d) Final Summary: START:** Longest ORF, same call in Glimmer and Genemark (final score: -3.613, z-value: 2.783, spacer distance: 11), strong coding potential in +2 frame from ∼11300–12100, 1:1 target:query for top 3 genes, most manual called start in Starterator (called 99.8% of the time when present), conserved synteny. 99 bp gap with previous gene—ORF present in gap however no coding potential and no close relatives call a gene in the gap. **FUNCTION:** Major capsid pentamer protein called by 8 of top 10 genes in BLAST, called by 83.8% (424 out of 506) of phages in same pham (227117), called by three closest whole-genome relatives, function not called by HHPred.

**Gene 13**

**1. Final Gene Information:**

**(a) FINAL Gene Co-ordinates: Start: 12118, Stop: 12954**

**(b) Is it a protein-coding gene? Yes**

**(c) What is its function? Hypothetical protein**

**(d) Final Summary: START:** Not longest ORF—longest ORF overlaps previous gene by 286 bp—called ORF is first ORF without significant overlap (11 bp gap), same call in Glimmer and Genemark (final score: -4.525, z-value: 2.345, spacer distance: 12). Strong coding potential in +3 frame from ∼1220–12920 save for ∼12600-12700, 1:1 target:query for top 3 genes, most manual called start in Starterator (called 100% of the time when present), conserved synteny. **FUNCTION:** Hypothetical protein called by 9 of top 10 genes in BLAST, called by 100% (458 out of 458) of phages in same pham (228180), called by three closest whole-genome relatives, function not called by HHPred.

**Gene 14**

**1. Final Gene Information:**

**(a) FINAL Gene Co-ordinates: Start: 12954, Stop: 13343**

**(b) Is it a protein-coding gene? Yes**

**(c) What is its function? Hypothetical protein**

**(d) Final Summary: START:** Longest ORF, 1 bp overlap with previous gene, same call in Glimmer and Genemark (final score: -3.558, z-value: 2.834, spacer distance: 12). Strong coding potential in +1 frame from ∼12980–13300, 1:1 target:query for top 3 genes, most manual called start in Starterator (called 97.4% of the time when present), conserved synteny. **FUNCTION:** Hypothetical protein called by 3 of top 10 genes in BLAST, called by 86% (314 out of 367) of phages in same pham (228207), called by three closest whole-genome relatives, function not called by HHPred. No transmembrane domains, therefore, not holin which was called by 14% of phages in same pham.

**Gene 15**

**1. Final Gene Information:**

**(a) FINAL Gene Co-ordinates: Start: 13402, Stop: 13701**

**(b) Is it a protein-coding gene? Yes**

**(c) What is its function? Hypothetical protein**

**(d) Final Summary: START:** Not longest ORF, same call in Glimmer and Genemark (final score: -6.174, z-score: 1.629, spacer: 14), most manual called start in Starterator (called 100% of the time when present), strong coding potential in +3 frame from ∼13400–13650, 1:1 target:query for top 3 genes, conserved synteny. 58 bp gap with previous gene however no coding potential within gap and no close relatives call a gene in gap. **FUNCTION:** Hypothetical protein called by 10 of top 10 genes in BLAST, called by 100% (282 out of 282) of phages in called pham, called by 1 of 3 closest whole-genome relatives, function not called by HHPred.

**Gene 16**

**1. Final Gene Information:**

**(a) FINAL Gene Co-ordinates: Start: 13698, Stop: 14312**

**(b) Is it a protein-coding gene? Yes**

**(c) What is its function? Hypothetical protein**

**(d) Final Summary: START:** Longest ORF, 4 bp overlap with previous gene, same call in Glimmer and Genemark (final score: -4.558, z-value: 2.305, spacer distance: 11). Strong coding potential in +1 frame from ∼13680–14300 save for dip, 1:1 target:query for top 3 genes, most manual called start in Starterator (called 97.7% of the time when present), conserved synteny. **FUNCTION:** Function called by HHPred, hypothetical protein called by 9 of top 10 genes in BLAST, called by 100% (452 out of 452) of phages in same pham (227151), called by 3 most-related whole-genome relatives.

**Gene 17**

**1. Final Gene Information:**

**(a) FINAL Gene Co-ordinates: Start: 14400, Stop: 15200**

**(b) Is it a protein-coding gene? Yes**

**(c) What is its function? Major tail protein**

 **(d) Final Summary: START:** Longest ORF, same call in Glimmer and Genemark (final score: -3.927, z-value: 2.624, spacer distance: 11), strong coding potential in +1 frame from ∼ 14450–15150, 1:1 target:query for top 20 genes, most manual called start in Starterator (called 100% of the time when present), conserved synteny. 87 bp gap with previous gene—ORF present in gap however no coding potential and no close relatives call a gene within the gap. **FUNCTION:** Major tail protein called by 10 of 10 top genes in BLAST, called by 84% (644 out of 765) of phages in same pham (225750), called by three closest whole-genome relatives, function not called by HHPred.

**Gene 18**

**1. Final Gene Information:**

**(a) FINAL Gene Co-ordinates: Start: 15490, Stop: 15260 (Reverse)**

**(b) Is it a protein-coding gene? Yes**

**(c) What is its function? Hypothetical protein**

**(d) Final Summary: START:** Longest ORF, call supported by Genemark (final score: -3.043, z-value: 3.184, spacer distance: 7), most manual called start in Starterator (called 97.5% of the time when present), strong coding potential in -1 frame, 1:1 target:query for top 10 genes, conserved synteny. 77 bp gap with previous gene, however no coding potential within the gap and no close relatives call a gene in the gap. Start not called by Glimmer. **FUNCTION:** Hypothetical protein called by 10 of 10 top genes in BLAST, called by 100% (537 out of 537) of phages in same pham (76668), called by three closest whole-genome relatives, function not called by HHPred.

 **Gene 19**

**1. Final Gene Information:**

**(a) FINAL Gene Co-ordinates: Start: 16344, Stop: 15568 (Reverse)**

**(b) Is it a protein-coding gene? Yes**

**(c) What is its function? DpdA-like tRNA-guanine transglycosylase**

**(d) Final Summary: START:** Not longest ORF—longest ORF overlaps previous gene by 316 bp, called ORF is first ORF without significant overlap (56 bp gap, close relatives have similarly sized gap with no gene called within), same call in Glimmer and Genemark (final score: -3.423, z-value: 2.079, spacer distance: 11). Strong coding potential in -3 frame from ∼16320–15600, 1:1 target:query for top 21 genes, most manual called start in Starterator (called 98.9% of the time when present), conserved synteny. **FUNCTION:** Function called by HHPred, dpdA-like tRNA-guanine transglycosylase called by 10 of top 10 genes in BLAST (9 use improper name—queuine tRNA-ribosyltransferase), called by 89% (328 out of 369) of phages in same pham (208288) (289 under improper name of queuine tRNA-ribosyltransferase), called by three closest whole-genome relatives (under proper name).

**Gene 20**

**1. Final Gene Information:**

**(a) FINAL Gene Co-ordinates: Start: 16664, Stop: 16401 (Reverse)**

**(b) Is it a protein-coding gene? Yes**

**(c) What is its function? Hypothetical protein**

**(d) Final Summary: START:** Not longest ORF with gap of 109 bp, but same call in Glimmer and Genemark (final score: -3.307, z-score: 3.143, spacer distance: 7), most manual called start in Starterator (called 65.8% of the time when present), strong coding potential in -2 frame from ∼16650–16400, 1:1 target:query for 6 of top 10 genes, conserved synteny. **FUNCTION:** Hypothetical protein called by 10 of top 10 genes in BLAST, called by 100% (284 out of 284) of phages in same pham (252), called by three closest whole-genome relatives. Function not called by HHPred.

**Gene 21**

**1. Final Gene Information:**

**(a) FINAL Gene Co-ordinates: Start: 16774, Stop: 17511**

**(b) Is it a protein-coding gene? Yes**

**(c) What is its function? Head-to-tail adaptor**

**(d) Final Summary: START:** Longest ORF, same call in Glimmer and Genemark (final score: -6.383, z-value: 1.494, spacer distance: 11), most manual called start in Starterator (called 86.0% of the time when present), strong coding potential in +3 frame from ∼16800–17500, 1:1 target:query for top 10 genes, conserved synteny, 73 bp gap with previous gene—no ORF present or coding potential within gap and no close relatives call a gene in gap. **FUNCTION:** Function called by HHPred, head-to-tail adaptor called by 10 of top 10 genes in BLAST, called by 80.4% (640 out of 796) of phages in same pham (217440), called by three closest whole-genome relatives.

**Gene 22**

**1. Final Gene Information:**

**(a) FINAL Gene Co-ordinates: Start: 17511, Stop: 18029**

**(b) Is it a protein-coding gene? Yes**

**(c) What is its function? Hypothetical protein**

**(d) Final Summary: START:** Not longest ORF—longest ORF overlaps with previous gene over 500 bases—called ORF is first ORF without significant overlap (1 bp overlap), same call in Glimmer and Genemark (final score: -4.583, z-value: 2.136, spacer distance: 9). Most manual called start in Starterator (called 97.9% of the time when present), strong coding potential in +1 frame from ∼17500–18000, 1:1 target:query for top 9 genes, conserved synteny. **FUNCTION:** Hypothetical protein called by 10 of top 10 genes in BLAST, called by 100% (506 out of 506) of phages in same pham (224150), called by three closest whole-genome relatives, function not called by HHPred.

**Gene 23**

**1. Final Gene Information:**

**(a) FINAL Gene Co-ordinates: Start: 18045, Stop: 18374**

**(b) Is it a protein-coding gene? Yes**

**(c) What is its function? Hypothetical protein**

**(d) Final Summary: START:** Not longest ORF, but supported by both Glimmer and Genemark (final score: -3.899, z-score: 2.467, spacer distance: 11), most manual called start in Starterator (called 77.2% of the time when present), 15 bp gap. Strong coding potential in +1 frame from ∼18050–18350, 1:1 target:query for 5 of top 10 genes, conserved synteny. **FUNCTION:** Function called by HHPred, hypothetical protein called by 10 of top 10 genes in BLAST, called by 99% (577 out of 580) of phages in same pham (225763), called by three closest whole-genome relatives.

**Gene 24**

**1. Final Gene Information:**

**(a) FINAL Gene Co-ordinates: Start: 18379, Stop: 18801**

**(b) Is it a protein-coding gene? Yes**

**(c) What is its function? Hypothetical Protein**

**(d) Final Summary: START:** Not longest ORF—longest ORF overlaps by 62 bp with previous gene—called ORF is first ORF without significant overlap (4 base gap), same call in Glimmer and Genemark (final score: -5.587, z-value: 1.847, spacer distance: 8), most manual called start in Starterator (called 99.6% of the time when present). Strong coding potential in +3 frame from ∼18400–18750 save for some reduction in coding potential at ∼18700-18750, 1:1 target:query for top 3 genes, conserved synteny. **FUNCTION:** Hypothetical protein called by 1 of top 10 genes in BLAST, called by 46% (325 out of 707) of phages in same pham (228154), called by 2 of 3 closest whole-genome relatives. No translational frameshift; therefore, no evidence for tail assembly chaperone function. Function not called by HHPred.

**Gene 25**

**1. Final Gene Information:**

**(a) FINAL Gene Co-ordinates: Start: 18852, Stop: 19415**

**(b) Is it a protein-coding gene? Yes**

**(c) What is its function? Hypothetical protein**

**(d) Final Summary: START:** Not longest ORF—longest ORF overlaps with previous gene by 121 bp—called ORF is first ORF without significant overlap (50 bp gap), same call in Glimmer and Genemark (final score: -3.030, z-value: 2.899, spacer distance: 11). Strong coding potential in +1 frame from ∼18850–19350, 1:1 target:query for top 25 genes, most manual called start in Starterator (called 100% of the time when present), conserved synteny.

**FUNCTION:** Hypothetical protein called by 8 of top 10 genes in BLAST, called by 95% (446 out of 468) of phages in same pham (225780), called by three closest whole-genome relatives. No translational frameshift; therefore, no evidence for tail assembly chaperone function. Function not called by HHPred.

**Gene 26**

**1. Final Gene Information:**

**(a) FINAL Gene Co-ordinates: Start: 19819, Stop: 19478 (Reverse)**

**(b) Is it a protein-coding gene? Yes**

**(c) What is its function? Hypothetical protein.**

**(d) Final Summary: START:** Not longest ORF—longest ORF has poor RBS and encompasses non-coding regions, same call in Glimmer and Genemark (final score: -3.030, z-value: 2.899, spacer distance: 11), most manual called start in Starterator (called 81.9% of the time when present),119 bp gap with previous gene—no ORF in any reading frame and no close relatives call a gene in the gap. Strong coding potential in -1 frame from ∼19800-19500 (minor dip at ∼19600), 1:1 target:query for 5 of top 10 genes, conserved synteny. **FUNCTION:** Hypothetical protein called by 10 of top 10 genes in BLAST, called by 100% (293 out of 293) of phages in same pham (84764), called by three closest whole-genome relatives.Function not called by HHPred.

**Gene 27**

**1. Final Gene Information:**

**(a) FINAL Gene Co-ordinates: Start: 19939, Stop: 25923**

**(b) Is it a protein-coding gene? Yes**

**(c) What is its function? Tape measure protein**

**(d) Final Summary: START:** Longest ORF, same call in Glimmer and Genemark (final score: - 7.392, z-value: 0.983, spacer distance: 7), most manual called start in Starterator (called 99.8% of the time when 1present), strong coding potential in +3 frame from ∼20000–26000 (save for few dips and minor coding potential in -3 frame at ∼20000–20400), 1:1 target:query for top 10 genes, conserved synteny. 119 bp gap with previous gene—ORF present in gap however no coding potential and no close relative call a gene within the gap. **FUNCTION:** Function not supported by HHPred score (low coverage), tape measure protein called by 10 of top 10 genes in BLAST, called by 87% (412 out of 472) of phages in same pham (100), called by three closest whole-genome relatives.

**Gene 28**

**1. Final Gene Information:**

**(a) FINAL Gene Co-ordinates: Start: 25933, Stop: 27366**

**(b) Is it a protein-coding gene? Yes**

**(c) What is its function? Minor tail protein**

**(d) Final Summary: START:** Longest ORF, same call in Glimmer and Genemark (final score: -4.112, z-value: 2.510, spacer distance: 10), most manual called start in Starterator (called 99.4% of the time when present), 9 bp gap with previous gene. Moderate to strong coding potential in +3 frame from ∼26000–27350 (many dips, but tall, clear mesa otherwise), 1:1 target:query for top 30 genes, conserved synteny. **FUNCTION:** Minor tail protein called by 10 of top 10 genes in BLAST (called by 30+ genes total), called by 86% (427 out of 498) of phages in same pham (219275), called by three closest whole-genome relatives. Function not called by HHPred.

**Gene 29**

**1. Final Gene Information:**

**(a) FINAL Gene Co-ordinates: Start: 27363, Stop: 28475**

**(b) Is it a protein-coding gene? Yes**

**(c) What is its function? Minor tail protein**

**(d) Final Summary: START:** Longest ORF, same call in Glimmer and Genemark (final score: -4.470, z-value: 2.372, spacer distance: 12), most manual called start in Starterator (called 96.9% of the time when present), 4 bp overlap with previous gene. Strong coding potential in +1 frame from ∼27400–28450 (dip at ∼27650 and ∼28350), 1:1 target:query for top 30 genes, conserved synteny. **FUNCTION:** Function called by HHPred, minor tail protein called by 10 of top 10 genes in BLAST (called by 30+ genes total), called by 78% (507 out of 654) of phages in same pham (187775), called by three closest whole-genome relatives.

**Gene 30**

**1. Final Gene Information:**

**(a) FINAL Gene Co-ordinates: Start: 28472, Stop: 30727**

**(b) Is it a protein-coding gene? Yes**

**(c) What is its function? Minor tail protein**

**(d) Final Summary: START:** Longest ORF, same call in Glimmer and Genemark (final score: -5.791, z-value: 1.730, spacer distance: 14), most manual called start in Starterator (called 99.7% of the time when present), 4 base overlap with previous gene. Strong coding potential in +2 frame from ∼28500–30700, 1:1 target:query for top 25 genes, conserved synteny. **FUNCTION:** Minor tail protein called by 7 of top 10 genes in BLAST (called by 50+ genes total), called by 78% (313 out of 400) of phages in same pham (219305), called by three closest whole-genome relatives. Function called by HHPred.

**Gene 31**

**1. Final Gene Information:**

**(a) FINAL Gene Co-ordinates: Start: 30731, Stop: 32077**

**(b) Is it a protein-coding gene? Yes**

**(c) What is its function? Minor tail protein**

**(d) Final Summary: START:** Longest ORF, same call in Glimmer and Genemark (final score: -3.906, z-value: 2.614, spacer distance: 10), most manual called start in Starterator (called 100% of the time when present), 3 bp gap with previous gene. Strong coding potential in +2 frame from ∼30700–32050, 1:1 target:query for top 35 genes, conserved synteny. **FUNCTION:** Minor tail protein called by 10 of top 10 genes in BLAST (called by 50+ genes total), called by 87% (346 out of 400) of phages in same pham (141), called by three closest whole-genome relatives. Function called by HHPred.

**Gene 32**

**1a. FINAL Gene Co-ordinates:Start: 32080, Stop: 33240**

**1b. Is it a protein-coding gene*?* Yes**

**1c. What is its function?Minor tail protein**

**1d.Final Summary: START:** Longest ORF, same call in Glimmer and Genemark (final score: -3.751, z-value: 2.55, spacer distance: 9), most annotated start in Starterator (called 97.1% of the time when present), 2 bp gap with previous gene. Strong coding potential in +3 frame from ~32100-33200, 1:1 target:query for top 10 genes, conserved synteny. **FUNCTION:** Minor tail protein function was called by 10 of top 10 genes in BLAST, called by 80% (271 out of 340) of phages in same pham (225762), called by three closest whole-genome relatives. Function called by HHPred.

**Gene 33**

**1a. FINAL Gene Co-ordinates: Start: 33271, Stop: 33660**

**1b. Is it a protein-coding gene? Yes**

**1c. What is its function? Hypothetical protein**

**1d.Final Summary: START:** Longest ORF, same call in Glimmer and Genemark (final score: -6.520, z-value: 1.584, spacer distance: 15), most annotated start in Starterator (called 95.6% of the time when present), 30 bp gap with previous gene—ORF present however no coding potential and no close relatives call a gene within the gap. Strong coding potential in +3 frame from ~33300-33600, 1:1 target:query for top 10 genes, conserved synteny. **FUNCTION:** Hypothetical protein was called by 9 of top 10 genes in BLAST, called by 99% (610 out of 614) of phages in same pham (84641), called by three closest whole-genome relatives. Function not called by HHPred.

**Gene 34**

**1a. FINAL Gene Co-ordinates: Start: 33669, Stop: 34055**

**1b. Is it a protein-coding gene? Yes**

**1c. What is its function? Hypothetical protein**

**1d.Final Summary: START:** Not longest ORF—longest ORF overlaps with previous gene over 100 bases—called ORF is first ORF without significant overlap (8 bp gap), same call in Glimmer and Genemark (final score: -3.524, z-value: 2.885, spacer distance: 8). Most manual called start in Starterator (called 98.5% of the time when present), strong coding potential in +1 frame from ∼33700–34000, 1:1 target:query for top 20 genes, conserved synteny. **FUNCTION:** Hypothetical protein called by 20 of top 20 genes in BLAST, called by 98% (466 out of 474) of phages in same pham (228173), called by three closest whole-genome relatives, function not called by HHPred.

**Gene 35**

**1a. FINAL Gene Co-ordinates: Start: 34052, Stop: 34702**

**1b. Is it a protein-coding gene? Yes**

**1c. What is its function? Hypothetical protein**

**1d.Final Summary: START:** Not the longest ORF—longest ORF overlaps with previous gene by 40 bp—called ORF is first ORF without significant overlap (4 bp overlap), most manual called start in Starterator (called 74.8% of time when present), strong coding potential in +2 frame from ~34050-34670, 1:1 target:query for 7 of top 10 genes, conserved synteny. Start not called by Glimmer or Genemark. **FUNCTION:** Hypothetical protein called by top 20 similar genes in BLAST, called by 100% (994 out of 994) of phages in same pham (227106), called by three closest whole-genome relatives, function not called by HHPred.

**Gene 36**

**1a. FINAL Gene Co-ordinates: Start: 34699, Stop: 35298**

**1b. Is it a protein-coding gene? Yes**

**1c. What is its function? Hypothetical protein**

**1d.Final Summary: START:** Not the longest ORF—longest ORF overlaps with previous gene by 208 bp—called ORF is first ORF without significant overlap (4 bp overlap), same call by Glimmer and Genemark (final score: -3.489, z-value: 2.71, spacer distance: 12), most manual called start in Starterator (called 74.8% of time when present). Strong coding potential in the +3 frame from ~34720-35250 with a dip from ~34950-3520, 1:1 target:query for 10 of top 10 genes, conserved synteny. **FUNCTION:** Hypothetical protein called by top 30 similar genes in BLAST, called by 100% (994 out of 994) of phages in same pham (227106), called by three closest whole-genome relatives, function not called by HHPred.

**Gene 37**

**1a. FINAL Gene Co-ordinates: Start: 35295, Stop: 35918**

**1b. Is it a protein-coding gene? Yes**

**1c. What is its function? Hypothetical protein**

**1d.Final Summary: START:** Not the longest ORF—longest ORF overlaps with previous gene by 46 bp—called ORF is first ORF without significant overlap (4 bp overlap), same call by Glimmer and Genemark (final score: -4.516, z-value: 2.2, spacer distance: 12), most manual called start in Starterator (called 97.6% of time when present). Strong coding potential in the +1 frame from ~35310-35890, 1:1 target:query for top 20 genes, conserved synteny. **FUNCTION:** Hypothetical protein called by top 30 similar genes in BLAST, called by 100% (384 out of 384) of phages in same pham (172562), called by three closest whole-genome relatives, function not called by HHPred.

**Gene 38**

**1a. FINAL Gene Co-ordinates: Start: 35952, Stop: 36320**

**1b. Is it a protein-coding gene? Yes**

**1c. What is its function? Hypothetical protein**

**1d.Final Summary: START:** Not the longest ORF, however same call by Glimmer and Genemark (final score: -4.292, z-value: 2.272, spacer distance: 11) and most manual called start in Starterator (called 95.1% of time when present). Strong coding potential in the +1 frame from ~35990-36250, 1:1 target:query for 7 of top 10 genes, conserved synteny. 33 bp overlap, no coding potential in gap and no close relatives call a gene within the gap. **FUNCTION:** Hypothetical protein called by top 20 similar genes in BLAST, called by 100% (491 out of 491) of phages in same pham (87), called by three closest whole-genome relatives, function not called by HHPred.

**Gene 39**

**1a. FINAL Gene Co-ordinates: Start: 36320, Stop: 36631**

**1b. Is it a protein-coding gene? Yes**

**1c. What is its function? Hypothetical protein**

**1d.Final Summary: START:** Not the longest ORF, however same call by Glimmer and Genemark (final score: -7.729, z-value: 0.855, spacer distance: 14) and most manual called start in Starterator (called 67.8% of time when present), 1 bp overlap. Strong coding potential in the +2 frame from ~36300-36580, 1:1 target:query for 6 of top 10 genes, conserved synteny. **FUNCTION:** Hypothetical protein called by top 20 similar genes in BLAST, called by 100% (495 out of 495) of phages in same pham (202564), called by three closest whole-genome relatives, function not called by HHPred.

**Gene 40**

**1a. FINAL Gene Co-ordinates: Start: 36671, Stop: 37078**

**1b. Is it a protein-coding gene? Yes**

**1c. What is its function? Hypothetical protein**

**1d.Final Summary: START:** Longest ORF, same call by Glimmer and Genemark (final score: -7.581, z-value: 0.867, spacer distance: 8) and most manual called start in Starterator (called 93.4% of time when present), strong coding potential in the +2 frame from ~36700-37050, 1:1 target:query for 19 of top 20 genes, conserved synteny. 39 bp gap, no coding potential within gap and no closest relatives call a gene within the gap. **FUNCTION:** Hypothetical protein called by top 20 similar genes in BLAST, called by 99% (492 out of 497) of phages in same pham (81), called by three closest whole-genome relatives, function not called by HHPred.

**Gene 41**

**1a. FINAL Gene Co-ordinates: Start: 37078, Stop: 38199**

**1b. Is it a protein-coding gene? Yes**

**1c. What is its function? Hypothetical protein**

**1d.Final Summary:** **START:** Longest ORF, same call by Glimmer and Genemark (final score: -5.832, z-value: 2.274, spacer distance: 18) and most manual called start in Starterator (called 99.7% of time when present), 1 bp overlap. Strong coding potential in the +3 frame from ~37100-38180, 1:1 target:query for top 13 genes, conserved synteny. **FUNCTION:** Hypothetical protein called by 9 top 10 similar genes in BLAST, called by 98.6% (362 out of 367) of phages in same pham (162), called by three closest whole-genome relatives, function not called by HHPred.

**Gene 42**

**1a. FINAL Gene Co-ordinates: Start: 38399, Stop: 38181 (reverse)**

**1b. Is it a protein-coding gene? Yes**

**1c. What is its function? Hypothetical protein**

**1d.Final Summary: START:** Not the longest ORF—longest ORF overlaps with previous gene by 146 bp—called ORF is first ORF without significant overlap (1 bp gap), same call by Glimmer and Genemark (final score: -4.649, z-value: 2.133, spacer distance: 12), most manual called start in Starterator (called 98.3% of time when present). Strong coding potential in the -2 frame from ~38220-38350, 1:1 target:query for 9 of top 20 genes, conserved synteny. **FUNCTION:** Hypothetical protein called by top 10 similar genes in BLAST, called by 100% (291 out of 291) of phages in same pham (84773), called by three closest whole-genome relatives, function not called by HHPred.

**Gene 43**

**1a. FINAL Gene Co-ordinates: Start: 38694, Stop: 38401 (reverse)**

**1b. Is it a protein-coding gene? Yes**

**1c. What is its function? Hypothetical protein**

**1d.Final Summary: START:** Not the longest ORF, however is most called start in Starterator (called 80.6% of time when present), 4 bp overlap, strongest RBS scores (final score: -4.513, z-value: 2.131, spacer distance: 10), strong coding potential in the -3 frame from ~38450-38700, 1:1 target:query for 8 of top 10 genes, conserved synteny. Starts called by Glimmer and Genemark are not frequently called in Starterator and have poorer RBS scores. **FUNCTION:** Hypothetical protein called by 9 of top 10 similar genes in BLAST, called by 99% (285 out of 288) of phages in same pham (240), called by three closest whole-genome relatives, function not called by HHPred.

**DELETED THIS AUTO-CALLED GENE**

**1a. Auto-called Gene Co-ordinates: 38,882-40,036**

**1b. Is it a protein-coding gene? No, this is not a gene**

**1c. What is its function? N/A**

**1d. Final Summary:** This gene was chosen to be deleted as it is an orpham, having zero BLAST results. Additionally, there is no coding potential in any forward frames where this gene was auto-called to be and this gene is not called in any close relatives.

**Gene 44**

**1a. FINAL Gene Co-ordinates: Start: 38906, Stop: 38691 (reverse)**

**1b. Is it a protein-coding gene? Yes**

**1c. What is its function? Helix-turn-helix DNA binding domain**

**1d.Final Summary: START:** Not the longest ORF—longest ORF overlaps with previous gene by 106 bp—called ORF is first ORF without significant overlap (4 bp overlap), same call by Glimmer and Genemark (final score: -4.299, z-value: 2.237, spacer distance: 10), most manual called start in Starterator (called 100% of time when present). Strong coding potential in the -2 frame from ~38720-38900, 1:1 target:query for 3 of top 10 genes, conserved synteny. **FUNCTION:** Helix-turn-helix DNA binding domain protein called by 5 of top 10 similar genes in BLAST, called by 82% (276 out of 338) of phages in same pham (222813), called by 1 of 3 closest whole-genome relatives (other two call similar DNA binding protein and DNA binding domain protein functions), function not called by HHPred.

**Gene 45**

**1a. FINAL Gene Co-ordinates: Start: 39433, Stop: 38903 (reverse)**

**1b. Is it a protein-coding gene? Yes**

**1c. What is its function? Helix-turn-helix DNA binding domain**

**1d.Final Summary:** **START:** Longest ORF, same call in Genemark (final score: -5.199, z-value, 1.789, spacer distance: 10), most called start in Starterator (called 100% of time when present), 0 bp gap. Strong coding potential in the -1 frame from ~38950-39400, 1:1 target:query for top 30 genes, conserved synteny. Start not called by Glimmer. **FUNCTION:** Helix-turn-helix DNA binding domain protein called by top 20 similar genes in BLAST, called by 76% (239 out of 315) of phages in same pham (228227), called by three closest whole-genome relatives. Function not called by HHPred.

**Gene 46**

**1a. FINAL Gene Co-ordinates: Start: 39988, Stop: 39434 (reverse)**

**1b. Is it a protein-coding gene? Yes**

**1c. What is its function? Hypothetical protein**

**1d.Final Summary: START:** Not the longest ORF, however is the most called start in Starterator (called 75.3% of time when present), same call by Genemark (final score: -3.212, z-value: 2.848, spacer distance: 12). Strong coding potential in the -1 frame from ~39450-40000, 1:1 target:query for 7 of top 10 genes, conserved synteny. 98 bp gap, no coding potential in gap and no close relatives call a gene within the gap. Start not called by Glimmer. **FUNCTION:** Hypothetical protein called by top 20 similar genes in BLAST, called by 99% (517 out of 5423) of phages in same pham (229809), called by three closest whole-genome relatives. Function not called by HHPred.

**Gene 47**

**1a. FINAL Gene Co-ordinates: Start: 40087, Stop: 41403**

**1b. Is it a protein-coding gene? Yes**

**1c. What is its function? Lysin A**

**1d.Final Summary: START:** Not the longest ORF, however is the most called start in Starterator (called 50.9% of the time when present), same call by Glimmer (final score: -6.691, z-value: 1.117, spacer: 12). Strong coding potential in the +3 frame from ~40100-41350, 1:1 target:query for 8 of top 10 genes, conserved synteny. 98 bp gap, no coding potential in gap and no close relatives call a gene within this gap. Start not called by Genemark. **FUNCTION:** Lysin A called by top 20 similar genes in BLAST, called by 91% (348 out of 381) of phages in same pham (229834), called by three closest whole-genome relatives. Function not called by HHPred.

**Gene 48**

**1a. FINAL Gene Co-ordinates: Start: 41413, Stop: 42768**

**1b. Is it a protein-coding gene? Yes**

**1c. What is its function? Lysin B**

**1d.Final Summary: START:** Longest ORF, same call in Glimmer and Genemark (final score: -4.326, z-value: 2.224, spacer: 10), most annotated start un Starterator (called 100% of time when present), 9 bp gap. Strong coding potential in the +3 frame from ~41450-42700, 1:1 target:query for top 20 genes, conserved synteny. **FUNCTION:** Lysin B called by top 20 similar genes in BLAST, called by 92% (460 out of 501) of phages in same pham (229811), called by three closest whole-genome relatives. Function not called by HHPred.

**Gene 49**

**1a. FINAL Gene Co-ordinates: Start: 43490, Stop: 42870 (reverse)**

**1b. Is it a protein-coding gene? Yes**

**1c. What is its function? Hypothetical protein**

**1d.Final Summary: START:** Longest ORF, same call by Glimmer and Genemark (final score: -2.305, z-value: 3.404, spacer: 13), most called start in Starterator (called 100% of time when present). Strong coding potential in the -2 frame from ~42900-43450, 1:1 target:query for top 12 genes, conserved synteny.57 bp gap, no coding potential within gap and no close relatives call a gene in the gap. **FUNCTION:** Hypothetical protein called by top 20 similar genes in BLAST, called by 100% (493 out of 493) of phages in same pham (221004), called by three closest whole-genome relatives. Function not called by HHPred.

**Gene 50**

**1a. FINAL Gene Co-ordinates: Start: 44951, Stop: 43548 (reverse)**

**1b. Is it a protein-coding gene? Yes**

**1c. What is its function? exonuclease**

**1d.Final Summary: START:** Not the longest ORF—longest ORF overlaps with previous gene by over 300 bp—called ORF is first ORF without significant overlap (2 bp gap), same call by Glimmer and Genemark (final score: -4.173, z-value: 2.848, spacer distance: 16), most manual called start in Starterator (called 84.3% of time when present). Strong coding potential in the -2 frame from ~43600-44920, 1:1 target:query for top 12 genes, conserved synteny. **FUNCTION:** Exonuclease called by 7 of top 10 similar genes in BLAST, called by 2 of 3 closest whole-genome relatives, called by 18% (90 out of 493) of phages in same pham (222760)—~4.7% called helicase function and remaining ~77% call hypothetical protein, function called by HHPred.

**Gene 51**

**1a. FINAL Gene Co-ordinates: Start: 46750, Stop: 44954 (reverse)**

**1b. Is it a protein-coding gene? Yes**

**1c. What is its function? DNA helicase**

**1d.Final Summary: START:** Longest ORF, same start called by three closest-whole genome relatives (final score: -4.863, z-value: 1.988, spacer: 11), second-most called start in Starterator (called 74.7% of time when present)—most called start is called 56% of time when present, 24 bp gap. Strong coding potential -1 frame from ~45000-46600, 1:1 target:query for top 20 genes, conserved synteny. Not called by Glimmer or Genemark. **FUNCTION:** DNA helicase called by top 20 similar genes in BLAST, called by 87% (463 out of 532) of phages in same pham (229807), called by three closest whole-genome relatives, function called by HHPred.

**Gene 52**

**1a. FINAL Gene Co-ordinates: Start: 47614, Stop: 46775 (reverse)**

**1b. Is it a protein-coding gene? Yes**

**1c. What is its function? Hypothetical protein**

**1d.Final Summary: START:** Not longest ORF, call supported by Glimmer but not GeneMark. Gap of 7 between previous gene (final score: -5.440; z-score: 1.844; spacer: 13). Most annotated start in Starterator (called 94.7% of time when present). Coding potential in –1 frame from around 47000-47500 bp, 1:1 query:target with 2 of top 10 top related genes, conserved synteny. **FUNCTION:** Hypothetical protein called by top 10 hits in BLAST, called by 100% (19 of 19) pham members (7040). Called by 3 of 3 top-related phages. Call not supported by HHPred (low coverage, e-values >12).

**Gene 53**

**1a. FINAL Gene Co-ordinates:** *47,837-47,622 bp (reverse)*

**1b. Is it a protein-coding gene?** *Yes*

**1c. What is its function?** *Hypothetical protein*

**1d.Final Summary: START:** Not longest ORF, same call by Glimmer and GeneMark. Overlap of 4 bp (final score: -2.377; z-score: 3.224; spacer: 11). Most annotated start in Starterator (called 100% of time when present). Strong coding potential in -2 frame. 1:1 query:target alignment with 8 of top 10 related genes. Conserved synteny for one of three top-related phages. **FUNCTION:** Hypothetical protein called by 10 of top 10 hits in BLAST, called by 100% (303 of 303) pham members (84757). Not supported by HHPred. No CD or TMD.

**Gene 54**

**1a. FINAL Gene Co-ordinates:** *47,956-47,834 bp (reverse)*

**1b. Is it a protein-coding gene?** *Yes*

**1c. What is its function?** *Hypothetical protein*

**1d.Final Summary: START:** Notlongest ORF, same call by Glimmer and GeneMark. Overlap of 4 bp (final score: -4.760; z-score: 2.078; spacer: 12). Most annotated start in Starterator (called 97.5% of time when present). Strong coding potential in -1 frame. 1:1 query:target alignment with 2 of top 10 related genes. Synteny conserved. **FUNCTION:** Hypothetical protein called by 9 of top 10 hits in BLAST (1 of 10 called gp56), called by 100% (119 of 119) pham members (1074). Not supported by HHPred. No CD or TMD.

**Gene 55**

**1a. FINAL Gene Co-ordinates:** *48,324-47,953 bp (reverse)*

**1b. Is it a protein-coding gene?** *Yes*

**1c. What is its function?** *Hypothetical protein*

**1d.Final Summary: START:** Not longest ORF, same call by Glimmer and GeneMark. Overlap of 7 bp (final score: -7.153; z-score: 1.269; spacer: 15). Does not have most annotated start in Starterator (start called 59.8% of time when present). Strong coding potential in -3 frame. 1:1 query:target alignment with 4 of top 10 related genes. Synteny conserved with 1 of 3 top-related phages. **FUNCTION:** Hypothetical protein called by 10 of top 10 hits in BLAST, called by 99.6% (773 of 775) pham members (217441). Not supported by HHPred. No CD or TMD.

**Gene 56**

**1a. FINAL Gene Co-ordinates:***48,608-48,318 bp (reverse)*

**1b. Is it a protein-coding gene*?*** *Yes*

**1c. What is its function?***Hypothetical protein*

**1d.Final Summary: START:** Not longest ORF, but same call by Glimmer and GeneMark. Gap of 75 bp, but no ORF in gap (final score: -2.016; z-score: 3.404; spacer: 11). Most annotated start in Starterator (called 77.4% of time when present). Strong coding potential in -2 frame. 1:1 query:target alignment with 1 of top 10 related genes. Synteny conserved for 1 of top 3 related phages. **FUNCTION:** Hypothetical protein called by 10 of top 10 hits in BLAST, called by 99.1% (864 of 871) pham members (163226). Not supported by HHPred. No CD or TMD.

**Gene 57**

**1a. FINAL Gene Co-ordinates:***51,431-48,684 bp (reverse)*

**1b. Is it a protein-coding gene*?*** *Yes*

**1c. What is its function?***DNA polymerase I*

**1d.Final Summary: START:** Not longest ORF, same call by GeneMark but not Glimmer. Gap of 78 bp, but no ORF within gap (final score: -5.549; z-score: 1.655; spacer: 9). Most annotated start in Starterator (called 74.1% of time when present). Strong coding potential in -2 frame. 1:1 query:target alignment with 2 of top 10 related genes. Conserved synteny. **FUNCTION:** DNA polymerase I called by 9 of top 10 hits in BLAST, called by 19% (95 of 495) pham members (84). Supported by CDD. TMD predicts intracellular localization. Not supported by HHPred.

**Gene 58**

**1a. FINAL Gene Co-ordinates:***52,199-51,570 bp (reverse)*

**1b. Is it a protein-coding gene*?*** *Yes*

**1c. What is its function?***Hypothetical protein*

**1d.Final Summary: START:** Longest ORF, same call by Glimmer and GeneMark. Overlap of 4 bp (final score: -3.662; z-score: 2.729; spacer: 13). Does not have most annotated start in Starterator (start called 21.3% of time when present). Strong coding potential in -2 frame from 51600-52100 bp. 1:1 query:target alignment with 2 of top 10 related genes. Synteny conserved with one of top-related phages. **FUNCTION:** Hypothetical protein called by 9 of top 10 hits in BLAST (1 call of gp57), called by 100% (290 of 290) pham members (228236). Supported by HHPred (top hit is domain of unknown function).

**Gene 59**

**1a. FINAL Gene Co-ordinates:***52,711-52,196 bp (reverse)*

**1b. Is it a protein-coding gene*?*** *Yes*

**1c. What is its function?***HNH endonuclease*

**1d.Final Summary: START:** Not longest ORF, same call by Glimmer and GeneMark. Gap of 8 bp (final score: -5.627; z-score: 1.901; spacer: 14). Does not have most annotated start in Starterator (start called 60.0% of time when present). Weak coding potential in -1 frame from 52350-52500 bp. 1:1 query:target alignment with 1 of top 10 related genes. Synteny conserved with one of three top-related phages. **FUNCTION:** HNH endonuclease or NUMOD4 motif-containing NHN endonuclease called by 9 of top 10 hits in BLAST, called by 81% (162 of 200) pham members (228309). Call supported by HHPred and CDD.

**Gene 60**

**1a. FINAL Gene Co-ordinates:***54,579-52,720 bp (reverse)*

**1b. Is it a protein-coding gene*?*** *Yes*

**1c. What is its function?***DNA polymerase I*

**1d.Final Summary: START:** Not longest ORF, same call by Glimmer and GeneMark. Overlap of 4 bp (final score: -4.888; z-score: 1.975; spacer: 11). Does not call most annotated start in Starterator (when present, start called 97.9% of time). Strong coding potential in -3 frame. 1:1 query:target alignment with 9 of top 10 related genes. Synteny conserved with one of top-three related phages. **FUNCTION:** DNA polymerase I called by 9 of top 10 hits in BLAST, called by 83% (1642 of 1972) pham members (229768). Supported by HHPred and CDD.

**Gene 61**

**1a. FINAL Gene Co-ordinates:***55,034-54,576 bp (reverse)*

**1b. Is it a protein-coding gene*?*** *Yes*

**1c. What is its function?***Hypothetical protein*

**1d.Final Summary: START:** Longest ORF, same call by Glimmer and GeneMark. Overlap of 8 bp (final score: -5.613; z-score: 1.846; spacer: 8). Most annotated start in Starterator (called 95.4% of time when present). Strong coding potential in -2 frame. 1:1 query:target alignment with 10 of top 10 related genes. Conserved synteny with one of three top-related phages. **FUNCTION:** Hypothetical protein called by 10 of top 10 hits in BLAST, called by 100% (417 of 417) pham members (228187). Not supported by HHPred.

**Gene 62**

**1a. FINAL Gene Co-ordinates:***55,279-55,127 bp (reverse)*

**1b. Is it a protein-coding gene*?*** *Yes*

**1c. What is its function?***Hypothetical protein*

**1d.Final Summary: START:** Not longest ORF, same call by GeneMark (no Glimmer call). Overlap of 1 bp (final score: -3.751; z-score: 2.55; spacer: 9). Most annotated start in Starterator (called 81.9% of time when present). (Strong) Coding potential in -1 frame. 1:1 query:target alignment with 5 of top 10 related genes. Synteny not conserved. **FUNCTION:** Hypothetical protein called by 10 of top 10 hits in BLAST, called by 100% (271 of 271) pham members (274). Not supported by HHPred.

**Gene 63**

**1a. FINAL Gene Co-ordinates:***55,635-55,279 bp (reverse)*

**1b. Is it a protein-coding gene*?*** *Yes*

**1c. What is its function?***Hypothetical protein*

**1d.Final Summary: START:** Not longest ORF, same call by Glimmer and GeneMark. Overlap of 1 bp (final score: -4.161; z-score: 2.63; spacer: 14). Most annotated start in Starterator (called 81.6% of time when present). Strong coding potential in -3 frame. 1:1 query:target alignment with 5 of top 10 related genes. Synteny not conserved. **FUNCTION:** Hypothetical protein called by 10 of top 10 hits in BLAST, called by 100% (289 of 289) pham members (224230). Not supported by HHPred.

**Gene 64**

**1a. FINAL Gene Co-ordinates:***56,107-55,748 bp (reverse)*

**1b. Is it a protein-coding gene*?*** *Yes*

**1c. What is its function?***Hypothetical protein*

**1d.Final Summary: START:** Not longest ORF, same call by Glimmer and GeneMark. Gap of 192 bp (final score: -4.575; z-score: 2.131; spacer: 11). Most annotated start in Starterator (called 83.2% of time when present). Strong coding potential in -1 frame. 1:1 query:target alignment with 6 of top 10 related genes. Synteny not conserved. **FUNCTION:** Hypothetical protein called by 10 of top 10 hits in BLAST, called by 100% (456 of 456) pham members (pham #). Not supported by HHPred.

**Gene 65**

**1a. FINAL Gene Co-ordinates:** *56,378-56,494 bp (forward)*

**1b. Is it a protein-coding gene?** *Yes*

**1c. What is its function?** *Hypothetical protein*

**1d. Final Summary: START**: Not longest ORF, call supported by Glimmer and GeneMark. Gap of 270 bp, including direction change (final score: -6.182; z-score: 1.563; spacer: 8). Same length as 43/106 (40.5%) of pham, highest percentage. Not most annotated start in Starterator (called 38.7% of time when present, most annotated start is 41.5%). Weak coding potential in +2 frame. 1:1 query:target alignment with 2 of top 10 related genes. Synteny conserved. **FUNCTION:** Hypothetical protein called by 10 of top 10 hits in BLAST, called by 100% (106 of 106) pham members (85366). Not supported by HHPred.

**Gene 66**

**1a. Final Gene Coordinates:** *56,573-57,334 bp (forward)*

**1b. Is it a protein coding gene?** *Yes*

**1c. What is its function?** *Hypothetical protein*

**1d. Final Summary: START:** Longest ORF, call supported by Glimmer and GeneMark. Gap of 78 bp (final score: -7.017; z-score: 1.684; spacer: 18). Most annotated start in Starterator (called 55.6% of time when present). Strong coding potential in +2 frame. 1:1 query:target alignment with X of top 10 related genes. Synteny conserved. **FUNCTION:** Hypothetical protein called by 9 of top 10 hits in BLAST (one call of gp65), called by 99.6% (473 of 475) pham members (229818). Not supported by HHPred.

**Gene 67**

**1a. FINAL Gene Co-ordinates:** *57,347-57,619 bp (forward)*

**1b. Is it a protein-coding gene?**  *Yes*

**1c. What is its function?** *Hypothetical protein*

**1d. Final Summary: START:** Added gene. Longest ORF, no call by Glimmer or GeneMark. Gap of 12 bp (final score: -4.513; z-score: 2.201; spacer: 12). Weak coding potential in +2 frame. 1:1 query:target alignment with 4 of top 10 related genes. Conserved synteny. No Starterator data. **FUNCTION:** Hypothetical protein called by 7 of top 10 hits in BLAST (three hits of ribbon-helix-helix DNA binding domain protein). HHPred shows alignments of up to 52% coverage. TMHMM predicts intracellular localization.

**Gene 68**

**1a. Final Gene Coordinates:** *57,651-57,833 bp (forward)*

**1b. Is it a protein coding gene?** *Yes*

**1c. What is its function?** *Hypothetical protein*

**1d. Final Summary: START:** Not longest ORF, call supported by GeneMark but not Glimmer. Gap/Overlap (final score: -4.874; z-score: 2.214; spacer: 8). Most annotated start in Starterator (called 79.7% of time when present). Same length as 79.1% (219/277) pham members. Strong coding potential in +3 frame. 1:1 query:target alignment with 7 of top 10 related genes. Downstream synteny conserved with three top related phages. **FUNCTION:** Hypothetical protein called by 10 of top 10 hits in BLAST, called by 100% (277 of 277) pham members (84796). Not supported by HHPred.

**Gene 69**

**1a. Final Gene Coordinates:** *57,894-58,589 bp (forward)*

**1b. Is it a protein coding gene?** *Yes*

**1c. What is its function?** *Helix-turn-helix DNA-binding protein*

**1d. Final Summary: START:** Longest ORF, call supported by Glimmer and GeneMark. Gap of 60 bp (final score: -4.426; z-score: 2.498; spacer: 14). Most annotated start in Starterator (called 100% of time when present). (Strong) Coding potential in XX frame from XX-XX bp. 1:1 query:target alignment with X of top 10 related genes. Synteny conserved with two of three top-related phages. **FUNCTION:** Helix-turn-helix DNA-binding protien called by 3 of top 10 hits in BLAST (2 calls RNase E, 5 calls hypothetical protein), called by 19.5% (94 of 482) pham members (229816). Supported by HHPred and CDD.

**Gene 70**

**1a. Final Gene Coordinates:** *58,582-59,163 bp (forward)*

**1b. Is it a protein coding gene?** *Yes*

**1c. What is its function?** *Hypothetical protein*

**1d. Final Summary: START:** Not longest ORF, call supported by Glimmer and GeneMark. Overlap of 8 bp (final score: -6.182; z-score: 1.563; spacer: 8). Most annotated start in Starterator (called 77.9% of time when present). Coding potential in +1 frame. 1:1 query:target alignment with 7 of top 10 related genes. Synteny conserved. **FUNCTION:** Hypothetical protein called by 9 of top 10 hits in BLAST (one call gp71), called by 99.7% (293 of 294) pham members (229863). Not supported by HHPred.

**Gene 71**

**1a. Final Gene Coordinates:** *59,160-59,582 bp(forward)*

**1b. Is it a protein coding gene?** *Yes*

**1c. What is its function?** *Hypothetical protein*

**1d. Final Summary: START:** Longest ORF, call supported by Glimmer and GeneMark. Overlap of 4 bp (final score: -3.915; z-score: 2.498; spacer: 12). Most annotated start in Starterator (called 100% of time when present). Strong coding potential in +3 frame. 1:1 query:target alignment with 10 of top 10 related genes. Upstream synteny conserved with three most related phages, downstream synteny conserved with one most-related phage. **FUNCTION:** Hypothetical protein called by 10 of top 10 hits in BLAST, called by 100% (288 of 288) pham members (84781). Not supported by HHPred.

**Gene 72**

**1a. Final Gene Coordinates:** *59,623-59,751 bp (forward)*

**1b. Is it a protein coding gene?** *Yes*

**1c. What is its function?** *Hypothetical protein*

**1d. Final Summary: START:** Not longest ORF, call supported by Glimmer (no call by GeneMark). Gap of 40 bp (final score: -6.780; z-score: 1.265; spacer: 8). Most annotated start in Starterator (called 100% of time when present). Strong coding potential in +1 frame. 1:1 query:target alignment with 3 of top 10 related genes. Synteny conserved. **FUNCTION:** Hypothetical protein called by 9 of top 10 hits in BLAST (one call gp71), called by 100% 66 of 66) pham members (85754). Not supported by HHPred.

**Gene 73**

**1a. Final Gene Coordinates:** *59,786-60,127 bp (forward)*

**1b. Is it a protein coding gene?** *Yes*

**1c. What is its function?** *Hypothetical protein*

**1d. Final Summary: START:** Not longest ORF, call supported by Glimmer but not GeneMark. Gap of 34 bp (final score: -5.043; z-score: 1.898; spacer: 11). Most annotated start in Starterator (called 87.9% of time when present). Strong coding potential in +2 frame. 1:1 query:target alignment with 9 of top 10 related genes. Synteny conserved. **FUNCTION:** Hypothetical protein called by 10 of top 10 hits in BLAST, called by 100% (293 of 293) pham members (223). Supported by HHPred.

**Gene 74**

**1a. Final Gene Coordinates:** *60,129-60,731 bp (forward)*

**1b. Is it a protein coding gene?** *Yes*

**1c. What is its function?** *Hypothetical protein*

**1d. Final Summary: START:** Longest ORF, call supported by Glimmer and GeneMark. Gap of 1bp (final score: -6.893; z-score: 1.495; spacer: 16). Most annotated start in Starterator (called 100% of time when present). Strong coding potential in +3 frame. 1:1 query:target alignment with 10 of top 10 related genes. Synteny conserved. **FUNCTION:** Hypothetical protein called by 10 of top 10 hits in BLAST, called by 9.8% (48 of 491) pham members (224151). Not supported by HHPred.

**Gene 75**

**1a. Final Gene Coordinates:** *60,790-61,026 bp (forward)*

**1b. Is it a protein coding gene?** *Yes*

**1c. What is its function?** *Hypothetical protein*

**1d. Final Summary: START:** Not longest ORF, call supported by Glimmer and GeneMark. Gap of 58 bp (final score: -6.150; z-score: 1.579; spacer: 8). Most annotated start in Starterator (called 99.6% of time when present). Strong coding potential in +1 frame. 1:1 query:target alignment with 8 of top 10 related genes. Synteny conserved. **FUNCTION:** Hypothetical protein called by 9 of top 10 hits in BLAST (one call of gp74), called by 100% (278 of 178) pham members (225810). Not supported by HHPred.

**Gene 76**

**1a. FINAL Gene Co-ordinates:** *61,311-61,102 bp (reverse)*

**1b. Is it a protein-coding gene?** *Yes*

**1c. What is its function?***Hypothetical protein*

**1d. Final Summary: START**: Not longest ORF, call not supported by Glimmer or GeneMark. Gap of 81 bp (final score: -4.731; z-score: 2.053; spacer: 11). Most annotated start in Starterator (called 84.4% of time when present). Strong coding potential in -3 frame. 1:1 query:target alignment with 9 of top 10 related genes. Synteny conserved. **FUNCTION:** Hypothetical protein called by 10 of top 10 hits in BLAST, called by 100% (289 of 289) pham members (235). Not supported by HHPred.

**Gene 77**

**1a. Final Gene Coordinates:** *61,812-61,393 bp (reverse)*

**1b. Is it a protein coding gene?** *Yes*

**1c. What is its function?** *Hypothetical protein*

**1d. Final Summary: START:** Longest ORF, call not supported by Glimmer and GeneMark. Gap of 80 bp (final score: -4.845 ; z-score: 1.996; spacer: 11). Most annotated start in Starterator (called 85.8% of time when present). Strong coding potential in -3 frame. 1:1 query:target alignment with 10 of top 10 related genes. Synteny conserved. **FUNCTION:** Hypothetical protein called by 10 of top 10 hits in BLAST, called by 100% (291 of 291) pham members (226). Supported by HHPred.

**Gene 78**

**1a. Final Gene Coordinates:** *61,991-61,848 bp (reverse)*

**1b. Is it a protein coding gene?** *Yes*

**1c. What is its function?** *Hypothetical protein*

**1d. Final Summary: START:** Longest ORF, same call by GeneMark (no call by Glimmer). Gap of 54 bp (final score: -4.669; z-score: 2.053; spacer: 10). Most annotated start in Starterator (called 100% of time when present). Strong coding potential in -2 frame. 1:1 query:target alignment with 7 of top 10 related genes. Synteny conserved. **FUNCTION:** Hypothetical protein called by 9 of the 9 BLAST hits, called by 100% (353 of 353) pham members (84733). Supported by HHPred.

**Gene 79**

**1a. Final Gene Coordinates:** *62,309-62,046 bp (reverse)*

**1b. Is it a protein coding gene?** *Yes*

**1c. What is its function?** *Hypothetical protein*

**1d. Final Summary: START:** Not longest ORF, call supported by GeneMark (no call by Glimmer). Overlap of 4 bp (final score: -2.953; z-score: 2,907; spacer: 10). Most annotated start in Starterator (called 100% of time when present). Strong coding potential in -2 frame . 1:1 query:target alignment with 10 of top 10 related genes. Synteny conserved. **FUNCTION:** Hypothetical protein called by 10 of the top 10 hits in BLAST, called by 100% (289 of 289) pham members (84777). Not supported by HHPred.

**DELETED THIS AUTO-CALLED GENE**

**1a. Final Gene Coordinates:** *61,893-62,279 bp (forward)*

**1b. Is it a protein coding gene?** No

**1c. What is its function?** *N/A*

**1d. Final Summary:** No coding potential. No BLAST results. Deleting gene preserves synteny with three closest-relative phages.

**Gene 80**

**1a. Final Gene Coordinates:** *62,467-62,306 bp (reverse)*

**1b. Is it a protein coding gene?** *Yes*

**1c. What is its function?** *Hypothetical protein*

**1d. Final Summary: START:** Not longest ORF, call supported by Glimmer and GeneMark. Overlap of 1 bp (final score: -4.905; z-score: 1.975; spacer: 9). Most annotated start in Starterator (called 100% of time when present). Strong coding potential in -3 frame. 1:1 query:target alignment with 8 of top 10 related genes. Synteny conserved. **FUNCTION:** Hypothetical protein called by 9 of top 10 hits in BLAST (one call for recombination directionality factor), called by 100% (238 of 238) pham members (324). Supported by HHPred.

**Gene 81**

**1a. Final Gene Coordinates:** *62,697-62,467 bp (reverse)*

**1b. Is it a protein coding gene?** *Yes*

**1c. What is its function?** *Hypothetical protein*

**1d. Final Summary: START:** Longest ORF, call supported by Glimmer and GeneMark. Overlap of 8 bp (final score: -4.198; z-score: 2.55; spacer: 8). Most annotated start in Starterator (called 98.6% of time when present). Strong coding potential in -3 frame. 1:1 query:target alignment with 9 of top 10 related genes. Synteny conserved. **FUNCTION:** Hypothetical called by 10 of top 10 hits in BLAST, called by 100% (289 of 289) pham members (233). Not supported by HHPred.

**Gene 82**

**1a. Final Gene Coordinates:** *63,076-62,690 bp (reverse)*

**1b. Is it a protein coding gene?** *Yes*

**1c. What is its function?** *Hypothetical protein*

**1d. Final Summary: START:** Not longest ORF, call supported by Glimmer and GeneMark. Gap of 43 bp (final score: -4.846; z-score: 2.513; spacer: 16). Most annotated start in Starterator (called 85.5% of time when present). Strong coding potential in -1 frame. 1:1 query:target alignment with 4 of top 10 related genes. Synteny conserved. **FUNCTION:** Hypothetical protein called by 9 of top 10 hits in BLAST (one hit for gp81), called by 100% (285 of 285) pham members (227200). Supported by HHPred.

**Gene 83**

**1a. Final Gene Coordinates:** *63,422-63,120 bp (reverse)*

**1b. Is it a protein coding gene?** *Yes*

**1c. What is its function?** *HNH endonuclease*

**1d. Final Summary: START:** Not longest ORF, same call by GeneMark (no call by Glimmer). Overlap of 4 bp (final score: -7.384; z-score: 0.734; spacer: 11). Most annotated start in Starterator (called 87.2% of time when present). Strong coding potential in -2 frame. 1:1 query:target alignment with 5 of top 10 related genes. Synteny conserved. **FUNCTION:** HNH endonuclease called by 9 of top 10 hits in BLAST (one call putative HNH protein), called by 75.3% (311 of 413) pham members (228188). Supported by HHPred.

**Gene 84**

**1a. Final Gene Coordinates:** *63,529-63,419 bp (reverse)*

**1b. Is it a protein coding gene?** *Yes*

**1c. What is its function?** *Hypothetical protein*

**1d. Final Summary: START:** Longest ORF, call supported by GeneMark (no call by Glimmer). Overlap of 4 bp (final score: -4.696; z-score: 2.11; spacer: 12). Most annotated start in Starterator (called 100% of time when present). Strong coding potential in -1 frame. 1:1 query:target alignment with 10 of top 10 related genes. Synteny conserved. **FUNCTION:** Hypothetical protein called by 8 of the 8 BLAST hits, called by 100% (225 of 225) pham members (84865). Supported by HHPred.

**Gene 85**

**1a. Final Gene Coordinates:** *63,705-63,526 bp (reverse)*

**1b. Is it a protein coding gene?** *Yes*

**1c. What is its function?** *Hypothetical protein*

**1d. Final Summary: START:** Not longest ORF, call supported by Glimmer and GeneMark. Gap of 11 bp (final score: -3.059; z-score: 2.885; spacer: 11). Most annotated start in Starterator (called 100% of time when present). Strong coding potential in -3 frame. 1:1 query:target alignment with 9 of top 10 related genes. Synteny conserved. **FUNCTION:** Hypothetical protein called by 10 of top 10 hits in BLAST, called by 100% (290 of 290) pham members (192283). Supported by top hit on HHPred.

**Gene 86**

**1a. FINAL Gene Co-ordinates:** *63,977-63,717 bp (reverse)*

**1b. Is it a protein-coding gene?** *Yes*

**1c. What is its function?** *Hypothetical Protein*

**1d. Final Summary: START:** Not longest ORF, call not supported by Glimmer or GeneMark. Overlap of 4 bp (final score: -4.257; z-score: 2.258; spacer: 10). Most annotated start in Starterator (called 84.4% of time when present). Strong coding potential in -2 frame. 1:1 query:target alignment with 4 of top 10 related genes. Synteny conserved. **FUNCTION:** Hypothetical called by 10 of top 10 hits in BLAST, called by 99.6% (256 of 257) pham members (225864). Not supported by HHPred.

**Gene 87**

**1a. Final Gene Coordinates:** *64,171-63,974 bp (reverse)*

**1b. Is it a protein coding gene?** *Yes*

**1c. What is its function?** *Hypothetical protein*

**1d. Final Summary: START:** Not longest ORF, call supported by Glimmer and GeneMark. Overlap of 4 bp (final score: -5.605; z-score: 2.11; spacer: 6). Most annotated start in Starterator (called 100% of time when present). Coding potential in -1 frame. 1:1 query:target alignment with 4 of top 4 related genes (only 4 hits). Synteny conserved. **FUNCTION:** Hypothetical protein called by 4 of the 6 BLAST hits (one call NAD-dependent protein deacylase, one call histidine-tRNA ligase), called by 100% (255 of 255) pham members (302). Supported by HHPred.

**Gene 88**

**1a. Final Gene Coordinates:** *64,392-64,168 bp (reverse)*

**1b. Is it a protein coding gene?** *Yes*

**1c. What is its function?** *Hypothetical protein*

**1d. Final Summary: START:** Not longest ORF, same call by Glimmer and GeneMark. Overlap of 7 bp (final score: -5.352; z-score: 1.975; spacer: 8). Most annotated start in Starterator (called 62.8% of time when present). Strong coding potential in -3 frame. 1:1 query:target alignment with 8 of top 10 related genes. Synteny conserved. **FUNCTION:** Hypothetical protein called by 10 of top 10 hits in BLAST, called by 100% (286 of 286) pham members (84783). Not supported by HHPred.

**Gene 89**

**1a. Final Gene Coordinates:** *64,571-64,386 bp (reverse)*

**1b. Is it a protein coding gene?** *Yes*

**1c. What is its function?** *Hypothetical protein*

**1d. Final Summary: START:** Not longest ORF, supported by Glimmer and GeneMark. Overlap of 4 bp (final score: -5.526; z-score: 1.697; spacer: 12). Most annotated start in Starterator (called 99.7% of time when present). Strong coding potential in -2 frame. 1:1 query:target alignment with 10 of top 10 related genes. Synteny conserved. **FUNCTION:** Hypothetical protein called by 10 of top 10 hits in BLAST, called by 100% (290 of 290) pham members (84771). Not supported by HHPred.

**Gene 90**

**1a. Final Gene Coordinates:** *64,789-64,568 bp (reverse)*

**1b. Is it a protein coding gene?** *Yes*

**1c. What is its function?** *Hypothetical protein*

**1d. Final Summary: START:** Longest ORF, call supported by Glimmer and GeneMark. Gap of 120 bp (final score: -4.101; z-score: 2.367; spacer: 11). Most annotated start in Starterator (called 99.7% of time when present). Strong coding potential in -1 frame. 1:1 query:target alignment with 8 of top 10 related genes. Downstream synteny conserved with top three related phages. **FUNCTION:** Hypothetical protein called by 10 of top 10 hits in BLAST, called by 100% (297 of 297) pham members (196245). Not supported by HHPred.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Gene 91**

**1a. FINAL Gene Co-ordinates:** *64,889-64,803 bp (reverse)*

**1b. Is it a protein-coding gene?** *Yes*

**1c. What is its function?** *Hypothetical protein*

**1d. Final Summary: START:** Longest ORF. No Glimmer, GeneMark, or Starterator data available; however, this gene is present in top three related phages. Gap of 20 bp (final score: -5.888; z-score: 1.721; spacer: 8). Weak coding potential in -2 frame bp. 1:1 query:target alignment with 2 BLAST hits. Synteny conserved. **FUNCTION:** Hypothetical protein conserved in top three closest-related phages. Two BLAST hits, one call hypothetical protein and one call gp90. No PhagesDB or HHPred data available.

**Gene 92**

**1a. Final Gene Coordinates:** *65,212-64,910 bp (reverse)*

**1b. Is it a protein coding gene?** *Yes*

**1c. What is its function?** *Hypothetical protein*

**1d. Final Summary: START:** Longest ORF, call supported by Glimmer and GeneMark. Gap of 31 bp (final score: -2.804; z-score: 2.981; spacer: 10). Most annotated start in Starterator (called 99.1% of time when present). Strong coding potential in -1 frame. 1:1 query:target alignment with 10 of top 10 related genes. Downstream synteny conserved with two of three top-related phages. **FUNCTION:** Hypothetical protein called by 10 of top 10 hits in BLAST, called by 100% (353 of 353) pham members (84730). Not supported by HHPred.

**Gene 93**

**1a. Final Gene Coordinates:** *65,441-65,244 bp (reverse)*

**1b. Is it a protein coding gene?** *Yes*

**1c. What is its function?** *Hypothetical protein*

**1d. Final Summary: START:** Longest ORF, call supported by Glimmer and GeneMark. Gap of 67 bp (final score: -3.110; z-score: 3.091; spacer: 8). Most annotated start in Starterator (called 100% of time when present). Strong coding potential in -2 frame. 1:1 query:target alignment with 10 of top 10 related genes. Synteny conserved. **FUNCTION:** Hypothetical protein called by 9 of top 10 hits in BLAST (one call gp92), called by 100% (339 of 339) pham members (197781). Not supported by HHPred.

**Gene 94**

**1a. Final Gene Coordinates:** *65,784-65,509 bp (reverse)*

**1b. Is it a protein coding gene?** *Yes*

**1c. What is its function?** *Hypothetical protein*

**1d. Final Summary: START:** Longest ORF, call supported by Glimmer but not GeneMark. Overlap of -4 bp (final score: -5.606; z-score: 2.135; spacer: 16). Most annotated start in Starterator (called 95.4% of time when present). Strong coding potential in -3 frame. 1:1 query:target alignment with 10 of top 10 related genes. Synteny conserved with two of three top-related phages. **FUNCTION:** Hypothetical protein called by 10 of top 10 hits in BLAST, called by 100% (263 of 263) pham members (282). Not supported by HHPred.

**Gene 95**

**1a. Final Gene Coordinates:** *66,026-65,781 bp (reverse)*

**1b. Is it a protein coding gene?** *Yes*

**1c. What is its function?** *Hypothetical protein*

**1d. Final Summary: START:** Longest ORF, call supported by Glimmer and GeneMark. Gap of 38 bp (final score: -2.884; z-score: 2.981; spacer: 9). Most annotated start in Starterator (called 100% of time when present). Strong coding potential in -2 frame. 1:1 query:target alignment with 9 of top 10 related genes. Synteny conserved. **FUNCTION:** Hypothetical protein called by 10 of top 10 hits in BLAST, called by 100% 242 of 242) pham members (307). Supported by HHPred.

**Gene 96**

**1a. Final Gene Coordinates:** *66,421-66,065 bp (reverse)*

**1b. Is it a protein coding gene?** *Yes*

**1c. What is its function?** *Hypothetical protein*

**1d. Final Summary: START:** Not longest ORF, call supported by Glimmer and GeneMark. Gap of 162 bp (final score: -2.297; z-score: 3.303; spacer: 12). Most annotated start in Starterator (called 73.4% of time when present). Strong coding potential in -1 frame. 1:1 query:target alignment with 8 of top 10 related genes. Synteny conserved. **FUNCTION:** Hypothetical protein called by 10 of top 10 hits in BLAST, called by 100% (290 of 290) pham members (230). Not supported by HHPred.

**Gene 97**

**1a. Final Gene Coordinates:** *67,315-66,584 bp (reverse)*

**1b. Is it a protein coding gene?** *Yes*

**1c. What is its function?** *Hypothetical protein*

**1d. Final Summary: START:** Longest ORF, call supported by Glimmer and GeneMark. Overlap of 4 bp (final score: -4.905; z-score: 1.975; spacer: 9). Most annotated start in Starterator (called 100% of time when present). Strong coding potential in -1 frame. 1:1 query:target alignment with 10 of top 10 related genes. Synteny conserved. **FUNCTION:** Hypothetical protein called by 10 of top 10 hits in BLAST, called by 100% (290 of 290) pham members (229867). Not supported by HHPred.

**Gene 98**

**1a. Final Gene Coordinates:** *67,491-67,312 bp (reverse)*

**1b. Is it a protein coding gene?** *Yes*

**1c. What is its function?** *Hypothetical protein*

**1d. Final Summary: START:** Longest ORF, call supported by Glimmer and GeneMark. Overlap of 4 bp (final score: -5.972; z-score: 1.445; spacer: 9). Most annotated start in Starterator (called 100% of time when present). Strong coding potential in -3 frame. 1:1 query:target alignment with 8 of top 10 related genes. Synteny conserved. **FUNCTION:** Hypothetical protein called by 7 of top 10 hits in BLAST (3 calls of membrane protein), called by 99.7% (288 of 289) pham members (224206). Not supported by HHPred.

**Gene 99**

**1a. Final Gene Coordinates:** *67,724-67,488 bp (reverse)*

**1b. Is it a protein coding gene?***Yes*

**1c. What is its function?** *Hypothetical protein*

**1d. Final Summary: START:** Not longest ORF, call supported by Glimmer and GeneMark. Gap of 52 bp (final score: -3.222; z-score: 3.224; spacer: 15). Most annotated start in Starterator (called 85.9% of time when present). Strong coding potential in -2 frame. 1:1 query:target alignment with 4 of top 10 related genes. Synteny conserved. **FUNCTION:** Hypothetical protein called by 10 of top 10 hits in BLAST, called by 100% (325 of 325) pham members (195). Not supported by HHPred.

**Gene 100**

**1a. Final Gene Coordinates:** *68,178-67,777 bp (reverse)*

**1b. Is it a protein coding gene?** *Yes*

**1c. What is its function?** *Hypothetical protein*

**1d. Final Summary: START:** Not longest ORF, call supported by Glimmer and GeneMark. Gap of 2 bp (final score: -3.769; z-score: 2.571; spacer: 12). Most annotated start in Starterator (called 99.7% of time when present). Strong coding potential in -3 frame. 1:1 query:target alignment with 10 of top 10 related genes. Synteny conserved. **FUNCTION:** Hypothetical protein called by 10 of top 10 hits in BLAST, called by 100% (360 of 360) pham members (170). Not supported by HHPred.

**Gene 101**

**1a. Final Gene Coordinates:** 68,519-68*,181 bp (reverse)*

**1b. Is it a protein coding gene?** *Yes*

**1c. What is its function?** *Hypothetical protein*

**1d. Final Summary: START:** Longest ORF, call supported by Glimmer and GeneMark. Gap of 64 bp (final score: -3.632; z-score: 2.981; spacer: 7). Most annotated start in Starterator (called 90.0% of time when present). Strong coding potential in –2 frame. 1:1 query:target alignment with 10 of top 10 related genes. Upstream synteny conserved. **FUNCTION:** Hypothetical protein called by 10 of top 10 hits in BLAST, called by 100% (294 of 294) pham members (229864). Not supported by HHPred.

**Gene 102**

**1a. Final Gene Coordinates:***68,781-68,584 bp (reverse)*

**1b. Is it a protein coding gene?** *Yes*

**1c. What is its function?** *Hypothetical protein*

**1d. Final Summary: START:** Not longest ORF, call supported by Glimmer and GeneMark. Final score: -3.306; z-score: 2.731; spacer: 10. Most annotated start in Starterator (called 88.1% of time when present). Strong coding potential in -3 frame. 1:1 query:target alignment with 3 of top 10 related genes. Synteny conserved. **FUNCTION:** Hypothetical protein called by 10 of top 10 hits in BLAST, called by 100% (201 of 201) pham members (391). Not supported by HHPred.