**MANEATER – CLUSTER EE – ANNOTATION NOTES**

**GENE 1** **hypothetical protein**

Original Glimmer call @bp 41 has strength 11.64

SSC: 41 to 379 (Forward)

CP: yes

SCS: both

ST: SS, start 4 [41], found in 96 of 97 ( 99.0% ) of genes in pham, called 100.0% of time when present

BLAST-start: aligns with Microbacterium phage Vanisius 1, function unknown, query 1-112 subject 1-112, E value = 2e-59, Identity = 100%, Similarity = 100%

GAP: 0 [1st gene]

LO: NA

RBS: Kibler7, spacer 13, Z score 2.053, Final -5.036, yes

F: NKF

SIF-blast: aligns with hypothetical protein from Microbacterium phage Scamander, ID YP\_010050599.1, query 1-112 subject 1-112, E value = 1e-71, Identity = 100%, Similarity = 100%

SIF-HHPred: NKF, no matches with a probability above 90%

SF-Syn: NKF

**GENE 2 hypothetical protein**

Original Glimmer call @bp 376 has strength 8.76

SSC: 376 to 1059 (Forward)

CP: yes

SCS: both

ST: SS, start 2 [376], found in 94 of 94 ( 100.0% ) of genes in pham, called 95.7% of time when present

BLAST-start: aligns with Microbacterium phage TimoTea\_2, function unknown, query 1-227 subject 1-227, E value = e-130, Identity = 100%, Similarity = 100%

GAP: ATGA overlap

LO: NA

RBS: Kibler7, spacer 13, Z score 2.403, Final -4.334, no

F: NKF

SIF-blast: aligns with hypothetical protein SEA\_TIMOTEA\_2 from Microbacterium phage TimoTea, ID QBI97315.1, query 1-227 subject 1-227, E value = 4e-158, Identity = 100%, Similarity = 100%

SIF-HHPred: NKF, no matches with a probability above 90%

SF-Syn: NKF

**GENE 3 terminase**

Original Glimmer call @bp 1062 has strength 11.60

SSC: 1062 to 2522 (Forward)

CP: yes

SCS: both

ST: NI, start 94 [1062], found in 124 of 1153 ( 10.8% ) of genes in pham, called 76.6% of time when present

BLAST-start: aligns with Microbacterium phage Vanisius\_3, terminase, query 1-486 subject 1-486, E value = 0.0, Identity = 100%, Similarity = 100%

GAP: 3bp

LO: NA

RBS: Kibler7, spacer 15, Z score 1.856, Final -5.987, no

F: terminase

SIF-blast: aligns with terminase from Microbacterium phage Scamander, ID YP\_010050601.1, query 1-486 subject 1-486, E value = 0.0, Identity = 100%, Similarity = 100%

SIF-HHPred: Terminase large subunit; genome packaging, bacteriophage, ATPase, nuclease, HET: BR; 2.2A Enterobacteria phage HK97, Probability: 100%, E-value: 2.7e-42

SF-Syn: This gene is pham 66930 [terminase], upstream is pham 1094, downstream is pham 67043 [portal protein], as seen in Microbacterium phage Scamander

**GENE 4 portal protein**

Original Glimmer call @bp 2758 has strength 12.09; GeneMark calls start at 2716

SSC: 2758 to 3774 (Forward)

CP: yes

SCS: both - Glimmer

ST: NI, start 54 [2758], found in 96 of 224 ( 42.9% ) of genes in pham, called 11.5% of time when present Start 25 [2647], found in 97 of 224 ( 43.3% ) of genes in pham, called 85.6% of time when present

BLAST-start: aligns with Microbacterium phage Azizam\_4, portal protein, query 1-138 subject 1-138, E value = 0.0, Identity = 100%, Similarity = 100%

GAP: 235bp

LO: NO

RBS: Kibler7, spacer 11, Z score 1.777, Final -5.300, no

F: portal protein

SIF-blast: aligns with portal protein from Microbacterium phage Azizam, ID QGJ97464.2, query 1-138 subject 1-138, E value = 0.0, Identity = 100%, Similarity = 100%

SIF-HHPred: Phage portal protein, HK97 family; "neck", "portal", "capsid", "tail tube", VIRUS; 3.58A, Rhodobacter capsulatus, Probability = 100%, E-value = 5.3e-40

SF-Syn: This gene is pham 67043 [portal protein], upstream is pham 66930 [terminase], downstream is pham 1062 [major capsid and protease fusion protein], as seen in Microbacterium phage Scamander

**GENE 5 major capsid and protease fusion protein**

Original Glimmer call @bp 3771 has strength 13.77

SSC: 3771 to 5357 (Forward)

CP: yes

SCS: both

ST: SS, start 2 [3771], found in 107 of 109 ( 98.2% ) of genes in pham, called 100.0% of time when present

BLAST-start: aligns with Microbacterium phage TimoTea\_5, major capsid and protease fusion protein, query 1-528 subject 1-528, E value = 0.0, Identity = 100%, Similarity = 100%

GAP: ATGA overlap

LO: NA

RBS: Kibler7, spacer 15, Z score 3.093, Final -3.509, yes

F: major capsid and protease fusion protein

SIF-blast: aligns with major capsid and protease fusion protein from Microbacterium phage Loca, ID URM86189.1, query 1-528 subject 1-529, E value = 0.0, Identity = 98%, Similarity = 99%

SIF-HHPred: 3JB5\_F major capsid protein, bacteriophage, HK97-like, 3.7A, Propionibacterium phage PA6, Probability = 99.9%, E-value = 4.1e-22,

SF-Syn: This gene is pham 1062 [major capsid and protease fusion protein], upstream is pham 70193 [portal protein], downstream is pham 1065 [head-to-tail adaptor], as seen in Microbacterium phage Luxx

**GENE 6 head-to-tail adaptor**

Original Glimmer call @bp 5361 has strength 17.42

SSC: 5361 to 5714 (Forward)

CP: yes

SCS: both

ST: SS, start 2 [5361], found in 97 of 97 ( 100.0% ) of genes in pham, called 100.0% of time when present

BLAST-start: aligns with Microbacterium phage TimoTea\_6, head-to-tail adaptor, query 1-117 subject 1-117, E value = 3e-60, Identity = 100%, Similarity = 100%

GAP: 3bp

LO: NA

RBS: Kibler7, spacer 16, Z score 2.668, Final -4.553, yes

F: head-to-tail adaptor

SIF-blast: aligns with head-tail connector protein from Microbacterium phage Bri160, ID YP\_009996641.1, query 1-117 subject 1-117, E value = 8e-78, Identity = 99%, Similarity = 99%

SIF-HHPred: 5A21\_C 15 PROTEIN, TAILED BACTERIOPHAGE, SIPHOVIRIDAE, SPP1, VIRAL ASSEMBLY, HEAD-TO-TAIL INTERFACE, 7.2A, BACILLUS PHAGE SPP1, Probability = 99.24%, E-value = 1.8e-10

SF-Syn: This gene is pham 1065 [head-to-tail adaptor], upstream is pham 1062 [major capsid and protease fusion protein], downstream is pham 1057 [tail terminator], as seen in Microbacterium phage Scamander

**GENE 7 tail terminator**

Original Glimmer call @bp 5711 has strength 19.90

SSC: 5711 to 6091 (Forward)

CP: yes

SCS: both

ST: SS, start 2 [5711], found in 97 of 97 ( 100.0% ) of genes in pham, called 100.0% of time when present

BLAST-start: aligns with Microbacterium phage TimoTea\_7, tail terminator, query 1-126 subject 1-126, E value = 2e-65, Identity = 100%, Similarity = 100%

GAP: ATGA overlap

LO: NA

RBS: Kibler7, spacer 7, Z score 1.515, Final -6.591, no

F: tail terminator

SIF-blast: aligns with tail terminator from Microbacterium phage TimoTea, ID QBI97320.1, query 1-126 subject 1-126, E value = 6e-84, Identity = 100%, Similarity = 100%

SIF-HHPred: 6TE9\_F Tail terminator protein Rcc01690, 3.58A, Rhodobacter capsulatus

Probability: 97.46%, E-value: 0.014,

SF-Syn: This gene is pham 1057 [tail terminator], upstream is pham 1065 [head-to-tail adaptor], downstream is pham 618 [major tail protein], as seen in Microbacterium phage Scamander

**GENE 8 major tail protein**

Original Glimmer call @bp 6130 has strength 16.50

SSC: 6130 to 6564 (Forward)

CP: yes

SCS: both

ST: NI, start 5 [6130], found in 97 of 160 ( 60.6% ) of genes in pham, called 100.0% of time when present

BLAST-start: aligns with Microbacterium phage Yubaba\_8, major tail protein, query subject , E value = 7e-82, Identity = 100%, Similarity = 100%

GAP: 39bp

LO: yes

RBS: Kibler7, spacer 11, Z score 1.529, Final -5.797, yes

F: major tail protein

SIF-blast: aligns with major tail protein from Microbacterium phage PaoPu, ID YP\_009996593.1, query 1-144 subject 1-144, E value = 6e-100, Identity = 100%, Similarity = 100%

SIF-HHPred: 6TE9\_G Phage major tail protein, TP901-1 family, 3.58A, Rhodobacter capsulatus, Probability = 99.28%, E value = 1.5e-10,

SF-Syn: This gene is pham 618 [major tail protein], upstream is pham 1057 [tail terminator], downstream is pham 58402 [hypothetical protein], as seen in Microbacterium phage Luxx

**GENE 9 hypothetical protein**

Original Glimmer call @bp 6577 has strength 14.55

SSC: 6577 to 6960 (Forward)

CP: yes

SCS: both

ST: NI, start 3 [6577], found in 97 of 146 ( 66.4% ) of genes in pham, called 100.0% of time when present

BLAST-start: aligns with Microbacterium phage YertPhresh\_9, function unknown, query 1-127 subject 1-127, E value = 2e-66, Identity = 100%, Similarity = 100%

GAP: 13bp

LO: yes

RBS: Kibler7, spacer 10, Z score 2.248, Final -4.294, yes

F: NKF

SIF-blast: aligns with hypothetical protein JTF51\_gp09 from Microbacterium phage PaoPu, ID YP\_009996594.1, query 1-127 subject 1-127, E value = 3e-85, Identity = 100%, Similarity = 100%

SIF-HHPred: NKF, no matches with a probability above 90%

SF-Syn: NKF

Phages DB best match was a minor tail protein but this was not corroborated by NCBI blast and HHPred

**GENE 10 tail assembly chaperone**

Original Glimmer call @bp 6974 has strength 17.49

SSC: 6974 to 7294 (Forward)

CP: yes

SCS: both

ST: SS, start 6 [6974], found in 109 of 109 ( 100.0% ) of genes in pham, called 100.0% of time when present

BLAST-start: aligns with Microbacterium phage Vanisius\_10, tail assembly chaperone, query 1-106 subject 1-106, E value = 1e-53, Identity = 100%, Similarity = 100%

GAP: 14bp

LO: yes

RBS: Kibler7, spacer 10, Z score 1.482, Final -3.826, yes

F: tail assembly chaperone

SIF-blast: aligns with tail assembly chaperone from Microbacterium phage Scamander, ID YP\_010050609.1, query 1-106 subject 1-106, E value = 7e-68, Identity = 100%, Similarity = 100%

SIF-HHPred: NKF, no matches with a probability above 90%

SF-Syn: This gene is pham 1058 [tail assembly chaperone], upstream is pham 58402, downstream is pham 1223 [tail assembly chaperone – genes 10 and 11 are part of a translational frameshift], as seen in Microbacterium phage Luxx

**GENE 11 tail assembly chaperone**

Original GeneMark call @bp 7321

SSC: 6577 to 6960 (Forward) – to be changed due to being part of a frameshift – now becomes 6974 to 7410

CP: no, this is part of the translational frameshift from gene 10, the tail assembly chaperone

SCS: GeneMark-cs

ST: NA, start is same as previous gene;

BLAST-start: start is same as previous gene;

GAP: overlap with previous gene – translational frameshift;

LO: yes

RBS: Kibler7, spacer , Z score , Final -, yes

F: tail assembly chaperone

SIF-blast: matches the tail assembly chaperone from the Microbacterium phage Luxx

SIF-HHPred: NKF, no matches with a probability above 90%

SF-Syn: gene 10 + gene 11 translational frameshift to give larger gene 11 synteny seen in other EE phage;

GENE 11 PART 1 = 6974 – 7258 GAGGGA

PART 2 = 7258 – 7410 TAG

A [7258] OF GGA [END OF PART 1] USED AS START [AAA - K] OF PART 2

FRAME 2 TO FRAME 1 THEREFORE IT IS A -1 FRAMESHIFT

**GENE 12 tape measure protein**

Original Glimmer call @bp 7528 has strength 14.88

SSC: 7528 to 9633 (Forward)

CP: yes

SCS: both

ST: NI, start 11[7528], found in 109 of 254 ( 42.9% ) of genes in pham, called 99.1% of time when present

BLAST-start: aligns with Microbacterium phage TimoTea\_12, tape measure protein, query 1-701 subject 1-701, E value = 0.0, Identity = 100%, Similarity = 100%

GAP: 118bp

LO: yes

RBS: Kibler7, spacer 16, Z score 2.752, Final -4.386, yes

F: tape measure protein

SIF-blast: aligns with tail length tape measure protein from Microbacterium phage Scamander, ID YP\_010050610.1, query 12-712 subject 1-701, E value = 0.0, Identity = 99%, Similarity = 100%

SIF-HHPred: 6V8I\_CF Tape Measure Protein, gp57; phage tail, tail tip, tape measure protein, Probability: 99.93%, E-value: 2.3e-17

SF-Syn: This gene is pham 125517 [tape measure protein], upstream is the tail assembly chaperone, downstream is pham 76903 [minor tail protein], as seen in Microbacterium phage Luxx

SIF-Mem: SIF-Mem: 10 domains found by TMHMM, 8 domains found by SOSUI

**GENE 13 minor tail protein**

Original Glimmer call @bp 9630 has strength 10.34

SSC: 9630 to 10655 (Forward)

CP: yes

SCS: both

ST: SS, Start 1[9630], Found in 93 of 93 ( 100.0% ) of genes in pham, Called 100.0% of time when present

BLAST-start: aligns with Microbacterium phage TimoTea\_13, minor tail protein, query 1-341 subject 1-341, E value = 0.0, Identity = 100%, Similarity = 100%

GAP: GTGA overlap

LO: NA

RBS: Kibler7, spacer 12, Z score 2.259, Final -4.412, yes

F: minor tail protein

SIF-blast: aligns with minor tail protein from Microbacterium phage Jannah, ID QYW07172.1, query 1-341 subject 1-341, E value = 0.0, Identity = 99%, Similarity = 100%

SIF-HHPred: 6ZPV\_AAA MgGH51; arabinofuranosidse, glycosidase, Probability: 98.65%, E-value: 0.0000013,

SF-Syn: This gene is pham 76903 [minor tail protein], upstream is pham 125517 [tape measure protein], downstream is pham 1068 [minor tail protein], as seen in Microbacterium phage Luxx

**GENE 14 minor tail protein**

Original Glimmer call @bp 10655 has strength 10.61

SSC: 10655 to 12724 (Forward)

CP: yes

SCS: both

ST: SS, Start 1[10655], Found in 113 of 113 ( 100.0% ) of genes in pham, Called 100.0% of time when present

BLAST-start: aligns with Microbacterium phage TeddyBoy\_14, minor tail protein, query 1-689 subject 1-689, E value = 0.0, Identity = 99%, Similarity = 100%

GAP: TAATG overlap

LO: NA

RBS: Kibler7, spacer 15, Z score 2.300, Final -5.098, no

F: minor tail protein

SIF-blast: aligns with minor tail protein from Microbacterium phage TimoTea, ID QBI97327.1, query 12-700 subject 1-689, E value = 0.0, Identity 100%= , Similarity = 100%

SIF-HHPred: 6TPW\_A, Receptor-type tyrosine-protein phosphatase F; Fibronectin type-III, adhesion protein, Probability: 99.7%, E-value: 7.3e-13

SF-Syn: This gene is pham 1068 [minor tail protein], upstream is pham 76903 [minor tail protein], downstream is pham 85094 [minor tail protein], as seen in Microbacterium phage Luxx

**GENE 15 minor tail protein**

Original Glimmer call @bp 12726 has strength 15.91

SSC: 12726 to 13274 (Forward)

CP: yes

SCS: both

ST: SS, Start 3[12726], Found in 110 of 137 ( 80.3% ) of genes in pham, called 99.1% of time when present

BLAST-start: aligns with Microbacterium phage , Yubaba\_15, minor tail protein, query 1-182 subject 1-182, E value = e-101, Identity = 100%, Similarity = 100%

GAP: 1bp

LO: NA

RBS: Kibler7, spacer 12, Z score 3.226, Final -2.394, yes

F: minor tail protein

SIF-blast: aligns with minor tail protein from Microbacterium phage EdElric , ID UUG69681.1, query 1-182 subject 1-182, E value = 1e-126, Identity = 99%, Similarity = 99%

SIF-HHPred: NKF, no matches with a probability above 90%

SF-Syn: This gene is pham 85094 [minor tail protein], upstream is pham 1068 [minor tail protein], downstream is pham 1097, as seen in Microbacterium phage Luxx

**GENE 16 hypothetical protein**

Original Glimmer call @bp 13311 has strength 14.29; GeneMark calls start at 13287

SSC: 13311 to 13598 (Forward)

CP: yes

SCS: both – Glimmer only

ST: SS, Start 5[13311], Found in 108 of 109 ( 99.1% ) of genes in pham, Called 100.0% of time when present

BLAST-start: aligns with Microbacterium phage Vanisius\_16, function unknown, query 1-95 subject 1-95, E value = 9e-46, Identity = 100%, Similarity = 100%

GAP: 37bp

LO: NA

RBS: Kibler7, spacer 9, Z score 2.640, Final -3.589, yes

F: NKF

SIF-blast: aligns with hypothetical protein from Microbacterium BURTONTHEPUP\_16, ID AVR56193.2, query 1-95 subject 1-95, E value = 3e-55, Identity = 100%, Similarity = 100%

SIF-HHPred: NKF, no matches with a probability above 90%

SF-Syn: NKF

**GENE 17 endolysin**

Original Glimmer call @bp 13617 has strength 8.94

SSC: 13617 to 14312 (Forward)

CP: yes

SCS: both

ST: SS, Start 4[13617], Found in 111 of 113 ( 98.2% ) of genes in pham, Called 98.2% of time when present

BLAST-start: aligns with Microbacterium phage TimoTea\_17, endolysin, query 1-231 subject 1-231, E value = e-136, Identity = 100%, Similarity = 100%

GAP: 19bp

LO: yes

RBS: Kibler7, spacer 12, Z score 2.476, Final -3.978, yes

F: endolysin

SIF-blast: aligns with endolysin from Microbacterium phage VitulaEligans, ID AXC34900.1, query 1-231 subject 1-231, E value = 4e-166, Identity = 100%, Similarity = 100%

SIF-HHPred: 6UE4\_A, ShyA endopeptidase; Endopeptidase M23, Probability: 99.38%, E-value: 4.2e-11

SF-Syn: This gene is pham 122577 [endolysin], upstream is pham 1057, downstream is pham 1063, as seen in Microbacterium phage Luxx

**GENE 18 membrane protein**

Original Glimmer call @bp 14279 has strength 9.95

SSC: 14279 to 14542 (Forward)

CP: yes

SCS: both

ST: SS, Start 7 [14279], Found in 111 of 114 ( 97.4% ) of genes in pham, Called 85.6% of time when present

BLAST-start: aligns with Microbacterium phage TimoTea\_18, function unknown, query 1-87 subject 1-87, E value = 2e-47, Identity = 100%, Similarity = 100%

GAP: 33bp overlap – 14309 gives a better RBS score but the option of this position does not appear anywhere else

LO: yes

RBS: Kibler7, spacer 11, Z score 1.397, Final -6.062, no [14309]

F: membrane protein

SIF-blast: aligns with membrane protein from Microbacterium phage BurtonThePup, ID AVR56195.2, query 1-87 subject 1-87, E value = 1e-56, Identity = 99%, Similarity = 100%

SIF-HHPred: 8HHF\_B, Cell division protein FtsB; membrane protein complex, heterotrimer, MEMBRANE PROTEIN, Probability: 95.84%, E-value: 0.27

SF-Syn: This gene is pham 1063, upstream is pham 122577 [endolysin], downstream is pham 1053, as seen in Microbacterium phage Luxx

SIF-Mem: 1 domain found by DEEPTMHMM, 1 domain found by SOSUI

**GENE 19 holin**

Original Glimmer call @bp 14539 has strength 11.67

SSC: 14539 to 14763 (Forward)

CP: yes

SCS: both

ST: SS, Start 5 [14539], Found in 112 of 114 ( 98.2% ) of genes in pham, Called 100.0% of time when present

BLAST-start: aligns with Microbacterium phage Yubaba\_19, function unknown, query 1-74 subject 1-74, E value = 4e-36, Identity = 100%, Similarity = 100%

GAP: ATGA overlap

LO: NA

RBS: Kibler7, spacer 11, Z score 2.840, Final -3.171, yes

F: NKF

SIF-blast: aligns with holin from Microbacterium phage PaoPu, ID YP\_009996604.1, query 1-74 subject 1-74, E value = 2e-43, Identity = 100%, Similarity = 100% [top matches are holins or membrane proteins]

SIF-HHPred: NKF, no matches with a probability above 90% [several at 50% and above are membrane proteins]

SF-Syn: holin/membrane protein

SIF-Mem: 1 domain found by DEEPTMHMM, 1 domain found by SOSUI

**GENE 20 Lsr2-like DNA bridging protein**

Original Glimmer call @bp 15044 has strength 11.61

SSC: 15044 to 14832 (Reverse)

CP: yes

SCS: both

ST: SS, Start 12 [15044], Found in 111 of 146 ( 76.0% ) of genes in pham, Called 100.0% of time when present

BLAST-start: aligns with Microbacterium phage TimoTea\_20, Lsr2-like DNA bridging protein, query 1-70 subject 1-70, E value = 6e-34, Identity = 100%, Similarity = 100%

GAP: 3bp

LO: NA

RBS: Kibler7, spacer 12, Z score 2.651, Final -3.627, yes

F: Lsr2-like DNA bridging protein

SIF-blast: aligns with Lsr2-like DNA bridging protein from Microbacterium phage TimoTea, ID QBI97333.1, query 1-70 subject 1-70, E value = 1e-41, Identity = 100%, Similarity = 100%

SIF-HHPred: 4E1P\_B, Protein lsr2; anti-parallel beta sheet, dimer, DNA BINDING PROTEIN, Probability: 99.86%, E-value: 6.7e-21,

SF-Syn: This gene is pham 85067 [Lsr2-like DNA bridging protein], upstream is pham 1053, downstream is pham 1059 [helix-turn-helix DNA binding domain protein], as seen in Microbacterium phage Luxx

**GENE 21 helix-turn-helix DNA binding domain protein**

Original Glimmer call @bp 15484 has strength 11.58; GeneMark calls start at **15547**

SSC: 15484 to 15047 (Reverse) – change start to 15547

CP: Genemark shows coding potential starting at 15547

SCS: both - GM

ST: NI, Start 12[15484], Found in 110 of 113 ( 97.3% ) of genes in pham, Called 14.5% of time when present. Start 6 [15547], Found in 104 of 113 ( 92.0% ) of genes in pham, Called 75.0% of time when present

BLAST-start: aligns with Microbacterium phage TimoTea\_21, helix-turn-helix DNA binding domain protein, query 1-145 subject 22-166, E value = 2e-76, Identity = 100%, Similarity = 100%

GAP: 144bp [using 15484] – 81bp [using 15547]

LO: no, if using 15484

RBS: Kibler7, spacer 13, Z score 2.992, Final -3.155, yes [if using 15484]

F: helix-turn-helix DNA binding domain protein

SIF-blast: aligns with helix-turn-helix DNA binding domain protein from Microbacterium phage Luxx, ID QKY78668.2, query 1-145 subject 22-166, E value = 6e-97, Identity = 99%, Similarity = 99%

SIF-HHPred: 4P0Z\_A AM32, HELIX-TURN-HELIX, DSDNA BINDING, DNA BINDING PROTEIN; HET: MSE; 1.35A {Enterococcus faecalis}, Probability: 99.13%, E-value: 9.2e-10

SF-Syn: This gene is pham 1059 [helix-turn-helix DNA binding domain protein], upstream is 85067 [Lsr2-like DNA bridging protein], downstream is pham 85194 [helix-turn-helix DNA binding domain protein], as seen in Microbacterium phage Luxx

The best RBS score is undoubtedly for position 15484 but all the other evidence [Genemark, alignment, Starterator] suggests 15547 as the start site.

**GENE 22 helix-turn-helix DNA binding domain protein**

Original Glimmer call @bp 15858 has strength 8.92

SSC: 15858 to 15628 (Reverse)

CP: yes

SCS: both

ST: SS, Start 5 [15858], Found in 113 of 113 ( 100.0% ) of genes in pham, Called 100.0% of time when present

BLAST-start: aligns with Microbacterium phage YertPhresh\_22, helix-turn-helix DNA binding domain protein, query 1-76 subject 1-76, E value = 1e-36, Identity = 100%, Similarity = 100%

GAP: 715bp

LO: no

RBS: Kibler7, spacer 15, Z score 0.932, Final -6.237, no [15888 has a far superior RBS score but no other evidence for this as start]

F: helix-turn-helix DNA binding domain protein

SIF-blast: aligns with helix-turn-helix DNA binding domain protein from Microbacterium phage Gardevoir, ID UDL14934.1, query 1-76 subject 1-76, E value = 9e-45, Identity = 99%, Similarity = 98%

SIF-HHPred: 2EF8\_B, Putative transcription factor; helix-turn-helix, DNA binding protein, TRANSCRIPTION REGULATOR, 1.95A {Enterobacteria phage P2}, Probability: 97.9%, E-value: 0.00011

SF-Syn: This gene is pham 85194 [helix-turn-helix DNA binding domain protein], upstream is pham 1059 [helix-turn-helix DNA binding domain protein], downstream is pham 3206, as seen in Microbacterium phage Luxx

**GENE 23 hypothetical protein**

Original GeneMark call @bp 16573

SSC: 16573 to 16755 (Forward)

CP: yes

SCS: both

ST: SS, Start 3 [16573], Found in 26 of 26 ( 100.0% ) of genes in pham, Called 96.2% of time when present

BLAST-start: aligns with Microbacterium phage Minima\_23, function unknown, query 1-60 subject 1-60, E value = 3e-28, Identity = 100%, Similarity = 100%

GAP: 715 bp

LO: yes

RBS: Kibler7, spacer 7, Z score 1.954, Final -5.712, yes

F: NKF

SIF-blast: aligns with hypothetical protein SEA\_MINIMA\_23 from Microbacterium phage Minima, ID AXQ64371.3, query 1-60 subject 1-60, E value = 2e-33, Identity = 100%, Similarity = 100%

SIF-HHPred: NKF, no matches with a probability above 90% [many proteins 70+ and 80+ % showing as helicases]

SF-Syn: NKF

**GENE 24 helix-turn-helix DNA binding domain protein**

Original Glimmer call @bp 16839 has strength 8.79; GeneMark calls start at 16848

SSC: 16839 to 17057 (Forward)

CP: yes

SCS: both - GL

ST: SS, Start 12 [16839], Found in 104 of 112 ( 92.9% ) of genes in pham, Called 98.1% of time when present

BLAST-start: aligns with Microbacterium phage TimoTea\_24, helix-turn-helix DNA binding domain protein, query 1-72 subject 1-72, E value = 4e-36 , Identity = 100%, Similarity = 100%

GAP: 84bp

LO: yes

RBS: Kibler7, spacer 9, Z score 1.960, Final -4.952, yes

F: helix-turn-helix DNA binding domain protein

SIF-blast: aligns with helix-turn-helix DNA binding domain protein from Microbacterium phage MrGreen, ID UVK59396.1, query 1-72 subject 1-72, E value = 9e-43, Identity = 97%, Similarity = 98%

SIF-HHPred: 4LHF\_A, Regulatory protein cox; helix-turn-helix, 2.401A {Enterobacteria phage P2}, Probability: 99%, E-value: 1.4e-8

SF-Syn: This gene is pham 117549 [helix-turn-helix DNA binding domain protein], upstream is pham 3206, downstream is pham 125593 [HNH endonuclease], as seen in Microbacterium phage Luxx

**GENE 25 HNH endonuclease**

Original Glimmer call @bp 17054 has strength 8.15

SSC: 17054 to 17353 (Forward)

CP: yes

SCS: both

ST: NI, Start 22 [17054], found in 116 of 120 ( 96.7% ) of genes in pham, Called 100.0% of time when present

BLAST-start: aligns with Microbacterium phage Yubaba\_25, HNH endonuclease, query 1-99 subject 1-99, E value = 3e-56, Identity = 100%, Similarity = 100%

GAP: ATGA overlap

LO: yes

RBS: Kibler7, spacer 9, Z score 1.412, Final -6.051, no but not significantly better

F: HNH endonuclease

SIF-blast: aligns with HNH endonuclease from Microbacterium phage PaoPu, ID YP\_009996610.1, query 1-99 subject 1-99, E value = 8e-68, Identity = 100%, Similarity = 100%

SIF-HHPred: 5H0M\_A, HNH endonuclease; Thermophilic bacteriophage, 1.52A {Geobacillus virus E2}, Probability: 97.92%, E-value: 0.000029

SF-Syn: This gene is pham 125593 [HNH endonuclease], upstream is pham 117549 [helix-turn-helix DNA binding domain protein], as seen in Microbacterium phage Luxx