**STEM – EE CLUSTER – ANNOTATION NOTES**

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

Original Glimmer call @bp 41 has strength 10.17

SSC: 41 to 379 (FWD)

CP: yes

SCS: both

ST: SS, Start 3 [41] Found in 128 of 128 ( 100.0% ) of genes in pham Called 100.0% of time when present

BLAST-start: aligns with Microbacterium phage Bullzi2019 , function unknown, query 1-112 subject 1-112 , E value = 4e-59 , Identity = 100% , Similarity = 100%

GAP: N/A

LO: NA

RBS: Kibler7, spacer 13 , Z score 2.064 , Final – 5.082, yes

F: NKF

SIF-blast: aligns with hypothetical protein from Microbacterium phage Bullzi2019 , ID QKY80439.1 , query 1-112 subject 1-112 , E value = 3e-73, Identity = 100%, Similarity = 100%

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

SF-Syn: NKF

**GENE 2 [CHANGE START-376] hypothetical protein**

Original Glimmer call @bp 361 has strength 13.18

SSC: 376 to 1110 (FWD) Change to start 2 at 376 [starterator, RBS, BLAST alignment favors nt376]

CP: yes

SCS: both

ST: NA, Start 1 [361] Found in 98 of 125 ( 78.4% ) of genes in pham,Called 6.1% of time when present - Change Start to Start 2 [376], Found in 125 of 125 ( 100.0% ) of genes in pham, Called 95.2% of time when present

BLAST-start: aligns with Microbacterium phage Bullzi2019 , function unknown, query 6-249 subject 1-238 , E value = e-129, Identity = 91% , Similarity = 94%

GAP: ATGA Overlap

LO: NA

RBS: Kibler7, spacer 13 , Z score 2.546 , Final – 4.125, NO

F: NKF

SIF-blast: aligns with hypothetical protein from Microbacterium phage Bullzi2019 , ID QKY80440.1 , query 6-244 subject 1-238 , E value = 4e-144, Identity = 92% , Similarity = 94%

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

SF-Syn: NKF

**GENE 3 [change start-1113] Terminase**

Original Glimmer call @bp 1119 has strength 12.80; GeneMark calls start at 1113

SSC: 1113 to 2573 (FWD)

CP: yes

SCS: both – GM [starterator, RBS, alignment favors 1113]

ST: NI, Found in 195 of 1272 ( 15.3% ) of genes in pham Called 36.9% of time when present

BLAST-start: aligns with Microbacterium phage Trireme , function Terminase, query 1-484 subject 3-486 , E value = 0.0, Identity = 100% , Similarity = 100%

GAP: 3 bp

LO: NA

RBS: Kibler7, spacer 17 , Z score 2.295, Final -4.854, yes

F: Terminase

SIF-blast: aligns with Terminase from Microbacterium phage Nobel , ID YP\_009996689.1 , query 1-484 subject 3-486 , E value = 0.0 , Identity = 100% , Similarity = 100%

SIF-HHPred: Terminase large subunit; genome packaging, bacteriophage, ATPase, nuclease, VIRAL PROTEIN; HET: BR; 2.2A {Enterobacteria Probability :100% E-Value: 3.2 e-42

SF-Syn: This gene is Pham 130372 [Terminase] upstream is Pham 85210 [NKF] and downstream is Pham 139605 [Portal Protein] as seen in Microbacterium Phage Noelani

**GENE 4 Portal Protein**

Original Glimmer call @bp 2698 has strength 13.27; GeneMark calls start at 2803

SSC: 2698 to 3825 (FWD)

CP: yes

SCS: both, GL [starterator, RBS, alignment, etc favors 2698]

ST: SS, Start 9 [2698] Found in 130 of 185 ( 70.3% ) of genes in pham Called 82.3% of time when present

BLAST-start: aligns with Microbacterium phage Trireme , function Portal Protein, query 1-375 subject 1-375, E value = 0.0 , Identity = 100%, Similarity = 100%

GAP: 125 bp

LO: Yes

RBS: Kibler7, spacer 15 , Z score 1.968, Final – 5.356, yes

F: Portal Protein

SIF-blast: aligns with Portal Protein from Microbacterium phage Nobel , ID YP\_009996690.1 , query 1-375 subject 1-375 , E value = 0.0 , Identity = 100% , Similarity = 100%

SIF-HHPred: Portal protein; Prohead I, icosahedral symmetry, HK97, phage, capsid, VIRUS; 3.6A {Escherichia phage HK97} Probability:100 E-Value – 5.2 E-38

SF-Syn: This gene is Pham 143864 [Portal Protein] upstream is Pham 1062 [Major Capsid and Protease Fusion Protein] downstream is Pham 140856 [Terminase] as seen in Microbacterium Phage Bullzi2019

**GENE 5 Major Capsid and Protease Fusion Protein**

Original Glimmer call @bp 3822 has strength 15.75

SSC: 3822 to 5408 (FWD)

CP: yes

SCS: both

ST: SS, Start 2 [3822] Found in 129 of 129 ( 100.0% ) of genes in pham Called 100.0% of time when present

BLAST-start: aligns with Microbacterium phage Bullzi2019 , function 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 3.139 , Final -3.034, yes

F: Major Capsid and Protease Fusion Protein

SIF-blast: aligns with hypothetical protein from Microbacterium phage Noelani , ID YP\_009996716.1 , query 1-528 subject 1-528 , E value = 0.0 , Identity = 99% , Similarity = 99%

SIF-HHPred: Scaffolding domain delta; Prohead I, icosahedral symmetry, HK97, phage, capsid, VIRUS; 3.5A {Escherichia phage HK97} Probability – 99.92 E-Value – 1.5 E-23

SF-Syn: This gene is Pham 1062 [Major Capsid and Protease Fusion Protein] upstream is Pham 143864 [Portal Protein] and downstream is Pham 1065 [Head to Tail Adaptor] as seen in Microbacterium Phage Trireme

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

Original Glimmer call @bp 5412 has strength 15.83

SSC: 5412 to 5762 (FWD)

CP: yes

SCS: both

ST: SS, Start 2 [5412] Found in 129 of 129 ( 100.0% ) of genes in pham Called 100.0% of time when present

BLAST-start: aligns with Microbacterium phage Publix, function Head to Tail Adaptor, query 1-116 subject 1-116 , E value = 7E-60 , Identity = 100% , Similarity = 100%

GAP: 4 bp

LO: NA

RBS: Kibler7, spacer 16 , Z score 2.687 , Final -3.997, yes

F: Head to Tail Adaptor

SIF-blast: aligns with Head to Tail Connector Protein from Microbacterium phage Noelani , ID YP\_009996717.1 , query 1-116 subject 1-116 , E value = 7e-78 , Identity = 100%, Similarity = 100%

SIF-HHPred: Gp6; 13-membered ring, VIRAL PROTEIN; HET: MPD, MSE; 2.1A {Enterobacteria phage HK97} Probability – 99.52 E-Value – 1.8 e-13

SF-Syn: This gene is Pham 1065 [Head to Tail Adaptor] upstream is Pham 1062 [Major Capsid and Protease Fusion Protein] and downstream is Pham 1057 [Tail Terminator] as seen in Microbacterium Phage Trireme

**GENE 7 Tail Terminator**

Original Glimmer call @bp 5759 has strength 18.21

SSC: 5759 to 6139 (FWD)

CP: yes

SCS: both

ST: SS, Start 2 [5759] Found in 129 of 129 ( 100.0% ) of genes in pham Called 99.2% of time when present

BLAST-start: aligns with Microbacterium phage Upsilon , function Tail Terminator, query 1-126 subject 1-126 , E value = 2e-65 , Identity = 100% , Similarity = 100%

GAP: ATGA Overlap

LO: NA

RBS: Kibler7, spacer 9 , Z score 1.196 , Final -6.778, yes

F: Tail Terminator

SIF-blast: aligns with Minor Tail Protein from Microbacterium phage Nobel, ID YP\_009996693.1 , query 1-126 subject 1-126 , E value = 4e-84 , Identity = 100% , Similarity = 100%

SIF-HHPred: Minor tail protein U; Mixed Alpha-Beta fold, VIRAL PROTEIN; HET: SO4, MSE; 2.7A {Enterobacteria phage lambda} SCOP: d.32 Probability – 98.65 E-Value – 7.4 e-6

SF-Syn: This gene is Pham 1057 [Tail Terminator] upstream is Pham 1065 [Head to Tail Adaptor] and downstream is Pham 139535 [Major Tail Protein] as seen in Microbacterium Phage Publix

**GENE 8 Major Tail Protein**

Original Glimmer call @bp 6178 has strength 19.89

SSC: 6178 to 6615 (FWD)

CP: yes

SCS: both

ST: SS, Start 5 [6178] Found in 130 of 194 ( 67.0% ) of genes in pham Called 100.0% of time when present

BLAST-start: aligns with Microbacterium phage Trireme , function Major Tail Protein, query 1-145 subject 1-145 , E value = 2e-82 , Identity = 100% , Similarity =100%

GAP: 39 bp

LO: NA

RBS: Kibler7, spacer 10 , Z score 2.656 , Final -3.797, yes

F: Major Tail Protein

SIF-blast: aligns with Major Tail Protein from Microbacterium phage Nobel , ID YP\_009996694.1 , query 1-145 subject 1-145 , E value = 1e-100 , Identity = 100%, Similarity = 100%

SIF-HHPred: Major tail protein; Marine bacteriophage, Cryo-EM, Siphophage, Baseplate, Megatron protein, Tail fibre protein, Distal t Probability: 99.37 E-Value – 1.4e11

SF-Syn: This gene is Pham 139535 [Major Tail Protein] upstream is Pham 1057 [Tail Terminator] and downstream is Pham 143854 [Tail Assembly Chaperone] as seen in Microbacterium Phage TinyTruffula

**GENE 9 hypothetical protein**

Original Glimmer call @bp 6628 has strength 12.57

SSC: 6628 to 7011 (FWD)

CP: yes

SCS: both

ST: SS, Start 7 [6628] Found in 136 of 190 ( 71.6% ) of genes in pham Called 100.0% of time when present

BLAST-start: aligns with Microbacterium phage Yubaba , function Minor Tail Protein, query 1-127 subject 1-127 , E value = 2e-66 , Identity = 100%, Similarity = 100%

GAP: 13 bp

LO: NA

RBS: Kibler7, spacer 10 , Z score 2.298 , Final -4.506, yes

F: NKF – no extra evidence, HHPred, etc., to assign a function

SIF-blast: aligns with hypothetical protein from Microbacterium PaoPu , ID YP\_009996594.1 , query 1-127 subject 1-127 , E value =5e-85 , Identity = 100%, Similarity =100%

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

SF-Syn: This gene is Pham 143854 [Tail Assembly Chaperone] upstream is Pham 139535 [Major Tail Protein] and downstream is Pham 85197 [Tail Assembly Chaperone] as seen in Microbacterium Phage TinyTruffula

**GENE 10 Tail Assembly Chaperone**

Original Glimmer call @bp 7025 has strength 19.79

SSC: 7025 to 7345 (FWD)

CP: yes

SCS: both

ST: SS, Start 6 [7025] Found in 130 of 130 ( 100.0% ) of genes in pham Called 100.0% of time when present

BLAST-start: aligns with Microbacterium phage Upsilon , function Tail Assembly Chaperone, query 1-106 subject 1-106 , E value = 3e-54 , Identity = 100%, Similarity = 100%

GAP: 14 bp

LO: NA

RBS: Kibler7, spacer 11 , Z score 2.830 , Final -3.493, yes

F: Tail Assembly Chaperone

SIF-blast: aligns with Tail Assembly Chaperone from Microbacterium phage Nobel , ID YP\_009996697.1 , query 1-106 subject 1-106, E value = 2e-68 , Identity = 100% , Similarity = 100%

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

SF-Syn: This gene is Pham 85197 [Tail Assembly Chaperone] upstream is Pham 143854 [Minor Tail Protein] and downstream is Pham 143915 [Tape Measure Protein] as seen in Microbacterium Phage TinyTruffula

This is the first of the tail assembly chaperone proteins the next gene [11] will use gene 10 and a frameshift to form the larger tail assembly chaperone protein

**GENE 11 Tail Assembly Chaperone**

To be changed due to presence of -1 frameshift

>TomBrady12 gp10 MDTTTKPKKFDLEQLTLGEVAAIEDLSGVAIGGVTTETPQGKFLAALYMVAKRRNGEPTFTFNAALQASMGEAQSFLGLDASDDEADEAESSAEGNGDGSPESAPA

>TomBrady12 gp11

MDTTTKPKKFDLEQLTLGEVAAIEDLSGVAIGGVTTETPQGKFLAALYMVAKRRNGEPTFTFNAALQASMGEAQSFLGLDASDDEADEAESSAEGKRRRLARERARLKAQFIVQLGMDPAAYERLTIAERDAIVRELNKRSRRRR

A close up of letters

Description automatically generated

>Stem\_Draft gp10

MDTTTKPKKFDLEQLTLGEVAAIEDLSGVAIGGVTTETPQGKFLAALYMVAKRRNGEPTFTFNAALQASMGEAQSFLGLDASDDEADEAESSAEGNGDGSPESAPA

Part 1 [7025 to 7309] GAGGGA

Part 2 [7309 to 7461] AAACGG

Frameshifts from frame 2 to frame 1 (-1 frameshift) Gene 10 amino acid G from GGA ends on A which then frameshifts and starts the AAA sequence for amino acid K

**GENE 12 Tape Measure Protein**

Original Glimmer call @bp 7579 has strength 15.68

SSC: 7579 to 9684 (FWD)

CP: yes

SCS: both

ST: SS, Start 6 [7579] Found in 130 of 139 ( 93.5% ) of genes in pham Called 97.7% of time when present

BLAST-start: aligns with Microbacterium phage Trireme , function Tape Measure Protein, query 1-701 subject 1-701 , E value = 0.0 , Identity = 100%, Similarity = 100%

GAP: 234 bp

LO: Yes

RBS: Kibler7, spacer 16 , Z score 2.769 , Final -3.835, yes

F: Tape Measure Protein

SIF-blast: aligns with Tape Measure Protein from Microbacterium phage Miaurora , ID AYD84507.1 , query 1-701 subject 1-701 , E value = 0.0 , Identity = 100%, Similarity = 100%

SIF-HHPred: Tape Measure Protein, gp57; phage tail, tail tip, tape measure protein, VIRAL PROTEIN; 3.7A {Staphylococcus virus 80alph Probability: 99.94 E-Value- 4.2e-18

SF-Syn: This gene is Pham 143915 [Tape Measure Protein] upstream is Pham 4838 [Minor Tail Protein] and downstream is Pham 85197 [Tail Assembly Chaperone] as seen in Microbacterium Phage TinyTruffula

**GENE 13 Minor Tail Protein**

Original Glimmer call @bp 9681 has strength 15.68

SSC: 9681 to 10187 (FWD)

CP: yes

SCS: both

ST: SS, Start 1 [9681] Found in 22 of 22 ( 100.0% ) of genes in pham Called 100.0% of time when present

BLAST-start: aligns with Microbacterium phage Trireme , function Minor Tail Protein, query 1-168 subject 1-168 , E value =6e-91 , Identity = 100% , Similarity = 100%

GAP: GTGA Overlap

LO: NA

RBS: Kibler7, spacer 18 , Z score 2.535 , Final -4.476, yes

F: Minor Tail Protein

SIF-blast: aligns with Minor Tail Protein from Microbacterium phage Nobel , ID YP\_009996699.1, query 1-168 subject 1-168 , E value = 3e-114, Identity = 100% , Similarity = 100%

SIF-HHPred: Baseplate hub gp41; Pseudomonas, phage, baseplate, VIRUS;{Pseudomonas phage vB\_PaeM\_E217} Probability: 97.87 E-Value- 4.1e-4

SF-Syn: This gene is Pham 4838 [Minor Tail Protein] upstream is Pham 143915 [Tape Measure Protein] and downstream is Pham 1068 [Minor Tail Protein] as seen in Microbacterium Phage PuffyCat

**GENE 14 Minor Tail Protein**

Original Glimmer call @bp 10187 has strength 16.40

SSC: 10187 to 12244 (FWD)

CP: yes

SCS: both

ST: SS, Start 1 [10187] Found in 130 of 130 ( 100.0% ) of genes in pham Called 100.0% of time when present

BLAST-start: aligns with Microbacterium phage TinyTruffula , function Minor Tail Protein, query 1-685 subject 1-685 , E value = 0.0 , Identity = 100% , Similarity = 100%

GAP: A Overlap

LO: NA

RBS: Kibler7, spacer 15 , Z score 1.798 , Final -5.694, yes

F: Minor Tail Protein

SIF-blast: aligns with Minor Tail Protein from Microbacterium phage TinyTruffula, ID QKY79168.1 , query 1-685 subject 1-685 , E value = 0.0 , Identity = 100% , Similarity = 100%

SIF-HHPred: Receptor-type tyrosine-protein phosphatase F; Fibronectin type-III, adhesion protein, CELL ADHESION; HET: SO4; 2.9A {Hom Probability : 99.73 E-Value-7.4e-14

SF-Syn: This gene is Pham 1068 [Minor Tail Protein] upstream is Pham 4838 [Minor Tail Protein] and downstream is Pham 85094 [Minor Tail Protein] as seen in Microbacterium Phage PuffyCat

**GENE 15 Minor Tail Protein**

Original Glimmer call @bp 12246 has strength 18.94

SSC: 12246 to 12797 (FWD)

CP: yes

SCS: both

ST: SS, Start 3 [12246] Found in 127 of 159 ( 79.9% ) of genes in pham Called 98.4% of time when present

BLAST-start: aligns with Microbacterium phage Niffler , function Minor Tail Protein, query 1-183 subject 1-183 , E value =e-103 , Identity = 100% , Similarity = 100%

GAP: 2 bp

LO: NA

RBS: Kibler7, spacer 12 , Z score 2.497 , Final -4.199, yes

F: Minor Tail Protein

SIF-blast: aligns with Minor Tail Protein from Microbacterium phage HarperAnne , ID QGJ91604.1 , query 1-183 subject 1-183 , E value = 2e-131, Identity = 100% , Similarity = 100%

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] and downstream is Pham 137548 [NKF] as seen in Microbacterium Phage PuffyCat

**GENE 16 Hypothetical Protein**

Original Glimmer call @bp 12834 has strength 15.28; GeneMark calls start at 12810

SSC: 12834 to 13133

CP: yes

SCS: both; GL – starterator, RBS and alignment suggest 12834 as the start

ST: SS, Start 5 [12834] Found in 54 of 55 ( 98.2% ) of genes in pham Called 100.0% of time when present

BLAST-start: aligns with Microbacterium phage Publix, function unknown, query 1-99 subject 1-99 , E value = 2e-48 , Identity = 100%, Similarity = 100%

GAP: 37 bp

LO: NA

RBS: Kibler7, spacer 9, Z score 2.721, Final -3.755, yes

F: NKF

SIF-blast: aligns with hypothetical protein from Microbacterium phage HarperAnne , ID QGJ91605.1 , query 1-99 subject 1-99, E value = 1e-57, Identity = 100%, Similarity = 100%

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

SF-Syn: NKF

**GENE 17 Endolysin**

Original Glimmer call @bp 13159 has strength 14.77; GeneMark calls start at 13153

SSC: 13153 to 13848 (FWD)

CP: yes

SCS: both; GM -starterator, RBS and alignment suggest 13153 as the start

ST: NA Start 7 [13153]: Found in 127 of 131 ( 96.9% ) of genes in pham Called 96.1% of time when present

BLAST-start: aligns with Microbacterium phage Publix, function Endolysin, query 1-229 subject 3-231 , E value = e-136, Identity = 100% , Similarity = 100%

GAP: 20 bp

LO: NA

RBS: Kibler7, spacer 11 , Z score 2.491 , Final -4.165, yes

F: Endolysin

SIF-blast: aligns with Endolysin from Microbacterium phage HarperAnne , ID QGJ91606.1 , query 1-229 subject 3-231, E value = 7e-164 , Identity = 100% , Similarity = 100%

SIF-HHPred: Peptidase M23; membrane protein, enzyme, TRANSPORT PROTEIN; HET: ADP;{Vibrio cholerae} Probability : 99.49 E-Value – 2e-12

SF-Syn: This gene is Pham 141078 [Endolysin] upstream is Pham 137948 [NKF] and downstream is Pham 1063 [NKF] as seen in Microbacterium Phage Upsilon

**GENE 18 membrane protein**

Original Glimmer call @bp 13827 has strength 10.89; GeneMark calls start at 13815

SSC: 13827 to 14078 (FWD)

CP: yes

SCS: both; GL – RBS and alignment suggest 13827, starterator suggests 13815

ST: NA, Called at Start 8 at 13827 - Found in 72 of 131 ( 55.0% ) of genes in pham Called 2.8% of time when present

BLAST-start: aligns with Microbacterium phage Namsahir , function unknown, query 1-83 subject 5-87 , E value = 4e-44 , Identity = 98% , Similarity = 100%

GAP: 33 bp Overlap

LO: NA

RBS: Kibler7, spacer 16 , Z score 2.299 , Final – 4.767, yes

F: membrane protein

SIF-blast: aligns with membrane protein from Microbacterium phage HarperAnne, ID QGJ91607.2, query 1-83 subject 5-87 , E value = 2e-53 , Identity =99% , Similarity = 100%

SIF-HHPred: 8BH1\_E, Cell division protein FtsB; bacterial cell division, peptidoglycan synthesis, membrane protein complex, MEMBRANE PROTEIN, Probability = 95.43%, E value = 0.33

SF-Syn: NKF

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

**GENE 19** **Hypothetical Protein**

Original Glimmer call @bp 14075 has strength 10.99

SSC: 14075 to 14299 (FWD)

CP: yes

SCS: both

ST: SS, Start 5 [14075] Found in 129 of 131 ( 98.5% ) of genes in pham Called 100.0% of time when present

BLAST-start: aligns with Microbacterium phage Upsilon , function unknown, query 1-74 subject 1-74, E value = 1e-35 , Identity = 98%, Similarity = 98%

GAP: ATGA Overlap

LO: NA

RBS: Kibler7, spacer 11 , Z score 2.894 , Final -3.366, yes

F: NKF – one BLAST suggests holin, it aligns perfectly with the holin from other phage e.g. Kaijohn – membrane protein regions found but no matches in HHPred with > 90%

SIF-blast: aligns with Holin from Microbacterium phage Upsilon , ID YP\_009996679.1 , query 1-74 subject 1-74 , E value = 4e-42 , Identity = 99% , Similarity = 98%

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

SF-Syn: NKF

SIF-Mem: 2 domains found by TMHMM, 1 domain found by SOSUI

**GENE 20 Lsr2-like DNA Bridging Protein**

Original Glimmer call @bp 14577 has strength 12.63

SSC: 14577 to 14368 (REV)

CP: yes

SCS: both

ST: SS, Start 12 [14577] Found in 128 of 163 ( 78.5% ) of genes in pham Called 100.0% of time when present

BLAST-start: aligns with Microbacterium phage Trireme, function Lsr2-like DNA Bridging Protein, query 1-69 subject 1-69 , E value = 2e-32, Identity = 98% , Similarity = 98%

GAP: 2 bp

LO: NA

RBS: Kibler7, spacer 12 , Z score 2.834 , Final -3.530, yes

F: Lsr2-like DNA Bridging Protein

SIF-blast: aligns with Lsr2-like DNA Bridging Protein from Microbacterium phage Miaurora , ID AYD84515.1 , query 1-69 subject 1-69 , E value = 6e-40 , Identity = 99% , Similarity = 98%

SIF-HHPred: Protein lsr2; anti-parallel beta sheet, dimer, DNA BINDING PROTEIN; 1.728A {Mycobacterium tuberculosis} Probability : 99.86 E-Value- 7.5e-21

SF-Syn: This gene is Pham 85067 [Lsr2-like DNA Bridging Protein] upstream is Pham 1053 [NKF] and downstream is Pham 1059 [Helix-Turn-Helix DNA Binding Protein] as seen in Microbacterium Phage Trireme

**GENE 21 Helix-Turn-Helix DNA Binding Domain Protein**

Original Glimmer call @bp 15170 has strength 13.86; GeneMark calls start at 15086

SSC: 15170 to 14580 (REV)

CP: yes

SCS: both; GL – RBS favors 15086 but alignment favors 15170

ST: N/A – Called at Start 3 15170 - Found in 47 of 130 ( 36.2% ) of genes in pham Called 10.6% of time when present; Using Start 6 at 15086 Found in 120 of 130 ( 92.3% ) of genes in pham Called 66.7% of time when present

BLAST-start: aligns with Microbacterium phage Clownery , function Helix-Turn-Helix DNA Binding Protein, query 1-196 subject 1-196 , E value = e-107 , Identity = 100%, Similarity = 100%

GAP: GTGA Overlap

LO: Yes

RBS: Kibler7, spacer 15 , Z score 1.499 , Final – 6.287, no

F: Helix-Turn-Helix DNA Binding Protein

SIF-blast: aligns with DNA Binding Protein from Microbacterium Nobel , ID YP\_009996707.1, query 1-196 subject 1-196 , E value = 3e-138, Identity = 100%, Similarity = 100%

SIF-HHPred: AM32; Repressor, Protein-DNA complex, Type IV secretion system, DNA BINDING PROTEIN; HET: EPE; 1.93A {Enterococcus faeca Probability: 99.47 E-Value- 6.4e-13

SF-Syn: This gene is Pham 1059 [Helix-Turn-Helix DNA Binding Protein] upstream is Pham 85067 [Lsr2-like DNA Binding Protein] and downstream is Pham 85194 [Helix-Turn-Helix DNA Binding Domain Protein] as seen in Microbacterium Phage Trireme

**GENE 22 Helix-Turn-Helix DNA Binding Domain Protein**

Original Glimmer call @bp 15397 has strength 4.79

SSC: 15397 to 15167 (REV)

CP: yes

SCS: both

ST: SS, Start 5 [15397] Found in 130 of 130 ( 100.0% ) of genes in pham Called 95.4% of time when present

BLAST-start: aligns with Microbacterium phage Trireme , function Helix-Turn-Helix DNA Binding Domain Protein, query 1-76 subject 1-76 , E value = 6e-37, Identity = 100% , Similarity = 100%

GAP: N/A No matches

LO: NA

RBS: Kibler7, spacer 8 , Z score 0.956, Final -7.333, yes

F: Helix-Turn-Helix DNA Binding Domain Protein

SIF-blast: aligns with Transcriptional Repressor from Microbacterium Kaijohn, ID YP\_009996683.1 , query 1-76 subject 1-76 , E value = 3e-45 , Identity = 100%, Similarity = 100%

SIF-HHPred: ComR; Streptococcus, Competence, Quorum sensing, ComR, TRANSCRIPTION REGULATOR; 2.9A {Streptococcus suis (strain 05ZYH33 Probability: 99.1 E-Value - 6e-9

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] and downstream is Pham 1383 [NKF] as seen in Microbacterium Phage Upsilon

**GENE 23 Hypothetical Protein**

Original Glimmer call @bp 15911 has strength 6.53; GeneMark calls start at 16133

SSC: 15911 to 16294 (FWD)

CP: yes

SCS: both; GL – RBS, starterator and alignment all favor 15911

ST: SS, Start 2 [15911] Found in 93 of 95 ( 97.9% ) of genes in pham Called 100.0% of time when present

BLAST-start: aligns with Microbacterium phage Upsilon , function unknown, query 1-127 subject 1-127, E value = 4e-65 , Identity = 100% , Similarity = 100%

GAP: 111 bp – No coding potential observed

LO: Yes

RBS: Kibler7, spacer 8 , Z score 1.962 , Final – 5.339, yes

F: NKF

SIF-blast: aligns with hypothetical protein from Microbacterium phage PoRanda , ID YP\_009996759.1 , query 1-127 subject 1-127 , E value = 7e-83 , Identity = 100% , Similarity = 100%

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

SF-Syn: NKF

**GENE 24 Helix-Turn-Helix DNA Binding Domain Protein**

Original Glimmer call @bp 16393 has strength 10.84

SSC: 16393 to 16611 (FWD)

CP: yes

SCS: both

ST: SS, Start 12 [16393] Found in 108 of 129 ( 83.7% ) of genes in pham Called 98.1% of time when present

BLAST-start: aligns with Microbacterium phage Upsilon , function Helix-Turn-Helix DNA Binding Domain Protein, query 1-72 subject 1-72 , E value = 2e-35 , Identity = 100% , Similarity = 100%

GAP: 99 bp

LO: Yes

RBS: Kibler7, spacer 16 , Z score 1.587 , Final -6.179, yes

F: Helix-Turn-Helix DNA Binding Domain Protein

SIF-blast: aligns with MerR-like-Helix-Turn-Helix DNA Binding Domain Protein from Microbacterium phage Nobel, ID YP\_009996710.1, query 1-72 subject 1-72 , E value = 2e-42 , Identity = 100% , Similarity = 100%

SIF-HHPred: Recombination Directionality Factor RdfS; Excisionase, Recombination Directionality Factor, winged helix-turn-helix, sup Probability : 99.43 E-Value- 4.3e-13

SF-Syn: This gene is Pham 117549 [Helix-Turn-Helix DNA Binding Domain Protein] upstream is Pham 1382 [NKF] and downstream is Pham 143877 [HNH Endonuclease] as seen in Microbacterium Phage Upsilon

**GENE 25 HNH - Endonuclease**

Original Glimmer call @bp 16608 has strength 11.01

SSC: 16608 to 16907 (FWD)

CP: yes

SCS: both

ST: SS, Start 33 [16608] Found in 130 of 173 ( 75.1% ) of genes in Called 100.0% of time when present

BLAST-start: aligns with Microbacterium phage Trireme , function HNH Endonuclease, query 1-99 subject 1-99 , E value = 3e-56 , Identity = 100% , Similarity = 100%

GAP: ATGA Overlap

LO: NA

RBS: Kibler7, spacer 5 , Z score 2.009 , Final – 5.519, yes

F: HNH Endonuclease

SIF-blast: aligns with HNH Endonuclease from Microbacterium phage Nobel , ID YP\_009996711.1, query 1-99 subject 1-99 , E value = 7e-68 , Identity = 100% , Similarity = 100%

SIF-HHPred: CRISPR-associated endonuclease Cas9; Inhibitor, Complex, VIRAL PROTEIN; HET: NI; 2.097A {Staphylococcus aureus} Probability: 97.81 E-Value- 2.4e-5

SF-Syn: This gene is Pham 143877 [HNH Endonuclease] upstream is Pham 117549 [Helix-Turn-Helix DNA Binding Domain Protein as seen in Microbacterium Phage Upsilon