Phage Biscayne:

Annotation Documentation

Head Instructor

Jaime Mayoral

Adjunct Lecturer

Alexis Mas

Students

[Appolon,Jazlyn](mailto:jappo008@fiu.edu)

[Artiles,Katherine](mailto:karti009@fiu.edu)

[Becerra Reymundo,Lara .](mailto:lbece014@fiu.edu)

[Briceno,Nicole Angelica](mailto:nbric008@fiu.edu)

[Burunat,Julianna Laura](mailto:jburu002@fiu.edu)

[Camacho,Alejandra Maria](mailto:acama102@fiu.edu)

[Lad,Pooja](mailto:plad002@fiu.edu)

[Phan,Nhan Minh](mailto:nphan009@fiu.edu)

[Purroy,Xavier Solans](mailto:xpurr001@fiu.edu)

[Ramirez Ramirez,Erika](mailto:erami056@fiu.edu)

[Rodriguez Leiva,Alison](mailto:arodr1625@fiu.edu)

[Serradet,Victoria Marlen](mailto:vserr022@fiu.edu)

|  |  | | |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |

Start Annotation Documentation

| Gene # 1 | Original 5’ End Start  41 | Original 3’ End Stop  379 | | Original Length  339 | Original Start Codon  ATG / GTG / TTG | |
| --- | --- | --- | --- | --- | --- | --- |
| Gene Direction  FWD / REV | Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?  Yes / No (if no, answer below) | | | | | | |
|  | Revised 5’ End Start |  | New Length | | | New Start Codon  ATG / GTG / TTG |

| Start called by both Glimmer and GeneMark? |
| --- |

Glimmer and GeneMark Calls

| Does this Gene have Coding Potential |
| --- |

Ribosome Binding Site

| RBS Final Score  -5.036 | Is this Score the Closest to 0?  Yes / No (if no) | Score Closest to 0: |
| --- | --- | --- |
| Start #: |
| ORF Length: |

| Starterator Start  41 (with 100 MA’s, all MA are only in this position) |
| --- |
|
|

Starterator

| Function **(F)**  Hypothetical Protein | Gene No:1 | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Hypothetical protein  Values Shown in many of the results | % Query  99% | % Identity  100% | | E-value  (if < 10-7)  5e-73 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from BLAST(Protein) **(BLAST-phagesDB)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Function Unknown  Values Shown in many of the results |  | % Identity  100% | | E-value  (if < 10-7)  3e-59 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from HHpred **(HHpred)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  N/A | Name / Description  N/A | Hit  N/A | | Probability (if > 90%)  All probabilities below 90 | | E-value  (if < 1)  N/A |
| Supporting Information for Function, from Syntney **(SIF-Syn)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF) | Pham number and number of members  Members (130) of Pham 1064 | | upstream pham conserved ?  Yes or No  If yes, what pham # or function ?  First Gene of the phage | | Downstream pham conserved ?  Yes or No  If yes, what pham # or function ?  85210 | |
| Transmembrane domains by TmHmm using Phamerator | | | | N/A | | | |

| Gene # 2 | Original 5’ End Start  376 | Original 3’ End Stop  1098 | | Original Length  723 | Original Start Codon  ATG / GTG / TTG | |
| --- | --- | --- | --- | --- | --- | --- |
| Gene Direction  FWD / REV | Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?  Yes / No (if no, answer below) | | | | | | |
|  | Revised 5’ End Start |  | New Length | | | New Start Codon  ATG / GTG / TTG |

| Start called by both Glimmer and GeneMark? |
| --- |

Glimmer and GeneMark Calls

| Does this Gene have Coding Potential |
| --- |

Ribosome Binding Site

| RBS Final Score  -4.334 | Is this Score the Closest to 0?  Yes / No (if no) | Score Closest to 0: |
| --- | --- | --- |
| Start #: |
| ORF Length: |

Starterator

| Starterator Start  376 (with 97 MA’s, all MA are only in this position) |
| --- |
|
|

| Function **(F)**  Hypothetical Protein | Gene No:2 | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Hypothetical Protein  Values Shown in many of the results | % Query  99% | % Identity  95.44% | | E-value  (if < 10-7)  1e-160 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from BLAST(Protein) **(BLAST-phagesDB)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Function Unknown  Values Shown in many of the results |  | % Identity  95% | | E-value  (if < 10-7)  e-137 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from HHpred **(HHpred)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  N/A | Name / Description  N/A | Hit  N/A | | Probability (if > 90%)  All probabilities below 90 | | E-value  (if < 1)  N/A |
| Supporting Information for Function, from Syntney **(SIF-Syn)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF) | Pham number and number of members  Members (127) of Pham 85210 | | upstream pham conserved ?  Yes or No  If yes, what pham # or function ?  1064 | | Downstream pham conserved ?  Yes or No  If yes, what pham # or function ?  149912 | |
| Transmembrane domains by TmHmm using Phamerator | | | | N/A | | | |

| Gene # 3 | Original 5’ End Start  1101 | Original 3’ End Stop  2561 | | Original Length  1461 | Original Start Codon  ATG / GTG / TTG | |
| --- | --- | --- | --- | --- | --- | --- |
| Gene Direction  FWD / REV | Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?  Yes / No (if no, answer below) | | | | | | |
|  | Revised 5’ End Start |  | New Length | | | New Start Codon  ATG / GTG / TTG |

| Start called by both Glimmer and GeneMark? |
| --- |

Glimmer and GeneMark Calls

| Does this Gene have Coding Potential |
| --- |

Ribosome Binding Site

| RBS Final Score  -5.987 | Is this Score the Closest to 0?  Yes / No (if no) | Score Closest to 0: -4.679 |
| --- | --- | --- |
| Start #: 1947 |
| ORF Length:615 ( this position would cuts off too much of the gene) |

Starterator

| Starterator Start  1101 (with 100 MA’s, all MA are only in this position) |
| --- |
|
|

| Function **(F)**  terminase | Gene No:3 | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Terminase  Function was shown in many results | % Query  99% | % Identity  96.50% | | E-value  (if < 10-7)  **0.0** | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from BLAST(Protein) **(BLAST-phagesDB)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Terminase  Function was shown in many results |  | % Identity  97% | | E-value  (if < 10-7)  0.0 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from HHpred **(HHpred)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Terminase large subunit | Name / Description  Terminase large subunit; genome packaging, bacteriophage, ATPase, nuclease, VIRAL PROTEIN; HET: BR; 2.2A {Enterobacteria phage HK97} | Hit  6Z6D\_A | | Probability (if > 90%)  100 | | E-value  (if < 1)  1.70006e-41 |
| Supporting Information for Function, from Syntney **(SIF-Syn)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  terminase  Function shown in many results | Pham number and number of members  Members (1278) of Pham 149912 | | upstream pham conserved ?  Yes or No  If yes, what pham # or function ?  85210 | | Downstream pham conserved ?  Yes or No  If yes, what pham # or function ?  151755 | |
| Transmembrane domains by TmHmm using Phamerator | | | | N/A | | | |



| Gene # 4 | Original 5’ End Start  2687 | Original 3’ End Stop  3844 | | Original Length  1158 | Original Start Codon  ATG / GTG / TTG | |
| --- | --- | --- | --- | --- | --- | --- |
| Gene Direction  FWD / REV | Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?  Yes / No (if no, answer below) | | | | | | |
|  | Revised 5’ End Start |  | New Length | | | New Start Codon  ATG / GTG / TTG |

| Start called by both Glimmer and GeneMark? |
| --- |

Glimmer and GeneMark Calls

| Does this Gene have Coding Potential |
| --- |



Ribosome Binding Site

| RBS Final Score  -6.283 | Is this Score the Closest to 0?  Yes / No (if no) | Score Closest to 0: -4.486 |
| --- | --- | --- |
| Start #: 2978 |
| ORF Length:867 |

Starterator

| Starterator Start  2687 with 88 MA’s |
| --- |
|
|



An

Anotation D

| Function **(F)**  Portal Protein | Gene No:4 | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Portal Protein  Function was shown in many results | % Query  95% | % Identity  96.46% | | E-value  (if < 10-7)  0.0 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from BLAST(Protein) **(BLAST-phagesDB)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Portal Protein  Function was shown in many results |  | % Identity  92% | | E-value  (if < 10-7)  0.0 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from HHpred **(HHpred)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Portal Protein | Name / Descr (incl DUF)  Portal protein; Prohead I, icosahedral symmetry, HK97, phage, capsid, VIRUS; 3.6A {Escherichia phage HK97} | Hit  8FQL\_L | | Probability (if > 90%)  100 | | E-value  (if < 1)  5e-38 |
| Supporting Information for Function, from Syntney **(SIF-Syn)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Portal Protein  Function shown in many results | Pham number and number of members  Members (175) of Pham 150121 | | upstream pham conserved ?  Yes or No  If yes, what pham # or function ?  149912/terminase | | Downstream pham conserved ?  Yes or No  If yes, what pham # or function ?  1062/ Major capsid and protease fusion protein | |
| Transmembrane domains by TmHmm using Phamerator | | | | N/A | | | |

| Gene # 5 | Original 5’ End Start  3841 | Original 3’ End Stop  5427 | | Original Length  1587 | Original Start Codon  ATG / GTG / TTG | |
| --- | --- | --- | --- | --- | --- | --- |
| Gene Direction  FWD / REV | Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?  Yes / No (if no, answer below) | | | | | | |
|  | Revised 5’ End Start |  | New Length | | | New Start Codon  ATG / GTG / TTG |

| Start called by both Glimmer and GeneMark? |
| --- |

Glimmer and GeneMark Calls

| Does this Gene have Coding Potential |
| --- |

Ribosome Binding Site

| RBS Final Score  -6.286 | Is this Score the Closest to 0?  Yes / No (if no) | Score Closest to 0: -4.331 |
| --- | --- | --- |
| Start #: 4408 |
| ORF Length:1020 (this position would cuts off too much of the gene) |

Starterator

| Starterator Start  3841 (with 100 MA’s, all MA are only in this position) |
| --- |
|
|

| Function **(F)**  major capsid and protease fusion protein | Gene No:5 | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  major capsid and protease fusion protein  Function was shown in many results | % Query  99% | % Identity  94.90% | | E-value  (if < 10-7)  0.0 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from BLAST(Protein) **(BLAST-phagesDB)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  major capsid and protease fusion protein  Function was shown in many results |  | % Identity  95% | | E-value  (if < 10-7)  0.0 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from HHpred **(HHpred)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Major Capsid Protein | Name / Description  Major capsid protein Rcc01687; "capsid", "jelly roll", "spike", "HK97", VIRUS; 3.42A {Rhodobacter capsulatus DE442} | Hit  6TSU\_A4 | | Probability (if > 90%)  99.9 | | E-value  (if < 1)  3.5e-23 |
| Supporting Information for Function, from Syntney **(SIF-Syn)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  major capsid and protease fusion protein  Function highly conserved | Pham number and number of members  Members (130) of Pham 1062 | | upstream pham conserved ?  Yes or No  If yes, what pham # or function ?  151755/portal protein | | Downstream pham conserved ?  Yes or No  If yes, what pham # or function ?  1065/head to tail adaptor | |
| Transmembrane domains by TmHmm using Phamerator | | | | N/A | | | |

| Gene # 6 | Original 5’ End Start  5431 | Original 3’ End Stop  5784 | | Original Length  354 | Original Start Codon  ATG / GTG / TTG | |
| --- | --- | --- | --- | --- | --- | --- |
| Gene Direction  FWD / REV | Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?  Yes / No (if no, answer below) | | | | | | |
|  | Revised 5’ End Start |  | New Length | | | New Start Codon  ATG / GTG / TTG |

| Start called by both Glimmer and GeneMark? |
| --- |

Glimmer and GeneMark Calls

| Does this Gene have Coding Potential |
| --- |



Ribosome Binding Site

| RBS Final Score  -4.553 | Is this Score the Closest to 0?  Yes / No (if no) | Score Closest to 0: |
| --- | --- | --- |
| Start #: |
| ORF Length: |

Starterator

| Starterator Start  5431 (with 100 MA’s, all MA are only in this position) |
| --- |
|
|

| Function **(F)**  Head-to-tail adaptor | Gene No:6 | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Head-to-tail adaptor  Function was shown in many results | % Query  99% | % Identity  98.29% | | E-value  (if < 10-7)  **3e-75** | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from BLAST(Protein) **(BLAST-phagesDB)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Head-to-tail adaptor  Function was shown in many results |  | % Identity  98% | | E-value  (if < 10-7)  1e-57 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from HHpred **(HHpred)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Head-to-tail adaptor  Function was shown in many results | Name / Description  gp6; Head-Tail Connector Protein gp6 of Bacteriophage HK97 and similar proteins. | Hit  Cd08054 | | Probability (if > 90%)  99.3% | | E-value  (if < 1)  9.6e-11 |
| Supporting Information for Function, from Syntney **(SIF-Syn)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Head-to-tail adaptor  Function was shown in many results | Pham number and number of members  Members (130) of Pham 1065 | | upstream pham conserved ?  Yes or No  If yes, what pham # or function ?  Major Capsid and protease fusion protein | | Downstream pham conserved ?  Yes or No  If yes, what pham # or function ?  Tail Terminator | |
| Transmembrane domains by TmHmm using Phamerator | | | | N/A | | | |

| Gene # 7 | Original 5’ End Start  5781 | Original 3’ End Stop  6161 | | Original Length  381 | Original Start Codon  ATG / GTG / TTG | |
| --- | --- | --- | --- | --- | --- | --- |
| Gene Direction  FWD / REV | Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?  Yes / No (if no, answer below) | | | | | | |
|  | Revised 5’ End Start |  | New Length | | | New Start Codon  ATG / GTG / TTG |

| Start called by both Glimmer and GeneMark? |
| --- |

Glimmer and GeneMark Calls

| Does this Gene have Coding Potential |
| --- |

Ribosome Binding Site

| RBS Final Score  -6.591 | Is this Score the Closest to 0?  Yes / No (if no) | Score Closest to 0: -4.122 |
| --- | --- | --- |
| Start #: 6030 |
| ORF Length:132 (this position would cuts off too much of the gene) |

Starterator

| Starterator Start  5781 (with 100 MA’s, all MA are only in this position) |
| --- |
|
|

| Function **(F)**  tail Terminator | Gene No:7 | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  tail Terminator | % Query  99 | % Identity  97.67% | | E-value  (if < 10-7)  1x10^-81 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from BLAST(Protein) **(BLAST-phagesDB)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  tail Terminator |  | % Identity  99 | | E-value  (if < 10-7)  2x10^64 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from HHpred **(HHpred)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  tail Terminator protein | Name / Descr (incl DUF)  Minor tail protein U; Mixed Alpha-Beta fold, VIRAL PROTEIN; HET: SO4, MSE; 2.7A {Enterobacteria phage lambda} SCOP: d.323.1.1 | Hit  [3FZ2\_I](https://www.rcsb.org/pdb/explore.do?structureId=3FZ2) | | Probability (if > 90%)  98.6 | | E-value  (if < 1)  0.000016 |
| Supporting Information for Function, from Syntney **(SIF-Syn)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  tail Terminator | Pham number and number of members  1057(128) | | upstream pham conserved ?  Yes or No  If yes, what pham # or function ?  1065 | | Downstream pham conserved ?  Yes or No  If yes, what pham # or function ?  139535 | |
| Transmembrane domains by TmHmm using Phamerator | | | | N/A | | | |

| Gene #8 | Original 5’ End Start  6202 | Original 3’ End Stop  6636 | | Original Length  435 | Original Start Codon  ATG / GTG / TTG | |
| --- | --- | --- | --- | --- | --- | --- |
| Gene Direction  FWD / REV | Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?  Yes / No (if no, answer below) | | | | | | |
|  | Revised 5’ End Start |  | New Length | | | New Start Codon  ATG / GTG / TTG |

| Start called by both Glimmer and GeneMark? |
| --- |

Glimmer and GeneMark Calls



| Does this Gene have Coding Potential |
| --- |

Ribosome Binding Site

| RBS Final Score  -5.797 | Is this Score the Closest to 0?  Yes / No (if no) | Score Closest to 0: -3.594 |
| --- | --- | --- |
| Start #: 6448 |
| ORF Length: 189 (this position would cuts off too much of the gene) |

Starterator

| Starterator Start  6202 (with 100 MA’s, all MA are only in this position) |
| --- |
|
|

| Function **(F)**  Major Tail Protein | Gene No:8 | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Major Tail Protein | % Query  99 | % Identity  100 | | E-value  (if < 10-7)  6x10^-100 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from BLAST(Protein) **(BLAST-phagesDB)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Major Tail Protein |  | % Identity  100 | | E-value  (if < 10-7)  1x10^81 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from HHpred **(HHpred)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Major Tail Protein | Name / Descr (incl DUF)  Major tail protein; Marine bacteriophage, Cryo-EM, Siphophage, Baseplate, Megatron protein, Tail fibre protein, Distal tail protein, Hub protein, VIRAL; 4.5A {Dinoroseobacter phage vB\_DshS-R4C} | Hit  [8GTC\_F](https://www.rcsb.org/pdb/explore.do?structureId=8GTC) | | Probability (if > 90%)  99.4 | | E-value  (if < 1)  3.4e-11 |
| Supporting Information for Function, from Syntney **(SIF-Syn)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Major Tail Protein | Pham number and number of members  139535(193) | | upstream pham conserved ?  Yes or No  If yes, what pham # or function ?  1057 | | Downstream pham conserved ?  Yes or No  If yes, what pham # or function ?  139555 | |
| Transmembrane domains by TmHmm using Phamerator | | | | N/A | | | |

| Gene # 9 | Original 5’ End Start  6649 | Original 3’ End Stop  7032 | | Original Length  384 | Original Start Codon  ATG / GTG / TTG | |
| --- | --- | --- | --- | --- | --- | --- |
| Gene Direction  FWD / REV | Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?  Yes / No (if no, answer below) | | | | | | |
|  | Revised 5’ End Start |  | New Length | | | New Start Codon  ATG / GTG / TTG |

| Start called by both Glimmer and GeneMark? |
| --- |

Glimmer and GeneMark Calls

| Does this Gene have Coding Potential |
| --- |



Ribosome Binding Site

| RBS Final Score  -4.294 | Is this Score the Closest to 0?  Yes / No (if no) | Score Closest to 0: |
| --- | --- | --- |
| Start #: |
| ORF Length: |

Starterator

| Starterator Start  6649 (with 104 MA’s, all MA are only in this position) |
| --- |
|
|

| Function **(F)** | Gene No:9 | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Hypothetical Protein | % Query  99 | % Identity  100 | | E-value  (if < 10-7)  5x10^-85 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from BLAST(Protein) **(BLAST-phagesDB)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Hypothetical Protein |  | % Identity  100 | | E-value  (if < 10-7)  2x10^66 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from HHpred **(HHpred)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Minor Capsid Protein | Name / Descr (incl DUF)  Minor\_capsid\_2 ; Minor capsid protein | Hit    [PF11114.12](http://pfam-legacy.xfam.org/family/PF11114.12) | | Probability (if > 90%)  99.4 | | E-value  (if < 1)  4.5e-12 |
| Supporting Information for Function, from Syntney **(SIF-Syn)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Minor tail | Pham number and number of members  139555(129) | | upstream pham conserved ?  Yes or No  If yes, what pham # or function ?  139535 | | Downstream pham conserved ?  Yes or No  If yes, what pham # or function ?  85197 | |
| Transmembrane domains by TmHmm using Phamerator | | | | N/A | | | |

| Gene # 10 | Original 5’ End Start  7046 | Original 3’ End Stop  7375 | | Original Length  330 | Original Start Codon  ATG / GTG / TTG | |
| --- | --- | --- | --- | --- | --- | --- |
| Gene Direction  FWD / REV | Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?  Yes / No (if no, answer below) | | | | | | |
|  | Revised 5’ End Start |  | New Length | | | New Start Codon  ATG / GTG / TTG |

| Start called by both Glimmer and GeneMark? |
| --- |

Glimmer and GeneMark Calls



| Does this Gene have Coding Potential |
| --- |

Ribosome Binding Site

| RBS Final Score  -4.148 | Is this Score the Closest to 0?  Yes / No (if no) | Score Closest to 0: |
| --- | --- | --- |
| Start #: |
| ORF Length: |

Starterator

| Starterator Start  7046 (with 100 MA’s, all MA are only in this position) |
| --- |
|
|

| Function **(F)**  Tail Assembly Chaperone | Gene No:10 | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Tail Assembly Chaperone | % Query  99 | % Identity  94.5% | | E-value  (if < 10-7)  4x10^-64 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from BLAST(Protein) **(BLAST-phagesDB)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Tail Assembly Chaperone |  | % Identity  95 | | E-value  (if < 10-7)  1x10^50 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from HHpred **(HHpred)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Tail Assembly Chaperone | Name / Descr (incl DUF)  Phage\_TAC\_12 ; Phage tail assembly chaperone protein, TAC | Hit  [PF12363.12](http://pfam.xfam.org/family/PF12363.12#tabview=tab0) | | Probability (if > 90%)  93.85 | | E-value  (if < 1)  0.39 |
| Supporting Information for Function, from Syntney **(SIF-Syn)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Tail Assembly Chaperone | Pham number and number of members  85197(129) | | upstream pham conserved ?  Yes or No  If yes, what pham # or function ?  139555 | | Downstream pham conserved ?  Yes or No  If yes, what pham # or function ? | |
| Transmembrane domains by TmHmm using Phamerator | | | | N/A | | | |

| Gene #  11 | Original 5’ End Start  7402 | Original 3’ End Stop  7491 | | Original Length  90 | Original Start Codon  ATG / GTG / TTG | |
| --- | --- | --- | --- | --- | --- | --- |
| Gene Direction  FWD / REV | Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?  Yes / No (if no, answer below) | | | | | | |
|  | Revised 5’ End Start  7046 |  | New Length  447 | | | New Start Codon  ATG / GTG / TTG |

| Start called by both Glimmer and GeneMark? |
| --- |

Glimmer and GeneMark Calls

| Does this Gene have Coding Potential |
| --- |



Ribosome Binding Site

| RBS Final Score  N/A | Is this Score the Closest to 0?  Yes / No (if no) | Score Closest to 0: |
| --- | --- | --- |
| Start #: |
| ORF Length: |

Starterator

| Starterator Start  N/A |
| --- |
|
|



| Function **(F)**  Tail Assemble Chaperone | Gene No: 11 | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Tail Assemble Chaperone  Function was shown in many results | % Query  99% | % Identity  94.59% | | E-value  (if < 10-7)  3e-92 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from BLAST(Protein) **(BLAST-phagesDB)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Tail Assemble Chaperone  Function was shown in many results |  | % Identity  94% | | E-value  (if < 10-7)  1e-70 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from HHpred **(HHpred)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Tail Assemble Chaperone | Name / Descr (incl DUF)  Phage\_TAC\_12 ; Phage tail assembly chaperone protein, TAC | Hit  PF12363.12 | | Probability (if > 90%)  93.04 | | E-value  (if < 1)  0.39 |
| Supporting Information for Function, from Syntney **(SIF-Syn)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Tail Assemble Chaperone  Function was shown in many results | Pham number and number of members  Members (17) of Pham 5553 | | upstream pham conserved ?  Yes or No  If yes, what pham # or function ?  85197/ Tail Assemble Chaperone | | Downstream pham conserved ?  Yes or No  If yes, what pham # or function ?  151697/ tape measure protein | |
| Transmembrane domains by TmHmm using Phamerator | | | | N/A | | | |

| Gene # 12 | Original 5’ End Start  7608 | Original 3’ End Stop  9713 | | Original Length  2106 | Original Start Codon  ATG / GTG / TTG | |
| --- | --- | --- | --- | --- | --- | --- |
| Gene Direction  FWD / REV | Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?  Yes / No (if no, answer below) | | | | | | |
|  | Revised 5’ End Start |  | New Length | | | New Start Codon  ATG / GTG / TTG |

| Start called by both Glimmer and GeneMark? |
| --- |

Glimmer and GeneMark Calls

| Does this Gene have Coding Potential |
| --- |

Ribosome Binding Site

| RBS Final Score  -4.386 | Is this Score the Closest to 0?  Yes / No (if no) | Score Closest to 0: |
| --- | --- | --- |
| Start #: |
| ORF Length: |

Starterator

| Starterator Start  7608 (with 100 MA’s, all MA are only in this position) |
| --- |
|
|

| Function **(F)**  Tape measure protein | Gene No: 12 | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Tape measure protein  Function was shown in many results | % Query  **99%** | % Identity  **99.29%** | | E-value  (if < 10-7)  **0** | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from BLAST(Protein) **(BLAST-phagesDB)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Tape measure protein  Function was shown in many results |  | % Identity  **100%** | | E-value  (if < 10-7)  **0** | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from HHpred **(HHpred)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Tape measure protein  Function was shown in many results | Name / Descr (incl DUF)  Composite atomic model of the Staphylococcus aureus phage 80alpha baseplate | Hit  6V8I\_BF | | Probability (if > 90%)  99.93 | | E-value  (if < 1)  8.7e-17 |
| Supporting Information for Function, from Syntney **(SIF-Syn)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Tape measure protein  Function was shown in many results | Pham number and number of members  **76903**  **106** | | upstream pham conserved ?  Yes or No  If yes, what pham # or function ?  **5553**  **17** | | Downstream pham conserved ?  Yes or No  If yes, what pham # or function ?  **76903**  **106** | |
| Transmembrane domains by TmHmm using Phamerator | | | | N/A | | | |

| Gene #13 | Original 5’ End Start  9710 | Original 3’ End Stop  10738 | | Original Length  1029 | Original Start Codon  ATG / GTG / TTG | |
| --- | --- | --- | --- | --- | --- | --- |
| Gene Direction  FWD / REV | Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?  Yes / No (if no, answer below) | | | | | | |
|  | Revised 5’ End Start |  | New Length | | | New Start Codon  ATG / GTG / TTG |

| Start called by both Glimmer and GeneMark? |
| --- |

Glimmer and GeneMark Calls

| Does this Gene have Coding Potential |
| --- |

Ribosome Binding Site

| RBS Final Score  -4.924 | Is this Score the Closest to 0?  Yes / No (if no) | Score Closest to 0: -4.473 |
| --- | --- | --- |
| Start #: 9968 |
| ORF Length:771 (this position would cut off too much of the gene) |

Starterator

| Starterator Start  9710 (with 85 MA’s, all MA are only in this position) |
| --- |
|
|

| Function **(F)**  Minor tail protein | Gene No: 13 | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Minor tail protein  Function was shown in many results | % Query  99% | % Identity  85.71% | | E-value  (if < 10-7)  0 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from BLAST(Protein) **(BLAST-phagesDB)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Minor tail protein  Function was shown in many results |  | % Identity  100% | | E-value  (if < 10-7)  0 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from HHpred **(HHpred)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  phage tail fiber | Name / Descr (incl DUF)  phage tail fiber; Phage type II tail fiber, VIRAL PROTEIN; 3.8A {Klebsiella phage Kp7} | Hit  7XYC\_B | | Probability (if > 90%)  97% | | E-value  (if < 1)  0.2 |
| Supporting Information for Function, from Syntney **(SIF-Syn)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Minor tail protein  Function was shown in many results | Pham number and number of members  76903  106 | | upstream pham conserved ?  Yes or No  If yes, what pham # or function ?  140967  223 | | Downstream pham conserved ?  Yes or No  If yes, what pham # or function ?  1068  129 | |
| Transmembrane domains by TmHmm using Phamerator | | | | N/A | | | |

| Gene #14 | Original 5’ End Start  10738 | Original 3’ End Stop  12798 | | Original Length  2061 | Original Start Codon  ATG / GTG / TTG | |
| --- | --- | --- | --- | --- | --- | --- |
| Gene Direction  FWD / REV | Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?  Yes / No (if no, answer below) | | | | | | |
|  | Revised 5’ End Start |  | New Length | | | New Start Codon  ATG / GTG / TTG |

| Start called by both Glimmer and GeneMark? |
| --- |

Glimmer and GeneMark Calls

| Does this Gene have Coding Potential |
| --- |

Ribosome Binding Site

| RBS Final Score  -4.089 | Is this Score the Closest to 0?  Yes / No (if no) | Score Closest to 0: |
| --- | --- | --- |
| Start #: |
| ORF Length: |

Starterator

| Starterator Start  10738 (with 100 MA’s, all MA are only in this position) |
| --- |
|
|

| Function **(F)**  Minor Tail Protein | Gene No:14 | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Minor Tail Protein | % Query  99 | % Identity  95.63% | | E-value  (if < 10-7)  0.0 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from BLAST(Protein) **(BLAST-phagesDB)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Minor Tail Protein |  | % Identity  95 | | E-value  (if < 10-7)  0.0 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from HHpred **(HHpred)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Function not on the list | Name / Descr (incl DUF)  Receptor-type tyrosine-protein phosphatase F; Fibronectin type-III | Hit    [6TPW\_A](http://www.rcsb.org/pdb/explore/explore.do?structureId=6TPW) | | Probability (if > 90%)  99.48 | | E-value  (if < 1)  1.2e-9 |
| Supporting Information for Function, from Syntney **(SIF-Syn)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Minor Tail Protein | Pham number and number of members  1068(130) | | upstream pham conserved ?  Yes or No  If yes, what pham # or function ?  76903 | | Downstream pham conserved ?  Yes or No  If yes, what pham # or function ?  85094 | |
| Transmembrane domains by TmHmm using Phamerator | | | | N/A | | | |

| Gene #15 | Original 5’ End Start  12800 | Original 3’ End Stop  13378 | | Original Length  579 | Original Start Codon  ATG / GTG / TTG | |
| --- | --- | --- | --- | --- | --- | --- |
| Gene Direction  FWD / REV | Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?  Yes / No (if no, answer below) | | | | | | |
|  | Revised 5’ End Start |  | New Length | | | New Start Codon  ATG / GTG / TTG |

| Start called by both Glimmer and GeneMark? |
| --- |

Glimmer and GeneMark Calls

| Does this Gene have Coding Potential |
| --- |



Ribosome Binding Site

| RBS Final Score  -5.109 | Is this Score the Closest to 0?  Yes / No (if no) | Score Closest to 0: -4.778 |
| --- | --- | --- |
| Start #: 13055 |
| ORF Length:324 (this position would cut off too much of the gene) |

Starterator

| Starterator Start  12800 (with 97 MA’s, all MA are only in this position) |
| --- |
|
|

| Function **(F)**  Minor tail protein | Gene No: 15 | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Minor tail protein  Function was shown in many results | % Query  99% | % Identity  93.75% | | E-value  (if < 10-7)  **3e-127** | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from BLAST(Protein) **(BLAST-phagesDB)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Minor tail protein  Function was shown in many results |  | % Identity  100% | | E-value  (if < 10-7)  e-108 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from HHpred **(HHpred)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  **\*too high e-value** | Name / Descr (incl DUF)  **too high e-value** | Hit  **too high e-value** | | Probability (if > 90%)  **too high e-value** | | E-value  (if < 1)  **too high e-value** |
| Supporting Information for Function, from Syntney **(SIF-Syn)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Minor tail protein  Function was shown in many results | Pham number and number of members  85094  158 | | upstream pham conserved ?  Yes or No  If yes, what pham # or function ?  1068  129 | | Downstream pham conserved ?  Yes or No  If yes, what pham # or function ?  85094 159 | |
| Transmembrane domains by TmHmm using Phamerator | | | | N/A | | | |

| Gene #16 | Original 5’ End Start  13423 | Original 3’ End Stop  13665 | | Original Length  243 | Original Start Codon  ATG / GTG / TTG | |
| --- | --- | --- | --- | --- | --- | --- |
| Gene Direction  FWD / REV | Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?  Yes / No (if no, answer below) | | | | | | |
|  | Revised 5’ End Start |  | New Length | | | New Start Codon  ATG / GTG / TTG |

| Start called by both Glimmer and GeneMark? |
| --- |

Glimmer and GeneMark Calls

| Does this Gene have Coding Potential |
| --- |

Ribosome Binding Site

| RBS Final Score  -3.365 | Is this Score the Closest to 0?  Yes / No (if no) | Score Closest to 0: |
| --- | --- | --- |
| Start #: |
| ORF Length: |

Starterator

| Starterator Start  13423 (with 37 MA’s, all MA are only in this position) |
| --- |
|
|

| Function **(F)**  Hypothetical Protein | Gene No:16 | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Hypothetical Protein | % Query  100 | % Identity  78.16 | | E-value  (if < 10-7)  2e-19 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from BLAST(Protein) **(BLAST-phagesDB)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Hypothetical Protein |  | % Identity  79 | | E-value  (if < 10-7)  9e-28 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from HHpred **(HHpred)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF) | Name / Descr (incl DUF) | Hit | | Probability (if > 90%)  All functions with probabilities below 90 | | E-value  (if < 1) |
| Supporting Information for Function, from Syntney **(SIF-Syn)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF) | Pham number and number of members  137948(55) | | upstream pham conserved ?  Yes or No  If yes, what pham # or function ?  85094 | | Downstream pham conserved ?  Yes or No  If yes, what pham # or function ?  141078 | |
| Transmembrane domains by TmHmm using Phamerator | | | | N/A | | | |

| Gene #17 | Original 5’ End Start  13696 | Original 3’ End Stop  14391 | | Original Length  696 | Original Start Codon  ATG / GTG / TTG | |
| --- | --- | --- | --- | --- | --- | --- |
| Gene Direction  FWD / REV | Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?  Yes / No (if no, answer below) | | | | | | |
|  | Revised 5’ End Start |  | New Length | | | New Start Codon  ATG / GTG / TTG |

| Start called by both Glimmer and GeneMark? |
| --- |

Glimmer and GeneMark Calls

| Does this Gene have Coding Potential |
| --- |

Ribosome Binding Site

| RBS Final Score  -4.989 (second closest to 0) | Is this Score the Closest to 0?  Yes / No (if no) | Score Closest to 0: -3.934 |
| --- | --- | --- |
| Start #: 14332 |
| ORF Length:60 (this position would cut off too much of the gene) |

Starterator

| Starterator Start  13696 (with 98 MA’s, all MA are only in this position) |
| --- |
|
|

| Function **(F)**  endolysin | Gene No: 17 | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Endolysin  Function was shown in many results | % Query  99% | % Identity  92.64% | | E-value  (if < 10-7)  0.0 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from BLAST(Protein) **(BLAST-phagesDB)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Endolysin  Function was shown in many results |  | % Identity  100% | | E-value  (if < 10-7)  e-137 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from HHpred **(HHpred)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Peptidase | Name / Descr (incl DUF)  Peptidase M23; membrane protein, enzyme, TRANSPORT PROTEIN; HET: ADP;{Vibrio cholerae} | Hit  8TZL\_E | | Probability (if > 90%)  99.32% | | E-value  (if < 1)  2.4e-10 |
| Supporting Information for Function, from Syntney **(SIF-Syn)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Endolysin  Function was shown in many results | Pham number and number of members  151805 (131) | | upstream pham conserved ?  Yes or No  137948 | | Downstream pham conserved ?  Yes or No  1063 | |
| Transmembrane domains by TmHmm using Phamerator | | | |  | | | |

Despite HHpred having results for peptidase, most of the evidence points out to this protein being endolysin.

| Gene #18 | Original 5’ End Start  14388 | Original 3’ End Stop  14621 | | Original Length  234 | Original Start Codon  ATG / GTG / TTG | |
| --- | --- | --- | --- | --- | --- | --- |
| Gene Direction  FWD / REV | Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?  Yes / No (if no, answer below) | | | | | | |
|  | Revised 5’ End Start  14358 |  | New Length  264 | | | New Start Codon  ATG / GTG / TTG |

| Start called by both Glimmer and GeneMark? |
| --- |

Glimmer and GeneMark Calls

| Does this Gene have Coding Potential |
| --- |

Ribosome Binding Site

| RBS Final Score  -6.724 | Is this Score the Closest to 0?  Yes / No (if no) | Score Closest to 0: -5.054 |
| --- | --- | --- |
| Start #: 14571 |
| ORF Length:51(this position would cut off too much of the gene) |

Starterator

| Starterator Start  14358 with 84 MA’s, as opposed to position 14388 with 16 MA’s. |
| --- |
|
|

| Function **(F)**  membrane protein | Gene No:18\* | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  membrane protein  function was shown in many results | % Query  98% | % Identity  94.32% | | E-value  (if < 10-7)  5e-52 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from BLAST(Protein) **(BLAST-phagesDB)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Function Unknown  No function was shown in many results |  | % Identity  94% | | E-value  (if < 10-7)  4e-43 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from HHpred **(HHpred)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  E-values too high | Name / Descr (incl DUF)  E values too high | Hit  E-values too high | | Probability (if > 90%)  E-values too high | | E-value  (if < 1)  E-values too high |
| Supporting Information for Function, from Syntney **(SIF-Syn)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Function Unknown  No function was shown in many results | Pham number and number of members  1063 (131) | | upstream pham conserved ?  Yes or No  If yes, what pham # or function ?  151805 | | Downstream pham conserved ?  Yes or No  If yes, what pham # or function ?  1053 | |
| Transmembrane domains by TmHmm using Phamerator | | | | Transmembrane found | | | |



| Gene #19 | Original 5’ End Start  14618 | Original 3’ End Stop  14842 | | Original Length  225 | Original Start Codon  ATG / GTG / TTG | |
| --- | --- | --- | --- | --- | --- | --- |
| Gene Direction  FWD / REV | Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?  Yes / No (if no, answer below) | | | | | | |
|  | Revised 5’ End Start |  | New Length | | | New Start Codon  ATG / GTG / TTG |

| Start called by both Glimmer and GeneMark? |
| --- |

Glimmer and GeneMark Calls

| Does this Gene have Coding Potential |
| --- |

Ribosome Binding Site

| RBS Final Score  -2.901 | Is this Score the Closest to 0?  Yes / No (if no) | Score Closest to 0: |
| --- | --- | --- |
| Start #: |
| ORF Length: |

Starterator

| Starterator Start  14618 (with 98 MA’s, all MA are only in this position) |
| --- |
|
|

| Function **(F)**  holin | Gene No:19 | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Holin  Function found in some of the results | % Query  98% | % Identity  97.30% | | E-value  (if < 10-7)  2e-42 | | AA # conserved in the first hits?  Yes or No |
| Supporting Information for Function, from BLAST(Protein) **(BLAST-phagesDB)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Holin  Only a few had this function |  | % Identity  97% | | E-value  (if < 10-7)  3e-35 | | AA # conserved in the first hits?  Yes or No |
| Supporting Information for Function, from HHpred **(HHpred)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Holin  Only one result found | Name / Descr (incl DUF)  Phage\_holin\_7\_1 ; Mycobacterial 2 TMS Phage Holin (M2 Hol) Family | Hit  [PF16081.9](http://pfam-legacy.xfam.org/family/PF16081.9) | | Probability (if > 90%)  98.3 | | E-value  (if < 1)  0.0000037 |
| Supporting Information for Function, from Syntney **(SIF-Syn)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Holin | Pham number and number of members  Members (132) of Pham 1053 | | upstream pham conserved ?  Yes or No  If yes, what pham # or function ?  1063 | | Downstream pham conserved ?  Yes or No  If yes, what pham # or function ?  85067 | |
| Transmembrane domains by TmHmm using Phamerator | | | | 2 transmembrane domains found | | | |



| Gene #20 | Original 5’ End Start  15120 | Original 3’ End Stop  14911 | | Original Length  210 | Original Start Codon  ATG / GTG / TTG | |
| --- | --- | --- | --- | --- | --- | --- |
| Gene Direction  FWD / REV | Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?  Yes / No (if no, answer below) | | | | | | |
|  | Revised 5’ End Start |  | New Length | | | New Start Codon  ATG / GTG / TTG |

| Start called by both Glimmer and GeneMark? |
| --- |

Glimmer and GeneMark Calls

| Does this Gene have Coding Potential |
| --- |

Ribosome Binding Site

| RBS Final Score  -3.627 | Is this Score the Closest to 0?  Yes / No (if no) | Score Closest to 0: |
| --- | --- | --- |
| Start #: |
| ORF Length: |

Starterator

| Starterator Start  15120 (with 98 MA’s, all MA are only in this position) |
| --- |
|
|

| Function **(F)**  Lsr2-like DNA bridging protein | Gene No: 20 | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Lsr2-like DNA bridging protein | % Query  98% | % Identity  100% | | E-value  (if < 10-7)  7e-42 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from BLAST(Protein) **(BLAST-phagesDB)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Lsr2-like DNA bridging protein |  | % Identity  100% | | E-value  (if < 10-7)  1e-24 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from HHpred **(HHpred)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  DNA binding protein | Name / Descr (incl DUF)  Crystal structure of the dimerization domain of Lsr2 from Mycobacterium tuberculosis in the P 1 21 1 space group | Hit  4E1P\_A | | Probability (if > 90%)  99.88% | | E-value  (if < 1)  2.5e-21 |
| Supporting Information for Function, from Syntney **(SIF-Syn)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Lsr2-like DNA bridging protein | Pham number and number of members  Members (164) of Pham 85067 | | upstream pham conserved ?  Yes or No  If yes, what pham # or function ?  1053 | | Downstream pham conserved ?  Yes or No  If yes, what pham # or function ?  1059 | |
| Transmembrane domains by TmHmm using Phamerator | | | | N/A | | | |

| Gene #21 | Original 5’ End Start  15566 | Original 3’ End Stop  15123 | | Original Length  444 | Original Start Codon  ATG / GTG / TTG | |
| --- | --- | --- | --- | --- | --- | --- |
| Gene Direction  FWD / REV | Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?  Yes / No (if no, answer below) | | | | | | |
|  | Revised 5’ End Start  15629 |  | New Length  507 | | | New Start Codon  ATG / GTG / TTG |

| Start called by both Glimmer and GeneMark? |
| --- |

Glimmer and GeneMark Calls

| Does this Gene have Coding Potential |
| --- |

Ribosome Binding Site

| RBS Final Score  -6.782 | Is this Score the Closest to 0?  Yes / No (if no) | Score Closest to 0: -3.155 |
| --- | --- | --- |
| Start #: 15566 |
| ORF Length:444 (gene is shorten using this start) |

Starterator

| Starterator Start  15629 has 78 MA’s while 15566 has 5 MA’s |
| --- |
|
|

| Function **(F)**  helix-turn-helix DNA binding domain | Gene No:21 | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  helix-turn-helix DNA binding domain  Function found in many results | % Query  99% | % Identity  91.07% | | E-value  (if < 10-7)  7e-107 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from BLAST(Protein) **(BLAST-phagesDB)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  helix-turn-helix DNA binding domain  Function found in many results |  | % Identity  92% | | E-value  (if < 10-7)  1e-84 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from HHpred **(HHpred)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  helix-turn-helix DNA binding domain | Name / Descr (incl DUF)  AM32; GRAM-POSITIVE, CONJUGATION, HELIX-TURN-HELIX, DSDNA BINDING, DNA BINDING PROTEIN; HET: MSE; 1.35A {Enterococcus faecalis} | Hit  4P0Z\_A | | Probability (if > 90%)  99.1 | | E-value  (if < 1)  4.8e-9 |
| Supporting Information for Function, from Syntney **(SIF-Syn)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  helix-turn-helix DNA binding domain  Function found in many results | Pham number and number of members  Members (131) of Pham 1059 | | upstream pham conserved ?  Yes or No  If yes, what pham # or function ?  85067 | | Downstream pham conserved ?  Yes or No  If yes, what pham # or function ?  85194 | |
| Transmembrane domains by TmHmm using Phamerator | | | | N/A | | | |

| Gene # 22 | Original 5’ End Start  15941 | Original 3’ End Stop  15711 | | Original Length  231 | Original Start Codon  ATG / GTG / TTG | |
| --- | --- | --- | --- | --- | --- | --- |
| Gene Direction  FWD / REV | Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?  Yes / No (if no, answer below) | | | | | | |
|  | Revised 5’ End Start |  | New Length | | | New Start Codon  ATG / GTG / TTG |

| Start called by both Glimmer and GeneMark? |
| --- |

Glimmer and GeneMark Calls

| Does this Gene have Coding Potential |
| --- |

Ribosome Binding Site

| RBS Final Score  -7.839 | Is this Score the Closest to 0?  Yes / No (if no) | Score Closest to 0: -6.079 |
| --- | --- | --- |
| Start #: 15857 |
| ORF Length:147 (this position would cut off too much of the gene) |

Starterator

| Starterator Start  15941 (with 100 MA’s, all MA are only in this position) |
| --- |
|
|

| Function **(F)**  helix-turn-helix DNA binding domain | Gene No:22 | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  helix-turn-helix DNA binding domain  Function found in many results | % Query  97% | % Identity  93.33% | | E-value  (if < 10-7)  9e-42 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from BLAST(Protein) **(BLAST-phagesDB)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  helix-turn-helix DNA binding domain  Function found in many results |  | % Identity  100% | | E-value  (if < 10-7)  6e-37 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from HHpred **(HHpred)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  helix-turn-helix DNA binding domain | Name / Descr (incl DUF)  Stl; Mobile Genetic Element, Repressor, SaPI, HTH domain, DNA BINDING PROTEIN; HET: MSE; 2.901A {Staphylococcus aureus} | Hit  7P4A\_A | | Probability (if > 90%)  98.6 | | E-value  (if < 1)  0.000015 |
| Supporting Information for Function, from Syntney **(SIF-Syn)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  helix-turn-helix DNA binding domain  Function found in many results | Pham number and number of members  Members (131) of Pham 85194 | | upstream pham conserved ?  Yes or No  If yes, what pham # or function ?  1059 | | Downstream pham conserved ?  Yes or No  If yes, what pham # or function ?  3206 | |
| Transmembrane domains by TmHmm using Phamerator | | | | N/A | | | |

| Gene # 23 | Original 5’ End Start  16657 | Original 3’ End Stop  16839 | | Original Length  183 | Original Start Codon  ATG / GTG / TTG | |
| --- | --- | --- | --- | --- | --- | --- |
| Gene Direction  FWD / REV | Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?  Yes / No (if no, answer below) | | | | | | |
|  | Revised 5’ End Start |  | New Length | | | New Start Codon  ATG / GTG / TTG |

| Start called by both Glimmer and GeneMark? |
| --- |

Glimmer and GeneMark Calls

| Does this Gene have Coding Potential |
| --- |

Ribosome Binding Site

| RBS Final Score  -5.914 | Is this Score the Closest to 0?  Yes / No (if no) | Score Closest to 0: -5.399 |
| --- | --- | --- |
| Start #: 16453 |
| ORF Length: 387 |

Starterator

| Starterator Start  16657 (with 22 MA’s, all MA are in this position but 1) |
| --- |
|
|

| Function **(F)**  Hypothetical Protein | Gene No: 23 | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Hypothetical Protein | % Query  100% | % Identity  98.33% | | E-value  (if < 10-7)  2e-33 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from BLAST(Protein) **(BLAST-phagesDB)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Hypothetical protein |  | % Identity  100% | | E-value  (if < 10-7)  3e-28 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from HHpred **(HHpred)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)    All results below 90% | Name / Descr (incl DUF)  All results below 90% | Hit  All results below 90% | | Probability (if > 90%)  All results below 90% | | E-value  (if < 1)  All results below 90% |
| Supporting Information for Function, from Syntney **(SIF-Syn)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  Hypothetical Protein | Pham number and number of members  Members (34) of Pham 3206 | | upstream pham conserved ?  Yes or No  If yes, what pham # or function ?  Pham 85194 | | Downstream pham conserved ?  Yes or No  If yes, what pham # or function ?  Pham 154556 | |
| Transmembrane domains by TmHmm using Phamerator | | | | N/A | | | |

| Gene #24 | Original 5’ End Start  16940 | Original 3’ End Stop  17158 | | Original Length  219 | Original Start Codon  ATG / GTG / TTG | |
| --- | --- | --- | --- | --- | --- | --- |
| Gene Direction  FWD / REV | Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?  Yes / No (if no, answer below) | | | | | | |
|  | Revised 5’ End Start |  | New Length | | | New Start Codon  ATG / GTG / TTG |

| Start called by both Glimmer and GeneMark? |
| --- |

Glimmer and GeneMark Calls

| Does this Gene have Coding Potential |
| --- |

Ribosome Binding Site

| RBS Final Score  -7.139 | Is this Score the Closest to 0?  Yes / No (if no) | Score Closest to 0: -5.685 |
| --- | --- | --- |
| Start #: 17081 |
| ORF Length:78 (this position would cut off too much of the gene) |

Starterator

| Starterator Start  16940 with 78 MA’s |
| --- |
|
|

| Function **(F)**  helix-turn-helix DNA binding domain | Gene No:24 | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  helix-turn-helix DNA binding protein  Many results show MerR HTH | % Query  98% | % Identity  94.44% | | E-value  (if < 10-7)  2e-41 | | AA # conserved in first 10 hits?  Yes or No  N/A |
| Supporting Information for Function, from BLAST(Protein) **(BLAST-phagesDB)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  helix-turn-helix DNA binding domain protein  Function shown in many results |  | % Identity  97% | | E-value  (if < 10-7)  2e-34 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from HHpred **(HHpred)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  helix-turn-helix DNA binding Protein | Name / Descr (incl DUF)  Recombination Directionality Factor RdfS; Excisionase, Recombination Directionality Factor, winged helix-turn-helix, superhelix, DNA BINDING PROTEIN; HET: GOL; 2.45A {Mesorhizobium japonicum R7A} | Hit  8DGL\_A | | Probability (if > 90%)  99.5 | | E-value  (if < 1)  2.8e-13 |
| Supporting Information for Function, from Syntney **(SIF-Syn)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  helix-turn-helix DNA binding domain | Pham number and number of members  Members (130) of Pham 154556 | | upstream pham conserved ?  Yes or No  If yes, what pham # or function ?  3206 | | Downstream pham conserved ?  Yes or No  If yes, what pham # or function ?  154440 | |
| Transmembrane domains by TmHmm using Phamerator | | | | N/A | | | |

| Gene # 25 | Original 5’ End Start  17155 | Original 3’ End Stop  17454 | | Original Length  300 | Original Start Codon  ATG / GTG / TTG | |
| --- | --- | --- | --- | --- | --- | --- |
| Gene Direction  FWD / REV | Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?  Yes / No (if no, answer below) | | | | | | |
|  | Revised 5’ End Start |  | New Length | | | New Start Codon  ATG / GTG / TTG |

| Start called by both Glimmer and GeneMark? |
| --- |
| Does this Gene have Coding Potential |

Glimmer and GeneMark Calls

Ribosome Binding Site

| RBS Final Score  -6.910 | Is this Score the Closest to 0?  Yes / No (if no) | Score Closest to 0: -5.036 |
| --- | --- | --- |
| Start #: 17398 |
| ORF Length:57 (this position would cut off too much of the gene) |

Starterator

| Starterator Start  17155 (with 100 MA’s, all MA are only in this position) |
| --- |
|
|

| Function **(F)**  HNH endonuclease | Gene No:25 | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  HNH endonuclease  Function found for many results | % Query  99% | % Identity  95.96% | | E-value  (if < 10-7)  2e-64 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from BLAST(Protein) **(BLAST-phagesDB)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  HNH endonuclease  Function found for many results |  | % Identity  95% | | E-value  (if < 10-7)  5e-53 | | AA # conserved in first 10 hits?  Yes or No |
| Supporting Information for Function, from HHpred **(HHpred)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  HNH endonuclease | Name / Description DUF)  HNH endonuclease; Thermophilic bacteriophage, HNH Endonuclease, DNA nicking, HYDROLASE; 1.52A {Geobacillus virus E2} | Hit  5H0M\_A | | Probability (if > 90%)  97.2 | | E-value  (if < 1)  0.0017 |
| Supporting Information for Function, from Syntney **(SIF-Syn)**  (All three lines of SIF evidence must be filled in) | Function  (or NKF)  HNH endonuclease  Function found for many results | Pham number and number of members  Members (281) of Pham 154440 | | upstream pham conserved ?  Yes or No  If yes, what pham # or function ?  154556 | | Downstream pham conserved ?  Yes or No  If yes, what pham # or function ?  Last gene in the genome | |
| Transmembrane domains by TmHmm using Phamerator | | | | N/A | | | |

Notes on tRNA genes:

tRNA scans show no results. Aragorn made a call at [9471, 9542], which overlaps with gene 12 [7608, 9713] by 72 bp. We have determined that Biscayne doesn’t have any tRNA genes.