Phage Bush:
Annotation Documentation

Head Instructor

Jaime Mayoral

Students

Victor Adolpho de Melo

Nicole Gonzalez Giliberti

Grace Intrator

Kristen Mclean

Akram M Mikhail

Carlos E Rodriguez

Arianna M Ruiz

Betty R Sierra

David Vega

Kiryl Yasinski

Gene # 1	Original 5' End Start 1	Original 3' End Stop 522	Original Length 522	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

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes]	Start:1
	No	Glimmer Start:
		GeneMark Start:

Does this Gene have Coding Potential	[✓ Yes] No	Estimated Start: 1

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 1 (With 7 MA's, all MA are in this start.)

Function **(F)** Terminase small subunit	Gene No:1		Student name: Betty Sierra/Grace Intrator Date:2/6/24		
Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)** (All three lines of SIF evidence must be filled in)	Function (or NKF) Terminase small subunit	% Query 99	% Identity 95.95	E-value (if < 10^{-7}) 1e-111	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 small subunit		% Identity 97	E-value (if < 10^{-7}) 2e-90	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, small subunit	Name / Descr (incl DUF) Small terminase subunit	Hit 7JQQ_ii	Probability (if > 90%) 97.12	E-value (if < 1) 0.033
Supporting Information for Function, from Syntney **(SIF-Syn)** (All three lines of SIF evidence must be filled in)	Function (or NKF) Terminase, small subunit	Pham number and number of members 87960 (9)	upstream pham conserved ? Yes or No If yes, what pham # or function ? N/A first gene of the genome		Downstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 8570
Transmembrane domains by TmHmm using Phamerator			N/A		

Notes : All evidence led to it being terminase small subunit. In addition, there is a terminase large subunit in the genome, so we think this could be a small subunit.

Gene # 2	Original 5’ End Start 519	Original 3’ End Stop 656	Original Length 138	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

Glimmer and GeneMark Calls

Does this Gene have Coding Potential	[✓ Yes] No

Estimated Start:505

Start called by both Glimmer and GeneMark?	[✓ Yes]	Start:519
	No	Glimmer Start:
		GeneMark Start:

Ribosome Binding Site

RBS Final Score	Is this Score the Closest to 0?	Score Closest to 0: -6.179
-6.770	Yes / [✓ No] (if no)	Start #: 585
		ORF Length:72 (this start cuts off too much of the gene)

Starterator

Starterator Start

519 (With 7 MA’s, all MA are in this start.)

Function **(F)** Membrane protein	Gene No: 2	Student name: Carlos Rodriguez Date: 02/25/2024			
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 95	% Identity 95.56	E-value (if < 10^{-7}) 2e-20	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 95	E-value (if < 10^{-7}) 1e-17	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) NKF	Name / Descr (incl DUF) DUF3789 ; Protein of unknown function (DUF3789)	Hit PF12664.11	Probability (if > 90%) 95.39	E-value (if < 1) 0.04
Supporting Information for Function, from Syntney **(SIF-Syn)** (All three lines of SIF evidence must be filled in)	Function (or NKF) NKF	Pham number and number of members 8570(9)	upstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 87960	Downstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 8381	

Transmembrane domains by TmHmm using Phamerator	Yes, one was found by two programs

Notes : The protein does not have a known function.

Gene #3	Original 5’ End Start 653	Original 3’ End Stop 880	Original Length 228	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 h	New Start Codon ATG / GTG / TTG

Does this Gene have Coding Potential	[✓ Yes] No

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes]	Start:653
	No	Glimmer Start:
		GeneMark Start:

Estimated Start:653

Ribosome Binding Site

RBS Final Score -5.621 (second closest to 0)	Is this Score the Closest to 0? Yes / [✓ No] (if no)	Score Closest to 0: -5.531
		Start #: 854
		ORF Length:27 (this start cuts off too much of the gene)

Starterator

Starterator Start

653 (With 7 MA's, all MA are in this start.)

Function **(F)** Hypothetical Protein	Gene No: 3			Student name: Carlos Rodriguez Date: 02/25/2024	
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 98%	% Identity 100.0%	E-value (if < 10^{-7}) 3e-45	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)		% Identity 100%	E-value (if < 10^{-7}) 4e-38	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 probablities 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 8381(9)	upstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 8570	Downstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 143300	
Transmembrane domains by TmHmm using Phamerator	N/A				

Does this Gene have Coding Potential	[✓ Yes] No

Gene # 4	Original 5' End Start 883	Original 3' End Stop 1059	Original Length 177	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

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes]	Start:883
	No	Glimmer Start:
		GeneMark Start:

Estimated Start:883

Ribosome Binding Site

RBS Final Score -2.034	Is this Score the Closest to 0? [✓ Yes] / No (if no)	Score Closest to 0:
		Start #:
		ORF Length:

Starterator

Starterator Start 883 (With 2 MA's, all MA are in this start.)

Function **(F)** Hypothetical Protein	Gene No: 4	Student name: Betty Sierra / Grace Intrator Date: 2/13/24			
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 98%	% Identity 84.48%	E-value (if < 10^{-7}) 2e-27	AA # conserved in first 10 hits? [✓ Yes] or No
Supporting Information for Function, from BLAST (Protein) **(BLAST-phagesDB)**	Function (or [✓ NKF])		% Identity 100%	E-value (if < 10^{-7}) 8e-29	AA # conserved in first 10 hits? [✓ Yes] or No

(All three lines of SIF evidence must be filled in)					
Supporting Information for Function, from HHpred **(HHpred)** (All three lines of SIF evidence must be filled in)	Function (or [✓ NKF])	Name / Descr (incl DUF) DUF6225 ; Family of unknown function	Hit PF19735.3	Probability (if > 90%) 94.63	E-value (if < 1) 0.13
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 146389 (4)	upstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 8381	Downstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 142641	
Transmembrane domains by TmHmm using Phamerator			N/A		

Gene # 5	Original 5’ End Start 1059	Original 3’ End Stop 1172	Original Length 114	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

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes] No	Start:1059 Glimmer Start: GeneMark Start:

Estimated Start:1050

Ribosome Binding Site

RBS Final Score -3.474	Is this Score the Closest to 0? [✓ Yes] / No (if no)	Score Closest to 0: Start #: ORF Length:

Does this Gene have Coding Potential	[✓ Yes] No

Starterator

Starterator Start

1059 (With 3 MA's, all MA are in this start.)

Function **(F)** Hypothetical protein	Gene No: 5		Student name: Betty Sierra / Grace Intrator Date: 2/13/24		
Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)**	Function (or NKF) Hypothetical protein	% Query 97	% Identity 100	E-value (if < 10^{-7}) 2e-18	AA # conserved in first 10 hits? [✓ Yes] or No

(All three lines of SIF evidence must be filled in)					
Supporting Information for Function, from BLAST (Protein) **(BLAST-phagesDB)** (All three lines of SIF evidence must be filled in)	Function (or [✓ NKF])		% Identity 100	E-value (if < 10^{-7}) 1e-17	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]) Protein not on the list, also it has just one with probability above 90	Name / Descr (incl DUF) DUF1660 ; Prophage protein	Hit PF07874.15	Probability (if > 90%) 98.2	E-value (if < 1) 0.000003
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 147239(43)	upstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 143300	Downstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 8812	
Transmembrane domains by TmHmm using Phamerator	N/A				

Gene # 6	Original 5' End Start 1172	Original 3' End Stop 1405	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

New Length	New Start Codon ATG / GTG / TTG

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes]	Start:1172
	No	Glimmer Start:
		GeneMark Start:

Estimated Start:1130

Ribosome Binding Site

Does this Gene have Coding Potential	[✓ Yes] No

RBS Final Score -4.189	Is this Score the Closest to 0? [✓ Yes] / No (if no)	Score Closest to 0:
		Start #:
		ORF Length:

Starterator

Starterator Start 1172 (With 6 MA’s, all MA are in this start.)

Function **(F)** Hypothetical protein	Gene No: 6	Student name: Carlos Rodriguez/ Betty Sierra / Grace Intrator Date: 02/25/2024

<table>
<tr><td>Supporting Information for Function, from BLAST (Protein) (BLAST-NCBI) (All three lines of SIF evidence must be filled in)</td><td>Function
(or NKF)
Hypothetical protein</td><td>% Query
98</td><td>% Identity
88.31</td><td>E-value
(if < 10^{-7})
4e-39</td><td>AA # conserved in first 10 hits?
Yes or No</td></tr>
<tr><td>Supporting Information for Function, from BLAST (Protein) (BLAST-phagesDB) (All three lines of SIF evidence must be filled in)</td><td>Function
(or NKF)</td><td></td><td>% Identity
94</td><td>E-value
(if < 10^{-7})
5e-37</td><td>AA # conserved in first 10 hits?
Yes or No</td></tr>
<tr><td>Supporting Information for Function, from HHpred (HHpred) (All three lines of SIF evidence must be filled in)</td><td>Function
(or NKF)</td><td>Name / Descr (incl DUF)</td><td>Hit</td><td>Probability (if > 90%)
All probabilities below 90</td><td>E-value
(if < 1)</td></tr>
<tr><td>Supporting Information for Function, from Syntney (SIF-Syn) (All three lines of SIF evidence must be filled in)</td><td>Function
(or NKF)</td><td colspan="2">Pham number and number of members
8812(9)</td><td>upstream pham conserved ?
Yes or No
If yes, what pham # or function ?
142641</td><td>Downstream pham conserved ?
Yes or No
If yes, what pham # or function ?
1724</td></tr>
<tr><td colspan="4">Transmembrane domains by TmHmm using Phamerator</td><td colspan="2">N/A</td></tr>
</table>

Gene # 7	Original 5’ End Start 1527	Original 3’ End Stop 2855	Original Length 1329	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

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	Yes [✓ No]	Start: Glimmer Start:1527 GeneMark Start:1518

Does this Gene have Coding Potential	[✓ Yes] No	Estimated Start:1527

Ribosome Binding Site

RBS Final Score -6.565	Is this Score the Closest to 0? Yes / [✓ No] (if no)	Score Closest to 0: -4.315 Start #: 1704 ORF Length:1152 (this start cuts off too much of the gene)

Starterator

Starterator Start 1527 with 16 MA’s

Function **(F)** Terminase, large subunit	Gene No: 7	Student name: Carlos Rodriguez/ Betty Sierra / Grace Intrator Date: 02/25/2024			
Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)** (All three lines of SIF evidence must be filled in)	Function (or NKF) Terminase, large subunit	% Query 99%	% Identity 99.78%	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) Terminase, large subunit		% Identity 99%	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) Terminase, large subunit	Name / Descr (incl DUF) Large subunit terminase; large terminase, VIRAL PROTEIN; 2.2A {Deep-sea thermophilic phage D6E}	Hit 5OE8_B	Probability (if > 90%) 100	E-value (if < 1) 1.1e-42
Supporting Information for Function, from Syntney **(SIF-Syn)** (All three lines of SIF evidence must be filled in)	Function (or NKF) Terminase large subunit	Pham number and number of members 1724(62)	upstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 8812	Downstream pham conserved ? Yes or [✓ No] If yes, what pham # or function ?	
Transmembrane domains by TmHmm using Phamerator	N/A				

Gene # 8	Original 5' End Start 2830	Original 3' End Stop 2964	Original Length 135	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 Gene deleted		New Length	New Start Codon ATG / GTG / TTG

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	Yes [✓ No]	Start:
		Glimmer Start:2830
		GeneMark Start: not called

Does this Gene have Coding Potential	Yes [✓ No]	Estimated Start:

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 No start reported

Notes: This gene has no coding potential and the staterator report is all for draft, also it overlaps with gene 7 (upstream) and gene 8 (downstream). Therefore, it was deleted.

Function **(F)** Gene Deleted	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)	% Query	% Identity	E-value (if < 10^{-7})	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)		% Identity	E-value (if < 10^{-7})	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%)	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		upstream pham conserved ? Yes or No If yes, what pham # or function ?	Downstream pham conserved ? Yes or No If yes, what pham # or function ?
Transmembrane domains by TmHmm using Phamerator					

Gene # 9	Original 5’ End Start 2921	Original 3’ End Stop 4535	Original Length 1614	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

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	Yes [✓ No]	Start: Glimmer Start:2921 GeneMark Start:2984

Does this Gene have Coding Potential	[✓ Yes] No	Estimated Start:2927

Ribosome Binding Site

RBS Final Score -4.475	Is this Score the Closest to 0? Yes / [✓ No] (if no)	Score Closest to 0: -3.449 Start #: 3182 ORF Length:1353 (this start cuts off too much of the gene)

Starterator

Starterator Start 2921 (With 10 MA’s, all MA are in this start.)

Function **(F)** portal protein	Gene No: 9	Student name: Kiryl Yasinski Date: 2/13/24			
Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)** (All three lines of SIF evidence must be filled in)	Function (or NKF) portal protein	% Query 99	% Identity 98.51	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) portal protein		% Identity 98	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) portal protein	Name / Descr (incl DUF) Portal protein; Bacteriophage, SPP1, Portal Protein, Head completion proteins, Connector Complex	Hit 7Z4W_B	Probability (if > 90%) 100	E-value (if < 1) 1.6e^-31
Supporting Information for Function, from Syntney **(SIF-Syn)** (All three lines of SIF evidence must be filled in)	Function (or NKF) portal protein	Pham number and number of members 1736(62)	upstream pham conserved ? Yes or [✓ No] If yes, what pham # or function ?	Downstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 141467	
Transmembrane domains by TmHmm using Phamerator	N/A				

Gene # 10	Original 5’ End Start 4537	Original 3’ End Stop 5652	Original Length 1116	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

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes]	Start:4537
	No	Glimmer Start: GeneMark Start:
Does this Gene have Coding Potential	[✓ Yes] No	

Estimated Start:4560

Ribosome Binding Site

RBS Final Score -4.247	Is this Score the Closest to 0? [✓ Yes] / No (if no)	Score Closest to 0: Start #: ORF Length:

Starterator

Starterator Start
4537 (With 7 MA’s, all MA are in this start.)

Function **(F)** Hypothetical Protein	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)	% Query	% Identity	E-value (if < 10^{-7})	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)		% Identity	E-value (if < 10^{-7})	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%)	E-value (if < 1)
Supporting Information for Function, from Syntney **(SIF-Syn)**	Function (or NKF)	Pham number and number of members	upstream pham conserved ? Yes or No If yes, what pham # or function ?	Downstream pham conserved ? Yes or No If yes, what pham # or function ?	

(All three lines of SIF evidence must be filled in)				
Transmembrane domains by TmHmm using Phamerator				

Gene # 11	Original 5' End Start 5682	Original 3' End Stop 5960	Original Length 279	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 5649	Does this Gene have Coding Potential [✓ Yes] No	Ne th	New Start Codon ATG / [✓ GTG] / TTG

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes]	Start:5682
	No	Glimmer Start:
		GeneMark Start:

Estimated Start:5682

Ribosome Binding Site

RBS Final Score -6.464	Is this Score the Closest to 0? Yes / [✓ No] (if no)	Score Closest to 0: -4.069
		Start #: 5715
		ORF Length: 246

Starterator

Starterator Start

5682 has no other MA's

Function **(F)** Hypothetical protein	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) Hypothetical protein	% Query 98	% Identity 45.26	E-value (if < 10^{-7}) 1e-16	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])		% Identity 48	E-value (if < 10^{-7}) 1e-15	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) Immunity protein	Name / Descr (incl DUF) Imm35 ; Immunity protein 35	Hit PF15567.10	Probability (if > 90%) 99.67	E-value (if < 1) 2.1e-16
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 137745 (114)	upstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 141467	Downstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 8957	
Transmembrane domains by TmHmm using Phamerator	N/A				

Gene #12	Original 5' End Start 5965	Original 3' End Stop 6156	Original Length 192	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 6007		New Length 150	New Start Codon [✓ ATG] / GTG / TTG

Start called by both

mmer and GeneMark Calls

Glimmer and GeneMark?	Yes [✓ No]	Start: Glimmer Start:5965 GeneMark Start:6007

[✓ Yes] No	Estimated Start:6001

Ribosome Binding Site

RBS Final Score -5.994	Is this Score the Closest to 0? Yes [✓ / No] (if no)	Score Closest to 0: -5.555 Start #: 6145 ORF Length:12

Starterator

Starterator Start 6007 has no MA's

Notes: Despite there not being any MA's in starterator for this position, coding potential begins around 6001, which is closer to 6007 than 6965. Also, start 6007 has a lower RBS than 6965.

Function **(F)** Hypothetical protein	Gene No: 12	Student name: Carlos Rodriguez Date: 02/25/2024

Does this Gene have Coding Potential

Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)** (All three lines of SIF evidence must be filled in)	Function (or NKF)	% Query	% Identity	E-value (if < 10^{-7})	AA # conserved in first 10 hits?
	Hypothetical protein	98	93.88	1e-24	[✓ 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])		% Identity 93	E-value (if < 10^{-7}) 7e-20	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 results are 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 8957 (9)	upstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 137745	Downstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 138124

Transmembrane domains by TmHmm using Phamerator	N/A

Gene #13*	Original 5' End Start 6255	Original 3' End Stop 7058	Original Length 804	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

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	Yes No	Start: Glimmer Start:6255 GeneMark Start:6282

Estimated Start:

Ribosome Binding Site

Does this Gene have Coding Potential	Yes No

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

Starterator

Starterator Start

Function **(F)** Scaffolding protein	Gene No: 13	Student name: Carlos Rodriguez Date: 02/25/2024			
Supporting Information for Function,	Function (or NKF) Scaffolding protein	% Query 99%	% Identity 99.88	E-value (if < 10^{-7})	AA # conserved in first 10 hits?

from BLAST (Protein) **(BLAST-NCBI)** (All three lines of SIF evidence must be filled in)				0	[✓ 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) Scaffolding protein		% Identity 98	E-value (if < 10^{-7}) e-150	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) Scaffolding protein	Name / Description) Scaffold protein; major capsid protein, HK97-like fold, scaffolding protein, procapsid, VIRUS; 3.72A {Staphylococcus pha	Hit 6B0X_b	Probability (if > 90%) 97.15	E-value (if < 1) 0.092
Supporting Information for Function, from Syntney **(SIF-Syn)** (All three lines of SIF evidence must be filled in)	Function (or NKF) Scaffolding protein	Pham number and number of members 138124 (29)		upstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 8957	Downstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 2582
Transmembrane domains by TmHmm using Phamerator				No	

Gene #14	Original 5' End Start 7075	Original 3' End Stop 7947	Original Length 873	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

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes] No	Start: 7075 Glimmer Start: GeneMark Start:

Does this Gene have Coding Potential	[✓ Yes] No	Estimated Start: 7080

Ribosome Binding Site

RBS Final Score -4.626	Is this Score the Closest to 0? Yes / [✓ No] (if no)	Score Closest to 0: -4.473 Start #: 7156 ORF Length:792

Starterator

Starterator Start 7075 has 33 MA's

Function **(F)** Major capsid protein	Gene No: 14		Student name: Carlos Rodriguez Date: 02/25/2024		
Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)** (All three lines of SIF evidence must be filled in)	Function (or NKF) Major capsid protein	% Query 99%	% Identity 95.86%	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) Major capsid protein		% Identity 95	E-value (if < 10^{-7}) e-154	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 / Descr (incl DUF) Major capsid protein; capsid, HK97-fold, dsDNA-phage, icosahedral, VIRUS; 3.8A {Escherichia phage T5}	Hit 6OMC_E	Probability (if > 90%) 99.96	E-value (if < 1) 1e-25
Supporting Information for Function, from Syntney **(SIF-Syn)** (All three lines of SIF evidence must be filled in)	Function (or NKF) major capsid protein	Pham number and number of members 2582 (42)		upstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 138124	Downstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 8234
Transmembrane domains by TmHmm using Phamerator			N/A		

Gene # 15	Original 5’ End Start 8001	Original 3’ End Stop 8417	Original Length 417	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 7947		New Length 471	New Start Codon [✓ ATG] / GTG / TTG

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	Yes	Start:
	[✓ No]	Glimmer Start: 8001
		GeneMark Start:7947

Does this Gene have Coding Potential	[✓ Yes] No	Estimated Start: 7957

Ribosome Binding Site

RBS Final Score -2.845	Is this Score the Closest to 0? [✓ Yes] / No (if no)	Score Closest to 0:
		Start #: 7947
		ORF Length:

Starterator

Starterator Start

7947 has 6 MA and 8001 has none

Function **(F)** Hypothetical Protein	Gene No: 15		Student name: Carlos Rodriguez / Betty Sierra / Grace Intrator		
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 100.00	E-value (if < 10^{-7}) 4e-101	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])		% Identity 100	E-value (if < 10^{-7}) 9e-87	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 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 8234(9)		upstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 2582	Downstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 139922
Transmembrane domains by TmHmm using Phamerator			N/A		

Gene # 16	Original 5’ End Start 8488	Original 3’ End Stop 8982	Original Length 495	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

Glimmer and GeneMark Calls

Does this Gene have Coding Potential	[✓ Yes] No

Start called by both Glimmer and GeneMark?	[✓ Yes] No	Start: 8448
		Glimmer Start:
		GeneMark Start:

Estimated start: 8515

Ribosome Binding Site

RBS Final Score -4.001	Is this Score the Closest to 0? [✓ Yes] / No (if no)	Score Closest to 0:
		Start #:
		ORF Length:

Starterator

Starterator Start

8488 has 19 MA’s with all MA’s being on this start

Function **(F)** Head-to-tail adaptor	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) Head-to-tail adaptor	% Query 99	% Identity 98.78	E-value (if < 10^{-7}) 4e-110	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		% Identity 98	E-value (if < 10^{-7}) 5e-86	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-tail connector	Name / Description) Gp6; 13-membered ring, VIRAL PROTEIN; HET: MPD, MSE; 2.1A {Enterobacteria phage HK97}	Hit 3JVO_E	Probability (if > 90%) 96.5%	E-value (if < 1) 0.029
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	Pham number and number of members 147164(54)		upstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 8234	Downstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 143907

Transmembrane domains by TmHmm using Phamerator	N/A

Gene # 17	Original 5' End Start 8979	Original 3' End Stop 9302	Original Length 324	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		Nev th	New Start Codon ATG / GTG / TTG

Does this Gene have Coding Potential	[✓ Yes] No

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes]	Start: 8979
	No	Glimmer Start:
		GeneMark Start:

Estimated Start: 8950

Ribosome Binding Site

RBS Final Score -4.229	Is this Score the Closest to 0? [✓ Yes] / No (if no)	Score Closest to 0:
		Start #:
		ORF Length:

Starterator

Starterator Start

8979 has 19 MA's with all the MA in this start

Function **(F)** Head to tail stopper	Gene No: 17			Student name: Akram Mikhail Date: 2/13/2024	
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 stopper	% Query 99%	% Identity 99.07%	E-value (if < 10^{-7}) 7e-69	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 stopper		% Identity 99%	E-value (if < 10^{-7}) 2e-56	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 completion protein Contains an alignment to SPP1 16 chain E	Name / Description) Head completion protein gp16; Bacteriophage, SPP1, Portal Protein, Head completion proteins, Connector Complex, DNA Chan	Hit 7z4w_3	Probability (if > 90%) 99.11	E-value (if < 1) 6.9e-10

Supporting Information for Function, from Syntney **(SIF-Syn)** (All three lines of SIF evidence must be filled in)	Function (or NKF) Head to tail stopper	Pham number and number of members 143907(144)	upstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 139922	Downstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 764

Transmembrane domains by TmHmm using Phamerator	N/A

Gene # 18	Original 5’ End Start 9353	Original 3’ End Stop 9685	Original Length 306	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 9299		New Length 360	New Start Codon ATG / [✓ GTG] / TTG

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	Yes [✓ No]	Start: Glimmer Start:9353 GeneMark Start:9299

Estimated Start:9310

Ribosome Binding Site

RBS Final Score -4.256	Is this Score the Closest to 0? [✓ Yes] / No (if no)	Score Closest to 0: Start #: ORF Length:

Starterator

Does this Gene have Coding Potential	[✓ Yes] No

Starterator Start 9299 has 18 MA with all MA in this position

Function **(F)** Hypothetical protein	Gene No: 18		Student name: Akram Mikhail Date: 2/13/2024		
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}) 1e-78	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])		% Identity 100%	E-value (if < 10^{-7}) 4e-60	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 PF111114.12	Probability (if > 90%) 99.67	E-value (if < 1) 3.2e-15
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 764(144)		upstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 142413	Downstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 114826
Transmembrane domains by TmHmm using Phamerator			N/A		

Notes: Some results suggest that this might be a minor capsid protein, however most of the results suggest that it has no known function and the only HHpred result suggesting that this is a minor capsid protein is coming from a PF result.

Gene # 19	Original 5' End Start 9655	Original 3' End Stop 10098	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

New Length	New Start Codon ATG / GTG / TTG

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes] No	Start:9655 Glimmer Start: GeneMark Start:

Estimated Start:9655

Ribosome Binding Site

Does this Gene have Coding Potential? [✓ Yes] No

RBS Final Score -3.447	Is this Score the Closest to 0? [✓ Yes] / No (if no)	Score Closest t Start #: ORF Length:

Starterator

Starterator Start 9655 has 8 MA with all MA in this start.

Function **(F)** Tail terminator	Gene No: 19	Student name: Akram Mikhail Date: 2/13/2024

Supporting Information	Function (or NKF)	% Query 99	% Identity 100	E-value (if < 10^{-7})	AA # conserved in first 10 hits?

for Function, from BLAST (Protein) **(BLAST-NCBI)** (All three lines of SIF evidence must be filled in)	Tail terminator			2e-101	[✓ 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 100	E-value (if < 10^{-7}) 4e-82	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) Tail terminator protein RCC01690	Hit 6TE9-F	Probability (if > 90%) 99.02	E-value (if < 1) 1.6e-8
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 114826(145)		upstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 764	Downstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 85343
Transmembrane domains by TmHmm using Phamerator			N/A		

Gene # 20	Original 5' End Start 10114	Original 3' End Stop 10626	Original Length 513	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

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes]	Start:10114
	No	Glimmer Start:
		GeneMark Start:

Does this Gene have Coding Potential	[✓ Yes] No	Estimated Start: 10170

Ribosome Binding Site

RBS Final Score -2.786	Is this Score the Closest to 0? [✓ Yes] / No (if no)	Score Closest to 0:
		Start #:
		ORF Length:

Starterator

Starterator Start 10114 has 21 MA with all MA in this start.

Function **(F)** Major tail protein	Gene No: 20		Student name: Akram Mikhail Date: 2/13/2024		
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}) 7e-118	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}) 5e-97	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) Phage major tail, TP901-1F	Hit 6TE9-G	Probability (if > 90%) 99.73	E-value (if < 1) 1e-15
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 85343(95)		upstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 114826	Downstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 85343
Transmembrane domains by TmHmm using Phamerator			N/A		

Gene # 21	Original 5’ End Start 10701	Original 3’ End Stop 11189	Original Length 489	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

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes]	Start:10701
	No	Glimmer Start:
		GeneMark Start:

Does this Gene have Coding Potential	[✓ Yes] No	Estimated Start:10698

Ribosome Binding Site

RBS Final Score -6.054	Is this Score the Closest to 0? Yes / [✓ No] (if no)	Score Closest to 0: -2.090
		Start #: 10722
		ORF Length:468

Starterator

Starterator Start 10701 has 5 MA, 4 MA are in start 10722, but this start would cut off too much coding potential.

Function **(F)** Tail assembly chaperone	Gene No: 21		Student name: Akram Mikhail Date: 2/13/2024		
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 99.38	E-value (if < 10^{-7}) 2e-115	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 100	E-value (if < 10^{-7}) 1e-89	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 results 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) Tail assembly chaperone	Pham number and number of members 123060(28)	upstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 85343	Downstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 144294	
Transmembrane domains by TmHmm using Phamerator			N/A		

Gene #22	Original 5’ End Start 11186	Original 3’ End Stop 11476	Original Length 291	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 10701		New Length 776	New Start Codon ATG / [✓ GTG] / TTG

Glimmer and GeneMark Calls

Does this Gene have Coding Potential	Yes No

Start called by both Glimmer and GeneMark?	Yes	Start:N/A
	No	Glimmer Start:N/A
		GeneMark Start:N/A

Estimated Start:N/A

Ribosome Binding Site

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

Starterator

Starterator Start

N/A

Notes: This is where the translational frameshift occur, since it's the gene upstream of the biggest gene in the genome (tape measure protein), therefore it begins where the upstream gene starts which is 10701. The slippery sequence is GGGAAAA and the slippage happens on the first A of that sequence which is nucleotide 11135.

Function **(F)** tail assembly chaperone	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) tail assembly chaperone	% Query 99	% Identity 98.84%	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) tail assembly chaperone		% Identity 99	E-value (if < 10^{-7}) e-149	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 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) tail assembly chaperone	Pham number and number of members 87960 (9)	upstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ?	Downstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ?	

Transmembrane domains by TmHmm using Phamerator	N/A

Annotation Documentation

Phage Name: Bush
Student Name: Betty Sierra/Grace Int.
Date:1/30/24

Gene #	Original 5' End Start	Original 3' End Stop	Original Length	Original Start Codon
23	11501	13699	2199	[✓ ATG] / GTG / TTG

Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?
[✓ FWD] / REV	[✓ Yes] / No (if no, answer below)

Revised 5' End Start	New Length	New Start Codon
		ATG / GTG / TTG

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes] No	Start: 11501 Glimmer Start: 11501 GeneMark Start: 11501

Does this Gene have Coding Potential	[✓ Yes] No	Estimated Start:

Ribosome Binding Site

RBS Final Score	Is this Score the Closest to 0?	
-4.917	Yes / [✓ No] (if no)	Score Closest to 0: -3.594 Start #: 12530 ORF Length:1170

Starterator

Starterator Start	Is the starterator Start called on by other phages in this cluster/subcluster?			
11501	[✓ Yes] / No (if no)	Is the start in this cluster/subcluster Generally conserved?	Yes No	Conserved Start:

Annotation Documentation

Phage Name: Bush
Student Name: Betty Sierra/Grace Int.
Date:1/30/24

Gene #	Original 5’ End Start	Original 3’ End Stop	Original Length	Original Start Codon
24	13699	14559	861	[✓ ATG] / GTG / TTG

Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?
[✓ FWD] / REV	[✓ Yes] / No (if no, answer below)

Revised 5’ End Start	New Length	New Start Codon
		ATG / GTG / TTG

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?		
	[✓ Yes]	Start: 13699
	No	Glimmer Start: 13699 GeneMark Start: 13699

Does this Gene have Coding Potential		
	[✓ Yes] No	Estimated Start:

Ribosome Binding Site

RBS Final Score	Is this Score the Closest to 0?	
-3.819	[✓ Yes] / No (if no)	Score Closest to 0: Start #: ORF Length:

Starterator

Starterator Start	Is the starterator Start called on by other phages in this cluster/subcluster?	Is the start in this cluster/subcluster Generally conserved?		
13699	[✓ Yes] / No (if no)		Yes No	Conserved Start:

Annotation Documentation

Phage Name: Bush
Student Name: Betty Sierra/Grace Int.
Date:1/30/24

Gene #	Original 5' End Start	Original 3' End Stop	Original Length	Original Start Codon
25	14559	16421	1863	[✓ 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

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes] No	Start: 14559 Glimmer Start: 14559 GeneMark Start: 14559

Does this Gene have Coding Potential	[✓ **Yes**] No	Estimated Start:

Ribosome Binding Site

RBS Final Score	Is this Score the Closest to 0?	
-4.957	**Yes** / [✓ No] (if no)	Score Closest to 0: -4.219 Start #: 16413 ORF Length:9

Starterator

Starterator Start	Is the starterator Start called on by other phages in this cluster/subcluster?	Is the start in this cluster/subcluster Generally conserved?		
14559	[✓ Yes] / No (if no)		Yes No	Conserved Start:

Annotation Documentation

Phage Name: Bush
Student Name: Betty Sierra/Grace Int.
Date:1/30/24

Gene # 26	Original 5’ End Start 16428	Original 3’ End Stop 16736	Original Length 309	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

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes] No	Start: 16428 Glimmer Start: 16428 GeneMark Start: 16428

Does this Gene have Coding Potential	[✓ **Yes**] No	Estimated Start:

Ribosome Binding Site

RBS Final Score -3.345	Is this Score the Closest to 0? [✓ **Yes**] / No (if no)	Score Closest to 0: Start #: ORF Length:

Starterator

Starterator Start 16428	Is the starterator Start called on by other phages in this cluster/subcluster? [✓ Yes] / No (if no)	Is the start in this cluster/subcluster Generally conserved?	Yes No	Conserved Start:

Annotation Documentation

Phage Name: Bush
Student Name: Betty Sierra/Grace Int.
Date:1/30/24

Gene #	Original 5' End Start	Original 3' End Stop	Original Length	Original Start Codon
27	16747	17481	735	[✓ ATG] / GTG / TTG

Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?
[✓ FWD] / REV	**[✓ Yes]** / No (if no, answer below)

Revised 5' End Start	New Length	New Start Codon
		ATG / GTG / TTG

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes] No	Start: 16747 Glimmer Start: 16747 GeneMark Start: 16747

Does this Gene have Coding Potential	**[✓ Yes]** No	Estimated Start:

Ribosome Binding Site

RBS Final Score	Is this Score the Closest to 0?	
-5.913	**Yes** / [✓ No] (if no)	Score Closest to 0: -4.205 Start #: 16870 ORF Length: 612

Starterator

Starterator Start	Is the starterator Start called on by other phages in this cluster/subcluster?	Is the start in this cluster/subcluster Generally conserved?		
16747	[✓ Yes] / No (if no)		Yes No	Conserved Start:

Annotation Documentation

Phage Name: Bush
Student Name: Betty Sierra/Grace Int.
Date:1/30/24

Gene # 28	Original 5' End Start 17498	Original 3' End Stop 18319	Original Length 822	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

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes] No	Start: 17498 Glimmer Start: 17498 GeneMark Start: 17498

Does this Gene have Coding Potential	[✓ **Yes**] No	Estimated Start:

Ribosome Binding Site

RBS Final Score -5.500	Is this Score the Closest to 0? **Yes** / [✓ No] (if no)	Score Closest to 0: -4.836 Start #: 18005 ORF Length: 315

Starterator

Starterator Start 17498	Is the starterator Start called on by other phages in this cluster/subcluster? [✓ Yes] / No (if no)	Is the start in this cluster/subcluster Generally conserved?	Yes No	Conserved Start:

Annotation Documentation

Phage Name: Bush
Student Name: Betty Sierra/Grace Int.
Date:1/30/24

Gene #	Original 5' End Start	Original 3' End Stop	Original Length	Original Start Codon
29	18358	18618	261	[✓ ATG] / GTG / TTG

Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?
[✓ FWD] / REV	[✓ Yes] / No (if no, answer below)

Revised 5' End Start	New Length	New Start Codon
		ATG / GTG / TTG

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes] No	Start: 18358 Glimmer Start: 18358 GeneMark Start: 18358

Does this Gene have Coding Potential	[✓ Yes] No	Estimated Start:

Ribosome Binding Site

RBS Final Score	Is this Score the Closest to 0?	
-5.913	Yes / [✓ No] (if no)	Score Closest to 0: -5.115 Start #: 18316 ORF Length: 303

Starterator

Starterator Start	Is the starterator Start called on by other phages in this cluster/subcluster?	Is the start in this cluster/subcluster Generally conserved?		
18358	[✓ Yes] / No (if no)		Yes No	Conserved Start:

Annotation Documentation

Phage Name: Bush
Student Name: Arianna Marie Ruiz
Date: 1/23/2024

Gene # 30	Original 5' End Start 18611	Original 3' End Stop 19504	Original Length 894	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

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes] No	Start: 18611 Glimmer Start: GeneMark Start:

Does this Gene have Coding Potential	[✓ Yes] No	Estimated Start: 18611

Ribosome Binding Site

RBS Final Score -3.948	Is this Score the Closest to 0? Yes / [✓ No] (if no)	Score Closest to 0: -2.837 Start #: 19142 ORF Length: 363

Starterator

Starterator Start 18611	Is the starterator Start called on by other phages in this cluster/subcluster? [✓ Yes] / No (if no)	Is the start in this cluster/subcluster Generally conserved?	Yes No	Conserved Start:

Annotation Documentation

Phage Name: Bush
Student Name: Arianna Marie Ruiz
Date: 1/23/2024

Gene # 31	Original 5’ End Start 19563	Original 3’ End Stop 19877	Original Length 315	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

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes]	Start: 19563
	No	Glimmer Start:
		GeneMark Start:

Does this Gene have Coding Potential	[✓ Yes] No	Estimated Start: 19563

Ribosome Binding Site

RBS Final Score -3.749	Is this Score the Closest to 0? [✓ Yes] / No (if no)	Score Closest to 0:
		Start #:
		ORF Length:

Starterator

Starterator Start 19563	Is the starterator Start called on by other phages in this cluster/subcluster? [✓ Yes] / No (if no)	Is the start in this cluster/subcluster Generally conserved?	Yes No	Conserved Start:

Annotation Documentation

Phage Name: Bush
Student Name: Arianna Marie Ruiz
Date: 1/30/2024

Gene # 32	Original 5’ End Start 19874	Original 3’ End Stop 20539	Original Length 666	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

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	Yes [✓ No]	Start: Glimmer Start: 19874 GeneMark Start: 20539

Does this Gene have Coding Potential	[✓ Yes] No	Estimated Start:

Ribosome Binding Site

RBS Final Score -3.977	Is this Score the Closest to 0? Yes / [✓ No] (if no)	Score Closest to 0: -3.031 Start #: 20123 ORF Length: 417

Starterator

Starterator Start 19874	Is the starterator Start called on by other phages in this cluster/subcluster? [✓ Yes] / No (if no)	Is the start in this cluster/subcluster Generally conserved?	Yes No	Conserved Start:

Annotation Documentation

Phage Name: Bush
Student Name: Arianna Marie Ruiz
Date: 1/30/2024

Gene #	Original 5' End Start	Original 3' End Stop	Original Length	Original Start Codon
33	20554	21117	564	[✓ ATG] / GTG / TTG

Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?
[✓ FWD] / REV	Yes / [✓ No] (if no, answer below)

Revised 5' End Start	New Length	New Start Codon
20536	582	ATG / [✓ GTG] / TTG

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes]	Start: 20554
	No	Glimmer Start:
		GeneMark Start:

Does this Gene have Coding Potential	[✓ Yes] No	Estimated Start: 20536

Ribosome Binding Site

RBS Final Score	Is this Score the Closest to 0?	
-8.194	Yes / [✓ No] (if no)	Score Closest to 0: -2.473
		Start #: 20968
		ORF Length: 150

Starterator

Starterator Start	Is the starterator Start called on by other phages in this cluster/subcluster?
20536	[✓ Yes] / No (if no)

Is the start in this cluster/subcluster Generally conserved?	Yes No	Conserved Start:

Comments

The RBS score for start position 20536 is very high at -8.194, however when the Blast alignment is observed, the protein expresses 7 nucleotides are missing, accounting for ~20bp missing from the original call.

Annotation Documentation

Phage Name: Bush
Student Name: Arianna Marie Ruiz
Date: 1/30/2024

Gene #	Original 5’ End Start	Original 3’ End Stop	Original Length	Original Start Codon
34	21176	21301	126	ATG / [✓ GTG] / TTG

Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?
[✓ FWD] / REV	Yes / No (if no, answer below)

Revised 5’ End Start	New Length	New Start Codon
		ATG / GTG / TTG

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	Yes	Start:
	[✓ No]	Glimmer Start: 21176
		GeneMark Start: N/A

Does this Gene have Coding Potential	[✓ Yes]	Estimated Start:
	No	

Ribosome Binding Site

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

Starterator

Starterator Start	Is the starterator Start called on by other phages in this cluster/subcluster?
	Yes / No (if no)

Is the start in this cluster/subcluster Generally conserved?	Yes	Conserved Start:
	No	

Annotation Documentation

Phage Name: Bush
Student Name: Arianna Marie Ruiz
Date: 1/30/2024

Gene #	Original 5’ End Start	Original 3’ End Stop	Original Length	Original Start Codon
35	21302	21625	324	ATG / [✓ GTG] / TTG
Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?			
[✓ FWD] / REV	Yes / [✓ No] (if no, answer below)			
	Revised 5’ End Start		New Length	New Start Codon
	21326		300	[✓ ATG] / GTG / TTG

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	Yes	Start:
	[✓ No]	Glimmer Start: 21302
		GeneMark Start: 21326

Does this Gene have Coding Potential	[✓ Yes]	Estimated Start: 21326
	No	

Ribosome Binding Site

RBS Final Score	Is this Score the Closest to 0?	Score Closest to 0:
-3.463	[✓ Yes] / No (if no)	Start #:
		ORF Length:

Starterator

Starterator Start	Is the starterator Start called on by other phages in this cluster/subcluster?
21326	[✓ Yes] / No (if no)

Is the start in this cluster/subcluster Generally conserved?	Yes	Conserved Start:
	No	

Comments

According to GeneMaster, the coding potential begins around the range of ~21320, this accompanied by the “best” RBS score of -3.463 of the potential starts. Compared to the original start assumption of 21302, which has an RBS score of -5.917, the furthest from 0, and no coding potential, should prompt the change of the start position from 21302 to 2126.

2 people called on start 5 @21302, 1 person called on start 6 @21326

Annotation Documentation

Phage Name: Bush
Student Name: Arianna Marie Ruiz
Date: 1/30/2024

Gene #	Original 5’ End Start	Original 3’ End Stop	Original Length	Original Start Codon
36	21622	22062	441	[✓ ATG] / GTG / TTG

Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?
[✓ FWD] / REV	Yes / No (if no, answer below)

Revised 5’ End Start	New Length	New Start Codon
		ATG / GTG / TTG

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes] No	Start: 21622 Glimmer Start: GeneMark Start:

Does this Gene have Coding Potential	Yes No	Estimated Start:

Ribosome Binding Site

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

Starterator

Starterator Start	Is the starterator Start called on by other phages in this cluster/subcluster? Yes / No (if no)

Is the start in this cluster/subcluster Generally conserved?	Yes No	Conserved Start:

Comments

No data

Annotation Documentation

Phage Name: Bush
Student Name: Arianna Marie Ruiz
Date: 1/30/2024

Gene #	Original 5' End Start	Original 3' End Stop	Original Length	Original Start Codon
37	22080	22526	447	[✓ ATG] / GTG / TTG

Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?
[✓ FWD] / REV	[✓ Yes] / No (if no, answer below)

Revised 5' End Start	New Length	New Start Codon
		ATG / GTG / TTG

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes]	Start: 22080
	No	Glimmer Start: GeneMark Start:

Does this Gene have Coding Potential	[✓ Yes] No	Estimated Start: 22080

Ribosome Binding Site

RBS Final Score	Is this Score the Closest to 0?	
-5.345	Yes / [✓ No] (if no)	Score Closest to 0: -3.734 Start #: 22434 ORF Length: 93

Starterator

Starterator Start	Is the starterator Start called on by other phages in this cluster/subcluster?	Is the start in this cluster/subcluster Generally conserved?		
22080	[✓ Yes] / No (if no)		Yes No	Conserved Start:

Annotation Documentation

Phage Name: Bush
Student Name: Arianna Marie Ruiz
Date: 1/30/2024

Gene # 38	Original 5’ End Start 22523	Original 3’ End Stop 22681	Original Length 159	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

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes]	Start: 22523
	No	Glimmer Start:
		GeneMark Start:

Does this Gene have Coding Potential	[✓ Yes]	Estimated Start: 22523
	No	

Ribosome Binding Site

RBS Final Score -6.066	Is this Score the Closest to 0? Yes / [✓ No] (if no)	Score Closest to 0: -4.150
		Start #: 22637
		ORF Length: 45

Starterator

Starterator Start 22523	Is the starterator Start called on by other phages in this cluster/subcluster? [✓ Yes] / No (if no)	Is the start in this cluster/subcluster Generally conserved?	Yes No	Conserved Start:

Annotation Documentation

Phage Name: Bush
Student Name: Arianna Marie Ruiz
Date: 1/30/2024

Gene #	Original 5' End Start	Original 3' End Stop	Original Length	Original Start Codon
40	23256	24347	1092	[✓ ATG] / GTG / TTG

Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?
[✓ FWD] / REV	[✓ Yes] / No (if no, answer below)

Revised 5' End Start	New Length	New Start Codon
		ATG / GTG / TTG

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes]	Start: 23256
	No	Glimmer Start:
		GeneMark Start:

Does this Gene have Coding Potential	[✓ Yes] No	Estimated Start:

Ribosome Binding Site

RBS Final Score	Is this Score the Closest to 0?	
-2.812	[✓ Yes] / No (if no)	Score Closest to 0:
		Start #:
		ORF Length:

Starterator

Starterator Start	Is the starterator Start called on by other phages in this cluster/subcluster?
23256	[✓ Yes] / No (if no)

Is the start in this cluster/subcluster Generally conserved?	Yes No	Conserved Start:

Annotation Documentation

Phage Name: Bush
Student Name: Arianna Marie Ruiz
Date: 1/30/2024

Gene # 41	Original 5' End Start 24344	Original 3' End Stop 25393	Original Length 1050	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

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes]	Start: 24344
	No	Glimmer Start:
		GeneMark Start:

Does this Gene have Coding Potential	[✓ Yes]	Estimated Start:
	No	

Ribosome Binding Site

RBS Final Score -2.812	Is this Score the Closest to 0? [✓ Yes] / No (if no)	Score Closest to 0:
		Start #:
		ORF Length:

Starterator

Starterator Start 24344	Is the starterator Start called on by other phages in this cluster/subcluster? [✓ Yes] / No (if no)	Is the start in this cluster/subcluster Generally conserved?	Yes No	Conserved Start:

Annotation Documentation

Phage Name: Bush
Student Name: Arianna Marie Ruiz
Date: 1/31/2024

Gene #	Original 5' End Start	Original 3' End Stop	Original Length	Original Start Codon
42	25393	25755	363	[✓ ATG] / GTG / TTG

Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?
[✓ FWD] / REV	[✓ Yes] / No (if no, answer below)

Revised 5' End Start	New Length	New Start Codon
		ATG / GTG / TTG

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes]	Start: 25393
	No	Glimmer Start: GeneMark Start:

Does this Gene have Coding Potential	[✓ Yes] No	Estimated Start: 25393

Ribosome Binding Site

RBS Final Score	Is this Score the Closest to 0?	
-5.721	Yes / [✓ No] (if no)	Score Closest to 0: -5.082 Start #: 25741 ORF Length: 15

Starterator

Starterator Start	Is the starterator Start called on by other phages in this cluster/subcluster?			
25393	[✓ Yes] / No (if no)	Is the start in this cluster/subcluster Generally conserved?	Yes No	Conserved Start:

Annotation Documentation

Phage Name: Bush
Student Name: Arianna Marie Ruiz
Date: 1/31/2024

Gene #	Original 5' End Start	Original 3' End Stop	Original Length	Original Start Codon
43	25748	26125	378	ATG / [✓ GTG] / TTG
Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?			
[✓ FWD] / REV	[✓ Yes] / No (if no, answer below)			
	Revised 5' End Start		New Length	New Start Codon
				ATG / GTG / TTG

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes]	Start: 25748
	No	Glimmer Start:
		GeneMark Start:

Does this Gene have Coding Potential	[✓ Yes]	Estimated Start: 25748
	No	

Ribosome Binding Site

RBS Final Score	Is this Score the Closest to 0?	Score Closest to 0: -4.687
-4.957	Yes / [✓ No] (if no)	Start #: 26039
		ORF Length: 87

Starterator

Starterator Start	Is the starterator Start called on by other phages in this cluster/subcluster?	Is the start in this cluster/subcluster Generally conserved?	Yes	Conserved Start:
25748	[✓ Yes] / No (if no)		No	

Annotation Documentation

Phage Name: Bush
Student Name: Kristen Mclean
Date: 1/30/2024

Gene #	Original 5’ End Start	Original 3’ End Stop	Original Length	Original Start Codon
43	25748	26125	378	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

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes]	Start: 25748
	No	Glimmer Start:
		GeneMark Start:

Does this Gene have Coding Potential	[✓ Yes] No	Estimated Start: 25748

Ribosome Binding Site

RBS Final Score	Is this Score the Closest to 0?	Score Closest to 0: -5.434
-5.435	Yes / [✓ No] (if no)	Start #: 26039
		ORF Length: 87

Starterator

Starterator Start	Is the starterator Start called on by other phages in this cluster/subcluster?	Is the start in this cluster/subcluster Generally conserved?	Yes No	Conserved Start:
25748	[✓ Yes] / No (if no)			

Annotation Documentation

Phage Name: Bush
Student Name: Kristen Mclean
Date: 1/30/2024

Gene #	Original 5’ End Start	Original 3’ End Stop	Original Length	Original Start Codon
44	26151	26984	834	[✓ ATG] / GTG / TTG

Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?
[✓ FWD] / REV	[✓ Yes] / No (if no, answer below)

Revised 5’ End Start	New Length	New Start Codon
		ATG / GTG / TTG

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes]	Start: 26151
	No	Glimmer Start: GeneMark Start:

Does this Gene have Coding Potential	[✓ Yes] No	Estimated Start: 26151

Ribosome Binding Site

RBS Final Score	Is this Score the Closest to 0?	
-3.469	[✓ Yes] / No (if no)	Score Closest to 0: Start #: ORF Length:

Starterator

Starterator Start	Is the starterator Start called on by other phages in this cluster/subcluster?
26151	[✓ Yes] / No (if no)

Is the start in this cluster/subcluster Generally conserved?	Yes No	Conserved Start:

Annotation Documentation

Phage Name: Bush
Student Name: Kristen Mclean
Date: 1/30/2024

Gene #	Original 5’ End Start	Original 3’ End Stop	Original Length	Original Start Codon
45	26981	27319	339	[✓ ATG] / GTG / TTG

Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?
[✓ FWD] / REV	[✓ Yes] / No (if no, answer below)

Revised 5’ End Start	New Length	New Start Codon
		ATG / GTG / TTG

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes]	Start: 26981
	No	Glimmer Start:
		GeneMark Start:

Does this Gene have Coding Potential	[✓ Yes] No	Estimated Start: 26981

Ribosome Binding Site

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

Starterator

Starterator Start	Is the starterator Start called on by other phages in this cluster/subcluster?	Is the start in this cluster/subcluster Generally conserved?		
26981	[✓ Yes] / No (if no) →		Yes No	Conserved Start:

Annotation Documentation

Phage Name: Bush
Student Name: Kristen Mclean
Date: 1/30/2024

Gene #	Original 5' End Start	Original 3' End Stop	Original Length	Original Start Codon
46	27316	28620	1305	ATG / [✓ GTG] / TTG

Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?
[✓ FWD] / REV	[✓ Yes] / No (if no, answer below)

Revised 5' End Start	New Length	New Start Codon
		ATG / GTG / TTG

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	Yes / No	
	[✓ Yes]	Start: 27316
	No	Glimmer Start:
		GeneMark Start:

Does this Gene have Coding Potential	Yes / No	
	[✓ Yes]	Estimated Start: 27316
	No	

Ribosome Binding Site

RBS Final Score	Is this Score the Closest to 0?	
-5.266	Yes / [✓ No] (if no)	Score Closest to 0: -2.918
		Start #: 27655
		ORF Length: 966

Starterator

Starterator Start	Is the starterator Start called on by other phages in this cluster/subcluster?	Is the start in this cluster/subcluster Generally conserved?	Yes / No	
27316	[✓ Yes] / No (if no)		Yes No	Conserved Start:

Annotation Documentation

Phage Name: Bush
Student Name: Kristen Mclean
Date: 1/30/2024

Gene #	Original 5' End Start	Original 3' End Stop	Original Length	Original Start Codon
47	28681	29172	492	[✓ ATG] / GTG / TTG

Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?
[✓ FWD] / REV	[✓ Yes] / No (if no, answer below)

Revised 5' End Start	New Length	New Start Codon
		ATG / GTG / TTG

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?		
[✓ Yes]	Start: 28681	
No	Glimmer Start:	GeneMark Start:

Does this Gene have Coding Potential		
[✓ Yes] / No	Estimated Start: 28681	

Ribosome Binding Site

RBS Final Score	Is this Score the Closest to 0?	
-2.548	[✓ Yes] / No (if no)	Score Closest to 0: Start #: ORF Length:

Starterator

Starterator Start	Is the starterator Start called on by other phages in this cluster/subcluster?	Is the start in this cluster/subcluster Generally conserved?		
28681	[✓ Yes] / No (if no)		Yes No	Conserved Start:

Annotation Documentation

Phage Name: Bush
Student Name: Kristen Mclean
Date: 1/30/2024

Gene # 48	Original 5' End Start 29269	Original 3' End Stop 29802	Original Length 534	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

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes] No	Start: 29269 Glimmer Start: GeneMark Start:

Does this Gene have Coding Potential	[✓ Yes] No	Estimated Start: 29269

Ribosome Binding Site

RBS Final Score -7.028	Is this Score the Closest to 0? Yes / [✓ No] (if no)	Score Closest to 0: -4.825 Start #: 29536 ORF Length: 267

Starterator

Starterator Start 29269	Is the starterator Start called on by other phages in this cluster/subcluster? [✓ Yes] / No (if no)	Is the start in this cluster/subcluster Generally conserved?	Yes No	Conserved Start:

Annotation Documentation

Phage Name: Bush
Student Name: Kristen Mclean
Date: 1/30/2024

Gene #	Original 5’ End Start	Original 3’ End Stop	Original Length	Original Start Codon
49	29795	30421	627	ATG / [✓ GTG] / TTG

Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?
[✓ FWD] / REV	[✓ Yes]/ No (if no, answer below)

Revised 5’ End Start	New Length	New Start Codon
		ATG / GTG / TTG

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?		
	[✓ Yes]	Start: 29795
	No	Glimmer Start: GeneMark Start:

Does this Gene have Coding Potential		
	[✓ Yes] No	Estimated Start: 29795

Ribosome Binding Site

RBS Final Score	Is this Score the Closest to 0?	
-4.323	[✓ Yes]/ No (if no)	Score Closest to 0: Start #: ORF Length:

Starterator

Starterator Start	Is the starterator Start called on by other phages in this cluster/subcluster?	Is the start in this cluster/subcluster Generally conserved?		
29795	[✓ Yes]/ No (if no)		Yes No	Conserved Start:

Annotation Documentation

Phage Name: Bush
Student Name: Kristen Mclean
Date: 1/30/2024

Gene # 50	Original 5' End Start 30579	Original 3' End Stop 30905	Original Length 327	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

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes]	Start: 30579
	No	Glimmer Start:
		GeneMark Start:

Does this Gene have Coding Potential	[✓ Yes] No	Estimated Start: 30579

Ribosome Binding Site

RBS Final Score -2.584	Is this Score the Closest to 0? [✓ Yes] / No (if no)	Score Closest to 0:
		Start #:
		ORF Length:

Starterator

Starterator Start 30579	Is the starterator Start called on by other phages in this cluster/subcluster? [✓ Yes] / No (if no)	Is the start in this cluster/subcluster Generally conserved?	Yes No	Conserved Start:

Annotation Documentation

Phage Name: Bush
Student Name: Kristen Mclean
Date: 1/30/2024

Gene #	Original 5' End Start	Original 3' End Stop	Original Length	Original Start Codon
52	32153	32530	378	[✓ ATG] / GTG / TTG

Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?
[✓ FWD] / REV	[✓ Yes] / No (if no, answer below)

Revised 5' End Start	New Length	New Start Codon
		ATG / GTG / TTG

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?		
	[✓ Yes]	Start: 32153
	No	Glimmer Start: GeneMark Start:

Does this Gene have Coding Potential		
	[✓ Yes]	Estimated Start: 32153
	No	

Ribosome Binding Site

RBS Final Score	Is this Score the Closest to 0?	
-3.432	[✓ Yes] / No (if no)	Score Closest to 0: Start #: ORF Length:

Starterator

Starterator Start	Is the starterator Start called on by other phages in this cluster/subcluster?
32153	[✓ Yes] / No (if no)

Is the start in this cluster/subcluster Generally conserved?		
	Yes No	Conserved Start:

*Annotation Documentation

Phage Name: Bush
Student Name: Kristen Mclean
Date: 1/30/2024

Gene #	Original 5’ End Start	Original 3’ End Stop	Original Length	Original Start Codon
53	32542	32817	276	[✓ ATG] / GTG / TTG

Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?
[✓ FWD] / REV	[✓ Yes] / No (if no, answer below)

Revised 5’ End Start	New Length	New Start Codon
		ATG / GTG / TTG

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	Yes	Start: 32542
	[✓ No]	Glimmer Start: 32542 GeneMark Start:

Does this Gene have Coding Potential	[✓ Yes] No	Estimated Start: 32510

*Not reported in GeneMark

Ribosome Binding Site

RBS Final Score -6.655	Is this Score the Closest to 0? Yes [✓ No] (if no)	Score Closest to 0: -5.280 Start #: 32608 ORF Length:210

Starterator

Starterator Start 32542	Is the starterator Start called on by other phages in this cluster/subcluster? [✓ Yes] / No (if no)	Is the start in this cluster/subcluster Generally conserved?	Yes No	Conserved Start:

Annotation Documentation

Phage Name: Bush
Student Name: Kristen Mclean
Date: 1/30/2024

Gene # 54	Original 5’ End Start 32814	Original 3’ End Stop 32957	Original Length 144	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

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes]	Start: 32814
	No	Glimmer Start:
		GeneMark Start:

Does this Gene have Coding Potential	[✓ Yes] No	Estimated Start: 32814

Ribosome Binding Site

RBS Final Score -6.747	Is this Score the Closest to 0? Yes / [✓ No] (if no)	Score Closest to 0: -5.040
		Start #: 32058
		ORF Length: 873

Starterator

Starterator Start 32814	Is the starterator Start called on by other phages in this cluster/subcluster? [✓ Yes] / No (if no)	Is the start in this cluster/subcluster Generally conserved?	Yes No	Conserved Start:

Annotation Documentation

Phage Name: Bush
Student Name: Kristen Mclean
Date: 1/30/2024

Gene #	Original 5’ End Start	Original 3’ End Stop	Original Length	Original Start Codon
55	32954	33304	351	[✓ ATG] / GTG / TTG

Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?
[✓ FWD] / REV	[✓ Yes] / No (if no, answer below)

Revised 5’ End Start	New Length	New Start Codon
		ATG / GTG / TTG

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?		
	[✓ Yes]	Start:
	No	Glimmer Start:
		GeneMark Start:

Does this Gene have Coding Potential		
	[✓ Yes]	Estimated Start: 32954
	No	

Ribosome Binding Site

RBS Final Score	Is this Score the Closest to 0?	
-5.533	Yes / [✓ No] (if no)	Score Closest to 0: -5.345
		Start #: 32618
		ORF Length: 687

Starterator

Starterator Start	Is the starterator Start called on by other phages in this cluster/subcluster?
32954	[✓ Yes] / No (if no)

Is the start in this cluster/subcluster Generally conserved?		
	Yes	Conserved Start:
	No	

Annotation Documentation

Phage Name: Bush
Student Name: Nicole Gonzalez G.
Date:1/23/24

Gene # 56	Original 5' End Start 33277	Original 3' End Stop 33627	Original Length 351	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 33301		New Length 327	New Start Codon [✓ ATG] / GTG / TTG

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	Yes [✓ No]	Start: Glimmer Start: 33277 GeneMark Start:33301

Does this Gene have Coding Potential	[✓ Yes] No	Estimated Start: 33301

Ribosome Binding Site

RBS Final Score -7.292	Is this Score the Closest to 0? Yes / [✓ No] (if no)	Score Closest to 0: -3.332 Start #: 33166 ORF Length: 462

Starterator

Starterator Start 33301	Is the starterator Start called on by other phages in this cluster/subcluster? [✓ Yes] / No (if no)	Is the start in this cluster/subcluster Generally conserved?	Yes No	Conserved Start:

Comments: Starting Codon (@bp33277) was revised to (@bp33301) in accordance with the following: The starting codon (@bp33277) was called by DNA Master for Gene 56, however, this does not have coding potential when looking at GeneMark and was not called by other phages in Starterator. The revised starting codon of (@bp33301) does have coding potential when looking at GeneMark and was called upon by 7 other MA's in Starterator.

Annotation Documentation

Phage Name: Bush
Student Name: Nicole Gonzalez G.
Date: 1/30/24

Gene #	Original 5' End Start	Original 3' End Stop	Original Length	Original Start Codon
57	33690	34085	396	[✓ 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

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes]	Start:33690
	No	Glimmer Start:
		GeneMark Start:

Does this Gene have Coding Potential	[✓ Yes]	Estimated Start:33690
	No	

Ribosome Binding Site

RBS Final Score	Is this Score the Closest to 0?	
-4.408	Yes / [✓ No] (if no)	Score Closest to 0: -4.390
		Start #: 34074
		ORF Length: 12

Starterator

Starterator Start	Is the starterator Start called on by other phages in this cluster/subcluster?
33690	[✓ Yes] / No (if no)

Is the start in this cluster/subcluster Generally conserved?	Yes No	Conserved Start:

Annotation Documentation

Phage Name: Bush
Student Name: Nicole Gonzalez G.
Date: 1/30/24

Gene #	Original 5' End Start	Original 3' End Stop	Original Length	Original Start Codon
58	34082	34519	438	[✓ 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

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes]	Start:34082
	No	Glimmer Start:
		GeneMark Start:

Does this Gene have Coding Potential	[✓ Yes] No	Estimated Start: 34082

Ribosome Binding Site

RBS Final Score	Is this Score the Closest to 0?	Score Closest to 0:
-4.487	[✓ Yes] / No (if no)	Start #:
		ORF Length:

Starterator

Starterator Start	Is the starterator Start called on by other phages in this cluster/subcluster?	Is the start in this cluster/subcluster Generally conserved?	Yes No	Conserved Start:
34082	[✓ Yes] / No (if no)			

Annotation Documentation

Phage Name: Bush
Student Name: Nicole Gonzalez G.
Date:1/30/24

Gene # 59	Original 5’ End Start 34494	Original 3’ End Stop 35009	Original Length 516	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

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes]	Start:34494
	No	Glimmer Start:
		GeneMark Start:

Does this Gene have Coding Potential	[✓ Yes]	Estimated Start:
	No	

Ribosome Binding Site

RBS Final Score -2.373	Is this Score the Closest to 0? [✓ Yes] / No (if no)	Score Closest to 0:
		Start #:
		ORF Length:

Starterator

Starterator Start 34494	Is the starterator Start called on by other phages in this cluster/subcluster? [✓ Yes] / No (if no)	Is the start in this cluster/subcluster Generally conserved?	Yes No	Conserved Start:

Annotation Documentation

Phage Name: Bush
Student Name: Nicole Gonzalez G.
Date: 1/30/24

Gene # 60	Original 5' End Start 35006	Original 3' End Stop 35482	Original Length 477	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

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes] No	Start: 35006 Glimmer Start: GeneMark Start:

Does this Gene have Coding Potential	[✓ Yes] No	Estimated Start: 35006

Ribosome Binding Site

RBS Final Score -5.589	Is this Score the Closest to 0? Yes / [✓ No] (if no)	Score Closest to 0: -3.807 Start #: 35192 ORF Length:291

Starterator

Starterator Start 35006	Is the starterator Start called on by other phages in this cluster/subcluster? [✓ Yes] / No (if no)	Is the start in this cluster/subcluster Generally conserved?	Yes No	Conserved Start:

Annotation Documentation

Phage Name: Bush
Student Name: Nicole Gonzalez G.
Date: 1/30/24

Gene #	Original 5’ End Start	Original 3’ End Stop	Original Length	Original Start Codon
61	35499	36332	834	[✓ 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

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes]	Start:35499
	No	Glimmer Start:
		GeneMark Start:

Does this Gene have Coding Potential	[✓ Yes]	Estimated Start:
	No	

Ribosome Binding Site

RBS Final Score	Is this Score the Closest to 0?	Score Closest to 0:
-2.661	[✓ Yes] / No (if no)	Start #:
		ORF Length:

Starterator

Starterator Start	Is the starterator Start called on by other phages in this cluster/subcluster?	Is the start in this cluster/subcluster Generally conserved?	Yes	Conserved Start:
35499	[✓ Yes] / No (if no)		No	

Annotation Documentation

Phage Name: Bush
Student Name: Nicole Gonzalez G.
Date:1/30/24

Gene #	Original 5' End Start	Original 3' End Stop	Original Length	Original Start Codon
62	36329	37051	723	[✓ ATG] / GTG / TTG

Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?
[✓ FWD] / REV	[✓ Yes] / No (if no, answer below)

Revised 5' End Start	New Length	New Start Codon
		ATG / GTG / TTG

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes] No	Start:36329 Glimmer Start: GeneMark Start:

Does this Gene have Coding Potential	[✓ Yes] No	Estimated Start:36329

Ribosome Binding Site

RBS Final Score	Is this Score the Closest to 0?	Score Closest to 0: -4.716
-4.844	Yes / [✓ No] (if no)	Start #: 36683 ORF Length:369

Starterator

Starterator Start	Is the starterator Start called on by other phages in this cluster/subcluster?	Is the start in this cluster/subcluster Generally conserved?		
36329	[✓ Yes] / No (if no)		Yes No	Conserved Start:

Annotation Documentation

Phage Name: Bush
Student Name: Nicole Gonzalez G.
Date:1/30/24

Gene # 63	Original 5' End Start 37036	Original 3' End Stop 37386	Original Length 351	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

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes]	Start:37036
	No	Glimmer Start:
		GeneMark Start:

Does this Gene have Coding Potential	[✓ Yes] No	Estimated Start:37036

Ribosome Binding Site

RBS Final Score	Is this Score the Closest to 0?	Score Closest to 0: -2.661
-6.094	Yes / [✓ No] (if no)	Start #: 37078
		ORF Length:309

Starterator

Starterator Start 37036	Is the starterator Start called on by other phages in this cluster/subcluster? [✓ Yes] / No (if no)	Is the start in this cluster/subcluster Generally conserved?	Yes No	Conserved Start:

Annotation Documentation

Phage Name: Bush
Student Name: Nicole Gonzalez G.
Date: 2/01/24

Gene #	Original 5' End Start	Original 3' End Stop	Original Length	Original Start Codon
64	37386	37856	471	[✓ ATG] / GTG / TTG

Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?
[✓ FWD] / REV	[✓ Yes] / No (if no, answer below)

Revised 5' End Start	New Length	New Start Codon
		ATG / GTG / TTG

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?		
	[✓ Yes]	Start:37386
	No	Glimmer Start:
		GeneMark Start:

Does this Gene have Coding Potential		
	[✓ Yes] No	Estimated Start:37386

Ribosome Binding Site

RBS Final Score	Is this Score the Closest to 0?	
-4.319	[✓ Yes] / No (if no)	Score Closest to 0:
		Start #:
		ORF Length:

Starterator

Starterator Start	Is the starterator Start called on by other phages in this cluster/subcluster?
37386	[✓ Yes] / No (if no)

Is the start in this cluster/subcluster Generally conserved?		
	Yes No	Conserved Start:

Annotation Documentation

Phage Name: Bush
Student Name: Nicole Gonzalez G.
Date: 2/01/24

Gene #	Original 5' End Start	Original 3' End Stop	Original Length	Original Start Codon
65	37906	38307	402	[✓ ATG] / GTG / TTG

Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?
[✓ FWD] / REV	Yes / [✓ No] (if no, answer below)

Revised 5' End Start	New Length	New Start Codon
37894	414	[✓ ATG] / GTG / TTG

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	Yes [✓ No]	Start: Glimmer Start: 37906 GeneMark Start: 37894

Does this Gene have Coding Potential	[✓ Yes] No	Estimated Start: 37894

Ribosome Binding Site

RBS Final Score	Is this Score the Closest to 0?	
-6.433	Yes / [✓ No] (if no)	Score Closest to 0: -4.558 Start #: 37894 ORF Length: 414

Starterator

Starterator Start	Is the starterator Start called on by other phages in this cluster/subcluster?			
	Yes / No (if no)	Is the start in this cluster/subcluster Generally conserved?	Yes No	Conserved Start:

Comments: There was no starterator report for this gene. There were two possible starting codons for this gene, the first starting codon (@bp37906) and the second starting codon (@bp37894).The reason for the revised starting codon to be at (@bp37894) was due to, there being a lower RBS score, there is more length to the chain, and it closes more of the gap between this gene and gene 64.

Annotation Documentation

Phage Name: Bush
Student Name: Nicole Gonzalez G.
Date: 2/01/24

Gene # 66	Original 5' End Start 38316	Original 3' End Stop 38591	Original Length 276	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

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	Yes [✓ No]	Start: Glimmer Start:38316 GeneMark Start:38304

Does this Gene have Coding Potential	[✓ Yes] No	Estimated Start:38316

Ribosome Binding Site

RBS Final Score -5.666	Is this Score the Closest to 0? [✓ Yes] / No (if no)	Score Closest to 0: Start #: ORF Length:

Starterator

Starterator Start 38316	Is the starterator Start called on by other phages in this cluster/subcluster? Yes / [✓ No] (if no)	Is the start in this cluster/subcluster Generally conserved?	Yes No	Conserved Start:

Comments: Although there was two different starting codons. The starting codon that was originally called by Glimmer was the deciding start codon to stay with due to that codon having a lower RBS score than the other.

Annotation Documentation

Phage Name: Bush
Student Name: Nicole Gonzalez G.
Date: 2/01/24

Gene #	Original 5’ End Start	Original 3’ End Stop	Original Length	Original Start Codon
67	38588	38815	228	[✓ ATG] / GTG / TTG

Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?
[✓ FWD] / REV	[✓ Yes] / No (if no, answer below)

Revised 5’ End Start	New Length	New Start Codon
		ATG / GTG / TTG

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	[✓ Yes]	Start:38588
	No	Glimmer Start: GeneMark Start:

Does this Gene have Coding Potential	[✓ Yes] No	Estimated Start:38588

Ribosome Binding Site

RBS Final Score	Is this Score the Closest to 0?	
-5.485	[✓ Yes] / No (if no)	Score Closest to 0: Start #: ORF Length:

Starterator

Starterator Start	Is the starterator Start called on by other phages in this cluster/subcluster?			
38588	[✓ Yes] / No (if no)	Is the start in this cluster/subcluster Generally conserved?	Yes No	Conserved Start:

Function **(F)** Tape measure protein	Gene No: 23		Student name: Betty Sierra / Grace Intrator Date: 2/13/24		
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	% Query 99	% Identity 92.62	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		% Identity 93	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	Name / Descr (incl DUF) Gp57 tape measure protein	Hit 6V8I_BF	Probability (if > 90%) 99.89	E-value (if < 1) 7.8e-16
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 1414009 (31)	upstream pham conserved ? Yes or [✓ No] If yes, what pham # or function ? 134673	Downstream pham conserved ? Yes or [✓ No] If yes, what pham # or function? 140921	
Transmembrane domains by TmHmm using Phamerator			no		

Notes: Most evidence suggests that it is a Tape measure protein, given that only phamerator leads to believe it is a minor tail protein.

Function **(F)** Minor tail protein	Gene No: 24		Student name: Betty Sierra / Grace Intrator Date: 2/13/24		
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 99.65	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		% Identity 99	E-value (if < 10^{-7}) e-162	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) Hypothetical protein	Name / Descr (incl DUF) Hypothetical protein 19.1	Hit 2X8K_C	Probability (if > 90%) 99.95	E-value (if < 1) 8.8e-26
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 140921 (339)		upstream pham conserved ? Yes or [✓ No] If yes, what pham # or function ? 141409	Downstream pham conserved ? Yes or [✓ No] If yes, what pham # or function? 142684
Transmembrane domains by TmHmm using Phamerator			no		

Notes: Most evidence leads to it being a minor tail protein given that only HHpred says it is a hypothetical protein.

Function **(F)** Minor tail protein	Gene No: 25		Student name: Betty Sierra / Grace Intrator Date: 2/13/24		
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 91.48	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		% Identity 91	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) Siphovirus protein	Name / Descr (incl DUF) Siphovirus ReqiPepy6	Hit PF14594.10	Probability (if > 90%) 99.95	E-value (if < 1) 9.5e-24
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 142684 (37)		upstream pham conserved ? Yes or [✓ No] If yes, what pham # or function ? 140921	Downstream pham conserved ? Yes or [✓ No] If yes, what pham # or function? 3528
Transmembrane domains by TmHmm using Phamerator			no		

Notes: Most evidence leads to it being a minor tail protein given that only HHpred says it is a syphovirus protein.

Function **(F)** Hypothetical protein	Gene No: 26			Student name: Betty Sierra / Grace Intrator Date: 2/13/24	
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 99.02	E-value (if < 10^{-7}) 1e-62	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])		% Identity 99	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)	Name / Descr (incl DUF)	Hit	Probability (if > 90%) All 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) Minor tail protein	Pham number and number of members 3528 (27)	upstream pham conserved ? Yes or [✓ No] If yes, what pham # or function ? 142684	Downstream pham conserved ? Yes or [✓ No] If yes, what pham # or function? 141462	
Transmembrane domains by TmHmm using Phamerator			no		

Notes: We decided it was a hypothetical protein as most evidence leads to suggest its function is unknown.

Function **(F)** Minor tail protein	Gene No: 27		Student name: Betty Sierra / Grace Intrator Date: 2/20/24		
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 99.18	E-value (if < 10^{-7}) 1e-169	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 99	E-value (if < 10^{-7}) e-134	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) Receptor binding protein	Name / Descr (incl DUF) Receptor binding protein	Hit 4L9B_A	Probability (if > 90%) 99.77	E-value (if < 1) 1e-17
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 141462 (27)	upstream pham conserved ? Yes or [✓ No] If yes, what pham # or function ?	Downstream pham conserved ? Yes or [✓ No] If yes, what pham # or function?	
Transmembrane domains by TmHmm using Phamerator			no		

Notes: We decided it was a minor tail protein since most evidence supports this, given that the other 2 program have different functions for it and there is more support for this function.

Function **(F)** Hypothetical protein	Gene No: 28		Student name: Betty Sierra / Grace Intrator Date: 2/20/24		
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}) 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])		% Identity 100	E-value (if < 10^{-7}) e-161	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) DUF859 sipho virus protein of unknown function	Hit PF05895.16	Probability (if > 90%) 99.44	E-value (if < 1) 5.2e-12
Supporting Information for Function, from Syntney **(SIF-Syn)** (All three lines of SIF evidence must be filled in)	Function (or NKF) Lysin	Pham number and number of members 143895 (152)	upstream pham conserved ? Yes or [✓ No] If yes, what pham # or function ?	Downstream pham conserved ? Yes or [✓ No] If yes, what pham # or function?	
Transmembrane domains by TmHmm using Phamerator			no		

Notes: We decided it is a hypothetical protein since more evidence leads to believe this, as only 1 program supports a different function.

Function **(F)** Hypothetical protein	Gene No: 29		Student name: Betty Sierra / Grace Intrator Date: 2/20/24		
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 98	% Identity 100	E-value (if < 10^{-7}) 7e-49	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])		% Identity 100	E-value (if < 10^{-7}) 9e-44	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 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 141167 (85)	upstream pham conserved ? Yes or [✓ No] If yes, what pham # or function ?	Downstream pham conserved ? Yes or [✓ No] If yes, what pham # or function?	
Transmembrane domains by TmHmm using Phamerator			no		

Notes: All evidence leads to believe it is a hypothetical protein as its function does not seem to be known yet.

Function **(F)** **endolysin**	Gene No: 30		Student name: Arianna Marie Ruiz Date: 2/6/2024		
Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)** (All three lines of SIF evidence must be filled in)	Function (or NKF) Lysin A	% Query 99%	% Identity 100.00%	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) Lysin A		% Identity 100%	E-value (if < 10^{-7}) e^-175	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) Lysin A	Name / Descr (incl DUF) D,D-dipeptidase/ D,D-carboxypeptidase; CENTER FOR STRUCTURAL GENOMICS OF INFECTIOUS DISEASES, CSGID, NATIONAL INSTITUTE O	Hit 4MUQ_A	Probability (if > 90%) 99.54%	E-value (if < 1) 1.9e-13
Supporting Information for Function, from Syntney **(SIF-Syn)**	Function (or NKF) Lysin A	Pham number and number of members 98619 (138)	upstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ?	Downstream pham conserved ? [✓ Yes] or No	

(All three lines of SIF evidence must be filled in)			147123 (63)	If yes, what pham # or function ? 8687 (10)

Transmembrane domains by TmHmm using Phamerator	1

Notes :

Function **(F)** **membrane protein**	Gene No: 31		Student name: Arianna Marie Ruiz Date: 2/13/2024		
Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)** (All three lines of SIF evidence must be filled in)	Function (or NKF) NKF	% Query 99%	% Identity 100.00%	E-value (if < 10^{-7}) 2e^-67	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) NKF		% Identity 100%	E-value (if < 10^{-7}) 2e^-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) NKF	Name / Descr (incl DUF) ; DUF6118 ; Family of unknown function (DUF6118)	Hit PF19613.3	Probability (if > 90%) 97.26	E-value (if < 1) 0.031
Supporting Information for Function, from Syntney **(SIF-Syn)** (All three lines of SIF evidence must be filled in)	Function (or NKF) NKF	Pham number and number of members 8687 (9)		upstream pham conserved? [✓ Yes] or No If yes, what pham # or function ? 98619 (138)	Downstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 8647 (9)
Transmembrane domains by TmHmm using Phamerator			yes, by 2 programs		

Notes :

Function **(F)** **membrane protein**	Gene No: 32		Student name: Arianna Marie Ruiz Date: 2/14/2024		
Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)** (All three lines of SIF evidence must be filled in)	Function (or NKF) NFK	% Query 96%	% Identity 99.53%	E-value (if < 10^{-7}) 2e-147	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) NFK		% Identity 98%	E-value (if < 10^{-7}) e-123	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%) *No probability above 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) NFK	Pham number and number of members 8647 (9)	upstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 8687 (9)	Downstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 5847 (14)	
Transmembrane domains by TmHmm using Phamerator	yes, by 2 programs				

Function **(F)** **holin**	Gene No: 33	Student name: Arianna Marie Ruiz Date: 2/13/2024			
Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)** (All three lines of SIF evidence must be filled in)	Function (or NKF) holin	% Query 99%	% Identity 90.67%	E-value (if < 10^{-7}) 3e^-123	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) holin		% Identity 90%	E-value (if < 10^{-7}) 1e^-91	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) holin	Name / Descr (incl DUF) ; Phage_holin_7_1 ; Mycobacterial 2 TMS Phage Holin (M2 Hol) Family	Hit PF16081.9	Probability (if > 90%) 98.14%	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) holin	Pham number and number of members 5847 (14)	upstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 8647 (9)	Downstream pham conserved ? Yes or [✓ No] If yes, what pham # or function ?	
Transmembrane domains by TmHmm using Phamerator					

Notes: Have to check for biochemical properties again. (if anything its NFK)

Function **(F) hypothetical protein**	Gene No: 34	Student name: Arianna Marie Ruiz Date: 2/14/2024

Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)** (All three lines of SIF evidence must be filled in)	Function (or NKF) NFK	% Query 97%	% Identity 75.61%	E-value (if < 10^{-7}) 3e-09	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) NFK		% Identity 75%	E-value (if < 10^{-7}) 4e-09	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%) *No probability above 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) NFK	Pham number and number of members 14633 (6)	upstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 5847 (14)	Downstream pham conserved ? Yes or [✓ No] If yes, what pham # or function ?

Transmembrane domains by TmHmm using Phamerator	

Notes :

Function **(F) Hypothetical protein**	Gene No: 35		Student name: Arianna Marie Ruiz Date: 2/14/2024		
Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)** (All three lines of SIF evidence must be filled in)	Function (or NKF) NFK	% Query 99%	% Identity 91.92%	E-value (if < 10^{-7}) 1e-57	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) NFK		% Identity 88%	E-value (if < 10^{-7}) 3e-49	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%) *No probability above 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) NFK	Pham number and number of members 88323 (8)	upstream pham conserved ? Yes or [✓ No] If yes, what pham # or function ?	Downstream pham conserved ? Yes or [✓ No] If yes, what pham # or function ?	
Transmembrane domains by TmHmm using Phamerator					

Notes : Only BLAST NCBI was calling hypothetical protein.

Function **(F)** **Hypothetical protein**	Gene No: 36			Student name: Arianna Marie Ruiz Date: 2/14/2024	
Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)** (All three lines of SIF evidence must be filled in)	Function (or NKF) N/A	% Query N/A	% Identity N/A	E-value (if < 10^{-7}) N/A	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) NFK		% Identity 43%	E-value (if < 10^{-7}) 0.11	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%) *No probability above 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) NFK	Pham number and number of members 137359 (1)	upstream pham conserved ? Yes or [✓ No] If yes, what pham # or function ?	Downstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 8611 (9)	
Transmembrane domains by TmHmm using Phamerator					

Notes : No similarities found in BLAST NCBI

Function **(F) hypothetical protein**	Gene No: 37		Student name: Arianna Marie Ruiz Date: 2/14/2024		
Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)** (All three lines of SIF evidence must be filled in)	Function (or NKF) NFK	% Query 99%	% Identity 68.46%	E-value (if < 10^{-7}) 2e-61	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) NFK		% Identity 70%	E-value (if < 10^{-7}) 6e-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)	Name / Descr (incl DUF)	Hit	Probability (if > 90%) *No probability above 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) NFK	Pham number and number of members 8611 (9)	upstream pham conserved ? Yes or [✓ No] If yes, what pham # or function ?	Downstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 8178 (9)	
Transmembrane domains by TmHmm using Phamerator					

Notes :

Function **(F) membrane protein**	Gene No: 38		Student name: Arianna Marie Ruiz Date: 2/14/2024		
Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)** (All three lines of SIF evidence must be filled in)	Function (or NKF) NFK	% Query 98%	% Identity 100%	E-value (if < 10^{-7}) 2e-24	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) NFK		% Identity 100%	E-value (if < 10^{-7}) 4e-21	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) NFK	Name / Descr (incl DUF) ; DUF3487 ; Protein of unknown function (DUF3487)	Hit PF11990.12	Probability (if > 90%) 91.1%	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) NFK	Pham number and number of members 8178 (9)		upstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 8611 (9)	Downstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 146719 (646)
Transmembrane domains by TmHmm using Phamerator			yes, 2 domains		

Notes :

Function **(F)** **RuvC-like resolvase**	Gene No: 39			Student name: Arianna Marie Ruiz Date: 2/14/2024	
Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)** (All three lines of SIF evidence must be filled in)	Function (or NKF) RuvC-like resolvase	% Query 99%	% Identity 98.96%	E-value (if < 10^{-7}) 1e-136	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) RuvC-like resolvase		% Identity 98%	E-value (if < 10^{-7}) e-104	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) RuvC-like resolvase	Name / Descr (incl DUF) Crossover junction endodeoxyribonuclease RuvC; nuclease, DNA repair, Homologous recombination, Holliday junction resolva	Hit 7XHJ_A	Probability (if > 90%) 99.93%	E-value (if < 1) 5.4e-23
Supporting Information for Function, from Syntney **(SIF-Syn)** (All three lines of SIF evidence must be filled in)	Function (or NKF) RuvC-like resolvase	Pham number and number of members 146719 (646)		upstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 8178 (9)	Downstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 45958 (59)

Transmembrane domains by TmHmm using Phamerator	

Notes :

<table>
<tr><td>Function (F)
exonuclease</td><td colspan="2">Gene No: 40</td><td colspan="3">Student name: Arianna Marie Ruiz
Date: 2/14/2024</td></tr>
<tr><td>Supporting Information for Function, from BLAST (Protein) (BLAST-NCBI)
(All three lines of SIF evidence must be filled in)</td><td>Function
(or NKF)
exonuclease</td><td>% Query
99%</td><td>% Identity
100.00%</td><td>E-value
(if $< 10^{-7}$)
0.0</td><td>AA # conserved in first 10 hits?
[✓ Yes] or No</td></tr>
<tr><td>Supporting Information for Function, from BLAST (Protein) (BLAST-phagesDB)
(All three lines of SIF evidence must be filled in)</td><td>Function
(or NKF)
exonuclease</td><td></td><td>% Identity
100%</td><td>E-value
(if $< 10^{-7}$)
0.0</td><td>AA # conserved in first 10 hits?
[✓ Yes] or No</td></tr>
<tr><td>Supporting Information for Function, from HHpred (HHpred)
(All three lines of SIF evidence must be filled in)</td><td>Function
(or NKF)
exonuclease</td><td>Name / Descr (incl DUF)
Exonuclease; alkaline exonuclease, digest double stranded DNA, strict 5-3-polarity, HYDROLASE; 1.9A {Laribacter hongkong</td><td>Hit
3SYY_A</td><td>Probability (if > 90%)
99.88%</td><td>E-value
(if < 1)
1.1e-20</td></tr>
<tr><td>Supporting Information for Function, from Syntney (SIF-Syn)
(All three lines of SIF evidence must be filled in)</td><td>Function
(or NKF)
exonuclease</td><td colspan="2">Pham number and number of members
45958 (59)</td><td>upstream pham conserved ?
[✓ Yes] or No
If yes, what pham # or function ?
146719 (646), RuvC-like resolvase</td><td>Downstream pham conserved ?
[✓ Yes] or No
If yes, what pham # or function ?
87972 (9)</td></tr>
</table>

Transmembrane domains by TmHmm using Phamerator	

Notes :

Function **(F)** **RrecT-like DNA pairing protein**	Gene No: 41		Student name: Arianna Marie Ruiz Date: 2/14/2024		
Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)** (All three lines of SIF evidence must be filled in)	Function (or NKF) RecT-like DNA pairing protein	% Query 99%	% Identity 99.43%	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) RecT-like DNA pairing protein		% Identity 99%	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) RecT-like DNA pairing protein	Name / Descr (incl DUF) Recombination protein bet; Annealase, Synaptase, SSAP, Single-strand annealing protein, DNA annealing intermediate, Reco	Hit 7UJL_A	Probability (if > 90%) 99.97%	E-value (if < 1) 2.9e-17
Supporting Information for Function, from Syntney **(SIF-Syn)** (All three lines of SIF evidence must be filled in)	Function (or NKF) RecT-like DNA pairing protein	Pham number and number of members 87972 (9)		upstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 45958 (59), exonuclease	Downstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 180 (312)

Transmembrane domains by TmHmm using Phamerator	

Notes :

Function **(F) hypothetical prrotein**	Gene No: 42		Student name: Arianna Marie Ruiz Date: 2/14/2024		
Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)** (All three lines of SIF evidence must be filled in)	Function (or NKF) NFK	% Query 99%	% Identity 99.17%	E-value (if < 10^{-7}) 1e-78	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) NFK		% Identity 100%	E-value (if < 10^{-7}) 1e-62	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) NFK	Name / Descr (incl DUF) Recombination protein uvsY; recombination, DNA binding, homo-heptamer, asymmetry, alpha barrel, VIRAL PROTEIN; 2.351A {E	Hit 4ZWQ_D	Probability (if > 90%) 95.36%	E-value (if < 1) N/A
Supporting Information for Function, from Syntney **(SIF-Syn)**	Function (or NKF) NFK	Pham number and number of members 180 (312)	upstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 87972 (9)	Downstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ?	

(All three lines of SIF evidence must be filled in)				146731 (507)
Transmembrane domains by TmHmm using Phamerator				

Notes :

Function **(F)** **HNH endonuclease**	Gene No: 43	Student name: Kristen Mclean Date: 2/25/2024			
Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)** (All three lines of SIF evidence must be filled in)	Function (or NKF) HNH endonuclease	% Query 99%	% Identity 98.88%	E-value (if < 10^{-7}) 1e^-83	AA # conserved in first 10 hits? Yes or [✓ No] It is conserved in the first 5 hits
Supporting Information for Function, from BLAST (Protein) **(BLAST-phagesDB)** (All three lines of SIF evidence must be filled in)	Function (or NKF) HNH endonuclease		% Identity 95%	E-value (if < 10^{-7}) 1e^-68	AA # conserved in first 10 hits? Yes or [✓ No] Only conserved in the first 6 hits
Supporting Information for Function, from HHpred **(HHpred)** (All three lines of SIF evidence must be filled in)	Function (or NKF) HNH endonuclease	Name / Descr (incl DUF) 5-methylcytosine-specific restriction enzyme A; HNH ENDONUCLEASE	Hit 6GHC_A	Probability (if > 90%) 93.65	E-value (if < 1) 0.18
Supporting Information for Function, from Syntney **(SIF-Syn)** (All three lines of SIF evidence must be filled in)	Function (or NKF) HNH endonuclease	Pham number and number of members 146731 (507)		upstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 145916 (26)	Downstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 146810 (265)
Transmembrane domains by TmHmm using Phamerator			N/A		

Notes :

Function **(F)** **Hypothetical prrotein**	Gene No: 44			Student name: Kristen Mclean Date: 2/13/2024	
Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)** (All three lines of SIF evidence must be filled in)	Function (or NKF) RepA-like replication initiator	% Query 99%	% Identity 99.64%	E-value (if < 10^{-7}) 0.0	AA # conserved in first 10 hits? Yes or [✓ No] Only conserved in the first 5 hits
Supporting Information for Function, from BLAST (Protein) **(BLAST-phagesDB)** (All three lines of SIF evidence must be filled in)	Function (or NKF) RepA-like replication initiator		% Identity 99%	E-value (if < 10^{-7}) e^-165	AA # conserved in first 10 hits? Yes or [✓ No] Conserved in the first 2 hits
Supporting Information for Function, from HHpred **(HHpred)** (All three lines of SIF evidence must be filled in)	Function (or NKF) Phage_rep_org_N ; N-terminal phage replisome organiser	Name / Descr (incl DUF) 5-methylcytosine-specific restriction enzyme A; HNH endonuclease, modification dependent restriction	Hit PF09681.14	Probability (if > 90%) 98.05	E-value (if < 1) 0.00034
Supporting Information for Function,	Function (or NKF)	Pham number and number of members 145916 (26)		upstream pham conserved ? [✓ Yes] or No	Downstream pham conserved ? [✓ Yes] or No

from Syntney **(SIF-Syn)** (All three lines of SIF evidence must be filled in)	RepA-like replication initiator		If yes, what pham # or function ? 146731 (507)	If yes, what pham # or function ? 144326 (29)
Transmembrane domains by TmHmm using Phamerator			N/A	

Notes :

Function **(F)** **helicase loader**	Gene No: 45		Student name: Kristen Mclean Date: 2/25/2024		
Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)** (All three lines of SIF evidence must be filled in)	Function (or NKF) Helicase loader	% Query 99%	% Identity 100%	E-value (if < 10^{-7}) 5e^-73	AA # conserved in first 10 hits? Yes or [✓ No] Conserved on the first hit
Supporting Information for Function, from BLAST (Protein) **(BLAST-phagesDB)** (All three lines of SIF evidence must be filled in)	Function (or NKF) Helicase loader		% Identity 100%	E-value (if < 10^{-7}) 1e^-56	AA # conserved in first 10 hits? Yes or [✓ No] Conserved in the first hit
Supporting Information for Function, from HHpred **(HHpred)** (All three lines of SIF evidence must be filled in)	Function (or NKF) replisome organizer; helical bipartite natively unfolded domain, replication; HET: MSE	Name / Descr (incl DUF) Phage_lambda_P; replication protein P	Hit 1NO1_A	Probability (if > 90%) 99.62	E-value (if < 1) 9.8e-14
Supporting Information for Function, from Syntney **(SIF-Syn)** (All three lines of SIF evidence must be filled in)	Function (or NKF) Helicase loader	Pham number and number of members 144326 (29)		upstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 145916 (26)	Downstream pham conserved ? Yes or [✓ No] If yes, what pham # or function ?
Transmembrane domains by TmHmm using Phamerator				N/A	

Notes :

Function **(F)** **DNAB-like dsDNA helicase**	Gene No: 46		Student name: Kristen Mclean Date: 2/25/2024		
Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)** (All three lines of SIF evidence must be filled in)	Function (or NKF) DnaB-like dsDNA helicase	% Query %99	% Identity %100	E-value (if < 10^{-7}) 0.0	AA # conserved in first 10 hits? Yes or [✓ No] Conserved on the first 5 hits.
Supporting Information for Function, from BLAST (Protein) **(BLAST-phagesDB)** (All three lines of SIF evidence must be filled in)	Function (or NKF) DnaB-like dsDNA helicase		% Identity 100%	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) DnaB-like dsDNA helicase	Name / Descr (incl DUF) Replicative helicase; Helicase, Primase, Replication, dnaB, dnaG; HET: MSE, SO4; 2.9	Hit 2R6A_A	Probability (if > 90%) 100	E-value (if < 1) 4.7e-44
Supporting Information for Function, from Syntney **(SIF-Syn)** (All three lines of SIF evidence must be filled in)	Function (or NKF) DnaB-like dsDNA helicase	Pham number and number of members 146810 (265)	upstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 144326 (29)	Downstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 145287 (506)	
Transmembrane domains by TmHmm using Phamerator	N/A				

Notes :

Function **(F)** SSB protein	Gene No: 47			Student name: Kristen Mclean Date: 2/25/2024	
Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)** (All three lines of SIF evidence must be filled in)	Function (or NKF) ssDNA binding protein	% Query %99	% Identity %96.32	E-value (if < 10^{-7}) 2e^-111	AA # conserved in first 10 hits? Yes or [✓ No] Conserved on the first 6 hits
Supporting Information for Function, from BLAST (Protein) **(BLAST-phagesDB)** (All three lines of SIF evidence must be filled in)	Function (or NKF) ssDNA binding protein		% Identity %95	E-value (if < 10^{-7}) 1e-88	AA # conserved in first 10 hits? Yes or [✓ No] Only conserved on the first 7 hits.
Supporting Information for Function, from HHpred **(HHpred)** (All three lines of SIF evidence must be filled in)	Function (or NKF) Single-stranded DNA-binding protein	Name / Descr (incl DUF) Single-stranded DNA-binding protein; DNA replication, TRANSFERASE; HET: MSE; 2.6001A	Hit 3TQY_C	Probability (if > 90%) 99.93	E-value (if < 1) 1.3e-22
Supporting Information for Function, from Syntney **(SIF-Syn)** (All three lines of SIF evidence must be filled in)	Function (or NKF) Single-stranded DNA-binding protein	Pham number and number of members 145287 (506)		upstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 146810 (265)	Downstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 146844 (215)
Transmembrane domains by TmHmm using Phamerator				N/A	

Notes :

Function **(F)** **PartB-like nuelcase domain**	Gene No: 48		Student name: Kristen Mclean Date: 2/25/2024		
Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)** (All three lines of SIF evidence must be filled in)	Function (or NKF)	% Query	% Identity	E-value (if < 10^{-7})	AA # conserved in first 10 hits?
	ParB-like nuclease domain protein	%99	%100	1e-124	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)		% Identity	E-value (if < 10^{-7})	AA # conserved in first 10 hits?
	ParB-like nuclease domain protein		%96	2e-95	Yes or [✓ No] Conserved on the first 8 hits
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%)	E-value (if < 1)
	ParB family protein; DNA-binding protein	ParB family protein; DNA-binding protein, CTP, Myxococcus, DNA-segregation, DNA BINDING PROTEIN; HET: UFQ, GOL; 1.7A	7BNR_B	99.52	1.3e-12
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	upstream pham conserved ? [✓ Yes] or No	Downstream pham conserved ? [✓ Yes] or No	
			If yes, what pham # or function ?	If yes, what pham # or function ?	
	ParB-like nuclease domain protein	146844 (215)	145287 (506)	147353 (28)	

Transmembrane domains by TmHmm using Phamerator	N/A

<table>
<tr><td>Function (F)
acetyltransferase</td><td colspan="2">Gene No: 49</td><td colspan="3">Student name: Kristen Mclean
Date: 2/25/2024</td></tr>
<tr><td>Supporting Information for Function, from BLAST (Protein) (BLAST-NCBI) (All three lines of SIF evidence must be filled in)</td><td>Function (or NKF)
acetyltransferase</td><td>% Query
%99</td><td>% Identity
%98.56</td><td>E-value (if < 10^{-7})
1e-150</td><td>AA # conserved in first 10 hits?
Yes or [✓ No]
Conserved on the first 6 hits</td></tr>
<tr><td>Supporting Information for Function, from BLAST (Protein) (BLAST-phagesDB) (All three lines of SIF evidence must be filled in)</td><td>Function (or NKF)
acetyltransferase</td><td></td><td>% Identity
%98</td><td>E-value (if < 10^{-7})
e-120</td><td>AA # conserved in first 10 hits?
Yes or [✓ No]
Conserved on the first 8 hits.</td></tr>
<tr><td>Supporting Information for Function, from HHpred (HHpred) (All three lines of SIF evidence must be filled in)</td><td>Function (or NKF)
acetyltransferase</td><td>Name / Descr (incl DUF)
ACETYLTRANSFERASE; Acyl Coenzyme A complex, TRANSFERASE; HET: ACO, MSE; 1.55A</td><td>Hit
1GHE_B</td><td>Probability (if > 90%)
97.77</td><td>E-value (if < 1)
0.0017</td></tr>
<tr><td>Supporting Information for Function, from Syntney (SIF-Syn) (All three lines of SIF evidence must be filled in)</td><td>Function (or NKF)
acetyltransferase</td><td>Pham number and number of members
147353 (28)</td><td colspan="2">upstream pham conserved ?
[✓ Yes] or No
If yes, what pham # or function ?
146844 (215)</td><td>Downstream pham conserved ?
[✓ Yes] or No
If yes, what pham # or function ?
8860 (9)</td></tr>
<tr><td colspan="3">Transmembrane domains by TmHmm using Phamerator</td><td colspan="3">N/A</td></tr>
</table>

Notes :

Function **(F)** **hypothetical prot**	Gene No: 50		Student name: Kristen Mclean Date: 2/25/2024		
Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)** (All three lines of SIF evidence must be filled in)	Function (or NKF) NFK	% Query %99	% Identity %99.07	E-value (if < 10^{-7}) 2e-71	AA # conserved in first 10 hits? Yes or [✓ No] Conserved on the first 3 hits.
Supporting Information for Function, from BLAST (Protein) **(BLAST-phagesDB)** (All three lines of SIF evidence must be filled in)	Function (or NKF) NKF		% Identity %99	E-value (if < 10^{-7}) 1e-58	AA # conserved in first 10 hits? Yes or [✓ No] Conserved in the first 5 hits.
Supporting Information for Function, from HHpred **(HHpred)** (All three lines of SIF evidence must be filled in)	Function (or NKF) Transforming growth factor beta	Name / Descr (incl DUF) Transforming growth factor beta-1; pro-complex, latency, homodimer, PROTEIN BINDING; HET: NAG, BMA; 2.9A	Hit 5VQF_C	Probability (if > 90%) 67.92	E-value (if < 1) 27
Supporting Information for Function, from Syntney **(SIF-Syn)** (All three lines of SIF evidence must be filled in)	Function (or NKF) NKF	Pham number and number of members 8860 (9)	upstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 147353 (28)		Downstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 142271 (208)
Transmembrane domains by TmHmm using Phamerator			N/A		

Notes :

Function **(F)** **hypothetical prot**	Gene No: 51		Student name: Kristen Mclean Date: 2/25/2024		
Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)** (All three lines of SIF evidence must be filled in)	Function (or NKF) NKF	% Query %99	% Identity %100	E-value (if < 10^{-7}) 0.0	AA # conserved in first 10 hits? Yes or [✓ No] Conserved on the first 5 hits
Supporting Information for Function, from BLAST (Protein) **(BLAST-phagesDB)** (All three lines of SIF evidence must be filled in)	Function (or NKF) NKF		% Identity %100	E-value (if < 10^{-7}) 0.0	AA # conserved in first 10 hits? Yes or [✓ No] Conserved on the first 8 hits
Supporting Information for Function, from HHpred **(HHpred)** (All three lines of SIF evidence must be filled in)	Function (or NKF) NKF	Name / Descr (incl DUF) DUF3871 ; Domain of unknown function, B. Theta Gene description	Hit PF12987.11	Probability (if > 90%) 99.93	E-value (if < 1) 2.4e-24
Supporting Information for Function, from Syntney **(SIF-Syn)** (All three lines of SIF evidence must be filled in)	Function (or NKF) NKF	Pham number and number of members 142271 (208)	upstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 8860 (9)	Downstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 8491 (9)	
Transmembrane domains by TmHmm using Phamerator	N/A				

Notes:

Function **(F)** **hypothetical prote**	Gene No: 52			Student name: Kristen Mclean Date: 2/25/2024	
Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)** (All three lines of SIF evidence must be filled in)	Function (or NKF) NKF	% Query %99	% Identity %96.80	E-value (if < 10^{-7}) 2e-79	AA # conserved in first 10 hits? Yes or [✓ No] Conserved on the first 4 hits
Supporting Information for Function, from BLAST (Protein) **(BLAST-phagesDB)** (All three lines of SIF evidence must be filled in)	Function (or NKF) NKF		% Identity %96	E-value (if < 10^{-7}) 3e-62	AA # conserved in first 10 hits? Yes or [✓ No] Conserved on the first 7 hits
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%)	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) NKF	Pham number and number of members 8491 (9)	upstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 142271 (208)	Downstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 147042 (86)	
Transmembrane domains by TmHmm using Phamerator			N/A		

Notes:

Function **(F)** **helix-turn helix dna binding protein**	Gene No: 53			Student name: Kristen Mclean Date: 2/25/2024	
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 protein	% Query % 95	% Identity %100	E-value (if < 10^{-7}) 1e-55	AA # conserved in first 10 hits? Yes or [✓ No] Conserved on the first 2 hits
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		% Identity %100	E-value (if < 10^{-7}) 2e-46	AA # conserved in first 10 hits? Yes or [✓ No] Conserved on the first 7 hits
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	Name / Descr (incl DUF) Regulatory protein cox; helix-turn-helix, DNA binding, VIRAL PROTEIN; 2.401A	Hit 4LHF_A	Probability (if > 90%) 98.28	E-value (if < 1) 0.000012
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 protein	Pham number and number of members 147042 (86)		upstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ? 8491 (9)	Downstream pham conserved ? Yes or [✓ No] If yes, what pham # or function ?
Transmembrane domains by TmHmm using Phamerator				N/A	

Notes:

Function **(F)** **Hypothetical prot**	Gene No: 54			Student name: Kristen Mclean Date: 2/25/2024	
Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)** (All three lines of SIF evidence must be filled in)	Function (or NKF) Membrane protein	% Query %97	% Identity %97.87	E-value (if < 10^{-7}) 1e-21	AA # conserved in first 10 hits? Yes or [✓ No] Conserved on the first 3 hits
Supporting Information for Function, from BLAST (Protein) **(BLAST-phagesDB)** (All three lines of SIF evidence must be filled in)	Function (or NKF) NKF		% Identity %97	E-value (if < 10^{-7}) 2e-19	AA # conserved in first 10 hits? Yes or [✓ No] Conserved on the first 5 hits
Supporting Information for Function, from HHpred **(HHpred)** (All three lines of SIF evidence must be filled in)	Function (or NKF) NKF	Name / Descr (incl DUF) DUF2970 ; Protein of unknown function	Hit PF11174.12	Probability (if > 90%) 81.89	E-value (if < 1) 4
Supporting Information for Function, from Syntney **(SIF-Syn)** (All three lines of SIF evidence must be filled in)	Function (or NKF) NKF	Pham number and number of members 10316 (7)		upstream pham conserved ? [✓ Yes] or No If yes, what pham # or function ?	Downstream pham conserved ? Yes or [✓ No] If yes, what pham # or function ?
Transmembrane domains by TmHmm using Phamerator			N/A		

Notes:

Function **(F)** Hypothetical protein	Gene No: 55		Student name: Betty Sierra / Grace Intrator Date: 2/20/24		
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 99.14	E-value (if < 10^{-7}) 6e-77	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])		% Identity 99	E-value (if < 10^{-7}) 3e-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)	Name / Descr (incl DUF)	Hit	Probability (if > 90%) All 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 144027 (87)	upstream pham conserved ? Yes or [✓ No] If yes, what pham # or function ?	Downstream pham conserved ? Yes or [✓ No] If yes, what pham # or function?	
Transmembrane domains by TmHmm using Phamerator			no		

Notes: All evidence leads to it being a hypothetical protein with unknown function as of now.

Function **(F) hypothetical protein**	Gene No: 56			Student name: Nicole Gonzalez Giliberti Date: 02/13/24	
Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)** (All three lines of SIF evidence must be filled in)	Function (or NKF) DNA binding protein	% Query 99%	% Identity 98.15%	E-value (if < 10^{-7}) 7e-69	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) DNA binding protein		% Identity 98%	E-value (if < 10^{-7}) 7e-55	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) Putative AphA-like transcription factor	Name / Descr (incl DUF) Putative AphA-like transcription factor; ZP_00208345.1, Putative AphA-like Transcription Factor, Structural Genomics, Jo	Hit 2RKH_A	Probability (if > 90%) 92.24%	E-value (if < 1) 0.52

Supporting Information for Function, from Syntney **(SIF-Syn)** (All three lines of SIF evidence must be filled in)	Function (or NKF)
	DNA binding protein
		Pham number and number of members
		88022(9)
			upstream pham conserved ?
			[✓ Yes] or No
			If yes, what pham # or function ?
			8461(10)
				Downstream pham conserved ?
				Yes or [✓ No]
				If yes, what pham # or function ?	

Transmembrane domains by TmHmm using Phamerator					

Notes: Across three databases, there is a consensus that the protein is identified to be a DNA Binding protein. HHpred is the only other database to state otherwise by saying it's a transcription factor.

Function **(F)** **HP**	Gene No: 57		Student name: Nicole Gonzalez Giliberti Date: 02/13/24		
Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)** (All three lines of SIF evidence must be filled in)	Function (or NKF) NKF	% Query 99%	% Identity 100.00%	E-value (if < 10^{-7}) 8e-90	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) NKF		% Identity 100%	E-value (if < 10^{-7}) 1e-68	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 / Descr (incl DUF) N/A	Hit N/A	Probability (if > 90%) N/A	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)
	NKF
		Pham number and number of members
		8216(9)
			upstream pham conserved ?
			Yes or [✓ No]
			If yes, what pham # or function ?
					Downstream pham conserved ?
					[✓ Yes] or No
					If yes, what pham # or function ?
					143748(371)

Transmembrane domains by TmHmm using Phamerator					

Notes : There was no known function to be found across all the databases, nothing to add.

Function **(F)** **HTH DNA bind domain**	Gene No: 58			Student name: Nicole Gonzalez Giliberti Date: 02/13/24	
Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)** (All three lines of SIF evidence must be filled in)	Function (or NKF) DNA binding protein	% Query 98%	% Identity 99.31%	E-value (if < 10^{-7}) 5e-97	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 protein		% Identity 99%	E-value (if < 10^{-7}) 8e-79	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) HTH_58;Helix-turn-helix domain	Name / Descr (incl DUF) HTH_58;Helix-turn-helix domain	Hit PF19575.3	Probability (if > 90%) 98.35%	E-value (if < 1) 0.000012

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
		142178 (454)
				upstream pham conserved ?
				[✓ Yes] or No
				If yes, what pham # or function ?
				8216 (9)
					Downstream pham conserved ?
					[✓ Yes] or No
					If yes, what pham # or function ?
					85852(50)

Transmembrane domains by TmHmm using Phamerator					

Notes : Protein has been identified to be a helix-turn-helix DNA binding domain protein across all the databases.

Function **(F)** **HTH DNA bind domain**	Gene No: 59		Student name: Nicole Gonzalez Giliberti Date: 02/13/24		
Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)** (All three lines of SIF evidence must be filled in)	Function (or NKF) DNA binding protein	% Query 99%	% Identity 90.70%	E-value (if < 10^{-7}) 1e-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) NKF		% Identity 88%	E-value (if < 10^{-7}) 1e-82	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) MarR family transcriptional regulator SArX	Hit 6PCP_C	Probability (if > 90%) 99.06%	E-value (if < 1) 2.8e-9

Supporting Information for Function, from Syntney **(SIF-Syn)** (All three lines of SIF evidence must be filled in)	Function (or NKF)
	NKF
		Pham number and number of members
		85852(50)
				upstream pham conserved ?
				[✓ Yes] or No
				If yes, what pham # or function ?
				143748(371)
					Downstream pham conserved ?
					Yes or [✓ No]
					If yes, what pham # or function ?

Transmembrane domains by TmHmm using Phamerator					

Function **(F)** **HP**	Gene No: 60		Student name: Nicole Gonzalez Giliberti Date: 2/20/24		
Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)** (All three lines of SIF evidence must be filled in)	Function (or NKF) NKF (hypothetical protein)	% Query 99%	% Identity 96.20%	E-value (if < 10^{-7}) 2e-105	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) NKF		% Identity 96%	E-value (if < 10^{-7}) 5e-85	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 / Descr (incl DUF) N/A	Hit N/A	Probability (if > 90%) N/A	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) NKF	Pham number and number of members 8740(9)	upstream pham conserved ? Yes or [✓ No] If yes, what pham # or function ?	Downstream pham conserved ? Yes or [✓ No] If yes, what pham # or function ?	
Transmembrane domains by TmHmm using Phamerator					

Notes : There was no known function to be found, the HHpred database had probability scores way below 90% and therefore no option was viable, therefore a "non applicable" answer was put down.

Function **(F)** **HP**	Gene No: 61		Student name: Nicole Gonzalez Giliberti Date: 2/20/24		
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 71.67%	E-value (if < 10^{-7}) 3e-126	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		% Identity 71%	E-value (if < 10^{-7}) e-106	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 / Descr (incl DUF) N/A	Hit N/A	Probability (if > 90%) N/A	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) NKF	Pham number and number of members 88148(8)	upstream pham conserved ? Yes or [✓ No] If yes, what pham # or function ?		Downstream pham conserved ? Yes or [✓ No] If yes, what pham # or function ?
Transmembrane domains by TmHmm using Phamerator					

Notes : No known function across a majority of the databases. the HHpred database had probability scores way below 90% and therefore no option was viable, therefore a “non applicable” answer was put down.

Function **(F)** **HP**	Gene No: 62		Student name: Nicole Gonzalez Giliberti Date: 2/23/2024		
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 80.42%	E-value (if < 10^{-7}) 9e-121	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) NKF		% Identity 80%	E-value (if < 10^{-7}) e-107	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 / Descr (incl DUF) N/A	Hit N/A	Probability (if > 90%) N/A	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) NKF	Pham number and number of members 8845(9)		upstream pham conserved ? Yes or [✓ No] If yes, what pham # or function ?	Downstream pham conserved ? Yes or [✓ No] If yes, what pham # or function ?
Transmembrane domains by TmHmm using Phamerator					

Notes : No known function across all databases, the HHpred database had probability scores way below 90% and therefore no option was viable, therefore a “non applicable” answer was put down.

Function **(F)** **HP**	Gene No: 63	Student name: Nicole Gonzalez Giliberti Date: 02/23/24			
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 94.64%	E-value (if < 10^{-7}) 7e-70	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) NKF		% Identity 94%	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) N/A	Name / Descr (incl DUF) N/A	Hit N/A	Probability (if > 90%) N/A	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) NKF	Pham number and number of members 87999(8)	upstream pham conserved ? Yes or [✓ No] If yes, what pham # or function ?	Downstream pham conserved ? Yes or [✓ No] If yes, what pham # or function ?	
Transmembrane domains by TmHmm using Phamerator					

Notes : No known function across databases, the HHpred database had probability scores way below 90% and therefore no option was viable, therefore a “non applicable” answer was put down.

Function **(F)** **HP**	Gene No: 64		Student name: Nicole Gonzalez Giliberti Date: 02/23/24		
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 86.54%	E-value (if < 10^{-7}) 3e-91	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) NKF		% Identity 86%	E-value (if < 10^{-7}) 1e-76	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 / Descr (incl DUF) N/A	Hit N/A	Probability (if > 90%) N/A	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) NKF	Pham number and number of members 8415(9)		upstream pham conserved ? Yes or [✓ No] If yes, what pham # or function ?	Downstream pham conserved ? Yes or [✓ No] If yes, what pham # or function ?
Transmembrane domains by TmHmm using Phamerator					

Notes : No known function across all databases, the HHpred database had probability scores way below 90% and therefore no option was viable, therefore a “non applicable” answer was put down.

Function **(F)** **HP**	Gene No: 65		Student name: Nicole Gonzalez Giliberti Date: 02/23/24		
Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)** (All three lines of SIF evidence must be filled in)	Function (or NKF) NKF	% Query 78%	% Identity 53.10%	E-value (if < 10^{-7}) 5e-25	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) NKF		% Identity 80%	E-value (if < 10^{-7}) 2e-08	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 / Descr (incl DUF) N/A	Hit N/A	Probability (if > 90%) N/A	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) NKF	Pham number and number of members 136905(1)	upstream pham conserved ? Yes or [✓ No] If yes, what pham # or function ?	Downstream pham conserved ? Yes or [✓ No] If yes, what pham # or function ?	
Transmembrane domains by TmHmm using Phamerator					

Notes : No known function across databases, the HHpred database had probability scores way below 90% and therefore no option was viable, therefore a “non applicable” answer was put down.

Function **(F)** **HP**	Gene No: 66		Student name: Nicole Gonzalez Giliberti Date: 02/23/24		
Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)** (All three lines of SIF evidence must be filled in)	Function (or NKF) NKF	% Query 98%	% Identity 96.70%	E-value (if < 10^{-7}) 7e-57	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) NKF		% Identity 96%	E-value (if < 10^{-7}) 9e-47	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 / Descr (incl DUF) N/A	Hit N/A	Probability (if > 90%) N/A	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) NKF	Pham number and number of members 140288(9)	upstream pham conserved ? Yes or [✓ No] If yes, what pham # or function ?		Downstream pham conserved ? Yes or [✓ No] If yes, what pham # or function ?
Transmembrane domains by TmHmm using Phamerator					

Notes : No known function across the databases, the HHpred database had probability scores way below 90% and therefore no option was viable, therefore a “non applicable” answer was put down.

Function **(F)** **HP**	Gene No: 67		Student name: Nicole Gonzalez Giliberti Date: 02/23/24		
Supporting Information for Function, from BLAST (Protein) **(BLAST-NCBI)** (All three lines of SIF evidence must be filled in)	Function (or NKF) NKF	% Query 98%	% Identity 98.67%	E-value (if < 10^{-7}) 2e-45	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) NKF		% Identity 98%	E-value (if < 10^{-7}) 5e-41	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 / Descr (incl DUF) N/A	Hit N/A	Probability (if > 90%) N/A	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) NKF	Pham number and number of members 90209(3)	upstream pham conserved ? Yes or [✓ No] If yes, what pham # or function ?		Downstream pham conserved ? Yes or [✓ No] If yes, what pham # or function ?
Transmembrane domains by TmHmm using Phamerator					

Notes : No function found across databases, the HHpred database had probability scores way below 90% and therefore no option was viable, therefore a “non applicable” answer was put down.

tRNA 1, starting at 32043 and ending at 32117, is not valid because its infernal score is 18.6 (below 35), as found by Aragorn and tRNA Scan SE's infernal score.

tRNA 2:

Starts at 30528

Ends at 30454

Type: Gly

Anticodon: CCC

Infernal Score: 46.4

The infernal score is above 35, so it is valid, as found by Aragorn and tRNA Scan SE's infernal score.