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	Original 3' End Stop	Original Length	Original Start Codon		
	Start 1	522	522	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		

Start called by both	Yes	Start:1	Does this Gene have Coding Potential	<mark>Yes</mark> No	Estimated Start: 1
Glimmer and	No	Glimmer Start:	<i>G</i>		
GeneMark?		GeneMark Start:			

Ribosome Binding Site

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

Starterator

Starterator Start

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

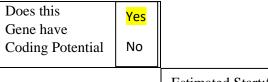
Function (F) Terminase small subunit	Gene No:1			Student name: Betty Sierra/Grace Intrator Date:2/6/24		
Supporting Information for Function, from BLAST (Protein) (BLAST-NCBI) (All three lines of SIF evidence must be filled in)	Function (or NKF) Terminase small subunit	<u>% Query</u> 99	<u>% Iden</u>	-	<u>E-value</u> (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		<u>% Iden</u> 97	tity	<u>E-value</u> (<u>if < 10⁻⁷)</u> 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	<u>Hit</u> 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 numb number of m 87960 (of members		upstream pham conserved? Yes or No es, what pham # or function? A first gene of the genome	Downstream pham conserved? Yes or No If yes, what pham # or function? 8570
Transmembrane	domains by Tm	Hmm using Pha	merator	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	Original 3' End Stop	Original Length	Original Start Codon	
	Start	656	138	ATG / GTG /	
	519			TTG	
Gene Direction	Gene Start, Leng	gth, and Start Codon Sar	me 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	<mark>Yes</mark>	Start:519
by both Glimmer and	No	Glimmer Start:
GeneMark?		GeneMark Start:



Estimated Start:505

Ribosome Binding Site

RBS Final		Score Closest to 0: -6.179
Score	Closest to 0?	Start #: 585
-6.770	Yes / No (if no)	Start #. 363
0.770	Test to (ii iio)	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 <mark>NKF</mark>) Hypothetical Protein	<u>% Query</u> 95	<u>% Iden</u> 95.5		<u>E-value</u> (<u>if < 10⁻⁷)</u> 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		<u>% Iden</u> 95		<u>E-value</u> (<u>if < 10⁻⁷)</u> 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)	<u>Function</u> (or <mark>NKF</mark>) <mark>NKF</mark>	Name / Descr (incl DUF) DUF3789; Protein of unknown function (DUF3789)	Hit PF1266		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)	<u>Function</u> (or <mark>NKF</mark>) <mark>NKF</mark>	Pham number of m	nembers	_	pstream pham conserved ? Yes or No s, 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	Original 3' End Stop	Original Length	Original Start Codon	
	Start	880	228	ATG / GTG /	
	653			TTG	
Gene Direction	Gene Start, Leng	gth, and Start Codon Sar	me as original (Au	to Annotated) call?	
FWD / REV	Yes / No (if no, answer below)				
	Revised 5' End Start	Does this	Nev h	New Start Codon	
		Gene have Coding Pot		ATG / GTG / TTG	
		Coding Pot	ential No		

Glimmer and GeneMark Calls

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

Estimated Start:653

Ribosome Binding Site

RBS Final	Is this Score the	Score Closest to 0: -5.531
Score	Closest to 0?	Start #: 854
-5.621 (second closest to 0)	Yes / No (if no) ☐	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			nt name: Carlos Rodrig	guez
Supporting Information for Function, from BLAST (Protein) (BLAST-NCBI) (All three lines of SIF evidence must be filled in)	Function (or <mark>NKF</mark>) Hypothetical Protein	<u>% Query</u> 98%	<u>% Ide</u> i	 <u>E-value</u> (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)	<u>Function</u> (or <mark>NKF</mark>)		<u>% Ide</u>	 <u>E-value</u> (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)	<u>Function</u> (or <mark>NKF</mark>)	Name / Descr (incl DUF)	<u>Hit</u>		<u>Probabi</u> <u>90</u> <u>All prob</u> <u>belo</u>	%) ablities	<u>E-value</u> (if < 1)
Supporting Information for Function, from Syntney (SIF-Syn) (All three lines of SIF evidence must be filled in)	<u>Function</u> (or <mark>NKF</mark>)	Pham number of m	nembers	<u>If yes</u>	pstream phoson conserved Yes or Notes, what phates function is 8570 Yes	<u>?</u> m # or	Downstream pham conserved? Yes or No If yes, what pham # or function? 143300
Transmembrane domains by TmHmm using Phamer Are no have No No							

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

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

Estimated Start:883

Ribosome Binding Site

RBS Final	Is this Score the	Score Closest to 0:
Score	Closest to 0?	
		Start #:
-2.034	Yes / No (if no)	
-2.034	1 cs / 140 (11 110)—	ORF Length:
		Old Leligiti.

Starterator

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

Function (F) Hypothetical Protein	Gene No: 4			nt name: Betty Sierra 2/13/24	/ 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	<u>% Query</u> 98%	<u>% Ide</u> 84.4	 <u>E-value</u> (<u>if < 10⁻⁷)</u> 2e-27	AA # conserved in first 10 hits? Yes or No
Supporting Information for Function, from BLAST (Protein) (BLAST- phagesDB)	<u>Function</u> (or <mark>NKF)</mark>		<u>% Ide</u> 100	 <u>E-value</u> (if < 10 ⁻⁷) 8e-29	AA # conserved in first 10 hits? Yes or No

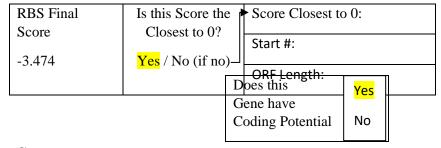
(All three lines of SIF evidence must be filled in)						
Supporting Information for Function, from HHpred (HHpred) (All three lines of SIF evidence must be filled in)	<u>Function</u> (or <mark>NKF</mark>)	Name / Descr (incl DUF) DUF6225; Family of unknown function	<u>Hit</u> PF1973		Probability (if > 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)	<u>Function</u> (or <mark>NKF</mark>)	Pham numb number of m 146389	nembers		pstream pham conserved ? Yes or No s, 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	Original 3' End Stop	Original Length	Original Start Codon	
	Start 1059	1172	114	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	

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

Estimated Start:1050

Ribosome Binding Site



Starterator

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

Function (F) Hypothetical protein	Gene No: 5			nt name: Betty Sierra , 2/13/24	/ Grace Intrator
Supporting Information for Function,	Function (or NKF) Hypothetical	<u>% Query</u> 97	<u>% Idei</u> 10	 <u>E-value</u> (if < 10 ⁻⁷)	AA # conserved in first 10 hits?
from BLAST (Protein) (BLAST-NCBI)	protein			2e-18	<mark>Yes</mark> 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 <mark>NKF</mark>)		<u>% Iden</u> 100		<u>E-value</u> (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 / <u>Descr (incl</u> <u>DUF)</u> DUF1660 ; Prophage protein	<u>Hit</u> PF07874	4.15	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)	<u>Function</u> (or <mark>NKF</mark>)	Pham numb number of m 147239(<u>embers</u>		pstream pham conserved ? Yes or No s, what pham # or function ? 143300	Downstream pham conserved? Yes or No If yes, what pham # or function? 8812
Transmembrane	Transmembrane domains by TmHmm using Phamerator					

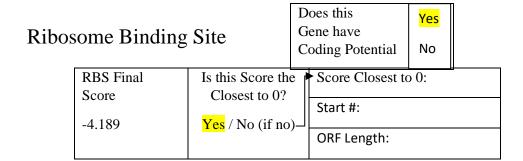
Gene # 6	Original 5' End	Original 3' End Stop	Original Length	Original Start Codon
	Start 1172	1405	234	ATG / GTG / TTG
Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?			
FWD / REV	Yes / No (if no, answer below)			

Revised 5' End Start

New Length	New Start Codon
	ATG / GTG / TTG

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

Estimated Start:1130



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

Supporting Information for Function,	Function (or NKF)	<u>% Query</u> 98	<u>% Iden</u> :		<u>E-value</u> (if < 10 ⁻⁷)	AA # conserved in first 10 hits?
from BLAST (Protein) (BLAST-NCBI) (All three lines of SIF evidence must be filled in)	Hypothetical protein				4e-39	<mark>Yes</mark> or No
Supporting Information for Function, from BLAST (Protein) (BLAST- phagesDB) (All three lines of SIF evidence must be filled in)	<u>Function</u> (or <u>NKF</u>)		<u>% Iden</u> 94	tity	<u>E-value</u> (if < 10 ⁻⁷) 5e-37	AA # conserved in first 10 hits? Yes or No
Supporting Information for Function, from HHpred (HHpred) (All three lines of SIF evidence must be filled in)	<u>Function</u> (or <mark>NKF</mark>)	<u>Name /</u> <u>Descr (incl</u> <u>DUF)</u>	<u>Hit</u>		Probability (if > 90%) All probabilities below 90	<u>E-value</u> (if < 1)
Supporting Information for Function, from Syntney (SIF-Syn) (All three lines of SIF evidence must be filled in)	<u>Function</u> (or <mark>NKF</mark>)	Pham number and number of members 8812(9)			pstream pham conserved ? Yes or No s, what pham # or function ? 142641	Downstream pham conserved? Yes or No If yes, what pham # or function? 1724
Transmembrane	Transmembrane domains by TmHmm using Phamerator					

Gene # 7	Original 5' End	Original 3' End Stop	Original Length	Original Start Codon
	Start	2855	1329	ATG / GTG /
	1527			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

			Does this	Yes	
Start called by both	Yes	Start:	Gene have Coding Potential	No	Estimated Start:1527
Glimmer and	No	Glimmer Start:1527	County Forential		
GeneMark?		GeneMark Start:1518			

Ribosome Binding Site

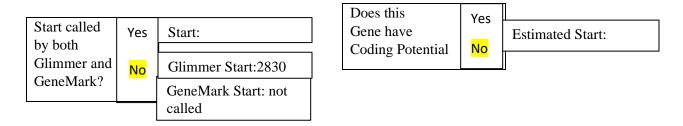
RBS Final	Is this Score the	Score Closest to 0: -4.315
Score	Closest to 0?	Start #: 1704
-6.565	Yes / <mark>No</mark> (if no)ᆜ	ORF Length:1152 (this start cuts off too much of the gene)

Starterator

Starterator Start 1527 with 16 MA's

Function (F) Terminase, large subunit	Gene No: 7			Intrato	_	uez/ Betty Sierra / Grace
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	<u>% Query</u> 99%	<u>% Ider</u> 99.78	-	<u>E-value</u> (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		<u>% Ider</u> 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}	<u>Hit</u>	_	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)			pstream pham conserved ? Yes or No s, what pham # or function ? 8812	Downstream pham conserved? Yes or No If yes, what pham # or function?
Transmembrane	domains by Tm	nHmm using Pha	merator	N/A		

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



Ribosome Binding Site

RBS Final	Is this Score the	Score Closest to 0:
Score	Closest to 0?	
20010		Start #:
N/A	Yes / No (if no)	
14/11	1 63 / 140 (11 110)—	ORF Length:
		om zengem

Starterator

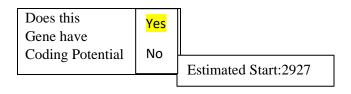
Starterator Start
No start
reported

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)	<u>% Query</u>	<u>% Iden</u>	tity	<u>E-value</u> (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)	<u>Function</u> (or NKF)		<u>% Iden</u>	tity	<u>E-value</u> (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)	<u>Hit</u>		Probability (if > 90%)	<u>E-value</u> (if < 1)
Supporting Information for Function, from Syntney (SIF-Syn) (All three lines of SIF evidence must be filled in)	Function (or NKF)	Pham numb number of m			pstream pham conserved ? Yes or No s, what pham # or function ?	Downstream pham conserved? Yes or No If yes, what pham # or function?
Transmembrane	domains by Tn	nHmm using Pha	merator			

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

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



Ribosome Binding Site

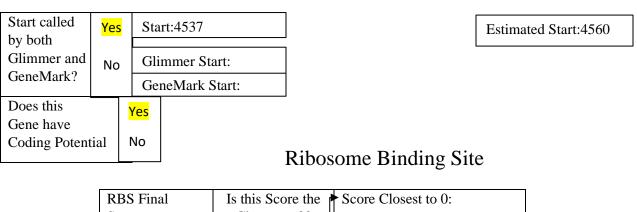
RBS Final	Is this Score the	Score Closest to 0: -3.449
Score	Closest to 0?	Start #: 3182
-4.475	Yes / No (if no)	
		ORF Length:1353 (this start
		cuts off too much of the
		gene)

Starterator

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

Function (F) portal protein	Gene No: 9				nt name: Kiryl Yasins 2/13/24	iki
Supporting Information for Function, from BLAST (Protein) (BLAST- NCBI) (All three lines of SIF evidence must be filled in)	Function (or NKF) portal protein	<u>% Query</u> 99	<u>% Iden</u> 98.5	-	<u>E-value</u> (if < 10 ⁻⁷) <u>0</u>	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		<u>% Iden</u> 98		<u>E-value</u> (<u>if < 10⁻⁷)</u> 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	<u>Hit</u> 7Z4W		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 numb number of m 1736(6)	embers		pstream pham conserved ? Yes or No s, what pham # or function ?	Downstream pham conserved? Yes or No If yes, what pham # or function? 141467
Transmembrane	domains by Tm	nHmm using Pha	merator	N/A		

Gene # 10	Original 5' End	Original 3' End Stop	Original Length	Original Start Codon	
	Start 4537	5652	1116	ATG / <mark>GTG</mark> / 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	



RBS Final	Is this Score the	Score Closest to 0:
Score	Closest to 0?	
		Start #:
-4.247	Yes / No (if no) →	
		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)	<u>% Query</u>	<u>% Identi</u>	ity	<u>E-value</u> (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)		% Identi	ity	<u>E-value</u> (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)	<u>Hit</u>		Probability (if > 90%)	<u>E-value</u> (if < 1)
Supporting Information for Function, from Syntney (SIF-Syn)	Function (or NKF)	Pham numb number of m			upstream pham conserved ? Yes or No es, 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 TmHmi	ım using Phamerator	

Gene # 11	Original 5' End	Original 3' End Stop	Original Length	Original Start Codon		
	Start	5960	279	ATG / GTG /		
	5682			TTG		
Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?					
FWD / REV	Yes / No (if no, answer below) Does this					
	Revised 5' End Start	Gene have	Ne Yes th	New Start Codon		
	5649	Coding Pot	ential No	ATG / GTG / TTG		

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

Estimated Start:5682

Ribosome Binding Site

RBS Final	Is this Score the	Score Closest to 0: -4.069
Score	Closest to 0?	
Beore	Closest to 0:	Start #: 5715
C 1C1	Man / NI (Cf)	36416 11. 37 23
-6.464	Yes / No (if no) ☐	
		ORF Length: 246

Starterator

Starterator Start

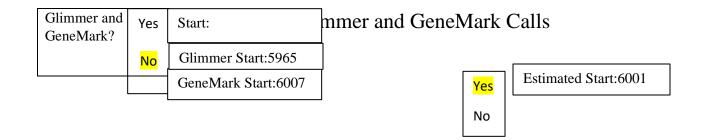
5682 has no other MA's

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	<u>% Query</u> 98	<u>% Identity</u> 45.26	<u>E-value</u> (<u>if < 10⁻⁷)</u> 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)	<u>Function</u> (or <mark>NKF</mark>)		<u>% Identity</u> 48	<u>E-value</u> (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	<u>Hit</u> PF1556	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) Transmembrane	Function (or <mark>NKF</mark>) domains by Tm	Pham numb number of m 137745 (2	<u>embers</u> 114)	pstream pham conserved ? Yes or No s, what pham # or function ? 141467	Downstream pham conserved? Yes or No If yes, what pham # or function? 8957

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

Start called by both



Ribosome Binding Site

RBS Final	Is this Score the	Score Closest to 0: -5.555
Score -5.994	Closest to 0? Yes / No (if no)	Start #: 6145
3.771	Tes to (II IIo)	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		Does this		nt name: Carlos Rodri 02/25/2024	guez
Supporting Information for Function, from BLAST (Protein) (BLAST-NCBI) (All three lines of SIF evidence must be filled in)	Function (or NKF) Hypothetical protein	<u>% Query</u> 98	Geneপ্রথপ্রধূদ্র Coding Pot 93.8	ential	E-value (if < 10 ⁻⁷) 1e-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)	<u>Function</u> (or <mark>NKF</mark>)		<u>% Iden</u> 93	tity	<u>E-value</u> (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)	<u>Hit</u>		Probability (if > 90%) All results are below 90%	<u>E-value</u> (<u>if < 1)</u>
Supporting Information for Function, from Syntney (SIF-Syn) (All three lines of SIF evidence must be filled in)	<u>Function</u> (or <mark>NKF</mark>)	Pham numb number of m 8957 (nembers		pstream pham conserved ? Yes or No s, what pham # or function ? 137745	Downstream pham conserved? Yes or No If yes, what pham # or function? 138124
Transmembrane	domains by Tml	Hmm using Pha	merator	N/A		

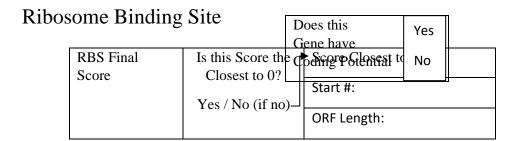
Gene #13*	Original 5' End	Original 3' End Stop	Original Length	Original Start Codon
	Start	7058	804	ATG / GTG /
	6255			TTG
Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?			
FWD / REV	Yes / No (if no, answer below)			

Revised 5' End Start

New Length	New Start Codon
	ATG / GTG / TTG

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

Estimated Start:



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	<u>% Query</u> 99%	<u>% Ider</u> 99.8		<u>E-value</u> (if < 10 ⁻⁷)	AA # conserved in first 10 hits?

from BLAST					0	Yes or No
(Protein) (BLAST-NCBI) (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) Scaffolding protein		<u>% Iden</u> 98	tity	<u>E-value</u> (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	<u>Hit</u> 6B0X <u>-</u>	_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)			pstream pham conserved ? Yes or No s, what pham # or function ? 8957	Downstream pham conserved? Yes or No If yes, what pham # or function? 2582
Transmembrane	Transmembrane domains by TmHmm using Phamerator					

Gene #14	Original 5' End	Original 3' End Stop	Original Length	Original Start Codon	
	Start	7947	873	ATG / GTG /	
	7075	,,,,,	070	TTG	
Gene Direction	Gene Start, Leng	th, and Start Codon Sar	me 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	

Start called	Yes	Start: 7075			Estimated Start: 7080
by both		2002017070	Does this	Yes	Estimated Start. 7000
Glimmer and	No	Glimmer Start:	Gene have		
GeneMark?		GeneMark Start:	Coding Potential	No	

Ribosome Binding Site

RBS Final	Is this Score the	Score Closest to 0: -4.473
Score	Closest to 0?	
20010		Start #: 7156
-4.626	Yes / No (if no)	
1.020	Test to (if ite)	ORF Length: 792

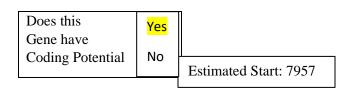
Starterator

Starterator Start 7075 has 33 MA's

Function (F) Major capsid	Gene No: 14 Student name: Carlos Rodrigu				iguez	
protein				Date:	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	<u>% Query</u> 99%	<u>% Ider</u> 95.86	-	<u>E-value</u> (<u>if < 10⁻⁷)</u> 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		<u>% Ider</u> 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	_	Probability (if > 90%) 99.96	<u>E-value</u> (<u>if < 1)</u> 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 numb number of m 2582 (4	<u>embers</u>		pstream pham conserved ? Yes or No s, what pham # or function ? 138124	Downstream pham conserved? Yes or No If yes, what pham # or function? 8234
Transmembrane	domains by Tm	nHmm using Pha	merator	N/A		

Gene #	Original 5' End	Original 3' End Stop	Original Length	Original Start Codon		
15	Start 8001	8417	417	ATG / <mark>GTG</mark> / 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		
	7947		471	ATG / GTG / TTG		

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



Ribosome Binding Site

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

Starterator

Starterator Start 7947 has 6 MA and 8001 has none

Function (F) Hypothetical Protein	Gene No: 15	o: 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	<u>% Query</u> 99	<u>% Iden</u>	-	<u>E-value</u> (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)	<u>Function</u> (or <mark>NKF</mark>)		% Identity 100		<u>E-value</u> (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)	<u>Function</u> (or <mark>NKF</mark>)	Name / Descr (incl DUF)	<u>Hit</u>		Probability (if > 90%) All probabilities below 90	<u>E-value</u> (<u>if < 1)</u>	
Supporting Information for Function, from Syntney (SIF-Syn) (All three lines of SIF evidence must be filled in)	<u>Function</u> (or <mark>NKF</mark>)	Pham numb number of m 8234(s	<u>nembers</u>		pstream pham conserved ? Yes or No s, what pham # or function ? 2582	Downstream pham conserved? Yes or No If yes, what pham # or function? 139922	
Transmembrane	Transmembrane domains by TmHmm using Phamerator N/A						

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

Does this	Yes
Gene have	
Coding Potential	No

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

Estimated start: 8515

Ribosome Binding Site

RBS Final	Is this Score the	Score Closest to 0:
Score	Closest to 0?	
		Start #:
-4.001	Yes / No (if no) ☐	
		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)	<u>Function</u> (or NKF) <mark>Head-to-tail</mark> <mark>adaptor</mark>	<u>% Query</u> 99	<u>% Iden</u> 98.78		<u>E-value</u> (<u>if < 10⁻⁷)</u> 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		<u>% Iden</u> 98	tity	<u>E-value</u> (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}	<u>Hit</u> 3JVO_	_E	Probability (if > 90%) 96.5%	<u>E-value</u> (<u>if < 1)</u> 0.029
Supporting Information for Function, from Syntney (SIF-Syn) (All three lines of SIF evidence must be filled in)	Function (or NKF) <mark>Head-to-tail</mark> <mark>adaptor</mark>	Pham numb number of m 147164(embers		pstream pham conserved ? Yes or No s, what pham # or function ? 8234	Downstream pham conserved? Yes or No If yes, what pham # or function? 143907

Transmembrane d	lomains l	bv TmHmn	using	Phamerator
Transmitting and		~,		· · · · · · · · · · · · · · · · · · ·

N/A

Gene #	Original 5' End	Original 3' End Stop	Original Length	Original Start Codon			
17	Start 8979	9302	324	ATG / GTG / TTG			
Gene Direction	Gana Start I and	eth and Start Codon Sar	ma as original (Aut	o Annotated) call?			
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	Does this	Nev Yes h	New Start Codon			
		Gene have Coding Pot		ATG / GTG / TTG			

Glimmer and GeneMark Calls

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

Estimated Start: 8950

Ribosome Binding Site

RBS Final	Is this Score the	Score Closest to 0:
Score	Closest to 0?	Start #:
-4.229	Yes / No (if no)	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			nt name: Akram Mikha 2/13/2024	ail
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	<u>% Query</u> 99%	<u>% Ide</u> 99.0	<u>E-value</u> (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		<u>% Ide</u>	 <u>E-value</u> (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	<u>Hit</u> 7z4w_	_3	Probability (if > 90%) 99.11	<u>E-value</u> (<u>if < 1)</u> 6.9e-10
Supporting Information for Function, from Syntney (SIF-Syn) (All three lines of SIF evidence must be filled in) Transmembrane	Function (or NKF) Head to tail stopper	Pham numb number of m 143907(1 nHmm using Pha	<u>embers</u> 44)		pstream pham conserved ? Yes or No s, what pham # or function ? 139922	Downstream pham conserved? Yes or No If yes, what pham # or function? 764

Gene # 18	Original 5' End	Original 3' End Stop	Original Length	Original Start Codon	
	9353	9685	306	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	
	9299		360	ATG / GTG / TTG	

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

Estimated Start:9310

Ribosome Binding Site

RBS Final	Is this Score the	Score Closest to 0:
Score	Closest to 0?	
Score		Start #:
-4.256	Yes / No (if no)	
- 1 .230	105 / 140 (II IIO)—	ORF Length:
		OIII LONGUII.

Starterator		Does the		<mark>Yes</mark> No
	Starterate 9299 has with all M this posit	or Start 18 MA MA in	- 335	

Function (F) Hypothetical protein	Gene No: 18				nt name: Akram Mikha 2/13/2024	ail
Supporting Information for Function, from BLAST (Protein) (BLAST-NCBI) (All three lines of SIF evidence must be filled in)	Function (or NKF) Hypothetical protein	<u>% Query</u> 99	<u>% Ide</u> 10	<u>-</u> _	<u>E-value</u> (<u>if < 10⁻⁷)</u> 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)	<u>Function</u> (or <mark>NKF)</mark>		<u>% Iden</u> 100%		<u>E-value</u> (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_capsi d_2; Minor capsid protein	<u>Hit</u> PF11111	.4.12	Probability (if > 90%) 99.67	<u>E-value</u> (<u>if < 1)</u> 3.2e-15
Supporting Information for Function, from Syntney (SIF-Syn) (All three lines of SIF evidence must be filled in)	<u>Function</u> (or <mark>NKF</mark>)	Pham number and number of members 764(144)			pstream pham conserved ? Yes or No s, 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.

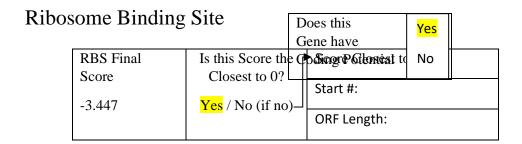
Gene # 19	Original 5' End	Original 3' End Stop	Original Length	Original Start Codon
	Start	10098	444	ATG / GTG /
	9655			TTG
Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?			
FWD / REV		Yes / No (if no, a	answer below)	

Revised 5' End Start

New Length	New Start Codon
	ATG / GTG / TTG

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

Estimated Start:9655



Starterator

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

Function (F)	Gene No: 19			Stude	nt name: Akram Mikha	ail
Tail terminator				Date:	2/13/2024	
Supporting Information	Function (or NKF)	<u>% Query</u> 99	<u>% Ide</u> 10		<u>E-value</u> (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	<mark>Yes</mark> 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		<u>% Iden</u> 100	-	<u>E-value</u> (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	<u>Hit</u> 6TE9-	·F	Probability (if > 90%) 99.02	<u>E-value</u> (<u>if < 1)</u> 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 numb number of m 114826(1	embers		pstream pham conserved ? Yes or No s, what pham # or function ? 764	Downstream pham conserved? Yes or No If yes, what pham # or function? 85343
Transmembrane	domains by Tm	nHmm using Pha	merator	N/A		

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

Start called	Yes	Start:10114			
by both					Estimated Start: 10170
Glimmer and	No	Glimmer Start:	Does this	Yes	I Estimated Start. 10170
GeneMark?		GeneMark Start:	Gene have Coding Potential	No	

Ribosome Binding Site

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

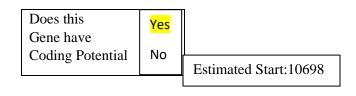
Starterator

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

Function (F) Major tail protein	Gene No: 20				nt name: Akram Mik 2/13/2024	hail
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	<u>% Query</u> 99	<u>% Ider</u>	-	<u>E-value</u> (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		<u>% Ider</u>	-	<u>E-value</u> (<u>if < 10⁻⁷)</u> 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	<u>Hit</u> 6TE9	_	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 numb number of m 85343(9	embers		pstream pham conserved ? Yes or No s, what pham # or function ? 114826	Downstream pham conserved? Yes or No If yes, what pham # or function? 85343
Transmembrane	domains by Tn	nHmm using Pha	merator	N/A		-

Gene # 21	Original 5' End	Original 3' End Stop	Original Length	Original Start Codon		
	Start	11189	489	ATG / <mark>GTG</mark> /		
	10701		,	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		

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



Ribosome Binding Site

RBS Final	Is this Score the	Score Closest to 0: -2.090
Score	Closest to 0?	Start #: 10722
-6.054	Yes / No (if no)	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				nt name: Akram Mik 2/13/2024	hail
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	<u>% Query</u> 99	<u>% Ider</u> 99.3		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		<u>% Ider</u>		<u>E-value</u> (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)	<u>Function</u> (or NKF)	Name / Descr (incl DUF)	<u>Hit</u>		Probability (if > 90%) All results below 90	<u>E-value</u> (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 numb number of m 123060(i	embers		pstream pham conserved ? Yes or No s, what pham # or function ? 85343	Downstream pham conserved? Yes or No If yes, what pham # or function? 144294
Transmembrane	domains by Tn	nHmm using Pha	merator	N/A	_	

Gene #22	Original 5' End	Original 3' End Stop	Original Length	Original Start Codon		
	Start 11186	11476	291	ATG / <mark>GTG</mark> / 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		
	10701		776	ATG / GTG / TTG		

Does this	Yes
Gene have	
Coding Potential	No

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

Estimated Start:N/A

Ribosome Binding Site

RBS Final	Is this Score the	Score Closest to 0: N/A
Score	Closest to 0?	
		Start #: N/A
N/A	Yes / No (if no)	
		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	<u>% Query</u> 99	<u>% Ident</u> 98.84%		<u>E-value</u> (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		<u>% Identi</u> 99	ity	<u>E-value</u> <u>(if < 10⁻⁷)</u> 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)	<u>Hit</u>		Probability (if > 90%) All probabilities below 90%	<u>E-value</u> (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 numb number of m 87960 (embers		upstream pham conserved ? Yes or No es, what pham # or function ?	Downstream pham conserved ? Yes or No If yes, what pham # or function ?

Transmembrane domains by TmHmm using Phamerator	N/A

Phage Name: Bush

Student Name: Betty Sierra/Grace Int.

Date:1/30/24

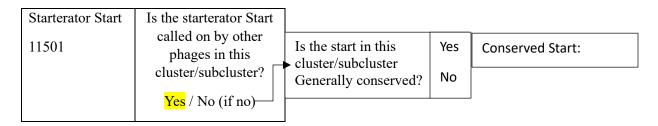
Gene #	Original 5' End Start	Original 3' End Stop	Original Length	Original Start Codon			
23	11501	13699	2199	ATG / GTG / TTG			
Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?						
FWD / REV	Yes / No (if no, answer below)						
	Revised 5' End Start		New Length	New Start Codon			
				ATG / GTG / TTG			

Glimmer and GeneMark Calls

Start called by both Glimmer and	Yes No	Start: 11501 Glimmer Start: 11501	Does this Gene have Coding Potential	Yes No	Estimated Start:
GeneMark?		GeneMark Start: 11501			

Ribosome Binding Site

RBS Final Score	Is this Score the	Score Closest to 0: -3.594
-4.917	Closest to 0? Yes / No (if no)	Start #: 12530
	100 / 110 (11 110)	ORF Length:1170



Phage Name: Bush

Student Name: Betty Sierra/Grace Int.

Date:1/30/24

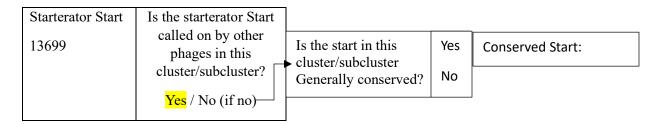
Gene #	Original 5' End Start	Original 3' End Stop	Original Length	Original Start Codon			
24	13699	14559	861	ATG / GTG / TTG			
Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?						
FWD / REV	Yes / No (if no, answer below)						
	Revised 5' End Start		New Length	New Start Codon			
				ATG / GTG / TTG			

Glimmer and GeneMark Calls

Start called by both	Yes	Start: 13699		Does this Gene have	Yes	Estimated Start:
Glimmer and	No	Glimmer Start: 13699	_	Coding Potential	No	
GeneMark?		GeneMark Start: 13699				

Ribosome Binding Site

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



Phage Name: Bush

Student Name: Betty Sierra/Grace Int.

Date:1/30/24

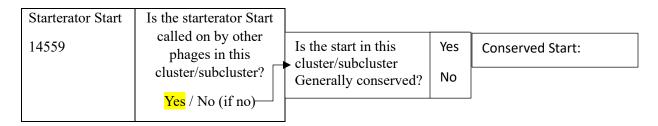
Gene #	Original 5' End Start	Original 3' End Stop	Original Length	Original Start Codon			
25	14559	16421	1863	ATG / GTG / TTG			
Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?						
FWD / REV	Yes / No (if no, answer below)						
	Revised 5' End Start		New Length	New Start Codon			
				ATG / GTG / TTG			

Glimmer and GeneMark Calls

Start called by both	Yes	Start: 14559	Does this Gene have	Yes	Estimated Start:
Glimmer and	No	Glimmer Start: 14559	Coding Potential	No	
GeneMark?		GeneMark Start: 14559			<u>]</u>

Ribosome Binding Site

RBS Final Score		Score Closest to 0: -4.219
-4.957	Closest to 0?	Start #: 16413
	Yes / No (if no)	
		ORF Length:9



Phage Name: Bush

Student Name: Betty Sierra/Grace Int.

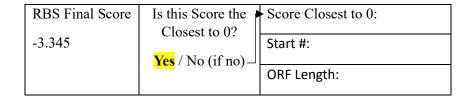
Date:1/30/24

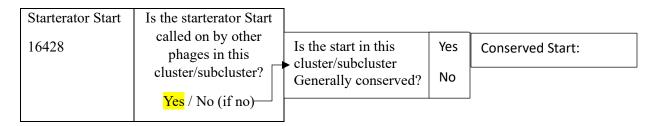
Gene #	Original 5' End Start	Original 3' End Stop	Original Length	Original Start Codon	
26	16428	16736	309	ATG / GTG / TTG	
Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?				
FWD / REV	Yes / No (if no, answer below)				
	Revised 5' End Start		New Length	New Start Codon	
				ATG / GTG / TTG	

Glimmer and GeneMark Calls

Start called by both	Yes	Start: 16428		Does this Gene have	Yes	Estimated Start:
Glimmer and	No	Glimmer Start: 16428	_	Coding Potential	No	
GeneMark?		GeneMark Start: 16428				

Ribosome Binding Site





Phage Name: Bush

Student Name: Betty Sierra/Grace Int.

Date:1/30/24

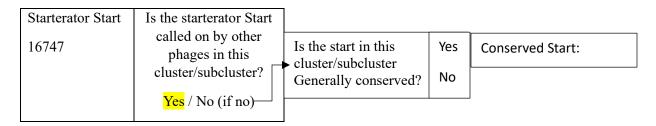
Gene #	Original 5' End Start	Original 3' End Stop	Original Length	Original Start Codon		
27	16747	17481	735	ATG / GTG / TTG		
Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?					
FWD / REV	Yes / No (if no, answer below)					
	Revised 5' End Start		New Length	New Start Codon		
				ATG / GTG / TTG		

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	<mark>Yes</mark> No	Start: 16747 Glimmer Start: 16747	7	Does this Gene have Coding Potential	<mark>Yes</mark> No	Estimated Start:
		GeneMark Start: 16747				

Ribosome Binding Site

RBS Final Score	Is this Score the	Score Closest to 0: -4.205
-5.913	Closest to 0? Yes / No (if no)	Start #: 16870
	105 / 110 (II IIe)	ORF Length: 612



Phage Name: Bush

Student Name: Betty Sierra/Grace Int.

Date:1/30/24

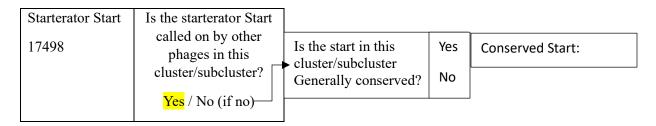
Gene #	Original 5' End Start	Original 3' End Stop	Original Length	Original Start Codon		
28	17498	18319	822	ATG / GTG / TTG		
Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?					
FWD / REV	Yes / No (if no, answer below)					
	Revised 5' End Start		New Length	New Start Codon		
				ATG / GTG / TTG		

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	Yes No	Start: 17498 Glimmer Start: 17498		Does this Gene have Coding Potential	<mark>Yes</mark> No	Estimated Start:
		GeneMark Start: 17498	'			_

Ribosome Binding Site

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



Phage Name: Bush

Student Name: Betty Sierra/Grace Int.

Date:1/30/24

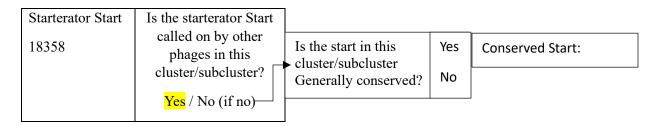
Gene #	Original 5' End Start	Original 3' End Stop	Original Length	Original Start Codon	
29	18358	18618	261	ATG / GTG / TTG	
Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?				
FWD / REV	Yes / No (if no, answer below)				
	Revised 5' End Start		New Length	New Start Codon	
				ATG / GTG / TTG	

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	Yes No	Start: 18358 Glimmer Start: 18358	, 	Does this Gene have Coding Potential	<mark>Yes</mark> No	Estimated Start:
		GeneMark Start: 18358				

Ribosome Binding Site

RBS Final Score	Is this Score the	Score Closest to 0: -5.115
-5.913	Closest to 0? Yes / No (if no)	Start #: 18316
	105 / 110 (II IIe)	ORF Length: 303



Phage Name: Bush

Student Name: Arianna Marie Ruiz

Date: 1/23/2024

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

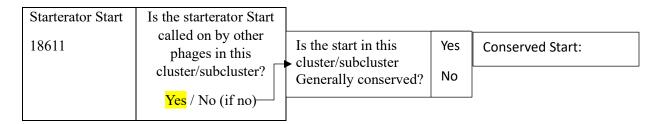
Glimmer and GeneMark Calls

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

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

Ribosome Binding Site

RBS Final Score	Is this Score the	Score Closest to 0: -2.837
-3.948	Closest to 0? Yes / No (if no)	Start #: 19142
	1637 <mark>110</mark> (II IIo) –	ORF Length: 363



Phage Name: Bush

Student Name: Arianna Marie Ruiz

Date: 1/23/2024

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

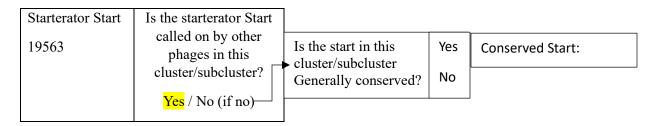
Glimmer and GeneMark Calls

Start called by both	Yes	Start: 19563	Does this Gene have
Glimmer and	No	Glimmer Start:	Coding Potentia
GeneMark?		GeneMark Start:	

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

Ribosome Binding Site

RBS Final Score		Score Closest to 0:
-3.749	Closest to 0?	Start #:
	Yes / No (if no) ☐	ORF Length:



Phage Name: Bush

Student Name: Arianna Marie Ruiz

Date: 1/30/2024

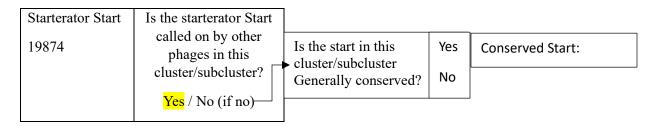
Gene #	Original 5' End Start	Original 3' End Stop	Original Length	Original Start Codon			
32	19874	20539	666	ATG / <mark>GTG</mark> / TTG			
Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?						
FWD / REV	Yes / No (if no, answer below)						
	Revised 5' End Start		New Length	New Start Codon			
				ATG / GTG / TTG			

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	Yes <mark>No</mark>	Start: Glimmer Start: 19874	٦	Does this Gene have Coding Potential	<mark>Yes</mark> No	Estimated Start:
		GeneMark Start: 20539				

Ribosome Binding Site

RBS Final Score	Is this Score the	Score Closest to 0: -3.031
-3.977	Closest to 0? Yes / No (if no)	Start #: 20123
	1657 (11 lie)	ORF Length: 417



Phage Name: Bush

Student Name: Arianna Marie Ruiz

Date: 1/30/2024

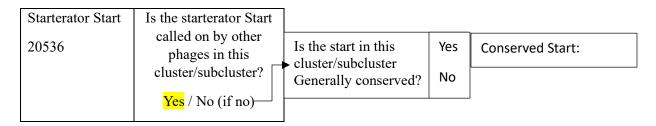
Gene #	Original 5' End Start	Original 3' End Stop	Original Length	Original Start Codon			
33	20554 21117		564	ATG / GTG / TTG			
Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?						
FWD / REV	Yes / No (if no, answer below)						
	Revised 5' End Start		New Length	New Start Codon			
	20536		582	ATG / GTG / TTG			

Glimmer and GeneMark Calls

Start called by both Glimmer and	Yes No	Start: 20554 Glimmer Start:	Does this Gene have Coding Potential	Yes No	Estimated Start: 20536
GeneMark?		GeneMark Start:			Ш

Ribosome Binding Site

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



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.

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		

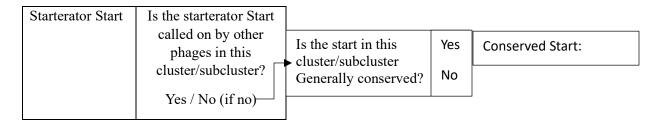
Glimmer and GeneMark Calls

Start called by both	Yes	Start:	Do Ge
Glimmer and	<mark>No</mark>	Glimmer Start: 21176	Со
GeneMark?		GeneMark Start: N/A	

Does this Gene have	Yes	Estimated Start:
Coding Potential	No	

Ribosome Binding Site

RBS Final Score		Score Closest to 0:
	Closest to 0?	Charles II
	77 / 27 (°C)	Start #:
	Yes / No (if no)	ODE Longthy
		ORF Length:



Phage Name: Bush

Student Name: Arianna Marie Ruiz

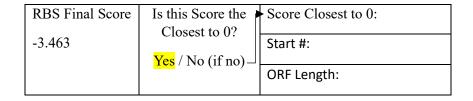
Date: 1/30/2024

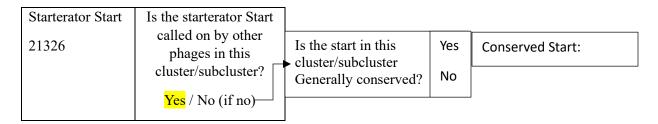
Gene #	Original 5' End Start	Original 3' End Stop	Original Length	Original Start Codon		
35	21302	21625	324	ATG / GTG / TTG		
Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?					
FWD / REV	Yes / No (if no, answer below)					
	Revised 5' End Start		New Length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

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?		Does this Gene have Coding Potential	<mark>Yes</mark> No	Estimated Start: 21326
--	--	--------------------------------------	------------------------	------------------------

Ribosome Binding Site





Comments

According to GeneMaster, the coding potential begins around the range of \sim 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

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		

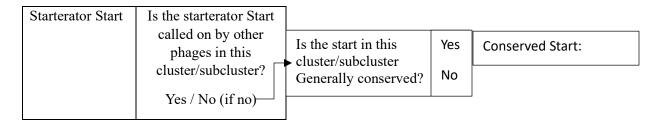
Glimmer and GeneMark Calls

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

Does this Gene have	Yes	Estimated Start:
Coding Potential	No	

Ribosome Binding Site

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



Comments

No data

Phage Name: Bush

Student Name: Arianna Marie Ruiz

Start: 22080

Date: 1/30/2024

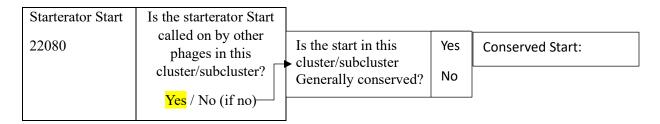
Gene #	Original 5' End Start	Original 3' End Stop	Original Length	Original Start Codon	
37	22080	22526	447	ATG / GTG / TTG	
Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?				
FWD / REV	Yes / No (if no, answer below)				
	Revised 5' End Start		New Length	New Start Codon	
				ATG / GTG / TTG	

Glimmer and GeneMark Calls

Start called by both	Yes	Start: 22080	Does this Gene have	Yes	Estimated
Glimmer and	No	Glimmer Start:	Coding Potential	No	
GeneMark?		GeneMark Start:			

Ribosome Binding Site

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



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		
38	22523	22681	159	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		

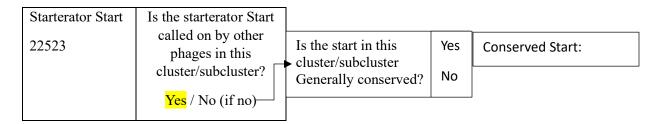
Glimmer and GeneMark Calls

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

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

Ribosome Binding Site

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



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		

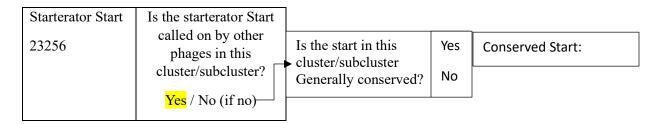
Glimmer and GeneMark Calls

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

Does this Gene have	Yes	Estimated Start:
Coding Potential	No	

Ribosome Binding Site

RBS Final Score		Score Closest to 0:
-2.812	Closest to 0?	Start #:
	Yes / No (if no) ☐	ORF Length:



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		
41	24344	25393	1050	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		

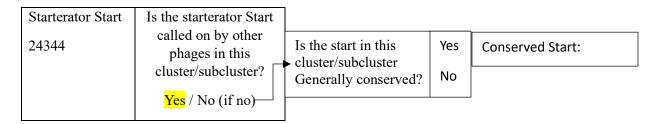
Glimmer and GeneMark Calls

Start called by both	Yes	Start: 24344	
Glimmer and	No	Glimmer Start:	
GeneMark?		GeneMark Start:	<u> </u>

Does this Gene have	Yes	Estimated Start:
Coding Potential	No	

Ribosome Binding Site

RBS Final Score		Score Closest to 0:
-2.812	Closest to 0?	Start #:
	Yes / No (if no) ☐	ORF Length:



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

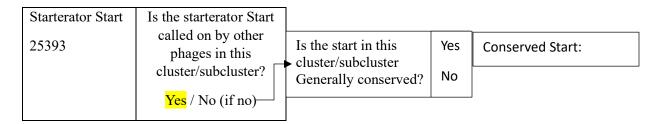
Glimmer and GeneMark Calls

Start called by both	Yes	Start: 25393	Does t
Glimmer and	No	Glimmer Start:	Codin
GeneMark?		GeneMark Start:	

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

Ribosome Binding Site

RBS Final Score	Is this Score the	Score Closest to 0: -5.082
-5.721	Closest to 0? Yes / No (if no)	Start #: 25741
	1657 rto (II lie)	ORF Length: 15



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

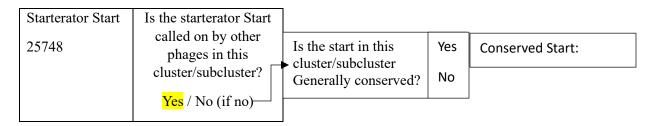
Glimmer and GeneMark Calls

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

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

Ribosome Binding Site

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



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	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	Yes	Start: 25748		
by both				
Glimmer and	No	Glimmer Start:		
GeneMark?		GeneMark Start:		

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

Ribosome Binding Site

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

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

Phage Name: Bush Student Name: Kristen Mclean

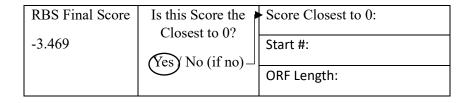
Date: 1/30/2024

Gene #	Original 5' End Start	Original 3' End Stop	Original Length	Original Start Codon		
44	26151	26984	834	ATO / 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	Yes	Start: 26151	Does this Gene have	Yes	Estimated Start: 26151
Glimmer and	No	Glimmer Start:	Coding Potential	No	
GeneMark?		GeneMark Start:			

Ribosome Binding Site



Starterator Start	Is the starterator Start			
26151	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:
	Yes/ No (II no)—			

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	Yes	Start: 26981	Does this Gene have	Yes	Estimated Start: 26981
Glimmer and	No	Glimmer Start:	Coding Potential	No	
GeneMark?		GeneMark Start:			

Ribosome Binding Site

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

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

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	Yes	Start: 27316	
Glimmer and	No	Glimmer Start:	
GeneMark?		GeneMark Start:	

Does this Gene have	Yes	Estimated Start: 27316
Coding Potential	No	

Ribosome Binding Site

RBS Final Score		Score Closest to 0: -2.918
-5.266	Closest to 0?	Start #: 27655
	Yes (No)(if no)	ORF Length: 966

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

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	Yes	Start: 28681	Does this Gene have	Yes	Estimated Start: 28681
Glimmer and	No	Glimmer Start:	Coding Potential	No	
GeneMark?		GeneMark Start:			

Ribosome Binding Site

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

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

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		
48	29269	29802	534	(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	Yes	Start: 29269	Does this Gene have	Yes	Estimated Start: 29269
Glimmer and	No	Glimmer Start:	Coding Potential	No	
GeneMark?		GeneMark Start:			

Ribosome Binding Site

RBS Final Score	Is this Score the	Score Closest to 0: -4.825
-7.028	Closest to 0? Yes (No)(if no)	Start #: 29536
	138 (119(11110)	ORF Length: 267

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

Phage Name: Bush

Student Name: Kristen Mclean

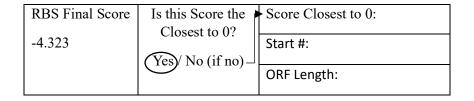
Date: 1/30/2024

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	Yes	Start: 29795	Does this Gene have	Yes	Estimated Start: 29795
Glimmer and	No	Glimmer Start:	Coding Potential	No	
GeneMark?		GeneMark Start:			

Ribosome Binding Site



Starterator Start	Is the starterator Start			
29795	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:
	Tes)/ No (II IIo)			

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
50	30579	30905	327	(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	Yes	Start: 30579	Does this Gene have	Yes	Estimated Start: 30579
Glimmer and	No	Glimmer Start:	Coding Potential	No	
GeneMark?		GeneMark Start:			

Ribosome Binding Site

RBS Final Score	Is this Score the	Score Closest to 0:
-2.584	Closest to 0? Yes/ No (if no)	Start #:
	110 (11110)	ORF Length:

Starterator Start	Is the starterator Start			
30579	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:
	Tes/ No (II IIo)			

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

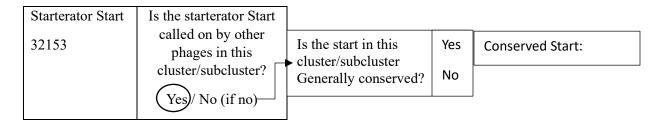
Glimmer and GeneMark Calls

Start called by both	Yes	Start: 32153		D
Glimmer and	No	Glimmer Start:		C
GeneMark?		GeneMark Start:	L	

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

Ribosome Binding Site

RBS Final Score	Is this Score the	Score Closest to 0:
2 422	Closest to 0?	
-3.432		Start #:
	Yes/ No (if no)	
		ORF Length:
		_



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	Yes	Start: 32542	Does this Gene have	Yes	Estimated Start: 32510
Glimmer and	(No)	Glimmer Start: 32542	Coding Potential	No	
GeneMark?		GeneMark Start:			

*Not reported in GeneMark

Ribosome Binding Site

RBS Final Score	Is this Score the	Score Closest to 0: -5.280
-6.655	Closest to 0? Yes (No)(if no)	Start #: 32608
	165 (119(11110)	ORF Length:210

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

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	
54	32814	32957	144	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	

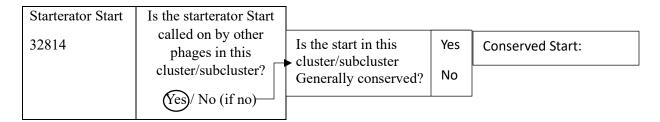
Glimmer and GeneMark Calls

Start called by both	Yes	Start: 32814		Does 1
Glimmer and	No	Glimmer Start:		Gene Codin
GeneMark?		GeneMark Start:	L	

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

Ribosome Binding Site

RBS Final Score		Score Closest to 0: -5.040
-6.747	Closest to 0?	Start #: 32058
0.7.17	Yes (No)(if no)	
		ORF Length: 873



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	(Yes)	Start:		Does this
by both			J 1	Gene have
Glimmer and	No	Glimmer Start:		Coding Potential
GeneMark?		GeneMark Start:		

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

Ribosome Binding Site

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

Starterator Start	Is the starterator Start			
32954	called on by other phages in this cluster/subcluster?	Is the start in this cluster/subcluster Generally conserved?	Yes No	Conserved Start:
	Yes / No (if no)			

Phage Name: Bush

Student Name: Nicole Gonzalez G.

Date:1/23/24

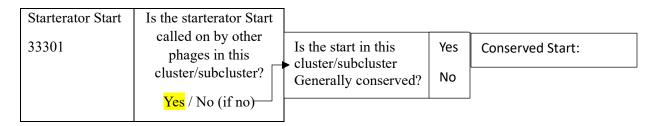
Gene #	Original 5' End Start	Original 3' End Stop	Original Length	Original Start Codon	
56	33277	33627	351	ATG / GTG / TTG	
Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?				
FWD / REV	Yes / No (if no, answer below)				
	Revised 5' End Start		New Length	New Start Codon	
	33301		327	ATG / GTG / TTG	

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	Yes <mark>No</mark>	Start: Glimmer Start: 33277	٦	Does this Gene have Coding Potential	<mark>Yes</mark> No	Estimated Start: 33301
		GeneMark Start:33301				

Ribosome Binding Site

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



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.

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	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	

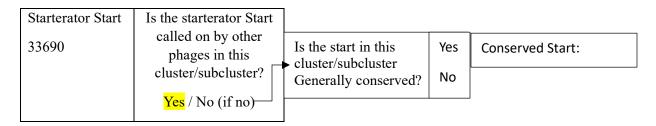
Glimmer and GeneMark Calls

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

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

Ribosome Binding Site

RBS Final Score	Is this Score the	Score Closest to 0: -4.390
-4.408	Closest to 0? Yes / No (if no)	Start #: 34074
	168 / <mark>140</mark> (11 110)—	ORF Length: 12



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	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	

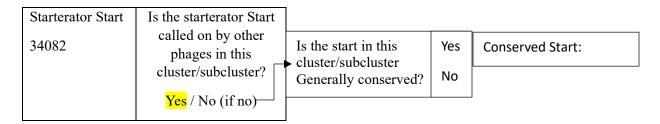
Glimmer and GeneMark Calls

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

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

Ribosome Binding Site

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



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
59	34494	35009	516	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

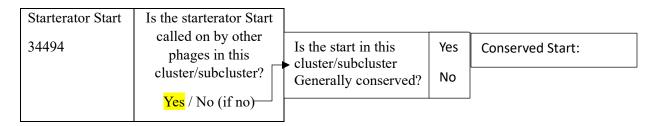
Glimmer and GeneMark Calls

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

Does this Gene have	Yes	Estimated Start:
Coding Potential	No	

Ribosome Binding Site

RBS Final Score	Is this Score the	Score Closest to 0:
-2.373	Closest to 0? Yes / No (if no)	Start #:
	res / Ive (II lie)	ORF Length:



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	
60	35006	35482	477	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	

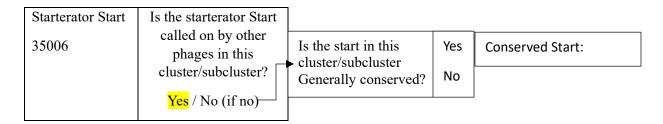
Glimmer and GeneMark Calls

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

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

Ribosome Binding Site

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



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	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

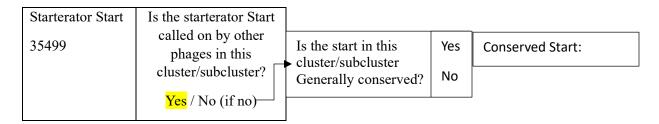
Glimmer and GeneMark Calls

Start called	<mark>Yes</mark>	Start:35499
by both		
Glimmer and	No	Glimmer Start:
GeneMark?		GeneMark Start:

Does this Gene have	Yes	Estimated Start:
Coding Potential	No	

Ribosome Binding Site

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



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	

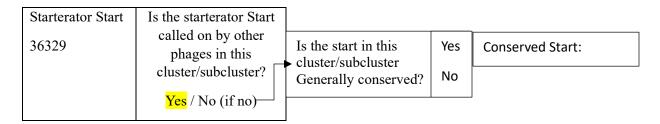
Glimmer and GeneMark Calls

Start called by both	Y <mark>es</mark>	Start:36329
Glimmer and	No	Glimmer Start:
GeneMark?		GeneMark Start:

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

Ribosome Binding Site

RBS Final Score	Is this Score the	Score Closest to 0: -4.716
-4.844	Closest to 0?	Start #: 36683
	Yes / <mark>No</mark> (if no) ☐	ORF Length:369



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	
63	37036	37386	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/G <mark>TG</mark> /TTG	

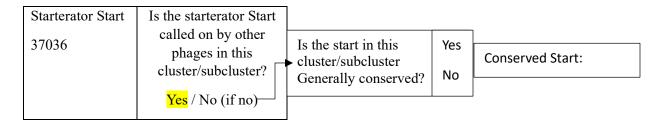
Glimmer and GeneMark Calls

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

Does this Gene have	Y <mark>es</mark>	Estimated Start:37036
Coding Potential	No	

Ribosome Binding Site

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



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	

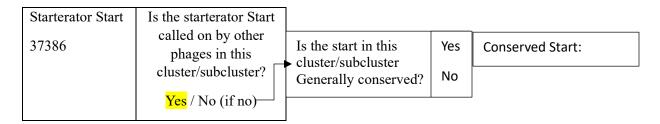
Glimmer and GeneMark Calls

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

Does this Gene have	Y <mark>es</mark>	Estimated Start:37386
Coding Potential	No	

Ribosome Binding Site

RBS Final Score	Is this Score the	Score Closest to 0:
-4.319	Closest to 0? Yes / No (if no)	Start #:
	105 / 110 (II IIO)—	ORF Length:



Phage Name: Bush

Student Name: Nicole Gonzalez G.

Date: 2/01/24

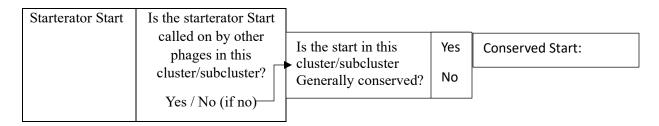
Gene #	Original 5' End Start	Original 3' End Stop	Original Length	Original Start Codon			
65	37906	37906 38307		ATG / GTG / TTG			
Gene Direction	Gene Start, Leng	gth, and Start Codon Sar	me as original (Auto	Annotated) call?			
FWD / REV	Yes / No (if no, answer below)						
	Revised 5' End Start		New Length	New Start Codon			
	37894		414	ATG / GTG / TTG			

Glimmer and GeneMark Calls

Start called by both Glimmer and GeneMark?	Yes <mark>No</mark>	Start: Glimmer Start: 37906 GeneMark Start: 37894		Does this Gene have Coding Potential	<mark>Yes</mark> No	Estimated Start: 378
---	------------------------	---	--	--------------------------------------	------------------------	----------------------

Ribosome Binding Site

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



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.

Phage Name: Bush

Student Name: Nicole Gonzalez G.

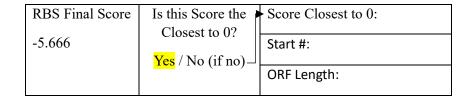
Date: 2/01/24

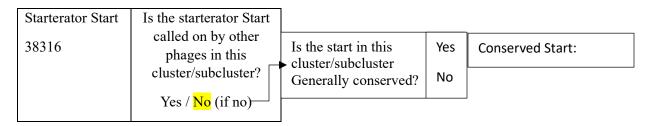
Gene #	Original 5' End Start	Original 3' End Stop	Original Length	Original Start Codon			
66	38316	38591	276	ATG / GTG / TTG			
Gene Direction	Gene Start, Length, and Start Codon Same as original (Auto Annotated) call?						
FWD / REV	Yes / No (if no, answer below)						
	Revised 5' End Start		New Length	New Start Codon			
				ATG / GTG / TTG			

Glimmer and GeneMark Calls

Start called by both Glimmer and	Yes No	Start: Glimmer Start:38316	Does this Gene have Coding Potential	<mark>Yes</mark> No	Estimated Start:38316
GeneMark?		GeneMark Start:38304		<u> </u>	

Ribosome Binding Site





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.

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			

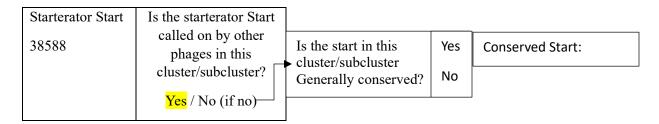
Glimmer and GeneMark Calls

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

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

Ribosome Binding Site

RBS Final Score		Score Closest to 0:
-5.485	Closest to 0? Yes / No (if no)	Start #:
	168 / No (II IIo) —	ORF Length:



Function (F) Tape measure	Gene No: 23			Stude	nt name: Betty Sierra	a / Grace Intrator
protein		Date: 2/13/			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	<u>% Query</u> <u>99</u>	<u>% Ider</u>		<u>E-value</u> (<u>if < 10⁻⁷)</u> <u>0</u>	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		<u>% Ider</u>		<u>E-value</u> (<u>if < 10⁻⁷)</u> <u>0</u>	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	<u>Hit</u> 6V8I_		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 numb number of m 1414009	<u>nembers</u>		pstream pham conserved ? Yes or No s, what pham # or function ? 73	Downstream pham conserved? Yes or No If yes, what pham # or function? 140921
Transmembrane	domains by Tml	Hmm using Pha	merator	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)	Gene No: 24		nt name: Betty Sierra	a / Grace Intrator			
Minor tail protein				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	<u>% Query</u> <u>99</u>	<u>% Ider</u>		<u>E-value</u> (<u>if < 10⁻⁷)</u> <u>0</u>	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		<u>% Ider</u> <u>99</u>	-	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) Hypothetica I protein 19.1	<u>Hit</u> 2X8K		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 numb number of m 140921 (nembers		pstream pham conserved ? Yes or No s, what pham # or function ?	Downstream pham conserved? Yes or No If yes, what pham # or function? 142684	
Transmembrane	domains by Tm	Hmm using Pha	merator	no			

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

Function (F)	Gene No: 25			Student name: Betty Sierra / Grace Intrator				
Minor tail protein			Date: 2/			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	<u>% Query</u> <u>99</u>	<u>% Ider</u> <u>91.4</u>		<u>E-value</u> (<u>if < 10⁻⁷)</u> <u>0</u>	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		<u>% Ider</u>		<u>E-value</u> (<u>if < 10⁻⁷)</u> <u>0</u>	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	<u>Hit</u> PF1459		Probability (if > 90%) 99.95	<u>E-value</u> (<u>if < 1)</u> <u>9.5e-24</u>		
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 numb number of m 142684	<u>nembers</u>		pstream pham conserved ? Yes or No s, what pham # or function ?	Downstream pham conserved? Yes or No If yes, what pham # or function? 3528		
Transmembrane	domains by Tml	Hmm using Pha	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	Gene No: 26			Student name: Betty Sierra / Grace Intrator		
protein		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	<u>% Query</u> <u>99</u>	<u>% Iden</u>		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)	<u>Function</u> (or <mark>NKF</mark>)		<u>% Iden</u> <u>99</u>		E-value (if < 10 ⁻⁷) <u>5e-53</u>	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)	<u>Function</u> (or NKF)	Name / Descr (incl DUF)	Hit		Probability (if > 90%) All Probabilities below 90	<u>E-value</u> (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 numb number of m 3528 (2	<u>nembers</u>		pstream pham conserved ? Yes or No s, what pham # or function ?	Downstream pham conserved? Yes or No If yes, what pham # or function? 141462
Transmembrane	Transmembrane domains by TmHmm using Phamerator					

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

Function (F)	Gene No: 27 Student name: Betty Sierra / Grace Intrator					
Minor tail protein			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	<u>% Query</u> <u>99</u>	<u>% Ider</u> <u>99.1</u>		<u>E-value</u> (<u>if < 10⁻⁷)</u> <u>1e-169</u>	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		<u>% Ider</u>	-	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	<u>Hit</u>		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)	<u>Function</u> (or <mark>NKF</mark>)	Pham numb number of m 141462	nembers		pstream pham conserved ? Yes or No s, 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	Gene No: 28				Student name: Betty Sierra / Grace Intrator		
protein				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	<u>% Query</u> <u>99</u>	<u>% Ider</u>		<u>E-value</u> (<u>if < 10⁻⁷)</u> <u>0</u>	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 <mark>NKF</mark>)		<u>% Ider</u>		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)	<u>Function</u> (or <mark>NKF</mark>)	Name / Descr (incl DUF) DUF859 sipho virus protein of unknown function	<u>Hit</u> PF0589		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 numb number of m 143895 (<u>nembers</u>		pstream pham conserved ? Yes or No s, 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	Gene No: 29				Student name: Betty Sierra / Grace Intrator		
protein				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	<u>% Query</u> <u>98</u>	<u>% Ider</u>		<u>E-value</u> (<u>if < 10⁻⁷)</u> <u>7e-49</u>	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 <mark>NKF)</mark>		<u>% Ider</u>	-	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)	<u>Hit</u>		Probability (if > 90%) All probabilities below 90	<u>E-value</u> (<u>if < 1)</u>	
Supporting Information for Function, from Syntney (SIF-Syn) (All three lines of SIF evidence must be filled in)	<u>Function</u> (or <mark>NKF</mark>)	Pham numb number of m 141167	<u>nembers</u>		pstream pham conserved ? Yes or No s, what pham # or function ?	Downstream pham conserved? Yes or No If yes, what pham # or function?	
Transmembrane	domains by Tml	Hmm using Pha	no				

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

Function (F)	Gene No: 30				Student name: Arianna Marie Ruiz			
endolysin		Date: 2/6/2024						
Supporting Information for Function, from BLAST (Protein) (BLAST-	Function (or NKF) Lysin A	<u>% Query</u> 99%	<u>% Iden</u> 100.0		<u>E-value</u> (if < 10 ⁻⁷) 0.0	AA # conserved in first 10 hits? Yes or No		
NCBI) (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) Lysin A		<u>% Iden</u>	-	<u>E-value</u> (<u>if < 10⁻⁷)</u> 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- carboxypepti dase; CENTER FOR STRUCTURAL GENOMICS OF INFECTIOUS DISEASES, CSGID, NATIONAL INSTITUTE O	Hit 4MUC		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 numb number of m 98619 (1	<u>embers</u>		pstream pham conserved ? Yes or No s, what pham # or function ?	<u>Downstream pham</u> <u>conserved ?</u> <u>Yes</u> or No		

(All three lines of SIF evidence must be filled in)			147123 (63)	If yes, what pham # or function ? 8687 (10)
Transmembrane	e domains by Tn	nHmm using Phamerator	1	

Function (F)	Gene No: 31			Stude	nt name: Arianna Ma	arie Ruiz
membrane protein				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	<u>% Query</u> 99%	<u>% Ider</u> 100.0		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		<u>% Ider</u>		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		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 numb number of m 8687 (9	<u>embers</u>		pstream pham conserved? Yes or No s, what pham # or function ? 98619 (138)	Downstream pham conserved? Yes or No If yes, what pham # or function? 8647 (9)
Transmembrane	domains by Tn	nHmm using Pha	merator	yes, b	y 2 programs	

Function (F)	Gene No: 32			Stude	nt name: Arianna Ma	arie Ruiz	
membrane protein				Date:	e: 2/14/2024		
Supporting	<u>Function</u>	<u>% Query</u>	<u>% Iden</u>	<u>itity</u>	<u>E-value</u>	AA # conserved in	
Information	(or NKF)				(if < 10 ⁻⁷)	first 10 hits?	
for Function,	NIEW	96%	99.53	3%			
from BLAST	NFK				2e-147	Yes or <mark>No</mark>	
(Protein)							
(BLAST-							
NCBI) (All three lines of SIF evidence must be filled in)							
Supporting	Function		<u>% Iden</u>	<u>itity</u>	<u>E-value</u>	AA # conserved in	
Information	(or NKF)		000	,	<u>(if < 10⁻⁷)</u>	first 10 hits?	
for Function, from BLAST	NFK		98%	o .	e-123	Yes or No	
(Protein)	IVIIX				6-125	TES OF NO	
(BLAST-							
phagesDB)							
(All three lines of SIF evidence							
must be filled							
in)							
Supporting	<u>Function</u>	Name / Descr	Hit	:	Probability (if >	<u>E-value</u>	
Information	(or NKF)	(incl DUF)			<u>90%)</u>	<u>(if < 1)</u>	
for Function,							
from HHpred					*No probability		
(HHpred)					above 90%		
(All three lines of SIF							
evidence must be filled in)							
,	Eumotion	Dham areas	orand		nstroom phose	Downstroom where	
Supporting Information	<u>Function</u> (or NKF)	Pham numb number of m		<u>u</u>	pstream pham conserved ?	<u>Downstream pham</u> <u>conserved?</u>	
for Function,	(3)	<u></u>			Yes or No	Yes or No	
from Syntney	NFK					<u> </u>	
(SIF-Syn)		8647 (9	9)	If yes	s, what pham # or	If yes, what pham # or	
(All three lines of SIF					function?	function?	
evidence must be					8687 (9)	5847 (14)	
filled in) Transmembrane	domains by Tn	l nHmm using Pha	ımerator	yes, b	y 2 programs		

Function (F)	Gene No: 33			Stude	nt name: Arianna Ma	arie Ruiz
holin				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	<u>% Query</u> 99%	<u>% lder</u> 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		<u>% Ider</u> 90%		<u>E-value</u> (<u>if < 10⁻⁷)</u> 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; Mycobacteri al 2 TMS Phage Holin (M2 Hol) Family	<u>Hit</u>		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 numb number of m 5847 (1	<u>embers</u>		pstream pham conserved ? Yes or No s, what pham # or function ? 8647 (9)	Downstream pham conserved? Yes or No If yes, what pham # or function?
Transmembrane	domains by Tn	nHmm using Pha	merator			

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

Function (F)	Gene No: 34			Stude	nt name: Arianna Ma	arie Ruiz
hypothetical protein				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	<u>% Query</u> 97%	<u>% Ider</u> 75.62		<u>E-value</u> (<u>if < 10⁻⁷)</u> 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		<u>% Ider</u> 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%	<u>E-value</u> (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 numb number of m 14633 (<u>embers</u>		pstream pham conserved ? Yes or No s, what pham # or function ? 5847 (14)	Downstream pham conserved? Yes or No If yes, what pham # or function?
Transmembrane	domains by Tn	nHmm using Pha	merator			

Function (F)	Gene No: 35			Stude	nt name: Arianna Ma	arie Ruiz
Hypothetical protein				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	<u>% Query</u> 99%	<u>% Ider</u> 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		<u>% Ider</u> 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)	<u>Hit</u>	:	Probability (if > 90%) *No probability above 90%	<u>E-value</u> (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 numb number of m 88323 (<u>embers</u>		pstream pham conserved ? Yes or No s, what pham # or function ?	Downstream pham conserved? Yes or No If yes, what pham # or function?
Transmembrane	domains by Tn	nHmm using Pha	merator			

Notes: Only BLAST NCBI was calling hypothetical protein.

Function (F)	Gene No: 36			Stude	nt name: Arianna Ma	arie Ruiz
Hypothetical protein				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	<u>% Query</u> N/A	<u>% Ider</u> N/ <i>F</i>		<u>E-value</u> (<u>if < 10⁻⁷)</u> 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		<u>% Ider</u> 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)	<u>Hit</u>	:	Probability (if > 90%) *No probability above 90%	<u>E-value</u> (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 numb number of m 137359	<u>embers</u>		pstream pham conserved ? Yes or No s, what pham # or function ?	Downstream pham conserved? Yes or No If yes, what pham # or function? 8611 (9)
Transmembrane	domains by Tn	nHmm using Pha	merator			

Notes: No similarities found in BLAST NCBI

Function (F)	Gene No: 37			Stude	nt name: Arianna Ma	arie Ruiz
hypothetical protein				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	<u>% Query</u> 99%	<u>% Ider</u> 68.46		<u>E-value</u> (<u>if < 10⁻⁷)</u> 2e-61	AA # conserved in first 10 hits? Yes or <mark>No</mark>
Supporting Information for Function, from BLAST (Protein) (BLAST- phagesDB) (All three lines of SIF evidence must be filled in)	Function (or NKF) NFK		<u>% Ider</u> 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%	<u>E-value</u> (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 numb number of m 8611 (9	<u>embers</u>		pstream pham conserved ? Yes or No s, what pham # or function ?	Downstream pham conserved? Yes or No If yes, what pham # or function? 8178 (9)
Transmembrane	domains by Tn	nHmm using Pha	merator			

Function (F)	Gene No: 38			Stude	nt name: Arianna Ma	arie Ruiz
membrane protein				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	<u>% Query</u> 98%	<u>% Ider</u>		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		<u>% Ider</u>		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		Probability (if > 90%) 91.1%	<u>E-value</u> (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 numb number of m 8178 (9	<u>embers</u>		pstream pham conserved ? Yes or No s, what pham # or function ? 8611 (9)	Downstream pham conserved? Yes or No If yes, what pham # or function? 146719 (646)
Transmembrane	domains by Tn	nHmm using Pha	merator	yes, 2	domains	

Function (F) RuvC-like	Gene No: 39				nt name: Arianna M 2/14/2024	arie Ruiz
resolvase		1				
Supporting	Function	<u>% Query</u>	<u>% Iden</u>	tity	<u>E-value</u>	AA # conserved in
Information	(or NKF)				(if < 10 ⁻⁷)	first 10 hits?
for Function,	RuvC-like	99%	98.96	5%		
from BLAST	resolvase				<u>1e-136</u>	Yes or <mark>No</mark>
(Protein)						
(BLAST-						
NCBI) (All three lines of SIF evidence must be filled in)						
Supporting	<u>Function</u>		<u>% Iden</u>	tity	<u>E-value</u>	AA # conserved in
Information	(or NKF)		000	,	(if < 10 ⁻⁷)	first 10 hits?
for Function, from BLAST	RuvC-like		98%	Ď	e-104	<u>Yes or <mark>No</mark></u>
(Protein)	resolvase				e-104	res or ivo
(BLAST-						
phagesDB)						
(All three lines						
of SIF evidence must be filled						
in)						
,	<u>Function</u>	Name / Descr	Hit		Probability (if >	<u>E-value</u>
	(or NKF)	(incl DUF)			90%)	(if < 1)
		Crossover	7XHJ_	_A	<u> </u>	<u>, , , , , , , , , , , , , , , , , , , </u>
6	RuvC-like	junction			99.93%	5.4e-23
Supporting	resolvase	endodeoxyri				
Information		bonuclease				
for Function,		RuvC;				
from HHpred		nuclease,				
(HHpred) (All three lines of SIF		DNA repair,				
evidence must be		Homologous				
filled in)		recombinatio				
		n, Holliday				
		junction				
		resolva				
Supporting	<u>Function</u>	Pham numb		u	pstream pham	Downstream pham
Information	(or NKF)	number of m	<u>embers</u>		conserved ?	conserved ?
for Function,		1467107	=46)		<mark>Yes</mark> or No	<mark>Yes</mark> or No
from Syntney	RuvC-like	146719 (6	040)	If ve	s, what pham # or	If you what above 4 co
(SIF-Syn) (All three lines of SIF evidence must be filled in)	resolvase			ii ye.	function ? 8178 (9)	<u>If yes, what pham # or</u> <u>function ?</u> 45958 (59)

Transmembrane domains by TmHmm using Phamerator	

Function (F)	Gene No: 40 Student name: Arianna Marie Ruiz						
exonuclease				Date:	2/14/2024		
Supporting Information for Function, from BLAST (Protein) (BLAST- NCBI) (All three lines of SIF evidence must be	Function (or NKF) exonuclease	<u>% Query</u> 99%	% Iden		<u>E-value</u> (if < 10 ⁻⁷) 0.0	AA # conserved in first 10 hits? Yes or No	
filled in) Supporting Information for Function, from BLAST (Protein) (BLAST- phagesDB) (All three lines of SIF evidence must be filled in)	Function (or NKF) exonuclease		<u>% Iden</u>		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) exonuclease	Name / Descr (incl DUF) Exonuclease; alkaline exonuclease, digest double stranded DNA, strict 5- 3-polarity, HYDROLASE; 1.9A {Laribacter hongkong	<u>Hit</u>		Probability (if > 90%) 99.88%	E-value (if < 1) 1.1e-20	
Supporting Information for Function, from Syntney (SIF-Syn) (All three lines of SIF evidence must be filled in)	Function (or NKF) exonuclease	Pham numb number of m 45958 (5	<u>embers</u>	<u>If ye</u>	pstream pham conserved ? Yes or No s, what pham # or function ? 19 (646), RuvC-like resolvase	Downstream pham conserved? Yes or No If yes, what pham # or function? 87972 (9)	

Transmembrane domains by TmHmm using Phamerator	

Function (F)	Gene No: 41				Student name: Arianna Marie Ruiz			
RrecT-like DNA	pairing protei	n		Date: 2/14/2024				
Supporting Information for Function, from BLAST	Function (or NKF) RecT-like DNA pairing	<u>% Query</u> 99%	<u>% Iden</u> 99.43		<u>E-value</u> (if < 10 ⁻⁷) 0.0	AA # conserved in first 10 hits? Yes or No		
(Protein) (BLAST- NCBI) (All three lines of SIF evidence must be filled in)	protein							
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		<u>% Iden</u> 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) Recombinati on protein bet; Annealase, Synaptase, SSAP, Single- strand annealing protein, DNA annealing intermediate, Reco	Hit 7UJL		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 numb number of m 87972 (embers	<u>If yes</u>	pstream pham conserved ? Yes or No s, what pham # or function ? 3 (59), exonuclease	Downstream pham conserved? Yes or No If yes, what pham # or function? 180 (312)		

Transmembrane domains by TmHmm using Phamerator	

	Gene No: 42			C+udo	nt name: Arianna M	orio Duiz	
Function (F)					Student name: Arianna Marie Ruiz		
hypothetical prrotein				Date:	2/14/2024		
Supporting	Function	<u>% Query</u>	<u>% Iden</u>	tity	<u>E-value</u>	AA # conserved in	
Information	(or NKF)				$(if < 10^{-7})$	first 10 hits?	
for Function,	NFK	99%	99.17	7 %			
from BLAST					1e-78	Yes or <mark>No</mark>	
(Protein)							
(BLAST- NCBI)							
(All three lines of SIF							
evidence must be							
filled in)	5		0/ 1-1		F .1 .	A A #	
Supporting Information	Function (or NKF)		<u>% Iden</u>	tity	<u>E-value</u> (if < 10 ⁻⁷)	AA # conserved in first 10 hits?	
for Function,	(OF IVICE)		1009	%	(11 < 10)	11130 10 11103:	
from BLAST	NFK				1e-62	Yes or No	
(Protein)							
(BLAST-							
phagesDB) (All three lines							
of SIF evidence							
must be filled							
in)							
	Function	Name / Descr	<u>Hit</u>		Probability (if >	<u>E-value</u>	
	(or NKF)	(incl DUF)	4ZWQ	. D	90%)	<u>(if < 1)</u>	
	NFK	Recombinati	42000		07.000/		
Supporting	IVIK	on protein			95.36%	N/A	
Supporting Information		uvsY; recombinatio					
for Function,		n, DNA					
from HHpred		binding,					
(HHpred)		homo-					
(All three lines of SIF		heptamer,					
evidence must be filled in)		asymmetry,					
illed III)		alpha barrel,					
		VIRAL					
		PROTEIN;					
		2.351A {E					
Cupporting	<u>Function</u>	Pham numb	er and	<u>u</u>	pstream pham	Downstream pham	
Supporting Information	(or NKF)	number of m	<u>embers</u>		conserved ?	conserved ?	
for Function,		180 (31	2)		<mark>Yes</mark> or No	<mark>Yes</mark> or No	
from Syntney	NFK	160 (31	-)	If ve	s, what pham # or	If yes, what pham # or	
(SIF-Syn)				<u>,s.</u>	function ?	function ?	
(0,)					87972 (9)		

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

Function (F) HNH endonuclease	Gene No: 43				nt name: Kristen Mc 2/25/2024	lean
Supporting Information for Function, from BLAST (Protein) (BLAST- NCBI) (All three lines of SIF evidence must be filled in)	Function (or NKF) HNH endonucleas e	<u>% Query</u> 99%	<u>% Ider</u> 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 endonucleas e		<u>% Ider</u> 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 endonucleas e	Name / Descr (incl DUF) 5- methylcytosi ne-specific restriction enzyme A; HNH ENDONUCLE ASE	<u>Hit</u>	•	Probability (if > 90%) 93.65	<u>E-value</u> (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 endonucleas e	Pham numb number of m 146731 (5	<u>embers</u>		pstream pham conserved ? Yes or No s, what pham # or function ? 145916 (26)	Downstream pham conserved? Yes or No If yes, what pham # or function? 146810 (265)
Transmembrane	domains by Tm	nHmm using Pha	merator	N/A		

Function (F)	Gene No: 44			Stude	nt name: Kristen Mcl	ean	
Hypothetical prrotein				Date:	2/13/2024		
Supporting Information for Function, from BLAST (Protein) (BLAST-	Function (or NKF) RepA-like replication initiator	<u>% Query</u> 99%	<u>% Iden</u> 99.64	-	<u>E-value</u> (if < 10 ⁻⁷) 0.0	AA # conserved in first 10 hits? Yes or No Only conserved in	
NCBI) (All three lines of SIF evidence must be filled 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		<u>% Iden</u> 99%	-	<u>E-value</u> (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- methylcytosi ne-specific restriction enzyme A; HNH endonucleas e, modification dependent restriction	<u>Hit</u> PF0968		Probability (if > 90%) 98.05	<u>E-value</u> (<u>if < 1)</u> 0.00034	
Supporting Information for Function,	<u>Function</u> (or NKF)	Pham numb number of m 145916 (embers	<u>u</u>	<u>pstream pham</u> conserved ? <mark>Yes</mark> or No	<u>Downstream pham</u> <u>conserved ?</u> <u>Yes</u> or No	

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

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	Function (or NKF) Helicase loader	<u>% Query</u> 99%	<u>% Iden</u> 100%		<u>E-value</u> (if < 10 ⁻⁷) 5e^-73	AA # conserved in first 10 hits? Yes or No Conserved on the first hit
filled in) Supporting Information for Function, from BLAST (Protein) (BLAST- phagesDB) (All three lines of SIF evidence must be filled in)	Function (or NKF) Helicase loader		<u>% Iden</u> 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_lambd a_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 numb number of m 144326 (embers		pstream pham conserved ? Yes or No s, what pham # or function ? 145916 (26)	Downstream pham conserved? Yes or No If yes, what pham # or function?
Transmembrane	domains by Tm	nHmm using Pha	merator	N/A		

Function (F) DNAB-like	Gene No: 46				Student name: Kristen Mclean		
dsDNA helicase				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	<u>% Query</u> %99	<u>% Iden</u> %10	-	<u>E-value</u> (<u>if < 10⁻⁷)</u> 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		<u>% Iden</u>	-	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	<u>Hit</u>		Probability (if > 90%) 100	<u>E-value</u> (<u>if < 1)</u> 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 numb number of m 146810 (2	<u>embers</u>		pstream pham conserved ? Yes or No s, what pham # or function ? 144326 (29)	Downstream pham conserved? Yes or No If yes, what pham # or function? 145287 (506)	
Transmembrane	domains by Tm	nHmm using Pha	merator	N/A			

Function (F)	Gene No: 47			Student name: Kristen Mclean		
SSB protein				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	<u>% Query</u> %99	<u>% Iden</u> %96.:	-	<u>E-value</u> (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		<u>% Iden</u> %99	-	<u>E-value</u> (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, TRANSFERAS E; HET: MSE; 2.6001A	Hit	-	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 numb number of m 145287 (5	<u>embers</u>	<u>If ye</u> :	pstream pham conserved ? Yes or No s, what pham # or function ? 146810 (265)	Downstream pham conserved? Yes or No If yes, what pham # or function? 146844 (215)
Transmembrane	domains by Tm	nHmm using Pha	merator	N/A		

Function (F) PartB-like nuelo	Gene No: 48 case				Student name: Kristen Mclean		
domain				Date: 2/25/2024			
Supporting Information	Function (or NKF)	<u>% Query</u>	<u>% Iden</u>	itity	<u>E-value</u> (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)	ParB-like nuclease domain protein	%99	%100		1e-124	Yes or <mark>No</mark>	
Supporting Information for Function, from BLAST (Protein) (BLAST- phagesDB) (All three lines of SIF evidence must be filled	Function (or NKF) ParB-like nuclease domain protein		<u>% Identity</u> %96		E-value (if < 10 ⁻⁷) 2e-95	AA # conserved in first 10 hits? Yes or No Conserved on the first 8 hits	
in)		/5			2 1 1 1111 /15		
Supporting Information for Function, from HHpred (HHpred) (All three lines of SIF evidence must be filled in)	Function (or NKF) ParB family protein; DNA- binding protein	Name / Descr (incl DUF) ParB family protein; DNA-binding protein, CTP, Myxococcus, DNA- segregation, DNA BINDING PROTEIN; HET: UFQ, GOL; 1.7A	Hit 7BNR_B		90%) 99.52	E-value (if < 1) 1.3e-12	
Supporting Information for Function,	(or NKF)	Pham numb number of m		<u>u</u>	pstream pham conserved ? Yes or No	<u>Downstream_pham</u> <u>conserved ?</u> <u>Yes</u> or No	
from Syntney (SIF-Syn) (All three lines of SIF	ParB-like nuclease domain	146844 (2	215)	If yes	s, what pham # or function ?	If yes, what pham # or function ?	
evidence must be filled in)	protein				145287 (506)	147353 (28)	

Transmembrane domains by TmHmm using Phamerator	N/A

Function (F)					Student name: Kristen Mclean		
acetyltransferas	e			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) acetyltransf erase	<u>% Query</u> %99	<u>% Iden</u> %98		E-value (if < 10 ⁻⁷) 1e-150	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) acetyltransf erase		<u>% Iden</u> %98		<u>E-value</u> (<u>if < 10⁻⁷)</u> e-120	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) acetyltransf erase	Name / Descr (incl DUF) ACETYLTRAN SFERASE; Acyl Coenzyme A complex, TRANSFERAS E; HET: ACO, MSE; 1.55A	<u>Hit</u> 1GHE		Probability (if > 90%) 97.77	<u>E-value</u> (<u>if < 1)</u> 0.0017	
Supporting Information for Function, from Syntney (SIF-Syn) (All three lines of SIF evidence must be filled in)	Function (or NKF) acetyltransf erase	Pham numb number of m 147353 (<u>embers</u>	<u>If ye</u>	pstream pham conserved ? Yes or No s, what pham # or function ? 146844 (215)	Downstream pham conserved? Yes or No If yes, what pham # or function? 8860 (9)	
Transmembrane	domains by Tm	nHmm using Pha	merator	N/A			

Function (F)	Gene No: 50				Student name: Kristen Mclean		
hypothetical prot				Date:	e: 2/25/2024		
Supporting Information	Function (or NKF)	<u>% Query</u>	<u>% Ider</u>	-	<u>E-value</u> (if < 10 ⁻⁷)	AA # conserved in first 10 hits?	
for Function, from BLAST (Protein)	NFK	%99	%99.	07	2e-71	Yes or <mark>No</mark>	
(BLAST- NCBI) (All three lines of SIF evidence must be filled in)						Conserved on the first 3 hits.	
Supporting Information for Function,	Function (or NKF)		<u>% Ider</u> %9		<u>E-value</u> (if < 10 ⁻⁷)	AA # conserved in first 10 hits?	
from BLAST (Protein) (BLAST- phagesDB) (All three lines of SIF evidence must be filled	NKF				1e-58	Yes or No Conserved in the first 5 hits.	
in)	<u>Function</u>	Name / Descr	Hit	<u> </u>	Probability (if >	<u>E-value</u>	
Supporting	(or NKF)	(incl DUF) Transforming	5VQF	_	90%)	<u>(if < 1)</u>	
Information for Function, from HHpred (HHpred) (All three lines of SIF	g growth factor beta	growth factor beta-1; pro- complex, latency, homodimer,			67.92	27	
evidence must be filled in)		PROTEIN BINDING; HET: NAG, BMA; 2.9A					
Supporting Information for Function,	Function (or NKF)	Pham numb number of m		<u>u</u>	pstream pham conserved ? Yes or No	<u>Downstream pham</u> <u>conserved ?</u> <u>Yes</u> or No	
from Syntney (SIF-Syn) (All three lines of SIF evidence must be filled in)	NKF	8860 (9	9)	If yes	s, what pham # or function ? 147353 (28)	If yes, what pham # or function? 142271 (208)	
Transmembrane	domains by Tm	nHmm using Pha	merator	N/A			

Function (F)	Gene No: 51				Student name: Kristen Mclean		
hypothetical prot				Date: 2	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	<u>% Query</u> %99	<u>% Ide</u> r		<u>E-value</u> (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		<u>% Ide</u> 1	-	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	<u>Hi</u>	_	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 numb number of m 142271 (2	nembers		pstream pham conserved ? Yes or No s, what pham # or function ? 8860 (9)	Downstream pham conserved? Yes or No If yes, what pham # or function? 8491 (9)	
Transmembrane	domains by Tm	nHmm using Pha	merator	N/A			

Function (F)	Gene No: 52				Student name: Kristen Mclean		
hypothetical prote				Date:	e: 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	<u>% Query</u> %99	<u>% Ider</u> %96.		<u>E-value</u> (<u>if < 10⁻⁷)</u> 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		<u>% Ider</u> %90	-	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)	<u>Hit</u>		Probability (if > 90%)	<u>E-value</u> (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 numb number of m 8491 (9	<u>embers</u>		pstream pham conserved ? Yes or No s, what pham # or function ? 142271 (208)	Downstream pham conserved? Yes or No If yes, what pham # or function? 147042 (86)	
Transmembrane	domains by Tn	nHmm using Pha	merator	N/A			

	Gene No: 53			ר+ייק~	nt name: Kristen Mal	oan
Function (F)	_			Student name: Kristen Mclean		
helix-turn helix binding protein	dna			Date: 2/25/2024		
Supporting	Function	% Query	<u>% Iden</u>	ititv	<u>E-value</u>	AA # conserved in
Information	(or NKF)	<u> 70 Que. y</u>	<u>/01001</u>	cicy	(if < 10 ⁻⁷)	first 10 hits?
for Function,		0/ 05	%10	Ω	<u>(20 /</u>	
from BLAST	helix-turn-	% 95	7010	U	1e-55	Yes or <mark>No</mark>
(Protein)	helix DNA					
(BLAST-	binding domain					Conserved on the first
NCBI)	protein					2 hits
(All three lines of SIF evidence must be filled in)						
Supporting	<u>Function</u>		<u>% Iden</u>	<u>itity</u>	E-value	AA # conserved in
Information	(or NKF)		0/40	0	(if < 10 ⁻⁷)	first 10 hits?
for Function, from BLAST	helix-turn-		%10	U	2e-46	Yes or N <mark>o</mark>
(Protein)	helix DNA				20 40	<u>163 01 14<mark>0</mark></u>
(BLAST-	binding					Conserved on the
phagesDB)	domain					first 7 hits
(All three lines of SIF evidence	protein					
must be filled						
in)						
	<u>Function</u>	Name / Descr	<u>Hit</u>	:	Probability (if >	<u>E-value</u>
Supporting	(or NKF)	(incl DUF)			<u>90%)</u>	<u>(if < 1)</u>
Information	la altri di con	Regulatory	4LHF	_A		
for Function,	helix-turn- helix, DNA	protein cox;			98.28	0.000012
from HHpred	binding	helix-turn-				
(HHpred)	J	helix, DNA				
(All three lines of SIF		binding,				
evidence must be filled in)		VIRAL				
filled iff)		PROTEIN;				
		2.401A		T		
Supporting	Function	Pham numb		<u>u</u>	pstream pham	<u>Downstream pham</u>
Information	(or NKF)	number of m	<u>empers</u>		conserved ? Yes or No	<u>conserved ?</u> Yes or <mark>No</mark>
for Function,	haller to co				ies of INO	162 OL <mark>INO</mark>
from Syntney	helix-turn-	147042 (86)	<u>If ye</u> :	s, what pham # or	If yes, what pham # or
(SIF-Syn)	helix DNA binding	·			function?	function ?
(All three lines of SIF evidence must be	domain				0.404 (0)	
filled in)	protein				8491 (9)	
	protein					
Transmembrane	domains by Tm	nHmm using Pha	merator	N/A		
	,					
I				1		

Function (F) Hypothetical	Gene No: 54			Student name: Kristen Mclean		
prot				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	<u>% Query</u> %97	<u>% Ider</u> %97.		<u>E-value</u> (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		<u>% Ider</u> %91		<u>E-value</u> (<u>if < 10⁻⁷)</u> 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	<u>Hit</u> PF1117		Probability (if > 90%) 81.89	<u>E-value</u> (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 numb number of m 10316 (nembers		pstream pham conserved ? Yes or No s, what pham # or function ?	Downstream pham conserved ? Yes or No If yes, what pham # or function ?
Transmembrane	domains by Tm	nHmm using Pha	merator	N/A		

Function (F) Hypothetical	Gene No: 55			Student name: Betty Sierra / Grace Intrator		
protein				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	<u>% Query</u> <u>99</u>	<u>% Ider</u> <u>99.1</u>		<u>E-value</u> (<u>if < 10⁻⁷)</u> <u>6e-77</u>	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)	<u>Function</u> (or <mark>NKF</mark>)		<u>% Ider</u>		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)	<u>Hit</u>		Probability (if > 90%) All probabilities below 90	<u>E-value</u> (if < 1)
Supporting Information for Function, from Syntney (SIF-Syn) (All three lines of SIF evidence must be filled in)	<u>Function</u> (or <mark>NKF</mark>)	Pham numb number of m 144027	<u>nembers</u>		pstream pham conserved ? Yes or No s, what pham # or function ?	Downstream pham conserved? Yes or No If yes, what pham # or function?
Transmembrane	domains by Tml	Hmm using Pha	merator	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	<u>% Query</u> <mark>99%</mark>	<u>% Iden</u>	 <u>E-value</u> (if < 10 ⁻⁷) <mark>7e-69</mark>	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		<u>% Iden</u> <u>98%</u>	E-value (if < 10 ⁻⁷) <mark>7e-55</mark>	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 transcriptio n factor	Name / Descr (incl DUF) Putative AphA- like transcription factor; ZP_00208345. 1, Putative AphA-like Transcription Factor, Structural Genomics, Jo	<u>Hit</u>	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 numb number of m 88022(s	embers	pstream pham conserved ? Yes or No s, what pham # or function ? 8461(10)	Downstream pham conserved? Yes or No If yes, what pham # or function?	
Transmembrane	domains by Tm	nHmm using Pha	merator			

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				nt name: Nicole Gor 02/13/24	nzalez Giliberti
Supporting Information for Function, from BLAST (Protein) (BLAST- NCBI) (All three lines of SIF evidence must be filled in)	<u>Function</u> (or NKF) <mark>NKF</mark>	<u>% Query</u> <mark>99%</mark>	<u>% Iden</u>		<u>E-value</u> (if < 10 ⁻⁷) <mark>8e-90</mark>	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		<u>% Iden</u>	-	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	<u>Hit</u>		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)	<u>Function</u> (or NKF) <mark>NKF</mark>	Pham numb number of m 8216(9	<u>embers</u>		pstream pham conserved ? Yes or No s, what pham # or function ?	Downstream pham conserved? Yes or No If yes, what pham # or function? 143748(371)
Transmembrane	domains by Tm	nHmm using Pha	merator			

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

Function (F) HTH DNA bind	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	<u>% Query</u> <mark>98%</mark>	<u>% Iden</u>	itity	E-value (if < 10 ⁻⁷) <u>5e-97</u>	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		<u>% Iden</u>	-	<u>E-value</u> (if < 10 ⁻⁷) <mark>8e-79</mark>	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;Heli x-turn-helix domain	Name / Descr (incl DUF) HTH_58;Helix -turn-helix domain	<u>Hit</u>		Probability (if > 90%) 98.35%	<u>E-value</u> (if < 1) <u>0.000012</u>	
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 numb number of m 142178 (4	<u>embers</u>		pstream pham conserved ? Yes or No s, what pham # or function ? 8216 (9)	Downstream_pham conserved ? Yes or No If yes, what pham # or function ? 85852(50)	
Transmembrane	domains by Tn	nHmm using Pha	merator				

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

Function (F) HTH DNA bind	Gene No: 59 domain				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	<u>% Query</u> 99%	<u>% Ider</u>	ntity_	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		<u>% Ider</u> <u>889</u>		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 transcription al regulator SArX	<u>Hit</u>	_	Probability (if > 90%) 99.06%	<u>E-value</u> (if < 1) <mark>2.8e-9</mark>	
Supporting Information for Function, from Syntney (SIF-Syn) (All three lines of SIF evidence must be filled in)	Function (or NKF) <mark>NKF</mark>	Pham numb number of m 85852(5	<u>embers</u>		pstream pham conserved ? Yes or No s, what pham # or function ? 143748(371)	Downstream pham conserved ? Yes or No If yes, what pham # or function ?	
Transmembrane	domains by Tm	nHmm using Pha	merator				

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 (hypothetic al protein)	<u>% Query</u> 99%	<u>% Iden</u>	tity	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		<u>% Iden</u> <mark>96%</mark>		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) <mark>N/A</mark>	Name / Descr (incl DUF) N/A	Hit N/#	_	Probability (if > 90%) N/A	<u>E-value</u> (if < 1) <u>N/A</u>
Supporting Information for Function, from Syntney (SIF-Syn) (All three lines of SIF evidence must be filled in)	<u>Function</u> (or NKF) <mark>NKF</mark>	Pham numb number of m 8740(9	<u>embers</u>		pstream pham conserved ? Yes or No s, what pham # or function ?	Downstream pham conserved ? Yes or No If yes, what pham # or function ?
Transmembrane	domains by Tm	nHmm using Pha	merator			

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)	Gene No: 61			Stude	Student name: Nicole Gonzalez Giliberti		
HP				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) Hypothetica I protein	<u>% Query</u> <mark>99%</mark>	<u>% Iden</u>		<u>E-value</u> (if < 10 ⁻⁷) <u>3e-126</u>	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		<u>% Iden</u> 71%	-	E-value (if < 10 ⁻⁷) <u>e-106</u>	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	<u>Hit</u>	_	Probability (if > 90%) N/A	<u>E-value</u> (if < 1) <u>N/A</u>	
Supporting Information for Function, from Syntney (SIF-Syn) (All three lines of SIF evidence must be filled in)	<u>Function</u> (or NKF) <mark>NKF</mark>	Pham numb number of m 88148(i	embers		pstream pham conserved ? Yes or No s, what pham # or function ?	Downstream pham conserved ? Yes or No If yes, what pham # or function ?	
Transmembrane	domains by Tm	nHmm using Pha	merator				

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) Hypothetica I protein	% Query <mark>99%</mark>	% Iden 80.42	tity	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		<u>% Iden</u>		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	<u>Hit</u>	_	Probability (if > 90%) N/A	<u>E-value</u> (if < 1) <u>N/A</u>	
Supporting Information for Function, from Syntney (SIF-Syn) (All three lines of SIF evidence must be filled in)	<u>Function</u> (or NKF) <mark>NKF</mark>	Pham numb number of m 8845(S	<u>embers</u>		pstream pham conserved ? Yes or No s, what pham # or function ?	Downstream pham conserved ? Yes or No If yes, what pham # or function ?	
Transmembrane	domains by Tm	nHmm using Pha	merator				

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				nt name: Nicole Gon	zalez Giliberti
Supporting Information for Function, from BLAST (Protein) (BLAST- NCBI) (All three lines of SIF	Function (or NKF) Hypothetica I protein	<u>% Query</u> 99%	<u>% Iden</u>	tity	02/23/24 <u>E-value</u> (if < 10 ⁻⁷) <u>7e-70</u>	AA # conserved in first 10 hits? Yes or No
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) NKF		<u>% Iden</u>		<u>E-value</u> (if < 10 ⁻⁷) <u>1e-57</u>	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) <mark>N/A</mark>	Name / Descr (incl DUF) N/A	<u>Hit</u>	_	Probability (if > 90%) N/A	<u>E-value</u> (if < 1) <u>N/A</u>
Supporting Information for Function, from Syntney (SIF-Syn) (All three lines of SIF evidence must be filled in)	<u>Function</u> (or NKF) <mark>NKF</mark>	number of members		pstream pham conserved ? Yes or No s, what pham # or function ?	Downstream pham conserved? Yes or No If yes, what pham # or function?	
Transmembrane	domains by Tm	nHmm using Pha	merator			

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				nt name: Nicole Gon	zalez Giliberti
Supporting Information for Function, from BLAST (Protein) (BLAST- NCBI) (All three lines of SIF evidence must be filled in)	Function (or NKF) Hypothetica I protein	<u>% Query</u> <mark>99%</mark>	<u>% Iden</u> <u>86.5</u> 4	tity	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		<u>% Iden</u> <u>86%</u>		E-value (if < 10 ⁻⁷) <mark>1e-76</mark>	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/ <i>F</i>	_	Probability (if > 90%) N/A	<u>E-value</u> (if < 1) <u>N/A</u>
Supporting Information for Function, from Syntney (SIF-Syn) (All three lines of SIF evidence must be filled in)	<u>Function</u> (or NKF) <mark>NKF</mark>	Pham number and number of members 8415(9)			pstream pham conserved ? Yes or No s, what pham # or function ?	Downstream pham conserved? Yes or No If yes, what pham # or function?
Transmembrane	domains by Tm	nHmm using Pha	merator			

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)	Gene No: 65				nt name: Nicole Gon	zalez Giliberti
Supporting Information for Function, from BLAST (Protein) (BLAST- NCBI) (All three lines of SIF evidence must be filled in)	Function (or NKF) NKF	<u>% Query</u> 78%	% Iden	tity	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		<u>% Iden</u>		<u>E-value</u> (if < 10 ⁻⁷) <u>2e-08</u>	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) <mark>N/A</mark>	Name / Descr (incl DUF) N/A	Hit N/ <i>E</i>		Probability (if > 90%) N/A	E-value (if < 1) <mark>N/A</mark>
Supporting Information for Function, from Syntney (SIF-Syn) (All three lines of SIF evidence must be filled in)	<u>Function</u> (or NKF) <mark>NKF</mark>	Pham number and number of members 136905(1)			pstream pham conserved ? Yes or No s, what pham # or function ?	Downstream pham conserved ? Yes or No If yes, what pham # or function ?
Transmembrane	domains by Tn	nHmm using Pha	merator			

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				nt name: Nicole Gon	zalez Giliberti
Supporting Information for Function, from BLAST (Protein) (BLAST- NCBI) (All three lines of SIF evidence must be filled in)	Function (or NKF) NKF	<u>% Query</u> <mark>98%</mark>	<u>% Iden</u>	tity	E-value (if < 10 ⁻⁷) <mark>7e-57</mark>	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		<u>%</u> Iden <u>96%</u>		<u>E-value</u> (if < 10 ⁻⁷) <u>9e-47</u>	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) <mark>N/A</mark>	Name / Descr (incl DUF) N/A	<u>Hit</u>	_	Probability (if > 90%) N/A	E-value (if < 1) <mark>N/A</mark>
Supporting Information for Function, from Syntney (SIF-Syn) (All three lines of SIF evidence must be filled in)	<u>Function</u> (or NKF) <mark>NKF</mark>	number of members		pstream pham conserved ? Yes or No s, what pham # or function ?	Downstream pham conserved ? Yes or No If yes, what pham # or function ?	
Transmembrane	domains by Tn	nHmm using Pha	merator			

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				nt name: Nicole Gon	zalez Giliberti
Supporting Information for Function, from BLAST (Protein) (BLAST- NCBI) (All three lines of SIF evidence must be filled in)	Function (or NKF) NKF	<u>% Query</u> <mark>98%</mark>	<u>% Iden</u>	tity	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		<u>% Iden</u>		E-value (if < 10 ⁻⁷) <u>5e-41</u>	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/ <i>F</i>	_	Probability (if > 90%) N/A	<u>E-value</u> (if < 1) <u>N/A</u>
Supporting Information for Function, from Syntney (SIF-Syn) (All three lines of SIF evidence must be filled in)	<u>Function</u> (or NKF) <mark>NKF</mark>	Pham number and number of members 90209(3)			pstream pham conserved ? Yes or No s, what pham # or function ?	Downstream pham conserved ? Yes or No If yes, what pham # or function ?
Transmembrane	domains by Tn	nHmm using Pha	merator			

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.