Basic Phage Information	
Phage Name	Akino08
Gene #	1
Stop Coordinate	359
Direction (For/Rev)	Rev
Gap (Overlap) with Previous Gene	Major Gap
Selected Start Coordinate	517
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes Glimmer, bp 517 (strength of 5.74)
(Glimmer, GeneMark)?	
Is there evidence for	Very little coding potential
coding potential?	
Is this gene present in	No other cenomes had the cene present
other annotated genomes?	no other genomes had the gene present.
Does the gene violate any	The gene violates major guiding principles such as overlapping other, stronger genes
major guiding principles?	
DECISION:	No

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and GeneMark suggest?	GeneMark Start Coordinate (type NA if not supported)::
Does the start site have an	List the final RBS score and Z-score of the currently predicted
associated Ribosome	start site using the Kiblerb/Karlin Medium scoring table. Indicate
Binding Sile with a high	In your response if this is the best score of not.
300101	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is the predicted start codon	Indicate the length of the ORF is with the predicted start and the
the longest ORF? If not,	gap/overlap to the nearest stop codon of the upstream ORF.
does the longest ORF	Does the proposed start site have a gap/overlap with the
overlap (>30bp)?	Principles?
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is this start site conserved	
In other phage genomes	Very will also pood to provide the following information from
Starterator?	Starterator: does the start match the consensus start site
	predicted from Starterator? If no, is the consensus start site not
	found in this ORF? If no, is there a better option for the
	consensus start site instead of the one predicted by Starterator?
	If Starterator doesn't reveal a consensus start site, you can
	record that Starterator was not informative.
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is this start site conserved	Provide the best BlastP match from NCBI, PhagesDB, and DNA
in other phage genomes	Master with alignment in the format of (Q#:S#), where Q (query)
as indicated by BlastP?	is the sequence you are analyzing and S (subject) is the
	database match. List the e-value and alignment of the best match for all three PlastP sources
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
DECISION	Record where you think the gene should start here and briefly
	explain your rationale.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
snows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
DUACES approved	Once you have arrived at a functional decision, check the <u>SEA-</u>
function list2	the guidelines for function List to ensure that you are following
	the guidelines for function naming. Functions that are not
	approval
	approval.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	3
Stop Coordinate	514
Direction (For/Rev)	Rev
Gap (Overlap) with Previous Gene	Major Overlap
Selected Start Coordinate	2049
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes Glimmer, bp 2049 (strength of 10.76)
(Glimmer, GeneMark)?	
Is there evidence for	Very little to none coding potential
coding potential?	
Is this gene present in	Yes – only in Loviatar_Draft
other annotated genomes?	
Does the gene violate any major guiding principles?	The gene violates major guiding principles such as overlapping other, stronger genes
DECISION:	No

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and GeneMark suggest?	GeneMark Start Coordinate (type NA if not supported)::
Does the start site have an	List the final RBS score and Z-score of the currently predicted
associated Ribosome	start site using the Kiblerb/Karlin Medium scoring table. Indicate
Binding Sile with a high	In your response if this is the best score of not.
300101	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is the predicted start codon	Indicate the length of the ORF is with the predicted start and the
the longest ORF? If not,	gap/overlap to the nearest stop codon of the upstream ORF.
does the longest ORF	Does the proposed start site have a gap/overlap with the
overlap (>30bp)?	Principles?
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is this start site conserved	
In other phage genomes	Very will also pood to provide the following information from
Starterator?	Starterator: does the start match the consensus start site
	predicted from Starterator? If no, is the consensus start site not
	found in this ORF? If no, is there a better option for the
	consensus start site instead of the one predicted by Starterator?
	If Starterator doesn't reveal a consensus start site, you can
	record that Starterator was not informative.
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is this start site conserved	Provide the best BlastP match from NCBI, PhagesDB, and DNA
in other phage genomes	Master with alignment in the format of (Q#:S#), where Q (query)
as indicated by BlastP?	is the sequence you are analyzing and S (subject) is the
	database match. List the e-value and alignment of the best match for all three PlastP sources
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
DECISION	Record where you think the gene should start here and briefly
	explain your rationale.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
shows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
DUACES approved	Once you have arrived at a functional decision, check the <u>SEA-</u>
Function list?	PHAGES Official Function List to ensure that you are following
Tunction list?	the guidelines for function naming. Functions that are not
	present on the approved list must be carefully velled for
	approval.
	If you believe this gene should be assigned, please write the
DEGICIONI	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	2
Stop Coordinate	625
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	Overlap
Selected Start Coordinate	395
Selected Function	Hypothetical Protein

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes Genemark, bp 395
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Yes – Didgeridoo and AvocadoMan both had this gene
other annotated genomes?	
Does the gene violate any	The gene follows all guiding principles.
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer and GeneMark suggest?	<i>Glimmer Start Coordinate (type NA if not supported)::</i> <i>GeneMark Start Coordinate (type NA if not supported)::</i> 395
Does the start site have an associated Ribosome Binding Site with a high score?	395 – large z-score (3.029) and appropriate spacer (9 bp)
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	No it is not the longest ORF, and haven't evaluated the previous gene so we do not know if there is major overlap with the longest ORF.
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, It was found in 60 of 62 genes in the pham, and called 96.8% of the time Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes it is highly conserved.
DECISION:	Start site is bp 395, based off of Glimmer/Genemark, Genemark coding potential, phamerator, starterator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	Yes, hypothetical protein
a protein having a	
functional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁻⁴ or	
smaller with appropriate	
coverage?	
Does this protein align with	No Quality Match - Clp protease ClpC,Heat shock survival AAA family
a protein having a	ATPase ClpK; ATPase associated with diverse cellular activities (AAA), 86.9 %
functional assignment in	
the PDB or other database	Hypothetical protein
in HHPred with a	
probability of 90% or	
greater with appropriate	
coverage?	
Is this gene located	Yes, terminase small subunit
adjacent to genes of	
known function and in a	
region of the genome that	
shows high conservation of	
gene order?	
Is this gene a possible	No
transmembrane protein?	
Is the proposed function	Yes, Hypothetical Protein
found on the SEA-	
PHAGES approved	
function list?	
DECISION:	Hypothetical Protein

Basic Phage Information	
Phage Name	Akino08
Gene #	4
Stop Coordinate	948
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	625
Selected Function	terminase, small subunit

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes GeneMark, bp 625
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Yes – BabyDaisy, BabyYoda, and Avocadoman all had this
other annotated genomes?	gene
Does the gene violate any	The gene follows all guiding principles.
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer and GeneMark suggest?	<i>Glimmer Start Coordinate (type NA if not supported)::</i> <i>GeneMark Start Coordinate (type NA if not supported)::</i> 625
Does the start site have an associated Ribosome Binding Site with a high score?	625 – small z-score (0.913) and appropriate spacer (<mark>16</mark> bp)
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	No it is not the longest ORF, and haven't evaluated the previous gene so we do not know if there is major overlap with the longest ORF.
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, It was found in 81 of 82 genes in the pham, and called 98.8% of the time Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes it is highly conserved.
DECISION:	Start site is bp 625, based off of Glimmer/Genemark, Genemark coding potential, phamerator, starterator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	Yes, terminase small subunit
a protein having a	
functional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁻⁴ or	
smaller with appropriate	
coverage?	
Does this protein align with	; Terminase_4 ; Phage terminase, small subunit – 95.04%
a protein having a	
functional assignment in	
the PDB or other database	
in HHPred with a	
probability of 90% or	
greater with appropriate	
coverage?	
Is this gene located	Yes, terminase small subunit
adjacent to genes of	
Known function and in a	
region of the genome that	
snows nigh conservation of	
gene order?	No
transmembrane pretein?	NO
liansmembrane protein?	Van tarminaan amall aubunit
found on the SEA	res, lerrinnase, sinali suburni
DUACES approved	
function list?	
	terminase, small subunit

Basic Phage Information	
Phage Name	Akino08
Gene #	5
Stop Coordinate	2577
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	952
Selected Function	terminase, large subunit

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes both, Glimmer bp 2044, Genemark bp 952
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves only in Loviator Draft
other annotated genomes?	res, only in Lovialar_Dran
Does the gene violate any	The gene follows all guiding principles
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer and GeneMark suggest?	Glimmer Start Coordinate (type NA if not supported):: 2044 GeneMark Start Coordinate (type NA if not supported):: 952
Does the start site have an associated Ribosome Binding Site with a high score?	952 – large z-score (3.029) and appropriate spacer (10 bp)
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	Yes it is the longest ORF.
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, It (2044) was found in 2 of 2 genes in the pham, and called 100% of the time Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes it is highly conserved.
DECISION:	Start site is bp 952, based off of Genemark, Genemark coding potential, phamerator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	Yes, terminase large subunit
a protein having a	
functional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁻⁴ or	
smaller with appropriate	
coverage?	
Does this protein align with	; TerL_nuclease ; Terminase large subunit, endonuclease domain – 99.02%
a protein having a	
functional assignment in	
the PDB or other database	
in HHPred with a	
probability of 90% or	
greater with appropriate	
coverage?	
Is this gene located	No Synteny Observed.
adjacent to genes of	
Known function and in a	
region of the genome that	
snows nigh conservation of	
gene order?	
Is this gene a possible	NO
la the proposed function	Van tarminaan large subunit
found on the SEA	res, terminase, large suburnit
DUACES approved	
function list?	
	terminase large subunit
DECISION.	

Basic Phage Information	
Phage Name	Akino08
Gene #	7
Stop Coordinate	2651
Direction (For/Rev)	Rev
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	3493
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	Original Glimmer call @bn 4270 has strength 6.81 ** not called
auto-annotation program	by GeneMark
(Glimmer, GeneMark)?	by Cenewark
Is there evidence for	No coding potential
coding potential?	
Is this gene present in	Ves only in Loviator Draft
other annotated genomes?	res – only in Lovialai_Drail
Does the gene violate any	The gene violates major guiding principles such as overlapping
major guiding principles?	other, stronger genes
DECISION:	Not a gene, deleted

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and GeneMark suggest?	GeneMark Start Coordinate (type NA if not supported)::
Does the start site have an	List the final RBS score and Z-score of the currently predicted
associated Ribosome	start site using the Kiblerb/Karlin Medium scoring table. Indicate
Binding Sile with a high	In your response if this is the best score of not.
300101	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is the predicted start codon	Indicate the length of the ORF is with the predicted start and the
the longest ORF? If not,	gap/overlap to the nearest stop codon of the upstream ORF.
does the longest ORF	Does the proposed start site have a gap/overlap with the
overlap (>30bp)?	Principles?
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is this start site conserved	
In other phage genomes	Very will also pood to provide the following information from
Starterator?	Starterator: does the start match the consensus start site
	predicted from Starterator? If no, is the consensus start site not
	found in this ORF? If no, is there a better option for the
	consensus start site instead of the one predicted by Starterator?
	If Starterator doesn't reveal a consensus start site, you can
	record that Starterator was not informative.
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is this start site conserved	Provide the best BlastP match from NCBI, PhagesDB, and DNA
in other phage genomes	Master with alignment in the format of (Q#:S#), where Q (query)
as indicated by BlastP?	is the sequence you are analyzing and S (subject) is the
	database match. List the e-value and alignment of the best match for all three PlastP sources
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
DECISION	Record where you think the gene should start here and briefly
	explain your rationale.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
snows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
DUACES approved	Once you have arrived at a functional decision, check the <u>SEA-</u>
function list2	the guidelines for function List to ensure that you are following
	the guidelines for function naming. Functions that are not
	approval
	approval.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	8
Stop Coordinate	3890
Direction (For/Rev)	Rev
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	4084
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	Original glimmer cell @bp 4094 has strength 5.52 not called by
auto-annotation program	gene mark
(Glimmer, GeneMark)?	
Is there evidence for	GeneMark did not show coding potential
coding potential?	Genemark and not show county potential
Is this gene present in	Voc. Loviator Droft had this gono
other annotated genomes?	Tes- Loviatar Drait riad tills gene
Does the gene violate any	The gape does not follow all guiding principles
major guiding principles?	
DECISION:	No

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and Genemark suggest?	Genemark Start Coordinate (type NA If not supported)::
Does the start site have an	List the final RBS score and Z-score of the currently predicted
Binding Site with a high	in your response if this is the best score or not.
score?	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is the predicted start codon	Indicate the length of the ORF is with the predicted start and the
does the longest ORF	Does the proposed start site have a gap/overlap with the
result in excessive gene overlap (>30bp)?	nearest upstream gene that does not violate the Guiding Principles?
	Notes if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is this start site conserved	
as indicated by Starterator?	You will also need to provide the following information from Starterator: does the start match the consensus start site predicted from Starterator? If no, is the consensus start site not found in this ORF? If no, is there a better option for the consensus start site instead of the one predicted by Starterator? If Starterator doesn't reveal a consensus start site, you can record that Starterator was not informative.
	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
Is this start site conserved in other phage genomes as indicated by BlastP?	Provide the best BlastP match from NCBI, PhagesDB, and DNA Master with alignment in the format of (Q#:S#), where Q (query) is the sequence you are analyzing and S (subject) is the database match. List the e-value and alignment of the best match for all three BlastP seurces
	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
DECISION:	Record where you think the gene should start here and briefly explain your rationale.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
snows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
DUACES approved	Once you have arrived at a functional decision, check the <u>SEA-</u>
function list2	the guidelines for function List to ensure that you are following
	the guidelines for function naming. Functions that are not
	approval
	approval.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	9
Stop Coordinate	4133
Direction (For/Rev)	Rev
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	4270
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	Original Glimmer call @bp 4270 has strength 6.81 not called by
auto-annotation program	Conomark
(Glimmer, GeneMark)?	Conemark
Is there evidence for	Genemark did not show coding potential
coding potential?	
Is this gene present in	Yes-Loviatar_Draft had this gene
other annotated genomes?	
Does the gene violate any	The gene does not follow all guiding principles
major guiding principles?	
DECISION:	No

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and Genemark suggest?	Genemark Start Coordinate (type NA If not supported)::
Does the start site have an	List the final RBS score and Z-score of the currently predicted
Binding Site with a high	in your response if this is the best score or not.
score?	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is the predicted start codon	Indicate the length of the ORF is with the predicted start and the
does the longest ORF	Does the proposed start site have a gap/overlap with the
result in excessive gene overlap (>30bp)?	nearest upstream gene that does not violate the Guiding Principles?
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is this start site conserved in other phage genomes	
as indicated by Starterator?	You will also need to provide the following information from Starterator: does the start match the consensus start site predicted from Starterator? If no, is the consensus start site not found in this ORF? If no, is there a better option for the consensus start site instead of the one predicted by Starterator? If Starterator doesn't reveal a consensus start site, you can record that Starterator was not informative.
	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
Is this start site conserved in other phage genomes as indicated by BlastP?	Provide the best BlastP match from NCBI, PhagesDB, and DNA Master with alignment in the format of (Q#:S#), where Q (query) is the sequence you are analyzing and S (subject) is the database match. List the e-value and alignment of the best
	match for all three BlastP sources.
	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
DECISION:	Record where you think the gene should start here and briefly explain your rationale.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
shows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
found on the SEA-	Once you have arrived at a functional decision, check the <u>SEA-</u>
PHAGES approved	<u>PHAGES Official Function List</u> to ensure that you are following
function list?	the guidelines for function naming. Functions that are not
	present on the approved list must be carefully vetted for
	approval.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	10
Stop Coordinate	4270
Direction (For/Rev)	Rev
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	6201
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	Original Climmer call @bn 6201 has strength 8 81 ** not called
auto-annotation program	by GeneMark
(Glimmer, GeneMark)?	by Cenewark
Is there evidence for	No coding potential
coding potential?	
Is this gene present in	No, the cone is an ornhan
other annotated genomes?	No, the gene is an orphan
Does the gene violate any	The gene violates major guiding principles such as overlapping
major guiding principles?	other, stronger genes
DECISION:	Not a gene, deleted

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and GeneMark suggest?	GeneMark Start Coordinate (type NA if not supported)::
Does the start site have an	List the final RBS score and Z-score of the currently predicted
associated Ribosome	start site using the Kibler6/Karlin Medium scoring table. Indicate
Binding Site with a high	in your response if this is the best score or not.
score?	
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is the predicted start codon	Indicate the length of the ORF is with the predicted start and the
the longest ORF? If not,	gap/overlap to the nearest stop codon of the upstream ORF.
does the longest ORF	Does the proposed start site have a gap/overlap with the
result in excessive gene	nearest upstream gene that does not violate the Guiding
overlap (>30bp)?	Principles?
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is this start site conserved	
in other phage genomes	
as indicated by	You will also need to provide the following information from
Starterator?	Starterator: does the start match the consensus start site
	predicted from Starterator? If no, is the consensus start site not
	found in this ORF? If no, is there a better option for the
	consensus start site instead of the one predicted by Starterator?
	If Starterator doesn't reveal a consensus start site, you can
	record that Starterator was not informative.
	Note: If you are considering more than 1 start site, provide the
	same information for each proposed start site.
le this start site conserved	Provide the best PlastP metab from NCPL PhagesDP and DNA
in other phage genemoe	Moster with alignment in the formet of (Ott:St), where O (quary)
in other phage genomes	is the sequence you are enclyzing and S (cubicet) is the
as indicated by blastr?	Is the sequence you are analyzing and S (Subject) is the
	meteb for all three PlactD sources
	Noto: if you are considering more than 1 start site, provide the
	some information for each proposed start site
	Record where you think the gene should start here and briefly
DECISION:	explain your rationale

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
snows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
DUACES approved	Once you have arrived at a functional decision, check the <u>SEA-</u>
function list2	the guidelines for function List to ensure that you are following
	the guidelines for function naming. Functions that are not
	approval
	approval.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	6
Stop Coordinate	4603
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	2588
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes GeneMark, bp 2588
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves – both BabyVoda and Avocadoman had this gane
other annotated genomes?	Tes – both baby foda and Avocadoman had this gene
Does the gene violate any	The gene follows all guiding principles.
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer and GeneMark suggest?	<i>Glimmer Start Coordinate (type NA if not supported)::</i> <i>GeneMark Start Coordinate (type NA if not supported)::</i> 2588
Does the start site have an associated Ribosome Binding Site with a high score?	2588 – large z-score (3.029) and appropriate spacer (9 bp)
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	No it is not the longest ORF, and does not overlap with the longest ORF Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, It was found in 81 of 234 genes in the pham, and called 34.6% of the time Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes it is highly conserved.
DECISION:	Start site is bp 2588, based off of Genemark, Genemark coding potential, phamerator, starterator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	Yes, Portal protein
a protein having a	
functional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁻⁴ or	
smaller with appropriate	
coverage?	
Does this protein align with	Portal protein; Archaeal virus, portal, portal capsid interface, Mg
a protein having a	ions, VIRUS; HET: MG, HIP; 2.342A {Haloferax tailed
functional assignment in	-100%
the PDB or other database	
in HHPred with a	
probability of 90% or	
greater with appropriate	
coverage?	
Is this gene located	Yes, portal protein
adjacent to genes of	
known function and in a	
region of the genome that	
shows high conservation of	
gene order?	
Is this gene a possible	No
transmembrane protein?	
Is the proposed function	Yes, portal protein
found on the SEA-	
PHAGES approved	
function list?	
DECISION:	portal protein

Basic Phage Information	
Phage Name	Akino08
Gene #	11
Stop Coordinate	5358
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	4603
Selected Function	capsid maturation protease

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes, Genemark bp 4603
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves – BabyVoda and Avocadoman both had this gane
other annotated genomes?	res – baby roda and Avocadoman both had this gene
Does the gene violate any	The gene follows all guiding principles
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer and GeneMark suggest?	<i>Glimmer Start Coordinate (type NA if not supported)::</i> <i>GeneMark Start Coordinate (type NA if not supported)::</i> 4603
Does the start site have an associated Ribosome Binding Site with a high score?	4603 – large z-score (3.029) and appropriate spacer (11 bp)
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	No it is not the longest ORF, and haven't evaluated the previous gene so we do not know if there is major overlap with the longest ORF.
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, It was found in 153 of 322 genes in the pham, and called 47.5% of the time Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes it is highly conserved.
DECISION:	Start site is bp 4603, based off of Genemark, Genemark coding potential, phamerator, starterator, and blast data

Gathering Evidence	Explain Your Rationale
Does this protein align with	Yes, capsid maturation protease
a protein having a	
functional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁻⁴ or	
smaller with appropriate	
coverage?	
Does this protein align with	;Peptidase_S78 ; Caudovirus prohead serine protease
a protein having a	99.85%
functional assignment in	
the PDB or other database	
in HHPred with a	
probability of 90% or	
greater with appropriate	
coverage?	
Is this gene located	YES
adjacent to genes of	
known function and in a	
region of the genome that	
shows high conservation of	
gene order?	
Is this gene a possible	No
transmembrane protein?	
Is the proposed function	YES
tound on the SEA-	
PHAGES approved	
function list?	
DECISION:	capsid maturation protease

Basic Phage Information	
Phage Name	Akino08
Gene #	13
Stop Coordinate	6312
Direction (For/Rev)	Rev
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	7052
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	Original Glimmer call @bn 7052 has strength 5.45 ** not called
auto-annotation program	by GeneMark
(Glimmer, GeneMark)?	by Cenewark
Is there evidence for	No coding potential
coding potential?	
Is this gene present in	No the cono is an ornhan
other annotated genomes?	No, the gene is an orphan
Does the gene violate any	The gene violates major guiding principles such as overlapping
major guiding principles?	other, stronger genes
DECISION:	Not a gene, deleted

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and GeneMark suggest?	GeneMark Start Coordinate (type NA if not supported)::
Does the start site have an	List the final RBS score and Z-score of the currently predicted
associated Ribosome	start site using the Kiblerb/Karlin Medium scoring table. Indicate
Binding Sile with a high	In your response if this is the best score of not.
300101	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is the predicted start codon	Indicate the length of the ORF is with the predicted start and the
the longest ORF? If not,	gap/overlap to the nearest stop codon of the upstream ORF.
does the longest ORF	Does the proposed start site have a gap/overlap with the
overlap (>30bp)?	Principles?
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is this start site conserved	
In other phage genomes	Very will also pood to provide the following information from
Starterator?	Starterator: does the start match the consensus start site
	predicted from Starterator? If no, is the consensus start site not
	found in this ORF? If no, is there a better option for the
	consensus start site instead of the one predicted by Starterator?
	If Starterator doesn't reveal a consensus start site, you can
	record that Starterator was not informative.
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is this start site conserved	Provide the best BlastP match from NCBI, PhagesDB, and DNA
in other phage genomes	Master with alignment in the format of (Q#:S#), where Q (query)
as indicated by BlastP?	is the sequence you are analyzing and S (subject) is the
	database match. List the e-value and alignment of the best match for all three PlastP sources
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
DECISION	Record where you think the gene should start here and briefly
	explain your rationale.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
shows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
found on the SEA-	Once you have arrived at a functional decision, check the <u>SEA-</u>
PHAGES approved	PHAGES Official Function List to ensure that you are following
function list?	the guidelines for function naming. Functions that are not
	present on the approved list must be carefully velled for
	approval.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.
Basic Phage Information	
----------------------------------	----------------------
Phage Name	Akino08
Gene #	12
Stop Coordinate	6579
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	5395
Selected Function	major capsid protein

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes GeneMark, bp 5395
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves BabyVada and Avacadaman both had this gene
other annotated genomes?	res – Baby roua, and Avocadoman both had this gene
Does the gene violate any	The gene follows all guiding principles.
major guiding principles?	
DECISION:	Yes

Cothoring Evidence	Eveloie Voue Dotionale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and GeneMark suggest?	GeneMark Start Coordinate (type NA if not supported):: 5395
Does the start site have an	5395 – large z-score (3.262) and appropriate spacer (10 bp)
associated Ribosome	
Binding Site with a high	
score?	
Is the predicted start codon	Yes it is the longest ORF.
the longest ORF? If not,	
does the longest ORF	
result in excessive gene	
overlap (>30bp)?	
Is this start site conserved	
in other phage genomes	
as indicated by	Yes, It (5395) was found in 184 of 348 genes in the pham, and
Starterator?	called 52.9% of the time
	Note: if you are considering more than 1 start site provide the
	same information for each proposed start site
Is this start site conserved	Yes it is highly conserved.
in other phage genomes	
as indicated by BlastP?	
	Start site is bp 5395, based off of Genemark. Genemark coding
DECISION:	potential, phamerator, starterator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	Yes, Major capsid protein
a protein having a	
functional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁻⁴ or	
smaller with appropriate	
coverage?	
Does this protein align with	Scaffolding domain delta; Prohead I, icosahedral symmetry, HK97, phage,
a protein having a	capsid, VIRUS; 3.5A {Escherichia phage HK97}
functional assignment in	100%
the PDB or other database	
in HHPred with a	
probability of 90% or	
greater with appropriate	
coverage?	
Is this gene located	YES
adjacent to genes of	
known function and in a	
region of the genome that	
shows high conservation of	
gene order?	
Is this gene a possible	No
transmembrane protein?	
Is the proposed function	YES
tound on the SEA-	
PHAGES approved	
function list?	
DECISION:	major capsid protein

Basic Phage Information	
Phage Name	Akino08
Gene #	14
Stop Coordinate	6977
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	6648
Selected Function	Hypothetical Protein

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes GeneMark, bp 6648
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves – BabyVoda and Avocadoman both had this gane
other annotated genomes?	res – Baby roda and Avocadoman both had this gene
Does the gene violate any	The gene follows all guiding principles.
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and GeneMark suggest?	GeneMark Start Coordinate (type NA if not supported):: 6648
Does the start site have an associated Ribosome Binding Site with a high score?	6648 – large z-score (3.029) and appropriate spacer (8 bp)
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	Yes it is the longest ORF.
Is this start site conserved in other phage genomes	
as indicated by Starterator?	Yes, It (6648) was found in 60 of 60 genes in the pham, and called 100% of the time
	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes, it is highly conserved.
DECISION:	Start site is bp 6648, based off of Genemark, Genemark coding potential, phamerator, starterator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	No, Hypothetical Protein
a protein having a	
functional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁻⁴ or	
smaller with appropriate	
coverage?	
Does this protein align with	GH94N_like_1; Glycoside hydrolase family 94 N-terminal-like domain of
a protein having a	uncharacterized function.
functional assignment in	34.14%
the PDB or other database	
in HHPred with a	
probability of 90% or	
greater with appropriate	
coverage?	
Is this gene located	Yes
adjacent to genes of	
known function and in a	
region of the genome that	
shows high conservation of	
gene order?	
Is this gene a possible	NO
transmembrane protein?	
Is the proposed function	Yes
Iound on the SEA-	
FIAGES approved	
	11 moderation Duratain
DECISION:	Hypothetical Protein

Basic Phage Information	
Phage Name	Akino08
Gene #	15
Stop Coordinate	7563
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	7012
Selected Function	head-to-tail adaptor

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes Genemark, bp 7012
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves – BabyVoda and Avocadoman both had this gane
other annotated genomes?	res – baby roda and Avocadoman both had this gene
Does the gene violate any	The gene follows all guiding principles
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and GeneMark suggest?	GeneMark Start Coordinate (type NA if not supported):: 7012
Does the start site have an associated Ribosome Binding Site with a high score?	7012 – large z-score (3.120) and slightly large spacer (13 bp)
Is the predicted start codon	Yes it is the longest ORF.
the longest ORF? If not,	
does the longest ORF	
result in excessive gene	
overlap (>30bp)?	
Is this start site conserved	
in other phage genomes	
as indicated by	Yes, It (7012) was found in 82 of 240 genes in the pham, and
Starterator?	called 34.2% of the time
Is this start site conserved	Yes, it is highly conserved.
in other phage genomes	
as indicated by BlastP?	
DECISION:	Start site is bp 7012, based off of Genemark, Genemark coding potential, phamerator, starterator, and blast data.
	perential, priamerator, stationator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	Yes, head-to-tail adaptor
a protein having a	
functional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁻⁴ or	
smaller with appropriate	
coverage?	
Does this protein align with	Head completion protein; Neck, Portal, T5, VIRUS, VIRAL PROTEIN; 3.2A
a protein having a	{Escherichia phage DT57C}
functional assignment in	99.95%
the PDB or other database	
in HHPred with a	
probability of 90% or	
greater with appropriate	
coverage?	
Is this gene located	YES
adjacent to genes of	
known function and in a	
region of the genome that	
shows high conservation of	
gene order?	
Is this gene a possible	No
transmembrane protein?	
Is the proposed function	Yes
found on the SEA-	
PHAGES approved	
function list?	
DECISION:	head-to-tail adaptor

Basic Phage Information	
Phage Name	Akino08
Gene #	17
Stop Coordinate	7673
Direction (For/Rev)	Rev
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	8554
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an auto-annotation program (Glimmer, GeneMark)?	Original Glimmer cell @bp 8554 has strength 5.40 not called by Genemark
Is there evidence for coding potential?	Genemark did not show coding potential
Is this gene present in other annotated genomes?	Yes- Loviatar_Draft had this gene
Does the gene violate any major guiding principles?	The gene does not follow all guiding prinicples
DECISION:	NO

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and Genemark suggest?	Genemark Start Coordinate (type NA If not supported)::
Does the start site have an associated Ribosome	List the final RBS score and Z-score of the currently predicted
Binding Site with a high	in your response if this is the best score or not.
score?	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is the predicted start codon	Indicate the length of the ORF is with the predicted start and the
does the longest ORF	Does the proposed start site have a gap/overlap with the
result in excessive gene overlap (>30bp)?	nearest upstream gene that does not violate the Guiding Principles?
	Notes if you are considering more then 4 start site, more ide the
	same information for each proposed start site.
Is this start site conserved	
as indicated by Starterator?	You will also need to provide the following information from Starterator: does the start match the consensus start site predicted from Starterator? If no, is the consensus start site not found in this ORF? If no, is there a better option for the consensus start site instead of the one predicted by Starterator? If Starterator doesn't reveal a consensus start site, you can record that Starterator was not informative.
	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
Is this start site conserved in other phage genomes as indicated by BlastP?	Provide the best BlastP match from NCBI, PhagesDB, and DNA Master with alignment in the format of (Q#:S#), where Q (query) is the sequence you are analyzing and S (subject) is the database match. List the e-value and alignment of the best match for all three BlastP sources.
	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
DECISION:	Record where you think the gene should start here and briefly explain your rationale.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
snows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
DUACES approved	Once you have arrived at a functional decision, check the <u>SEA-</u>
function list2	the guidelines for function List to ensure that you are following
	the guidelines for function naming. Functions that are not
	approval
	approval.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	16
Stop Coordinate	7892
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	7563
Selected Function	head-to-tail stopper

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes Genemark, bp 7563
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves — Albedo and Albright both bad this gene
other annotated genomes?	res – Albedo and Albright both had this gene
Does the gene violate any	The gene follows all guiding principles
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and GeneMark suggest?	GeneMark Start Coordinate (type NA if not supported):: 7563
Does the start site have an	7563 – large z-score (3.029) and appropriate spacer (9 bp)
associated Ribosome	
Binding Site with a high	
score?	
Is the predicted start codon	No it is not the longest ORF, but it does not have any maior
the longest ORF? If not.	overlap with other genes
does the longest ORF	genee
result in excessive gene	Note: if you are considering more than 1 start site, provide the
overlap (>30bp)?	same information for each proposed start site
Is this start site conserved	
in other phage genomes	
as indicated by	Yes It was found in 122 of 156 genes in the pham and called
Starterator?	78.2% of the time
Otarterator :	
Is this start site conserved	Yes it is highly conserved
in other phage genomes	
as indicated by BlastP2	
as indicated by DiaStP?	Chart site is by 7562 based off of Consumary, Consumerity and inc
DECISION:	Start site is pp 7563, based off of Genemark, Genemark coding
	potential, phamerator, starterator, and plast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	Yes, head-to-tail stopper
a protein having a	
functional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁻⁴ or	
smaller with appropriate	
coverage?	
Does this protein align with	Head-to-tail stopper; Bacteriophage, portal, VIRAL PROTEIN;{Mycobacterium
a protein having a	phage Bxb1}
functional assignment in	99.65%
the PDB or other database	
in HHPred with a	
probability of 90% or	
greater with appropriate	
coverage?	
Is this gene located	YES
adjacent to genes of	
known function and in a	
region of the genome that	
shows high conservation of	
gene order?	
Is this gene a possible	No
transmembrane protein?	
Is the proposed function	Yes
found on the SEA-	
PHAGES approved	
function list?	
DECISION:	head-to-tail stopper

Basic Phage Information	
Phage Name	Akino08
Gene #	18
Stop Coordinate	8216
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	7908
Selected Function	Hypothetical protein

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes GeneMark, bp 7908
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves – Albedo and Albright both bad this gene
other annotated genomes?	Tes – Albedo and Albright both had this gene
Does the gene violate any	The gene follows all guiding principles.
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer and GeneMark suggest?	<i>Glimmer Start Coordinate (type NA if not supported)::</i> <i>GeneMark Start Coordinate (type NA if not supported):: 7908</i>
Does the start site have an associated Ribosome Binding Site with a high score?	7908 – low z-score (1.145) and appropriate spacer (10 bp)
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	Not the longest ORF, but does not overlap with other genes
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, It (7908) was found in 2 of 363 genes in the pham, and called 0.6% of the time Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes, it is highly conserved.
DECISION:	Start site is bp 7908, based off of Genemark, Genemark coding potential, phamerator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	Yes, hypothetical protein QDW21_gp10
a protein having a	
functional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁻⁴ or	
smaller with appropriate	
coverage?	
Does this protein align with	; DUF5403 ; Family of unknown function (DUF5403)
a protein having a	94.99%
functional assignment in	
the PDB or other database	
in HHPred with a	
probability of 90% or	
greater with appropriate	
coverage?	
Is this gene located	No
adjacent to genes of	
known function and in a	
region of the genome that	
shows high conservation of	
gene order?	
Is this gene a possible	NO
transmembrane protein?	
is the proposed function	Yes
Tound on the SEA-	
PHAGES approved	
	L hundhatiaal Duatain
DECISION:	Hypothetical Protein

Basic Phage Information	
Phage Name	Akino08
Gene #	19
Stop Coordinate	8611
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	8216
Selected Function	Tail terminator

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes Genemark, bp 8216
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves — Albedo and Albright both bad this gene
other annotated genomes?	res – Albedo and Albright both had this gene
Does the gene violate any	The gene follows all guiding principles
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer and GeneMark suggest?	Glimmer Start Coordinate (type NA if not supported):: GeneMark Start Coordinate (type NA if not supported):: 8216
Does the start site have an associated Ribosome Binding Site with a high score?	8216 – moderate z-score (2.776) and appropriate spacer (10 bp)
Is the predicted start codon the longest ORF? If not, does the longest ORF	Not the longest ORF, but the longest has no overlap with other genomes
result in excessive gene overlap (>30bp)?	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, It (8216) was found in 77 of 419 genes in the pham, and called 18.4% of the time
	same information for each proposed start site.
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes, it is highly conserved.
DECISION:	Start site is bp 8216, based off of Genemark, Genemark coding potential, phamerator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	Yes, tail terminator
a protein having a	
functional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁻⁴ or	
smaller with appropriate	
coverage?	
Does this protein align with	Tail terminator; Bacteriophage, portal, VIRAL
a protein having a	PROTEIN;{Mycobacterium phage Bxb1}
functional assignment in	99.79%
the PDB or other database	
in HHPred with a	
probability of 90% or	
greater with appropriate	
coverage?	
Is this gene located	No
adjacent to genes of	
known function and in a	
region of the genome that	
shows high conservation of	
gene order?	
Is this gene a possible	NO
transmembrane protein?	
Is the proposed function	Yes
Tound on the SEA-	
PHAGES approved	
	Tail ta main atau
DECISION:	i all terminator

Basic Phage Information	
Phage Name	Akino08
Gene #	21
Stop Coordinate	8640
Direction (For/Rev)	Rev
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	9416
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an auto-annotation program (Glimmer, GeneMark)?	Original Glimmer all @bp 9416 has strength 10.05 not called by Genemark
Is there evidence for coding potential?	Genemark did not show coding potential
Is this gene present in other annotated genomes?	Yes- Lovitar_Draft had this gene
Does the gene violate any major guiding principles?	The gene does not follow all guiding principles
DECISION:	No

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and Genemark suggest?	Genemark Start Coordinate (type NA If not supported)::
Does the start site have an associated Ribosome	List the final RBS score and Z-score of the currently predicted
Binding Site with a high	in your response if this is the best score or not.
score?	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is the predicted start codon	Indicate the length of the ORF is with the predicted start and the
does the longest ORF	Does the proposed start site have a gap/overlap with the
result in excessive gene overlap (>30bp)?	nearest upstream gene that does not violate the Guiding Principles?
	Notes if you are considering more then 4 start site, more ide the
	same information for each proposed start site.
Is this start site conserved	
as indicated by Starterator?	You will also need to provide the following information from Starterator: does the start match the consensus start site predicted from Starterator? If no, is the consensus start site not found in this ORF? If no, is there a better option for the consensus start site instead of the one predicted by Starterator? If Starterator doesn't reveal a consensus start site, you can record that Starterator was not informative.
	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
Is this start site conserved in other phage genomes as indicated by BlastP?	Provide the best BlastP match from NCBI, PhagesDB, and DNA Master with alignment in the format of (Q#:S#), where Q (query) is the sequence you are analyzing and S (subject) is the database match. List the e-value and alignment of the best match for all three BlastP sources.
	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
DECISION:	Record where you think the gene should start here and briefly explain your rationale.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
snows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
DUACES approved	Once you have arrived at a functional decision, check the <u>SEA-</u>
function list2	the guidelines for function List to ensure that you are following
	the guidelines for function naming. Functions that are not
	approval
	approval.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	20
Stop Coordinate	9422
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	8625
Selected Function	Major tail protein

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes Genemark, bp 8625
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves – Albedo and Abigail both bad this gene
other annotated genomes?	res – Albedo and Abigali both had this gene
Does the gene violate any	The gene follows all guiding principles
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer and GeneMark suggest?	Glimmer Start Coordinate (type NA if not supported):: GeneMark Start Coordinate (type NA if not supported):: 8625
Does the start site have an associated Ribosome Binding Site with a high score?	8625 – large z-score (3.262) and appropriate spacer (11 bp)
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	Yes it is the longest ORF.
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, It (8625) was found in 248 of 569 genes in the pham, and called 43.6% of the time
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes, it is highly conserved.
DECISION:	Start site is bp 8625, based off of Genemark, Genemark coding potential, phamerator, starterator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	Yes, major tail protein
a protein having a	
functional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁻⁴ or	
smaller with appropriate	
coverage?	
Does this protein align with	Major tail protein; Bacteriophage, tail tube, VIRUS, VIRAL
a protein having a	PROTEIN;{Mycobacterium phage Bxb1}
functional assignment in	100%
the PDB or other database	
in HHPred with a	
probability of 90% or	
greater with appropriate	
coverage?	
Is this gene located	YES
adjacent to genes of	
known function and in a	
region of the genome that	
shows high conservation of	
gene order?	
Is this gene a possible	No
transmembrane protein?	
Is the proposed function	Yes
tound on the SEA-	
PHAGES approved	
tunction list?	
DECISION:	Major tail protein

Basic Phage Information	
Phage Name	Akino08
Gene #	22
Stop Coordinate	9826
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	9518
Selected Function	tail assembly chaperone

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes Genemark, bp 9518
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves — Albedo and Albright both bad this gene
other annotated genomes?	res – Albedo and Albright both had this gene
Does the gene violate any	The gene follows all guiding principles
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and GeneMark suggest?	GeneMark Start Coordinate (type NA if not supported):: 9518
Does the start site have an associated Ribosome Binding Site with a high score?	9518 – small z-score (1.953) and appropriate spacer (10 bp)
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	Yes it is the longest ORF.
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, It (9518) was found in 116 of 139 genes in the pham, and called 83.5% of the time
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes, it is highly conserved.
DECISION:	Start site is bp 9518, based off of Genemark, Genemark coding potential, phamerator, starterator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	Yes, tail assembly chaperone
a protein having a	
functional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁻⁴ or	
smaller with appropriate	
coverage?	
Does this protein align with	; GP24_25 ; Mycobacteriophage tail assembly protein
a protein having a	97%
functional assignment in	
the PDB or other database	
in HHPred with a	
probability of 90% or	
greater with appropriate	
coverage?	
Is this gene located	Yes, known functions include tail assembly chaperones
adjacent to genes of	
known function and in a	
region of the genome that	
shows high conservation of	
gene order?	
Is this gene a possible	NO
transmembrane protein?	Vac teil accomption at an arrange
Is the proposed function	Yes, tall assembly chaperone
Iound on the SEA-	
function list?	
	tail accombly changes
DECISION:	tall assembly chaperone

Basic Phage Information	
Phage Name	Akino08
Gene #	23
Stop Coordinate	9850
Direction (For/Rev)	Rev
Gap (Overlap) with Previous Gene	Yes
Selected Start Coordinate	10074
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes Genemark, bp 10074
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed no coding potential
coding potential?	
Is this gene present in	Ves_Loviatar. Draft had this gene
other annotated genomes?	Tes - Lovialar_Drait had this gene
Does the gene violate any	The gene does not follow all guiding principles
major guiding principles?	
DECISION:	No, deleted

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and GeneMark suggest?	GeneMark Start Coordinate (type NA if not supported)::
Does the start site have an	List the final RBS score and Z-score of the currently predicted
associated Ribosome	start site using the Kibler6/Karlin Medium scoring table. Indicate
Binding Site with a high	in your response if this is the best score or not.
score?	
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is the predicted start codon	Indicate the length of the ORF is with the predicted start and the
the longest ORF? If not,	gap/overlap to the nearest stop codon of the upstream ORF.
does the longest ORF	Does the proposed start site have a gap/overlap with the
result in excessive gene	nearest upstream gene that does not violate the Guiding
overlap (>30bp)?	Principles?
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is this start site conserved	
in other phage genomes	
as indicated by	You will also need to provide the following information from
Starterator?	Starterator: does the start match the consensus start site
	predicted from Starterator? If no, is the consensus start site not
	found in this ORF? If no, is there a better option for the
	consensus start site instead of the one predicted by Starterator?
	If Starterator doesn't reveal a consensus start site, you can
	record that Starterator was not informative.
	Note: If you are considering more than 1 start site, provide the
	same information for each proposed start site.
le this start site conserved	Provide the best PlastP metab from NCPL PhagesDP and DNA
in other phage genemoe	Moster with alignment in the formet of (Ott:St), where O (quary)
in other phage genomes	is the sequence you are enclyzing and S (cubicet) is the
as indicated by blastr?	Is the sequence you are analyzing and S (Subject) is the
	meteb for all three PlactD sources
	Noto: if you are considering more than 1 start site, provide the
	some information for each proposed start site
	Record where you think the gene should start here and briefly
DECISION:	explain your rationale

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
snows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
DUACES approved	Once you have arrived at a functional decision, check the <u>SEA-</u>
function list2	the guidelines for function List to ensure that you are following
	the guidelines for function naming. Functions that are not
	approval
	approval.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	30
Stop Coordinate	12975
Direction (For/Rev)	Rev
Gap (Overlap) with Previous Gene	Yes
Selected Start Coordinate	13718
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an auto-annotation program (Glimmer, GeneMark)?	Original Glimmer cell @bp 13718 has strength 8.47 not called by Genemark
Is there evidence for coding potential?	Genemark did not show coding potential
Is this gene present in other annotated genomes?	Yes-Loviatar_Draft had this gene
Does the gene violate any major guiding principles?	The gene does not follow all guiding principles
DECISION:	NO

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and Genemark suggest?	Genemark Start Coordinate (type NA If not supported)::
Does the start site have an	List the final RBS score and Z-score of the currently predicted
Binding Site with a high	in your response if this is the best score or not.
score?	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is the predicted start codon	Indicate the length of the ORF is with the predicted start and the
does the longest ORF	Does the proposed start site have a gap/overlap with the
result in excessive gene overlap (>30bp)?	nearest upstream gene that does not violate the Guiding Principles?
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is this start site conserved in other phage genomes	
as indicated by Starterator?	You will also need to provide the following information from Starterator: does the start match the consensus start site predicted from Starterator? If no, is the consensus start site not found in this ORF? If no, is there a better option for the consensus start site instead of the one predicted by Starterator? If Starterator doesn't reveal a consensus start site, you can record that Starterator was not informative.
	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
Is this start site conserved in other phage genomes as indicated by BlastP?	Provide the best BlastP match from NCBI, PhagesDB, and DNA Master with alignment in the format of (Q#:S#), where Q (query) is the sequence you are analyzing and S (subject) is the database match. List the e-value and alignment of the best
	match for all three BlastP sources.
	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
DECISION:	Record where you think the gene should start here and briefly explain your rationale.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
snows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
DUACES approved	Once you have arrived at a functional decision, check the <u>SEA-</u>
function list2	the guidelines for function List to ensure that you are following
	the guidelines for function naming. Functions that are not
	approval
	approval.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.
Basic Phage Information	
----------------------------------	----------------------
Phage Name	Akino08
Gene #	28
Stop Coordinate	13050
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	10420
Selected Function	tape measure protein

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes Genemark, bp 10420
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves – Albedo and Albright both bad this gene
other annotated genomes?	Tes – Albedo and Albright both had this gene
Does the gene violate any	The gene follows all guiding principles
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Your Pationale
What start site do Glimmer and GeneMark suggest?	Glimmer Start Coordinate (type NA if not supported):: GeneMark Start Coordinate (type NA if not supported):: 10420
Does the start site have an associated Ribosome Binding Site with a high score?	10420 – moderate z-score (2.178) and appropriate spacer (10 bp)
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	Yes It is the longest ORF
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, It (10420) was found in 65 of 71 genes in the pham, and called 91.5% of the time
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes, it is highly conserved.
DECISION:	Start site is bp 10420, based off of Genemark, Genemark coding potential, phamerator, starterator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	Yes, tape measure protein
a protein having a	
functional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁻⁴ or	
smaller with appropriate	
coverage?	
Does this protein align with	Tape measure protein (CD1366); Phage tail-like, bacteriocin, baseplate, pre-
a protein having a	contraction, VIRUS LIKE PARTICLE;{Clostridio
functional assignment in	00 540/
the PDB or other database	99.51%
in HHPred with a	
probability of 90% or	
greater with appropriate	
coverage?	
Is this gene located	Yes, tape measure protein
adjacent to genes of	
known function and in a	
region of the genome that	
shows high conservation of	
gene order?	
Is this gene a possible	No
transmembrane protein?	
Is the proposed function	Yes, tape measure protein
found on the SEA-	
PHAGES approved	
function list?	
DECISION:	tape measure protein

Basic Phage Information	
Phage Name	Akino08
Gene #	32
Stop Coordinate	13874
Direction (For/Rev)	Rev
Gap (Overlap) with Previous Gene	Yes
Selected Start Coordinate	15751
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	Original Glimmer call @bp 15751 has strength 9.25 not called
(Glimmer, GeneMark)?	by genemark
Is there evidence for coding potential?	Genemark did not show coding potential
Is this gene present in other annotated genomes?	Yes-Loviatar_Draft had this gene
Does the gene violate any major guiding principles?	The gene does not follow all guiding principles
DECISION:	No

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and Genemark suggest?	Genemark Start Coordinate (type NA If not supported)::
Does the start site have an associated Ribosome	List the final RBS score and Z-score of the currently predicted
Binding Site with a high	in your response if this is the best score or not.
score?	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is the predicted start codon	Indicate the length of the ORF is with the predicted start and the
does the longest ORF	Does the proposed start site have a gap/overlap with the
result in excessive gene overlap (>30bp)?	nearest upstream gene that does not violate the Guiding Principles?
	Notes if you are considering more then 4 start site, more ide the
	same information for each proposed start site.
Is this start site conserved	
as indicated by Starterator?	You will also need to provide the following information from Starterator: does the start match the consensus start site predicted from Starterator? If no, is the consensus start site not found in this ORF? If no, is there a better option for the consensus start site instead of the one predicted by Starterator? If Starterator doesn't reveal a consensus start site, you can record that Starterator was not informative.
	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
Is this start site conserved in other phage genomes as indicated by BlastP?	Provide the best BlastP match from NCBI, PhagesDB, and DNA Master with alignment in the format of (Q#:S#), where Q (query) is the sequence you are analyzing and S (subject) is the database match. List the e-value and alignment of the best match for all three BlastP sources.
	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
DECISION:	Record where you think the gene should start here and briefly explain your rationale.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
snows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
DUACES approved	Once you have arrived at a functional decision, check the <u>SEA-</u>
function list2	the guidelines for function List to ensure that you are following
	the guidelines for function naming. Functions that are not
	approval
	approval.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	31
Stop Coordinate	13904
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	13047
Selected Function	Minor Tail Protein

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes Genemark, bp 13047
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves Albedo and Abigail both had this gene
other annotated genomes?	res – Albedo and Abigali bolit nad litis gene
Does the gene violate any	The gene follows all guiding principles
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer and GeneMark suggest?	<i>Glimmer Start Coordinate (type NA if not supported)::</i> <i>GeneMark Start Coordinate (type NA if not supported)::</i> 13047
Does the start site have an associated Ribosome Binding Site with a high score?	13047 – moderate z-score (2.348) and appropriate spacer (12 bp)
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	Yes It is the longest ORF
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, It (13047) was found in 81 of 378 genes in the pham, and called 21.4% of the time
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes, it is highly conserved.
DECISION:	Start site is bp 13047, based off of Genemark, Genemark coding potential, phamerator, starterator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	Yes, minor tail protein
a protein having a	
functional assignment in	Hint: you may have already found this information from
BlastP (phagesDB and/or	annotation decision #2. Provide the alignment (q#:s#) and e-
GenBank) with an	value. It is only necessary to provide one match from each
alignment of 10 ⁻⁴ or	database.
smaller with appropriate	
coverage?	
Does this protein align with	HYPOTHETICAL PROTEIN 19.1; VIRAL PROTEIN, DISTAL TAIL PROTEIN; 2.95A
a protein having a	{BACILLUS PHAGE SPP1}
functional assignment in	99.96%
the PDB or other database	Note: If you believe there is not a quality HHPred match, type
In HHPred with a	No. Quality Match and list the data for the best match available
probability of 90% or	to affirm the noor quality of the result and to document that
greater with appropriate	HHPred was considered
	Vee
adjacent to genes of	res
known function and in a	
region of the genome that	
shows high conservation of	
dene order?	
Is this gene a possible	No
transmembrane protein?	
Is the proposed function	Yes
found on the SEA-	
PHAGES approved	
function list?	
DECISION:	Minor Tail Protein

Basic Phage Information	
Phage Name	Akino08
Gene #	34
Stop Coordinate	15808
Direction (For/Rev)	Rev
Gap (Overlap) with Previous Gene	Yes
Selected Start Coordinate	15966
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an auto-annotation program	Original Glimmer call @bp 15966 has strength 3.41 not called
(Glimmer, GeneMark)?	by Cenemark
Is there evidence for coding potential?	Genemark did not show coding potential
Is this gene present in other annotated genomes?	Yes- Loviatar_Draft had this gene
Does the gene violate any major guiding principles?	The gene does not follow all guiding prinicples
DECISION:	No

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and Genemark suggest?	Genemark Start Coordinate (type NA If not supported)::
Does the start site have an associated Ribosome	List the final RBS score and Z-score of the currently predicted
Binding Site with a high	in your response if this is the best score or not.
score?	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is the predicted start codon	Indicate the length of the ORF is with the predicted start and the
does the longest ORF	Does the proposed start site have a gap/overlap with the
result in excessive gene overlap (>30bp)?	nearest upstream gene that does not violate the Guiding Principles?
	Notes if you are considering more then 4 start site, more ide the
	same information for each proposed start site.
Is this start site conserved	
as indicated by Starterator?	You will also need to provide the following information from Starterator: does the start match the consensus start site predicted from Starterator? If no, is the consensus start site not found in this ORF? If no, is there a better option for the consensus start site instead of the one predicted by Starterator? If Starterator doesn't reveal a consensus start site, you can record that Starterator was not informative.
	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
Is this start site conserved in other phage genomes as indicated by BlastP?	Provide the best BlastP match from NCBI, PhagesDB, and DNA Master with alignment in the format of (Q#:S#), where Q (query) is the sequence you are analyzing and S (subject) is the database match. List the e-value and alignment of the best match for all three BlastP sources.
	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
DECISION:	Record where you think the gene should start here and briefly explain your rationale.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
snows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
DUACES approved	Once you have arrived at a functional decision, check the <u>SEA-</u>
function list2	the guidelines for function List to ensure that you are following
	the guidelines for function naming. Functions that are not
	approval
	approval.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	35
Stop Coordinate	15942
Direction (For/Rev)	Rev
Gap (Overlap) with Previous Gene	Yes
Selected Start Coordinate	19190
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes Glimmer, bp 19190
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed no coding potential
coding potential?	
Is this gene present in	Ves_Loviatar. Draft had this gene
other annotated genomes?	Tes - Lovialar_Drait had this gene
Does the gene violate any	The gene does not follow all guiding principles
major guiding principles?	
DECISION:	No deleted

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and GeneMark suggest?	GeneMark Start Coordinate (type NA if not supported)::
Does the start site have an	List the final RBS score and Z-score of the currently predicted
associated Ribosome	start site using the Kiblerb/Karlin Medium scoring table. Indicate
Binding Sile with a high	In your response if this is the best score of not.
300101	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is the predicted start codon	Indicate the length of the ORF is with the predicted start and the
the longest ORF? If not,	gap/overlap to the nearest stop codon of the upstream ORF.
does the longest ORF	Does the proposed start site have a gap/overlap with the
overlap (>30bp)?	Principles?
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is this start site conserved	
In other phage genomes	Very will also pood to provide the following information from
Starterator?	Starterator: does the start match the consensus start site
	predicted from Starterator? If no, is the consensus start site not
	found in this ORF? If no, is there a better option for the
	consensus start site instead of the one predicted by Starterator?
	If Starterator doesn't reveal a consensus start site, you can
	record that Starterator was not informative.
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is this start site conserved	Provide the best BlastP match from NCBI, PhagesDB, and DNA
in other phage genomes	Master with alignment in the format of (Q#:S#), where Q (query)
as indicated by BlastP?	is the sequence you are analyzing and S (subject) is the
	database match. List the e-value and alignment of the best match for all three PlastP sources
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
DECISION	Record where you think the gene should start here and briefly
	explain your rationale.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
shows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
Tound on the SEA-	Once you have arrived at a functional decision, check the <u>SEA-</u>
PHAGES approved	PHAGES Official Function List to ensure that you are following
	the guidelines for function harming. Functions that are not
	present on the approved list must be carefully velled for
	approval.
	If you believe this gene should be assigned, please write the
DEGICIONI	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	33
Stop Coordinate	16030
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	13904
Selected Function	Minor tail protein

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes Genemark, bp 13904
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves – BabyVoda and DirtyBubble both had this gene
other annotated genomes?	res – Daby roda and Dirty Bubble both had this gene
Does the gene violate any	The gene follows all guiding principles
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer and GeneMark suggest?	Glimmer Start Coordinate (type NA if not supported):: GeneMark Start Coordinate (type NA if not supported):: 13904
Does the start site have an associated Ribosome Binding Site with a high score?	13904 – low z-score (1.963) and appropriate spacer (12 bp)
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	Yes It is the longest ORF
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, It (13904) was found in 11 of 11 genes in the pham, and called 100% of the time
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes, it is highly conserved.
DECISION:	Start site is bp 13904, based off of Genemark, Genemark coding potential, phamerator, starterator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	Yes, minor tail protein
a protein having a	
functional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁴ or	
smaller with appropriate	
Coverage?	
Does this protein align with	Minor tail protein; Bacteriophage, tail tip, VIRAL
functional assignment in	PROTEIN;{Mycobacterium phage Bxb1}
the PDB or other database	99.96%
in HHPred with a	
probability of 90% or	
greater with appropriate	Note: If you believe there is not a quality HHPred match, type
coverage?	No Quality Match and list the data for the best match available
cororage.	to affirm the poor quality of the result and to document that
	HHPred was considered.
Is this gene located	Yes
adjacent to genes of	
known function and in a	
region of the genome that	
shows high conservation of	
gene order?	No
transmombrane protein?	NO
	Voc
found on the SEA	100
PHACES approved	
function list?	
DECISION	Minor Tail protein

Basic Phage Information	
Phage Name	Akino08
Gene #	36
Stop Coordinate	16880
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	16044
Selected Function	Minor Tail protein

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes Genemark, bp 16044
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves – Albedo and Abigail both bad this gene
other annotated genomes?	res – Albedo and Abigali both had this gene
Does the gene violate any	The gene follows all guiding principles
major guiding principles?	
DECISION:	Yes

Cothoring Evidence	Explain Vour Potionala
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and GeneMark suggest?	GeneMark Start Coordinate (type NA if not supported):: 16044
Does the start site have an	16044 – large z-score (3.273) and high spacer (13 bp)
associated Ribosome	5 () 5 7 ()
Binding Site with a high	
score?	
30016 !	
is the predicted start codon	Yes It is the longest ORF
the longest ORF? If not,	
does the longest ORF	
result in excessive gene	
overlap (>30bp)?	
Is this start site conserved	
in other phage genomes	
as indicated by	Ves It (16044) was found in 175 of 178 genes in the pham and
Startorator?	colled 08 29/ of the time
Starterators	
	Nee it is highly segreened
is this start site conserved	Yes, it is nignly conserved.
in other phage genomes	
as indicated by BlastP?	
DECISION	Start site is bp 16044, based off of Genemark, Genemark
DECISION.	coding potential, phamerator, starterator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	Yes, Minor tail protein
a protein having a	
functional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁻⁴ or	
smaller with appropriate	
coverage?	
Does this protein align with	Receptor Binding Protein; beta sandwich domain, phage
a protein having a	receptor binding protein, Lactococcus lactis pellicle cell wall
functional assignment in	polypnosphosaccharide, VIRAL PROTEIN; 1.75A (Lactococcus
ine PDB or other database	pnage 1358}
IN HHPIED WITH a	
greater with appropriate	
greater with appropriate	
ls this gene located	Vec
adjacent to denes of	103
known function and in a	
region of the genome that	
shows high conservation of	
gene order?	
Is this gene a possible	No
transmembrane protein?	
Is the proposed function	Yes
found on the SEA-	
PHAGES approved	
function list?	
DECISION:	Minor tail protein

Basic Phage Information	
Phage Name	Akin08
Gene #	37
Stop Coordinate	17693
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	16890
Selected Function	Hypothetical

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes Genemark, bp 16890
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves – Albedo and Abigail both bad this gene
other annotated genomes?	res – Albedo and Abigan both had this gene
Does the gene violate any	The gene follows all guiding principles
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer and GeneMark suggest?	Glimmer Start Coordinate (type NA if not supported):: NA GeneMark Start Coordinate (type NA if not supported):: 16890
Does the start site have an associated Ribosome Binding Site with a high score?	16890 – moderate z-score (2.862) and appropriate spacer (9 bp)
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	Yes it is the longest ORF
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, It (16890) was found in 118 of 221 genes in the pham, and called 91.5% of the time
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes, it is highly conserved.
DECISION:	Start site is bp 16890, based off of Genemark, Genemark coding potential, phamerator, starterator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	Yes, Hypothetical Protein QDW21_gp20
a protein having a	
functional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁻⁴ or	
smaller with appropriate	
coverage?	
Does this protein align with	; DUF859 ; Siphovirus protein of unknown function (DUF859)
a protein having a	99.3%
functional assignment in	
the PDB or other database	
in HHPred with a	
probability of 90% or	
greater with appropriate	
coverage?	
Is this gene located	No
adjacent to genes of	
known function and in a	
region of the genome that	
shows high conservation of	
gene order?	
Is this gene a possible	No
transmembrane protein?	
Is the proposed function	Yes
found on the SEA-	
PHAGES approved	
function list?	
DECISION:	Hypothetical

Basic Phage Information	
Phage Name	Akino08
Gene #	39
Stop Coordinate	19370
Direction (For/Rev)	Rev
Gap (Overlap) with Previous Gene	Yes
Selected Start Coordinate	20533
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an auto-annotation program	Original Glimmer call @bp 20533 has strength 3.78 not called
(Glimmer, GeneMark)?	by Genemark
Is there evidence for coding potential?	Genemark did not show coding potential
Is this gene present in other annotated genomes?	Yes-Loviatar_Draft had this gene
Does the gene violate any major guiding principles?	The gene does not follow all guiding prinicples
DECISION:	No

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and Genemark suggest?	Genemark Start Coordinate (type NA If not supported)::
Does the start site have an associated Ribosome	List the final RBS score and Z-score of the currently predicted
Binding Site with a high	in your response if this is the best score or not.
score?	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is the predicted start codon	Indicate the length of the ORF is with the predicted start and the
does the longest ORF	Does the proposed start site have a gap/overlap with the
result in excessive gene overlap (>30bp)?	nearest upstream gene that does not violate the Guiding Principles?
	Notes if you are considering more then 4 start site may ide the
	same information for each proposed start site.
Is this start site conserved	
as indicated by Starterator?	You will also need to provide the following information from Starterator: does the start match the consensus start site predicted from Starterator? If no, is the consensus start site not found in this ORF? If no, is there a better option for the consensus start site instead of the one predicted by Starterator? If Starterator doesn't reveal a consensus start site, you can record that Starterator was not informative.
	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
Is this start site conserved in other phage genomes as indicated by BlastP?	Provide the best BlastP match from NCBI, PhagesDB, and DNA Master with alignment in the format of (Q#:S#), where Q (query) is the sequence you are analyzing and S (subject) is the database match. List the e-value and alignment of the best match for all three BlastP sources.
	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
DECISION:	Record where you think the gene should start here and briefly explain your rationale.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
snows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
DUACES approved	Once you have arrived at a functional decision, check the <u>SEA-</u>
function list2	the guidelines for function List to ensure that you are following
	the guidelines for function naming. Functions that are not
	approval
	approval.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	38
Stop Coordinate	19409
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	17703
Selected Function	Hypothetical

Gathering Evidence	Explain Your Rationale
Was the gene called by an auto-annotation program	Yes Glimmer (bp 19248 strength of 8.83) and GeneMark (bp
(Glimmer, GeneMark)?	17703)
Is there evidence for	GeneMarkS showed coding potential at 17703
coding potential?	
Is this gene present in other annotated genomes?	Yes – Loviatar_Draft had this gene
Does the gene violate any major guiding principles?	The gene follows all guiding principles
DECISION:	Yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported):: 19248
and GeneMark suggest?	GeneMark Start Coordinate (type NA if not supported):: 17703
Does the start site have an associated Ribosome Binding Site with a high score?	17703 – large z-score (3.273) and appropriate spacer (9 bp)
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	Yes It is the longest ORF
Is this start site conserved	
In other phage genomes	Vec It (100,10) were found in 0 of 0 menos in the show and
Starterator?	called 100% of the time
Is this start site conserved	Yes, it is highly conserved.
in other phage genomes as indicated by BlastP?	
DECISION:	Start site is bp 17703, based off of Genemark, Genemark coding potential, phamerator, starterator, and blast data. We went with 17703 as the start position instead of 19248 because the coding potential is greatest at the start of 17703. Additionally, the coding potential is really consistent throughout the 18000's up until the end point. Starterator says 19248, however according to our previous gene, it would make more sense for the gene to start at 17703. Lastly, the RBS of the ORF showed the z value at 17703 to be the greatest, a good spacer distance, the lowest final score, and longest ORF length. For all of these reasons, we concluded gene 38 (19409) to start at 17703.

Gathering Evidence	Explain Your Rationale
Does this protein align with	Yes, endolysin
a protein having a	
functional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁻⁴ or	
smaller with appropriate	
coverage?	
Does this protein align with	No quality match; DUF1949 ; Domain of unknown function
a protein having a	(DUF1949)
functional assignment in	64.14%
the PDB or other database	
in HHPred with a	
probability of 90% or	
greater with appropriate	
coverage?	
Is this gene located	No
adjacent to genes of	
known function and in a	
region of the genome that	
shows high conservation of	
gene order?	
Is this gene a possible	No
transmembrane protein?	
Is the proposed function	Yes
tound on the SEA-	
PHAGES approved	
function list?	
DECISION:	Hypothetical

Basic Phage Information	
Phage Name	Akino08
Gene #	40
Stop Coordinate	19891
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	19409
Selected Function	Hypothetical Protein (Transmembrane?)

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes Genemark, bp 19409
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves Abigail and Arrovo both had this gene
other annotated genomes?	res – Abigali and Anoyo both had this gene
Does the gene violate any	The gene follows all guiding principles
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer and GeneMark suggest?	Glimmer Start Coordinate (type NA if not supported):: GeneMark Start Coordinate (type NA if not supported)::19409
Does the start site have an associated Ribosome Binding Site with a high score?	19409 – moderate z-score (2.549) and low spacer (6 bp)
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	Yes It is the longest ORF
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, It (19409) was found in 66 of 66 genes in the pham, and called 100% of the time
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes, it is highly conserved.
DECISION:	Start site is bp 19409, based off of Genemark, Genemark coding potential, phamerator, starterator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	Yes, hypothetical
a protein having a	
functional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁻⁴ or	
smaller with appropriate	
coverage?	
Does this protein align with	No quality match ; DUF6264 ; Family of unknown function
a protein having a	(DUF6264) 76.8%
functional assignment in	
the PDB or other database	
in HHPred with a	
probability of 90% or	
greater with appropriate	
coverage?	
Is this gene located	No
adjacent to genes of	
known function and in a	
region of the genome that	
shows high conservation of	
gene order?	
Is this gene a possible	Yes
transmembrane protein?	
is the proposed function	Yes
Iound on the SEA-	
FIAGES approved	
	I hundhatiaal but inaludaa a tuana wana wana aa waaat
DECISION:	nypolitelical but includes a transmembrane segment

Basic Phage Information	
Phage Name	Akino08
Gene #	41
Stop Coordinate	20243
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	19902
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes Genemark, bp 19902
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves — Albedo and Albright both bad this gene
other annotated genomes?	res – Albedo and Albright both had this gene
Does the gene violate any	The gene follows all guiding principles
major guiding principles?	
DECISION:	Yes

Cathoring Evidence	Explain Vour Pationalo
what start site do Glimmer	Glimmer Start Coordinate (type NA If not supported)::
and GeneMark suggest?	GeneMark Start Coordinate (type NA if not supported):: 19902
Does the start site have an associated Ribosome Binding Site with a high score?	19902 – moderate z-score (2.708) and appropriate spacer (10 bp)
Is the predicted start codon	No, not the longest ORF but does not overlap
the longest ORF? If not,	
does the longest ORF	
result in excessive gene	
overlap (>30bp)?	
Is this start site conserved	
in other phage genomes	
as indicated by	Yes It (19902) was found in 87 of 224 genes in the pham and
Starterator?	called 38.8 % of the time
Is this start site conserved	Yes, it is highly conserved.
in other phage genomes	
as indicated by BlastP?	
DECISION:	Start site is bp 19902, based off of Genemark, Genemark coding potential, phamerator, starterator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	Yes, membrane protein
a protein having a	
functional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁻⁴ or	
smaller with appropriate	
coverage?	
Does this protein align with	; MTLN ; Mitoregulin
a protein having a	
functional assignment in	probability 95.82
the PDB or other database	
in HHPred with a	
probability of 90% or	
greater with appropriate	
coverage?	
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
shows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
found on the SEA-	Once you have arrived at a functional decision, check the <u>SEA-</u>
PHAGES approved	<u>PHAGES Official Function List</u> to ensure that you are following
function list?	the guidelines for function naming. Functions that are not
	present on the approved list must be carefully vetted for
	approval.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.
Basic Phage Information	
----------------------------------	---------
Phage Name	Akino08
Gene #	42
Stop Coordinate	20497
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	20246
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes GeneMark, bp 20246
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves — Albedo and Albright both bad this gene
other annotated genomes?	res – Albedo and Albright both had this gene
Does the gene violate any	The gene follows all guiding principles
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer and GeneMark suggest?	Glimmer Start Coordinate (type NA if not supported):: GeneMark Start Coordinate (type NA if not supported):: 20246
Does the start site have an associated Ribosome Binding Site with a high score?	20246 – moderate z-score (2.581) and low spacer (7 bp)
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	Yes It is the longest ORF
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, It (20246) was found in 55 of 55 genes in the pham, and called 100% of the time
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes, it is highly conserved.
DECISION:	Start site is bp 20246, based off of Genemark, Genemark coding potential, phamerator, starterator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	Yes, membrane protein
a protein having a	
functional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁻⁴ or	
smaller with appropriate	
coverage?	
Does this protein align with	; Halogen_Hydrol ; 5-bromo-4-chloroindolyl phosphate hydrolysis protein
a protein having a	
functional assignment in	Probability 73.12
the PDB or other database	
in HHPred with a	
probability of 90% or	
greater with appropriate	
coverage?	
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
shows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
found on the SEA-	Once you have arrived at a functional decision, check the <u>SEA-</u>
PHAGES approved	<u>PHAGES Official Function List</u> to ensure that you are following
function list?	the guidelines for function naming. Functions that are not
	present on the approved list must be carefully vetted for
	approval.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	43
Stop Coordinate	20645
Direction (For/Rev)	Rev
Gap (Overlap) with Previous Gene	Yes
Selected Start Coordinate	21139
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	Original Glimmer call @bp 21139 has strength 5 78 ** not called
auto-annotation program	by GeneMark
(Glimmer, GeneMark)?	by Cenewark
Is there evidence for	No coding potential
coding potential?	
Is this gene present in	Vac only found in Loviator Draft
other annotated genomes?	res – only lound in Lovialar_Drait
Does the gene violate any	The gene violates major guiding principles such as overlapping
major guiding principles?	other, stronger genes
DECISION:	Not a gene, deleted

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and Genemark suggest?	Genemark Start Coordinate (type NA If not supported)::
Does the start site have an associated Ribosome	List the final RBS score and Z-score of the currently predicted
Binding Site with a high	in your response if this is the best score or not.
score?	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is the predicted start codon	Indicate the length of the ORF is with the predicted start and the
does the longest ORF	Does the proposed start site have a gap/overlap with the
result in excessive gene overlap (>30bp)?	nearest upstream gene that does not violate the Guiding Principles?
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is this start site conserved in other phage genomes	
as indicated by Starterator?	You will also need to provide the following information from Starterator: does the start match the consensus start site predicted from Starterator? If no, is the consensus start site not found in this ORF? If no, is there a better option for the consensus start site instead of the one predicted by Starterator? If Starterator doesn't reveal a consensus start site, you can record that Starterator was not informative.
	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
Is this start site conserved in other phage genomes	Provide the best BlastP match from NCBI, PhagesDB, and DNA Master with alignment in the format of (Q#:S#), where Q (query)
as indicated by BlastP?	is the sequence you are analyzing and S (subject) is the
	match for all three BlastP sources.
	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
DECISION:	Record where you think the gene should start here and briefly explain your rationale.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
shows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
found on the SEA-	Once you have arrived at a functional decision, check the <u>SEA-</u>
PHAGES approved	PHAGES Official Function List to ensure that you are following
function list?	the guidelines for function naming. Functions that are not
	present on the approved list must be carefully velled for
	approval.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	44
Stop Coordinate	21310
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	20708
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	Ves Glimmer (bp 21113 strength of 5 76) and GeneMark (bp
auto-annotation program	20708)
(Glimmer, GeneMark)?	20100)
Is there evidence for	GeneMarkS showed coding potential at 20708
coding potential?	Genemarks showed county potential at 20700
Is this gene present in	Ves – AvocadoMan and Albright both had this gene
other annotated genomes?	Tes – Avocadoman and Albright both had this gene
Does the gene violate any	The gene follows all guiding principles
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported):: 21113
and GeneMark suggest?	GeneMark Start Coordinate (type NA if not supported):: 20708
Does the start site have an associated Ribosome Binding Site with a high score?	20708 – small z-score (1.979) and low spacer (7 bp)
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	No It is the longest ORF, but the gene does not overlap with the previous gene.
Is this start site conserved	
in other phage genomes	
as indicated by Starterator?	Yes, It (21113) was found in 2 of 2 genes in the pham, and called 100% of the time
Is this start site conserved	Yes, it is highly conserved.
in other phage genomes as indicated by BlastP?	
DECISION:	Start site is bp 20708, based off of Genemark, Genemark coding potential, phamerator, starterator, and blast data. We went with 20708 as the start position instead of 21113 because the coding potential is greatest at the start of 20708. Additionally, the coding potential is really consistent throughout the 21000's up until the end point. Starterator says 21113, however according to our previous gene, it would make more sense for the gene to start at 20708.

Gathering Evidence	Explain Your Rationale
Does this protein align with	Yes, hypothetical
a protein having a	
functional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁻⁴ or	
smaller with appropriate	
coverage?	
Does this protein align with	Met_tRNA_FMT_C; C-terminal domain of Formyltransferase and other
a protein having a	enzymes. C-terminal domain of formyl transferase and ot
functional assignment in	
the PDB or other database	Probability 57.02
in HHPred with a	
probability of 90% or	
greater with appropriate	
coverage?	
Is this gene located	<i>If the answer is YES, evaluate the proposed function in the gene</i>
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
shows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
found on the SEA-	Once you have arrived at a functional decision, check the <u>SEA-</u>
PHAGES approved	<u>PHAGES Official Function List</u> to ensure that you are following
function list?	the guidelines for function naming. Functions that are not
	present on the approved list must be carefully vetted for
	approval.
	If you believe this gene should be assigned, please write the
	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	46
Stop Coordinate	22001
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	21321
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes Genemark, bp 21321
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves – Arrovo and Albedo both bad this gene
other annotated genomes?	Tes – Altoyo and Albedo bolit had this gene
Does the gene violate any	The gene follows all guiding principles
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer and GeneMark suggest?	Glimmer Start Coordinate (type NA if not supported):: GeneMark Start Coordinate (type NA if not supported):: 21321
Does the start site have an associated Ribosome Binding Site with a high score?	21321 – moderate z-score (2.793) and appropriate spacer (10 bp)
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	Yes It is the longest ORF
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, It (21321) was found in 59 of 64 genes in the pham, and called 92.2% of the time
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes, it is highly conserved.
DECISION:	Start site is bp 21321, based off of Genemark, Genemark coding potential, phamerator, starterator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	Yes, hypothetical
a protein having a	
functional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁻⁴ or	
smaller with appropriate	
coverage?	
Does this protein align with	; DUF6378 ; Domain of unknown function (DUF6378)
a protein having a	
functional assignment in	Probability 99.88
the PDB or other database	
in HHPred with a	
probability of 90% or	
greater with appropriate	
coverage?	
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
shows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
found on the SEA-	Once you have arrived at a functional decision, check the <u>SEA-</u>
PHAGES approved	<u>PHAGES Official Function List</u> to ensure that you are following
function list?	the guidelines for function naming. Functions that are not
	present on the approved list must be carefully vetted for
	approval.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	48
Stop Coordinate	22015
Direction (For/Rev)	Rev
Gap (Overlap) with Previous Gene	Yes
Selected Start Coordinate	22500
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	Original Climmer call @bn 2250 bas strength 11 13 not called
auto-annotation program	by Cenemark
(Glimmer, GeneMark)?	by Genemark
Is there evidence for	Genemark did not show coding potential
coding potential?	
Is this gene present in	Yes-Loviatar_Draft had this gene
other annotated genomes?	
Does the gene violate any	The gene does not follow all guiding principles
major guiding principles?	
DECISION:	No

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and Genemark suggest?	Genemark Start Coordinate (type NA If not supported)::
Does the start site have an associated Ribosome	List the final RBS score and Z-score of the currently predicted
Binding Site with a high	in your response if this is the best score or not.
score?	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is the predicted start codon	Indicate the length of the ORF is with the predicted start and the
does the longest ORF	Does the proposed start site have a gap/overlap with the
result in excessive gene overlap (>30bp)?	nearest upstream gene that does not violate the Guiding Principles?
	Notes if you are considering more then 4 start site, more ide the
	same information for each proposed start site.
Is this start site conserved	
as indicated by Starterator?	You will also need to provide the following information from Starterator: does the start match the consensus start site predicted from Starterator? If no, is the consensus start site not found in this ORF? If no, is there a better option for the consensus start site instead of the one predicted by Starterator? If Starterator doesn't reveal a consensus start site, you can record that Starterator was not informative.
	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
Is this start site conserved in other phage genomes as indicated by BlastP?	Provide the best BlastP match from NCBI, PhagesDB, and DNA Master with alignment in the format of (Q#:S#), where Q (query) is the sequence you are analyzing and S (subject) is the database match. List the e-value and alignment of the best match for all three BlastP sources.
	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
DECISION:	Record where you think the gene should start here and briefly explain your rationale.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
snows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
DUACES approved	Once you have arrived at a functional decision, check the <u>SEA-</u>
function list2	the guidelines for function List to ensure that you are following
	the guidelines for function naming. Functions that are not
	approval
	approval.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	47
Stop Coordinate	22144
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	3 bp overlap
Selected Start Coordinate	21998
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes GeneMark, bp 21998
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves - Annalie and Arrovo both had this gene
other annotated genomes?	res – Annalie and Arroyo both had this gene
Does the gene violate any	The gene follows all guiding principles
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer and GeneMark suggest?	Glimmer Start Coordinate (type NA if not supported):: GeneMark Start Coordinate (type NA if not supported):: 21998
Does the start site have an associated Ribosome Binding Site with a high score?	21998 – moderate z-score (2.690) and appropriate spacer (10 bp)
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	Yes It is the longest ORF
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, It (21998) was found in 46 of 46 genes in the pham, and called 100% of the time
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes, it is highly conserved.
DECISION:	Start site is bp 21998, based off of Genemark, Genemark coding potential, phamerator, starterator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
shows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
found on the SEA-	Once you have arrived at a functional decision, check the <u>SEA-</u>
PHAGES approved	<u>PHAGES Official Function List</u> to ensure that you are following
function list?	the guidelines for function naming. Functions that are not
	present on the approved list must be carefully vetted for
	approval.
	If you believe this gene should be assigned, please write the
DECISION:	name of the function here. If the evidence does not support a
	TUNCTIONAL CALL, RECORD TINK F" FOR NO KNOWN TUNCTION. 50-70% Of
	pnage genes tall into the INKIF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	49
Stop Coordinate	22635
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	22144
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an auto-annotation program (Glimmer, GeneMark)?	Yes Genemark, bp 22144 but glimmer said bp 22462 with a strength of 1.40
Is there evidence for coding potential?	GeneMarkS showed coding potential
Is this gene present in other annotated genomes?	Gene is not present in other annotated genomes
Does the gene violate any major guiding principles?	The gene follows all guiding principles
DECISION:	Yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported):: 22462
and GeneMark suggest?	GeneMark Start Coordinate (type NA if not supported):: 22144
Does the start site have an associated Ribosome Binding Site with a high score?	22144 – low z-score (1.260) and appropriate spacer (11 bp)
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	Yes It is the longest ORF
Is this start site conserved in other phage genomes as indicated by Starterator?	No, the start site is not conserved in other phage genomes.
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes it is highly conserved
DECISION:	Start site is bp 22144, based off of Genemark, Genemark coding potential, phamerator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
shows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
Tound on the SEA-	Once you have arrived at a functional decision, check the <u>SEA-</u>
PHAGES approved	PHAGES Official Function List to ensure that you are following
	the guidelines for function harming. Functions that are not
	present on the approved list must be carefully velled for
	approval.
	If you believe this gene should be assigned, please write the
DEGICIONI	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	51
Stop Coordinate	22638
Direction (For/Rev)	Rev
Gap (Overlap) with Previous Gene	Gap
Selected Start Coordinate	23432
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes, called by Glimmer bp 23432
(Glimmer, GeneMark)?	
Is there evidence for	NO coding potential
coding potential?	
Is this gene present in	Ves - Loviatar, Draft had this game
other annotated genomes?	Tes - Lovialar_Drait had this gene
Does the gene violate any	The gene does not follow all guiding principles
major guiding principles?	
DECISION:	NO deleted

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and GeneMark suggest?	GeneMark Start Coordinate (type NA if not supported)::
Does the start site have an	List the final RBS score and Z-score of the currently predicted
associated Ribosome	start site using the Kiblerb/Karlin Medium scoring table. Indicate
Binding Sile with a high	In your response if this is the best score of not.
300101	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is the predicted start codon	Indicate the length of the ORF is with the predicted start and the
the longest ORF? If not,	gap/overlap to the nearest stop codon of the upstream ORF.
does the longest ORF	Does the proposed start site have a gap/overlap with the
overlap (>30bp)?	Principles?
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is this start site conserved	
In other phage genomes	Very will also pood to provide the following information from
Starterator?	Starterator: does the start match the consensus start site
	predicted from Starterator? If no, is the consensus start site not
	found in this ORF? If no, is there a better option for the
	consensus start site instead of the one predicted by Starterator?
	If Starterator doesn't reveal a consensus start site, you can
	record that Starterator was not informative.
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is this start site conserved	Provide the best BlastP match from NCBI, PhagesDB, and DNA
in other phage genomes	Master with alignment in the format of (Q#:S#), where Q (query)
as indicated by BlastP?	is the sequence you are analyzing and S (subject) is the
	database match. List the e-value and alignment of the best match for all three PlastP sources
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
DECISION	Record where you think the gene should start here and briefly
	explain your rationale.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
shows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
Tound on the SEA-	Once you have arrived at a functional decision, check the <u>SEA-</u>
PHAGES approved	PHAGES Official Function List to ensure that you are following
function list?	the guidelines for function harming. Functions that are not
	present on the approved list must be carefully velled for
	αρριοναι.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	52
Stop Coordinate	23417
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	Yes
Selected Start Coordinate	23638
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	Original Glimmer call @bp 23638 has strength 6 52 ** not called
auto-annotation program	by GeneMark
(Glimmer, GeneMark)?	by Cenewark
Is there evidence for	No coding potential
coding potential?	
Is this gene present in	Ves only found in Loviator Draft
other annotated genomes?	res – only lound in Lovialar_Drait
Does the gene violate any	The gene violates major guiding principles such as overlapping
major guiding principles?	other, stronger genes
DECISION:	Not a gene, deleted

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and Genemark suggest?	Genemark Start Coordinate (type NA If not supported)::
Does the start site have an associated Ribosome	List the final RBS score and Z-score of the currently predicted
Binding Site with a high	in your response if this is the best score or not.
score?	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is the predicted start codon	Indicate the length of the ORF is with the predicted start and the
does the longest ORF	Does the proposed start site have a gap/overlap with the
result in excessive gene overlap (>30bp)?	nearest upstream gene that does not violate the Guiding Principles?
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is this start site conserved in other phage genomes	
as indicated by Starterator?	You will also need to provide the following information from Starterator: does the start match the consensus start site predicted from Starterator? If no, is the consensus start site not found in this ORF? If no, is there a better option for the consensus start site instead of the one predicted by Starterator? If Starterator doesn't reveal a consensus start site, you can record that Starterator was not informative.
	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
Is this start site conserved in other phage genomes	Provide the best BlastP match from NCBI, PhagesDB, and DNA Master with alignment in the format of (Q#:S#), where Q (query)
as indicated by BlastP?	is the sequence you are analyzing and S (subject) is the
	match for all three BlastP sources.
	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
DECISION:	Record where you think the gene should start here and briefly explain your rationale.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
shows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
found on the SEA-	Once you have arrived at a functional decision, check the <u>SEA-</u>
PHAGES approved	PHAGES Official Function List to ensure that you are following
function list?	the guidelines for function naming. Functions that are not
	present on the approved list must be carefully velled for
	approval.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	50
Stop Coordinate	23558
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	3 bp overlap
Selected Start Coordinate	22632
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes GeneMark, bp 22632
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves - Annalie and Arrovo both had this gene
other annotated genomes?	res – Annalle and Anoyo both had this gene
Does the gene violate any	The gene follows all guiding principles
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer and GeneMark suggest?	Glimmer Start Coordinate (type NA if not supported):: GeneMark Start Coordinate (type NA if not supported):: 22632
Does the start site have an associated Ribosome Binding Site with a high score?	22632 – low z-score (1.968) and appropriate spacer (9 bp)
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	Yes It is the longest ORF
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, It (22632) was found in 90 of 243 genes in the pham, and called 37.0% of the time
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes, it is highly conserved.
DECISION:	Start site is bp 22632, based off of Genemark, Genemark coding potential, phamerator, starterator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
shows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
found on the SEA-	Once you have arrived at a functional decision, check the <u>SEA-</u>
PHAGES approved	<u>PHAGES Official Function List</u> to ensure that you are following
function list?	the guidelines for function naming. Functions that are not
	present on the approved list must be carefully vetted for
	approval.
	If you believe this gene should be assigned, please write the
DECISION:	functional call report "NKE" for no known function 50,70% of
	nunctional call, record INNE for no known function. 50-70% of
	phage genes fail into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	54
Stop Coordinate	23632
Direction (For/Rev)	Rev
Gap (Overlap) with Previous Gene	Yes
Selected Start Coordinate	24141
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an auto-annotation program	Original Glimmer call @bp 24141 has strength 7.67 not called
(Glimmer, GeneMark)?	by genemark
Is there evidence for coding potential?	Genemark did not show coding potential
Is this gene present in other annotated genomes?	Yes-Loviatar_Draft had this gene
Does the gene violate any major guiding principles?	The gene does not follow all guiding principles
DECISION:	NO

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and Genemark suggest?	Genemark Start Coordinate (type NA If not supported)::
Does the start site have an associated Ribosome	List the final RBS score and Z-score of the currently predicted
Binding Site with a high	in your response if this is the best score or not.
score?	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is the predicted start codon	Indicate the length of the ORF is with the predicted start and the
does the longest ORF	Does the proposed start site have a gap/overlap with the
result in excessive gene overlap (>30bp)?	nearest upstream gene that does not violate the Guiding Principles?
	Notes if you are considering more then 4 start site, more ide the
	same information for each proposed start site.
Is this start site conserved	
as indicated by Starterator?	You will also need to provide the following information from Starterator: does the start match the consensus start site predicted from Starterator? If no, is the consensus start site not found in this ORF? If no, is there a better option for the consensus start site instead of the one predicted by Starterator? If Starterator doesn't reveal a consensus start site, you can record that Starterator was not informative.
	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
Is this start site conserved in other phage genomes as indicated by BlastP?	Provide the best BlastP match from NCBI, PhagesDB, and DNA Master with alignment in the format of (Q#:S#), where Q (query) is the sequence you are analyzing and S (subject) is the database match. List the e-value and alignment of the best match for all three BlastP sources.
	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
DECISION:	Record where you think the gene should start here and briefly explain your rationale.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
snows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
DUACES approved	Once you have arrived at a functional decision, check the <u>SEA-</u>
function list2	the guidelines for function List to ensure that you are following
	the guidelines for function naming. Functions that are not
	approval
	approval.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	53
Stop Coordinate	23893
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	23588
Selected Function	Hypothetical Protein

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes Genemark, bp 23588
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves – BabyVoda and SanaSana both had this gene
other annotated genomes?	res – Daby roua and SanaSana bour nad uns gene
Does the gene violate any	The gene follows all guiding principles
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer and GeneMark suggest?	Glimmer Start Coordinate (type NA if not supported):: GeneMark Start Coordinate (type NA if not supported):: 23588
Does the start site have an associated Ribosome Binding Site with a high score?	23588 – low z-score (1.433) and appropriate spacer (9 bp)
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	Not the longest ORF and does not result in excessive gene overlap
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, It (23588) was found in 6 of 6 genes in the pham, and called 100% of the time
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes, it is highly conserved.
DECISION:	Start site is bp 23588, based off of Genemark, Genemark coding potential, phamerator, starterator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	Yes, hypothetical protein
a protein having a	
functional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁻⁴ or	
smaller with appropriate	
coverage?	
Does this protein align with	DNA packaging protein Gp17; nucleotide-binding fold,
a protein having a	HYDROLASE; HET: ADP; 1.8A {Enterobacteria phage T4}
functional assignment in	99.35%
the PDB or other database	
in HHPred with a	
probability of 90% or	
greater with appropriate	
coverage?	
Is this gene located	No
adjacent to genes of	
known function and in a	
region of the genome that	
shows high conservation of	
gene order?	
Is this gene a possible	No
transmembrane protein?	
Is the proposed function	NO
found on the SEA-	
PHAGES approved	
function list?	
DECISION:	Hypothetical Protein
Basic Phage Information	
----------------------------------	----------------------
Phage Name	Akino08
Gene #	55
Stop Coordinate	24090
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	23899
Selected Function	Hypothetical Protein

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes GeneMark, bp 23899
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Yes – WalkingDead_Draft and Laviator_Draft both had this
other annotated genomes?	gene
Does the gene violate any	The gene follows all guiding principles
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer and GeneMark suggest?	Glimmer Start Coordinate (type NA if not supported):: GeneMark Start Coordinate (type NA if not supported):: 23899
Does the start site have an associated Ribosome Binding Site with a high score?	23899 – moderate z-score (2.207) and low spacer (5 bp)
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	Yes It is the longest ORF
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, It (23899) was found in 5 of 9 genes in the pham, and called 55.6% of the time
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes, it is highly conserved.
DECISION:	Start site is bp 23899, based off of Genemark, Genemark coding potential, phamerator, starterator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	Yes, hypothetical protein
a protein having a	
functional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁻⁴ or	
smaller with appropriate	
coverage?	
Does this protein align with	; HTH_Tnp_Tc3_2_like ; Transposable element Tc3
a protein having a	transposase, HTH
functional assignment in	84.35%
the PDB or other database	
In HHPred with a	
probability of 90% or	
greater with appropriate	
coverage?	
Is this gene located	NO
adjacent to genes of	
known function and in a	
region of the genome that	
snows nigh conservation of	
gene order?	No
Is this gene a possible	NO
liansmembrane protein?	No
found on the SEA	NO
DUACES approved	
function list?	
	Hunothetical Protein
DECISION:	

Basic Phage Information	
Phage Name	Akino08
Gene #	57
Stop Coordinate	24442
Direction (For/Rev)	Rev
Gap (Overlap) with Previous Gene	Gap
Selected Start Coordinate	25224
Selected Function	

Explain Your Rationale
Yes, called by Glimmer bp 25224
Genemark did not show coding potential
Ves_Loviatar. Draft had this gene
The gene does not follow all guiding principles
NO deleted

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and GeneMark suggest?	GeneMark Start Coordinate (type NA if not supported)::
Does the start site have an	List the final RBS score and Z-score of the currently predicted
associated Ribosome	start site using the Kibler6/Karlin Medium scoring table. Indicate
Binding Site with a high	in your response if this is the best score or not.
score?	
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is the predicted start codon	Indicate the length of the ORF is with the predicted start and the
the longest ORF? If not,	gap/overlap to the nearest stop codon of the upstream ORF.
does the longest ORF	Does the proposed start site have a gap/overlap with the
result in excessive gene	nearest upstream gene that does not violate the Guiding
overlap (>30bp)?	Principles?
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is this start site conserved	
in other phage genomes	
as indicated by	You will also need to provide the following information from
Starterator?	Starterator: does the start match the consensus start site
	predicted from Starterator? If no, is the consensus start site not
	found in this ORF? If no, is there a better option for the
	consensus start site instead of the one predicted by Starterator?
	If Starterator doesn't reveal a consensus start site, you can
	record that Starterator was not informative.
	Note: If you are considering more than 1 start site, provide the
	same information for each proposed start site.
le this start site conserved	Provide the best PlastP metab from NCPL PhagesDP and DNA
in other phage genemoe	Moster with alignment in the formet of (Ott:St), where O (quary)
in other phage genomes	is the sequence you are enclyzing and S (cubicet) is the
as indicated by blastr?	Is the sequence you are analyzing and S (Subject) is the
	meteb for all three PlactD sources
	Noto: if you are considering more than 1 start site, provide the
	some information for each proposed start site
	Record where you think the gene should start here and briefly
DECISION:	explain your rationale

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
snows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
is the proposed function	Indicate a response with a Yes or No response.
DUACES approved	Once you have arrived at a functional decision, check the <u>SEA-</u>
function list2	the guidelines for function List to ensure that you are following
	present on the approved list must be carefully vetted for
	approval
	approval.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	56
Stop Coordinate	24466
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	7 bp overlap
Selected Start Coordinate	24083
Selected Function	Hypothetical Protein

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes Genemark, bp 24083
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves – Albedo and SanaSana both had this gene
other annotated genomes?	Tes - Albedo and Sanasana both had this gene
Does the gene violate any	The gene follows all guiding principles
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer and GeneMark suggest?	Glimmer Start Coordinate (type NA if not supported):: GeneMark Start Coordinate (type NA if not supported):: 24083
Does the start site have an associated Ribosome Binding Site with a high score?	24083 – low z-score (1.944) and appropriate spacer (11 bp)
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	Yes has the longest ORF
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, It (24083) was found in 10 of 79 genes in the pham, and called 12.7% of the time
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes, it is highly conserved.
DECISION:	Start site is bp 24083, based off of Genemark, Genemark coding potential, phamerator, starterator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	Yes, hypothetical protein
a protein having a	
functional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁻⁴ or	
smaller with appropriate	
coverage?	
Does this protein align with	c.52.1.35 (A:) stNUC {Salmonella phage SETP3 [TaxId:
a protein having a	424944]] CLASS: Alpha and beta proteins (a/b), FOLD:
functional assignment in	Restriction
the PDB or other database	84.94%
in HHPred with a	
probability of 90% or	
greater with appropriate	
coverage?	
Is this gene located	NO
adjacent to genes of	
Known function and in a	
region of the genome that	
shows high conservation of	
gene order?	
Is this gene a possible	NO
transmembrane protein?	No
is the proposed function	NO
DUACES approved	
function list?	
	Hunothatical Protain
DECISION:	nypolitelical Fiolelli

Basic Phage Information	
Phage Name	Akino08
Gene #	58
Stop Coordinate	24624
Direction (For/Rev)	Rev
Gap (Overlap) with Previous Gene	Νο
Selected Start Coordinate	24821
Selected Function	toxin in toxin/antitoxin system, HicA-like

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes GeneMark, bp 24821
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves only in Louister Draft
other annotated genomes?	res – only in Lovialar_Drait
Does the gene violate any	The gene follows all guiding principles
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer and GeneMark suggest?	Glimmer Start Coordinate (type NA if not supported):: GeneMark Start Coordinate (type NA if not supported):: 24821
Does the start site have an associated Ribosome Binding Site with a high score?	24821 – moderate z-score (2.440) and low spacer (6 bp)
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	Yes It is the longest ORF
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, It (24821) was found in 32 of 197 genes in the pham, and called 16.2% of the time
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes, it is highly conserved.
DECISION:	Start site is bp 24821, based off of Genemark, Genemark coding potential, phamerator, starterator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	Yes, HicA-like toxin
a protein having a	
functional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁻⁴ or	
smaller with appropriate	
coverage?	
Does this protein align with	; HicA_toxin ; HicA toxin of bacterial toxin-antitoxin,
a protein having a	99.46%
functional assignment in	
the PDB or other database	
in HHPred with a	
probability of 90% or	
greater with appropriate	
coverage?	
Is this gene located	Yes
adjacent to genes of	
known function and in a	
region of the genome that	
shows high conservation of	
gene order?	
Is this gene a possible	No
transmembrane protein?	
Is the proposed function	YES
found on the SEA-	
PHAGES approved	
function list?	
DECISION:	toxin in toxin/antitoxin system, HicA-like

Basic Phage Information	
Phage Name	Akino08
Gene #	60
Stop Coordinate	25339
Direction (For/Rev)	Rev
Gap (Overlap) with Previous Gene	Yes
Selected Start Coordinate	25482
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an auto-annotation program	Original Glimmer call @bp 25482 has strength 8.05 not called
(Glimmer, GeneMark)?	by Cenemark
Is there evidence for coding potential?	Genemark did not show coding potential
Is this gene present in other annotated genomes?	Yes-Loviatar_Draft had this gene
Does the gene violate any major guiding principles?	The gene does not follow all guiding principles
DECISION:	No

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and Genemark suggest?	Genemark Start Coordinate (type NA If not supported)::
Does the start site have an associated Ribosome	List the final RBS score and Z-score of the currently predicted
Binding Site with a high	in your response if this is the best score or not.
score?	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is the predicted start codon	Indicate the length of the ORF is with the predicted start and the
does the longest ORF	Does the proposed start site have a gap/overlap with the
result in excessive gene overlap (>30bp)?	nearest upstream gene that does not violate the Guiding Principles?
	Notes if you are considering more then 4 start site, more ide the
	same information for each proposed start site.
Is this start site conserved	
as indicated by Starterator?	You will also need to provide the following information from Starterator: does the start match the consensus start site predicted from Starterator? If no, is the consensus start site not found in this ORF? If no, is there a better option for the consensus start site instead of the one predicted by Starterator? If Starterator doesn't reveal a consensus start site, you can record that Starterator was not informative.
	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
Is this start site conserved in other phage genomes as indicated by BlastP?	Provide the best BlastP match from NCBI, PhagesDB, and DNA Master with alignment in the format of (Q#:S#), where Q (query) is the sequence you are analyzing and S (subject) is the database match. List the e-value and alignment of the best match for all three BlastP sources.
	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
DECISION:	Record where you think the gene should start here and briefly explain your rationale.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
snows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
DUACES approved	Once you have arrived at a functional decision, check the <u>SEA-</u>
function list2	the guidelines for function List to ensure that you are following
	the guidelines for function naming. Functions that are not
	approval
	approval.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	61
Stop Coordinate	25440
Direction (For/Rev)	Rev
Gap (Overlap) with Previous Gene	Yes
Selected Start Coordinate	26033
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes, called by Glimmer bp 26033
(Glimmer, GeneMark)?	
Is there evidence for	Genemark did not show coding potential
coding potential?	
Is this gene present in	Ves - Loviatar, Draft had this gene
other annotated genomes?	Tes - Lovialar_Drait had this gene
Does the gene violate any	The gene does not follow all guiding principles
major guiding principles?	
DECISION:	NO deleted

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and GeneMark suggest?	GeneMark Start Coordinate (type NA if not supported)::
Does the start site have an	List the final RBS score and Z-score of the currently predicted
associated Ribosome	start site using the Kibler6/Karlin Medium scoring table. Indicate
Binding Site with a high	in your response if this is the best score or not.
score?	
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is the predicted start codon	Indicate the length of the ORF is with the predicted start and the
the longest ORF? If not,	gap/overlap to the nearest stop codon of the upstream ORF.
does the longest ORF	Does the proposed start site have a gap/overlap with the
result in excessive gene	nearest upstream gene that does not violate the Guiding
overlap (>30bp)?	Principles?
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is this start site conserved	
in other phage genomes	
as indicated by	You will also need to provide the following information from
Starterator?	Starterator: does the start match the consensus start site
	predicted from Starterator? If no, is the consensus start site not
	found in this ORF? If no, is there a better option for the
	consensus start site instead of the one predicted by Starterator?
	If Starterator doesn't reveal a consensus start site, you can
	record that Starterator was not informative.
	Note: If you are considering more than 1 start site, provide the
	same information for each proposed start site.
le this start site conserved	Provide the best PlastP metab from NCPL PhagesDP and DNA
in other phage genemoe	Moster with alignment in the formet of (Ott:St), where O (quary)
in other phage genomes	is the sequence you are enclyzing and S (cubicet) is the
as indicated by blastr?	Is the sequence you are analyzing and S (Subject) is the
	meteb for all three PlactD sources
	Noto: if you are considering more than 1 start site, provide the
	some information for each proposed start site
	Record where you think the gene should start here and briefly
DECISION:	explain your rationale

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
snows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
is the proposed function	Indicate a response with a Yes or No response.
DUACES approved	Once you have arrived at a functional decision, check the <u>SEA-</u>
function list2	the guidelines for function List to ensure that you are following
	present on the approved list must be carefully vetted for
	approval
	approval.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	59
Stop Coordinate	25521
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	24925
Selected Function	dUTPase

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes GeneMark, bp 24925
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves – Albright and Annalia both had this gana
other annotated genomes?	res – Albright and Annalle both had this gene
Does the gene violate any	The gene follows all guiding principles
major guiding principles?	
DECISION:	Yes

Catharing Evidence	Evaloia Vour Dotionalo
Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and GeneMark suggest?	GeneMark Start Coordinate (type NA if not supported):: 24925
Does the start site have an	24925 – moderate z-score (2.744) and low spacer (7 bp)
associated Ribosome	
Binding Site with a high	
score?	
Is the predicted start codon	No it is not the longest ORE but it is not overlapping with the
the longest ORE2 If not	nevious gene
doos the longest OPE	previous gene.
result in excessive gene	
overlap (>30bp)?	
Is this start site conserved	Yes, It (24925) was found in 77 of 85 genes in the pham, and
in other phage genomes	called 90.6% of the time
as indicated by	
Starterator?	
Is this start site conserved	Yes, it is highly conserved.
in other phage genomes	
as indicated by BlastP?	
DECISION	Start site is bp 24925, based off of Genemark, Genemark
DECISION:	coding potential, phamerator, starterator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	Yes, deoxyuridine triphosphatase
a protein having a	
functional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁻⁴ or	
smaller with appropriate	
coverage?	
Does this protein align with	dUTPase from DI S. aureus phage; Staphylococcus aureus,
a protein having a	pathogenicity island, SaPI, dUTPases, signalling, gene transfer,
functional assignment in	99.46%
the PDB or other database	
in HHPred with a	
probability of 90% or	
greater with appropriate	
coverage?	
Is this gene located	Yes
adjacent to genes of	
known function and in a	
region of the genome that	
shows high conservation of	
gene order?	
Is this gene a possible	No
transmembrane protein?	
Is the proposed function	YES
tound on the SEA-	
PHAGES approved	
function list?	
DECISION:	dUTPase

Basic Phage Information	
Phage Name	Akino08
Gene #	63
Stop Coordinate	26021
Direction (For/Rev)	Rev
Gap (Overlap) with Previous Gene	Yes
Selected Start Coordinate	26803
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes, called by Glimmer bp 26803
(Glimmer, GeneMark)?	
Is there evidence for	Genemark did not show coding potential
coding potential?	
Is this gene present in	Ves - Loviatar, Draft had this gene
other annotated genomes?	Tes - Lovialar_Drait had this gene
Does the gene violate any	The gene does not follow all guiding principles
major guiding principles?	
DECISION:	NO deleted

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and GeneMark suggest?	GeneMark Start Coordinate (type NA if not supported)::
Does the start site have an	List the final RBS score and Z-score of the currently predicted
associated Ribosome	start site using the Kiblerb/Karlin Medium scoring table. Indicate
Binding Sile with a high	In your response if this is the best score of not.
300101	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is the predicted start codon	Indicate the length of the ORF is with the predicted start and the
the longest ORF? If not,	gap/overlap to the nearest stop codon of the upstream ORF.
does the longest ORF	Does the proposed start site have a gap/overlap with the
overlap (>30bp)?	Principles?
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is this start site conserved	
In other phage genomes	Very will also pood to provide the following information from
Starterator?	Starterator: does the start match the consensus start site
	predicted from Starterator? If no, is the consensus start site not
	found in this ORF? If no, is there a better option for the
	consensus start site instead of the one predicted by Starterator?
	If Starterator doesn't reveal a consensus start site, you can
	record that Starterator was not informative.
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is this start site conserved	Provide the best BlastP match from NCBI, PhagesDB, and DNA
in other phage genomes	Master with alignment in the format of (Q#:S#), where Q (query)
as indicated by BlastP?	is the sequence you are analyzing and S (subject) is the
	database match. List the e-value and alignment of the best match for all three PlastP sources
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
DECISION	Record where you think the gene should start here and briefly
	explain your rationale.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
shows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
Tound on the SEA-	Once you have arrived at a functional decision, check the <u>SEA-</u>
PHAGES approved	PHAGES Official Function List to ensure that you are following
	the guidelines for function harming. Functions that are not
	present on the approved list must be carefully velled for
	approval.
	If you believe this gene should be assigned, please write the
DEGICIONI	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	62
Stop Coordinate	26024
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	25518
Selected Function	thymidylate kinase

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes Genemark, bp 25518
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves – Albedo and Albright both bad this gene
other annotated genomes?	Tes – Albedo and Albright both had this gene
Does the gene violate any	The gene follows all guiding principles
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and GeneMark suggest?	GeneMark Start Coordinate (type NA if not supported):: 25518
Does the start site have an associated Ribosome Binding Site with a high score?	25518 – moderate z-score (2.532) and large spacer (12 bp)
Is the predicted start codon	Not the longest ORF and does not result in excessive gene
the longest ORF? If not,	overlap
does the longest ORF	
result in excessive gene	
overlap (>300p)?	
in other phage genemos	
in other phage genomes	Vas It (22588) was found in 77 of 82 conos in the nham and
Starterator?	called 03 0% of the time
Is this start site conserved	Yes, it is highly conserved.
in other phage genomes	
as indicated by BlastP?	
DECISION:	Start site is bp 25518, based off of Genemark, Genemark coding potential, phamerator, starterator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	Yes, thymidylate kinase
a protein having a	
functional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁻⁴ or	
smaller with appropriate	
coverage?	
Does this protein align with	Thymidylate kinase; Nucleotide monophosphate kinase,
a protein having a	TRANSFERASE; HET: TMP; 1.19A {Thermus thermophilus
functional assignment in	(strain HB8 / ATCC 27634 / DSM 579)} SCOP: c.37.1.0
the PDB or other database	99.78%
in HHPred with a	
probability of 90% or	
greater with appropriate	
coverage?	
Is this gene located	Yes
adjacent to genes of	
known function and in a	
region of the genome that	
shows high conservation of	
gene order?	
Is this gene a possible	No
transmembrane protein?	
Is the proposed function	Yes
tound on the SEA-	
PHAGES approved	
function list?	
DECISION:	thymidylate kinase

Basic Phage Information	
Phage Name	Akino08
Gene #	64
Stop Coordinate	26788
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	26039
Selected Function	Recombination Directionality factor

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes GeneMark, bp 26039
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves learion and Sanasana both had this gene
other annotated genomes?	res – icanon and Sanasana both nad this gene
Does the gene violate any	The gene follows all guiding principles
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Your Pationalo
What start site do Glimmer and GeneMark suggest?	Glimmer Start Coordinate (type NA if not supported):: GeneMark Start Coordinate (type NA if not supported):: 26039
Does the start site have an associated Ribosome Binding Site with a high score?	26039 – moderate z-score (2.132) and high spacer (15 bp)
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	Yes it is the longest ORF
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, It (26039) was found in 81 of 200 genes in the pham, and called 40.5% of the time
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes, it is highly conserved.
DECISION:	Start site is bp 26039, based off of Genemark, Genemark coding potential, phamerator, starterator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	Yes, recombination directionality factor
a protein having a	
functional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁻⁴ or	
smaller with appropriate	
coverage?	
Does this protein align with	; Gp3-like ; Recombination directionality factor-like
a protein having a	99.93%
functional assignment in	
the PDB or other database	
in HHPred with a	
probability of 90% or	
greater with appropriate	
coverage?	
Is this gene located	Yes
adjacent to genes of	
known function and in a	
region of the genome that	
shows high conservation of	
gene order?	
Is this gene a possible	NO
transmembrane protein?	
is the proposed function	Yes
Tound on the SEA-	
PHAGES approved	
DECISION:	Recomplination Directionality factor

Basic Phage Information	
Phage Name	Akino08
Gene #	65
Stop Coordinate	26948
Direction (For/Rev)	Rev
Gap (Overlap) with Previous Gene	Yes
Selected Start Coordinate	27604
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an auto-annotation program	Original Glimmer call @bp 27604 has a strength 5.64 not called
(Glimmer, GeneMark)?	by genemark
Is there evidence for coding potential?	Genemark did not show coding potential
Is this gene present in other annotated genomes?	Yes –Loviatar_Draft had this gene
Does the gene violate any major guiding principles?	The gene does not follow all guiding principles
DECISION:	NO

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and Genemark suggest?	Genemark Start Coordinate (type NA If not supported)::
Does the start site have an	List the final RBS score and Z-score of the currently predicted
Binding Site with a high	in your response if this is the best score or not.
score?	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is the predicted start codon	Indicate the length of the ORF is with the predicted start and the
does the longest ORF	Does the proposed start site have a gap/overlap with the
result in excessive gene overlap (>30bp)?	nearest upstream gene that does not violate the Guiding Principles?
	Notes if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is this start site conserved	
as indicated by Starterator?	You will also need to provide the following information from Starterator: does the start match the consensus start site predicted from Starterator? If no, is the consensus start site not found in this ORF? If no, is there a better option for the consensus start site instead of the one predicted by Starterator? If Starterator doesn't reveal a consensus start site, you can record that Starterator was not informative.
	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
Is this start site conserved in other phage genomes as indicated by BlastP?	Provide the best BlastP match from NCBI, PhagesDB, and DNA Master with alignment in the format of (Q#:S#), where Q (query) is the sequence you are analyzing and S (subject) is the database match. List the e-value and alignment of the best
	match for all three BlastP sources.
	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
DECISION:	Record where you think the gene should start here and briefly explain your rationale.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
snows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
DUACES approved	Once you have arrived at a functional decision, check the <u>SEA-</u>
function list2	the guidelines for function List to ensure that you are following
	the guidelines for function naming. Functions that are not
	approval
	approval.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	66
Stop Coordinate	27043
Direction (For/Rev)	Rev
Gap (Overlap) with Previous Gene	Νο
Selected Start Coordinate	27342
Selected Function	Hypothetical Protein

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes GeneMark, bp 27342
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves - found in Loviatar, Draft only
other annotated genomes?	res – lound in Lovialar_Drait only
Does the gene violate any	The gene follows all guiding principles
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer and GeneMark suggest?	Glimmer Start Coordinate (type NA if not supported):: GeneMark Start Coordinate (type NA if not supported)::27342
Does the start site have an associated Ribosome Binding Site with a high score?	27342 – moderate z-score (2.062) and high spacer (13 bp)
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	Yes It is the longest ORF
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, It (27342) was found in 49 of 49 genes in the pham, and called 100% of the time
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes, it is highly conserved.
DECISION:	Start site is bp 27342, based off of Genemark, Genemark coding potential, phamerator, starterator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	Hypothetical Protein
a protein having a	
functional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁻⁴ or	
smaller with appropriate	
coverage?	
Does this protein align with	No quality match ; RNR_inhib ; Ribonucleotide reductase
a protein having a	inhibitor 76.87%
functional assignment in	
the PDB or other database	
IN HHPred with a	
probability of 90% or	
greater with appropriate	
coverage?	
Is this gene located	NO
adjacent to genes of	
Known function and in a	
region of the genome that	
snows nigh conservation of	
gene order?	No
transmombrane protein?	NO
	Vee
found on the SEA	100
DHACES approved	
function list?	
DECISION:	Hypothetical Protein
Basic Phage Information	
----------------------------------	------------------------
Phage Name	Akino08
Gene #	67
Stop Coordinate	27649
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	27416
Selected Function	NrdH-like glutaredoxin

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes Genemark, bp 27416
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves – BabyVoda and SanaSana both had this gene
other annotated genomes?	res – Daby roua and SanaSana bour nad unis gene
Does the gene violate any	The gene follows all guiding principles
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Vour Pationalo
what start site do Glimmer	Glimmer Start Coordinate (type INA If not supported)::
and GeneMark suggest?	GeneMark Start Coordinate (type NA if not supported):: 27416
Does the start site have an	27416 – moderate z-score (2.140) and high spacer (16 bp)
associated Ribosome	
Binding Site with a high	
score?	
300101	
is the predicted start codon	Not the longest ORF and does not result in excessive gene
the longest ORF? If not,	overlap
does the longest ORF	
result in excessive gene	
overlap (>30bp)?	
Is this start site conserved	
in other phage genomes	
as indicated by	Ves. It (23588) was found in 81 of 421 genes in the nham and
Stortorotor?	$res, re(20000)$ was round in or of $\pm 2r$ genes in the phani, and colled 10.2% of the time
Starterator ?	
is this start site conserved	Yes, it is highly conserved.
in other phage genomes	
as indicated by BlastP?	
DECISION	Start site is bp 27416, based off of Genemark, Genemark
DECISION:	coding potential, phamerator, starterator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	Yes, NrdH-like glutaredoxin
a protein having a	
functional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁻⁴ or	
smaller with appropriate	
coverage?	
Does this protein align with	c.47.1.0 (A:) automated matches {Baker's yeast
a protein having a	(Saccharomyces cerevisiae) [Taxld: 559292]} CLASS:
functional assignment in	Alpha and beta proteins (a/b), FOLD: Thioredoxin fold,
the PDB or other database	SUPFAM: Thioredoxin-like, FAM: automated matches
in HHPred with a	99.34%
probability of 90% or	
greater with appropriate	
coverage?	
Is this gene located	Yes
adjacent to genes of	
known function and in a	
region of the genome that	
shows high conservation of	
gene order?	
Is this gene a possible	No
transmembrane protein?	
Is the proposed function	Yes
found on the SEA-	
PHAGES approved	
function list?	
DECISION:	NrdH-like glutaredoxin

Basic Phage Information	
Phage Name	Akino08
Gene #	68
Stop Coordinate	27627
Direction (For/Rev)	Rev
Gap (Overlap) with Previous Gene	Yes
Selected Start Coordinate	28298
Selected Function	

Explain Your Rationale
Yes, called by Glimmer bp 28298
Genemark did not show coding potential
Ves_Loviatar. Draft had this game
Tes - Lovialar_Drait had this gene
The gene does not follow all guiding principles
NO deleted

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and GeneMark suggest?	GeneMark Start Coordinate (type NA if not supported)::
Does the start site have an	List the final RBS score and Z-score of the currently predicted
associated Ribosome	start site using the Kibler6/Karlin Medium scoring table. Indicate
Binding Site with a high	in your response if this is the best score or not.
score?	
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is the predicted start codon	Indicate the length of the ORF is with the predicted start and the
the longest ORF? If not,	gap/overlap to the nearest stop codon of the upstream ORF.
does the longest ORF	Does the proposed start site have a gap/overlap with the
result in excessive gene	nearest upstream gene that does not violate the Guiding
overlap (>30bp)?	Principles?
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is this start site conserved	
in other phage genomes	
as indicated by	You will also need to provide the following information from
Starterator?	Starterator: does the start match the consensus start site
	predicted from Starterator? If no, is the consensus start site not
	found in this ORF? If no, is there a better option for the
	consensus start site instead of the one predicted by Starterator?
	If Starterator doesn't reveal a consensus start site, you can
	record that Starterator was not informative.
	Note: If you are considering more than 1 start site, provide the
	same information for each proposed start site.
le this start site conserved	Provide the best PlastP metab from NCPL PhagesDP and DNA
in other phage genemoe	Moster with alignment in the formet of (Ott:St), where O (quary)
in other phage genomes	is the sequence you are enclyzing and S (cubicet) is the
as indicated by blastr?	Is the sequence you are analyzing and S (Subject) is the
	meteb for all three PlactD sources
	Noto: if you are considering more than 1 start site, provide the
	some information for each proposed start site
	Record where you think the gene should start here and briefly
DECISION:	explain your rationale

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
snows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
is the proposed function	Indicate a response with a Yes or No response.
DUACES approved	Once you have arrived at a functional decision, check the <u>SEA-</u>
function list2	the guidelines for function List to ensure that you are following
	present on the approved list must be carefully vetted for
	approval
	approval.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	69
Stop Coordinate	27975
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	27649
Selected Function	Hypothetical

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes GeneMark, bp 27649
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves – Icarian and Sanasana both had this gene
other annotated genomes?	res – leanan and Sanasana both nad this gene
Does the gene violate any	The gene follows all guiding principles
major guiding principles?	
DECISION:	Yes

Cathorizan Erridon co	Eveloie Voue Dotionala
Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and GeneMark suggest?	GeneMark Start Coordinate (type NA if not supported):: 27649
Does the start site have an	27649 – moderate z-score (2.167) and high spacer (16 bp)
associated Ribosome	
Binding Site with a high	
score?	
Is the predicted start codon	No, it is not the longest ORF but it is not overlapping with the
the longest ORF? If not,	previous gene.
does the longest ORF	
result in excessive gene	
overlap (>30bp)?	
Is this start site conserved	
in other phage genomes	Yes, It (27649) was found in 75 of 353 genes in the pham, and
as indicated by	called 21.2% of the time
Starterator?	
Is this start site conserved	Yes, it is highly conserved.
in other phage genomes	
as indicated by BlastP?	
DECISION	Start site is bp 27649, based off of Genemark, Genemark
DECISICIA.	coding potential, phamerator, starterator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	BlastP/ PhagesDB: Hypothetical protein/ unknown function
a protein having a	
functional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁻⁴ or	
smaller with appropriate	
coverage?	
Does this protein align with	This protein has a probability of 1.93 % in the HHPred with an
a protein having a	alignment of a Hypothetical protein/ unknown function
functional assignment in	
the PDB or other database	
in HHPred with a	
probability of 90% or	
greater with appropriate	
coverage?	
Is this gene located	No Synteny Observed.
adjacent to genes of	
known function and in a	
region of the genome that	
shows high conservation of	
gene order?	
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	yes
found on the SEA-	
PHAGES approved	
function list?	
DECISION:	NKF

Basic Phage Information	
Phage Name	Akino08
Gene #	70
Stop Coordinate	28313
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	None
Selected Start Coordinate	28005
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes GeneMark, bp 28005
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves — Albright and Albedo both bad this gene
other annotated genomes?	res – Albright and Albedo both had this gene
Does the gene violate any	The gene follows all guiding principles
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Your Pationalo
What start site do Climmer	Climmer Start Coordinate (type NA if not supported)::
what Start Sile do Gillinner	Ginniner Start Coordinate (type NA II not supported)
and Genemark suggest?	Genemark Start Coordinate (type INA If not supported).: 28005
Does the start site have an	28005 – low z-score (1.797) and high spacer (14 bp)
associated Ribosome	
Binding Site with a high	
score?	
Is the predicted start codon	Yes It is the longest ORF.
the longest ORF? If not,	
does the longest ORF	
result in excessive gene	
overlap (>30bp)?	
Is this start site conserved	
in other phage genomes	
as indicated by	Yes It (28005) was found in 61 of 61 genes in the pham and
Starterator?	called 100% of the time
Is this start site conserved	Yes it is highly conserved
in other phage denomes	
as indicated by BlastP?	
as indicated by Diastr?	Start aita ia ha 28005 haaad off of Canamark, Canamark
DECISION:	Start Site is up 20005, Dased On Or Generilark, Generilark
	coung potential, phamerator, starterator, and plast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
shows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
found on the SEA-	Once you have arrived at a functional decision, check the <u>SEA-</u>
PHAGES approved	<u>PHAGES Official Function List</u> to ensure that you are following
function list?	the guidelines for function naming. Functions that are not
	present on the approved list must be carefully vetted for
	approval.
	If you believe this some abould be assisted alasses write the
	n you believe this gene should be assigned, please write the
DECISION:	functional call record "NKE" for no known function 50 70% of
	nhage genes fall into the NKE category
	ן אומצע שנוונט ומו ווונט נווט זאוער טמנכעטוץ.

Basic Phage Information	
Phage Name	Akino08
Gene #	71
Stop Coordinate	28337
Direction (For/Rev)	Rev
Gap (Overlap) with Previous Gene	Yes
Selected Start Coordinate	28495
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	Original Climmer call @bn 28405 bas strength 7 24 not called
auto-annotation program	by Conemark
(Glimmer, GeneMark)?	by Genemark
Is there evidence for	GeneMarkS did not show coding potential
coding potential?	
Is this gene present in	Vas Louistar Draft had this gone
other annotated genomes?	res -Lovialar_Drait nau triis gene
Does the gene violate any	The gene does not follow all guiding principles
major guiding principles?	
DECISION:	No

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and Genemark suggest?	Genemark Start Coordinate (type NA If not supported)::
Does the start site have an associated Ribosome	List the final RBS score and Z-score of the currently predicted
Binding Site with a high	in your response if this is the best score or not.
score?	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is the predicted start codon	Indicate the length of the ORF is with the predicted start and the
does the longest ORF	Does the proposed start site have a gap/overlap with the
result in excessive gene overlap (>30bp)?	nearest upstream gene that does not violate the Guiding Principles?
	Notes if you are considering more then 4 start site, more ide the
	same information for each proposed start site.
Is this start site conserved	
as indicated by Starterator?	You will also need to provide the following information from Starterator: does the start match the consensus start site predicted from Starterator? If no, is the consensus start site not found in this ORF? If no, is there a better option for the consensus start site instead of the one predicted by Starterator? If Starterator doesn't reveal a consensus start site, you can record that Starterator was not informative.
	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
Is this start site conserved in other phage genomes as indicated by BlastP?	Provide the best BlastP match from NCBI, PhagesDB, and DNA Master with alignment in the format of (Q#:S#), where Q (query) is the sequence you are analyzing and S (subject) is the database match. List the e-value and alignment of the best match for all three BlastP sources.
	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
DECISION:	Record where you think the gene should start here and briefly explain your rationale.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
shows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
found on the SEA-	Once you have arrived at a functional decision, check the <u>SEA-</u>
PHAGES approved	PHAGES Official Function List to ensure that you are following
function list?	the guidelines for function naming. Functions that are not
	present on the approved list must be carefully velled for
	approval.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	72
Stop Coordinate	28717
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	28394
Selected Function	Hypothetical Protein

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes Genemark, bp 28394
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves - Loviatar. Draft had this gene
other annotated genomes?	Tes - Eoviatal_Drait had this gene
Does the gene violate any	The gene follows all guiding principles
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer and GeneMark suggest?	<i>Glimmer Start Coordinate (type NA if not supported)::</i> <i>GeneMark Start Coordinate (type NA if not supported)::</i> 28394
Does the start site have an associated Ribosome Binding Site with a high score?	28394 – low z-score (1.889) and appropriate spacer (10 bp)
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	Yes it is the longest ORF
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, It (28394) was found in 2 of 2 genes in the pham, and called 100% of the time
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes, it is highly conserved.
DECISION:	Start site is bp 28394, based off of Genemark, Genemark coding potential, phamerator, starterator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	Yes, Hypothetical Protein
a protein having a	
functional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁻⁴ or	
smaller with appropriate	
coverage?	
Does this protein align with	Hypothetical Protein
a protein having a	40.37%
functional assignment in	
the PDB or other database	
in HHPred with a	
probability of 90% or	
greater with appropriate	
coverage?	
Is this gene located	No
adjacent to genes of	
known function and in a	
region of the genome that	
shows high conservation of	
gene order?	
Is this gene a possible	No
transmembrane protein?	
Is the proposed function	Yes
found on the SEA-	
PHAGES approved	
DECISION:	Hypotnetical Protein

Basic Phage Information	
Phage Name	Akino08
Gene #	73
Stop Coordinate	28926
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	None
Selected Start Coordinate	28717
Selected Function	Hypothetical Protein

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes GeneMark, bp 28717
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves — Albright and Albedo both bad this gene
other annotated genomes?	res – Albright and Albedo both had this gene
Does the gene violate any	The gene follows all guiding principles
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer and GeneMark suggest?	Glimmer Start Coordinate (type NA if not supported):: GeneMark Start Coordinate (type NA if not supported):: 28717
Does the start site have an associated Ribosome Binding Site with a high score?	28717 – low z-score (1.926) and appropriate spacer (10 bp)
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	Not the longest ORF. The longest ORF does have an excessive overlap.
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, It (28717) was found in 43 of 63 genes in the pham, and called 68.3% of the time
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes, it is highly conserved.
DECISION:	Start site is bp 28717, based off of Genemark, Genemark coding potential, phamerator, starterator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	Yes, Hypothetical Protein
a protein having a	
functional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁻⁴ or	
smaller with appropriate	
coverage?	
Does this protein align with	Hypothetical Protein
a protein having a	63.7%
functional assignment in	
the PDB or other database	
in HHPred with a	
probability of 90% or	
greater with appropriate	
coverage?	
Is this gene located	No
adjacent to genes of	
known function and in a	
region of the genome that	
shows high conservation of	
gene order?	
Is this gene a possible	INO
transmembrane protein?	
is the proposed function	Yes
Tound on the SEA-	
PHAGES approved	
	11 moderation Durate in
DECISION:	Hypothetical Protein

Basic Phage Information	
Phage Name	Akino08
Gene #	74
Stop Coordinate	29007
Direction (For/Rev)	Rev
Gap (Overlap) with Previous Gene	Yes
Selected Start Coordinate	29168
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes, called by Glimmer bp 29168
(Glimmer, GeneMark)?	
Is there evidence for	Genemark did not show coding potential
coding potential?	
Is this gene present in	Ves_Loviatar. Draft had this gene
other annotated genomes?	res – Eoviatar_Drait had this gene
Does the gene violate any	The gene does not follow all guiding principles
major guiding principles?	
DECISION:	NO deleted

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and GeneMark suggest?	GeneMark Start Coordinate (type NA if not supported)::
Does the start site have an	List the final RBS score and Z-score of the currently predicted
associated Ribosome	start site using the Kiblerb/Karlin Medium scoring table. Indicate
Binding Sile with a high	In your response if this is the best score of not.
300101	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is the predicted start codon	Indicate the length of the ORF is with the predicted start and the
the longest ORF? If not,	gap/overlap to the nearest stop codon of the upstream ORF.
does the longest ORF	Does the proposed start site have a gap/overlap with the
overlap (>30bp)?	Principles?
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is this start site conserved	
In other phage genomes	Very will also pood to provide the following information from
Starterator?	Starterator: does the start match the consensus start site
	predicted from Starterator? If no, is the consensus start site not
	found in this ORF? If no, is there a better option for the
	consensus start site instead of the one predicted by Starterator?
	If Starterator doesn't reveal a consensus start site, you can
	record that Starterator was not informative.
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is this start site conserved	Provide the best BlastP match from NCBI, PhagesDB, and DNA
in other phage genomes	Master with alignment in the format of (Q#:S#), where Q (query)
as indicated by BlastP?	is the sequence you are analyzing and S (subject) is the
	database match. List the e-value and alignment of the best match for all three PlastP sources
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
DECISION	Record where you think the gene should start here and briefly
	explain your rationale.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
snows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
DUACES approved	Once you have arrived at a functional decision, check the <u>SEA-</u>
function list2	the guidelines for function List to ensure that you are following
	the guidelines for function naming. Functions that are not
	approval
	approval.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	76
Stop Coordinate	29202
Direction (For/Rev)	Rev
Gap (Overlap) with Previous Gene	YES
Selected Start Coordinate	29330
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	Original Glimmer call @bp 29330 has strength 15 13 not called
auto-annotation program	by Genemark
(Glimmer, GeneMark)?	by cononiant
Is there evidence for	GeneMarkS did not show coding potential
coding potential?	
Is this gene present in	Ves_Loviatar Draft had this gene
other annotated genomes?	Tes -Loviatal_Drait flad this gene
Does the gene violate any	The gene does not follow all guiding principles
major guiding principles?	
DECISION:	No

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and Genemark suggest?	Genemark Start Coordinate (type NA If not supported)::
Does the start site have an	List the final RBS score and Z-score of the currently predicted
Binding Site with a high	in your response if this is the best score or not.
score?	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is the predicted start codon	Indicate the length of the ORF is with the predicted start and the
does the longest ORF	Does the proposed start site have a gap/overlap with the
result in excessive gene overlap (>30bp)?	nearest upstream gene that does not violate the Guiding Principles?
	Notes if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is this start site conserved	
as indicated by Starterator?	You will also need to provide the following information from Starterator: does the start match the consensus start site predicted from Starterator? If no, is the consensus start site not found in this ORF? If no, is there a better option for the consensus start site instead of the one predicted by Starterator? If Starterator doesn't reveal a consensus start site, you can record that Starterator was not informative.
	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
Is this start site conserved in other phage genomes as indicated by BlastP?	Provide the best BlastP match from NCBI, PhagesDB, and DNA Master with alignment in the format of (Q#:S#), where Q (query) is the sequence you are analyzing and S (subject) is the database match. List the e-value and alignment of the best
	match for all three BlastP sources.
	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
DECISION:	Record where you think the gene should start here and briefly explain your rationale.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
snows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
DUACES approved	Once you have arrived at a functional decision, check the <u>SEA-</u>
function list2	the guidelines for function List to ensure that you are following
	the guidelines for function naming. Functions that are not
	approval
	approval.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	77
Stop Coordinate	29772
Direction (For/Rev)	Rev
Gap (Overlap) with Previous Gene	Yes
Selected Start Coordinate	31034
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by	
an auto-annotation	Yes, called by Glimmer bp 31034
program (Glimmer,	
GeneMark)?	
Is there evidence for	Genemark did not show coding potential
coding potential?	
Is this gene present in	
other annotated	Yes – Loviatar_Draft had this gene
genomes?	
Does the gene violate any	The gene does not follow all guiding principles
major guiding principles?	
DECISION:	NO deleted

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and GeneMark suggest?	GeneMark Start Coordinate (type NA if not supported)::
Does the start site have an	List the final RBS score and Z-score of the currently predicted
associated Ribosome	start site using the Kiblerb/Karlin Medium scoring table. Indicate
Binding Sile with a high	In your response if this is the best score of not.
300101	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is the predicted start codon	Indicate the length of the ORF is with the predicted start and the
the longest ORF? If not,	gap/overlap to the nearest stop codon of the upstream ORF.
does the longest ORF	Does the proposed start site have a gap/overlap with the
overlap (>30bp)?	Principles?
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is this start site conserved	
In other phage genomes	Very will also pood to provide the following information from
Starterator?	Starterator: does the start match the consensus start site
	predicted from Starterator? If no, is the consensus start site not
	found in this ORF? If no, is there a better option for the
	consensus start site instead of the one predicted by Starterator?
	If Starterator doesn't reveal a consensus start site, you can
	record that Starterator was not informative.
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is this start site conserved	Provide the best BlastP match from NCBI, PhagesDB, and DNA
in other phage genomes	Master with alignment in the format of (Q#:S#), where Q (query)
as indicated by BlastP?	is the sequence you are analyzing and S (subject) is the
	database match. List the e-value and alignment of the best match for all three PlastP sources
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
DECISION	Record where you think the gene should start here and briefly
	explain your rationale.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
snows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
DUACES approved	Once you have arrived at a functional decision, check the <u>SEA-</u>
function list2	the guidelines for function List to ensure that you are following
	the guidelines for function naming. Functions that are not
	approval
	approval.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	78
Stop Coordinate	31355
Direction (For/Rev)	Rev
Gap (Overlap) with Previous Gene	Yes
Selected Start Coordinate	33319
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	Original Climmer call @bn 33319 has strength 9 98 ** not called
auto-annotation program	by GeneMark
(Glimmer, GeneMark)?	by Cenewark
Is there evidence for	No coding potential
coding potential?	
Is this gene present in	Ves found only in Loviator. Draft
other annotated genomes?	Tes, lound only in Lovialar_Drait
Does the gene violate any	The gene violates major guiding principles such as overlapping
major guiding principles?	other, stronger genes
DECISION:	Not a gene, deleted

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and Genemark suggest?	Genemark Start Coordinate (type NA If not supported)::
Does the start site have an associated Ribosome	List the final RBS score and Z-score of the currently predicted
Binding Site with a high	in your response if this is the best score or not.
score?	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is the predicted start codon	Indicate the length of the ORF is with the predicted start and the
does the longest ORF	Does the proposed start site have a gap/overlap with the
result in excessive gene overlap (>30bp)?	nearest upstream gene that does not violate the Guiding Principles?
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is this start site conserved in other phage genomes	
as indicated by Starterator?	You will also need to provide the following information from Starterator: does the start match the consensus start site predicted from Starterator? If no, is the consensus start site not found in this ORF? If no, is there a better option for the consensus start site instead of the one predicted by Starterator? If Starterator doesn't reveal a consensus start site, you can record that Starterator was not informative.
	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
Is this start site conserved in other phage genomes	Provide the best BlastP match from NCBI, PhagesDB, and DNA Master with alignment in the format of (Q#:S#), where Q (query)
as indicated by BlastP?	is the sequence you are analyzing and S (subject) is the
	match for all three BlastP sources.
	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
DECISION:	Record where you think the gene should start here and briefly explain your rationale.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
shows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
found on the SEA-	Once you have arrived at a functional decision, check the <u>SEA-</u>
PHAGES approved	PHAGES Official Function List to ensure that you are following
function list?	the guidelines for function naming. Functions that are not
	present on the approved list must be carefully velled for
	approval.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	75
Stop Coordinate	31517
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	29043
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes Genemark, bp 29043
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves — Albedo and Albright both bad this gene
other annotated genomes?	res – Albedo and Albright both had this gene
Does the gene violate any	The gene follows all guiding principles
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and GeneMark suggest?	GeneMark Start Coordinate (type NA if not supported):: 29043
Does the start site have an	29043 – moderate z-score (2.782) and appropriate spacer (7
associated Ribosome	bp)
Binding Site with a high	
score?	
Is the predicted start codon	Yes the longest ORF
the longest ORF? If not,	
does the longest ORF	
result in excessive gene	
overlap (>30bp)?	
Is this start site conserved	
in other phage genomes	
as indicated by	Yes, It (29043) was found in 81 of 198 genes in the pham, and
Starterator?	called 40.9% of the time
Is this start site conserved	Yes, it is highly conserved.
in other phage genomes	
as indicated by BlastP?	
DECISION	Start site is bp 29043, based off of Genemark, Genemark
DECISION.	coding potential, phamerator, starterator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
shows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
found on the SEA-	Once you have arrived at a functional decision, check the <u>SEA-</u>
PHAGES approved	<u>PHAGES Official Function List</u> to ensure that you are following
function list?	the guidelines for function naming. Functions that are not
	present on the approved list must be carefully vetted for
	approval.
	If you believe this gene should be assigned, please write the
DECISION:	name of the function here. If the evidence does not support a
	TUNCTIONAL CALL, RECORD TINK F" FOR NO KNOWN TUNCTION. 50-70% Of
	pnage genes tall into the INKIF category.
Basic Phage Information	
----------------------------------	---------
Phage Name	Akino08
Gene #	79
Stop Coordinate	33388
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	33272
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes GeneMark, bp 33272
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves - Loviatar. Draft had this gene
other annotated genomes?	Tes - Eoviatal_Drait had this gene
Does the gene violate any	The gene follows all guiding principles
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer and GeneMark suggest?	Glimmer Start Coordinate (type NA if not supported):: GeneMark Start Coordinate (type NA if not supported)::33272
Does the start site have an associated Ribosome Binding Site with a high score?	27649 – moderate z-score (2.395) and appropriate spacer (12 bp)
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	Yes it is the longest ORF.
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, It (33272) was found in 2 of 2 genes in the pham, and called 100% of the time
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes, it is highly conserved.
DECISION:	Start site is bp 33272, based off of Genemark, Genemark coding potential, phamerator, starterator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
shows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
Tound on the SEA-	Once you have arrived at a functional decision, check the <u>SEA-</u>
PHAGES approved	PHAGES Official Function List to ensure that you are following
	the guidelines for function harming. Functions that are not
	present on the approved list must be carefully velled for
	approval.
	If you believe this gene should be assigned, please write the
DEGICIONI	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	80
Stop Coordinate	33585
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	33388
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes Genemark, bp 33388
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves — Albedo and Albright both bad this gene
other annotated genomes?	res – Albedo and Albright both had this gene
Does the gene violate any	The gene follows all guiding principles
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and GeneMark suggest?	GeneMark Start Coordinate (type NA if not supported):: 33388
Does the start site have an associated Ribosome Binding Site with a high score?	33388 – moderate z-score (2.479) and high spacer (15 bp)
Is the predicted start codon the longest ORF? If not, does the longest ORF	Not the longest ORF and does not result in excessive gene overlap
overlap (>30bp)?	
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, It (33388) was found in 78 of 82 genes in the pham, and called 95.1% of the time
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes, it is highly conserved.
DECISION:	Start site is bp 33388, based off of Genemark, Genemark coding potential, phamerator, starterator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
shows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
Tound on the SEA-	Once you have arrived at a functional decision, check the <u>SEA-</u>
PHAGES approved	PHAGES Official Function List to ensure that you are following
function list?	the guidelines for function harming. Functions that are not
	present on the approved list must be carefully velled for
	αρριοναι.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	82
Stop Coordinate	33623
Direction (For/Rev)	Rev
Gap (Overlap) with Previous Gene	Yes
Selected Start Coordinate	34675
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	Original Glimmer call @bp 34675 has strength 7.00 not called
auto-annotation program	by genemark
(Gilmmer, Genemark)?	
Is there evidence for	GeneMarkS did not show coding potential
coding potential?	
Is this gene present in	Ves_Loviatar Draft had this gene
other annotated genomes?	
Does the gene violate any	The gene does not follow all guiding principles
major guiding principles?	
DECISION:	No

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and Genemark suggest?	Genemark Start Coordinate (type NA If not supported)::
Does the start site have an associated Ribosome	List the final RBS score and Z-score of the currently predicted
Binding Site with a high	in your response if this is the best score or not.
score?	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is the predicted start codon	Indicate the length of the ORF is with the predicted start and the
does the longest ORF	Does the proposed start site have a gap/overlap with the
result in excessive gene overlap (>30bp)?	nearest upstream gene that does not violate the Guiding Principles?
	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is this start site conserved in other phage genomes	
as indicated by Starterator?	You will also need to provide the following information from Starterator: does the start match the consensus start site predicted from Starterator? If no, is the consensus start site not found in this ORF? If no, is there a better option for the consensus start site instead of the one predicted by Starterator? If Starterator doesn't reveal a consensus start site, you can record that Starterator was not informative.
	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
Is this start site conserved in other phage genomes as indicated by BlastP?	Provide the best BlastP match from NCBI, PhagesDB, and DNA Master with alignment in the format of (Q#:S#), where Q (query) is the sequence you are analyzing and S (subject) is the database match. List the e-value and alignment of the best match for all three BlastP sources.
	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
DECISION:	Record where you think the gene should start here and briefly explain your rationale.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
snows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
is the proposed function	Indicate a response with a Yes or No response.
DUACES approved	Once you have arrived at a functional decision, check the <u>SEA-</u>
function list2	the guidelines for function List to ensure that you are following
	the guidelines for function naming. Functions that are not
	approval
	approval.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	81
Stop Coordinate	33884
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	33582
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes GeneMark, bp 33582
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves — Albright and Albedo both bad this gene
other annotated genomes?	res – Albright and Albedo both had this gene
Does the gene violate any	The gene follows all guiding principles
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer and GeneMark suggest?	Glimmer Start Coordinate (type NA if not supported):: GeneMark Start Coordinate (type NA if not supported)::33582
Does the start site have an associated Ribosome Binding Site with a high score?	33582 – high z-score (3.029) and appropriate spacer (12 bp)
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	No, it is not the longest ORF. It excessively overlaps with the last gene.
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, It (33582) was found in 101 of 181 genes in the pham, and called 55.8% of the time
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes, it is highly conserved.
DECISION:	Start site is bp 33582, based off of Genemark, Genemark coding potential, phamerator, starterator, and blast data.

Gathering Evidence	Explain Your Pationale
Doos this protoin align with	PlactP/ Phagap DP unknown function
Does this protein anyth with	Diasir/ FliagesDB ulikilowil luliciloli
a protein naving a	
Tunctional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁻⁴ or	
smaller with appropriate	
coverage?	
Does this protein align with	This protein has a probability of 1.93 % in the HHPred with an
a protein having a	alignment of a Hypothetical protein/ unknown function
functional assignment in	
the PDB or other database	
in HHPred with a	
probability of 90% or	
greater with appropriate	
coverage?	
Is this gene located	No Synteny Observed.
adjacent to genes of	
known function and in a	
region of the genome that	
shows high conservation of	
gene order?	
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
found on the SEA-	Once you have arrived at a functional decision, check the <u>SEA-</u>
PHAGES approved	PHAGES Official Function List to ensure that you are following
function list?	the guidelines for function naming. Functions that are not
	present on the approved list must be carefully vetted for
	approval.
DECISION:	NKF

Basic Phage Information	
Phage Name	Akino08
Gene #	83
Stop Coordinate	34603
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	33884
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes Genemark, bp 33884
(Glimmer, Genewark)?	
Is there evidence for coding potential?	GeneMarkS showed coding potential
Is this gene present in other annotated genomes?	Yes – Albedo and Albright both had this gene
Does the gene violate any maior guiding principles?	The gene follows all guiding principles
DECISION:	Yes

Cathoring Evidence	Explain Your Pationalo
What start site do Climmer	Climmer Start Coordinate (type NA if not supported)::
what Start Sile do Gillinner	Ginniner Start Coordinate (type NA II not supported)
and Genemark suggest?	Genemark Start Coordinate (type NA II not supported) 33664
Does the start site have an	33884 – low z-score (1.581) and appropriate spacer (10 bp)
associated Ribosome	
Binding Site with a high	
score?	
Is the predicted start codon	Not the longest ORF and does not result in excessive gene
the longest ORF? If not,	overlap
does the longest ORF	
result in excessive gene	
overlap (>30bp)?	
Is this start site conserved	
in other phage genomes	
as indicated by	Yes It (33884) was found in 19 of 189 genes in the pham and
Starterator?	called 10 1% of the time
Is this start site conserved	Yes it is highly conserved
in other phage genomes	
as indicated by BlastP?	
	Start site is bn 22881 based off of Conomark Conomark
DECISION:	start site is up 55004, based on or Generilark, Generilark

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
shows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
found on the SEA-	Once you have arrived at a functional decision, check the <u>SEA-</u>
PHAGES approved	PHAGES Official Function List to ensure that you are following
function list?	the guidelines for function naming. Functions that are not
	present on the approved list must be carefully velled for
	approval.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	85
Stop Coordinate	34854
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	34648
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes Genemark, bp 34648
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS did not show coding potential
coding potential?	
Is this gene present in	Ves_Loviatar. Draft had this gene
other annotated genomes?	
Does the gene violate any	The gene does not follow all guiding principles
major guiding principles?	
DECISION:	No

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and Genemark suggest?	Genemark Start Coordinate (type NA If not supported)::
Does the start site have an	List the final RBS score and Z-score of the currently predicted
Binding Site with a high	in your response if this is the best score or not.
score?	Note: if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is the predicted start codon	Indicate the length of the ORF is with the predicted start and the
does the longest ORF	Does the proposed start site have a gap/overlap with the
result in excessive gene overlap (>30bp)?	nearest upstream gene that does not violate the Guiding Principles?
	Notes if you are considering more than 1 start site, provide the
	same information for each proposed start site.
Is this start site conserved	
as indicated by Starterator?	You will also need to provide the following information from Starterator: does the start match the consensus start site predicted from Starterator? If no, is the consensus start site not found in this ORF? If no, is there a better option for the consensus start site instead of the one predicted by Starterator? If Starterator doesn't reveal a consensus start site, you can record that Starterator was not informative.
	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
Is this start site conserved in other phage genomes as indicated by BlastP?	Provide the best BlastP match from NCBI, PhagesDB, and DNA Master with alignment in the format of (Q#:S#), where Q (query) is the sequence you are analyzing and S (subject) is the database match. List the e-value and alignment of the best match for all three BlastP sources
	Note: if you are considering more than 1 start site, provide the same information for each proposed start site.
DECISION:	Record where you think the gene should start here and briefly explain your rationale.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
snows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
is the proposed function	Indicate a response with a Yes or No response.
DUACES approved	Once you have arrived at a functional decision, check the <u>SEA-</u>
function list2	the guidelines for function List to ensure that you are following
	present on the approved list must be carefully vetted for
	approval
	approval.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	88
Stop Coordinate	35280
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	35116
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes GeneMark, bp 35116
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves – Loviatar, Draft both had this gene
other annotated genomes?	res – Eovialar_Drait both had this gene
Does the gene violate any	The gene follows all guiding principles
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer	Glimmer Start Coordinate (type NA if not supported)::
and GeneMark suggest?	GeneMark Start Coordinate (type NA if not supported):: 35116
Does the start site have an associated Ribosome Binding Site with a high score?	35116 – low z-score (1.797) and high spacer (15 bp)
Is the predicted start codon	Not the langest OPE, the langest OPE dags have evenesive
does the longest ORF	overlap.
result in excessive gene	
overlap (>30bp)?	
Is this start site conserved in other phage genomes	
as indicated by	Yes, It (35116) was found in 52 of 52 genes in the pham, and
Statterator	
Is this start site conserved	Yes, it is highly conserved, only one blast available.
in other phage genomes as indicated by BlastP?	
	Start site is bn 35116 based off of Genemark Genemark
DECISION:	coding potential, phamerator, starterator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
snows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
DUACES approved	Once you have arrived at a functional decision, check the <u>SEA-</u>
function list2	the guidelines for function List to ensure that you are following
	the guidelines for function naming. Functions that are not
	approval
	approval.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	87
Stop Coordinate	35702
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	
Selected Start Coordinate	35280
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	Yes, Glimmer called 34863 and GeneMark called calls start at
(Glimmer, GeneMark)?	35820
Is there evidence for	GeneMarkS did show coding potential
coding potential?	
Is this gene present in	Yes –Loviatar Draft had this gene
other annotated genomes?	
Does the gene violate any	Gene does not violate any major guiding principles
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer and GeneMark suggest?	<i>Glimmer Start Coordinate (type NA if not supported):: 34863</i> <i>GeneMark Start Coordinate (type NA if not supported):: 35280</i>
Does the start site have an associated Ribosome Binding Site with a high score?	35280 – moderate z-score (2.779) and appropriate spacer (9 bp)
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	Not the longest ORF and does not result in excessive gene overlap
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, It (34863) was found in 2 of 2 genes in the pham, and called 100% of the time
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes, it is highly conserved.
DECISION:	The reason we went with 35280 as the start position instead of 34863, is because the coding potential is greatest at the start of 35280. Starterator says 34683, however according to other genes it would make sense for the gene to start at 35280. For this reason, we have concluded that moving gene 87 (34683) to start at (35280)

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
shows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
found on the SEA-	Once you have arrived at a functional decision, check the <u>SEA-</u>
PHAGES approved	<u>PHAGES Official Function List</u> to ensure that you are following
function list?	the guidelines for function naming. Functions that are not
	present on the approved list must be carefully vetted for
	approval.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	91
Stop Coordinate	36629
Direction (For/Rev)	for
Gap (Overlap) with Previous Gene	overlap
Selected Start Coordinate	36189
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes, Genemark, bp 36189
(Glimmer, GeneMark)?	
Is there evidence for	GeneMarkS showed coding potential
coding potential?	
Is this gene present in	Ves – DirtyBubble and Icarian
other annotated genomes?	
Does the gene violate any	The gene follows all guiding principles
major guiding principles?	
DECISION:	Yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer and GeneMark suggest?	Glimmer Start Coordinate (type NA if not supported):: n/a GeneMark Start Coordinate (type NA if not supported):: 36189
Does the start site have an associated Ribosome Binding Site with a high score?	36189– highest z-score (3.273) and appropriate spacer (13bp)
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	Yes, it the length of 441. There is overlapping with other genes.
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, It (36189) was found in 50/338 genes in the pham, and called 100% of the time
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes, it is highly conserved.
DECISION:	Start site is bp 36189, based off of Genemark, Genemark coding potential, phamerator, starterator, and blast data.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
shows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
found on the SEA-	Once you have arrived at a functional decision, check the <u>SEA-</u>
PHAGES approved	<u>PHAGES Official Function List</u> to ensure that you are following
function list?	the guidelines for function naming. Functions that are not
	present on the approved list must be carefully vetted for
	approval.
	If you believe this gene should be assigned, please write the
DECISION:	name of the function here. If the evidence does not support a
	TUNCTIONAL CALL, RECORD TINK F" FOR NO KNOWN TUNCTION. 50-70% Of
	pnage genes tall into the INKIF category.

37718

Basic Phage Information	
Phage Name	Akino08
Gene #	96
Stop Coordinate	37718
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	Overlap
Selected Start Coordinate	37557
Selected Function	unknown

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes GeneMark call @bp 37557
(Glimmer, GeneMark)?	
Is there evidence for	Yes, GeneMark shows coding potential
coding potential?	
Is this gene present in	Ves in RAiuniner, RehyDeisy, and Rechero
other annotated genomes?	res, in DAjuniper, DabyDaisy, and Dachaco.
Does the gene violate any	No
major guiding principles?	NO
DECISION:	yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer and GeneMark suggest?	Glimmer Start Coordinate (type NA if not supported):: N/A GeneMark Start Coordinate (type NA if not supported):: 37557
Does the start site have an associated Ribosome Binding Site with a high score?	37557 - Large z-score of 2.444 and a appropriate spacer 9bp
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	No, the longest ORF would be start 1 with a length of 330
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, it was 75/82 the pham and called 91.5 % of the time
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes, it is highly conserved
DECISION:	Start site is bp 37718, based off Genemark, phamerator, starterator.

Gathering Evidence	Explain Your Rationale
Does this protein align with	Yes, an unknown function/ hypothetical protein
a protein having a	
functional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁻⁴ or	
smaller with appropriate	
coverage?	
Does this protein align with	This protein has a probability of 82.46% in the HHPred with an
a protein having a	alignment of uknown/hypothetical protein.
functional assignment in	
the PDB or other database	
in HHPred with a	
probability of 90% or	
greater with appropriate	
coverage?	
Is this gene located	No Synteny Observed.
adjacent to genes of	
known function and in a	
region of the genome that	
shows high conservation of	
gene order?	If the ensure is VEQ indicate summarian data from at least Q
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Yes
Iound on the SEA-	
FIAGES approved	
DECISION:	NKF

Basic Phage Information	
Phage Name	Akino08
Gene #	98
Stop Coordinate	38090
Direction (For/Rev)	Rev
Gap (Overlap) with Previous Gene	Overlap
Selected Start Coordinate	38587
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an auto-annotation program (Glimmer, GeneMark)?	Original Glimmer call @bp 38587 has strength 1.92 ** not called by GeneMark
Is there evidence for	GeneMark shows no coding potential
is this gene present in	Yes in Loviatar Draft
other annotated genomes?	
Does the gene violate any	Vaa
major guiding principles?	res
DECISION:	No

Gathering Evidence	Explain Your Rationale
What start site do Glimmer and GeneMark suggest?	Glimmer Start Coordinate (type NA if not supported):: 38587 GeneMark Start Coordinate (type NA if not supported):: N/A
Does the start site have an associated Ribosome Binding Site with a high score?	38587 - Large z-score of 2.281 and a appropriate spacer 7bp
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	No, the longest ORF would be start 1 with a length of 747
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, it was 2/2 the pham and called 100.0 % of the time
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes, it is highly conserved
DECISION:	Start site is bp 38587, based off Genemark, phamerator, starterator.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
snows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
DUACES approved	Once you have arrived at a functional decision, check the <u>SEA-</u>
function list2	the guidelines for function List to ensure that you are following
	the guidelines for function naming. Functions that are not
	approval
	approval.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

38261

Basic Phage Information	
Phage Name	Akino08
Gene #	97
Stop Coordinate	38261
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	Overlap
Selected Start Coordinate	37776
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes GeneMark call @bp 37776
(Glimmer, GeneMark)?	
Is there evidence for	Yes, GeneMark shows coding potential
coding potential?	
Is this gene present in	Ves in Annal ie Arrovo and BahyDaisy
other annotated genomes?	res, in Annalie, Anoyo, and DabyDaisy.
Does the gene violate any	No
major guiding principles?	
DECISION:	

Gathering Evidence	Explain Your Rationale
What start site do Glimmer and GeneMark suggest?	Glimmer Start Coordinate (type NA if not supported):: N/A GeneMark Start Coordinate (type NA if not supported):: 37776
Does the start site have an associated Ribosome Binding Site with a high score?	37776 - Large z-score of 1.700 and a appropriate spacer 13bp
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	Yes, it is the longest ORF with a length of 486
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, it was 75/82 the pham and called 91.5 % of the time
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes it is highly conserved
DECISION:	Start site is bp 37776, based off Genemark, phamerator, starterator.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
shows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
Tound on the SEA-	Once you have arrived at a functional decision, check the <u>SEA-</u>
PHAGES approved	PHAGES Official Function List to ensure that you are following
	the guidelines for function harming. Functions that are not
	present on the approved list must be carefully velled for
	approval.
	If you believe this gene should be assigned, please write the
DECICION	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.
Basic Phage Information	
----------------------------------	------------------
Phage Name	Akino08
Gene #	97
Stop Coordinate	38261
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	Overlap
Selected Start Coordinate	37776
Selected Function	Unknown function

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes GeneMark call @bp 37557
(Glimmer, GeneMark)?	
Is there evidence for	Yes, GeneMark shows coding potential
coding potential?	
Is this gene present in	Ves in Annal ie Arrovo and BahyDaisy
other annotated genomes?	res, in Annalle, Anoyo, and Babybaisy.
Does the gene violate any	No
major guiding principles?	NO
DECISION:	yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer and GeneMark suggest?	Glimmer Start Coordinate (type NA if not supported):: N/A GeneMark Start Coordinate (type NA if not supported):: 37557
Does the start site have an associated Ribosome Binding Site with a high score?	Large z-score of 2.295 and a appropriate spacer 13bp
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	No, the longest ORF would be Start 2
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, it was 2/83 the pham and called 2.4 % of the time
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes, it is highly conserved
DECISION:	Start site is bp 42028, based off Genemark, phamerator, starterator.

Gathering Evidence	Explain Your Rationale
Does this protein align with	Yes, an unknown function/ hypothetical protein
a protein having a	
functional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁻⁴ or	
smaller with appropriate	
coverage?	
Does this protein align with	This protein has a probability of 60.64% in the HHPred with an
a protein having a	alignment of uknown/hypothetical protein.
functional assignment in	
the PDB or other database	
in HHPred with a	
probability of 90% or	
greater with appropriate	
coverage?	
Is this gene located	No Synteny Observed.
adjacent to genes of	
known function and in a	
region of the genome that	
shows high conservation of	
gene order?	
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	yes.
found on the SEA-	
PHAGES approved	
function list?	
DECISION:	NKF

Basic Phage Information	
Phage Name	Akino08
Gene #	99
Stop Coordinate	38584
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	Overlap
Selected Start Coordinate	38324
Selected Function	Hypothetical Protein

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes GeneMark call @bp 38324
(Glimmer, GeneMark)?	
Is there evidence for	Yes, GeneMark shows coding potential
coding potential?	
Is this gene present in	Ves in Armstrong Celsens ChiliPenner
other annotated genomes?	res, in Annstrong, Celaena, Chill epper.
Does the gene violate any	No
major guiding principles?	NO
DECISION:	yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer and GeneMark suggest?	Glimmer Start Coordinate (type NA if not supported):: N/A GeneMark Start Coordinate (type NA if not supported):: 38324
Does the start site have an associated Ribosome Binding Site with a high score?	38324 - Large z-score of 1.832 and a appropriate spacer 10bp
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	Yes, it is the only start.
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, it was 13/39 the pham and called 33.3 % of the time
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes, it is highly conserved
DECISION:	Start site is bp 38324, based off Genemark, phamerator, starterator.

Gathering Evidence	Explain Your Rationale
Does this protein align with	Yes, an unknown function/ hypothetical protein
a protein having a	
functional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁻⁴ or	
smaller with appropriate	
coverage?	
Does this protein align with	This protein has a probability of 90.55% in the HHPred with an
a protein having a	alignment of uknown/hypothetical protein.
functional assignment in	
the PDB or other database	
in HHPred with a	
probability of 90% or	
greater with appropriate	
coverage?	
Is this gene located	No Synteny Observed.
adjacent to genes of	
known function and in a	
region of the genome that	
shows high conservation of	
gene order?	
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	yes
found on the SEA-	
PHAGES approved	
function list?	
DECISION:	NKF

39149

Basic Phage Information	
Phage Name	Akino08
Gene #	100
Stop Coordinate	39149
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	Overlap
Selected Start Coordinate	38574
Selected Function	Dihydrofolate reductase

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes GeneMark call @bp 38574
(Glimmer, GeneMark)?	
Is there evidence for	Yes, GeneMark shows coding potential
coding potential?	
Is this gene present in	Ves in Bernstein, Brahms and BubbaBear
other annotated genomes?	res, in Demstein, Drainns, and Dubbabear.
Does the gene violate any	No
major guiding principles?	NO
DECISION:	yes

Gathering Evidence	Explain Your Rationale
What start site do Glimmer and GeneMark suggest?	Glimmer Start Coordinate (type NA if not supported):: N/A GeneMark Start Coordinate (type NA if not supported):: 38574
Does the start site have an associated Ribosome Binding Site with a high score?	38574 - Large z-score of 2.132 and a appropriate spacer 10p
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	Yes, it is the longest ORF with a length of 576
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, it was 75/75 the pham and called 100.0% of the time
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes, it is highly conserved
DECISION:	Start site is bp 38574, based off Genemark, phamerator, starterator.

Gathering Evidence	Explain Your Rationale
Does this protein align with	Yes, Dihydrofolate reductase
a protein having a	
functional assignment in	
BlastP (phagesDB and/or	
GenBank) with an	
alignment of 10 ⁻⁴ or	
smaller with appropriate	
coverage?	
Does this protein align with	This protein has a probability of 99.91% in the HHPred with an
a protein having a	alignment of Dihydrofolate reductase.
functional assignment in	
the PDB or other database	
in HHPred with a	
probability of 90% or	
greater with appropriate	
coverage?	
Is this gene located	Yes, with Dihydrofolate reductase
adjacent to genes of	
known function and in a	
region of the genome that	
shows high conservation of	
gene order?	
Is this gene a possible	<i>If the answer is YES, indicate supporting data from at least 2</i>
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Yes
found on the SEA-	
PHAGES approved	
function list?	
DECISION:	Yes, this gene should align with Dihydrofolate reductase

Basic Phage Information	
Phage Name	Akino08
Gene #	113
Stop Coordinate	42025
Direction (For/Rev)	Rev
Gap (Overlap) with Previous Gene	Overlap
Selected Start Coordinate	42324
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes Glimmer, bp 42324, no GeneMark call
(Glimmer, GeneMark)?	
Is there evidence for	No, GeneMark shows no coding potential
coding potential?	
Is this gene present in	
other annotated genomes?	
Does the gene violate any	
major guiding principles?	
DECISION:	

Gathering Evidence	Explain Your Rationale
What start site do Glimmer and GeneMark suggest?	Glimmer Start Coordinate (type NA if not supported):: N/A GeneMark Start Coordinate (type NA if not supported):: 42324
Does the start site have an associated Ribosome Binding Site with a high score?	Large z-score of 1.801 and a appropriate spacer 13bp
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	No, the longest ORF would be Start 1
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, it was 2/2 in the pham and called 100% of the time
Is this start site conserved in other phage genomes as indicated by BlastP?	There is no BlastP for this gene
DECISION:	Start site is bp 42324, based off of Glimmer, no Genemark coding potential, phamerator, starterator.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
shows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
found on the SEA-	Once you have arrived at a functional decision, check the <u>SEA-</u>
PHAGES approved	<u>PHAGES Official Function List</u> to ensure that you are following
function list?	the guidelines for function naming. Functions that are not
	present on the approved list must be carefully vetted for
	approval.
	If you believe this gene should be assigned, please write the
DECISION:	name of the function here. If the evidence does not support a
	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	114
Stop Coordinate	42363
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	Overlap
Selected Start Coordinate	42028
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes GeneMark call @ bp 42363
(Glimmer, GeneMark)?	
Is there evidence for	Yes, GeneMark shows coding potential
coding potential?	
Is this gene present in	
other annotated genomes?	
Does the gene violate any	No
major guiding principles?	
DECISION:	

Gathering Evidence	Explain Your Rationale
What start site do Glimmer and GeneMark suggest?	Glimmer Start Coordinate (type NA if not supported):: N/A GeneMark Start Coordinate (type NA if not supported):: 42363
Does the start site have an associated Ribosome Binding Site with a high score?	Large z-score of 2.295 and a appropriate spacer 13bp
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	No, the longest ORF would be Start 2
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, it was 54/224 in the pham and called 9.3% of the time
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes it is highly conserved
DECISION:	Start site is bp 42028, based off Genemark, phamerator, starterator.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
shows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
found on the SEA-	Once you have arrived at a functional decision, check the <u>SEA-</u>
PHAGES approved	PHAGES Official Function List to ensure that you are following
function list?	the guidelines for function naming. Functions that are not
	present on the approved list must be carefully velled for
	approval.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.

Basic Phage Information	
Phage Name	Akino08
Gene #	114
Stop Coordinate	42491
Direction (For/Rev)	For
Gap (Overlap) with Previous Gene	Yes
Selected Start Coordinate	42363
Selected Function	

Gathering Evidence	Explain Your Rationale
Was the gene called by an	
auto-annotation program	Yes GeneMark call @ bp 42363
(Glimmer, GeneMark)?	
Is there evidence for	Yes, GeneMark shows coding potential
coding potential?	
Is this gene present in	
other annotated genomes?	
Does the gene violate any	No
major guiding principles?	
DECISION:	Yes, it is a gene

Gathering Evidence	Explain Your Rationale
What start site do Glimmer and GeneMark suggest?	Glimmer Start Coordinate (type NA if not supported):: N/A GeneMark Start Coordinate (type NA if not supported):: 42363
Does the start site have an associated Ribosome Binding Site with a high score?	Large z-score of 2.876 and a appropriate spacer 15bp
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	No, the longest ORF would be Start 1
Is this start site conserved in other phage genomes as indicated by Starterator?	Yes, it was 51/51 in the pham and called 100% of the time
Is this start site conserved in other phage genomes as indicated by BlastP?	Yes it is highly conserved
DECISION:	Start site is bp 42363, based off Genemark, phamerator, starterator.

Gathering Evidence	Explain Your Rationale
Does this protein align with	List the most informative BlastP match from each source
a protein having a	PhagesDB:
functional assignment in	NCBI:
BlastP (phagesDB and/or	DNA Master:
GenBank) with an	
alignment of 10 ⁻⁴ or	Hint: you may have already found this information from
smaller with appropriate	annotation decision #2. Provide the alignment (q#:s#) and e-
coverage?	value. It is only necessary to provide one match from each
	database.
Does this protein align with	List the most informative HHPred match, including database
a protein having a	source and probability score. It is only necessary to provide the
functional assignment in	best match.
the PDB or other database	
in HHPred with a	Note: If you believe there is not a quality HHPred match, type
probability of 90% or	No Quality Match and list the data for the best match available
greater with appropriate	to affirm the poor quality of the result and to document that
coverage?	HHPred was considered.
Is this gene located	If the answer is YES, evaluate the proposed function in the gene
adjacent to genes of	order. Examine the adjacent genes found in the most closely
known function and in a	related annotated phage (hint: use Phamerator) and record the
region of the genome that	function of the genes found on each side of the gene in the
snows high conservation of	same pham in the most closely related phage. If the answer is
gene order?	NO, enter No Synteny Observed.
Is this gene a possible	If the answer is YES, indicate supporting data from at least 2
transmembrane protein?	different transmembrane prediction programs.
Is the proposed function	Indicate a response with a Yes or No response.
DUACES approved	Once you have arrived at a functional decision, check the <u>SEA-</u>
function list2	the guidelines for function List to ensure that you are following
	the guidelines for function naming. Functions that are not
	approval
	approval.
	If you believe this gene should be assigned, please write the
DECISION	name of the function here. If the evidence does not support a
DECISION:	functional call, record "NKF" for no known function. 50-70% of
	phage genes fall into the NKF category.