

Draft Annotation

An annotation in need of refinement

What do we know?

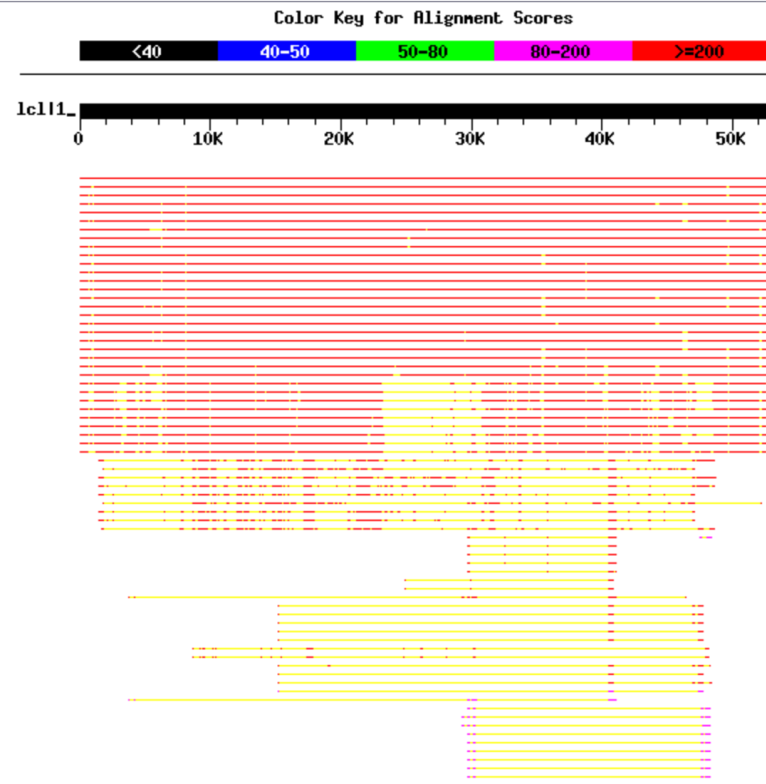
- Go to phageDB (home page and phage page) and answer these questions:
 - What is the host of this phage?
 - What do we know about the host?
 - What do we know about the phage?
 - There are 514 sequenced *Microbacterium* phages on phagesDB (4018 found)
 - There are 19 clusters of *Microbacterium* phages (EA -> EM, GA -> GF)
 - What cluster is Jefe in?
 - How many members of Cluster EC are there?
 - What papers are written about *Microbacterium* phages?
 - 26 Microbial Resource Announcements
 - Jacobs-Sera et al, 2020 – “Genomic diversity of bacteriophages infecting *Microbacterium* spp”

BlastN Jefe.fasta

Query= Microbacterium phage Jefe complete sequence, 53125 bp,
circularly permuted, Cluster EC
(53,125 letters)

[Distribution of 3175 Blast Hits on the Query Sequence](#)

Mouse-over to show defline and scores. Click to show alignments



Sequences producing significant alignments:	Score (bits)	E Value
Microbacterium phage Jefe complete sequence, 53125 bp, circularl...	1.053e+05	0.0
Microbacterium phage Scumberland complete sequence, 53276 bp, ci...	3.958e+04	0.0
Microbacterium phage Cranjis complete sequence, 53222 bp, circul...	3.504e+04	0.0
Microbacterium phage CrazyRich complete sequence, 52950 bp, circ...	3.447e+04	0.0
Microbacterium phage Antares complete sequence, 53362 bp, circul...	3.333e+04	0.0
Microbacterium phage Savannah complete sequence, 53396 bp, circu...	3.306e+04	0.0
Microbacterium phage Quhwah complete sequence, 53549 bp, Cluster EC	3.093e+04	0.0
Microbacterium phage PiperSansNom complete sequence, 53277 bp, c...	2.564e+04	0.0
Microbacterium phage Paschalis complete sequence, 52935 bp, Clus...	2.563e+04	0.0
Microbacterium phage Piperis complete sequence, 53241 bp, circul...	2.545e+04	0.0
Microbacterium phage Selwyn23 complete sequence, 53205 bp, circu...	2.407e+04	0.0
Microbacterium phage Ramiel05 complete sequence, 53250 bp, circu...	2.398e+04	0.0
Microbacterium phage Kowalski complete sequence, 53250 bp, circu...	2.396e+04	0.0
Microbacterium phage BrazzalePHS complete sequence, 53255 bp, ci...	2.389e+04	0.0
Microbacterium phage Honeyfin complete sequence, 53154 bp, circu...	2.345e+04	0.0
Microbacterium phage Busephilis complete sequence, 52986 bp, cir...	2.202e+04	0.0
Microbacterium phage Phrancesco complete sequence, 53255 bp, cir...	2.195e+04	0.0
Microbacterium phage Shotgun complete sequence, 52659 bp, circul...	1.990e+04	0.0
Microbacterium phage KaiHaiDragon complete sequence, 52992 bp, c...	1.855e+04	0.0
Microbacterium phage EarickHC complete sequence, 52992 bp, circu...	1.855e+04	0.0
Microbacterium phage Pulchra complete sequence, 53312 bp, circul...	1.630e+04	0.0
Microbacterium phage Phorgeous complete sequence, 53194 bp, circ...	1.250e+04	0.0
Microbacterium phage NoodlelyBoi complete sequence, 53614 bp, ci...	7586	0.0

DHA
Jefe_Draft

Overview | Features | References | Sequence | Documentation

Sort By Index ▼

Select Features Direct SQL

Type is All ▼

Name like

GenelD =

Locus like

Start >

Length >

Regions >

% GC <

CAI >

EC# like

Product like

Function like

FeatureID =

Tag like

Hide Ignored Features

Tag	Name	5' End	3' End	Length
DNAM_1	1	1	645	645
DNAM_2	2	642	830	189
DNAM_3	3	839	1126	288
DNAM_4	4	1123	2745	1623
DNAM_5	5	2820	3053	234
DNAM_6	6	3050	3457	408
DNAM_7	7	3457	3621	165
DNAM_8	8	3682	4422	741
DNAM_9	9	4422	4748	327
DNAM_10	10	4745	5122	378
DNAM_11	11	5192	5404	213
DNAM_12	12	5401	5811	411
DNAM_13	13	5811	6179	369
DNAM_14	14	6176	6406	231
DNAM_15	15	6403	6747	345
DNAM_16	16	6840	7385	546
DNAM_17	17	7451	7621	171
DNAM_18	18	7624	8016	393
DNAM_19	19	8013	8201	189
DNAM_20	20	8427	10094	1668
DNAM_21	21	10091	10537	447
DNAM_22	22	10518	10616	99
DNAM_23	23	10609	11412	804
DNAM_24	24	11488	11772	285
DNAM_25	25	11891	13360	1470
DNAM_26	26	13357	13599	243
DNAM_27	27	13737	15596	1860

Description | Sequence | Product | Regions | Blast | Context

Name [GenelD](#)

Type CDS ▼ [GI](#)

5' End Locus Tag

3' End Regions

Length Tag

Direction Forward Feature ID

Translation Table Standard Code ▼

EC Number

Product gp1 ▼

Function

Notes

1 - 50000
Position : 31182
 Controls >> Map
 Map >> Controls

91 Features
Live
53125

What happens next?

- Once a sequence is “**finalized**”, an automated draft annotation is sent to Phamerator and PhagesDB.
 - Phamerator has 2 components:



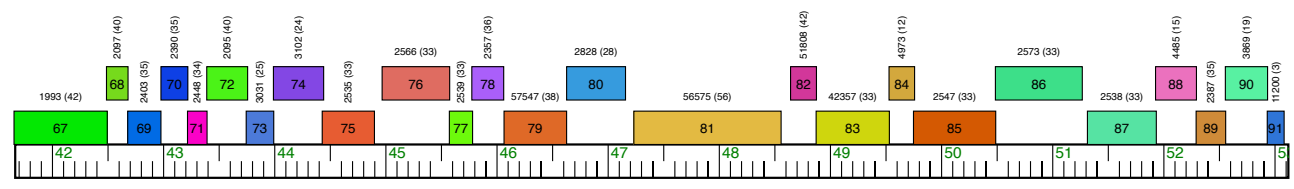
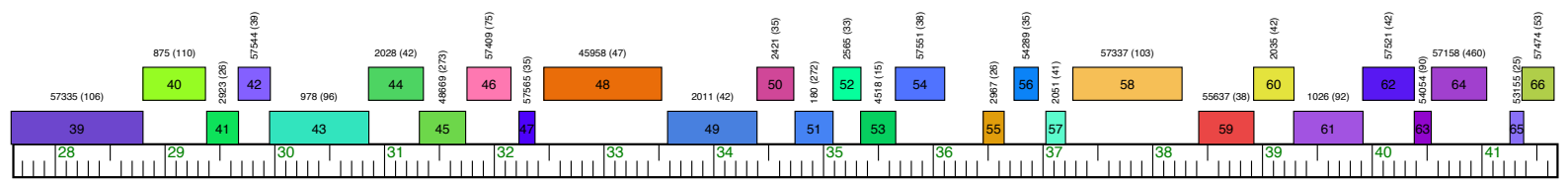
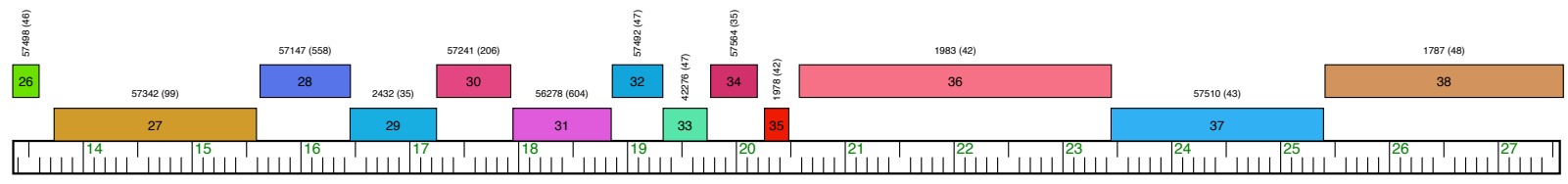
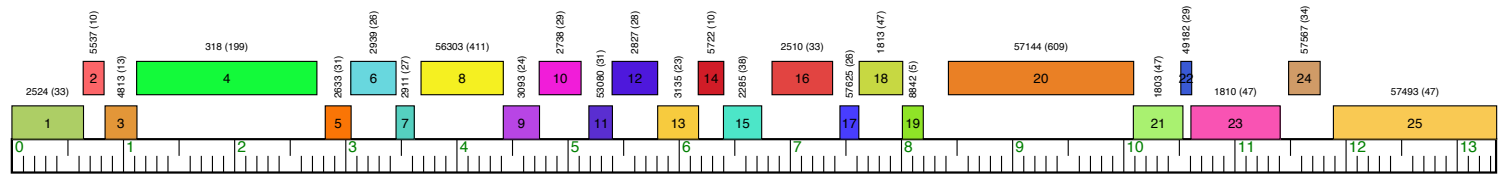
Steve Cresawn

- Database creation – a PECAAN server and Pitt set of actions
- Database display – the web display from James Madison University
- PhagesDB display- lists



Dan Russell

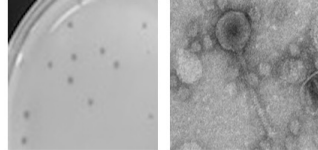
Jefe Draft (Cluster EC)



The Actinobacteriophage Database
at PhagesDB.org

Phages Hosts Data BLAST Publications Resources Software Social About

Microbacterium phage Jefe



Locally BLAST this genome Run GeneMarkS Run GeneMark (M. folio B-24224)

Know something about this phage that we don't? [Modify its data.](#)

Detailed Information for Phage Jefe

Discovery Information	
Isolation Host	<i>Microbacterium foliorum</i> NRRL B-24224
Former names	EJefe
Found By	Salvador Simpson
Year Found	2020
Location Found	Radford, VA United States
Finding Institution	Radford University
Program	Science Education Alliance-Phage Hunters Advancing Genomics and Evolutionary Science
From enriched soil sample?	Yes
Isolation Temperature	30°C
GPS Coordinates	37.102003 N, 80.567825 W Map
Discovery Notes	Grassy hillside with lots of leaf litter About 5cm depth
Naming Notes	From the movie 22 Jump Street
Sequencing Information	
Sequencing Complete?	Yes
Date Sequencing Completed	Aug 18, 2021
Sequencing Facility	Pittsburgh Bacteriophage Institute
Shotgun Sequencing Method	Illumina Sequencing
Approximate Shotgun Coverage	57
Genome length (bp)	53125
Character of genome ends	Circularly Permuted
GC Content	68.9%
Fasta file available?	Yes: Download fasta file
Characterization	
Cluster	EC
Subcluster	--
Cluster Life Cycle	Lytic
Other Cluster Members	Click to View
Annotating Institution	University of Pittsburgh
Annotation Status	Being Annotated (Expected completion by 5/1/2023)
Plaque Notes	On average the plaque size was 0.169cm in diameter
Has been Pharmed?	Yes
Gene List	Click to View

Gene Details

[BLAST the gene product on PhagesDB](#)

[BLAST the gene product on NCBI](#)

[Next Gene >](#)

Details for gene Jefe_Draft_1

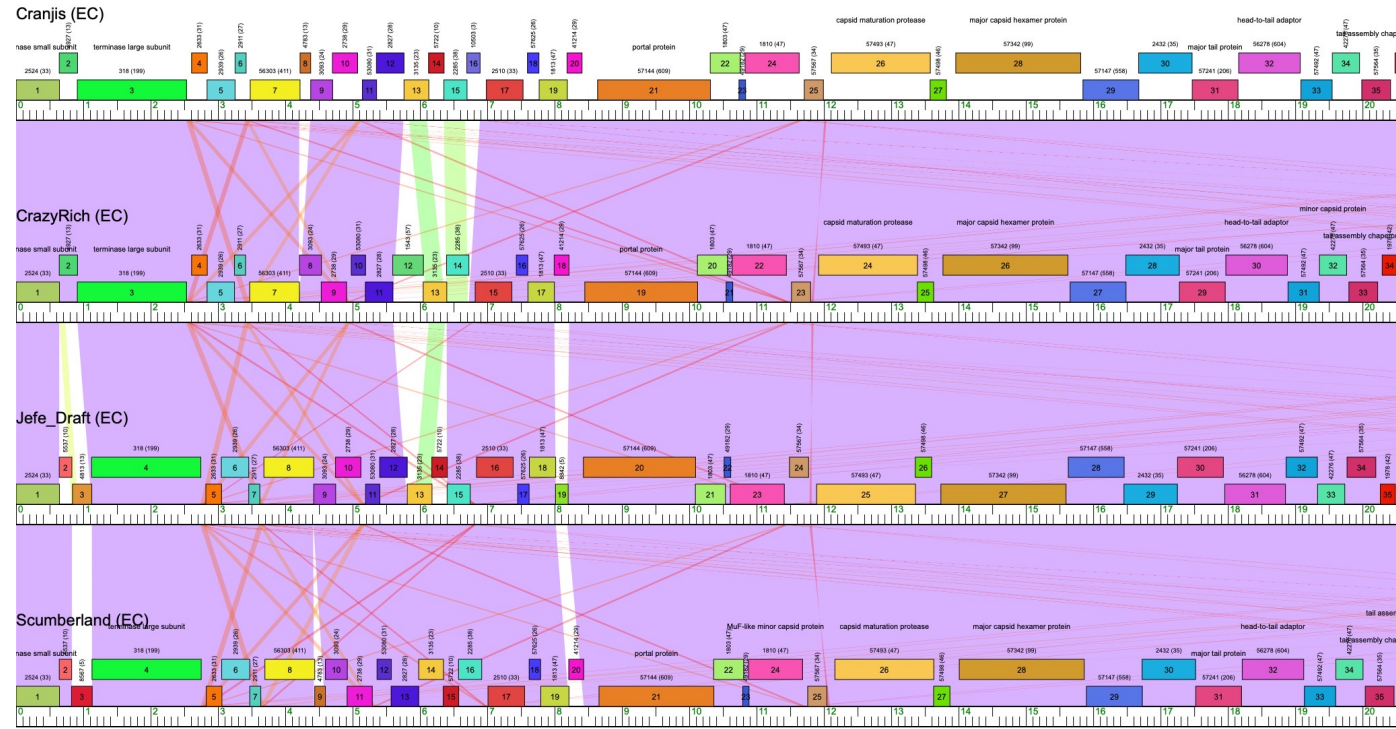
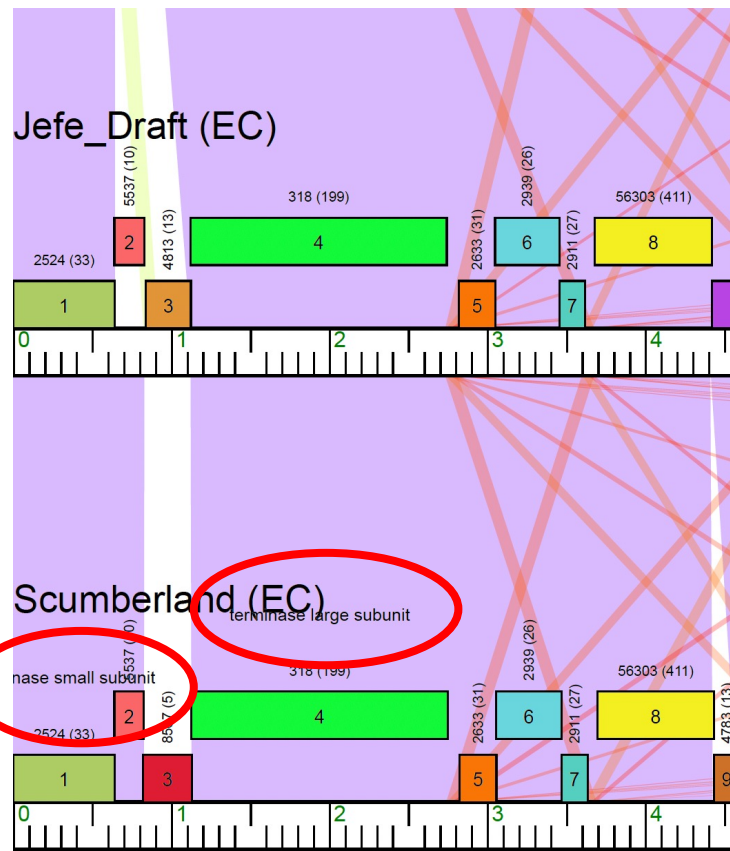
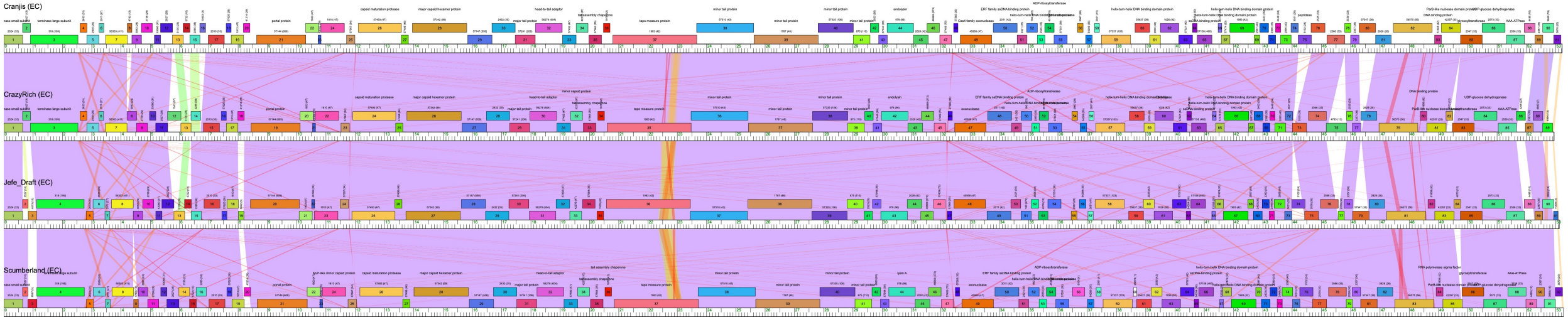
Phage	Jefe · Cluster EC · 53125 bp
Gene	Jefe_Draft_1
Pham (click for Pham view →)	2524
Starterator	Pham 2524 report
Genome Position	1 to 645 (Forward)
Length	645 base pairs 214 amino acids
Amino Acid Sequence	Click to View
Notes	

Members (33) of Pham 2524

Antares_1 (terminase small subunit)	Ariadne_1 (terminase small subunit)
BrazzalePHS_Draft_1	Busephilis_1 (terminase small subunit)
ClearAsMud_1 (terminase small subunit)	Cranjis_1 (terminase small subunit)
CrazyRich_1 (terminase small subunit)	EarickHC_1 (terminase small subunit)
Fireman_1 (terminase small subunit)	Gina_1 (terminase small subunit)
Honeyfin_1 (terminase small subunit)	Jayden_1 (terminase small subunit)
Jefe_Draft_1	KaiHaiDragon_1 (terminase small subunit)
Kowalski_1 (terminase small subunit)	Metamorphoo_1 (terminase small subunit)
NoodlelyBoi_1 (terminase small subunit)	Paschalis_1 (terminase small subunit)
Phorgeous_1 (terminase small subunit)	Phrancesco_1 (terminase small subunit)
Piperis_1 (terminase small subunit)	PiperSansNom_1 (terminase small subunit)
Pulchra_1 (terminase small subunit)	Quhwah_1 (terminase small subunit)

I have Jefe-Draft surrounded by finished annotations

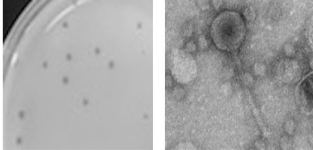
- How do I know? Jefe_**Draft** vs Cranjis, Crzyrich, and Scumberland
- What to do next:
 - Find the differences
 - Why are there differences?
 - Assign Draft FUNCTIONS
 - Why? To help with context
 - What is the source of these functional calls?



The Actinobacteriophage Database
at PhagesDB.org

Phages Hosts Data BLAST Publications Resources Software Social About

Microbacterium phage Jefe



Locally BLAST this genome Run GeneMarkS Run GeneMark (M. folio B-24224)

Know something about this phage that we don't? [Modify its data.](#)

Detailed Information for Phage Jefe

Discovery Information	
Isolation Host	<i>Microbacterium foliorum</i> NRRL B-24224
Former names	EJefe
Found By	Salvador Simpson
Year Found	2020
Location Found	Radford, VA United States
Finding Institution	Radford University
Program	Science Education Alliance-Phage Hunters Advancing Genomics and Evolutionary Science
From enriched soil sample?	Yes
Isolation Temperature	30°C
GPS Coordinates	37.102003 N, 80.567825 W Map
Discovery Notes	Grassy hillside with lots of leaf litter About 5cm depth
Naming Notes	From the movie 22 Jump Street
Sequencing Information	
Sequencing Complete?	Yes
Date Sequencing Completed	Aug 18, 2021
Sequencing Facility	Pittsburgh Bacteriophage Institute
Shotgun Sequencing Method	Illumina Sequencing
Approximate Shotgun Coverage	57
Genome length (bp)	53125
Character of genome ends	Circularly Permuted
GC Content	68.9%
Fasta file available?	Yes: Download fasta file
Characterization	
Cluster	EC
Subcluster	--
Cluster Life Cycle	Lytic
Other Cluster Members	Click to View
Annotating Institution	University of Pittsburgh
Annotation Status	Being Annotated (Expected completion by 5/1/2023)
Plaque Notes	On average the plaque size was 0.169cm in diameter
Has been Pharmed?	Yes
Gene List	Click to View

Gene Details

[BLAST the gene product on PhagesDB](#)

[BLAST the gene product on NCBI](#)

[Next Gene >](#)

Details for gene Jefe_Draft_1

Phage	Jefe · Cluster EC · 53125 bp
Gene	Jefe_Draft_1
Pham (click for Pham view →)	2524
Starterator	Pham 2524 report
Genome Position	1 to 645 (Forward)
Length	645 base pairs 214 amino acids
Amino Acid Sequence	Click to View
Notes	

Members (33) of Pham 2524

Antares_1 (terminase small subunit)	Ariadne_1 (terminase small subunit)
BrazzalePHS_Draft_1	Busephilis_1 (terminase small subunit)
ClearAsMud_1 (terminase small subunit)	Cranjis_1 (terminase small subunit)
CrazyRich_1 (terminase small subunit)	EarickHC_1 (terminase small subunit)
Fireman_1 (terminase small subunit)	Gina_1 (terminase small subunit)
Honeyfin_1 (terminase small subunit)	Jayden_1 (terminase small subunit)
Jefe_Draft_1	KaiHaiDragon_1 (terminase small subunit)
Kowalski_1 (terminase small subunit)	Metamorphoo_1 (terminase small subunit)
NoodlelyBoi_1 (terminase small subunit)	Paschalis_1 (terminase small subunit)
Phorgeous_1 (terminase small subunit)	Phrancesco_1 (terminase small subunit)
Piperis_1 (terminase small subunit)	PiperSansNom_1 (terminase small subunit)
Pulchra_1 (terminase small subunit)	Quhwah_1 (terminase small subunit)

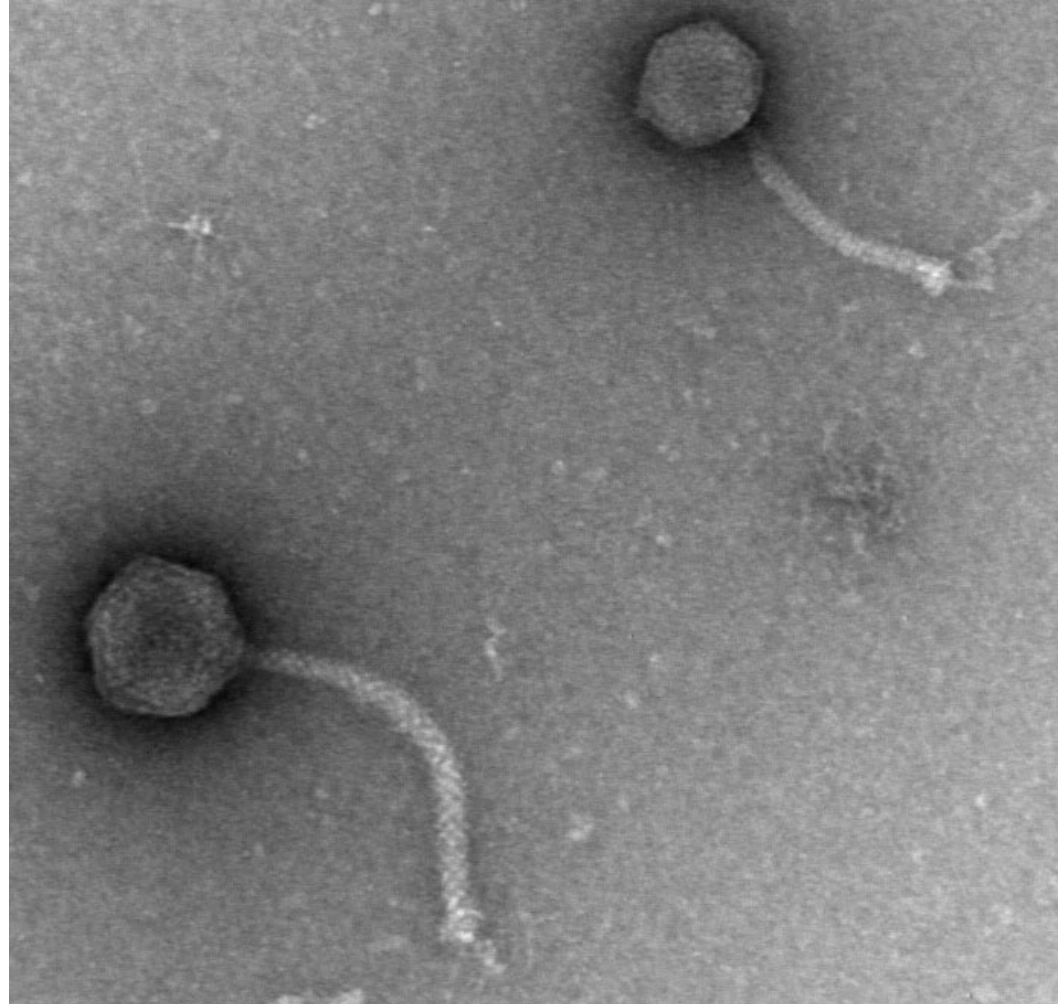
To Do:

- Assign functions according to Phamerator/phagesDB
 - Record on your phage map
 - Record on classroom map
- Add any insights or questions as you go?

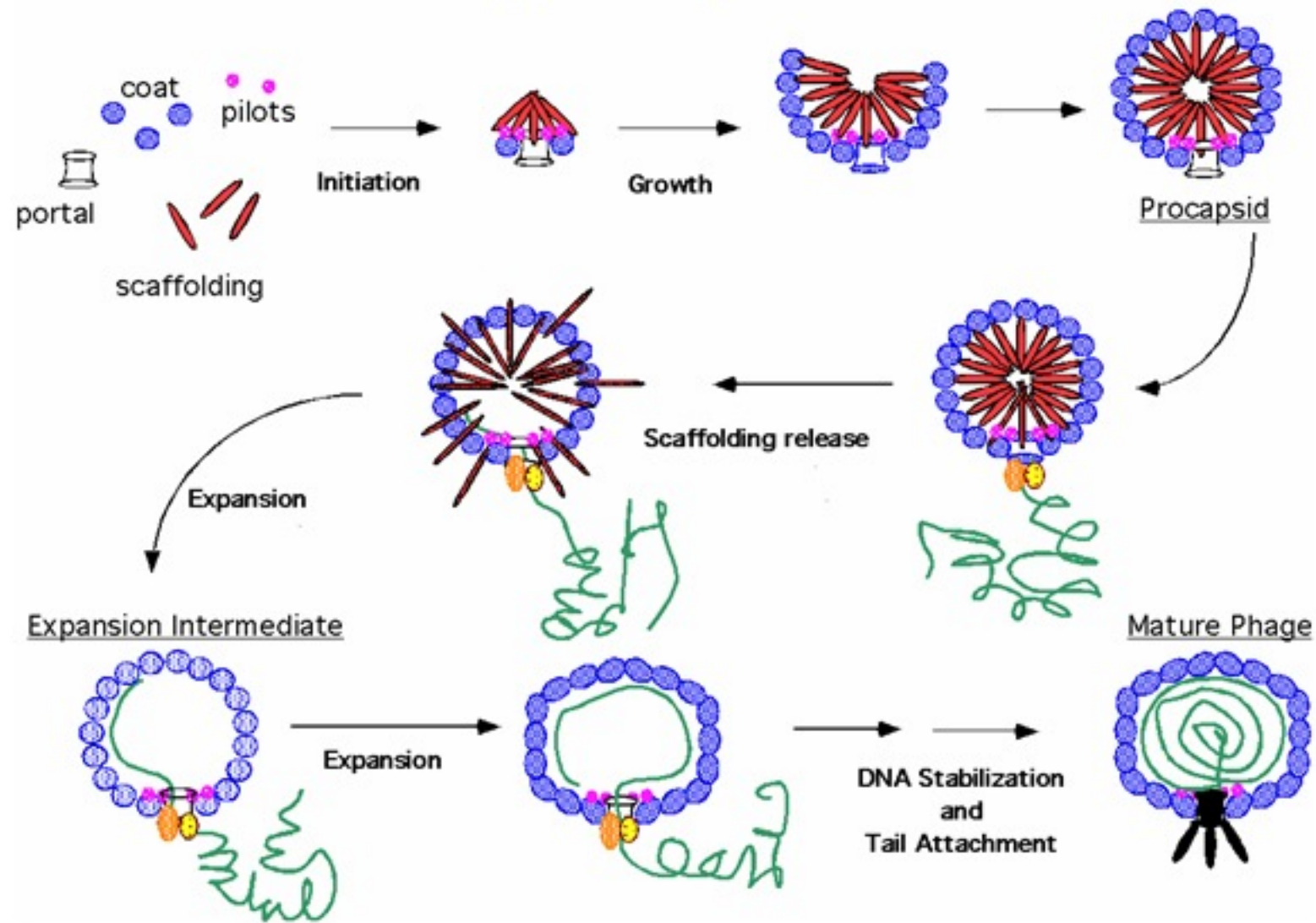
What functions did we find that we expected?

What functions did we find that we did not expect?

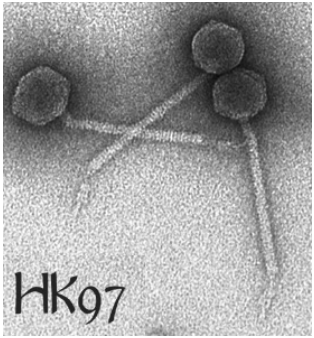
Structural genes



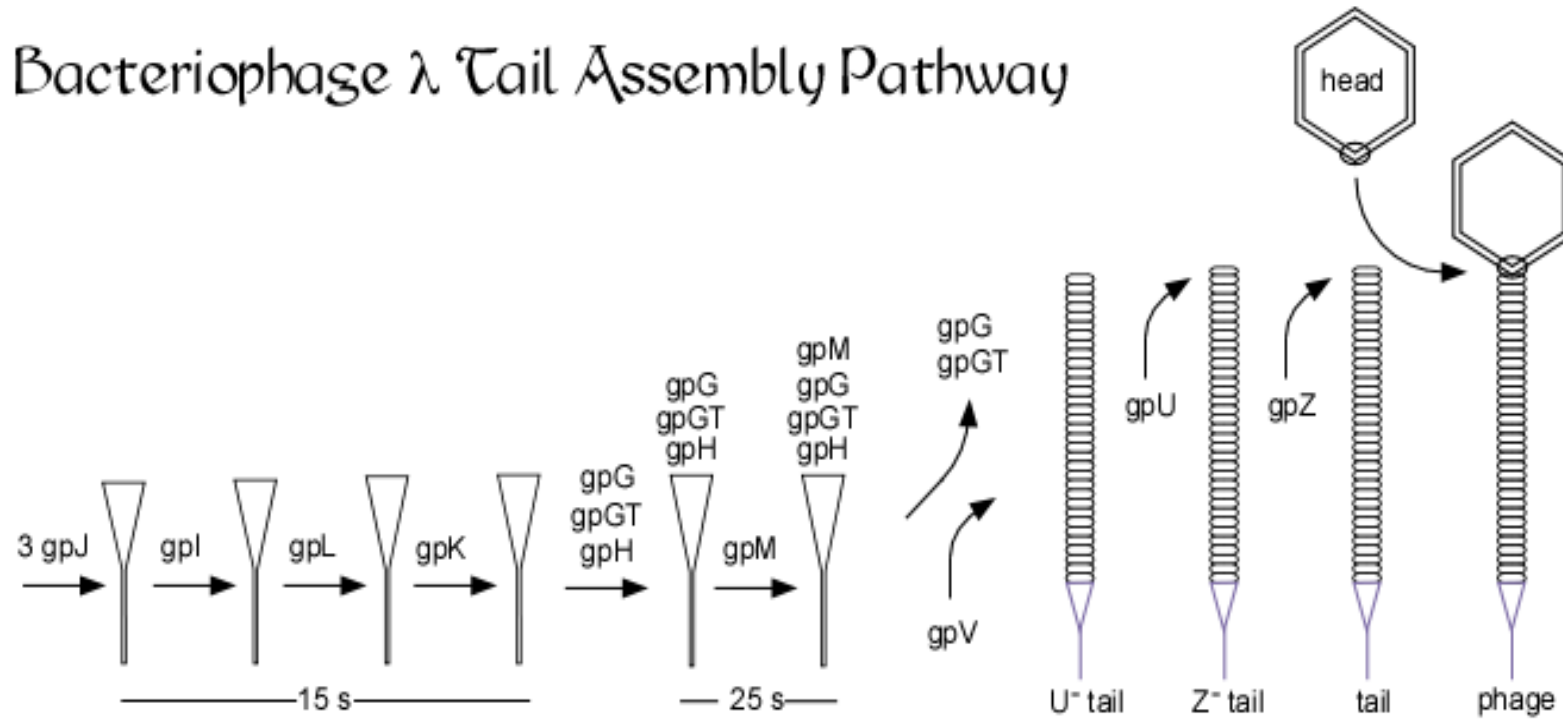
Assembly pathway of P22



Adapted from B. Greene

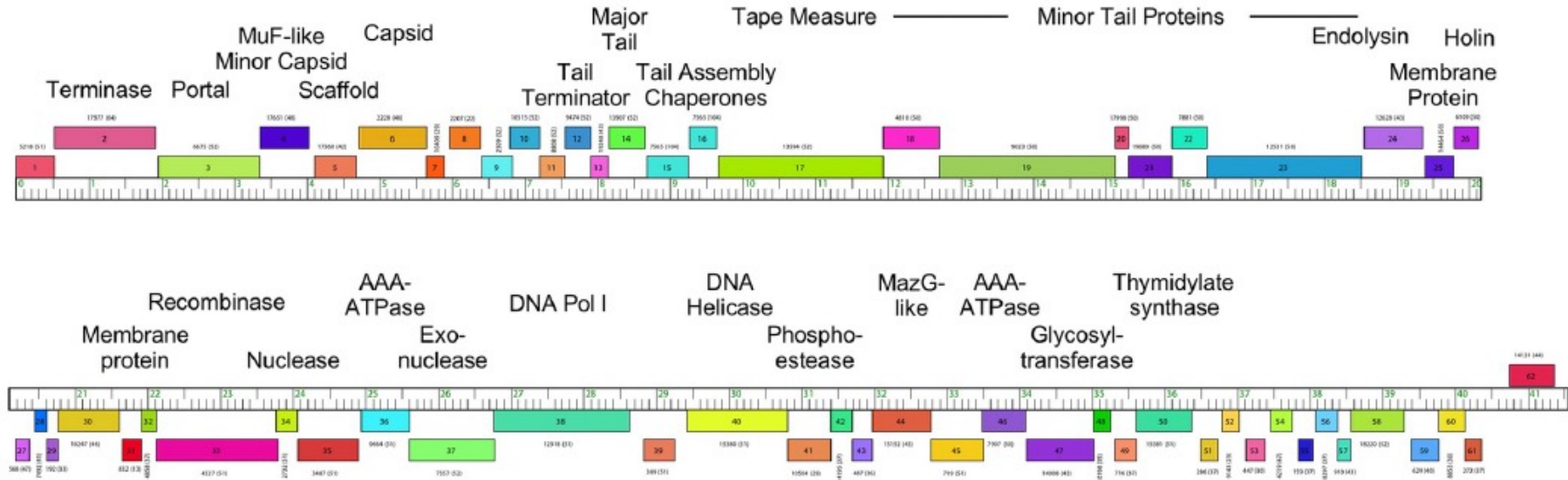


Bacteriophage λ Tail Assembly Pathway



Bob Duda

A TeddyBear (EA1)



- **Virion structural and assembly genes**, i.e. those encoding proteins that are either components of virion particles or assist in their formation. These include genes encoding the terminase, portal, capsid maturation protease, scaffolding protein, major capsid protein, head to tail connectors, major tail subunit, tail assembly chaperones, tape measure protein, and minor tail proteins.
- **Genes involved in phage DNA replication.** These include DNA polymerase, DNA primase, DNA helicase, nucleotide metabolism genes, and ssDNA binding proteins.
- **Genes involved in life cycle regulation.** These include various regulators such as repressors and activators, integrases, recombination directionality factors, etc.
- **Genes involved in lysis**, including endolysins (referred to as Lysin A in the mycobacteriophages), Lysin B, and Holins.
- **Other genes**, including transcription factors, toxin/anti-toxin systems, peptidases, phosphatases, host gene homologues, methylases, nucleases, and DNA binding proteins, among others.

What was the point of this exercise?

- Provides context for the work that we will do

What is the work that we will do?

- Call the genes of the genome as precisely as possible by answering 3 questions:
 - Is it a gene?
 - What is its start?
 - What is its function?