

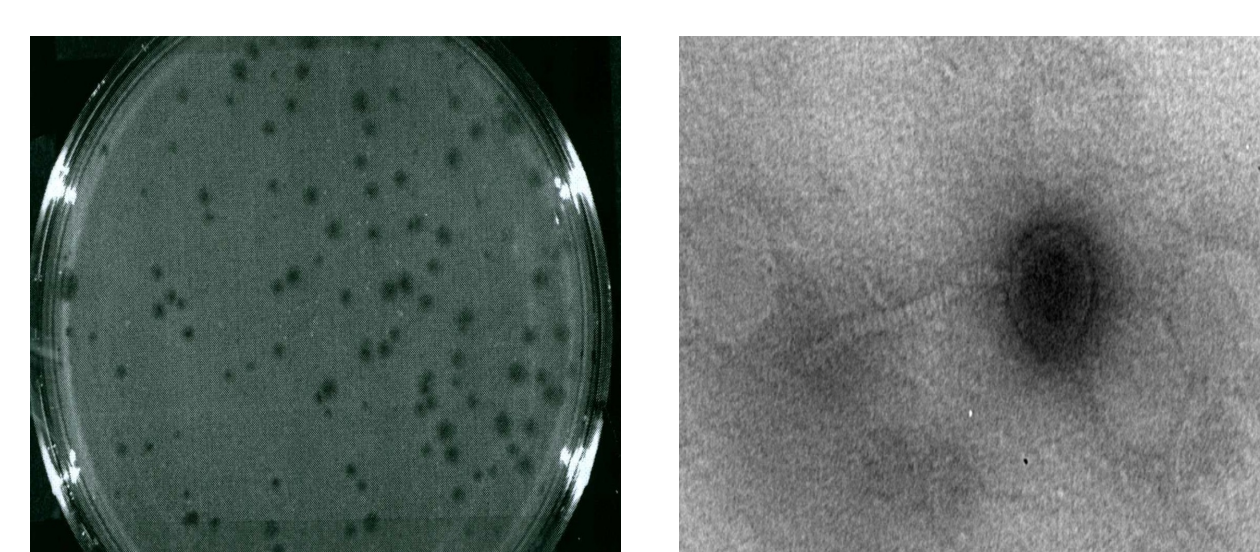
Angela V Albrecht, Lauren E Schlegel, Jennifer G Bateman, Jessie A Brill, Juliet S Chung, Lisa M Glover, Chelsea M Hipwell, Audrey K Hla, Allison B Kelliher, Nicole A Lando, Idowu D Olugbade, Garrett T Santini, Emily R Seier, Divya G Sirdeshpande, Barbara E Tsaousis, Juliana N Young, Elise C Esposito, Emily L Heckman, Kenneth J Brill, Mikala A Kowal, Catherine M Magee, Javier A Buceta, Margaret A Kenna, Vassie C Ware
Department of Biological Sciences, Lehigh University, Bethlehem, PA 18015

Abstract

The SEA-PHAGES program at Lehigh University is a collaborative undergraduate research enterprise focused on isolating and characterizing phages that infect Actinobacter hosts to gain a better understanding of phage genome structure, gene function, and phage biology in general. In addition to uncovering new phages for comparative genome analysis, our program has focused on characterizing phage gene products that function in lytic infection to evaluate their potential as biocontrol agents to target pathogenic bacteria. Here we report on the genomic characterization of four of our most recently discovered novel phages (Derek [C1], Annyong [A4], James [F1], and Taptic [W]) that infect *Mycobacterium smegmatis*. We also present an update on the functional characterization of one of several phage gene products encoded by orphans found in the N cluster phage, *Mycobacterium* phage Butters. Of particular interest among our newly discovered phages is *Mycobacterium* phage Taptic, a third member of the W cluster. With a genome size of 60,973 bp, Taptic has a 91 ORFs, all positioned on the forward strand – an organization common in genomes of only a few mycobacteriophage clusters. Taptic plaques are minuscule (<1mm) and clear – the latter suggestive of a lytic phage. Annotation of a putative function for *gp38* as a helix-turn-helix repressor (>99% probability in HHPred) for transcriptional regulation suggests several possibilities, including a putative lysogenic cycle for Taptic. To explore this possibility, lysogeny experiments were performed. The presence of Taptic mesas and positive patch tests were suggestive of Taptic lysogens; however, stable lysogens were not formed as host cells were consistently lysed over time. The role of *gp38* in the Taptic life cycle remains to be explored. In other experiments, the computational prediction that Butters *gp31* is a transmembrane protein was tested by expressing a C terminal tetracycline-tagged *gp31* and control ORFs within *E. coli* and imaging tagged proteins using fluorescence microscopy. Data show that Butters *gp31* resides within the *E. coli* membrane coincident with a membrane marker. Further analysis of Butters *gp31* function as a membrane protein is underway to determine how this membrane protein may function in a mycobacterial host. In summary, we will present a comparative analysis of four novel mycobacteriophage genomes and provide an update on functional studies of a newly characterized membrane protein encoded by Butters *gp31*. Overall, our bioinformatics efforts continue to highlight genomic features of interest for laboratory exploration.

I. Genome Annotation of Novel Mycobacteriophages

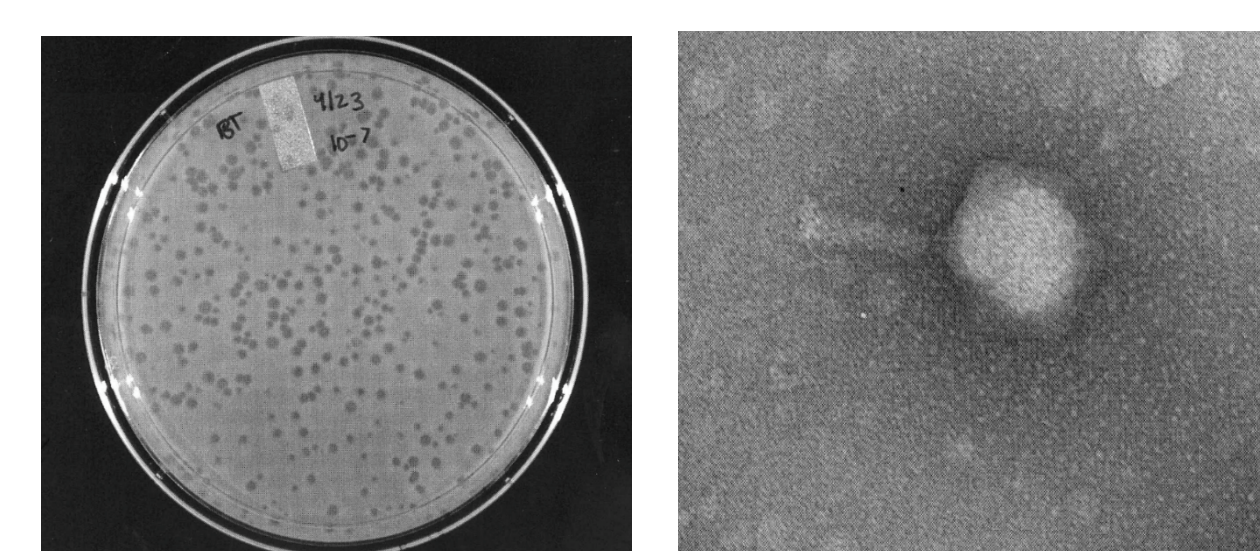
Mycobacterium phage Annyong



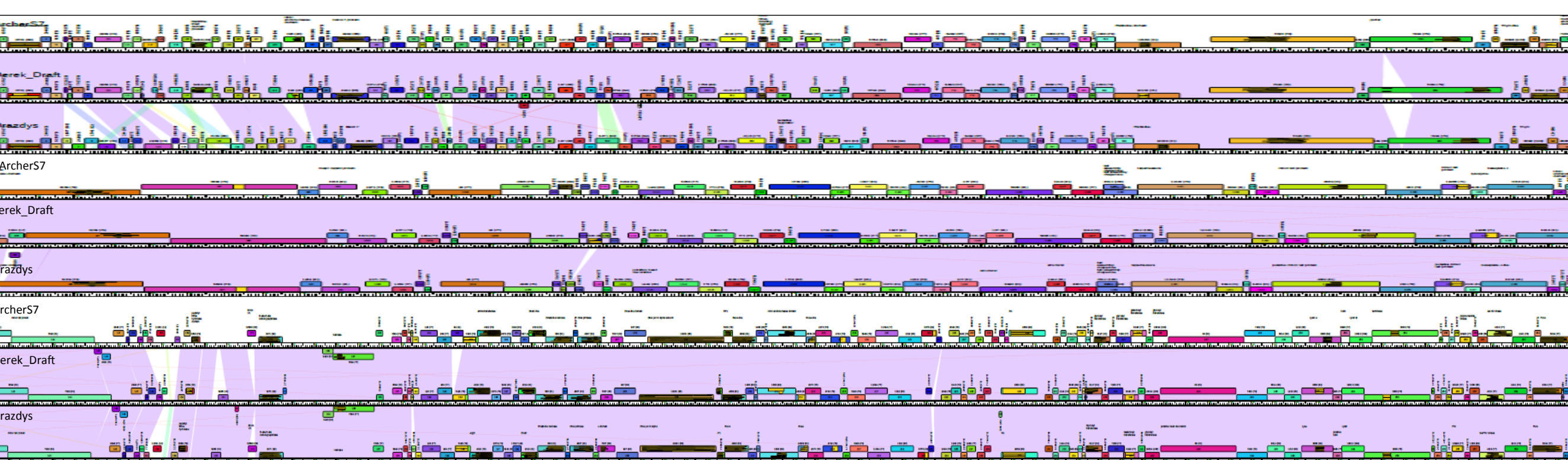
- Founder: Audrey Hla
- Location: Avon, CT
- Characteristics:
 - A4 cluster, 85 ORFs
 - 51,418 bp linear genome with 3' sticky overhang
 - Turbid plaques
 - Temperate Phage
 - Siphoviridae (1:4 head:tail ratio)



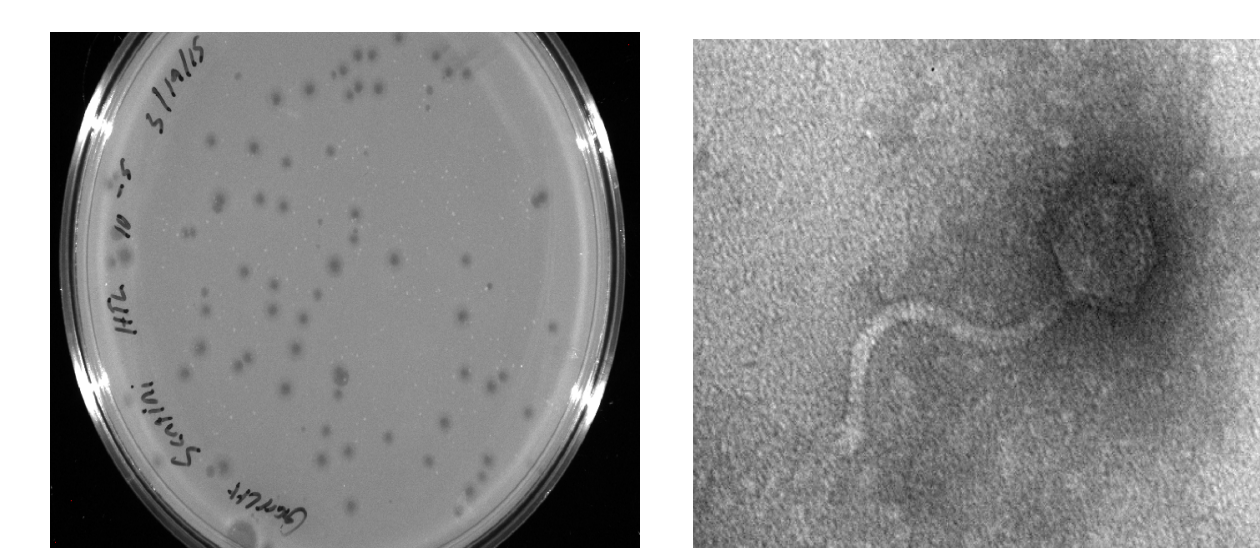
Mycobacterium phage Derek



- Founder: Barbara Tsaousis
- Location: Bethlehem, PA
- Characteristics:
 - C cluster, 229 ORFs, 31 tRNAs, 1 tmRNA
 - 156,199 bp circularly permuted genome
 - Turbid plaques (1mm after 24hr)
 - Temperate Phage
 - Myoviridae (1:1 head:tail ratio)



Mycobacterium phage James



- Founder: Garrett Santini
- Location: Wayne, NJ
- Characteristics:
 - F1 subcluster, 108 ORFs
 - 59,617 bp linear genome with 3' sticky overhang
 - Bullseye plaques (2mm after 24hr)
 - Temperate Phage
 - Siphoviridae (~1:2.5 head:tail ratio)



Genomic Features of Interest:

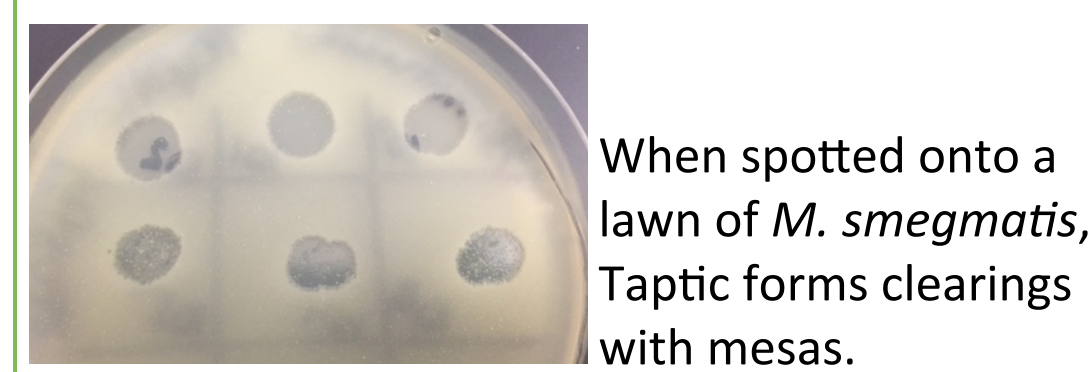
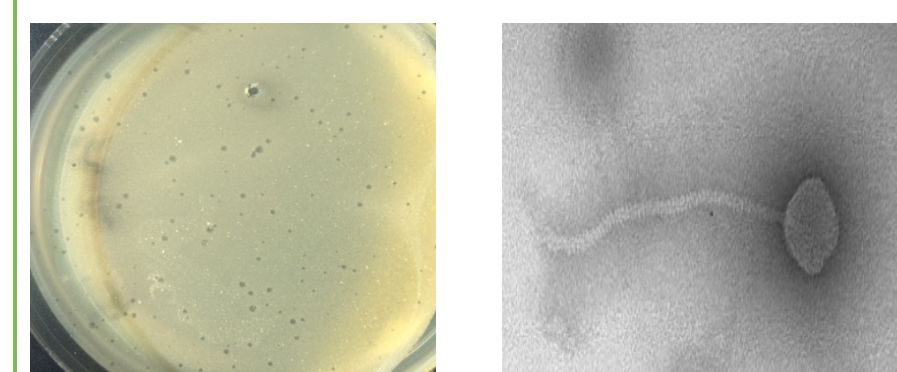
Annyong: 87 ORFs; no orphans; 99% identical to MeeZee
Derek: 230 ORFs, 31 tRNAs, 1 tmRNA; no orphans; 99% identical to SmallFry
James: 108 ORFs; no orphans; 99% identical to Saal

Future Work:

- Determining if Derek lysogens can be formed
- Exploring the use of DNA fragmentation and Bacteriophage Recombining using Electroporated DNA (BRED) to genetically manipulate large phage genomes

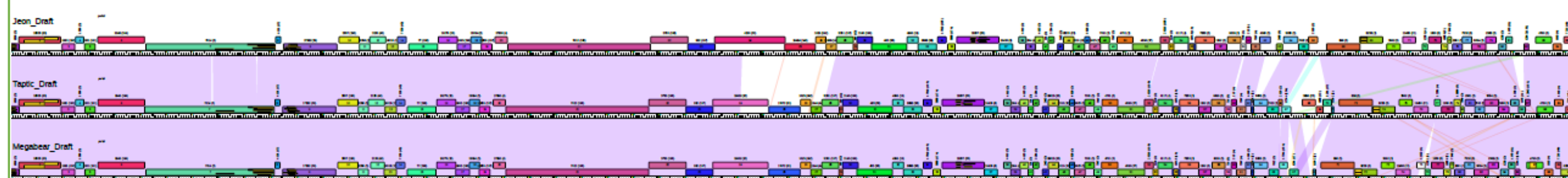
II. Exploring the Taptic life cycle

Mycobacterium phage Taptic



- Founder: Emily Seier
- Location: Northampton, PA
- Characteristics:
 - W cluster, 91 ORFs, 1 tRNA
 - 60,973 bp circularly permuted genome
 - Clear plaques (pinprick)
 - Forms mesas when spotted onto *M. smegmatis* lawn (see below)
 - Siphoviridae (1:5 head:tail ratio)

Phamerator map of Taptic Genome



Genomic Features of Interest:

93 ORFs, 1 tRNA-Gly; Repressor (*gp38*), *gp1*: HTH DNA binding protein; *P1* ParB, plasmid partition, antitoxin component of toxin-antitoxin system; *gp42*: HTH DNA Binding protein; toxin-antitoxin system antidote; *gp75*: *P1* ParB plasmid partition protein; 84% of gene function unknown.

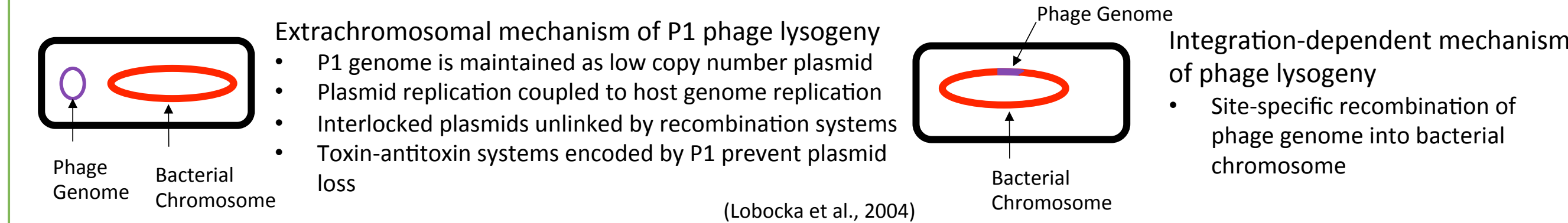
Interesting gene functions with low probability scores (<90%) in HHPred: *gp43*: RepA, replication initiator protein A; *gp45*: toxin type I toxin anti-toxin system; *gp52*: P1-like phage RecA-dependent nuclease; *gp79*: antitoxin of the type IV toxin-antitoxin; *gp87*: P1-like DOC (death on curing)

Hypothesis:

Several lines of evidence support the hypothesis that Taptic may be a temperate phage:

- Overgrowth of bacterial cells within Taptic clearings forming mesas
- Bioinformatic analysis predicting *gp38* function as a putative repressor
 - Note no clear bioinformatics evidence for a canonical-type integrase.
- Several functions related to P1 prophage maintenance replication and toxin-antitoxin systems

It may be temperate and model a mechanism of P1 phage lysogeny.

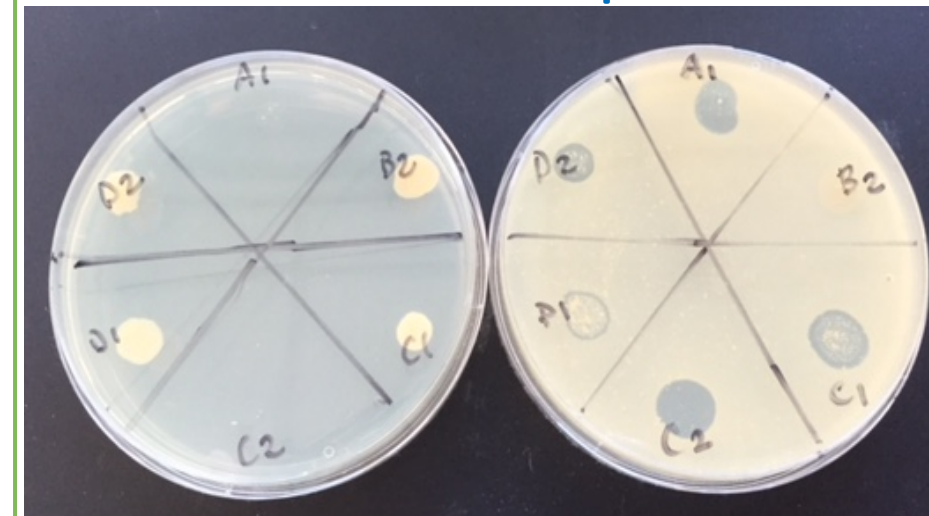


Experimental Plan:

Evaluate bacterial cells from mesas as possible lysogens (as described at phagesdb.org) and test putative lysogens for phage release and immunity response to Taptic.

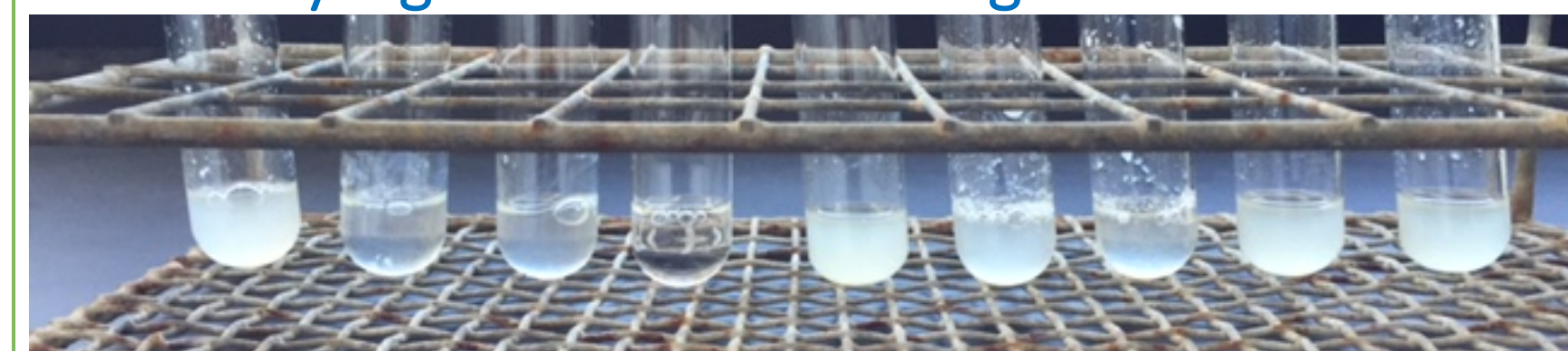
Results:

Patch tests reveal putative Taptic lysogens



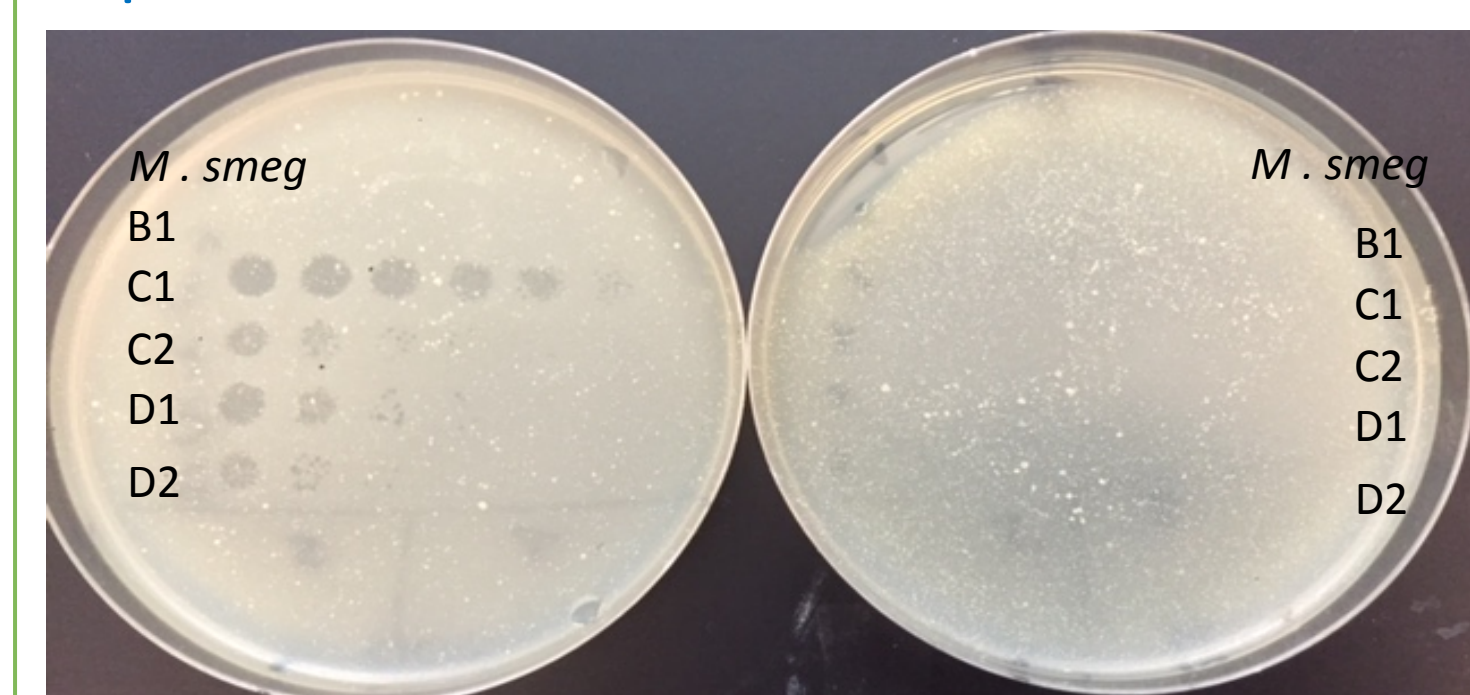
Putative Taptic lysogens from modified patch tests. Single bacterial colonies (A1, B2, C1, C2, D1, D2) were resuspended in 7H9 and an equal volume spotted onto each of two plates. Left plate shows bacterial spots grown in the absence of a top agar overlay. Right plate shows bacterial cells spotted onto *M. smeg*/Top agar. Phage clearings are observed in putative lysogens A1, C1, C2, D1, and D2. Note the minimal bacterial growth of lysogens A1 and C2, but the presence of large clearings.

Putative lysogens show bacterial growth variation in liquid culture



Putative Taptic lysogens grown in liquid culture. Variable levels of growth are observed, possibly due to instability of a Taptic lysogen.

Supernatant release from cells re-streaked after patch tests show phage



Supernatants from each Taptic lysogen were spotted onto *M. smegmatis* or C1 lysogen lawns. Supernatants from lysogens C1, C2, D1, D2 but not from B1 or an *M. smegmatis* culture contain phage. C1 lysogen immunity appears established due to reduced infectivity of all supernatants on the C1 lawn.

Putative Taptic Lysogen	Bacterial Spot on Patch Test	Clearing from Patch Test	Liquid Cultures Growth	Supernatant Release on <i>M. smeg</i> Lawn	Supernatant Release on C1 Lawn
<i>M. Smegmatis mc</i> ² 155	N/A	N/A	Very Turbid	No clearings	No clearings
A1	None	7mm	Minimal Turbidity	N/A	N/A
A2	N/A	N/A	Minimal Turbidity	N/A	N/A
B1	N/A	N/A	No growth	No clearings	No clearings
B2	8mm	None	Very Turbid	N/A	N/A
C1	5mm	10mm	Moderate Turbidity	Countable plaques 10 ³	Clearing 10 ⁰
C2	None	10mm	Minimal Turbidity	Countable plaques 10 ²	Clearing 10 ⁰
D1	9mm	2mm	Very Turbid	Countable plaques 10 ²	Clearing 10 ⁰
D2	6mm	6mm	Very Turbid	Countable plaques 10 ⁴	Clearing 10 ⁰

Unusual characteristics of Taptic lysogen isolation

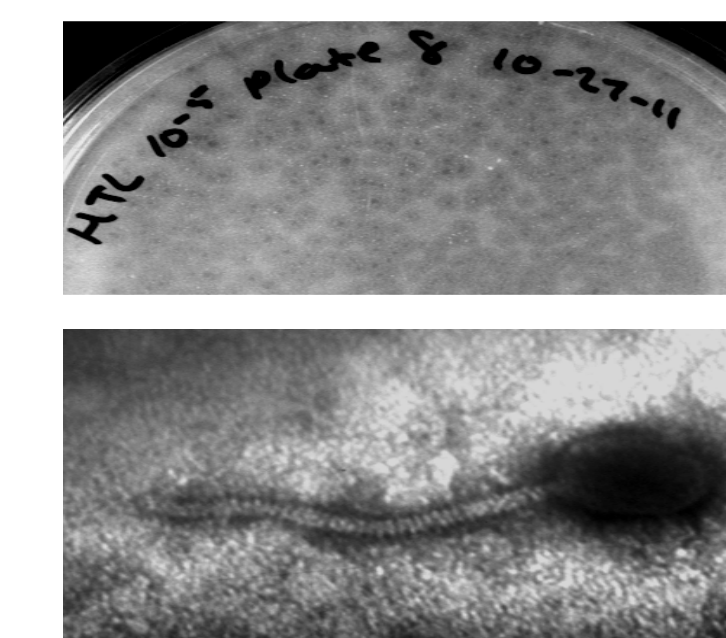
- Some colonies releasing the most phage in patch test show the smallest amount of bacterial growth
- Supernatants from bacterial re-growth show significantly reduced phage titers compared to original supernatants

Future Work:

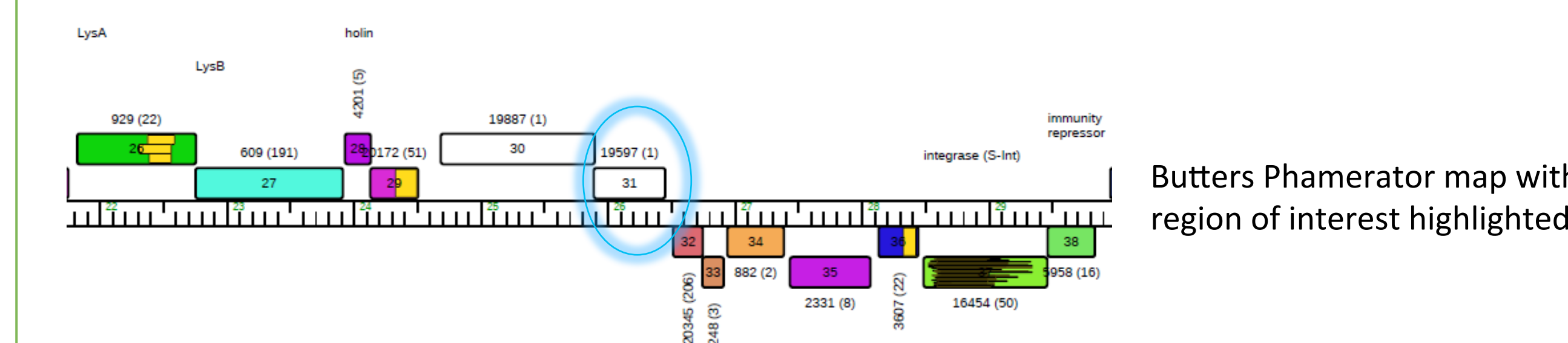
- Confirm presence of Taptic in putative lysogens by PCR
- Complete immunity testing and further lysogen purification
- Explore possibility that putative Taptic prophage is maintained extrachromosomally within *M. smegmatis*. (e.g., similar to a phage P1 mechanism)
- Clone *gp38* into pLAM12 and transform into *M. smegmatis* to ask if immunity relationships are preserved with repressor alone

III. Investigating GP31 in Mycobacteriophage Butters

An Introduction to *Mycobacterium* phage Butters



- Founder: Lena Ma SEA 2011-12
- Characteristics:
 - Turbid plaques (0.5-3mm 24-72 hours)
 - Temperate phage
 - N cluster, 41,491bp linear genome, defined physical ends
 - 66 ORFs, includes 3 orphans of unknown function, and N specific ORF (phamerator map 6/2016)



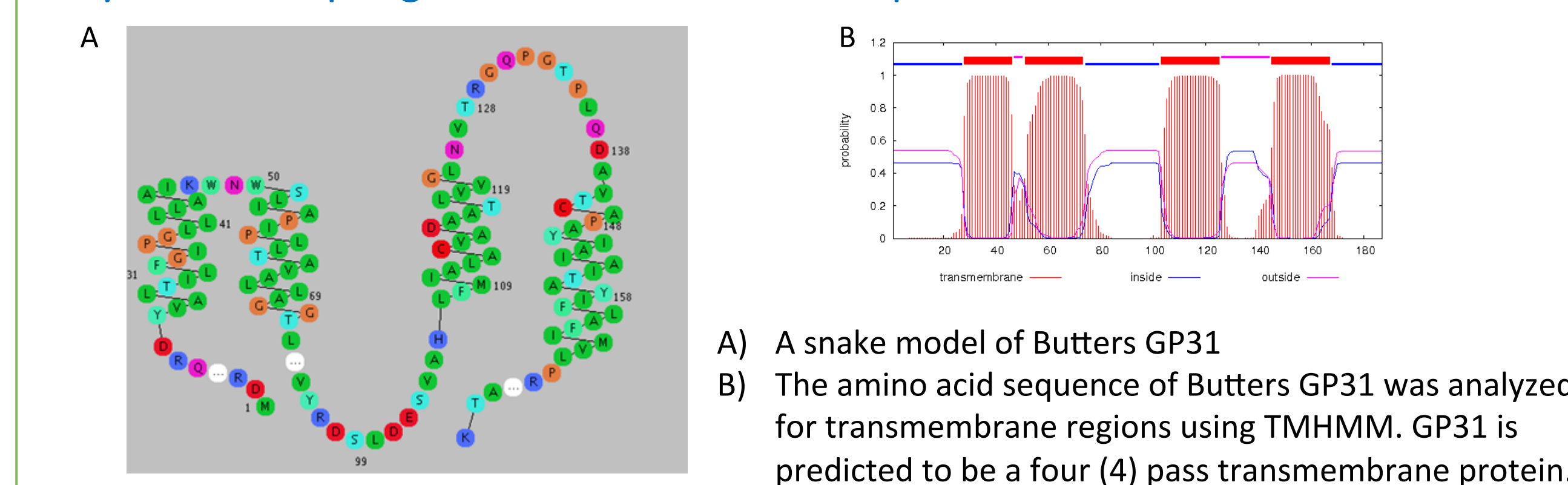
Hypothesis:

Mycobacteriophage Butters GP31 is a membrane protein.

Experimental Plan:

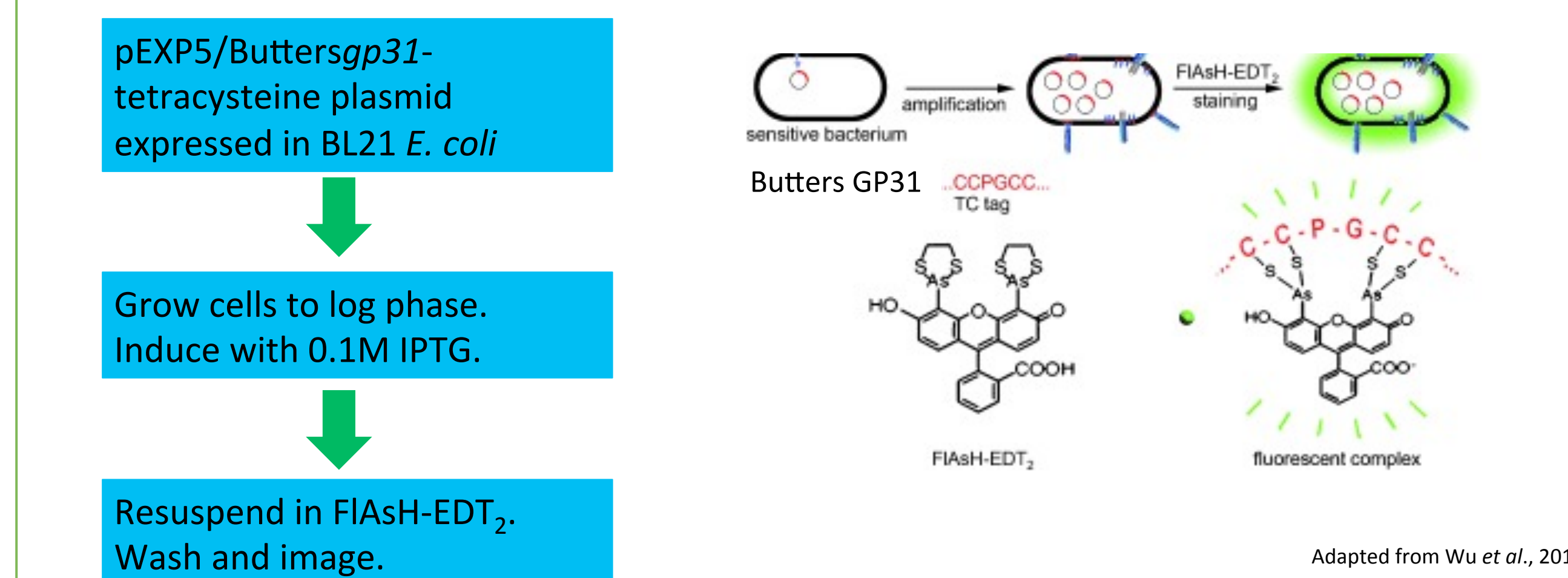
Use TMHMM to computationally model Butters GP31. Clone *gp31*-tetracycline into pEXP5. Using a molecular probe, FIAsh, determine Butters GP31 localization within BL21 *E. coli*.

Mycobacteriophage Butters GP31 has four predicted transmembrane domains



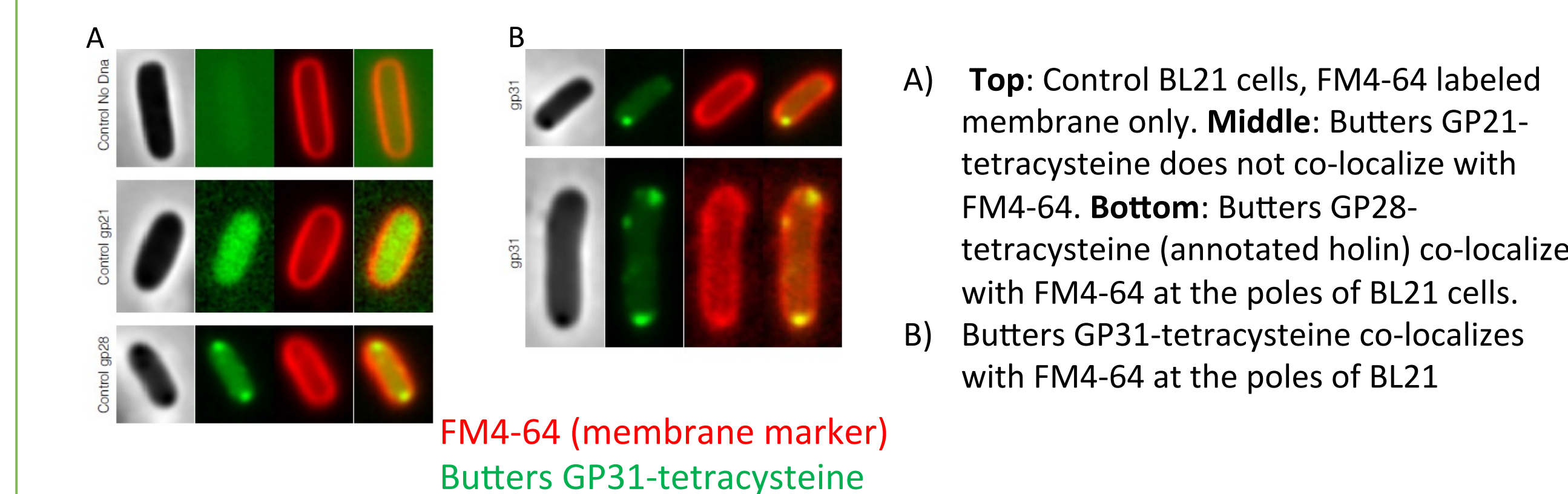
- A snake model of Butters GP31
- The amino acid sequence of Butters GP31 was analyzed for transmembrane regions using TMHMM. GP31 is predicted to be a four (4) pass transmembrane protein.

Experimental Approach for Imaging Butters GP31

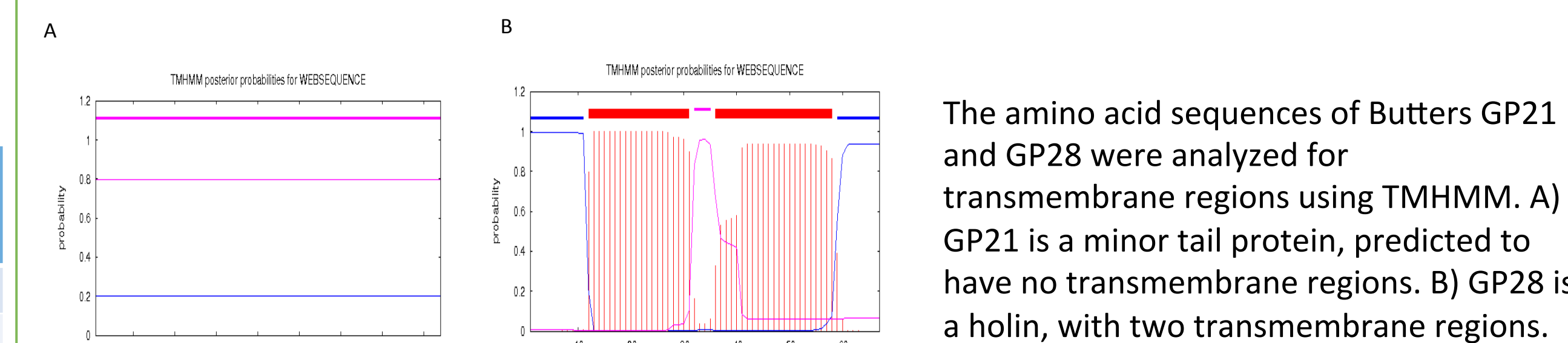


The protein of interest (GP31) is tagged with the tetracycline tag. The BL21 cells are allowed to amplify. The FIAsh-EDT₂ is added. Cells expressing GP31-tetracycline will fluoresce, while those without will not.

Mycobacteriophage Butters GP31-tetracycline localizes to the BL21 *E. coli* membranes primarily at the poles



- Top:** Control BL21 cells, FM4-64 labeled membrane only. **Middle:** Butters GP21-tetracycline does not co-localize with FM4-64. **Bottom:** Butters GP28-tetracycline (annotated holin) co-localizes with FM4-64 at the poles of BL21 cells.
- Butters GP31-tetracycline co-localizes with FM4-64 at the poles of BL21



The amino acid sequences of Butters GP21 and GP28 were analyzed for transmembrane regions using TMHMM. A) GP21 is a minor tail protein, predicted to have no transmembrane regions. B) GP28 is a holin, with two transmembrane regions.

Future Work:

- Create *M. smegmatis* strain(s) expressing Butters GP31 and GP30-31
- Determine how the incorporation of GP31 into the host membrane alters host membrane physiology, susceptibility to phage infection, susceptibility to antibiotics
- Investigate if Butters GP30 and GP31 interact using differentially tagged proteins and fluorescence microscopy

ACKNOWLEDGMENTS: HHMI and Lehigh's Department of Biological Sciences for the financial support; HHMI SEA Staff for infrastructure assistance; The Hatfull lab including Welkin Pope for QC'ing the Butters' genome; Maria Bracia for SEA Program support; Lee Graham for laboratory support and EM imaging; Bryan Berger for Butters GP31 snake model; Patricia Morales for additional computational analysis.

REFERENCES:
 Cresawn SG, Bogel M, Day N, Jacobs-Sera D, Hendrix RW, Hatfull GF. (2011). Phamerator: a bioinformatics tool for comparative bacteriophage genomes. BMC bioinformatics 12:395.
 Lobočka, Malgorzata B.; Debra J. Rose; Guy Plunkett; Marek Rusin; Arkadiusz Samojedny; Hansjörg Lehnher; Michael B. Yarmolinsky; Frederick R. Blattner (November 2004). "Genome of Bacteriophage P1". Journal of Bacteriology 186 (21): 7032–7068.
 Molecular Probes. (2011). TC-FIAsh TC-ReAsH III-cell tetracycline tag detection kits. Eugene, OR: Invitrogen.
 Wu, L., Huany T., Yang L., Pan J., Zhu S., Yan H. (2011). Sensitive and selective bacteriophage detection using tetracycline-tagged phages in conjunction with biarsenic dye. Angewandte Chemie International Edition; 50(26):5873-5877.
 http://phagesdb.org/media/workflow/protocols/pdfs/LysogenyProtocol_3.19.13.pdf
 www.cbs.dtu.dk/services/TMHMM/