

What's New for School Year 2016-17 in Phage Genome Annotation

Created by DJS & WHP, November 2016.

The purpose of this document is to highlight the changes and/or updates to DNA Master Annotation Guide and the Annotation Process. Happy Annotating!

We are beginning an overhaul of the entire Guide, replacing it with individual chapters. Software changes are still in the works, so a complete comprehensive manual update is not available at this time. We will be working on it throughout 2017 so look for updates on the Faculty Information page at seaphages.org.

The overall workflow of The DNA Master Annotation Guide is the same, however, some of the screenshots are out-of-date as the menus now have additional tabs or windows. DNA Master has been updated in November, 2016. As of November 2016, DNA Master is at version 5.23.0 Build 2487 15 Nov 2016.

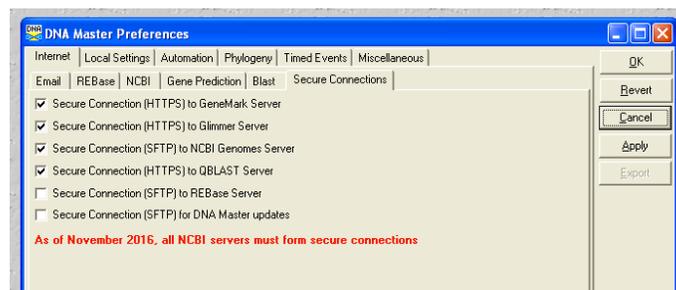
Note: The build used for the current annotation guide was Version 5.22.19 Build 2448, dated 13 June 2014. You can determine the current version of your installation of the program by going to Help -> About.

Changes

DNA Master

The reason for the latest update to DNA Master was an NCBI change that required that all of their servers to be encrypted. This caused an inability of DNA to communicate (both Auto-Annotation and BLAST) with NCBI.

You will find a new window in Preferences that addresses Secure Connections. Shown are the default (correct) settings.



Once you update DNA Master, upon restarting you will receive this message (just once):



If you find that you cannot Auto-Annotate or Blast, please try your update again. Upon updating, if you receive a message that you are up-to-date (but no updating occurred), check your build number. If it isn't Build 2487 (or later), you have not connected to update properly. Please contact us if you have any additional DNA Master difficulties.

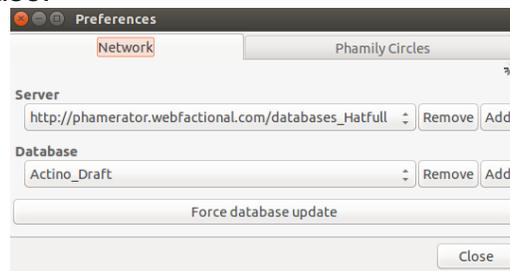
Software Installation

There is a new Ubuntu SEAVM for 2017. Please use this virtual machine as it has an updated 64-bit operating system and updated SEA programs. It will work best in the latest version of Virtual Box (as of November 2016). We are best prepared to troubleshoot this machine (rather than last year's), so please upgrade!

Phamerator

In addition to the Ubuntu-version of Phamerator, Steve Cresawn has a web version available at phamerator.org. At this time, the web-version has many, but not all, of the capabilities of the Ubuntu-Phamerator. Steve plans to continue adding functionality to the web-based version throughout the semester, so check back often.

Phamerator will be updated systematically. If you cannot find your actinobacteriophage genome in Phamerator, make sure you have updated. If you still can't find your actinobacteriophage genome in Phamerator, check that you have it pointing to the correct server and database:



tRNA Investigations

The Mobyl Pasteur Website that hosts tRNAscan-SE will be turned off on December 30, 2016. We are using this opportunity to upgrade our tRNA finding software, and we now endorse tRNAscan-SE v2.0 found at <http://trna.ucsc.edu/tRNAscan-SE/>

tRNAscan-SE 2.0 is more straightforward to use than the old version, as there are fewer options available than at the Pasteur Moby site. However, the interpretation of the results is a bit more involved. You will find the updated protocol for tRNA analysis at the Faculty Information page on seaphages.org.

Note: The settings and interpretations remain unchanged for the Aragorn site data.

Starterator

Chris Shaffer (of Washington University at St. Louis) is now managing Starterator and its output. Changes to Starterator include:

- Changes to the graphic display on page 1. Now, each track is labeled with a phage name and gene number.
- Instead of displaying all of the current start in the Phamerator database in blue, starts that are part of a final annotation are green, and starts that are part of a draft annotation are yellow. All of the rest of the colors are chosen at random.
- Starterator no longer suggests a single specific start. Instead, the text summarizes the findings of the alignment in such a way as to allow you to draw your own conclusion. An updated Starterator Guide is available at seaphages.org.
- All Starterator reports for phams with more than two members in the Actinobacteriophage_Draft database will be available as .pdfs online. See the whole index at: <http://phages.wustl.edu/starterator/>
- Starterator pham reports are linked to the gene pages in PECAAN (Phage Evidence Collection and Annotation Network).

You will find the new Starterator instructions and links to the Starterator guide at Faculty Information Page on seaphages.org.

PECAAN (Phage Evidence Collection and Annotation Network)

Claire Rinehart, from Western Kentucky University, introduced us to PECAAN at last summer's Retreat. PECAAN is a companion program to DNA Master (*not a replacement*), streamlining the complicated decision-making process of Gene Prediction by displaying relevant evidence for each gene on its own web page. It also provides a record-keeping system that can be used as input to create your final DNA Master files. Using PECAAN is optional. You may want to add PECAAN into your students' workflow once they get the hang of using DNA Master and evaluating the data from the various data sources. Or you may want to use it as a tool for the final review of your genomes because it stores BLAST (PhagesDB and NCBI), HHPRED, and Starterator data for the entire genome and can output this data into DNA Master as a functional file or as a fully annotated file.

To gain access to PECAAN, instructors should email claire.rinehart@wku.edu a request along with:

Firstname, MI, Lastname

Email

Full institution name

Approximate number of students in their group.

The PECAAN user manual is available on seaphages.org on the Faculty Information page. The user manual contains instructions on how to load phages into PECAAN and how to add your students as users of the site.

Official List of Functions

By far, the most corrected part of submitted annotations is the functions. To address this:

- 1) Please check your functional annotation against recently annotated phages of the same cluster to ensure that you have found the majority of the functions.
- 2) Please use the Official SEA-PHAGES Function list to ensure the function names you are using are consistent with the rest of the annotation across the SEA. If you find a function that is not on the list, please email Welkin or Debbie prior to including it in an annotation. We approve and add new functions to the list all the time, so always look for the newest version online on the Faculty Information page at seaphages.org.

Debbie's Recommendations for 2017

Comparative analyses

A most valuable component of genome annotations is the comparative analyses. Two scenarios come to mind:

- Annotation of a Cluster A4 genome, where there are 89 cluster A4 genomes. It will be difficult to review all 89 pieces of data, but don't ignore all of it. Starterator and Phamerator graphic displays may help you to make sense of the lists of blast hits that you get from PhagesDB or NCBI.
- Annotation of your first Microbacterium phage may be daunting because your phage is one of the first discovered on that host. Do not ignore or dismiss Draft annotations to uncover relationships with other phage proteins. You may need to evaluate all of the Draft-Annotated members of a particular Pham as you make the decision for your phage gene.

Happy Annotating!