BLASTable Mycobacteriophage Database

Purpose: To have easy access to current known protein function through use of a DNA Master created protein database.

Background: When determining protein function, a variety of tools including BLASTP data, Phamerator maps, conserved domain identification, and HHpred are used. This is another tool to help. This database was created in DNA Master (procedure to follow) using 243 of the most recent mycobacteriophage genomes. This database was created December 2012.

Overview: This protocol consists of three parts:

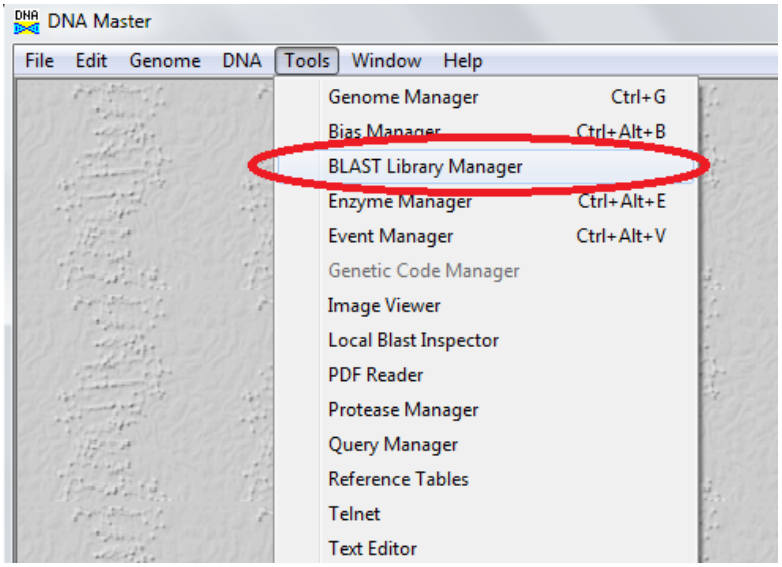
- **Creation of local database** (supplied, see addendum for procedure)
- **Import of database**
- **Local Blast of database**

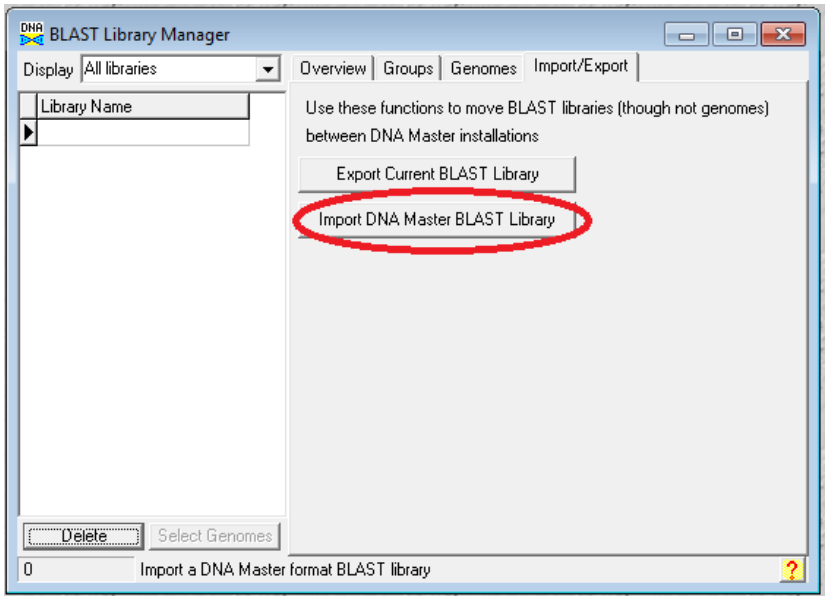
Procedure:

Import of Database

The library of completed mycobacteriophages with functions is currently a 3.41 MB file named Proteins_243MP_12.2012.dnamblib. This is a static file that will be periodically updated.

1. Get Proteins_243MP_12.2012.dnamblib from the PhagesDB website. Click on Resources→ Documents →Annotation Resources. Save this file to a known location.
2. Go to Tools → BLAST Library Manager



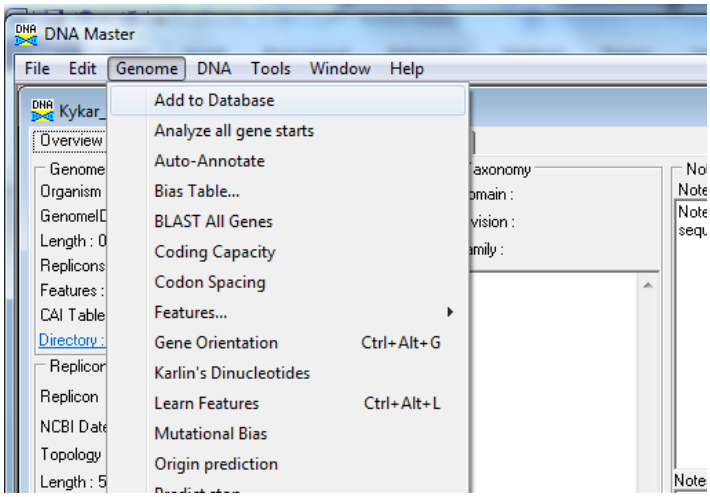


3. Click the Import/Export tab
4. Click the Import DNA Master BLAST Library button
5. Browse to the where the .dnamblib file is saved and select it.

You have now installed the Blast Library of proteins of 243 genomes. The next step is to blast your genome against these.

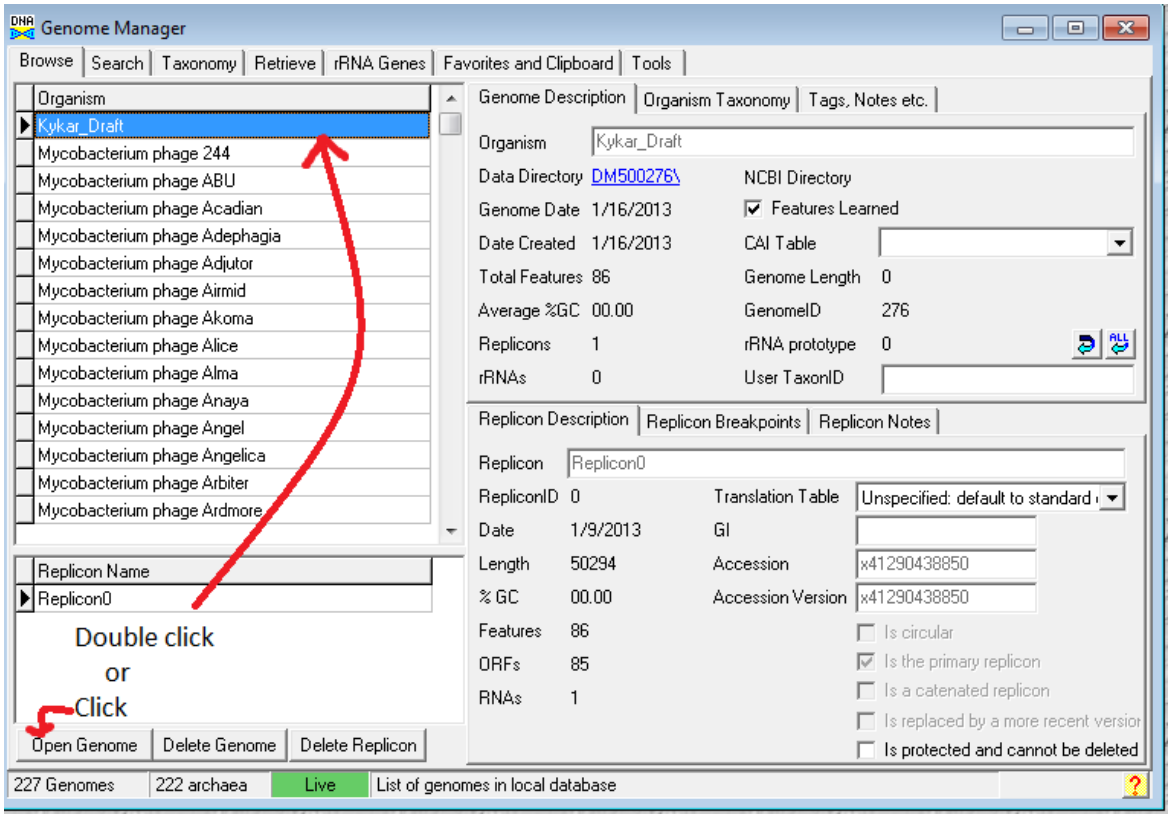
Local Blast of Database

1. The genome must be added to the Genome Manager. To do this, open your genome in DNA Master. Click on Main menu→Genome→Add to Database. You can now close your genome file.

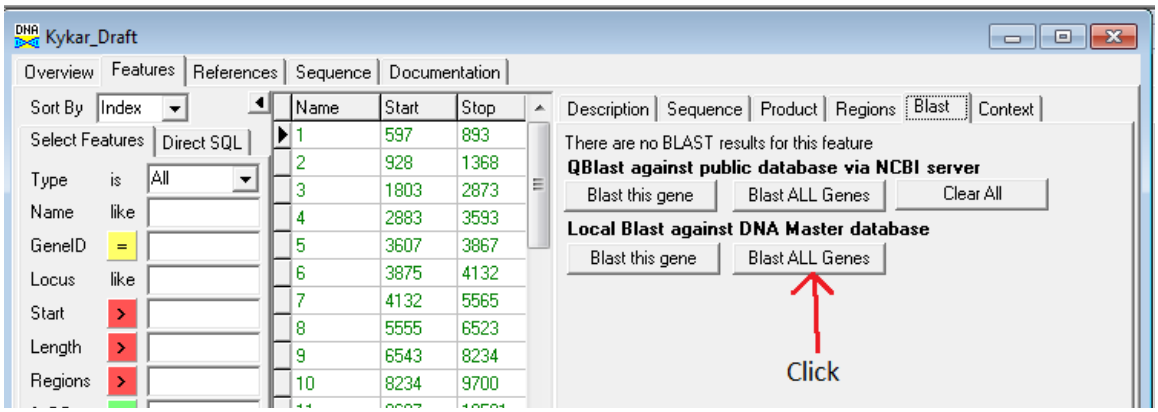


2. Now the sequence must be opened from the Genome Manager.
3. To open your file in Genome Manager: click on Tools→ Genome Manager

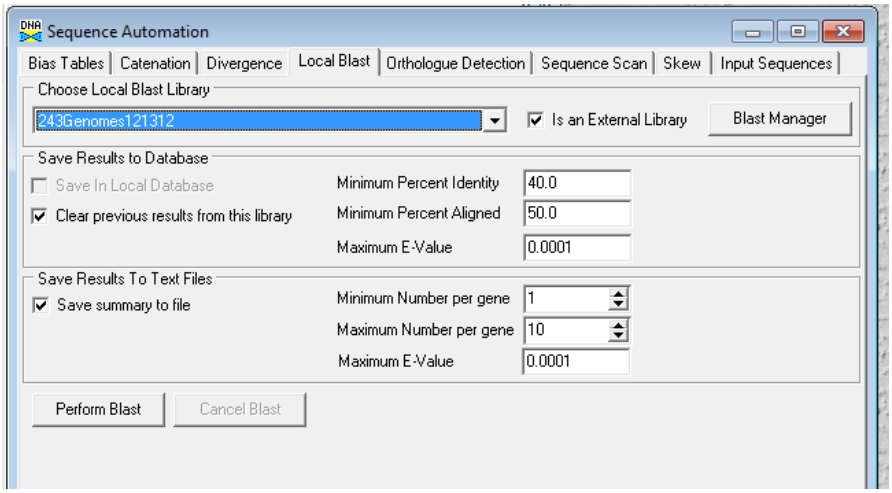
4. Double click on your genome to open the file. To blast this genome, it must be in the Genome Manager.



5. Now in this genome, go to the Feature tables and find a gene that has no blast data.
6. Click on the Blast tab and the following window appears.



7. Find Local Blast against DNA Master database and select the button Blast ALL Genes.
8. The Sequence Automation window pops up. Choose the following features.



- Choose the Local Blast library: 243Genomes121312.
- Under Save Results to Database: Use default settings.
- Under Save Results to Text Files: Checkmark the "Save summary to file" option. Use default settings for the remaining features.
- Click the Perform Blast button and choose where to save the output file.

9. The output file is a .csv file that can be opened in Excel.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Idx	QueryIdx	TargetIdx	QueryCoo	QueryN	QueryLoc	TargetOrganism	TargetC	TargetGer	TargetLoc	TargetProduct	PercentId	Alignmen	PercentAl	EValue
2	1	1	1	597 - 893	1	DNAM1	Mycobacterium_phage_Solon	821	1	PBI_SOLO	Hypothetical_Protein	98	98	98.99	1.20E-54
3	2	1	2	597 - 893	1	DNAM1	Mycobacterium_phage_KSSJEB	742	1	PBI_KSSJE	WhiB	95.9	98	98.99	3.50E-54
4	3	1	3	597 - 893	1	DNAM1	Mycobacterium_phage_MrGordo	762	1	PBI_MRGC	HNH_endonuclease	97.9	97	97.98	6.00E-54
5	4	1	4	597 - 893	1	DNAM1	Mycobacterium_phage_Doom	693	1	PBI_DOOM	HNH_endonuclease	96.9	98	98.99	6.00E-54
6	5	1	5	597 - 893	1	DNAM1	Mycobacterium_phage_Museum	764	3	PBI_MUSE	WhiB	96.9	98	98.99	1.30E-53
7	6	1	6	597 - 893	1	DNAM1	Mycobacterium_phage_JC27	732	2	PBI_JC27_	HNH-like_domain	95.9	98	98.99	1.30E-53
8	7	1	7	597 - 893	1	DNAM1	PattyP_ED	855	2	PBI_PATT	HNH_endonuclease_domain	95.9	98	98.99	1.70E-53
9	8	1	8	597 - 893	1	DNAM1	Mycobacterium_phage_Violet	843	2	PBI_VIOLE	HNH_endonuclease_domain	94.9	98	98.99	3.90E-53
10	9	1	9	597 - 893	1	DNAM1	Mycobacterium_phage_RidgeCB	803	2	PBI_RIDGE	HNH_endonuclease_domain	95.9	98	98.99	3.90E-53
11	10	1	10	597 - 893	1	DNAM1	Mycobacterium_phage_Perseus	781	1	PBI_PERSE	HNH_endonuclease_domain	95.9	98	98.99	3.90E-53
12	11	2	1	928 - 1368	2	DNAM11	Mycobacterium_phage_Violet	843	3	PBI_VIOLE	Hypothetical_Protein	100	146	99.32	3.20E-80
13	12	2	2	928 - 1368	2	DNAM11	Mycobacterium_phage_RidgeCB	803	3	PBI_RIDGE	Hypothetical_Protein	100	146	99.32	3.20E-80
14	13	2	3	928 - 1368	2	DNAM11	Mycobacterium_phage_Museum	764	4	PBI_MUSE	Hypothetical_Protein	100	146	99.32	3.20E-80
15	14	2	4	928 - 1368	2	DNAM11	Mycobacterium_phage_MrGordo	762	3	PBI_MRGC	tail_protein	100	146	99.32	3.20E-80
16	15	2	5	928 - 1368	2	DNAM11	Mycobacterium_phage_KSSJEB	742	2	PBI_KSSJE	Hypothetical_Protein	100	146	99.32	3.20E-80
17	16	2	6	928 - 1368	2	DNAM11	Mycobacterium_phage_Doom	693	3	PBI_DOOM	Hypothetical_Protein	100	146	99.32	3.20E-80
18	17	2	7	928 - 1368	2	DNAM11	Mycobacterium_phage_Bxb1	667	2	PBI_BXB1_	related_to_L5_gp5_(41%);_16.3_	100	146	99.32	3.20E-80
19	18	2	8	928 - 1368	2	DNAM11	Mycobacterium_phage_Lesedi	747	2	PBI_LESED	tail	99.3	146	99.32	1.20E-79
20	19	2	9	928 - 1368	2	DNAM11	Mycobacterium_phage_Lesedi	747	2	PBI_LESED	tail	99.3	146	99.32	1.20E-79
21	20	2	10	928 - 1368	2	DNAM11	Mycobacterium_phage_Solon	821	2	PBI_SOLO	Hypothetical_Protein	98.6	146	99.32	1.30E-78
22	21	3	1	1803 - 287	3	DNAM21	Mycobacterium_phage_Perseus	781	6	PBI_PERSE	Hypothetical_Protein	98.3	356	99.72	0.00E+00
23	22	3	2	1803 - 287	3	DNAM21	Mycobacterium_phage_Lockley	754	5	PBI_Lockl	tail_fiber-like_protein	98	356	99.72	0.00E+00
24	23	3	3	1803 - 287	3	DNAM21	Mycobacterium_phage_Lesedi	747	4	PBI_LESED	Hypothetical_Protein	98	356	99.72	0.00E+00
25	24	3	4	1803 - 287	3	DNAM21	Mycobacterium_phage_Lesedi	747	4	PBI_LESED	Hypothetical_Protein	98	356	99.72	0.00E+00
26	25	3	5	1803 - 287	3	DNAM21	Mycobacterium_phage_JC27	732	6	PBI_JC27	tail	97.5	356	99.72	0.00E+00

Kykar_Draft

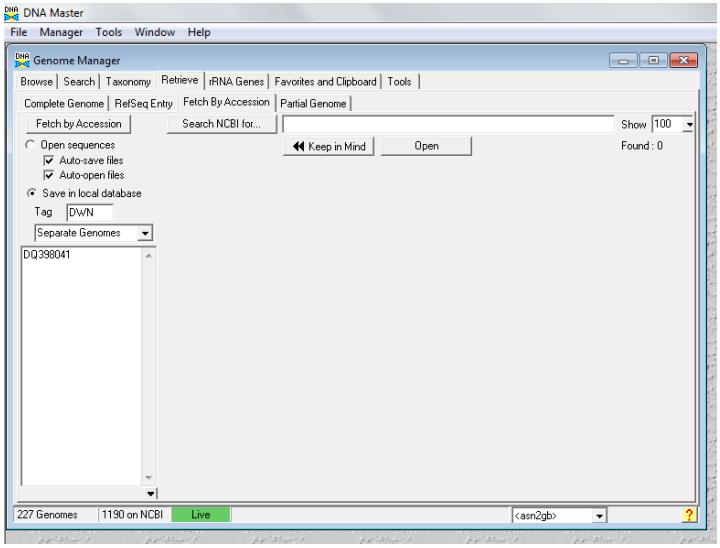
Notes:

- This is a static database created December 2012 with 243 mycobacteriophage genomes.
- At present, you can only retrieve the best 10 hits.
- Functions are present in this database as called by Phage Hunters. Information obtained should still undergo evaluation.

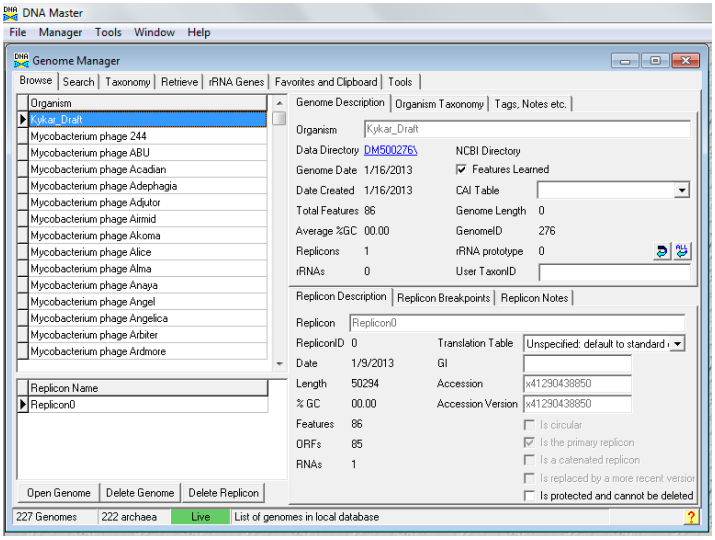
Addendum:

Creation of Local Database

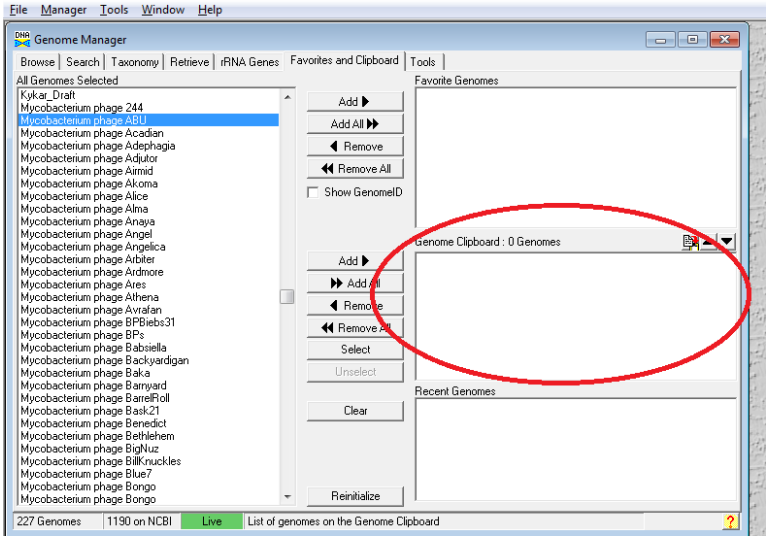
1. To create a local database, go to Tools→Genome Manager.
2. Select the genomes you wish to place in the library by importing desired genomes into your version of DNA Master Genome Manager. Any genome with a DNA Master file or one found in GenBank can be added. To add a DNA Master file, see first step of Local Blast of Database. To add a file from GenBank, go to the Retrieve tab in Genome Manager. From the submenu there, select the Fetch by Accession tab. If you know the accession number(s), place that number(s) in the window on the left side of the window. Once your list is complete, click Save in local database.



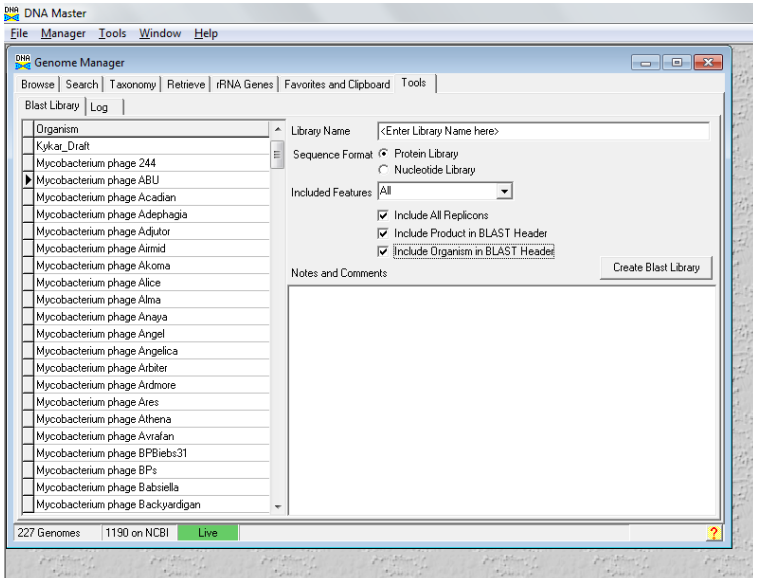
3. Click on the Browse tab. This will populate your Genome Manager with as many genomes as you choose.



4. Go to Favorites and Clipboard tab and add all genomes of interest to the clipboard (middle window).



5. Go to the Tools tab→ Blast Library
6. Enter a name of the library, choose a protein library, specify that the Products and Organisms by included in the BLAST header.



7. Create the library
8. This will create a .dnamblib file with the name you chose in step 5.
9. Proceed to Import of Database.