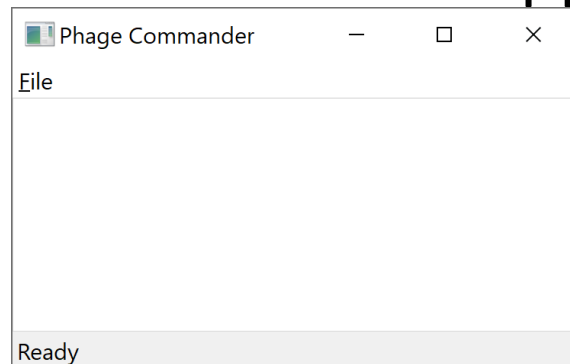


# Introduction to Phage Commander

- **Phage Commander** is a software tool for identifying genes in phage genomes
- Phage Commander runs a phage's DNA sequence through gene identification tools and outputs a list of potential genes. These tools include:
  - Glimmer, Genemark, Genemark.hmm, Genemark S, Genemark S2, Genemark Heuristic, Prodigal, RAST, Metagene, and Aragorn (for tRNA genes)
- Phage Commander's output can be exported in Excel format (.xlsx) or NCBI GenBank format (.gb)
- A draft paper describing Phage Commander in detail is here:  
<https://www.biorxiv.org/content/10.1101/2020.11.11.378802v1>

# How to run Phage Commander

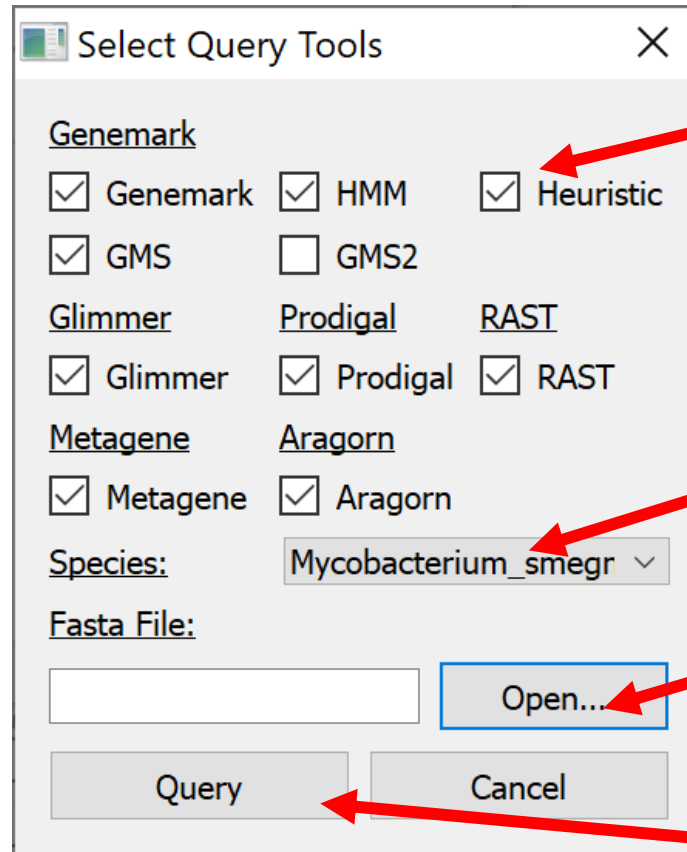
- Download Phage Commander from GitHub repository:  
<https://github.com/sarah-harris/PhageCommander>
- Run Phage Commander:
  - **Windows:** Navigate to phagecommander/bin. Click on the executable: phagecom-windows.exe
  - **Linux or Mac:** Open a shell. Navigate to folder where you have downloaded phagecommander. Type `py phagecom.py`. If you need to install any of the supporting packages, first type: `'pip install .'` (in the PhageCommander-master directory – the directory that contains setup.py)
- A small Phage Commander window will appear:



# How to use Phage Commander

- In the Phage Commander Window, click on the **File** menu, and select **New**
- A Select Query Tools window will open (see next slide). Select which tools you would like to use
  - If using **GeneMark.hmm** you will need to select your phage's bacterial host from the drop down menu under **Species**
  - To use **RAST**, you will need to create a login and password at <https://rast.nmpdr.org>
- Press **Open** and select your phage's fasta file - an example fasta file is included in the GitHub repository: Patience.fasta
- Press **Query** to run Phage Commander. Phage Commander will now run the phage genome through the selected tools – this may take several minutes
- When Phage Commander completes, press **OK**

# Select Query Tools Window



The screenshot shows a window titled "Select Query Tools" with a close button (X) in the top right corner. The window contains several sections of tool selection:

- Genemark**
  - ☒ Genemark
  - ☒ HMM
  - ☒ Heuristic
  - ☒ GMS
  - ☐ GMS2
- Glimmer**
  - ☒ Glimmer
  - ☒ Prodigal
  - ☒ RAST
- Metagene**
  - ☒ Metagene
  - ☒ Aragorn

Below the tool sections, there is a "Species:" label and a dropdown menu showing "Mycobacterium\_smegr". Below that is a "Fasta File:" label and an empty text box. To the right of the text box is an "Open..." button. At the bottom of the window are two buttons: "Query" and "Cancel".

Red arrows point from the following text to the interface:

- "Check the boxes of the tools you want to use" points to the checked boxes for Genemark, HMM, Heuristic, GMS, and Glimmer.
- "Select host from drop down menu if 'HMM' box is checked (GeneMark.hmm)" points to the "Species:" dropdown menu.
- "Press open select phage fasta file" points to the "Open..." button.
- "Press 'Query'" points to the "Query" button.

Check the boxes  
of the tools you  
want to use

Select host from  
drop down menu if  
"HMM" box is  
checked  
(GeneMark.hmm)

Press open select  
phage fasta file

Press "Query"

# In the Event of an Error

- If Phage Commander runs successfully, you will see a window saying **“Done! Query Successful”**
- Press **OK** and you will see the Phage Commander output in the form of a spreadsheet (expand the window)
- If phage commander encounters an error, the error message will display which program is causing the error (e.g. Glimmer, Aragorn, GMS2, etc.)
- To work around this, re-run Phage Commander without including the program causing the error

Total Calls column  
Is the number of  
programs identifying  
a gene

“GM” is the genes  
identified by  
GeneMark

“HMM” is the  
genes identified by  
GeneMark.hmm

“Heuristic” is the  
genes identified by  
GeneMark Heuristic

Genes																			
TRNA																			
TOTAL CALLS	ALL	ONE	GM	GM	GM	GM		HMM	HMM	HMM	HMM		HEURISTIC	HEURISTIC	HEURISTIC	HEURISTIC		GMS	
2								+	2	94	93								
8	X		+	138	431	294		+	138	431	294		+	138	431	294		+	
8	X		+	616	2076	1461		+	616	2076	1461		+	616	2076	1461		+	
8	X		+	2135	3547	1413		+	2093	3547	1455		+	2135	3547	1413		+	
8	X		+	3538	4875	1338		+	3538	4875	1338		+	3538	4875	1338		+	
8	X		+	4879	5244	366		+	4879	5244	366		+	4879	5244	366		+	
8	X		+	5281	6225	945		+	5281	6225	945		+	5281	6225	945		+	
8	X		+	6225	6575	351		+	6225	6575	351		+	6225	6575	351		+	
8	X		+	6579	6941	363		+	6579	6941	363		+	6579	6941	363		+	
8	X		+	6941	7291	351		+	6941	7291	351		+	6941	7291	351		+	
8	X		+	7307	7654	348		+	7307	7654	348		+	7307	7654	348		+	
8	X		+	7654	7923	270		+	7654	7923	270		+	7654	7923	270		+	
8	X		+	7923	8336	414		+	7923	8336	414		+	7923	8336	414		+	
8	X		+	8333	8770	438		+	8333	8770	438		+	8333	8770	438		+	
8	X		+	8862	9860	999		+	8862	9860	999		+	8862	9860	999		+	
8	X		+	9987	10586	600		+	9987	10586	600		+	9987	10586	600		+	
8	X		+	10687	10905	219		+	10687	10905	219		+	10687	10905	219		+	
8	X		+	10919	14569	3651		+	10919	14569	3651		+	10919	14569	3651		+	
8	X		+	14566	16311	1746		+	14566	16311	1746		+	14566	16311	1746		+	
8	X		+	16311	18047	1737		+	16311	18047	1737		+	16311	18047	1737		+	
8	X		+	18070	18963	894		+	18070	18963	894		+	18070	18963	894		+	
8	X		+	18964	21225	2262		+	18964	21225	2262		+	18964	21225	2262		+	
8	X		+	21225	22292	1068		+	21225	22292	1068		+	21225	22292	1068		+	
8	X		+	22289	22465	177		+	22289	22465	177		+	22289	22465	177		+	
1		X																	
7			+	22477	22839	363		+	22477	22839	363		+	22477	22839	363		+	
1		X																	
8	X		+	22836	24017	1182		+	22839	24017	1179		+	22836	24017	1182		+	
8	X		+	24014	25180	1167		+	24014	25180	1167		+	24014	25180	1167		+	
8	X		+	25202	25450	249		+	25202	25450	249		+	25202	25450	249		+	
8	X		+	25447	25806	360		+	25447	25806	360		+	25447	25806	360		+	

Each row is a gene

Row shading is  
proportional to  
how many  
programs identify a  
gene (darker =  
more programs,  
white = only one  
program)

**“GM” is the genes identified by GeneMark S**

**“Glimmer” is the genes identified by Glimmer**

**“Prodigal” is the genes identified by Prodigal**

**“Metagene” is the genes identified by Metagene**

Phage Commander - C:\Users\Administrator\Dropbox\209X\Hoot Aragorn.gq

Genes	GMS				GLIMMER				PRODIGAL				METAGENE			
	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-
1	+	487	1176	690	+	1089	1925	837	+	1089	1925	837	+	418	1176	759
2	+	1089	1925	837	+	1925	2146	222	+	1925	2146	222	+	1925	2146	222
3	+	1925	2146	222	+	2482	2952	471	+	2482	2952	471	+	2482	2952	471
4	+	2482	2952	471	+	3037	3999	963	+	3037	3999	963	+	3037	3999	963
5	+	3037	3999	963												
6																
7	+	3999	4664	666	+	3999	4664	666	+	3999	4664	666	+	3999	4664	666
8	+	4661	5359	699	+	4661	5359	699	+	4661	5359	699	+	4661	5359	699
9	+	5696	5833	138	+	5696	5833	138	+	5696	5833	138	+	5696	5833	138
10	+	5833	7044	1212	+	5833	7044	1212	+	5833	7044	1212	+	5833	7044	1212
11	+	7044	7475	432	+	7044	7475	432	+	7044	7475	432	+	7044	7475	432
12	+	7512	9266	1755	+	7512	9266	1755	+	7512	9266	1755	+	7512	9266	1755
13	+	9284	10726	1443	+	9284	10726	1443	+	9323	10726	1404	+	9284	10726	1443
14	+	10801	11607	807	+	10801	11607	807	+	10801	11607	807	+	10801	11607	807
15	+	11656	12186	531	+	11656	12186	531	+	11656	12186	531	+	11656	12186	531
16	+	12219	13190	972	+	12219	13190	972	+	12219	13190	972	+	12219	13190	972
17	+	13263	13460	198	+	13263	13460	198	+	13263	13460	198	+	13263	13460	198
18	+	13463	13840	378	+	13463	13840	378	+	13463	13840	378	+	13463	13840	378
19	+	13837	14028	192	+	13837	14028	192	+	13837	14028	192	+	13837	14028	192
20	+	14028	14396	369	+	14028	14396	369	+	14028	14396	369	+	14028	14396	369
21	+	14396	14749	354	+	14396	14749	354	+	14396	14749	354	+	14396	14749	354
22	+	14746	15165	420	+	14746	15165	420	+	14746	15165	420	+	14773	15165	393
23	+	15188	15790	603	+	15188	15790	603	+	15188	15790	603	+	15188	15790	603
24	+	15901	16305	405	+	15901	16305	405	+	15901	16305	405	+	15901	16305	405
25	+	16302	16721	420	+	16302	16721	420	+	16257	16721	465	+	16302	16721	420
26	+	16714	18918	2205	+	16714	18918	2205	+	16714	18918	2205	+	16714	18918	2205

**DNA strand  
(+ or -)**

**Gene start (stop  
if strand “-”)**

**Gene stop (start  
if strand is “-”)**

**Gene length**

tRNA genes  
Identified by Aragorn  
in “TRNA” tab

Phage Commander - C:/Users/Administrator/Dropbox/209X/Hoot Aragorn.gq

File

Genes		TRNA					
	TOTAL CALLS	ALL	ONE	ARAGORN	ARAGORN	ARAGORN	ARAGORN
1	1	X		+	5403	5475	73
2	1	X		+	5480	5553	74
3	1	X		+	5595	5667	73

Total calls is always 1  
since Aragorn is the only  
program for identifying  
tRNA genes



# Exporting Phage Commander Results

- To export as excel spreadsheet, select:
  - **File** → **Export** → **Excel**
- To export as GenBank (.gb) format file, select:
  - **File** → **Export** → **GenBank**
  - Set the threshold number of programs for exporting genes
  - “Less than or equal to” will export those genes identified by an equal or lower number of programs than the threshold (e.g. genes identified by 3 or fewer programs. Set the threshold to maximum to export all genes identified)
  - “Greater than” will export those genes identified by a number of programs greater than the threshold (e.g. genes identified by more than 2 programs. Use “0” as the threshold and this setting to export all genes identified)
  - Press “Save as” and enter desired filename
  - Press “Export”

# Exporting Phage Commander results in .gb format

User-set threshold  
number of programs  
For gene exporting

Export genes identified  
by less than or equal the  
set number of programs

Export genes identified  
by more than the set  
number of programs

Export genes identified  
only by set number of  
programs

51		+	6225	6575	351
53		+	6579	6941	363
51		+	6941	7291	351
48		+	7307	7654	348
70				7922	270
14				8336	414
50				8770	438
99				9860	999
00				10586	600
19				10905	219
51				14569	3651
46		+	14566	16311	1746
37		+	16311	18047	1737
94		+	18070	18963	894
62		+	18964	21225	2252
68		+	21225	22292	1068

Export to Genbank

Select genes with:

8 Calls

☒ Less than or equal to ☐ ALL

☐ Greater than ☐ NONE

☐ Exactly

Save as...

Export Cancel

Export only those  
genes identified by  
all programs (max)

Export genes  
identified by  
only one program

Press "Save as"  
and enter desired  
filename

Press "Export"