



OAT

Orthologous Average Nucleotide Identity Tool
a similarity measurement tool for genomes

User Manual

version 0.90

July 10 2015

Laboratory of Evolutionary Bioinformatics
Seoul National University

CONTENTS

1.	General Information	2
1.1	Purpose of the Application	2
1.2	Developers	2
1.3	Overview	3
2.	Configuring	4
2.1	Switching the Blast Program	4
2.2	Multithreading	6
3.	Preparing the Calculation Step	7
3.1	Adding Genome Data	7
3.2	Removing Genome Data	8
3.3	Run	9
4.	Calculation Step	10
4.1	Calculation Step	10
4.2	Result screen	11
4.2.1	Table View	11
4.2.2	Matrix View	12
4.2.3	Heatmap	13
5.	Support	13

1. General Information

1.1 Purpose of the Application

OAT (**O**rthologous **A**verage Nucleotide Identity **T**ool) uses OrthoANI to measure overall similarity between two genome sequences. Unlike the original ANI algorithm, it produces almost identical reciprocal similarities. It has been shown by a large comparison study, values generated by the original ANI and OrthoANI are comparable. The proposed cut-off for species demarcation is 95~96% for both OrthoANI and the original ANI. The detailed algorithm is given in Lee *et al.*

OAT has an easy to follow Graphical User Interface that allow researchers to calculate OrthoANI values between genomes of their interest without the need of dealing with an unfamiliar Command Line Environment.

1.2 Developers

OAT was developed by members of the Laboratory of Evolutionary Bioinformatics at Seoul National University.

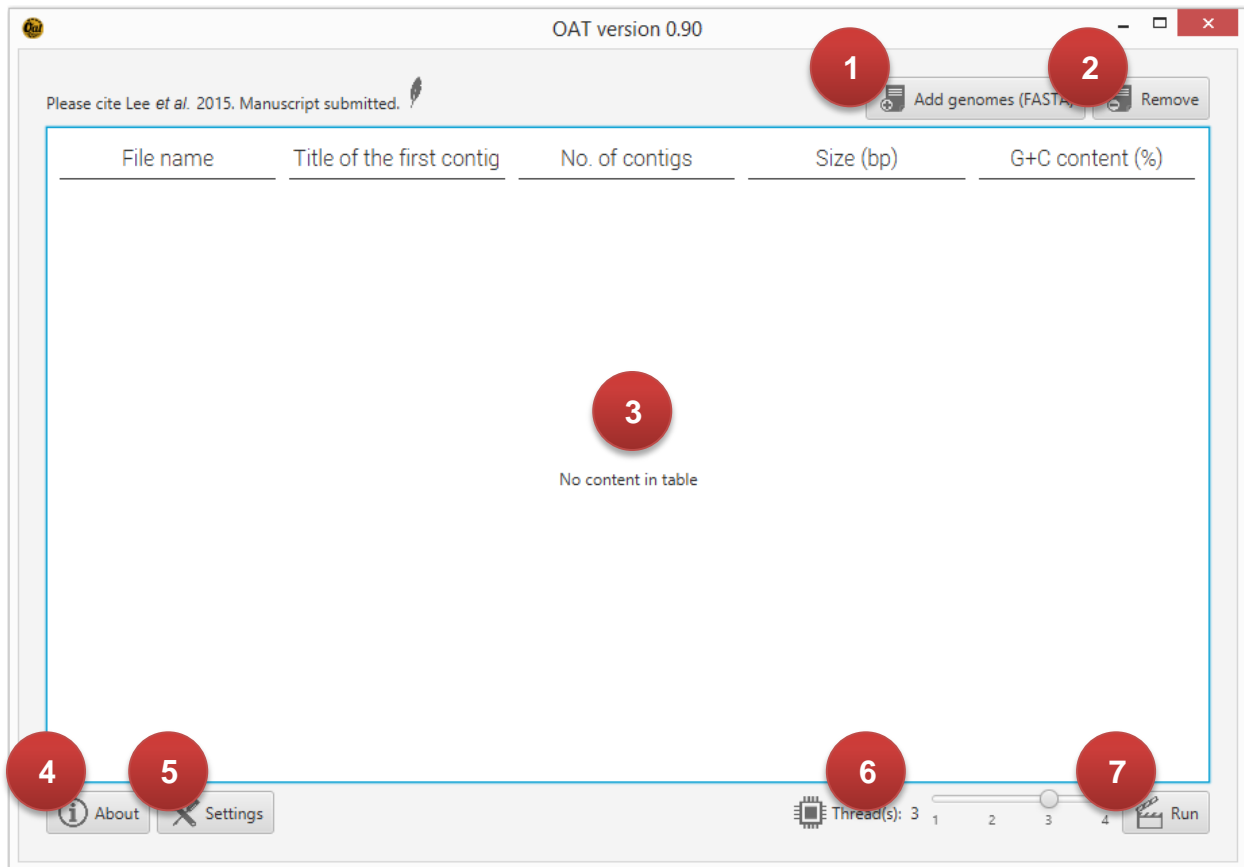
Yeong Ouk Kim, Seoul National Univ.

Imchang Lee, Seoul National Univ.

Sang-Cheol Park, Seoul National Univ.

Jongsik Chun, Seoul National Univ. & ChunLab, Inc.

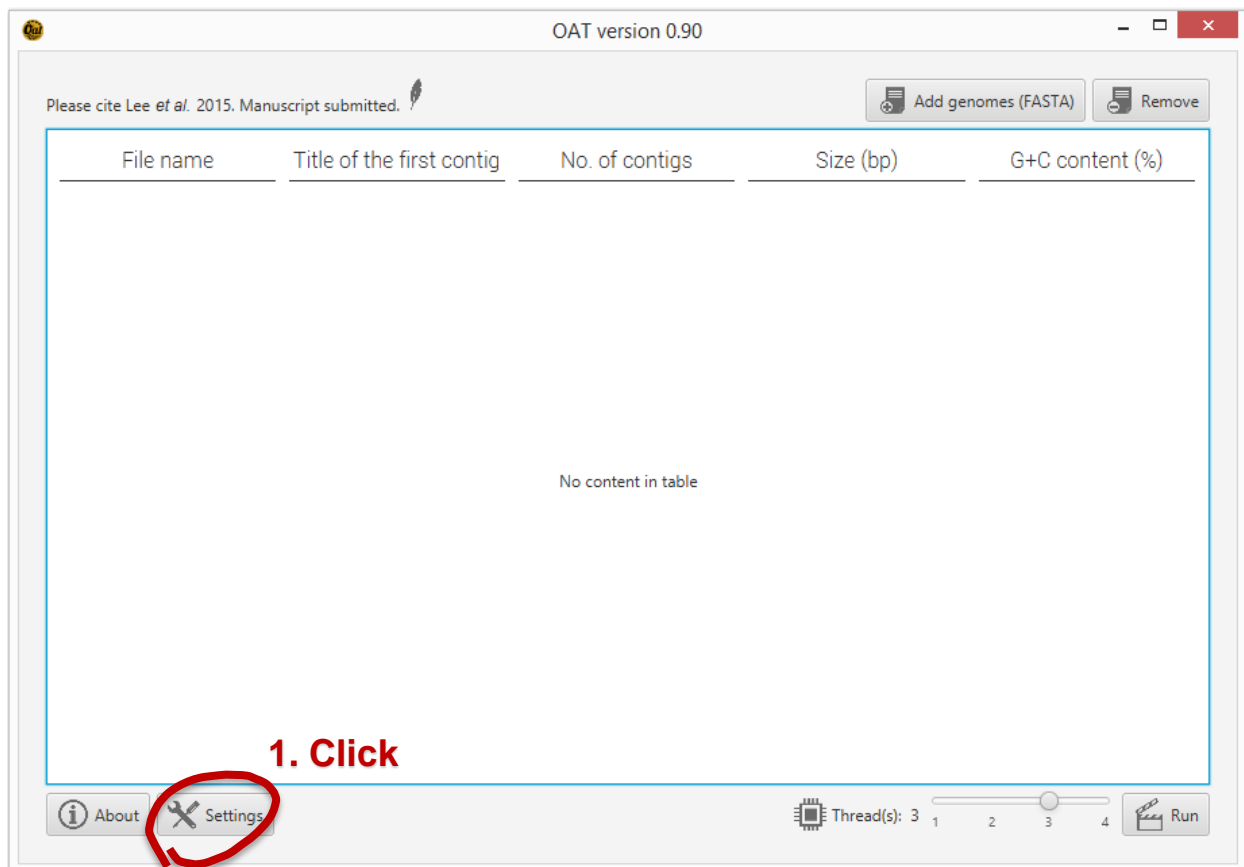
1.3 Overview



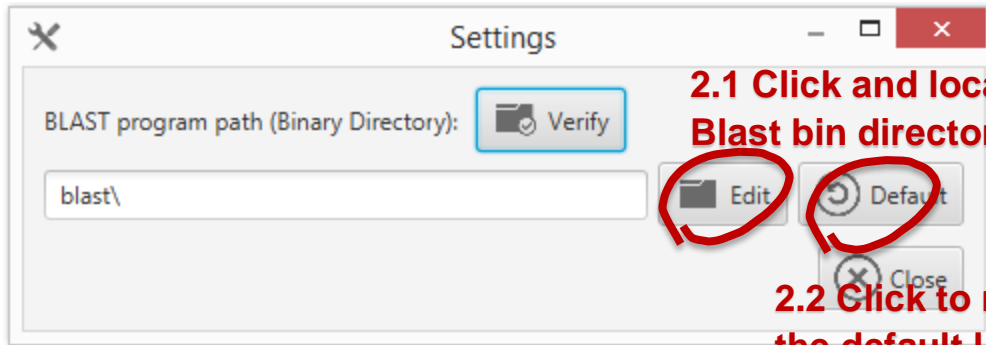
2. Configuration

2.1 Switching the Blast Program

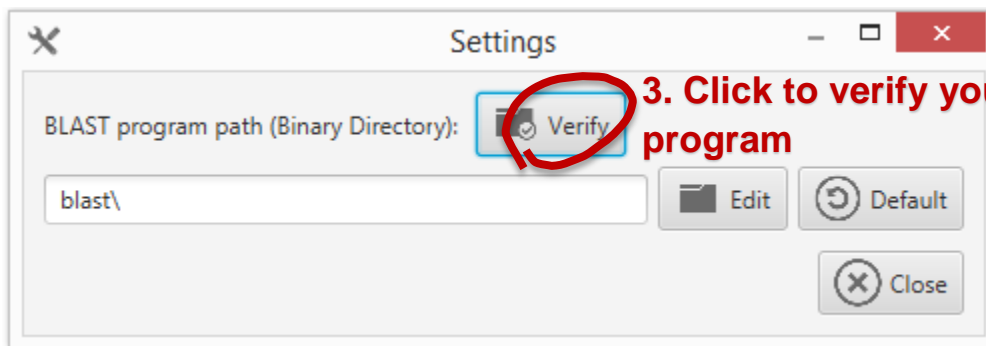
If you wish to use a different Blast version or if you downloaded the Runnable Jar version of the OAT application (which does not include a Blast program), follow the instructions provided below. We recommend users to stick with the provided version of Blast (ncbi-blast-2.2.30+) or a higher version since our application was tested with that version and versions lower than it may potentially cause malfunctions. If you are using the Runnable Jar version of the OAT application you can find the appropriate Blast program for your OS from their website ([link](#)).



1. Click the “Settings” button



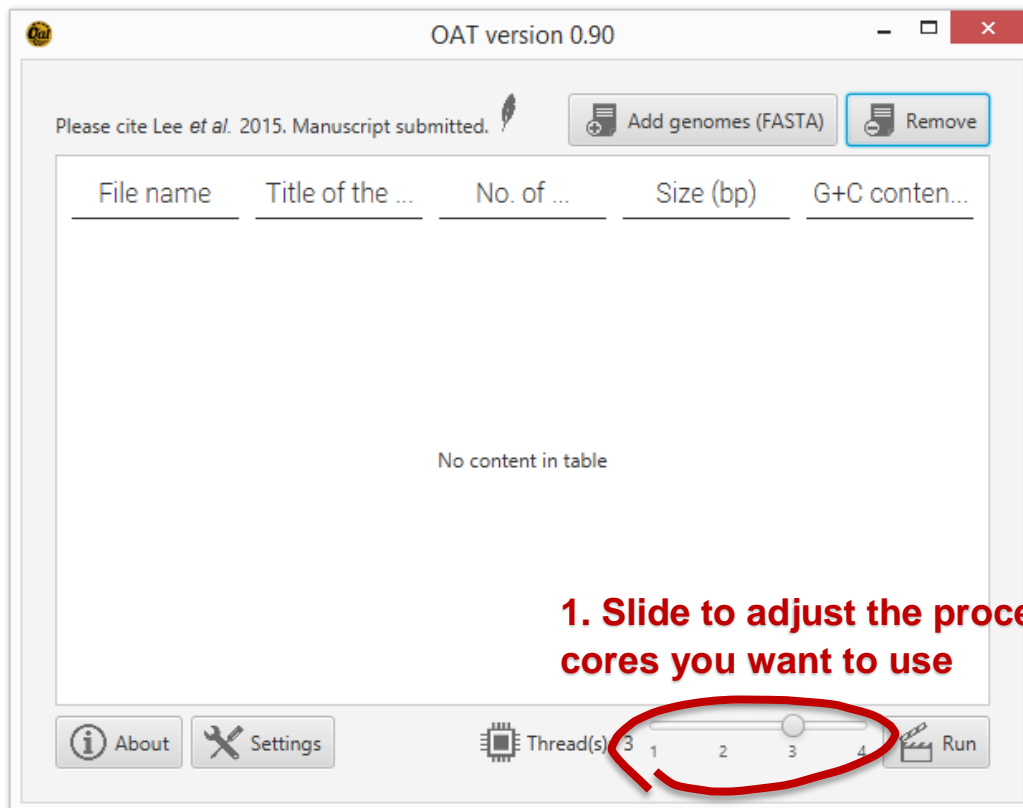
2. Set the Blast program bin directory by...
 - 2.1 Click the “Edit” button to locate your Blast program’s bin directory
 - 2.2 Click the “Default” button to set the Blast program’s path to the default location



3. Verify your blast program by clicking the “Verify” button

2.2 Multithreading

If you more than 1 processing cores on your computer you may increase the number of threads the application uses to speed up the calculation process. If you are running other programs on your computer you should avoid setting the threads equal to the number of available processing cores since it will hinder the performance of the other programs.

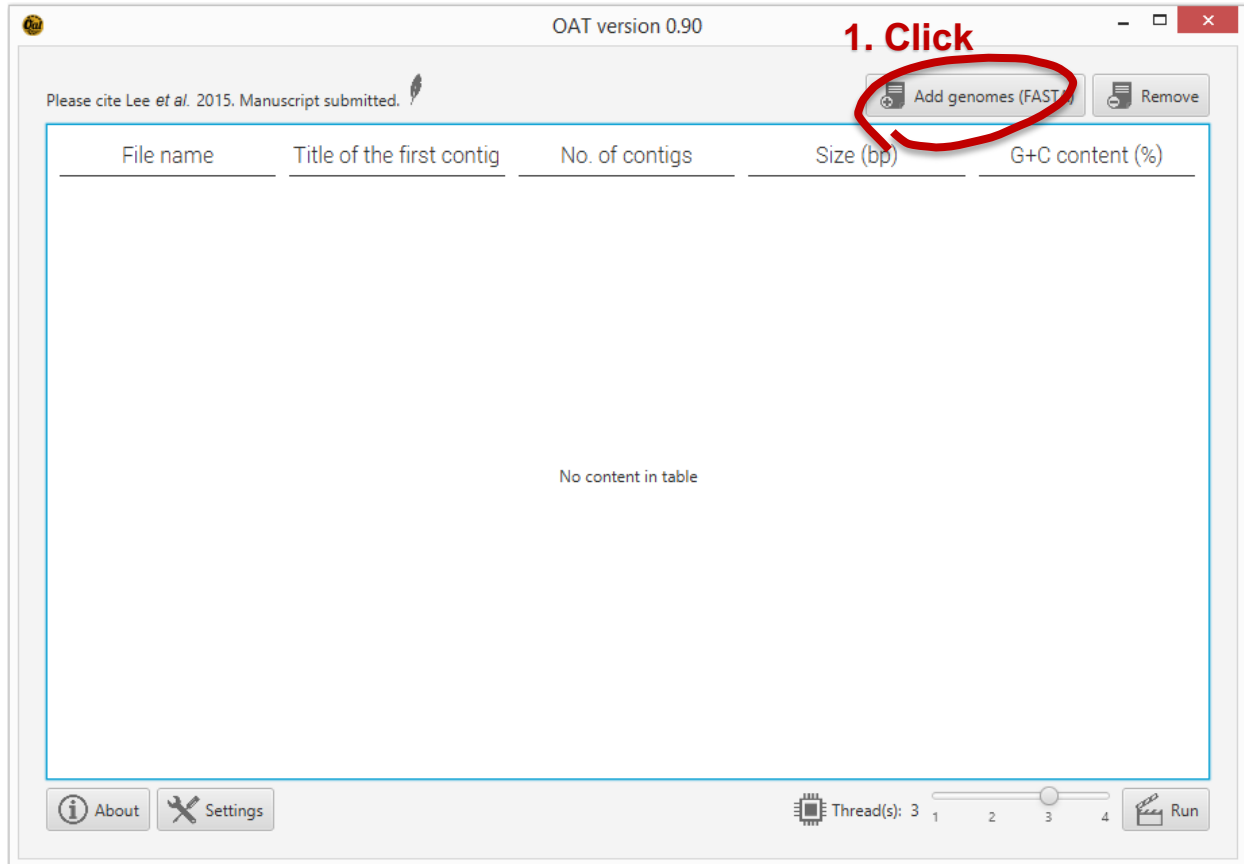


1. Use the slider to adjust the number of threads you wish to use for the calculation process. The maximum value is automatically set to the number of processing cores the computer has.

3. Preparing the Calculation Step

3.1 Adding Genome Data

Before starting the calculation, you need to provide genomes that you wish to compare to the Application. The application accepts FASTA files as genome data. Adding genomes to the application can be done easily by locating your FASTA files in your system. You can add up to 10 genomes to the program.



1. Click on the “Add genomes (FASTA)” button to easily add genome data you wish to compare

3.2 Removing Genome Data

Removing genomes from the application can be done easily by selecting the FASTA files and clicking the “Remove” button

OAT version 0.90

Please cite Lee *et al.* 2015. Manuscript submitted.

2. Click to remove the selected genomes

Add genomes (FASTA) Remove

File name	Title of the first contig	No. of contigs	Size (bp)	G+C content (%)
M. bovis CQ-W70_PRJNA1...	CP005933.1 Mycoplasma bovis ...	1	948,516	29.3%
M. bovis HB0801_PRJNA16...	NC_018077.1 Mycoplasma bovi...	1	991,702	29.3%
M. bovis PG45_PRJNA6085...	NC_014760.1 Mycoplasma bovi...	1	1,003,404	29.3%
M. hyorhinitis DBS 1050_PRJ...	NC_022807.1 Mycoplasma hyor...	1	837,447	25.9%
M. hyorhinitis GDL-1_PRJNA...	NC_016829.1 Mycoplasma hyor...	1	837,480	25.9%
M. hyorhinitis HUB-1_PRJNA...	NC_014448.1 Mycoplasma hyor...	1	839,615	25.9%
M. hyorhinitis MCLD_PRJNA...	NC_017519.1 Mycoplasma hyor...	1	829,709	25.9%
M. hyorhinitis SK76_PRJNA1...	NC_019552.1 Mycoplasma hyor...	1	836,897	25.9%
M. suis KI3806_PRJNA6366...	NC_015153.1 Mycoplasma suis ...	1	709,270	31.1%
M. suis str. Illinois_PRJNA6...	NC_015155.1 Mycoplasma suis ...	1	742,431	31.1%

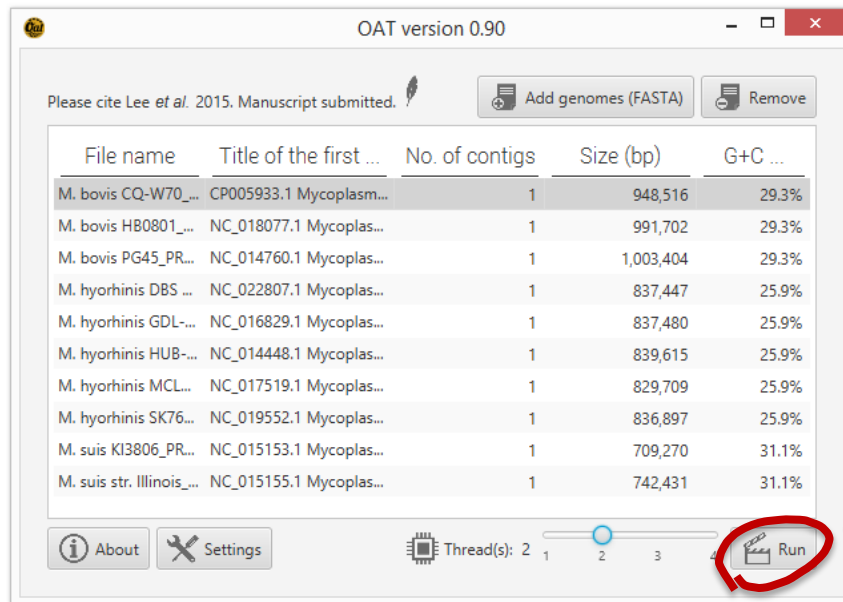
1. Select the genomes you want to remove

About Settings Thread(s): 3 1 2 3 4 Run

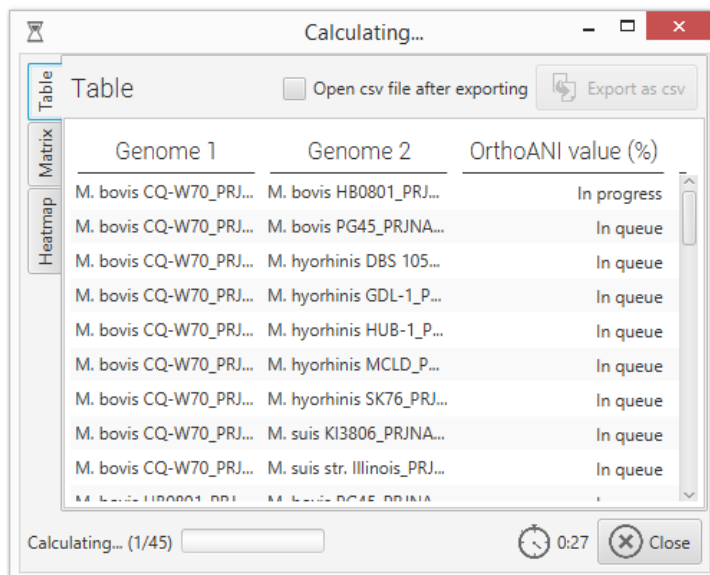
1. Click on the genomes you want remove from the application (click while pressing the shift or control key to select multiple genomes)
2. Click on the “Remove” button to remove the genomes selected from the application

3.3 Run

After managing your input data (genome files) you are now ready to run the calculation step. Simply press the “Run” button and a new window will popup showing the progress and eventually the results of the comparisons. You may start multiple runs at the same time but you should be careful not to exceed the number of processing cores available. Because you are most likely to see no performance (speed wise) gain while overwhelming your system.



1. Click the “Run” button



A new window will popup displaying the progress of the run

4. Calculation Step

4.1 Calculation Step

The calculation step will most likely take a while so it is possible to view the progress of the run and also view partial results throughout the calculation process. There will be a total of $N(N-1)/2$ comparisons (Where N is the number of genomes).
i.e.) $N = 10 \rightarrow 45$ comparisons

The screenshot shows a window titled "Calculating..." with a table of comparison results. The table has three columns: "Genome 1", "Genome 2", and "OrthoANI value (%)". The rows show various comparisons between genomes, with some completed and others in progress or in queue. Red annotations highlight specific parts of the interface:

- Comparison pairs:** A bracket on the left side of the table highlights the first two columns.
- Completed:** A bracket on the right side of the table highlights the first five rows, which have numerical OrthoANI values.
- In progress:** A bracket on the right side of the table highlights the row with the value "In progress".
- In queue:** A bracket on the right side of the table highlights the three rows with the value "In queue".
- Current pair being calculated / Total pairs:** An arrow points to the status bar at the bottom left, which shows "Calculating... (11/45)".
- Elapsed time:** An arrow points to the status bar at the bottom right, which shows a clock icon and the time "8:34".

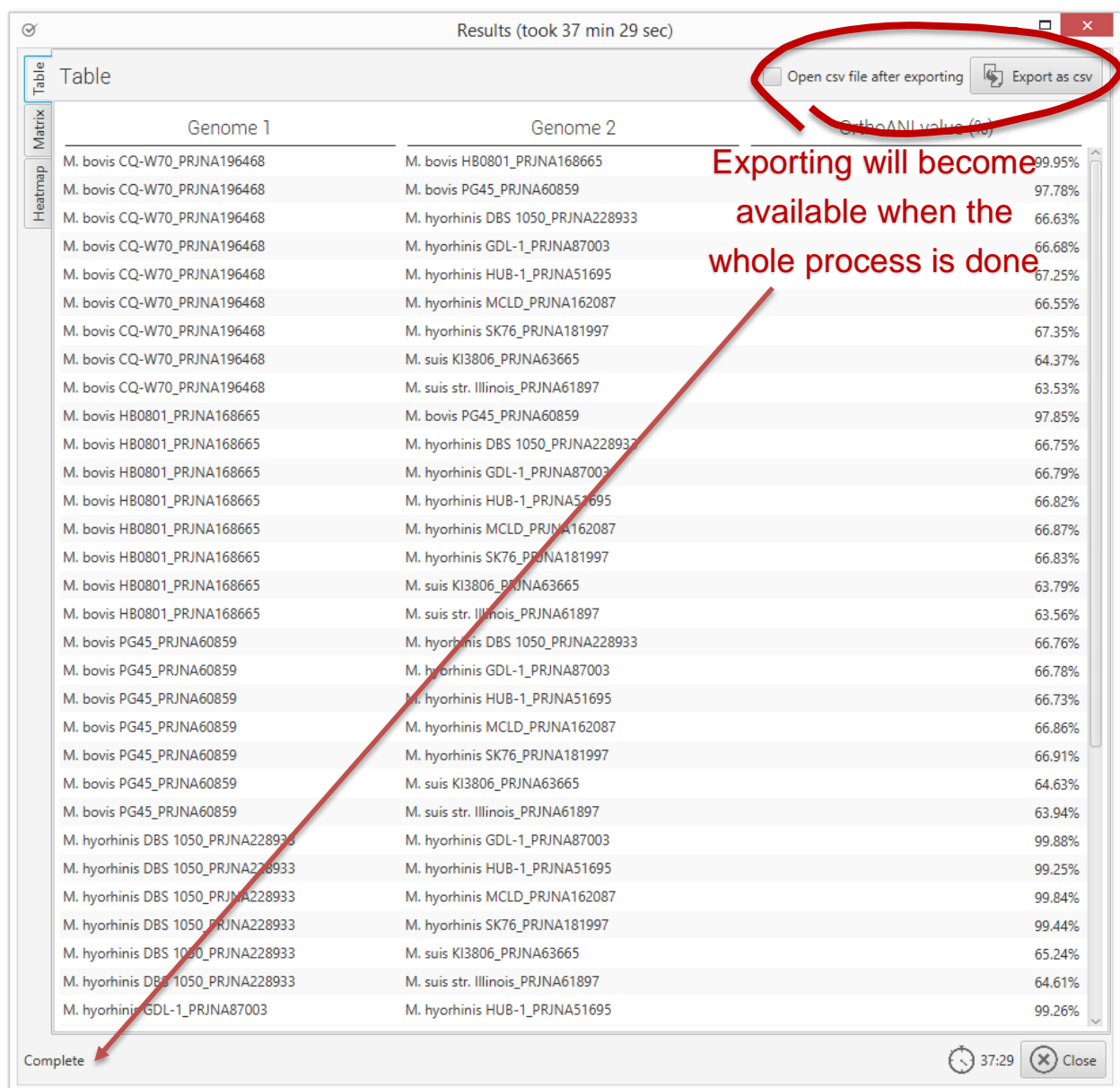
Genome 1	Genome 2	OrthoANI value (%)
M. bovis CQ-W70_PRJ...	M. hyorhinitis HUB-1_P...	67.25%
M. bovis CQ-W70_PRJ...	M. hyorhinitis MCLD_PR...	66.55%
M. bovis CQ-W70_PRJ...	M. hyorhinitis SK76_PRJ...	67.35%
M. bovis CQ-W70_PRJ...	M. suis KI3806_PRJNA...	64.37%
M. bovis CQ-W70_PRJ...	M. suis str. Illinois_PRJ...	63.53%
M. bovis HB0801_PRJ...	M. bovis PG45_PRJNA...	97.85%
M. bovis HB0801_PRJ...	M. hyorhinitis DBS 1050...	In progress
M. bovis HB0801_PRJ...	M. hyorhinitis GDL-1_P...	In queue
M. bovis HB0801_PRJ...	M. hyorhinitis HUB-1_P...	In queue
M. bovis HB0801_PRJ...	M. hyorhinitis MCLD_PR...	In queue

4.2 Result Screen

There are 3 result formats (all populated with the same data) which could be navigated through the tabs. All of them will dynamically update themselves as the OrthoANI values are being computed.

4.2.1 Table View

The table view has a basic format with 3 columns; name of the two genomes being compared and their OrthoANI values. There will be $N*(N-1)/2$ rows (N being the number of genomes) with a single header row. You can export the results as a CSV file once the whole calculations process is complete.



Results (took 37 min 29 sec)

Table

Open csv file after exporting Export as csv

Genome 1	Genome 2	OrthoANI value (%)
M. bovis CQ-W70_PRJNA196468	M. bovis HB0801_PRJNA168665	99.95%
M. bovis CQ-W70_PRJNA196468	M. bovis PG45_PRJNA60859	97.78%
M. bovis CQ-W70_PRJNA196468	M. hyorhinis DBS 1050_PRJNA228933	66.63%
M. bovis CQ-W70_PRJNA196468	M. hyorhinis GDL-1_PRJNA87003	66.68%
M. bovis CQ-W70_PRJNA196468	M. hyorhinis HUB-1_PRJNA51695	67.25%
M. bovis CQ-W70_PRJNA196468	M. hyorhinis MCLD_PRJNA162087	66.55%
M. bovis CQ-W70_PRJNA196468	M. hyorhinis SK76_PRJNA181997	67.35%
M. bovis CQ-W70_PRJNA196468	M. suis KI3806_PRJNA63665	64.37%
M. bovis CQ-W70_PRJNA196468	M. suis str. Illinois_PRJNA61897	63.53%
M. bovis HB0801_PRJNA168665	M. bovis PG45_PRJNA60859	97.85%
M. bovis HB0801_PRJNA168665	M. hyorhinis DBS 1050_PRJNA228933	66.75%
M. bovis HB0801_PRJNA168665	M. hyorhinis GDL-1_PRJNA87003	66.79%
M. bovis HB0801_PRJNA168665	M. hyorhinis HUB-1_PRJNA51695	66.82%
M. bovis HB0801_PRJNA168665	M. hyorhinis MCLD_PRJNA162087	66.87%
M. bovis HB0801_PRJNA168665	M. hyorhinis SK76_PRJNA181997	66.83%
M. bovis HB0801_PRJNA168665	M. suis KI3806_PRJNA63665	63.79%
M. bovis HB0801_PRJNA168665	M. suis str. Illinois_PRJNA61897	63.56%
M. bovis PG45_PRJNA60859	M. hyorhinis DBS 1050_PRJNA228933	66.76%
M. bovis PG45_PRJNA60859	M. hyorhinis GDL-1_PRJNA87003	66.78%
M. bovis PG45_PRJNA60859	M. hyorhinis HUB-1_PRJNA51695	66.73%
M. bovis PG45_PRJNA60859	M. hyorhinis MCLD_PRJNA162087	66.86%
M. bovis PG45_PRJNA60859	M. hyorhinis SK76_PRJNA181997	66.91%
M. bovis PG45_PRJNA60859	M. suis KI3806_PRJNA63665	64.63%
M. bovis PG45_PRJNA60859	M. suis str. Illinois_PRJNA61897	63.94%
M. hyorhinis DBS 1050_PRJNA228933	M. hyorhinis GDL-1_PRJNA87003	99.88%
M. hyorhinis DBS 1050_PRJNA228933	M. hyorhinis HUB-1_PRJNA51695	99.25%
M. hyorhinis DBS 1050_PRJNA228933	M. hyorhinis MCLD_PRJNA162087	99.84%
M. hyorhinis DBS 1050_PRJNA228933	M. hyorhinis SK76_PRJNA181997	99.44%
M. hyorhinis DBS 1050_PRJNA228933	M. suis KI3806_PRJNA63665	65.24%
M. hyorhinis DBS 1050_PRJNA228933	M. suis str. Illinois_PRJNA61897	64.61%
M. hyorhinis GDL-1_PRJNA87003	M. hyorhinis HUB-1_PRJNA51695	99.26%

Complete

37:29 Close

Exporting will become available when the whole process is done

4.2.2 Matrix View

The matrix view has columns and rows for each genomes making it a N by N table (N being the number genomes). Each cell will represent the OrthoANI value for its column (Genome # 2) and row (Genome # 1). You can export the results as a CSV file once the whole calculations process is complete.

Results (took 37 min 29 sec)

☐ Open csv file after exporting

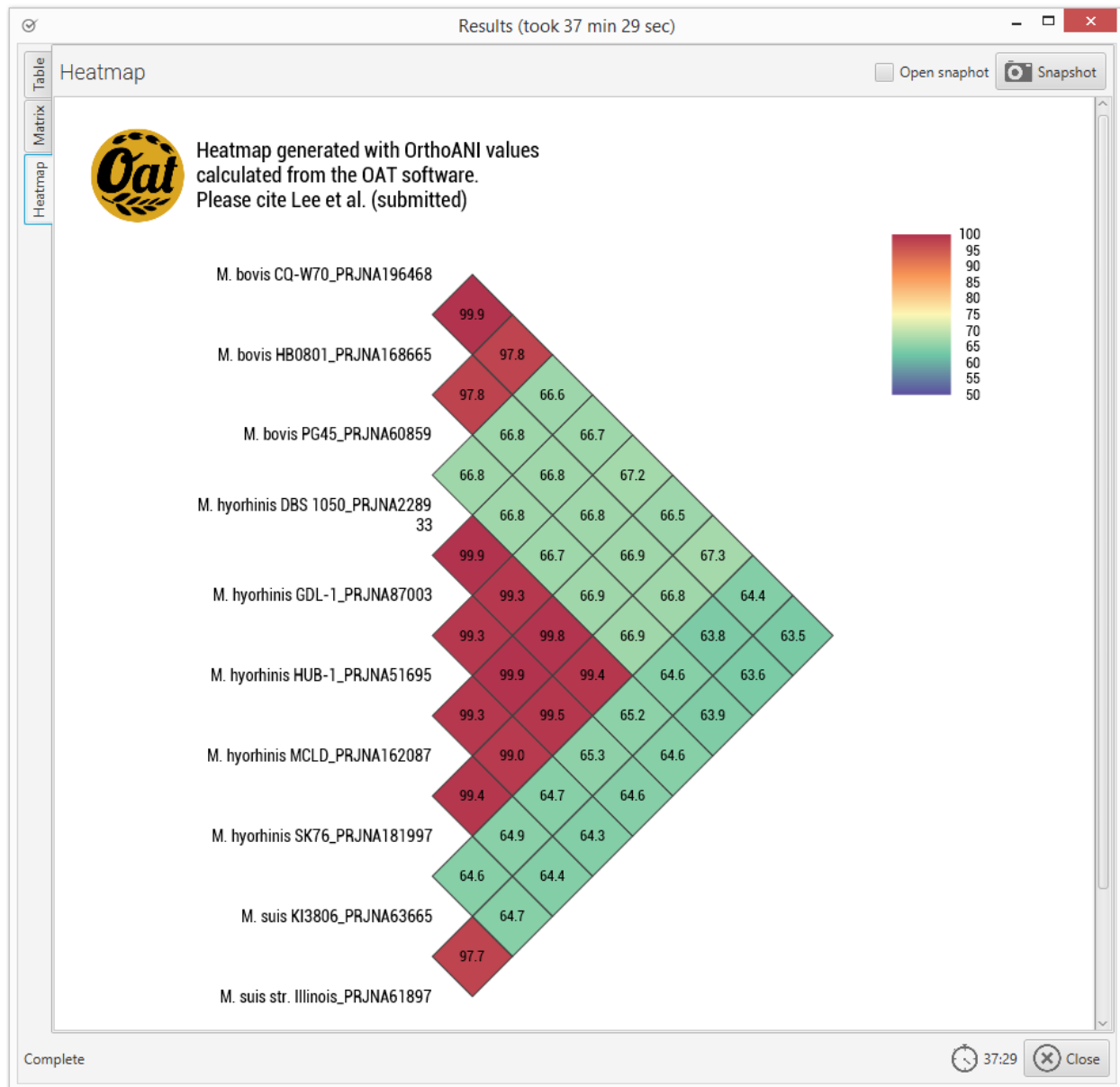
Matrix	M. bovis C...	M. bovis H...	M. bovis P...	M. hyorhin...	M. hyorhini...	M. hyorhin...	M. hyorhin...	M. hyorhin...	M. suis KI3...	M. suis str. ...
M. bovis CQ...		99.95%	97.78%	66.63%	66.68%	67.25%	66.55%	67.35%	64.37%	63.53%
M. bovis HB...			97.85%	66.75%	66.79%	66.82%	66.87%	66.83%	63.79%	63.56%
M. bovis PG...				66.76%	66.78%	66.73%	66.86%	66.91%	64.63%	63.94%
M. hyorhinis...					99.88%	99.25%	99.84%	99.44%	65.24%	64.61%
M. hyorhinis...						99.26%	99.92%	99.47%	65.26%	64.56%
M. hyorhinis...							99.26%	99.01%	64.65%	64.31%
M. hyorhinis...								99.43%	64.86%	64.43%
M. hyorhinis...									64.63%	64.72%
M. suis KI38...										97.73%
M. suis str. II...										

OrthoANI value
between the 7th genome
and the 8th genome

Complete 37:29 Close

4.2.3 Heatmap

The Heatmap is very similar to the matrix view but uses colors to indicate the OrthoANI values. This results in a more intuitive view. You can take a snapshot (PNG file) of the heatmap at any point of the process.



4.3 Support

For any difficulties or questions that you have for the OAT application, please reach out to us via lebmaster.snu+oat.help@gmail.com