PARGT User’s Guide for Windows

(version 1.0)

PARGT Documentation (last updated: October 06, 2019)

School of Electrical Engineering and Computer Science,

Washington State University, Pullman, USA

For more information, please contact Abu Sayed Chowdhury (Email: abu.chowdhury@wsu.edu, abu.chowdhury2005@gmail.com)

**PARGT** is a standalone software package developed to predict antimicrobial resistance proteins in bacteria. It is written mainly in python3 though some modules are implemented in R and called from main python script.

1. To install and use PARGT on Windows:

* Download “PARGT\_Windows.zip” from Github link- <https://github.com/abu034004/PARGT> . Unzip it and save it in a drive.
* Download “Materials\_PARGT\_Windows.zip” from <https://drive.google.com/file/d/1pUFyEKdh2DjuzMB4j0mK28DQM8RzH3i4/view?usp=sharing> . Unzip it and save it in the “PARGT\_Windows” folder. Please note that this zip file contains legacy BLAST that can be downloaded from <ftp://ftp.ncbi.nlm.nih.gov/blast/executables/legacy.NOTSUPPORTED/2.2.26/> if needed.
* To install “Jupyter Notebook”, visit the link- <https://www.anaconda.com/download/> and install anaconda for python 3. To install it from command prompt, please visit the link- <https://jupyter.org/install> .
* Provide the test fasta sequences in the file “input\_seq.fasta”. Please note that there are some example test sequences available in the folder- “testing data”.
* Now, to start Jupyter Notebook, open “Anaconda Prompt (Anaconda 3)”. Then write command to go the the directory where “PARGT\_Windows” folder is saved and then type command- “Jupyter notebook”. For example, if the folder “PARGT\_Windows” is saved in “F:\PARGT\_Windows” location then the run the commands given below.

A screenshot of a cell phone

Description automatically generated

Then after couple of seconds, a new page for Jupyter Notebook will appear in your default browser. A screenshot is given below.

A screenshot of a social media post

Description automatically generated

Now, click on PARGT.ipynb (marked by red circle in the above figure). Then it will be open up in a new browser. A screenshot is given below.

A screenshot of a social media post

Description automatically generated

**Before running the script, please install rpy2 package using command “conda install -c r rpy2” from Anaconda command prompt. All of the remaining packages should come with Jupyter Notebook by default. In the worse case, if any package is missing, please use ‘conda install’ command to install the required packages.**

Now, click inside a code cell and then click on “Run” button (red marked in the following screenshot).

A screenshot of a cell phone

Description automatically generated

If you want to stop running the script, please click on the “Interrupt the kernel” (i.e., stop button) (red marked in the following figure).

A screenshot of a cell phone

Description automatically generated

If you want to restart the kernel after facing an error, please click on “restart the kernel” button (red marked in the following figure).

A screenshot of a cell phone

Description automatically generated

After running the script, a GUI should appear as below.

A screenshot of a cell phone

Description automatically generated

Please click on the option menu to see all options available in the tool (please see below the figure).

A screenshot of a cell phone

Description automatically generated

Brief description of the options is given below.

1. **Predict aac/bla/dfr/bac/van resistance sequences**: This option is for predicting acetyltransferase (*aac*), beta-lactamase (*bla*), dihydrofolate reductase (*dfr*), bacitracin (*bac*) and vancomycin (*van*) antimicrobial resistance (AMR) proteins from the input fasta sequences in the “input\_seq.fasta” file.
2. **Include new aac/bla/dfr/bac/van resistance sequences:** These options are for a user who wants to include new known AMR sequences to the original training data comes with this tool. Again the sequences that a user wants to add need to be given in the “input\_seq.fasta” file.
3. **Include new aac/bla/dfr/bac/van non-resistance sequences:** These options are for a user who wants to include new known non-AMR sequences to the training data that comes with this tool. Again the sequences that a user wants to add need to be given in the “input\_seq.fasta” file.
4. **Restore training sets:** This option is to reset all training datasets back to the original version came with this tool.

If you find our tool useful, please cite our following papers.

**Citations:**

1. Chowdhury, A.S., Call, D.R and and Broschat, S.L., 2019. Antimicrobial Resistance Prediction for Gram-Negative Bacteria via Game Theory-Based Feature Evaluation. Scientific Reports.
2. Chowdhury, A.S., Khaledian, E. and Broschat, S.L., 2019. Capreomycin resistance prediction in two species of Mycobacterium using a stacked ensemble method. Journal of applied microbiology.