PARGT User's Guide for Linux

(Version 1.0)

PARGT Documentation (last updated: December 17, 2019)

Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA, USA

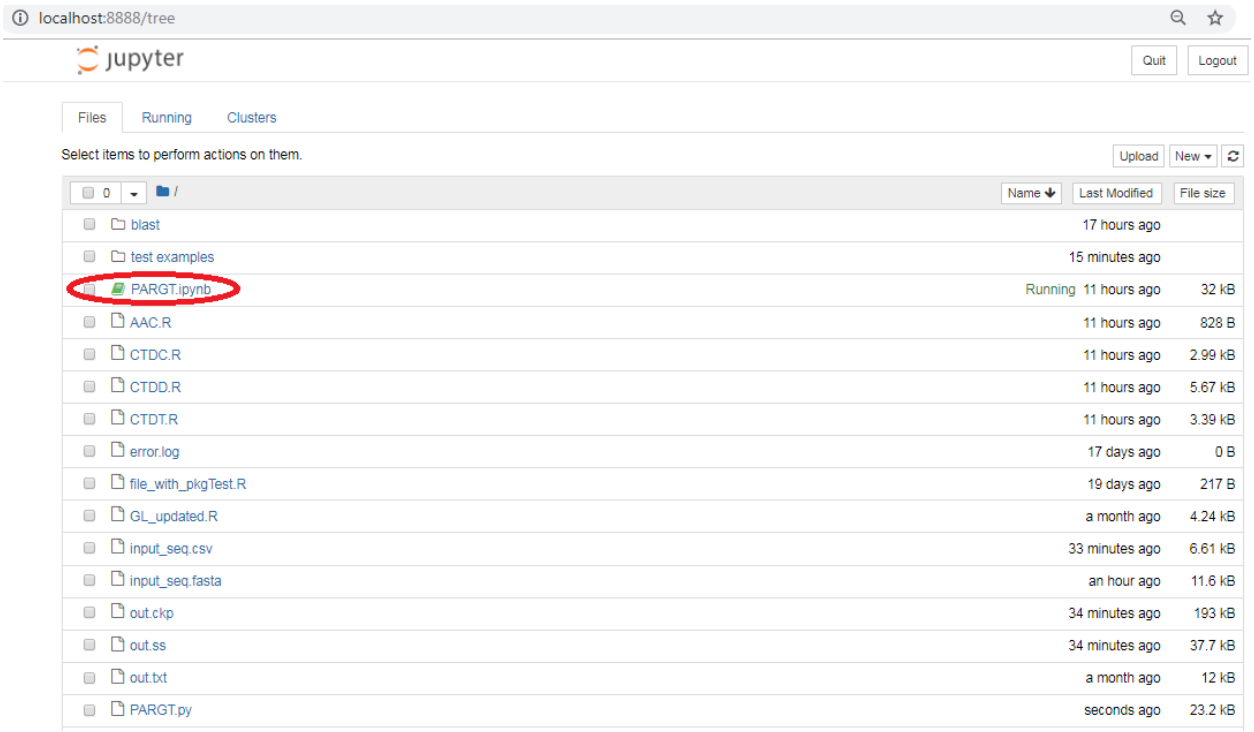
School of Electrical Engineering and Computer Science, Washington State University, Pullman, WA, USA

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

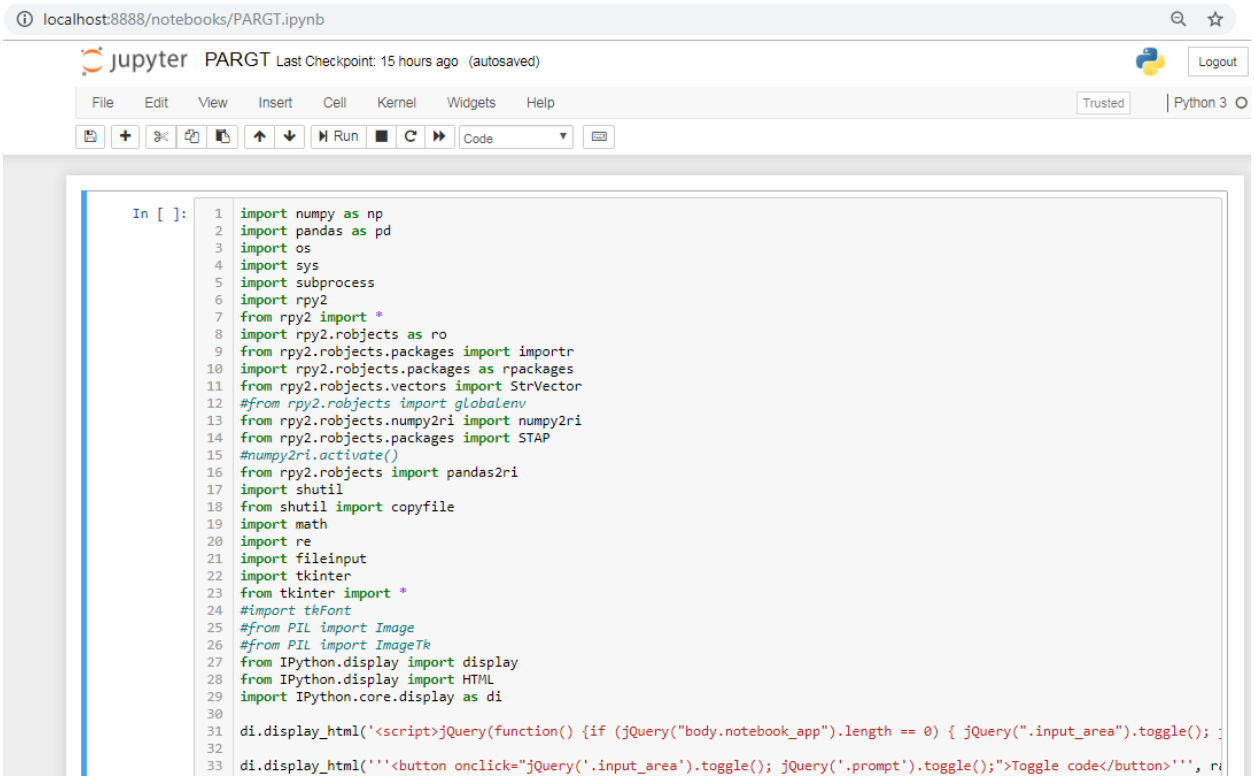
PARGT is a standalone software package developed to predict antimicrobial resistance proteins in bacteria. It is written mainly in python3 though some modules are implemented in R and called from main python script.

1. To install and use PARGT on Windows:
   - Download "PARGT_Linux.zip" from Github link- https://github.com/abu034004/PARGT . Unzip it and save it in a directory.
   - Download "Materials_PARGT_Linux.zip" from https://drive.google.com/file/d/12kHalfWWntaFyyzfY7m0YwYyj_LFDwWo/view?usp=sharing . Unzip it and save it in the "PARGT_Linux" folder. Please note that this zip file contains legacy BLAST and PSIPRED that also can be downloaded from ftp://ftp.ncbi.nlm.nih.gov/blast/executables/legacy.NOTSUPPORTED/2.2.26/ and https://drive.google.com/file/d/1cK09elzh6_SqscCtmc74cGeRvj8_Dpam/view?usp=sharing , respectively if different versions of BLAST and PSIPRED needed to install in your machine.
   - To install "Jupyter Notebook", visit the link- https://www.anaconda.com/download/ and install anaconda for python 3. To install it from terminal, please visit the link- https://jupyter.org/install .
   - It may happen that you may get an error noticing permission denied while running legacy BLAST and PSIPRED executive files. In this situation, please consider the following steps.
     Step-1: Run the command "sudo chmod 755 psipred".
     Step-2: use 'cd' command to go the directory where "PARGT_Linux/blast/bin" folder is saved and run the command "sudo chmod 755 blastpgp" and "sudo chmod 755 makemat".
     Note that you need administrative permissions to perform these two steps.
   - Edit the fasta file "input_seq.fasta" to include sequences you want to test. Note that there are some default example test sequences in this file, and you can also use example test sequences available in the folder "test examples".
   - Now, use 'cd' command to go the directory where "PARGT_Linux" folder is saved and then type command- "Jupyter notebook".

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

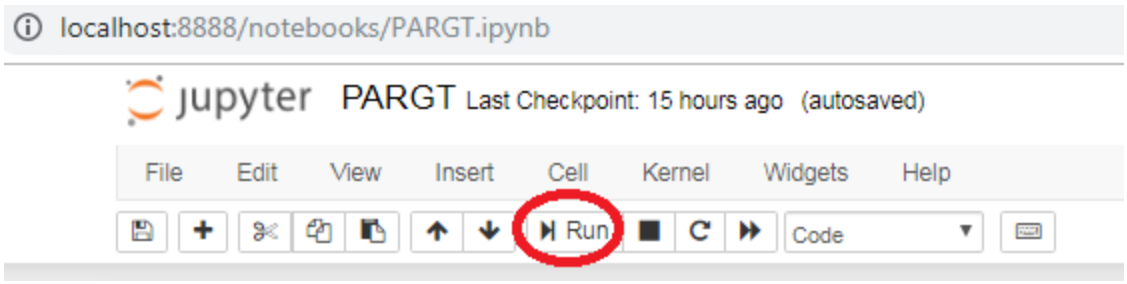


Now, click on PARGT.ipynb (marked by red circle in the above figure). Then it will be open up in a new browser. A screenshot is given below. Please note that PARGT.py script is given if a user wants to use it instead of notebook (.ipynb) version.

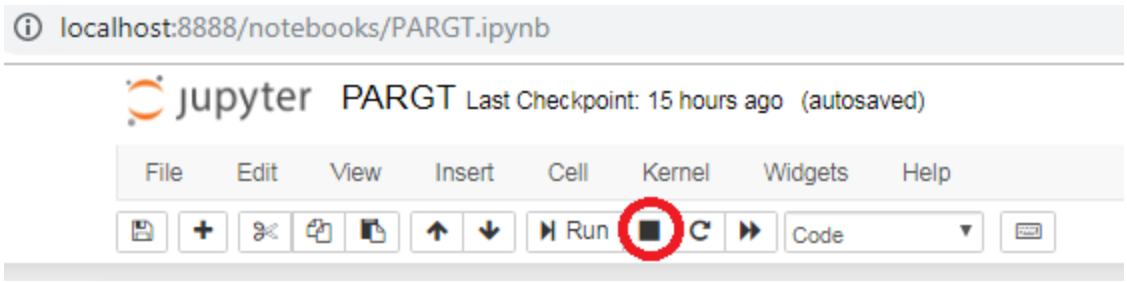


Before running the script, please install rpy2 package using command "conda install -c r rpy2" from the command prompt. Also, please install tzlocal package using command "conda install -c conda-forge tzlocal" if needed. For details, please visit https://anaconda.org/r/rpy2 **and** https://anaconda.org/conda-forge/tzlocal **for rpy2 and tzlocal packages, respectively. It may happen that your system cannot find 'psipred' executable file. In that case, please install psipred package using command "conda install -c biocore psipred". Also, if your system cannot locate e1071 package, please install e1071 package using command "conda install -c r r-e1071". For details, please visit** https://anaconda.org/biocore/psipred **and** https://anaconda.org/r/r-e1071 **for psipred and e1071 packages, respectively. All of the remaining packages should come with Jupyter Notebook by default. In the worst case, if any package is missing, please use 'conda install' command to install the required packages. Also, users can consider 'pip install' command to install necessary packages as an alternative option.**

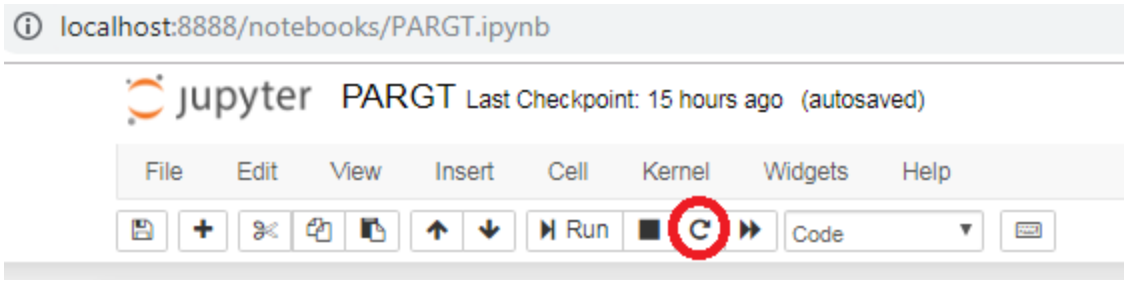
Now, click inside a code cell and then click on "Run" button (red marked in the following screenshot).



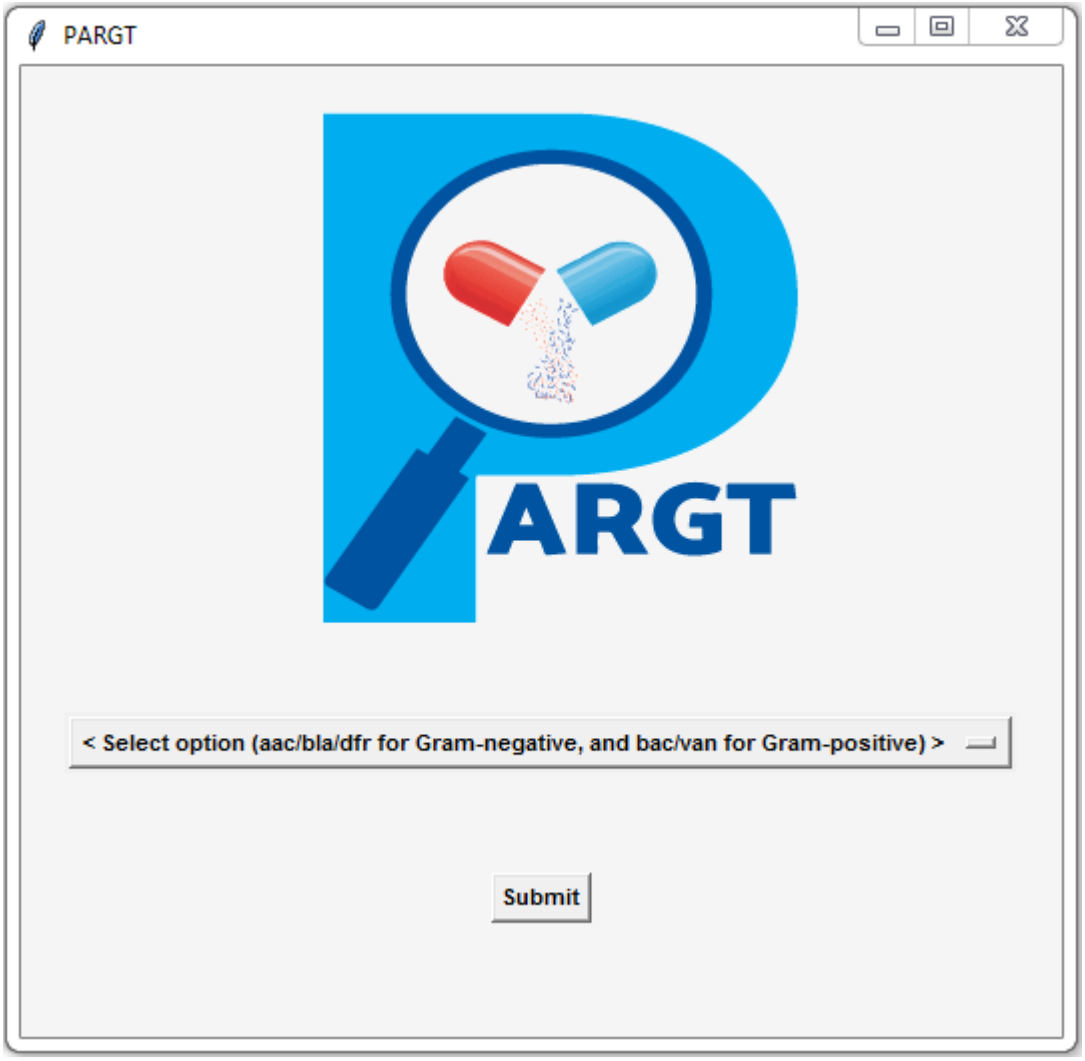
If you want to stop running the script, please click on the "Interrupt the kernel" (i.e., stop button) (red marked in the following figure).



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



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



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

Brief description of the options is given below.

(a) **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.

(b) **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.

(c) **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.

(d) **Restore training sets:** This option is to reset all training datasets back to the original version came with this tool.

After successful execution, the tool will show a message- “The operation is successful!” and the predicted output will be available in the “predicted_resistance_sequences.fasta” file located in the “PARGT_Linux” folder.

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

Citations:

1. Chowdhury, A.S., Call, D.R. and Broschat, S.L., 2020. PARGT: A Software Tool for Predicting Antimicrobial Resistance in Bacteria. Scientific Reports.
2. Chowdhury, A.S., Call, D.R. and Broschat, S.L., 2019. Antimicrobial Resistance Prediction for Gram-Negative Bacteria via Game Theory-Based Feature Evaluation. Scientific Reports.
3. 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.