PARGT User's Guide for macOS

(Version 1.0)

PARGT Documentation (last updated: December 17, 2019)

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

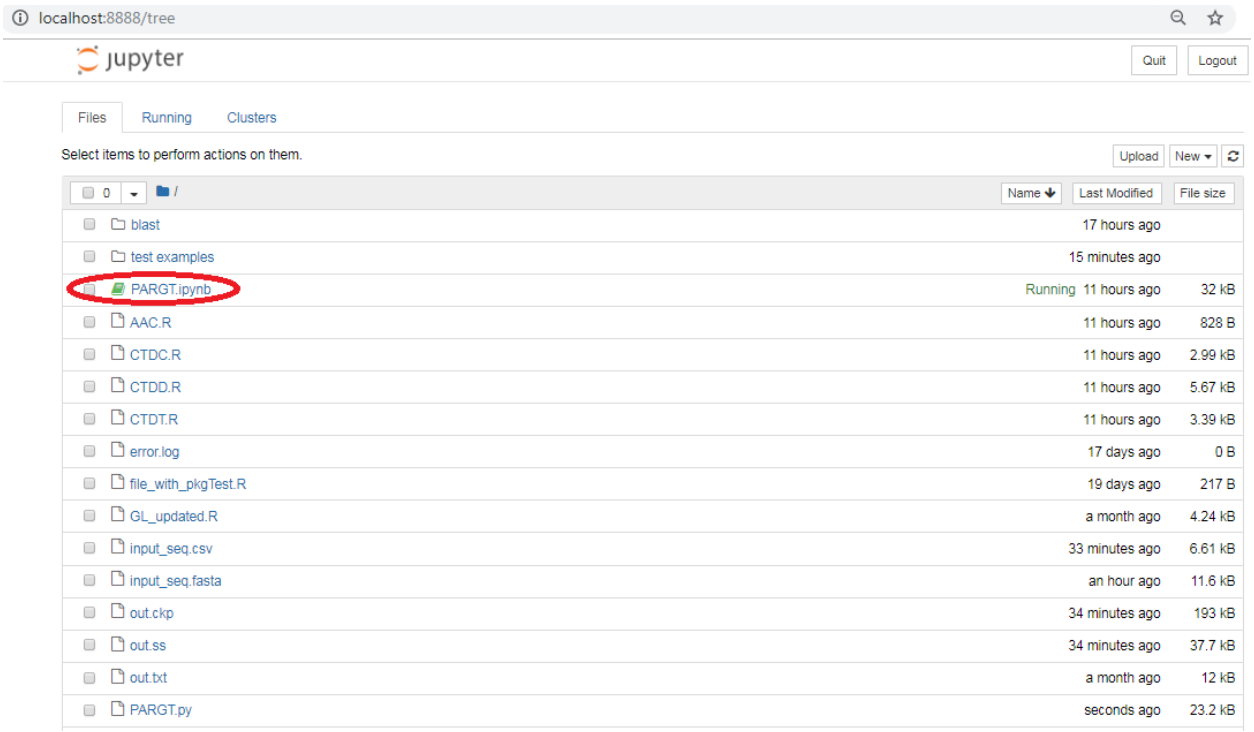
School of Electrical Engineering and Computer Science, Washington State University, Pullman, WA, USA

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

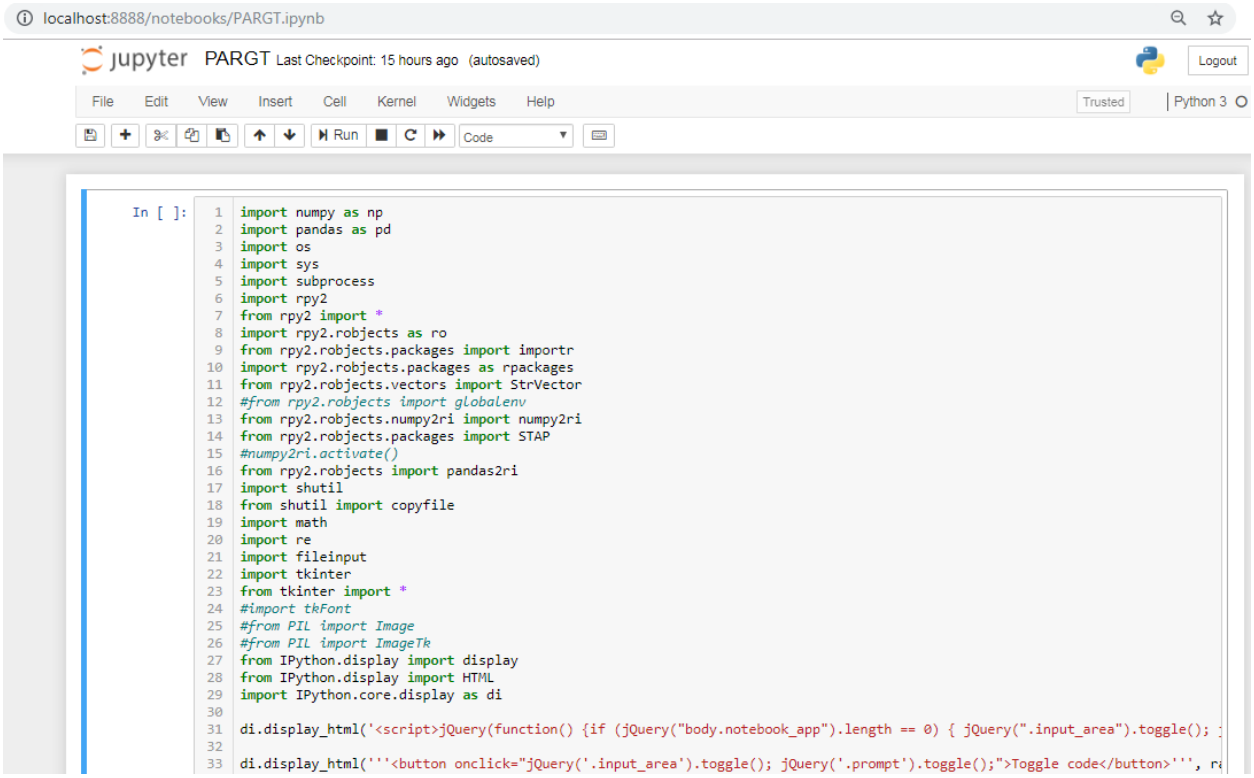
PARGT is a standalone software package developed to predict antimicrobial resistance proteins in bacteria. It is written mainly in python3 though some modules are implemented in R and called from the main python script.

1. To install and use PARGT on a Mac:
   - Download "PARGT_Mac.zip" from Github link- https://github.com/abu034004/PARGT . Unzip it and save it in a directory.
   - Download "Materials_PARGT_Mac.zip" from https://drive.google.com/file/d/1TqT9G4M70GjuPnV4UCCpTIuRaFTcOdxO/view?usp=sharing . Unzip it and save the unzipped files in the "PARGT_Mac" folder. Please note that this zip file contains PSIBLAST, legacy BLAST and PSIPRED that can also be downloaded from ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/LATEST/ , ftp://ftp.ncbi.nlm.nih.gov/blast/executables/legacy.NOTSUPPORTED/2.2.26/ and https://drive.google.com/file/d/14WY4aoR3TZz2p4_Xb6vGBcjcZmvGnfYX/view?usp=sharing , respectively, if different versions of BLAST and PSIPRED need to be installed on your machine.
   - You will need to install "Jupyter Notebooks" if you want to use the notebook version of this program. For installation, visit the link: https://www.anaconda.com/download/ and install anaconda for python 3. To install it from a terminal window, please visit the link: https://jupyter.org/install .
   - You may get an error indicating permission has been denied while running legacy BLAST and PSIPRED files. If this occurs, please follow the steps below:
     Step 1: Run the command "sudo chmod 755 psipred".
     Step 2: Use the 'cd' command to go the directory where the "PARGT_Mac/blast/bin" folder has been saved and run the commands "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".
   - Next, use the 'cd' command to go to the directory where the "PARGT_Mac" folder has been saved and then type the command "jupyter notebook".

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

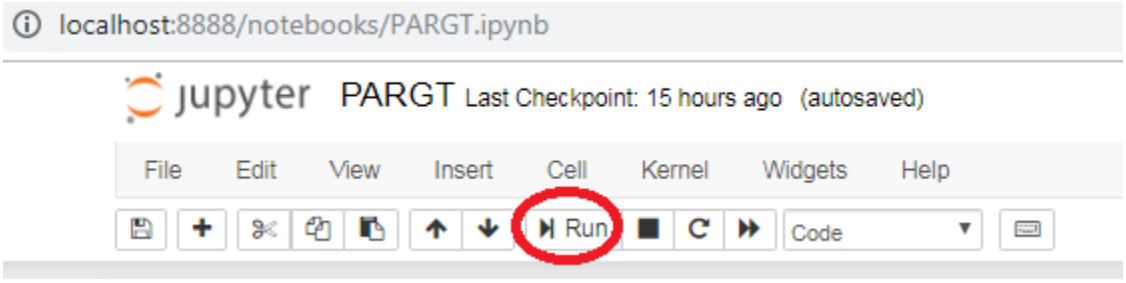


Click on PARGT.ipynb (marked by the red circle in the figure above). PARGT will open up in a new browser window or tab. A screenshot is shown below. Please note that the PARGT.py script is included if a user wants to use it rather than the notebook (.ipynb) version.

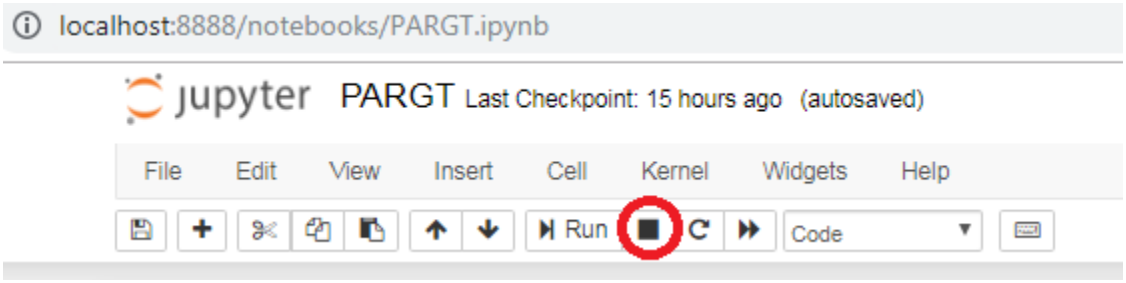


Before running the notebook, please install the rpy2 package using the command "conda install -c r rpy2" from the command prompt. Also, please install the tzlocal package using command "conda install -c conda-forge tzlocal" if needed. For details, please visit https://anaconda.org/r/rpy2 and https://anaconda.org/conda-forge/tzlocal for the rpy2 and tzlocal packages, respectively. If your system cannot find the 'psipred' executable file, please install the psipred package using the command "conda install -c biocore psipred". Also, if your system cannot locate the e1071 package, please install the e1071 package using the command "conda install -c r r-e1071". For details, please visit https://anaconda.org/biocore/psipred and https://anaconda.org/r/r-e1071 for the 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 the 'conda install' command to install the required packages. Alternatively, users can use the 'pip install' command to install the necessary packages rather than using the 'conda install' command.

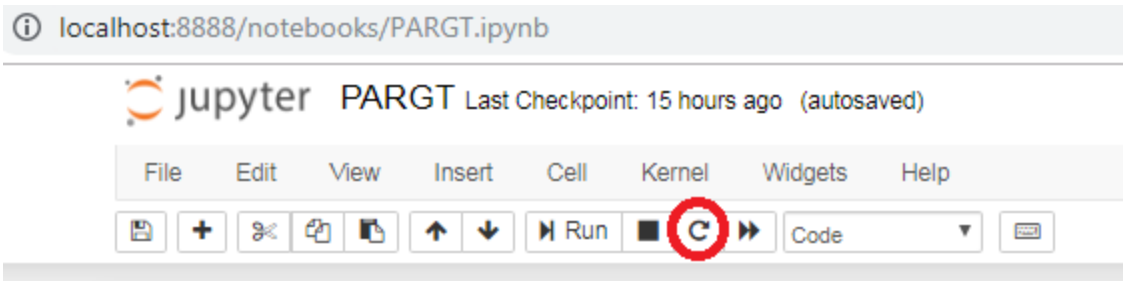
Next, click inside a code cell and then click on the "Run" button (circled in red in the following screenshot).



If you want to stop running the notebook, please click on "Interrupt the kernel" (i.e., the stop button) (circled in red in the following screenshot).



If an error occurs, you can restart the kernel. Click on the "restart the kernel" button (circled in red in the following screenshot).



After starting the notebook or script, a GUI should appear as in the following screenshot.

Please click on the option menu to see all options available in the tool as shown in the following screenshot.

A 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 default "input_seq.fasta" file.

(b) **Include new aac/bla/dfr/bac/van resistance sequences:** These options are given for a user who wants to include new known AMR sequences to the original training data that comes with this tool. Again the sequences that a user wants to add need to be placed in the "input_seq.fasta" file.

(c) **Include new aac/bla/dfr/bac/van non-resistance sequences:** These options are given 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 placed in the "input_seq.fasta" file.

(d) **Restore training sets:** This option allows a user to reset all training datasets back to the original versions that 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_Mac” folder.

If you find our tool useful, please cite the 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.