

TRAPID documentation

Quick start

This quick start section will give, through a limited set of steps to follow, a basic insight in how to process your transcriptome datasets with the TRAPID platform. For more detailed documentation, please refer to the “General Documentation” section.

The TRAPID platform is available at

<http://bioinformatics.psb.ugent.be/webtools/trapid>

1. User authentication

The TRAPID platform for the fast analysis of transcriptome data sets stores the data of users at the server. To limit access to this data, each user needs to register with the TRAPID platform in order to receive a username and password. Using these credentials, unauthorized access is impossible.

1. Go to the website
2. Click on the “register” link at the bottom
3. Provide a valid email-address and other necessary information
4. An email will be send to you, with your password. Your email-address will be used as login
5. Use this email-address and password to login into the TRAPID platform.

2. Create a TRAPID transcriptome experiment

Each user can have up to five different transcriptome data sets at the same time within the TRAPID platform. In order to differentiate between these data sets, the user must create a so-called “TRAPID experiment”.

1. Login into the TRAPID platform
2. Provide a name and description for your data set
3. Indicate which reference database you would like to use. If the transcriptome data contains transcripts from a plant species, we recommend to use the PLAZA resource as reference, otherwise the OrthoMCLDB resource.
4. Create the experiment.
5. The newly created experiment is now present in the table. Click on the name to gain access to the experiment page.

3. Uploading transcriptome data

The TRAPID experiment does not contain transcriptome data yet. The TRAPID platform supports upload of transcriptome data in multi-FASTA format. The potential large size of the transcriptome data files is offset by the ability of users to upload compressed files as well, using either zip or gzip.

1. Login into the TRAPID platform, and select the desired experiment
2. Click on the “import data” link
3. Select the file containing the transcriptome data set.
4. Click on “import transcript sequences”

In case the transcriptome data set is split across several multi-fasta files, the user has the ability to perform this process within the same experiment several times, by following steps 2 to 4 again.

4. Process the transcriptome data

In order for the platform to let the user perform custom analyzes, a pre-processing step needs to be performed by the platform itself, during which the user will not be able to view or change the experiment. This pre-processing steps consists of (among others) similarity searches, gene family association and functional annotation.

1. Login into the TRAPID platform, and select the desired experiment
2. Click on the link “Perform transcript processing”, under the header “Initial Processing”
3. Select the necessary options:
 - a. Similarity search database type: if the organism from which the transcripts originate has a very close relative species or clade available from within the list, choose this. If unknown, choose the gene family representatives.
 - b. Similarity search database: see previous point
 - c. E-value: E-value to be used during the similarity search
 - d. Gene family type: If a single species is selected as similarity search database type (using PLAZA as reference database), then the Integrative Orthology groups will provide the same or better coverage, but with smaller groups, than the otherwise selected gene family clusters.
4. Click on the “start transcriptome pipeline” button.

From this point on, the experiment is not available until the server has processed all transcriptome data from the selected experiment.

5. Perform custom analyses

With the processing having been completed by the TRAPID platform, the user now has the ability to perform many custom analyses, or view statistics, etc...

For example:

Define and use subsets within the experiment

Performing functional comparisons, GO enrichment,... can be achieved by defining subsets within the experiment (for example based on tissues, stress-conditions or developmental stage).

To define a subset within an experiment:

1. Login into the TRAPID platform, and select the desired experiment
2. Click on the link "import data" under the "import/export" header
3. Give a file in the desired format, and give a proper name for the subset.

Afterwards, you will be notified how many transcripts have been annotated with the indicated label.

Create phylogenetic trees

1. Login into the TRAPID platform, and select the desired experiment
2. Select the transcript, either through the search function, or through any link in the platform
3. On the transcript page, select the associated gene family
4. On the gene family page, select the "create phylogenetic tree" from the toolbox
5. Select the reference species from the reference database you want to include in the tree.