Use iPATH3 on uproc outputs

Alise Ponsero[1]

[1]University of Arizona

dx.doi.org/10.17504/protocols.io.u3neyme

iMicrobe Metafunc course 2018

Alise Ponsero
University of Arizona

ABSTRACT

Step-by-step description on the use of iPATH3 on Uproc-DNA outputs.

Path3.0 (http://pathways.embl.de) is a web-application for the visualization and analysis of cellular pathways. It is freely available and open to everyone. Currently it is based on four KEGG global maps, which summarize up to 158 traditional KEGG pathway maps, 192 KEGG modules and other metabolic elements into one connected and manually curated metabolic network. Users can fully customize these networks and interactively explore them through its redesigned, fast and lightweight interface, which highlights general metabolic trends in multi-omics data. It also offers navigation at various levels of details to help users further investigate those trends and ultimately uncover novel biological insights.

More information about iPATH3 can be found here : https://academic.oup.com/nar/article/46/W1/W510/4990021

TAGS

metagenomics

EXTERNAL LINK

https://pathways.embl.de/

SRS013542.denovo_duplicates_marked.trimmed.1.fastq.uproc.kegg.counts.txt

SRS143565.denovo_duplicates_marked.trimmed.1.fastq.uproc.kegg.counts.txt

PROTOCOL STATUS

Working

We use this protocol in our group and it is working

BEFORE STARTING

This protocol requires you to have a "counts" output file from Uproc. To learn more on Uproc and how to run this tool using the iMicrobe platform, read this protocol.

Create a pathway map for a unique sample

1 Note : this protocol uses as an example set the sample *SRS143565* from the HMP project. *SRS143565 is a right cubital fossa WGS available here*. The uproc-dna result for this sample is available in the description of this protocol.

In order to create a map showing the protein hits found in the sample by Uproc, go to the iPATH3 portal and select "Tools" on the top of the screen. Choose the "Numeric data" tool.

In the tool interface, use the browser to select your imput file from your computer. Note that you need to provide here the FILE_NAME.kegg.counts. The pfam28 annotation cannot be used in iPATH3.
Select the following parameters :

- **Field separator** : comma
- **Customization parameter :** width (the line width range can be further customized in the "Line width" fields).
- **Ignore unknown IDs** : yes
- **Vizualize on map** : Microbial metabolism in diverse environments.

In fixed values, the user can customize the color and opacity of the pathway map.
When ready, the user can hit "Vizualize in iPATH".

After submitting the customization data, the current map will be adjusted according to your selection. A new dialog will be displayed, giving an overview of how your Uproc annotation list matched the map.
You can zoom in/out in that map. By clicking on the edges of a pathway of interest, you can obtain additional information about the selected pathway. In this "Detailed information" window, howering the cursor on the reaction, pathway, enzyme will provide you with additional informations.

selection matching statistics window

You should also find a "Selecting Matching" window in this interface. The table lists all submitted IDs, with 3 numbers for each :

1. **Hits:** shows the total number of edges/nodes which were matched by each submitted ID.
2. **Indirect hits:** shows the number of edges/nodes matched through "Select whole pathways", "Select whole modules" or "Query reaction compounds" options.
3. **Conflicts:** When multiple IDs with different parameters in your selection match the same element in the map, a confilct occurs.

Click on any number in the table to highlight the corresponding edges/nodes.

controls window

In order to download your customized map, go to the "Controls window", select "EXPORT". You can then export your map under your favorite file format.

EXPECTED RESULT

map1.pdf

Map obtained for the sample SRS143565 . The map settings were the following :
Field separator: comma
Calculation : simple average
Customization parameter : width (1 to 50px)
Fixed values: color =red ; opacity =1
Ignore unknown IDs: yes
Visualize on map: Microbial metabolism in diverse environments

Create a pathway map comparing two samples

2 *Note : This protocol uses as example sets the samples SRS143565 and SRS013542 from the HMP project. SRS143565 is a right cubital fossa WGS available here. SRS013542 is a fornix WGS available here. The results of uproc-dna for those samples can be found in the description part of this protocol.*

In order to get a list a Kegg IDs from the uproc "kegg.counts" output, download the file from your Cyverse datastore. You will need to customize and run the following command. It will create a new file named "FILE_NAME.kegg.list" containing the list you will need for creating the map.

COMMAND

```
DOWNLOAD_DIR="Path/to/Your/Download/directory"
cd $DOWNLOAD_DIR

LIST="Name_of_your_uproc_file.kegg.counts"
OUTPUT="Name_of_your_uproc_file.kegg.list"

cat $LIST | while read LINE ; do
 IFS=', ' read -r -a array <<< "$LINE"
  echo "${array[0]}" >> $OUTPUT
done
```

Bash script to create a list of Kegg ID from the uproc counts output.

In order to create a map showing the difference between the protein hits found in two samples, go to the iPATH3 portal and select "Tools" on the top of the screen. Choose the "ID overlap" tool.

Open the list of kegg ID obtained by the small script, paste it in the ID list inputs. The user can the choose the color and width of the pathways found in either list or in both.

Choose to vizualize the hits on the "Microbial metabolism in diverse environments" map.

You should obtain a map as in step 1, showing the different hits in the two submitted lists. As describe in step 1, you can interact with this map to learn more about the hits.

EXPECTED RESULT

map.pdf

map obtained for the samples SRS143565 (blue) and SRS013542 (red). The common pathways are displayed in green. The width of the pathway hit was set as 10px and opacity of 1.