

User Manual for TIME v1

1. Input Data Formats:

Two files are required to be uploaded to start an analysis. The first file referred to as the “data file” contains the data pertaining to the abundances of taxa. The second file containing the information about the conditions and time points of the samples (referred to as the “metadata file”). Files to be uploaded need to be in very specific formats.

Data File:

The formats of these two files need to be in one of the following formats:

1. QIIME Output TSV Format: QIIME is a commonly used tool to find out, among other things, the abundance and the taxonomic information of the microbes which are present in a sample. QIIME offers the advantage that several samples can be processed and analysed in one run of their workflow. The output of QIIME can have a number of formats, the most common one being the BIOM format. The BIOM format is not readable by a normal text editor, but it can be easily converted to a TSV format file using the following command:

```
biom convert -i table.biom -o table.from_biom_w_taxonomy.txt --to-tsv --header-key taxonomy
```

The input file (“table.biom”) is converted to TSV format and the output TSV file (“table.from_biom_w_taxonomy.txt”) can be directly uploaded into TIME by selecting the appropriate option.

2. TSV File: In case the user has analysed the sequencing data using some software other than QIIME, they can upload their data in a slightly different format. This format is similar to the above format, with a few changes in order to provide flexibility to the user. The exact format (Sample File) can be downloaded from the TIME “Upload and Analyse Data” page.
3. Single Taxonomic Level TSV: In cases when only the taxa information at a specific taxonomic level are available, the user can upload the exact format as described in the Sample file provided in TIME “Upload” page.

Metadata File:

The metadata file needs to have the four columns in the following order:

- i. **Sample Name:** This column contains all names, which corresponds to the sample names in the Data File. Please ensure that the sample names in the metadata file and the data file are exactly identical.
- ii. **Date/Timedelta:** The user can provide the dates either in this specific format: “mm/dd/yy”, or as timedeltas, i.e. number of days/time-periods the sample was collected after the first sample. In either case, the algorithm converts all the time information into Timedeltas for ease of processing and analysis.
- iii. **Source:** This refers to the source of Microbiome. For example microbiome can be collected at various time points from the guts of two individuals. In this case the source will be Individual 1 and 2. Please note that several samples can be collected from each source, and it is not necessary that the number of samples collected from each source to be the same.
- iv. **Condition:** If any underlying condition is present across a time stretch for any Source, the information can be supplied in the Condition column. In case no information is available about the condition, the column needs to be filled up with dummy information. For example all the elements of the column can be filled with “None” values.

Note :

- Users have to select the appropriate option “Timedelta” or “Date” appropriately depending on which format the time information is provided.
- The sample names in the input data file and metadata file should be identical.

2. Data input

The various available features in TIME can be explored using the provided three example case studies (“Explore Examples”). A user can select the “Upload and Analyze data” button for custom data analysis using the tool. Under the “Explore Examples”, clicking on the “View Paper” button will direct the user to the website containing the journal publication from which the data was taken. It is recommended to go over the paper in order to understand the data and the metadata. The “Download” button provides a link for downloading the data set (Data and Metadata) used for the case studies used in each of the examples. The “Analyze” button allows the user to start the analysis process. The user will be redirected to a page containing available options for analysing the data.

TIME

Temporal Insights into Microbial Ecology

TIME About User Manual

TIME is a web application which enables microbiome researchers to visually explore, make logical inferences, and gather insights from time series microbiome datasets. TIME has been designed with a user friendly graphical interface. The intuitive and interactive mouse operations allow the end users to visually explore the datasets and discern the underlying biological interactions of the habitat. In addition, the generated plots as well as the tables can be easily downloaded for publication or further analysis.

Explore Examples

Upload and Analyse Data

Important

Select only one from the List of Examples

Click on "Analyse" to Start analysis

After starting an analysis, to analyse a different example dataset, please come back to this page and start afresh

Example 1 :	Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation	View Paper	Download	▶ Analyse
Example 2 :	Moving Picture of the Human Microbiome	View Paper	Download	▶ Analyse
Example 3 :	Temporal Dynamics of the Human Vaginal Microbiota	View Paper	Download	▶ Analyse

Click here to upload a custom dataset and proceed analysis

Click here to explore preloaded example datasets

Click on the 'Analysis' button beside an example study to proceed analysing it

Clicking on the 'Upload and Analyze Data' button opens the page where user can upload and analyse own data. A "Sample data file" is provided with each option for understanding the format of his/her input files. Files in any other format will not be processed.

Custom data input for analysis

Explore Examples

Upload and Analyse Data

Upload Data File

Click Here for more information on Input Data Formats

Select the type of data you want to input

- QIIME Output Tab Separated File Sample Data File 1
- Taxonomy Hierarchy Tab Separated File Sample Data File 2
- Single Taxonomic Level TSV Sample Data File 3

Choose an abundance table file to upload

No file chosen

Choose whether to perform normalization

Rarefaction/Normalization

Rarefy or Normalize Abundance values

Do Not Normalize Abundance values

Number of samples to generate the rarefaction curve?

Number of times Resampling is done to generate the rarefaction curve

Upload Metadata File

Click Here for more information on Input MetaData Format

Sample Metadata File 1

Sample Metadata File 2

Sample Metadata File 3

Choose a metadata file to upload

No file chosen

Select the format of the time column

Timedelta Date

Perform analysis keeping the 'dd/mm/yy' form or convert to timedelta

Enter a job Title

Must be less than 25 characters. Characters can be alpha-numeric or "_"

Choose the desire file format and then click on the "Choose file" button to upload the file

Choose the desire file format and then click on the "Choose file" button to upload the file

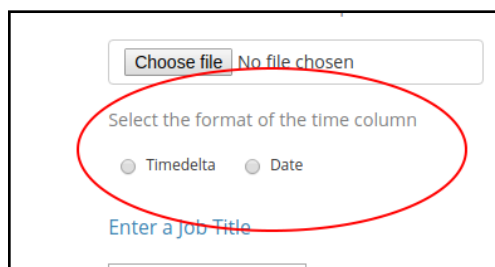
For uploading the abundance data files, user can choose to upload either the QIIME output file (TSV file created from the BIOM file), or other TSV formats (as shown in the Sample files). If the taxonomic lineages of all the taxa are available along with the abundance data, then the second type of data input must be selected. In case when only the abundance at one specific taxonomic level is available, then the third option must be selected.

Note: For uploading QIIME output files, user must first convert the BIOM file (QIIME output) to the proper TSV format by running the following command:

```
biom convert -i table.biom -o table.from_biom_w_taxonomy.txt --to-tsv --header-key taxonomy
```

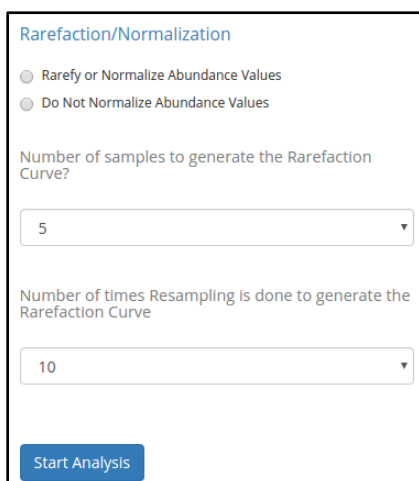
where “table.biom” is the name of the QIIME output file and “table.from_biom_w_taxonomy.txt” is the name of the TSV file which is generated by the command.

Metadata files have to be in the format as specified in the Sample File, with 4 columns corresponding to the Sample_Name, Date, Source and Condition. The Condition column should not be left blank, and should be filled up with any constant string even if no such information is available. The Date column can contain either raw dates or time elapsed from the starting of the experiment. In case of the former, the Date option, and in case of the latter, the Timedelta option should be selected as shown below.



The screenshot shows a web interface with a file upload button labeled "Choose file" and "No file chosen". Below it, there is a section titled "Select the format of the time column" with two radio buttons: "Timedelta" and "Date". A red oval highlights the "Timedelta" radio button. Below the radio buttons, there is a text input field with the placeholder "Enter a Job Title".

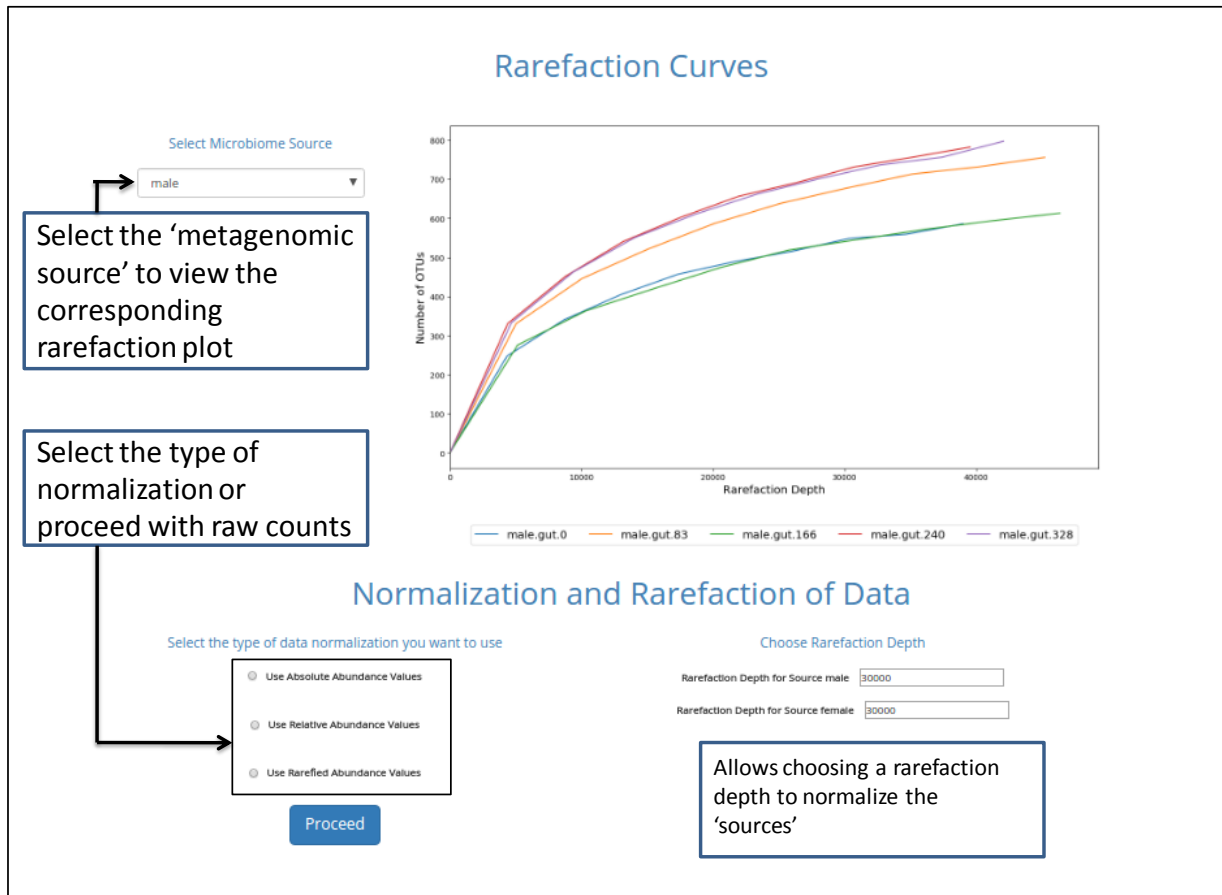
Also, user can choose to analyse the raw abundance counts or apply a rarefaction or relative normalization. If a rarefaction is required, users can select two parameters. The first one allows selecting the number of samples (time points) that would be used to generate the rarefaction plot. User can either select all the points or a small representative set of equidistant time points to generate the plots. Selecting the ‘All’ options might take a considerable long time to generate the plots depending on the size of the dataset. Additionally, the second parameter allows choosing the re-sampling count to generate the rarefaction plots. Please note that ‘rarefaction curves’ and ‘rarefaction based normalization’ can only be applied to OTU table format inputs only.



The screenshot shows a web interface for "Rarefaction/Normalization" settings. It has two radio buttons: "Rarefy or Normalize Abundance Values" and "Do Not Normalize Abundance Values". Below the radio buttons, there are two dropdown menus. The first dropdown is labeled "Number of samples to generate the Rarefaction Curve?" and has the value "5" selected. The second dropdown is labeled "Number of times Resampling is done to generate the Rarefaction Curve" and has the value "10" selected. At the bottom of the form, there is a blue button labeled "Start Analysis".

A Job Title must also be typed into the corresponding field. This text should contain only alphanumeric characters and “_”.

After filling up the form, the “Start Analysis” button should be clicked which generates the rarefaction plots as shown below.



Based on the rarefaction plot, user can either select a suitable rarefaction depth for normalization, use relative normalization or skip the normalization and “Proceed” with the absolute count data. The rarefaction plots for the different ‘metagenomic sources’ can be viewed by selecting the required option from the dropdown at the top left.

N.B: If the rarefaction plot is not generated or the ‘Proceed’ button gives a blank result page please check the following:

1. The correct radio-button selection is made for the dataset format
2. The metadata is uploaded in the correct format
3. Please make sure that the ‘Timedelta’ or ‘Date’ option is select correctly. If the metadata contains the sampling times in data and the user selects ‘Timedelta’ in the radiobutton option, there will be no output (blank page).
4. Please make sure the normalization option is chosen carefully. Users should refer to the ‘TIME’ manuscript for more details.

Analysis Home

Options to specify the metagenomic sample source to analyze and at what phylogenetic level.

Once this options are selected and the 'View' button is clicked all the workflows will be valid only for the current selection.

Please modify selection and click 'View' again to perform analysis on a different source or phylogenetic level



A plot of the 'Richness' and 'Shannon diversity' for the selected sample at the chosen phylogenetic level

Click on the 'Analyse' button beside a workflow to proceed analysing the same

The "Analysis Homepage" is the control panel of TIME analysis. After selecting the "Microbiome Source" and the "Taxonomic Level" user can click on the "Start" button to start the analysis with the selected parameters. The plot generated on the right corresponds to the time series plot of the diversity measures (Taxa Richness and Shannon Diversity). User can explore the plot using various intuitive mouse operations listed in the help box and can also download the results as images or as TSV for further analysis. In every workflow in TIME, the Help Section of the page explains all the interactive operations and other useful features.

4. Workflows

In the Analysis Home Page, user needs to scroll down to find all the workflows. User has the option to choose from the workflows that they can choose from. A short description of the utility and objective of each of the workflows can be obtained by simply clicking on it. Please go through the original article of the TIME to get more information about the workflows and the backend algorithms.

Select a Workflow

Click on "Analyse" to start the workflow.

Workflow 1: Identify abundance based variations in taxonomic groups over time	More Information	Analyse
Workflow 2: Compare temporal trends between selected taxa	More Information	Analyse
Workflow 3: Identify temporally stable/unstable taxa	More Information	Analyse
Workflow 4: Identify variations in taxonomic groups between two time ranges	More Information	Analyse
Workflow 5a: Cluster groups of taxa having similar behavior over time	More Information	Analyse
Workflow 5b: Explore pair-wise relationship among taxonomic groups	More Information	Analyse
Workflow 6: Explore inter taxa interactions using causality network	More Information	Analyse
Workflow 7: Cluster time points based on similar community patterns	More Information	Analyse

In order to start analysis using a chosen workflow, the user has to click on the “Analyse” button, which will open a new tab for that workflow.

Note: The Analysis Home page should not be closed while an analysis is underway. The user should open a new tab for new analysis. In case the page is closed, the user will lose the current analysis and will have to start afresh.

Workflow 1: Identify abundance based variations in taxonomic groups over time

Abundance Plot

Please Choose the desired Parameters and click on the 'Plot' Button

Microbiome Source:

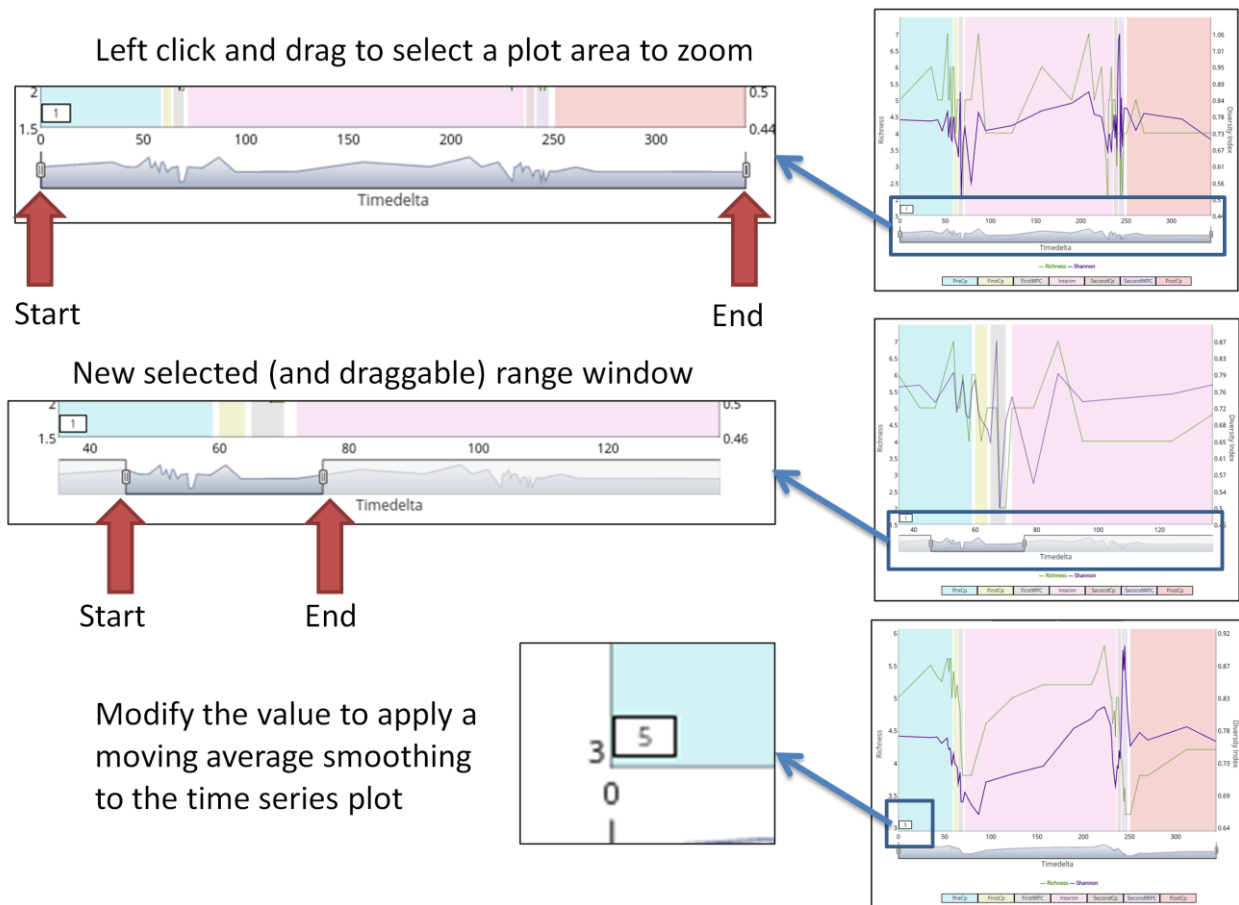
Taxonomic Level:

Select Abundance:

The user has to select the taxa which one wishes to view. There are seven options: “Very High”, “High”, “Low”, “Very Low”, “Core”, Persistent” and “Transient”. For more information about these options, please refer to the paper. It must be noted that the “Microbiome Source” and the “Taxonomic Level” cannot be changed from this page. If the user wants to change any of these, it should be done from the Analysis Home Page only. Clicking the ‘Plot’ button will generate the time series plot of the selected group of taxa.



Interacting with the plot



Once the plot is generated, the user has the following options. These options are common in all the visualizations where time series plots are generated:

1. **Scaling the Y axis:** In order to view the plots of taxa which differ in abundances in several orders of magnitude, “Log Scaling” the Y axis has been provided. The user has the option to toggle between “Log Scale” and “Linear Scale” the Y axis.
2. **Gridlines:** The Gridlines in the plot can be turned off or on by the “View Gridlines” checkbox.
3. **Download Results:** For further analysis, the users can download the results as an image (.png file) and also as text files.
4. **Moving Average:** The moving average of every timepoint is calculated. The text box allows the user to choose any value over which the moving average is calculated.
5. **Zooming the Timescale:** The user have the option to zoom in to a sub stretch of the time by moving the appropriate sliding bars. It can also be clicked and dragged.
6. **Conditions:** If provided in the metadata, the underlying conditions across timepoints can be highlighted and the legend will provide information about them.

Workflow 2: Compare temporal trends between selected taxa

The user has to click on the “Plot” button to generate the workflow for the selected options. A blank plot will be generated, along with a text box above it. Entering the query taxa name in the text box either by typing or by selecting from the dropdown, enables one to view the corresponding trend plot.

Subsequently, users can also add or delete one or more taxa to append or remove the respective trend plot from the plot. All the mouse operations described in Workflow 1 will also work in this workflow.

Microbiome Source:

E ▼

Taxonomic Level:

Phylum ▼

Select and View Taxa

Source: E; Taxonomic Level: Phylum

Log Scale Y Axis | Linear Scale Y Axis

View Gridlines

Download Graph PNG

To modify these options, please go back to the 'Analysis Home'

Type taxa names or select taxa from the dropdown to view their time series plot

➤ This dropdown allows to directly select a taxa or enter the query search string.

➤ Once a valid taxon name is selected, the corresponding trend plot will be displayed in the canvas.

➤ Multiple taxa can be appended or removed from the plot

Help with Mouse Operations

Operation	Action
Hover on Graph	Highlights nearest timeseries and nearest value of timeseries
Click and Drag Timescale	Zooms in on the selected Timestretch
Click and Type on the Dropdown	Names of the taxa present will be suggested, clicking on it will show the corresponding time series plot
Click on the cross (X) option next to taxon name	Removes the plot of the corresponding taxon, while the other taxa timeplots remain.
Click on Log or Linear Scale Buttons	Linear or Log Scale the Y axis to assist time series comparison

Microbiome Source:

E

Taxonomic Level:

Phylum

Help with Mouse Operations

Operation	Action
Hover on Graph	Highlights nearest timeseries and nearest value of timeseries
Click and Drag Timescale	Zooms in on the selected Timestretch
Click and Type on the Dropdown	Names of the taxa present will be suggested, clicking on it will show the corresponding time series plot
Click on the cross (X) option next to taxon name	Removes the plot of the corresponding taxon, while the other taxa timeplots remain.
Click on Log or Linear Scale Buttons	Linear or Log Scale the Y axis to assist time series comparison

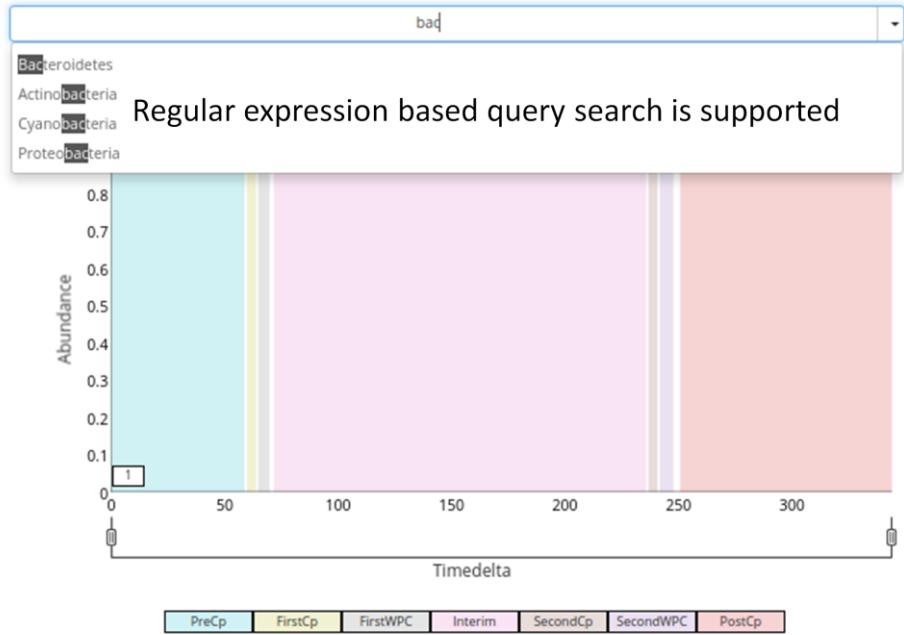
Select and View Taxa

Source: E; Taxonomic Level: Phylum

Log Scale Y Axis Linear Scale Y Axis

View Gridlines

Download Graph PNG



Microbiome Source:

E

Taxonomic Level:

Phylum

Help with Mouse Operations

Operation	Action
Hover on Graph	Highlights nearest timeseries and nearest value of timeseries
Click and Drag Timescale	Zooms in on the selected Timestretch
Click and Type on the Dropdown	Names of the taxa present will be suggested, clicking on it will show the corresponding time series plot
Click on the cross (X) option next to taxon name	Removes the plot of the corresponding taxon, while the other taxa timeplots remain.
Click on Log or Linear Scale Buttons	Linear or Log Scale the Y axis to assist time series comparison

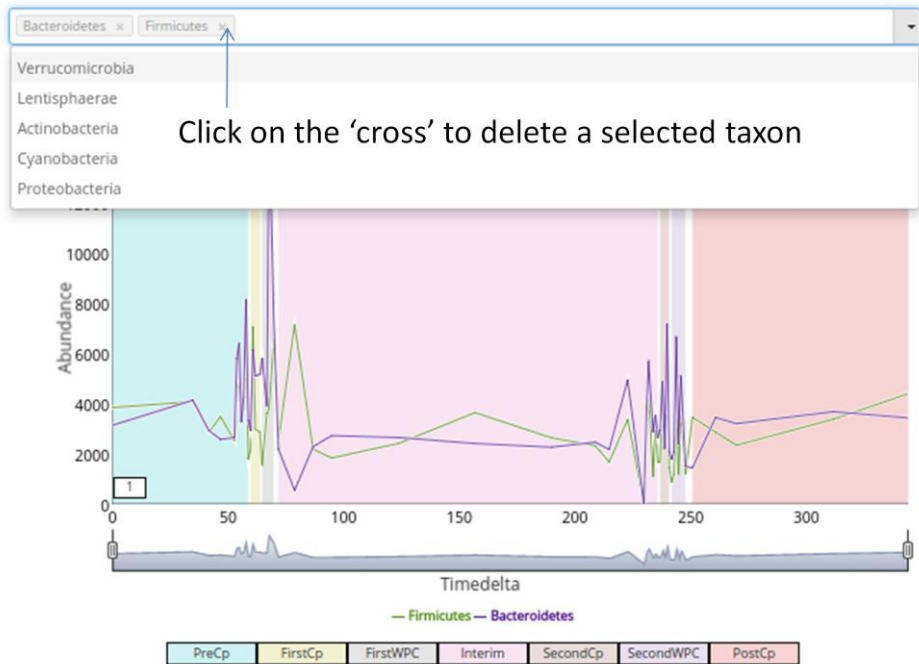
Select and View Taxa

Source: E; Taxonomic Level: Phylum

Log Scale Y Axis Linear Scale Y Axis

View Gridlines

Download Graph PNG



Workflow 3: Identify temporally stable/unstable taxa

Identify Stationary Taxa

Source: E; Taxonomic Level: Genus

Results

Click on the small arrow to sort the contents in the category

Modify the parameters and click on "Analyze" button to recalculate the table. The 'Microbiome source' and 'Taxonomic level' can be changed from the 'Analysis home';

Taxon	Stationary	Type
[Ruminococcus]	Yes	Transient
Akkermansia	Yes	Transient
Anaerostipes	Yes	Core
Bacteroides	Yes	Core
Bilophila	Yes	Transient
Blautia	Yes	Core
Butyricimonas	Yes	Core
Clostridium	Yes	Transient
Coprococcus	Yes	Core
Dorea	Yes	Core

Showing 1 to 10 of 22 entries

Download Results Table

Previous 1 2 3 Next

Help with Mouse Operations

Operation	Action
Click on Column Headings	Sorts the corresponding columns in ascending or descending order
Type in the Search field	Searches for a taxa with similar names

The user has to specify the Significance Level of the Augmented Dickey Fuller statistical test for identifying whether the time series is stationary or not. Along with the significance, a "Cutoff" needs to be selected from a dropdown. The taxa which are present in less than the Cutoff * 100% of the timepoints will not be considered for this analysis and will be neglected. After selecting the above parameters as per requirements, clicking the "Plot" button will generate the results in a tabular format. Any column can be chosen to sort the table. Simply clicking on the column header sorts the column in ascending or descending order. The user can also search for a particular taxa in the "Search" box. Additionally, clicking on the "Download Results Table" downloads the results as a TSV (tab delimited) file.

Workflow 4: Identify variations in taxonomic groups between two time ranges

In this workflow, once the user chooses two time periods, the behaviour of every taxa across those two periods will be compared. The start and the end timepoints of every period has to be selected from the dropdowns (as shown below). Choosing the timepoints also highlights the selection in the adjacent plots, from which the user can easily discern the conditions which the periods correspond to. The colors correspond to the condition information available in the metadata.

Microbiome Source:
E

Taxonomic Level:
Genus

Choose Period 1 Start

Choose Period 1 End

Choose Period 2 Start

Choose Period 2 End

Calculate

Variations in taxa between two time ranges

Source: E; Taxonomic Level: Genus

Please select the two desired timeranges (Period 1 and Period 2) by selecting the Start and End time points of each period using the dropdowns and click on the Calculate Button

After selection, click on the 'Calculate' button to generate the results

Clicking the "Plot" button generates a tabular output, which like the previous workflow's results have sortable columns, search fields, and options to download the results as TSV files.

Result Summary

Period 1 Start Time: 0, End Time: 59
Period 2 Start Time: 65, End Time: 124

Taxon	Period 1 Mean	Period 2 Mean	Period 1 Percentage Zeros	Period 2 Percentage Zeros	Period 1 Correlation with Time	Period 2 Correlation with Time	Log (Period 1 Mean / Period 2 Mean)	P Value of Wilcoxon Rank Sum Test	Corrected P Value (Using Benjamini Hochberg)
Parabacteroides	16.18	6.0	0.0	22.22	0.26	0.64	0.99	0.01	0.57
Streptococcus	1.09	5.22	45.45	11.11	0.26	-0.4	-1.57	0.03	0.7
Bilophila	12.82	11.78	0.0	55.56	0.35	0.09	0.08	0.04	0.7
Butyrivibrio	2.91	0.89	18.18	55.56	-0.01	-0.02	1.19	0.05	0.7
Phascolarctobacterium	375.0	265.89	0.0	0.0	0.37	-0.03	0.34	0.14	1.0
Haemophilus	9.73	2.33	27.27	66.67	0.27	0.43	1.43	0.16	1.0
Staphylococcus	0.36	0.0	63.64	100.0	0.03	nan	inf	0.17	1.0
Prevotella	27.45	20.33	9.09	33.33	0.43	0.08	0.3	0.18	1.0
[Eubacterium]	1.36	0.44	45.45	77.78	-0.19	0.62	1.12	0.2	1.0
Faecalibacterium	98.73	74.89	0.0	0.0	-0.5	-0.07	0.28	0.31	1.0

Showing 1 to 10 of 50 entries
Previous **1** 2 3 4 5 Next

Workflow 5a: Cluster groups of taxa having similar behaviour over time

The current workflow allows user to select two parameters as shown below:

Select DTW Constraint

2

Cutoff for Ignoring Rare Taxa

0.1

Plot

More information about the first parameter “Select DTW Constraint” can be found in the original TIME article. Along with the significance, a “Cutoff” needs to be selected from a dropdown. The taxa which are present in less than the Cutoff * 100% of the timepoints will not be considered for this analysis and will be neglected.

The screenshot displays the TIME web application interface. At the top, the title "TIME Temporal Insights into Microbial Ecology" is visible. The main navigation bar includes "HOME" and "Workflow 5".

The interface is divided into several sections:

- Microbiome Source:** A dropdown menu set to "E".
- Taxonomic Level:** A dropdown menu set to "Genus".
- Select DTW Constraint:** A dropdown menu set to "2".
- Cutoff for Ignoring Rare Taxa:** A dropdown menu set to "0.5".
- Plot:** A blue button to generate the visualization.
- Time Series Similarity Trees:** The main visualization area, showing a hierarchical tree structure. A "Choose a Condition" dropdown is set to "Default: All". A "Radial layout" checkbox is checked.
- Trend plot:** A line graph showing abundance over time (0 to 300). The y-axis is labeled "Abundance" and ranges from 0 to 15000. The x-axis is labeled "Timepoints".
- Legend:** A color-coded legend for the taxa, including Bifidobacterium, Lactobacterium, and others.

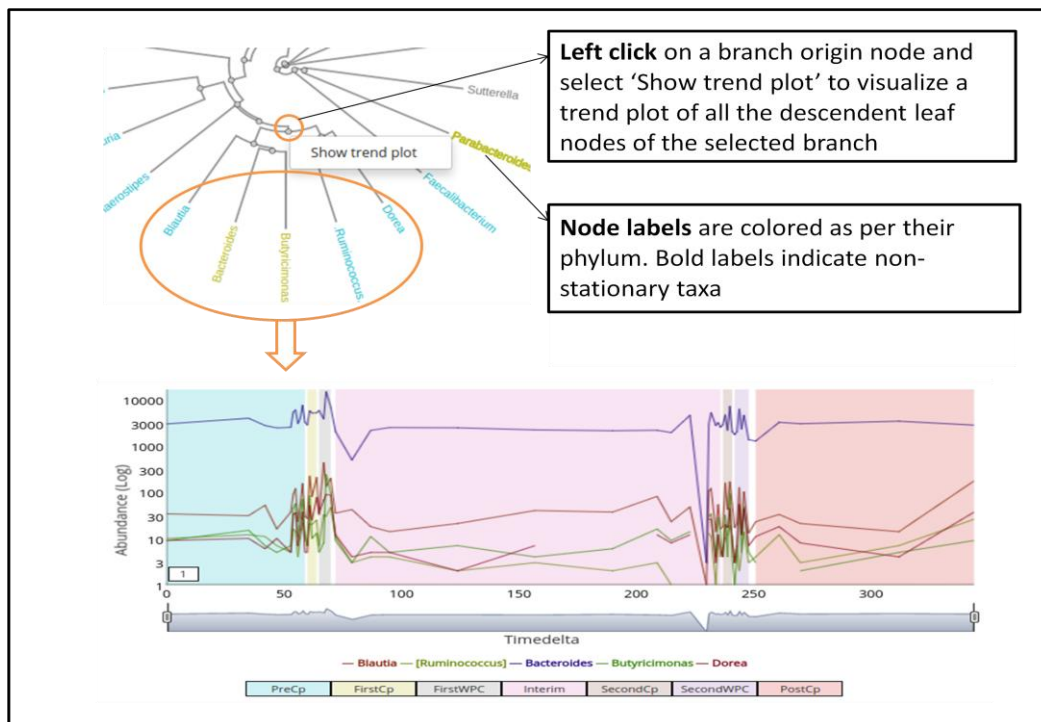
Annotations and callouts provide additional information:

- Select a desired time range to generate the tree. This range corresponds to the time ranges described in the metadata** (points to the "Choose a Condition" dropdown).
- Select the 'Radial layout' checkbox to view the same tree in a radial form** (points to the "Radial layout" checkbox).
- Circular plot with phylum (label colors) and non-stationarity information (bold labels) overlaid** (points to the radial plot).
- Modify the 'DTW constraint' and 'rare taxa cutoff' using the dropdown. To modify 'source' and 'taxonomic level', please go back to the 'Analysis Home'** (points to the parameter selection area).

The output visualization has two parts. The first part consists of a hierarchical tree, which can be viewed in form of a radial or a linear tree. The default option renders the entire time period's in the displayed tree. Clicking on the “Choose a Condition” button will provide the user the option to generate the hierarchical tree for time periods bounded by the specific condition. The user also has

the option to view the tree as a radial tree, which is especially useful in analysing trees with a large number of nodes. Users can also find out the Phylum which the taxa belong to, and whether they are stationary (by clicking the “Overlay Phylum and Stationarity Information” button). The color of the taxa refers to the Phylum it belongs to; the Phylum and their corresponding colors are displayed below the tree. Additionally, the Non-Stationary taxa are easily discernible as their names are in a bolder font compared to the rest. Please note that some non standard text labels (containing symbols like '[' or ']') are automatically replaced by a dot (‘.’).

Right click on a branch origin node and select ‘Show trend plot’ to visualize a trend plot of all the descendent leaf nodes of the selected branch



The operations of the time series plot are similar to the one in Workflow 1.

Workflow 5b: Explore pair-wise relationship among taxonomic groups

The pair-wise similarity or dissimilarity among the time series of every pair of taxa can be visualised using this workflow (with the help of a heatmap). The user has to choose either the TIME DTW Distance or Correlation Coefficient as distance in order to quantify the similarity/dissimilarity. The three parameters the user has to select are the following three parameters: Algorithm, i.e. which distance measure should be used; “DTW Constraint” (as described in Workflow 5, required only when TIME DTW Distance is used) and “Cutoff” (as in other workflows). For more information, the user is requested to refer to the TIME paper. The output generated is shown below:

TIME

Temporal Insights into Microbial Ecology

HOME
Workflow 7
Antibiotic

Microbiome Source: E

Taxonomic Level: Genus

Select Algorithm

TIME DTW Distance

Select DTW Constraint (Only for Modified DTW)

Cutoff for Ignoring Rare Taxa

Plot

Pair-wise relationship among taxa with Heatmaps

Source: E; Taxonomic Level: Genus

Choose a Condition: Default: All

Download Result as CSV File

Download Image

Source: E; Taxonomic Level: Genus

Choose a Condition: Default: All

- SecondCp
- PreCp
- FirstCp
- Interim
- SecondWPC
- FirstWPC
- PostCp
- All

Select a desired time range to generate the tree. This range corresponds to the time ranges described in the metadata

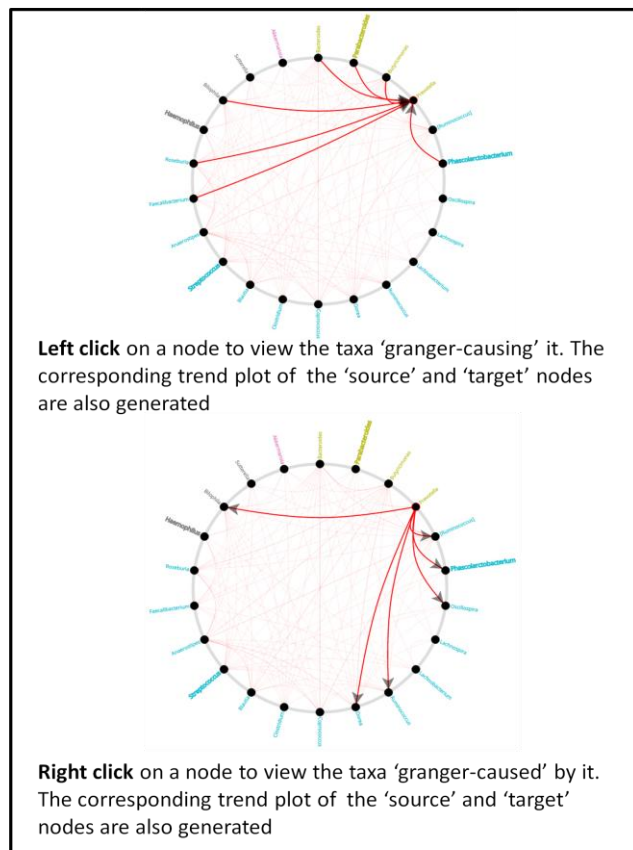
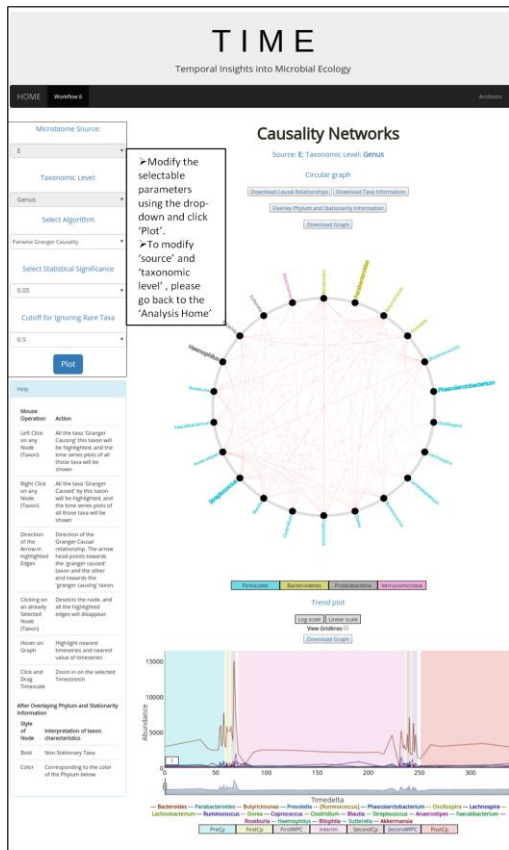
Similar to Workflow 5, the user can choose to focus the analysis on a selected Condition's time period. It should be noted that if there are only a few timepoints corresponding to a condition, then the results might lead to incorrect inferences. The columns of the heatmap are clustered hierarchically using the distance vectors, such that all the taxa showing similar distance with others are placed closer. The rows are ordered by the Phylum, i.e. the taxa belonging to the same Phylum are placed together. On hovering the mouse on any particular cell, the value of the distance measure pops up.

Workflow 6: Explore inter taxa interactions using causality network

In this workflow, three parameters need to be selected. The first parameter corresponds to the choice of algorithm for finding Causal relationships. There are three algorithm options, more information on which can be found in the TIME paper. The second parameter pertains to the statistical significance of the Granger Causality test. The third parameter refers to the “Cutoff” (as described in previous workflows).

Note: For large datasets, “Granger Lasso” and the “Granger Lasso Filtered” algorithms may take a very long time to run. The users should not close the tab while running the same. Once the parameters are selected and the “Plot” button is clicked, a circular graph is generated as shown in

the figure below. Similar to the previous workflow, the Phylum and Stationary information can be overlaid by clicking on the corresponding button. Left clicking on any node (taxon) shows the directed edges connecting the nodes (taxa) which ‘cause’ the selected taxa. Right clicking shows all the edges connecting to the nodes which are ‘caused’ by the selected taxa. Direction of the edges shows the causation event. For example if there is an edge between A and B and it is directed towards A, it means that the taxon B causes the taxon A. For more information about 'Causality', user is recommended to read the original TIME article.



Additionally, clicking on any node generates the time series plots of the node (taxon) and the connected (highlighted) nodes (taxa). The operations of the time series plot are similar to the previous ones.

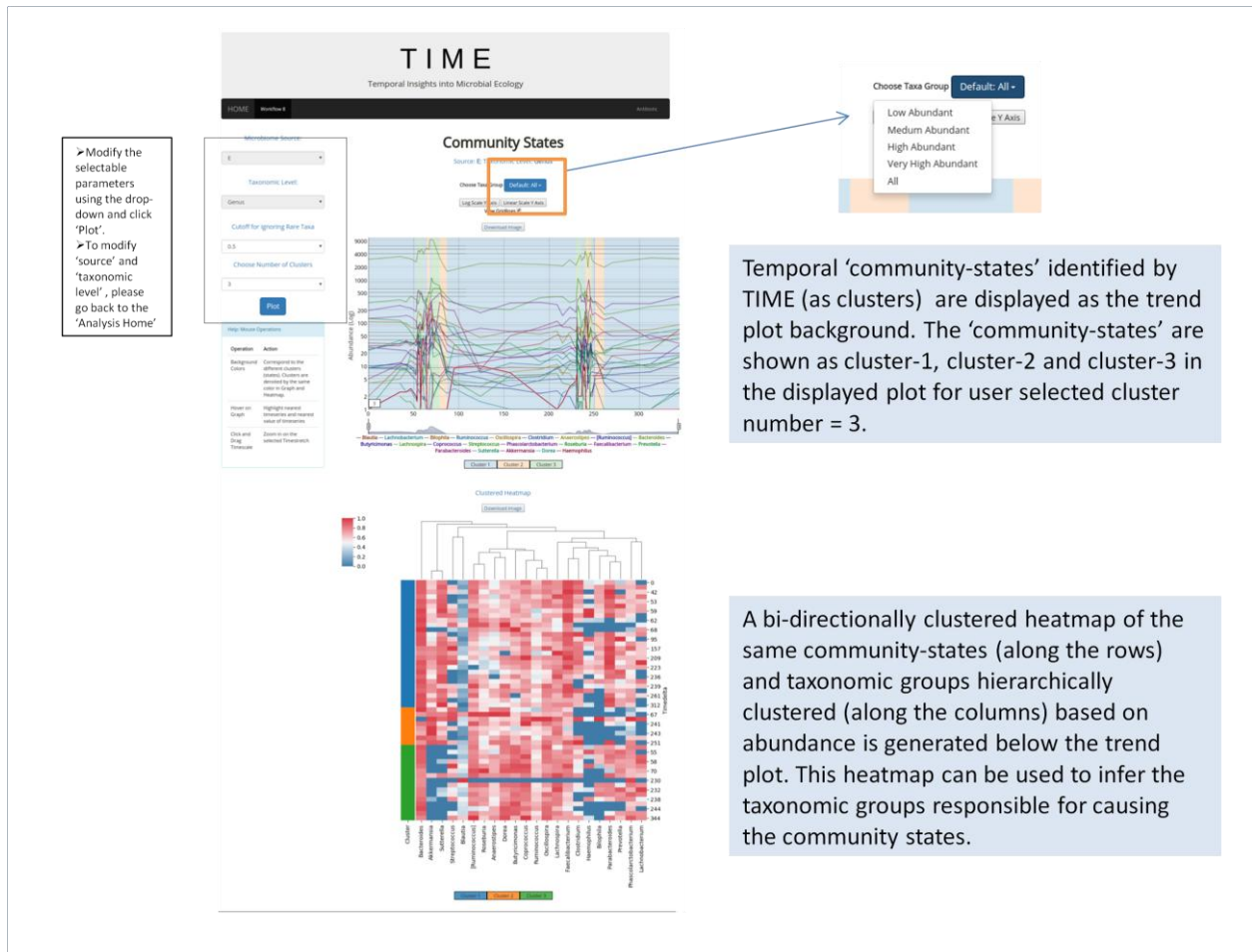
Workflow 7: Cluster time points based on similar community patterns

In this workflow, the user has to choose two parameters, namely “Cutoff” and “Number of Clusters” (number of community states). Clicking the “Plot” button generates a time series plot as

Cutoff for ignoring Rare Taxa

Choose Number of Clusters

shown below.



Temporal 'community-states' identified by TIME (as clusters) are displayed as the trend plot background. The 'community-states' are shown as cluster-1, cluster-2 and cluster-3 in the displayed plot for user selected cluster number = 3.

A bi-directionally clustered heatmap of the same community-states (along the rows) and taxonomic groups hierarchically clustered (along the columns) based on abundance is generated below the trend plot. This heatmap can be used to infer the taxonomic groups responsible for causing the community states.

The user has the option to view the "Community States" for taxa grouped according to their abundance into four groups: "Low", "Medium", "High" and "Very High", which lets the user easily explore the visualization and draw inferences. The Community States information is displayed in the time series plots and the corresponding information is available below. The user can also view the determinant taxa of each community state with the help of the heatmap. The heatmap helps in the identification of taxa which are differently abundant in the different Community States. The colors of the Clusters correspond to the color of the Clusters in the time series plot. Please note that both the cluster legends corresponds to the same 'community states' identified by TIME.

Contact

For any queries/troubleshoot or bug report please direct your mails to krishanu.dasbaksi@tcs.com or kuntal@atc.tcs.com