a) Required fields for submission

Refer Next Slide

File Type

Specify the type of file. TSV or CSV for Table format Min Biom or Rich Biom for Biom format

DATA TROUBLESHOOTING MODULE

Job Submission Wizard

Follow the step by step wizard for submitting a job to MetagenoNets

Upload Files

Upload required files in the respective fields to proceed with submission process

Info tags

Access the (i) tags to get information of type of file to be uploaded in respective section and also find a sample data file for reference

| STEP 1: DATA UPLOAD | STEP 2: SUMMARY | STEP 3: NORMALIZATION | FINALLY: APPROVE DASHBOARD |

Input Type

Options

TABLE   BIOM

○ TSV   ○ CSV

Job Label

Assign a job label to each job to be executed, for the convenience of recognising the context of the job.

Specify Job Label, like: IB

Node meta data available?   YES

Secondary input data available?   YES

PRIMARY **INPUT**

PRIMARY **METADATA**

NODE **METADATA**

SECONDARY **INPUT**

Status Terminal

Table (TSV) format is default selection. Please ensure that:

- **Abundance Data: Rows are features, Columns are samples.**
- **Metadata: Rows are Samples, Columns are Classes of Metadata.**

Please refer the tool tips against each file type for further assistance

Status terminal

Status terminal displays basic file requirements, summary of the uplaoded files and successful job submission

Data availability options

Select option "YES" only if the type of additional data is available to proceed with upload process. More information about importance of these data is provided in the respective (i) tags

⬆ UPLOAD AND PROCESS

Process Files

Click "UPLOAD AND PROCESS" button once required files are selected to process and validate the files for further analysis

NEXT

Job History

Access Jobs run by a user in chronological order (oldest first), wherein each JOB ID is prefixed with the JOB LABEL assigned by the user for the JOB (i.e Job Label: Job ID). Click on a specific job to recreate the dashboard

Data Troubleshooting Module of MetagenoNets facilitates the following features:

1. Handling special characters in the datasets
2. Extracting data form Biom format files
3. Imputing NA/missing/ NaN values

Upload file to process

Upload any file which needs management for format requirements

Troubleshoot Data

y manage special characters and NA (or missing value) issues pertaining to your data here.

Missing/ NA Value Imputation

Selected the option with which the Na/ missing values in the file to be imputed.

In case of MetaData, Missing or NA values are imputed as 'Unknown' for categorical data, and as Median, Mean or Zero for continuous data.

MANAGE SPECIAL CHARACTERS AND MISSING DATA IMPUTATION

✓ MEDIAN
MEAN
ZERO

DATA

ype?

MEDIAN

Imputation?

✓ ABUNDANCE DATA
METADATA
BIOM

SELECT INPUT

Go back to Job Submission Process

Status Terminal

UPLOAD AND PROCESS

Status terminal

Status terminal displays successful file processing status, summary of the uploaded files, and link to download the files

Data Type: Abundance Data/ Metadata

Files created when Abundance Data/ Metadata is selected to be processed for Special character handling/ Na value imputation

Data Type: Biom

Files created when Biom is selected to extract tab-delimited Data from the biom format file

Status Terminal

Data formatted!
Primary data uploaded
Data Processed: DOWNLOAD DATA

SUBMIT NEW

Note: Data will be auto purged within 2 minutes!

Status Terminal

Data formatted!
Primary data uploaded
Data Processed: Click below buttons to download

ABUNDANCE DATA

META DATA

NODE DATA

SUBMIT NEW

b) Sample Upload Page with all types of files available for network creation

Job Submission Wizard

step by step wizard for submitting a job to MetagenoNets

DATA TROUBLESHOOTING MODULE

Job Label

Relevant Job label is specified

STEP 1: DATA UPLOAD | STEP 2: SUMMARY | STEP 3: NORMALIZATION | FINALLY: APPROVE DASHBOARD

Input Type

TABLE | BIOM

Specify Job Label, like: **IBDnetwork** (required)

Community_network

Node meta data available? — YES

Secondary input data available? — YES

Options

● TSV ○ CSV

Primary_abun.txt — CHANGE

Metadata.txt — CHANGE

Node_metadata.txt — CHANGE

Secondary_Input.txt — CHANGE

Status terminal

Status terminal displays:

1. Job Label and Job ID
2. Type of error and the filename containing the error if any
3. Upload status
4. Statistics of the abundance files uploaded
5. Number as well as names of categorical and numerical categories in given metadata

Status Terminal

Job Label: **Community_network**

JOB ID: **d2311786**
Primary data uploaded
Primary Metadata uploaded
Node Metadata uploaded
Secondary Input uploaded
Data Statistics: (displayed on **Step 2** as well)

Input Type	Features	Samples
Primary	110	155
Secondary	306	155

Next step

Click next to proceed to next step In the upload process.

NEXT

Uploaded Files

All the four types of files are available to upload and generate network

Hence 'YES' option is selected for last two categories of files

c) Sample Upload Page with Secondary Input data not available for network creation

DATA TROUBLESHOOTING MODULE

Job Submission Wizard

Follow the step by step wizard for submitting a job to MetagenoNets

| STEP 1: DATA UPLOAD | STEP 2: SUMMARY | STEP 3: NORMALIZATION | FINALLY: APPROVE DASHBOARD |

Input Type

Options

● TSV   ○ CSV

TABLE   BIOM

Specify Job Label, like: **IBDnetwork** (required)

Community_net

Primary_abun.txt   📂 CHANGE ⓘ

Node meta data available?   **YES**

Metadata.txt   📂 CHANGE ⓘ

Secondary input data available?   **NO**

Node_metadata.txt   📂 CHANGE ⓘ

Status Terminal

Job Label: **Community_net**

JOB ID: **73aea99f**
Primary data uploaded
Primary Metadata uploaded
Node Metadata uploaded
Secondary Input not uploaded
Data Statistics:(displayed on **Step 2** as well)

Input Type	Features	Samples
Primary	110	155
Secondary	n.a	n.a

NEXT

Status terminal

Status terminal displaying the names of the files in red colour, which are not provided by the user by choice

Data availability options

"NO" option is selected if Secondary Input data is not available.

Simmilarly, if Node metadata is not available, "NO" option needs to be selected to proceed with submission process.

To get more information on the importance of these file categories click on the (i) tags

Job Submission Wizard

Follow the step by step wizard for submitting a job to MetagenoNets

| STEP 1: DATA UPLOAD | STEP 2: SUMMARY | STEP 3: NORMALIZATION | FINALLY: APPROVE DASHBOARD |

Summary of uploaded data

Total Samples: 155

InputData	Filename	Features
Primary	Primary_abun.txt	110
Secondary	Secondary_Input.txt	306

Total Metadata Headers: 5

MetaData	Number	Headers
Continous	2	BMI_values,Age
Categorical	3	Tribal_Urban,Geography,Gender

Summary of Metadata categories

Type of metadata is intelligently seggregated into categorical and continous data types.

Total no. Metadata fields, number of Continuos and Categorical metadata in the uploaded file is detected and displayed alongwith the column headers

PREVIOUS

Summary of Abundance data files uploaded

Summary of Primary and/ or Secondary Input data uploaded alongwith the number of samples and number of features present in each file.

NEXT

Upload again

Click 'PREVIOUS' button to go back and upload the files again

Next step

Click 'NEXT' to proceed to next step In the upload process.

JOB SUBMISSION GUIDE | Step3: Normalization

Job Submission Wizard

Follow the step by step wizard for submitting a job to MetagenoNets

STEP 1: DATA UPLOAD	STEP 2: SUMMARY	STEP 3: NORMALIZATION

Data Normalization, Filtration, Transformation

Normalization methods

Choose a normalization method to be applied on the input data if required or else choose "Don't Normalize" option

TOTAL SUM SCALING
- Don't Normalize
- Total Sum Scaling
- Cumulative Sum Scaling
- Quartile Normalization

Normalization

DON'T NORMALIZE

Normalization — You may choose to avoid normalization of data. Or MetaNets provides three popular data normalization techniques: Total Sum Scaling (TSS), Cumulative Sum Scaling (CSS) and Upper Quartile Normalization.

Prevalence threshold

Prevalence threshold — Prevalence refers to abundance of a feature (taxon) in a sample. In the context of a metagenomic sample collection, prevalence threshold refers to the minimum abundance of the taxon that must be there in the samples of the environment. MetaNets allows relative computation of prevalence (in terms of fraction of the highest normalized abundance) to avoid biases due to arbitrary choice of a minimum abundance. **You'll be able to modify this threshold later in your dashboard as well.**

Occurence threshold

Occurence threshold — Occurence threshold is closely connected to prevalence threshold. It refers to the minimum percentage of samples of the population in which a given taxon must prevail at the defined minimum prevalence threshold, to be considered for analysis. **You'll be able to modify this threshold later in your dashboard as well.**

Data Filtration

Choose a prevalence threshold for a feature to be selected between 0 to 0.1 of maximum abundance

Choose an occurrence threshold value for a feature between 0 to 100 percent of samples

Transformation

DON'T TRANSFORM

Transformation — A user may choose not to transform the data (especially when Normalization is already being considered). MetaNets provides for three well known data transformation techniques applied to microbiome data (considering the compositional nature of microbial abundance profiles): Centered log ratio (CLR), DeSeq2 / Relative Log Expression and EdgeR's Trimmed Mean (TMM)

Transformation methods

Choose a transformation method to be applied on the input data if required or else choose "Don't Transform" option

DON'T TRANSFORM
- Don't Transform
- Centered log ratio (CLR)
- DeSeq2 / RLE
- EdgeR's Trimmed Mean (TMM)

PREVIOUS

NEXT

Next step

Click 'NEXT' to proceed to next step In the upload process.

Job Submission Wizard

Follow the step by step wizard for submitting a job to MetagenoNets

| STEP 1: DATA UPLOAD | STEP 2: SUMMARY | STEP 3: NORMALIZATION | FINALLY: APPROVE DASHBOARD |

Following is a consolidated summary of your submission. Approve the submission for creating your personal dashboard.

Total Samples: 155

Total Metadata Headers: 5

InputData	Filename	Features
Primary	Primary_abun.txt	110
Secondary	Secondary_Input.txt	306

MetaData	Number	Headers
Continous	2	BMI_values,Age
Categorical	3	Tribal_Urban,Geography,Gender

If both Normalization and Transformation were selected, only Normalization will be performed.

Parameter	Choice
Normalization	Total Sum Scaling
Prevalence	atleast 0.00119 of max
Occurence	in atleast 10% of samples
Transformation	Don't Transform

PREVIOUS

APPROVE & SUBMIT

Uploaded File Summary

Summary of files upload and number of features and categories in the uploaded data files

Input Data Modification Summary

Summary of the normalization and transformation methods applied as well as filtration cutoff applied on the data

Submit Data

Click this button after approving the information provided as summary for reference, related to the data to be used for network creation

JOB SUBMISSION GUIDE | FINAL SUBMISSION DIALOG

Successful Submission Dialog

Once user clicks on "APPROVE & SUBMIT" button and data is successfully submitted, a pop-up with unique Dashboard ID is prompted

Note the unique Dashboard ID or it can be accessed in the Job history page along with the corresponding specified Job label.

This 8 character unique alphanumeric Job ID can be used to recreate the dashboard as and when required.

STEP 1: DATA UPLOAD   STEP 2: SUMMARY   STEP 3: NORMALIZATION   FINALLY: APPROVE DASHBOARD

Following is a consolidated summary of your submission. Approve the submission for creating your personal dashboard.

Total Samples: 155

InputData	Filename	
Primary	Primary_ab	values,Age
Secondary	Secondary_	I_Urban,Geography,Gender

If both Normalization and Transf

Parameter

Normalization	
Prevalence	atleast 0.00119 of max
Occurence	in atleast 10% of samples
Transformation	Don't Transform

Dashboard: d2311786 Created!

Your personal dashboard has now been created. Once you press OK, you will be taken into your Dashboard. d2311786is your Dashboard ID, which may be accessed in your job history as well. You can access your dashboard for upto 7 days using this unique ID. Use your job history page or the search widget in the menu above, to access your dashboard any time later.

OK

PREVIOUS

APPROVE & SUBMIT