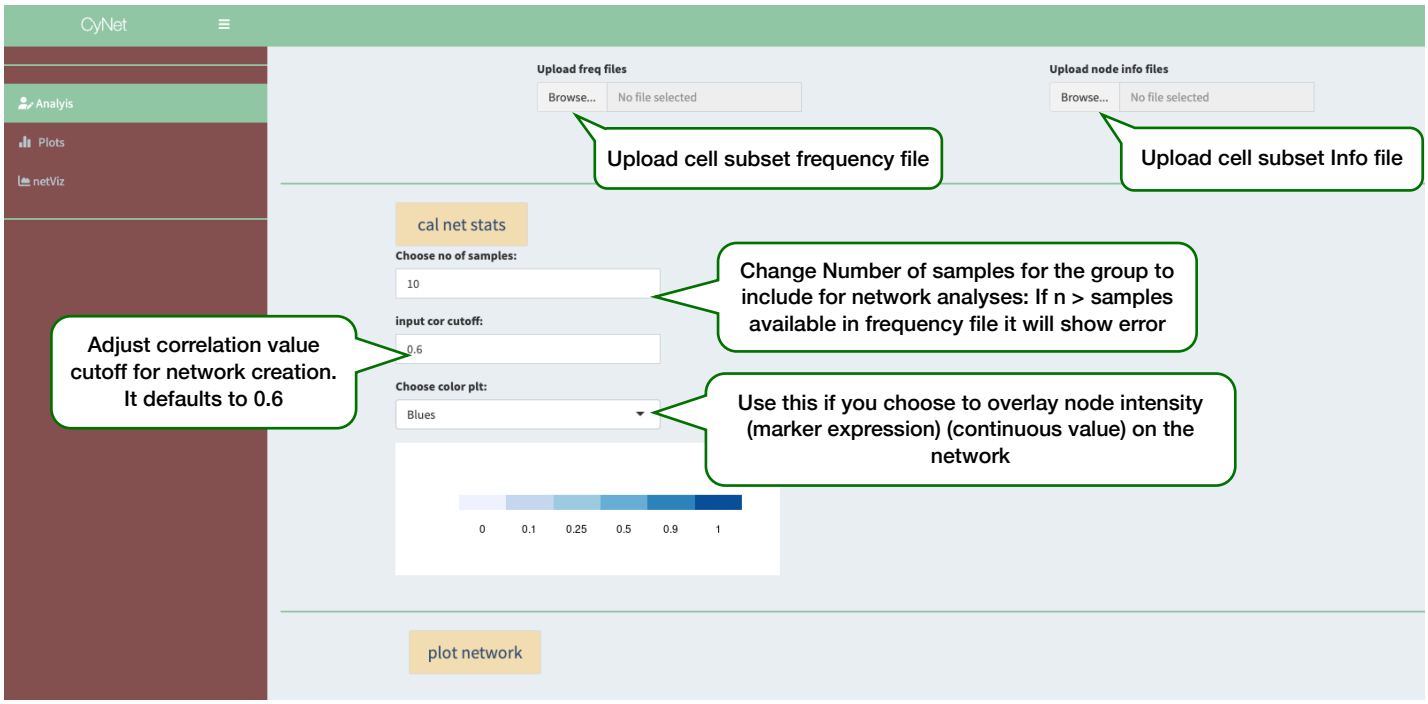
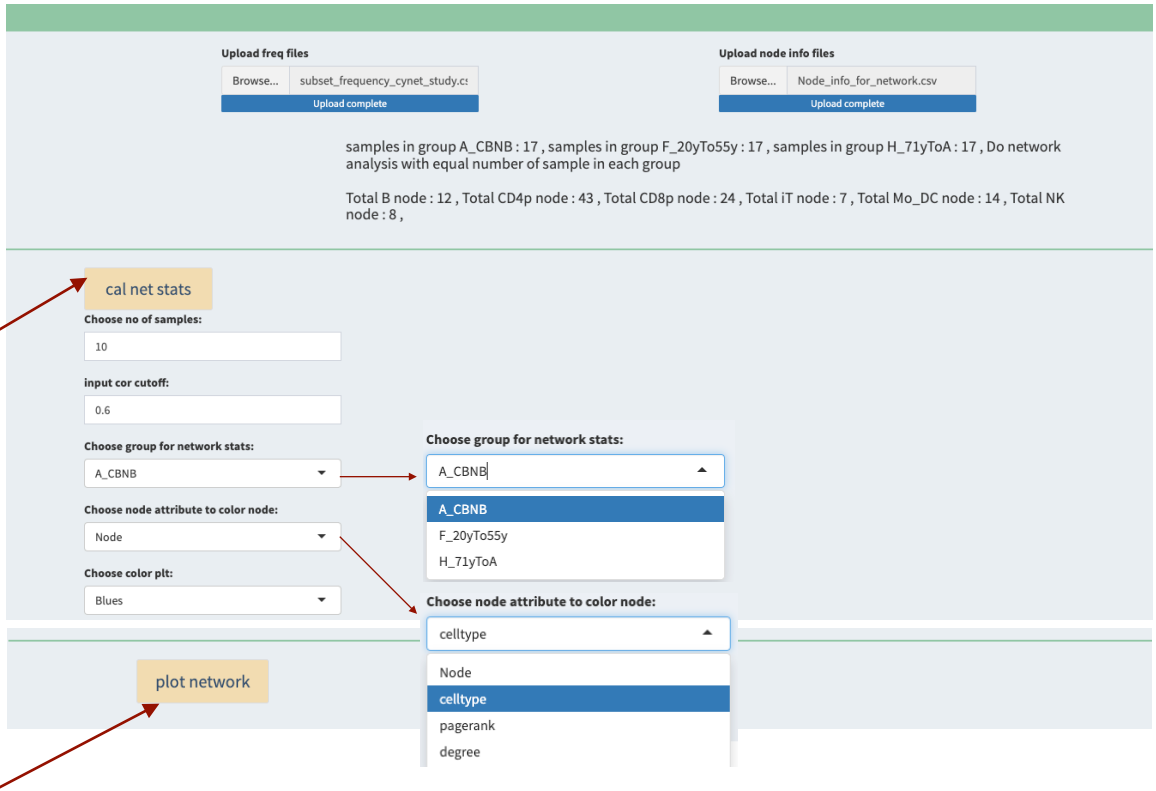
CyNET UI

https://epicimmuneatlas.org





Cal net stats: will calculate the network properties and will be shown as table.

Plot network: will plot the network. It can be downloaded as html file where node position can be adjusted for declustering some node

CyNET: requires two file for the network analysis of the cytometry data

1: Node frequency file
2: Node information file

Node frequency file:

Node frequency file can be generated from flow-cytometry as well as mass cytometry data. Usually mass cytometry panels are more than 35 marker that include most lineage markers. However if lab just have regular flow-cytometer with 12 color, they can stain using multiple panels each panel focussed on a single major lineage and then merge the cell frequency in one file from each lineage to create network using CyNET. Usually when using high dimensional mass/flow cytometry data we use clustering algorithms to group the cells in an unbiased manner to allow previously unknown cell subsets. Node frequency file must be formated as specified on page 3 of this mannual and should be in csv file format.

Node Information file:

Node information file allows overlaying information of nodes (cell subset) on the network diagram. Any additional information can be visualised on network. Node and celltype information is required columns. Node frequency file must be formated as specified on page 3 of this mannual. Node information file must be formated as specified on page 3 of this mannual and should be in csv file format.

Correlation cutoff value and number of samples in each group.

Correlation (R) value cutoff: it defaults to 0.6 to include only strongly correlated nodes. However different values can be tested. Suggest to keep it between 0.5-0.8. For number of sample in each group n<7 may not be very robust for correlation values. Statistically significant correlations are used only and the p-value is set to 0.05.

Positively correlated edges in the netwrok are shown in red color and negatively correlated edges are shown in green color

Outputs:

1: Network property table can be copied and saved.
2: Network diagram overlayed with node information can be saved and downlaoded as html file.
3: Node centrality bar graph and csv file can downloaded
4: Node pageRank centrality figure and csv file can be downloaded

Data format : for CyNET App

Cell subset Frequency file

group	sampleName	Node_1	Node_2	Node_3	Node_4	Node_5
A_CBNB	EPICC011	0.0064	0.008	0.0318	0.0559	0.0559
A_CBNB	EPICP010	0.0175	0.0047	0.0273	0.0787	0.0787
F_20yTo55y	EPICP005	0.0029	0.0104	0.01	0.0372	0.0372
F_20yTo55y	EPICP002	0.0044	0.0063	0.018	0.0528	0.0528
H_71yToA	EPICC004	0.0136	0.0119	0.0171	0.0449	0.0449
H_71yToA	EPICC005	0.0136	0.0119	0.0171	0.0449	0.0449

Format for frequency file:

The file must have group and sampleName columns. Names of subsets (nodes) must follow the convention Node_{node_id}. Once the file is uploaded group dropdown (Fig 2.) will appear on the app. Choose the group for network creation and analysis. Input the number of samples that you want to include. If sample number exceed the number of sample present in the uploaded file it will show error. If you are analysing multiple group it's better to keep number of samples balance in all groups. For eg. If you have three group with n number of sample, Healthy Control(n=20), Disease 1(n=17), Disease 2(n=15). We suggest to use n =15 to have equal number of samples in each group.

Node (cell subset) Information file

Node	celltype
Node_1	CD4p
Node_2	CD4p
Node_3	CD8p
Node_4	CD8p
Node_5	NK
Node_6	NK
Node_7	Bell

Node information file must have Node and celltype columns. Node columns should have all the nodes that are in Node frequency file. Celltype column should have major cell lineage for Nodes in Nodes columns Additional information can also be included in this file to overlay the node information onto the network. Marker expression values can also be plotted on network. All the columns in this file will have those columns in dropdown menu to choose. If the column values are continuous values then continuous color scale will be applied to overlay node information. However if the column values are categorical then discrete color for each category will be overlayed.