

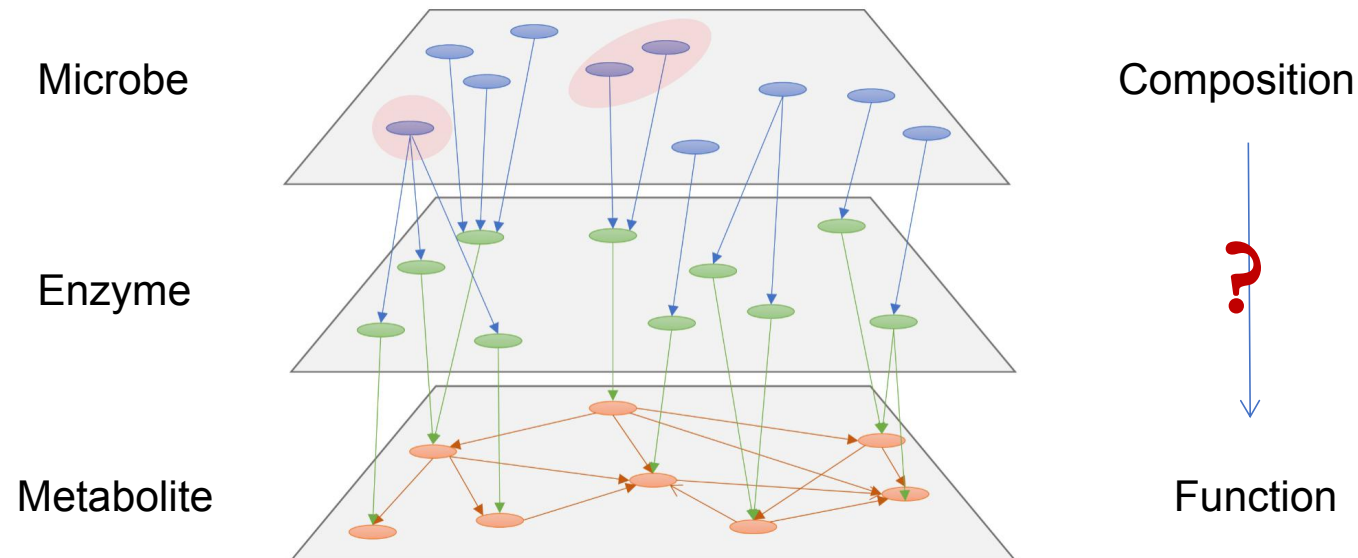
OmicNet Tutorial: Build network from a taxon list

Computer Requirement

- Modern browser supporting WebGL
- Chrome 50+, Firefox 47+, Safari 10.1+ and Edge 12+
- Please make sure WebGL is enabled in your browser
 - Please consult this web page to verify: <https://get.webgl.org/>
- If not enabled, please consult our FAQ page for instructions
- For best performance and visualization, use:
 - Latest version of Google Chrome
- A modern computer with at least 4GB of physical RAM
- A 15-inch screen or bigger (larger is better)
- Retina Display is supported

Motivation

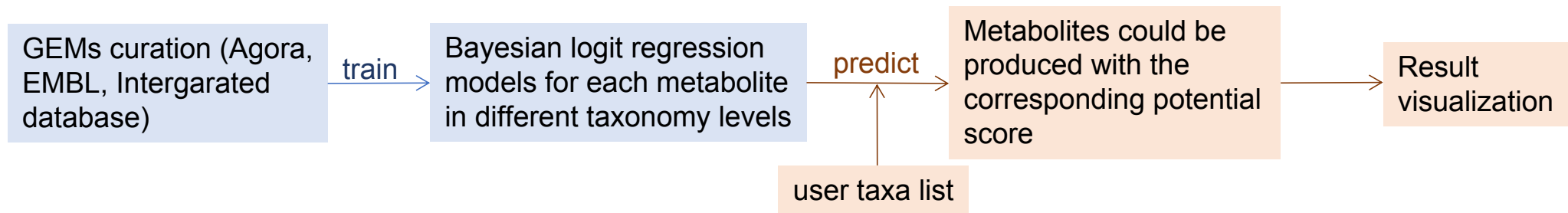
Microbiome is known to be involved in numerous host health effects ranging from host responses to cancer immunotherapy to metabolic disease and obesity. Metabolites produced by microbes are direct players in host-microbe interaction triggering downstream signals. Inferring the microbial metabolite features can facilitate the discovery of novel taxon-metabolite associations or serve as testable hypotheses for microbiome-metabolomics integration thus help to understand the underlying mechanism between host-microbiome crosstalk. However, tools that can predict microbial metabolites from a given microbiome are still lacking.



Overview

- **Goal:** This module is designed to decipher the metabolite production potential of the user interested taxa list to help understand the role of microbiome (currently focusing on human gut microbiome) in specific biological conditions.

- **Method and Procedure:**



- GEMs: Genome-scale metabolic models, describe the detailed microbial metabolism at strain level that can serve as knowledge background for model training.
- Bayesian logit regression model: provide the probabilities for classification of certain taxon as a producer of given metabolite
- Potential score: a prediction value based on Bayesian logit regression models indicating the possibility of a specific taxon to produce a given metabolite.

- **Result:** The prediction result will be present by either a composite network or an interaction heatmap for visualization and analysis

Data Upload

Multi-omics Integration via Biological Networks

Objective	Click on a panel below to start				
Explore networks in 2D or 3D space			A Graph File		
Annotate SNPs, taxa, or LC-MS peaks for network analysis		SNPs	Microbial Taxa		
Network analysis of one or more list(s) of molecules	Genes	Proteins	Transcription Factors	miRNAs	Metabolites

▶ Proceed ↺ Reset

Upload a list of taxon names

Enter your data below: ?

Taxon Name Level: Species Name ▾

Faecalibacterium_prausnitzii
Bacteroides_uniformis
Eubacterium_rectale
Alistipes_putredinis
Subdoligranulum_unclassified
Escherichia_coli
Bacteroides_vulgatus
Clostridium_clostridioforme
Klebsiella_pneumoniae
Clostridium_hathewayi
Alistipes_shahii
Ruminococcus_obeum
Roseburia_inulinivorans
Bacteroides_thetaiotaomicron

Use our example data

▲ Upload Cancel

Users can upload a **taxa list** from a specific taxonomy, e.g. a list of dyregulated taxa or the taxonomy annotation of the sequencing samples.

By selecting from the drop-down menu, users need to specified the **taxonomy level** which is necessary for the function prediction. 7 different levels are provided ranging from phylum to strain.

The **abundance** is optional. If it is not provided, 1 will be used as default value.

Database Selection

To construct the metabolic functional network from the input taxa list, user need to specify parameters to decide the metabolites involved including:

- **GEMs database:** used to predict the metabolite produced and their corresponding potential scores
- **Threshold for potential score:** score over 0.5 indicates the taxon is more likely to produce the given metabolite and the increasing score value means the greater production possibility.
- **Exclude metabolites:** users can choose to exclude currency metabolites, universal metabolites or the metabolites without pathway annotation

Database Selection

Databases are organized under different tabs. Please choose proper database(s) for network creation based on your analysis objectives. Multiple types of networks will be merged (based on shared nodes) and customized in the next page

Metabolite-protein **Taxon-metabolite**

Predicting Metabolic Potential of Microbial Taxa

The prediction is obtained based on logistic regression models trained based on high-quality genome-scale metabolic models (GEMs). The taxon-metabolite network could be further enriched by introducing protein-metabolite to find out potential enzymes. In the network viewer, you can click on icon located in the vertical toolbar for overview of the potential scores across all metabolites for your input taxa.

Potential score 0.9

AGORA AGORA GEMs (potential scores for 1110 metabolites)

EMBL EMBL GEMs (potential scores for 930 metabolites)

Excluding:

- Currency metabolites
- Universal metabolites
- Metabolites without pathway annotation

ally contains several disconnected subnetworks.

	Browse	Download	Delete
Microbial taxa			
Metabolite			

<< Previous < < 1 > >> Proceed >>

Currency metabolites: abundant substances such as water and carbon dioxide known to occur in normal functioning cells.

Universal metabolites: include currency metabolites and other metabolites shared across all taxa based on the GEMs databases.

Records of the individually generated network that users can trace, browse download or deleted accordingly

Network expansion

In the **database selection page**, users can also expand the taxon-metabolite network:

The taxon-metabolite network can be extended to include Metabolite-protein interactions by querying the related microbial databases.

Users can trace the networks built in different types individually.

Database Selection

Databases are organized under different tabs. Please choose proper database(s) for network creation based on your analysis objectives. Multiple types of networks will be merged (based on shared nodes) and customized in the next page

Metabolite-protein Taxon-metabolite

- [KEGG \(Organism-specific\)](#) Metabolite-protein interaction data based on all KEGG reactions (updated on 01/04/2022)
- [Recon3](#) High-quality genome-scale metabolic reconstruction (human) (updated on 01/04/2022)
- [AGORA](#) Agora based microbial metabolic reactions (updated on 01/12/2022)
- [EMBL](#) EMBL GEMs based microbial metabolic reactions (updated on 01/12/2022)
- [KEGG Generic](#) Non-organism specific metabolic reactions from KEGG metabolic network (updated on 01/04/2022)

[▶ Submit](#)

Individual Omics Networks

Each network is created independently by searching input list against a selected database. The network usually contains several disconnected subnetworks.

Input Type	Network Type	Sizes (node# - edge# - seed#)	Browse	Download	Delete
Microbial taxa	Taxon-metabolite	55 - 379 - 28	🗒	↓	🗑
Metabolite	Metabolite-protein	124 - 109 - 0	🗒	↓	🗑

<< < 1 > >>

[<< Previous](#) [Proceed >>](#)

Network interaction table

For each individual network, an interaction table is provided for users to dig into the details of their query results in :

Individual Omics Networks

Each network is created independently by searching input list against a selected database. The network usually contains several disconnected subnetworks.

Input Type	Network Type	Sizes (node# - edge# - seed#)	Browse	Download	Delete
Microbial taxa	Taxon-metabolite	55 - 379 - 28			
Metabolite	Metabolite-protein	124 - 109 - 0			

Note: Query ID list is the input taxon for the first network and result metabolites can be used as queries in the following networks, and so forth.

Individual Network Interaction Table

You can browse, search or manual delete an interaction (edge), or use the **Advanced Filter** to exclude a node (and its all associated edges).

Advanced Filter

Id1 ↓	Query ID	Id2 ↓	Result ID	Name1 ↓	Query Name	Name2 ↓	Result Name	Action
Roseburia hominis		C04691		Roseburia hominis		2-Dehydro-3-deoxy-D-arabino-heptonate 7-phosphate		Delete
Roseburia hominis		C02637		Roseburia hominis		3-Dehydroshikimate		Delete
Roseburia hominis		C03232		Roseburia hominis		Phosphohydroxypyruvic acid		Delete
Roseburia hominis		C01269		Roseburia hominis		5-O-(1-Carboxyvinyl)-3-phosphoshikimate		Delete
Roseburia hominis		C06056		Roseburia hominis		4-Hydroxy-L-threonine		Delete
Roseburia hominis		C00979		Roseburia hominis		O-Acetyl-L-serine		Delete
Roseburia hominis		C05774		Roseburia hominis		Cobinamide		Delete
Roseburia hominis		C00522		Roseburia hominis		(R)-Pantoate		Delete
Roseburia hominis		pgp161		Roseburia hominis		Phosphatidylglycerophosphate (dihexadec-9-enoyl, n-C16:1)		Delete
Roseburia hominis		C01102		Roseburia hominis		O-Phosphohomoserine		Delete
Roseburia hominis		C02876		Roseburia hominis		Propanoyl phosphate		Delete
Roseburia hominis		C00493		Roseburia hominis		Shikimate		Delete
Roseburia hominis		C09332		Roseburia hominis		Tetrahydrofolyl-[Glu](2)		Delete
Bacteroides uniformis		C04691		Bacteroides uniformis		2-Dehydro-3-deoxy-D-arabino-heptonate 7-phosphate		Delete
Bacteroides uniformis		C02637		Bacteroides uniformis		3-Dehydroshikimate		Delete

Navigation: << Previous (1 of 26) 1 2 3 4 5 6 7 8 9 10 >> 15 Proceed >>

Users can tailor their network by deleting the results they are not interested in

Network Building

After generating the network individually, users will be lead to the summary of the integrated network:

The screenshot shows the 'Multi-omics Network Building' interface. On the left, a 'Network Tools' sidebar contains buttons for: Degree Filter, Betweenness Filter, Minimum Network, Steiner Forest (PCSF), Tissue Filter, P-value Filter, Zero-order Network, and Reset to First-order. The main content area has a title 'Multi-omics Network Building' and two paragraphs of text. Below the text is a table with columns: Networks, Sizes (node# - edge# - seed#), Topology, and Download (edge list). The table contains one row for 'subnetwork1' with sizes '420 - 966 - 28' and a 'Details' link. Below the table are navigation buttons: '<< Previous', '1', '>>', and 'Proceed >>'. Three callout boxes are present: one pointing to the 'Network Tools' sidebar, one pointing to the 'Details' link in the table, and one pointing to the 'Proceed >>' button.

Network Tools ⓘ

- Degree Filter
- Betweenness Filter
- Minimum Network
- Steiner Forest (PCSF)
- Tissue Filter
- P-value Filter
- Zero-order Network
- Reset to First-order

Multi-omics Network Building

If more than one network was generated in the previous page, they are merged together to form multi-omics network through shared nodes. In some cases, the result will contain a larger subnetwork containing most of queries, and several smaller subnetworks containing one or a few queries. These subnetworks will be available for visual analysis in the next page.

If the network is too large, it is recommended to trim the network to a smaller size that is suitable for visual analytics (less than 2000 nodes), you can use network tools located on the left hand side for trimming and filtering.

Networks	Sizes (node# - edge# - seed#)	Topology	Download (edge list)
subnetwork1	420 - 966 - 28	Details	Download

<< < 1 > >>

<< Previous

Proceed >>

Methods for pruning and tailoring the network for better visualization and enhanced functional interpretation

Detail topology information of the final integrated network.

Proceed for visualization

Network Visualization

Network: subnetwork1 Background: Black Layout: Concentric circle Styling: - Specify - Scope: -- Specify -- Download: -- Specify -- More Options

Global Node Styles

Type	Size	Color
Microbe	<input type="range"/>	<input type="checkbox"/> Green
Metabolite	<input type="range"/>	<input type="checkbox"/> Yellow
Gene/Protein	<input type="range"/>	<input type="checkbox"/> Red

Node Explorer

ID	Name	Degree	Betweenness	Exp
<input type="checkbox"/>	C06056	4-Hydroxy-L-threonine	43	2245.137
<input type="checkbox"/>	C02637	3-Dehydroshikimate	34	831.7906
<input type="checkbox"/>	C00493	Shikimate	33	610.8937
<input type="checkbox"/>	C00979	O-Acetylserine	33	910.3452
<input type="checkbox"/>	C01102	O-Phosphoserine	32	648.7705
<input type="checkbox"/>	C01269	5-O-(1-Ca)phosphoserine	30	469.3452
<input type="checkbox"/>	C00522	(R)-Pantoic acid	29	596.7222
<input type="checkbox"/>	C03232	Phosphoserine	29	285.2545
<input type="checkbox"/>	C04691	2-Dehydroshikimate	29	380.9539
<input type="checkbox"/>	C09332	Tetrahydroshikimate	29	320.3452
<input type="checkbox"/>	C02876	Propanoic acid	28	450.8099
<input type="checkbox"/>	Escherichia coli	Escherichia coli	27	4331.905
<input type="checkbox"/>	C05774	Cobinamide	26	293.9515
<input type="checkbox"/>	pgp161	Phosphoserine	24	698.9325
<input type="checkbox"/>	Faecalibacterium prausnitzii	Faecalibacterium prausnitzii	22	2559.189
<input type="checkbox"/>	C00254	Prephenate	18	2280.579

Current Selections

Name: 4-Hydroxy-L-threonine
ID: C06056

- Vellonella parvula:0.9142775
- Alistipes shahii:0.9142775
- Bacteroides ovatus:0.9196845
- Alistipes putredinis:0.9142775
- Vellonella dispar:0.9142775
- Bacteroides uniformis:0.9142775
- Bacteroides finegoldii:0.9142775

Network Visualization

Functional Pathway Enrichment

Query: All nodes
Database: KEGG (Metabolite) Submit Save

Name	Hits	P-val	P-val(adj.)	Color
Phenylalanine, tyrosine and tryptophan biosynthesis	5	2.61e-7	0.0000952	
Folate biosynthesis	5	0.0000031	0.000577	
Arginine biosynthesis	3	0.000111	0.0135	
Biosynthesis of secondary metabolites	4	0.000313	0.0286	
Glycine, serine and threonine biosynthesis	3	0.00114	0.0829	
Cysteine and methionine metabolism	3	0.00222	0.135	
Biosynthesis of plant secondary products	4	0.00262	0.137	
Vitamin B6 metabolism	2	0.00601	0.274	
Biosynthesis of phenylpropanoids	3	0.00884	0.358	
mTOR signaling pathway	1	0.0167	0.555	

Module Explorer

Algorithm: InfoMap Submit Save

Module Size P-value Color

Result Panel

Name: 4-Hydroxy-L-threonine
ID: C06056

- Vellonella parvula:0.9142775
- Alistipes shahii:0.9142775
- Bacteroides ovatus:0.9196845
- Alistipes putredinis:0.9142775
- Vellonella dispar:0.9142775
- Bacteroides uniformis:0.9142775
- Bacteroides finegoldii:0.9142775

Navigation

Previous Proceed

Node type in the network

Node list in the network

Result panel provides the microbial producers of the selected metabolite with their potential scores.

Bigger taxon or metabolite nodes indicate the underlying taxon or metabolite that have higher probability to produce or to be produced.

Functional pathway enrichment

The End
