

Patient - patient tutorial

Graph Edges (must be uploaded): A file containing the interactions between nodes.

Graph Nodes (must be uploaded): A file containing node attributes.

Node Group: A file containing node groups for single or multiple network visualizations.

Data 1: A file containing continuous values for the outermost ring.

Data 2: A file containing continuous values for the second ring.

Data 3: A file containing integer values for the third ring.

Data 4: A file containing integer values for the innermost ring.

Sample Group: A file containing group information for individual bars in the ring (e.g., genes in a patient similarity network or patients in a gene-gene interaction network).

1. Overview

The Patient–Patient interaction network is used to model similarity relationships between patient samples (TCGA samples).

Each node represents one patient, and each edge represents the similarity between two patients based on multi-omics data.

If patient-to-patient interaction data are available, users should follow the Patient–Patient input file template described below.

The required input files include:

- patients_example_edges_patients.csv
- patients_example_nodes.csv
- patients_example_megList.csv
- patients_example_expression.csv
- patients_example_methylation.csv
- patients_example_cnv.csv
- patients_example_snv.csv

2. Input File Descriptions

A. patients_example_edges_patients.csv — Patient Similarity Edges

This file defines pairwise similarity relationships between patients.

Each row represents one directed edge between two patient samples.

Columns:

- **from**
The source patient TCGA sample ID in the network.
- **to**
The target patient TCGA sample ID.
The edge direction is defined as from → to.
- **weight**
The similarity score between the two patients.
This value is calculated using a kernel-based similarity algorithm.
Higher values indicate greater similarity between patients.

This file determines which patients are connected and how similar they are.

from	to	weight
TCGA-E2-A1IN-01A	TCGA-AQ-A0Y5-01A	0.001202
TCGA-BH-A0H9-01A	TCGA-AQ-A0Y5-01A	0.001773
TCGA-A2-A0YT-01A	TCGA-AQ-A0Y5-01A	0.002035
TCGA-D8-A1JA-01A	TCGA-AQ-A0Y5-01A	0.002168
TCGA-BH-A204-01A	TCGA-AQ-A0Y5-01A	0.001091
TCGA-D8-A1XZ-01A	TCGA-AQ-A0Y5-01A	0.001216

B. patients_example_nodes.csv — Patient Nodes

This file defines the set of patient samples included as nodes in the network.

Each row corresponds to one patient.

Columns:

- **name**
The TCGA sample ID of the patient.

All patient IDs appearing in the edge file should also be listed in this file.

name
TCGA-3C-AAAU-01A
TCGA-3C-AALI-01A
TCGA-3C-AALJ-01A
TCGA-3C-AALK-01A
TCGA-4H-AAAK-01A
TCGA-5L-AAT0-01A
TCGA-5T-A9QA-01A
TCGA-A1-A0SB-01A
TCGA-A1-A0SE-01A

C. patients_example_megList.csv — Patient Community Assignment

This file assigns patients to user-defined communities (groups).

Each row represents one patient.

Columns:

- **name**
Patient TCGA sample ID.
- **community**
Community identifier (integer).

If no prior community or group information is available, users may assign arbitrary integer values as placeholders (e.g., all patients assigned to community 1). But if all the gene names come from one network without groups separation, please just assign one arbitrary integer.

name	community
TCGA-3C-AAAU-01A	1
TCGA-3C-AALI-01A	1
TCGA-3C-AALJ-01A	1
TCGA-3C-AALK-01A	1
TCGA-4H-AAAK-01A	1
TCGA-5L-AAT0-01A	1

D. patients_example_expression.csv — Gene Expression Matrix

This file provides gene expression profiles for each patient.

- **Rows:** Gene symbols
- **Columns:** TCGA sample IDs (patients)
- **Values:** Log2-transformed, normalized gene expression values

Each row represents the expression level of one gene across multiple patients.

	TCGA-AQ-A0Y5-01A	TCGA-C8-A274-01A	TCGA-B6-A1KC-01B	TCGA-AC-A62V-01A	TCGA-A2-A0YM-01A
TSPAN6	0.642019	1.435314	0.905893	0.680979	1.27382
DPM1	1.590688	1.520711	1.46089	1.614744	1.436365
SCYL3	0.731065	1.010347	0.716954	0.339591	0.4929
FGR	0.327584	0.295589	0.23573	0.393838	0.695902
GCLC	0.729861	0.5792	0.634024	0.541067	0.931717
NFYA	1.189431	1.274742	1.395498	0.81909	1.55423
SEMA3F	1.342211	1.309907	1.946764	1.7679	0.962966

E. patients_example_methylation.csv — DNA Methylation Matrix

This file contains **gene-level DNA methylation data for patients**.

The format is **identical to the gene expression file**:

- **Rows:** Gene symbols
- **Columns:** TCGA sample IDs
- **Values:** Normalized methylation levels

	TCGA-AQ-A0Y5-01A	TCGA-C8-A274-01A	TCGA-B6-A1KC-01B	TCGA-AC-A62V-01A	TCGA-A2-A0YM-01A
TSPAN6	0.627609	0.44052	0.2535	0.522068	0.243916
DPM1	0.048339	0.048328	0.068088	0.043229	0.041871
SCYL3	0.186311	0.174254	0.216154	0.209671	0.205538
FGR	0.64619	0.734972	0.69399	0.723035	0.646808
GCLC	0.323994	0.337164	0.330322	0.36901	0.365856
NFYA	0.105899	0.113769	0.14041	0.109719	0.105975
SEMA3F	0.407646	0.437245	0.39831	0.462195	0.447582

F. patients_example_cnv.csv — Copy Number Variation (CNV)

This file describes copy number variation states of genes in patient samples.

- **Rows:** Gene symbols
- **Columns:** TCGA sample IDs

Values:

- 0 : Copy-number neutral
- 1 : Copy-number gain
- -1: Copy-number loss

Each row represents CNV profiles of one gene across multiple patients.

	TCGA-AQ-A0Y5-01A	TCGA-C8-A274-01A	TCGA-B6-A1KC-01B	TCGA-AC-A62V-01A
UPF1	-1	1	1	-1
KRIT1	0	0	0	-1
CREBBP	1	1	-1	1
MPO	-1	0	1	1
POLDIP2	1	0	1	0
KDM1A	-1	-1	-1	0
SCYL3	1	1	1	0
ALS2	-1	1	-1	-1
LAP3	-1	0	-1	1
HOXA11	0	-1	0	1

G. patients_example_snv.csv — Single Nucleotide Variations (SNV)

This file records mutation information for genes in patient samples.

The file format is the same as the CNV file:

- **Rows:** Gene symbols
- **Columns:** TCGA sample IDs

Values:

- 0 : No mutation
- 1 : One mutation
- 2 : Two mutations

In downstream analysis, any value greater than 0 is treated as mutation present, and the gene is labeled as a mutation gene.