

# Gene - gene 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

This template supports two types of network analyses: Gene–Gene interaction networks and Patients–Patients interaction networks.

If gene–gene interaction data are available, users should follow the Gene–Gene input file template described below.

The Gene–Gene interaction analysis requires the following input files:

- example\_graph\_edges.csv
- example\_graph\_nodes.csv
- example\_megList.csv
- example\_gene\_expression.csv
- example\_methylation.csv
- example\_cnv.csv
- example\_snv.csv
- example\_stage.csv (optional)

## 2. Input File Descriptions

### 2.1 Gene–Gene Interaction Network Files

#### A. example\_graph\_edges.csv — Gene Interaction Edges

This file defines the directed interactions between genes in the network.

Each row represents one interaction (edge) between two genes.

Columns:

- **from**  
The source gene symbol (HGNC-approved gene name).
- **to**  
The target gene symbol. The interaction direction is defined as from → to.
- **interact**  
A label indicating the interaction type.  
This column is optional and can be used for categorization or visualization.
- **weight**  
The interaction strength between the two genes. Values are typically calculated using Pearson correlation-weighted measures. Positive values indicate positive association, while negative values indicate inverse association.

| from    | to    | interact | weight   |
|---------|-------|----------|----------|
| DNAJC14 | AGTR1 | 0        | 0.09132  |
| DNAJC14 | DRD1  | 0        | 0.02288  |
| DBN1    | PTEN  | 0        | -0.1533  |
| NCOA3   | PTEN  | 0        | 0.05806  |
| UBE2I   | PTEN  | 0        | -0.24979 |
| DLC1    | PTEN  | 0        | 0.52956  |
| PRDX1   | PTEN  | 0        | -0.2439  |
| PTEN    | MME   | 1        | 0.40162  |
| NDFIP1  | PTEN  | 0        | 0.08153  |
| SHARPIN | PTEN  | 0        | -0.24187 |
| PARK7   | PTEN  | 0        | -0.10611 |
| MVP     | PTEN  | 0        | -0.1284  |
| USP10   | PTEN  | 0        | -0.16491 |
| MCRS1   | PTEN  | 0        | -0.26287 |
| USP13   | PTEN  | 1        | 0.0504   |
| PTEN    | WWP2  | 0        | -0.08772 |

## B. example\_graph\_nodes.csv — Gene Nodes

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

Each row corresponds to one gene.

Columns:

- **name**  
The gene symbol used as a node identifier in the network.

All genes appearing in example\_graph\_edges.csv should also be listed in this file.

| name    |
|---------|
| DNAJC14 |
| DBN1    |

|       |
|-------|
| NCOA3 |
| UBE2I |
| DLC1  |
| PRDX1 |

### C. example\_megList.csv — Gene Community Assignment

This file assigns each gene to a functional subnetwork (community).

Each row represents one gene.

Columns:

- **gene**  
Gene symbol.
- **community**  
Community identifier (integer).  
Communities represent smaller functional modules within the full gene network.

Communities can be generated using algorithms such as the spin-glass method. If no predefined community information is available, users may assign arbitrary integer value as placeholders. But if all the gene names come from one network without groups separation, please just assign one arbitrary integer.

| gene     | community |
|----------|-----------|
| DNAJC14  | 1         |
| IGFBP2   | 1         |
| GGA3     | 1         |
| NMT1     | 1         |
| SYPL1    | 1         |
| SLC9A3R2 | 1         |
| AGTR1    | 1         |

### D. example\_gene\_expression.csv — Gene Expression Matrix

This file provides gene expression data across patient samples.

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

Each row represents the expression profile of one patient across multiple genes.

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

## E. example\_methylation.csv — DNA Methylation Matrix

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

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

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

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

## F. example\_cnv.csv — Copy Number Variation (CNV)

This file describes copy number variation states for genes across patient samples.

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

Values:

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

Each row represents CNV profiles for one patient.

|                  | ACAP3 | ACTRT2 | AGRN | ANKRD65 | ATAD3A | ATAD3B |
|------------------|-------|--------|------|---------|--------|--------|
| TCGA-3C-AAAU-01A | 0     | 0      | 0    | 0       | 0      | 0      |
| TCGA-3C-AALI-01A | -1    | -1     | -1   | -1      | -1     | -1     |
| TCGA-3C-AALJ-01A | -1    | -1     | -1   | -1      | -1     | -1     |
| TCGA-3C-AALK-01A | 0     | 0      | 0    | 0       | 0      | 0      |
| TCGA-4H-AAAK-01A | 0     | 0      | 0    | 0       | 0      | 0      |
| TCGA-5L-AAT0-01A | 0     | 0      | 0    | 0       | 0      | 0      |

## G. example\_snv.csv — Single Nucleotide Variations (SNV)

This file records mutation counts per gene for each patient.

The format is the same as the CNV file:

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

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.

|                  | PIK3CA | DSP | PITPNM1 | USP15 | ACTG1 | NRCAM |
|------------------|--------|-----|---------|-------|-------|-------|
| TCGA-EW-A2FW-01A | 0      | 0   | 0       | 0     | 0     | 1     |
| TCGA-AR-A0U0-01A | 0      | 0   | 0       | 0     | 0     | 0     |
| TCGA-AR-A24Q-01A | 1      | 0   | 0       | 0     | 1     | 0     |
| TCGA-D8-A27R-01A | 0      | 0   | 0       | 0     | 0     | 0     |
| TCGA-D8-A1XR-01A | 0      | 0   | 0       | 0     | 0     | 0     |
| TCGA-B6-A0WV-01A | 0      | 0   | 0       | 0     | 0     | 0     |

## H. example\_stage.csv — Clinical Stage Information (Optional)

This file defines clinical stage or category information for patient samples.

- Each TCGA sample ID is assigned to a cancer stage group.
- Multiple samples may belong to the same stage.

This file is primarily used for visualization and stratified analysis.