

## TABLE OF CONTENTS

Cat-E: A Comprehensive Web Tool for Exploring Cancer Targeting Strategies.....	3
Introduction.....	3
Oncolytic Virus Database.....	5
Differential Expression Analysis.....	7
Metabolic Pathway Analysis.....	11
Metabolic Flux Analysis.....	14
GO Enrichment Analysis.....	18
KEGG Enrichment Analysis.....	20
Clinical Survival Analysis.....	22
Immune Signature Analysis.....	24
Single Nucleotide Variation Analysis.....	25
Gene Regulatory Network Analysis with Jimena.....	27
Protein Structure- AphaFold.....	30
Protein Structure- 3Dmol.js.....	31
Cell line Database.....	32
Cell type Database.....	33
Drug Description.....	34
Drug-Tissue Database.....	35
Drug-Gene Database.....	35
Immune Modulatory Database.....	36
Bispesific Antibody.....	36
CAR-T Cell Therapy.....	36
Checkpoint.....	37
Cytostatic Therapy.....	37
Oncolytic Virus Therapy.....	38
HOTLIST.....	38
HOT PAPERS.....	39
References.....	40

# Cat-E: A Comprehensive Web Tool for Exploring Cancer Targeting Strategies

## Introduction

Cancer remains a central area of focus for human geneticists, persisting as a subject of intense inquiry due to its intricate nature. Furthermore, despite the prevalence of cancer diagnoses, there are still numerous cancers for which the etiology remains unknown, and the optimal therapeutic approaches remain elusive.

Prominent resources in the field of cancer-related RNA sequencing, such as The Cancer Genome Atlas (TCGA) and Genotype-Tissue Expression (GTEx), have significantly expanded opportunities for data analysis and an enhanced comprehension of gene functionalities (<https://tcga-data.nci.nih.gov/tcga> ; <https://gtexportal.org/home/>). With the availability of these data to the research community, a multitude of methodologies and protocols for scrutinizing extensive gene expression datasets have emerged, with some subsequently attaining status as conventional approaches for interrogating such information. For instance, survival analysis has emerged as a valuable tool for discerning associations between gene expression levels and prognostic outcomes, thereby finding widespread application in the evaluation of a given gene's clinical relevance. For instance, in the context of differential gene expression analysis using TCGA data, web tools such as TSVdb (Sun et al., 2018), Gepia2 (Tan et al., 2019), and cBioPortal (Gao et al., 2013) have been developed, allowing researchers to explore gene expression variations in cancer samples.

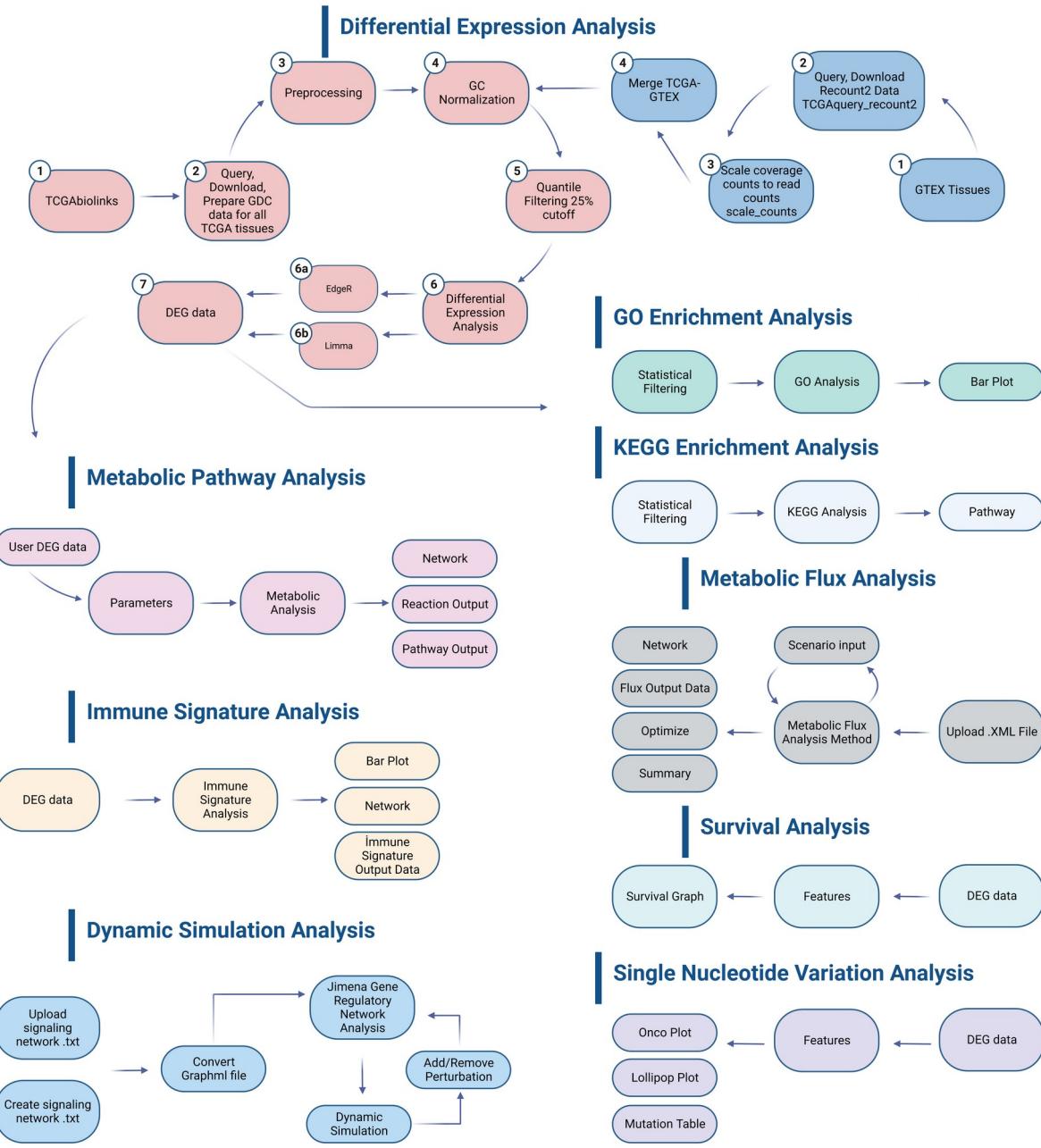
Additionally, TCGA data serves as a valuable resource for investigating gene mutations in the context of cancer research. Web tools like IntOGen (Gonzalez-Perez et al., 2013) and Cancer3D (Porta-Pardo et al., 2015) have been instrumental in identifying and characterizing genetic mutations associated with various cancer types. For example, Shiny GATOM (Emelianova et al., 2022) serves as an illustration of a Shiny web application designed for conducting metabolic analysis, while MIMOSA2 (Noecker et al., 2022) takes a comprehensive approach, simplifying the modeling and evaluation of relationships between microbiota members and their metabolic products.

In order to streamline the user-friendly and interactive application of these methodologies to TCGA and GTEx datasets, Cat-E was developed.

*Cat-E Basic functionality:* A comprehensive database was meticulously constructed, encompassing data pertaining to oncolytic viruses and their corresponding host cells. This resource also presents an extensive array of protein-protein interactions between human proteins and various types of oncolytic viruses. Users are afforded the convenience of straightforwardly searching by cell line and cell types within this database. Additionally, Cat-E involves information and illustrations of protein structures, drug descriptions, drug-gene, and drug-tissue relationships. Moreover, the available clinical trials of immune modulatory interventions.

*Cat-E Additional functionalities:* Equipped with features such as differential expression analysis, GO (Gene Ontology) enrichment analysis, metabolic pathway analysis, metabolic flux analysis, immune signature analysis, survival analysis, single nucleotide variation analysis the identification of genes with similar profiles, and dynamic gene regulatory network analysis (Jimena).

*Use cases:* Cat-E offers a robust database and functionalities aimed at streamlining the utilization of critical data (e.g., bispecific antibody therapies, CAR-T cell therapies, checkpoint inhibitors, cytostatic agents, and oncologic treatments) by experimental and clinical biologists. This software has played a pivotal role in the exploration of various genes across diverse cancer types, ultimately contributing to the identification of potential biomarkers and therapeutic targets.



# Oncolytic Virus Database

## ① Accessing the Oncolytic Viruses Tab (Fig. 1)

- To begin, follow these steps:
- On the left-hand side of the screen, locate the tab marked with the number ①
- Click on this tab to access the oncolytic viruses section.

## ② Selecting the Virus Type of Interest (Fig. 1)

- In the oncolytic viruses section, you will find a selection input area numbered ② Here's how to proceed:
- Navigate to the selection input area, which allows you to choose the specific virus type you are interested in.
- Click on the dropdown menu and select the virus type that you want to explore further.

## ③ Accessing Relevant Data (Fig. 1)

- Once you have selected your desired virus type, the web app will present you with relevant data in a table format. Now, let's explore how to download this data:
- Look for the section labeled ③
- Here, you will find options to download the table in various formats, including CSV, PDF, and Excel.
- To download the data in your preferred format, simply click on the corresponding download option.

## ④ Searching Within the Table (Fig. 1)

- To make your experience more efficient, our web app provides a search functionality within the table. Here's how to use it:
- Locate the search section labeled ④
- Enter your search query in the provided search bar.
- As you type, the table will dynamically filter the results to match your search criteria.

ID	Virus Name	Strain	Type	Family	Modification	Alone	Combination	Immune Gene	Source	Origin	Cell Line	Concentration	Viability
1	1	Adenovirus	oAd/DCN/LRP	DNA	Adenoviridae	Insertional mutant for DCN and LRP gene	Yes	No	No	ATCC	Human pancreatic cancer cell line	PANC-1	Grown to 70% confluence
2	2	Adenovirus	oAd/DCN/LRP	DNA	Adenoviridae	Insertional mutant for DCN and LRP gene	Yes	No	No	ATCC	Human pancreatic cancer cell line	PANC-1	Grown to 70% confluence
3	3	Adenovirus	oAd/DCN/LRP	DNA	Adenoviridae	Insertional mutant for DCN and LRP gene	Yes	No	No	ATCC	Human pancreatic cancer cell line	PANC-1	Grown to 70% confluence
4	4	Adenovirus	oAd/DCN	DNA	Adenoviridae	Insertional	Yes	No	No	ATCC	Human	PANC-1	Grown to 70% NA

**Fig. 1.** Overview of therapeutically important oncolytic viruses and cancer cell lines

In the same tab, you can see protein protein interactions between *Homo sapiens* and oncolytic viruses. To do this, follow the guidance below.

### ① Selecting the Oncolytic Virus Type for Protein-Protein Interactions (Fig. 2)

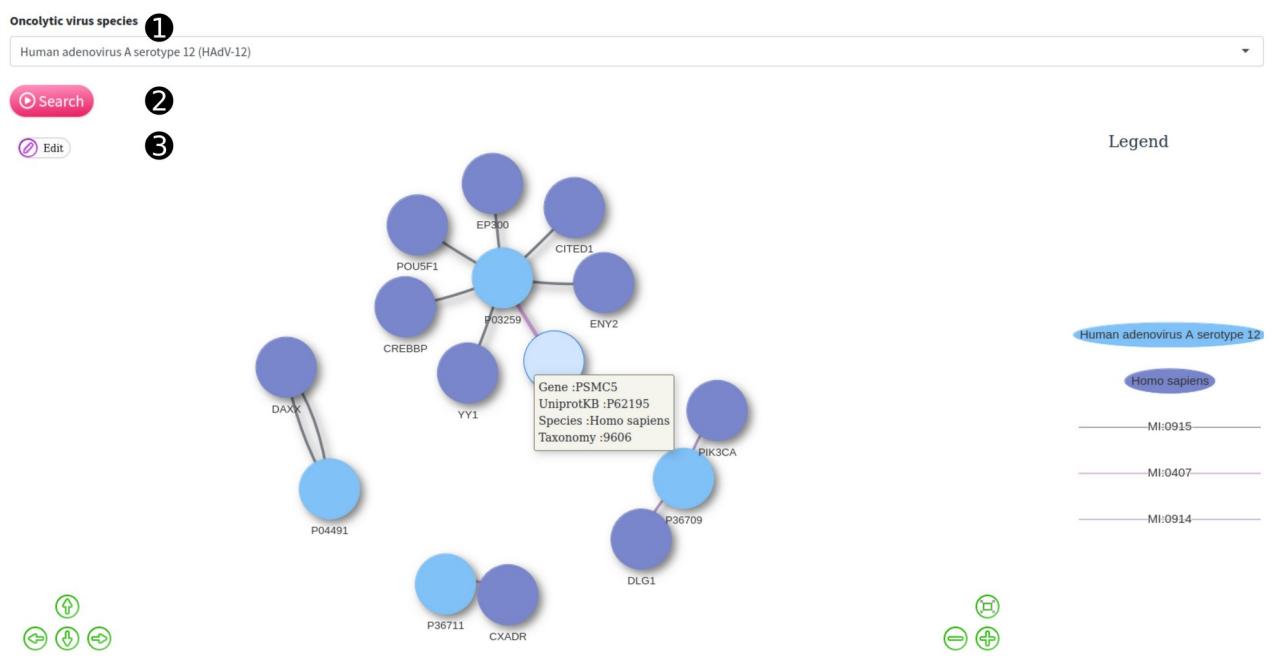
- To begin exploring protein-protein interactions, follow these steps:
- Locate the select input section numbered ①
- Within this section, you can choose the specific oncolytic virus type you are interested in by selecting it from the dropdown menu.

### ② Initiating the Search (Fig. 2)

- Once you have selected your desired oncolytic virus type, you can proceed to create a protein-protein network:
- Look for the "Search" button numbered ②
- Click on the "Search" button to initiate the process.

### ③ Customizing the Protein-Protein Network (Fig. 2)

- Now that you have generated a protein-protein network, you can customize it according to your preferences. Here's how to do it:
- Locate the "edit" button numbered ③ (Fig. 2)
- Click on the "edit" button to access the customization options.
- Within the customization interface, you can:
- Add Proteins: If there are specific proteins you'd like to include in the network, use the "add node" function to incorporate them.
- Delete Proteins: To remove proteins from the network, use the "delete selected" function.



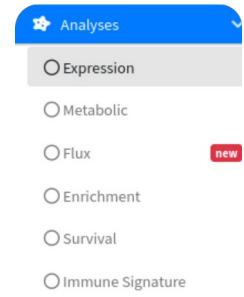
**Fig 2.** PPI network of *Homo sapiens*-Human adenovirus A serotype 12. Detailed information about proteins are shown in pop-up window

# Differential Expression Analysis

① You can perform expression analysis of the data in Cat-E by selecting the relevant cancer type. Alternatively, you can create a volcano plot using your own DEG data or compare it with the data in Cat-E for the chosen cancer type using the Venn diagram option. Please follow the appropriate steps based on your specific interest.

## ①a Selecting the Cancer Type Data (Fig. 3)

- To begin your analysis, follow these steps:
- Locate the "Select Input" section numbered ①a
- In this section, choose the cancer type data you are interested in by selecting it from the available options.



## ② Setting the LogFC Threshold (Fig. 3)

- To refine your analysis, set a LogFC (Log-Fold Change) value threshold:
- Find the LogFC threshold input labeled ②
- Enter the desired LogFC value that will be used as the threshold for your analysis.

## ③ Setting the P-Value Limit (Fig. 3)

- Now, let's set the limit for the p-value, which is used to determine the statistical significance of differential expression:
- Locate the p-value limit input marked ③
- Enter the p-value value you are interested in to define your significance threshold.

## ④ Selecting the Analysis Method (Fig. 3)

- You have the option to choose the method you prefer for performing the analysis:
- In the "Select Method" section numbered ④ choose the analysis method that best suits your research needs.

## ⑤ Running the Analysis (Fig. 3)

- Once you have configured your analysis parameters, you can initiate the analysis:
- Look for the "Run" button labeled ⑤
- Click the "Run" button to start the analysis process with the selected parameters.

## ⑥ Downloading the Analysis Results (Fig. 3)

- After completing the analysis, you can download the results for further exploration:
- Find the download options section labeled ⑥

Here, you can download the analysis results table in PDF, CSV, or Excel format, depending on your preferences.

Ensembl	logFC	logCPM	LR	Pvalue	FDR
1151 ENSG0000011016	-2.6227314051277	9.2212016457698	888.014999297001	3.958779234817e-195	1.0871624405389e-190
3302 ENSG00000168309	-4.39476382206032	5.1849351021794	878.270158538623	5.19626358035960e-193	7.1372784080305e-189
1835 ENSG00000135604	-2.93209147766058	4.14865985671478	874.935505672372	2.7582016306012e-192	2.52568523314603e-188
4012 ENSG00000182010	-3.93773275997591	5.52712545518767	871.495505271962	1.5433559931246e-191	1.05993828895681e-187
684 ENSG00000102760	-2.7154236262123	6.19349008581472	834.88848886693	1.40241098363645e-183	7.705126437936e-180
3110 ENSG00000165312	-2.01178288198767	4.6219738006295	828.941264119108	2.75326125744696e-182	1.26058066672209e-178
76 ENSG00000010319	-3.12697712266931	3.97462838465264	819.445933592535	3.13921339463146e-180	1.25315378805601e-176
1330 ENSG00000122679	-3.116962813948374	4.58215930963767	817.835608947104	7.15042084073762e-180	2.45536513644879e-176

Fig. 3. Differential expression analysis with LUAD data

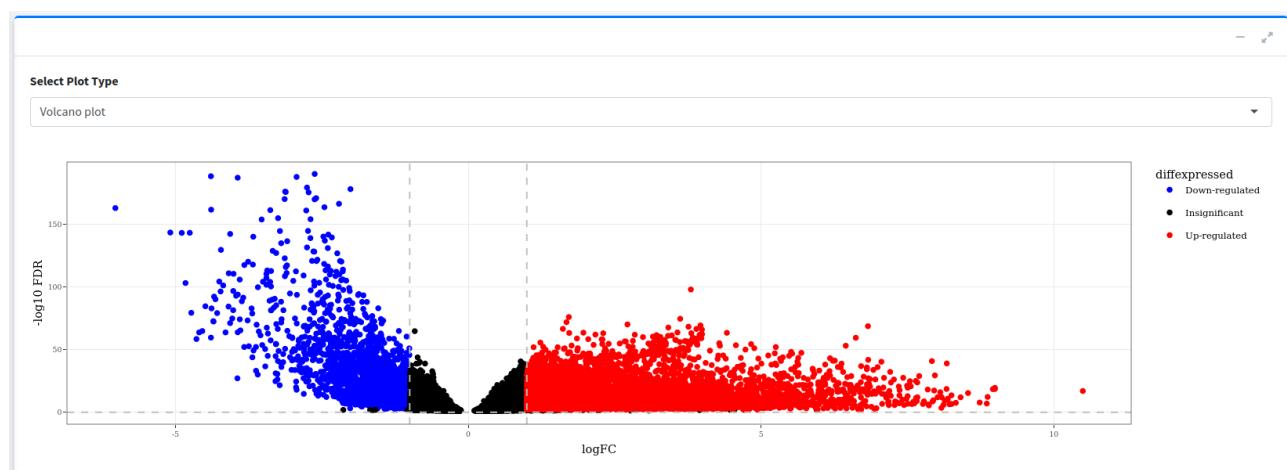
## 7 Searching Within the Analysis Tables (Fig. 3)

To facilitate data exploration, you can easily search within the analysis tables:

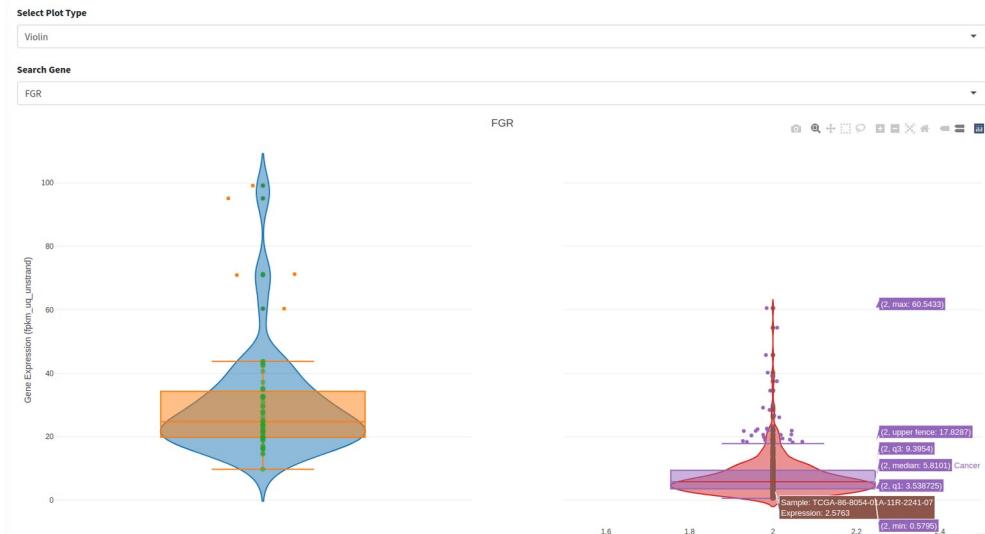
- Locate the search section marked 7
- Enter your search queries to quickly locate specific information within the tables.

Locate the "Select Plot Type" section (Fig. 4)

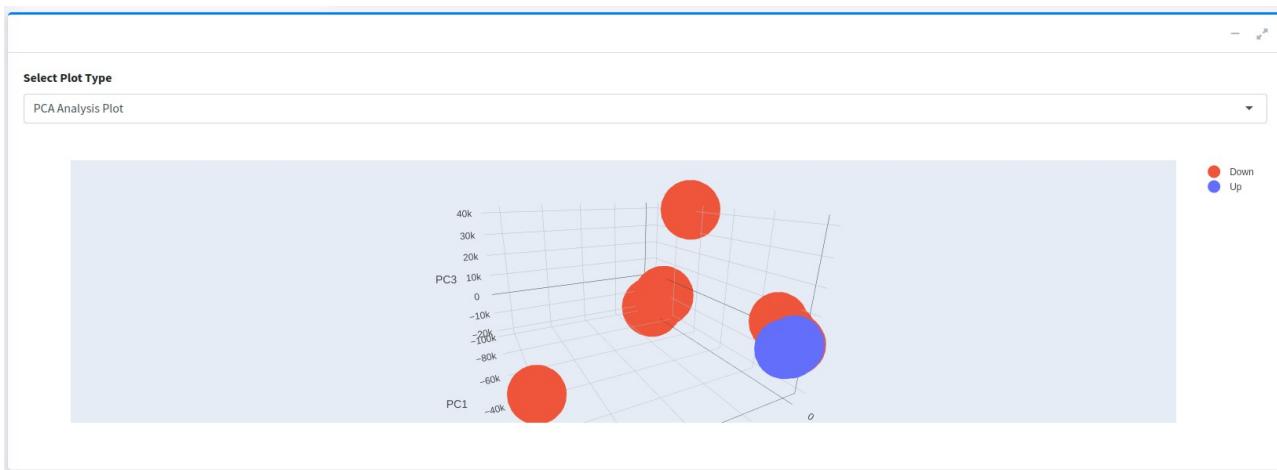
- Here, you will find three different plot options: Volcano Plot, Violin Plot, PCA (Principal Component Analysis) Plot
- Choose one of the available plot types based on your analysis needs. Each type offers unique insights into your data:
- Volcano Plot: Use this plot to visualize the relationship between statistical significance and fold change. It's great for identifying significantly different features.
- Violin Plot: This plot is ideal for visualizing the distribution of data and comparing multiple variables at once. It helps you understand data spread and density.
- PCA Plot: Principal Component Analysis (PCA) is useful for reducing dimensionality and visualizing the variance in your data. It helps you identify patterns and clusters within your data.
- After selecting your desired plot type, you can access the corresponding graphics:
- Refer to Fig.4 and Fig.5 to view the results of the chosen plot type.
- For example, if you selected the Volcano Plot, refer to Fig.4 for the Volcano Plot graphic.
- Additionally, you can access Fig.6 to view other related graphics or visualizations associated with your analysis.



**Fig. 4.** Volcano plot of LUAD cancer DEGs data



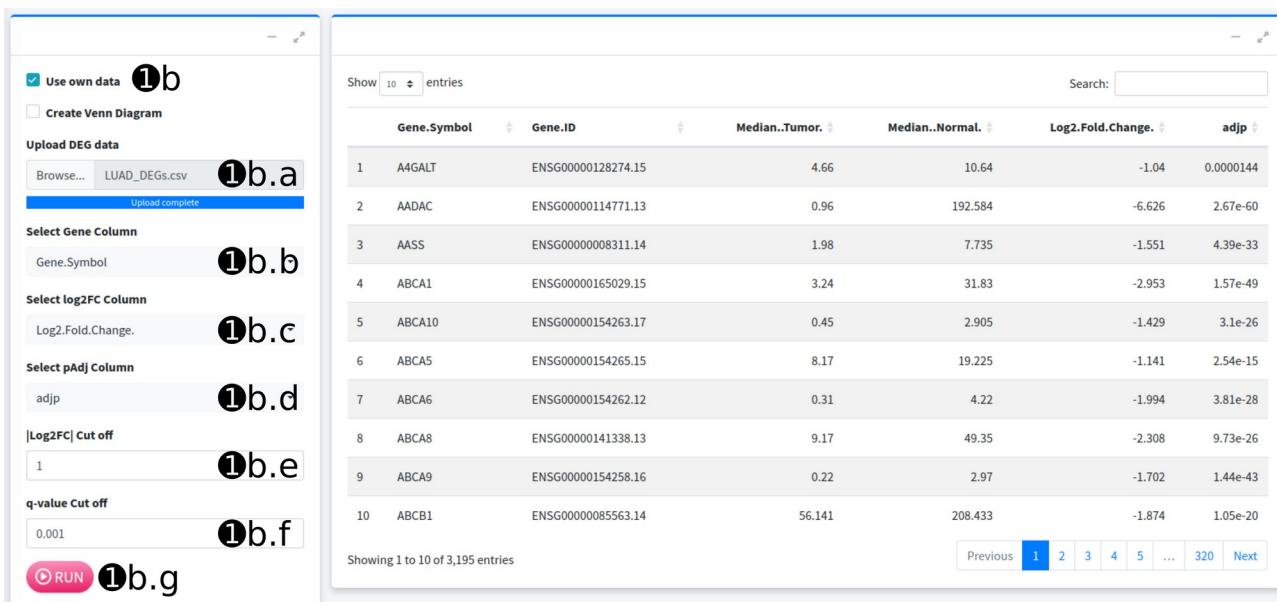
**Fig. 5.** Violin plot of FGR gene in LUAD normal and cancer samples



**Fig. 6.** PCA 3D plot of LUAD cancer DEGs data

Follow the steps below to create a volcano plot by loading **user data ①b** (Fig.7) :

- ①b.a: Upload your DEG data in CSV format.
- ①b.b: Select the "gene name" column in your data.
- ①b.c: Select the "logFC" column in your data.
- ①b.d: Select the "p-value" column in your data.
- ①b.e: Set a threshold for LogFC.
- ①b.f: Set a threshold for p-value.
- ①b.g: Press the "Run" button to create the volcano plot.

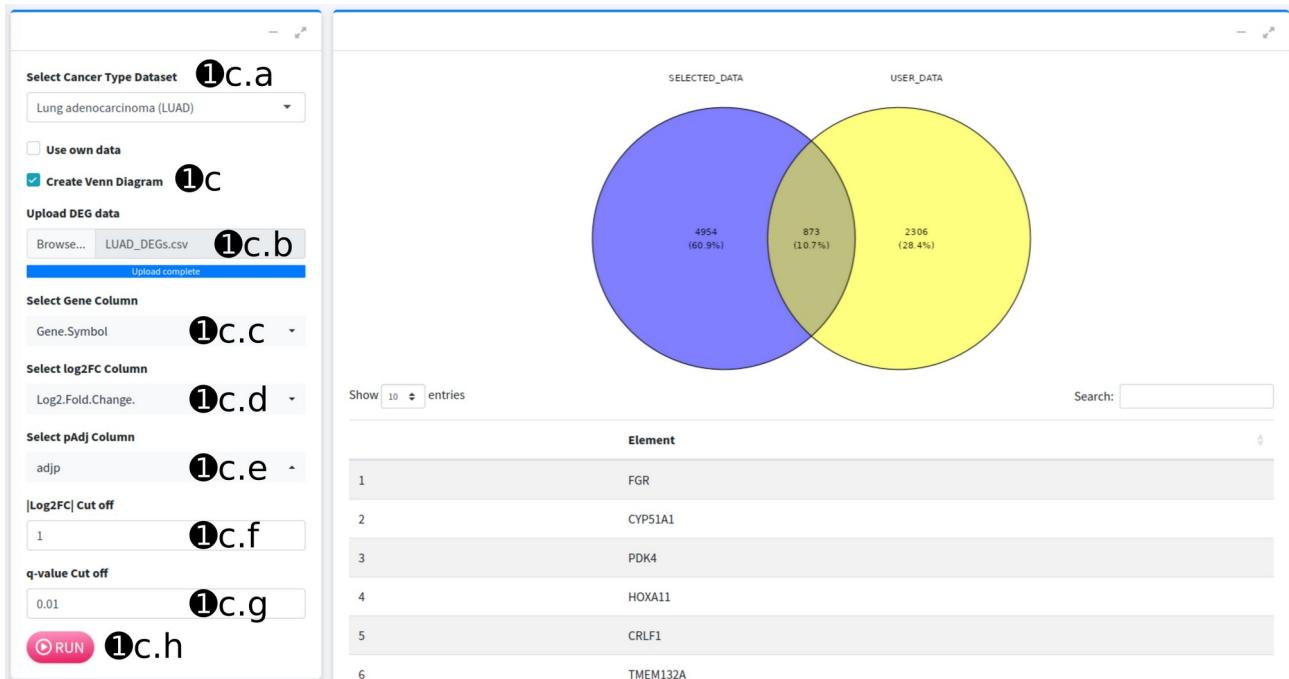


**Fig. 7.** Upload own DEG data

**①c** Select the "Create Venn Diagram" option (Fig.8):

- ①c.a: Choose the type of cancer with which you want to compare your data.
- ①c.b: Upload your DEG data in CSV format.
- ①c.c: Select your “gene” column.
- ①c.d: Select the “logFC” column.
- ①c.e: Specify the “p-value or padjust” column.
- ①c.f: Set a threshold for LogFC.
- ①c.g: Set a threshold for p-value.
- ①c.h: Create the Venn diagram by clicking the “Run” button.

The Venn diagram is interactive, allowing you to click on regions and view the unique and intersecting genes.



**Fig. 8.** Comparison of selected data and user data with Venn diagram

# Metabolic Pathway Analysis

## ① Accessing the Metabolic Tab (Fig. 9)

- On the left sidebar, find and click on the "Metabolic" tab numbered ①



## ② Selecting the Cancer Type (Fig. 9)

- Once you're in the Metabolic section, you will need to specify the type of cancer you are interested in. This helps tailor the analysis to your research needs.
- Locate the "Select Cancer Type" option numbered ②
- Choose the specific type of cancer you want to focus on from the available options.

## ③ Loading Your Data (Optional) (Fig. 9)

- If you have your own data that you'd like to incorporate into the analysis, follow these steps:
- Continue to the "Use Own Data" option numbered ③

## ④ Selecting the Network (Fig. 9)

- In the "Select Network" input section numbered ④ choose the network that aligns with your research interests. This network will serve as the basis for your analysis.

## ⑤ Setting Topology Parameters (Fig. 9)

- To refine your analysis, select the appropriate topology options by locating the "Select Topology" input section numbered ⑤ These settings help determine the relationships between metabolic pathways.

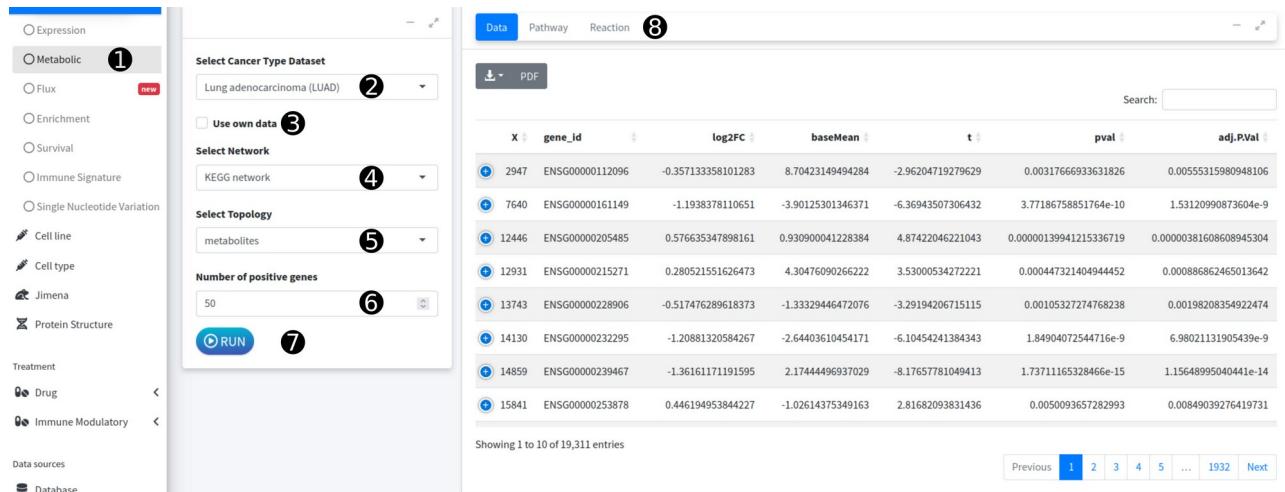
## ⑥ Determining the Number of Positive Genes (Fig. 9)

- Specify the number of positive genes relevant to your analysis. You can do this by adjusting the "Number of Positive Genes" setting numbered ⑥

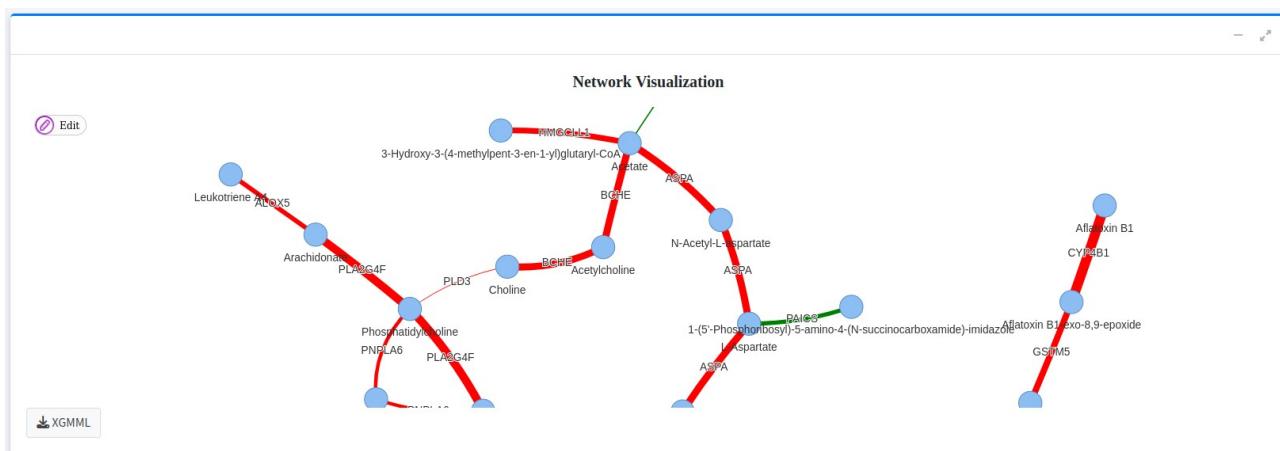
## ⑦ Initiating the Analysis (Fig. 9)

- Once you have configured your analysis parameters, you are ready to start the analysis. Locate the "Run" button numbered ⑦
- Click the "Run" button to initiate the metabolic pathway analysis with the selected settings.

On the page, you will find a metabolic network visualization (Fig. 10). This graphical representation provides a visual overview of the metabolic pathways associated with the selected cancer type.

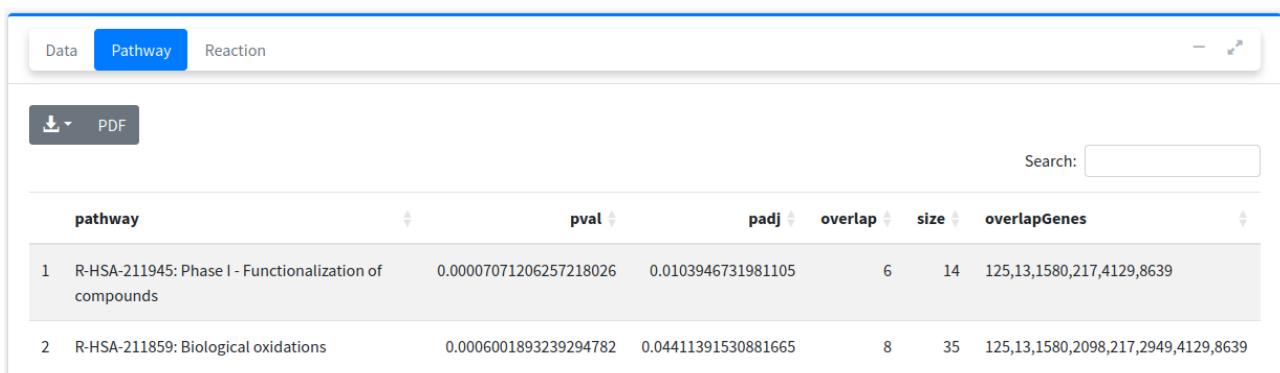


**Fig. 9.** Metabolic pathway analysis using LUAD DEG data



**Fig. 10.** Metabolic pathway network of LUAD data

Below the metabolic network visualization, you will find a section containing detailed pathway information (Fig. 11). This includes a list of identified pathways relevant to your analysis. Each pathway will be labeled and described, allowing you to gain insights into their significance.



**Fig. 11.** Metabolic pathway data of LUAD data

Additionally, you will have access to tables that list the reactions associated with the identified pathways. These tables provide a comprehensive overview of the biochemical reactions involved in the metabolic pathways (Fig. 12).

from	to	label	pval	origin	Symbol	gene	enzyme	reaction_name	reaction_equation
2	40	MGAT3	5.27008024093613e-99	6184	MGAT3	4248	2.4.1.144	UDP-N-acetyl-D-glucosamine:beta-D-mannosyl-glycoprotein 4-beta-N-acetyl-D-glucosaminyltransferase	UDP-N-acetyl-alpha-D-glucosamine + beta-D-Mannosyl-R => UDP + 4-(N-Acetyl-beta-D-glucosaminyl)-beta-D-mannosyl-R
35	40	MGAT3	5.27008024093613e-99	6184	MGAT3	4248	2.4.1.144	UDP-N-acetyl-D-glucosamine:beta-D-mannosyl-glycoprotein 4-beta-N-acetyl-D-glucosaminyltransferase	UDP-N-acetyl-alpha-D-glucosamine + beta-D-Mannosyl-R => UDP + 4-(N-Acetyl-beta-D-glucosaminyl)-beta-D-mannosyl-R
4	38	GSTM5	1.39460117596508e-74	4323	GSTM5	2949	2.5.1.18	glutathione-S-transferase	Aflatoxin B1-exo-8,9-epoxide + Glutathione <=> Aflatoxin

Showing 1 to 10 of 43 entries

Previous 1 2 3 4 5 Next

**Fig. 12.** Metabolic pathway reaction data of LUAD data

If you want to perform metabolic pathway analysis by uploading your own data, follow the steps mentioned above after uploading your data (Fig. 13).

- ① Choose "Use Your Own Data." (Fig. 13)
- ② Select the organism to which your data belongs.
- ③ Upload your Differential Expression Gene (DEG) data or ④ Metabolic data in CSV format.
- ⑤ Choose the network database you want to use.
- ⑥ Define the topology.
- ⑦ Specify the gene size.
- ⑧ Initiate the analysis by clicking the "Run" button.

**Select Cancer Type Dataset**

Use own data ①

**Select Organism**  
 ②

**DE Genomic Data**  
  ③  
Upload complete

**DE Metabolic Data**  
  ④

**Select Network**  
 ⑤

**Select Topology**  
 ⑥

**Number of positive genes**  
 ⑦

⑧

**Data** Pathway Reaction

Search:

V1	gene_id	log2FC	baseMean	t	pval	adj.PVal
+ 12141	ENSG00000205485	0.781634315292702	0.689068687968623	6.38187359862878	3.84364556762045e-10	1.3168466175457e-09
+ 12859	ENSG00000226380	2.46790924417285	0.0190737728630822	9.34657400217279	2.57414026220073e-19	1.92378512394499e-18
+ 13792	ENSG00000239467	1.83967766447932	0.772563495213589	7.64825596416162	9.77253814677348e-14	4.58357246440653e-13
+ 14589	ENSG00000255823	-1.19942056294	-1.11170424260587	-4.78115848310571	0.00000226590556609807	5.30084574856883e-06
+ 14861	ENSG00000259820	0.459620922668518	2.02212897088885	2.49755705791981	0.0128098106985266	0.0185660579389082
+ 14863	ENSG00000259834	-2.8888747521795	-0.529852180307348	-12.980963369511	1.27866447202597e-33	2.56222001033839e-32
+ 15049	ENSG00000261490	-0.51317115917093	0.637895648123273	-3.15132518008131	0.00171792387858234	0.0028148466768371
+ 15670	ENSG00000273888	-0.584856304715007	-2.98384920282837	-2.645137993839	0.00841005036323998	0.0125321996675927

Showing 1 to 10 of 16,612 entries

Previous 1 2 3 4 5 ... 1662 Next

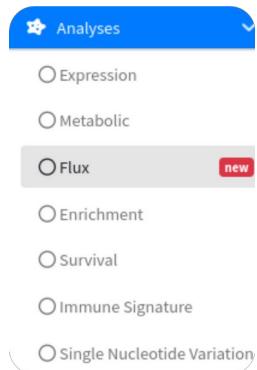
**Fig. 13.** Metabolic pathway analysis using own data

13

# Metabolic Flux Analysis

## ① Selecting Analysis Type (Fig.16)

- Begin by selecting the appropriate analysis type. Locate and choose the desired analysis type labeled ①



Do you have the XML/SBML file? If you already have an SBML file, follow ②a. If not, Cat-E will allow you to create your own SBML file, so follow ②b.

## ②a Uploading Your XML/SBML File (Fig.16)

- To perform metabolic flux analysis, you will need to provide an XML file that describes the metabolic interactions. Follow these steps:
  - Click on the "Upload" button numbered ③
  - Select and upload your XML/SBML file, which should contain information about metabolic reactions and their interactions.

(you can find the sample sbml file in the link:

[https://github.com/salihoglu/Cat-E/blob/1760c8cdf5ded4aecdc0bea4c87b893ece88c541/example\\_data/model.xml](https://github.com/salihoglu/Cat-E/blob/1760c8cdf5ded4aecdc0bea4c87b893ece88c541/example_data/model.xml))

- The 'XML' file you created in CellDesigner should contain the elements 'nodes' and 'reactions,' as shown Fig.14, for example.

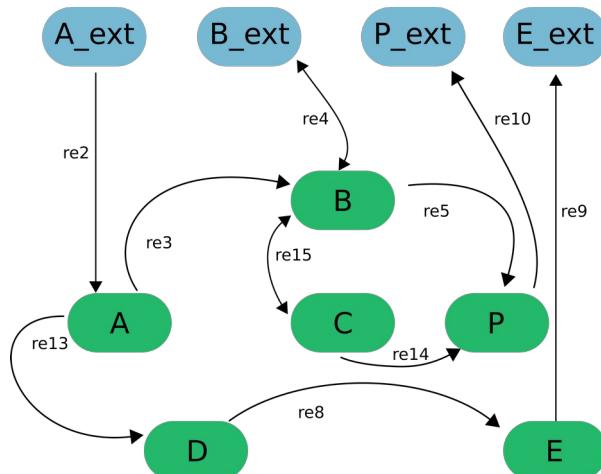


Fig. 14. Sample reaction chart created in CellDesigner

## ②b If you do not have any sbml file, you can create one by following the commands below. (Fig.15)

If you don't possess an SBML file, you can generate a CSV file instead. This CSV file should include columns for "type," "name," and "equation." In the "equation" column, please make sure that you use the same names as you specified in the "name" column. State the directions of the reactions. You can access the sample file and image from the link below. You can obtain a sample CSV file by following this link:

<https://github.com/salihoglu/Cat-E>

③ Upload your CSV file (Fig.15)

④ Click on the "Create SBML" button (Fig.15)

⑤ You can see the created pathway on the right. Allows visual zoom-in/zoom-out (Fig.15)

⑥ You can save your image in SVG format by clicking the "Download SVG" button (Fig.15)

The figure consists of two side-by-side screenshots of the Cat-E software interface.

**Left Panel (Create SBML file):**

- Choose:**  Upload own SBML/XML,  Create SBML (2b)
- Upload CSV file:** Browse... metabolites\_example.csv (3)
- Table:**

type	name	equation
metabolite	D	
metabolite	E	
metabolite	P	
reaction	R1	$\rightarrow A$
reaction	R2	$\leftrightarrow B$
reaction	R3	$P \rightarrow$
reaction	R4	$E \rightarrow$
reaction	R5	$A \rightarrow B$
- Buttons:** Create SBML (4)

**Right Panel (Created Model):**

Diagram illustrating a metabolic network with nodes A, B, C, D, E, P and arrows indicating fluxes:

```

graph TD
    A((A)) --> B((B))
    B --> C((C))
    C --> D((D))
    D --> E((E))
    E --> P((P))
    P --> A
    A --> C
    C --> B
    B --> D
    D --> E
    E --> P
    P --> A
    A <--> B
    C <--> D
    D <--> E
    
```

**Buttons:** Download SVG (6)

**Fig. 15.** Create SBML file in Cat-E

#### ④ Adjusting Reaction Bounds (Optional) (Fig.16)

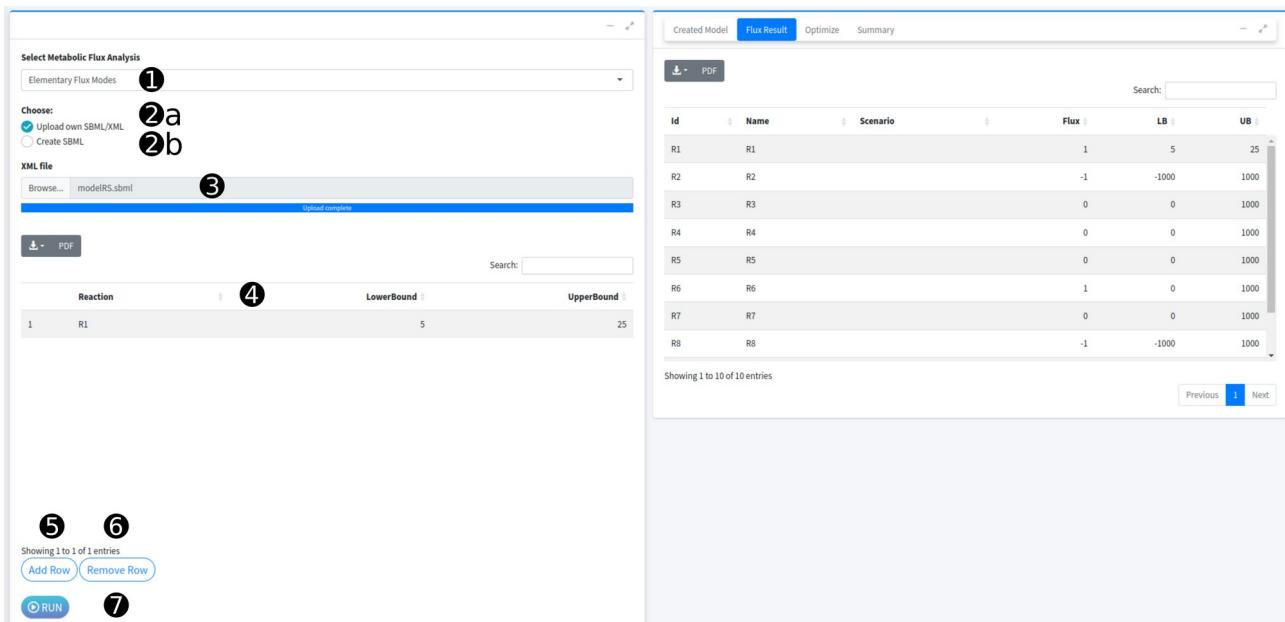
- If you wish to customize the lower and upper bound values for specific reactions in your analysis, you can do so by following these steps:
- Click the "Add Row" button labeled ⑤ to create a new row for the reaction you want to modify. This is illustrated in ④

#### ⑥ Removing Reaction Bound Values (Optional) (Fig.16)

- If you decide to cancel the values you have entered for a specific reaction, follow these steps:
- Click on the row you want to delete.
- Then click the "Remove Row" button to remove the values for that specific reaction.

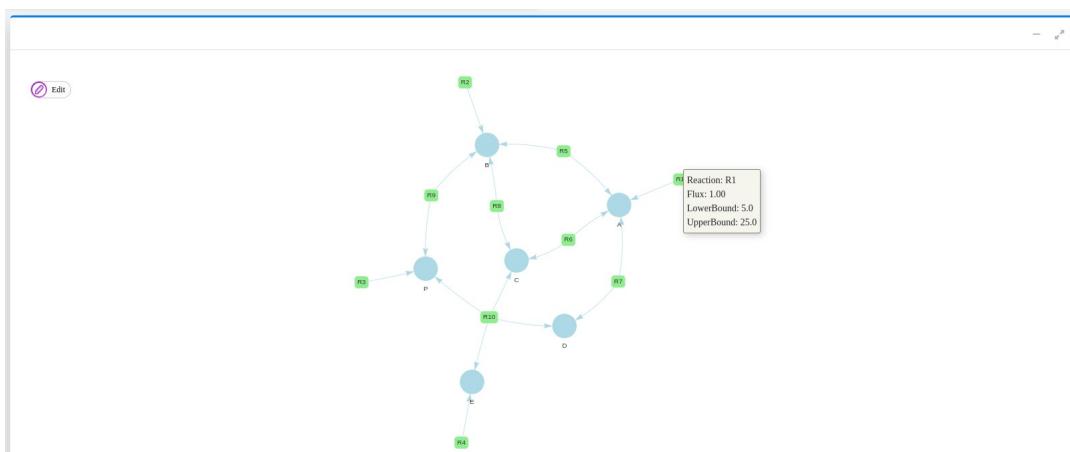
#### ⑦ Initiating the Analysis (Fig.16)

- Once you have configured your analysis parameters and selected the appropriate reactions, you are ready to start the metabolic flux analysis.
- Locate the "Run" button numbered ⑦
- Click the "Run" button to begin the metabolic flux analysis with the specified settings.



**Fig. 16.** Metabolic flux analysis with uploaded example data

After running your metabolic flux analysis, you will be presented with various result components. As a result of your metabolic flux analysis, you will receive a network visualization where flux values are presented in a graphical format. This visualization allows you to visually explore the metabolic pathways and the flow of metabolites within your system (Fig. 17).



**Fig. 17.** Elementary flux modes analysis results

To access detailed information about flux values for individual reactions and pathways:

Look for the "Flux Result" section or file (Fig. 16). This will provide you with a comprehensive dataset containing flux values for each reactions.

Explore the "Optimize" table section or files. These tables typically contain information on the optimized flux distributions, which are crucial for understanding the metabolic activity in your system (Fig. 18).

The figure shows an 'Optimize' table with rows for reactions R1 through R8. The flux values are all 0.0.

Reaction	Flux
R1	0
R2	0
R3	0
R4	0
R5	0
R6	0
R7	0
R8	0

**Fig. 18.** Optimize table of analysis

For a concise overview of your metabolic flux analysis results:

- Review the "Summary" section or files. These tables provide a summary of key findings and important metrics from your analysis (Fig. 19).
- By examining these various components, you can gain a comprehensive understanding of the metabolic flux patterns, optimized values, and summary statistics resulting from your analysis.

**Fig. 19.** Summary table of metabolic flux analysis

Created Model	Flux Result	Optimize	Summary	—	v <sup>2</sup>
<input type="button" value="PDF"/>				Search: <input type="text"/>	
<b>Content</b>					
Objective					
=====					
nan Expression = nan					
Uptake					
=====					
Metabolite Reaction Flux C-Number C-Flux					
A R1 5 0 0.00%					
Showing 1 to 10 of 13 entries				Previous	1 2 Next

# GO Enrichment Analysis

## ① Selecting the Cancer Type (Fig. 20)

- To initiate GO enrichment analysis, begin by selecting the cancer type you are interested in.
- Locate the "Select Cancer Type" option numbered ①
- Choose the specific cancer type you want to focus on from the available options.



## ② Setting LogFC Threshold (Fig. 20)

- Next, set the LogFC (Log-Fold Change) threshold to refine your analysis:
- Find the LogFC threshold input labeled ②
- Enter the desired LogFC value that will be used as the threshold for your analysis.

## ③ Specifying P-value Range (Fig. 20)

- Define the range for the p-value, which is used to assess the statistical significance of GO enrichments:
- Locate the p-value range input marked ③
- Specify the range of p-values that align with your research interests.

## ④ Initiating the Analysis (Fig. 20)

- Once you have configured your analysis parameters, you are ready to start the GO enrichment analysis:
- Look for the "Run" button labeled ④
- Click the "Run" button to initiate the analysis process with the selected parameters.

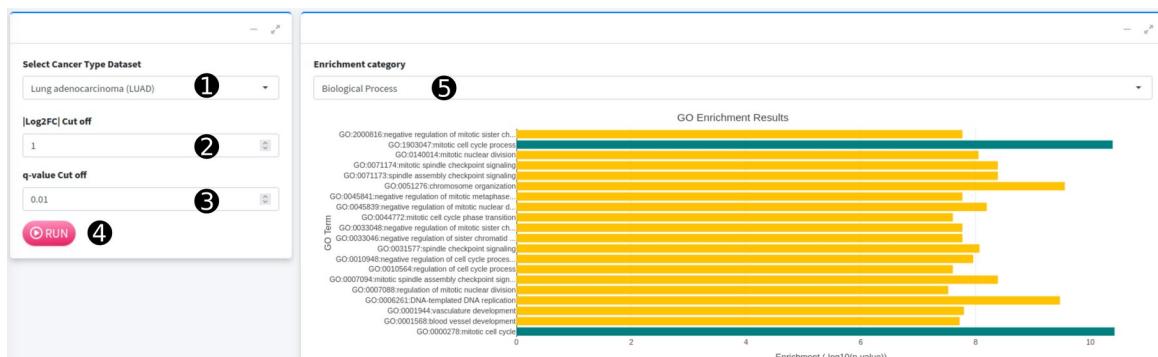


Fig. 20. GO Enrichment analysis biological process result of LUAD DEG data

## ⑤ Exploring GO Enrichment Categories (Fig. 20)

After running the analysis, you can explore GO enrichment categories to gain insights into gene functions.

Locate the "Select GO Enrichment Category" section numbered ⑤

You can choose from categories such as Biological Process, Molecular Function, and Cellular Component, depending on your research focus.

To visualize the GO enrichment results, examine the chart containing the top 20 pathways (Fig. 21)

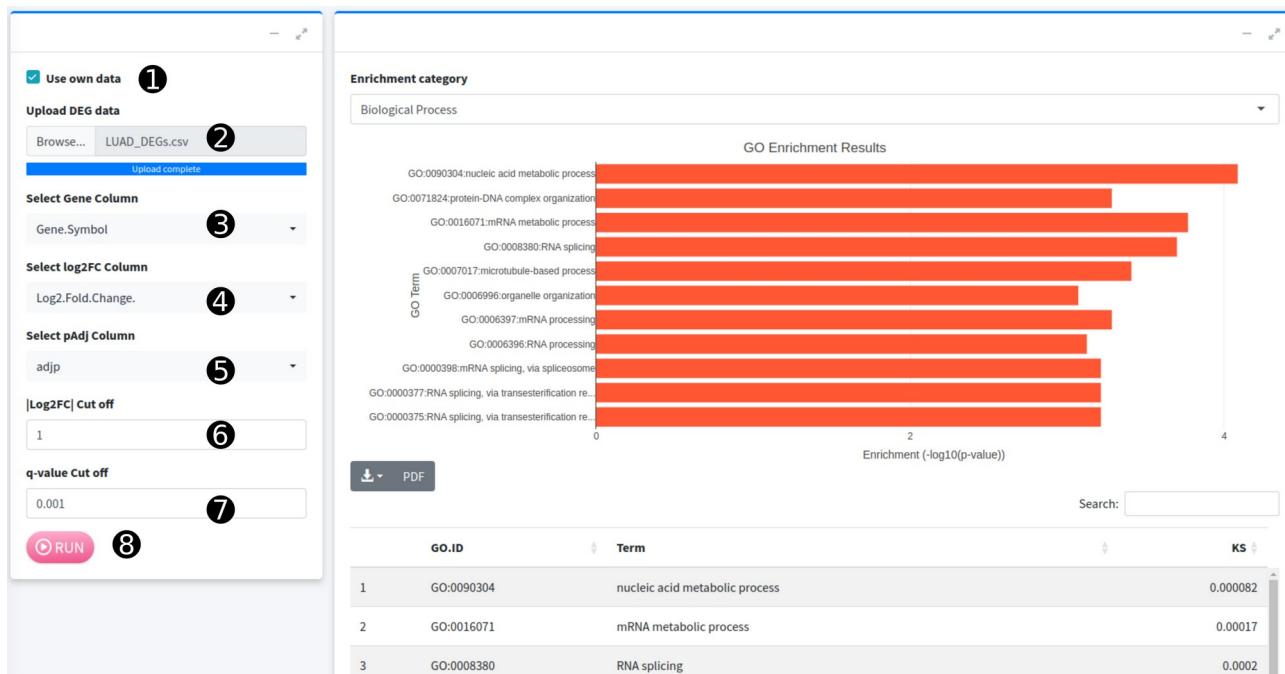
GO.ID	Term	KS
1	GO:0000278	mitotic cell cycle
2	GO:1903047	mitotic cell cycle process
3	GO:0051276	chromosome organization
4	GO:0000261	DNA-templated DNA replication
5	GO:0007094	mitotic spindle assembly checkpoint sign...
6	GO:0071173	spindle assembly checkpoint signaling
7	GO:0071174	mitotic spindle checkpoint signaling
8	GO:0045839	negative regulation of mitotic nuclear d...

Fig. 21. Top 20 biological process pathway of LUAD DEG data

Cat-E allows users to perform GO enrichment analysis on own DEG data. Follow the steps below ① (Fig.22):

- ② Upload your DEG data in CSV format.
- ③ Select the "gene name" column in your data.
- ④ Choose the "logFC" column in your data.
- ⑤ Select the "p-value" column in your data.
- ⑥ Set a threshold for LogFC.
- ⑦ Set a threshold for p-value.

⑧ Click the "Run" button to generate the top 20 significantly expressed terms in a bar plot and table format in the GO enrichment analysis results.



**Fig. 22.** Top 20 biological process pathway of own LUAD DEG data

# KEGG Enrichment Analysis

GO KEGG

Parameters

Choose:  
 Cancer Data (1)  
 PPI Data

Select Cancer Type Dataset:  
Lung adenocarcinoma (LUAD) (2)

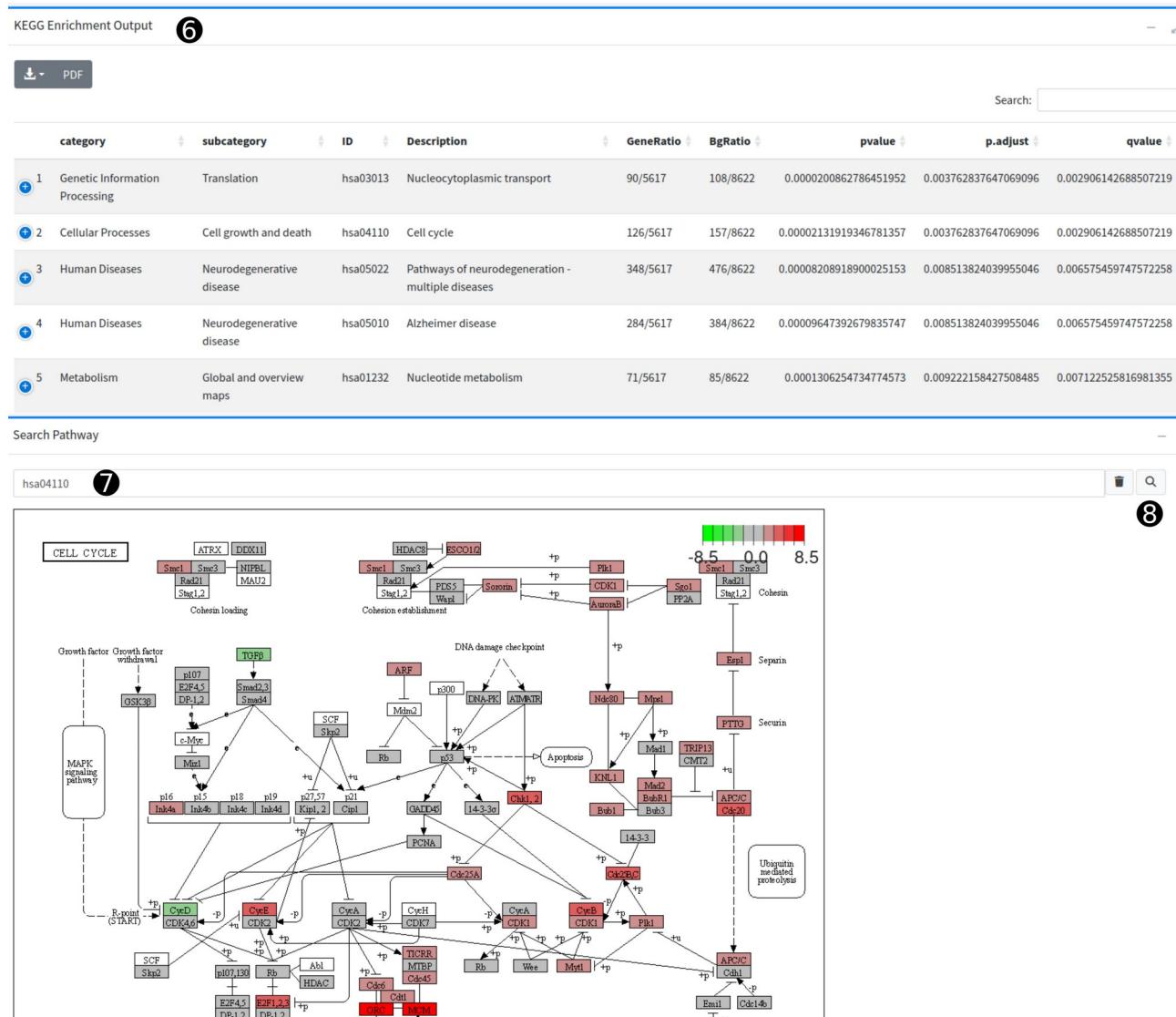
qValue:  
0.01 (3)

pValue:  
0.01 (4)

**Run KEGG** (5)

- ① Select the "Cancer Data" option to perform KEGG analysis on the cancer data of your choice (Fig. 23).
- ② Choose the specific cancer data you are interested in (Fig. 23).
- ③ Set a "q-value" threshold (Fig. 23).
- ④ Set a "p-value" threshold (Fig. 23).
- ⑤ Click the "Run KEGG" button (Fig. 23).
- ⑥ The output obtained from the enrichment analysis is displayed in a table (Fig. 24).
- ⑦ To find the pathway you're looking for, select the pathway ID from the "ID" column in the table, and paste it into the "Search Pathway" section (Fig. 24).
- ⑧ Click the search button, and you'll see a color-coded pathway generated using the logFC values of the selected cancer data (Fig. 24).

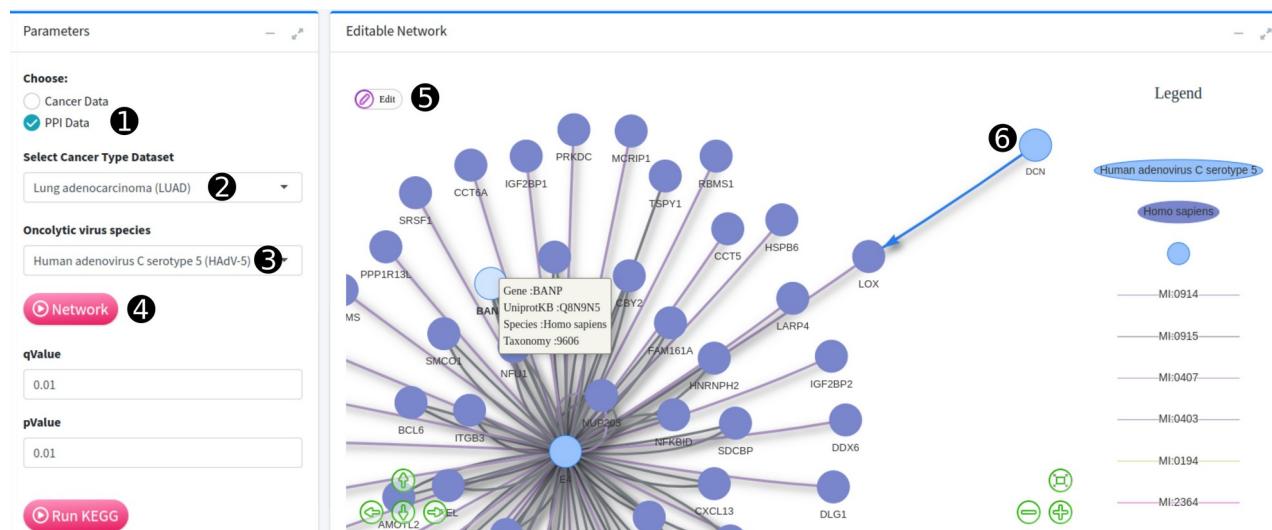
**Fig. 23.** KEGG analysis with cancer data



**Fig. 24.** Table and colored pathway of the KEGG analysis result using LUAD data

Cat-E allows you to modify virus-human PPI data and perform KEGG analysis with human proteins found in this network. Follow the steps below to conduct a pathway analysis using relevant cancer data for the oncolytic virus model you are designing.

- ① Select the "PPI Data" option (Fig. 25).
- ② Choose the cancer data you are interested in to utilize logFC values (Fig. 25).
- ③ For Oncolytic virus, select the virus you want to modify (Fig. 25).
- ④ Visualize PPIs by clicking the "Network" button (Fig. 25).
- ⑤ The PPI network is editable. You can add a "node" by clicking the "Edit" button and connect this node to a protein of your choice using the "edge" button (Fig. 25).
- ⑥ The node you added will appear like this. The human proteins you include here will be used in the KEGG analysis (Fig. 25).
- ⑦ For KEGG analysis, follow the steps outlined in Figure 23 (qvalue, pvalue).

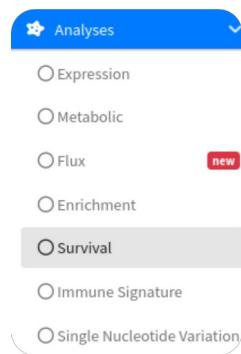


**Fig. 25.** KEGG analysis with Human proteins in the Network

# Clinical Survival Analysis

To initiate clinical survival analysis, begin by selecting the specific cancer type you are interested in (Fig. 26):

- Locate the "Select Cancer Type" option.
- Choose the cancer type you want to focus on from the available options.



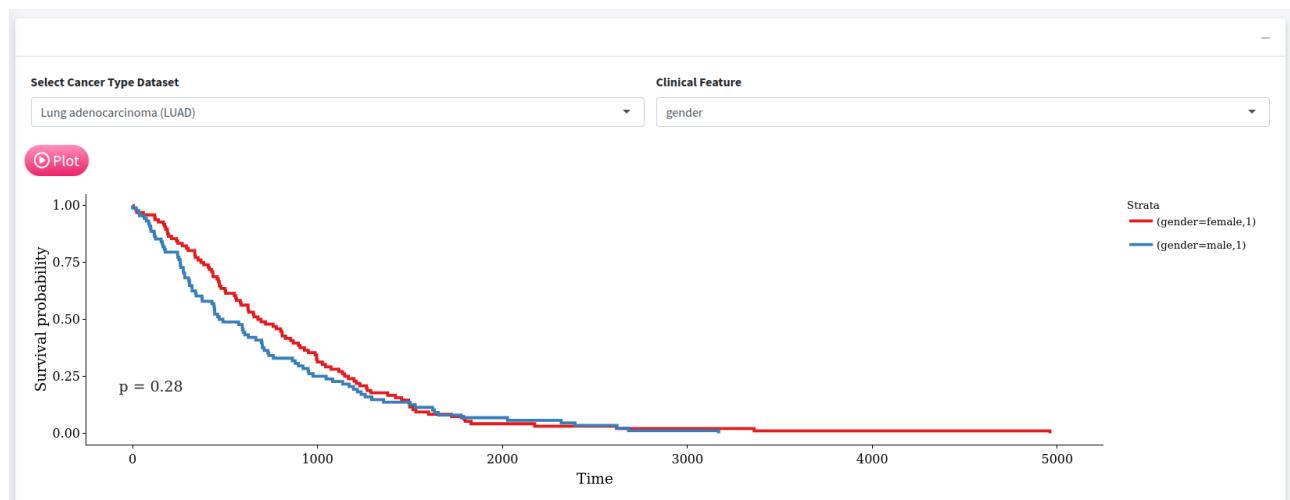
Once you have selected the cancer type, it's time to determine the clinical feature you want to analyze (Fig. 26):

- Identify the "Clinical Feature" option.
- Choose the clinical feature you want to investigate from the available list.

After selecting the cancer type and clinical feature, you are ready to run the clinical survival analysis (Fig. 26):

- Locate the "Plot" button.
- Click the "Plot" button to initiate the analysis with the specified cancer type and clinical feature.

Once the analysis is complete, you will be presented with the results of the clinical survival analysis, typically in the form of survival curves and clinical data (Fig. 26, 27). You can now view and interpret the survival analysis results to gain insights into the relationship between the clinical feature and patient survival outcomes.



**Fig. 26.** Survival Curves for Lung Adenocarcinoma (LUAD) Patients

project	submitter_id	NA.	synchronous_malignancy	ajcc_pathologic_stage	days_to_diagnosis	created_datetime	last_known_disease_status	tissue_or_organ_of_origin
1	TCGA-LUAD	TCGA-55-8513	No	Stage IIIB	0	not reported	Lower lobe, lung	
2	TCGA-LUAD	TCGA-83-5908	No	Stage IA	0	not reported	Upper lobe, lung	
3	TCGA-LUAD	TCGA-95-7948	Not Reported	Stage IB	0	not reported	Lower lobe, lung	
4	TCGA-LUAD	TCGA-44-5643	No	Stage IIIA	0	not reported	Lower lobe, lung	
5	TCGA-LUAD	TCGA-86-8279	No	Stage IIA	0	not reported	Upper lobe, lung	
6	TCGA-LUAD	TCGA-64-5815	No	Stage IIIB	0	not reported	Upper lobe, lung	
7	TCGA-LUAD	TCGA-73-A9RS	No	Stage IIIB	0	not reported	Upper lobe, lung	
8	TCGA-LUAD	TCGA-L9-A50W	No	Stage IIA	0	not reported	Lower lobe, lung	

Showing 1 to 10 of 585 entries

Previous 1 2 3 4 5 ... 59 Next

**Fig. 27.** Clinical Data for LUAD patients

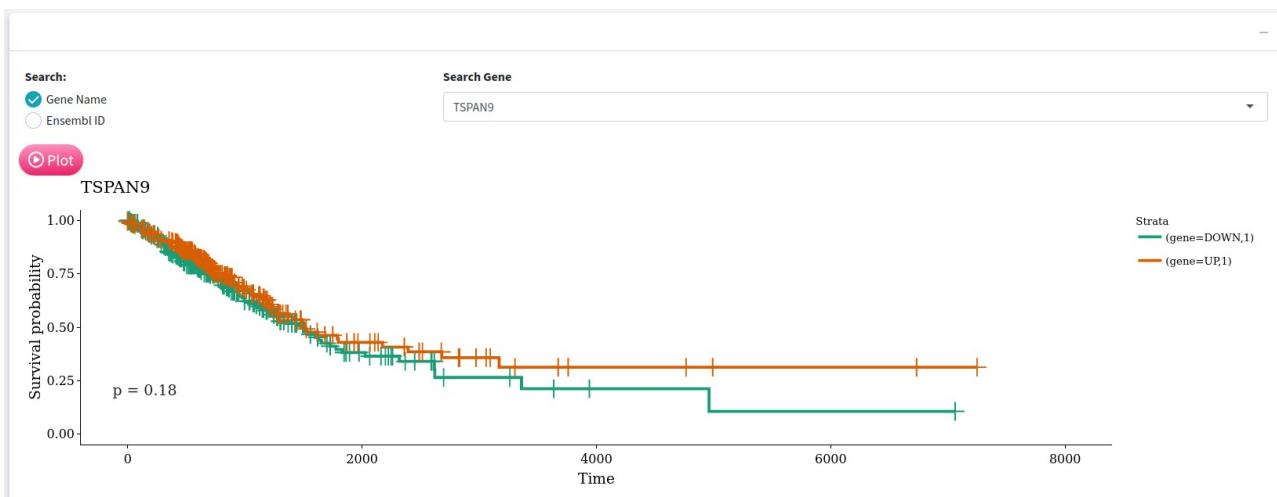
Once you have selected the cancer type, you will need to choose the gene you want to investigate (Fig. 28):

- Identify the "Select Gene" option.
- Choose the gene you want to analyze from the available list of genes.

After selecting the cancer type and gene of interest, you are ready to run the gene survival analysis (Fig. 28):

- Locate the "Plot" button. Click the "Plot" button to initiate the analysis with the specified cancer type and gene.

Once the analysis is complete, you will be presented with the results of the gene survival analysis, typically in the form of survival curves plot (Fig. 28). You can now view and interpret the survival curves to gain insights into how the up/down regulation of the selected gene is associated with patient survival within the context of the chosen cancer type.



**Fig. 28.** Gene Survival Curve for TSPAN9 in LUAD data

# Immune Signature Analysis

To initiate immune signature analysis, begin by selecting the specific cancer type you are interested in (Fig. 29):

- Locate the "Select Cancer Type" option numbered ①
- Choose the cancer type you want to focus on from the available options.

After selecting the cancer type, you can proceed to run the immune signature analysis (Fig. 29):

- Locate the "Run" button numbered ②
- Click the "Run" button to initiate the analysis with the specified cancer type.

As a result of the analysis, barplots will be generated to visualize the relationship between cancer samples and immune cells ③ (Fig. 29)



Additionally, a network of immune signature genes will be created ④ (Fig. 30)

- This network will help you explore the interactions and relationships among immune signature genes, providing valuable insights into their roles in the immune response within the context of the specified cancer type.

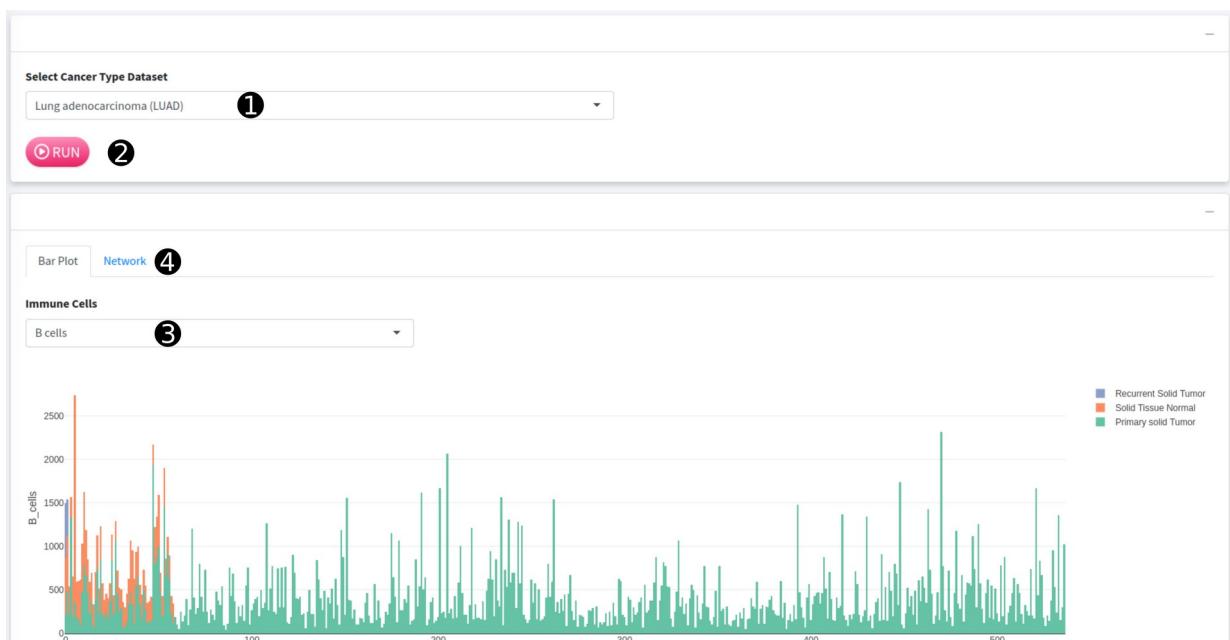


Fig. 29. B cells in LUAD patients samples

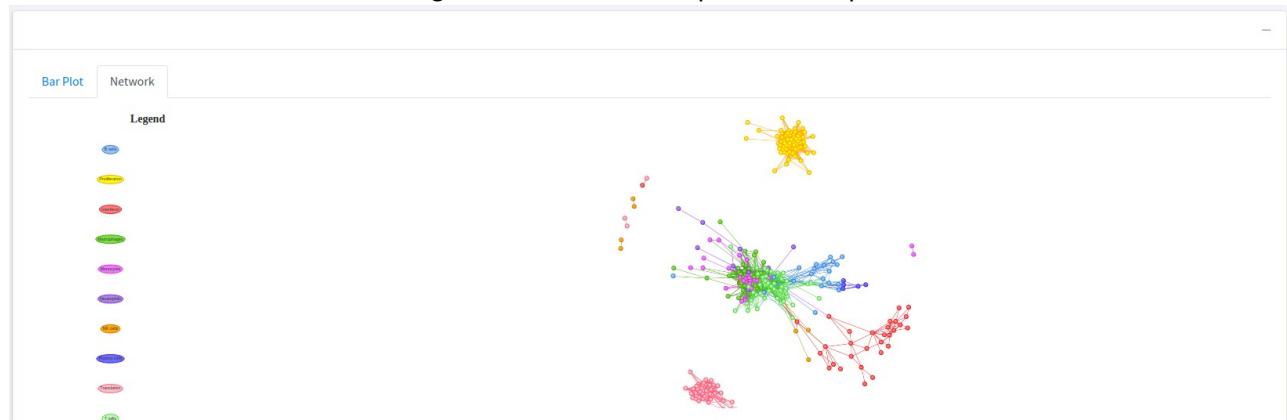


Fig. 30. Network of LUAD immune signature genes

# Single Nucleotide Variation Analysis

## ① Selecting the Cancer Type (Fig. 31)

- To initiate SNV analysis, begin by selecting the specific cancer type you are interested in:
- Locate the "Select Cancer Type" option numbered ①
- Choose the cancer type you want to focus on from the available options.



## ② Selecting the Plot Type (Fig. 31)

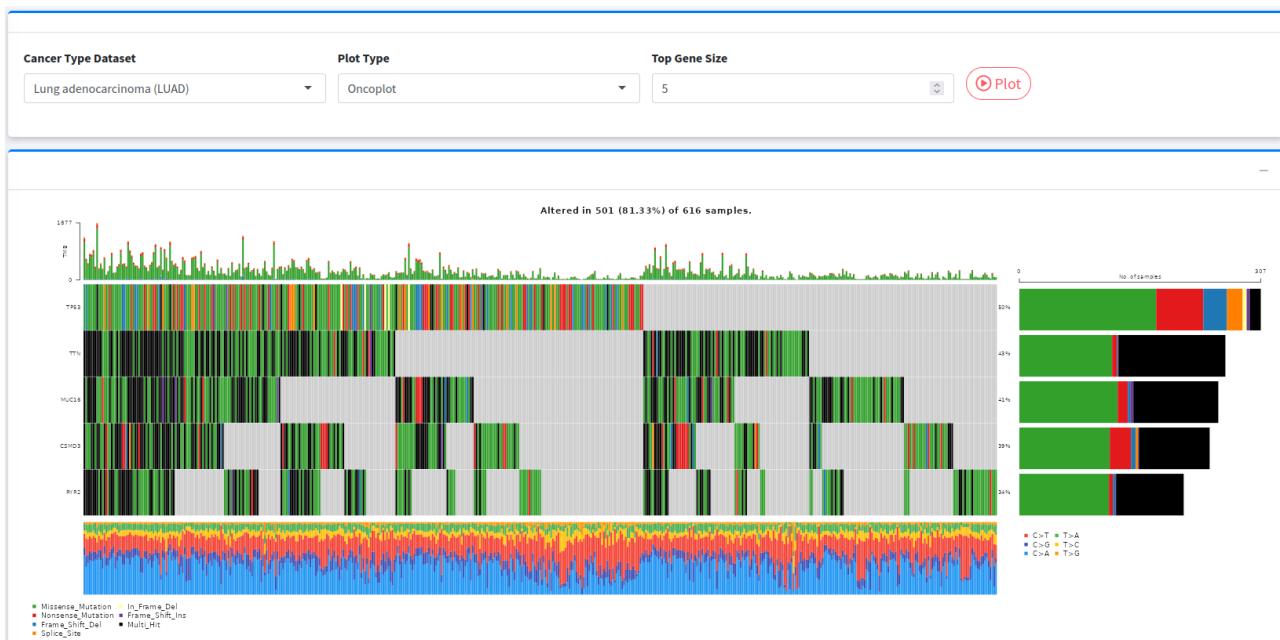
- Once you have chosen the cancer type, it is time to select the type of SNV plot you want to generate:
- Click the "Plot Type" button number ②
- Choose from the available plot types, which may include Oncoplot, Heatmap, or Lollipop plot, depending on your preference (Fig. 31,32,33,34).

## ③ Determining the Gene Size (Fig. 31)

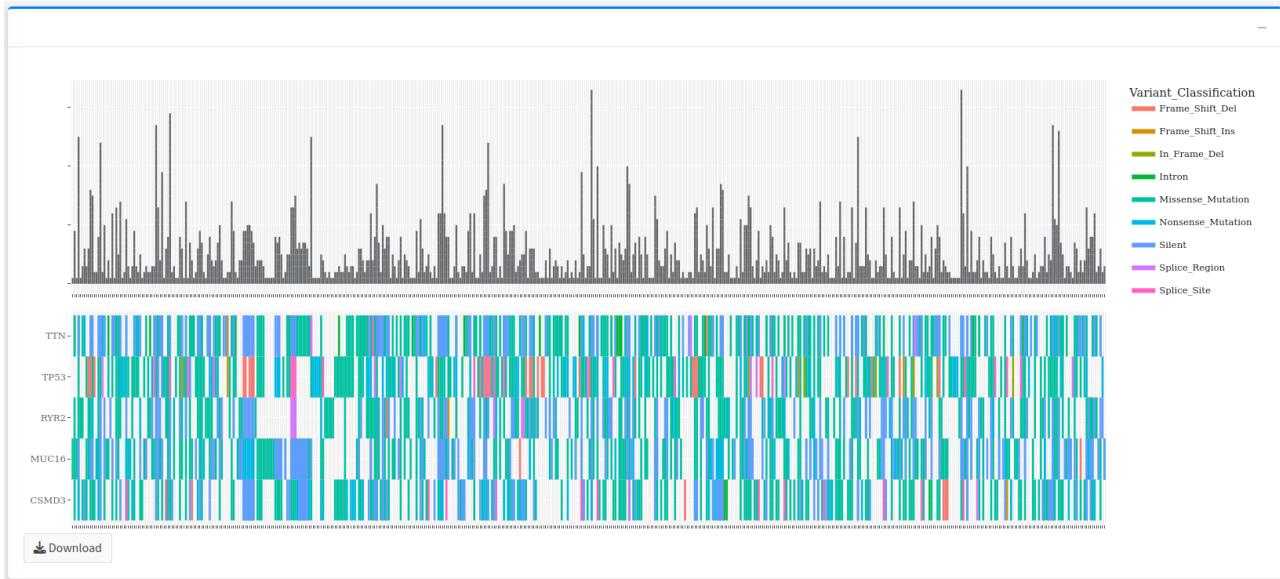
- Specify the gene size you want to include in the SNV analysis:
- Find the "Gene Size" input labeled ③
- Enter the desired gene size or range of gene sizes based on your analysis goals.

## ④ Generating the Plot (Fig. 31)

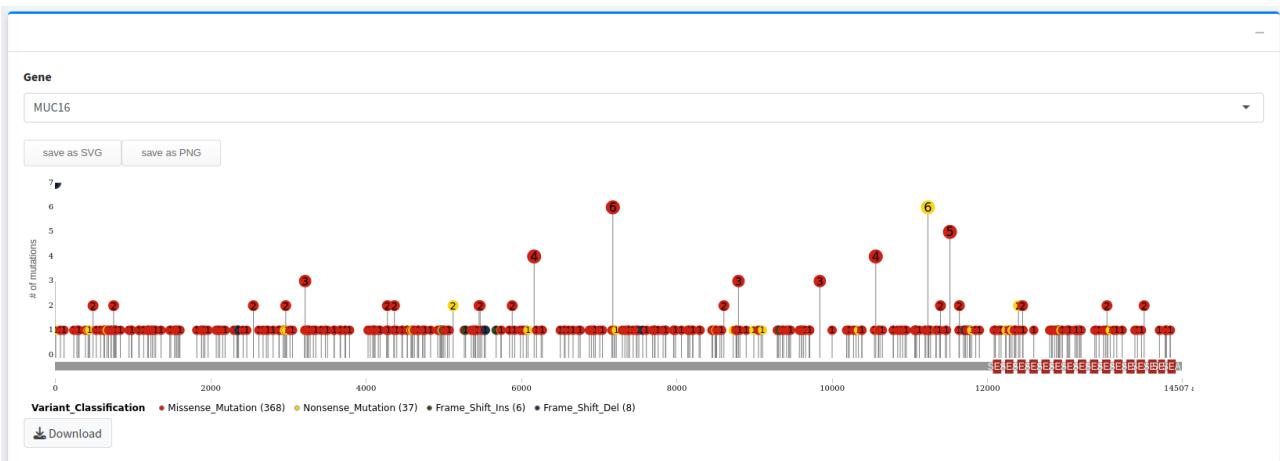
- After configuring your analysis parameters, you are ready to generate the SNV plot:
- Locate the "Plot" button number ④
- Click the "Plot" button to initiate the analysis with the specified cancer type, plot type, and gene size.



**Fig. 31.** Oncoplot of LUAD data



**Fig. 32.** Heatmap of LUAD data



**Fig. 33.** Lollipop plot of MUC16 gene in LUAD patients

A table showing mutation data for the MUC16 gene. The columns include Hugo\_Symbol, Entrez\_Gene\_Id, Center, NCBI\_Build, Chromosome, Start\_Position, End\_Position, Strand, Variant\_Classification, Variant\_Type, and Reference\_Allele. The data shows 8 entries for MUC16, all on chr19, with various mutation types (Missense\_Mutation, Nonsense\_Mutation) and reference alleles (T, A, G, C). A 'PDF' download button and a search bar are at the top, and a navigation bar with page numbers 1-42 is at the bottom.

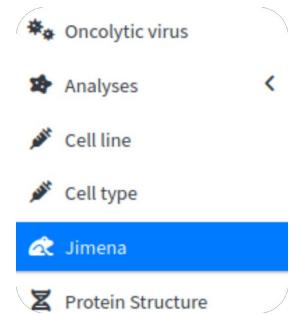
	Hugo_Symbol	Entrez_Gene_Id	Center	NCBI_Build	Chromosome	Start_Position	End_Position	Strand	Variant_Classification	Variant_Type	Reference_Allele
1	MUC16	94025	BI	GRCh38	chr19	8939097	8939097	+	Missense_Mutation	SNP	T
2	MUC16	94025	BI	GRCh38	chr19	8955964	8955964	+	Missense_Mutation	SNP	A
3	MUC16	94025	BI	GRCh38	chr19	8937140	8937140	+	Missense_Mutation	SNP	T
4	MUC16	94025	BI	GRCh38	chr19	8953476	8953476	+	Missense_Mutation	SNP	G
5	MUC16	94025	BI	GRCh38	chr19	8937575	8937575	+	Missense_Mutation	SNP	G
6	MUC16	94025	BI	GRCh38	chr19	8979071	8979071	+	Missense_Mutation	SNP	A
7	MUC16	94025	BI	GRCh38	chr19	8954224	8954224	+	Missense_Mutation	SNP	C
8	MUC16	94025	BI	GRCh38	chr19	8935553	8935553	+	Nonsense_Mutation	SNP	G

**Fig. 34** Mutation data of MUC16 in LUAD patients

# Gene Regulatory Network Analysis with Jimena

## ① Creating Signaling Network Data (Fig. 35)

- To begin, select the option to create signaling network data:
- Locate and click "Create Data," labeled ①



## ② Searching for Proteins of Interest (Fig. 35)

- You can start by searching for the protein(s) you are interested in using the "Search Protein" section:
- In the "Search Protein" section labeled ② enter the name of the protein you want to investigate.

## ③ Selecting Rows and Creating Your Table (Fig. 35)

- From the resulting table, select the rows that correspond to the signaling network data you want to use:
- Once you have selected the rows of interest, they will be transferred to a designated area labeled ④ creating your custom table.
- You can repeat this process by searching for other proteins and selecting additional signaling networks as needed.

## ④ Removing Rows (Optional) (Fig. 35)

- If you need to remove a row from your custom table:
- Select the row you want to delete.
- Click the "Remove Selected Rows" button labeled ⑤

## ⑥ Converting to GraphML Format (Fig. 35)

- To use the data you have prepared with Jimena, you need to convert it into GraphML format:
- Click the "Convert GraphML" button labeled ⑥

## ⑦ Initiating the Analysis (Fig. 35)

After converting the data, you are ready to initiate the dynamic gene regulatory network analysis with Jimena:

Click the "Run Jimena" button labeled ⑦ to start the analysis.

Choose:  
① Create data  
② Use example data  
③ Upload your data

Search Protein:  
BCL2  
②

Show 10 entries  
node1 label node2

node1	label	node2
21	BCL2	③ inhibition BAD
22	BCL2	inhibition NLRP1
23	BCL2	inhibition TP53AIP1
24	BCL2	inhibition ITM2B
25	BCL2L1	activation BCL2
26	BCL2L11	inhibition BCL2
27	BIK	inhibition BCL2
28	BMF	inhibition BCL2
29	BNIP3	inhibition BCL2
30	BNIP3L	inhibition BCL2

Showing 21 to 30 of 60 entries  
⑥ Convert GraphML  
⑦ Run Jimena

Show 10 entries  
④ node1 label node2

node1	label	node2
6	BAD	activation TP53
7	BCL2	inhibition TP53
11	BRCA1	inhibition TP53
288	TP53	activation GDF15
1	GDF15	activation ERBB2
3	BAX	inhibition BCL2
8	BCL2	activation MYC
21	BCL2	inhibition BAD

Showing 1 to 8 of 8 entries  
⑤ Remove Selected Rows  
Previous Next

Fig. 35. Dynamic gene regulatory network analysis with Jimena

Begin by initiating the dynamic gene regulatory network analysis using Jimena as previously described.

After initiating the analysis, you can introduce perturbations to the network to simulate changes in gene activation levels. Follow these steps:

Find the "Add Perturbation" selection input box ① (Fig. 36)

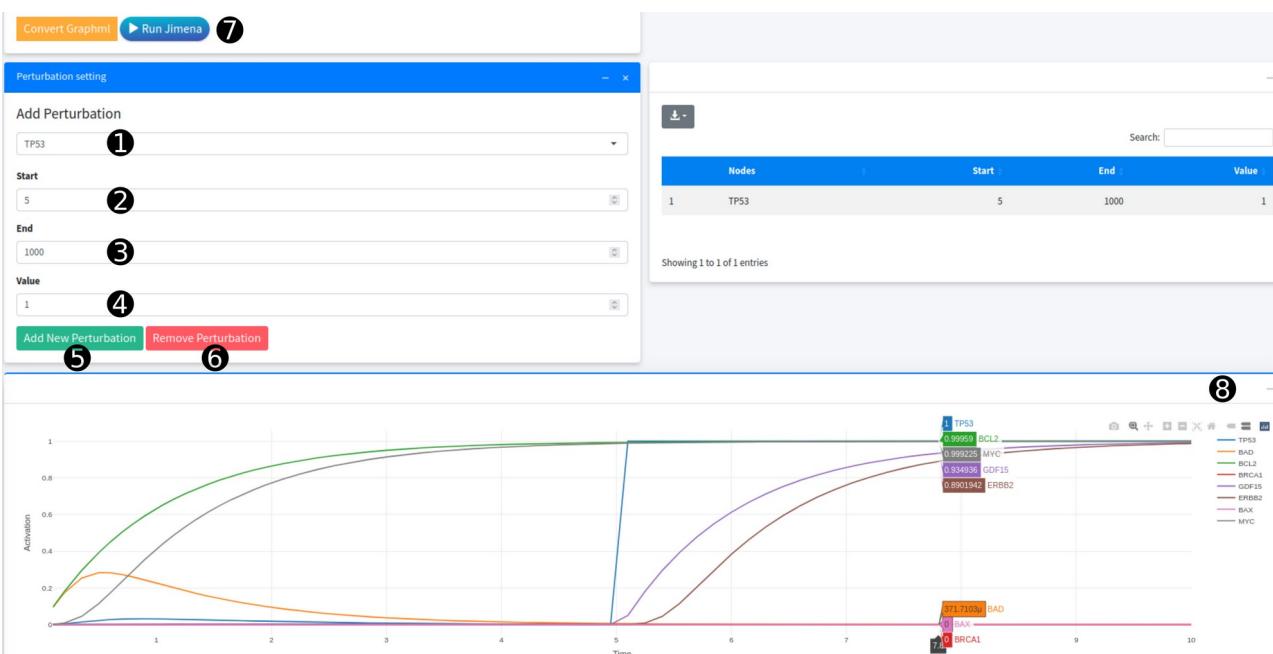
- Select the node you want to perturb (change its activation level) from the available options.
- Specify the time interval during which you want this node to change in the simulation (start ②- end ③).
- Choose the activation level for the selected node (a value between 0 and 1). You can use values like 0.5 to indicate partial activation ④.
- After configuring the perturbation, click the "Add New Perturbation" button to apply it ⑤.

⑥ Removing Perturbations (Optional) (Fig. 36)

- If you need to remove a perturbation you have added, follow these steps: Select the node for which you want to remove the perturbation from the analysis. Click the "Remove Perturbation" button to remove the selected perturbation ⑥. The perturbed node and values will be removed from the analysis table.

⑦ Running the Analysis with Perturbations (Fig. 36)

- Once you have added or removed perturbations as needed, you can re-run the Jimena analysis with the updated network: Click the "Run Jimena" button again to perform the analysis based on the perturbations you have made. The interactive chart will be updated to reflect the changes in gene activation levels due to the perturbations. ⑧ The interactive chart and data will provide insights into how changes in gene activation levels impact the network dynamics over time (Fig. 33, 34).



**Fig. 36.** Cumulative results of the Jimena dynamic simulation, showcasing the effects of introduced perturbations.

The screenshot shows a table titled "Jimena output table" with a header row containing columns for "time", "TP53", "BAG", "BCL2", "BRCA1", "GDF15", "ERBB2", "BAX", and "MYC". Below the header, there are 8 data rows, each corresponding to a time point from 1 to 8. The data includes numerical values for each gene at each time point. A search bar is located at the top right of the table area.

time	TP53	BAG	BCL2	BRCA1	GDF15	ERBB2	BAX	MYC
1	0.1	0.0009592365436761785	0.09400430901281065	0.09516257705071343	0	0.00005494592560660389	3.310640399407163e-8	0
2	0.2	0.004626411608020169	0.1735273436473686	0.1812692380304939	0	0.00004387477357276553	3.644651895390331e-7	0
3	0.35	0.01415635248593018	0.253816896975857	0.295311896885263	0	0.0002412281178336296	0.000003003911211264468	0
4	0.4999999999999999	0.02329768300965781	0.2831452295172636	0.3934693238198506	0	0.0006468328097853363	0.00001125486754692076	0
5	0.6	0.0274292238544481	0.2826444252941251	0.4511883460254726	0	0.00100385778572482	0.00002125877822775899	0
6	0.7000000000000001	0.0299393786612215	0.273244306883248	0.503414677333153	0	0.001391552635002802	0.00003523117503884472	0
7	0.8000000000000002	0.03114684339474045	0.2591081240077262	0.5506710163636903	0	0.001778978395924895	0.00005298566570939976	0
8	0.9000000000000002	0.0313968578930243	0.2427660867989604	0.5934303203900995	0	0.002143047308459863	0.00007400693824306367	0

Showing 1 to 9 of 78 entries

**Fig. 37.** Jimena output table

If you have your own data in the form of a **txt** file containing "**node1**," "**label**," and "**node2**" columns, you can follow these steps to use your custom data: Locate the option for uploading your own data. Click the "**Upload your data**" button or a similar option to select and upload your custom text file. Ensure that your custom data file adheres to the format of "node1," "label," and "node2" columns as required (Fig. 38).

Once your data is uploaded successfully, you can proceed to "**Convert Graphml**", "**Run Jimena**" and add or remove perturbations and run the analysis as described in the previous steps using your own data. By following these steps, you can integrate your custom data into the dynamic gene regulatory network analysis with Jimena, allowing you to analyze and visualize your specific regulatory network.

The screenshot shows the Jimena interface with a "Choose" section containing three options: "Create data" (radio button), "Use example data" (radio button), and "Upload your data" (checkbox, checked). Below this is a "Choose Txt File" section with a "Browse..." button and a "No file selected" text input field. At the bottom are two buttons: "Convert Graphml" (orange) and "Run Jimena" (blue).

**Fig. 38.** Upload your signaling network file

# Protein Structure- AlphaFold

## ① Searching for a Protein (Fig. 39)

- Begin by entering the Uniprot ID of the protein you want to search for. Locate the input box labeled ① and type in the Uniprot ID.
- After entering the Uniprot ID, click the "Search" button labeled ② to initiate the search.
- Once the search is complete, the structure of the protein you have selected will be displayed.

## ③ Selecting Regions of Interest (Fig. 39)

- To focus on specific regions of the protein, locate and use the selection parties labeled ③

## ④ Choosing a Structure Type (Fig. 39)

- You have the option to select different representations of the protein's structure. Locate the options for structure representation labeled ④ and choose the one that best suits your analysis.

## ⑤ Exploring Surface and Ligand Structure (Fig. 39)

- You can interact with the surface ⑤ and ligand ⑥ structures to gain further insights into the protein's characteristics.

## ⑦ Customizing Ligand Color (Fig. 39)

- If desired, you can change the color of the ligand. Locate the option for adjusting ligand color labeled ⑦

## ⑧ Modifying Background Color (Fig. 39)

- You can also change the background color on which the protein structure is presented. Locate the option for adjusting the background color labeled ⑧

## ⑨ Saving Images (Fig. 39)

- If you want to save an image of the protein structure, you can do so using the "Save" option labeled ⑨

## ⑩ Creating Animations (Fig. 39)

- To create animations of the protein structure, find and utilize the animation creation option labeled ⑩

The screenshot shows the AlphaFold Protein Structure Prediction interface. On the left, there is a sidebar with various controls:

- Options**:
  - Enter PDB name (UniProt ID): Q5VSL9 (labeled ①)
  - Search (labeled ②)
  - Selection: 1-20 (labeled ③)
  - Structure: cartoon (labeled ④)
  - Surface: hide (labeled ⑤)
  - Ligand: hide (labeled ⑥)
  - Select colour: #FFFF00 (labeled ⑦)
  - Background color: #000000 (labeled ⑧)
  - Add, Remove, Save (labeled ⑨)
  - Fullscreen (labeled ⑩)
  - Animation: None, Spin, Rock

The main area displays a protein structure as a ribbon model against a black background. A legend is visible on the left side of the main window, mapping colors to pLDDT values:

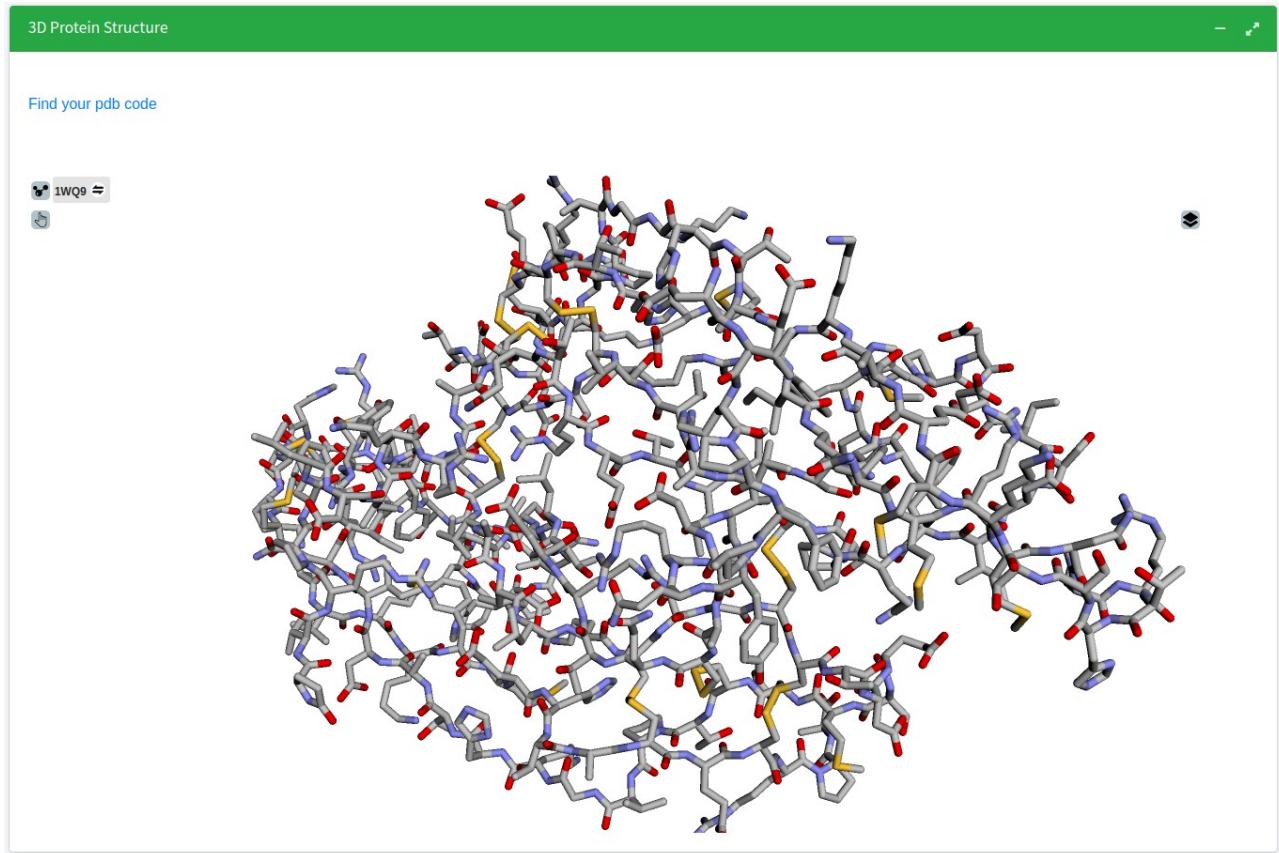
Color	pLDDT Range
Red	Very low (pLDDT < 50)
Yellow	Low (70 > pLDDT > 50)
Cyan	Confident (90 > pLDDT > 70)
Blue	Very high (pLDDT > 90)

The protein structure is composed of blue and yellow ribbons, indicating higher confidence regions.

**Fig. 39.** Protein structure prediction with AlphaFold

## Protein Structure- 3Dmol.js

Please enter the PDB ID to access the protein structure. In case you are unsure about the ID of the protein you are interested in, you can locate it on the respective website by selecting the "[Find Your PDB Code](#)" option (Fig. 40).



**Fig. 40.** Protein structure of 1WQ9

# Cell line Database

To see information about cell lines, type the cell line you want to search in the "search" section and click on the search icon (Fig. 41).

The screenshot shows the Cell line Database interface. At the top, there is a search bar with the text "panc-1" and a search icon. Below the search bar, there are several filter options: "Oncolytic virus", "Analyses", "Cell line" (which is highlighted in blue), "Cell type", "Jimena", and "Protein Structure". The main content area is titled "Cell line information" and contains a table with one row for the PANC-1 cell line. The table columns include: DepMap\_ID, cell\_line\_name, stripped\_cell\_line\_name, CCLE\_Name, alias, COSMICID, sex, source, RRID, WTSI\_Master\_Cell\_ID, and sample\_collection\_site. The row data is: 1, ACH-000164, PANC-1, PANC1, PANC1\_PANCREAS, Male, ATCC, CVCL\_0480, pancreas. Below the table, there are detailed annotations: "primary\_or\_metastasis" (Primary), "primary\_disease" (Pancreatic Cancer), "Subtype" (Ductal Adenocarcinoma, exocrine), "age" (56), "Sanger\_Model\_ID" (SIDM00610), "depmap\_public\_comments", "lineage" (pancreas), and "lineage\_subtype" (exocrine). At the bottom of the table, it says "Showing 1 to 1 of 1 entries". On the right side of the table, there are buttons for "Previous", "1", and "Next".

DepMap_ID	cell_line_name	stripped_cell_line_name	CCLE_Name	alias	COSMICID	sex	source	RRID	WTSI_Master_Cell_ID	sample_collection_site
1	ACH-000164	PANC-1	PANC1	PANC1_PANCREAS		Male	ATCC	CVCL_0480		pancreas

**Fig. 41.** Information of PANC-1 cell line

# Cell type Database

Select the tissue to see the cell types of the tissue of interest (Fig. 42).

The screenshot shows a search interface for 'pancreas'. At the top, there are several navigation links: 'Oncolytic virus', 'Analyses', 'Cell line', 'Cell type' (which is highlighted in blue), and 'Jimena'. Below this is a search bar with the term 'pancreas'. The main area is a table titled 'Select Cell Type for Gene Markers' with columns: species, tissue\_class, tissue\_type, uberonontology\_id, cancer\_type, cell\_type, cell\_name, cellontology\_id, marker, symbol, GenelD, and Genotype. The table contains 6 rows of data, each with a blue circular icon followed by numerical values. At the bottom of the table, it says 'Showing 1 to 10 of 537 entries (filtered from 15,107 total entries)' and includes a page navigation bar with buttons for 'Previous', '1', '2', '3', '4', '5', '...', '54', and 'Next'.

**Fig. 42.** Cell types of pancreas tissue

Gene markers of cell types are listed according to Cancer or Normal tissue (Fig. 43).

The screenshot shows a search interface for 'Pancreas'. At the top, there is a dropdown menu labeled 'Select Tissue For Cell Type Samples' with 'Pancreas' selected. Below this is a search bar with the term 'Pancreas'. The main area is a table with columns: SRA, SRS, Tissue, Protocol, Species, Cluster, Cells, and Cell Type. The table contains 10 rows of data, each with numerical values. At the bottom of the table, it says 'Showing 1 to 10 of 293 entries' and includes a page navigation bar with buttons for 'Previous', '1', '2', '3', '4', '5', '...', '30', and 'Next'.

**Fig. 43.** Gene markers of pancreas pancer cells

## Drug Description

- To begin, select the "Drug" option labeled ① from the available menu (Fig. 44).
- In the "Search" section labeled ② enter the name of the drug you want to search for (Fig. 44).
- After entering the drug name, the drug and its structure will be displayed, typically in a visual format, as shown in ③ (Fig. 44)
- To explore additional drug-related information, you can search for specific terms within the tables. Look for the "Search" in table option numbered ④ (Fig. 44)
- If you want to retain the drug-related data for future reference, you can save the tables in various formats such as PDF, CSV, or Excel. Locate the option for saving tables labeled ⑤ (Fig. 44)

The screenshot shows the DrugBank search results for Aspirin. At the top left, under 'Choose:', 'Drug' is selected (labeled ①). In the search bar (labeled ②), 'aspirin' is entered. Below the search bar, the text 'Click search icon to update or hit 'Enter'' is displayed. To the right, a large chemical structure of Aspirin (labeled ③) is shown, consisting of a benzene ring with a hydroxyl group (-OH) and an acetoxy group (-OCH<sub>3</sub>). Below the structure, its InChI string, PubChem CID, Canonical SMILES, Standard InChI, Standard InChI Key, Molecular Formula, and Molecular Weight are listed.

On the right side, a table (labeled ⑤) provides detailed information about Aspirin. The table has columns for DrugBank\_Accession\_Number, Generic\_Name, Summary, Background, and Indication. The first entry (DB00945) for Aspirin is shown. The 'Summary' column contains a detailed description of Aspirin's uses and properties. The 'Indication' column lists its common uses: pain, fever, and inflammation. A 'Search:' input field (labeled ④) is located at the top right of the table area. Navigation buttons 'Previous', '1', and 'Next' are at the bottom right of the table.

**Fig. 44.** Information and structure of Aspirin

You can also search for a term you want by selecting "indication" (Fig. 45).

The screenshot shows the DrugBank search results for 'pancreas'. Under 'Choose:', 'Indication' is selected (labeled ①). In the search bar (labeled ②), 'pancreas' is entered. Below the search bar, the text 'Click search icon to update or hit 'Enter'' is displayed.

On the right side, a table (labeled ⑤) provides information about Pancrelipase. The table has columns for DrugBank\_Accession\_Number, Generic\_Name, Summary, and Background. The first entry (DB00085) for Pancrelipase is shown. The 'Summary' column contains a detailed description of Pancrelipase's composition and uses. A 'Search:' input field (labeled ④) is located at the top right of the table area. Navigation buttons 'Previous', '1', '2', and 'Next' are at the bottom right of the table.

**Fig. 45.** Indication of pancreas

## Drug-Tissue Database

Drug

- » Description
- » Tissues**
- » Genes
- » STITCH

The screenshot shows a table with the following columns: Drug.name, Drug.Id, Cell.line.name, Cosmic.sample.Id, TCGA.classification, Tissue, Tissue.sub.type, IC50, AUC, and Max.conc. The first row is highlighted in blue and contains: Camptothecin, 1003, Calu-3, 687777, LUAD, lung, lung\_NSCLC\_adenocarcinoma, -1.99331798052966, 0.852331146598283, 0.1. Below the table, there are RMSE and Z.score values: 0.134727053386585 and 0.057655362688808 respectively. The dataset version is GDSC2.

Drug.name	Drug.Id	Cell.line.name	Cosmic.sample.Id	TCGA.classification	Tissue	Tissue.sub.type	IC50	AUC	Max.conc	
Camptothecin	1003	Calu-3	687777	LUAD	lung	lung_NSCLC_adenocarcinoma	-1.99331798052966	0.852331146598283	0.1	
RMSE	0.134727053386585									
Z.score	0.057655362688808									
Dataset.version GDSC2										
(1) 2	Camptothecin	1003	NCI-H1623	687798	LUAD	lung	lung_NSCLC_adenocarcinoma	-4.17608648770136	0.669750000619124	0.1
(2) 3	Camptothecin	1003	NCI-H1648	687799	LUAD	lung	lung_NSCLC_adenocarcinoma	-2.81789067780435	0.823600404927073	0.1
(3) 4	Camptothecin	1003	NCI-H1650	687800	LUAD	lung	lung_NSCLC_adenocarcinoma	0.736512593614056	0.923567963019469	0.1
(4) 5	Camptothecin	1003	NCI-H1693	687802	LUAD	lung	lung_NSCLC_adenocarcinoma	0.889164021950653	0.937692035004125	0.1

Showing 1 to 10 of 28,660 entries

Previous 1 2 3 4 5 ... 2866 Next

Fig. 46. Drug list of LUAD

## Drug-Gene Database

Drug

- » Description
- » Tissues
- » Genes**
- » STITCH

The screenshot shows a table with the following columns: entrez\_id, gene\_name, gene\_claim\_name, interaction\_claim\_source, interaction\_types, drug\_claim\_name, drug\_claim\_primary\_name, drug\_name, and drug\_concept\_id. The first row is highlighted in blue and contains: 7157, TP53, TP53, DTC, CAULIBUGULONE D, CAULIBUGULONE D, CAULIBUGULONE D, chembl:CHEMBL447876. Below the table, there are 8 rows of data corresponding to TP53 interactions.

entrez_id	gene_name	gene_claim_name	interaction_claim_source	interaction_types	drug_claim_name	drug_claim_primary_name	drug_name	drug_concept_id
(1) 1	7157	TP53	TP53	DTC	CAULIBUGULONE D	CAULIBUGULONE D	CAULIBUGULONE D	chembl:CHEMBL447876
(2) 2	7157	TP53	TP53	PharmGKB	fluorouracil	fluorouracil	FLUOROURACIL	chembl:CHEMBL185
(3) 3	7157	TP53	TP53	DTC	CID 867101	CID 867101	CHEMBL406557	chembl:CHEMBL406557
(4) 4	7157	TP53	TP53	DTC	SJ000276790	SJ000276790	CHEMBL602922	chembl:CHEMBL602922
(5) 5	7157	TP53	TP53	DTC	OLEOYL DOPAMINE	OLEOYL DOPAMINE	OLEOYL DOPAMINE	chembl:CHEMBL250711
(6) 6	7157	TP53	TP53	DTC	DNDI1417085	DNDI1417085	CHEMBL1522984	chembl:CHEMBL1522984
(7) 7	7157	TP53	TP53	DTC	DNDI1340257	DNDI1340257	CHEMBL1702181	chembl:CHEMBL1702181
(8) 8	7157	TP53	TP53	DTC	TEBUFENPYRAD	TEBUFENPYRAD	TEBUFENPYRAD	chembl:CHEMBL1897070

Showing 1 to 10 of 497 entries

Previous 1 2 3 4 5 ... 50 Next

Fig. 47. Drug list of TP53

# Immune Modulatory Database

**Bispecific Antibody**

You can access clinical trials data of bispecific antibody studies (Fig. 48).

The screenshot shows a table of clinical trials for bispecific antibodies. The columns include NCT.Number, Title, Status, Study.Results, Conditions, Interventions, Outcome.Measures, Phases, and Study.Type. One study is listed:

NCT.Number	Title	Status	Study.Results	Conditions	Interventions	Outcome.Measures	Phases	Study.Type
1	Individual Dietary Counseling Based on Taste-tests in Patients With Hematological Cancer in Cytostatic Therapy	Enrolling by invitation	No Results Available	Hematological Malignancy	Dietary Supplement: Taste-test Other: Standard care	Consumption of energy as a percentage of estimated need Fat-free-mass and fat-mass in percent (bioimpedance measurement) Satisfaction (measured with EORTC QLQ-C30 version 3.0) Hand-grip-strength (measured with a dynamometer) Body weight Protein intake Burning/pain in the mouth (measured on a scale)	Not Applicable	Interventional

**Study.Designs** Allocation: Randomized|Intervention Model: Parallel Assignment|Masking: Single (Participant)|Primary Purpose: Other

Showing 1 to 10 of 189 entries

**Fig. 48.** Bispecific antibody table

**CAR-T Cell Therapy**

You can access clinical trial data of CAR-T Cell therapy studies (Fig. 49).

The screenshot shows a table of clinical trials for CAR-T cell therapy. The columns include NCT.Number, Title, Status, Study.Results, Conditions, Interventions, Outcome.Measures, Phases, and Study.Type. One study is listed:

NCT.Number	Title	Status	Study.Results	Conditions	Interventions	Outcome.Measures	Phases	Study.Type
1	GPC3/Mesothelin /Claudin18.2/GUCY2C /BT-H3/PSCA /PSMA/MUC1 /TGFβ/HER2/Lewis-Y/AXL/EGFR-CAR-T Cells Against Cancers	Recruiting	No Results Available	Lung Cancer/Cancer Immunotherapy CAR-T Cell	Biological: CAR-T cells targeting GPC3, Mesothelin, Claudin18.2, GUCY2C, BT-H3, PSCA, PSMA, MUC1, TGF $\beta$ , HER2, Lewis-Y, AXL, or EGFR	Number of Patients with Dose Limiting Toxicity Percent of Patients with best response as either complete remission or partial remission. Median CAR-T cell persistence	Phase 1	Interventional

**Study.Designs** Allocation: N/A|Intervention Model: Single Group Assignment|Masking: None (Open Label)|Primary Purpose: Treatment

Showing 1 to 10 of 28 entries

**Fig. 49.** CAR-T Cell therapy table

## Immune Modulatory

- » Bispecific Antibody
- » CAR-T cell Therapy new
- » Checkpoint**
- » Cytostatic Therapy
- » Oncolytic virus Therapy

# Checkpoint

You can access checkpoint data (Fig. 50).

The screenshot shows a table with the following columns: Name, Full name, Tumor types, Primary References, Journal...5, Score, Crystal Structure, Binding Site, Entrez Gene ID, and Uniprot ID. The first row contains data for IL-23, with a blue background highlighting the entire row.

Name	Full name	Tumor types	Primary References	Journal...5	Score	Crystal Structure	Binding Site	Entrez Gene ID	Uniprot ID	
1	IL-23	Interleukin-23	Kidney Cancer	Tumor-associated Macrophage-derived Interleukin-23 Interlinks Kidney Cancer Glutamine Addiction with Immune Evasion.	Eur Urol.2018 Oct 4 pii:S0302-2838(18)30718-8	32	PDB code: 5Mxa	FTMap/FTSite	51561	Q9NPF7

**Function:** This gene encodes a subunit of the heterodimeric cytokine interleukin 23 (IL23). IL23 is composed of this protein and the p40 subunit of interleukin 12 (IL12B). The receptor of IL23 is formed by the beta 1 subunit of IL12 (IL12R $\beta$ 1) and an IL23 specific subunit, IL23R. Both IL23 and IL12 can activate the transcription activator STAT4, and stimulate the production of interferon-gamma (IFNG). In contrast to IL12, which acts mainly on naive CD4(+) T cells, IL23 preferentially acts on memory CD4(+) T cells.

**GeneCards Link:** <https://www.genecards.org/cgi-bin/carddisp.pl?gene=IL23A&keywords=IL-23>

**Complete Relevant Tumor Types:** Leukemia; Bone Cancer; Lung Cancer; Bladder Cancer; Lymphoma; Prostate Cancer; Kidney Cancer

**2nd reference:** Tumor-associated Macrophage-derived Interleukin-23 Interlinks Kidney Cancer Glutamine Addiction with Immune Evasion

**Journal...15:** Eur Urol. 2019 May;75(5):752-763.

**2nd reference:** Now insights into the IL-12 and IL-23: From a molecular basic to clinical application in immune-mediated inflammation and cancer

Showing 1 to 10 of 105 entries

Previous 1 2 3 4 5 ... 11 Next

Fig. 50. Checkpoint table

# Cytostatic Therapy

## Immune Modulatory

- » Bispecific Antibody
- » CAR-T cell Therapy new
- » Checkpoint**
- » Cytostatic Therapy**
- » Oncolytic virus Therapy

You can access clinical trial data of cytostatic therapy studies (Fig. 51).

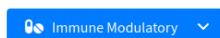
The screenshot shows a table with the following columns: NCT.Number, Title, Status, Study.Results, Conditions, Interventions, Outcome.Measures, Phases, and Study.Type. The first two rows are highlighted with blue backgrounds.

NCT.Number	Title	Status	Study.Results	Conditions	Interventions	Outcome.Measures	Phases	Study.Type
1 NCT05364359	Individual Dietary Counseling Based on Taste-tests in Patients With Hematological Cancer in Cytostatic Therapy	Enrolling by invitation	No Results Available	Hematological Malignancy	Dietary Supplement: Taste-test Other: Standard care	Consumption of energy as a percentage of estimated need Fat-free-mass and fat-mass in percent (bioimpedance measurement) Satisfaction (measured with EORTC QLQ-C30 version 3.0) Hand-grip strength (measured with a dynamometer) Body weight Protein intake Burning/pain in the mouth (measured on a scale)	Not Applicable	Interventional
2 NCT00541801	Acoustic Cardiographic Assessment of Heart Function in Comparison to	Completed	No Results Available	Heart Failure, Congestive Antineoplastic Agents	Device: Doppler-echocardiography and acoustic	Changes of echocardiographic parameters and changes of electromechanical activation time	Not Applicable	Interventional

Showing 1 to 10 of 189 entries

Previous 1 2 3 4 5 ... 19 Next

Fig. 51. Cytostatic therapy table



## Oncolytic Virus Therapy

» Bispecific Antibody

» CAR-T cell Therapy new

» Checkpoint

» Cytostatic Therapy

» Oncolytic virus Therapy

You can access clinical trial data of oncolytic virus therapy studies (Fig. 52).

NCT.Number	Title	Status	Study.Results	Conditions	Interventions	Outcome.Measures	Phases	Study.Type
1 NCT03004183	SBRT and Oncolytic Virus Therapy Before Pembrolizumab for Metastatic TNBC and NSCLC	Active, not recruiting	Has Results	Metastatic Non-small Cell Lung Cancer Metastatic Triple-negative Breast Cancer	Biological: ADV/HSV-tk Drug: Valacyclovir Radiation: SBRT Drug: Pembrolizumab	Objective Response Rate Duration of Response Overall Survival Rate Progression-free Survival Rate Number of Participants With Treatment-related Adverse Events Antitumor Activity Clinical Benefit Rate	Phase 2	Interventional
2 NCT03282916	Anti-viral Therapy in Alzheimer's Disease	Active, not recruiting	No Results Available	Alzheimer Disease Herpes Simplex 1 Herpes Simplex 2	Drug: Valacyclovir Drug: Placebo	Change in Alzheimer's Disease Assessment Scale - Cognition (ADAS-COG11, modified version) scores from baseline to 78 weeks. Change in Alzheimer's Disease Cooperative Study - Activities of Daily Living (ADCS-ADL) scores from baseline to 78 weeks. Change in total 18F-Florbetapir brain uptake from baseline to 78 weeks. Change in total 18F-MK-6240 brain uptake from baseline to 78 weeks.	Phase 2	Interventional

Showing 1 to 10 of 85 entries

Previous 1 2 3 4 5 ... 9 Next

**Fig. 52.** Oncolytic Virus Therapy table

## HOTLIST

Below are some crucial databases. Access to certain ones, like the Cochrane Library, is typically provided by institutions.

**Cochrane Library** - This renowned online source is a hub for evidence-based healthcare research and systematic reviews. <https://www.cochranelibrary.com>

**DrumPID** - The DrumPID database is an extensive resource that offers tailored data about medications and their protein networks. It includes information on indications, targets, and side-effects, making it indispensable for drug development, predicting adverse reactions, and studying structure-activity relationships. <https://drumpid.bioapps.biozentrum.uni-wuerzburg.de>

**Europe PMC** - Europe PMC is a valuable online platform that grants easy access to a wide array of biomedical literature and research articles. It simplifies the retrieval of scientific information for researchers and healthcare professionals. <https://europepmc.org>

**Peptide to SMILES Conversion** - NovoPro Labs provides a user-friendly online tool for converting peptide sequences into the simplified molecular input line entry system (SMILES) notation. This notation is used for representing and analyzing chemical structures. <https://www.novoprolabs.com/tools/convert-peptide-to-smiles-string>

## HOT PAPERS

Peindl M, Göttlich C, Crouch S, Hoff N, Lüttgens T, Schmitt F, Pereira JGN, May C, Schliermann A, Kronenthaler C, Cheufou D, Reu-Hofer S, Rosenwald A, Weigl E, Walles T, Schüler J, Dandekar T, Nietzer S, Dandekar G. EMT, Stemness, and Drug Resistance in Biological Context: A 3D Tumor Tissue/In Silico Platform for Analysis of Combinatorial Treatment in NSCLC with Aggressive KRAS-Biomarker Signatures. *Cancers (Basel)*. 2022 Apr 27;14(9):2176. doi: 10.3390/cancers14092176. PMID: 35565305; PMCID: PMC9099837.

Kühnemundt J, Leifeld H, Scherg F, Schmitt M, Nelke LC, Schmitt T, Baur F, Göttlich C, Fuchs M, Kunz M, Peindl M, Brähler C, Kronenthaler C, Wischhusen J, Prelog M, Walles H, Dandekar T, Dandekar G, Nietzer SL. Modular micro-physiological human tumor/tissue models based on decellularized tissue for improved preclinical testing. *ALTEX*. 2020 Dec 11;38(2):289-306. doi: 10.14573/altex.2008141. Epub ahead of print. PMID: 33313956.

Baur F, Nietzer SL, Kunz M, Saal F, Jeromin J, Matschos S, Linnebacher M, Walles H, Dandekar T, Dandekar G. Connecting Cancer Pathways to Tumor Engines: A Stratification Tool for Colorectal Cancer Combining Human In Vitro Tissue Models with Boolean In Silico Models. *Cancers (Basel)*. 2019 Dec 20;12(1):28. doi: 10.3390/cancers12010028. PMID: 31861874; PMCID: PMC7017315.

Göttlich C, Kunz M, Zapp C, Nietzer SL, Walles H, Dandekar T, Dandekar G. A combined tissue-engineered/in silico signature tool patient stratification in lung cancer. *Mol Oncol*. 2018 Aug;12(8):1264-1285. Doi: 10.1002/1878-0261.12323. Epub 2018 Jun 22. PMID: 29797762; PMCID: PMC6068345.

Göttlich C, Müller LC, Kunz M, Schmitt F, Walles H, Walles T, Dandekar T, Dandekar G, Nietzer SL. A Combined 3D Tissue Engineered In Vitro/In Silico Lung Tumor Model for Predicting Drug Effectiveness in Specific Mutational Backgrounds. *J Vis Exp*. 2016 Apr 6;(110):e53885. doi: 10.3791/53885. PMID: 27077967; PMCID: PMC4841364.

Stratmann AT, Fecher D, Wangorsch G, Göttlich C, Walles T, Walles H, Dandekar T, Dandekar G, Nietzer SL. Establishment of a human 3D lung cancer model based on a biological tissue matrix combined with a Boolean in silico model. *Mol Oncol*. 2014 Mar;8(2):351-65. doi: 10.1016/j.molonc.2013.11.009. Epub 2013 Dec 18. PMID: 24388494; PMCID: PMC5528544.

## References

- Colaprico, A., Silva, T. C., Olsen, C., Garofano, L., Cava, C., Carolini, D., ... & Noushmehr, H. (2016). TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic acids research*, 44(8), e71-e71.
- Emelianova, M., Gainullina, A., Poperechnyi, N., Loboda, A., Artyomov, M., & Sergushichev, A. (2022). Shiny GATOM: omics-based identification of regulated metabolic modules in atom transition networks. *Nucleic acids research*, 50(W1), W690-W696.
- Gao, J., Aksoy, B. A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S. O., ... & Schultz, N. (2013). Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Science signaling*, 6(269), pl1-pl1.
- Gonzalez-Perez, A., Perez-Llamas, C., Deu-Pons, J., Tamborero, D., Schroeder, M. P., Jene-Sanz, A., ... & Lopez-Bigas, N. (2013). IntOGen-mutations identifies cancer drivers across tumor types. *Nature methods*, 10(11), 1081-1082.
- Noecker, C., Eng, A., Muller, E., & Borenstein, E. (2022). MIMOSA2: a metabolic network-based tool for inferring mechanism-supported relationships in microbiome-metabolome data. *Bioinformatics*, 38(6), 1615-1623.
- Porta-Pardo, E., Hrabe, T., & Godzik, A. (2015). Cancer3D: understanding cancer mutations through protein structures. *Nucleic acids research*, 43(D1), D968-D973.
- Sun, W., Duan, T., Ye, P., Chen, K., Zhang, G., Lai, M., & Zhang, H. (2018). TSVdb: a web-tool for TCGA splicing variants analysis. *BMC genomics*, 19, 1-7.
- Tang, Z., Kang, B., Li, C., Chen, T., & Zhang, Z. (2019). GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic acids research*, 47(W1), W556-W560.