



# **GTEx Visualizer User Guide**

Version 1.0

Pietro Hiram Guzzi, Lomoio Ugo, Pierangelo Veltri  
Magna Graecia University of Catanzaro  
December 4, 2022

# Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Methods</b>	<b>2</b>
<b>3</b>	<b>Examples</b>	<b>3</b>
3.1	Select all expression data of a specific gene from all tissues .	3
3.1.1	No filters . . . . .	4
3.1.2	Filtered by Gender . . . . .	5
3.2	Gene expression of a specific gene in a specific tissue . . . . .	7
3.2.1	With no filters . . . . .	8
3.2.2	With Filter by Gender . . . . .	9
3.2.3	With Filter by Age . . . . .	10
3.2.4	With Filter by Age and Gender . . . . .	11
3.3	PPI subnetwork visualization and analysis of a specific gene that transcript to a protein related to a certain disease . . . . .	12
3.3.1	Method: Betweenness centrality . . . . .	14
3.3.2	Method: Closeness centrality . . . . .	15
3.3.3	Method: Degree centrality . . . . .	16
3.3.4	Method: Eigenvector centrality . . . . .	17
3.3.5	Method: Community extraction with Louvain . . . . .	18
3.3.6	Method: Community extraction with Leiden . . . . .	19
3.3.7	Method: Clusters extraction with Spectral Clustering .	20

# 1 Introduction

Study ageing processes and the possible gender differences at molecular level is currently a hot research area. Such studies require data, models, algorithms and tools for elucidating molecular mechanisms. The number of databases containing data useful for ageing studies is continuously growing, as well as data sources. The Genotype-Tissue Expression (GTEx) data portal is one of the prominent resources. GTEx stores transcriptomic, genomic and epigenomic data of individuals, annotated with information related to tissues, gender and human age. It offers many analysis and data extraction possibilities through a web interface. Despite this, the web interface lacks on the analysis at sex/age level, hence the researcher has to download data and then write queries by hand (e.g. for retrieving the expression of a given gene on different age-class in many tissues). It also lacks in the analysis of Protein-Protein Interaction networks of proteins related to ageing and diseases. Therefore, the need for the introduction of tools enabling easy access and powerful analysis capabilities arises. Nevertheless, disease studies based on age and genes analysis, may require to access data by using filtering mechanisms before downloading entire sections of GTEx databases.

This user guide present what can be used as a web portal allowing to support user in filtering GTEx datasets by using simple interface providing powerful and flexible data analysis and results representation modules. This web portal can also be used to visualize and analyze Protein-Protein Interaction subnetworks of genes that transcript to proteins involved in ageing processes or certain diseases.

# 2 Methods

GTExVisualizer is a web based platform hosted by Heroku and written in the Python programming language for ageing studies on the GTEx data. The platform is freely available at <https://gtexvisualizer.herokuapp.com/>.

Here we found a list of python libraries used for the development:

- dash, used for the realization of the web application layout;
- plotly, used for the realization of interactive Violin, Pie and Graph plots;
- scipy, used to compute some statistical tests to study the null hypothesis such as: the Shapiro test, the One way ANOVA test and the Kruskal test;

- sklearn, used to apply the spectral clustering algorithm on the PPI subnetwork to extract clusters;
- networkx, used to create the PPI subnetwork from the edgelist file and compute node centralities measures on this PPI;
- cdlib, used to apply Louvain and Leiden algorithm for extracting communities on the PPI;
- requests, used to make APIs http requests to obtain data, for example: GTex API is used to obtain gene data, STRING API is used to obtain the PPI of a specific protein.

The user can use a dropdown menu to set parameters such gene name, tissue name and other filters, for example age and sex, to apply to the GTEx data. Once finished, the user can update plots using the "Update Plots" button.

Here we found a list of all supported type of plots:

- Violin plot;
- Pie plot;
- Graph plot.

The lists of all possible genes and tissues are taken from the GTex database. The list of all possible filters to apply to the GTex data are:

- No filters;
- Filter by Age;
- Filter by Gender;
- Filter by Age and Gender.

## 3 Examples

Here we found some useful examples of usage of our web application to analyze gene and protein data related to a certain disease and ageing.

### 3.1 Select all expression data of a specific gene from all tissues

First Dropdown: GENE NAME

Second Dropdown: All

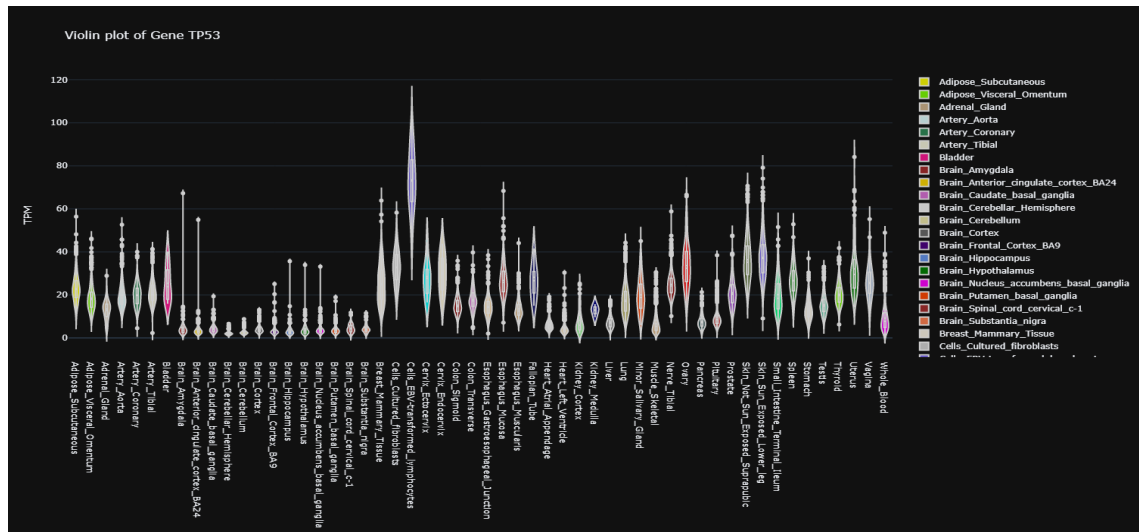


Figure 1: Violin plot for TP53 gene expression for all tissues. The web app provides to the user a slider for selecting the part of the violin plot to visualize.

### 3.1.1 No filters

Third Dropdown: No filters

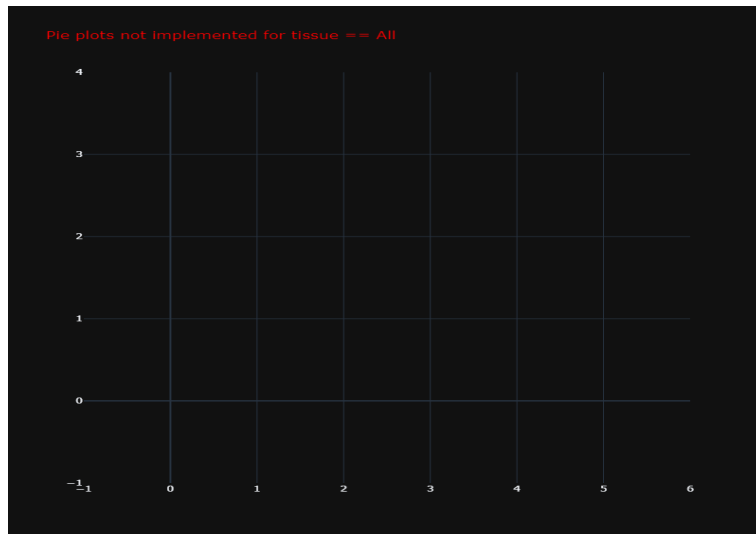


Figure 2: Pie plot for TP53 gene expression in all tissues. Not implemented so the web app returns an error.

### 3.1.2 Filtered by Gender

Third Dropdown: Filter by Gender

Violin plot for a specific gene expression in all tissues with divide by age filter is not implemented due to the amount of violin to insert (up to 500 violins).

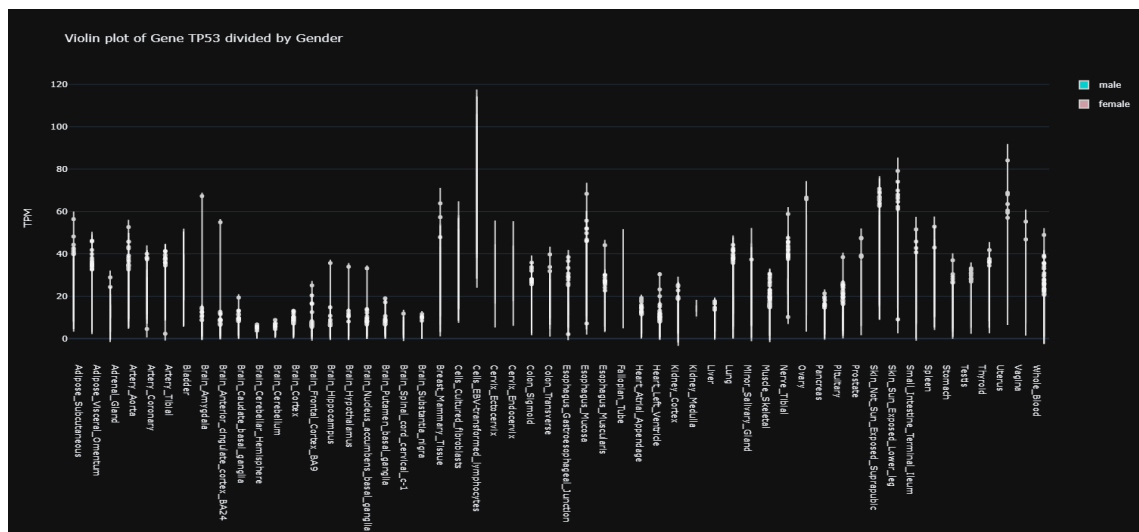


Figure 3: Violin plot for TP53 gene gene expression in all tissues and divided by gender. To much data in a single plot, we provided a box zoom feature to better visualize single violin data.

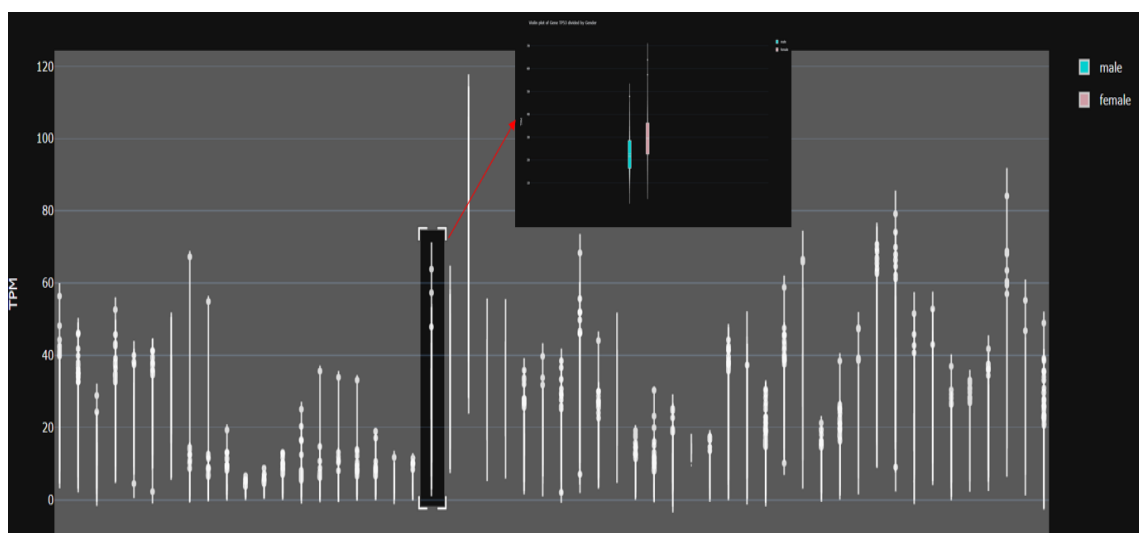


Figure 4: Violin plot for TP53 gene expression in all tissues divided by gender after using the box zoom feature.

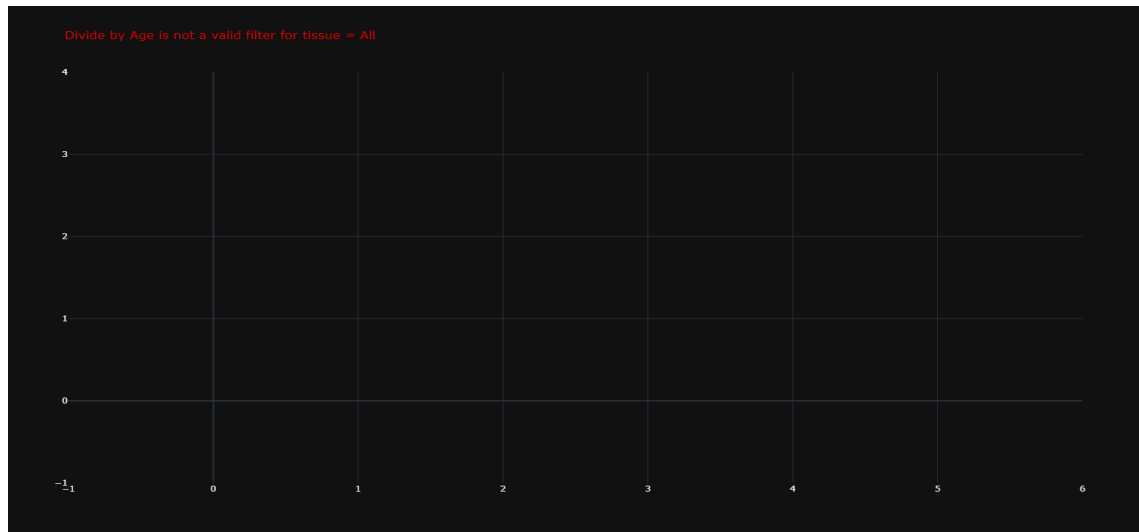


Figure 5: Violin plot for TP53 gene expression in all tissues divided by Age is not implemented error. Same error occurs with the Divide by Age and Gender filter.

### 3.2 Gene expression of a specific gene in a specific tissue

Second Dropdown: TISSUE



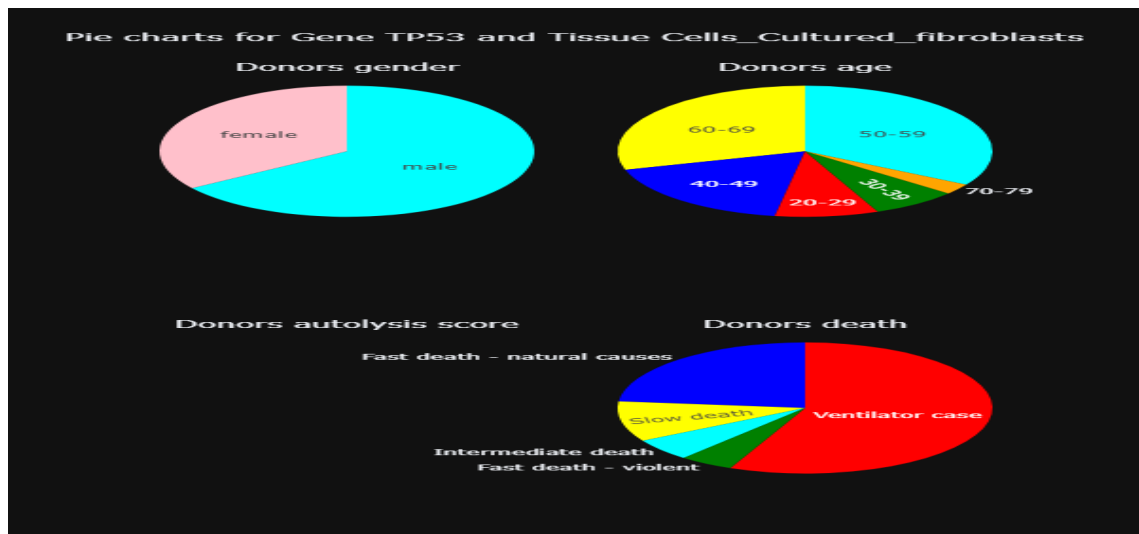


Figure 6: Pie plot for TP53 gene expression in tissue Cell Cultured Fibroblast

### 3.2.1 With no filters

Third Dropdown: No filters

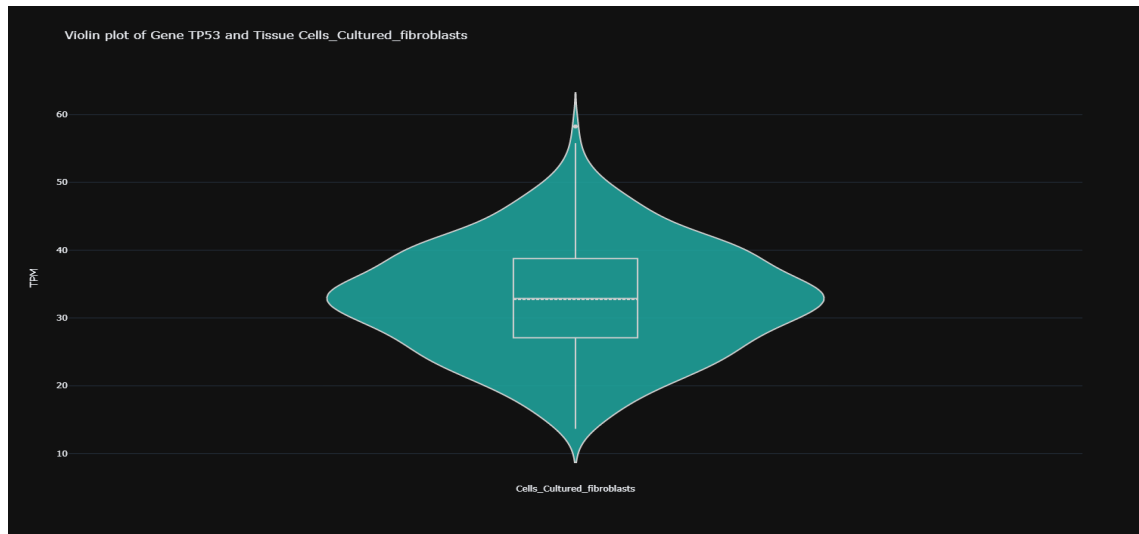


Figure 7: Violin plot for TP53 gene expression in tissue Cell Cultured Fibroblast and no filter.

### 3.2.2 With Filter by Gender

Third Dropdown: Filter by Gender

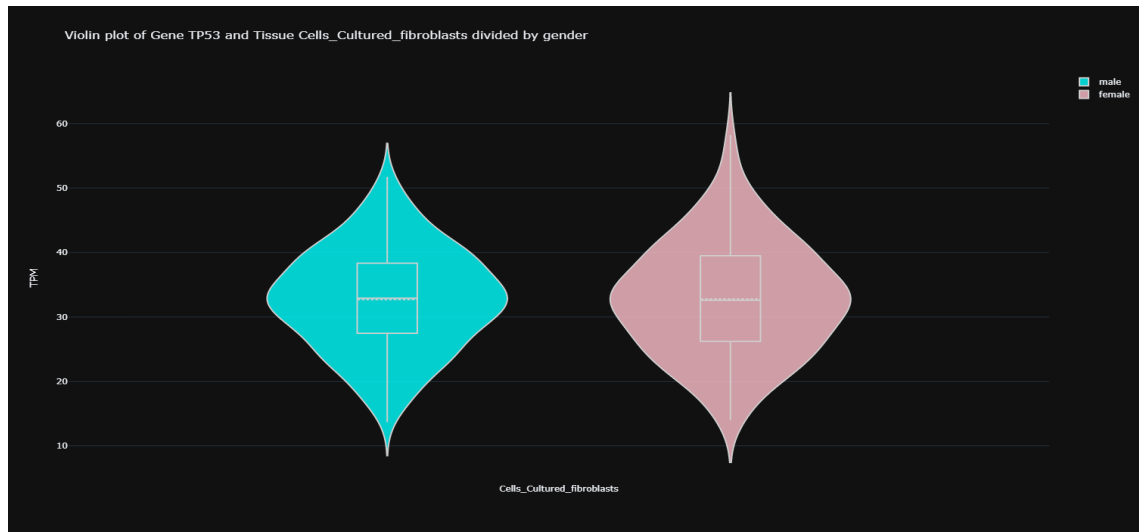


Figure 8: Violin plot for TP53 gene expression in tissue Cell Cultured Fibroblast and divided by gender

### 3.2.3 With Filter by Age

Third Dropdown: Filter by Age

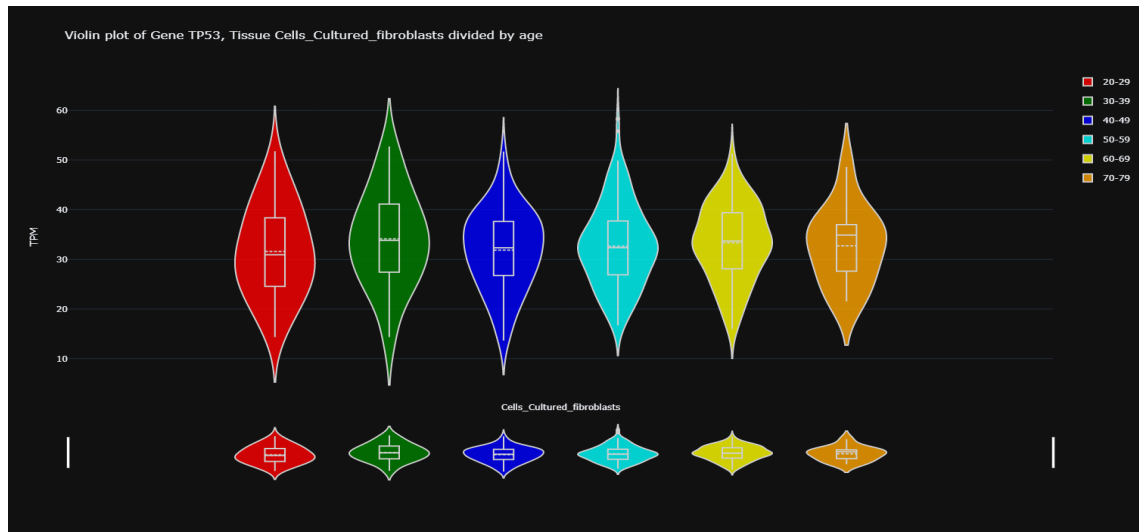


Figure 9: Violin plot for TP53 gene expression in tissue Cell Cultured Fibroblast divided by Age.

### 3.2.4 With Filter by Age and Gender

Third Dropdown: Filter by Age and Gender

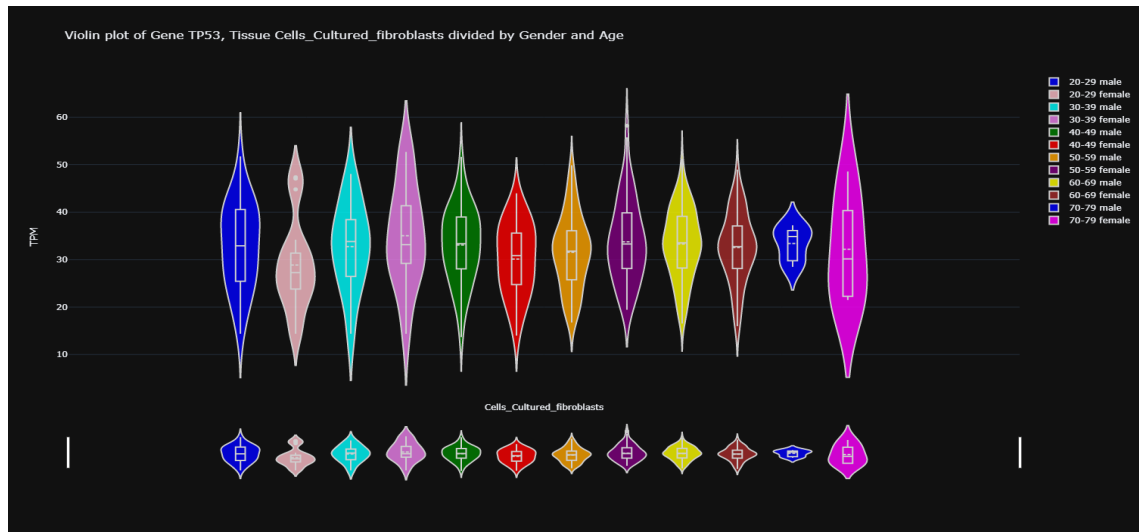


Figure 10: Violin plot for TP53 gene expression in tissue Cell Cultured Fibroblast divided by age and gender.

### 3.3 PPI subnetwork visualization and analysis of a specific gene that transcript to a protein related to a certain disease

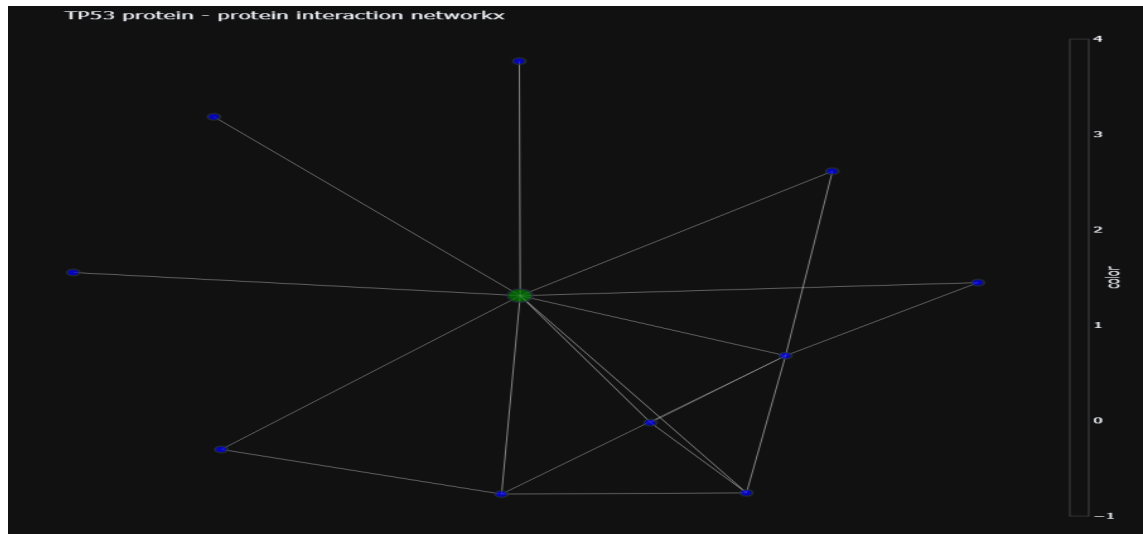


Figure 11: PPI plot for protein TP53.

In case of gene that not transcript to a protein, the web application returns the following error:

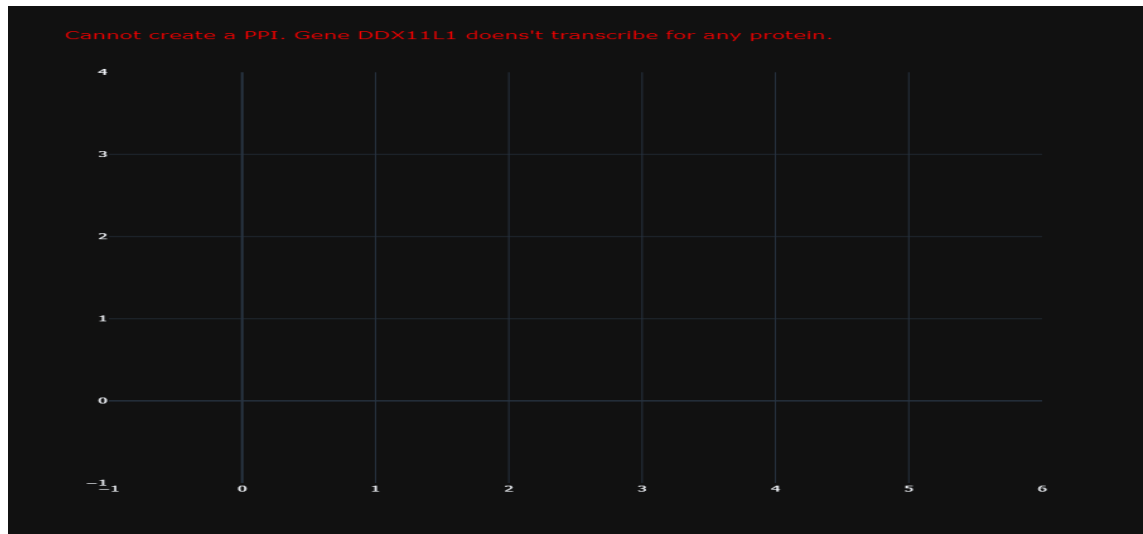


Figure 12: Gene DDX11L1 doesn't transcribe to a protein

### 3.3.1 Method: Betweenness centrality

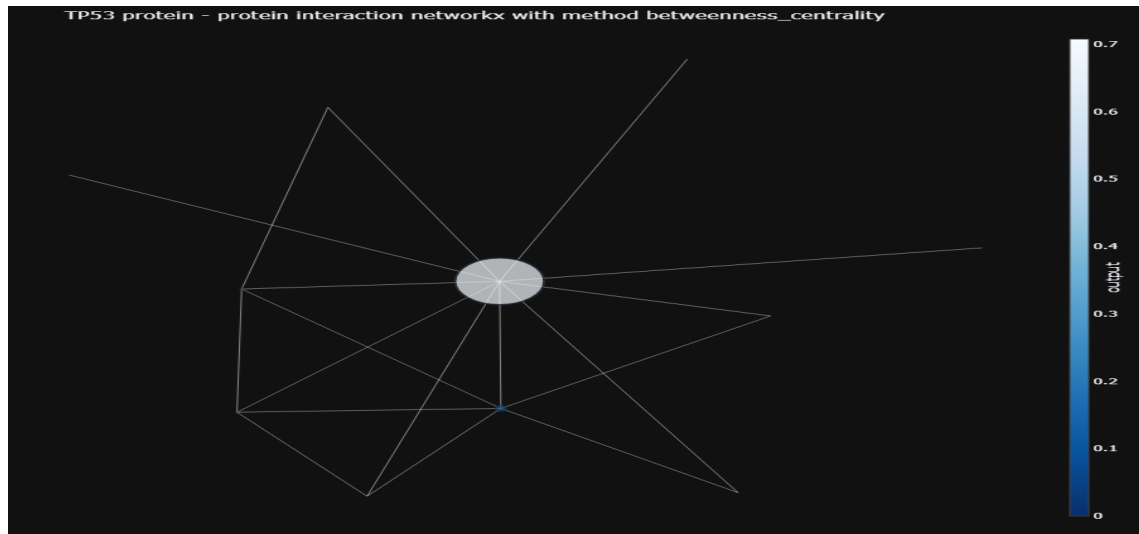


Figure 13: PPI subnetwork plot for protein TP53 colored by betweenness centrality values.

### 3.3.2 Method: Closeness centrality



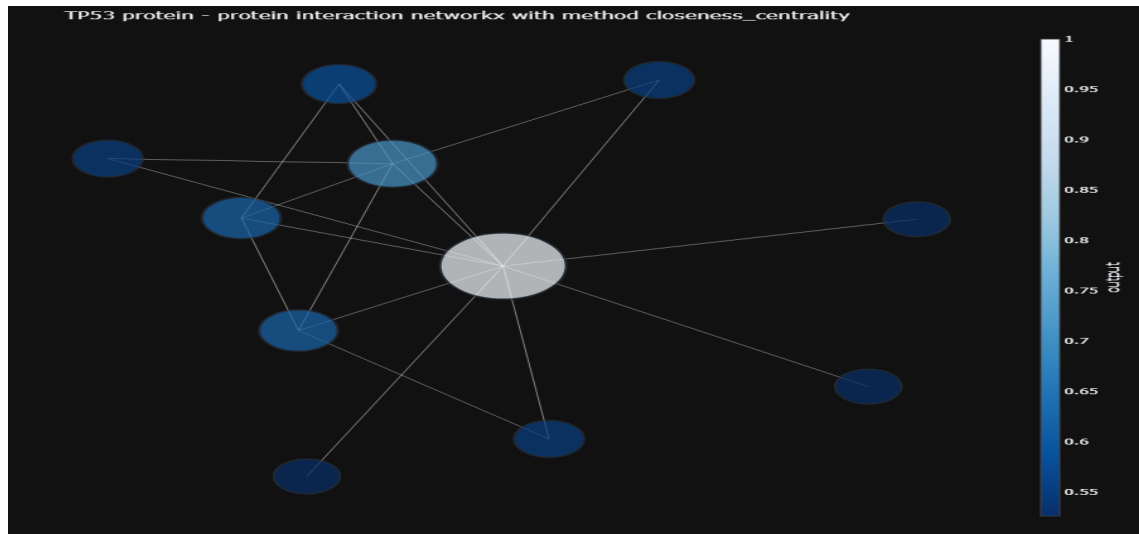


Figure 14: PPI plot for gene TP53 colored by closeness centrality values.

### 3.3.3 Method: Degree centrality

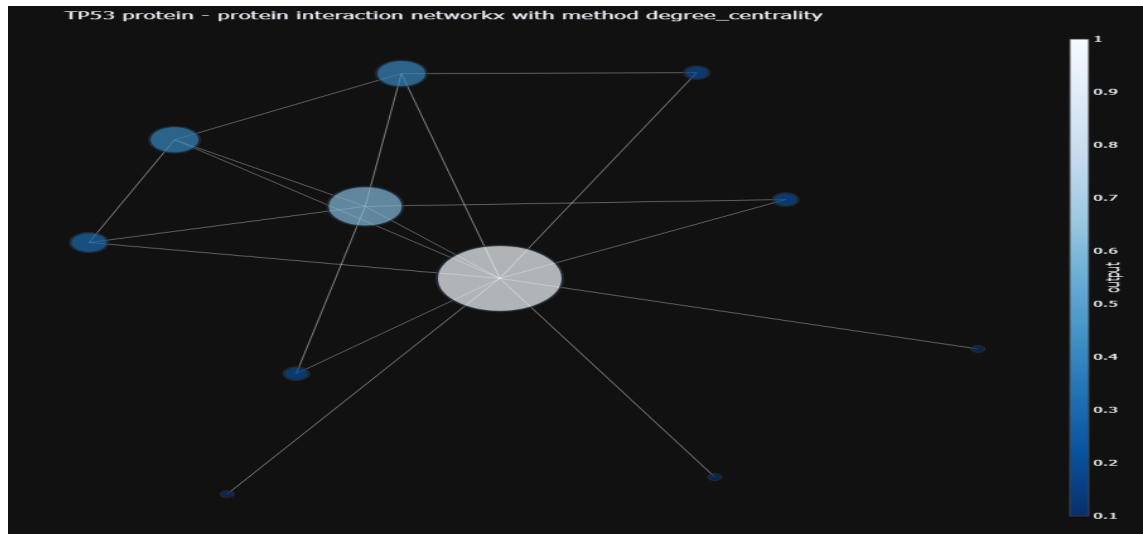


Figure 15: PPI plot for gene TP53 colored by degree centrality values.

#### 3.3.4 Method: Eigenvector centrality

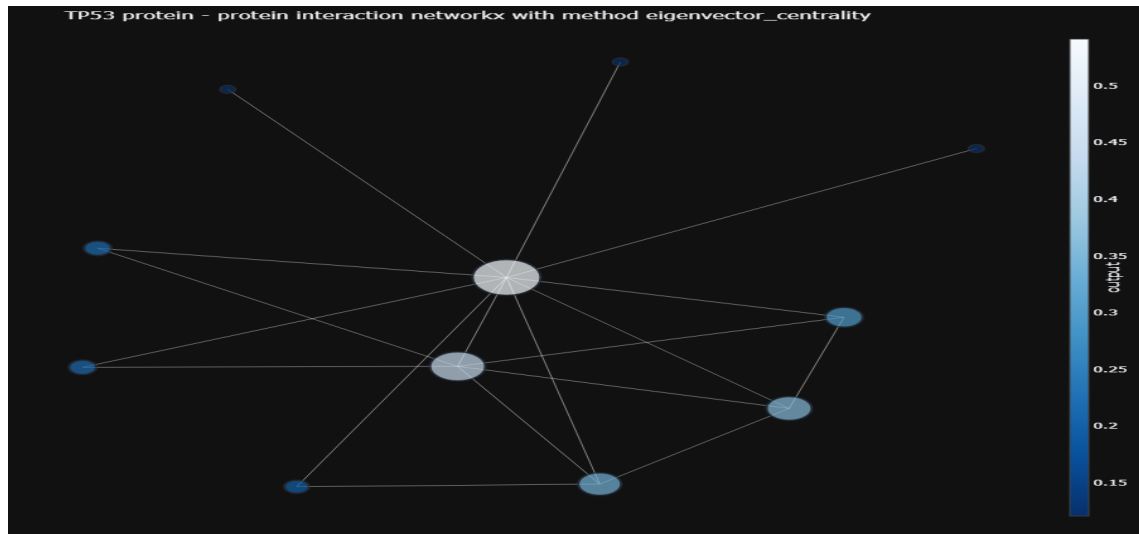


Figure 16: PPI plot for gene TP53 colored by eigenvector centrality values.

### 3.3.5 Method: Community extraction with Louvain



Figure 17: PPI plot of gene TP53 colored by communities extracted with the Louvain algorithm.

### 3.3.6 Method: Community extraction with Leiden



Figure 18: PPI plot of gene TP53 colored by communities extracted with the Louvain algorithm.

### 3.3.7 Method: Clusters extraction with Spectral Clustering

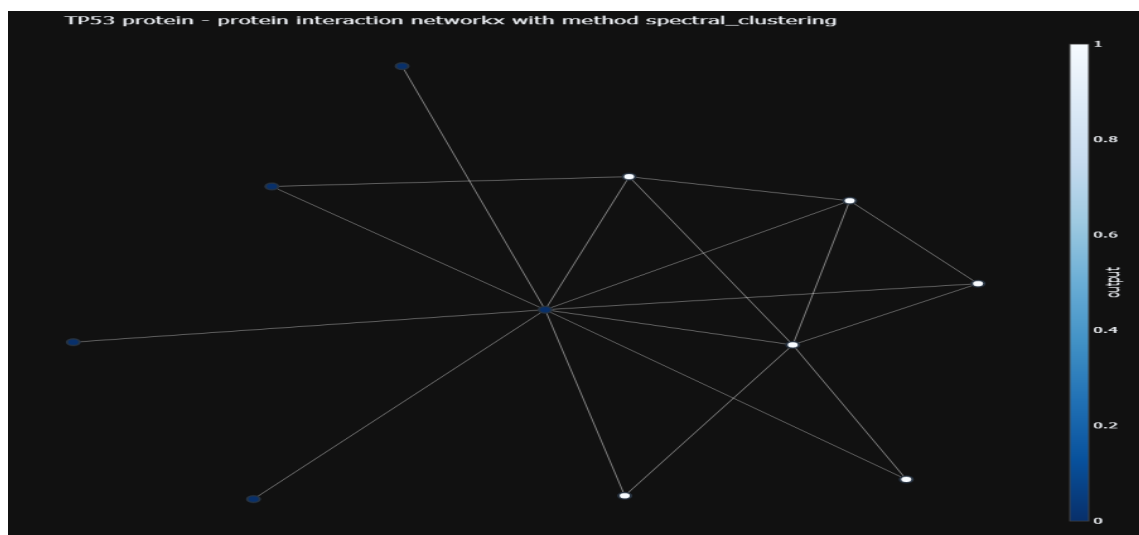


Figure 19: PPI plot of gene TP53 colored by clusters extracted with the Spectral Clustering algorithm.