

Purpose:

Biomedical literature is a rich but also complex source of information. To help researchers summarize literature on a topic in the form of a network, we developed a literature network application for cellular landscape that draws inspiration from two existing applications, agilentliteraturesearch and StringApp. The user query provides a summary of the topics described and selects the organisms for which the network should be generated. The application then queries the weekly database of precomputed text mining results through API to retrieve the list of genes mentioned in the summary, as well as the sentences in the excerpt that refer to any two genes. According to this information, literature network should program to construct a Cytoscape network, in which nodes and edges represent genes and their co-display, respectively. The application allows users to view sentences that support any edge, and all retrieved information is stored as a source in the Cytoscape session file.

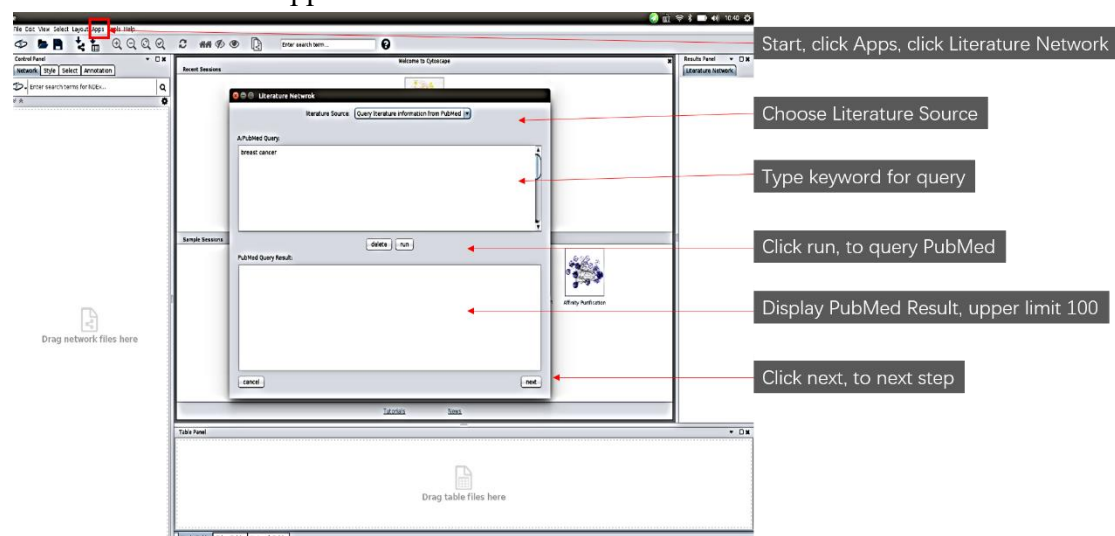
Operation:

Step 1: the document list

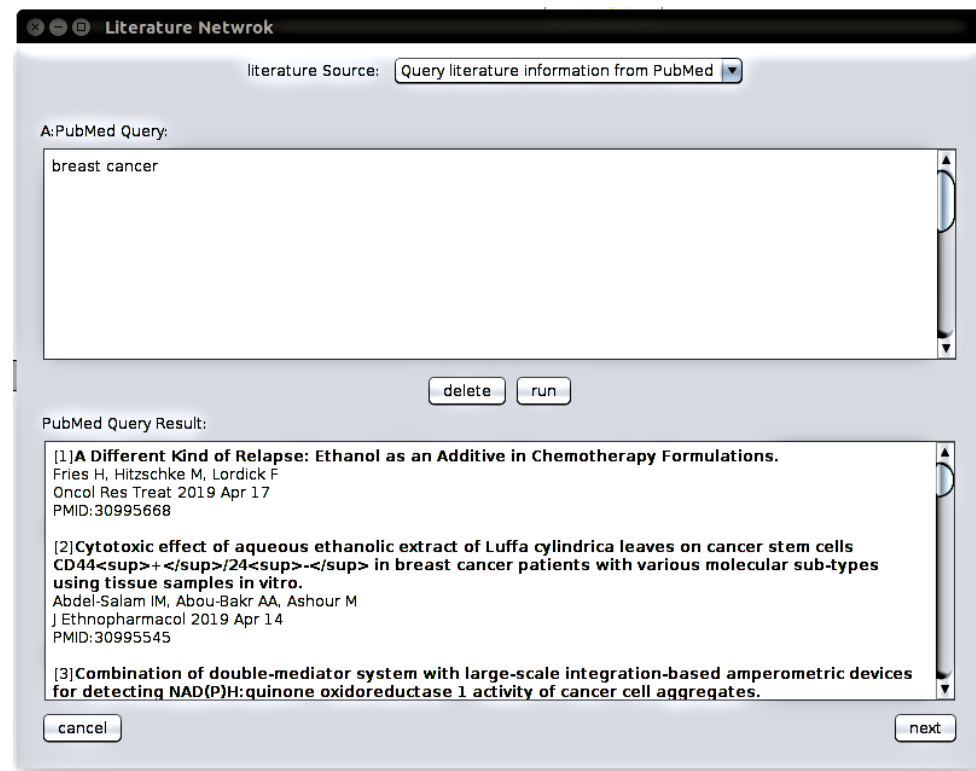
in the form of PMID, for example:30995668

two ways of selecting: 1. retrieve related documents in PubMed with subject words; 2, the user uploads the PMID list file or copies the PMID list content to input box

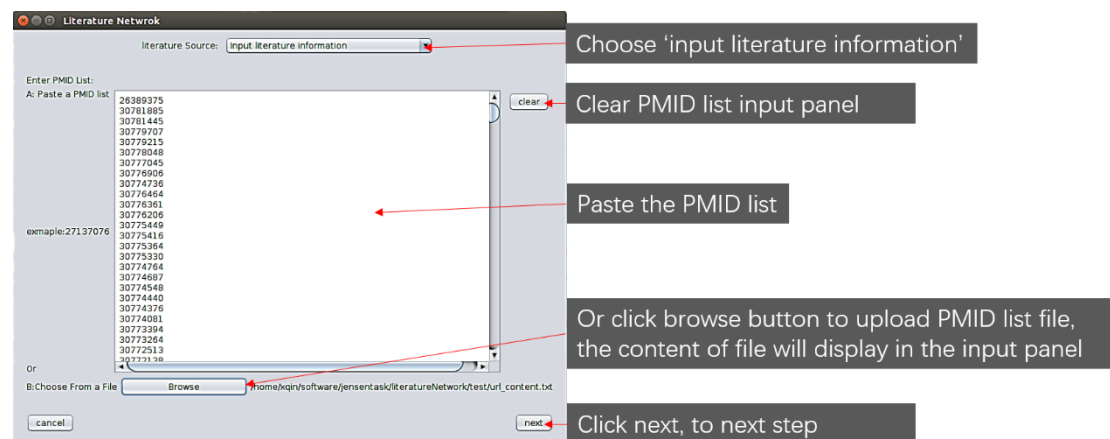
1. to retrieve the relevant literature PMID with a theme word, such as: breast cancer
input breast cancer, click the run button and the app will return the literature-related content with an upper limit of 100 in the result box



After click run button, the result interface:



2. User uploads the PMID list file himself or copies the PMID list content to the input box



Click next button

Step 2: the protein list

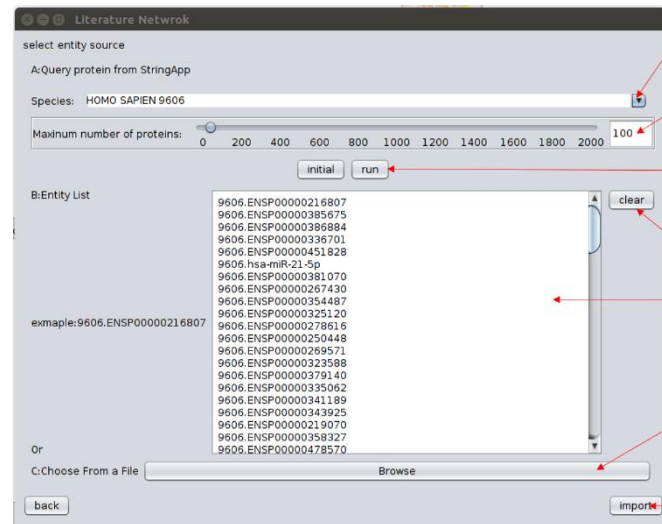
In form of speciesID.proteinID, for example: 9606.ENSEP00000216807

Select species in the species box

two ways of selecting: 1. retrieve related protein in String; 2, the user uploads the protein list file or copies the protein list content to input box

1, retrieve in string, set the maximum number of returned proteins n, The protein list of the top n in PMID was retrieved from string database. Example: the number of proteins is 100.

2, users themselves upload the speciesID.proteinID list file or copy the list content to the input box

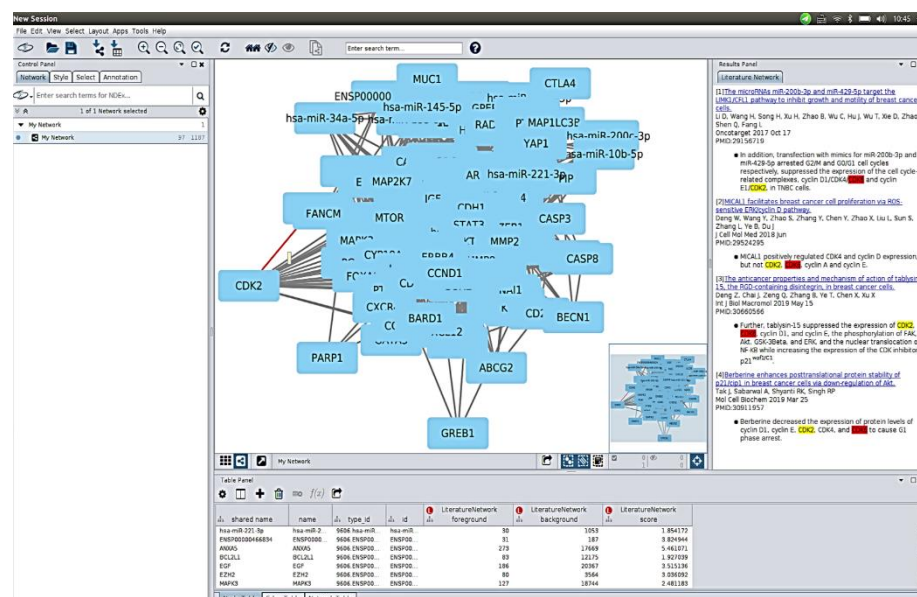


The screenshot shows the 'Literature Network' application window. It includes a 'select entity source' section with a 'Species' dropdown set to 'HOMO SAPIEN 9606' and a 'Maximum number of proteins' slider set to 100. Below this is a 'B:Entity List' containing a list of protein IDs (e.g., 9606. ENSP00000216807). At the bottom, there is a 'Browse' button and an 'import' button. Red arrows point from text boxes to these specific UI elements.

- Choose species of the protein
- Select the max protein number of String Retrieve
- Click run, start query, the result will display in the input panel
- Clear the content of input panel
- Paste protein list in form of speciesID.proteinID
- Or click browse button to upload protein list file, the content of file will display in the input panel, protein in form of speciesID.proteinID
- Click import button, build literature network

Step3 click import button to build network

Step4 click the edge in the network



The screenshot shows the Cytoscape software interface. A network of nodes (proteins) is displayed, with edges representing interactions. Nodes are labeled with names like MUC1, CTLA4, and various miRNAs. A 'Results Panel' on the right shows a list of sentences supporting the network. A text box with an arrow points to an edge in the network.

Click on any of the edges to show the sentence information supporting this edge

shared name	name	type_id	id	literature network foreground	literature network background	literature network score
hsa-miR-21-3p	hsa-miR-21-3p	miRNA	9606	10	1053	1.854177
ENSP00000216807	ENSP00000216807	ENSP00000216807	9606	10	1053	1.854177
AR	AR	AR	9606	273	17689	5.461071
EGF	EGF	EGF	9606	186	20367	3.531136
MAPK3	MAPK3	MAPK3	9606	127	18744	2.481183

Display sentences that support any edge, and all retrieved information is stored as a source in the Cytoscape session file