Community Detection App for Cytoscape

Cytoscape is an open source software platform for visualizing complex networks and integrating these with any type of attribute data. The Community Detection App is a Java plugin for Cytoscape which provides the user interface to execute several community detection algorithms and gene-set enrichment tools. All these algorithms and tools run on a remote server which exposes REST API. The app uses these API to execute the algorithms and tools.

1 Installation

1.1 Cytoscape

- If you don't have Cytoscape installed on your system, please follow the instructions at https://cytoscape.org/download.html.
- More information on Cytoscape can be found at https://cytoscape.org.

1.2 Community Detection App

1. Download the JAR from https://cytoscape-builds.ucsd.edu/cytoscape-builds/community_detection/.
2. Click on "Apps" from the menu options at the top of the screen → click on "App Manager" → click on "Install from File..." under the "Install Apps" tab → select the JAR → click on "Open" → once the app is installed, close the dialog box.

2 User Options

The community detection app adds menu options in the menu bar and the context menu. A menu bar is the list of options on the top-left panel of Cytoscape window (Figure 1). A context menu gets populated when a user right-clicks on the network view panel (Figure 8). In general, the menu bar options of this app apply to the whole network and those in the context menu apply to the selected node(s).
The app currently provides three major functionalities, explained below.

2.1 Community Detection algorithms

There are several algorithms which can be used for community detection of a network. However, most of these algorithms are written in languages other than Java and hence cannot be used via Cytoscape. The CD service exposes REST API to execute community detection algorithms and the application acts as a simplified interface between a user and CD service.
A sample workflow to execute an algorithm can be found in the images below. A user can select one of the algorithms from the menu (Figure 1) and an edge attribute from a drop-down list (Figure 2) which would be used as a weight for the network. The resultant hierarchy network can be seen in Figure 3.

An added feature for the community detection algorithms is the option to tune resolution parameters. Each of these algorithms has a parameter that determines the size of resulting clusters.
The "Settings..." option (Figure 4) opens a dialog box to set the appropriate resolution parameter. Since each algorithm implements the resolution parameter differently, a brief description for each algorithm's resolution parameter will appear when clicked on the parameter name or the "i" icon (Figure 5).

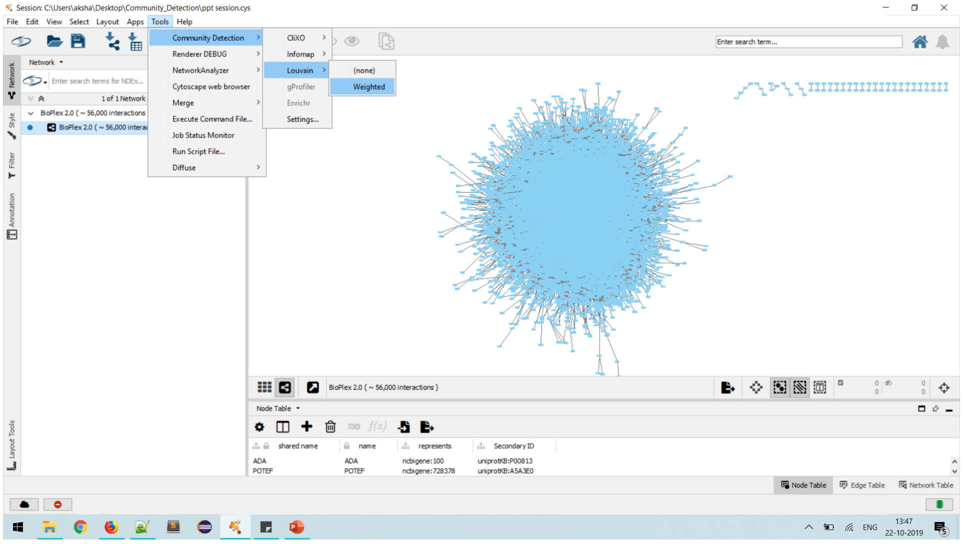
2.2 Gene-set Enrichment

One of the main reasons to apply a community detection algorithm on a protein-protein interaction network is to find out previously unknown interactions in known complexes. In order to find known complexes associated with the communities in a hierarchy network, the app currently supports two enrichment tools called GProfiler and Enrichr. There are two ways of running these tools:

- Menu bar - The menu bar option runs for the whole hierarchy network. Figures 6 and 7 show the menu bar option and GO enrichment result. Depending upon the size of the hierarchy network, GO enrichment could take several minutes or hours to complete.
- Context Menu - Another way to call an enrichment tool is through the context menu. However, this option works for selected nodes. Figure 8 shows the context menu option.

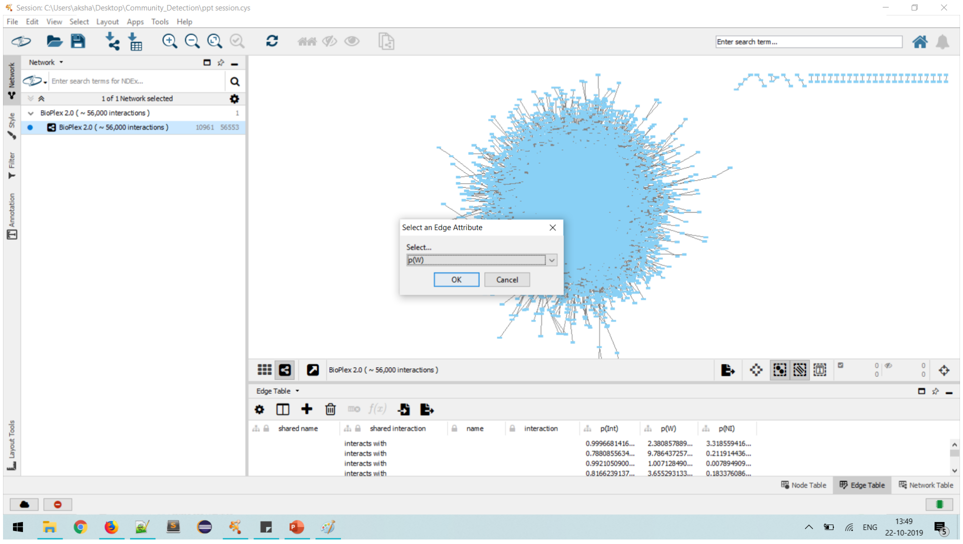
2.3 Interaction Sub-Networks

Another useful feature of this application is to view interactions among the genes of a community. This is achieved by creating a sub-network of the original interaction network. The sub-network contains genes of the selected community node and the edges connecting those genes. Figures 9 and 10 provide a snapshot of the context menu option and the resulting sub-network respectively. The context menu option, however, works only if 1 node is selected. An important point to note about the sub-network creation task is that it depends on network table (CD_OriginalNetwork) and node table (CD_MemberList) entries. To create a correct sub-network both these entries should exist and entry under CD_OriginalNetwork should point to a valid network.



shared name	name	represents	Secondary ID
ADA	ADA	ncbigene:100	uniprotKB:P00813
POTEF	POTEF	ncbigene:728378	uniprotKB:A5A3E0

Figure 1: An example to choose a community detection algorithm.

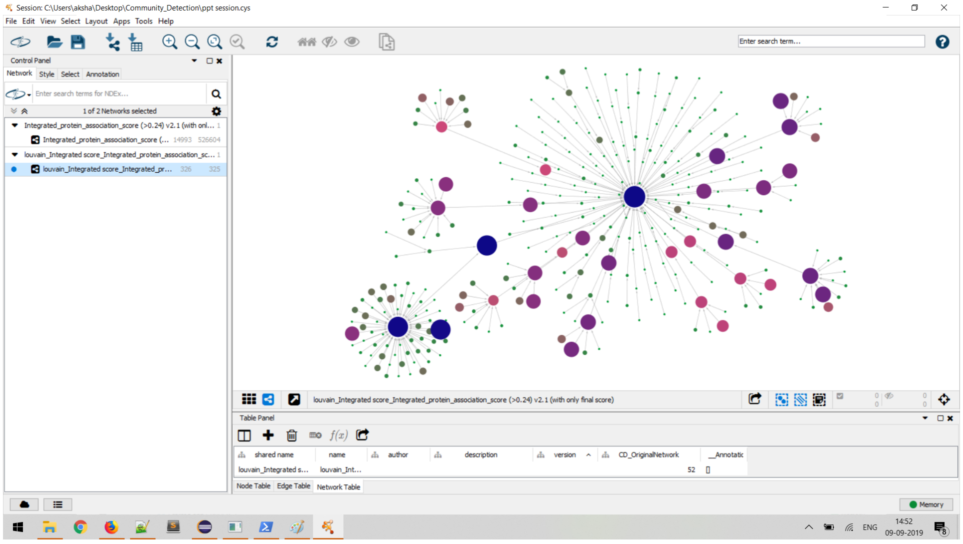

Session: C:\Users\aksha\Desktop\Community_Detection\ppt session.cys
File
Edit
View
Select
Layout
Apps
Tools
Help
Network
Enter search terms for NDEx...
1 of 1 Network selected
BioPlex 2.0 (~ 56,000 interactions)
1
BioPlex 2.0 (~ 56,000 interactions)
10961
56553
Style
Filter
Annotation
Layout Tools
Enter search term...
OK
Cancel
BioPlex 2.0 (~ 56,000 interactions)
Edge Table
Node Table
Edge Table
Network Table
ENG
13:49
22-10-2019

Select an Edge Attribute

Select... p(W)

shared name	shared interaction	name	interaction	p(Int)	p(W)	p(NI)
	interacts with			0.9996681416...	2.380857889...	3.318559416...
	interacts with			0.7880855634...	9.786437257...	0.211914436...
	interacts with			0.9921050900...	1.007128490...	0.007894909...
	interacts with			0.8166239137...	3.655293133...	0.183376086...

Figure 2: Selecting an edge attribute as a weight.



Session: C:\Users\aksha\Desktop\Community_Detection\ppt session.cys

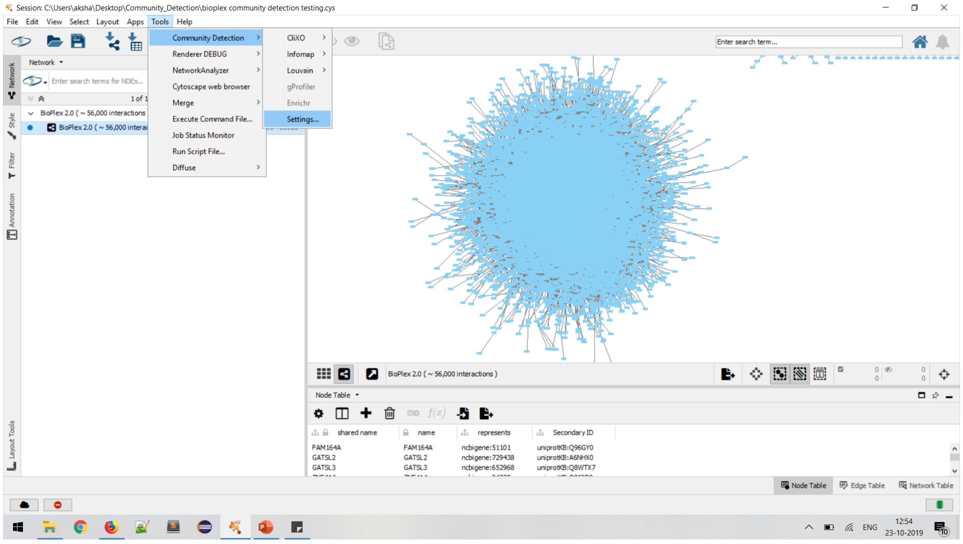
1 of 2 Networks selected

- Integrated_protein_association_score (>0.24) v2.1 (with onl... 1
  - Integrated_protein_association_score (... 14993 526604
- louvain_Integrated score_Integrated_protein_association_sc... 1
  - louvain_Integrated score_Integrated_pr... 326 325

louvain_Integrated score_Integrated_protein_association_score (>0.24) v2.1 (with only final score)

shared name	name	author	description	version	CD_OriginalNetwork	__Annotatic
louvain_Integrated s...	louvain_Int...				52	[]

Figure 3: Resultant hierarchy network.



Session: C:\Users\aksha\Desktop\Community_Detection\bioplex community detection testing.cys

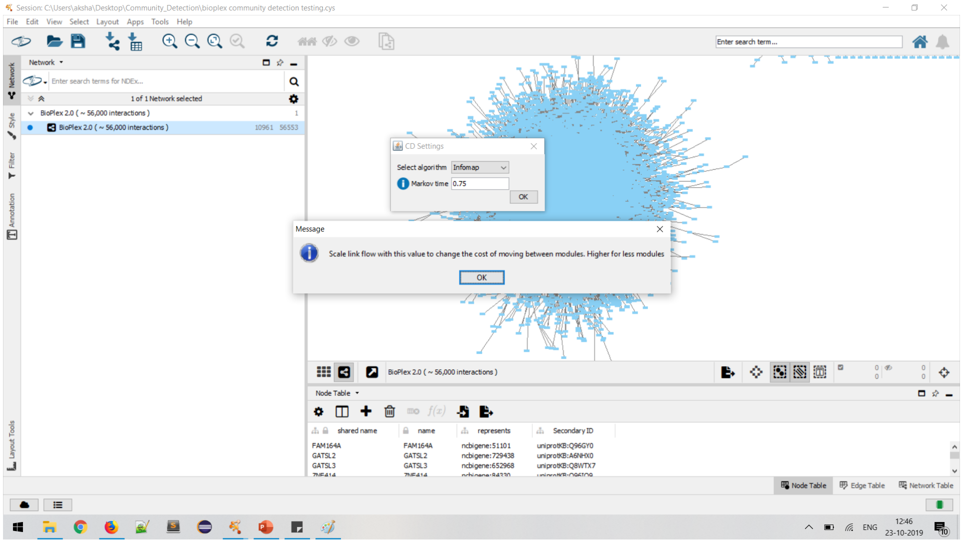
- BioPlex 2.0 (~ 56,000 interactions)
  - BioPlex 2.0 (~ 56,000 interac

BioPlex 2.0 (~ 56,000 interactions)

Node Table

shared name	name	represents	Secondary ID
FAM164A	FAM164A	ncbigene:51101	uniprotKB:Q96GY0
GATSL2	GATSL2	ncbigene:729438	uniprotKB:A6NHX0
GATSL3	GATSL3	ncbigene:652968	uniprotKB:Q8WTX7

Figure 4: Menu option to select ”Settings...”



Session: C:\Users\aksha\Desktop\Community_Detection\bioplex community detection testing.cys
File Edit View Select Layout Apps Tools Help
Enter search term...
Network
Enter search terms for NDEx...
1 of 1 Network selected
BioPlex 2.0 (~ 56,000 interactions) 1
BioPlex 2.0 (~ 56,000 interactions) 10961 56553
Network
Style
Filter
Annotation
Layout Tools
BioPlex 2.0 (~ 56,000 interactions)
Node Table
Node Table
Edge Table
Network Table
ENG
12:46
23-10-2019

CD Settings

Select algorithm: Infomap

Markov time: 0.75

OK

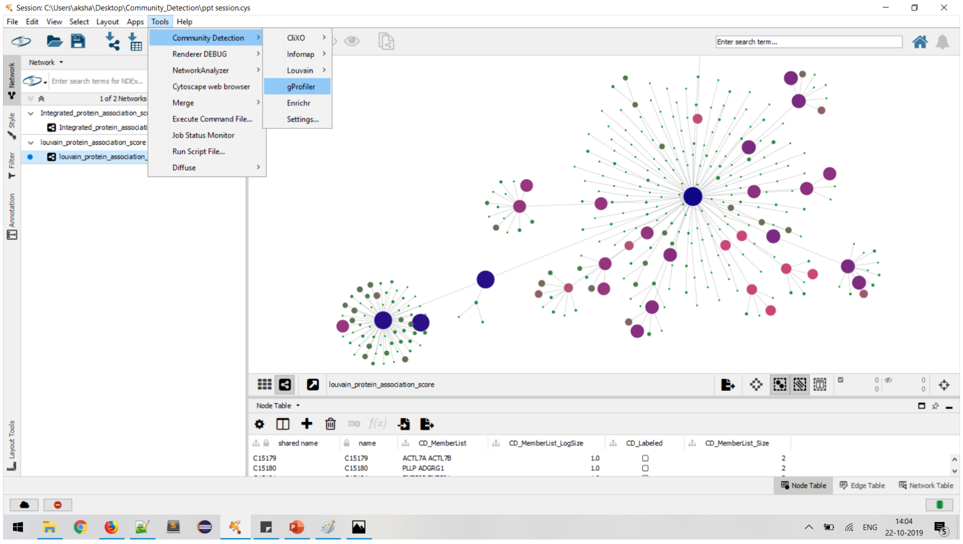
Message

Scale link flow with this value to change the cost of moving between modules. Higher for less modules

OK

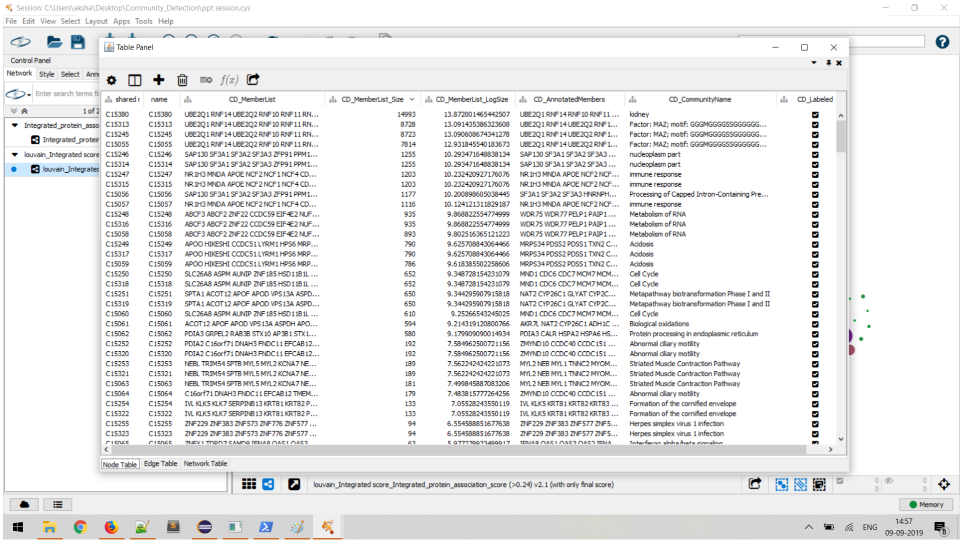
shared name	name	represents	Secondary ID
FAM164A	FAM164A	ncbigene:51101	uniprotKB:Q96GY0
GATSL2	GATSL2	ncbigene:729438	uniprotKB:A6NHX0
GATSL3	GATSL3	ncbigene:652968	uniprotKB:Q8WTX7

Figure 5: Dialog box with a brief description of the resolution parameter



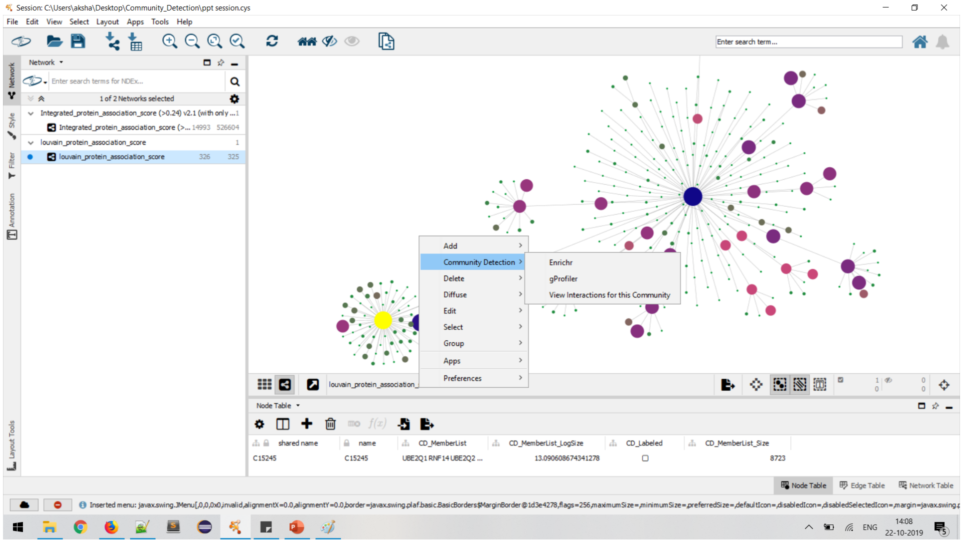
shared name	name	CD_MemberList	CD_MemberList_LogSize	CD_Labeled	CD_MemberList_Size
C15179	C15179	ACTL7A ACTL7B	1.0	□	2
C15180	C15180	PLLP ADGRG1	1.0	□	2

Figure 6: The menu bar option to run GProfiler.


Session: C:\Users\aksha\Desktop\Community_Detection\ppt session.cys
File Edit View Select Layout Apps Tools Help
Table Panel
Control Panel
Network Style Select Anno
Node Table
Edge Table
Network Table
louvain_Integrated score_Integrated_protein_association_score (>0.24) v2.1 (with only final score)
Memory
ENG
14:57
09-09-2019

shared name	name	CD_MemberList	CD_MemberList_Size	CD_MemberList_LogSize	CD_AnnotatedMembers	CD_CommunityName	CD_Labeled
C15380	C15380	UBE2Q1 RNF14 UBE2Q2 RNF10 RNF11 RN...	14993	13.872001465442507	UBE2Q1 RNF14 RNF10 RNF11 ...	kidney	☑
C15313	C15313	UBE2Q1 RNF14 UBE2Q2 RNF10 RNF11 RN...	8728	13.091435386323608	UBE2Q1 RNF14 UBE2Q2 RNF1...	Factor: MAZ; motif: GGGMGGGGSSGGGGG...	☑
C15245	C15245	UBE2Q1 RNF14 UBE2Q2 RNF10 RNF11 RN...	8723	13.090608674341278	UBE2Q1 RNF14 UBE2Q2 RNF1...	Factor: MAZ; motif: GGGMGGGGSSGGGGG...	☑
C15055	C15055	UBE2Q1 RNF14 UBE2Q2 RNF10 RNF11 RN...	7814	12.931845540183673	UBE2Q1 RNF14 UBE2Q2 RNF1...	Factor: MAZ; motif: GGGMGGGGSSGGGGG...	☑
C15246	C15246	SAP130 SF3A1 SF3A2 SF3A3 ZFP91 PPM1...	1255	10.293471648838134	SAP130 SF3A1 SF3A2 SF3A3 ...	nucleoplasm part	☑
C15314	C15314	SAP130 SF3A1 SF3A2 SF3A3 ZFP91 PPM1...	1255	10.293471648838134	SAP130 SF3A1 SF3A2 SF3A3 ...	nucleoplasm part	☑
C15247	C15247	NR1H3 MNDA APOE NCF2 NCF1 NCF4 CD...	1203	10.232420927176076	NR1H3 MNDA APOE NCF2 NCF...	immune response	☑
C15315	C15315	NR1H3 MNDA APOE NCF2 NCF1 NCF4 CD...	1203	10.232420927176076	NR1H3 MNDA APOE NCF2 NCF...	immune response	☑
C15056	C15056	SAP130 SF3A1 SF3A2 SF3A3 ZFP91 PPM1...	1177	10.200898605038445	SF3A1 SF3A2 SF3A3 HNRNPH...	Processing of Capped Intron-Containing Pre...	☑
C15057	C15057	NR1H3 MNDA APOE NCF2 NCF1 NCF4 CD...	1116	10.124121311829187	NR1H3 MNDA APOE NCF2 NCF...	immune response	☑
C15248	C15248	ABCF3 ABCF2 ZNF22 CCDC59 EIF4E2 NUF...	935	9.868822554774999	WDR75 WDR77 PELP1 PAIP1 ...	Metabolism of RNA	☑
C15316	C15316	ABCF3 ABCF2 ZNF22 CCDC59 EIF4E2 NUF...	935	9.868822554774999	WDR75 WDR77 PELP1 PAIP1 ...	Metabolism of RNA	☑
C15058	C15058	ABCF3 ABCF2 ZNF22 CCDC59 EIF4E2 NUF...	893	9.802516365121223	WDR75 WDR77 PELP1 PAIP1 ...	Metabolism of RNA	☑
C15249	C15249	APOO HIKESHI CCDC51 LYRM1 HPS6 MRP...	790	9.625708843064466	MRPS34 PDSS2 PDSS1 TXN2 C...	Acidosis	☑
C15317	C15317	APOO HIKESHI CCDC51 LYRM1 HPS6 MRP...	790	9.625708843064466	MRPS34 PDSS2 PDSS1 TXN2 C...	Acidosis	☑
C15059	C15059	APOO HIKESHI CCDC51 LYRM1 HPS6 MRP...	786	9.618385502258606	MRPS34 PDSS2 PDSS1 TXN2 C...	Acidosis	☑
C15250	C15250	SLC26A8 ASPM AUNIP ZNF185 HSD11B1L ...	652	9.348728154231079	MND1 CDC6 CDC7 MCM7 MCM...	Cell Cycle	☑
C15318	C15318	SLC26A8 ASPM AUNIP ZNF185 HSD11B1L ...	652	9.348728154231079	MND1 CDC6 CDC7 MCM7 MCM...	Cell Cycle	☑
C15251	C15251	SPTA1 ACOT12 APOF APOD VPS13A ASPD...	650	9.344295907915818	NAT2 CYP26C1 GLYAT CYP2C...	Metapathway biotransformation Phase I and II	☑
C15319	C15319	SPTA1 ACOT12 APOF APOD VPS13A ASPD...	650	9.344295907915818	NAT2 CYP26C1 GLYAT CYP2C...	Metapathway biotransformation Phase I and II	☑
C15060	C15060	SLC26A8 ASPM AUNIP ZNF185 HSD11B1L ...	610	9.25266543245025	MND1 CDC6 CDC7 MCM7 MCM...	Cell Cycle	☑
C15061	C15061	ACOT12 APOF APOD VPS13A ASPDH APO...	594	9.214319120800766	AKR7L NAT2 CYP26C1 ADH1C ...	Biological oxidations	☑
C15062	C15062	PDIA3 GRPEL2 RAB3B STX10 AP3B1 STX1...	580	9.179909090014934	PDIA3 CALR HSPA2 HSPA6 HS...	Protein processing in endoplasmic reticulum	☑
C15252	C15252	PDIA2 C16orf71 DNAH3 FNDC11 EFCAB12...	192	7.584962500721156	ZMYND10 CCDC40 CCDC151 ...	Abnormal ciliary motility	☑
C15320	C15320	PDIA2 C16orf71 DNAH3 FNDC11 EFCAB12...	192	7.584962500721156	ZMYND10 CCDC40 CCDC151 ...	Abnormal ciliary motility	☑
C15253	C15253	NEBL TRIM54 SPTB MYL5 MYL2 KCNA7 NE...	189	7.562242424221073	MYL2 NEB MYL1 TNNC2 MYOM...	Striated Muscle Contraction Pathway	☑
C15321	C15321	NEBL TRIM54 SPTB MYL5 MYL2 KCNA7 NE...	189	7.562242424221073	MYL2 NEB MYL1 TNNC2 MYOM...	Striated Muscle Contraction Pathway	☑
C15063	C15063	NEBL TRIM54 SPTB MYL5 MYL2 KCNA7 NE...	181	7.499845887083206	MYL2 NEB MYL1 TNNC2 MYOM...	Striated Muscle Contraction Pathway	☑
C15064	C15064	C16orf71 DNAH3 FNDC11 EFCAB12 TMEM...	179	7.483815777264256	ZMYND10 CCDC40 CCDC151 ...	Abnormal ciliary motility	☑
C15254	C15254	IVL KLK5 KLK7 SERPINB13 KRT81 KRT82 P...	133	7.05528243550119	IVL KLK5 KRT81 KRT82 KRT83 ...	Formation of the cornified envelope	☑
C15322	C15322	IVL KLK5 KLK7 SERPINB13 KRT81 KRT82 P...	133	7.05528243550119	IVL KLK5 KRT81 KRT82 KRT83 ...	Formation of the cornified envelope	☑
C15255	C15255	ZNF229 ZNF383 ZNF573 ZNF776 ZNF577 ...	94	6.554588851677638	ZNF229 ZNF383 ZNF577 ZNF5...	Herpes simplex virus 1 infection	☑
C15323	C15323	ZNF229 ZNF383 ZNF573 ZNF776 ZNF577 ...	94	6.554588851677638	ZNF229 ZNF383 ZNF577 ZNF5...	Herpes simplex virus 1 infection	☑
C15065	C15065	ZNFX1 TDRD7 SAMD9 IFNA8 OAS1 OAS3	63	5.977279923499917	IFNA8 OAS1 OAS3 OAS2 IFNA	Interferon alpha/beta signaling	☑

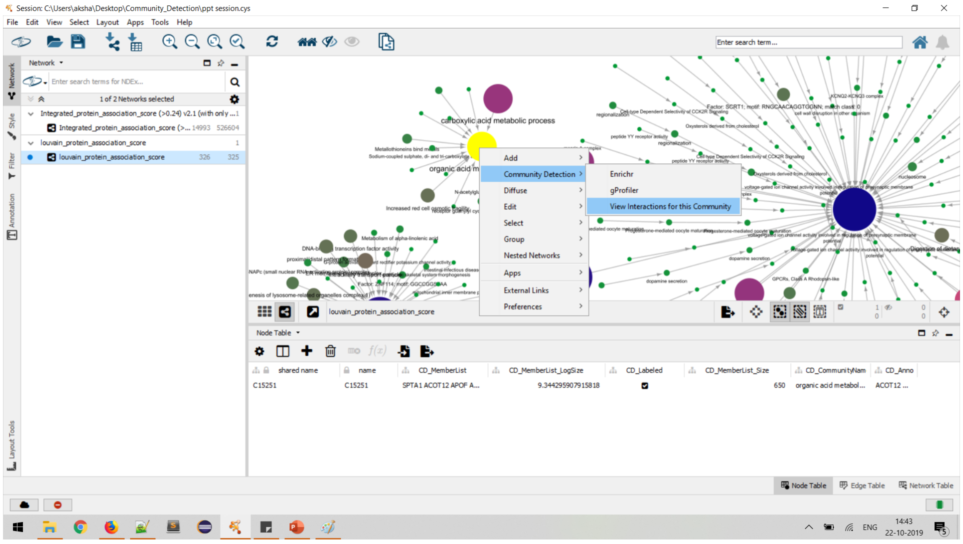
Figure 7: GO enrichment result for the whole network.


Session: C:\Users\aksha\Desktop\Community_Detection\ppt session.cys
File Edit View Select Layout Apps Tools Help
Enter search term...
Network
Enter search terms for NDEx...
1 of 2 Networks selected
Integrated_protein_association_score (>0.24) v2.1 (with only ...
Integrated_protein_association_score (>... 14993 526604
louvain_protein_association_score
louvain_protein_association_score 326 325
Network
Style
Filter
Annotation
Layout Tools
Add
Community Detection
Delete
Diffuse
Edit
Select
Group
Apps
Preferences
Enrichr
gProfiler
View Interactions for this Community
louvain_protein_association_
Node Table
Edge Table
Network Table
ENG
14:08
22-10-2019

shared name	name	CD_MemberList	CD_MemberList_LogSize	CD_Labeled	CD_MemberList_Size
C15245	C15245	UBE2Q1 RNF14 UBE2Q2 ...	13.090608674341278	☐	8723

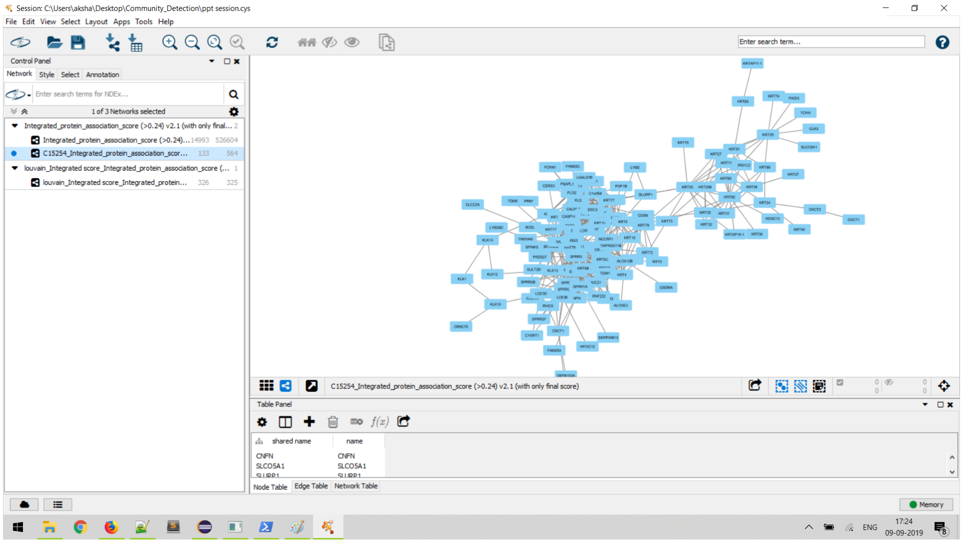
```
Inserted menu: javax.swing.JMenu[,0,0,0x0,invalid,alignmentX=0.0,alignmentY=0.0,border=javax.swing.plaf.basic.BasicBorders$MarginBorder@1d3e4278,flags=256,maximumSize=,minimumSize=,preferredSize=,defaultIcon=,disabledIcon=,disabledSelectedIcon=,margin=javax.swing.p
```

Figure 8: Context menu option for the selected nodes.


Session: C:\Users\aksha\Desktop\Community_Detection\ppt session.cys
File Edit View Select Layout Apps Tools Help
Enter search term...
Network
1 of 2 Networks selected
Integrated_protein_association_score (>0.24) v2.1 (with only ...
Integrated_protein_association_score (>... 14993 526604
louvain_protein_association_score
louvain_protein_association_score 326 325
carboxylic acid metabolic process
organic acid m
Add
Community Detection
Diffuse
Edit
Select
Group
Nested Networks
Apps
External Links
Preferences
Enrichr
gProfiler
View Interactions for this Community
louvain_protein_association_score
Node Table
Node Table
Edge Table
Network Table

shared name	name	CD_MemberList	CD_MemberList_LogSize	CD_Labeled	CD_MemberList_Size	CD_CommunityNam	CD_Anno
C15251	C15251	SPTA1 ACOT12 APOF A...	9.344295907915818	☑	650	organic acid metabol...	ACOT12 ...

Figure 9: Context menu option to view interactions of the selected community node's genes.


Session: C:\Users\aksha\Desktop\Community_Detection\ppt session.cys
File Edit View Select Layout Apps Tools Help
Enter search term...
Control Panel
Network
Style
Select
Annotation
Enter search terms for NDEx...
1 of 3 Networks selected
Integrated_protein_association_score (>0.24) v2.1 (with only final... 2
Integrated_protein_association_score (>0.24)... 14993 526604
C15254_Integrated_protein_association_scor... 133 564
louvain_Integrated score_Integrated_protein_association_score (... 1
louvain_Integrated score_Integrated_protein... 326 325
C15254_Integrated_protein_association_score (>0.24) v2.1 (with only final score)
Table Panel
Node Table
Edge Table
Network Table
Memory

shared name	name
CNFN	CNFN
SLCO5A1	SLCO5A1
SLURP1	SLURP1

Figure 10: Resultant interaction network as a sub-network of the original interaction network.