bioinformatics.ca

Canadian Bioinformatics Workshops

www.bioinformatics.ca

bioinformaticsdotca.github.io

Creative Commons
This page is available in the following languages:
Afrikaans български Català Dansk Deutsch Ελληνικά English English (CA) English (GB) English (US) Esperanto
Castellano Castellano (AR) Español (CL) Castellano (CO) Español (Ecuador) Castellano (MX) Castellano (PE)
Euskara Suomeksi français français (CA) Galego עברית hrvatski Magyar Italiano 日本語 한국어 Macedonian Melayu
Nederlands Norsk Sesotho sa Leboa polski Português română slovenski jezik српски srpski (latinica) Sotho svenska
中文 華語 (台灣) isiZulu
creative commons
Attribution-Share Alike 2.5 Canada
You are free:
to Share — to copy, distribute and transmit the work
to Remix — to adapt the work
Free Cultural
APPROVED FOR
Works
Under the following conditions:
Attribution. You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work).
Share Alike. If you alter, transform, or build upon this work, you may distribute the resulting work only under the same or similar licence to this one.
For any reuse or distribution, you must make clear to others the licence terms of this work.
Any of the above conditions can be waived if you get permission from the copyright holder.
The author's moral rights are retained in this licence.
Disclaimer
Your fair dealing and other rights are in no way affected by the above.
This is a human-readable summary of the Legal Code (the full licence) available in the following languages:
English French
Learn how to distribute your work using this licence

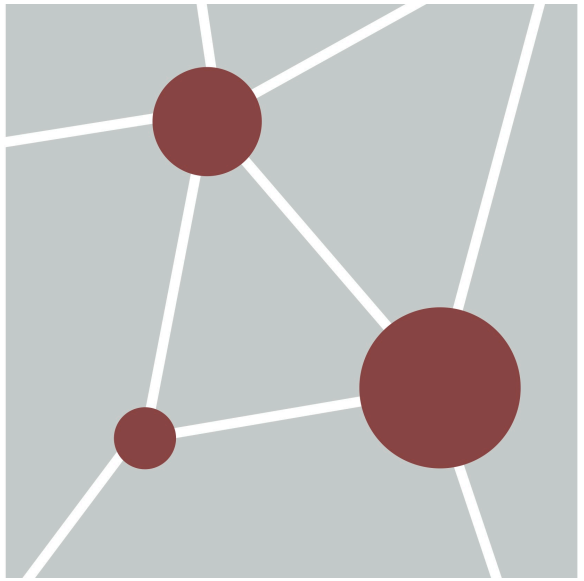
More Depth on Pathway & Network Analysis
Reactome FI practical lab

Veronique Voisin

Pathway and Network Analysis

June 5-7, 2023

bioinformatics.ca



reactome

OICR

Cytoscape

Learning Objectives of Module

- Be able to perform pathway and network-based data analysis using the ReactomeFIViz app.

Major Features in ReactomeFIViz

The ReactomeFIViz app is designed to find pathways and network patterns related to cancer and other types of diseases.

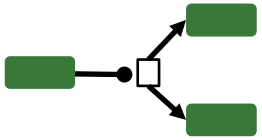
This app accesses the Reactome pathways stored in the database. It can:

- help you to do pathway enrichment analysis for a set of genes
- visualize hit pathways using manually laid-out pathway diagrams directly in Cytoscape
- investigate functional relationships among genes in hit pathways.

The app can also access the Reactome Functional Interaction (FI) network. It can:

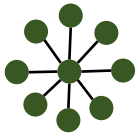
- construct a FI network based on a set of genes, query the FI data source for the underlying evidence for the interaction
- build and analyze network modules of highly-interacting groups of genes
- perform functional enrichment analysis to annotate the modules
- expand the network by finding genes related to the experimental data set, and overlay with a variety of information sources such as cancer gene index annotations or FDA-approved cancer drugs.

Reactome Functional Interaction Network



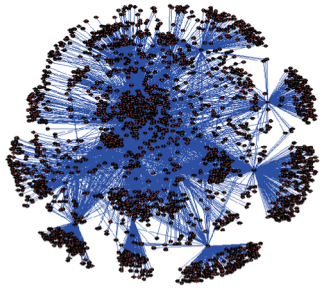
Curated Pathway DBs

+



Uncurated Interaction Evidence

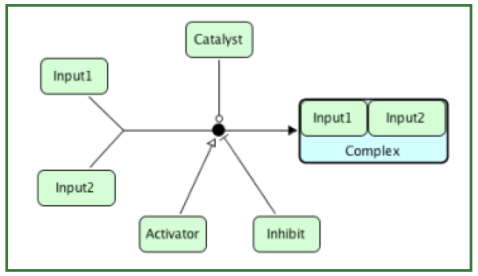
Machine Learning

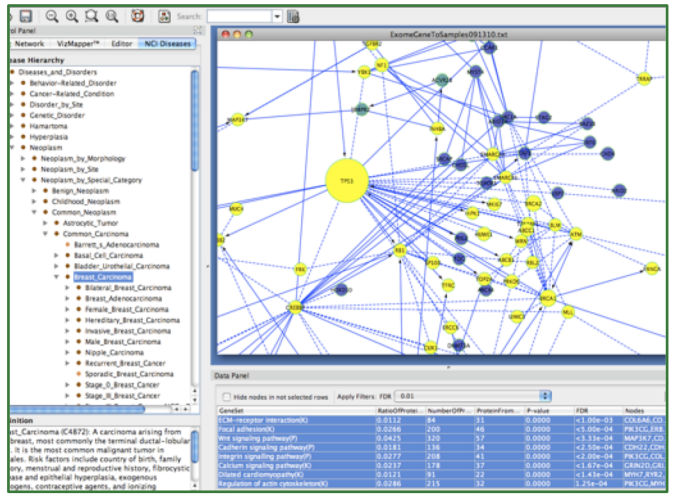


Reactome Functional Interaction Network
(~13K proteins; ~436K interactions)

+

Project your data into Reactome FI Network



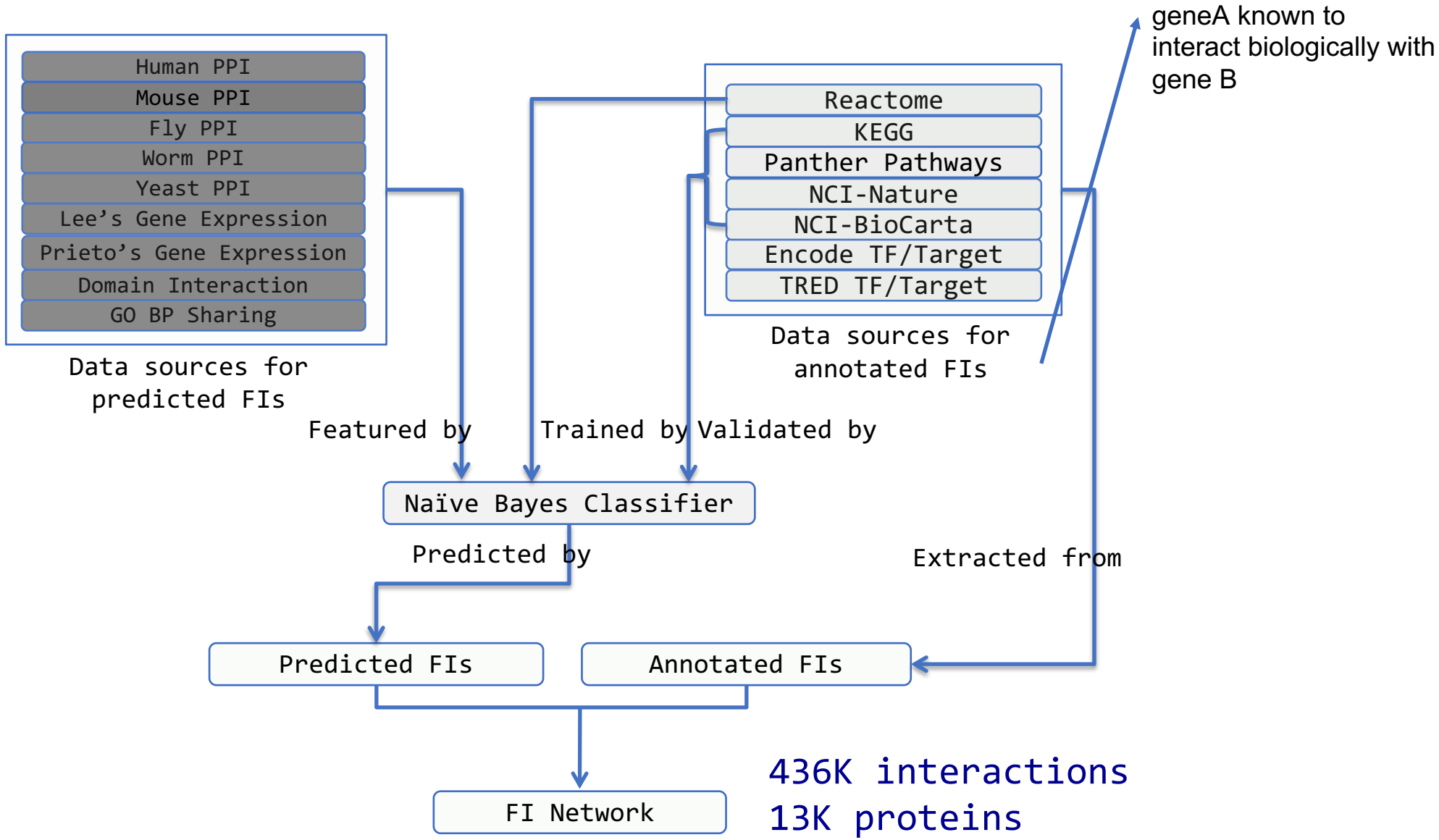


Cytoscape using ReactomeFIViz app

What is a Functional Interaction?

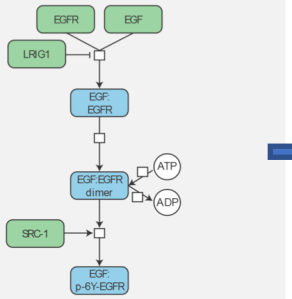
- Convert reactions in pathways into pair-wise relationships
  - **Functional Interaction**: an interaction in which two proteins are involved in the same reaction as input, catalyst, activator and/or inhibitor, or as components in a complex

Construction of the FI Network

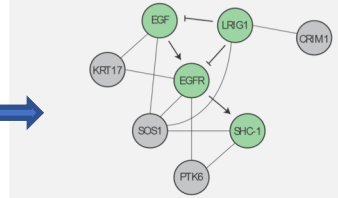


ReactomeFIVIz (ReactomeFIPlugIn): pathway AND network analysis

Pathway diagram



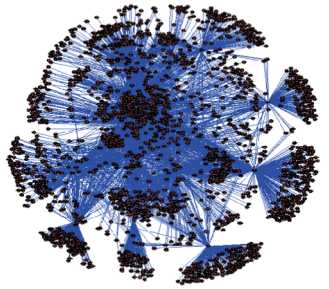
Network



- Browse the collections of the Reactome pathways
- Convert these pathways into a Cytoscape network

Your gene list

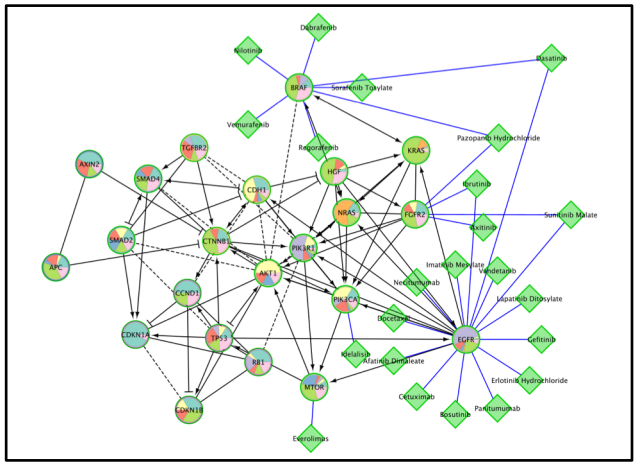
- Import your gene list.
- Reactome pathways enriched in your gene list will be returned.
- Genes that in your gene list will be highlighted in the pathway diagrams and network



Reactome Functional Interaction Network
(~13K proteins; ~436K interactions)

Your gene list

Create a FI network

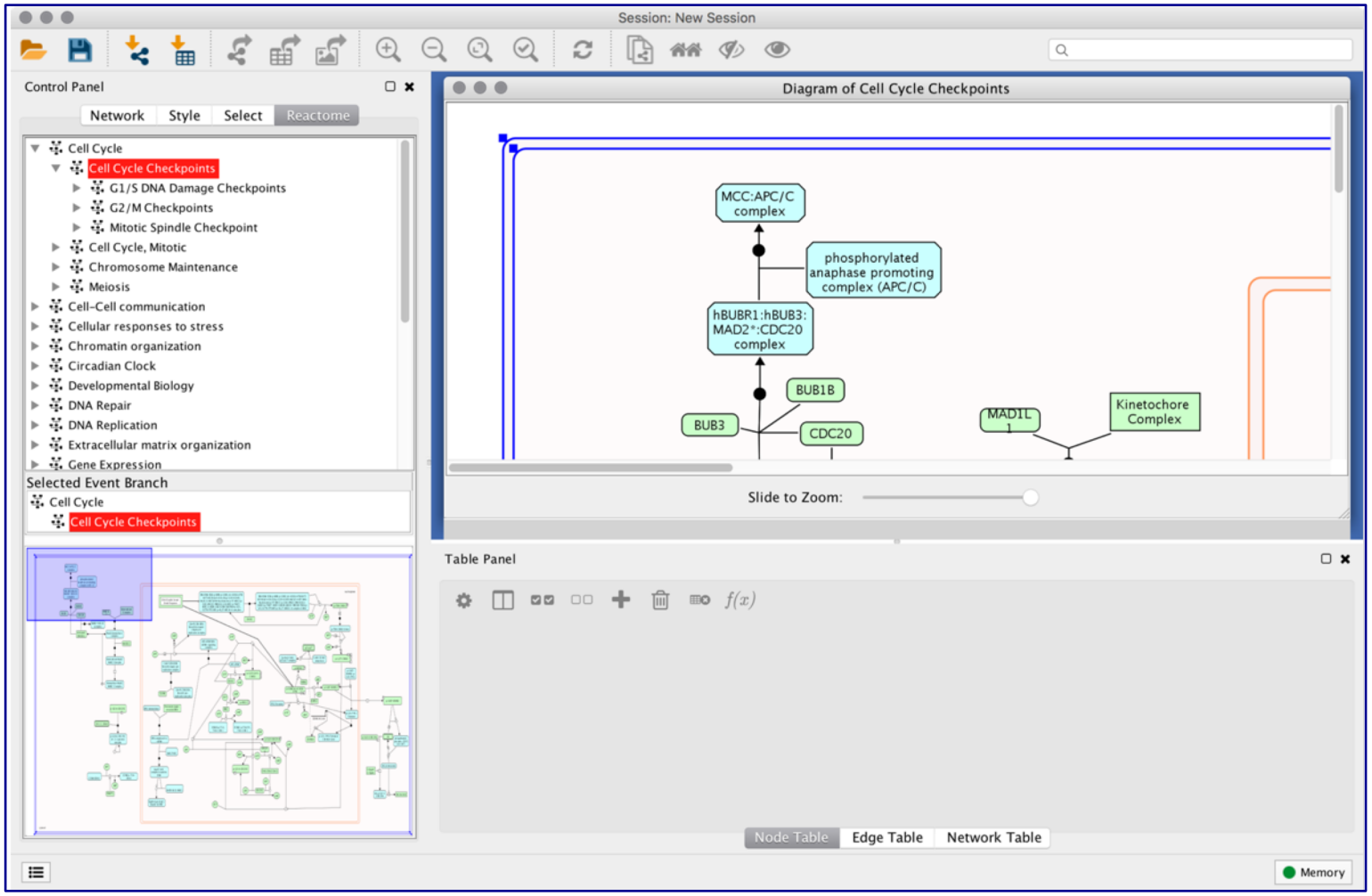


- Create a FI network using your gene list :
- Genes in your list will be connected by functional interactions.

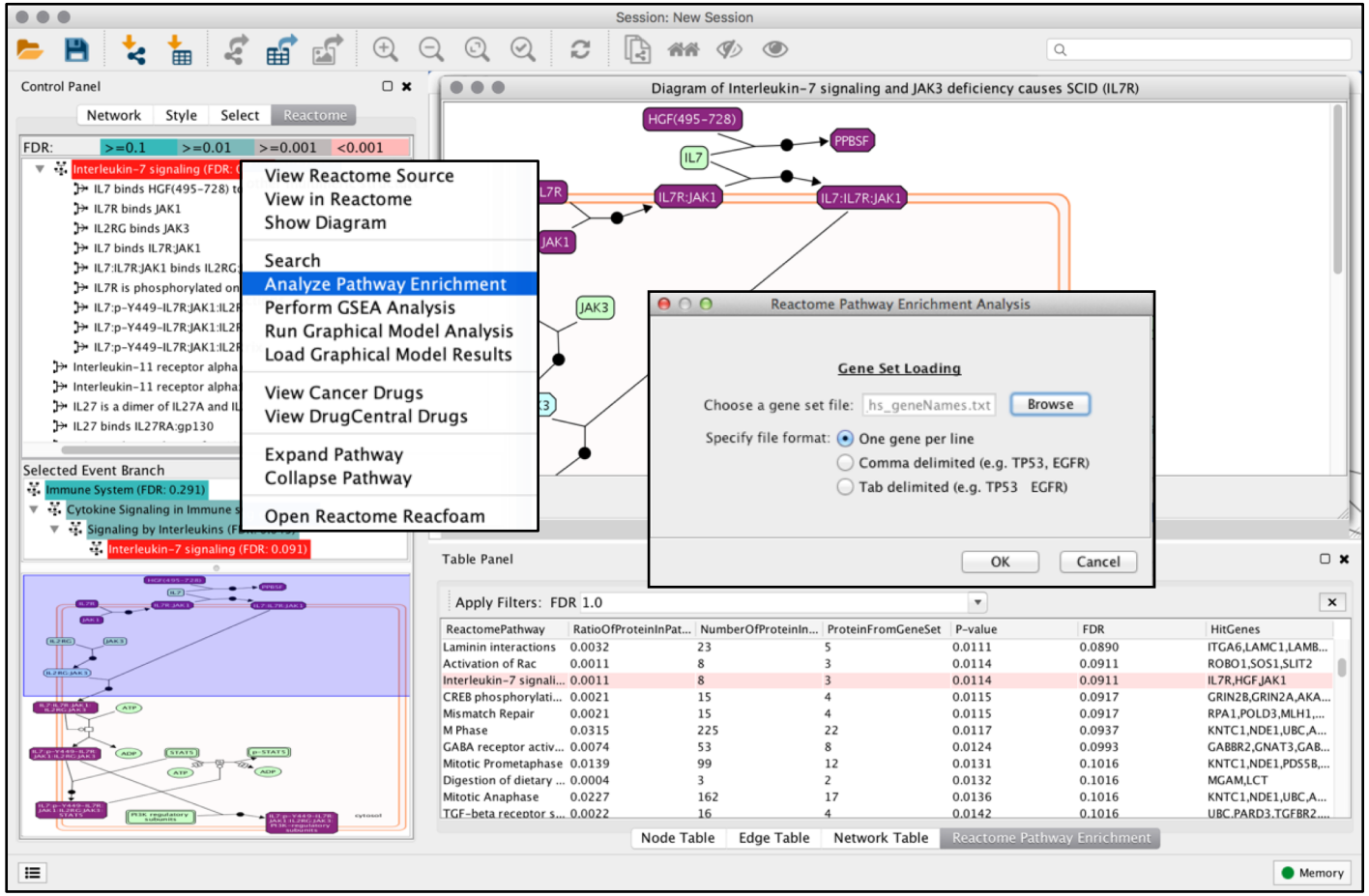
- You can: further cluster your FI network to find protein complexes
- You can perform pathway analysis on the whole network.
- You can perform pathway analysis on the individual modules and it helps to add a biological label.
- You can add drug targets to the network.

And more…

Reactome: Collection of Pathways

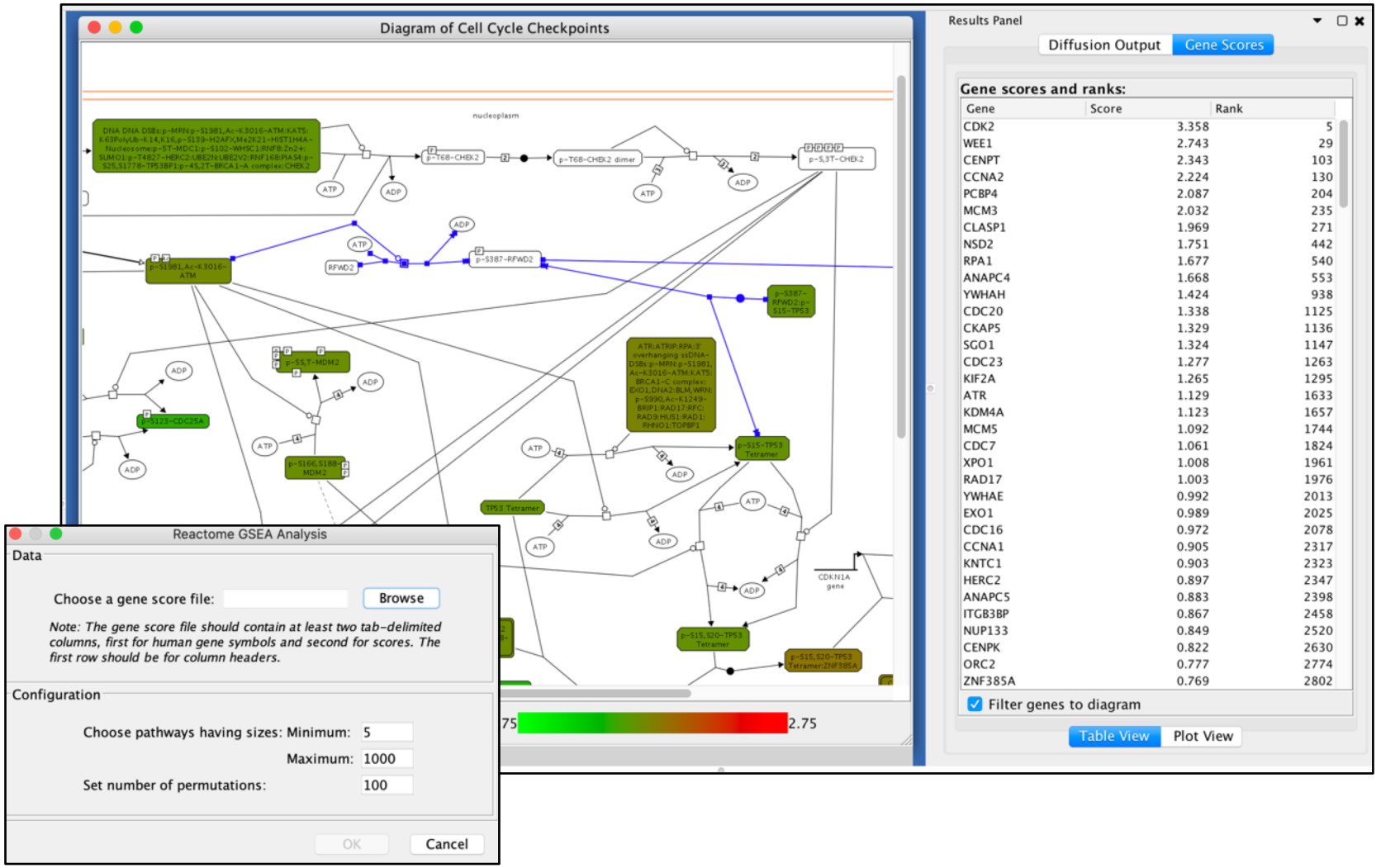


Pathway Enrichment Analysis (gene list) binomial test

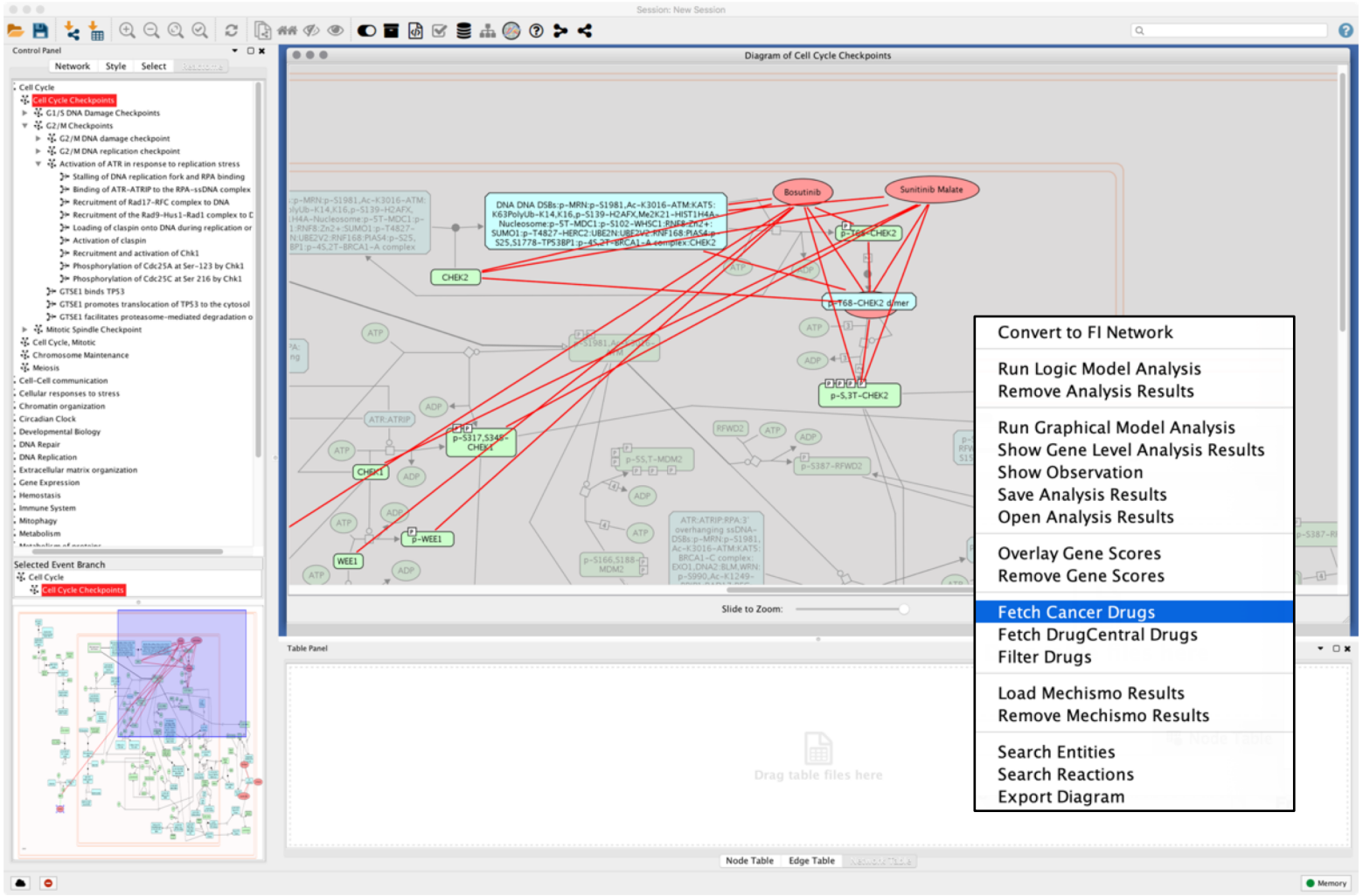


A binomial test uses sample data to determine if the population proportion of one level in a binary (or dichotomous) variable equals a specific claimed value.

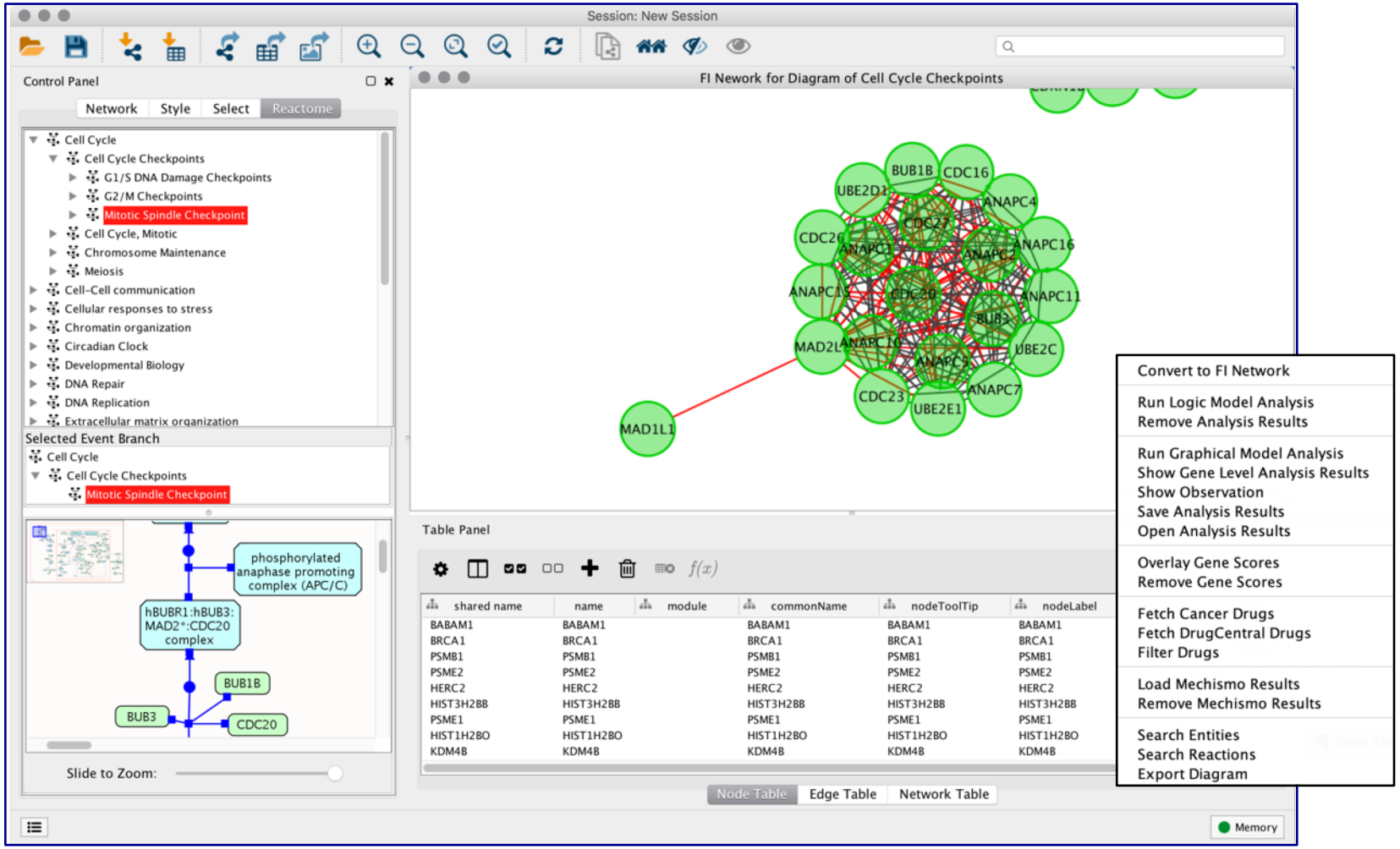
Pathway Enrichment Analysis (ranked list) GSEA



Visualize Cancer Targetome in Reactome Pathways

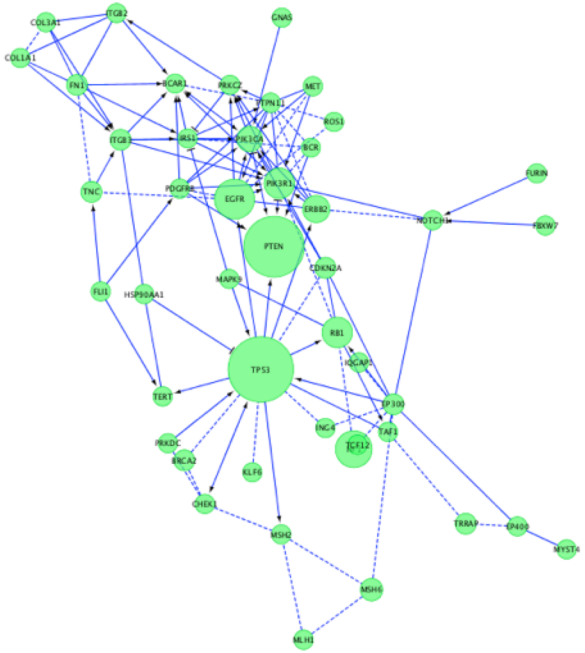


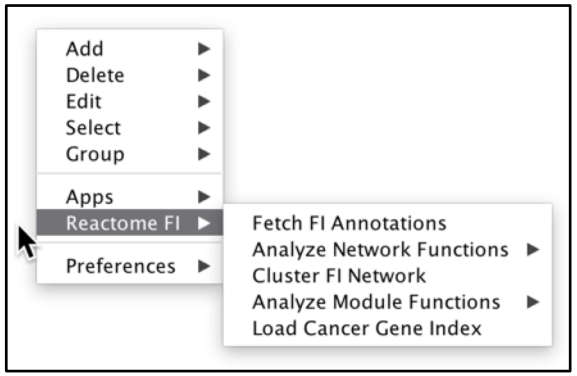
Displaying Reactome Pathways in the FI Network View

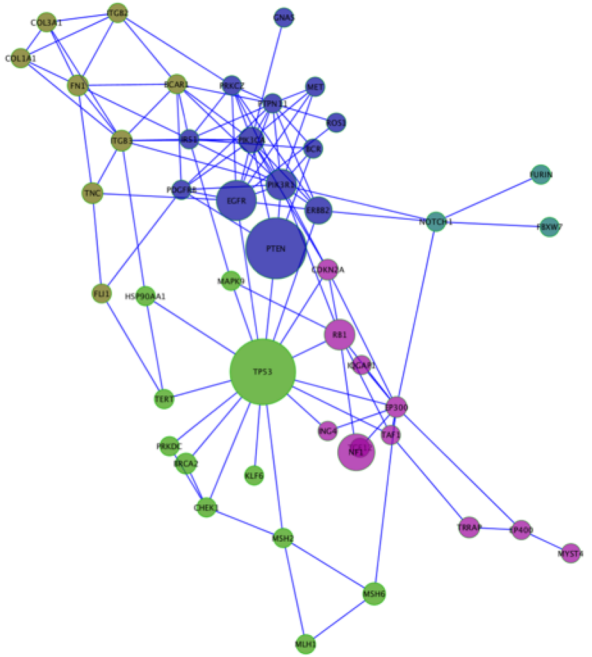


Create and Cluster FI Network using a gene list

- Create the network
- Cluster the network
- Perform pathway analysis







File Formats to Create the FI Network

- Choose Plugins, Reactome FIs.

Simple Gene List

MSI2
PTPRT
PELO
SLC18A1
TACC2
FAM148B
PRC1
MSTN
ATP6V1G2
APOE
IMPA2
AGER
XPO5
MEST
RREB1
BAT1
WIPI1

NCI MAF (mutation annotation file)

Hugo_Symbol	Entrez_Gene_	Center	NCBI_Build	Chromosome	Start_position	End_position	Strand
BAI2	576	broad.mit.edu	36	1	31974808	31974808	+
LRRC41	10489	broad.mit.edu	36	1	46524584	46524584	+
C1orf173	127254	broad.mit.edu	36	1	74828082	74828082	+
SSX2IP	117178	broad.mit.edu	36	1	84900740	84900740	+
NTNG1	22854	broad.mit.edu	36	1	107668441	107668441	+
PGLYRP3	114771	broad.mit.edu	36	1	151541551	151541551	+
FCRL5	83416	broad.mit.edu	36	1	155783467	155783467	+
FBXO28	23219	broad.mit.edu	36	1	222388418	222388418	+
OR2M2	391194	broad.mit.edu	36	1	246410342	246410342	+
MKX	283078	broad.mit.edu	36	10	28063690	28063690	+
.....							

Variant_Classification	Variant_Type	Reference_Allele	Tumor_Seq_Allele1	Tumor_Seq_	dbSNP_RS	dbSNP_Val_S	Tumor_Samp
Missense_Mutation	SNP	C	C	G	novel	none	TCGA-04-133
Missense_Mutation	SNP	C	C	T	novel	none	TCGA-04-133
Missense_Mutation	SNP	T	T	G	novel	none	TCGA-04-133
Silent	SNP	A	A	G	novel	none	TCGA-04-133
Missense_Mutation	SNP	G	G	A	novel	none	TCGA-04-133
Missense_Mutation	SNP	G	G	T	novel	none	TCGA-04-133
Missense_Mutation	SNP	A	A	C	novel	none	TCGA-04-133
Missense_Mutation	SNP	A	A	T	novel	none	TCGA-04-133
Missense_Mutation	SNP	G	G	A	novel	none	TCGA-04-133
Missense_Mutation	SNP	C	C	A	novel	none	TCGA-04-133

Matched_No	Match_Norm	Match_Norm	Tumor_Valid	Tumor_Valid	Match_Norm	Match_Norm	Verification_	Validation_S	Mutation_St	Sequencing_	Sequence_Sc	Validation_M	Score	BAM_file	Sequencer
TCGA-04-1331-10A-01W							Unknown	Unknown	Unknown	x	x	---			x
TCGA-04-1331-10A-01W							Unknown	Unknown	Unknown	x	x	---			x
TCGA-04-1331-10A-01W							Unknown	Unknown	Unknown	x	x	---			x
TCGA-04-1331-10A-01W							Unknown	Unknown	Unknown	x	x	---			x
TCGA-04-1331-10A-01W							Unknown	Unknown	Unknown	x	x	---			x
TCGA-04-1331-10A-01W							Unknown	Unknown	Unknown	x	x	---			x
TCGA-04-1331-10A-01W							Unknown	Unknown	Unknown	x	x	---			x
TCGA-04-1331-10A-01W							Unknown	Unknown	Unknown	x	x	---			x
TCGA-04-1331-10A-01W							Unknown	Unknown	Unknown	x	x	---			x
TCGA-04-1331-10A-01W							Unknown	Unknown	Unknown	x	x	---			x

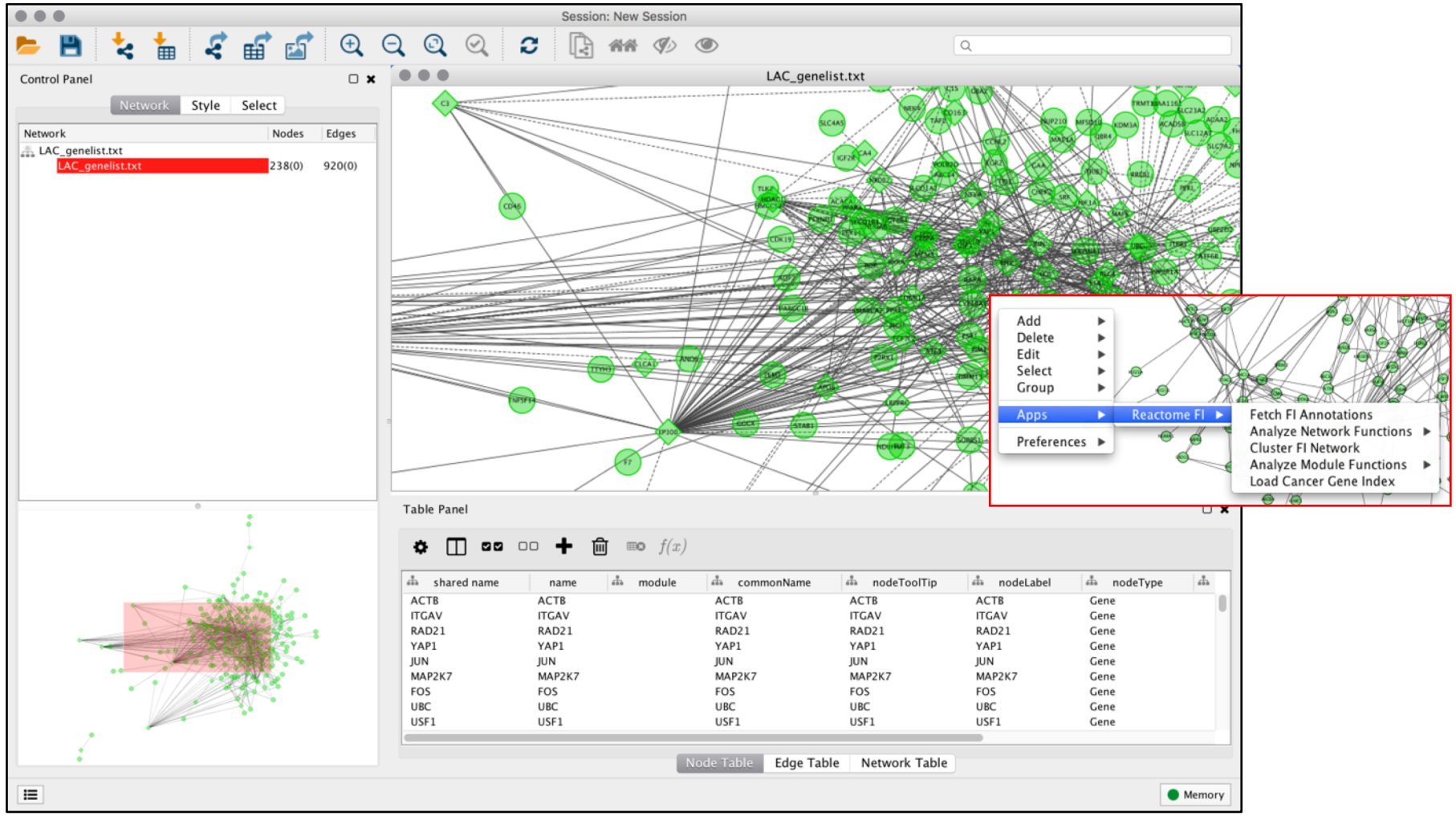
Gene/Sample Number Pairs

Gene	Sample_Number	Samples
IFT88	1	TCGA-10-0937
TACC2	5	TCGA-24-2288;TCGA-09-1662;TCGA-13-0755;TCGA-13-0920;TCGA-24-2035
ADAD2	1	TCGA-24-1419
PODXL2	1	TCGA-24-2290
RTN3	1	TCGA-13-0726
TADA2B	1	TCGA-25-1313
CD274	1	TCGA-25-1627
SULF1	2	TCGA-24-2262;TCGA-09-2056
SYCP2	3	TCGA-24-1563;TCGA-61-2008;TCGA-24-1427
CD200R1L	1	TCGA-13-1488
RCBTB2	1	TCGA-61-2113
TOP2B	1	TCGA-13-0923
C1orf50	1	TCGA-25-2392
ATP9A	5	TCGA-25-1329;TCGA-04-1530;TCGA-24-2262;TCGA-10-0933;TCGA-13-0795
NUDT5	1	TCGA-61-2003
ZNF189	1	TCGA-24-2254
KCNB1	1	TCGA-23-1028
FERMT3	1	TCGA-24-1464
ZCCHC2	1	TCGA-61-1998
DDX18	1	TCGA-13-1509
TGM1	2	TCGA-24-0975;TCGA-25-1321
SAMD7	1	TCGA-09-2051
OR2F2	3	TCGA-13-0723;TCGA-24-2267;TCGA-25-2392
ZSWIM3	2	TCGA-24-1466;TCGA-23-1118
EIF2AK2	2	TCGA-59-2355;TCGA-24-1463
KIAA2018	4	TCGA-23-1024;TCGA-04-1367;TCGA-24-2288;TCGA-24-2289
.....		

Microarray (array) data file

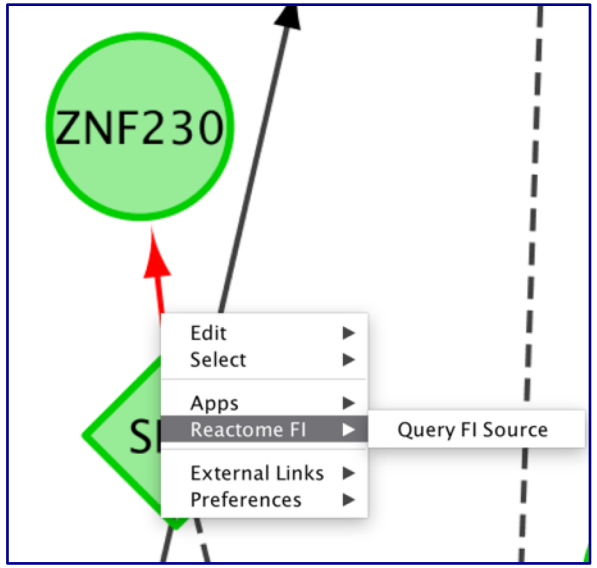
gene	Sample 122	Sample 123	Sample 124	Sample 125	Sample 126	Sample 127	Sample 128	Sample 129	Sample 130
IFT88	1.433357937	0.749571365	0.204170171	1.376375723	0.049504161	0.912377692	0.464660294	0.106486375	0.40767808
KIF9	0.349692236	-0.093947168	-0.568902295	0.177455762	-0.30793794	-0.574121582	-0.772454492	0.850743799	0.19833291
CP5F3	-0.628846878	-0.519774777	0.200101088	-0.388888256	-0.094393584	0.549131811	0.44005971	-0.247094525	1.421708617
TACC2	0.985299156	-0.732555355	1.013646591	-0.057886422	0.730172249	0.554418157	1.280112472	0.463706367	0.956951722
C14orf132	1.249528811	0.104876045	0.524096007	-0.43464182	-1.528259113	-0.467450338	-0.715336925	-0.525776594	-1.06893985
RTN3	0.10643433	-0.286920286	-0.329065423	0.548958273	1.054699921	-0.827782882	-0.462525025	1.39888521	-0.588960437
PODXL2	-0.031430138	-0.390413822	-0.075387324	-0.412392415	0.085789024	-0.097365917	-0.053408731	-0.104692114	1.43380939
CORO1A	-0.247037235	0.664838693	-0.423713196	-0.001970579	0.488162732	-0.435111645	0.596447999	-0.463607768	0.055021666
RBM17	-0.360070659	-0.142165087	-0.233914802	0.018396914	-0.130696373	-0.394476802	-0.818819233	-0.692663375	0.993237633
WT1	-0.768498156	-0.669419285	-0.763665041	-0.674252401	-0.618671571	-0.393931693	-0.906241952	-0.35285021	-0.538925163
SULF1	-0.863622291	1.588947769	0.272263522	-0.419487008	0.401966747	-0.081472544	-1.17805435	0.327289133	-0.21903657
SYCP2	-0.365259803	1.246319117	-0.107066702	0.891658263	0.914356558	0.091543376	-1.037696782	0.727095627	0.009262058
PHYH	-0.252934465	-0.340129465	0.084219531	-0.613340462	-1.85732245	0.479503528	0.22954453	-1.130697457	1.037551523
RCBTB2	-1.216044791	1.137367735	-0.741461602	0.636779988	-0.695953625	-0.46841374	-0.344892088	-0.130354482	-0.572431973
TOP2B	1.329376011	1.272684609	0.636481097	0.611284918	0.466406891	0.982928554	-0.749308731	0.397117399	0.050669942
SEC22B	-0.712050988	0.661526326	-0.33943161	0.917245507	0.427725933	0.413113408	-0.638988365	0.632301277	0.486176031
ATP9A	1.547427859	0.444111328	1.415509578	-0.14352465	0.540051896	0.510070468	-0.233468933	-0.149520936	-0.047584083
TGFBR1	-0.782212802	-0.264862477	-0.123008356	0.027190126	-0.239829397	0.427719409	-0.773868442	0.110633726	1.270499777
NUDT5	-0.616643202	-0.117594702	0.072831699	-0.025664715	0.37488737	0.433985219	-0.89899959	0.020300278	0.112230265
VTI1B	1.252187159	0.515526231	0.073529674	0.770902019	-0.368466883	-1.684634408	0.191395423	0.967344934	2.116535981
RPS6KB1	-0.387163384	0.639591396	-0.597648114	-0.392297158	0.932216509	1.563670699	-0.782463975	0.465043084	-0.674654723
RAPGEF4	0.202545341	0.240920292	0.106607963	1.564856103	0.557513638	-0.027704365	-0.651297319	-0.612922368	-1.178952896
PRO1596	-0.010404297	0.228155239	0.228155239	0.445509484	0.148635394	0.079718194	-0.376195587	0.138032748	0.03730761
KCNB1	0.093447081	-0.061401218	0.135678435	0.628377569	0.656531805	1.458927537	1.318156356	-0.469637643	0.346835207
ALDH9A1	0.702468807	-0.111692273	0.963729452	-1.181038168	-1.083824904	1.698904754	1.844724649	-0.038782325	-1.545587905
ZCCHC2	0.55864339	0.120590563	0.845270548	-0.479704052	0.477522496	-1.144895382	-0.398583158	0.55323533	-0.095731821
LOC51274	-0.291416641	-0.264217345	0.134705668	-0.527143876	-0.073822271	-0.699406086	-0.971399049	-0.218885185	-0.146353728
DDX18	0.34967191	0.294984327	-1.00840307	1.015037505	0.140036175	0.076233995	-0.160745532	0.51373466	1.069725088

FI Results Display

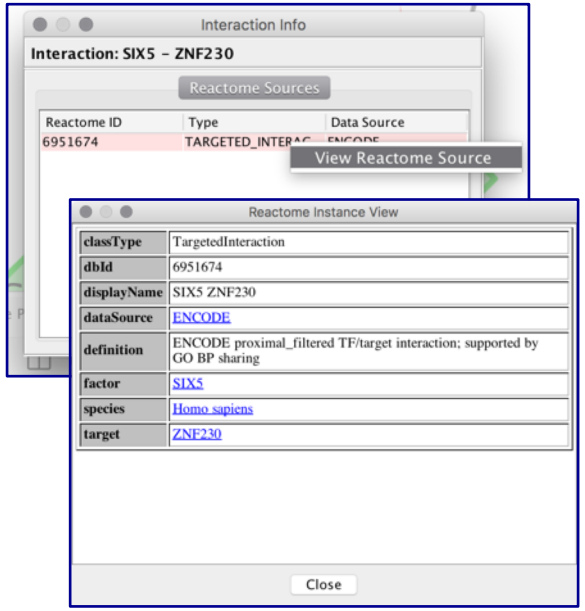


- Constructed network is displayed in the Network View panel using an FI specific visual style

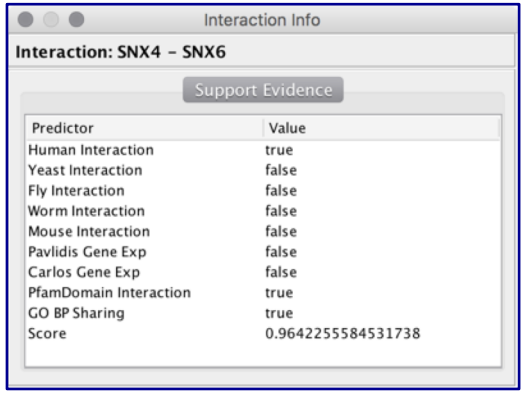
Query FI Source



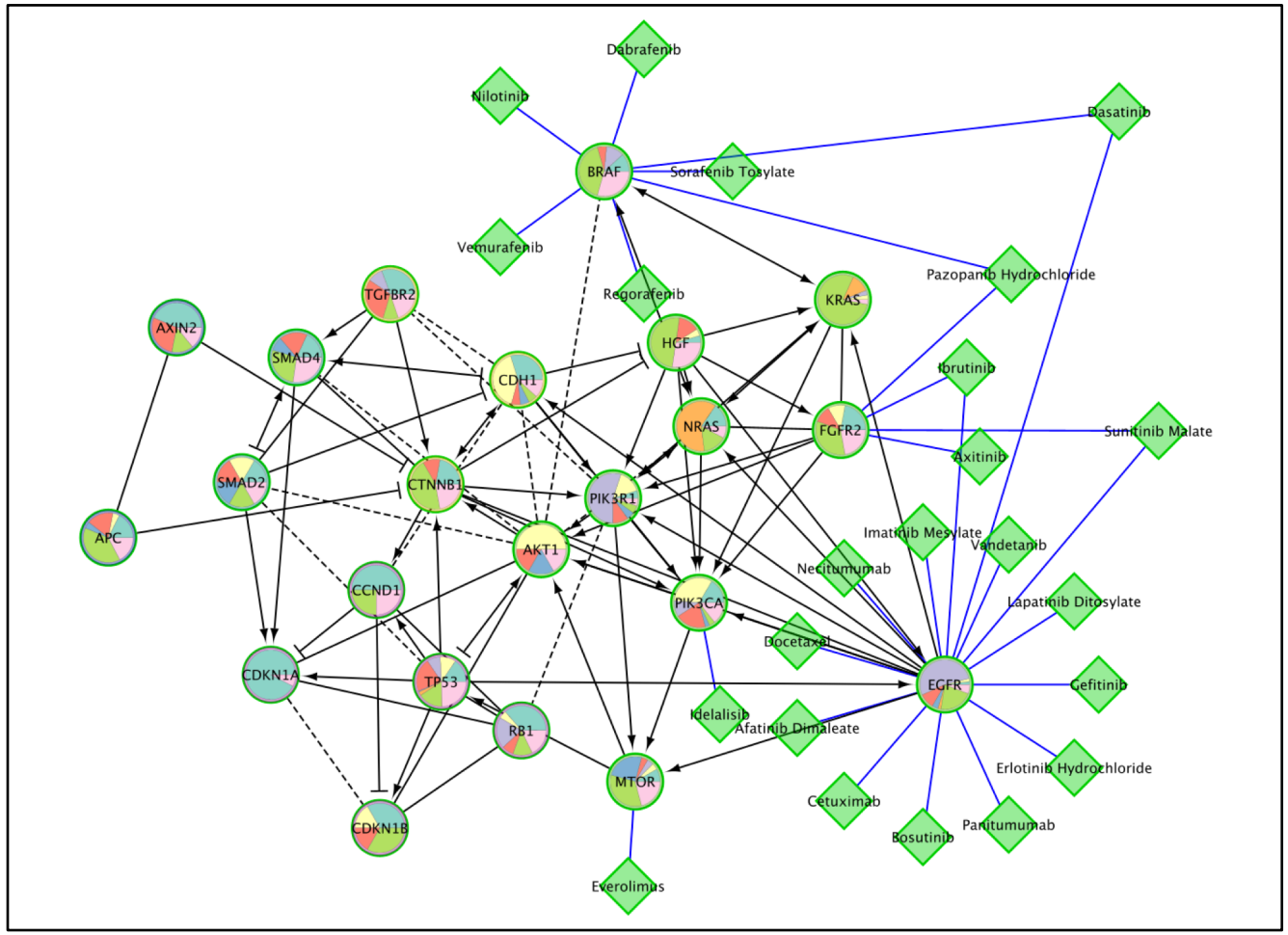
Annotated FIs



Predicted FIs

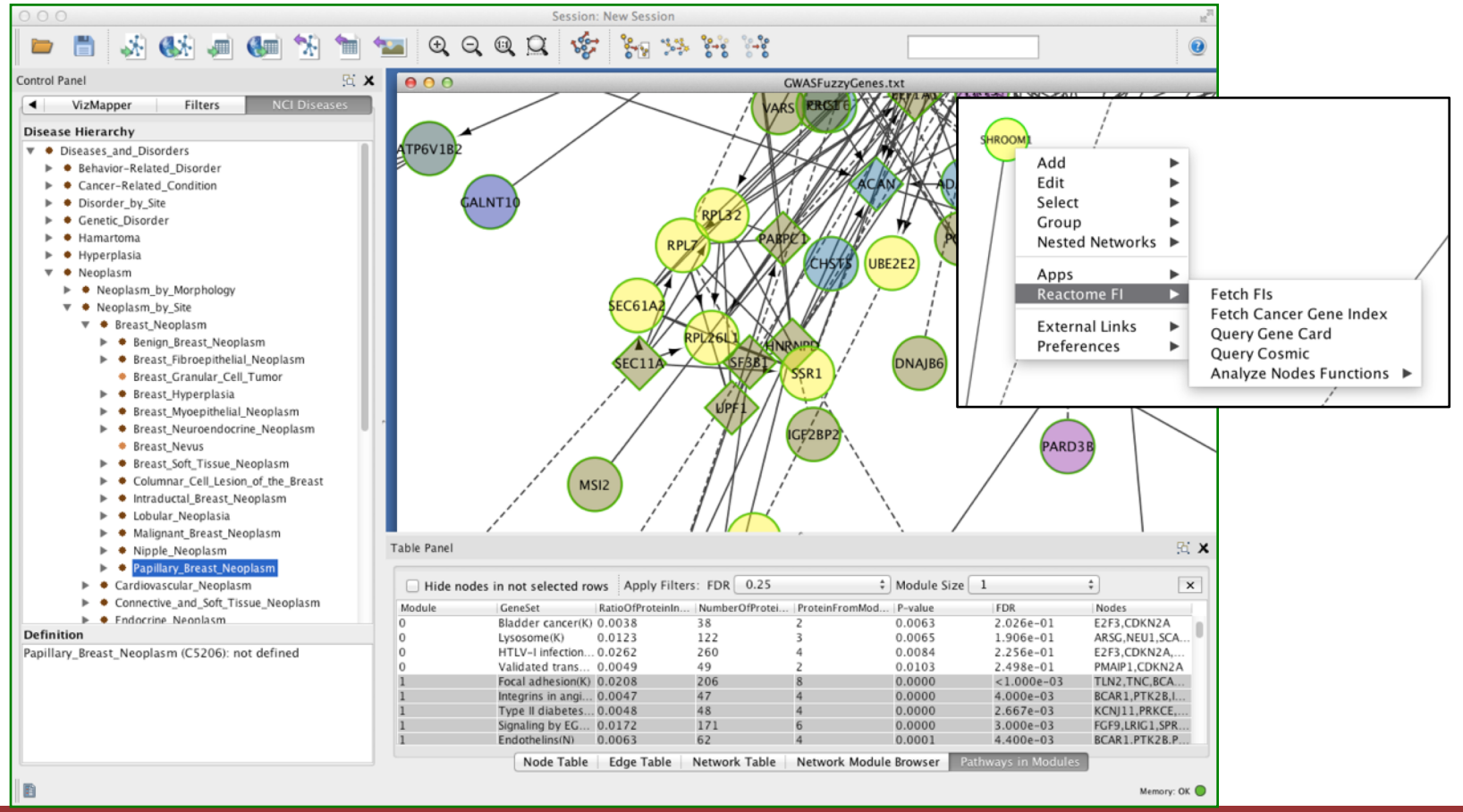


Overlay FDR approved Cancer Drugs



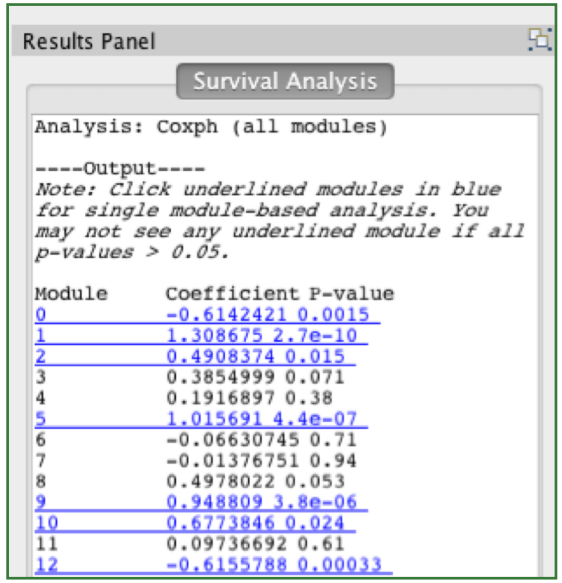
Overlay Cancer Gene Index

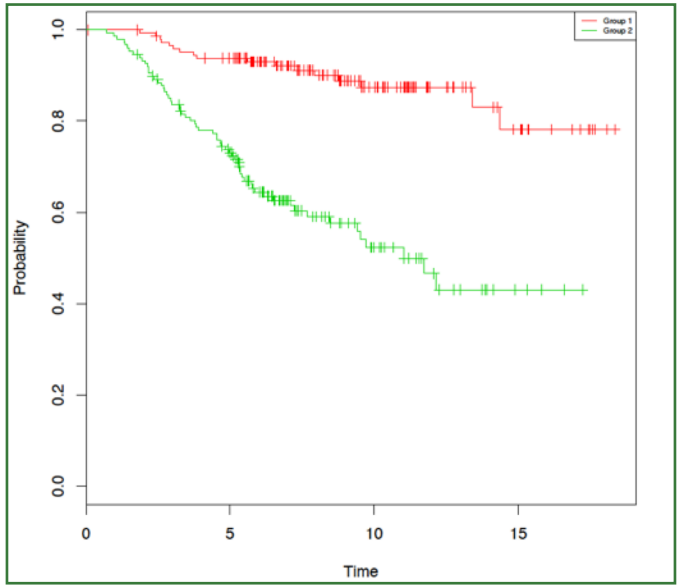
- Load the NCI disease terms hierarchy in the left panel.
- Select a disease term in the tree to select all nodes that have this annotation or one of its sub-terms.



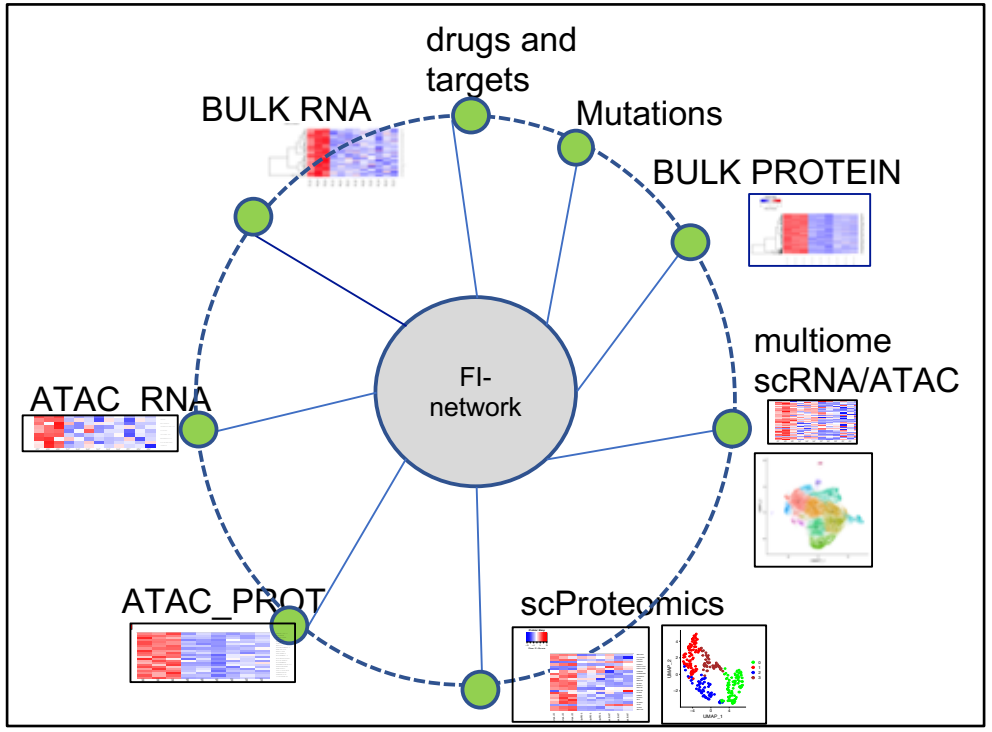
Module Based Survival Analysis

- Discover Prognostic Signatures in Disease Module Datasets.
- Based on a server-side R script that runs either CoxPH or Kaplan-Meyer survival analysis.
- Requires appropriate clinical data file.



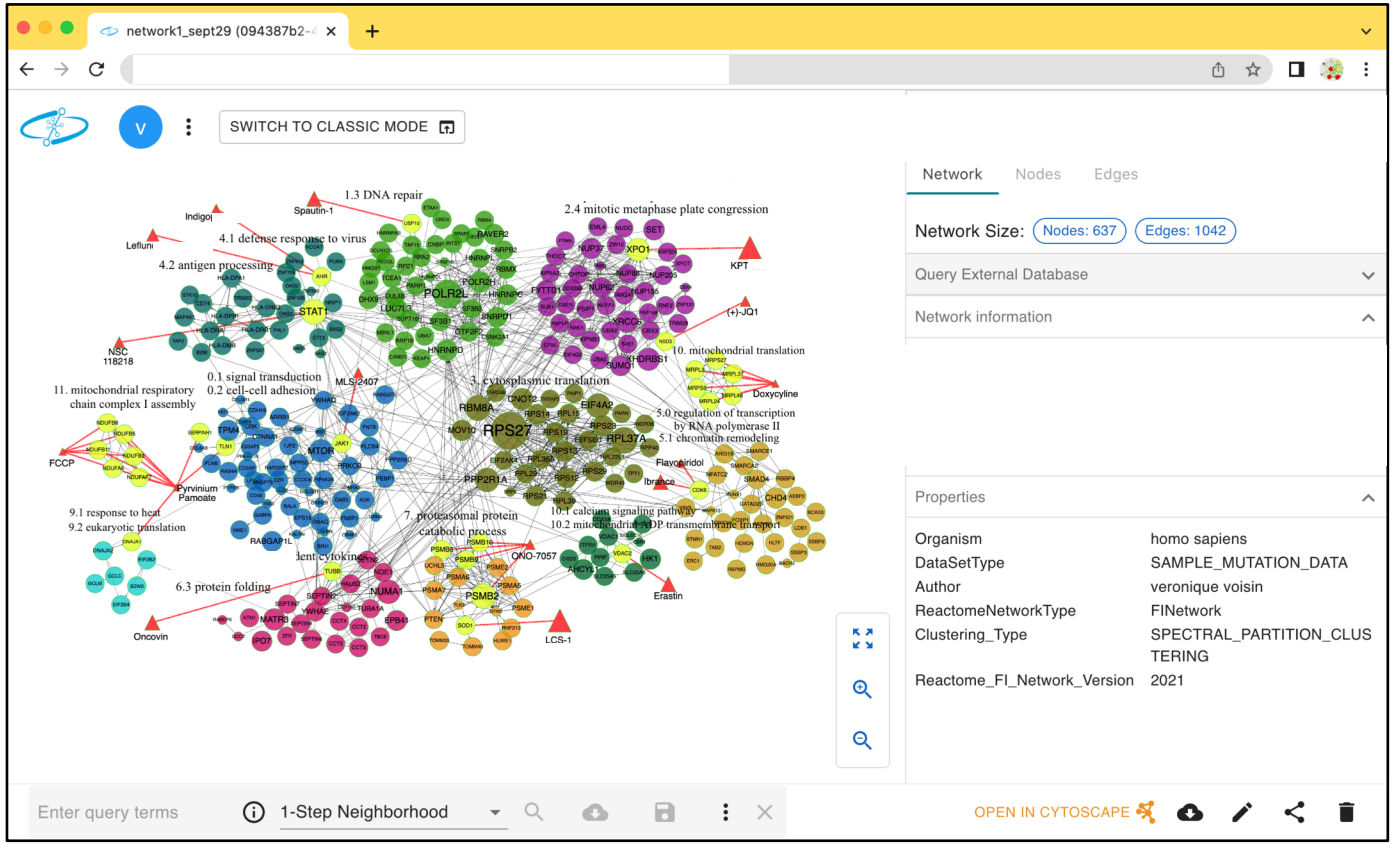


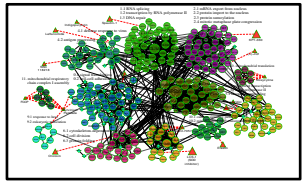
One example of use of Reactome FI VIZ

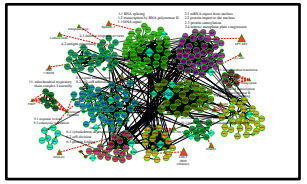


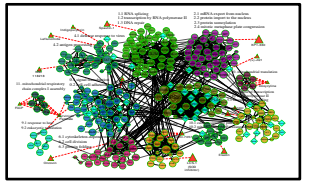
Get gene lists specific to a pure population extracted from the same experiment from different omics data:

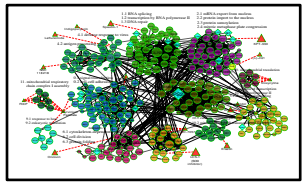
- the omics lists have few direct overlaps but the genes might connect with each other if they are put in the same FI network.




scRNA


scPROT


ATAC


RNA

Workshop Sponsors:

Canadian Centre for Computational Genomics

OICR
Ontario Institute for Cancer Research

GenomeCanada

HPC4Health