

Canadian Bioinformatics Workshops

www.bioinformatics.ca

bioinformaticsdotca.github.io



CC BY-SA 4.0 DEED

Attribution-ShareAlike 4.0 International

Canonical URL : <https://creativecommons.org/licenses/by-sa/4.0/>

[See the legal code](#)


You are free to:


Share — copy and redistribute the material in any medium or format for any purpose, even commercially.

Adapt — remix, transform, and build upon the material for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

 **Attribution** — You must give [appropriate credit](#), provide a link to the license, and [indicate if changes were made](#). You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

 **ShareAlike** — If you remix, transform, or build upon the material, you must distribute your contributions under the [same license](#) as the original.

No additional restrictions — You may not apply legal terms or [technological measures](#) that legally restrict others from doing anything the license permits.

Notices:

You do not have to comply with the license for elements of the material in the public domain or where your use is permitted by an applicable [exception or limitation](#).

No warranties are given. The license may not give you all of the permissions necessary for your intended use. For example, other rights such as [publicity, privacy, or moral rights](#) may limit how you use the material.

Learning Objectives

- By the end of this lecture, you will:
 - Understand the advantages of network visualization.
 - Understand how to choose and install Cytoscape Apps
 - Be able to use basic Cytoscape features
 - Be able to create and optimize a network in Cytoscape

Network Visualization and Analysis Outline

- Introduction
- Network Basics
- Network Visualization
- Network Analysis

- Demo: Cytoscape software tool for Network visualization and analysis

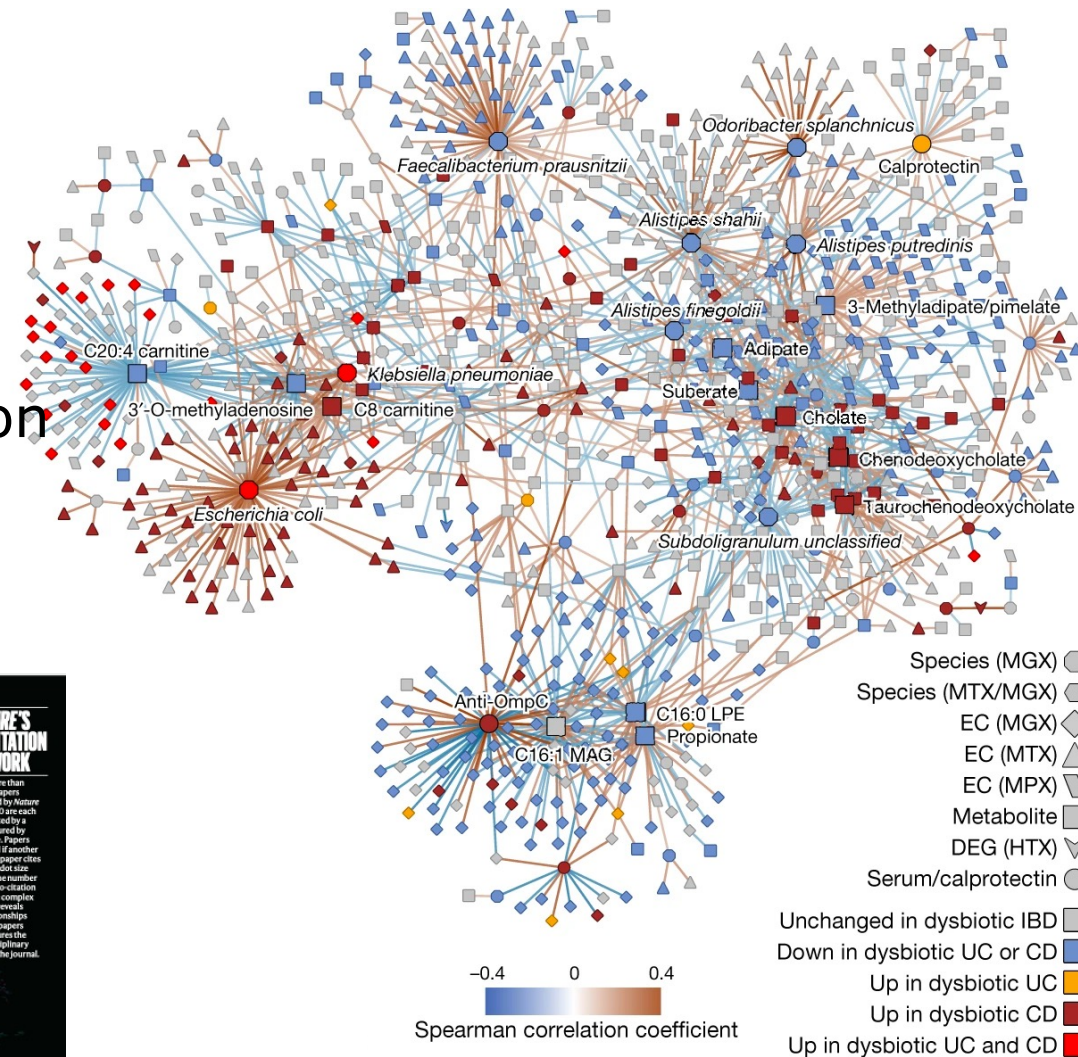
Six Degrees of Separation

- Everyone in the world is connected by at most six links.
- Originally demonstrated by Stanley Milgram at Harvard in 1960s in an experiment to find how many random acquaintances it would take to connect two people.... sometimes defined as the the **social distance** between two people.
- Popularized in the 1990s with six degrees of Kevin Bacon.
- Demonstrated the interconnectivity of our social structures.
- Networks exists everywhere – power grids, cell tower networks, Fed ex delivery networks, supply chain networks
- Is this Biologically relevant?

Why Networks?

Networks are everywhere....

- Molecular networks
- Cell-Cell Communication
- Nervous systems
- Social Networks

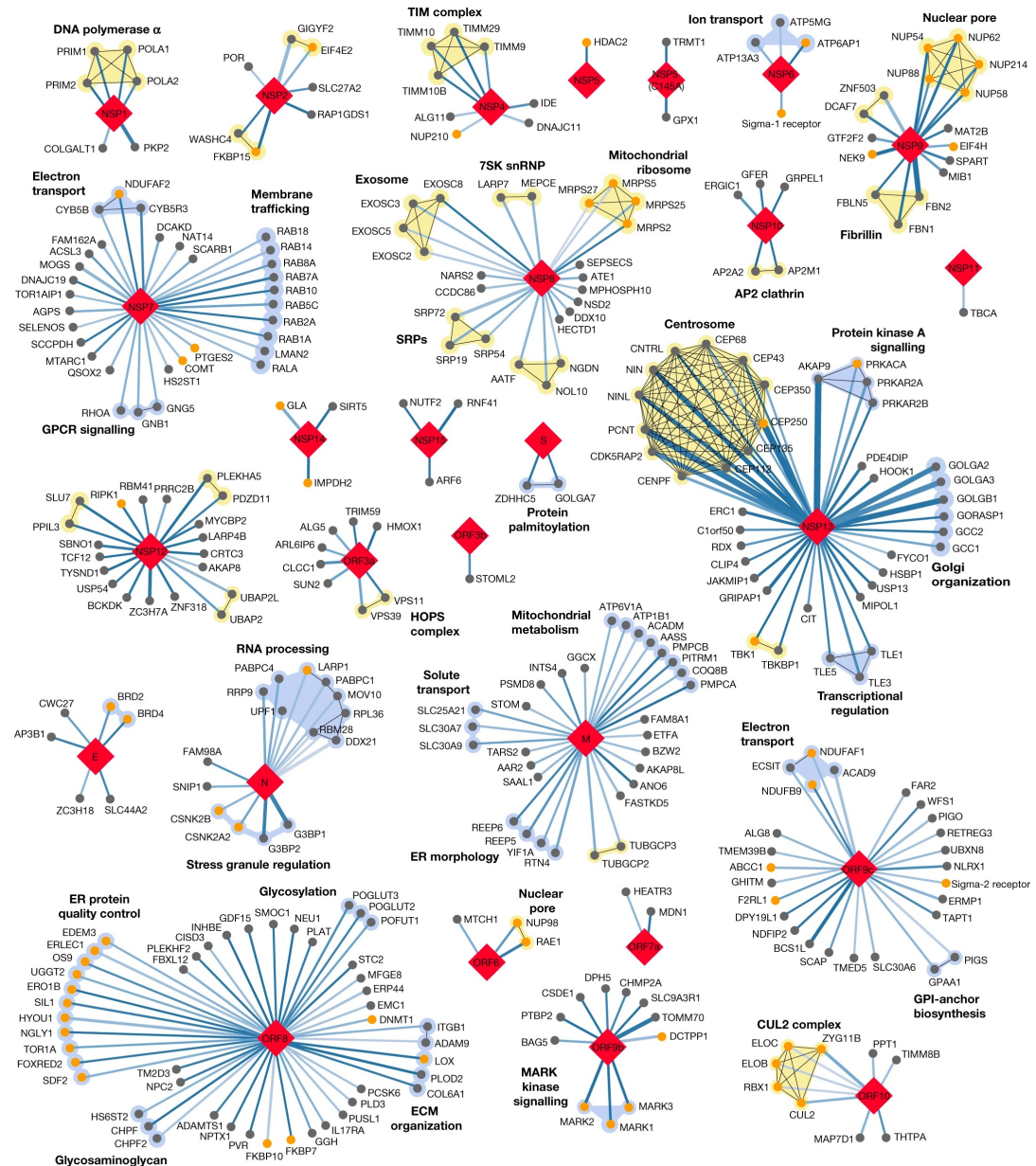
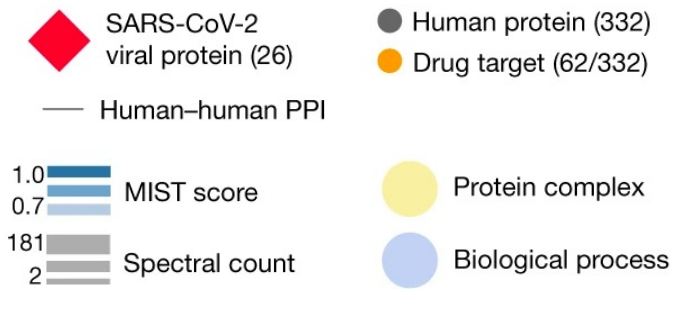


Lloyd-Price, J., Arze, C., Ananthakrishnan, A.N. *et al.* Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. *Nature* **569**, 655–662 (2019)

Why Networks?

Networks are powerful tools....

- Reduce complexity
- More efficient than tables
- Great for data integration
- Intuitive visualization



Gordon, D.E., Jang, G.M., Bouhaddou, M. *et al.* A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* **583**, 459–468 (2020).

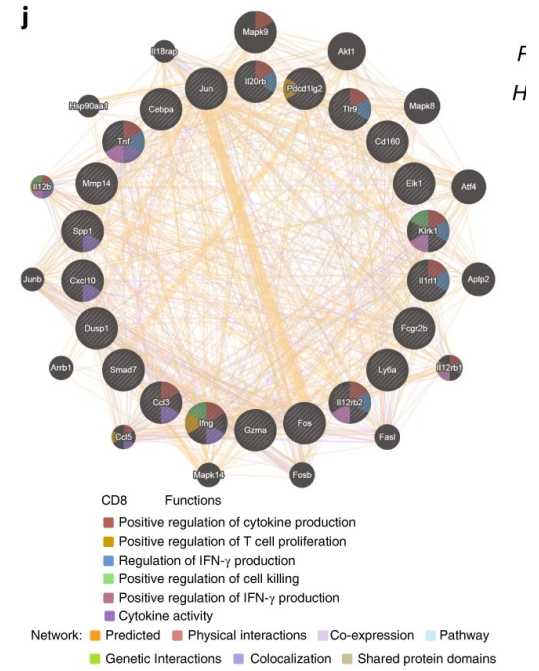
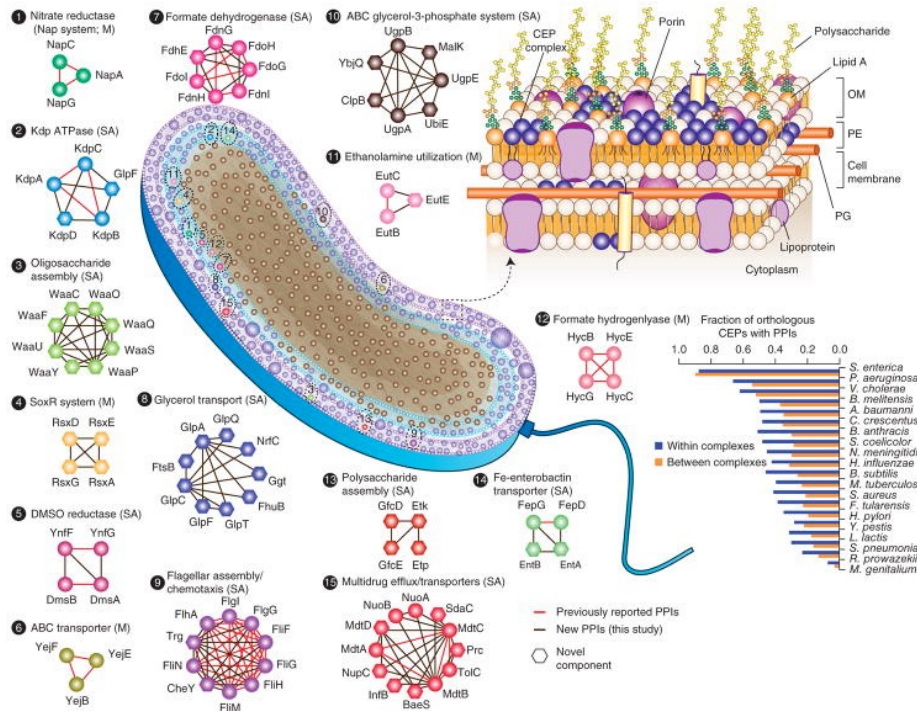
Why Would We Use Network Visualization for Biological Data?

- Represent relationships of biological molecules
 - Physical, regulatory, genetic, functional interactions
- Useful for discovering relationships in large data sets
 - Better than tables in Excel
- Visualize multiple data types together
 - Discover interesting patterns
- Network analysis
 - Finding sub-networks with certain properties (densely connected, co-expressed, frequently mutated, clinical characteristics)
 - Finding paths between nodes (or other network “motifs”)
 - Finding central nodes in network topology (“hub” genes)

Applications of Network Biology

Detection of protein complexes/other modular structures – identification of subnetworks of interconnected nodes that might work together to perform a specific function, i.e. molecular complexes.

Gene Function Prediction – associate novel gene to existing function or pathway



Wang, Z., Aguilar, E.G., Luna, J.I. et al. Paradoxical effects of obesity on T cell function during tumor progression and PD-1 checkpoint blockade. *Nat Med* **25**, 141–151 (2019)

Babu M, Bundalovic-Torma C, Calmettes C, et al. Global landscape of cell envelope protein complexes in *Escherichia coli*. *Nat Biotechnol.* 2018;36(1):103-112

Applications of Network Biology

Subnetwork based diagnosis – biomarker or biomarker panel. Identification of subnetworks activated in disease

Network Alignment and comparison– analyze evolution of networks between different species to identify similarities and evolutionary relationships

Motif analysis– discover higher order organization such motifs, feedback loops (small network patterns that are over-represented when compared to a randomized version of the same network)

Pathway based gene association – uncovering common pathways disturbed during disease.

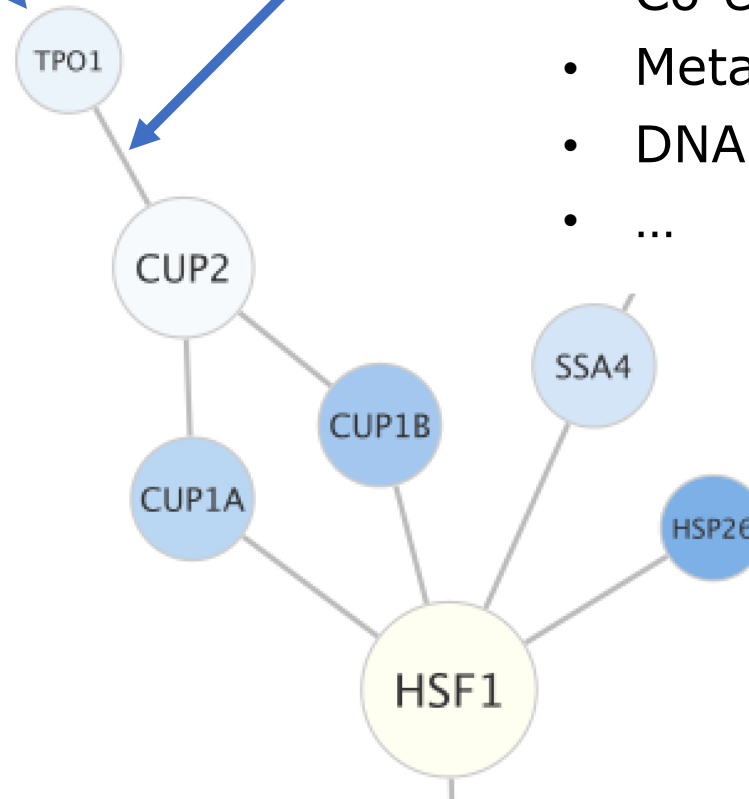
Network Basics

Node (molecule/entity)

- Gene
- Protein
- Transcript
- Drug
- MicroRNA
- ...

Edge (interaction/relationship)

- Genetic interaction
- Physical protein interaction
- Co-expression
- Metabolic reaction
- DNA-binding
- ...



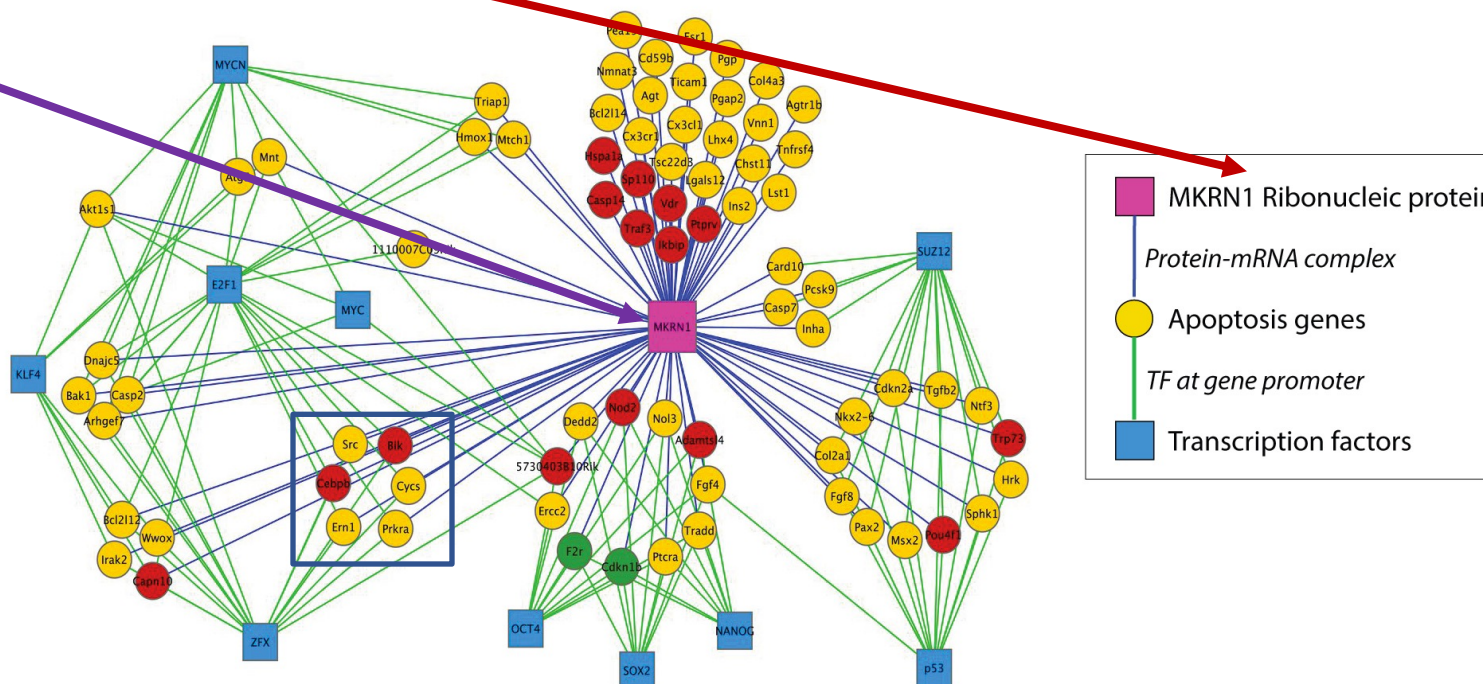
Networks as Tools

Analysis

- Topological properties
- Hubs and subnetworks
- Classify, cluster, diffuse and over-representation
- Data integration

Topological properties:

- Number of nodes and edges
- Node degree
- Degree distribution
- Clustering coefficients

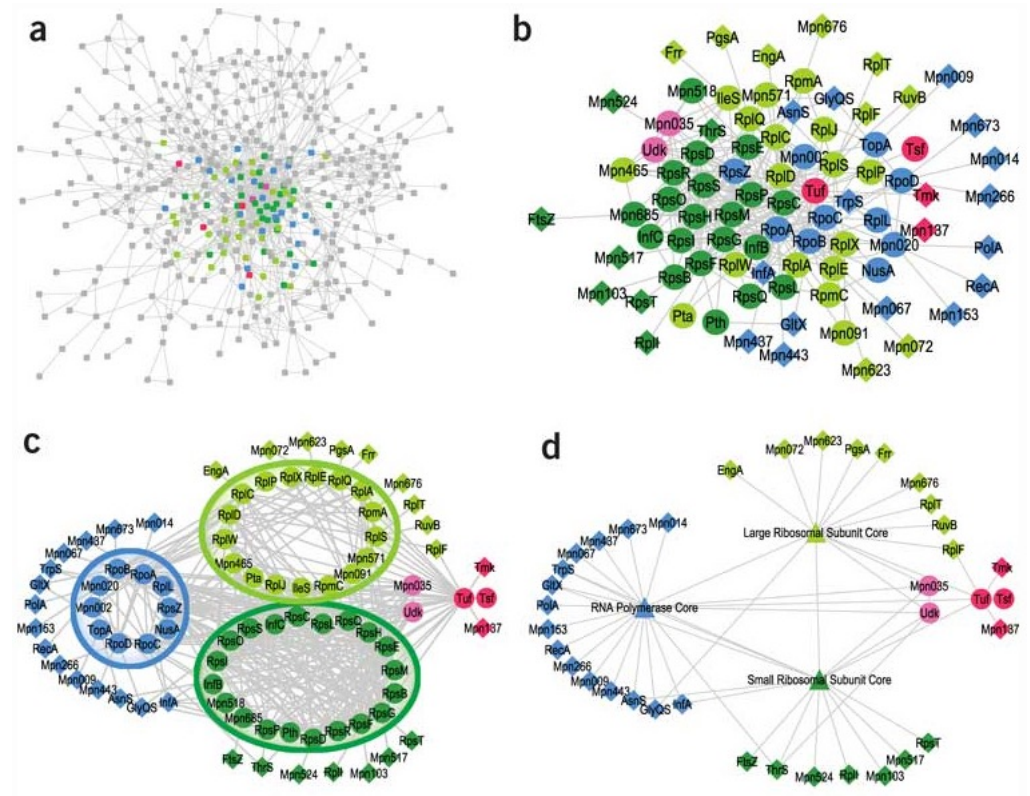


Networks as Tools

Visualization

- Data overlays – integrate multiple data types
- Layouts
- Exploratory analysis
- Context and interpretation

- (a) Mass spectrometry analysis: 400 PPI in pneumonia microbe.
- (b) Subnetwork with functional annotations
- (c) Same network as (b) with added knowledge of complexes
- (d) Same network as (c) with complexes collapsed to group nodes to reduce complexity



Gehlenborg N, O'Donoghue SI, Baliga NS, et al. Visualization of omics data for systems biology. *Nat Methods*. 2010;7(3 Suppl):S56-S68

What's Missing?

- Dynamics
 - Pathways/networks represented as static processes
 - Difficult to represent a calcium wave or a feedback loop
 - More detailed mathematical representations exist that handle these e.g. Stoichiometric modeling, Kinetic modeling (VirtualCell, E-cell, ...)
 - Need to accumulate or estimate comprehensive kinetic information
- Detail – atomic structures
- Context – cell type, developmental stage

What Have We Learned?

- Networks are useful for seeing relationships in large data sets
- Important to understand what the nodes and edges mean
- Important to define the biological question - know what you want to do with your gene list or network
- Many methods available for gene list and network analysis
 - Good to determine your question and search for a solution
 - Or get to know many methods and see how they can be applied to your data

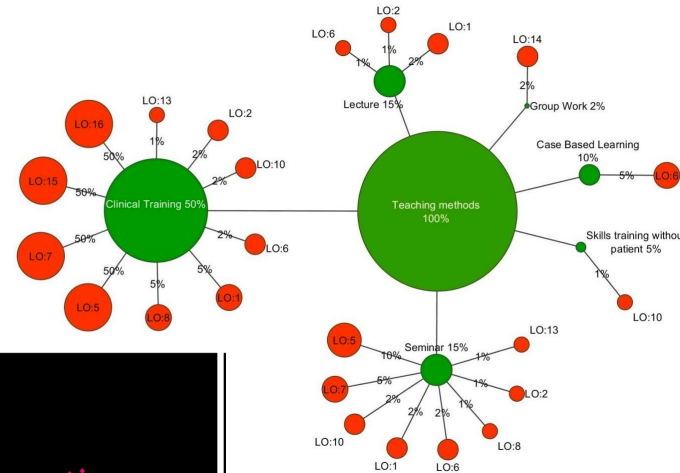
Network Visualization and Analysis using Cytoscape

- Network visualization and analysis using Cytoscape software
- Cytoscape basics
- Cytoscape network analysis examples

Cytoscape (cytoscape.org)



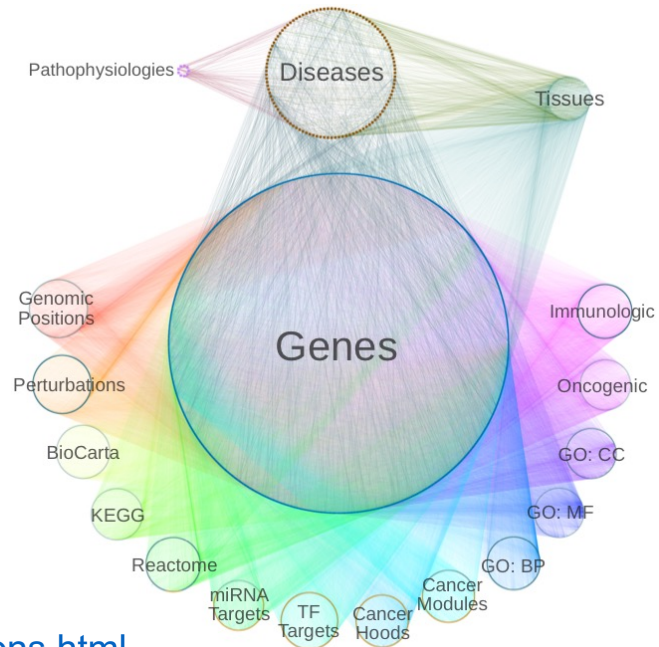
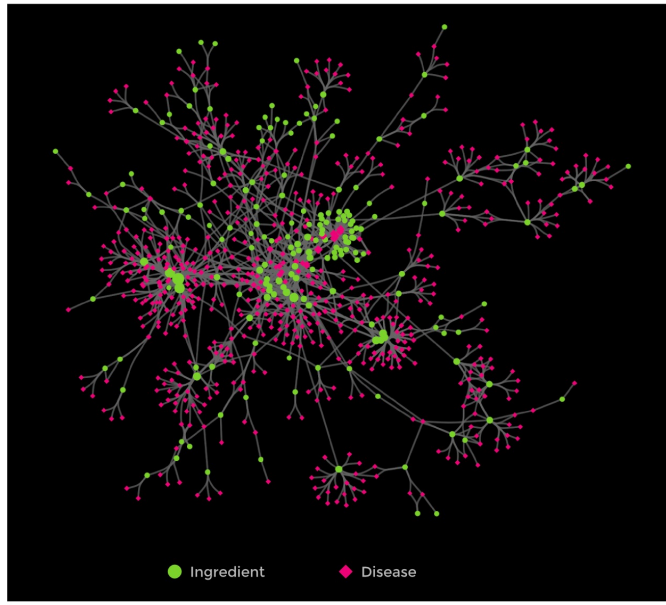
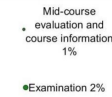
- An open source software platform for visualizing and analyzing complex networks.



Non-taught learning outcomes



Non-teaching oriented activities



<https://nrnb.org/competitions.html>

Cytoscape (cytoscape.org)



Partners and Collaborators

GLADSTONE
INSTITUTES

UCSF University of California
San Francisco

 **UCSD**

 **HARVARD MEDICAL SCHOOL**

 **DANA-FARBER**
CANCER INSTITUTE

 **UNIVERSITY OF**
TORONTO

Google


INDIANA UNIVERSITY


UNIVERSITY OF
COPENHAGEN

 **Cancer Genome**
COLLABORATORY
Cloud computing for collaborative research

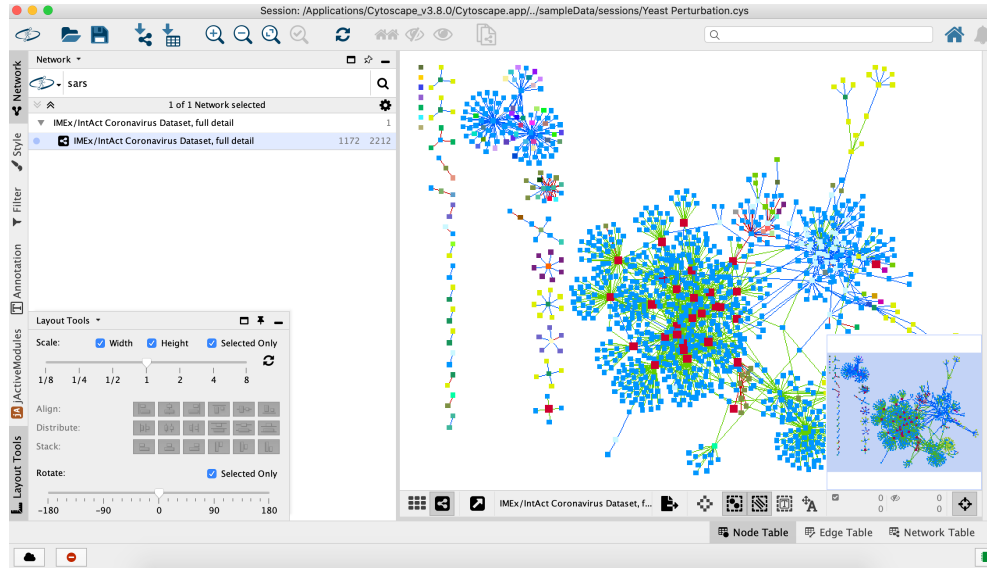
 **BROAD**
INSTITUTE

 **GENEONTOLOGY**
Unifying Biology

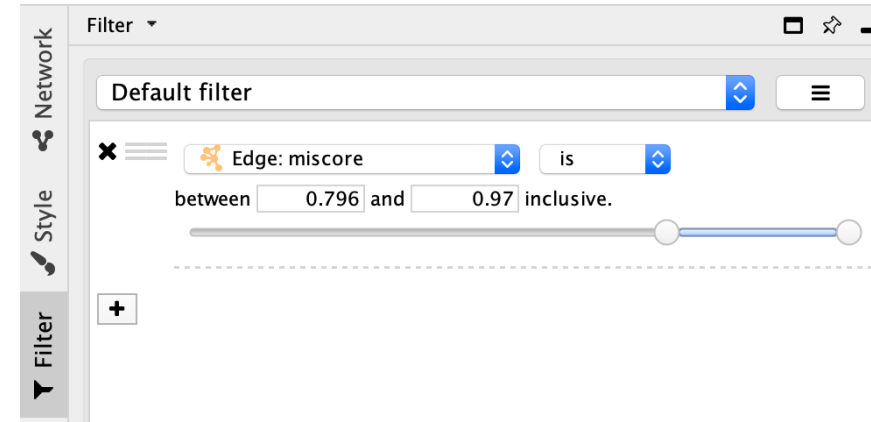




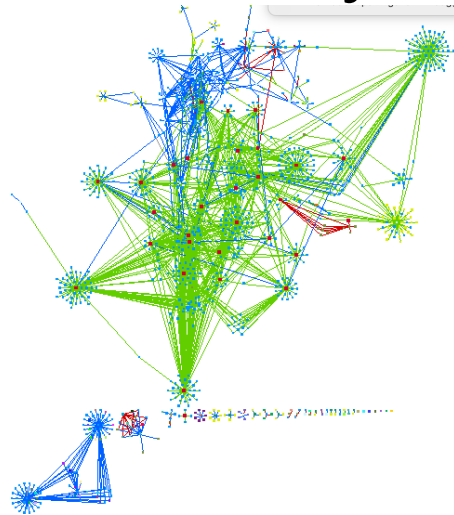
Manipulate Networks



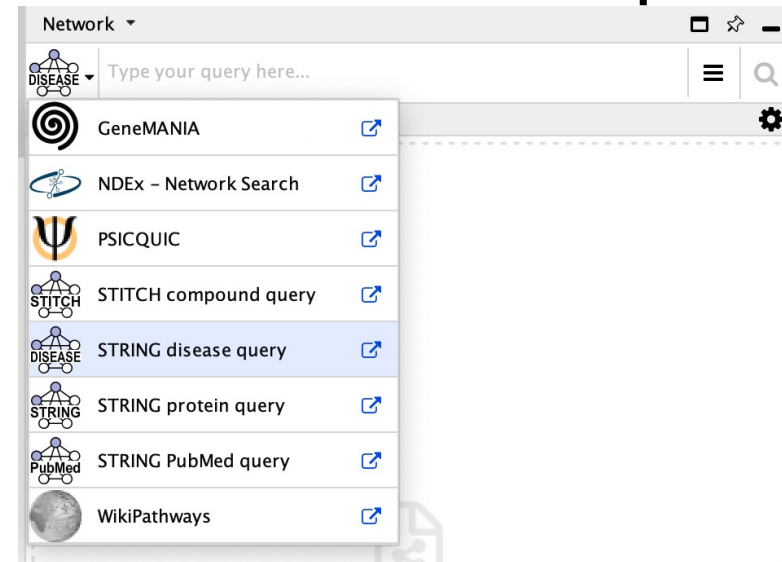
Filter/Query



Automatic Layout



Network Import



The Cytoscape App Store

Cytoscape App Store

Submit an App - Search the App Store Sign In

<http://apps.cytoscape.org>

Wall of Apps 383 total

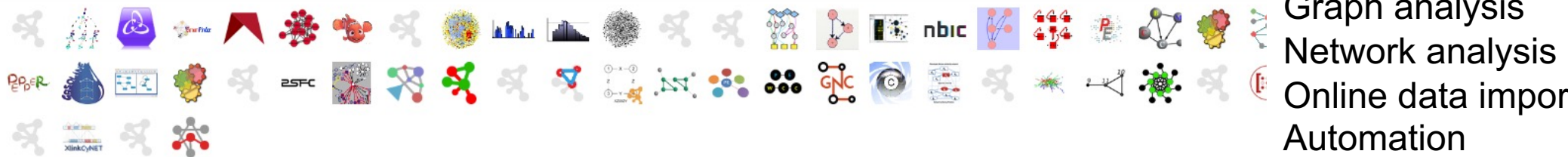
data visualization



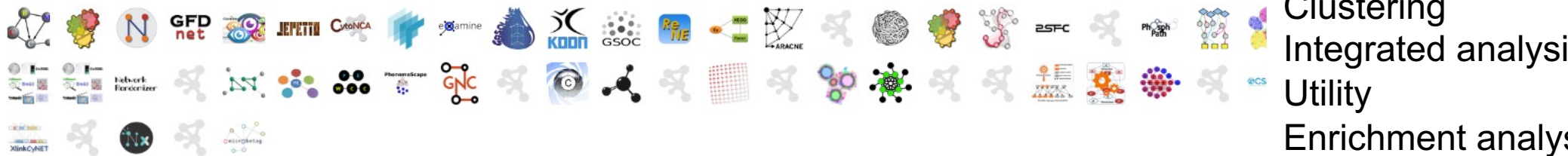
network generation



graph analysis



network analysis



- Categories:**
- Data visualization
 - Network generation
 - Graph analysis
 - Network analysis
 - Online data import
 - Automation
 - Clustering
 - Integrated analysis
 - Utility
 - Enrichment analysis
 - Layout
 - Network Comparison

Active Community

<http://www.cytoscape.org>

- 10,000s users, >17,600 downloads/month (May 2018)
- Help
 - Documentation, data sets
 - Mailing lists
 - <http://tutorials.cytoscape.org>
- Automatable through R or python using cyrest
- Cited 2745 (google scholar), 1,843 (ISI web of Science) (June 2024)
- >383 active Apps Extend Functionality (June 2024)
 - Build your own, requires programming

Cline MS et al. Integration of biological networks and gene expression data using Cytoscape Nat Protoc. 2007;2(10):2366-82

Top Cytoscape Apps

Published 2009
Citations: 5,856



ClueGO 3.0+
Creates and visualizes a functionally grouped network of




stringApp 3.0+
Import and augment Cytoscape networks from STRING

Published 2018
Citations: 1529

Published 2003
Citations: 5,928



MCODE 3.0+
Clusters a given network based on topology to find densely



yFiles Layout Algorithms 3.0+
Highly sophisticated algorithms for arranging networks.

Published 2013
Citations: 1,038



CluePedia 3.0+
CluePedia: A ClueGO plugin for pathway insights using integrated



BiNGO 3.0+
Calculates overrepresented GO terms in the network and display

Published 2005
Citations: 4,548

Core app
Published 2014
Citations: 46




enhancedGraphics 3.0+
A passthrough mapper for charts and gradients




clusterMaker2 3.0+
Multi-algorithm clustering app for Cytoscape

Published 2011
Citations: 610

Published 2010
Citations: 2,183*




EnrichmentMap 3.0+
Visualizes enrichments of pathways as an enrichment map.



cytoHubba 3.0+
Predicts and explores important nodes and subnetworks in a given

Published 2014
Citations: 4,153

Core app
Published 2015
Citations: 92




cyREST 3.0+
Core App: Language-agnostic RESTful API



GeneMANIA 3.0+
Imports interaction networks from public databases from a list of

Published 2010
Citations: 654*

Published 2010
Citations: 242*




ReactomeFIPlugin 3.0+
Explore Reactome pathways and search for diseases related



KEGGscape 3.0+
File reader and pathway visualizer for KEGG XML (KGML) files

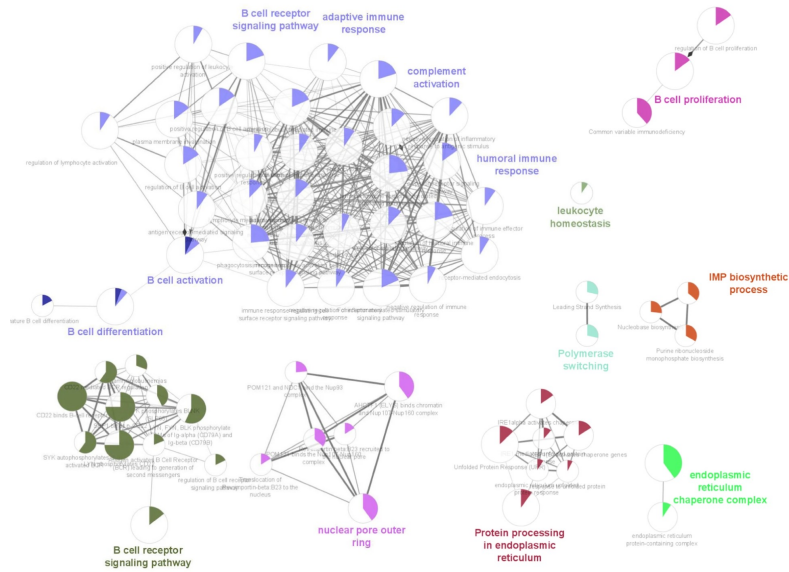
Published 2014
Citations: 79

* Citations based on single publication with multiple primary publications available

 Apps we covered in workshop
Google scholar
(Updated June 2024)

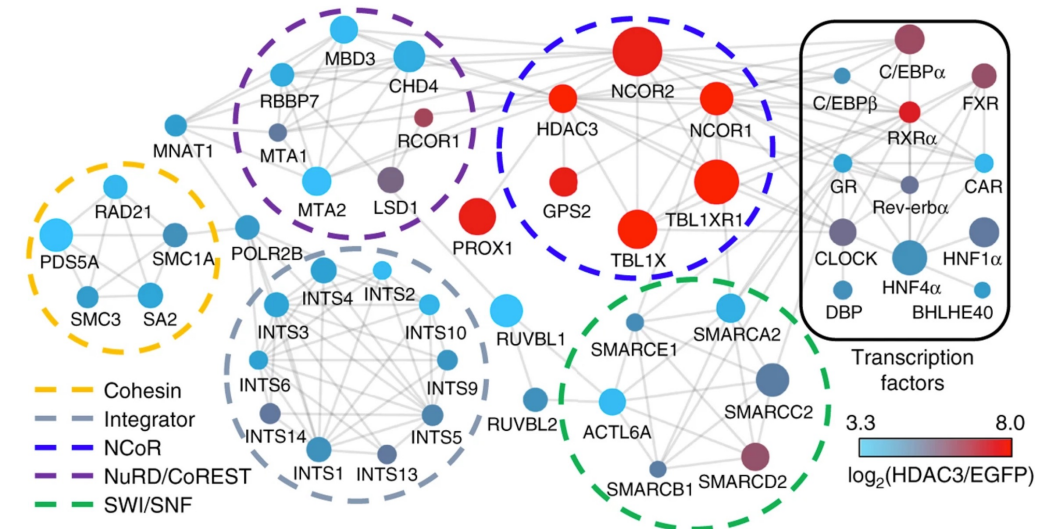
Example of Top Cytoscape Apps

not covered in workshop



Riese, P., Trittel, S., Akmatov, M.K. *et al.* Distinct immunological and molecular signatures underpinning influenza vaccine responsiveness in the elderly. *Nat Commun* **13**, 6894 (2022).

ClueGO Creates and visualizes a functionally grouped network of	stringApp Import and augment Cytoscape networks from STRING
MCODE Clusters a given network based on topology to find densely	yFiles Layout Algorithms Highly sophisticated algorithms for arranging networks.
CluePedia CluePedia: A ClueGO plugin for pathway insights using integrated	BiNGO Calculates overrepresented GO terms in the network and display
enhancedGraphics A passthrough mapper for charts and gradients	clusterMaker2 Multi-algorithm clustering app for Cytoscape...
EnrichmentMap Visualizes enrichments of pathways as an enrichment map,	cytoHubba Predicts and explores important nodes and subnetworks in a given
cyREST Core App: Language-agnostic RESTful API	GeneMANIA Imports interaction networks from public databases from a list of
ReactomeFIPPlugin Explore Reactome pathways and search for diseases related	KEGGscape File reader and pathway visualizer for KEGG XML (KGML) files



Armour, S.M., Remsberg, J.R., Damle, M. *et al.* An HDAC3-PROX1 corepressor module acts on HNF4α to control hepatic triglycerides. *Nat Commun* **8**, 549 (2017). <https://doi.org/10.1038/s41467-017-00772-5>

What Have We Learned?

- Cytoscape is a useful, free software tool for network visualization and analysis
- Provides basic network manipulation features
- Apps are available to extend the functionality

Cytoscape Demo

Version 3.8.0

www.cytoscape.org

Interface Overview

Results Panel

The screenshot shows a software interface for network analysis. At the top, there's a title bar 'Session: New Session' and a toolbar with icons for file operations, search, and navigation. On the left, a vertical sidebar contains 'Network manager', 'Style', 'Filter', 'Annotation', 'App Store', and 'Layout Tools'. The main area is divided into several sections:

- Network manager:** Located in the top-left, it shows a tree view of networks. The selected network is 'IMEx/IntAct Coronavirus Dataset, full detail' with 2233 nodes and 4477 edges.
- Canvas:** The central area displays a complex network graph with nodes (blue and red circles) and edges (green and blue lines). A smaller inset graph is visible in the bottom-right of the canvas area.
- Table Panel:** Located at the bottom, it contains a 'Node Table' with columns for 'shared name', 'name', 'alias', 'taxid', 'species', 'm.type', 'm.type.mi', and 'mut'. It lists various protein identifiers and their associated data.
- Control Panel:** Located at the bottom-left, it includes a 'Command Line' and tabs for 'Node Table', 'Edge Table', 'Network Table', 'Network Module Browser', and 'Enrichment Table'.
- Results Panel:** Located at the top-right, it contains a 'Node Summary' panel.

Network manager

Network overview

Canvas

Table Panel

Control Panel

Basic Navigation

Open or save network to NDex

Open a session

Save your session

Zoom in/out

Apply preferred layout

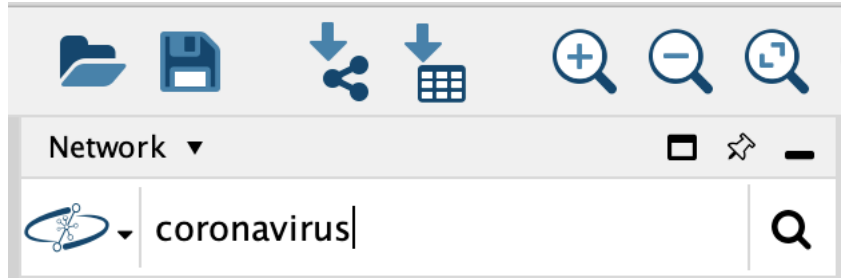
Hide selected nodes and edges

Import Network/table from file

Select first neighbours of selected nodes

The screenshot shows the jActiveModules software interface. On the left is a sidebar with 'jActiveModules' and 'Style' sections. The 'Network' section is active, showing a list of networks: 'sars', '1 of 1 Network selected', and 'IMEx/IntAct Coronavirus Dataset, full detail'. The 'Style' section shows 'IMEx/IntAct Coronavirus Dataset, full ... 1172 2212'. The main area displays a network graph with blue nodes and green edges. A toolbar at the top contains various icons for file operations, zooming, layout, and visibility. Red arrows point from text labels to these icons: 'Open or save network to NDex' points to the network icon; 'Open a session' points to the folder icon; 'Save your session' points to the save icon; 'Zoom in/out' points to the zoom icons; 'Apply preferred layout' points to the refresh icon; 'Hide selected nodes and edges' points to the eye icon; 'Import Network/table from file' points to the import icon; and 'Select first neighbours of selected nodes' points to the first neighbours icon.

Load a Network



Enter a search term into Network search bar. I am using Ndex- (<https://home.ndexbio.org/index/>)

Find Networks

NDEX Click on download arrow to load network into Cytoscape Sign in

coronavirus

Results


	name	owner	visibility	nodes	edges	modified
⬇	SARS CORONAVIRUS AND INNATE IMMUNITY	wikipathways	PUBLIC	30	0	2021-06-28, 9:10 p.m.
⬇	IMEx/IntAct Coronavirus Dataset, mutations de...	pporras	PUBLIC	2233	4477	2020-09-05, 5:19 p.m.
⬇	IMEx/IntAct Coronavirus Dataset, full detail	pporras	PUBLIC	2233	4477	2020-09-05, 5:19 p.m.
⬇	Coronavirus Proteases Sequence Similarity Net...	rudipillich	PUBLIC	57	341	2020-08-24, 1:14 p.m.
⬇	IMEx / IntAct Coronavirus Dataset, collapsed e...	pporras	PUBLIC	2233	3735	2020-09-05, 5:20 p.m.
⬇	IMEx/IntAct Coronavirus Dataset, binding regi...	pporras	PUBLIC	2233	4477	2020-09-05, 5:20 p.m.
⬇	Common cellular processes important in coro...	wikipathways	PUBLIC	56	0	2021-06-28, 8:07 p.m.
⬇	WP4912 - SARS coronavirus and innate immuni...	wikipathways	PUBLIC	74	20	2021-12-13, 7:16 p.m.
⬇	IMEx/IntAct Coronavirus Dataset: SARS-CoV-1...	pporras	PUBLIC	892	1523	2020-08-25, 10:41 a.m.
⬇	BioGRID Curation Projects: Coronavirus - Inter...	rudipillich	PUBLIC	4983	16762	2021-02-05, 7:02 p.m.
⬇	IMEx/IntAct Coronavirus Dataset: SARS-CoV-2...	pporras	PUBLIC	1583	2449	2020-08-25, 10:40 a.m.
⬇	WP4853 - Linoleic acid metabolism affected by ...	wikipathways	PUBLIC	52	20	2021-12-13, 7:12 p.m.

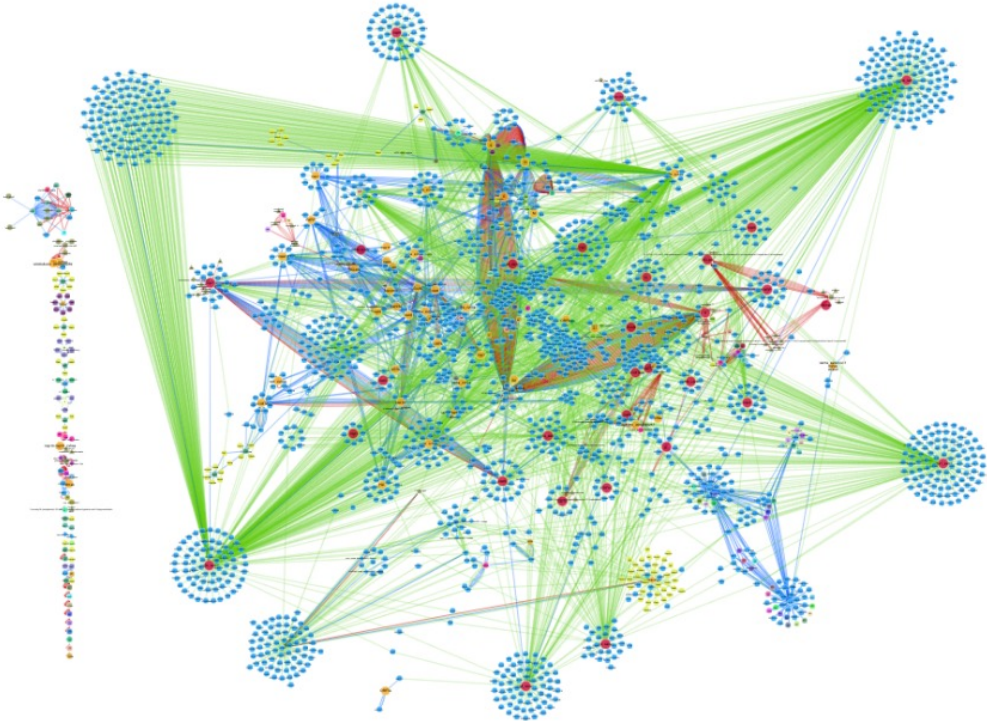
WARNING: In some cases, not all network information stored in NDEX will be available within Cytoscape after loading.

Close Dialog

About the network

<http://ndexbio.org/#/network/3e13a4de-a4d7-11ea-aaef-0ac135e8bacf>

About Docs Report Bug Contact Us FAQSearchLogin/Register



Network Info Nodes/Edges

IMEx/IntAct Coronavirus Dataset, full detail

DOI: [10.18119/N9RC8F](https://doi.org/10.18119/N9RC8F)

Nodes: 2233 **Edges:** 4477

PUBLIC [Q](#)

@context: [view namespaces](#)

Owner: Pablo Porras Millan

Created: Jun 2, 2020 9:44:55 AM

Last Modified: Sep 5, 2020 5:19:46 PM

UUID: 3e13a4de-a4d7-11ea-aaef-0ac135e8bacf

Description: Full Coronavirus dataset as extracted from IntAct on July 2020. Each edge represents a separate piece of evidence for an interaction, with edge colour representing different interaction types according to the PSI-MI Controlled Vocabulary (green = 'association' [MI:0914]; blue = 'physical association' [MI:0915]; red = 'direct interaction' [MI:0407] and children, including enzymatic reactions; grey = 'colocalization' [MI:0403]). Dashed lines represent spoke-expanded interactions. Node colour represents different species origin and node shape different molecule types. SARS-CoV-2 and SARS-CoV-1 proteins are represented as larger nodes for ease of identification. A short description of the dataset can be found here: <https://www.ebi.ac.uk/intact/resources/datasets>. Browse and download detailed interaction data on the IntAct website at: <https://www.ebi.ac.uk/intact/query/annot:%22dataset:coronavirus%22>

Rights Holder: IMEx Consortium
Rights: Attribution 4.0 International (CC BY 4.0)

Reference:
<https://doi.org/10.1101/2020.06.16.153817>










Version: 1.2

Type: [6](#) 1-step neighborhood [v](#) [Run Query](#)

[Open in Cytoscape](#) [Download](#) [Table](#) [Log in](#)









Visual Styles

Edge Stroke Color (Unselected) Mapping




Edge Stroke Color (Unselected)	int.type
	association
	cleavage reaction
	colocalization
	covalent binding
	direct interaction
	hydroxylation reaction
	physical association
	protein cleavage
	rna cleavage

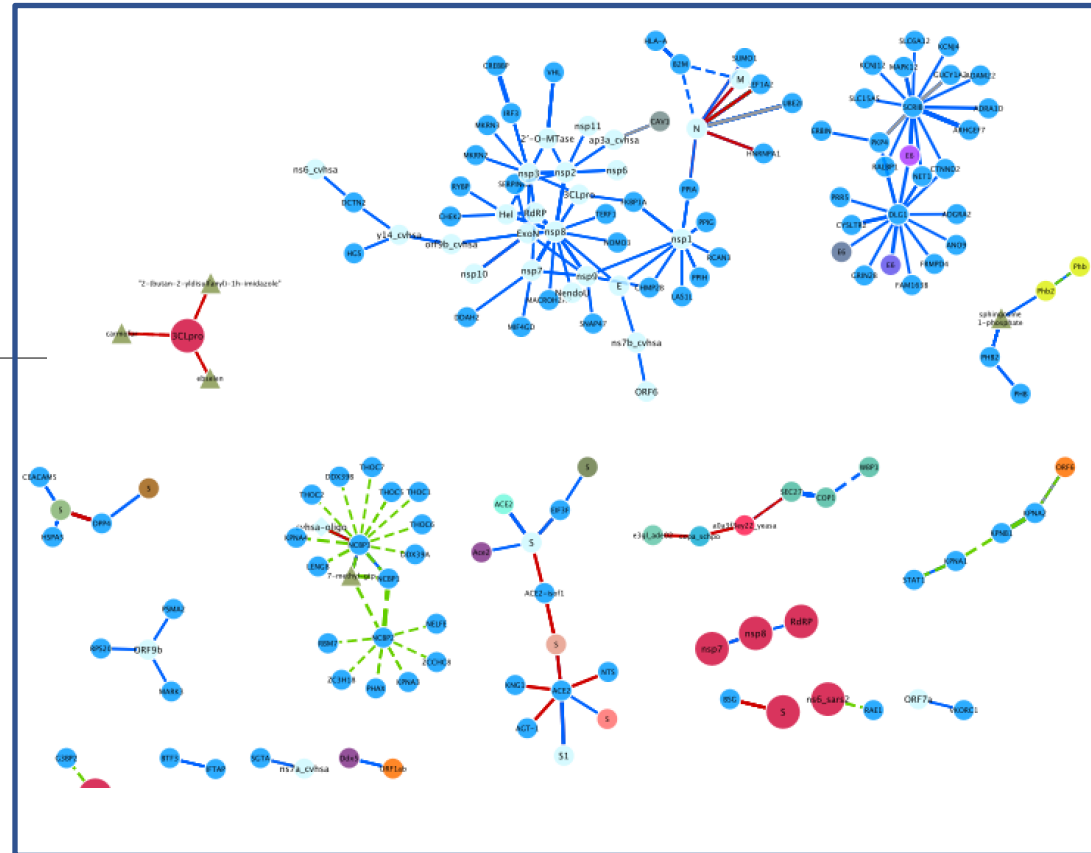
Node Properties

Node Shape Mapping

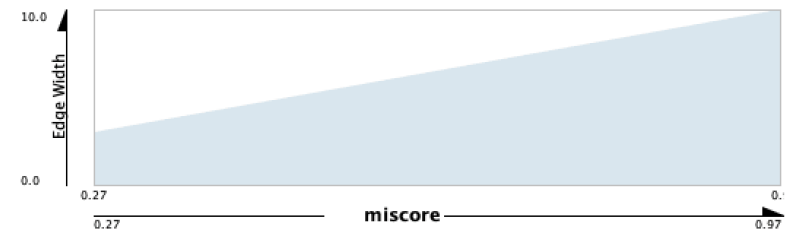
Node Shape	m.type
	messenger rna
	micro rna
	nucleic acid
	peptide
	protein
	ribonucleic acid
	small molecule
	small nuclear rna

Node Fill Color Mapping

Node Fill Color	species
	SARS-CoV-2
	human
	mouse
	rat

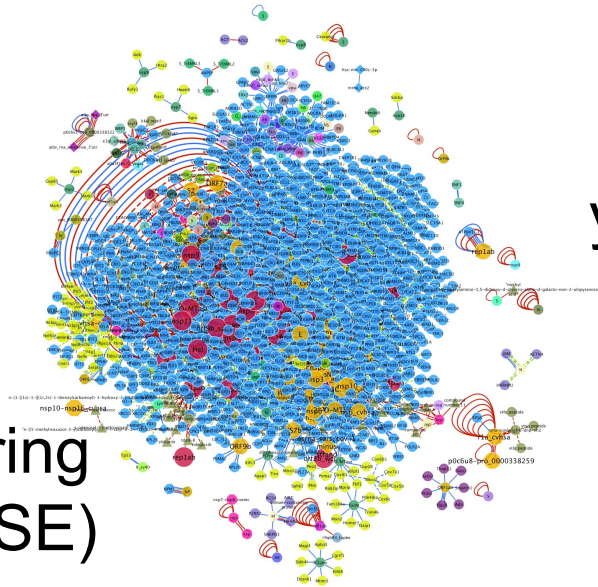


Edge Width Mapping

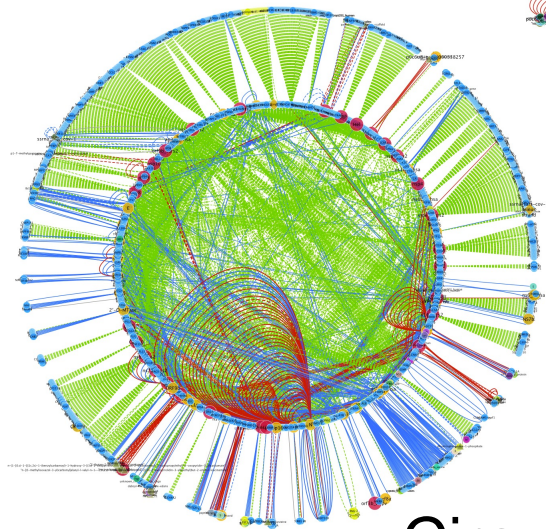
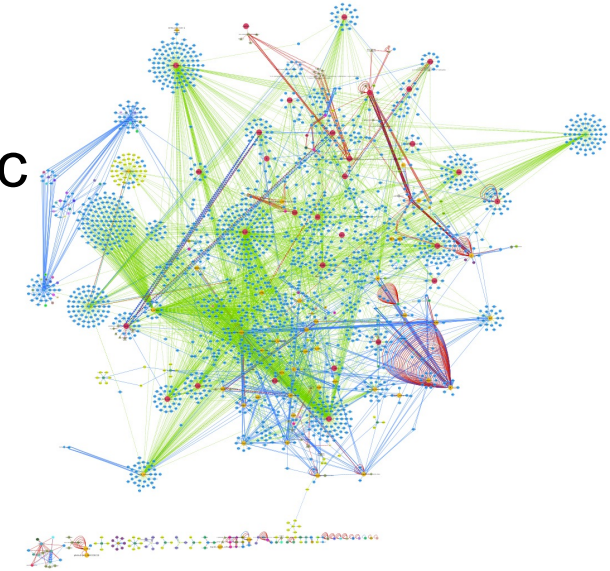


Experiment with different layouts

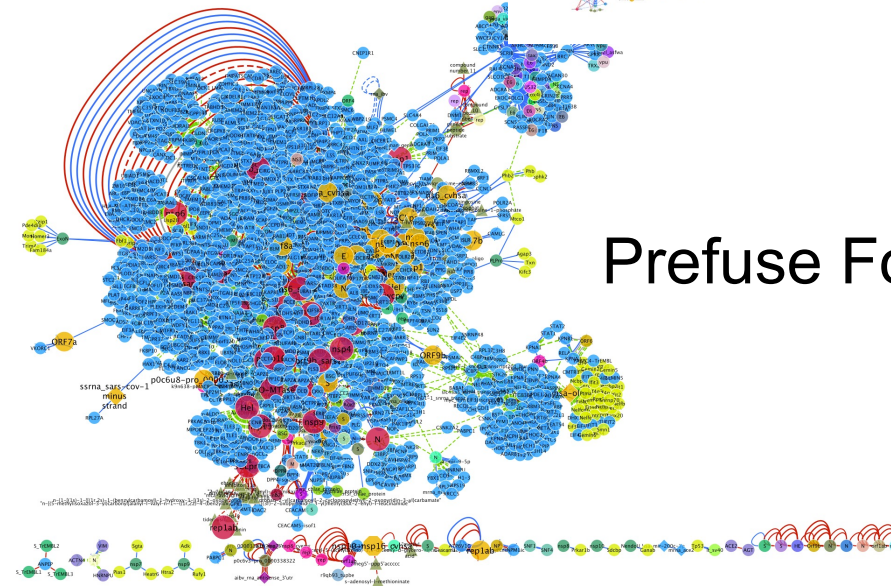
Compound Spring
Embedder (CoSE)



yfiles organic

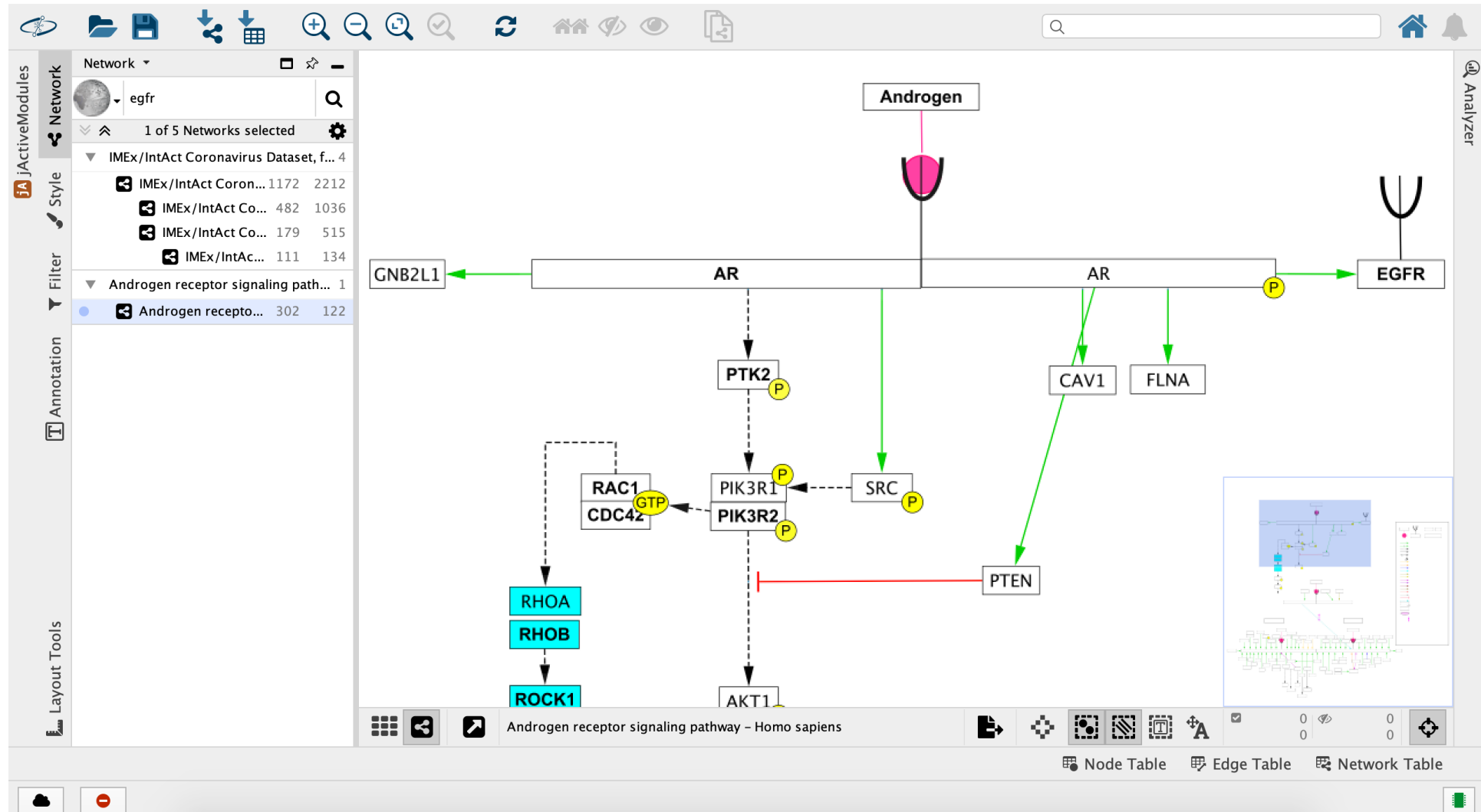


Circular Layout



Prefuse Force directed

Load different types of networks



Androgen Receptor Pathway from WikiPathways (WP138) =
<https://www.wikipathways.org/index.php/Pathway:WP138>

We are on a Coffee Break & Networking Session

Workshop Sponsors:



Canadian Centre for
Computational
Genomics



HPC4Health



Ontario
Genomics



GenomeCanada