

Canadian Bioinformatics Workshops

www.bioinformatics.ca

bioinformaticsdotca.github.io

Supported by



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 Suomi français français (CA) Galego עברית hrvatski Magyar Italiano 日本語 한국어 Macdonian Melayu
Nederlands Norsk Sesotho sa Leboa polski Português română slovenski jezik српски српски (latinica) Sotho svenska
中文 華語 (台灣) isiZulu



Attribution-Share Alike 2.5 Canada

You are free:



to Share — to copy, distribute and transmit the work



to Remix — to adapt the work



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](#)

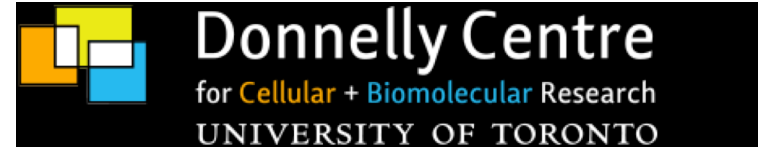
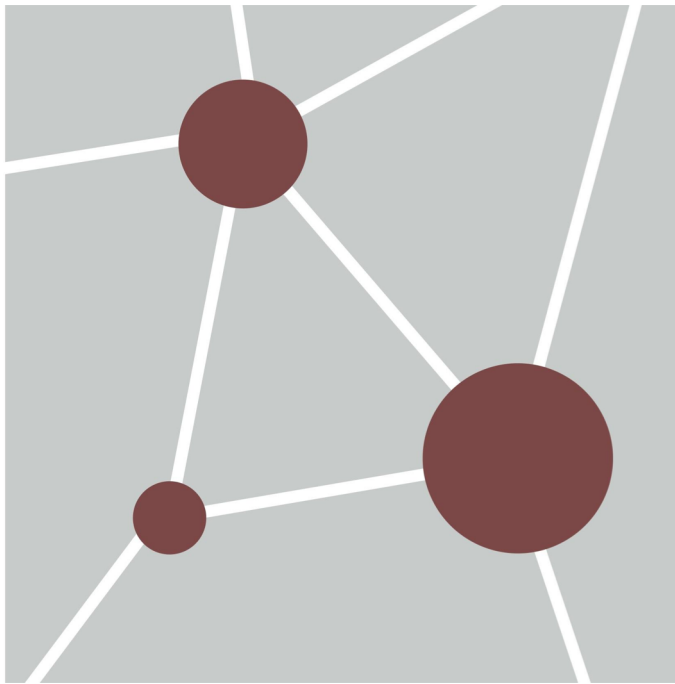
[Learn how to distribute your work using this licence](#)

Summary Slides

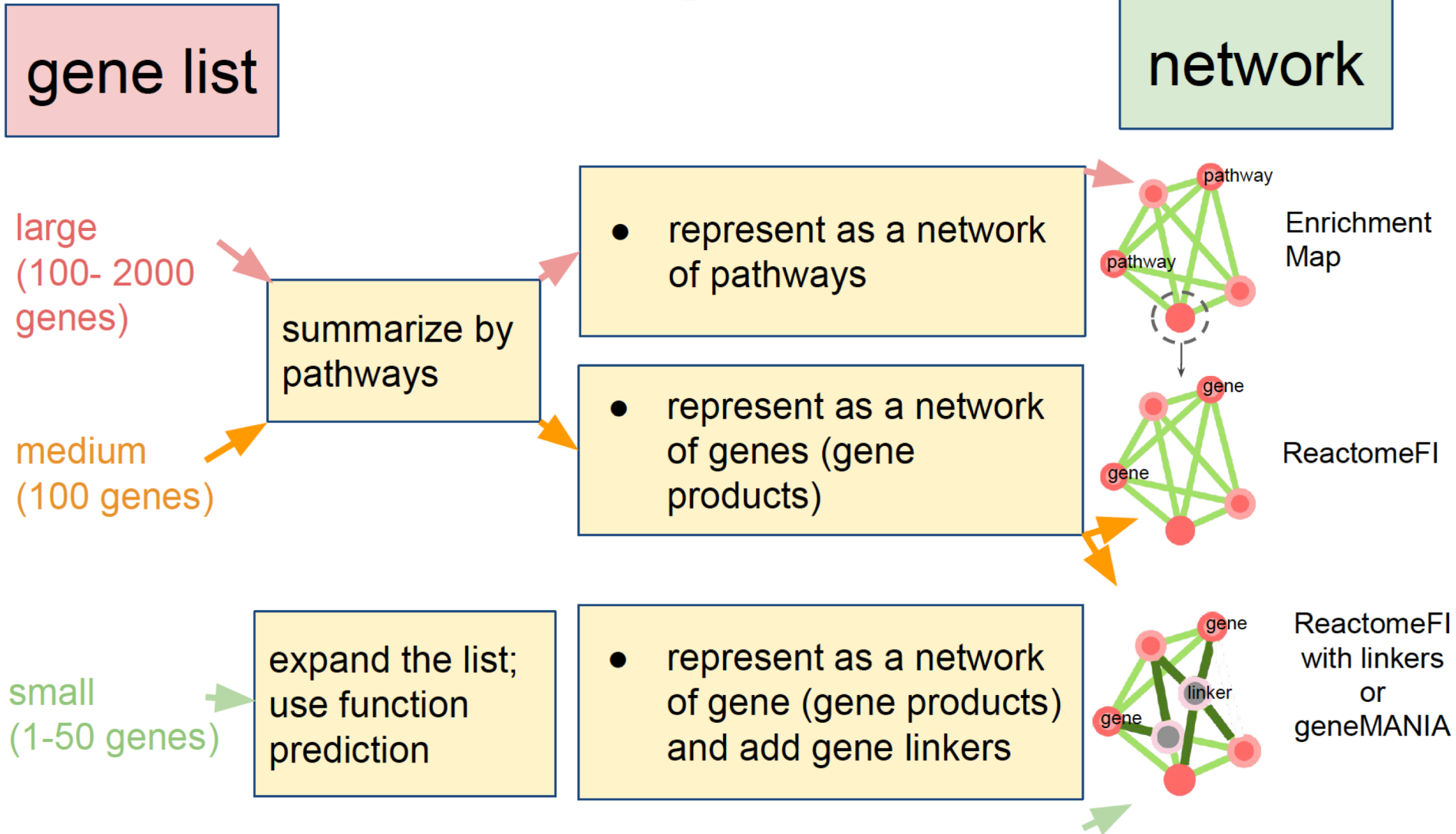
Veronique Voisin

Pathway and Network Analysis of -omics Data

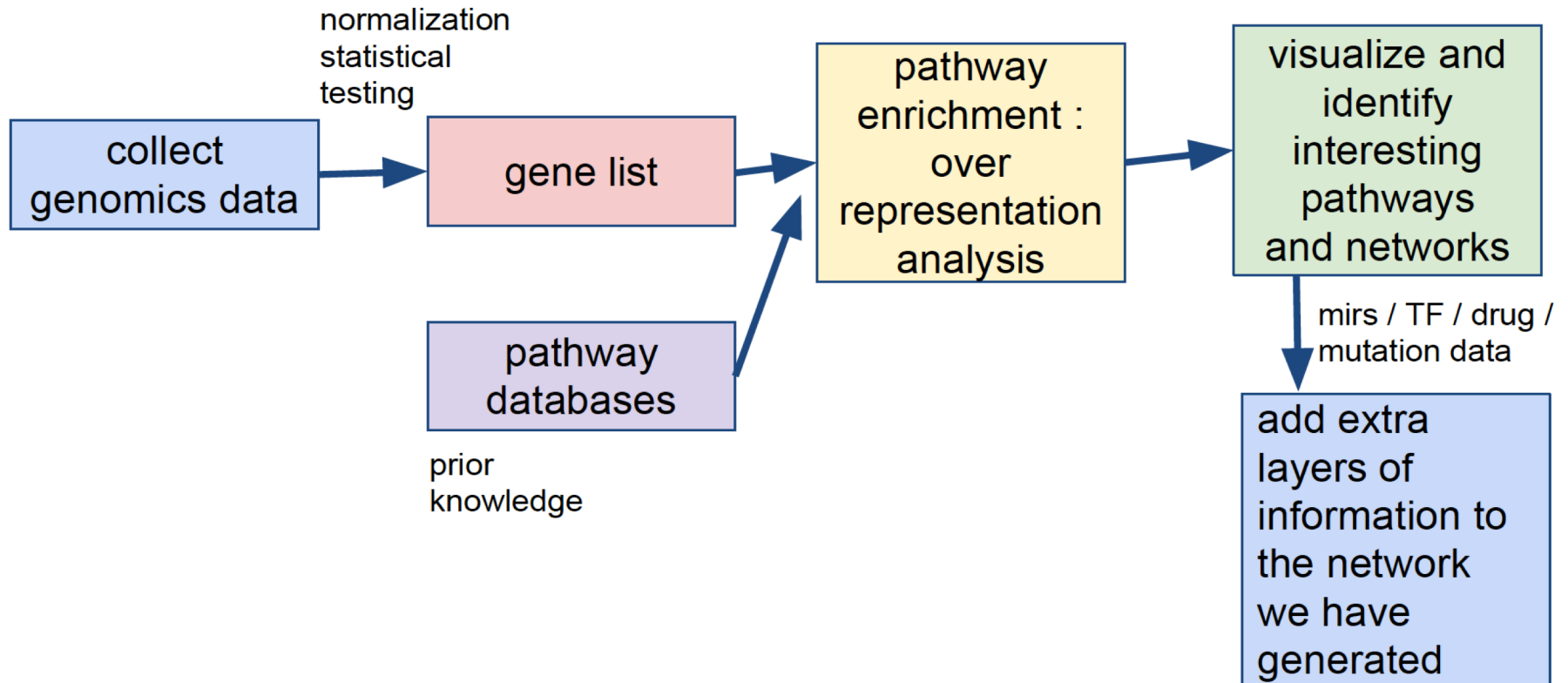
May, 10-12, 2021

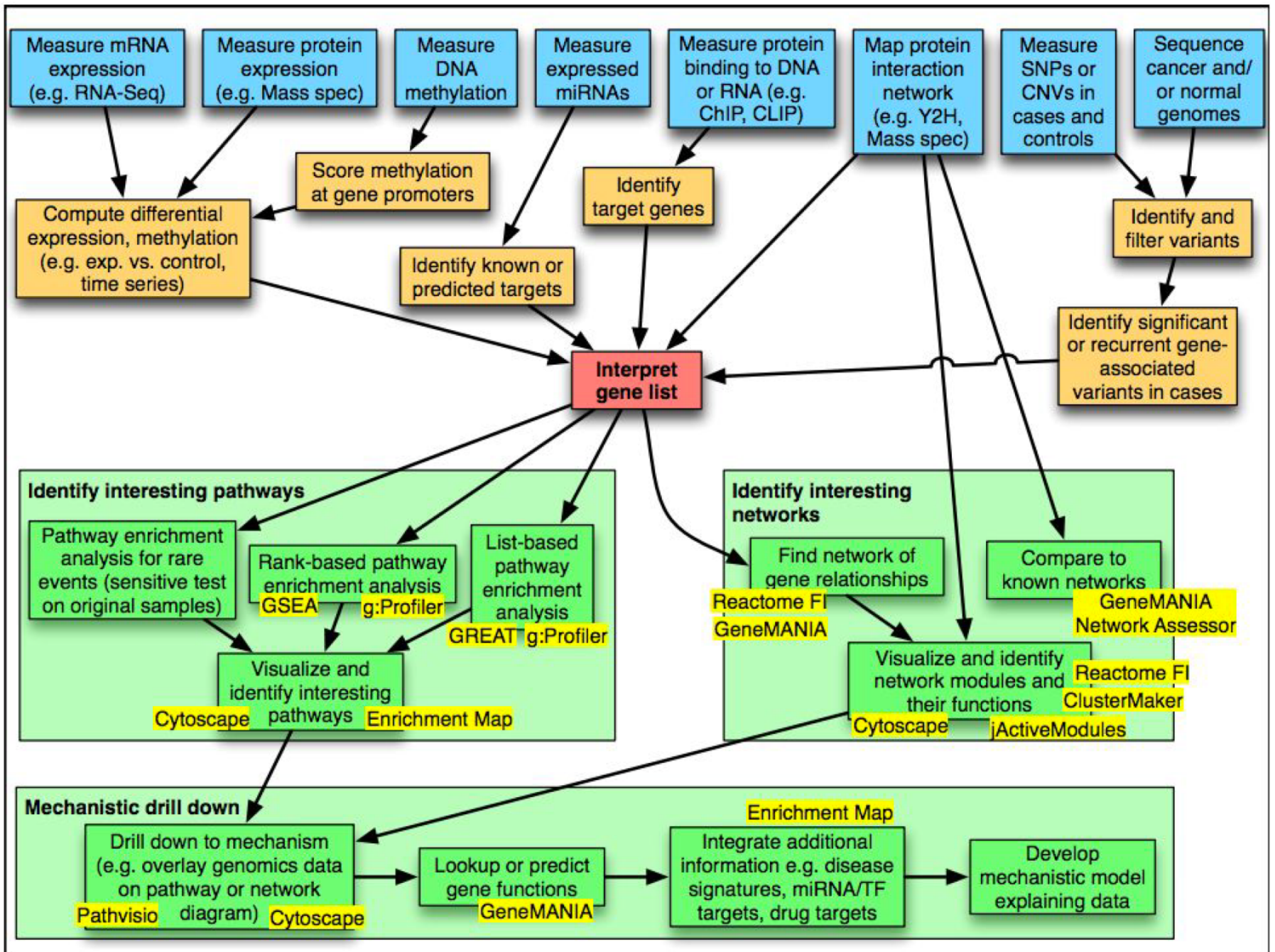


Creating Networks



Where are we in the workflow?





RNAseq

- Bulk RNAseq 2 class design:
 - GSEA
 - Enrichment Map

- Single cell Data:
 - GSEA
 - single sample GSEA **ssGSEA()**, **gsva()** **GSVA** in R
 - Wilcoxon Rank sum test (R, Panther)

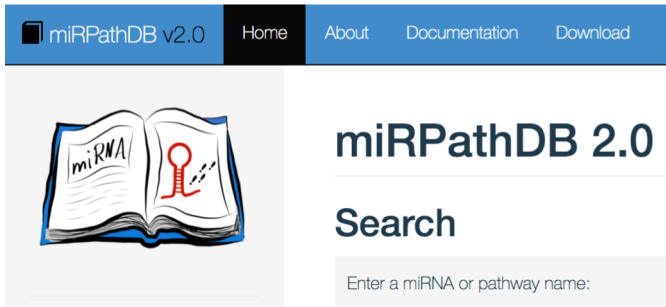
GWAS -- > MAGENTA

[https://software.broadinstitute.org/
mpg/magenta/](https://software.broadinstitute.org/mpg/magenta/)

The only **input** required is a table with variant association p-values and their chromosome positions taken from a genome-wide association study or meta-analysis. **Optional:** pathway/s or gene set/s of interest. Otherwise, we provide a set of pathways from public databases (see below).

The main **output** of MAGENTA is a nominal **gene set enrichment analysis (GSEA) p-value** and a **false discovery rate** for each gene set or pathway tested. There are several options, including running MAGENTA in the absence of a subset of genes, such as a predefined set of disease or trait genes. Additional information is provided, such as the expected and observed number of genes above the enrichment cutoff, and the number and name of genes in each tested gene set that lie near validated disease or trait SNPs if inputted by the user.

Mirs, Pathways and Targets



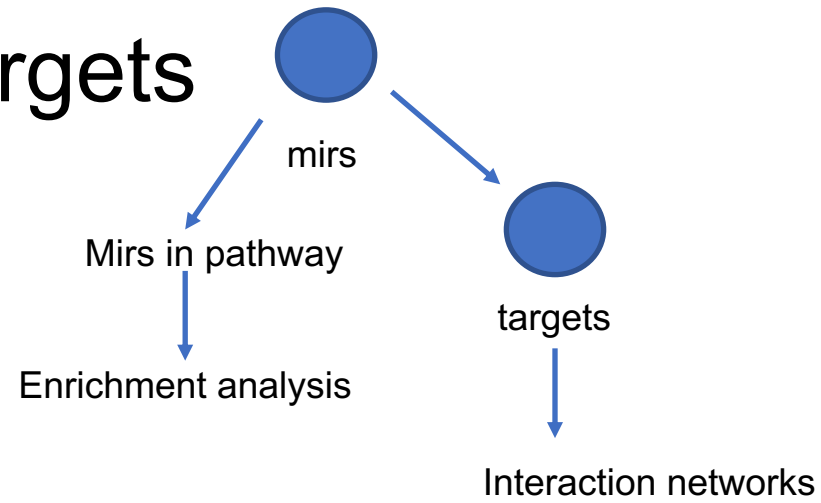
Fc gamma R-mediated phagocytosis

miRNAs that are significantly enriched for this pathway

In this table miRNAs are depicted that have significantly more targets in this pathway than expected by chance.

Show 10 entries

Database	miRNA	Evidence	Hits	Expected hits	P-value	Targets
miRBase	hsa-miR-126-3p	experimental (any)	7	0.261231	3.34e-8	AKT1, AKT2, CRK, CRKL, PKCGB, PKC
miRBase	hsa-miR-184	experimental (any)	5	0.297456	1.61e-4	AKT1, AKT2, INPPL1, PLFPG, PRKCB
miFCarta	m-5765	predicted (union)	56	36.0553	3.01e-4	AKT2, AMFH, APPC2, APPC3, APPC4
miFCarta	m-17942	predicted (intersection)	12	2.26763	3.86e-4	APPC2, CRKL, LAM, MAPK1, NDF1, PIP
miFCarta	m-152	predicted (union)	57	38.8709	4.14e-4	AKT2, AMFH, APPC1B, APPC2, APPC
miFCarta	m-12614	predicted (union)	55	36.7072	4.47e-4	AKT2, APPC1B, APPC2, APPC4, APP
miRBase	hsa-miR-184	experimental (strong)	5	0.336283	4.48e-4	AKT1, AKT2, INPPL1, PLFPG, PRKCB
miRBase	hsa-miR-550a-3p	experimental (strong)	2	0.0353982	6.65e-4	MAPK1, MAPK3



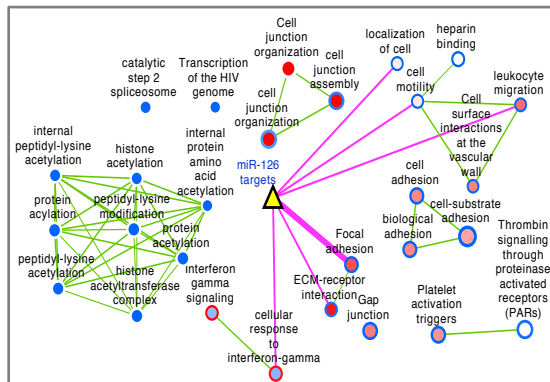
miEAA: microRNA enrichment analysis and annotation

http://www.ccb.uni-saarland.de/mieaa_tool/

<http://www.lirmed.com/tam2/>

EnrichmentMap
Post analysis
Mir targets

Post analysis:
Good for drug target
and transcription
factors



Result

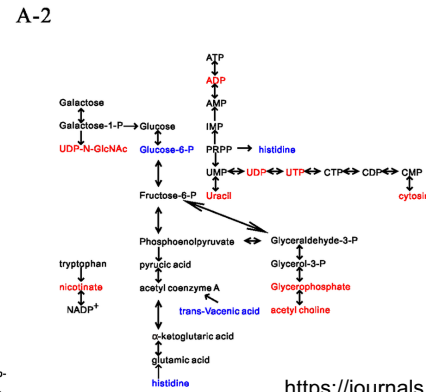
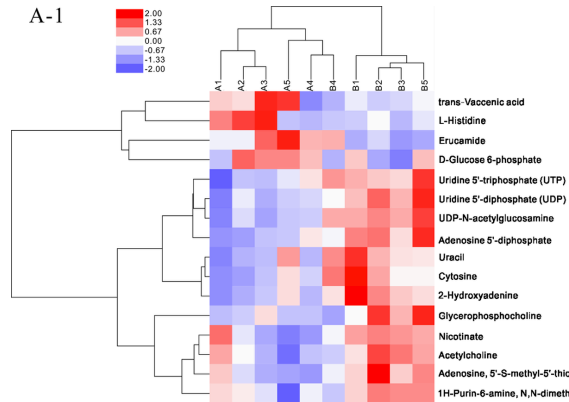
Enrichment analysis results

Text file of results Results Visualization

Term	Count	Percent	Fold	P-value	Bonferroni	FDR
Category: Cluster (4 Items)						
hsa-mir-106b cluster [details]	1	0.33333	33.63889	0.0295	1	0.3755
hsa-mir-17 cluster [details]	2	0.33333	33.63889	1.32e-3	0.3569	0.08
hsa-mir-423 cluster [details]	1	0.5	50.45833	0.0197	1	0.3365
hsa-mir-6081 cluster [details]	1	0.2	20.18333	0.0487	1	0.479
Category: Disease (194 Items)						
Acute Cerebral Infarction [details]	1	0.16667	16.81944	0.0581	1	0.5292
Acute Ischemic Stroke [details]	2	0.14286	14.41667	7.67e-3	1	0.1858
Acute Myocardial Infarction [details]	2	0.04348	4.38768	0.0731	1	0.5944
Acute Pancreatitis [details]	1	0.14286	14.41667	0.0675	1	0.5676
Adenocarcinoma, Colon [details]	2	0.08696	8.77536	0.0203	1	0.2926
Adenocarcinoma, Esophageal [details]	1	0.04545	4.58712	0.1983	1	1
Adenocarcinoma, Gastric [details]	1	0.02632	2.6557	0.3191	1	1
Adenocarcinoma, Lung [details]	2	0.0198	1.99835	0.2642	1	1
Adrenal Cortex Neoplasms [details]	1	0.08333	8.40972	0.1131	1	0.7828

Metabolomics

A) list of metabolites from your experiment

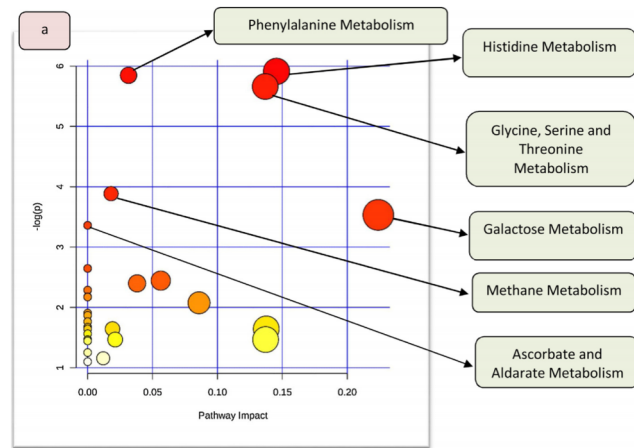


<https://journals.plos.org/plosone/article/figure?id=10.1371/journal.pone.0196231.g007>

B) pathway: a set of metabolites known to be involved in metabolic pathways

The colored compounds indicate potential matches from the user's input, with red colors indicating significant hits and blue colors indicating non-significant hits.

Pathway	Metabolites
Tryptophan metabolism	C00025; C00024; C00027; C00026; C00021; C00020; CE2119; C00028; C03722; C05647; C05645; CE1395; C05643; C05640; C00780; C05648; C00704; C01342; C00010; C00014; C00016; C00019; C15605; C00067; thbpt4acam; C05651; C03512; C05653; C02693; C05660; C00398; CE5982; C00643; C02220; C00078; C00978; C00877; C05642; CE5860; C01252; C05637; C02406; C00108; C00272; CE1918; C01652; C02470; C01144; C06212; C06213; CE2949; CE2948; CE3140; CE2122; C00479; CE2947; C01717; CE3092; CE6205; CE5899; C03161; C00268; C00632; CE2152; CE2153; C02700; CE2095; C00328; CE3087; CE3086; C00322; C04409; C01352; C00051; C00058; C00030; C00331; C00332; C00936; C03824; C05636; C05635; C05634; C05639; C05638; C10164; C00637; C03227; C01598; C00525; C00527; C00041; C03024; C00954

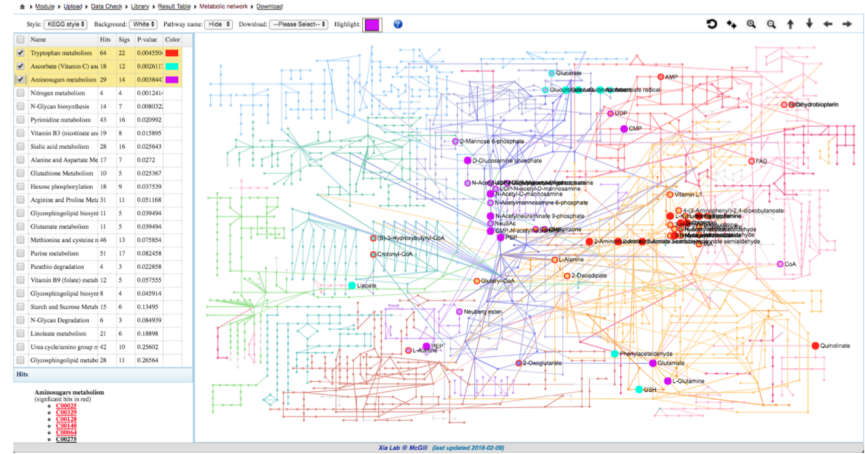


<https://sciex.com/applications/life-science-research/metabolomics/metabolomics-pathway-analysis>

Explore Results in Network

Pathway Name	Total	Hits (all)	Hits (sig.)	Fisher's P-value	EASE Score	Gamma P-value	Match Details
Tryptophan metabolism	94	64	21	0.0045504	0.0096086	0.0046682	View
Ascorbate (Vitamin C) and Aldarate Metabolism	29	18	9	0.0026117	0.010691	0.0046835	View
Aminosugars metabolism	69	29	12	0.0038443	0.011655	0.0047003	View
Nitrogen metabolism	6	4	4	0.0012414	0.022604	0.0048951	View
N-Glycan biosynthesis	48	14	7	0.0080322	0.032406	0.0050767	View
Pyrimidine metabolism	70	43	14	0.020992	0.045203	0.0053243	View
Vitamin B3 (nicotinate and nicotinamide) metabolism	28	19	8	0.015895	0.049225	0.0054047	View
Sialic acid metabolism	107	28	10	0.025643	0.062907	0.0056878	View
Alanine and Aspartate Metabolism	30	17	7	0.0272	0.08133	0.0060936	View
Glutathione Metabolism	19	10	5	0.025367	0.099714	0.0065288	View
Hexose phosphorylation	20	18	7	0.037539	0.10359	0.0066246	View
Arginine and Proline Metabolism	45	31	10	0.051168	0.10984	0.0067823	View
Glycosphingolipid biosynthesis - ganglioseries	62	11	5	0.039494	0.13447	0.0074426	View
Glutamate metabolism	15	11	5	0.039494	0.13447	0.0074426	View
Methionine and cysteine metabolism	94	46	13	0.075854	0.13715	0.0075184	View
Purine metabolism	80	51	14	0.082458	0.1436	0.007704	View
Parathio degradation	6	4	3	0.022858	0.16222	0.0082678	View
Vitamin B9 (folate) metabolism	33	12	5	0.057555	0.17336	0.0086255	View
Glycosphingolipid biosynthesis - globoseries	16	8	4	0.045914	0.17647	0.0087285	View
Starch and Sucrose Metabolism	33	15	5	0.13495	0.30626	0.014409	View

Download Result Tables: [Pathway Hits](#) [Compound Hits](#)



List of metabolites: Fishers' exact test
 If it is possible to rank all the metabolites: GSEA

Pathways

[LIPID MAPS® Wiki Pathways](#)

[Structure Tables](#)

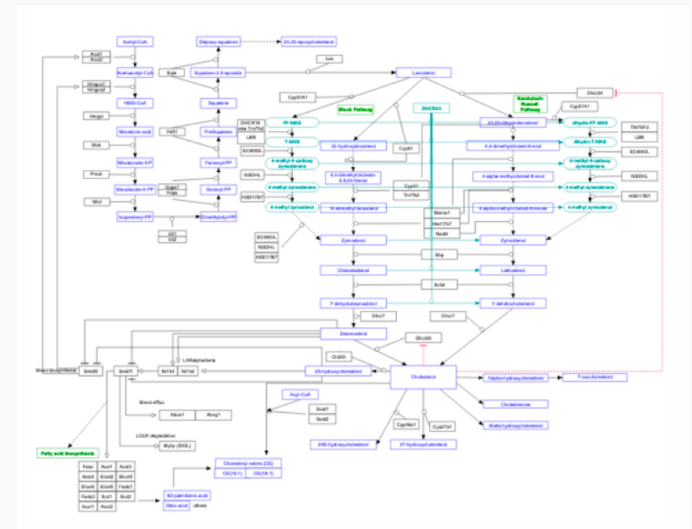
[SphinGOMAP®](#)

[External Resources](#)

LIPID MAPS® WikiPathways

LIPID MAPS® has contributed a set of 10 pathways to the WikiPathways project. Pathways include metabolism of

- cholesterol
- eicosanoids
- glycerolipids
- omega fatty acids
- sphingolipids



<https://www.lipidmaps.org/resources/pathways/index.php>

ATACseq / CHIPseq

- GREAT (Stanford) is a recommended tool (from chromosomal position to gene + enrichment analysis)
- It is compatible with EnrichmentMap

GREAT Overview News Use GREAT Demo Video How to Cite Help Forum

GREAT version 4.0.4 current (08/19/2019 to now)

GREAT predicts functions of cis-regulatory regions.

Many coding genes are well annotated with their biological functions. Non-coding regions typically lack such annotation. GREAT assigns biological meaning to a set of non-coding genomic regions by analyzing the annotations of the nearby genes. Thus, it is particularly useful in studying cis functions of sets of non-coding genomic regions. Cis-regulatory regions can be identified via both experimental methods (e.g. ChIP-seq) and by computational methods (e.g. [comparative genomics](#)). For more see our [Nature Biotech Paper](#).

News

- Aug. 19, 2019: GREAT version 4 [adds support for human hg38 assembly and updates ontology datasets for all supported assemblies](#).
- Sep. 8, 2018: GREAT has served over 1 million job submissions.
- Oct. 23, 2017: GREAT is moved to a VM to eliminate proxy errors.
- June 22, 2017: GREAT hardware upgrade to meet increasing submission volume.
- Nov. 16, 2015: The [GREAT user help forums](#) are frozen.
- Feb. 15, 2015: GREAT version 3 [switches to Ensembl genes, adds support for zebrafish danRer7 and mouse mm10 assemblies, and adds new ontologies](#).
- Apr. 3, 2012: GREAT version 2 [adds new annotations to human and mouse ontologies and visualization tools for data exploration](#).
- Feb. 18, 2012: The [GREAT user help forums](#) are opened.
- May 2, 2010: GREAT version 1 is launched, concurrent to [Nature Biotechnology publication \(reprint, Faculty of 1000 "Must Read"\)](#). [How to Cite GREAT?](#)

[More news items...](#)

Tips:

Do

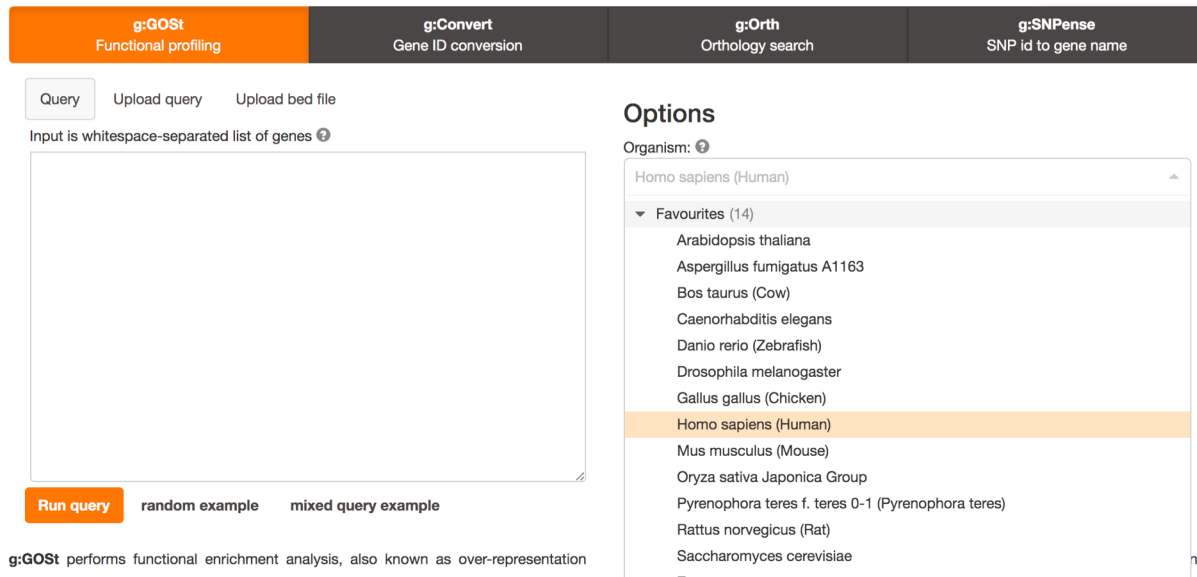
1) Proximal analysis (+-2kb around TSS of genes)

And

2) Distal analysis (+-50kb around genes, filter genomic regions using tools like Segway or BEHST)

Non model organisms

1. Find a pathway database/ gmt file which is the closest to your organism
2. Convert your gene identifier to the gene identifiers used in the gmt file that you found using g:Convert and g:Orth
3. Both GSEA and g:Profiler accept custom gmt file.
4. GeneMANIA offers several organisms and the option to build your interaction networks.



The screenshot displays the g:GOST web interface. At the top, there are four tabs: g:GOST (Functional profiling), g:Convert (Gene ID conversion), g:Orth (Orthology search), and g:SNPense (SNP id to gene name). The g:GOST tab is active. Below the tabs, there are buttons for 'Query', 'Upload query', and 'Upload bed file'. A text input field is labeled 'Input is whitespace-separated list of genes'. Below the input field, there are buttons for 'Run query', 'random example', and 'mixed query example'. To the right, the 'Options' section is visible, showing a dropdown menu for 'Organism:'. The dropdown is open, showing a list of organisms with 'Homo sapiens (Human)' selected and highlighted in orange. Other organisms listed include Arabidopsis thaliana, Aspergillus fumigatus A1163, Bos taurus (Cow), Caenorhabditis elegans, Danio rerio (Zebrafish), Drosophila melanogaster, Gallus gallus (Chicken), Mus musculus (Mouse), Oryza sativa Japonica Group, Pyrenophora teres f. teres 0-1 (Pyrenophora teres), Rattus norvegicus (Rat), and Saccharomyces cerevisiae.

Script to create a gmt file from the GO ontology:

<https://www.dropbox.com/s/wm3kq4lsdlfwcoq/creategmt.R?dl=0>

The Cytoscape App Store

 cytoscape app store




















Submit an App ▾

Search the App Store

Sign In

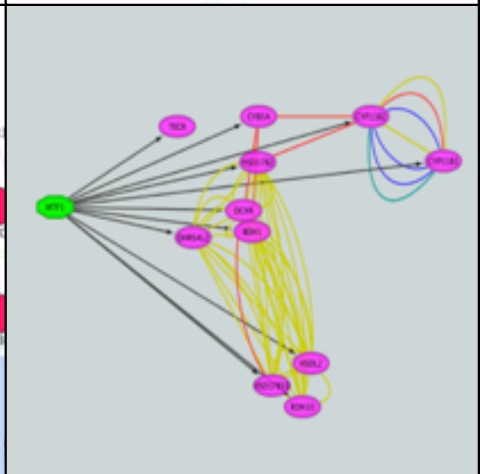
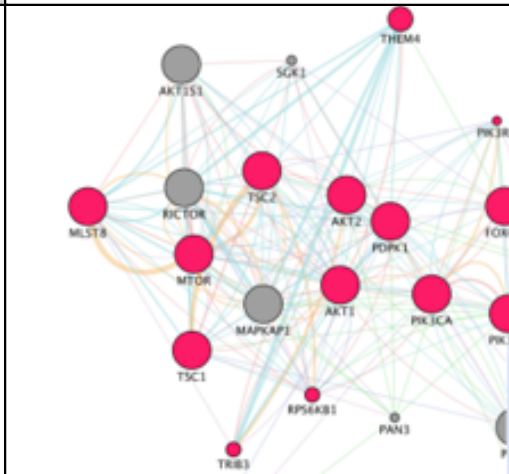
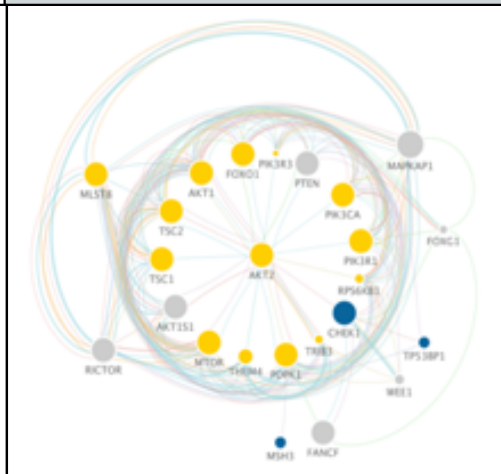
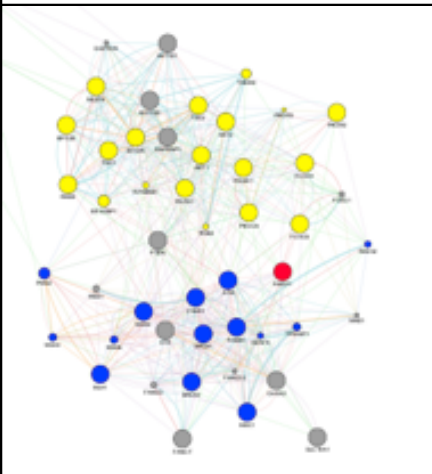
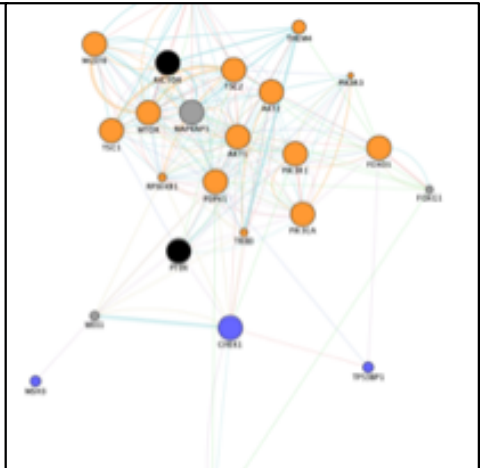
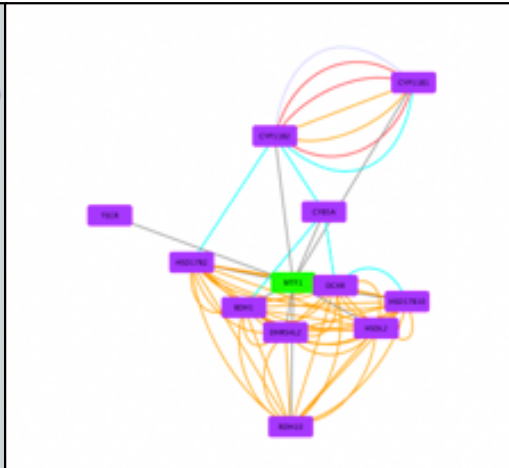
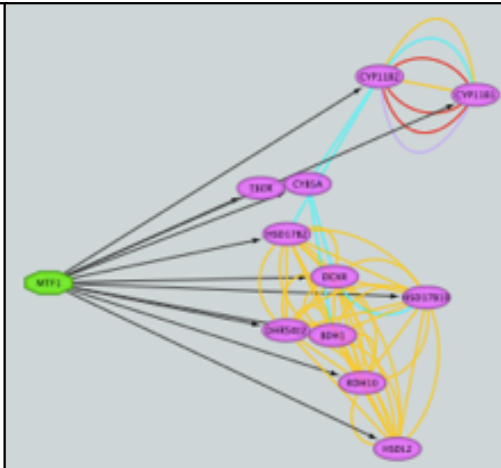
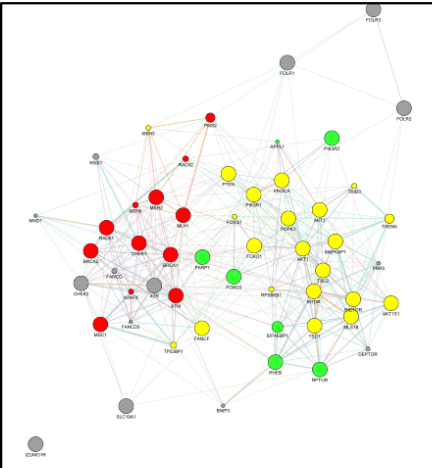
<http://apps.cytoscape.org>

Wall of Apps 184 total

network generation													
													
													
													
online data import													
													
													
													
graph analysis													

Pathway analysis
Gene expression analysis
Complex detection
Literature mining
Network motif search
Pathway comparison

Congratulations!!



We are on a Coffee Break & Networking Session

Workshop Sponsors:



Canadian Centre for
Computational
Genomics



HPC4Health



OICR
Ontario Institute
for Cancer Research

MiC:M McGill initiative in
Computational Medicine