IMIM Parc de Salut MAR Barcelona
Institut Hospital del Mar d'Investigacions Mèdiques

Functional analysis of omics data

february 2020

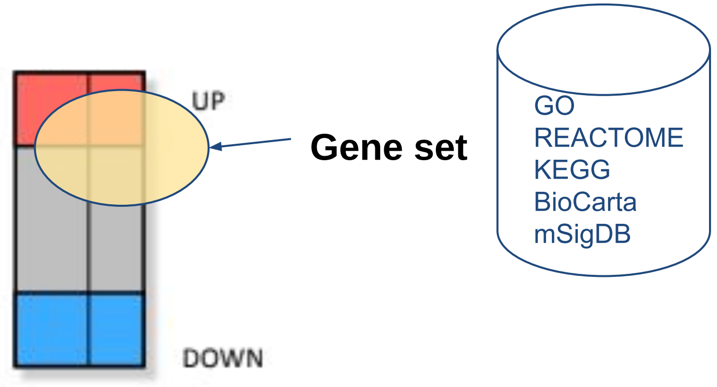
MARGenomics

Summary

1. Summary of previous days
2. Example datasets: He et al, Am J Transl Res, 2018
    a. GEO
    b. TCGA
3. Hands on
    a. Differential Expression Analysis of Public Data
    b. DAVID
    c. GSEA Preranked
    d. Cytoscape
    e. Public Resources to “validate” candidates

1. Summary of enrichment strategies

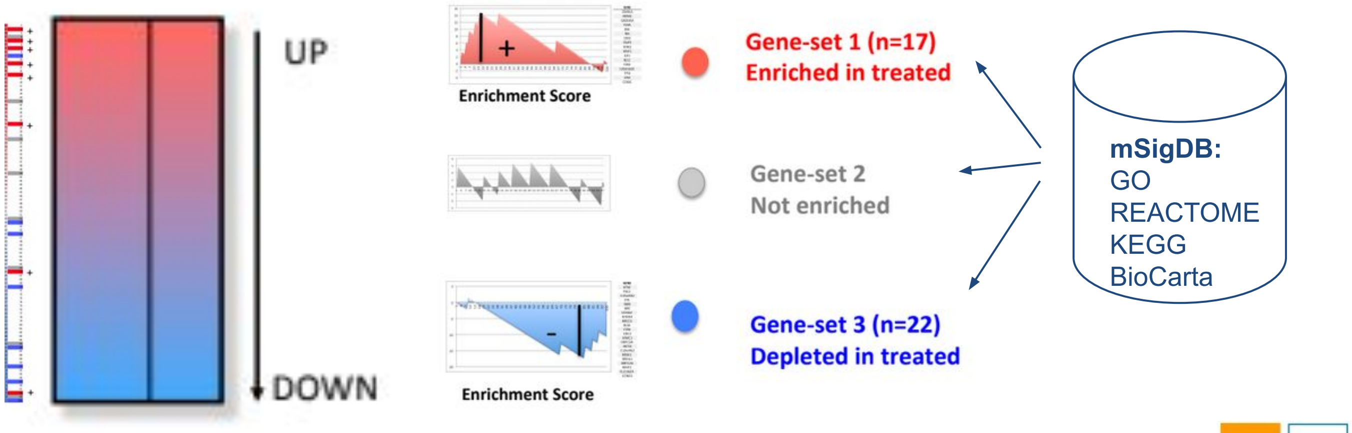
- Gene list (e.g. expression change > 2-fold AND FDR < 0.05)



Gene set

GO
REACTOME
KEGG
BioCarta
mSigDB

- Ranked list (e.g. by -log10(p)*sign(logFC))



Gene-set 1 (n=17)
Enriched in treated

Gene-set 2
Not enriched

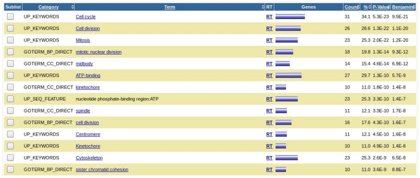
Gene-set 3 (n=22)
Depleted in treated

mSigDB:
GO
REACTOME
KEGG
BioCarta

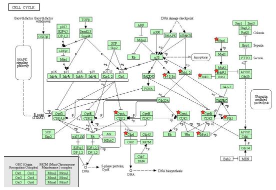
1. Summary of enrichment strategies

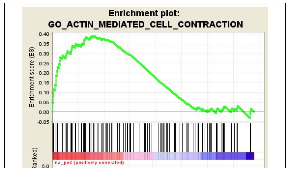
- Gene list (e.g. expression change > 2-fold AND FDR < 0.05)
  - **Answers the question**: Are any gene sets surprisingly enriched (or depleted) in my gene list?
  - **Statistical test**: Fisher's Exact Test (aka Hypergeometric test)
  - **Tools**: DAVID
  - **Benefits:** simple, you only need a list of gene names
  - **Problems**: Possible loss of statistical power due to thresholding. Different results at different threshold settings
- Ranked list (e.g. by -log10(p)*sign(logFC))
  - **Answers the question**: Are any gene sets ranked surprisingly high or low in my ranked list of genes?
  - **Statistical test**: GSEA
  - **Benefit**: use information of all genes measured. Increase of statistical power
  - **Problems**: more difficult to prepare files. You need the whole experiment

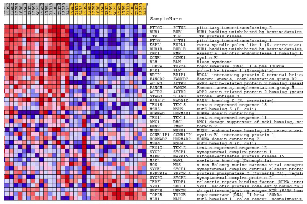
1. Summary of enrichment strategies

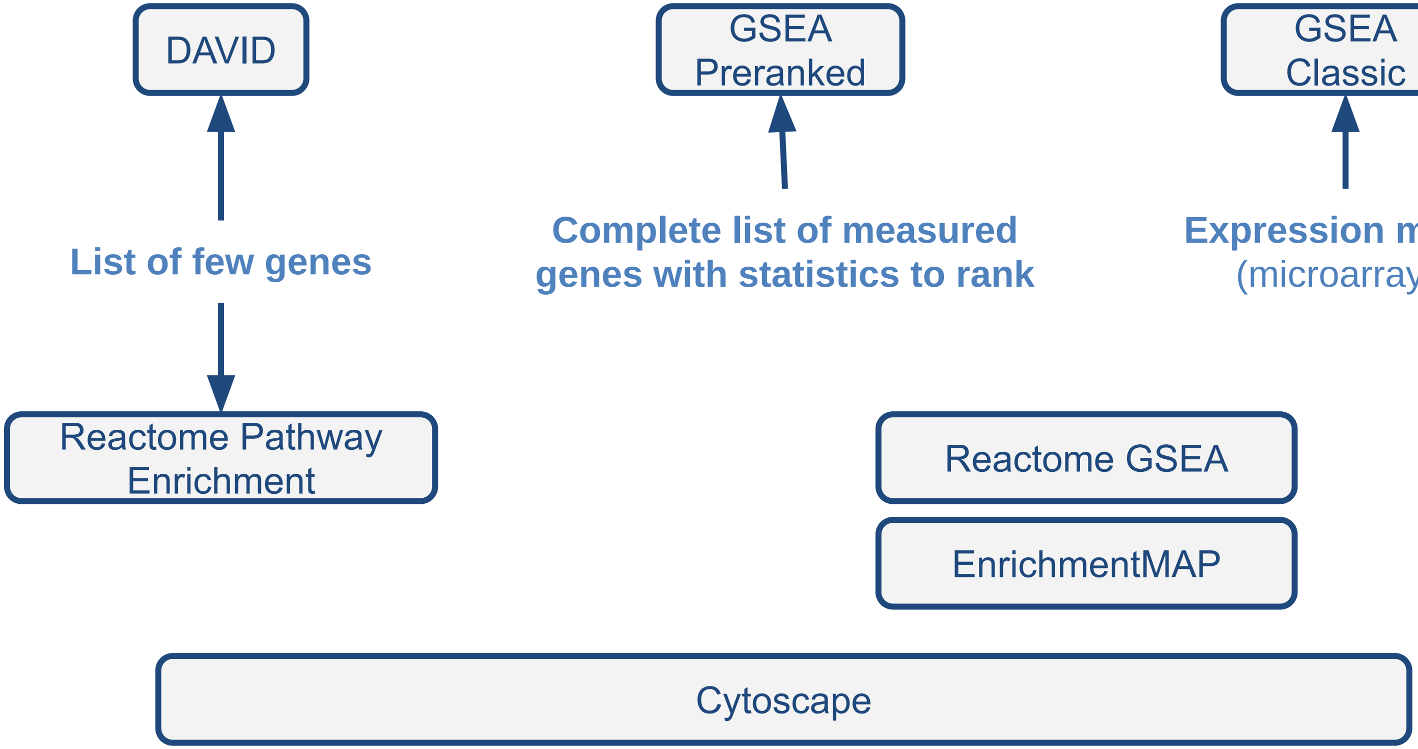


Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
□	UP_KEYWORDS	Cell cycle	RT		31	34.1	5.3E-23	9.5E-21
□	UP_KEYWORDS	Cell division	RT		26	28.6	1.3E-22	1.1E-20
□	UP_KEYWORDS	Mitosis	RT		23	25.3	2.0E-22	1.2E-20
□	GOTERM_BP_DIRECT	mitotic nuclear division	RT		18	19.8	1.3E-14	9.3E-12
□	GOTERM_CC_DIRECT	midbody	RT		14	15.4	4.6E-14	6.9E-12
□	UP_KEYWORDS	ATP-binding	RT		27	29.7	1.3E-10	5.7E-9
□	GOTERM_CC_DIRECT	kinetochore	RT		10	11.0	1.8E-10	1.4E-8
□	UP_SEQ_FEATURE	nucleotide phosphate-binding region:ATP	RT		23	25.3	3.3E-10	1.4E-7
□	GOTERM_CC_DIRECT	spindle	RT		11	12.1	3.3E-10	1.7E-8
□	GOTERM_BP_DIRECT	cell division	RT		16	17.6	4.3E-10	1.6E-7
□	UP_KEYWORDS	Centromere	RT		11	12.1	4.5E-10	1.6E-8
□	UP_KEYWORDS	Kinetochore	RT		10	11.0	4.9E-10	1.4E-8
□	UP_KEYWORDS	Cytoskeleton	RT		23	25.3	2.6E-9	6.5E-8
□	GOTERM_BP_DIRECT	sister chromatid cohesion	RT		10	11.0	3.6E-9	8.8E-7









DAVID

List of few genes

Reactome Pathway Enrichment

GSEA Preranked

Complete list of measured genes with statistics to rank

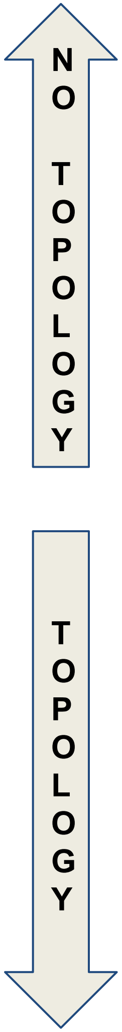
GSEA Classic

Expression matrix (microarrays)

Reactome GSEA

EnrichmentMAP

Cytoscape



1. Summary of public sources (open data)

Projects

ENCODE
TCGA
GTEx
Cancer Cell Line Encyclopedia
The Human Protein Atlas
FANTOM5
Researchers
Publications
...

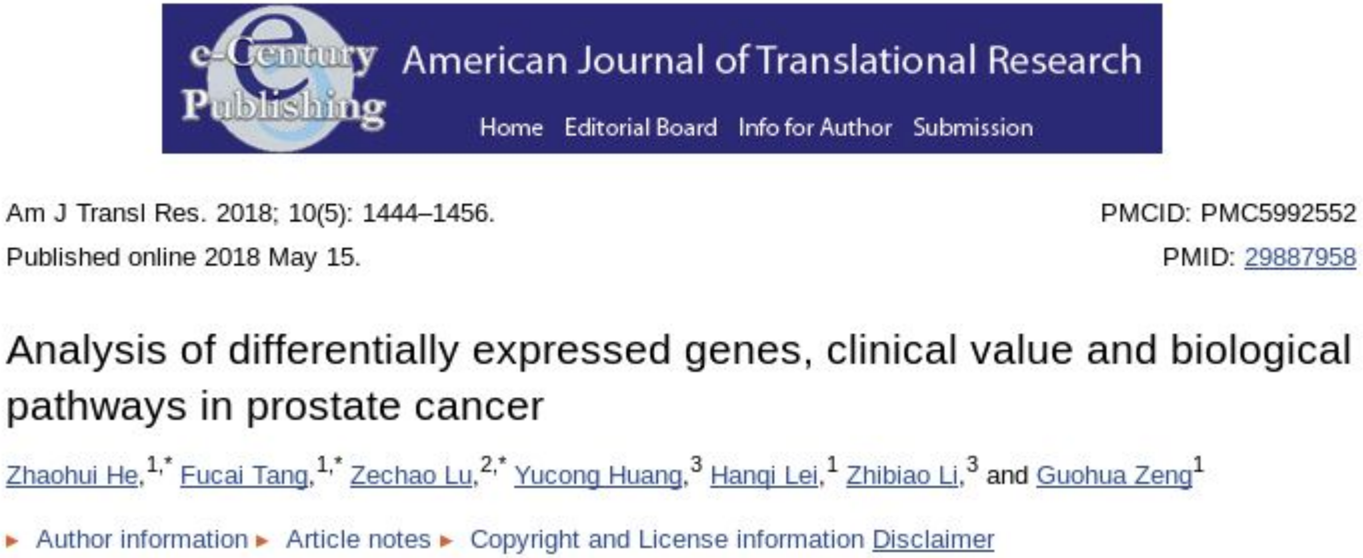
Web sources

GDAC iCGC
cBioPortal
GEPIA XENA
GEO SRA
dbGaP
...

Databases

GO
REACTOME
KEGG
BioCarta
mSigDB

2. Example datasets: He (2018) Am J Transl Res



Am J Transl Res. 2018; 10(5): 1444–1456.
Published online 2018 May 15.

PMCID: PMC5992552
PMID: 29887958

Analysis of differentially expressed genes, clinical value and biological pathways in prostate cancer

Zhaohui He,[1,*] Fucai Tang,[1,*] Zechao Lu,[2,*] Yucong Huang,[3] Hanqi Lei,[1] Zhibiao Li,[3] and Guohua Zeng[1]

▸ Author information ▸ Article notes ▸ Copyright and License information Disclaimer

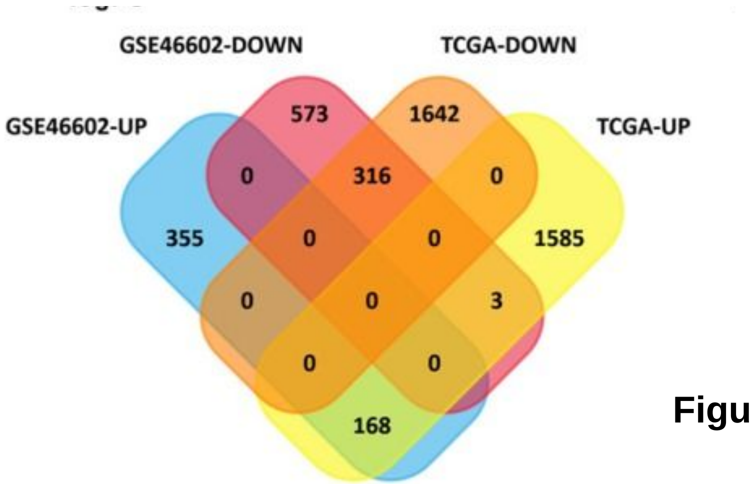


Figure 1

He_etal_2018.pdf

2. Example datasets: He (2018) Am J Transl Res

Figure 2. GO Enrichment Analysis

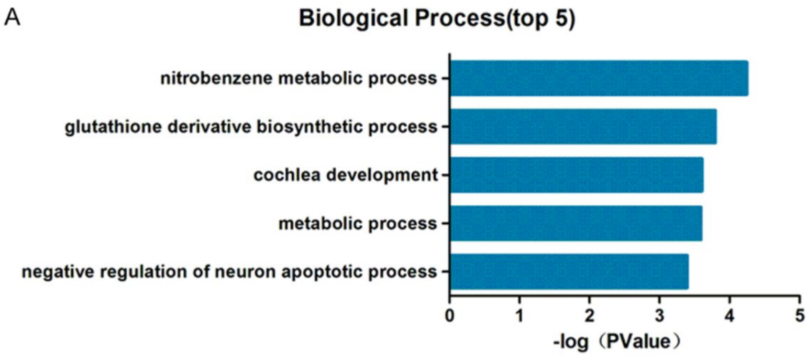


Figure 3. PPI network of DEGs

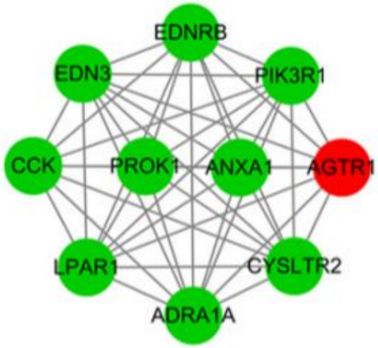


Figure 4. Validation of genes with RT-qPCR

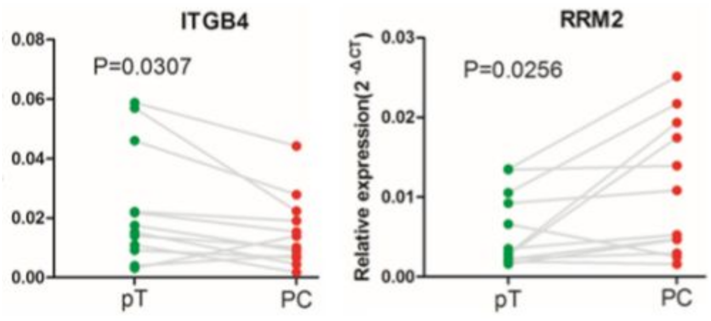
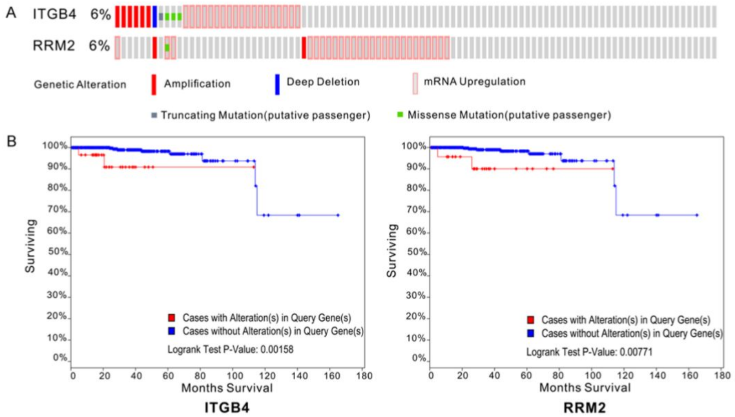


Figure 5. Validation of genes' prognostic value



He_etal_2018.pdf

2. Example datasets: He (2018) Am J Transl Res

Figure 2. GO Enrichment Analysis

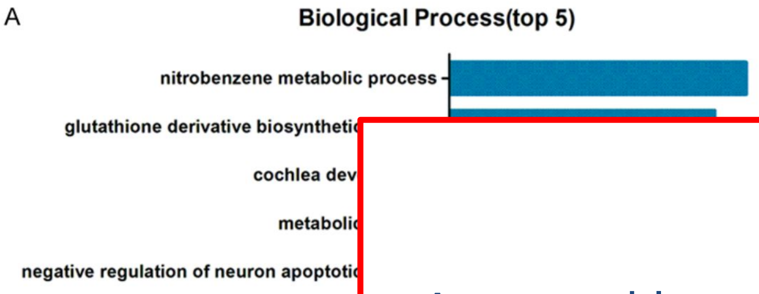
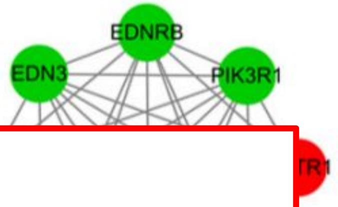
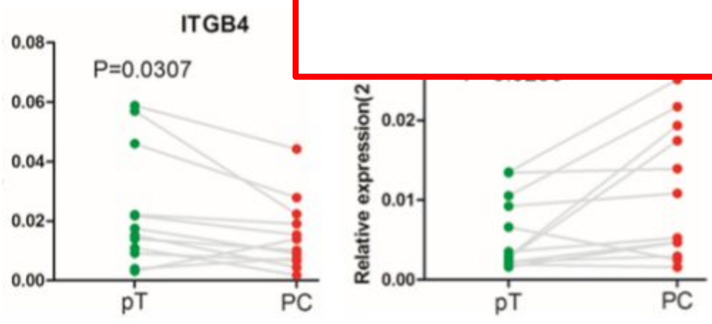


Figure 3. PPI network of DEGs

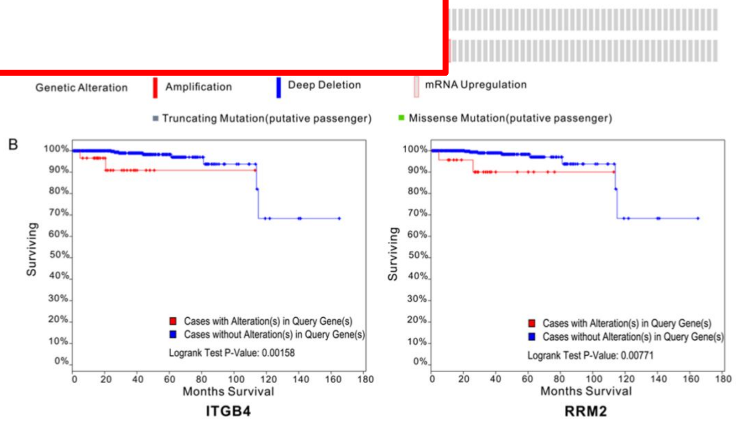


Are you able to reproduce this paper with the tools we have learnt during the course?

Figure 4. Validation

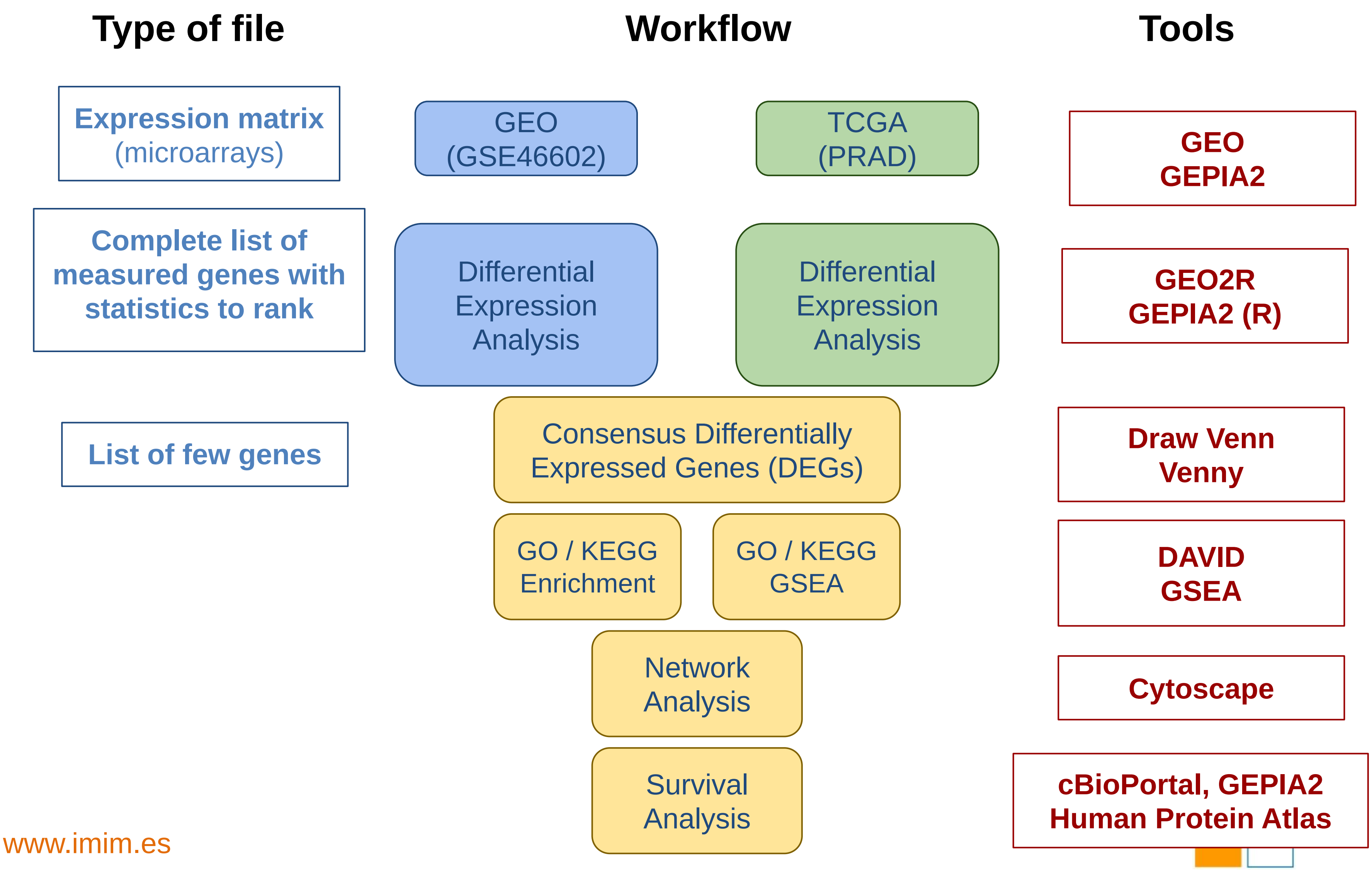


s’ prognostic value



He_etal_2018.pdf

3. Hands on: He et al (2018) Am J Transl Res



Type of file

Workflow

Tools

Expression matrix (microarrays)

GEO (GSE46602)

TCGA (PRAD)

GEO GEPIA2

Complete list of measured genes with statistics to rank

Differential Expression Analysis

Differential Expression Analysis

GEO2R GEPIA2 (R)

List of few genes

Consensus Differentially Expressed Genes (DEGs)

Draw Venn Venny

GO / KEGG Enrichment

GO / KEGG GSEA

DAVID GSEA

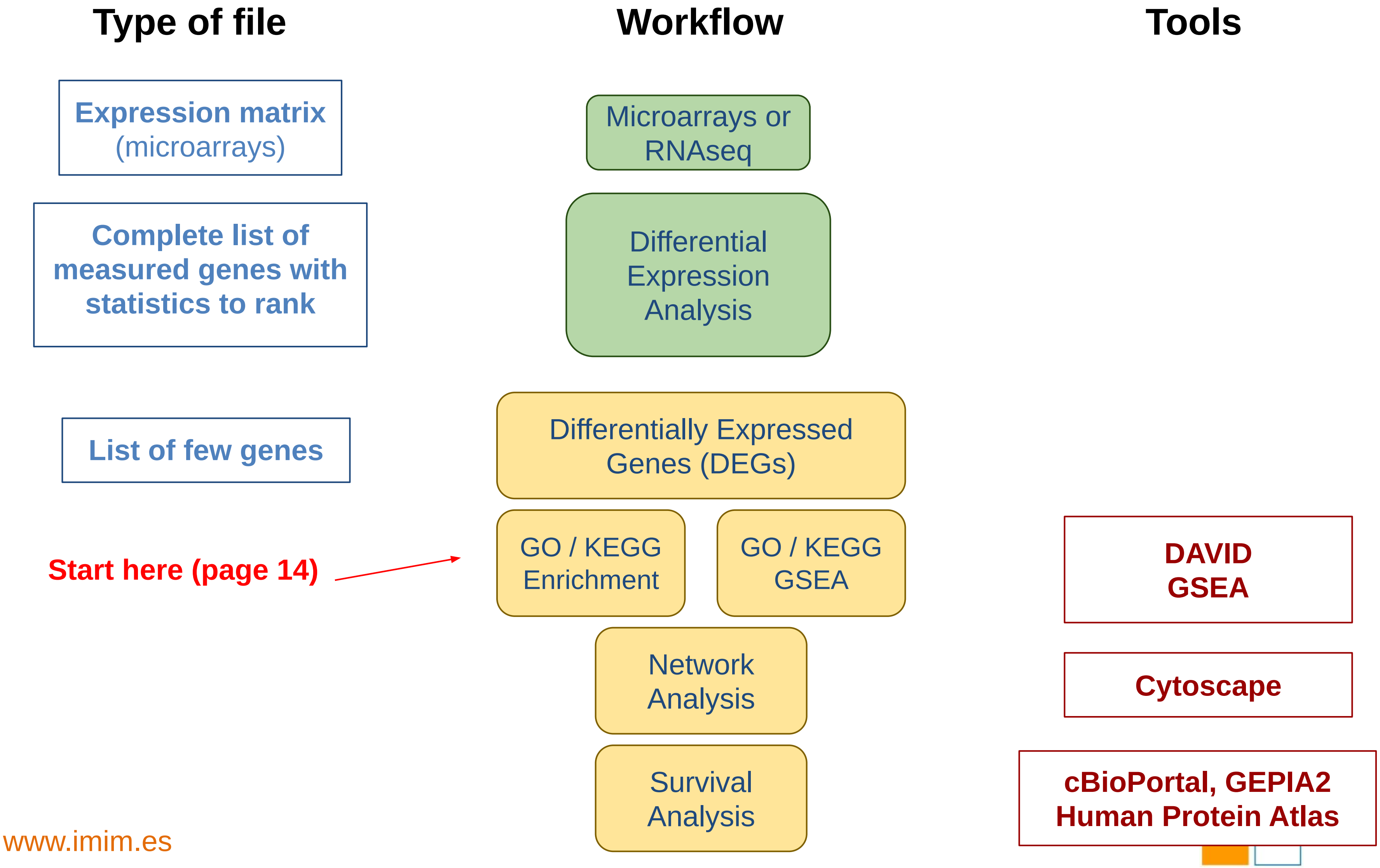
Network Analysis

Cytoscape

Survival Analysis

cBioPortal, GEPIA2 Human Protein Atlas

3. Hands on: Your own data



Type of file

Workflow

Tools

Expression matrix (microarrays)

Microarrays or RNAseq

Complete list of measured genes with statistics to rank

Differential Expression Analysis

List of few genes

Differentially Expressed Genes (DEGs)

Start here (page 14)

GO / KEGG Enrichment

GO / KEGG GSEA

DAVID GSEA

Network Analysis

Cytoscape

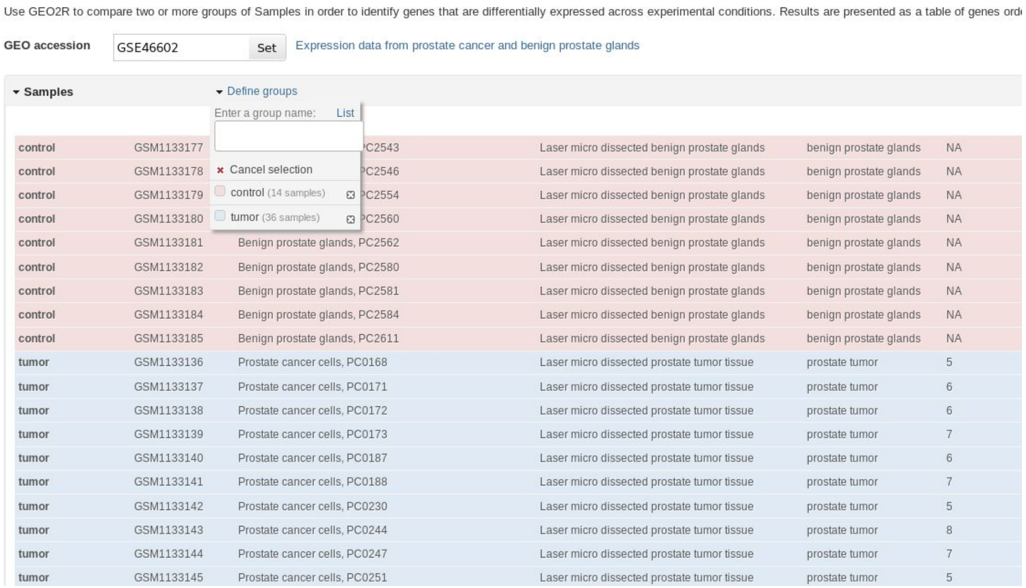
Survival Analysis

cBioPortal, GEPIA2 Human Protein Atlas

3. Hands on: Differential Expression

Imagine you are the authors of the paper, so you are interested in studying key genes and functions involved in prostate cancer. Take advantage of published datasets by using the tools we have learnt during this course.

- Perform differential expression analysis for GSE46602 using GEO2R
  - control: 14 samples
  - tumor: 36 samples
  - Results saved in file: GSE46602_GEO2R_all_results.txt



Use GEO2R to compare two or more groups of Samples in order to identify genes that are differentially expressed across experimental conditions. Results are presented as a table of genes ord

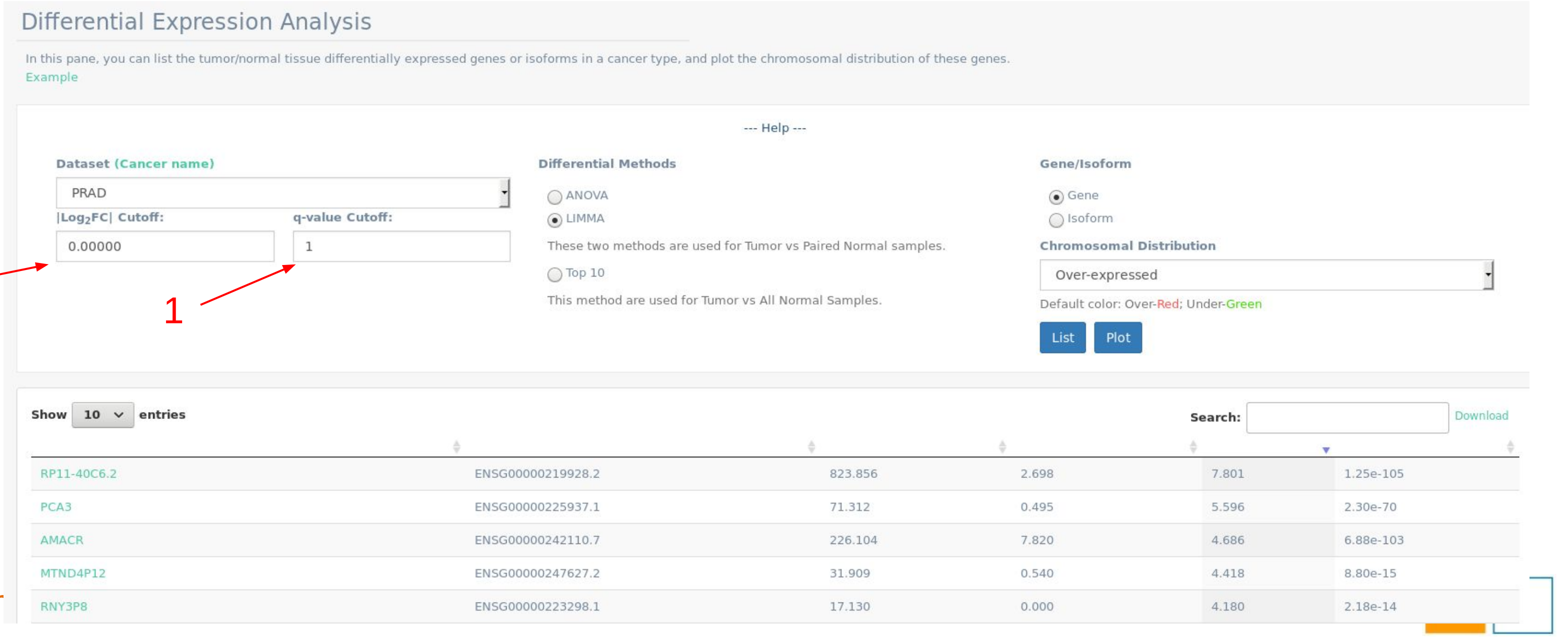
GEO accession: GSE46602 Set Expression data from prostate cancer and benign prostate glands

control	GSM1133177	C2543	Laser micro dissected benign prostate glands	benign prostate glands	NA
control	GSM1133178	C2546	Laser micro dissected benign prostate glands	benign prostate glands	NA
control	GSM1133179	C2554	Laser micro dissected benign prostate glands	benign prostate glands	NA
control	GSM1133180	C2560	Laser micro dissected benign prostate glands	benign prostate glands	NA
control	GSM1133181	Benign prostate glands, PC2562	Laser micro dissected benign prostate glands	benign prostate glands	NA
control	GSM1133182	Benign prostate glands, PC2580	Laser micro dissected benign prostate glands	benign prostate glands	NA
control	GSM1133183	Benign prostate glands, PC2581	Laser micro dissected benign prostate glands	benign prostate glands	NA
control	GSM1133184	Benign prostate glands, PC2584	Laser micro dissected benign prostate glands	benign prostate glands	NA
control	GSM1133185	Benign prostate glands, PC2611	Laser micro dissected benign prostate glands	benign prostate glands	NA
tumor	GSM1133136	Prostate cancer cells, PC0168	Laser micro dissected prostate tumor tissue	prostate tumor	5
tumor	GSM1133137	Prostate cancer cells, PC0171	Laser micro dissected prostate tumor tissue	prostate tumor	6
tumor	GSM1133138	Prostate cancer cells, PC0172	Laser micro dissected prostate tumor tissue	prostate tumor	6
tumor	GSM1133139	Prostate cancer cells, PC0173	Laser micro dissected prostate tumor tissue	prostate tumor	7
tumor	GSM1133140	Prostate cancer cells, PC0187	Laser micro dissected prostate tumor tissue	prostate tumor	6
tumor	GSM1133141	Prostate cancer cells, PC0188	Laser micro dissected prostate tumor tissue	prostate tumor	7
tumor	GSM1133142	Prostate cancer cells, PC0230	Laser micro dissected prostate tumor tissue	prostate tumor	5
tumor	GSM1133143	Prostate cancer cells, PC0244	Laser micro dissected prostate tumor tissue	prostate tumor	8
tumor	GSM1133144	Prostate cancer cells, PC0247	Laser micro dissected prostate tumor tissue	prostate tumor	7
tumor	GSM1133145	Prostate cancer cells, PC0251	Laser micro dissected prostate tumor tissue	prostate tumor	5

3. Hands on: Differential Expression

Imagine you are the authors of the paper, so you are interested in studying key genes and functions involved in prostate cancer. Take advantage of published datasets by using the tools we have learnt during this course.

- Use GEPIA2 to perform differential expression analysis of RNAseq data from TCGA
  - Results saved in file: PRAD_GEPIA2_table_degenes.txt
  - Add column names: Gene, ENSEMBL, median (tumor), median (normal), log2FC, adj.p.val



Differential Expression Analysis

In this pane, you can list the tumor/normal tissue differentially expressed genes or isoforms in a cancer type, and plot the chromosomal distribution of these genes.
Example

--- Help ---

Dataset (Cancer name): PRAD

|Log$_2$FC| Cutoff: 0.00000

q-value Cutoff: 1

Differential Methods

○ ANOVA
● LIMMA

These two methods are used for Tumor vs Paired Normal samples.

○ Top 10

This method are used for Tumor vs All Normal Samples.

Gene/Isoform

● Gene
○ Isoform

Chromosomal Distribution: Over-expressed

Default color: Over-Red; Under-Green

RP11-40C6.2	ENSG00000219928.2	823.856	2.698	7.801	1.25e-105
PCA3	ENSG00000225937.1	71.312	0.495	5.596	2.30e-70
AMACR	ENSG00000242110.7	226.104	7.820	4.686	6.88e-103
MTND4P12	ENSG00000247627.2	31.909	0.540	4.418	8.80e-15
RNY3P8	ENSG00000223298.1	17.130	0.000	4.180	2.18e-14

0.0000

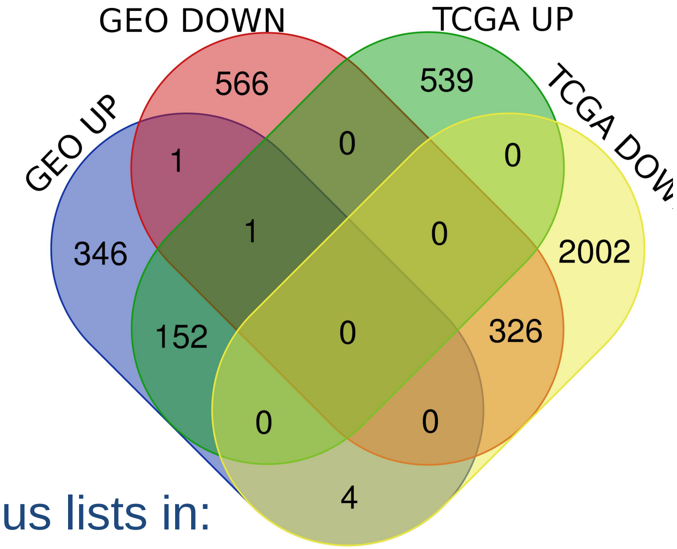
1

3. Hands on: Intersection of lists

Now, you can use the lists of genes to perform downstream functional analyses.

But, which list of genes are you going to work with? In the article they use adj.P.Val<0.05 AND |logFC|>1

- You can check the overlap using Draw Venn, Venny or VennDiagrams
- With the data I downloaded, I get this venn:
  - 326 consensus DOWN genes
  - 152 consensus UP genes
  - Total: 478 DEG

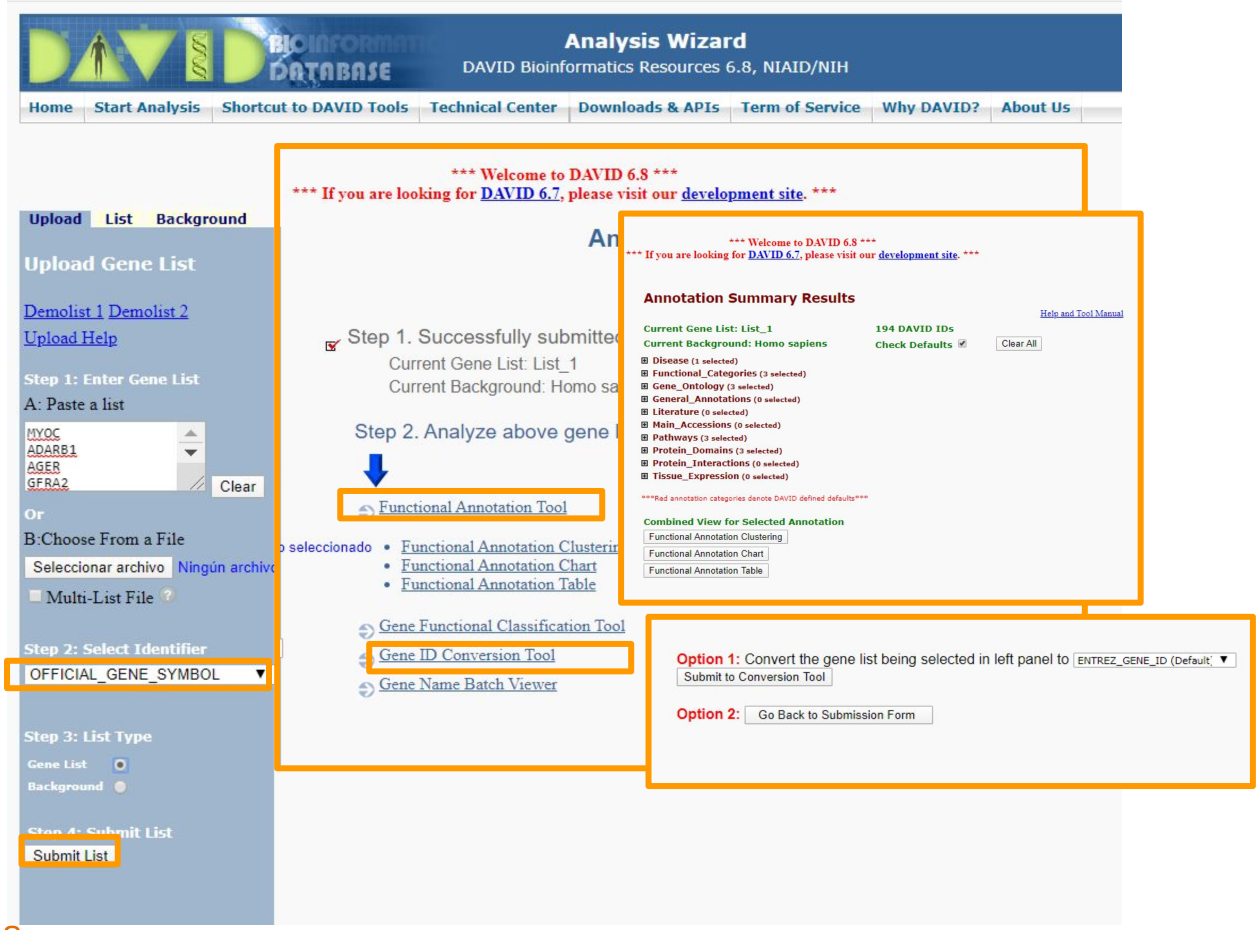


The numbers are not exactly the same as in the publication (168 UP, 316 DOWN). Any ideas why?

- You can find the UP,DOWN and ALL consensus lists in:
  - Consensus_UP.txt
  - Consensus_DOWN.txt
  - Consensus_ALL.txt

3. Hands on: Enrichment analysis

Using the list of genes, you can perform a simple gene set enrichment using DAVID. Do you find similar gene sets enriched as in the paper?



DAVID Bioinformatics Database
Analysis Wizard
DAVID Bioinformatics Resources 6.8, NIAID/NIH
Home | Start Analysis | Shortcut to DAVID Tools | Technical Center | Downloads & APIs | Term of Service | Why DAVID? | About Us
Upload | List | Background

Upload Gene List

Demolist 1 Demolist 2
Upload Help

Step 1: Enter Gene List

A: Paste a list

MYOC
ADARB1
AGER
GFRA2

Clear

Or

B:Choose From a File

Seleccionar archivo Ningún archivo

□ Multi-List File

Step 2: Select Identifier

OFFICIAL_GENE_SYMBOL ▼

Step 3: List Type

Gene List ●
Background ○

Step 4: Submit List

Submit List

*** Welcome to DAVID 6.8 ***
*** If you are looking for DAVID 6.7, please visit our development site. ***

An

Step 1. Successfully submitte
Current Gene List: List_1
Current Background: Homo sa

Step 2. Analyze above gene l

Functional Annotation Tool

o seleccionado

- Functional Annotation Clusterin
- Functional Annotation Chart
- Functional Annotation Table

Gene Functional Classification Tool
Gene ID Conversion Tool
Gene Name Batch Viewer

*** Welcome to DAVID 6.8 ***
*** If you are looking for DAVID 6.7, please visit our development site. ***

Annotation Summary Results

Help and Tool Manual

Current Gene List: List_1 — 194 DAVID IDs
Current Background: Homo sapiens — Check Defaults ☑ — Clear All

- ⊞ Disease (1 selected)
- ⊞ Functional_Categories (3 selected)
- ⊞ Gene_Ontology (3 selected)
- ⊞ General_Annotations (0 selected)
- ⊞ Literature (0 selected)
- ⊞ Main_Accessions (0 selected)
- ⊞ Pathways (3 selected)
- ⊞ Protein_Domains (3 selected)
- ⊞ Protein_Interactions (0 selected)
- ⊞ Tissue_Expression (0 selected)

Red annotation categories denote DAVID defined defaults

Combined View for Selected Annotation

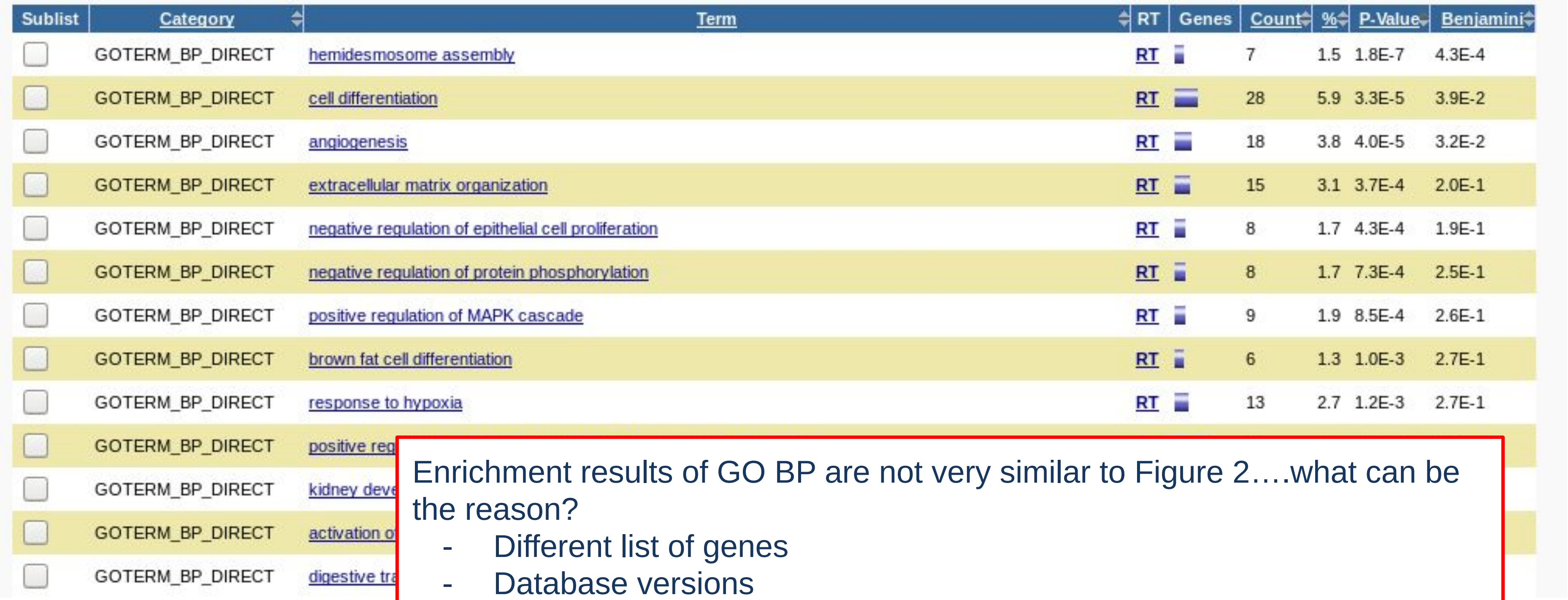
Functional Annotation Clustering
Functional Annotation Chart
Functional Annotation Table

Option 1: Convert the gene list being selected in left panel to ENTREZ_GENE_ID (Default) ▼
Submit to Conversion Tool

Option 2: Go Back to Submission Form

3. Hands on: Enrichment analysis

Using the list of genes, you can perform a simple gene set enrichment using DAVID. Do you find similar gene sets enriched as in the paper?



Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
□	GOTERM_BP_DIRECT	hemidesmosome assembly	RT	▬	7	1.5	1.8E-7	4.3E-4
□	GOTERM_BP_DIRECT	cell differentiation	RT	▬	28	5.9	3.3E-5	3.9E-2
□	GOTERM_BP_DIRECT	angiogenesis	RT	▬	18	3.8	4.0E-5	3.2E-2
□	GOTERM_BP_DIRECT	extracellular matrix organization	RT	▬	15	3.1	3.7E-4	2.0E-1
□	GOTERM_BP_DIRECT	negative regulation of epithelial cell proliferation	RT	▬	8	1.7	4.3E-4	1.9E-1
□	GOTERM_BP_DIRECT	negative regulation of protein phosphorylation	RT	▬	8	1.7	7.3E-4	2.5E-1
□	GOTERM_BP_DIRECT	positive regulation of MAPK cascade	RT	▬	9	1.9	8.5E-4	2.6E-1
□	GOTERM_BP_DIRECT	brown fat cell differentiation	RT	▬	6	1.3	1.0E-3	2.7E-1
□	GOTERM_BP_DIRECT	response to hypoxia	RT	▬	13	2.7	1.2E-3	2.7E-1
□	GOTERM_BP_DIRECT	positive reg						
□	GOTERM_BP_DIRECT	kidney deve						
□	GOTERM_BP_DIRECT	activation o						
□	GOTERM_BP_DIRECT	digestive tra						

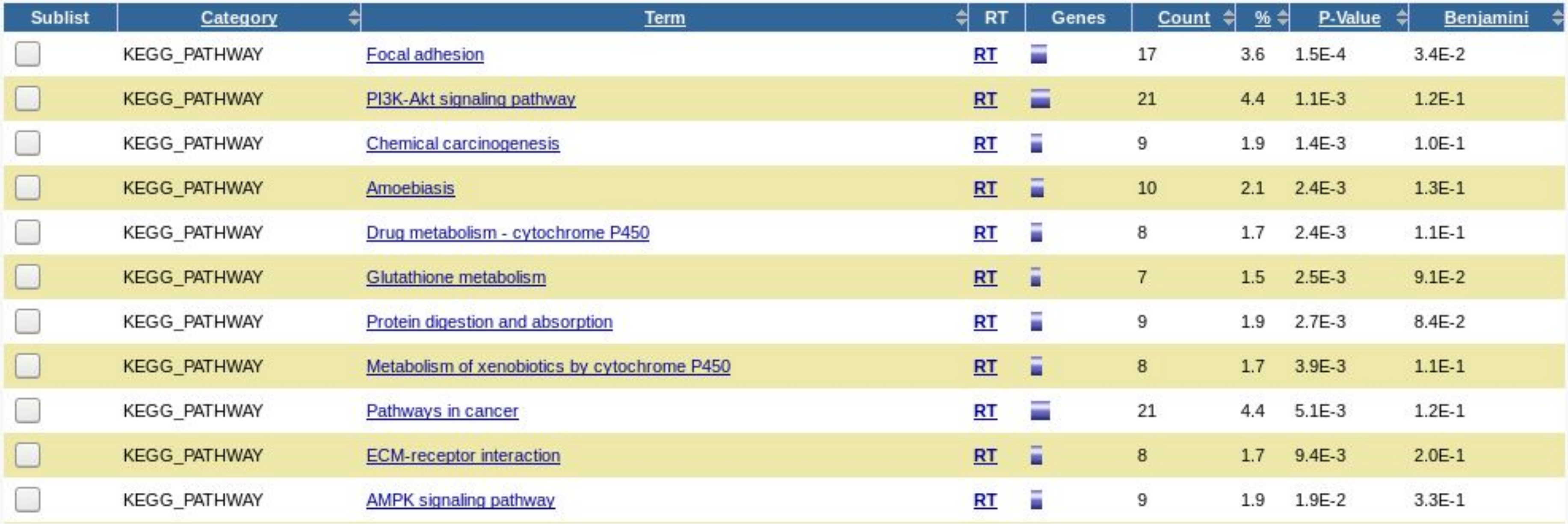
Enrichment results of GO BP are not very similar to Figure 2….what can be the reason?

- Different list of genes
- Database versions

Also, I can be advisable to do it separately for up and down-regulated genes

3. Hands on: Enrichment analysis

Using the list of genes, you can perform a simple gene set enrichment using DAVID. Do you find similar gene sets enriched as in the paper?



Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
□	KEGG_PATHWAY	Focal adhesion	RT	▬	17	3.6	1.5E-4	3.4E-2
□	KEGG_PATHWAY	PI3K-Akt signaling pathway	RT	▬	21	4.4	1.1E-3	1.2E-1
□	KEGG_PATHWAY	Chemical carcinogenesis	RT	▬	9	1.9	1.4E-3	1.0E-1
□	KEGG_PATHWAY	Amoebiasis	RT	▬	10	2.1	2.4E-3	1.3E-1
□	KEGG_PATHWAY	Drug metabolism - cytochrome P450	RT	▬	8	1.7	2.4E-3	1.1E-1
□	KEGG_PATHWAY	Glutathione metabolism	RT	▬	7	1.5	2.5E-3	9.1E-2
□	KEGG_PATHWAY	Protein digestion and absorption	RT	▬	9	1.9	2.7E-3	8.4E-2
□	KEGG_PATHWAY	Metabolism of xenobiotics by cytochrome P450	RT	▬	8	1.7	3.9E-3	1.1E-1
□	KEGG_PATHWAY	Pathways in cancer	RT	▬	21	4.4	5.1E-3	1.2E-1
□	KEGG_PATHWAY	ECM-receptor interaction	RT	▬	8	1.7	9.4E-3	2.0E-1
□	KEGG_PATHWAY	AMPK signaling pathway	RT	▬	9	1.9	1.9E-2	3.3E-1

KEGG seems to give more similar results

3. Hands on: GSEA

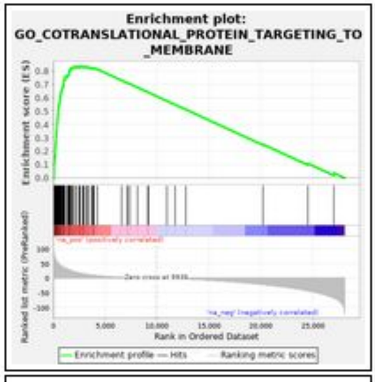
Use the complete ranked list of genes (TCGA data) to perform GSEA analysis.

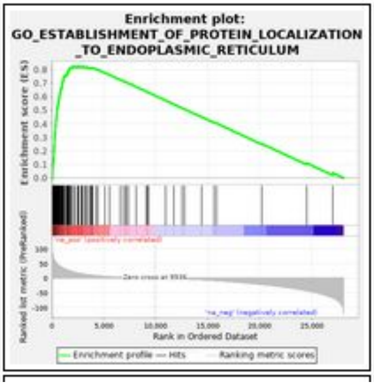
- GSEA classic
  - You need to generate two files (.gct and .cls), which can be a bit tricky. Also, it is only available for microarray data
- **GSEA Preranked**
  - You need to generate one file (.rnk). Easy to generate with a simple formula (-log10*sign(FC)) and suitable for microarrays and RNAseq data

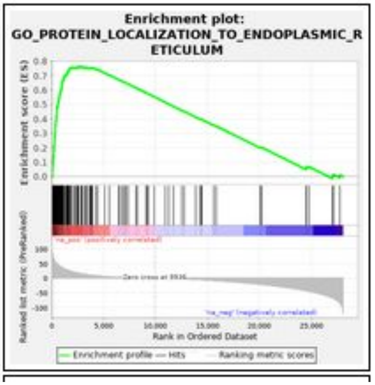
3. Hands on: GSEA

We will perform a GSEA Preranked using PRAD GEPIA2 data:

- Prepare rnk input file:
  - take column Gene and Rank from file PRAD_GEPIA2_table_degenes.txt and save it in a new file. Call it PRAD.rnk.
  - Or use the provided PRAD.rnk file
- Load data → Browse files → PRAD.rnk
- Run GSEAPreranked
- Select c5.bp.v7.0.symbols.gmt gene set database (or any other you want)
- Select “No collapse”
- Run and wait (output in folder my_analysis.GseaPreranked.1582108828269)

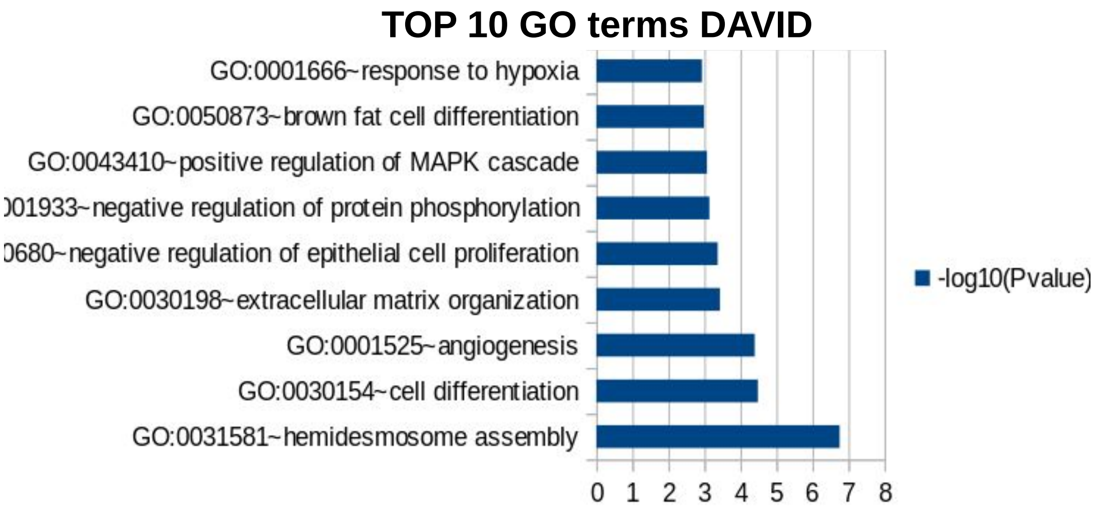






TOP 10 GO BP terms: DAVID vs GSEA

Can you make these plots in excel?

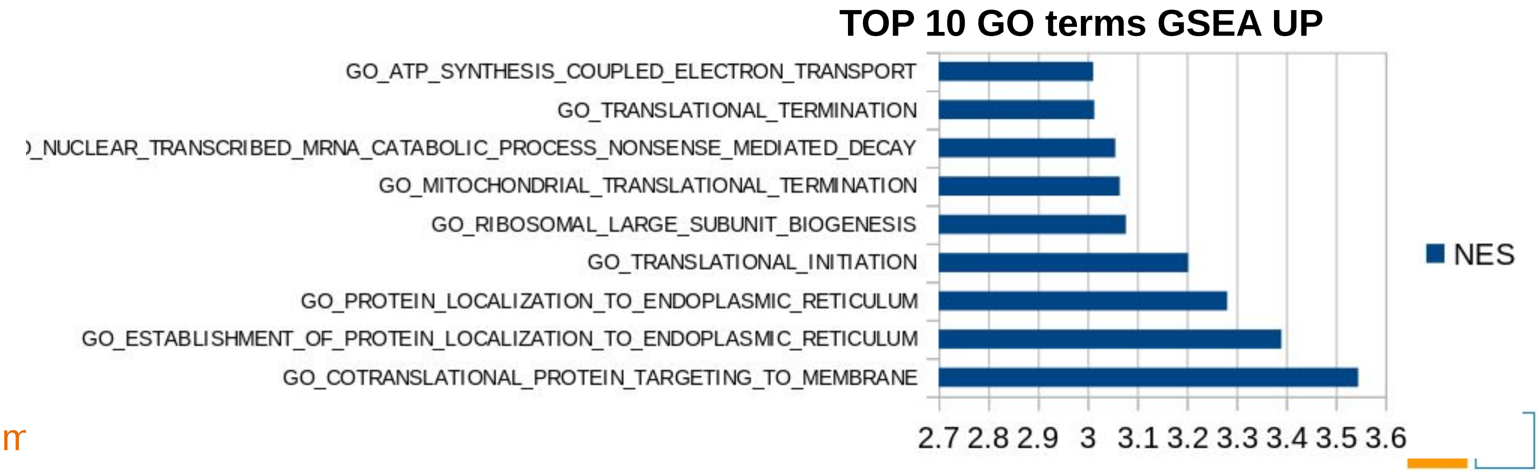


Observe the different top results between the two methods.

Are they the same as in Figure 3?

Reproducibility is an issue!!!

If you want, try to do this comparison with KEGG



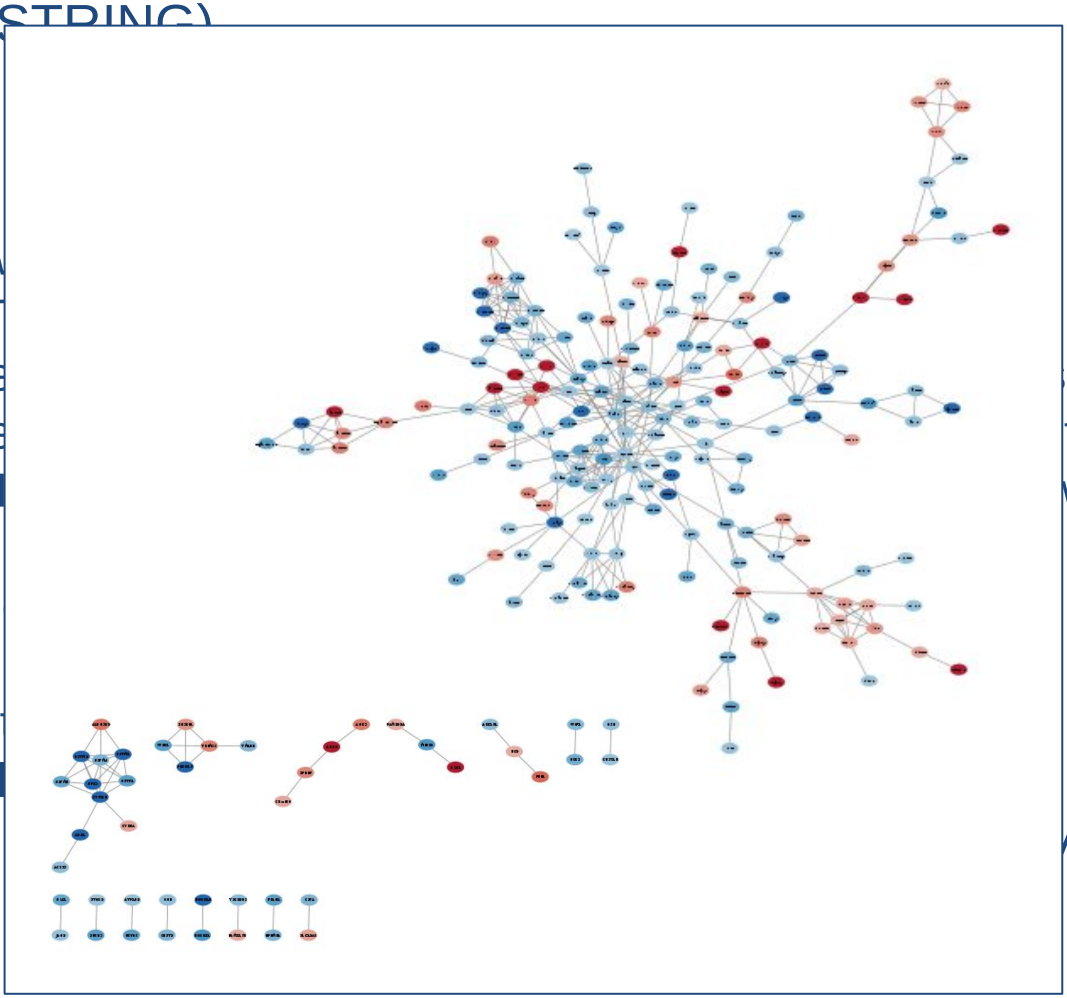
3. Hands on: Network analysis

Upload the list of 478 DEGs in Cytoscape and populate it with a database (eg.STRING):

1. Install stringApp and MCODE app
2. File → New Network → New Network from Public Databases → Data Source=STRING:protein query
3. Copy/Paste the list of DEG (Up and Down) AND confidence score > 0.7
4. Tools → NetworkAnalyzer → Network Analysis → Analyse network
5. Select → Node:Degree “is” between 1 and 1000 → New Network from Selection (how many genes have you filtered?)
6. Import Table from File: PRAD_GEPIA2_table_degenes.txt
7. Style
   a. Default
   b. Fill Color using log2FC column
8. Save this network as “Network_STRING_0.7_1degree_GEPIA.cys”

3. Hands on: Network analysis

Upload the list of 478 DEGs in Cytoscape and populate it with a database (eg.STRING)



1. Install STR
2. File → New → Data Source=ST
3. Copy/Paste core > 0.7
4. Tools → Ne twork
5. Select → N vork from Selection
6. Import Tabl
7. Style
   a. Defaul
   b. Fill Col
8. Save this n A.cys”

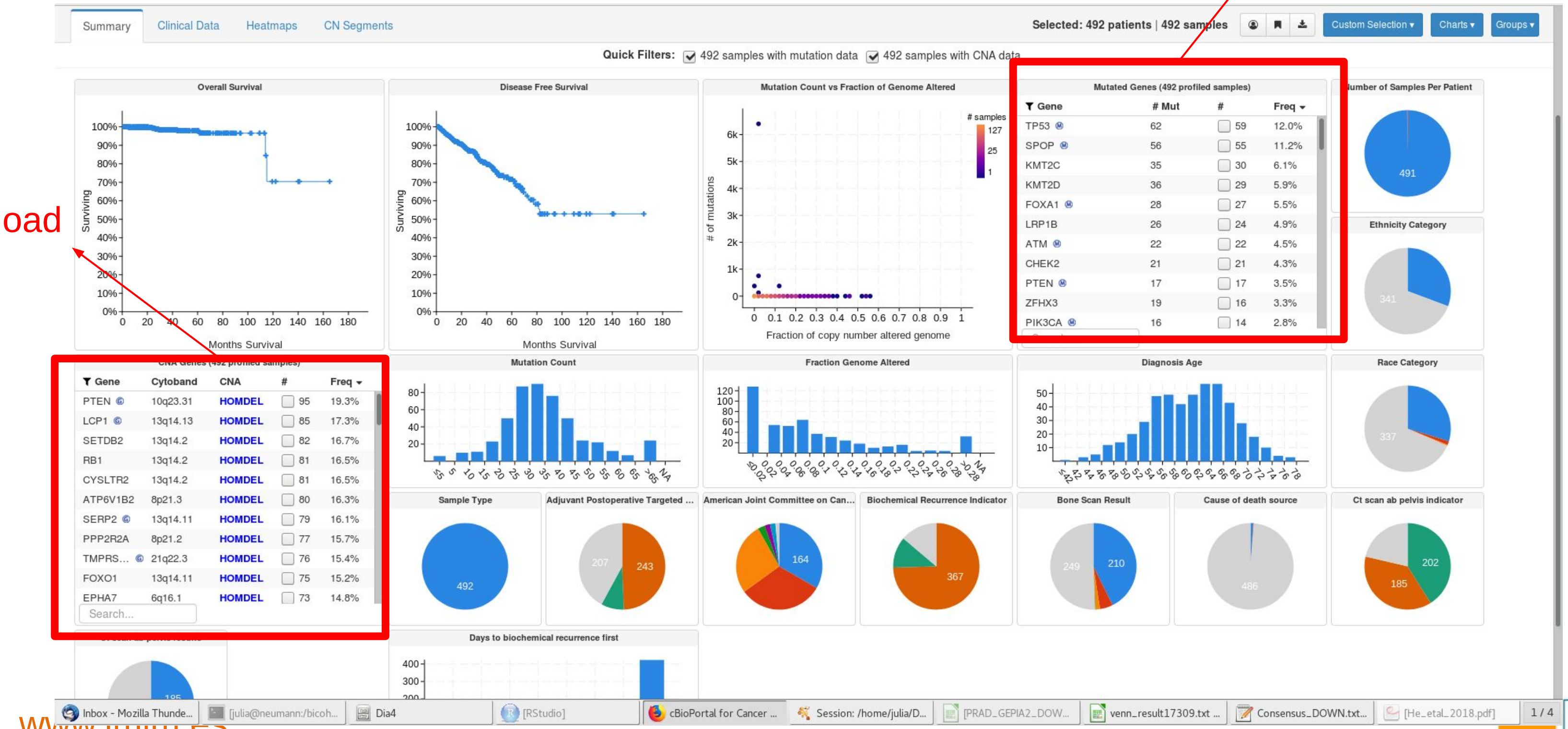
3. Hands on: Network analysis

Integrate Mutations and CNA:

- Search in cBioPortal the frequency of mutations and copy number in PRAD and integrate the information to the network
  - Select study: Prostate Adenocarcinoma (TCGA, Firehose Legacy) → Explore selected studies

download

download



Selected: 492 patients | 492 samples

Quick Filters: ☑ 492 samples with mutation data ☑ 492 samples with CNA data

Mutated Genes (492 profiled samples)

Gene	# Mut	#	Freq
TP53	62	□ 59	12.0%
SPOP	56	□ 55	11.2%
KMT2C	35	□ 30	6.1%
KMT2D	36	□ 29	5.9%
FOXA1	28	□ 27	5.5%
LRP1B	26	□ 24	4.9%
ATM	22	□ 22	4.5%
CHEK2	21	□ 21	4.3%
PTEN	17	□ 17	3.5%
ZFHX3	19	□ 16	3.3%
PIK3CA	16	□ 14	2.8%

CNA Genes (492 profiled samples)

Gene	Cytoband	CNA	#	Freq
PTEN	10q23.31	HOMDEL	□ 95	19.3%
LCP1	13q14.13	HOMDEL	□ 85	17.3%
SETDB2	13q14.2	HOMDEL	□ 82	16.7%
RB1	13q14.2	HOMDEL	□ 81	16.5%
CYSLTR2	13q14.2	HOMDEL	□ 81	16.5%
ATP6V1B2	8p21.3	HOMDEL	□ 80	16.3%
SERP2	13q14.11	HOMDEL	□ 79	16.1%
PPP2R2A	8p21.2	HOMDEL	□ 77	15.7%
TMPRS...	21q22.3	HOMDEL	□ 76	15.4%
FOXO1	13q14.11	HOMDEL	□ 75	15.2%
EPHA7	6q16.1	HOMDEL	□ 73	14.8%

Search...

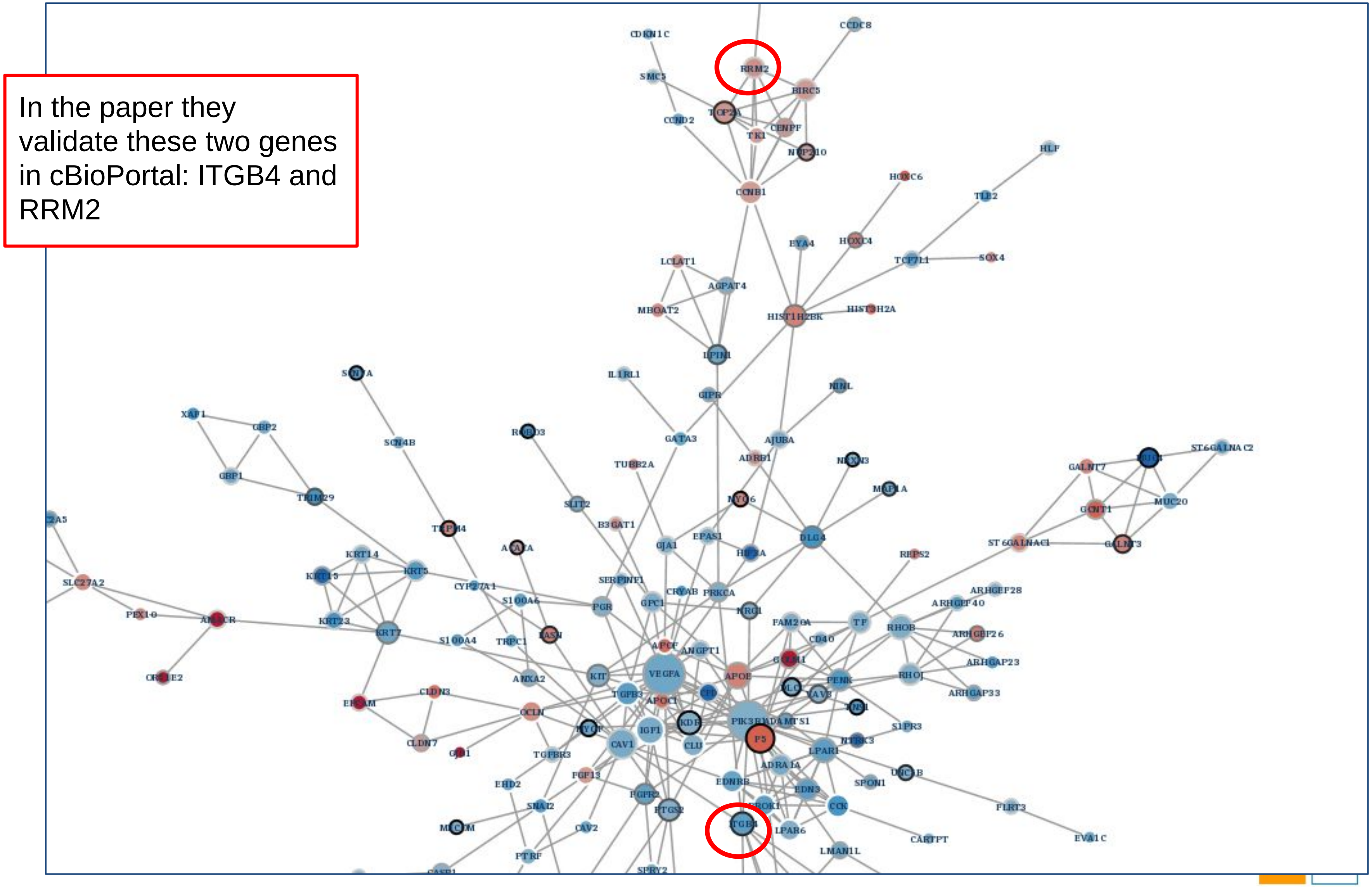
3. Hands on: Network analysis

Integrate Mutations and CNA:

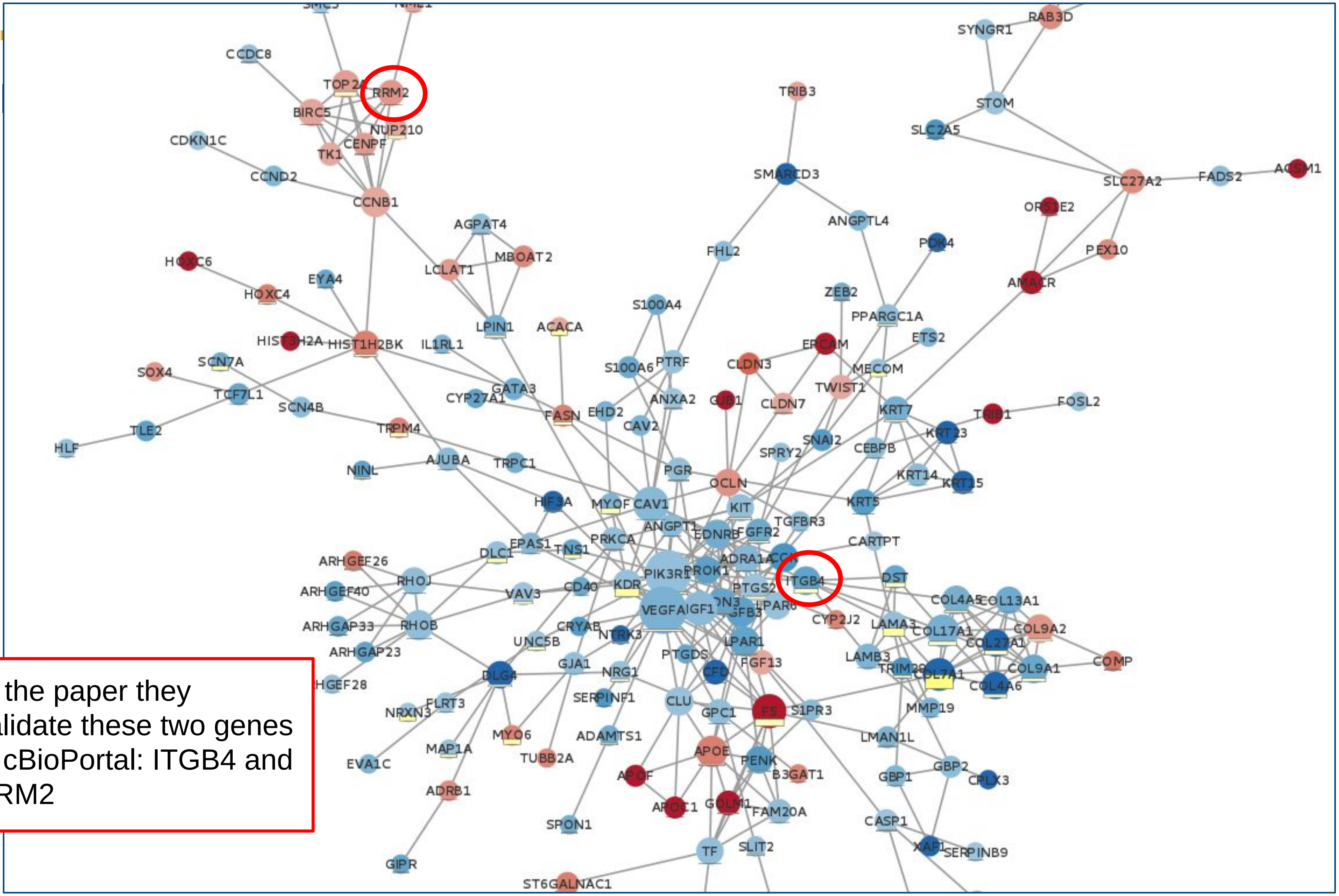
- Search in cBioPortal the frequency of mutations and copy number in PRAD and integrate the information to the network
  - Select study: Prostate Adenocarcinoma (TCGA, Firehose Legacy) → Explore selected studies
  - Download mutations and CN data
  - Import Table from File: PRAD_Mutated_Genes.txt
  - Import Table from File: PRAD_CNA_Genes.txt
  - If you don't manage, open the network provided in file: "Network_STRING_0.7_1degree_GEPIA_cBioPortal.cys"
- Try to style the network using:
  - Style: Ripple
  - Size → Degree
  - Border Paint: # Mut
  - Or Image/Chart → Bar → # Mut

3. Hands on: Network analysis

In the paper they validate these two genes in cBioPortal: ITGB4 and RRM2

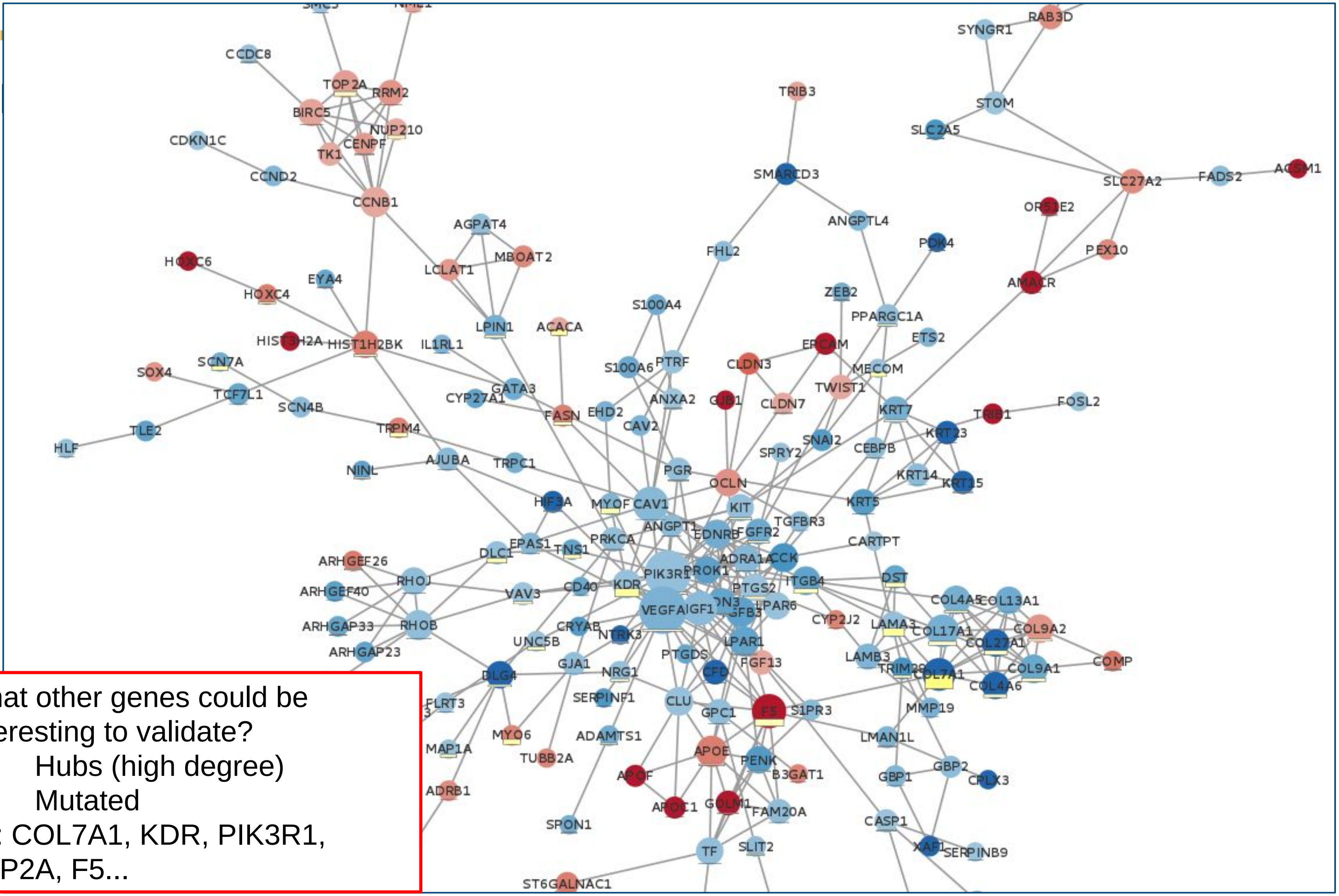


3. Hands on: Network analysis



In the paper they validate these two genes in cBioPortal: ITGB4 and RRM2

3. Hands on: Network analysis



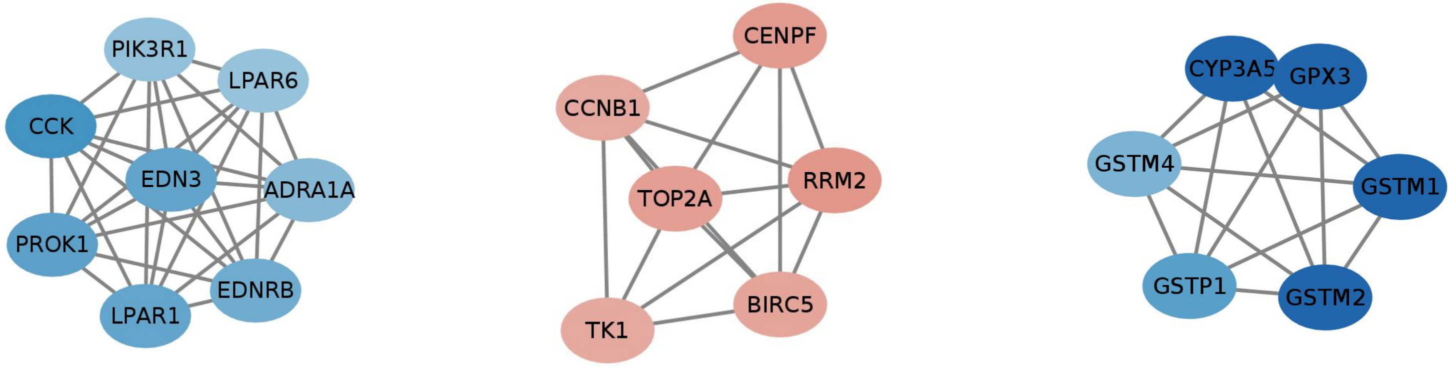
What other genes could be Interesting to validate?

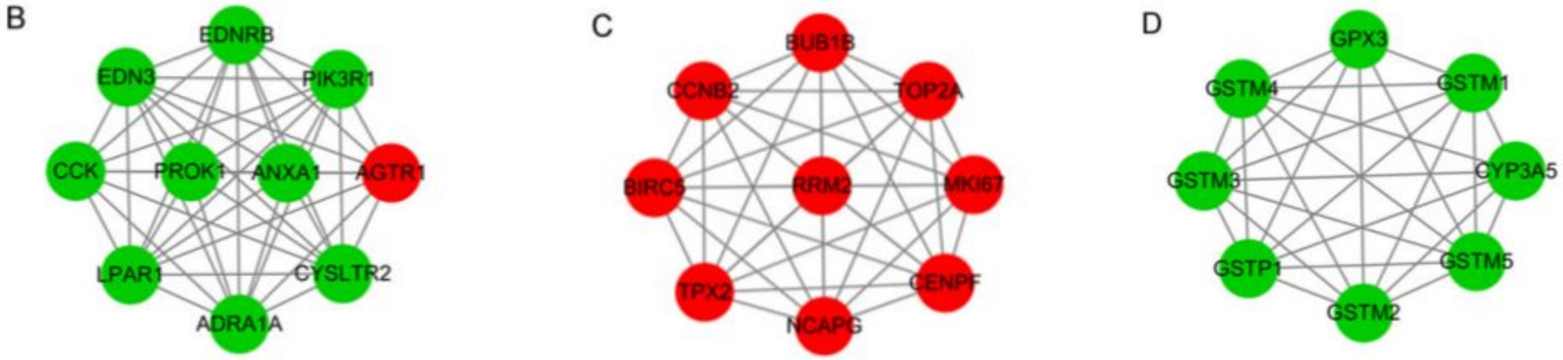
- Hubs (high degree)
- Mutated

Eg: COL7A1, KDR, PIK3R1, TOP2A, F5...

3. Hands on: Network analysis

- MCODE → + → Analyze current network (default parameters)
- Do you find the modules in Figure 3?

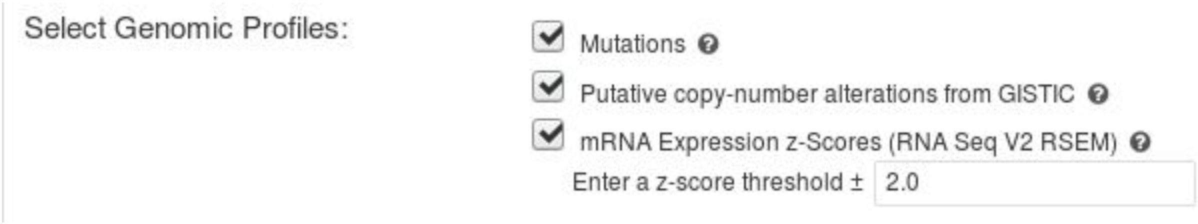




3. Hands on: Network analysis

Prognostic value of identified genes:

- Search in cBioPortal
  - Select study: Prostate Adenocarcinoma (TCGA, Firehose Legacy) → Query Genes → **RRM2**



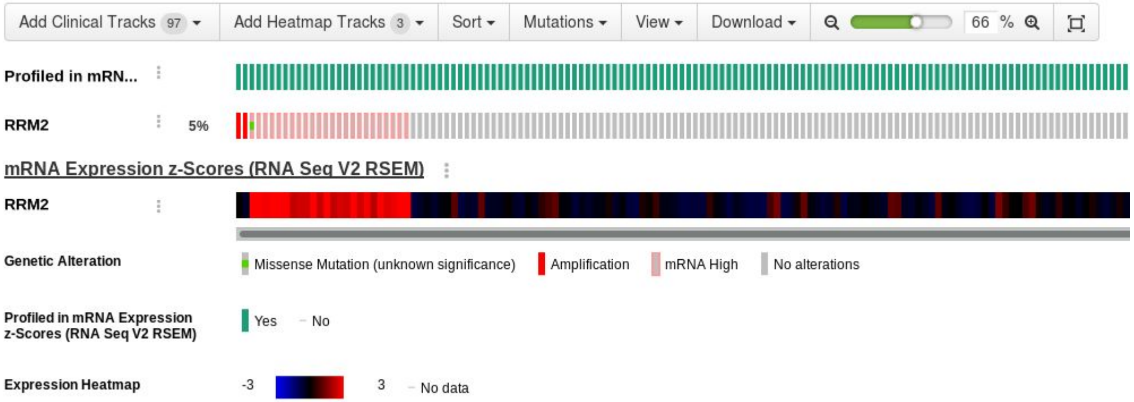
Select Genomic Profiles:

☑ Mutations

☑ Putative copy-number alterations from GISTIC

☑ mRNA Expression z-Scores (RNA Seq V2 RSEM)

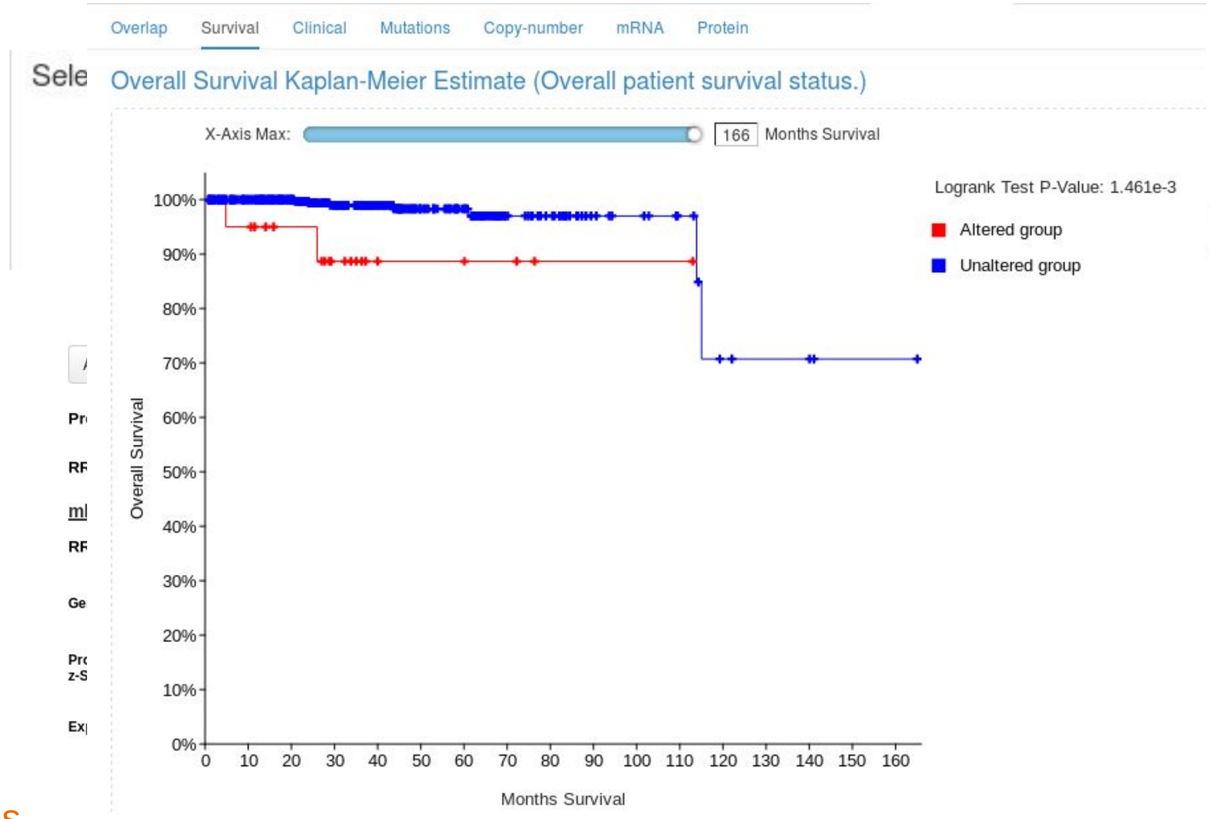
Enter a z-score threshold ± 2.0



3. Hands on: Network analysis

Prognostic value of identified genes:

- Search in cBioPortal
  - Go to Survival Tab



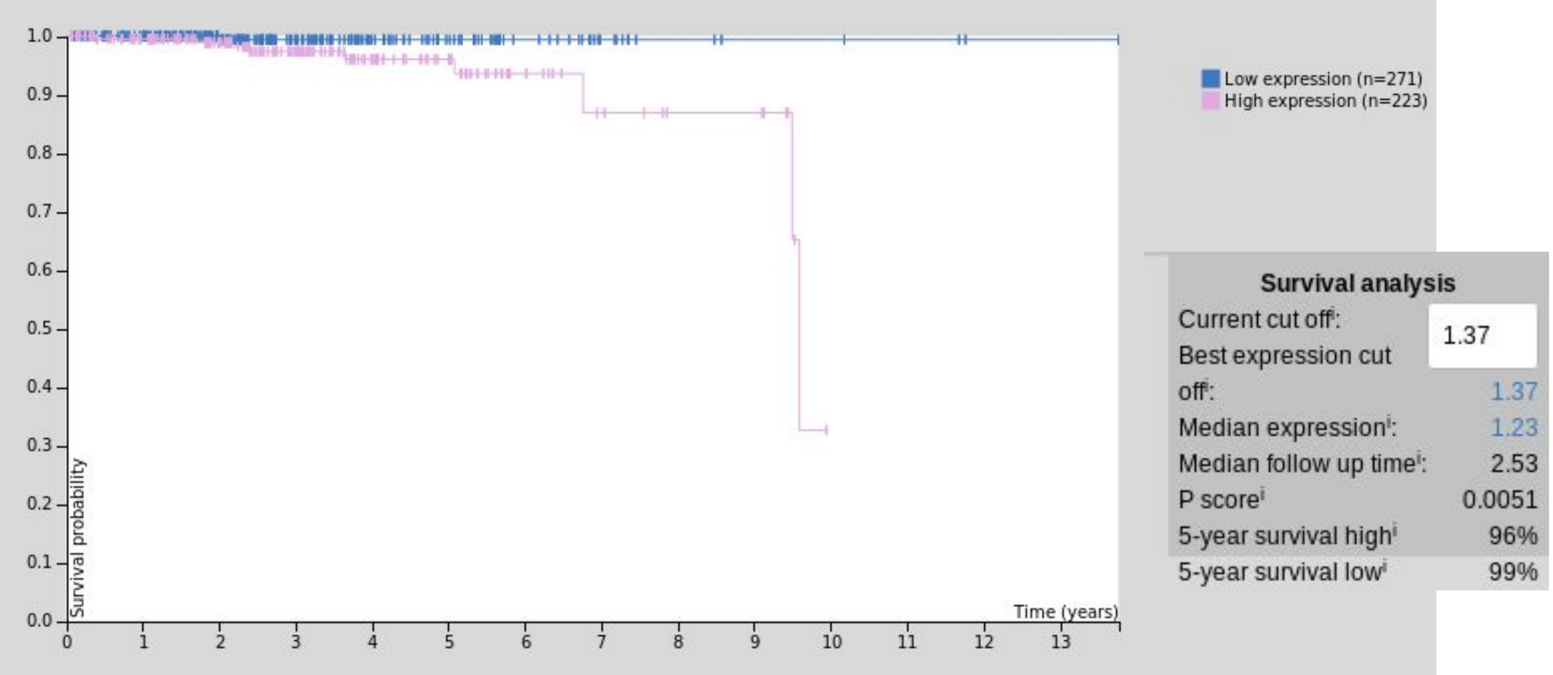
Overall Survival Kaplan-Meier Estimate (Overall patient survival status.)

X-Axis Max: 166 Months Survival

3. Hands on: Network analysis

Prognostic value of identified genes:

- Search in The Human Protein Atlas → RMM2 → Pathology → click on Prostate Cancer boxplot



Survival analysis

Current cut off[i]:	1.37
Best expression cut off[i]:	1.37
Median expression[i]:	1.23
Median follow up time[i]:	2.53
P score[i]	0.0051
5-year survival high[i]	96%
5-year survival low[i]	99%

What is the difference between the two Kaplan-Meier curves?

Summary

After the course you should be able to…

- select the appropriate enrichment test for your data
- be aware of the different gene sets databases
- perform an enrichment test with a list of genes using DAVID
- perform GSEA
- do basic network manipulation with Cytoscape
- use some apps to extend the functionality
- understand that reproducibility is a huge issue
- publish a bioinformatics paper such as He et al!