

# Functional enrichment analysis of high-throughput omics data

*Dr. Ludwig Geistlinger, 27 Jul 2017  
Waldron Lab, CUNY School of Public Health*

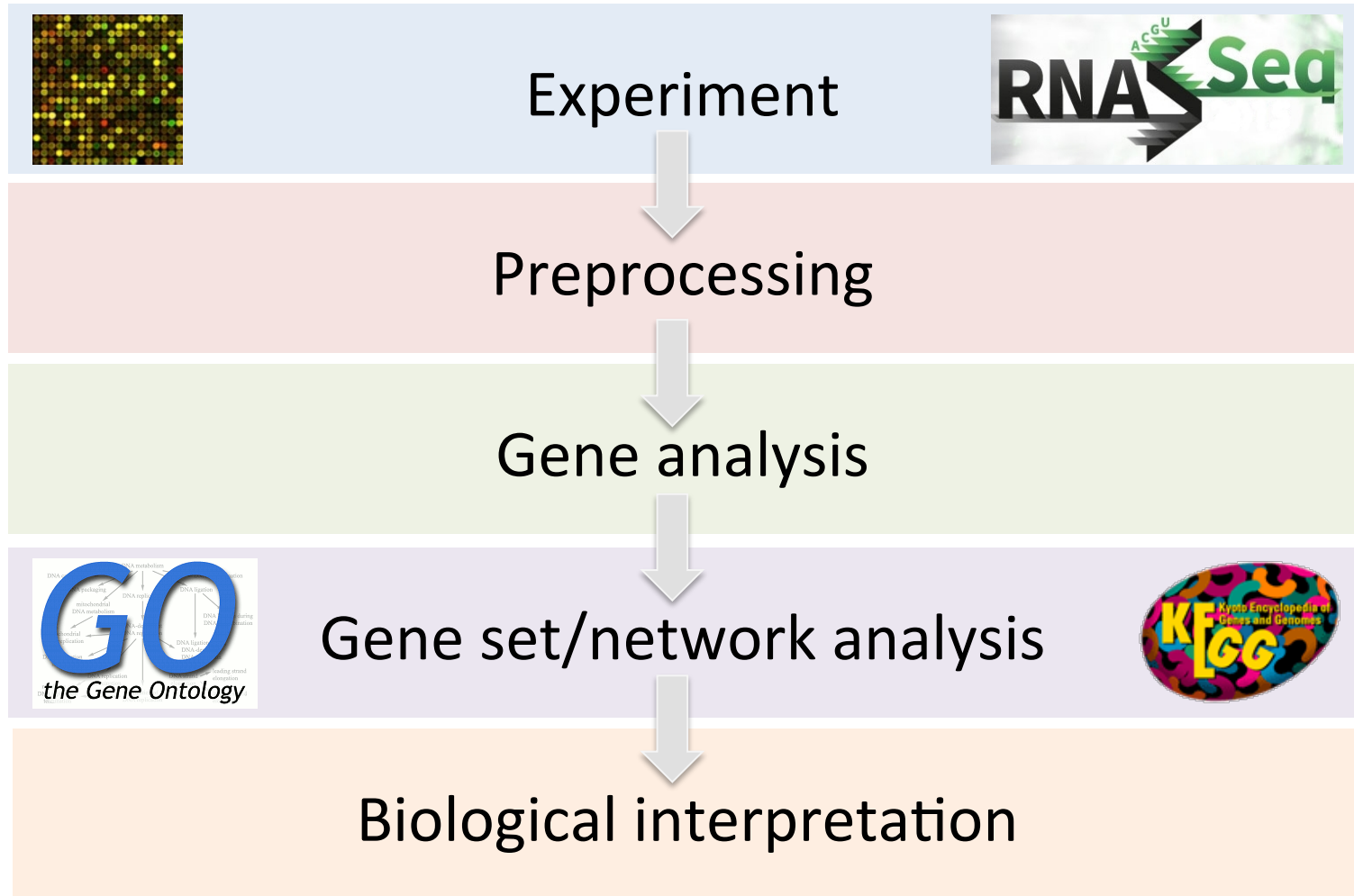
# Contents

- 1 Gene expression-based enrichment analysis
- 2 Genomic region enrichment analysis
- 3 Multi-omics enrichment analysis

# Contents

- 1 Gene expression-based enrichment analysis
- 2 Genomic region enrichment analysis
- 3 Multi-omics enrichment analysis

# Gene expression data analysis



Home » [Bioconductor 3.2](#) » [Software Packages](#) » [EnrichmentBrowser](#)

## EnrichmentBrowser

platforms **all** downloads **top 20%** posts **1 / 0 / 0 / 0** in Bioc **1.5 years**  
 build **ok** commits **1.83** test coverage **unknown**

Seamless navigation through combined results of set-based and network-based enrichment analysis

Bioconductor version: Release (3.2)

The EnrichmentBrowser package implements essential functionality for the enrichment analysis of gene expression data. The analysis combines the advantages of set-based and network-based enrichment analysis in order to derive high-confidence gene sets and biological pathways that are differentially regulated in the expression data under investigation. Besides, the package facilitates the visualization and exploration of such sets and pathways.

Author: Ludwig Geistlinger, Gergely Csaba, Ralf Zimmer

Maintainer: Ludwig Geistlinger <Ludwig.Geistlinger at bio.ifi.lmu.de>

Citation (from within R, enter `citation("EnrichmentBrowser")`):

Geistlinger L, Csaba G and Zimmer R (2016). "Bioconductor's EnrichmentBrowser: seamless navigation through combined results of set- & network-based enrichment analysis." *BMC Bioinformatics*, **17**, pp. 45. <http://doi.org/10.1186/s12859-016-0884-1>.

Month	Nb of distinct IPs	Nb of downloads
May/2015	132	190
Jun/2015	140	220
Jul/2015	147	226
Aug/2015	125	172
Sep/2015	124	200
Oct/2015	152	287
Nov/2015	125	205
Dec/2015	148	276
Jan/2016	196	339
Feb/2016	186	362
Mar/2016	210	345
Apr/2016	280	399
<b>All months</b>	<b>1535</b>	<b>3221</b>

the following locations:

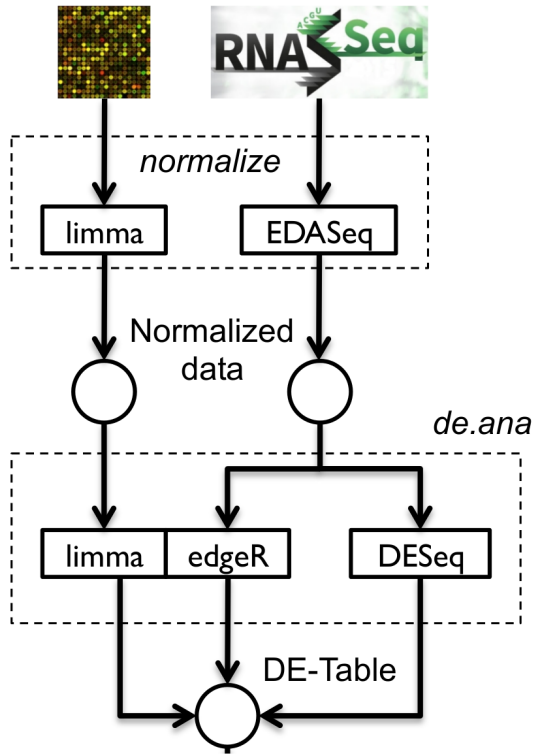
- [Support site](#) - for questions about Bioconductor packages
- [Bioc-devel](#) mailing list - for package developers

# Intro: EnrichmentBrowser

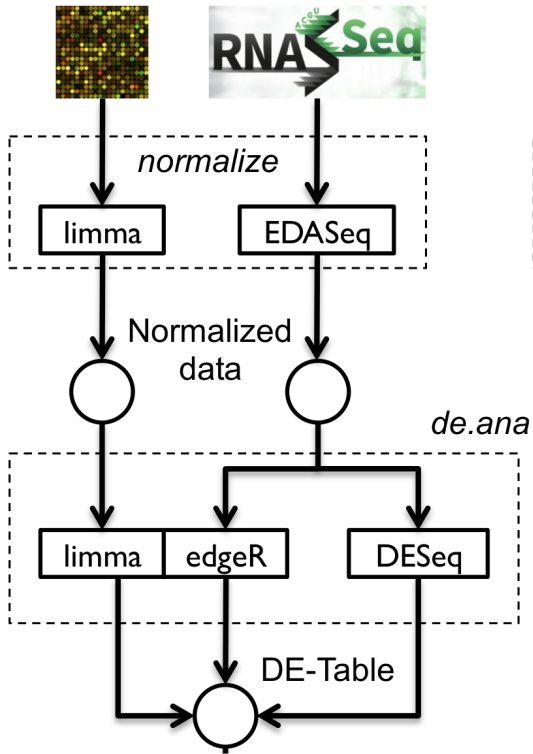
→ *Implements an end-to-end gene expression analysis pipeline*

- Preparing the data
- Preprocessing the data
- Differential expression (DE) analysis
- Defining gene sets and networks of interest
- Executing individual enrichment methods
- Combining results of different methods
- Visualization and exploration of results

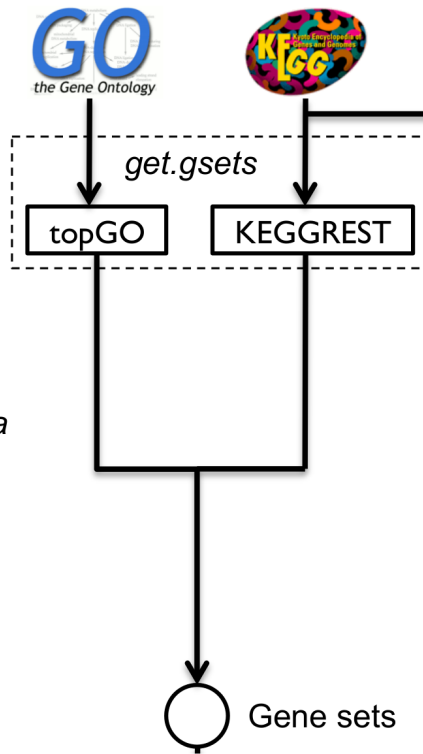
# EXPRESSION DATA



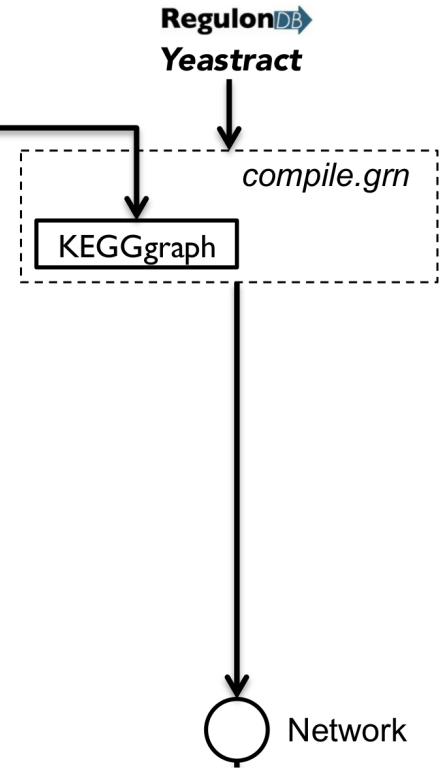
### EXPRESSION DATA



### GENE SETS / PATHWAYS

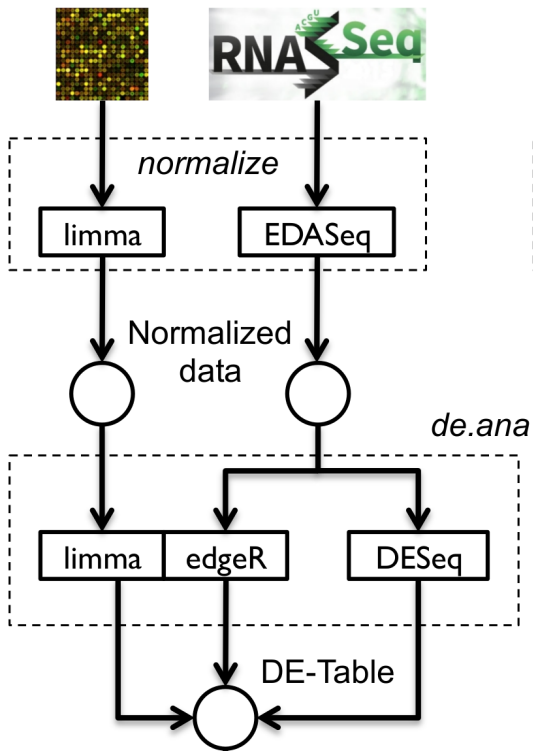


### REGULATORY NETWORK

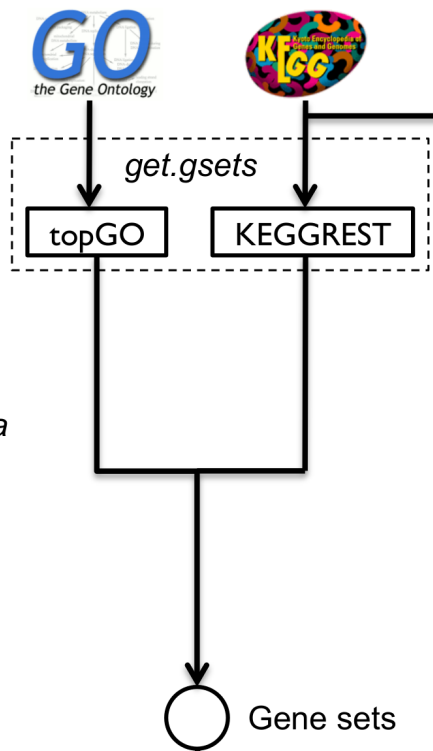




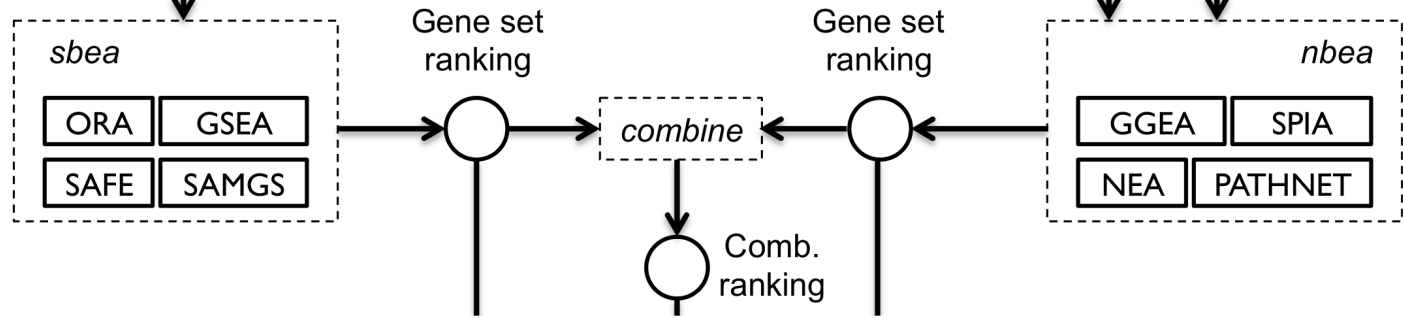
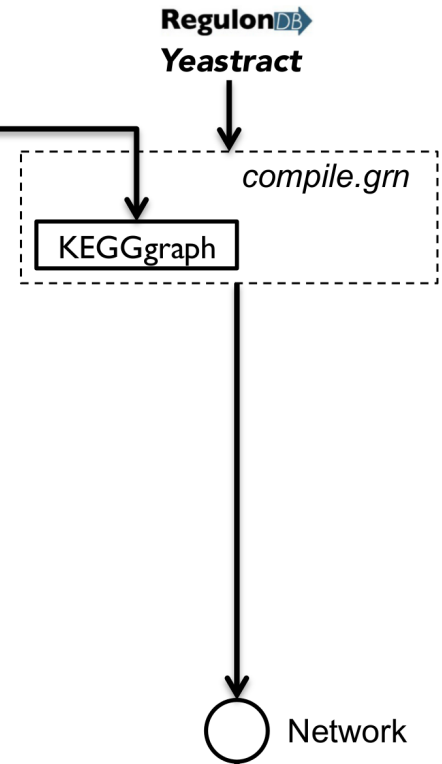
### EXPRESSION DATA



### GENE SETS / PATHWAYS



### REGULATORY NETWORK



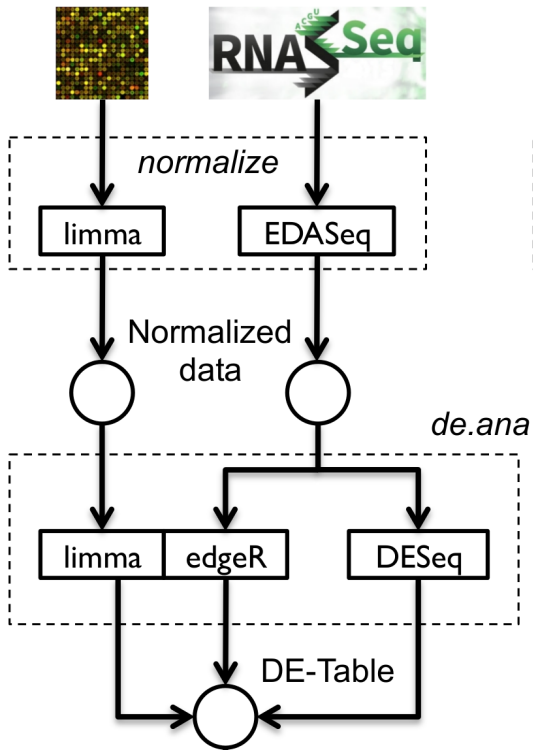
# A primer on gene set enrichment analysis

- GO/KEGG overrepresentation
  - Dozens of methods published, partly available (in BioC)
  - Development of additional methods ongoing
  - Lack of standards for method evaluation
    - No „best“ method
    - Existing methods have never been compared consistently
- ➔ *No easy solution*
- ➔ *Pragmatic approach: enable comparative exec of methods*
- ➔ *Detection of gene sets with evidence across methods*

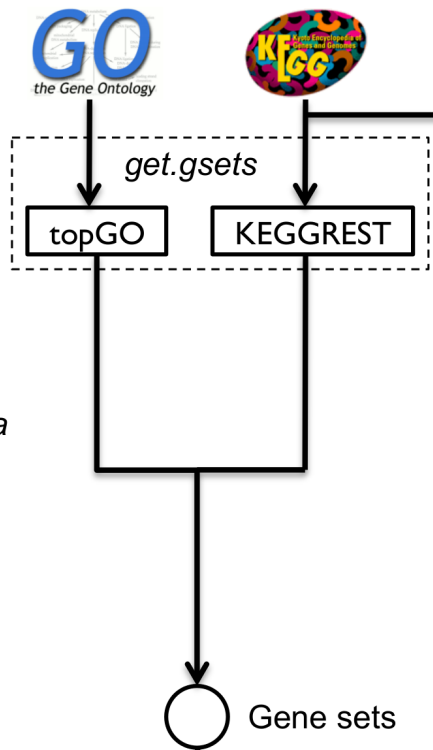
# Why combining results!?

- gene sets reported by multiple methods
- downgrades spurious hits of individual methods

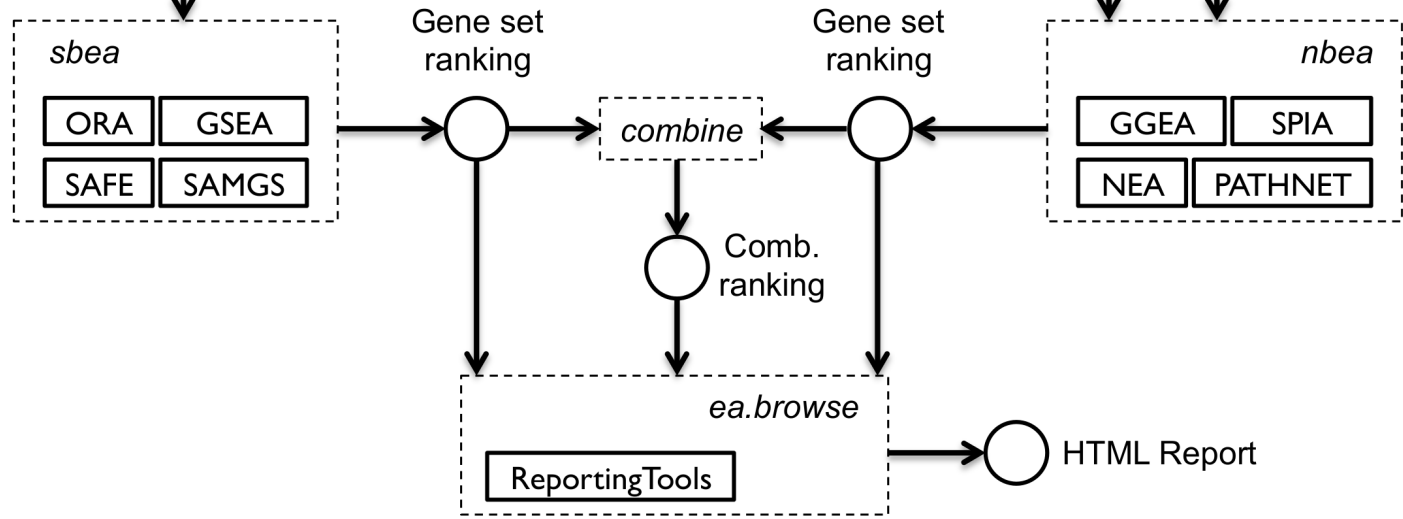
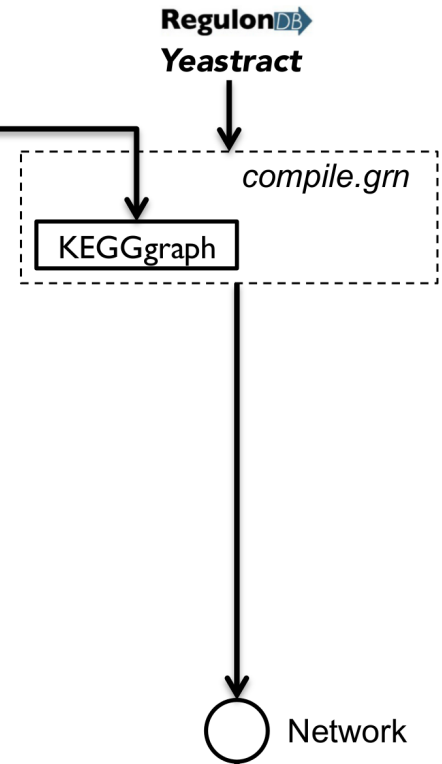
### EXPRESSION DATA



### GENE SETS / PATHWAYS



### REGULATORY NETWORK



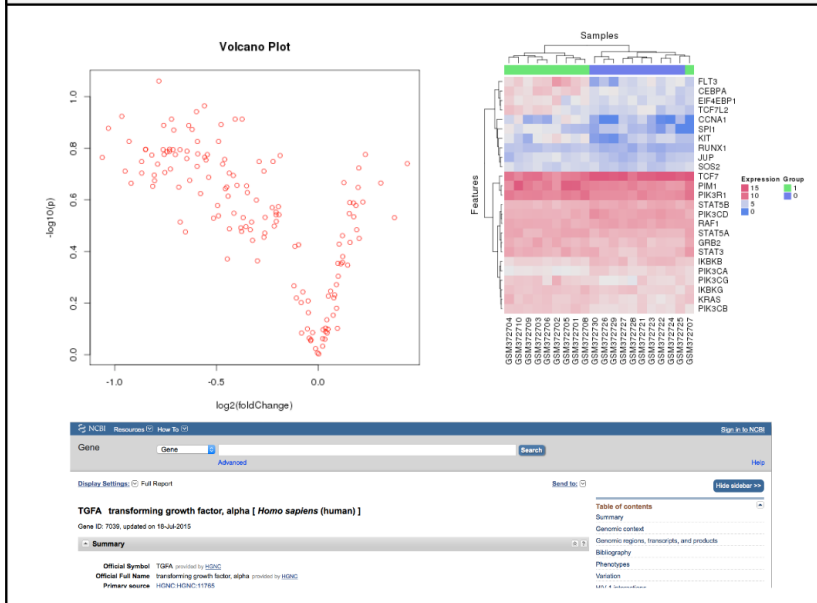
## Index page

- Overview graphics
- Gene report (DE)
- Full flat rankings

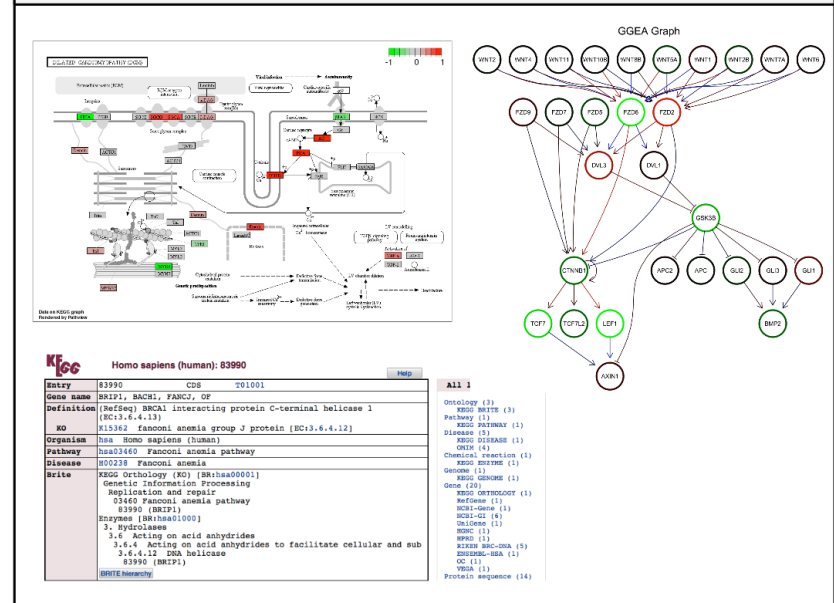
Enriched gene sets

Enriched gene sets

## SBEA page



## NBEA page



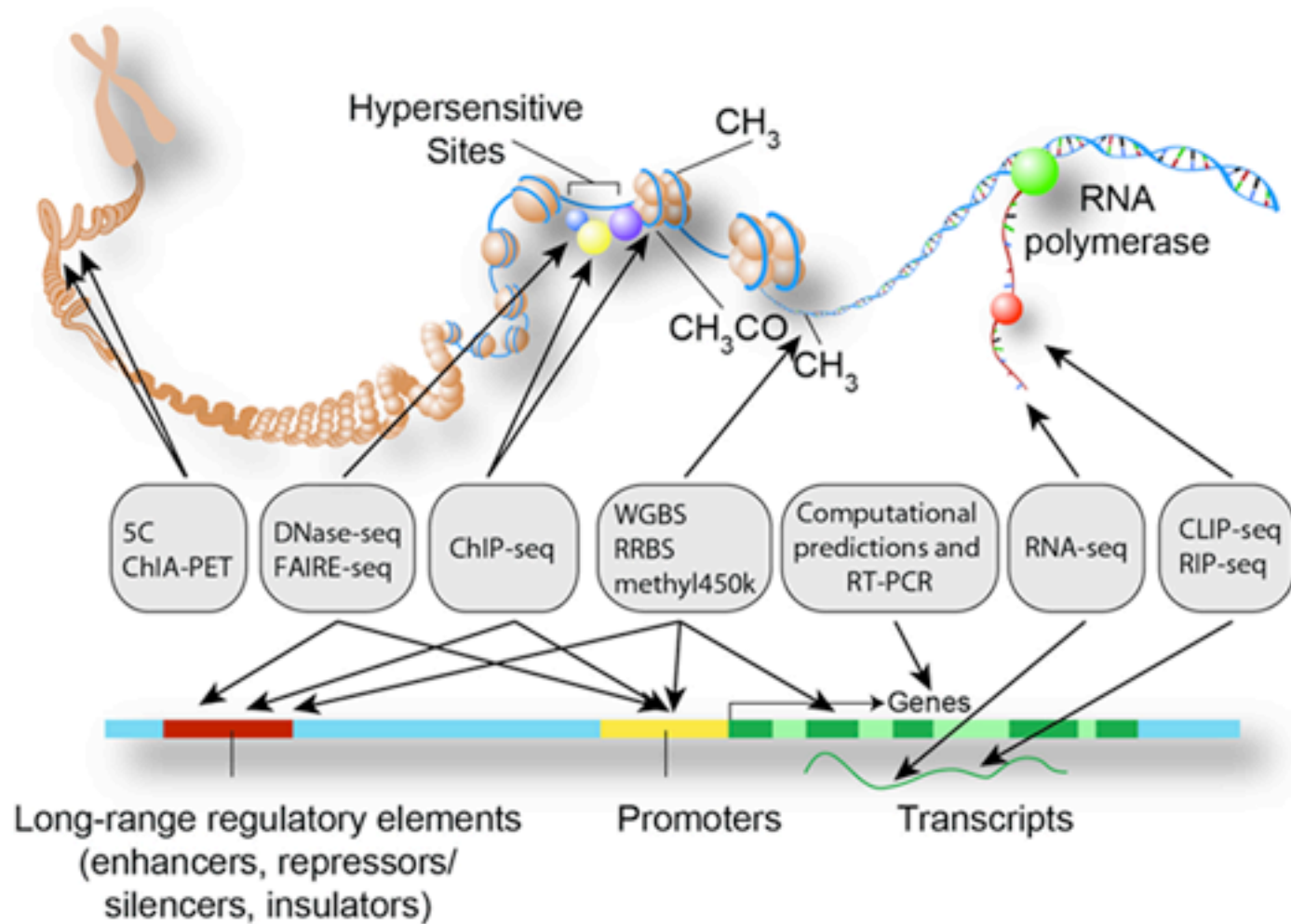
## Integrated view

- Individual & combined ranks
- Interactive sorting & filtering

# Contents

- 1 Gene expression-based enrichment analysis
- 2 Genomic region enrichment analysis
- 3 Multi-omics enrichment analysis

# ENCODE (Encyclopedia of DNA Elements)



# Setup: genomic region enrichment

- Experimentally-derived set of genomic regions
  - CNVs, CHIP-seq peaks, open chromatin, ...
- Functional feature set of genomic regions
  - Genes, QTLs, ...

→ *Enrichment of functional features in experimentally-derived regions?*

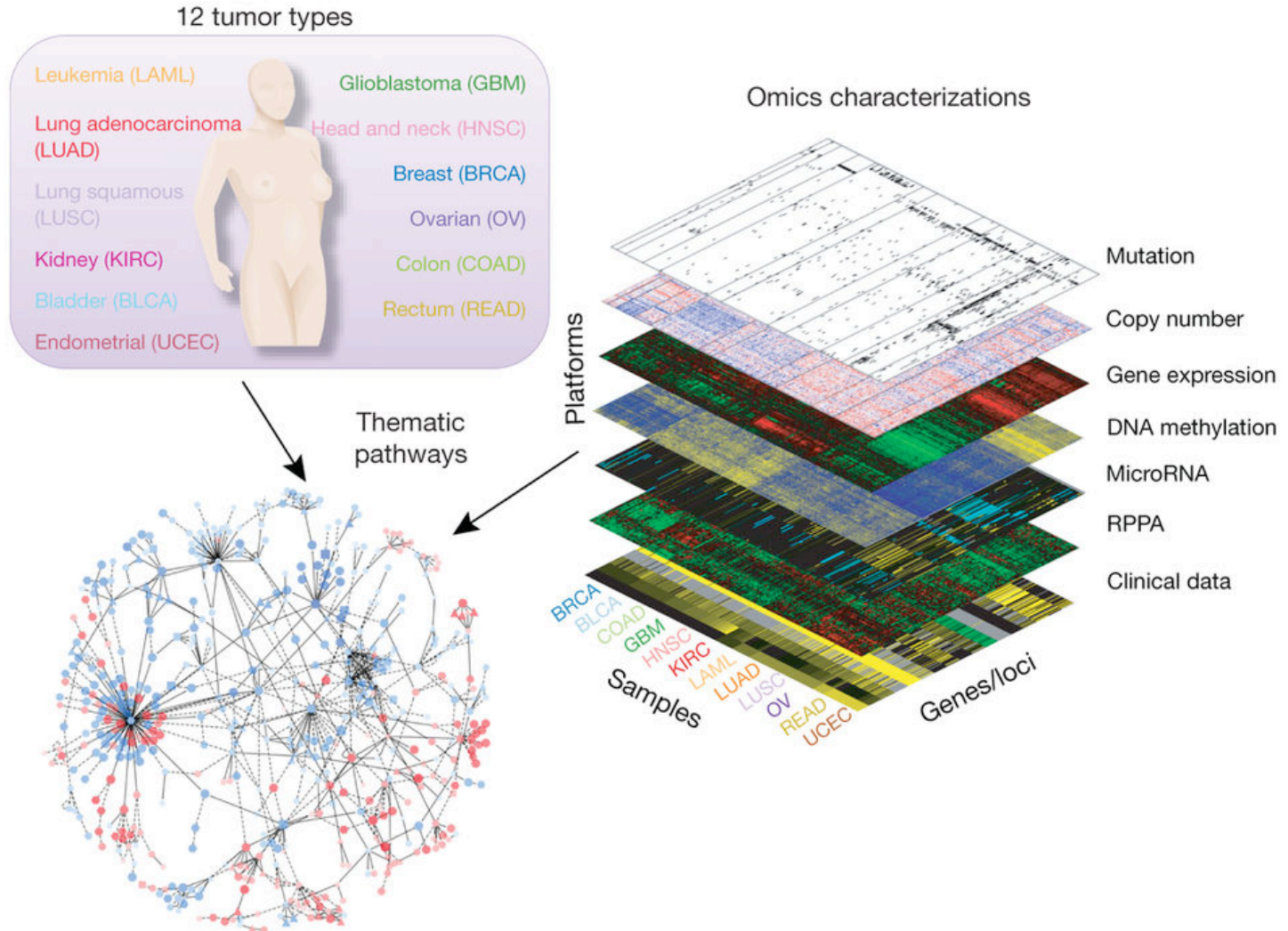
Packages: regioneR, LOLA, rGREAT, ...



# Contents

- 1 Gene expression-based enrichment analysis
- 2 Genomic region enrichment analysis
- 3 **Multi-omics enrichment analysis**

# TCGA (The Cancer Genome Atlas)



# Setup: multi-omics enrichment

- Multi-omics (enrichment) methods just begin to emerge
- Basic strategies
  - integrate & enrich vs. enrich & integrate
- *Relationships between data layers are not well understood*
  - *Cooperative (mRNA / protein), inhibitory (miRNA / mRNA), ...*

Packages: mogsa, RTopper, (...?)

Outside R/Bioc: PARADIGM, MONA