```
library(dplyr)

rladies_global %>%
  filter(city == 'Barcelona')
```

Bioinformatics with R

Or how R can help you and try to decipher omics data

Hello!

People call me Ania

I work at Institut Català d'Oncologia.

You can find me at @aniusbananius on Twitter, and @aniabmsi on GitHub

Agenda

1. Bioinformatics?
    a. Bioinformatics
    b. Omics
    c. Learning resources
2. What can R do?
    a. BioConductor
    b. Well-known packages
3. Time to play!
4. Final notes

1. Bioinformatics?

Bioinformatics

Biology + Informatics

Bioinformatics

Biology + Informatics

Genes

Bioinformatics

Statistics

Metabolomics

Biology + Informatics

Genes

Proteomics

Bioinformatics

Statistics

Genomics

Epigenomics

Metabolomics

GWAS

Biology + Informatics

Genes

Proteomics

Bioinformatics

EWAS

Sequencing

Phylogeny

Genomics

Statistics

Epigenomics

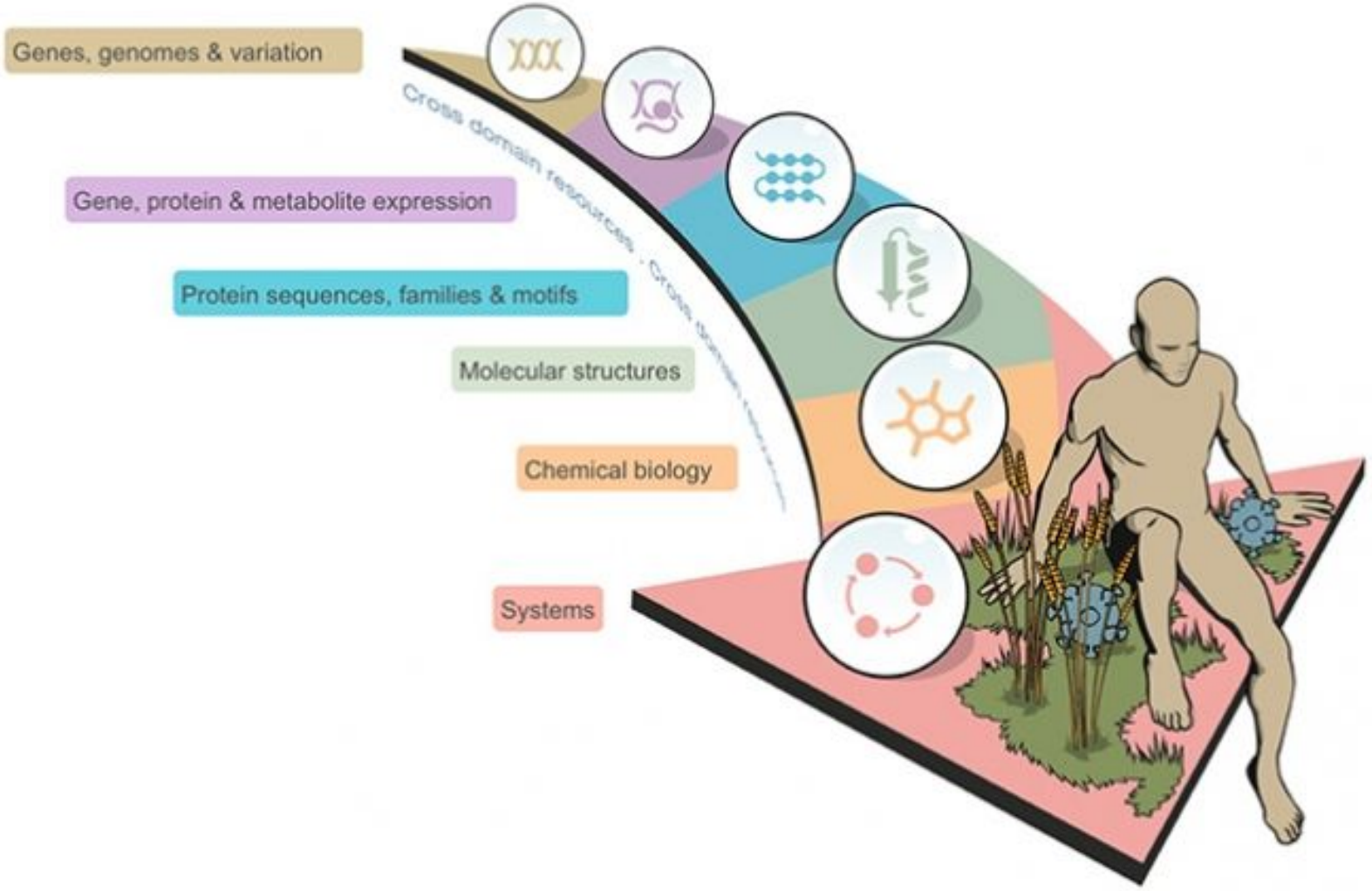
Biomarker discovery

Pathway analysis

Omics

Omics

The suffix -ome as used in molecular biology refers to a **totality of some sort**. *Omics aims at the collective characterization and quantification of pools of biological molecules that translate into the structure, function, and dynamics of an organism or organisms.*


Genes, genomes & variation
Gene, protein & metabolite expression
Protein sequences, families & motifs
Molecular structures
Chemical biology
Systems

Learning resources

Official (BCN)

- BSc Bioinformatics (UPF-UB-UPC)
- MSc Bioinformatics (UAB)
- MSc Bioinformatics for health sciences (UPF)
- MSc Statistics and operations research (UPC)
- MSc Omics data analysis (UVic)
- MSc Bioinformatics and biostatistics (UOC)
- MSc Big Data Processing for Life Sciences (UAB)
- MSc Medical Imaging and Applications (UdG)
- CFGS Desenvolupament d'aplicacions web
- Workshops (bioinformaticsBCN)

Unofficial

- Canadian Bioinformatics Workshops
- Long list of MOOCs
  - Bioinformatics specialisation (Coursera)
  - Genome data science specialisation (Coursera)
  - PH525.x Series (EdX)
  - Statistical analysis in bioinformatics (EdX)
- GitHub 🎉
  - https://github.com/open-source-society/bioinformatics
  - https://github.com/danielecook/Awesome-Bioinformatics
  - ...
- Bioconductor courses & conferences
- ...

2. What can R do for us?

Or at least it tries!

And if you don’t find a solution, you can always create it! 🐌

Bioconductor (BioC)

BioC provides tools for the analysis and comprehension of high-throughput genomic data.

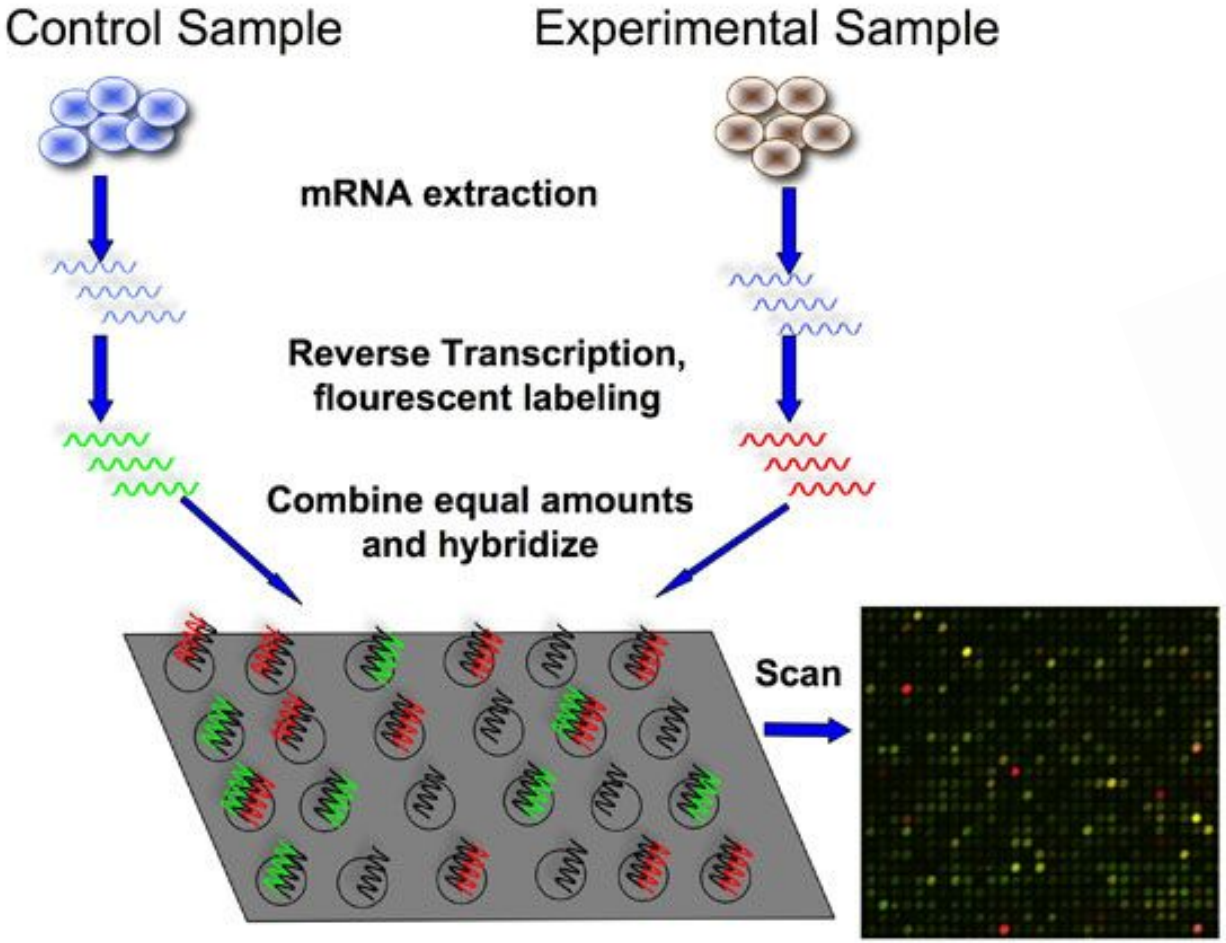
Bioconductor 3.5

- 1383 software packages
- 316 experiment data packages
- 911 annotation packages.

Well-known packages

- Microarray
  - Limma
  - Affy
  - Oligo
- Sequencing (Differential expression)
  - DESeq / DESeq2
  - EdgeR
- MVA: mixOmics
- Annotations: annotate
- Phylogeny: phyclust
- PPI: ppiData, ppiStats
- Networks: igraph, BioNet, dnet

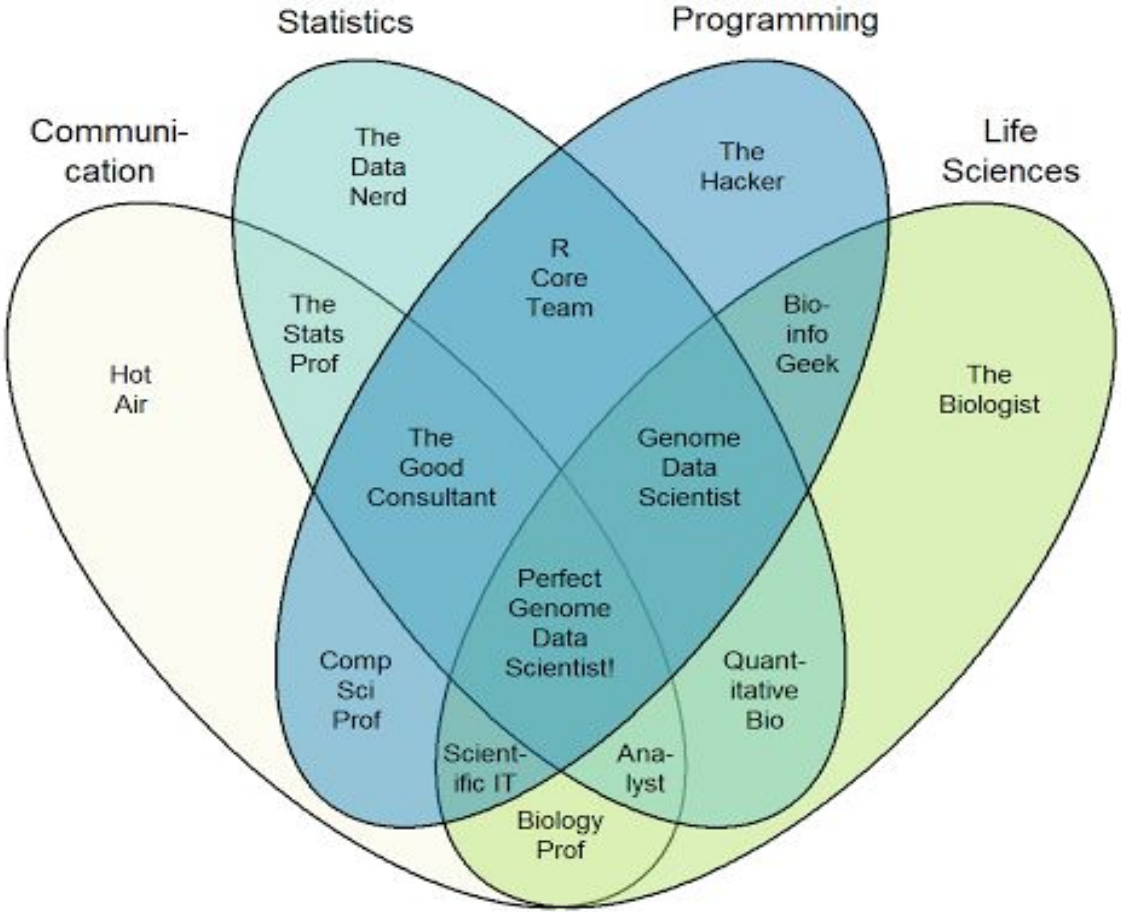
3. Time to play!


Control Sample
Experimental Sample
mRNA extraction
Reverse Transcription,
flourescent labeling
Combine equal amounts
and hybridize
Scan

GeneChip
Affymetrix

Switch to R

4. Final notes



http://datascience.stackexchange.com/questions/2403

© F. Supek

Want to share your experience with R?

bcn@rladies.org

“ *Tomorrow, 18h45, King Offices*

R Users Group:

Automatic tools for improving packages 🎉