

Complex Networks in Systems Biology

Introduction – Public Resources

Costas Bouyioukos

UMR7216, Paris Epigénétique et Destin Cellulaire
Université Paris Diderot

M2BI, November 2018



Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built

Proteomics
Transcriptomics
Bioinformatics

Databases

Graph Theory
Introduction
Analysis

Closing Remarks

Overview of the course

Introduction - Logistics

Practical

Course Contents

Structural elements and interacting partners in Complex Biological Networks

Protein interaction Networks

Protein Nucleic Acids networks

Metabolic Networks

Methods to generate and build networks

Proteomics

Transcriptomics

Bioinformatics

Public databases for network biology

Introduction to Graph Theory

Introduction

Analysis

Closing Remarks

Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built
Proteomics
Transcriptomics
Bioinformatics

Databases

Graph Theory
Introduction
Analysis

Closing Remarks

Complex Networks in Systems Biology

Organisation of the Module

- Instructors :

- Costas Bouyioukos, Paris-Diderot
(Complex Biomolecular Networks)
- Karine Auduze, Paris-Diderot
(Disease and toxicology networks)

- Duration 2 weeks full-time :

- 6 CMs with 5 TD-TPs.
- 2 means of evaluation : One presentation and one TP project.

- Contacts

- email costas.bouyioukos@univ-paris-diderot.fr
- email karine.auduze@univ-paris-diderot.fr

Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built

Proteomics
Transcriptomics
Bioinformatics

Databases

Graph Theory
Introduction
Analysis

Closing Remarks

Complex Networks in Systems Biology

CM1 Introduction, types and resources for generating biomolecular networks, (smooth) introduction to graph theory.

CM2 Network structure, global measures, topological analysis, local measures.

CM4 Modelling, clustering and modularity.

CM8 Network inference.

TPs

Practice with some well known software in network analysis and present a (really) small project.

Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built
Proteomics
Transcriptomics
Bioinformatics

Databases

Graph Theory
Introduction
Analysis

Closing Remarks

Types and resources for network generation and introduction to Graph theory

1. Interacting (biological) partners and their mode(s) of interaction.
2. Biological and computational techniques to acquire data to build networks.
3. On-line resources to acquire data.
4. Smooth introduction to graph theory.

Bonus TP, seance libre :

First experience with the most popular network analysis and visualisation program Cytoscape.

You can download it for free here :

<http://www.cytoscape.org/download.php>

Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built

Proteomics
Transcriptomics
Bioinformatics

Databases

Graph Theory

Introduction
Analysis

Closing Remarks

Interacting Partners

1. Proteins <→ Proteins

- Interactions between proteins :
 - Protein Complexes
 - Other (mostly transient) interactions, phosphorylations, signal transduction etc.
- studied by **Proteomics**

2. Enzymes <→ Metabolites

- Interactions between enzyme and reactants – products in biochemical reactions (metabolites).
- studied by : **Metabolomics**

3. Nucleic acids <→ Proteins

- Interactions between nucleic acids and proteins.
Sequences vs. structural and/or regulatory factors.
- studied by many-**omics**

4. Heterogeneous networks

- Combination and juxtaposition of one or more of the above interacting partners.

Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built

Proteomics
Transcriptomics
Bioinformatics

Databases

Graph Theory
Introduction
Analysis

Closing Remarks

Protein interaction networks

Protein-Protein Interaction (PPI) or Protein Interaction Networks (PIN)

Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built
Proteomics
Transcriptomics
Bioinformatics

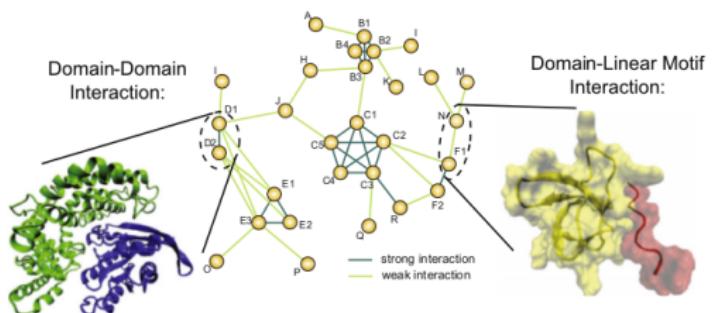
Databases

Graph Theory
Introduction
Analysis

Closing Remarks

- Proteins that interact to each other can be represented by a network.

- Two modes of interactions stable (complexes) and transient (de/activations).

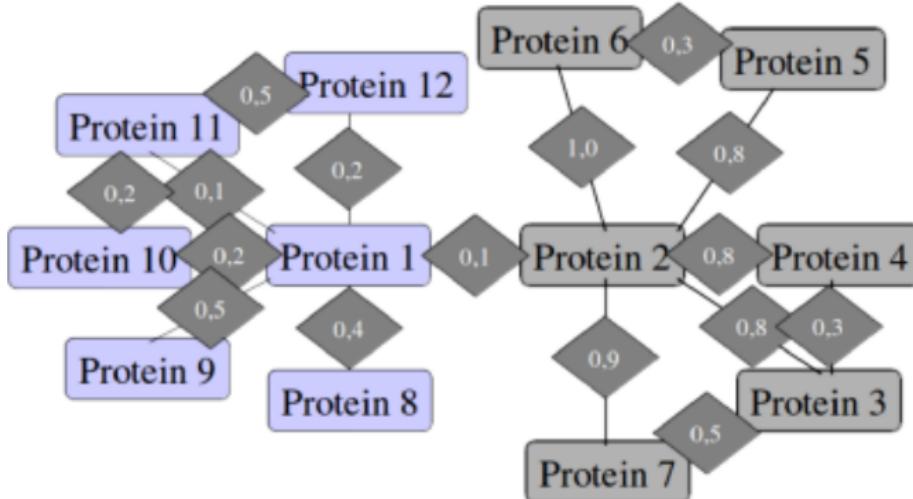


- Domain length from 25 - 500 AA
- Affinities: K_D nM to μ M
- Rather stable interactions
- Examples: BTB(POZ), Ras-GAP, CARD

- Motif length from 3 - 10 AA
- Affinities: $K_D \sim \mu$ M
- Rather transient interactions
- Examples: Sh3/PxxP, EVH1/FPPPP

Protein interaction networks

Protein-Protein Interaction (PPI) or Protein Interaction Networks (PIN)



- Strength of interactions -> Protein complexes
- PPIs are **weighted, undirected** networks.

Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built

Proteomics
Transcriptomics
Bioinformatics

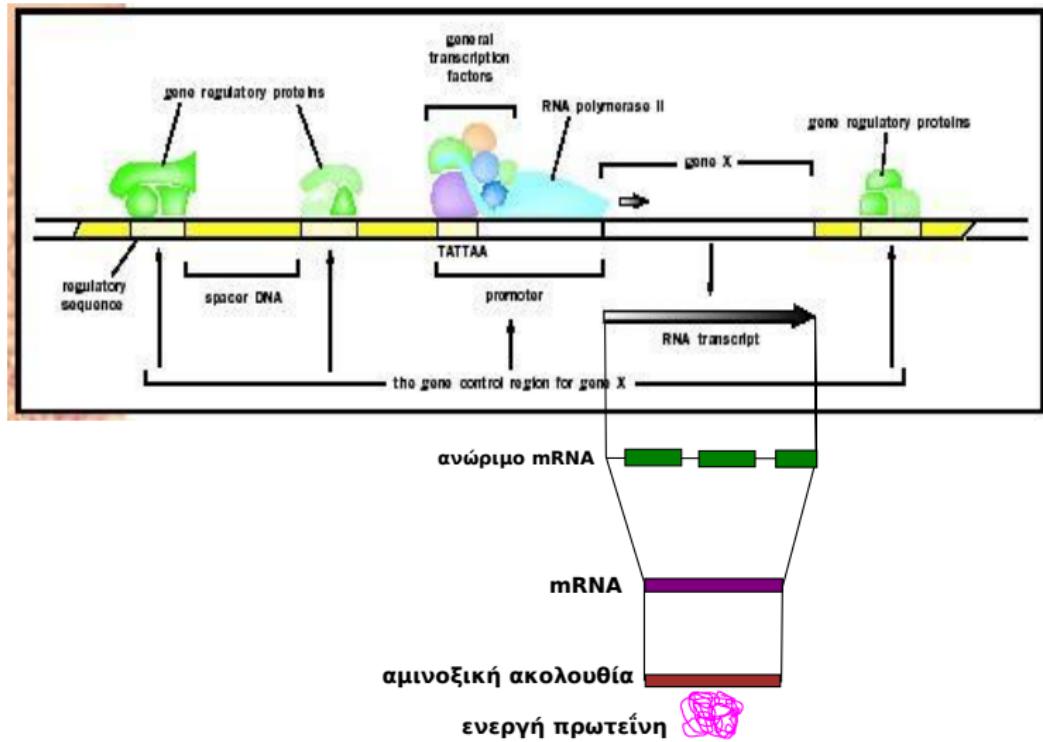
Databases

Graph Theory
Introduction
Analysis

Closing Remarks

Interactions between DNA and regulatory factors

Transcription Regulation



Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built
Proteomics
Transcriptomics
Bioinformatics

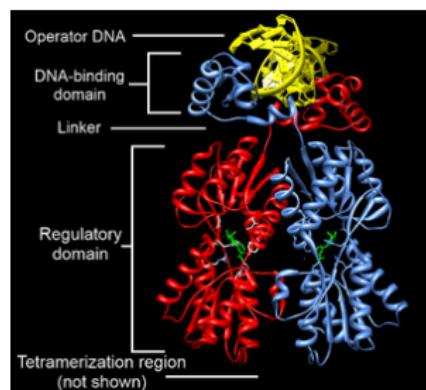
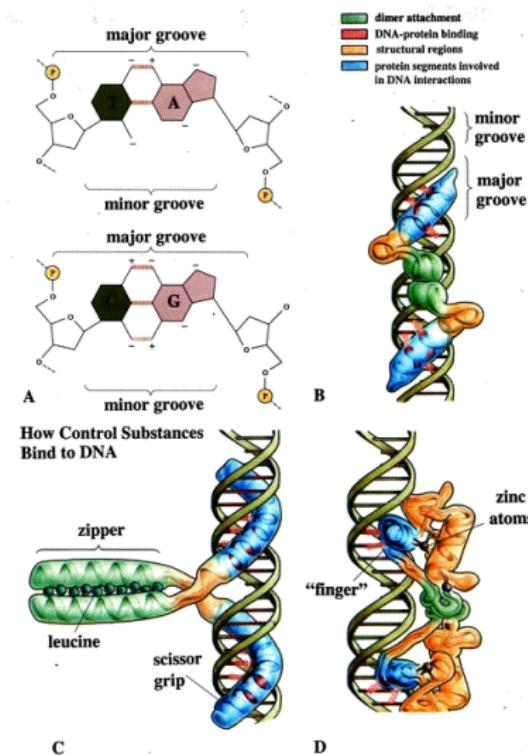
Databases

Graph Theory
Introduction
Analysis

Closing Remarks

Interactions between DNA and regulatory factors

Binding



Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built

Proteomics
Transcriptomics
Bioinformatics

Databases

Graph Theory
Introduction
Analysis

Closing Remarks

Interactions between DNA and regulatory factors

Updated Dogma of Transcription Regulation

Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

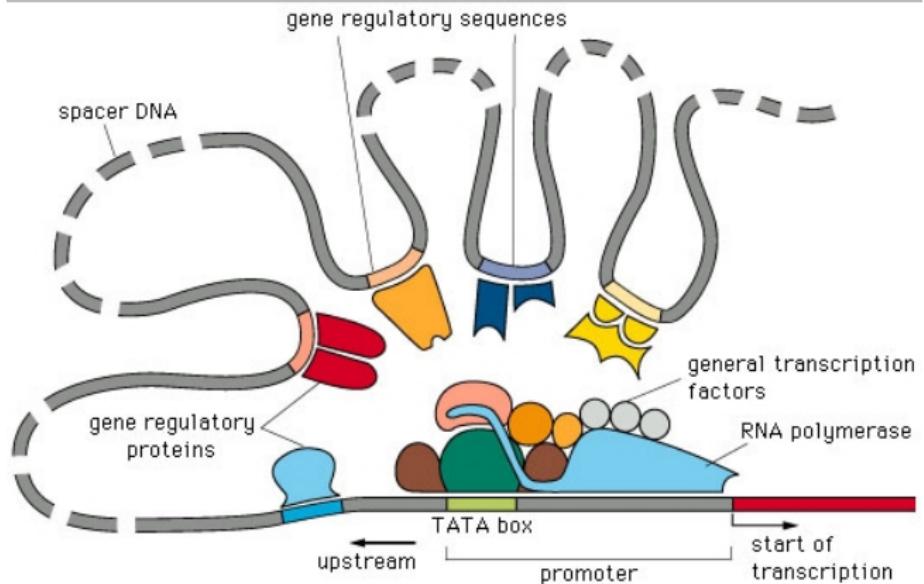
Network Built
Proteomics
Transcriptomics
Bioinformatics

Databases

Graph Theory
Introduction
Analysis

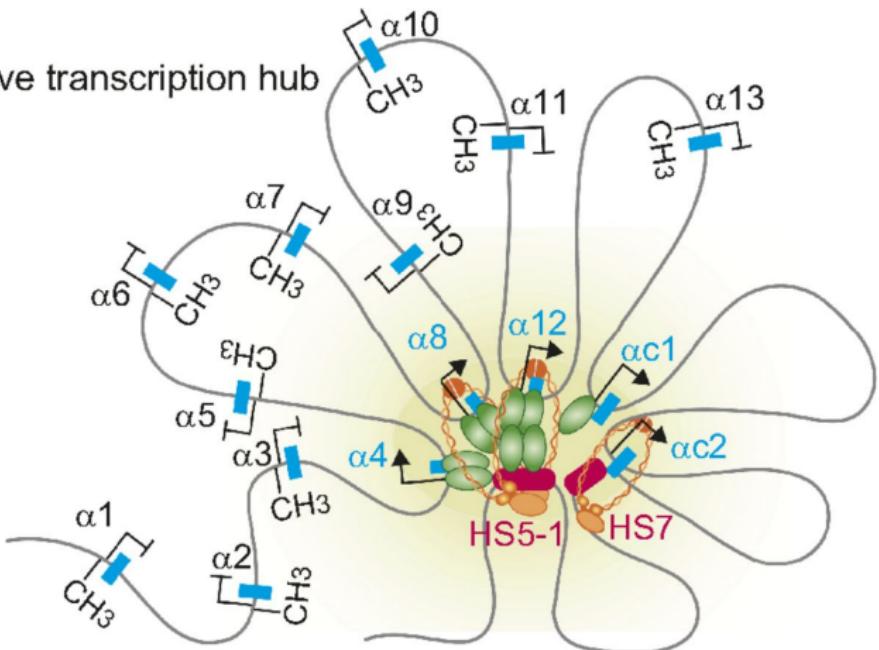
Closing Remarks

Revisiting gene regulation
... in reality this is what happens (?)



Transcription in 3 dimensions

Active transcription hub



Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built

Proteomics
Transcriptomics
Bioinformatics

Databases

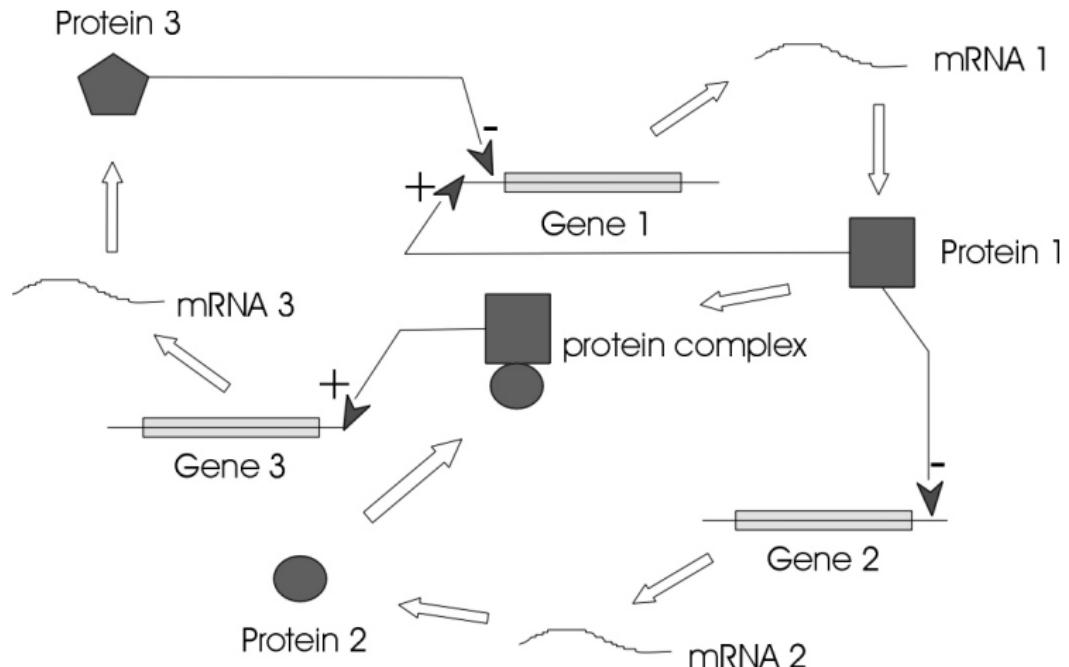
Graph Theory

Introduction
Analysis

Closing Remarks

Genetic Regulatory Networks

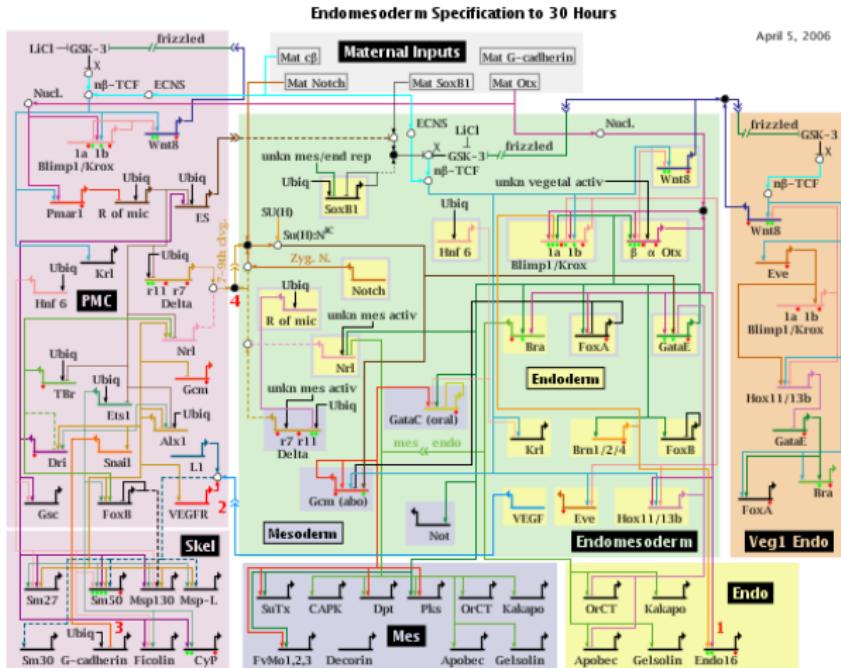
Composition



- GRNs consists of interactions between genes and their products.

Genetic Regulatory Networks (GRN)

Representation I, evo-devo



Ubiqu = ubiquitin; Mat = maternal; activ = activator; rep = repressor;
 unk = unknown; Nucl = nuclearization; X = β -catenin source;
 n β -TCF = nuclearized β -catenin-Tcf1; ES = early signal;
 ECNS = early cytoplasmic nuclearization system; Zyg. N. = zygotic Notch

Copyright © 2001–2006 Hamid Beloumi and Eric Davidson

- Developmental stages of the sea urchin embryo.

Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built
Proteomics
Transcriptomics
Bioinformatics

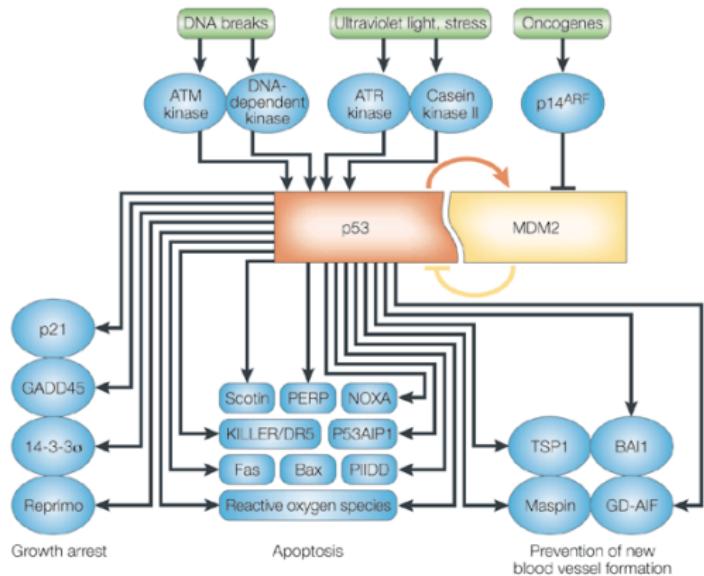
Databases

Graph Theory
Introduction
Analysis

Closing Remarks

Genetic Regulatory Networks (GRN)

Representation II, control



- The GRN of the p53 oncogene (“hub” or highly connected node).

Terminology

GRN refers to both “Gene Regulatory Networks” but also to “Genetic Regulatory Networks”. The terms are (almost) interchangeable, for this course I will use the former.

Introduction - Logistics

Practical
Contents

Structural Elements

PPI
GRNs
Metabolic

Network Built

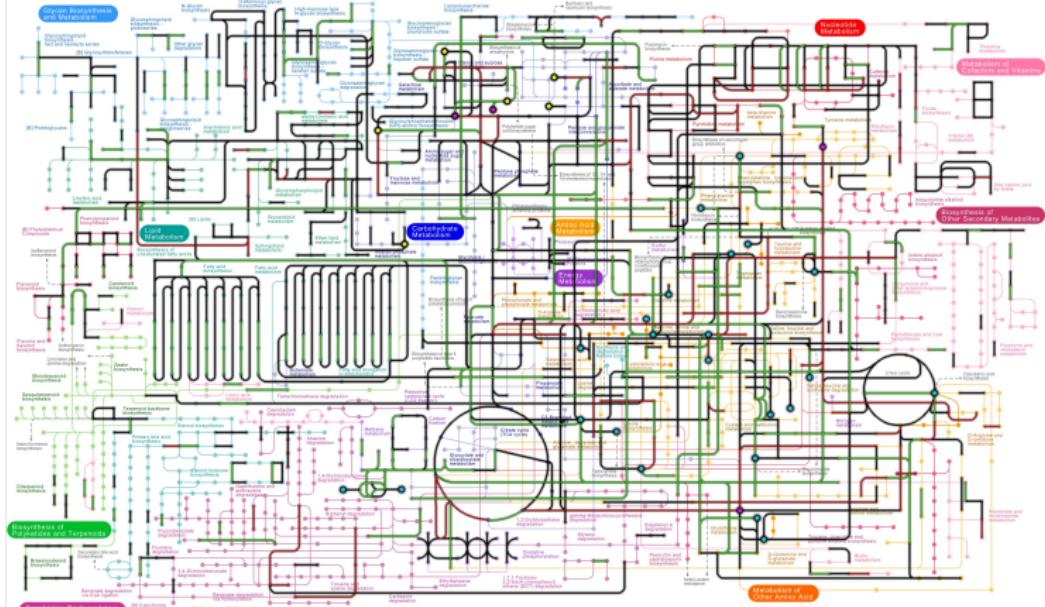
Proteomics
Transcriptomics
Bioinformatics

Databases

Graph Theory
Introduction
Analysis

Closing Remarks

Metabolic Networks – Metabolomics



Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built

Proteomics
Transcriptomics
Bioinformatics

Databases

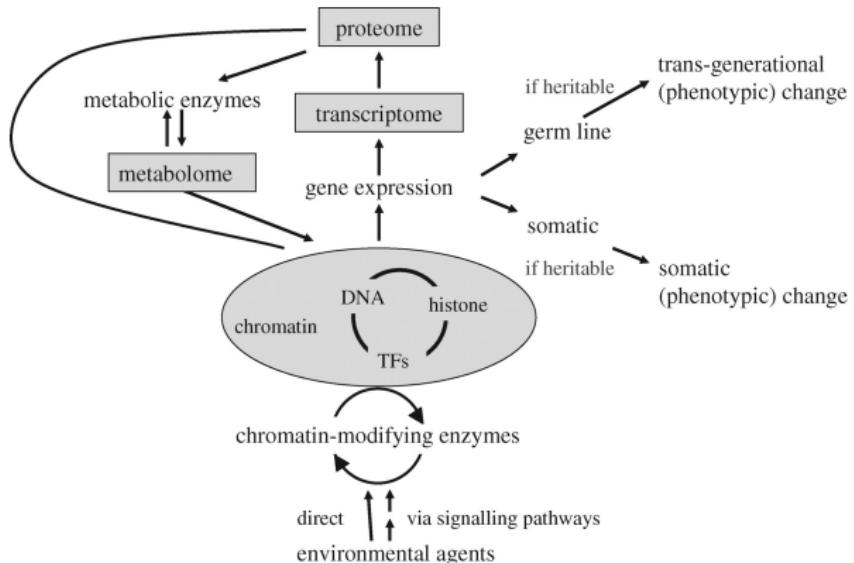
Graph Theory
Introduction
Analysis

Closing Remarks

If you are impressed you can enjoy this picture in extremely high resolution and zoom and inspect every metabolic reaction you like, the link is [here](#).

Information flow in biology

Epigenetics-epigenomics



“Central Dogma” in biology revisited !

Interwoven, multi-layered, epi-genomic, complex biomolecular networks control all biological processes in multiple levels of regulation.

Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built

Proteomics
Transcriptomics
Bioinformatics

Databases

Graph Theory

Introduction
Analysis

Closing Remarks

Data generation methods to build biological networks

...omics everywhere

x-omics technologies

The advancement of high-throughput methods to acquire biological data has generated vast amounts of data coming from :

1. Proteins – Proteomics
2. Genomes – Genomics
3. Metabolites, enzymes – Metabolomics
4. RNAs (mRNA, ncRNA etc.) – Transcriptomics

Omics and... badomics

In parallel, the new trend for the usage of the suffix -omics has led to some abuses, so be careful with the use of the term!

(e.g. have a look at the hashtag #badomics. Latest #badomics catchword *consciousome...*)

Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built

Proteomics
Transcriptomics
Bioinformatics

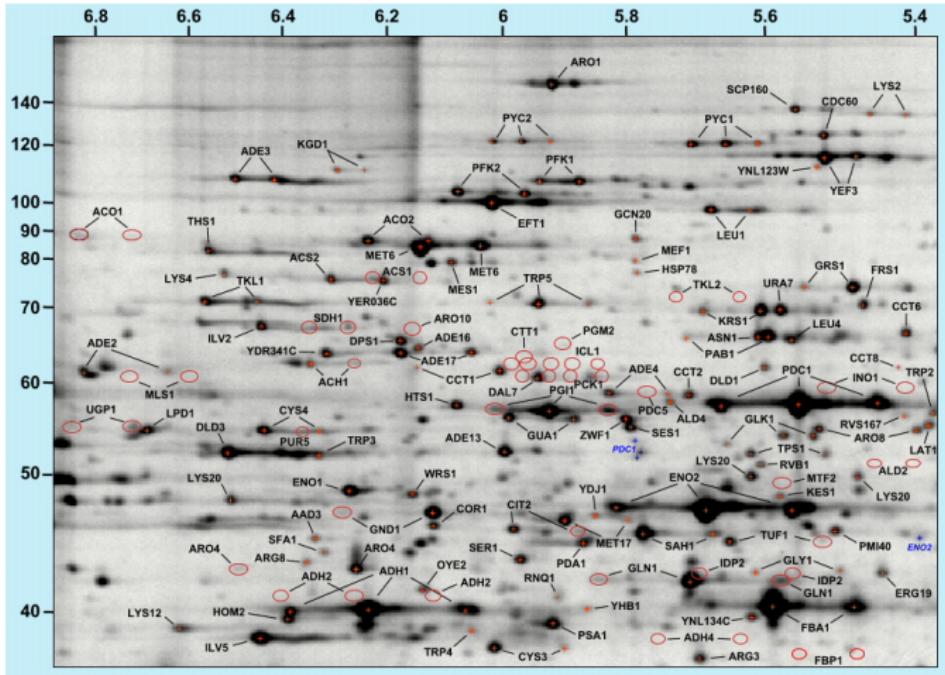
Databases

Graph Theory
Introduction
Analysis

Closing Remarks

Proteomics

2D protein gel electrophoresis



- Separation, identification (based on charge and size) and quantification of proteins in a biological sample.

Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built
Proteomics
Transcriptomics
Bioinformatics

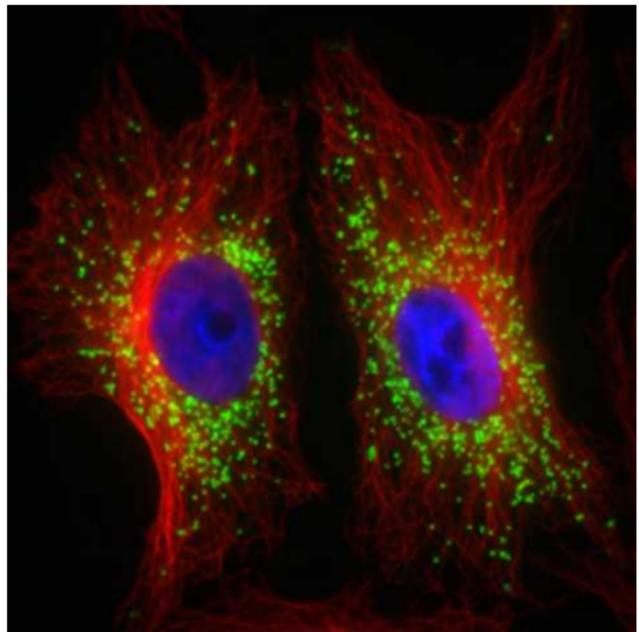
Databases

Graph Theory
Introduction
Analysis

Closing Remarks

Proteomics

Intra-cellular layout



- Interactions are realised in space, compartments of the cell.
- Labelled cDNA with GFP fusion.
- Image : 1057 labelled human ORFs.
<http://gfp-cdna.embl.de/>
- Fractionation techniques as well as Hi-C, DNA and RNA-FISH provides data for the spatial realisation of biomolecular net

Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built
Proteomics
Transcriptomics
Bioinformatics

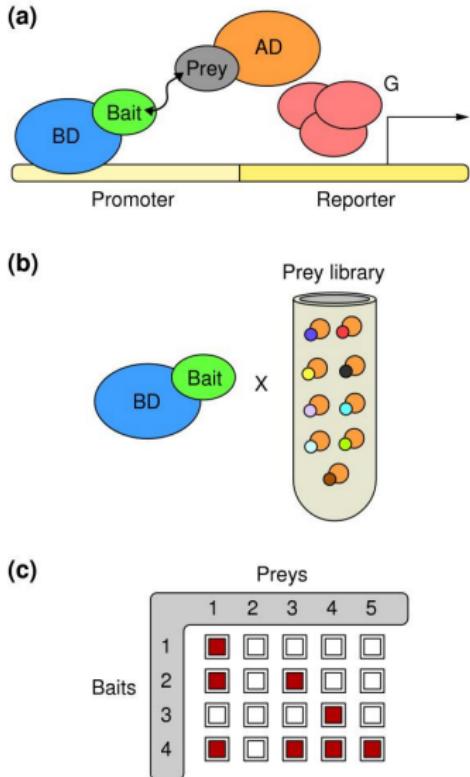
Databases

Graph Theory
Introduction
Analysis

Closing Remarks

Proteomics

Yeast 2-hybrids system



- Evidence of interaction between bait-prey from the reporter gene expression.
- Promoter UAS, Gal4 galactose metabolism transcription factor.
- High-throughput prey library, scan genome scale.
- False positives.

Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built

Proteomics
Transcriptomics
Bioinformatics

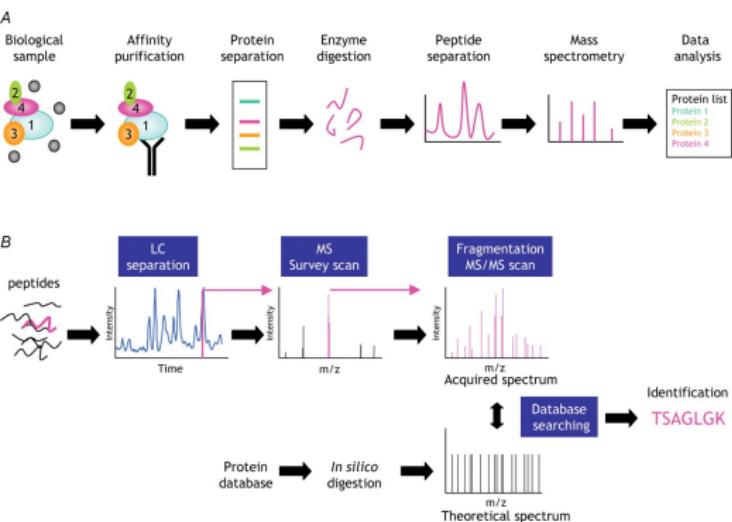
Databases

Graph Theory
Introduction
Analysis

Closing Remarks

Proteomics

High-throughput mass chromatography MS



Immunoprecipitation
-> MS, IP-MS.

■ Tandem Affinity
Purification -> MS,
TAP-MS.

■ Liquid Crystal
Chromatography ->
MS, LC-MS.

Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built

- Proteomics
- Transcriptomics
- Bioinformatics

Databases

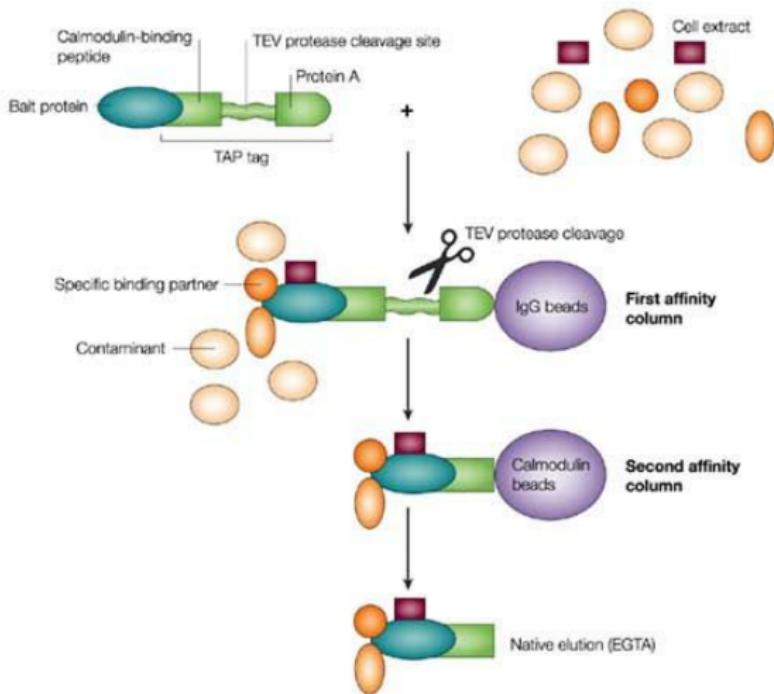
Graph Theory

- Introduction
- Analysis

Closing Remarks

Proteomics

Tandem Affinity Purification



Nature Reviews | Molecular Cell Biology

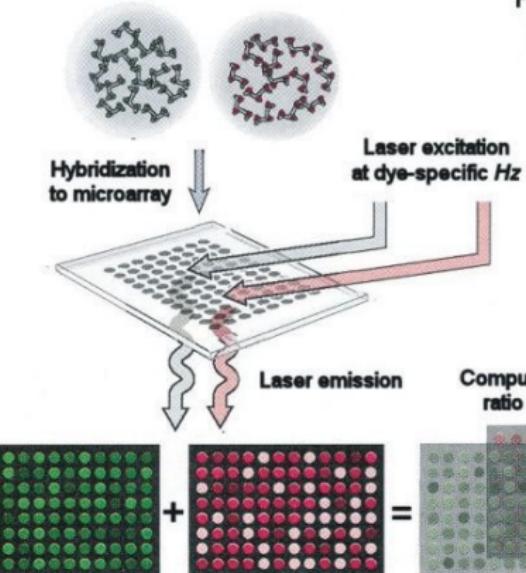
Transcriptomics

Complementary DNA microarrays

Make cDNA reverse transcript
Label cDNAs with fluorescent dyes

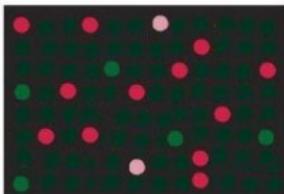
Control

Experimental



Principle of cDNA microarray assay for gene expression
(after Gibson & Muse 2002)

Red = "up-regulation"
Green = "down-regulation"
Black = constitutive expression



- Genes that are under- or over-expressed under ONE condition.

Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built

Proteomics
Transcriptomics
Bioinformatics

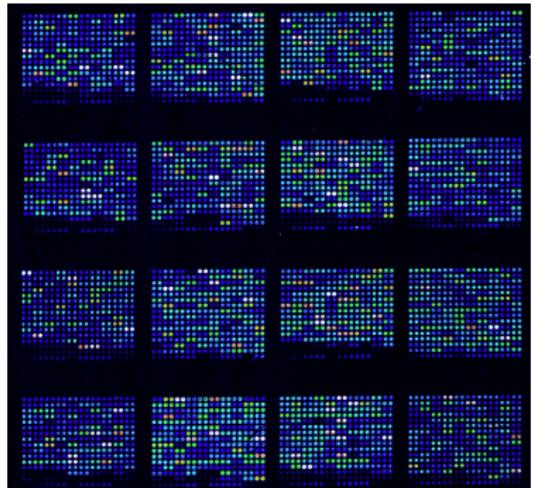
Databases

Graph Theory
Introduction
Analysis

Closing Remarks

Transcriptomics

Oligonucleotides biochip



1. Canonicalisation of measurements
2. Statistical significantly expressed genes.
3. Timeseries data.
4. Clustering (hierarchical).
5. Network reconstruction.

Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built

Proteomics
Transcriptomics
Bioinformatics

Databases

Graph Theory

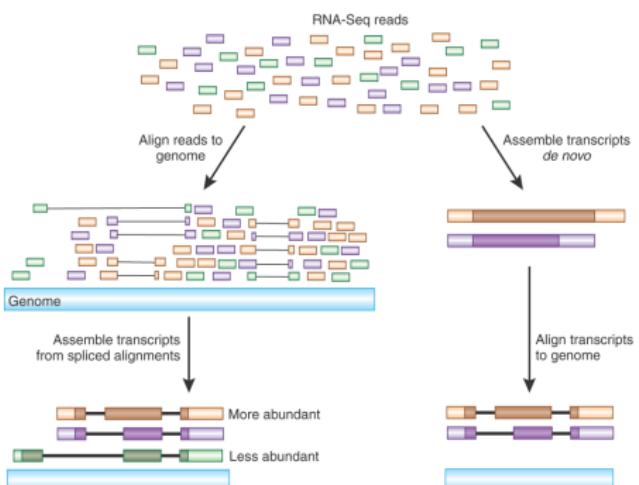
Introduction
Analysis

Closing Remarks

Transcriptomics

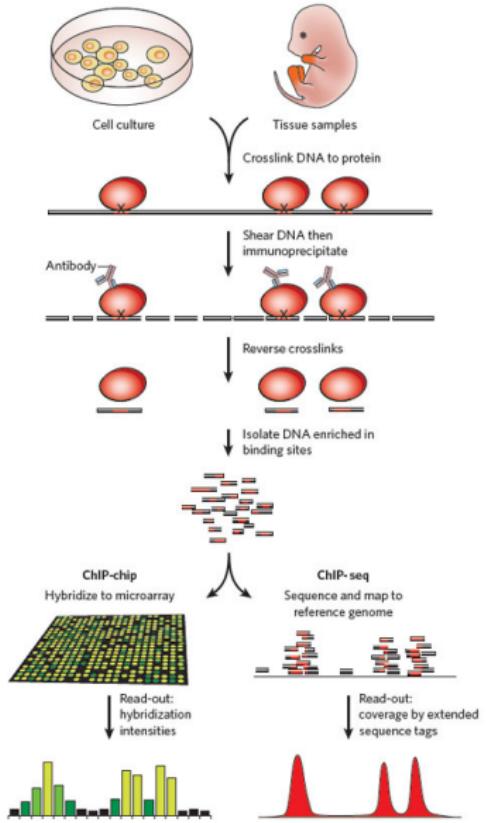
RNA NGS sequencing

The advent of next generation sequencing techniques (NGS) allows us to read an extremely huge amount of data.



- Extract of whole RNA.
- Preparation of cDNA and high-throughput sequencing.
- Alignment of short RNA reads on the reference genome. Estimation of gene expression level.
- More sensitive technique than the microarrays. Allows the study of “open questions”.

Transcriptomics - ChIP-Chip, ChIP-Seq



Chromatin Immunoprecipitation

- ChIP-Chip (microarray)
- ChIP-Seq (NGS sequencing)

Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built
Proteomics
Transcriptomics
Bioinformatics

Databases

Graph Theory
Introduction
Analysis

Closing Remarks

Transcriptomics

Other techniques

- RT-qPCR
Highly sensitive and reproducible !
- Sage
(Serial Analysis of Gene Expression)

Constrains

- Every experiment based on kinetics must be followed by a computational verification (inference) step, such that a statistically significant model of transcription interactions can be obtained.
- ...and followed ALSO by the suitable statistical tests to accept/reject the hypothesis of gene differential expression.

Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built

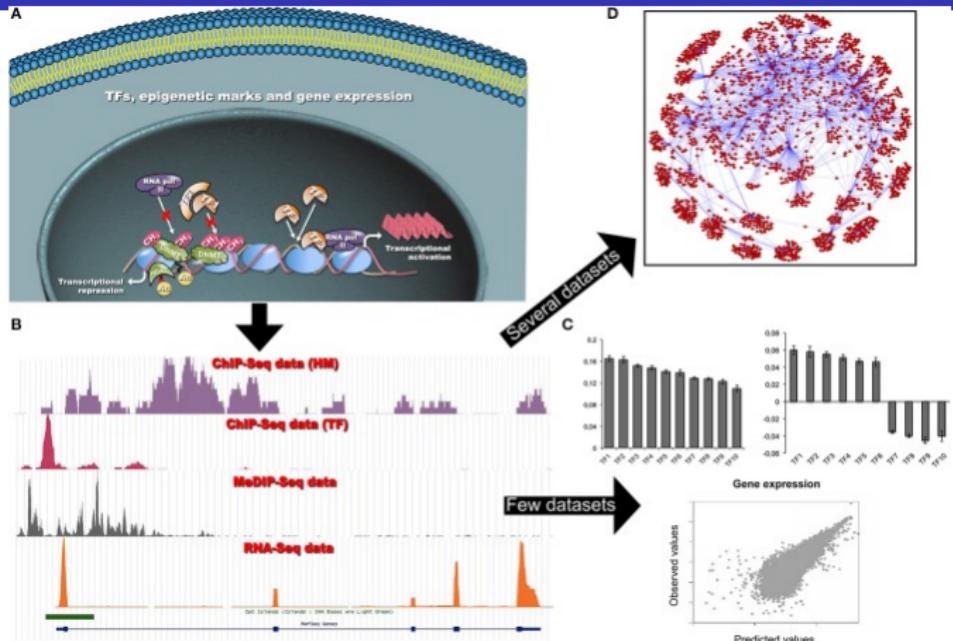
Proteomics
Transcriptomics
Bioinformatics

Databases

Graph Theory
Introduction
Analysis

Closing Remarks

Transcriptomics - Integration



Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built

Proteomics
Transcriptomics
Bioinformatics

Databases

Graph Theory

Introduction
Analysis

Closing Remarks

Integrating multi-Omics

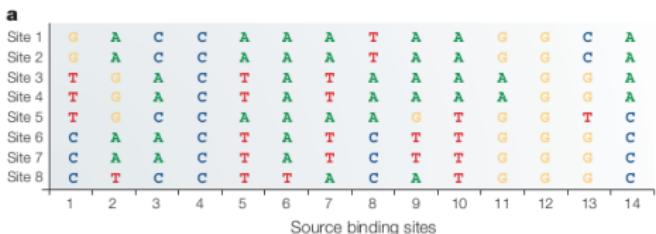
- Few datasets - PCA, regression models.
- Many datasets - Network Inference.

Computational Methods

Comparative and Computational Genomics I

Position Frequency Matrix PFM

- Different Transcription Factor Binding Sites (TFBS) alignment.



- Calculation of the frequency of appearance of each nucleotide in each position.

c Position frequency matrix (PFM)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
A	0	4	4	0	3	7	4	3	5	4	2	0	0	4
C	3	0	4	8	0	0	0	3	0	0	0	0	2	4
G	2	3	0	0	0	0	0	0	1	0	6	8	5	0
T	3	1	0	0	5	1	4	2	2	4	0	0	1	0

Introduction - Logistics

Practical Contents

Structural Elements

PPI

GRNs

Metabolic

Network Built

Proteomics

Transcriptomics

Bioinformatics

Databases

Graph Theory

Introduction

Analysis

Closing Remarks

Computational Methods

Comparative and Computational Genomics II

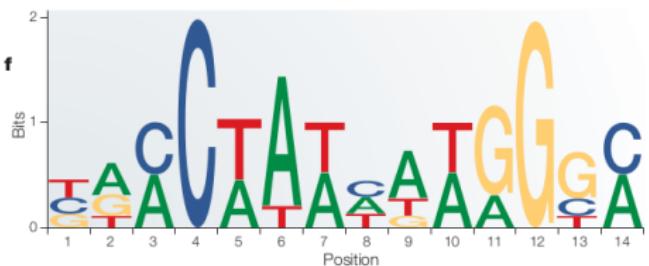
Position Weight Matrices PWM

- Log-transformation of the PFM.

d Position weight matrix (PWM)

	A	C	G	T	A	C	G	T	A	C	G	T	A	C	G	T
A	-1.93	0.79	0.79	-1.93	0.45	1.50	0.79	0.45	1.07	0.79	0.00	-1.93	-1.93	0.79		
C	0.45	-1.93	0.79	1.68	-1.93	-1.93	-1.93	0.45	-1.93	-1.93	-1.93	0.00	0.79			
G	0.00	0.45	-1.93	-1.93	-1.93	-1.93	-1.93	-1.93	0.66	-1.93	1.30	1.68	1.07	-1.93		
T	0.15	0.66	-1.93	-1.93	1.07	0.66	0.79	0.00	0.00	0.79	-1.93	-1.93	-0.66	-1.93		

- The information content (in bytes) of each nucleotide in each position is represented by sequence logos.



Introduction - Logistics

Practical Contents

Structural Elements

PPI

GRNs

Metabolic

Network Built

Proteomics

Transcriptomics

Bioinformatics

Databases

Graph Theory

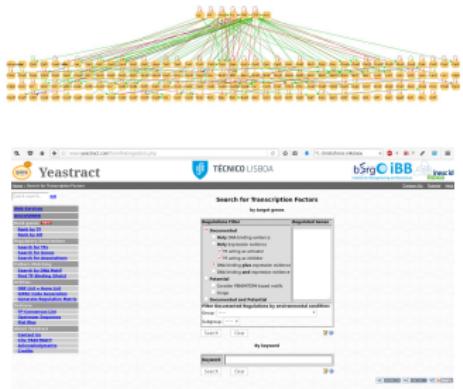
Introduction

Analysis

Closing Remarks

Genetic Regulatory networks databases

For model organisms yeast and *E.coli*



- The most well curated transcriptional regulatory information is stored in RegulonDB
regulondb.ccg.unam.mx/
- For yeast a bigger wealth of information is available, YEASTRACT contains its GRN
www.yeastRACT.com/
- For a collection of species Network Portal provides a good resource (13 species in Nov. 2016)
networks.systemsbiology.net/

Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built
Proteomics
Transcriptomics
Bioinformatics

Databases

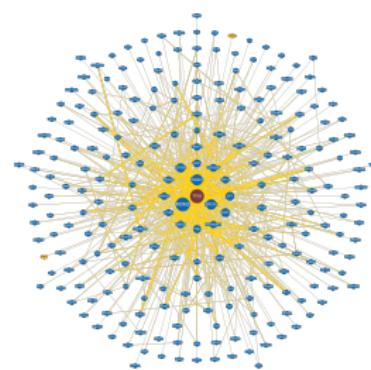
Graph Theory
Introduction
Analysis

Closing Remarks

PPI networks databases

For yeast and human

- Historical pioneers BIND and MINT.
- The Database of Interacting Proteins, experimental data highly curated.
dip.doe-mbi.ucla.edu/
- The most comprehensive is STRING 200 organisms, 2M interactions.
string-db.org/
- BioGRID a generic database comprising different types of interactions.
thebiogrid.org/



Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built
Proteomics
Transcriptomics
Bioinformatics

Databases

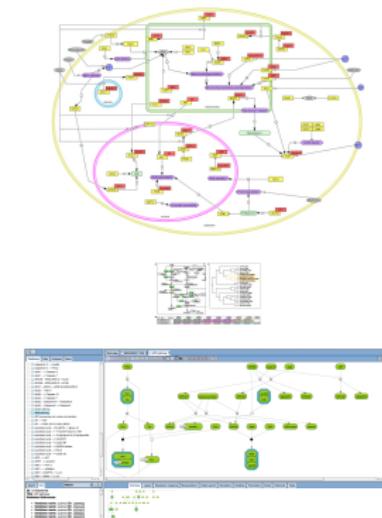
Graph Theory
Introduction
Analysis

Closing Remarks

Pathway and GRN databases

For many organisms

- The Kyoto Encyclopedia of Genes and Genomes ([KEGG](#))
- The [REACTOME](#) database. Manually curated pathway database.
- The [Pathguide](#) resource for pathways and molecular interactions in BioPAX, CellML, SVML etc.
- The [TransPath](#) database of signal transduction and GRNs contains half a million interactions.



Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built

Proteomics
Transcriptomics
Bioinformatics

Databases

Graph Theory

Introduction
Analysis

Closing Remarks

Introduction - What is a graph/network



- Is a collection of vertices (nodes) and edges (links).

- **NETWORK** : Is in fact a representation map of relationships between entities.

Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built

Proteomics
Transcriptomics
Bioinformatics

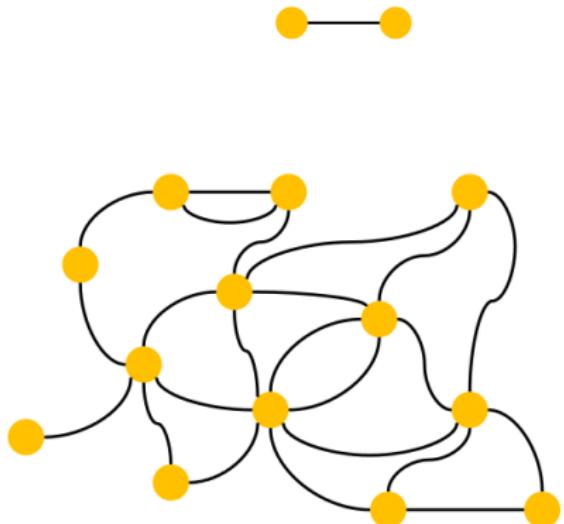
Databases

Graph Theory

Introduction
Analysis

Closing Remarks

Introduction - What is a graph/network



- Is a collection of vertices (nodes) and edges (links).
- **NETWORK** : Is in fact a representation map of relationships between entities.

Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built

Proteomics
Transcriptomics
Bioinformatics

Databases

Graph Theory

Introduction
Analysis

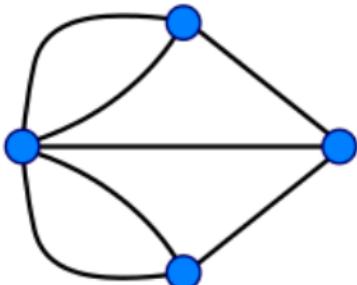
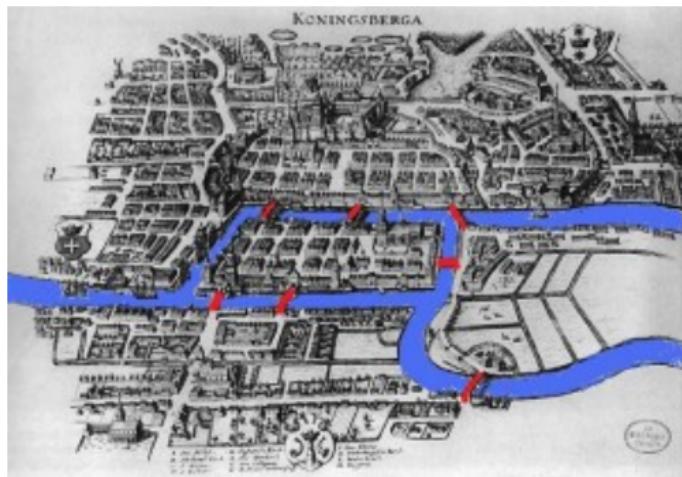
Closing Remarks

Introduction - History

The bridges of Königsberg

The 7 bridges of Königsberg

- The problem was set by Leonard Euler at 1753.
- It is considered the foundation of graph theory !



Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built

Proteomics
Transcriptomics
Bioinformatics

Databases

Graph Theory

Introduction
Analysis

Closing Remarks

Introduction - Formal definition of a graph

- A graph \mathcal{G} is the *combination* of a non-empty set of N nodes \mathcal{V} and a, not necessarily non-empty, set of M edges \mathcal{E}
 - $\mathcal{V} = (v_1, v_2, \dots, v_n)$, $n \in N$ is the vertex (or node) set.
A set of points in fact.
 - $\mathcal{E} = (\{v_2, v_3\}, \{v_3, v_1\}, \dots, \{v_m, v_n\})$, $1 \leq i \leq M$ and $1 \leq i, j \leq N$ is the set of edges (or links).
Ordered pairs of vertices.
- Formally we define a graph by the notation : $\mathcal{G} = (\mathcal{V}, \mathcal{E})$

-
- Graphs are among the most well studied subjects in mathematics (topology, discrete maths) and computer science.

MTH Graph colouring and coverage

MTH Symmetry, flows and search

CS Navigation, path finding

CS Travelling salesman problem, Hamiltonian

Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI

GRNs

Metabolic

Network Built

Proteomics

Transcriptomics

Bioinformatics

Databases

Graph Theory

Introduction

Analysis

Closing Remarks

Graphs and Networks

The terms are absolutely interchangeable

Netowrk Science

1. More popular in biology and social sciences.
2. Developed mostly on the 20-21st century.
3. Modelling of “real life” systems.
4. Measure structure and topology.

Graph Theory

1. More in Maths and Computer Science
2. Founded on the 18th century
3. Modelling “abstract” systems
4. Solve problems of graphs as objects themselves

Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built
Proteomics
Transcriptomics
Bioinformatics

Databases

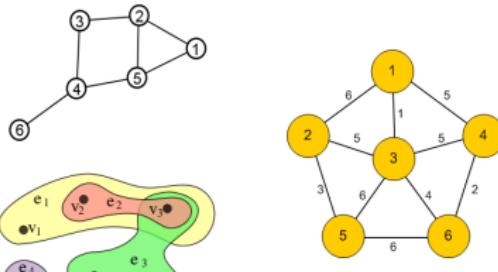
Graph Theory
Introduction
Analysis

Closing Remarks

Categories of graphs/networks

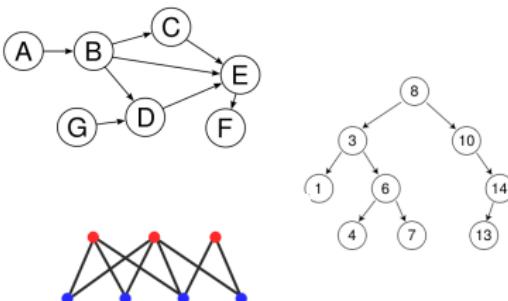
■ Based on their edges :

- Directed / undirected
- Weighted
- Simple or super-edges



■ Based on the structure

- Directed Acyclic Graphs (DAG)
- Binary trees
- Bipartite graphs



Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built

Proteomics
Transcriptomics
Bioinformatics

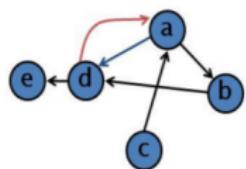
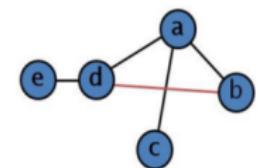
Databases

Graph Theory

Introduction
Algorithms

Closing Remarks

Data structure – Adjacency Matrix



	a	b	c	d	e
a	0	1	1	1	0
b	1	0	0	1	0
c	1	0	0	0	0
d	1	1	0	0	1
e	0	0	0	1	0

	a	b	c	d	e
a	0	1	0	1	0
b	0	0	0	1	0
c	1	0	0	0	0
d	1	0	0	0	1
e	0	0	0	0	0

- Complete mathematical representation of a graph.
- Demonstrates the relationships between edges in a quantitative way.
- Contains simply zeros and ones (for un-weighted graphs). Matrix elements can be any real number for representing weights.

The adjacency matrix IS the graph and vice-versa!

Introduction - Logistics

Practical Contents

Structural Elements

PPI

GRNs

Metabolic

Network Built

Proteomics

Transcriptomics

Bioinformatics

Databases

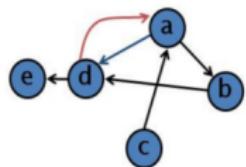
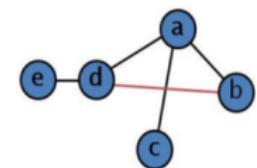
Graph Theory

Introduction

Algorithms

Closing Remarks

Data structure – Adjacency Matrix



	a	b	c	d	e
a	0	1	1	1	0
b	1	0	0	1	0
c	1	0	0	0	0
d	1	1	0	0	1
e	0	0	0	1	0

	a	b	c	d	e
a	0	1	0	1	0
b	0	0	0	1	0
c	1	0	0	0	0
d	1	0	0	0	1
e	0	0	0	0	0

- Complete mathematical representation of a graph.
- Demonstrates the relationships between edges in a quantitative way.
- Contains simply zeros and ones (for un-weighted graphs). Matrix elements can be any real number for representing weights.
- The adjacency matrix IS the graph and vice-versa!

Introduction - Logistics

Practical Contents

Structural Elements

PPI

GRNs

Metabolic

Network Built

Proteomics

Transcriptomics

Bioinformatics

Databases

Graph Theory

Introduction

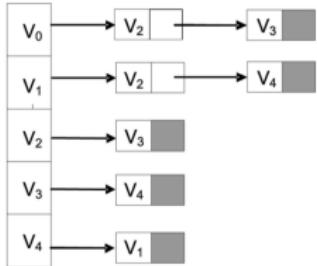
Algorithms

Closing Remarks

Data structures – Adjacency / Edges list

Adjacency List

- More “dense” representation of the network.
- Each node is written in the list together with the node(s) which it is connected.



Edges List

- Take into account only edges.
- It is equivalent with the edge set.
- For un-directed graphs $(v_i, v_j) = (v_j, v_i)$

Introduction - Logistics

Practical Contents

Structural Elements

PPI

GRNs

Metabolic

Network Built

Proteomics

Transcriptomics

Bioinformatics

Databases

Graph Theory

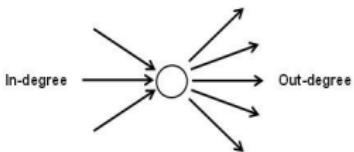
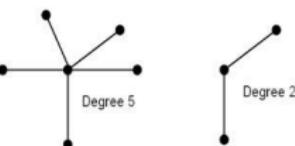
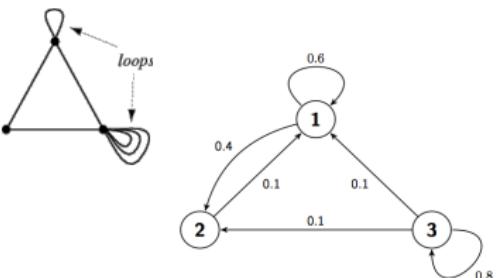
Introduction

Algorithms

Closing Remarks

Basic Definitions I

- **Loop** : An edge which starts and ends at the same node.
- **Degree** : The number of edges which are connected with a node.
 - Undirected : Degree is the sum of the edges.
 - Directed : Degree is divided to in-degree and out-degree.



Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built
Proteomics
Transcriptomics
Bioinformatics

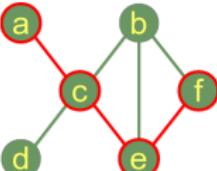
Databases

Graph Theory
Introduction
Algorithms

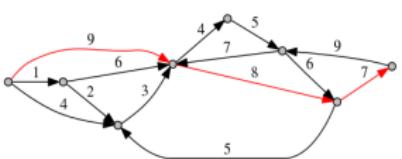
Closing Remarks

Basic definitions II

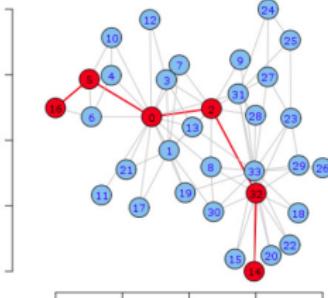
- **Path** : Is a sequence of edges which connects two nodes.
(for un-directed graphs we follow the directionality of the edges.)



- **Shortest Path** : The closest of all paths that connect two nodes.
(we calculate each individual weight for weighted networks)



- **Diameter** : The longest among all shortest paths in a network.
- **Cycle** : A path which starts and stops at the same node.



Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built
Proteomics
Transcriptomics
Bioinformatics

Databases

Graph Theory
Introduction
Algorithms

Closing Remarks

Network measures complexity

- Small introduction on the computational complexity of network measures.
- Graph theory is one of THE most well studied fields in the analysis of algorithms.

Property	Adjacency matrix	Adjacency list	Edge list
Memory requirement	N^2	$N + M$	M
Initialisation	N^2	N	1
Copying a node	N^2	M	M
Deleting an edge	N	M	1
Finding an edge	1	N	M
Is a node isolated	N	1	M
Testing for a path between two nodes	N^2	$M \log(N)$	$N + M$

More on :

<https://mitpress.mit.edu/books/introduction-algorithms>

Graph theory

Our first theorem

A theorem which associates the number of edges q in a graph, with the degree \deg of its nodes p .

$$\sum_{i=1}^p \deg_{v_i} = 2q$$

It does look difficult...

1. The degree of a node is the number of edges which are connected with it.
2. How many times an edge is “connected”... ;
3. Mais bien sur, two times, one from the outgoing and once to the incoming node.

Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built

Proteomics
Transcriptomics
Bioinformatics

Databases

Graph Theory
Introduction
Algorithms

Closing Remarks

Graph theory

Our first theorem

A theorem which associates the number of edges q in a graph, with the degree \deg of its nodes p .

$$\sum_{i=1}^p \deg_{v_i} = 2q$$

It does look difficult... but if you think about it a little bit is something trivial. Just remember :

1. The degree of a node is the number of edges which are connected with it.
2. How many times an edge is “connected”... ;
3. Mais bien sur, two times, one from the outgoing and once to the incoming node.

Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built

Proteomics
Transcriptomics
Bioinformatics

Databases

Graph Theory
Introduction
Algorithms

Closing Remarks

Graph theory

Our first theorem

A theorem which associates the number of edges q in a graph, with the degree \deg of its nodes p .

$$\sum_{i=1}^p \deg_{v_i} = 2q$$

It does look difficult... but if you think about it a little bit is something trivial. Just remember :

1. The degree of a node is the number of edges which are connected with it.
2. How many times an edge is “connected”... ;
3. Mais bien sur, two times, one from the outgoing and once to the incoming node.

Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built

Proteomics
Transcriptomics
Bioinformatics

Databases

Graph Theory
Introduction
Algorithms

Closing Remarks

Graph theory

Our first theorem

A theorem which associates the number of edges q in a graph, with the degree \deg of its nodes p .

$$\sum_{i=1}^p \deg_{v_i} = 2q$$

It does look difficult... but if you think about it a little bit is something trivial. Just remember :

1. The degree of a node is the number of edges which are connected with it.
2. How many times an edge is “connected”... ;
3. Mais bien sur, two times, one from the outgoing and once to the incoming node.

Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built

Proteomics
Transcriptomics
Bioinformatics

Databases

Graph Theory
Introduction

Closing Remarks

Graph theory

Our first theorem

A theorem which associates the number of edges q in a graph, with the degree \deg of its nodes p .

$$\sum_{i=1}^p \deg_{v_i} = 2q$$

It does look difficult... but if you think about it a little bit is something trivial. Just remember :

1. The degree of a node is the number of edges which are connected with it.
2. How many times an edge is “connected”... ;
3. Mais bien sur, two times, one from the outgoing and once to the incoming node.

In everyday english :

“The sum of all the degrees in a network equals two times the number of its edges.”

Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built

Proteomics
Transcriptomics
Bioinformatics

Databases

Graph Theory

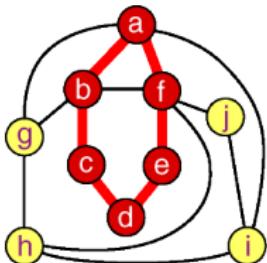
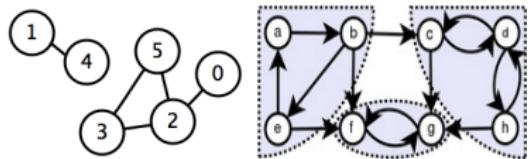
Introduction
Graphs

Closing Remarks

Network Characteristics

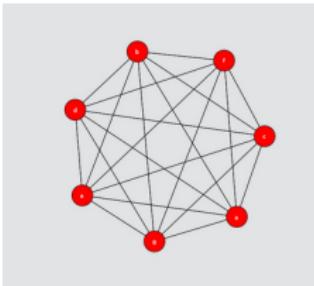
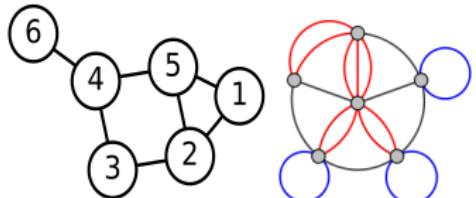
Sub-graphs

Connected component



Simple and complex networks

Fully connected networks.



Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built

Proteomics
Transcriptomics
Bioinformatics

Databases

Graph Theory

Introduction
Algorithms

Closing Remarks

Other Biological (and not) Networks

- Computer networks
 - WWW, internet.
 - Telecommunications, satellite.
- Social Networks
 - Friends Networks (Facebook)
 - Co-authorship networks (papers), Networks of collaborators (muoc, movies)
- Transportation
 - Railways, flight networks, logistics.
- Electronic circuits
- An a lot lot more...

Graphs/networks are such successful abstractions that allow us to represent (almost) everything with them... But why we really study them so hard ?

Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built

Proteomics
Transcriptomics
Bioinformatics

Databases

Graph Theory
Introduction
Analysis

Closing Remarks

Other Biological (and not) Networks

- Computer networks
 - WWW, internet.
 - Telecommunications, satellite.
- Social Networks
 - Friends Networks (Facebook)
 - Co-authorship networks (papers), Networks of collaborators (muoc, movies)
- Transportation
 - Railways, flight networks, logistics.
- Electronic circuits
- An a lot lot more...

Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built

Proteomics
Transcriptomics
Bioinformatics

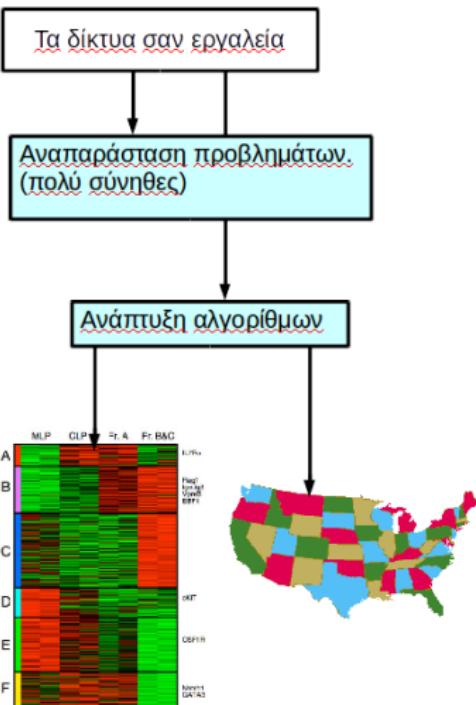
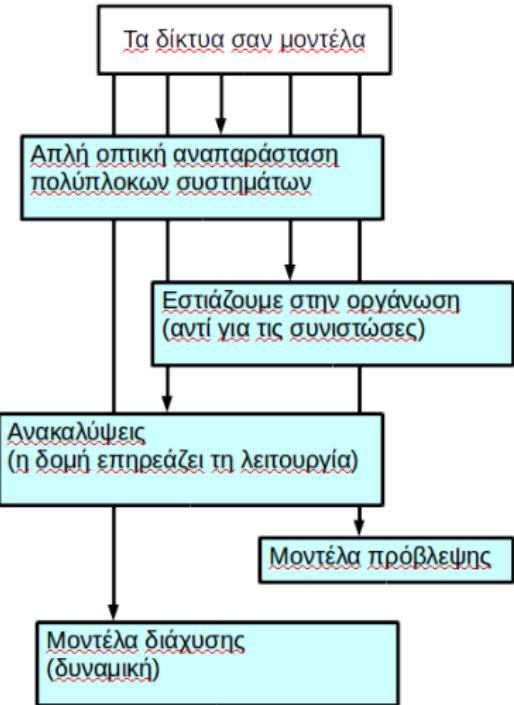
Databases

Graph Theory
Introduction
Analysis

Closing Remarks

Graphs/networks are such successful abstractions that allow us to represent (almost) everything with them... But why we really study them so hard ?

Why we study networks



Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built

Proteomics
Transcriptomics
Bioinformatics

Databases

Graph Theory

Introduction
Analysis

Closing Remarks

Closing relaxing figure...

Introduction -
Logistics

Practical
Contents

Structural
Elements

PPI
GRNs
Metabolic

Network Built

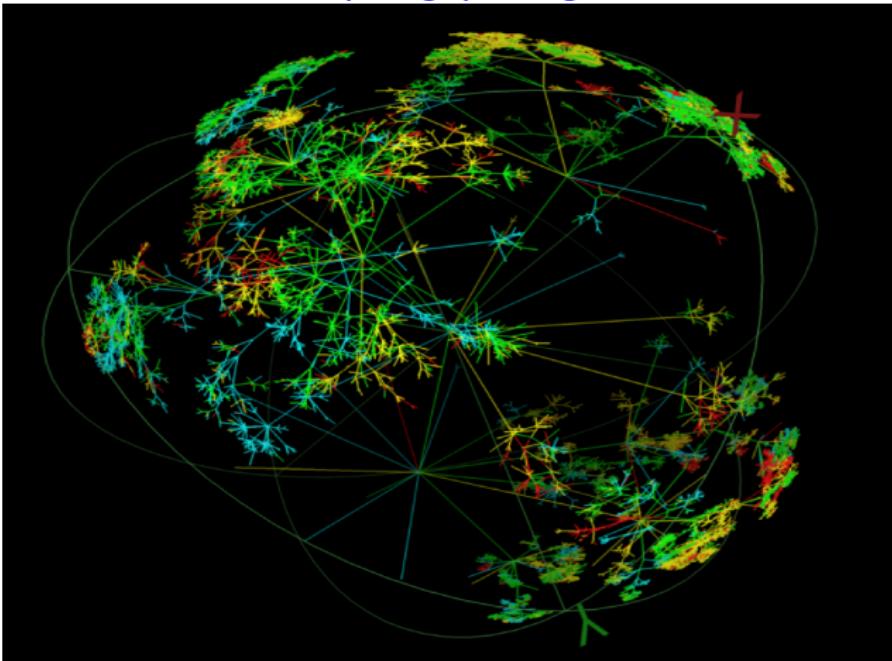
Proteomics
Transcriptomics
Bioinformatics

Databases

Graph Theory
Introduction
Analysis

Closing Remarks

<https://gephi.org/>



<http://cytoscape.org/>