Systems Biology in the Rabbit?
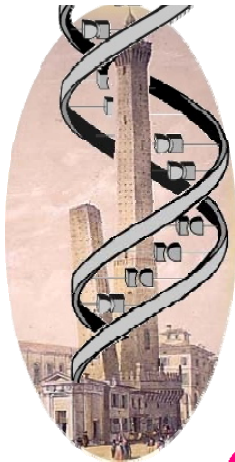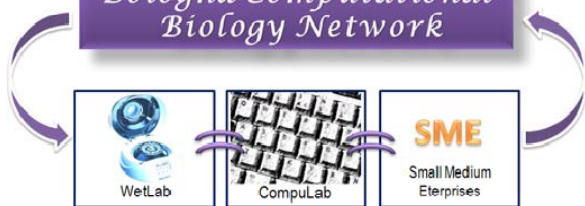
Rita Casadio

BIOCOMPUTING GROUP
University of Bologna, Italy
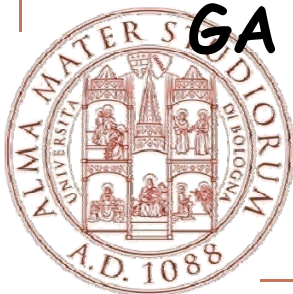
fiRBBC

Bologna Computational Biology Network

WetLab   CompuLab   SME Small Medium Eterprises
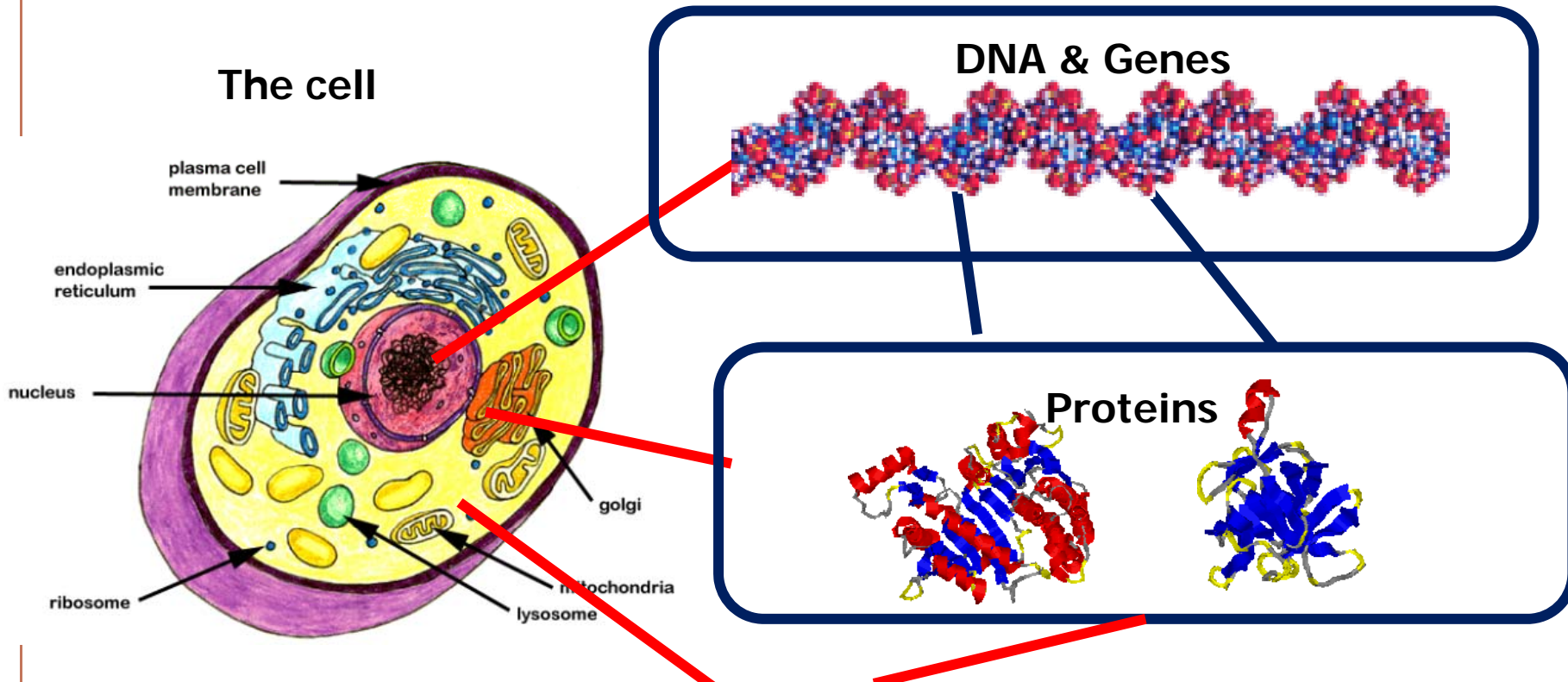
GB-Net

The "omic" revolution

The analysis of the components of a living organism in its entirety

GA EVANS, Nature Biotechnology 2000, 18, 127

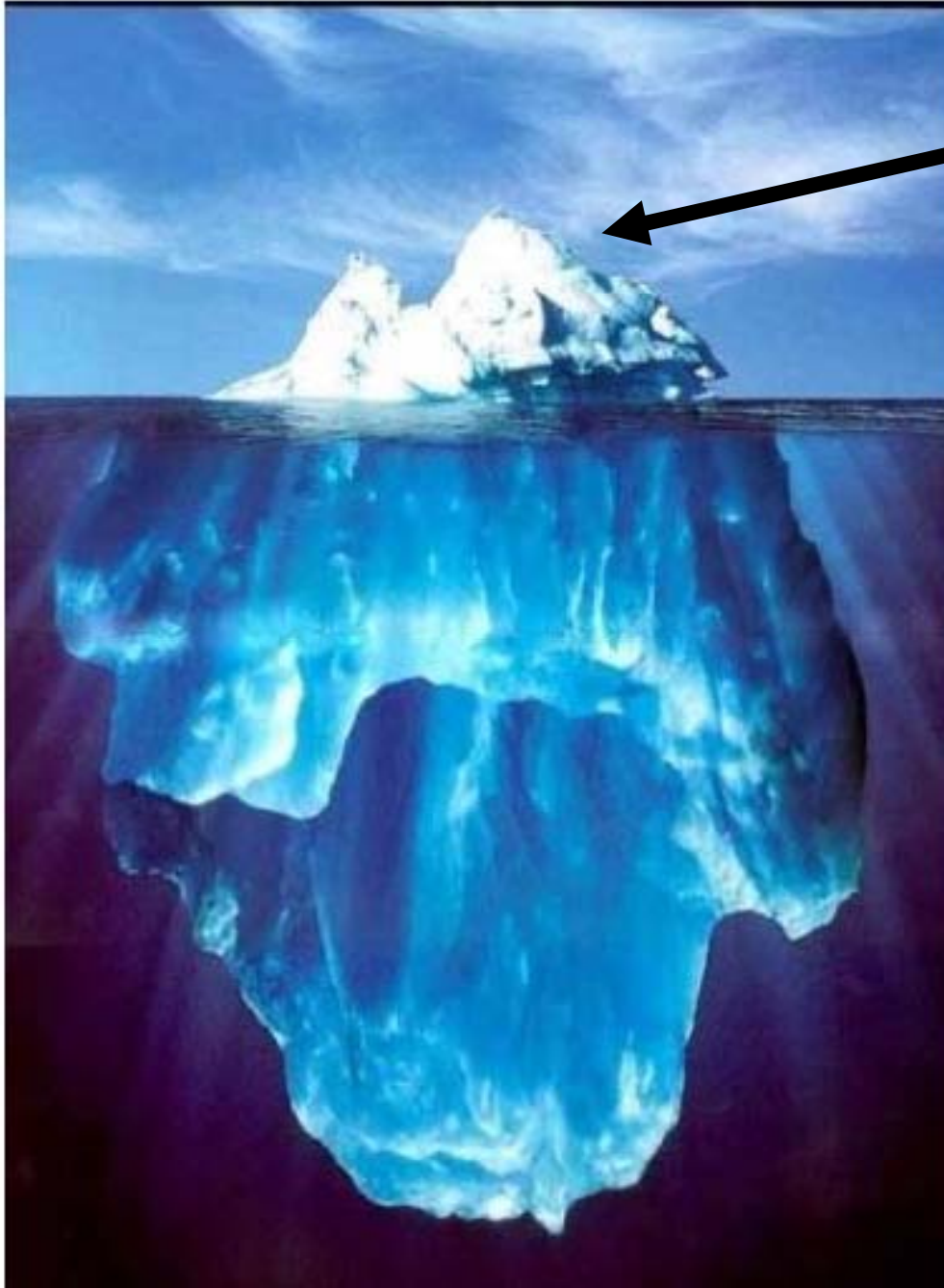# The ingredients of biological complexity at the cell level
From genes to proteins and their interaction

The cell

DNA & Genes

plasma cell membrane

endoplasmic reticulum

nucleus

Proteins

ribosome

golgi

mitochondria

lysosome

Macromolecular crawding

Molecular Interactions and Functions

The "omic" era

C
o
m
p
l
e
x
i
t
y

Genomics

Transcriptomics

**Proteomics**

Metabolomics

Regulomics

Systems Biology

Systems Biology at the organism level

Systems Biology : A philosophical definition

"Systems biology...is about putting together rather than taking apart, integration rather than reduction. It requires that we develop ways of thinking about integration that are as rigorous as our reductionist programmes, but different....It means changing our philosophy, in the full sense of the term"

Dennis Noble (2006). The Music of Life. Oxford University Press
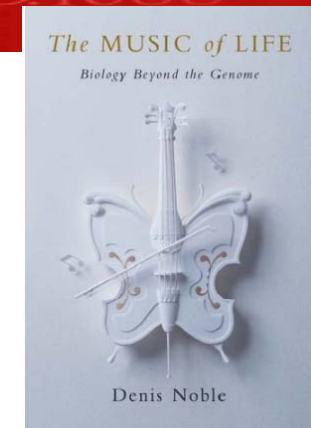
Systems Biology at the organism level:

Systems biology is the study of systems of biological components, **which may be molecules, cells, organisms or entire species**. **Systems problems are emerging as central to all areas of biology and medicine.**

•Living systems are dynamic and complex, and their behavior may be hard to predict from the properties of individual parts (Complex Systems).

•To study them quantitative measurements of the behavior of groups of interacting components are necessary.

•Measurement technologies such as **genomics**, **bioinformatics** and **proteomics**, and **mathematical and computational models to describe and predict dynamical and systemic behavior** are necessary.

Genomics: Organismal deoxyribonucleic acid (DNA) sequence, including intra-organisamal cell specific variation. (i.e. Telomere length variation etc.).

Variomics: The study of the genetic variation of the population (SNP, Copy Number Variations, translocations)

Epigenomics / Epigenetics: Organismal and corresponding cell specific transcriptomic regulating factors not empirically coded in the genomic sequence. (i.e. DNA methylation, Histone Acetylation etc.).

Transcriptomics: Organismal, tissue or whole cell gene expression measurements by DNA microarrays or serial analysis of gene expression

Translatomics / Proteomics: Organismal, tissue, or cell level measurements of proteins and peptides via two-dimensional gel electrophoresis, mass spectrometry or multi-dimensional protein identification techniques (advanced HPLC systems coupled with mass spectrometry). Sub disciplines include phosphoproteomics, glycoproteomics and other methods to detect chemically modified proteins.

Metabolomics: Organismal, tissue, or cell level measurements of all small-molecules known as metabolites.

Interactomics: Organismal, tissue, or cell level study of interactions between molecules. Currently the authoritative molecular discipline in this field of study is protein-protein interactions (PPI),
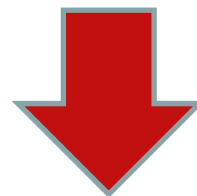
Phenomics: Organismal variation in phenotype as it changes during its life span
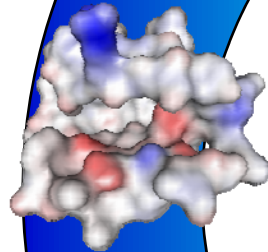
(about 20,876 genes and 181,744 transcripts in the human genome)

Genes in DNA...

>protein kinase

acctgttgatggcgacagggactgtatgctgatct
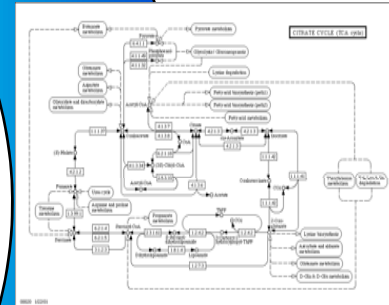atgctgatgcatgcatgctgactactgatgtggggg
gctattgacttgatgtctatc....

...with different effects depending on variability

Over 50 millions of single mutations are known

...code for proteins...

Overall: from Genotype to Phenotype

....in methabolic pathways

...proteins correspond to functions...

From 5000 to 10000 proteins per tissue

Proteins interact
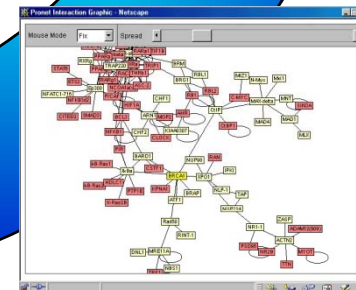
...when they are expressed
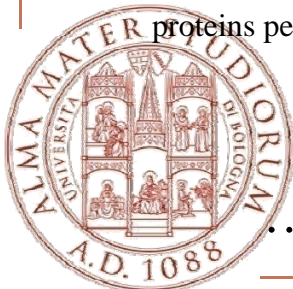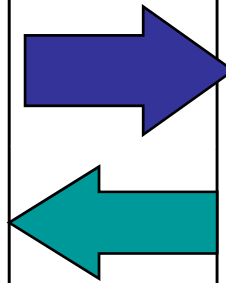
BIOINFORMATICS

Data Bases

(Biosequences, Structures, Genomes, DNA Chips, Proteomes, Interatomics, Literature)

- Implementation
- Data Mining
- Links

Computational Biology

Tools for:
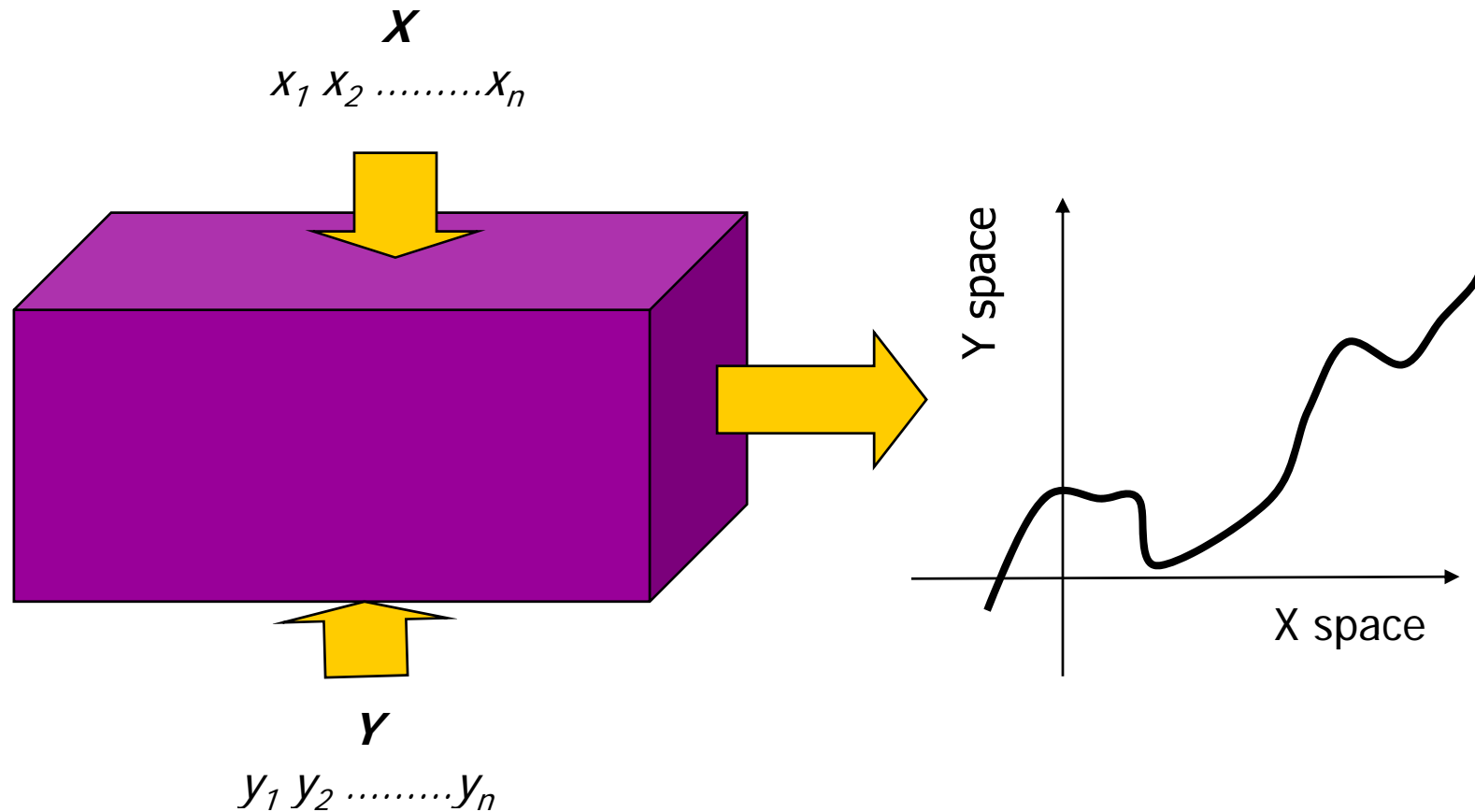- Sequence analysis
- Functional genomics
- Proteomics
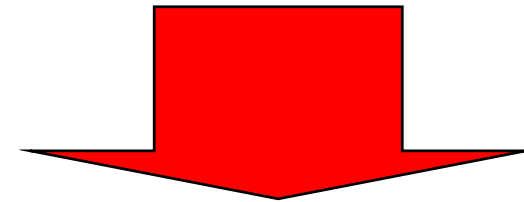
Systems Biology

Models for:
Interatomics, Methabolomics, Evolving complex biosystems (Cell, Organism,..)

General non-linear functional mapping

X

x_1 x_2 ....x_n

Y space

X space

Y

y_1 y_2 ....y_n

Machine learning approaches help data mining and, if existing, extraction of general rules relating inputs to outputs

Tools out of machine learning approaches
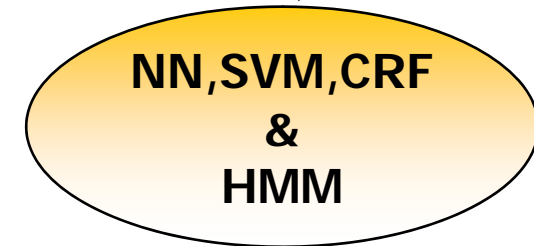
Training   *Testing*

Data Base Subset
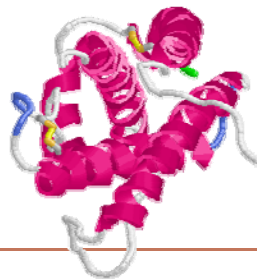
TTCCPSIVARSNFNVCRLPGTPEAICATYTGCIIIPGATCPGDYAN

New sequence

NN, SVM,CRF & HMM

General rules

NN,SVM,CRF & HMM
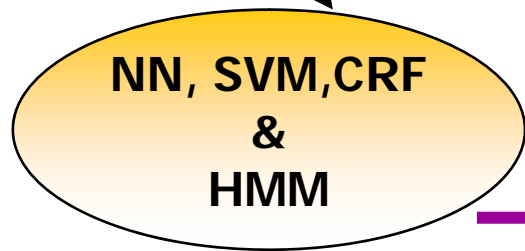
Known mapping/labelling

Known structures
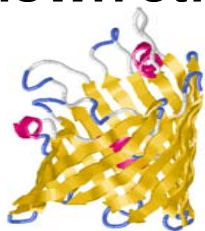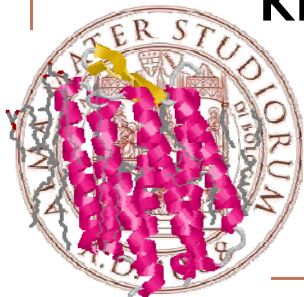
Prediction

Bologna Biocomputing Group
University of Bologna

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Predictors

- **BaCelLo** - Balanced subCellular Localization predictor
- **BAR+** - Bologna Annotation Resource
- **BetAware** - Detection of Prokaryotic outer-membrane betabarrel proteins
- **CCHMM** - Predictor of Coiled-Coils Regions in Proteins
- **CCHMMPROF** - Predictor of Coiled-Coils Regions in Proteins exploiting evolutionary information
- **CORNET** - Predictor of Residue Contacts in Proteins
- **DCON** - Predictor of Disulfide Connectivity in Proteins
- **DisLocate** - Find Disulfide bonds in Eukaryotes with predicted subcellular Localization
- **FT-COMAR** - Fault Tolerance Reconstruction of 3D Structure from Protein Contact Maps
- **HIPPIE** - Protease Inhibitor engine
- **I-MUTANT** - Neural Network based Predictor of Protein stability Changes upon Single Point Mutation from the Protein Structure

Application Servers

- **TRAMPLE**: the transmembrane protein labelling environment
- **PONGO**: a web server for multiple predictions of all-alpha transmembrane proteins

Databases

- **eSLDB** - eukaryotic Subcellular Localization DataBase
- **ZenPatches** - Database of predicted protein interaction sites
- **DBMFHS** - Data Base of Minimally-Frustrated Helical Segments

News and Announcement

September 19-22, 2011
2nd International Summer School on Gender Medicine in Sassari, Italy, from 19. - 22. Sept, organized by international experts from the EU and board members of the International Society of Gender Medicine and the University of Sassari. PDF

July 7-10th, 2011
IIIrd Italian-Hispano-Portuguese Workshop On The Molecular Biology And Biophysics Of Ion Channels And Transporters. Website and Playbill

June 21-24th, 2011
Cycle of seminars: "Elements of Graph Theory". Professor Ugur Sezerman, Sabanci University, Istanbul (Turkey). Laboratory C, Via San Giacomo 9/2 9.30AM-1.00PM
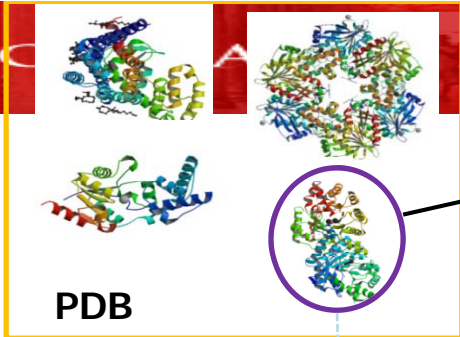
BaCelLo
Balanced subCellular Localization predictor

BAR

SNPs&GO

PDB

Template

Target

NN, HMM, SVM, GRHCRF
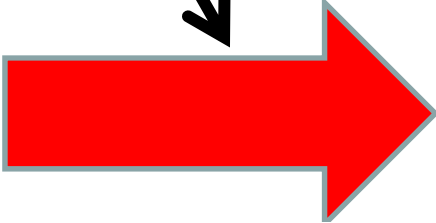
Systems biology at the cell level

Alternative splicing variant annotation

http://t.caspur.it/ASPicDB/
Martelli et al., 2011, NAR

The Bologna functional annotation resource

www.biocomp.unibo.it

Prediction of:
Secondary structure
Disulfide bonding state
Solvent accessibility
Active sites, Protein protein interaction sites, protein DNA/RNA interactions, GPI anchor
Functional domains

Networks

WDR20
KCNQ1
IGF2
PHLDA
MYOD1
CDKN1C
SKP2
CCNA2
CDK2
CCND1
CDK4
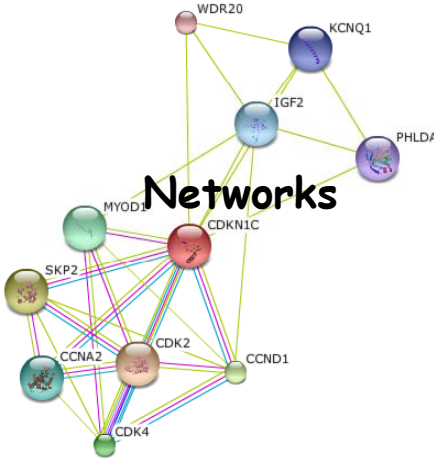
Feature-driven alignment

Modelling

Model

Check of model stereochemistry

Protein-Protein interaction

Structural location of variation
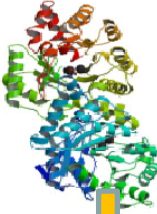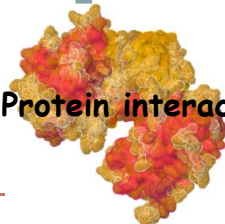
Systems Biology at the organism level

How to link all the omics to phenomics:
towards a rabbit systems biology

The case study of the RBG network:

- The RGB-Net as system comprising
  4 Working Groups WG (#people=96)
- The RGB-Net analysis to:
  - o identify strengths and weaknesses
  - o enable information streams and collaboration
  - o plan future goals

Current status of the GB-network

WG1=blue
WG2=red
WG3=green
WG4=yellow

GB-Net

Carneiro

Ferrand

Fontanesi

David
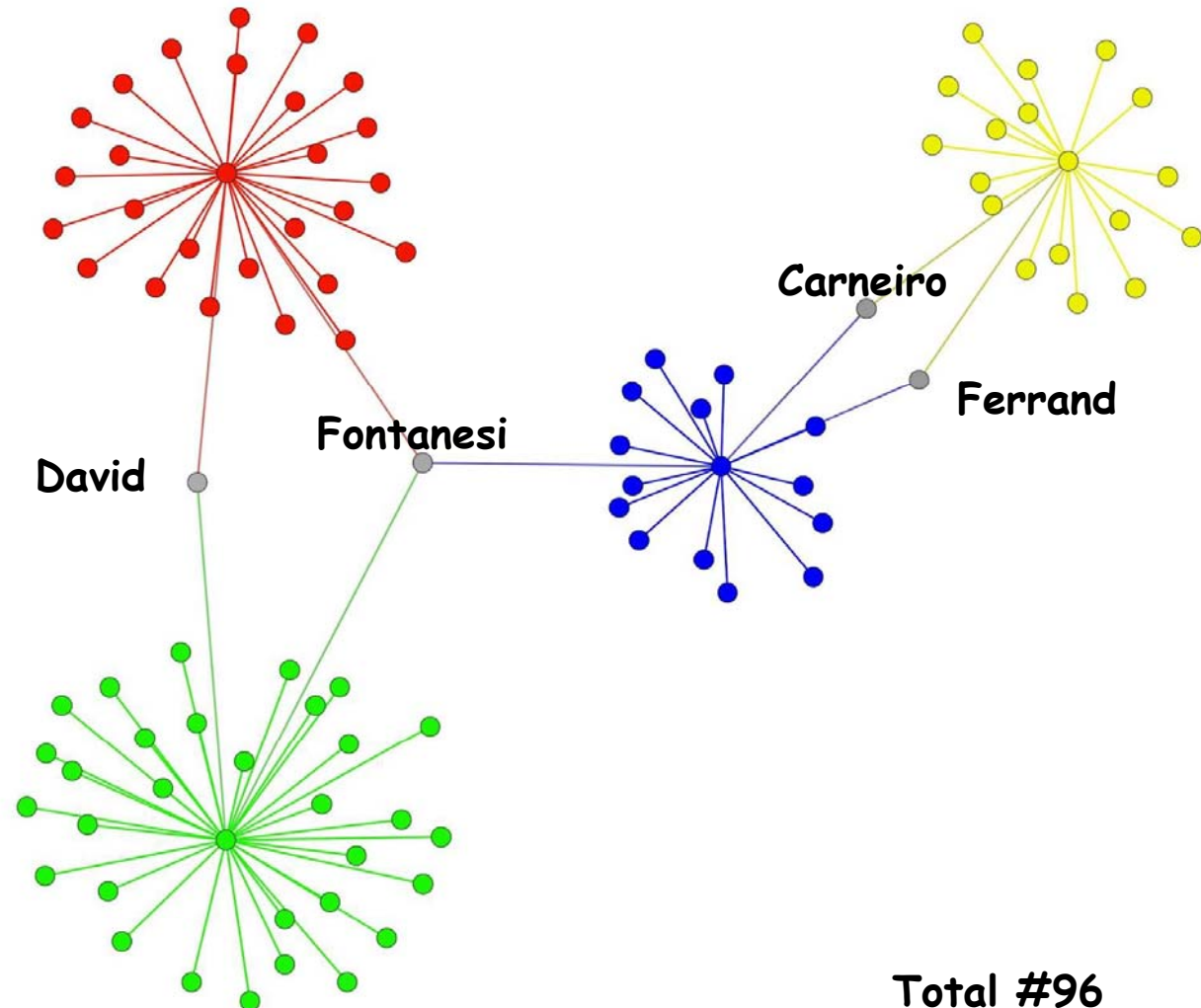
Total #96

Establishing connections among groups usuful for a systems biology of the rabbit

- Online survey
  - o 115 possible keywords
  - o user selected from 5 up to 10 keywords
  - o chance to type custom tags in a separate page
- Results
  - o 75 people answered
  - o 108 distinct keywords were selected

GB-Net

Keywords allow connecting people in different WGs

WG1=blue
WG2=red
WG3=green
WG4=yellow

108 Keywords=purple
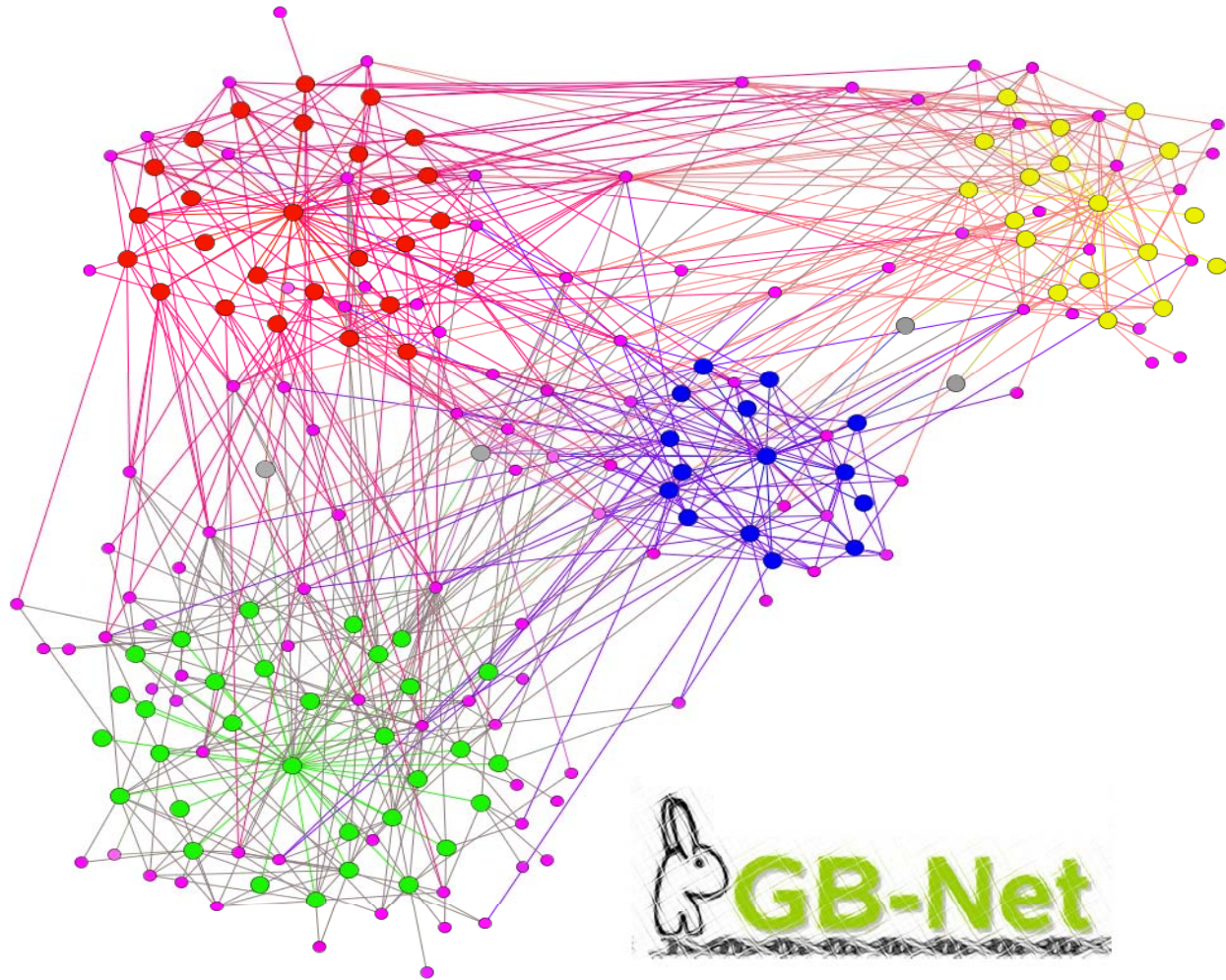
Clustering the 108 keywords into general topics of systems biology

Available

- **genome/genomics**
  - 7x
  - annotation (ENSEMBL)
- **transcriptomics**
  - RNA-Seq (Broad Institute)
- **phenomics**

Possibly available

- **variomics**
  - SNPs (available soon)
  - Resequencing (available soon)
- **epigenomics**
- **proteomics**
- **metabolomics**
- **interactomics**

Rabbit systems biology and the rabbit genome network: a collaborative systems biology network?
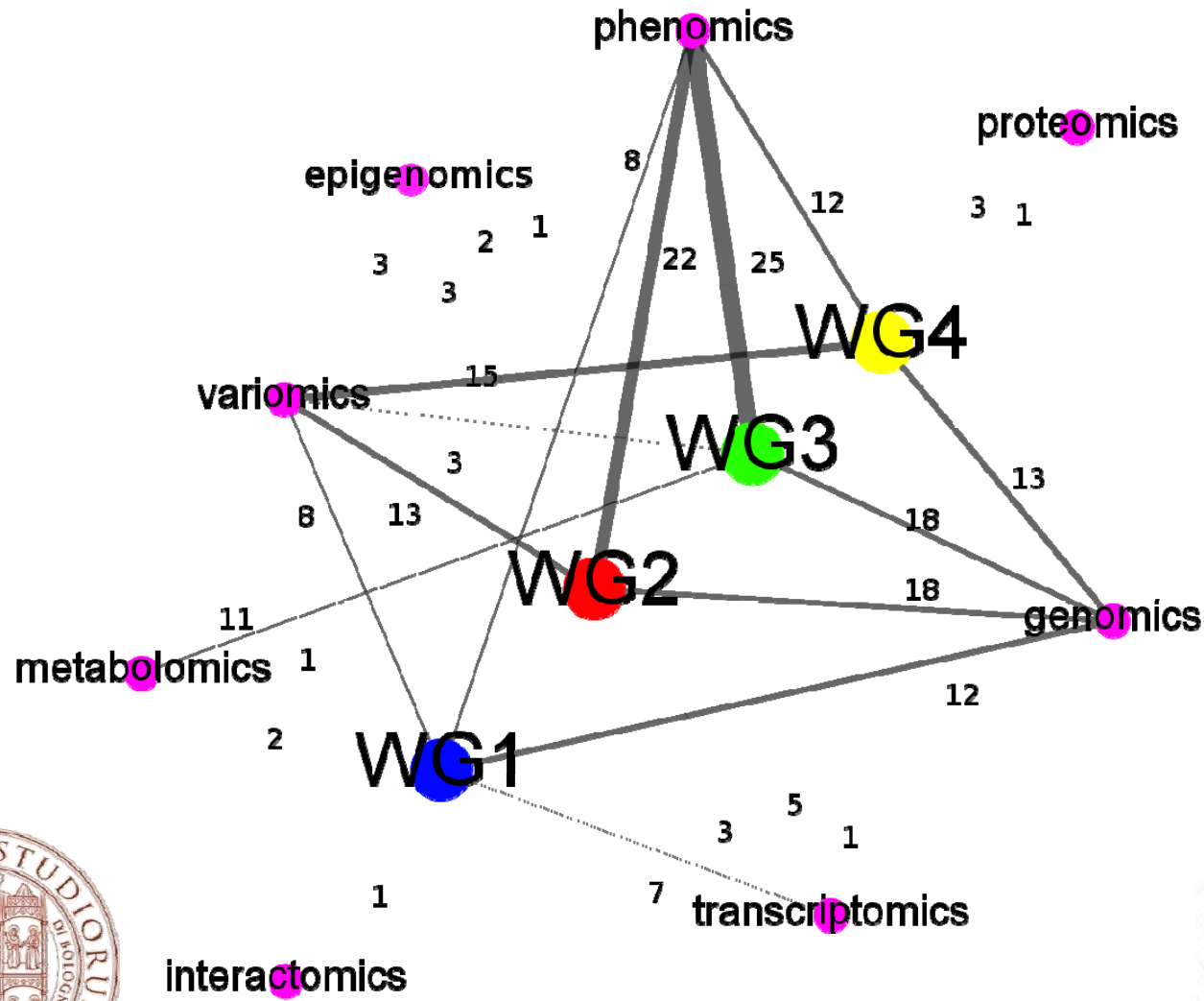
Conclusions

•**The collaborative rabbit genome network can become the framework for the rabbit systems biology**

•**We need however to strengthen the number of people who can provide data particularly for:**

Interactomics
Proteomics
Epigenomics

Bologna Biocomputing Group
University of Bologna

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

| Home | Members | Predictors/Databases | Publications | Training | BWS | Activities | WebMail |

Group Leader

Rita Casadio

Senior Researchers

Piero Fariselli
Pier Luigi Martelli

Technical Staff

Gianluca Tasco

Post-Doc

Pietro Di Lena
Marco Vassura

External Collaborators

Alberto Eusebi (BioDec)
Andrea Pierleoni (Externautics)
Giuseppe Profiti(CIRI LSHT)
Ivan Rossi (BioDec)

PhD Students

Raffaele Fronza
Damiano Piovesan
Deepak Rajan
Castrense Savojardo
Valentina Indio

Former Members

Lisa Bartoli
Remo Calabrese
Emidio Capriotti
Paola Marani
Ludovica Montanucci
Priyank Shukla
Shalinee Tiwari

Acknowledgements