



Istituto  
C.S.S. Mendel



Fondazione  
Casa Sollievo  
della Sofferenza

# IRCCS CASA SOLLIEVO DELLA SOFFERENZA

## Bioinformatics unit



Tommaso Mazza  
[t.mazza@css-mendel.it](mailto:t.mazza@css-mendel.it)

Roma Tre – 11 March 2014



**L'Istituto C.S.S. Mendel di Roma è stato fondato nel 1953 dal Prof. Luigi Gedda, che lo ha reso un Centro di riferimento per gli studi e le ricerche nel campo della Genetica Medica e della Gemellologia.**

**Nel 1997 l'Istituto è stato donato alla Fondazione Casa Sollievo della Sofferenza di San Giovanni Rotondo.**



# **Fondazione Casa Sollievo della Sofferenza Istituto C.S.S. Mendel**

**Attualmente l'Istituto C.S.S. Mendel è un presidio  
multifunzione autorizzato per attività di  
Laboratorio Specializzato di Genetica Medica e  
Poliambulatorio Specialistico**



**Il Laboratorio Specializzato è anche accreditato  
con il SSR/SSN relativamente alle prestazioni di  
Genetica Medica**

# Fondazione Casa Sollievo della Sofferenza

## Istituto C.S.S. Mendel

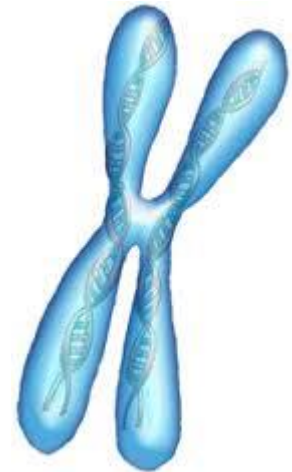
La struttura dispone di una superficie di oltre **2.600 mq<sup>2</sup>**, compresi circa **700 mq<sup>2</sup>** di laboratori. Sono presenti una biblioteca, una sala congressi, una sala riunioni, una stanza multimediale con servizi di supporto e alcuni ambulatori.

Attualmente l'Istituto accoglie circa **70** persone, tra **ricercatori, medici, biologi, biotecnologi, tecnici e personale amministrativo.**

# Laboratorio Specializzato di Genetica Medica

## *Diagnosi Citogenetica*

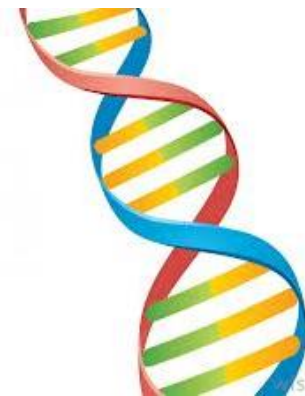
- Consulenza Genetica
- Test rapido Aneuploidie cromosomi 13, 18, 21, X, Y
- Cariotipo su sangue periferico (SP)
- Cariotipo su villi coriali (CVS)
- Cariotipo su liquido amniotico (LA)
- Cariotipo su materiale abortivo (MA o POC)
- Sindromi da microriarrangiamenti cromosomici
- Analisi su preparati fissati in paraffina
- Allestimento di linee linfoblastiche mediante EBV (Epstein Barr Virus)
- Studio dei riarrangiamenti cromosomici



# Laboratorio Specializzato di Genetica Medica

## *Diagnosi Genetica Molecolare*

- Acondroplasia
- Analisi forensi
- Angelman, Sindrome di
- Atassia-teleangiectasia
- Atrofia muscolare spinale
- Beckwith-Wiedemann, Sindrome di
- Branchio-Oto-Renale, Sindrome
- Disomia Uniparentale
- Displasia OculoDentoDigitale
- Distonia mioclonica
- Distonia primaria di torsione
- Distrofia miotonica di Steinert
- Distrofia muscolare dei cingoli di tipo 1C
- Ellis-Van Creveld, Sindrome di
- Esostosi multiple ereditarie
- Fibrosi cistica
- Hallervorden-Spatz, Sindrome di
- Insensibilità agli androgeni
- Ipocondroplasia
- Kennedy, Malattia di
- Leopard, Sindrome di
- Martin Bell, Sindrome di
- Microdelezioni cromosoma Y
- Microftalmia sindromica 3
- Miotonia congenita di Thomsen/Becker
- Neoplasie endocrine multiple tipo 2
- Neurofibromatosi tipo 1
- Noonan, Sindrome di
- Norrie, Sindrome di
- Prader Willi, Sindrome di
- Rene policistico dell'adulto
- Rene policistico infantile
- Retinite pigmentosa di tipo 2
- Retinite pigmentosa di tipo 3
- Retinite pigmentosa di tipo 7
- Silver-Russell, Sindrome di
- Sinpolidattilia 1
- Sordità indotta da streptomicina
- Sordità neurosensoriale AR1
- Talassemia beta
- Wolfram, Sindrome di
- Trombofilia



# Poliambulatorio Specialistico

## Ambulatorio di Ostetricia e Ginecologia

1. Ecografie
2. Amniocentesi

## Ambulatorio di Dermatologia

1. Visita specialistica
2. epiluminiscenza
3. Mappa nei
4. Crioterapia
5. Diatermocoagulazione
6. Dermatosi della gravidanza
7. Angiomi del neonato
8. Eczema seborroico del lattante

## Ambulatorio di Cardiologia

1. Visita specialistica
2. Elettrocardiogramma basale
3. Holter cardiaco e pressorio

## Ambulatorio di Pediatria

1. Visita specialistica
2. Screening audiometrico ed oculare

## Ambulatorio di Endocrinologia

Visita specialistica

## Ambulatorio di Neurologia

Visita specialistica

## Ambulatorio di Medicina Legale

Visita specialistica

# Unità di Ricerca

***Neurogenetica***

***Diabete***

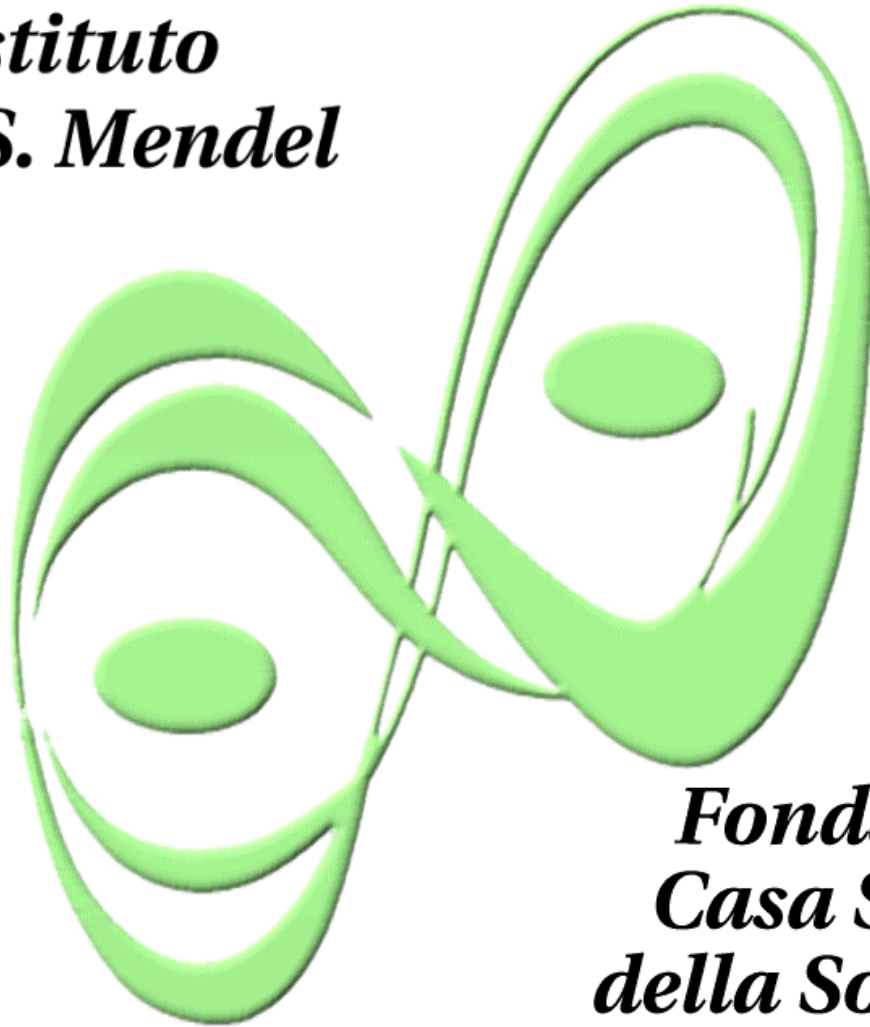
***Oncogenomica***

***Patologie psichiatriche***

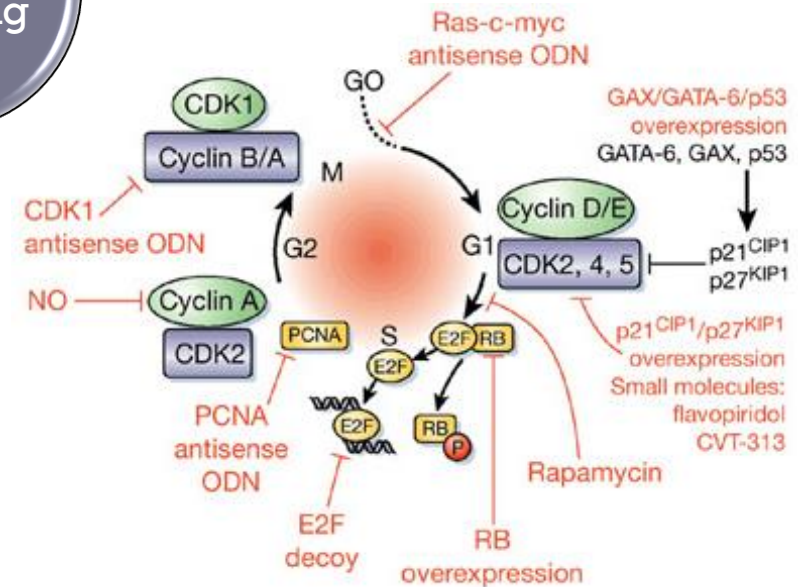
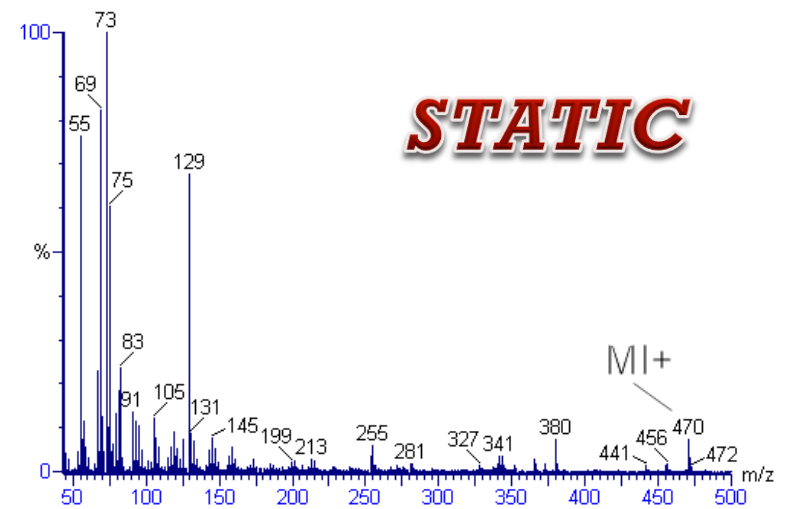
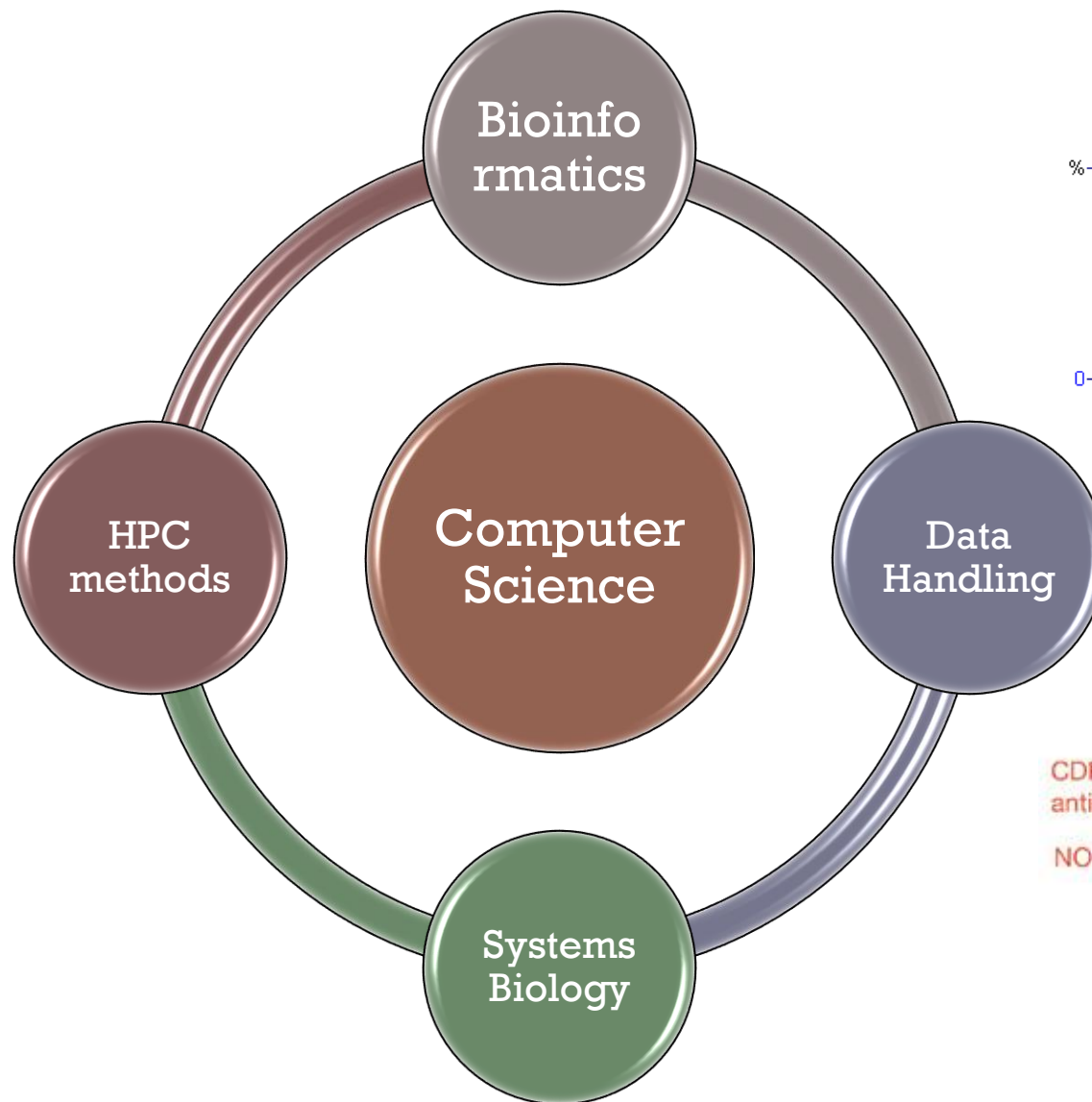
***Bioinformatica***



***Istituto  
C.S.S. Mendel***



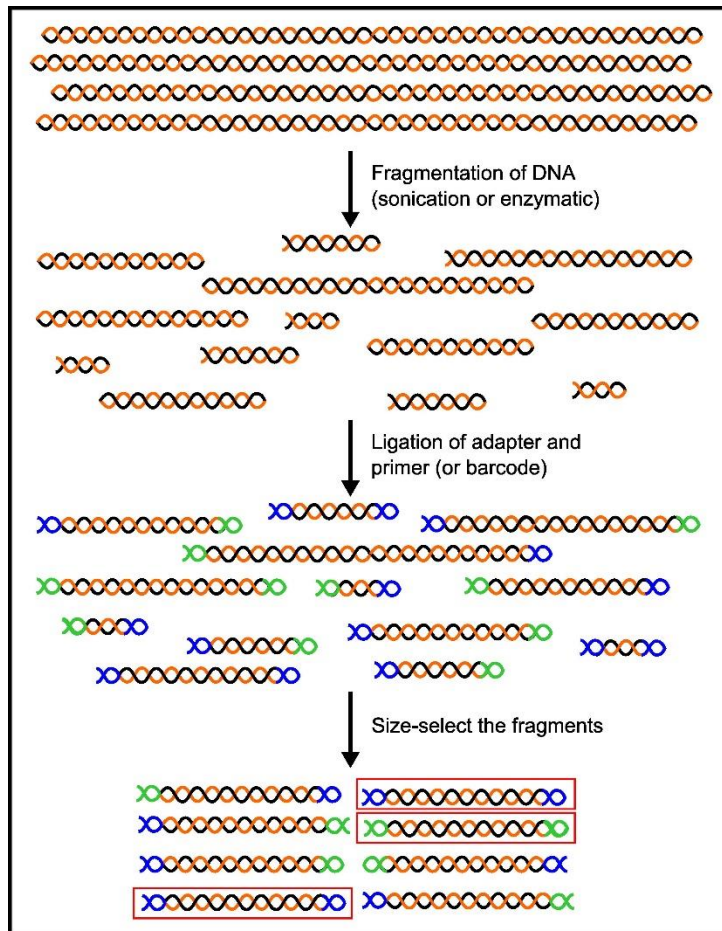
***Fondazione  
Casa Sollievo  
della Sofferenza***



**DYNAMIC**



# NEXT GENERATION SEQUENCING



# NGS PRIMARY ANALYSIS

Raw Reads

Mapping

Local  
realignment

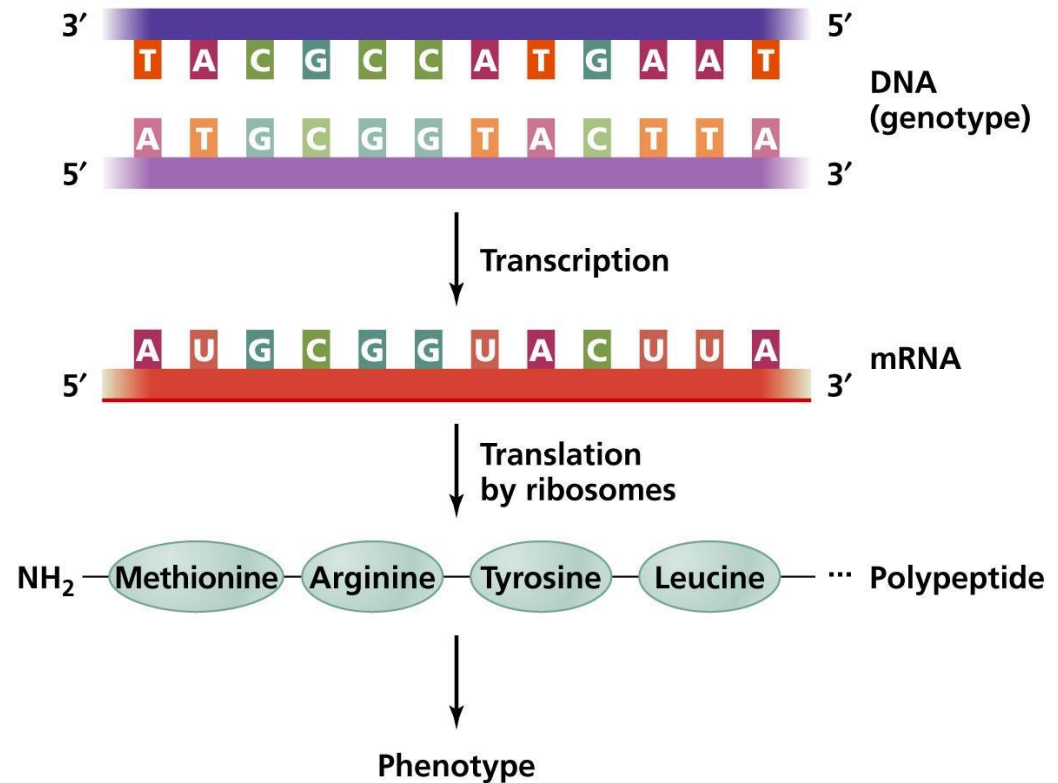
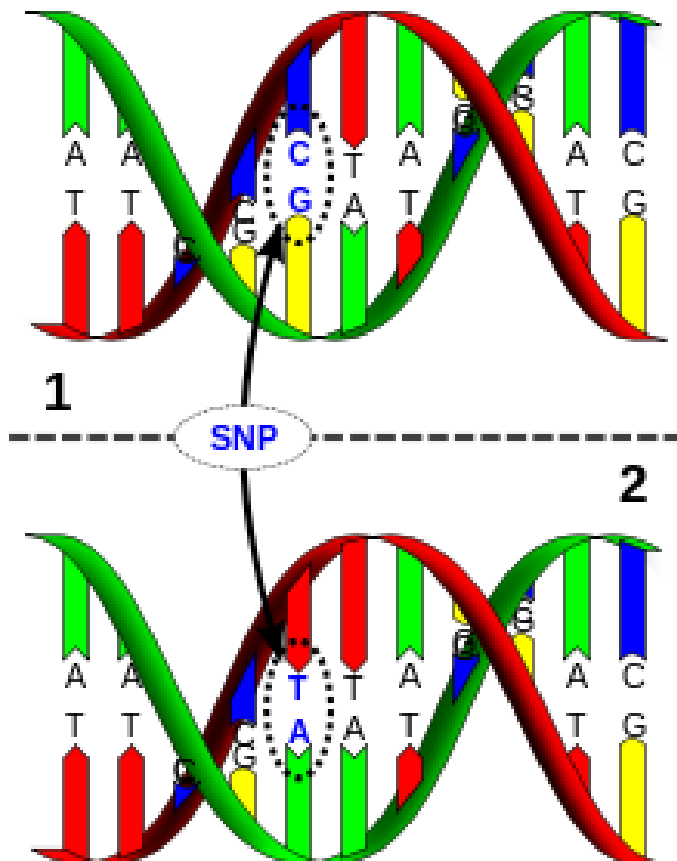
Duplicate  
marking

Base quality  
recalibration

Analysis-ready  
reads



# NGS SECONDARY ANALYSIS



Copyright © 2006 Pearson Education, Inc., publishing as Benjamin Cummings.

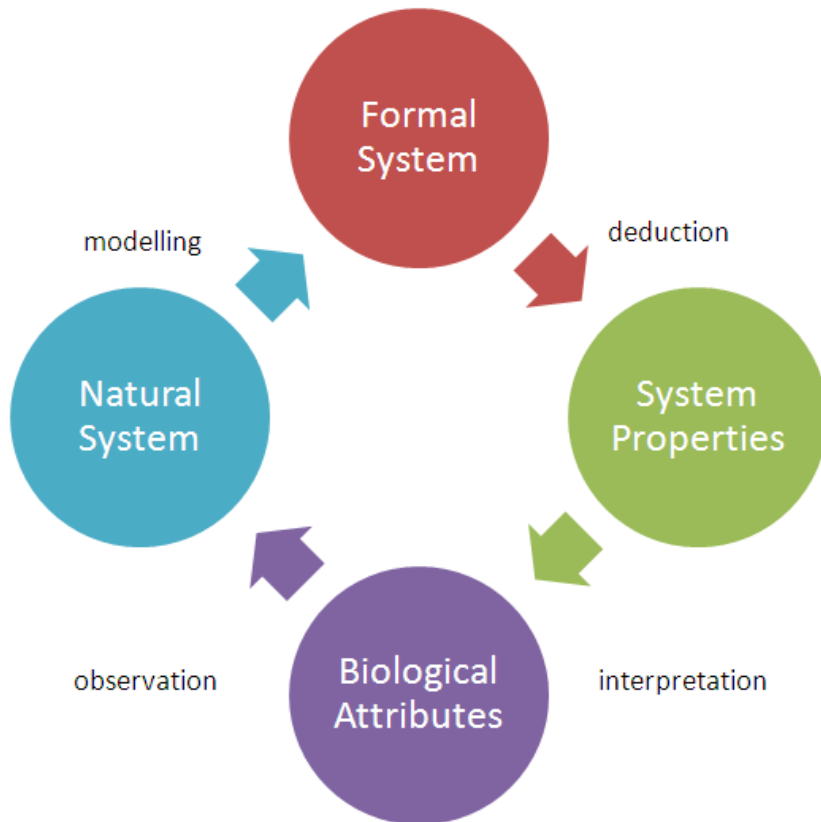


# ADVANCED ANALYSIS

- NGS data analysis
  - whole genome
  - whole exome
  - targeted re-sequencing
  - RNA-Seq
  - Gene fusion
  - CNV e SV detection
- Microarray data analysis
  - Gene expression
  - SNP call
  - CNV detection
  - Non-parametric linkage analysis
- Pathogenicity analysis of genetics mutations
- DNA-pattern matching and TFBS determination of model organisms
- Genetics and functional analyses of mutations
- Molecular targeting of microRNA in cancerous processes

# BIOLOGY AS A REACTIVE SYSTEM OF SYSTEMS

*When one can shift back and forth between seeing the trees and the forest one also sees which aspects of the trees are relevant to the description of the forest.*  
**(Yaneer Bar-Yam, 2000)**



$$S \subset \times \{S_j\}$$

$$R_\mu = l_{\mu 1}X_1 + l_{\mu 2}X_2 + \dots + l_{\mu n}X_n \longrightarrow^{k_\mu} \dots,$$

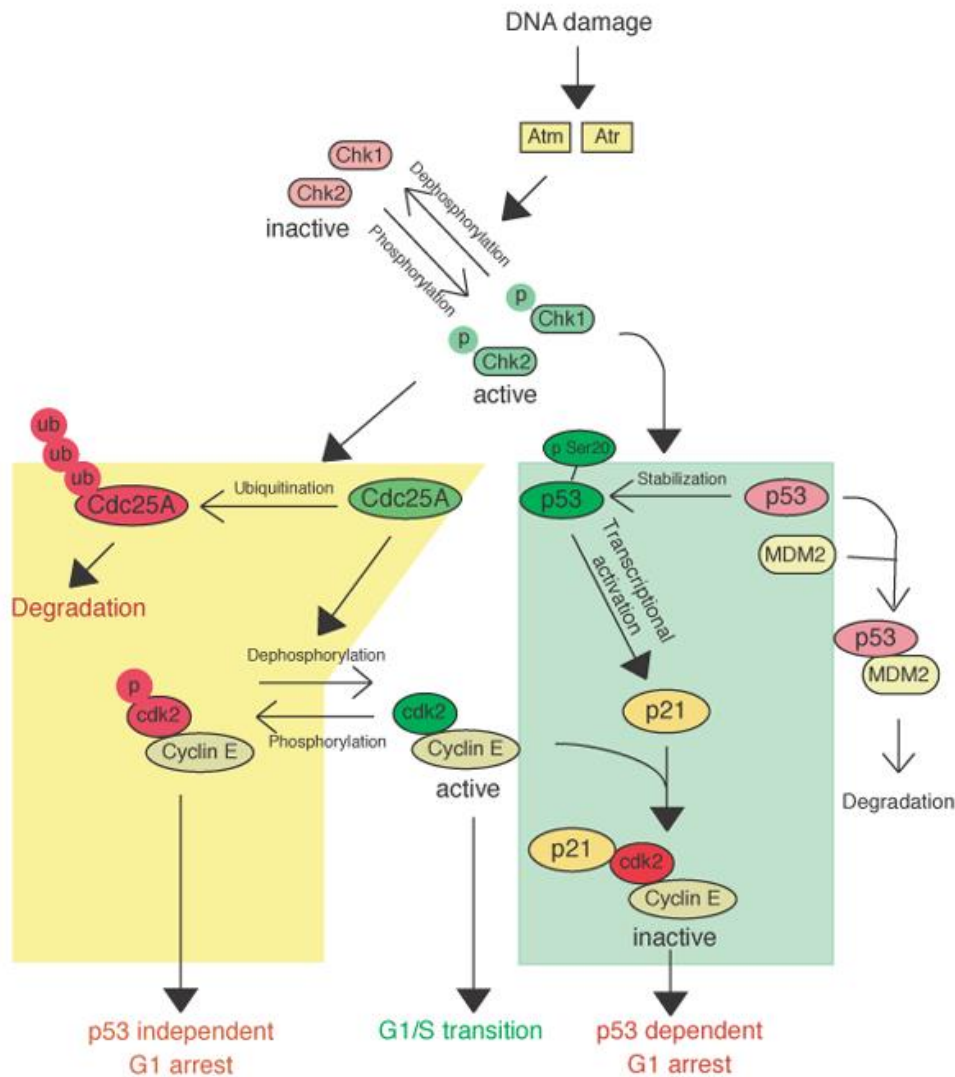
$$S \subseteq \Omega \times \Gamma,$$

$$\sigma : \Omega \rightarrow \Gamma \quad \omega \mapsto \gamma$$

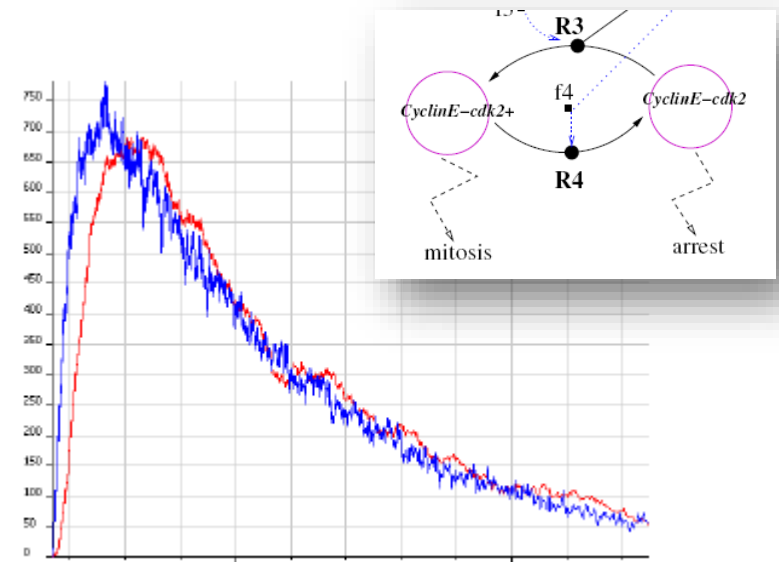
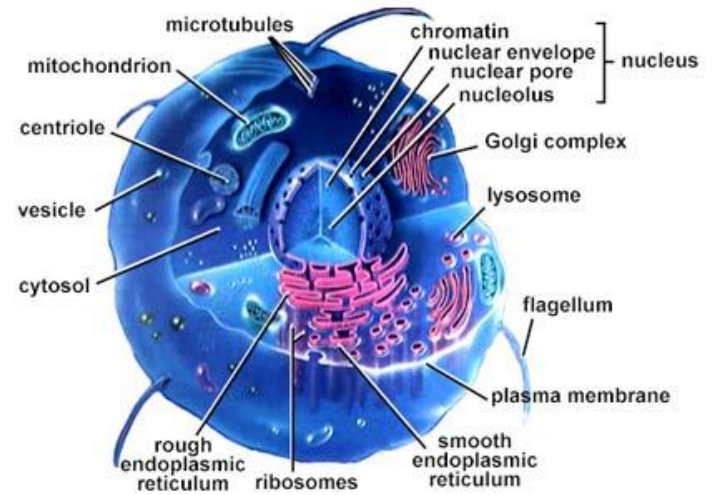
$$\Omega = \{\omega : I \rightarrow U\}$$

$$\Gamma = \{\gamma : I \rightarrow Y\}$$





Summary of p53 dependent and independent G1 arrest in response to stress



Expected behavior of two oscillators with high degradation activity (high amount of ubiquitin).





# MODELING (TEXTUAL)

```
object food, prey, predator
```

```
rule r1      food + prey 10 -> 2 prey + food
```

```
rule r2      prey + predator 0.01 -> 2 predator
```

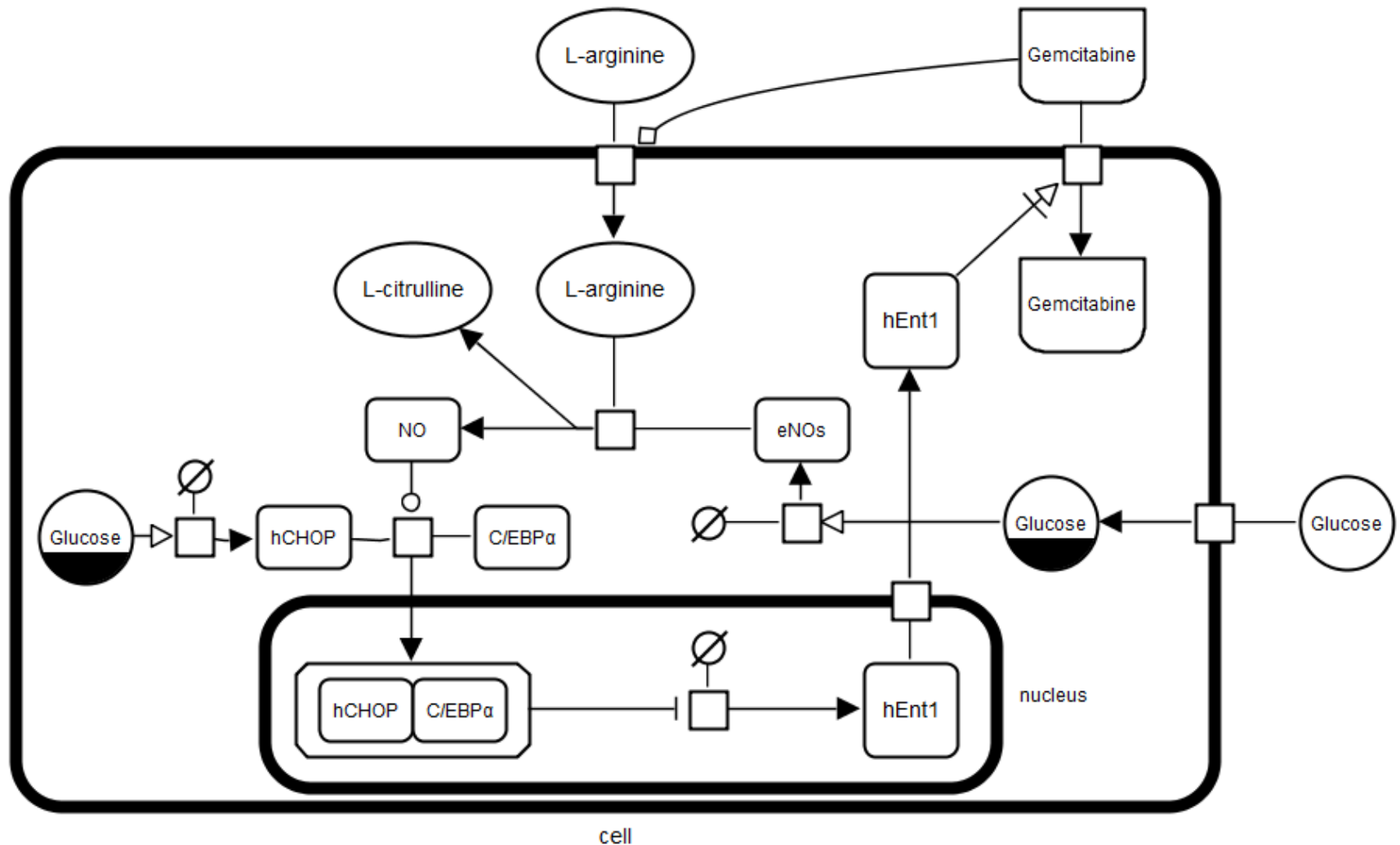
```
rule r3      predator 10 -> 0 predator
```

```
system 1 food, 1000 prey, 1000 predator, r1, r2, r3
```

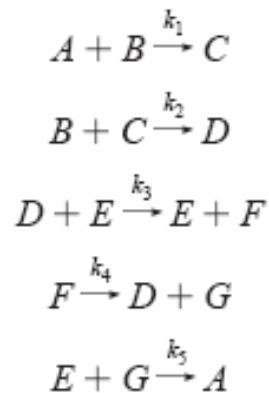
```
evolve 0-10000000
```

```
plot predator, prey
```

# MODELING (GRAPHICAL)

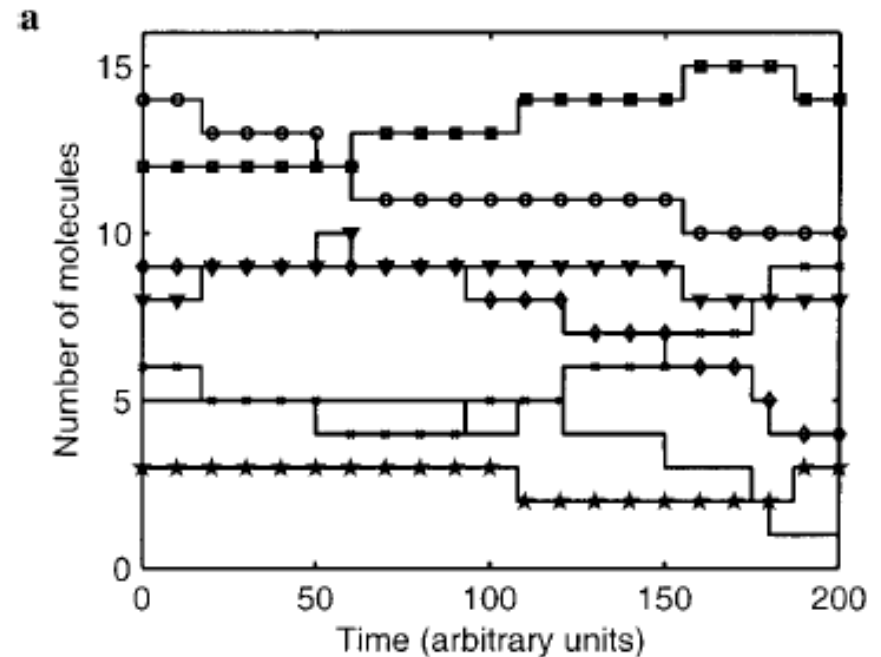


# STOCHASTIC SIMULATION



The state of the system in the stochastic framework is defined by the number of molecules of each species and changes discretely whenever one of the reactions is executed.

*Gibson-Bruck*

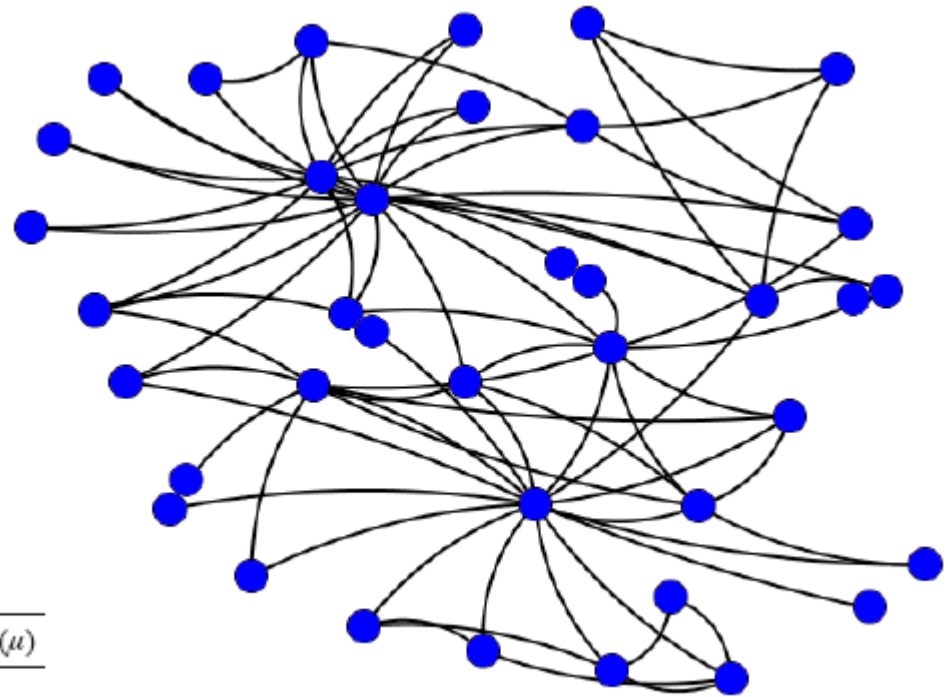
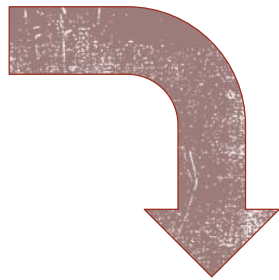
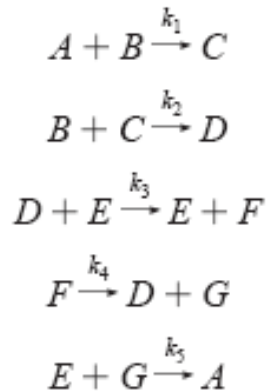


**b**

Time	0	17	50	60	93	108	121	150	155	175
#A	6	5	4	4	5	5	6	7	7	8
#B	14	13	12	11	11	11	11	11	10	10
#C	8	9	10	9	9	9	9	9	8	8
#D	12	12	12	13	13	14	14	14	15	15
#E	9	9	9	9	8	8	7	6	6	5
#F	3	3	3	3	3	2	2	2	2	2
#G	5	5	5	5	4	5	4	3	3	2
Reaction	—	1	1	2	5	4	5	5	2	5



# REACTIVE SYSTEMS AS GRAPHS

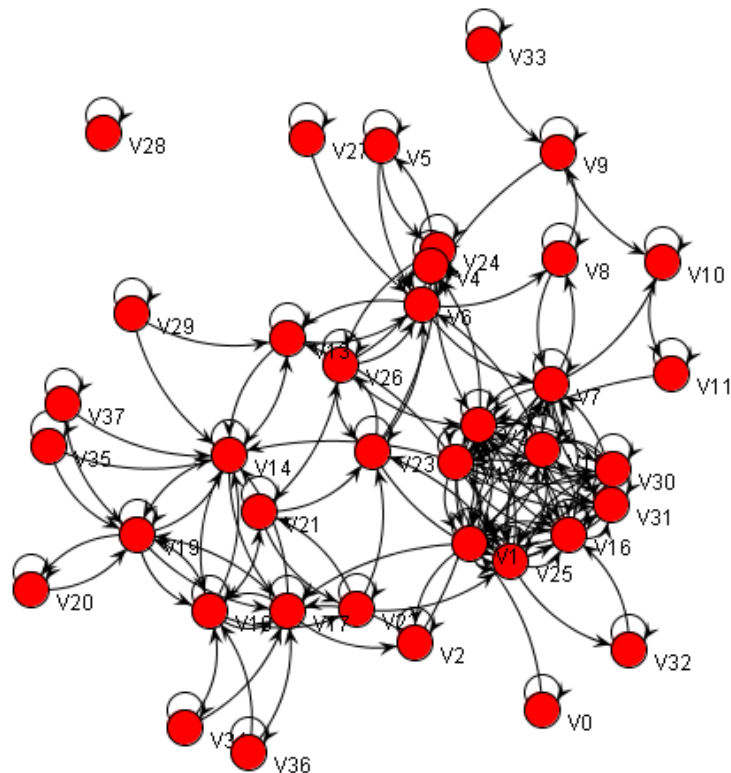


Workload  
Dependency  
Graph

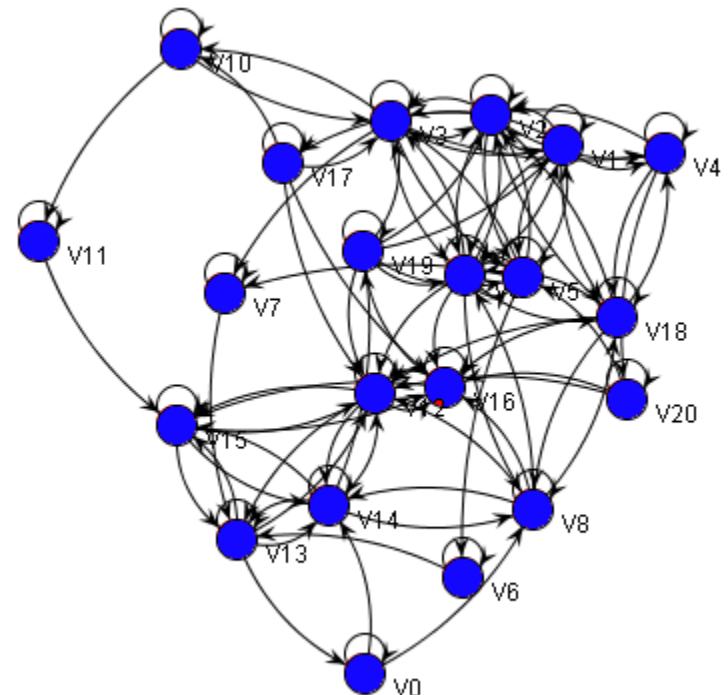
reaction	$a_\mu$	$DependsOn(a_\mu)$	$Affects(u)$
$A + B \xrightarrow{k_1} C$	$k_1 \times (\#A) \times (\#B)$	A, B	A, B, C
$B + C \xrightarrow{k_2} D$	$k_2 \times (\#B) \times (\#C)$	B, C	B, C, D
$D + E \xrightarrow{k_3} E + F$	$k_3 \times (\#D) \times (\#E)$	D, E	D, F
$F \xrightarrow{k_4} D + G$	$k_4 \times (\#F)$	F	D, F, G
$E + G \xrightarrow{k_5} A$	$k_5 \times (\#E) \times (\#G)$	E, G	A, E, G



# METABOLIC DEPENDENCY GRAPH



*Kinetic model of human erythrocytes*



*Metabolic model of plants metabolism*



# ORDER OF MAGNITUDE

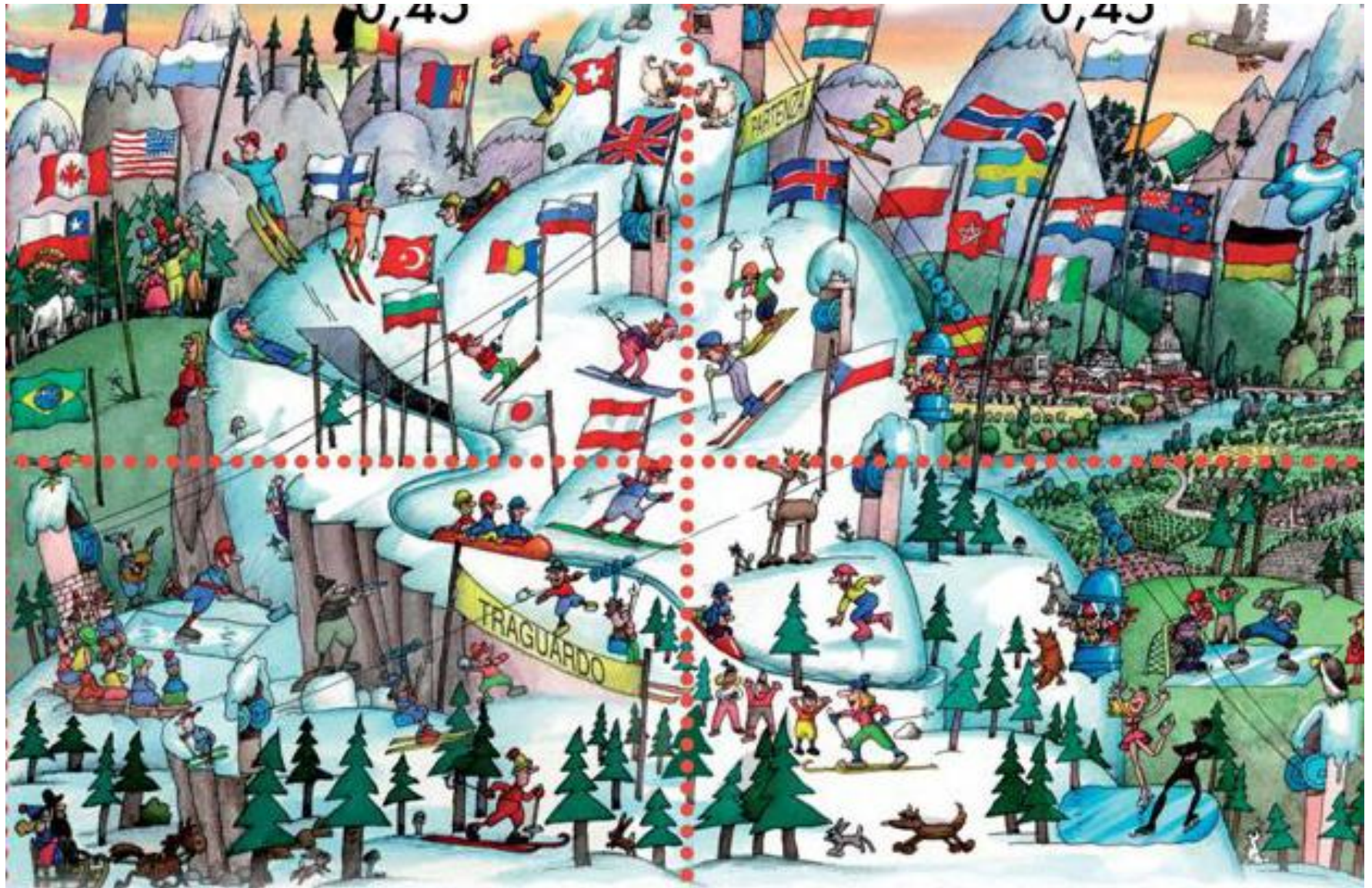
- How many genes? **23299 genes**
- Number of proteins in one cell:  **$\approx 100$  million**
- Different types of proteins in a cell:  **$\approx 20.000$**
- Number of cells in a young human being of average age and weight?  **$\approx 100.000$  billions**
- Number of cells (neurons) in human:  **$\approx 100$  billion**  
**(they die with a frequency of 500 per hour)**
- Average age of most cells:  **$\approx 1$  month**



# ORDER OF MAGNITUDE

- DNA length in a cell: **≈ 2 meters**
- Nucleotides number in a cell: **≈ 3.2 billion**
- DNA length in all your cells: **≈ 20 million Km**  
**(Earth's diameter: 40.075 km)**
- Human homologs genes in Drosophila: **60%**
- Human homologs genes in Mouse: **90%**
- Chemical functions of Banana that are similar to those occurred in humans: **50%**









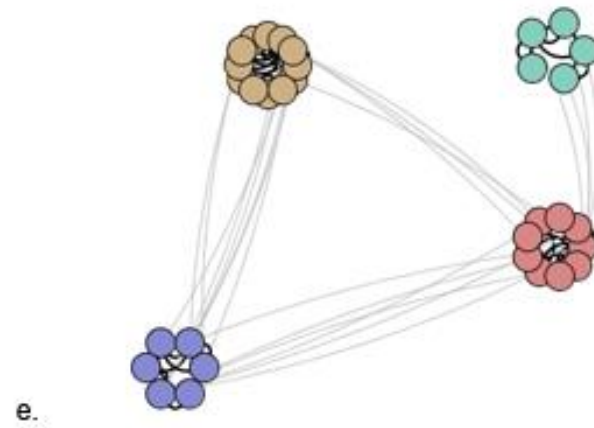
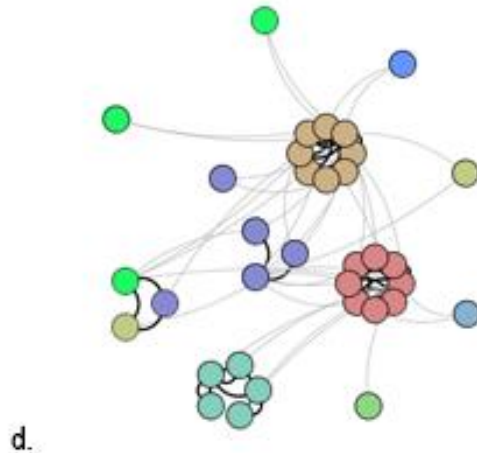
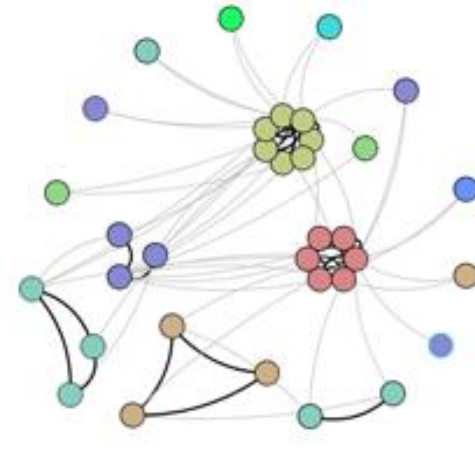
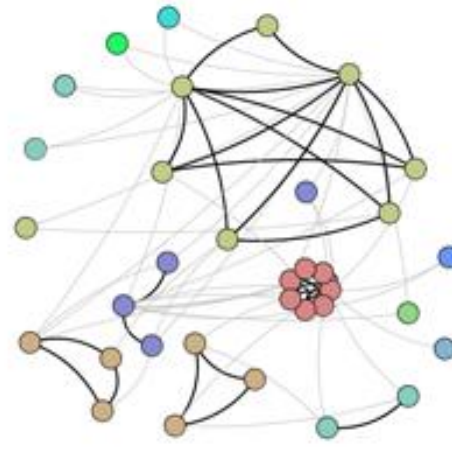
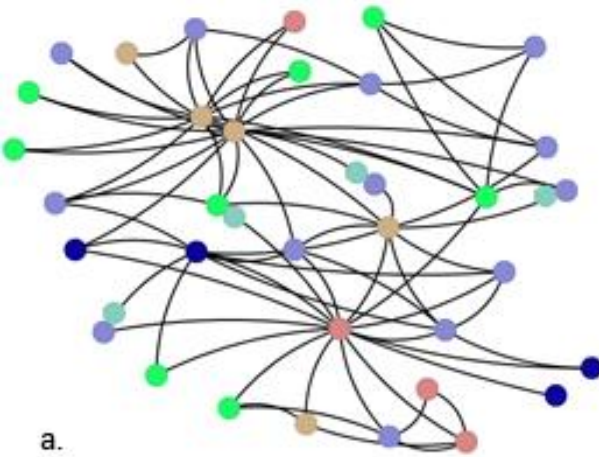




# CLUSTERING

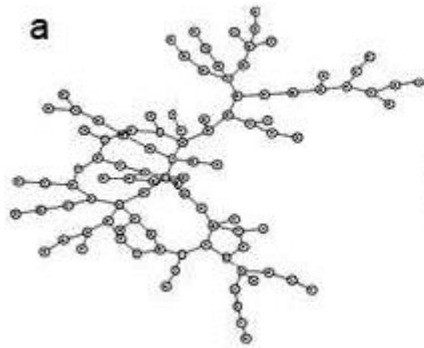
*The sole advantage of power is that  
you can do more good.*

**Baltasar Gracian (1601 - 1658)**

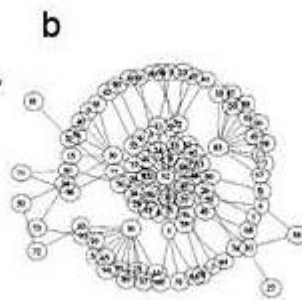


# GEOMETRIC PROPERTIES

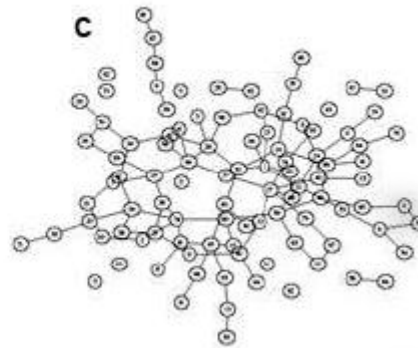
*The book of nature is written in  
the language of mathematics*  
**Galileo Galilei (1564 - 1642)**



betweenness



closeness



degree

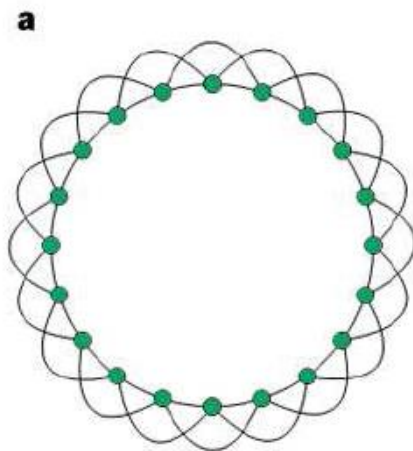
Degree:  $k_i = \sum_{j=1}^n A_{ij}$

Eigenvector:  $x_i = \frac{1}{\lambda} \sum_{j=1}^N A_{ij} x_j$

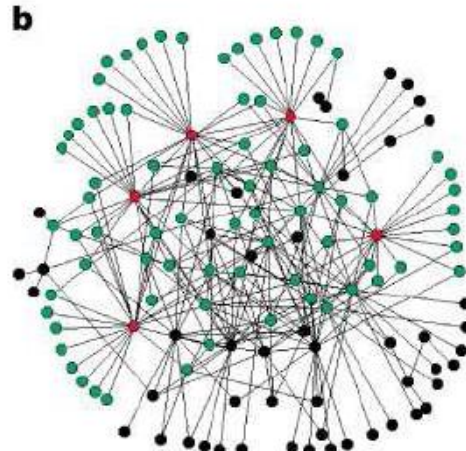
Betweenness:  $C_B(v) = \sum_{s \neq v \neq t \in V} \frac{\sigma_{st}(v)}{\sigma_{st}}$

Closeness:  $\frac{\sum_{t \in V} d_G(v, t)}{n}$

Clustering coeff.:  $C_i = \frac{|e_{jk}|}{k_i(k_i - 1)}$



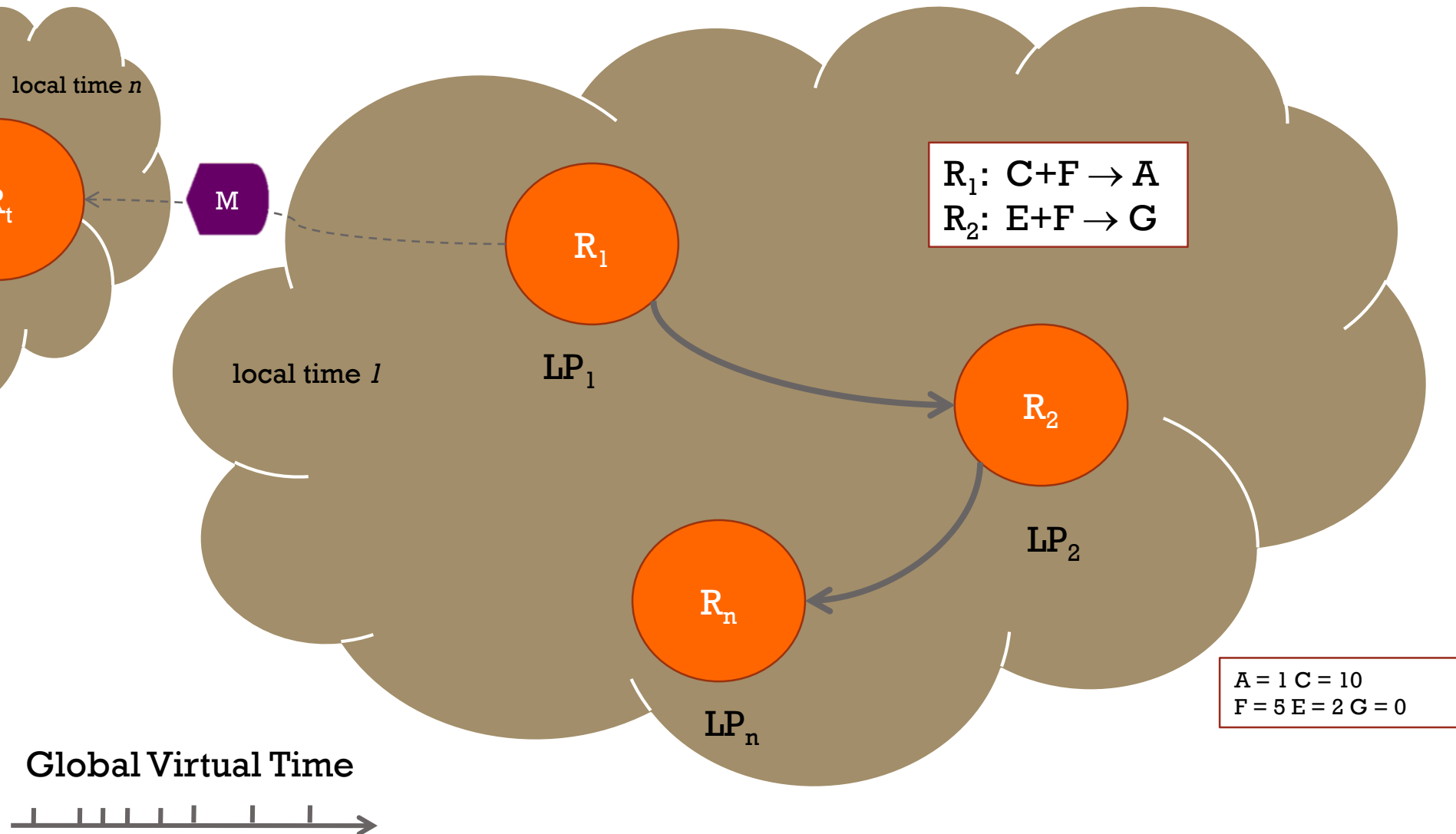
Small-world



Scale-free



# DISTRIBUTED SIMULATION ENVIRONMENT



# ARCHITECTURE AND POLICIES

Multi-core machine (256 cores)





# QUESTIONS?



**Bfx unit @CSS**

*Tommaso Mazza*

[t.mazza@css-mendel.it](mailto:t.mazza@css-mendel.it)

*Via Regina Margherita, 261  
Roma*

*Via dei Cappuccini, 1  
S. Giovanni Rotondo  
Foggia*