ePTRI
EUROPEAN PAEDIATRIC TRANSLATIONAL RESEARCH INFRASTRUCTURE
ITALIAN NODE

Proteomics Data Processing from Profiling to Network Analysis

Dario Di Silvestre
Institute for Biomedical Technologies (ITB-CNR)

EUROPEAN BIOTECH WEEK
INNOVATION IS IN OUR GENES

FEDERCHIMICA
ASSOBIOTEC
Associazione nazionale per lo sviluppo delle biotecnologie

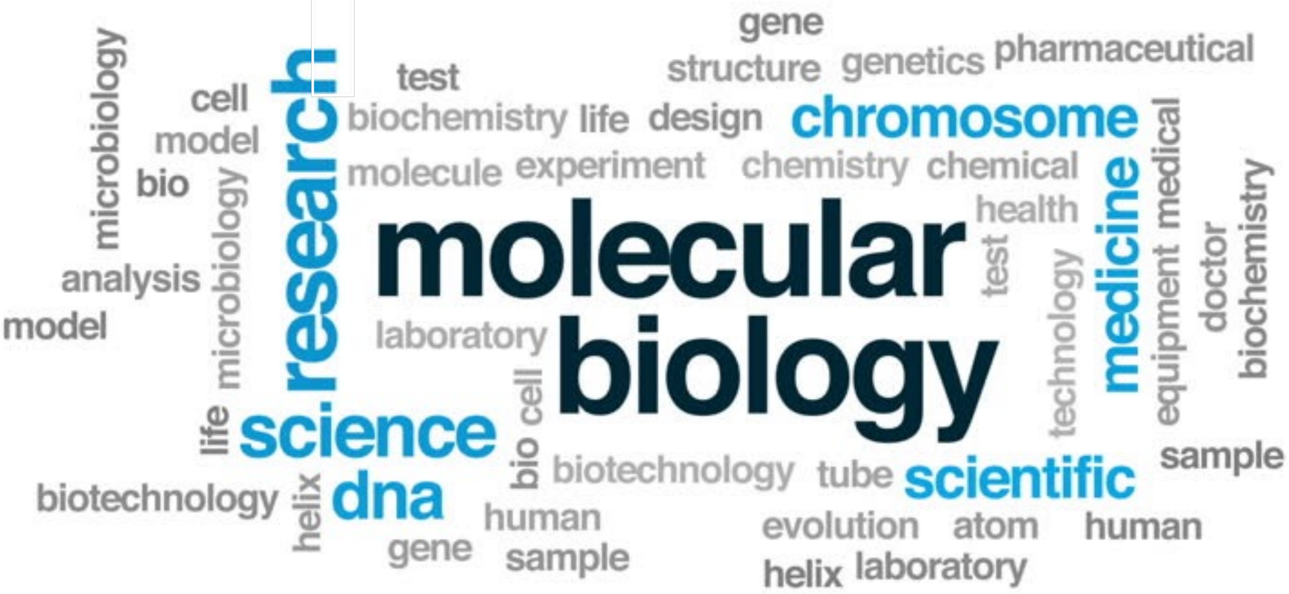
PART OF GLOBAL BIOTECH WEEK

Coordinated by
CVBF
Consorzio per Valutazioni Biologiche e Farmacologiche

General Assembly led by
National Research Council of Italy

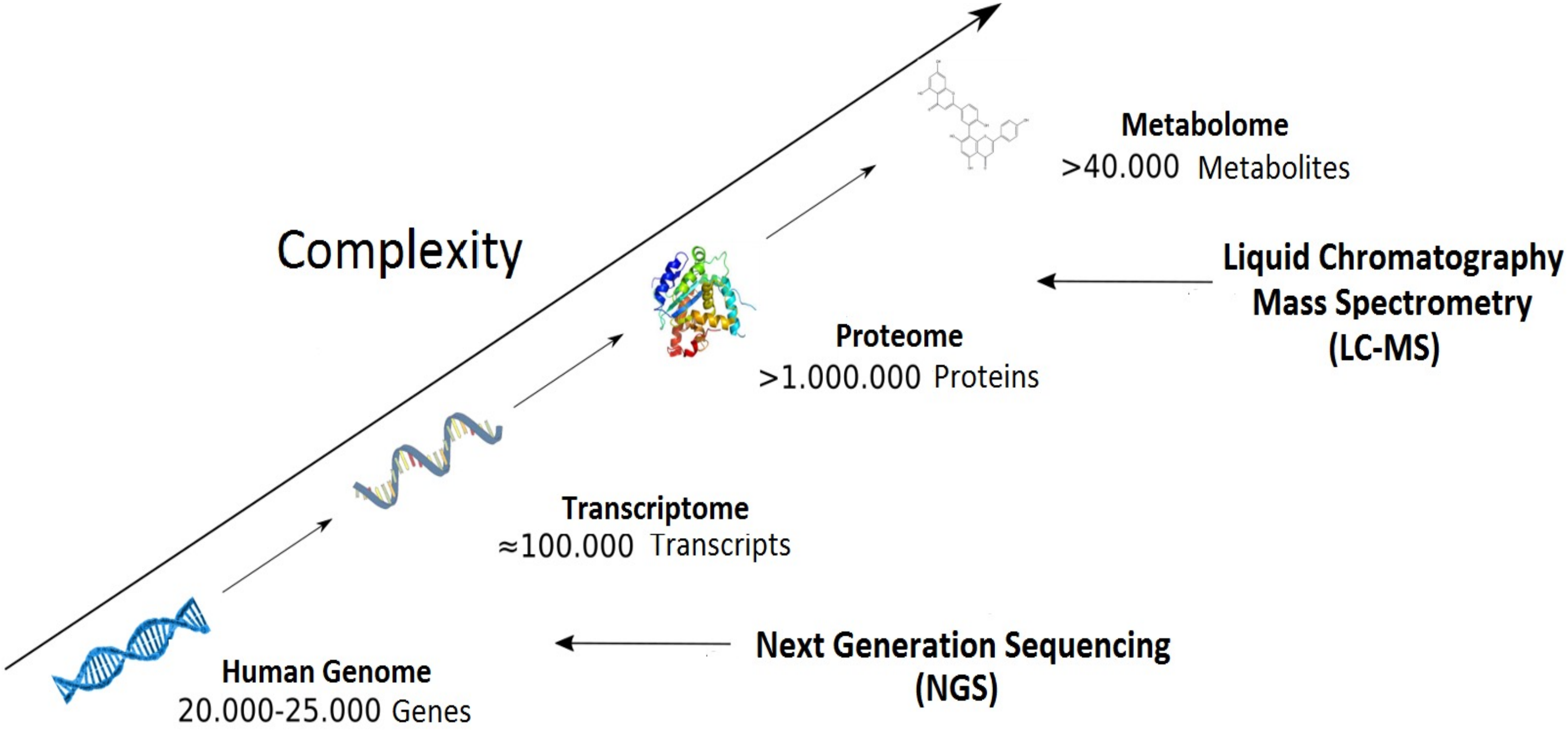
Scientific Committee led by
INCiPiT Italian Network for Paediatric Clinical Trials

20th century



Warren Weaver
(1894 – 1978)

Living organisms complexity



Complexity

Metabolome
>40.000 Metabolites

Liquid Chromatography Mass Spectrometry (LC-MS)

Proteome
>1.000.000 Proteins

Transcriptome
≈100.000 Transcripts

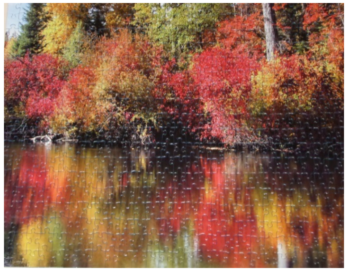
Human Genome
20.000-25.000 Genes

Next Generation Sequencing (NGS)

Systems biology and emerging properties


Reductionism

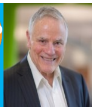
?


Holism

> “Science is built up with facts, as a house is with stones.
> But a collection of facts is no more a science than a heap of stones is a house”
>
> **Jules Henri Poincaré**
> (1854 – 1912)

<u>SYSTEMS BIOLOGY</u>
DEFINES AND ANALYSES THE INTERRELATIONSHIPS
OF ALL OF THE ELEMENTS IN A FUNCTIONING SYSTEM
IN ORDER TO UNDERSTAND HOW THE SYSTEM WORKS

Institute for Systems Biology


Leroy Hood

ePTRI
European Paediatric Translational Research Infrastructure
ITALIAN NODE

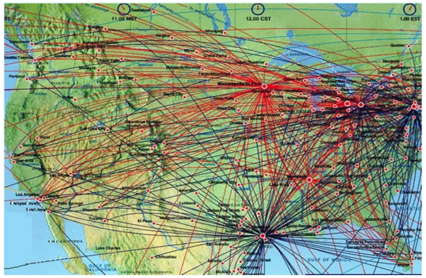
Networks as a universal language

Internet

Electronic Circuit

Social Network

Biological Network



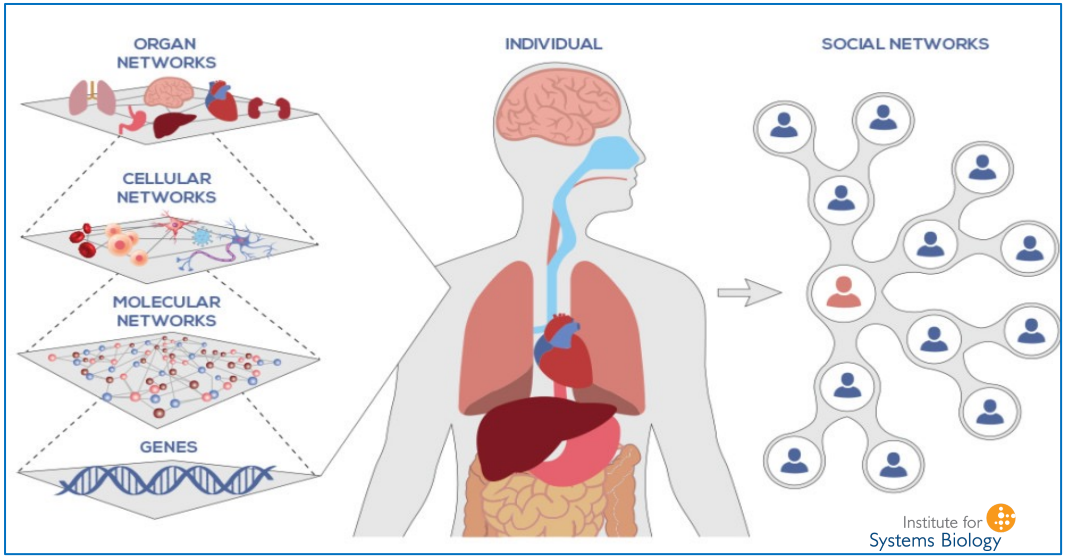
Transportation networks

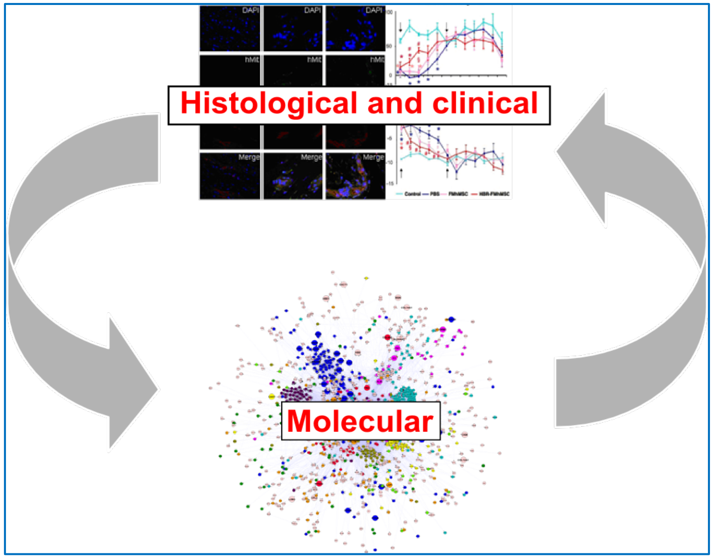
We are caught in an inescapable network of mutuality....
Whatever affects one directly, affects all indirectly.

EPTRI
EUROPEAN PAEDIATRIC TRANSLATIONAL RESEARCH INFRASTRUCTURE
ITALIAN NODE

Martin Luther King Jr.
(1929-1968)

Network of networks

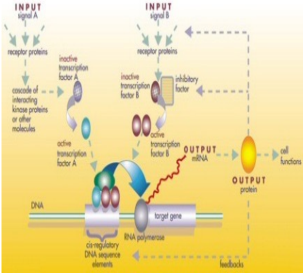




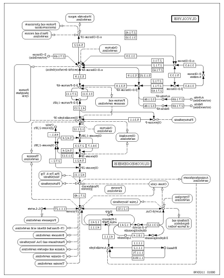
ePTRI
EUROPEAN PAEDIATRIC TRANSLATIONAL RESEARCH INFRASTRUCTURE
ITALIAN NODE

Biological Networks

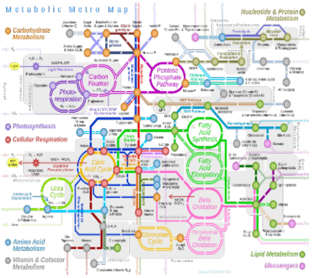
Gene regulatory



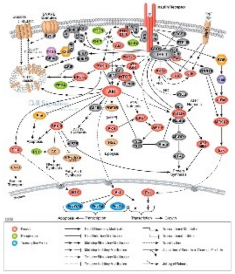
Pathway



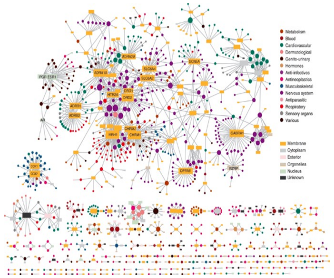
Metabolic



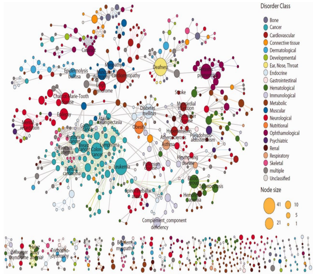
Cell Signaling



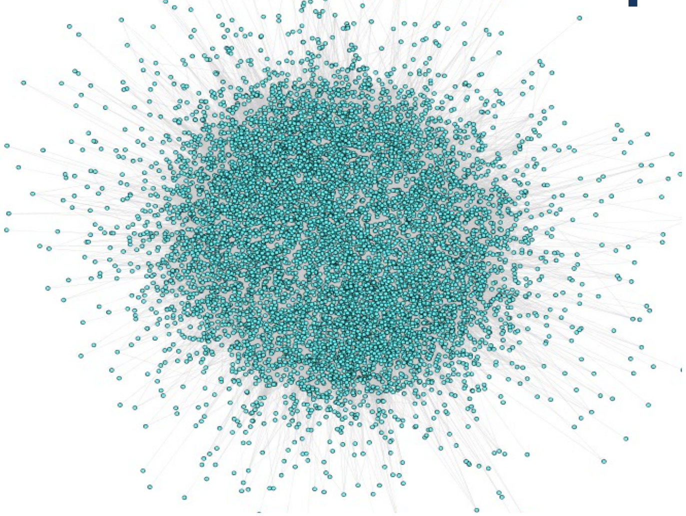
Drug-target



Human diseases



Protein-Protein Interaction and Co-Expression networks

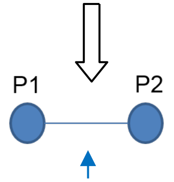


EPTRI
European Paediatric Translational Research Infrastructure
ITALIAN NODE

Protein-Protein Interaction (PPI) vs Co-Expression

Protein-Protein Interaction (PPI) networks

(by Literature and Experimental Data)



physical/functional interaction

(Protein complexes, Pathways etc.)

+

⇦

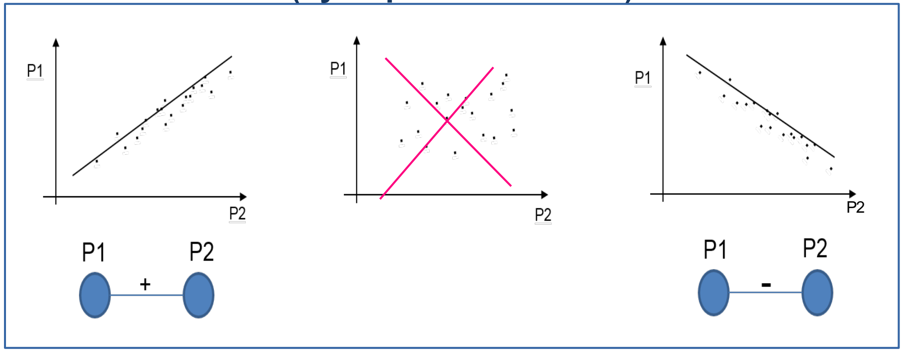
Protein Data Matrix

Sample	*Prot 1*	*Prot 2*	*Prot 3*	·	*Prot n*
Healthy 1	*SpC 1.1*	*SpC 1.2*	*SpC 1.3*	·	*SpC 1.n*
Healthy 2	*SpC 2.1*	*SpC 2.2*	*SpC 2.3*	·	*SpC 2.n*
Healthy 3	*SpC 3.1*	*SpC 3.2*	*SpC 3.3*	·	*SpC 3.n*
·	·	·	·	·	·
Healthy *m*	*SpC m.1*	*SpC m.2*	*SpC m.3*	·	*SpC m.n*
Diseased	*SpC 1.1*	*SpC 1.2*	*SpC 1.3*	.	*SpC 1.n*
Diseased	*SpC 2.1*	*SpC 2.2*	*SpC 2.3*	.	*SpC 2.n*
Diseased	*SpC 3.1*	*SpC 3.2*	*SpC 3.3*	.	*SpC 3.n*
·	·	·	·	·	·
Diseased *m*	*SpC m.1*	*SpC m.2*	*SpC m.3*	.	*SpC m.n*

WGCNA ⇩

Co-expression networks

(by Experimental Data)



ePTRI
European Paediatric Translational Research Infrastructure
ITALIAN NODE

What is a network?

Graph *G*=(*V*,*E*) is a set of vertices (or nodes) *V* and edges *E*

$N = \{n_1, n_2, n_{...}, n_m\}$

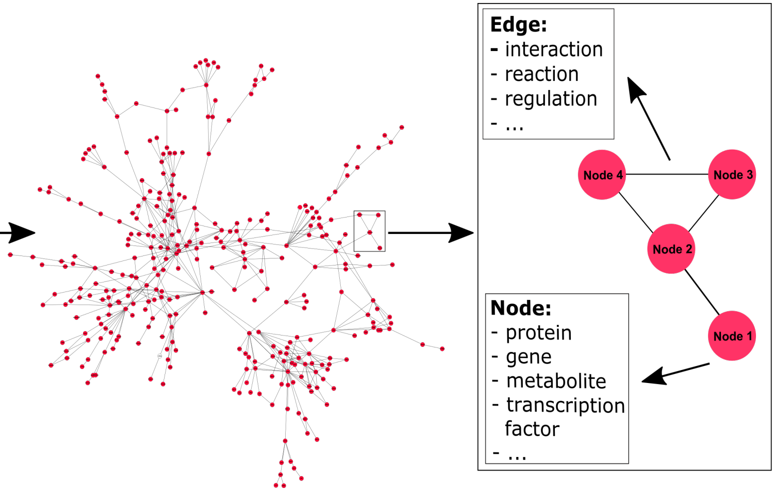
$E = \{e_1, e_2, e_{...}, e_k\}$

$A = \{a_{(ni,nj)}\}$

$i = j = 1, 2 \ldots m$

$$a_{(ni,nj)} = \begin{cases} 1 & \text{if } a_{(ni,nj)} \in E \\ 0 & \text{otherwise} \end{cases}$$

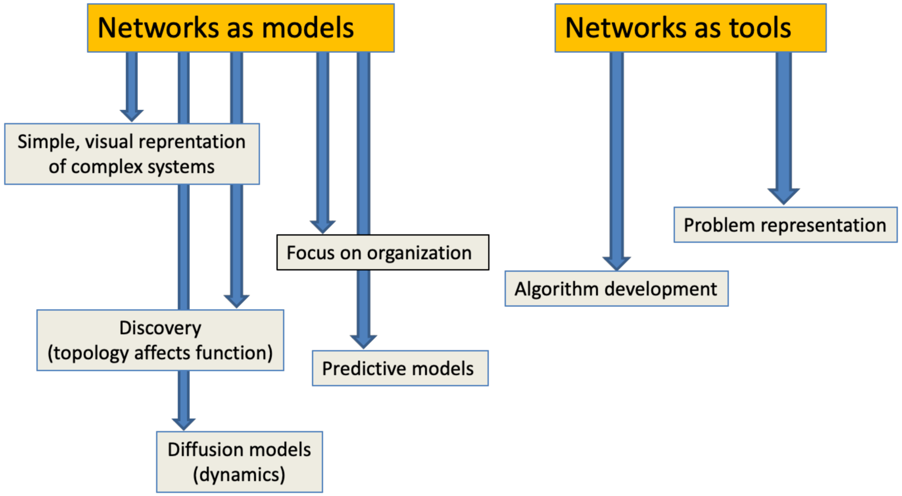
$$\begin{array}{c|c} & \begin{array}{cccc} n_1 & n_2 & n_{...} & n_m \end{array} \\ \begin{array}{c} n_1 \\ n_2 \\ n_{...} \\ n_m \end{array} & \begin{bmatrix} 0 & 1 & 0 & \ldots \\ 1 & 0 & 1 & \ldots \\ 0 & 1 & 0 & \ldots \\ \ldots & \ldots & \ldots & \ldots \end{bmatrix} \end{array}$$



Edge:
- interaction
- reaction
- regulation
- …

Node:
- protein
- gene
- metabolite
- transcription factor
- …

Vella et al. *EURASIP J Bioinform Syst Biol.* 2017(1):6.



Networks as models

Simple, visual reprentation of complex systems

Focus on organization

Discovery (topology affects function)

Predictive models

Diffusion models (dynamics)

Networks as tools

Problem representation

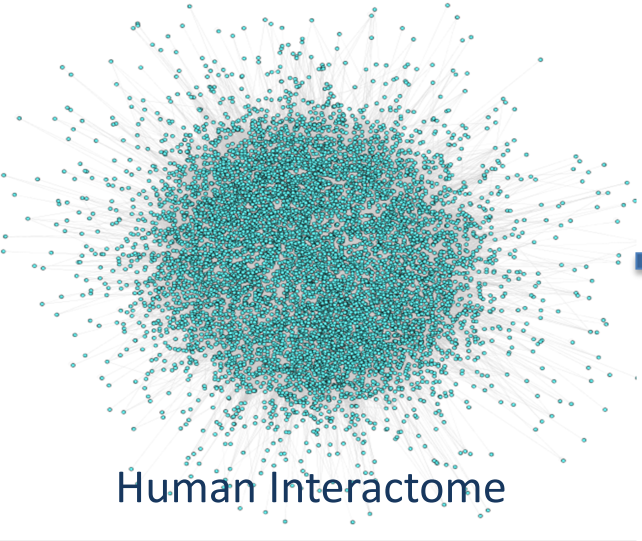
Algorithm development

ePTRI

European Paediatric Translational Research Infrastructure

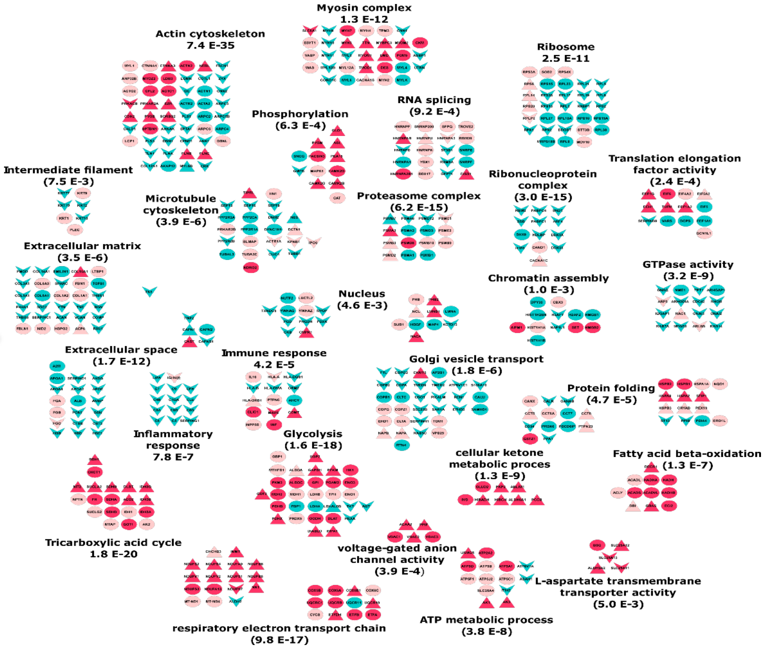
ITALIAN NODE

FUNCTIONAL MODULES



+

DEPs
and
GO ANNOTATIONS



BBA Gen Subj. 2017;1861(5 Pt A):1190-1199.

SCIENTIFIC REPORTS

OPEN

MTGO: PPI Network Analysis Via Topological and Functional Module Identification

d: 1 November 2017
d: 28 February 2018

Danila Vella[1,2], Simone Marini[3], Francesca Vitali[4,6,7,8], Dario Di Silvestre[9], Giancarlo Mauri[2] & Riccardo Bellazzi[1,4,5]

frontiers in Genetics

MTGO-SC, A Tool to Explore Gene Modules in Single-Cell RNA Sequencing Data

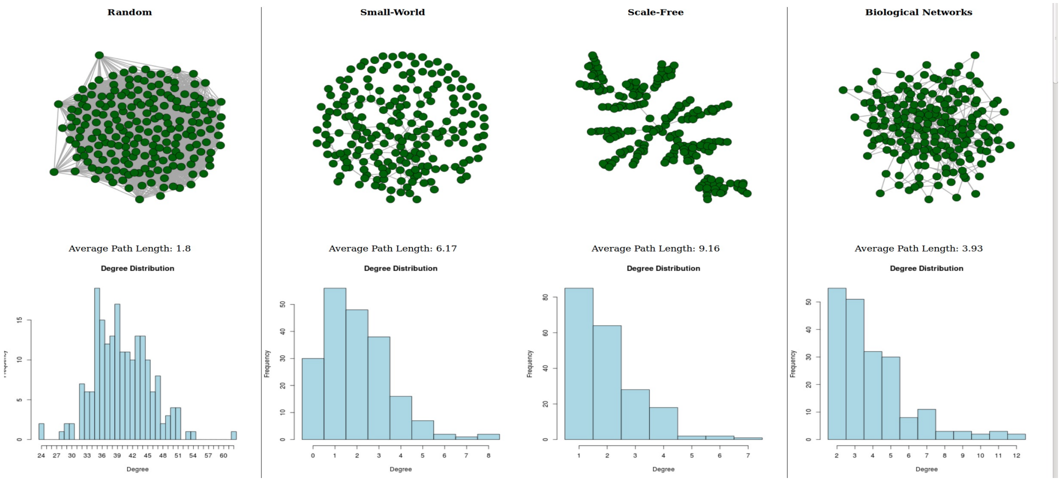
*Nelson Nazzicari[1], Danila Vella[2,3], Claudia Coronnello[4], Dario Di Silvestre[5], Riccardo Bellazzi[3,6] and Simone Marini[6,7]**

ePTRI
European Paediatric Translational Research Infrastructure
ITALIAN NODE

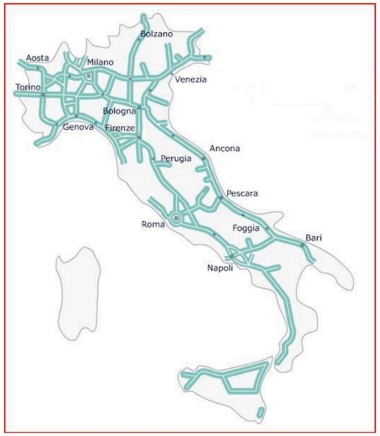
D. Vella

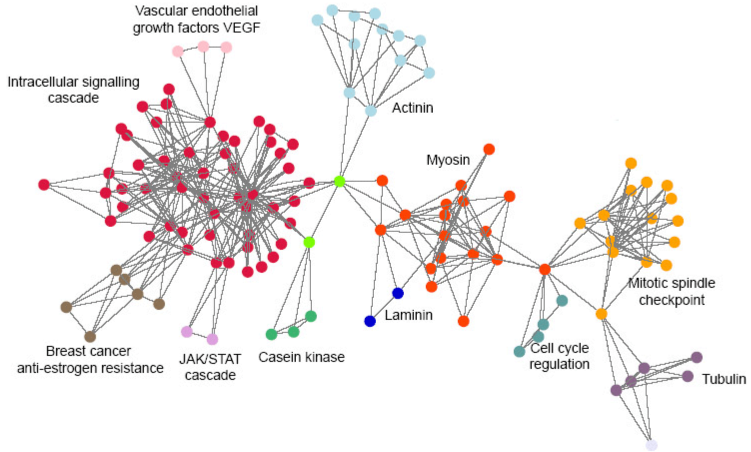
Fondazione Ri.MED

Structure affects function?



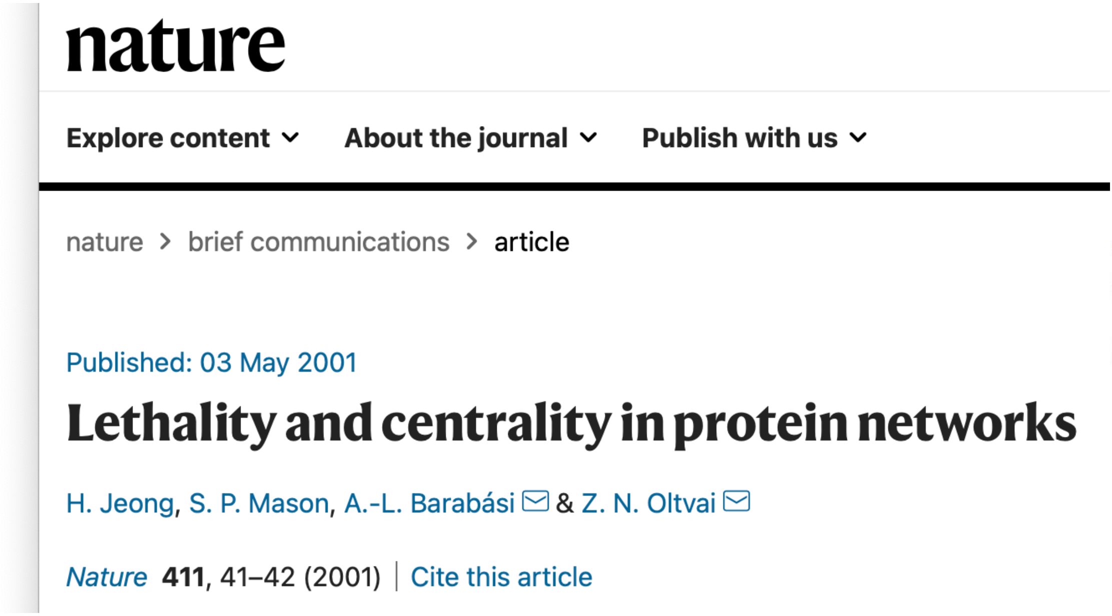
The degree distribution plots the number of nodes that have a given degree *k* as a function of *k*.
The shape of the degree distribution allows us to distinguish among types of networks.





EPTRI
European Paediatric Translational Research Infrastructure
ITALIAN NODE

Structure affects function?



nature

Explore content ⌄ About the journal ⌄ Publish with us ⌄

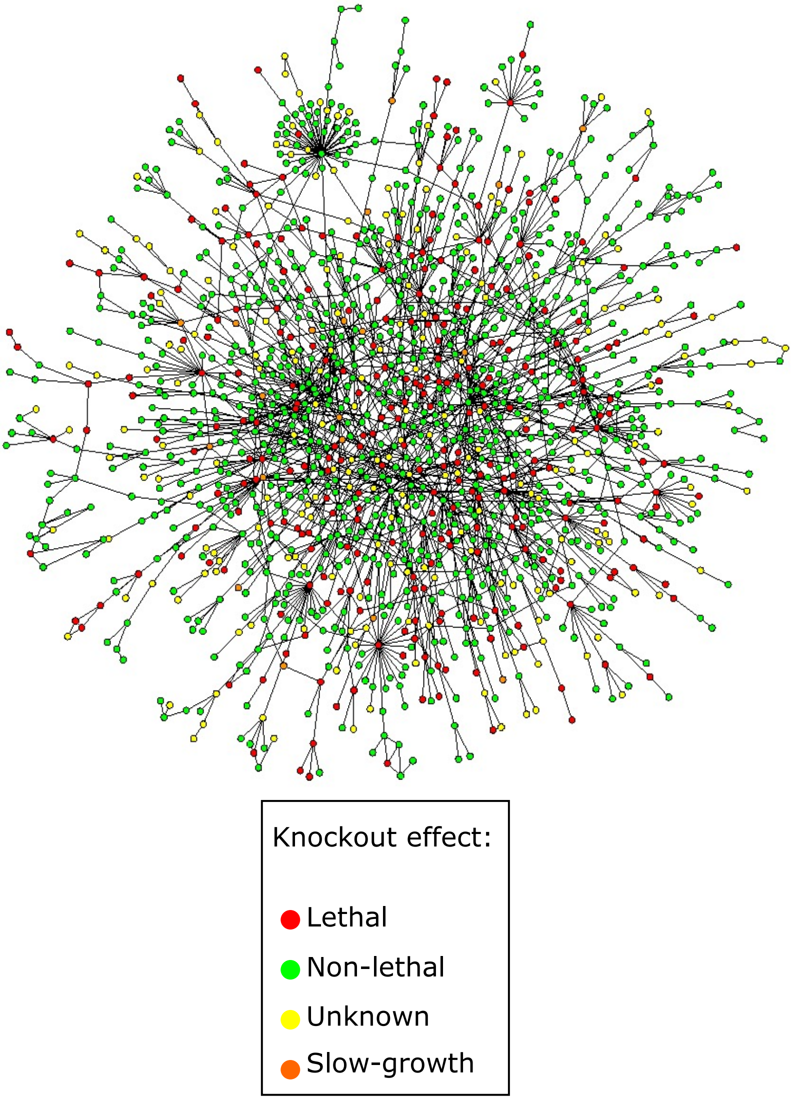
nature › brief communications › article

Published: 03 May 2001

Lethality and centrality in protein networks

H. Jeong, S. P. Mason, A.-L. Barabási ✉ & Z. N. Oltvai ✉

Nature **411**, 41–42 (2001) | Cite this article



EPTRI
European Paediatric Translational Research Infrastructure
ITALIAN NODE

Network Modules

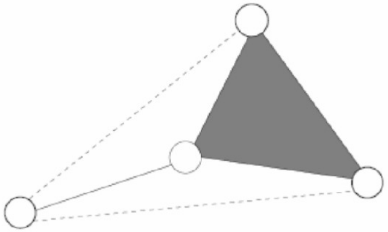
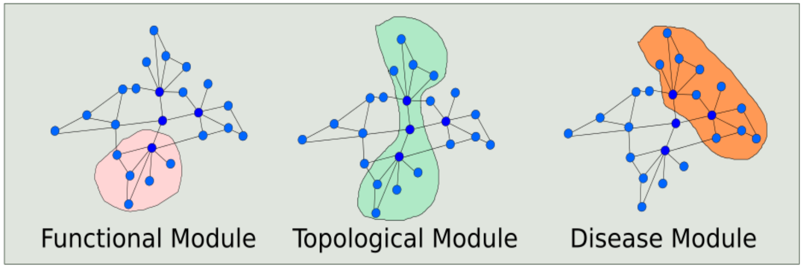


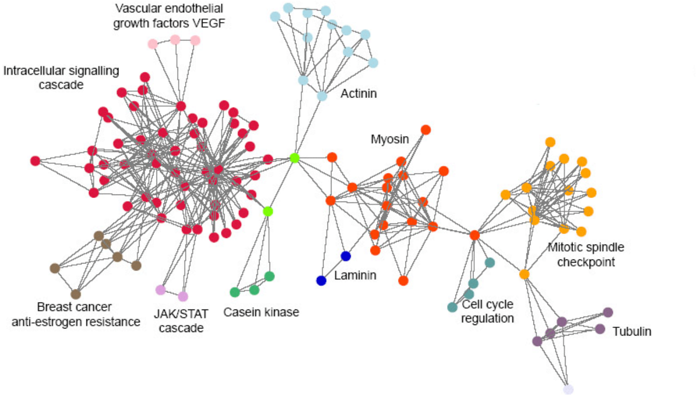
Fig. 1.5 The clustering coefficient of the central vertex is 1/3. This is because its degree is three and its neighbours can be connected each other in three different ways. Of these possibilities (dashed line) only one is actually realized (solid line) and therefore $C_i = 1/3$. The three connected vertices form the coloured triangle. For that reason, sometimes the clustering coefficient of a vertex is defined through the number of triangles it belongs to.

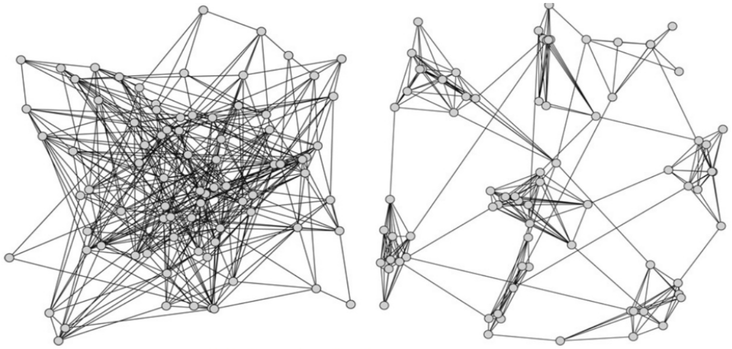
$$C_i = \frac{1}{k_i(k_i-1)/2} \sum_{jk} a_{ij} a_{ik} a_{jk}$$

Clustering coefficient



Module = highly interacting subset of nodes cooperating to perform functions





Courtesy of Elsevier, Inc., http://www.sciencedirect.com. Used with permission.
Source: Schaeffer, Satu Elisa. "Graph Clustering." *Computer Science Review* 1, no. 1 (2007): 27-64.

Goal: divide the graph into subgraphs each of which has lots of internal connections and few connections to the rest of the graph

Computer scince Review 1, 1 (2007);27-64.

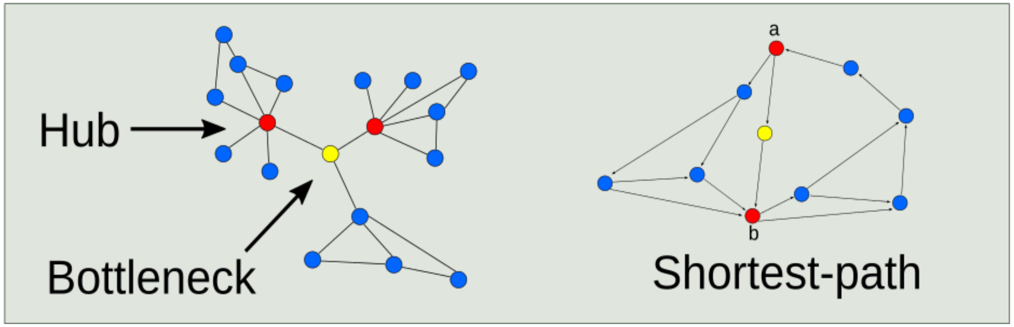
ePTRI
European Paediatric Translational Research Infrastructure
ITALIAN NODE

HUBS and Shortest-Path

How much is my Node *v* **central** or **peripheral**?

Bottlenecks
- Connects subnetworks
- Restricts information flow

Hubs
- Highly connected
- Central point of control

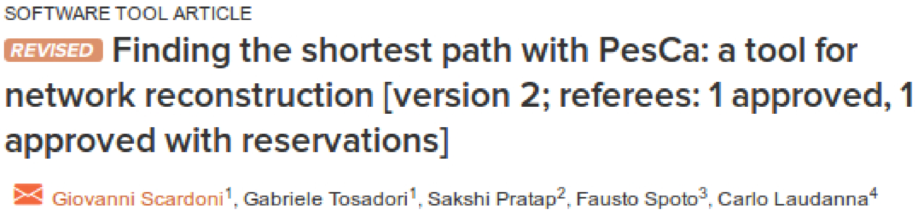


BIOINFORMATICS **APPLICATIONS NOTE** Vol. 25 no. 21 2009, pages 2857–2859
doi:10.1093/bioinformatics/btp517

Systems biology

Analyzing biological network parameters with CentiScaPe

Giovanni Scardoni[1,*], Michele Petterlini[1] and Carlo Laudanna[1,2]



SOFTWARE TOOL ARTICLE

REVISED **Finding the shortest path with PesCa: a tool for network reconstruction [version 2; referees: 1 approved, 1 approved with reservations]**

Giovanni Scardoni[1], Gabriele Tosadori[1], Sakshi Pratap[2], Fausto Spoto[3], Carlo Laudanna[4]

In biological terms:

Proteins possibly bringing in communication sets of regulatory proteins that interact with many other proteins, suggesting a central super-regulatory role or a critical target of a regulatory pathway.

EPTRI
European Paediatric Translational Research Infrastructure
ITALIAN NODE

Network Centralities

Diameter (*ΔG*):

the maximal distance (shortest path) amongst all the distances calculated between each couple of vertexes in graph G.

Average distance (*AvDG*):

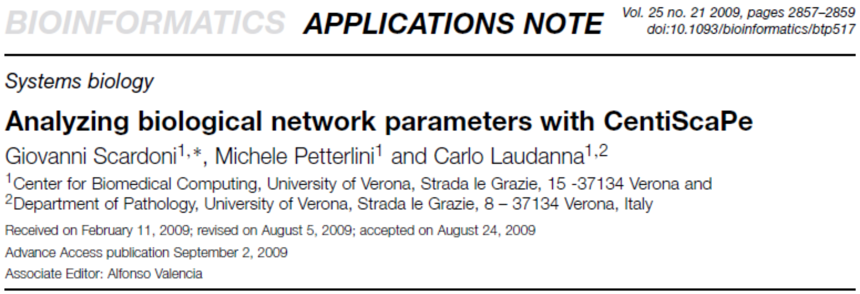
the average shortest path of a graph G, corresponding to the summa of all shortest paths between vertex couples divided for the total number of vertex couples.

High value: the network can be "disperse"
Low value: the network is "compact"

In biological terms:

- Easiness of the proteins to communicate
- Reciprocal influence between proteins
- Functional convergence

Node and Edge Centralities



BIOINFORMATICS **APPLICATIONS NOTE** *Vol. 25 no. 21 2009, pages 2857–2859*
doi:10.1093/bioinformatics/btp517

Systems biology

Analyzing biological network parameters with CentiScaPe

Giovanni Scardoni[1,*], Michele Petterlini[1] and Carlo Laudanna[1,2]

[1]Center for Biomedical Computing, University of Verona, Strada le Grazie, 15 -37134 Verona and
[2]Department of Pathology, University of Verona, Strada le Grazie, 8 – 37134 Verona, Italy

Received on February 11, 2009; revised on August 5, 2009; accepted on August 24, 2009
Advance Access publication September 2, 2009
Associate Editor: Alfonso Valencia

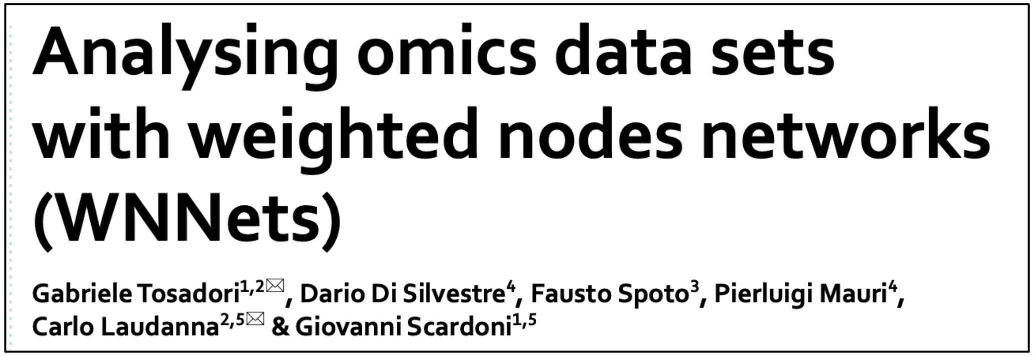
Node Centralities

- Degree
- Centroid
- Stress
- Closeness
- Betweenness
- Radiality
- Eccentricity
- Eigenvector
- Bridging

Edge Centrality

Edge Betweenness

Scientific Reports (2021) 11:14447



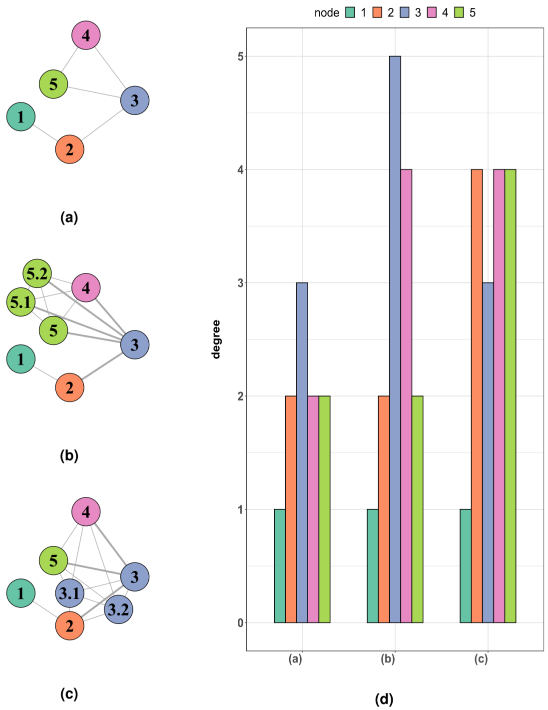
Analysing omics data sets with weighted nodes networks (WNNets)

Gabriele Tosadori[1,2]✉, Dario Di Silvestre[4], Fausto Spoto[3], Pierluigi Mauri[4], Carlo Laudanna[2,5]✉ & Giovanni Scardoni[1,5]

G. Scardoni

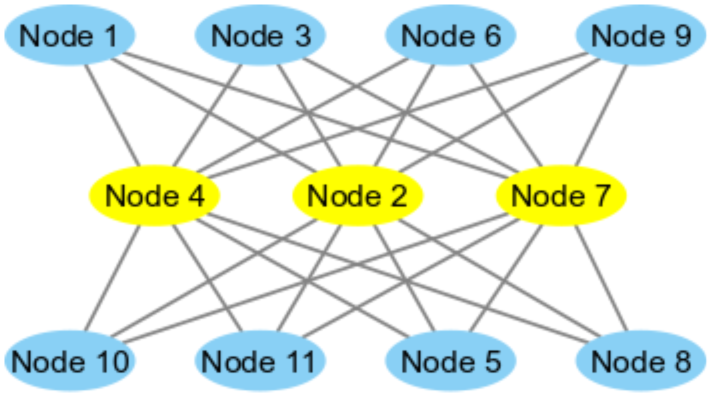
G. Tosadori

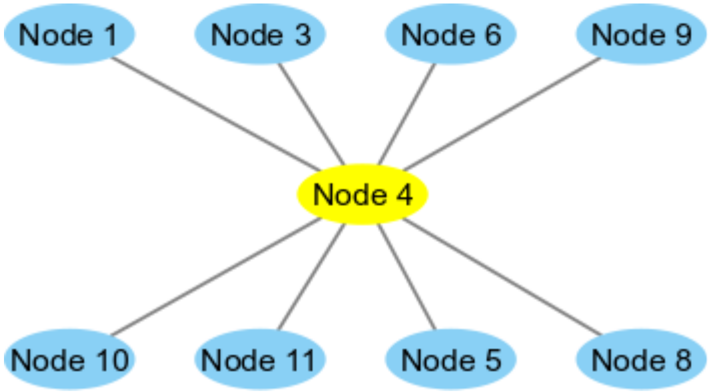
Center for Biomedical Computing



EPTRI
European Paediatric Translational Research Infrastructure
ITALIAN NODE

Node Centralities





Degree	Betweenness
8	18.7

Node 4

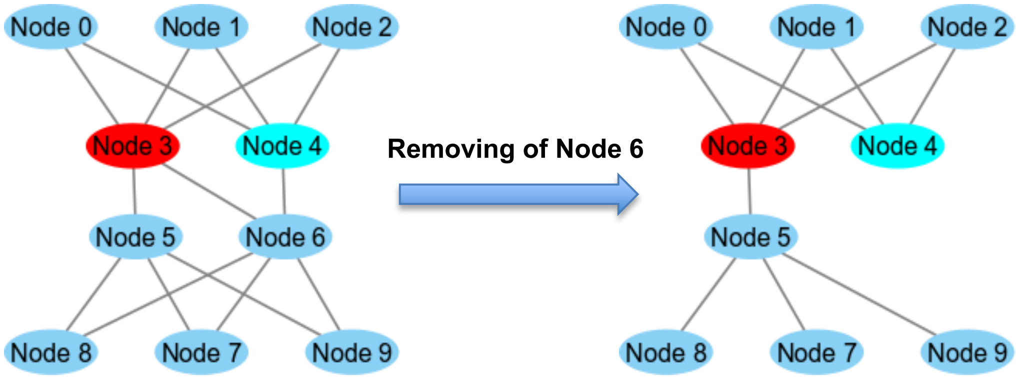
Degree	Betweenness
8	56

Degree $(deg(k))$	Number of nodes adjacent to a given node *v*
S.-P. Betweenness $(C_{spb}(v))$	It is calculated considering couples of nodes (v1, v2) and counting the number of shortest paths linking v1 and v2 and passing through a node n. Then, the value is related to the total number of shortest paths linking v1 and v2.

EPTRI
European Paediatric Translational Research Infrastructure
ITALIAN NODE

Networks structure is dynamic

Can the role of a node strongly affected by removing or adding a given node?



Removing of Node 6

Node 4 becomes a “peripheral” node
Node 5 and **Node 3** become more important communicating nodes
Node 3 is more robust than **Node 4**

Betweenness

Node	Net A	Net B	Interference (Node 6)
5	15	42.86	-27.86
3	32.05	41.67	-9.61
0	0.98	3.97	-2.99
1	0.98	3.97	-2.99
2	0.98	3.97	-2.99
7	0.98	0	0.98
8	0.98	0	0.98
9	0.98	0	0.98
4	15	3.57	11.43
6	32.05		

Node 3		Node 4	
Robustness	0.046	Robustness	0.036

Interference

$$Int_C(i,n,G) = \frac{C(G,n)}{\sum_{j\in N} C(G,j)} - \frac{C(G_{|i},n)}{\sum_{j\in N} C(G_{|i},j)}$$

Robustness

$$Rob_{Btw}(n,G) = \frac{1}{\max_{i\in N_{|n}}\{|Int_{Btw}(i,n,G)|\}}.$$

EPTRI
European Paediatric Translational Research Infrastructure
ITALIAN NODE

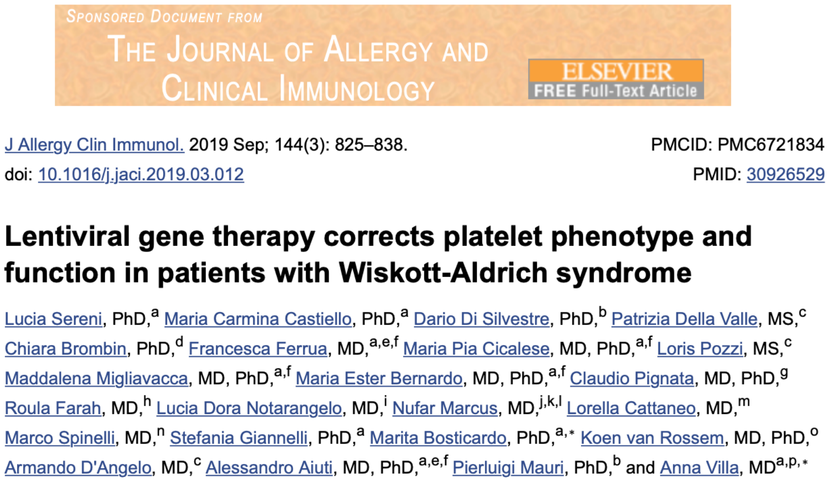
Case1- Wiskott-Aldrich syndrome

Sponsored Document from
The Journal of Allergy and Clinical Immunology
ELSEVIER FREE Full-Text Article

J Allergy Clin Immunol. 2018 Oct; 142(4): 1272–1284. PMID: 29421274
Published online 2018 Feb 6. doi: 10.1016/j.jaci.2017.12.1000

Autonomous role of Wiskott-Aldrich syndrome platelet deficiency in inducing autoimmunity and inflammation

Lucia Sereni, PhD,[a,b] Maria Carmina Castiello, PhD,[a] Francesco Marangoni, PhD,[c] Achille Anselmo, PhD,[d] Dario di Silvestre, PhD,[e] Sara Motta, PhD,[e] Elena Draghici, BS,[a] Stefano Mantero, BS,[d,f] Adrian J. Thrasher, MD, PhD,[g] Silvia Giliani, PhD,[h] Alessandro Aiuti, MD, PhD,[a,b,i] Pierluigi Mauri, PhD,[e] Luigi D. Notarangelo, MD,[j] Marita Bosticardo, PhD,[a] and Anna Villa, MD[a,f]



Sponsored Document from
The Journal of Allergy and Clinical Immunology
ELSEVIER FREE Full-Text Article

J Allergy Clin Immunol. 2019 Sep; 144(3): 825–838. PMCID: PMC6721834
doi: 10.1016/j.jaci.2019.03.012 PMID: 30926529

Lentiviral gene therapy corrects platelet phenotype and function in patients with Wiskott-Aldrich syndrome

Lucia Sereni, PhD,[a] Maria Carmina Castiello, PhD,[a] Dario Di Silvestre, PhD,[b] Patrizia Della Valle, MS,[c] Chiara Brombin, PhD,[d] Francesca Ferrua, MD,[a,e,f] Maria Pia Cicalese, MD, PhD,[a,f] Loris Pozzi, MS,[c] Maddalena Migliavacca, MD, PhD,[a,f] Maria Ester Bernardo, MD, PhD,[a,f] Claudio Pignata, MD, PhD,[g] Roula Farah, MD,[h] Lucia Dora Notarangelo, MD,[i] Nufar Marcus, MD,[j,k,l] Lorella Cattaneo, MD,[m] Marco Spinelli, MD,[n] Stefania Giannelli, PhD,[a] Marita Bosticardo, PhD,[a,*] Koen van Rossem, MD, PhD,[o] Armando D'Angelo, MD,[c] Alessandro Aiuti, MD, PhD,[a,e,f] Pierluigi Mauri, PhD,[b] and Anna Villa, MD[a,p,*]

Wiskott-Aldrich syndrome is a **rare** (the incidence in USA is **1 in 250,000** live male births) genetic disorder caused by **WAS gene** mutations. It is and characterized by **immune deficiency** and manifests itself since **childhood** with eczema, recurrent and relapsing infections and bleeding disorders. The syndrome is also associated with an **increased risk of autoimmune, lymphoma and leukemia**. Affected individuals show marked **deficiency of platelets** which are typically smaller than normal.

At the **San Raffaele-Telethon Institute for Gene Therapy (SR-Tiget)** in Milan, **the first study in the world** for the human experimentation of gene therapy with **lentiviral vectors** for the treatment of this disease was launched in April 2010, thanks to the strategic alliance between San Raffaele Hospital, the Telethon Foundation and GlaxoSmithKline, recently transferred to Orchard Therapeutics.

The "ex vivo" **gene therapy** therapeutic protocol involves the removal of **adult stem cells from the marrow or mobilized blood** of patients, their correction by introducing the vector containing the therapeutic gene and finally reinfusion into the patient.

A. Villa

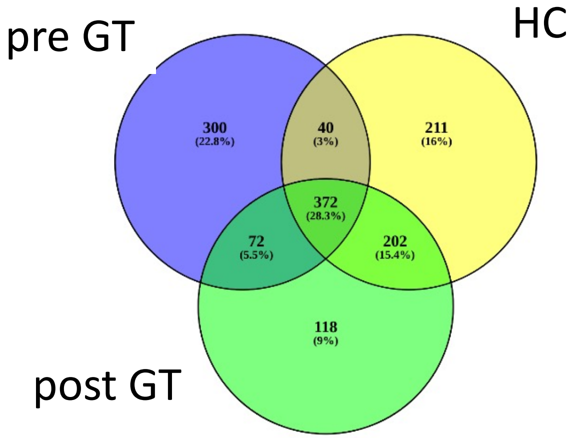
L. Sereni

OSPEDALE SAN RAFFAELE

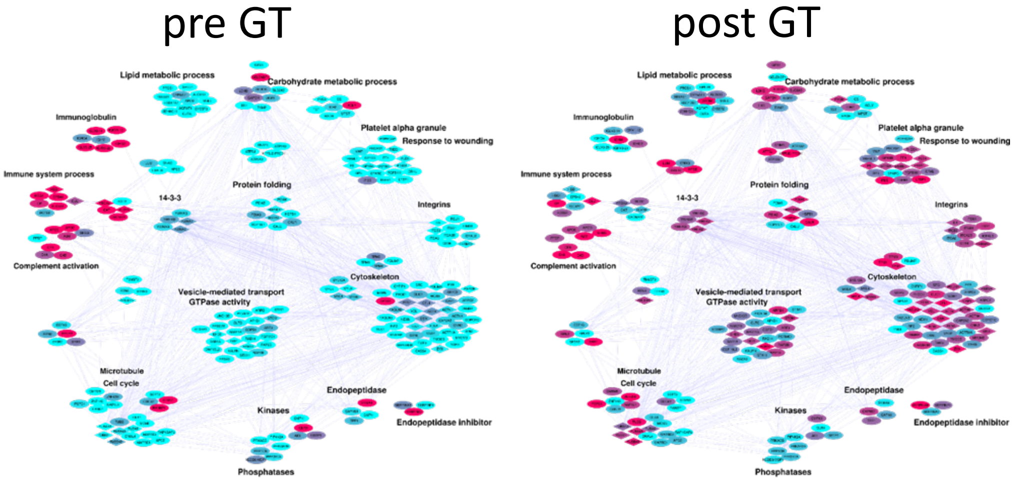
TeleThon

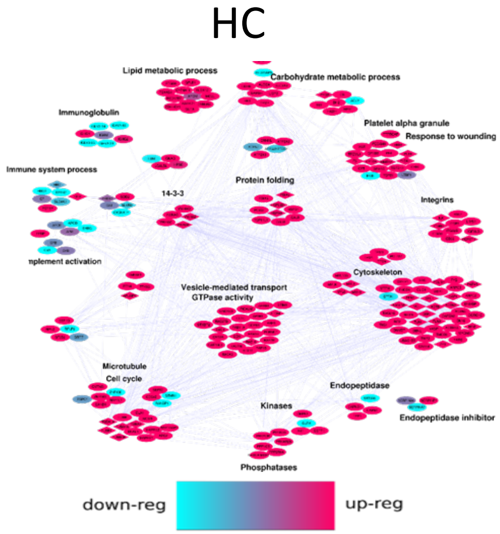
EPTRI
European Paediatric Translational Research Infrastructure
ITALIAN NODE

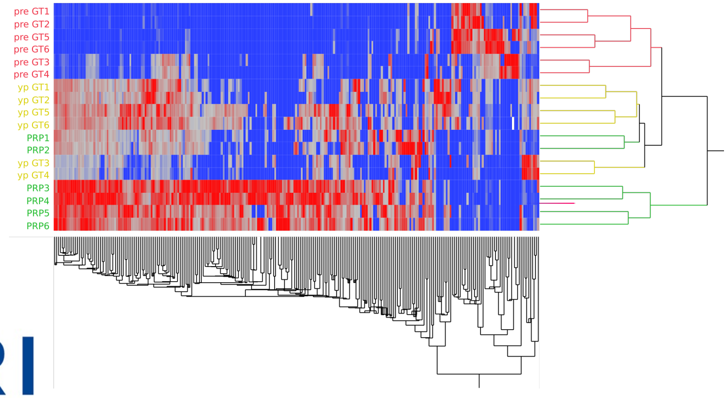
Case1- Wiskott-Aldrich syndrome



1314 nr proteins







255 DEPs

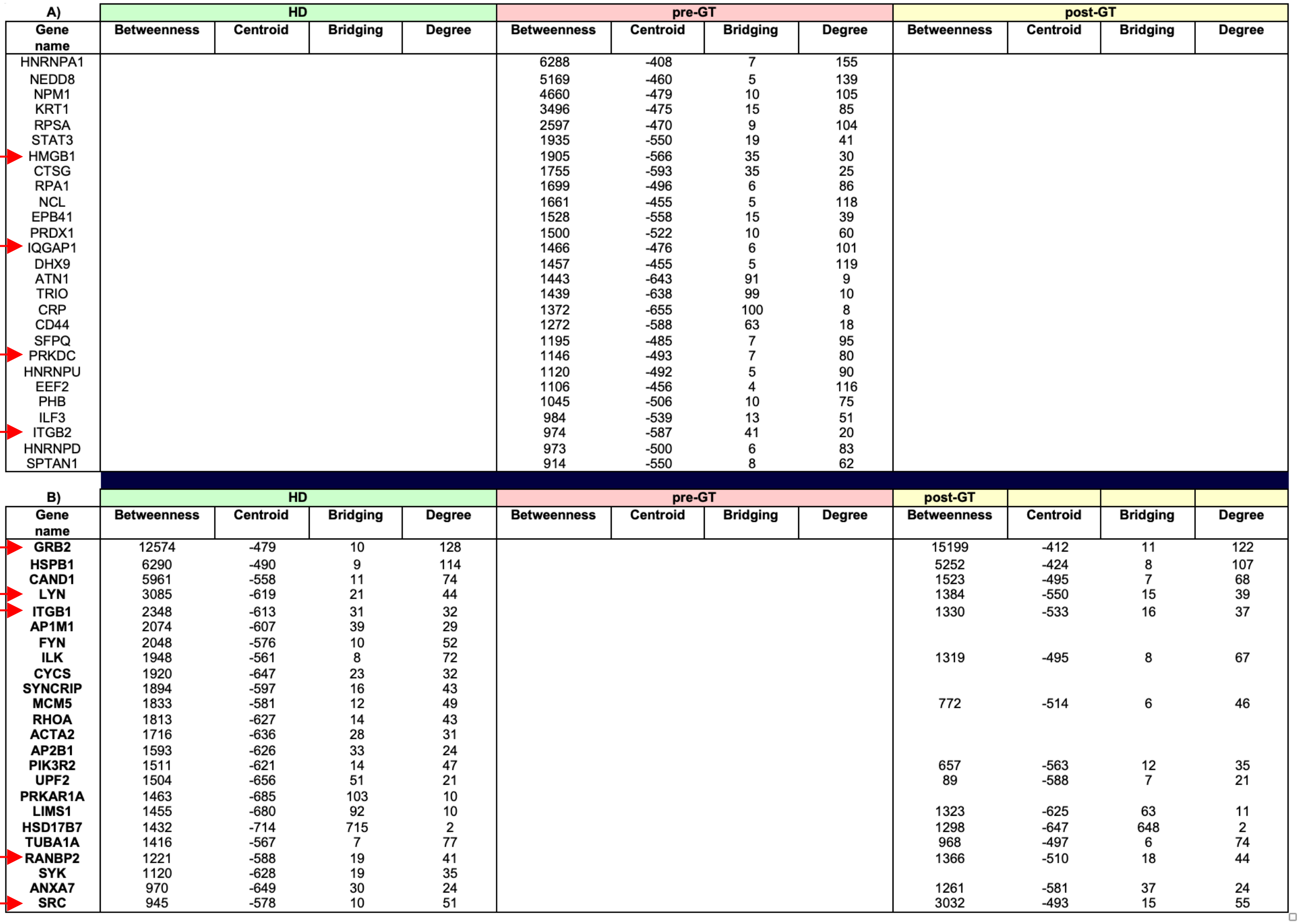
J Allergy Clin Immunol. 2019 Sep;144(3):825-838.

Case1- Wiskott-Aldrich syndrome

PPI Network Hubs

HC | pre GT | post GT

WAS interactors



A)

Gene name	HD Betweenness	HD Centroid	HD Bridging	HD Degree	pre-GT Betweenness	pre-GT Centroid	pre-GT Bridging	pre-GT Degree	post-GT Betweenness	post-GT Centroid	post-GT Bridging	post-GT Degree
HNRNPA1					6288	-408	7	155				
NEDD8					5169	-460	5	139				
NPM1					4660	-479	10	105				
KRT1					3496	-475	15	85				
RPSA					2597	-470	9	104				
STAT3					1935	-550	19	41				
HMGB1					1905	-566	35	30				
CTSG					1755	-593	35	25				
RPA1					1699	-496	6	86				
NCL					1661	-455	5	118				
EPB41					1528	-558	15	39				
PRDX1					1500	-522	10	60				
IQGAP1					1466	-476	6	101				
DHX9					1457	-455	5	119				
ATN1					1443	-643	91	9				
TRIO					1439	-638	99	10				
CRP					1372	-655	100	8				
CD44					1272	-588	63	18				
SFPQ					1195	-485	7	95				
PRKDC					1146	-493	7	80				
HNRNPU					1120	-492	5	90				
EEF2					1106	-456	4	116				
PHB					1045	-506	10	75				
ILF3					984	-539	13	51				
ITGB2					974	-587	41	20				
HNRNPD					973	-500	6	83				
SPTAN1					914	-550	8	62				

B)

Gene name	HD Betweenness	HD Centroid	HD Bridging	HD Degree	pre-GT Betweenness	pre-GT Centroid	pre-GT Bridging	pre-GT Degree	post-GT Betweenness	post-GT Centroid	post-GT Bridging	post-GT Degree
GRB2	12574	-479	10	128					15199	-412	11	122
HSPB1	6290	-490	9	114					5252	-424	8	107
CAND1	5961	-558	11	74					1523	-495	7	68
LYN	3085	-619	21	44					1384	-550	15	39
ITGB1	2348	-613	31	32					1330	-533	16	37
AP1M1	2074	-607	39	29								
FYN	2048	-576	10	52								
ILK	1948	-561	8	72					1319	-495	8	67
CYCS	1920	-647	23	32								
SYNCRIP	1894	-597	16	43								
MCM5	1833	-581	12	49					772	-514	6	46
RHOA	1813	-627	14	43								
ACTA2	1716	-636	28	31								
AP2B1	1593	-626	33	24								
PIK3R2	1511	-621	14	47					657	-563	12	35
UPF2	1504	-656	51	21					89	-588	7	21
PRKAR1A	1463	-685	103	10								
LIMS1	1455	-680	92	10					1323	-625	63	11
HSD17B7	1432	-714	715	2					1298	-647	648	2
TUBA1A	1416	-567	7	77					968	-497	6	74
RANBP2	1221	-588	19	41					1366	-510	18	44
SYK	1120	-628	19	35								
ANXA7	970	-649	30	24					1261	-581	37	24
SRC	945	-578	10	51					3032	-493	15	55

EPTRI
European Paediatric Translational Research Infrastructure
ITALIAN NODE

J Allergy Clin Immunol. 2019 Sep;144(3):825-838.

Case2- Drug Repurposing

Are drugs (drug targets) that are closer
to the disease (disease proteins) more effective than distant drugs?

- Genes associated with a disease tend to cluster in the same network neighborhood called the **Disease Module.**
- Moreover, the impact of drugs is typically local, restricted to disease proteins within **two steps** in the interactome.

Nat Commun. 2016 Feb 1;7:10331.

nature COMMUNICATIONS

ARTICLE

Received 7 May 2015 | Accepted 29 Nov 2015 | Published 1 Feb 2016 DOI: 10.1038/ncomms10331 OPEN

Network-based *in silico* drug efficacy screening

Emre Guney[1,2], Jörg Menche[1,3], Marc Vidal[2,4] & Albert-László Barábasi[1,2,3,5]

AL. Barabasi

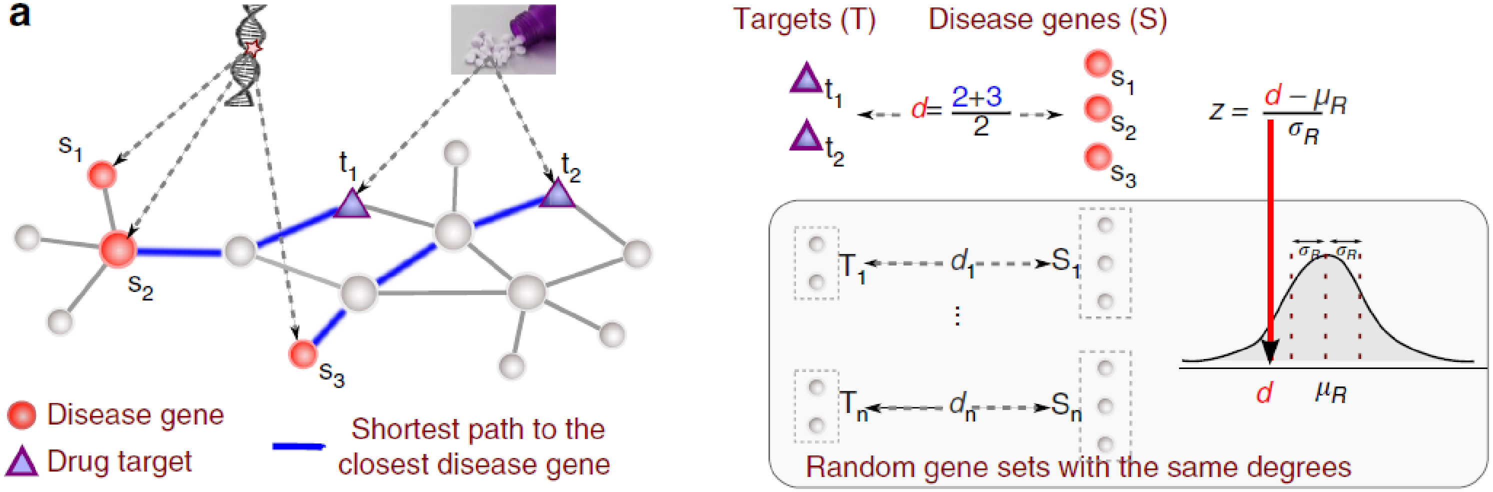
Case2- Drug Repurposing

Disease + Associated Genes (OMIM DB and GWAS catalog)
Drug + Target Association (DrugBANK)
Drug – Disease Indication (MEDI-HPS, Metab2MeSH)

238 drugs and 78 diseases
(402 combinations)

↓

Distance of proximity



Case2- Drug Repurposing

For 15% drug-disease association drug targets=disease proteins.
For 59%, drug targets are proximal to disease proteins.

Proximal drugs are more likely to be therapeutically beneficial than distant drugs that usually correspond to palliative treatments.

Proximal drugs tend to involve endocrine system and metabolic processes, whereas distant drugs are more enriched in anti-inflammatory and pain relief related categories.

Using proximity, it is possible repurpose existing drugs for other diseases (rare).

> Proc Natl Acad Sci U S A. 2021 May 11;118(19):e2025581118. doi: 10.1073/pnas.2025581118.

Network medicine framework for identifying drug-repurposing opportunities for COVID-19

Deisy Morselli Gysi [1] [2] [3], Ítalo do Valle [1] [2], Marinka Zitnik [4] [5], Asher Ameli [2] [6], Xiao Gan [1] [2] [3], Onur Varol [1] [2] [7], Susan Dina Ghiassian [6], J J Patten [8], Robert A Davey [8], Joseph Loscalzo [9], Albert-László Barabási [10] [2] [11]

Messages.....

PPI and co-expression networks are useful

- to visualize omics data
- to abstract the organization of complex systems
- to develop algorithms

The network structure is informative. Its analysis allows:

- the identification of hubs and modules correlating with phenotypes
- to shed light on disease/therapy mechanisms (diagnosis, prognosis, drug repurposing etc)

The extraction of relevant information and their significance may be improved by:

- higher number of molecular profiles
- accurate clinical stratification
- accurate and complete network models

ePTRI
EUROPEAN PAEDIATRIC TRANSLATIONAL RESEARCH INFRASTRUCTURE
ITALIAN NODE

Thank you.....

ePTRI
EUROPEAN PAEDIATRIC TRANSLATIONAL RESEARCH INFRASTRUCTURE
ITALIAN NODE

Institute for Biomedical Technologies
National Research Council (ITB-CNR)