NETWORK ANALYSIS OF PROTEIN-PROTEIN INTERACTIONS

Urszula Czeriwnska

U900 Computational Systems Biology for Cancer

OUTLINE

FROM EXTRATION OF KNOWLEDGE
TO INTELLIGENT LAYOUT

COMBING MUTIDIMENTIONAL DATA
AND NETWORK STRUCTURE

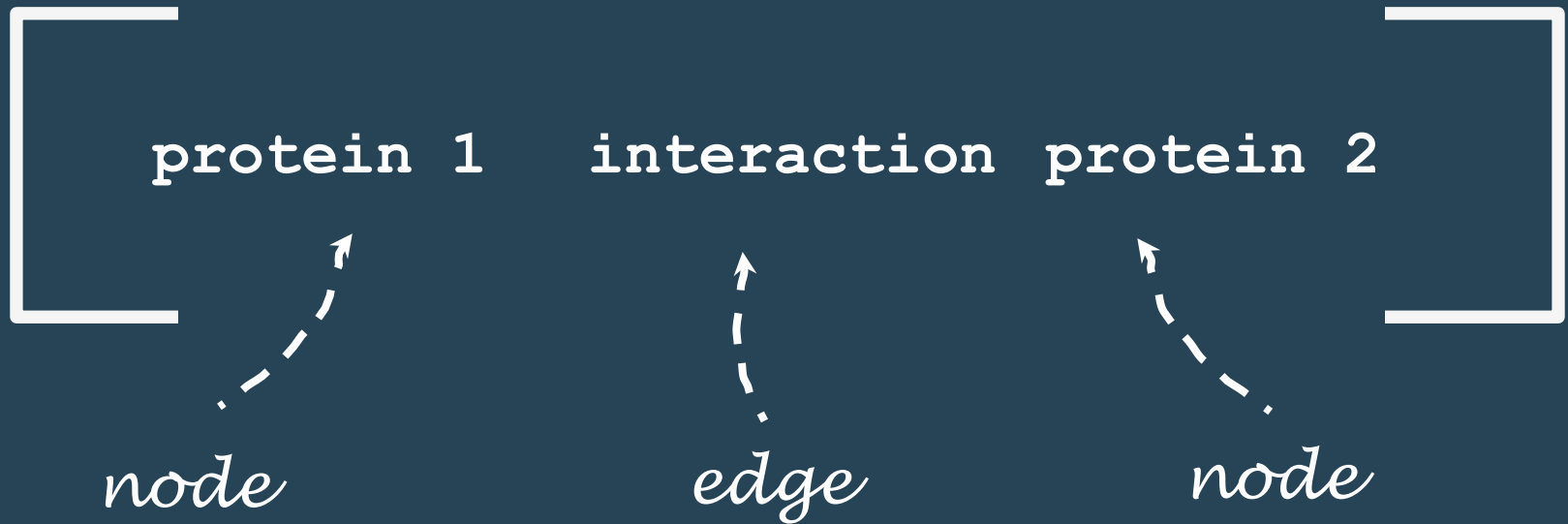DEDAL CYTOSCAPE 3.0 APP.

SUMMARY

FROM EXTRACTION OF KNOWLEDGE TO INTELLIGENT LAYOUT

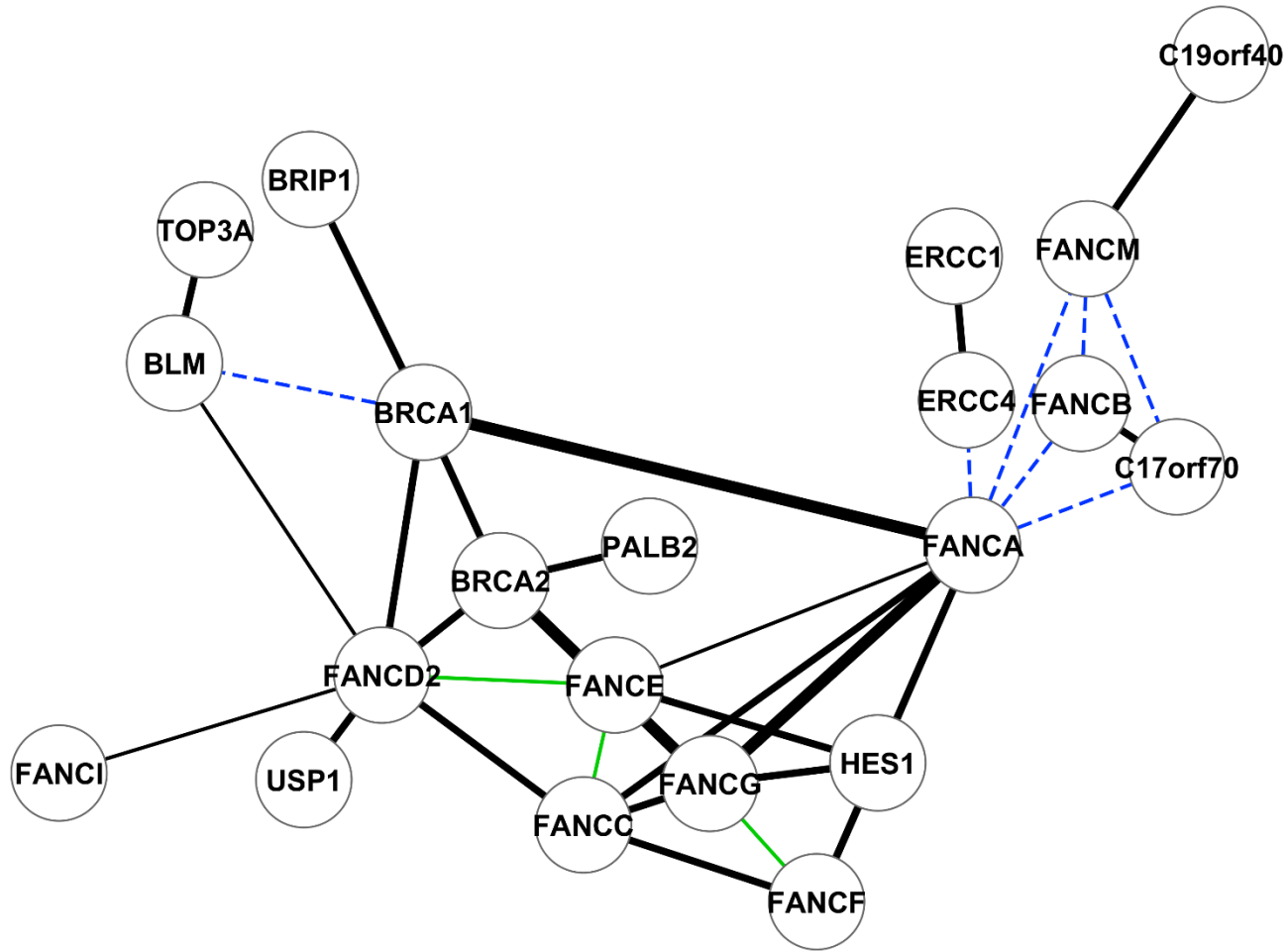# EXTRACTION OF KNOWLEDGE
IN BIOLOGY

: the creation of knowledge from structured and unstructured sources; the resulting knowledge needs to be in a machine-readable format;

molecular biology interactions -> networks

NETWORKS

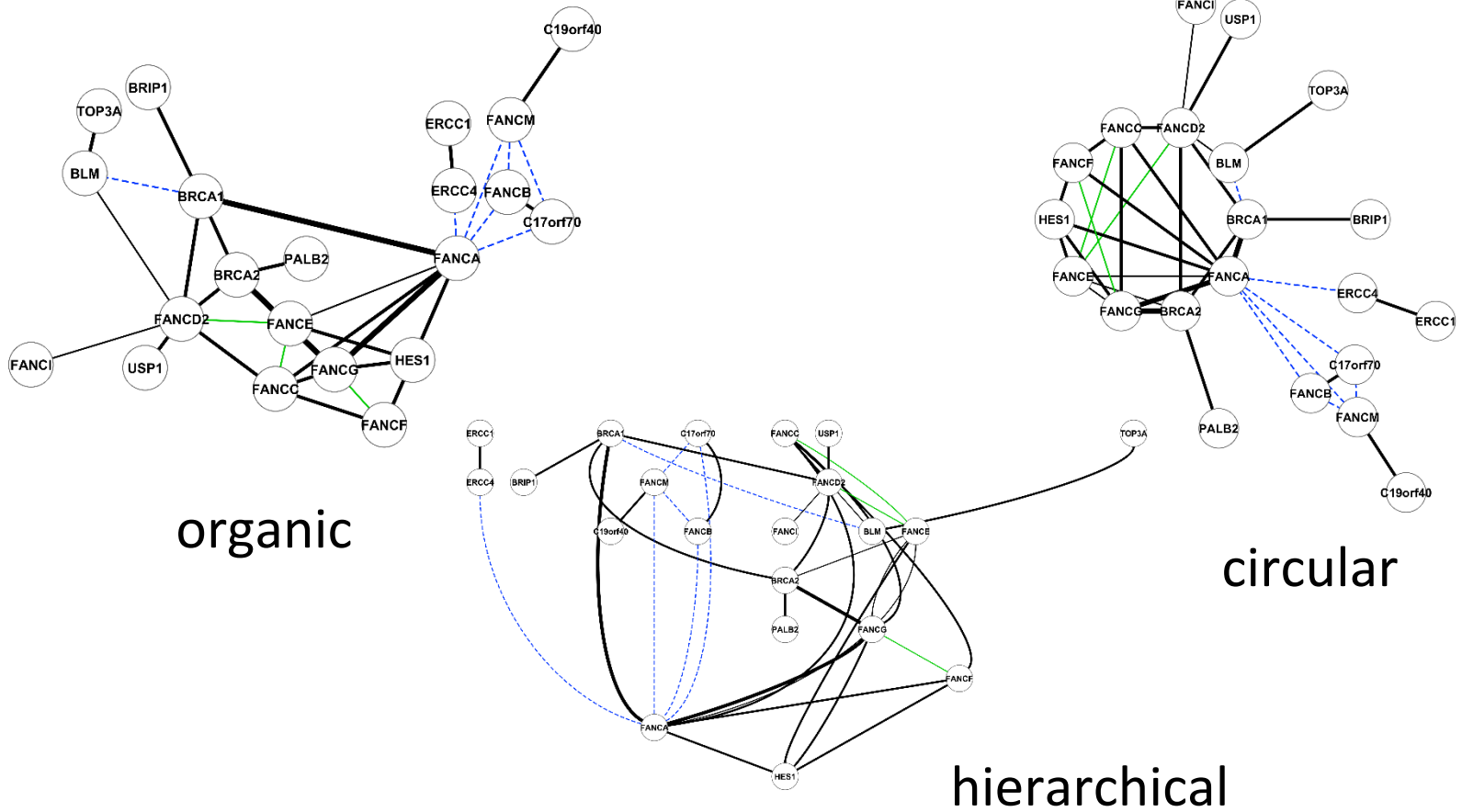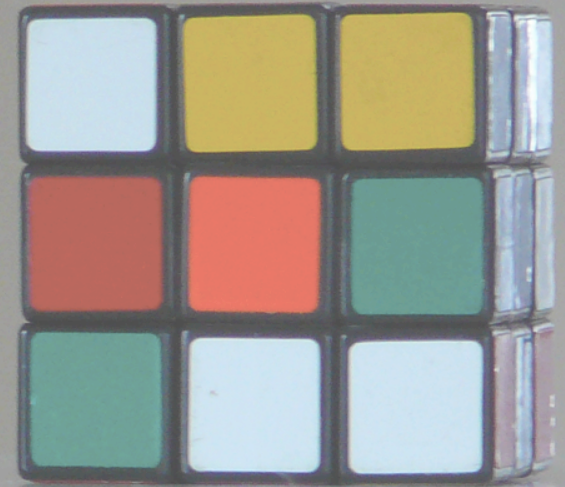$$[\quad \text{protein 1} \quad \text{interaction} \quad \text{protein 2} \quad]$$

node   *edge*   *node*

NETWORKS

Moldovan and D'Andrea 2009
Peri *et al.*, 2004

LAYOUT

VARIATIONS IN THE LOCATIONS OF POINTS COULD BE USED TO STRESS IMPORTANT STRUCTURAL PATTERNS IN THE DATA

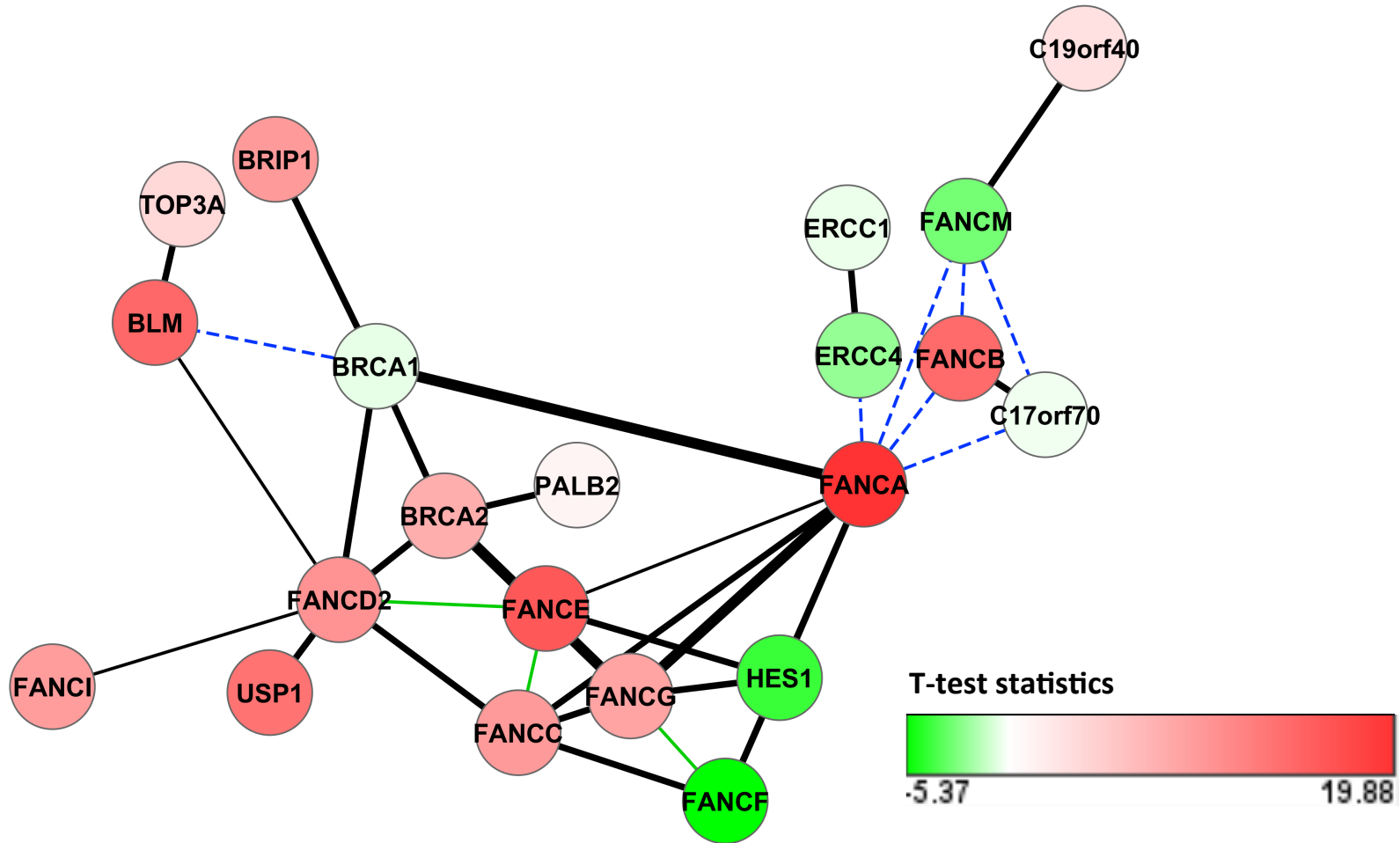*Freeman, 2000*

LAYOUTS

organic

hierarchical
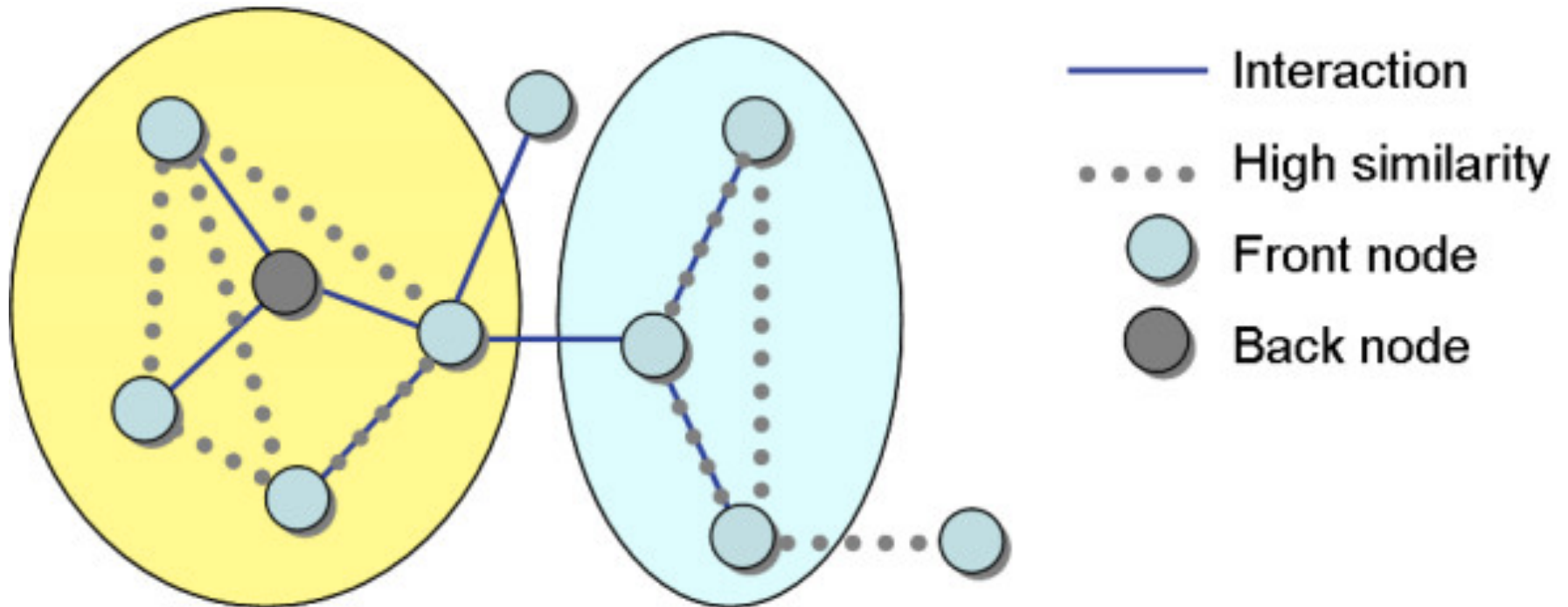
circular

COMBINING MULTIDIMENSIONAL DATA
AND NETWORK

THERE IS NOT A SINGLE LAYOUT

1 mapping data on the top of pre-defined biological network layout

2 identyfing subnetworks from a global network processing certain properities computed from the data

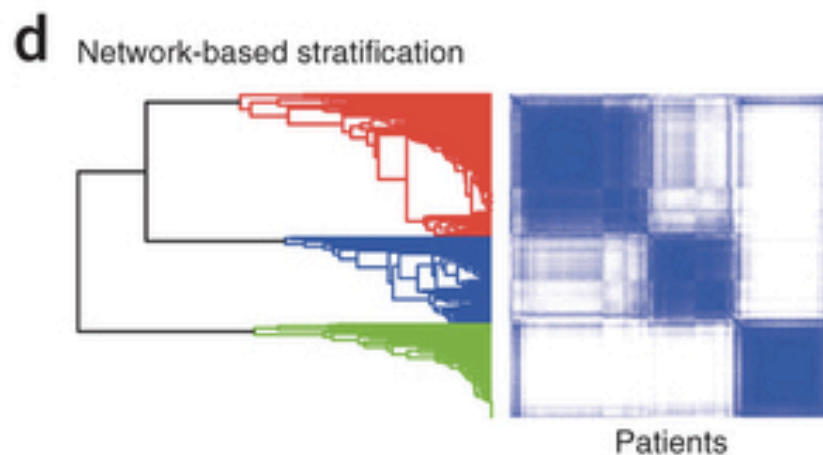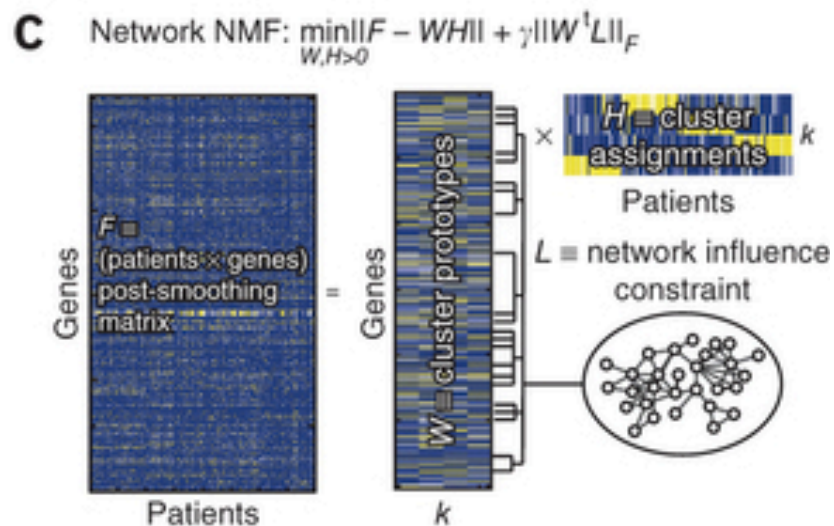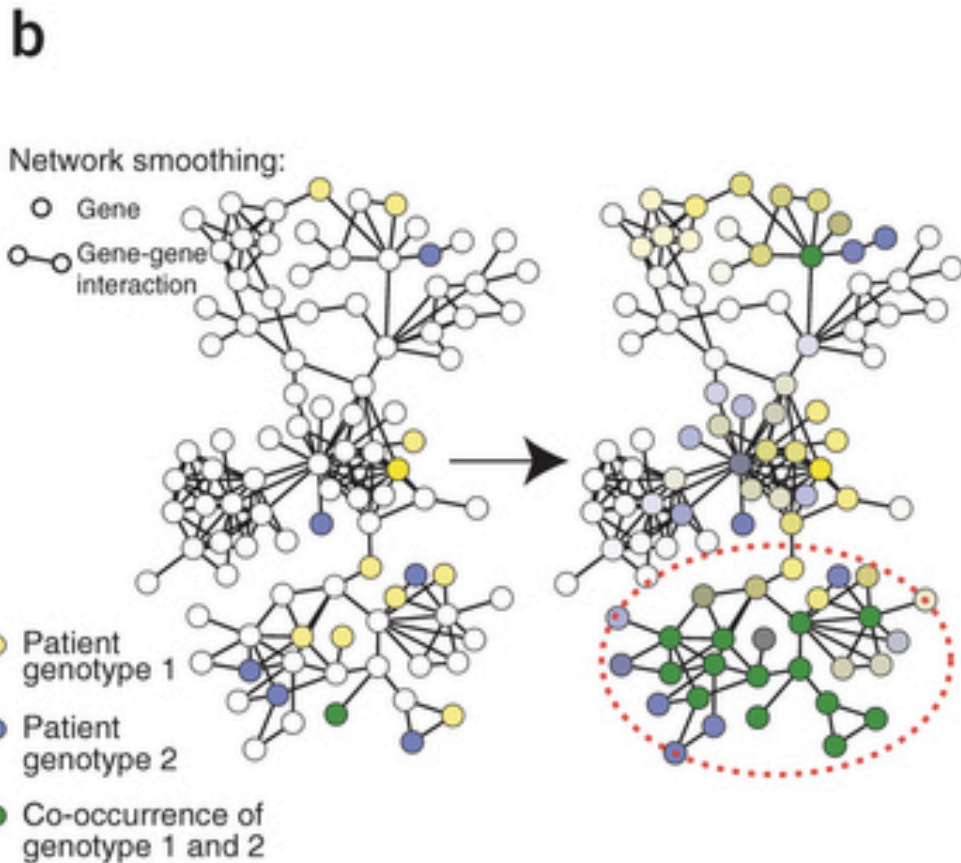*3* using biological network structure for pre-processing the high throughput data

1. MAPPING DATA ON THE TOP OF THE NETWORK

Moldovan and D'Andrea 2009
Peri *et al.*, 2004
TGCA, 2012

2. IDENTYFING SUBNETWORKS

Legend:
— Interaction
••••• High similarity
○ Front node
● Back node

3. USING BIOLOGICAL NETWORK STRUCTURE FOR PRE-PROCESSING THE HIGH THROUGHPUT DATA
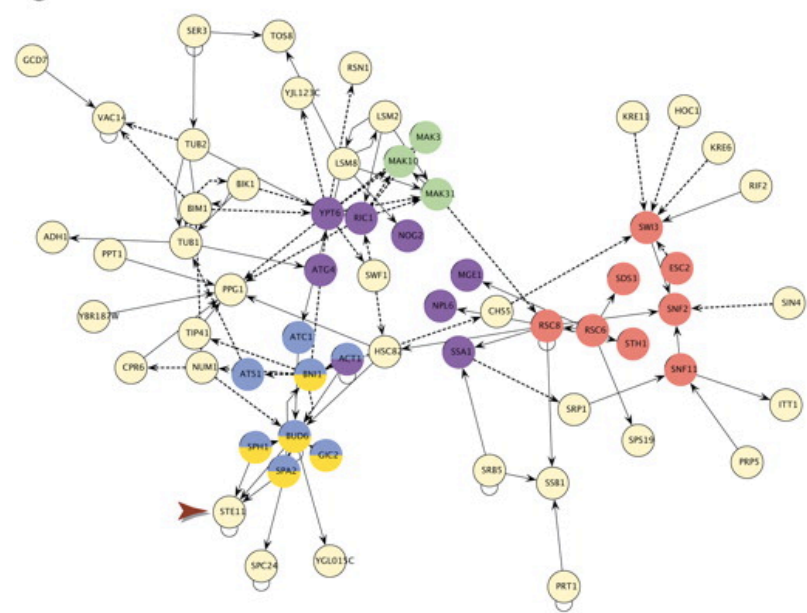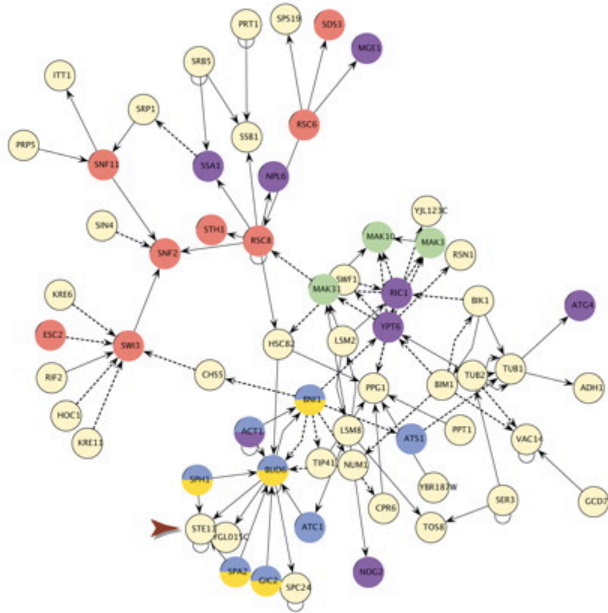
MULTIDIMENTIONAL SCALING

represent (dis)similarties as distances

dimension reduction

i.e. Principal Components Analysis

EXAMPLE: GOlorize Cytoscape App.

Garcia *et al*, 2007

DeDaL Cytoscape 3.0 App

DATA DRIVEN NETWORK LAYOUTS & MORPHING

OPEN SOURCE
CREATED AT THE INSTITUTE OF SYSTEMS BIOLOGY IN SEATTLE IN 2002
INPUT: FILE, NETWORK, DATABASE IMPORT
RICH VISUAL STYLES
ANALYSIS: PLUGINS AKA APPS

Layout

💬 Search for posts about layout →

❓ Ask a question about layout →

Categories

- **All Apps**
- network generation
- data visualization
- online data import
- graph analysis
- network analysis
- integrated analysis
- clustering
- utility
- enrichment analysis
- data integration
- systems biology
- **layout**
- ontology analysis
- visualization
- pathway database
- local data import
- network comparison
- core app
- scripting
- interaction database

Sort by [name ▲] [downloads] [votes] [newest release]

Adj Exporter `3.0+`
Cytoscape app that enables AdjacencyMatrixExport format

AllegroLayout `3.0+`
OpenCL-accelerated Interactive Force-Directed Layouts

Cy3D `3.0+`
Simple 3D Network Renderer App

CytoGEDEVO `3.0+`
Pairwise global alignment of PPI or other networks

DeDaL `3.0+`
Data-driven layout (PCA...), network alignement, network

DynNetwork `3.0+`
Visualize dynamic networks in Cytoscape 3.0

EntOptLayout `3.0+`
Relative entropy optimization based layout plugin

GOlorize `3.0+`
Uses Gene Ontology (GO) categories to direct the network

All Apps

Categories

network generation

data visualization

online data import

graph analysis

network analysis

integrated analysis

clustering

utility

enrichment analysis

data integration

systems biology

layout

ontology analysis

visualization

pathway database

local data import

network comparison

core app

scripting

interaction database

Layout

💬 Search for posts about layout →

❓ Ask a question about layout →

Sort by   name ▲   downloads   votes   newest release

Adj Exporter 3.0+
Cytoscape app that enables AdjacencyMatrixExport format

AllegroLayout 3.0+
OpenCL-accelerated Interactive Force-Directed Layouts

Cy3D 3.0+
Simple 3D Network Renderer App

CytoGEDEVO 3.0+
Pairwise global alignment of PPI or other networks

DeDaL 3.0+
Data-driven layout (PCA...), network alignement, network

DynNetwork 3.0+
Visualize dynamic networks in Cytoscape 3.0

EntOptLayout 3.0+
Relative entropy optimization based layout plugin

GOlorize 3.0+
Uses Gene Ontology (GO) categories to direct the network

# CYTOSCAPE: ORGANIC LAYOUT WITH MAPPED EXPRESSION DATA

T-test statistics

-5.37   19.88

Moldovan and D'Andrea 2009
Peri *et al.*, 2004
TGCA, 2012
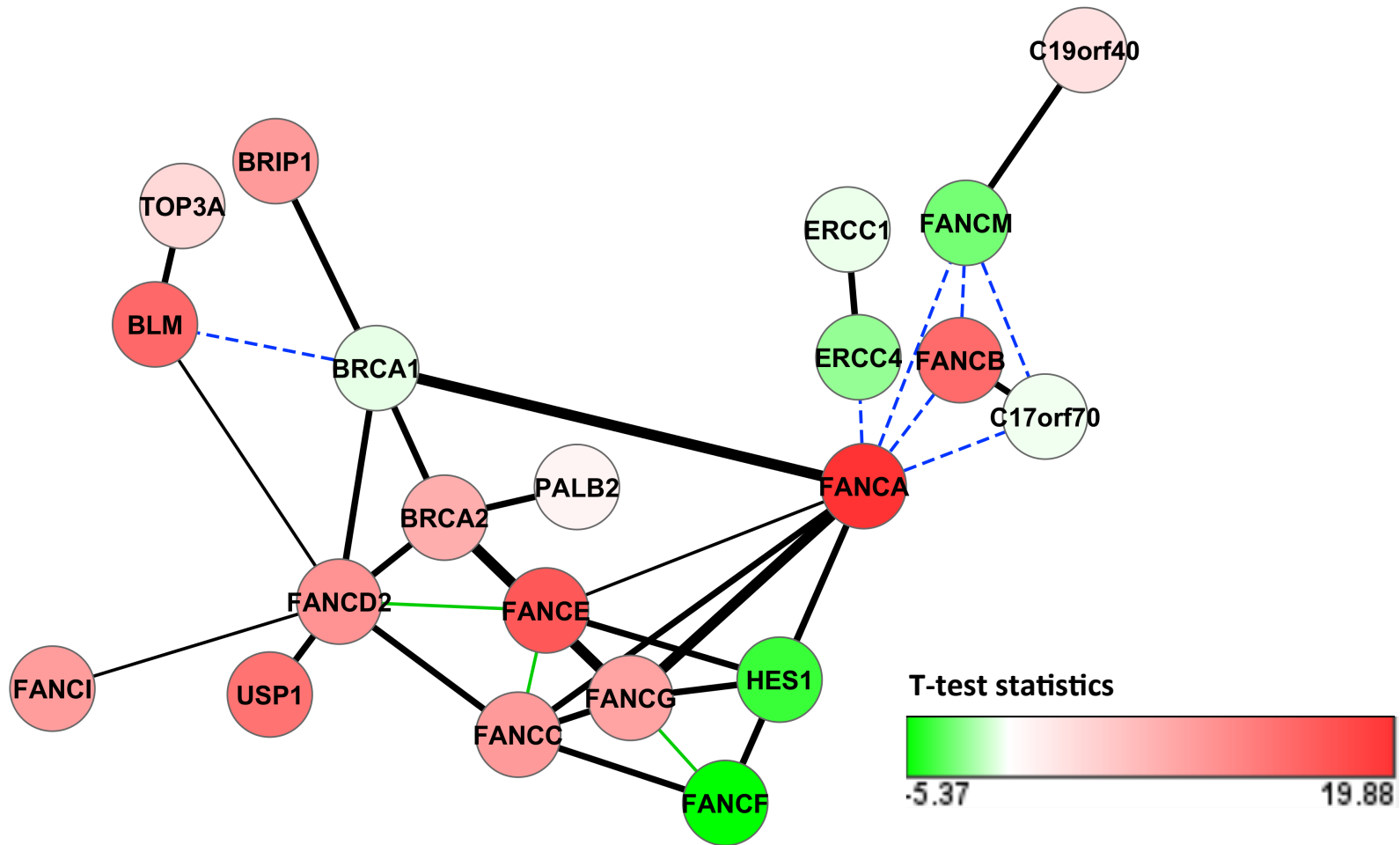
DEDAL: PRINCIPAL COMPONENT ANALYSIS DRIVEN LAYOUT
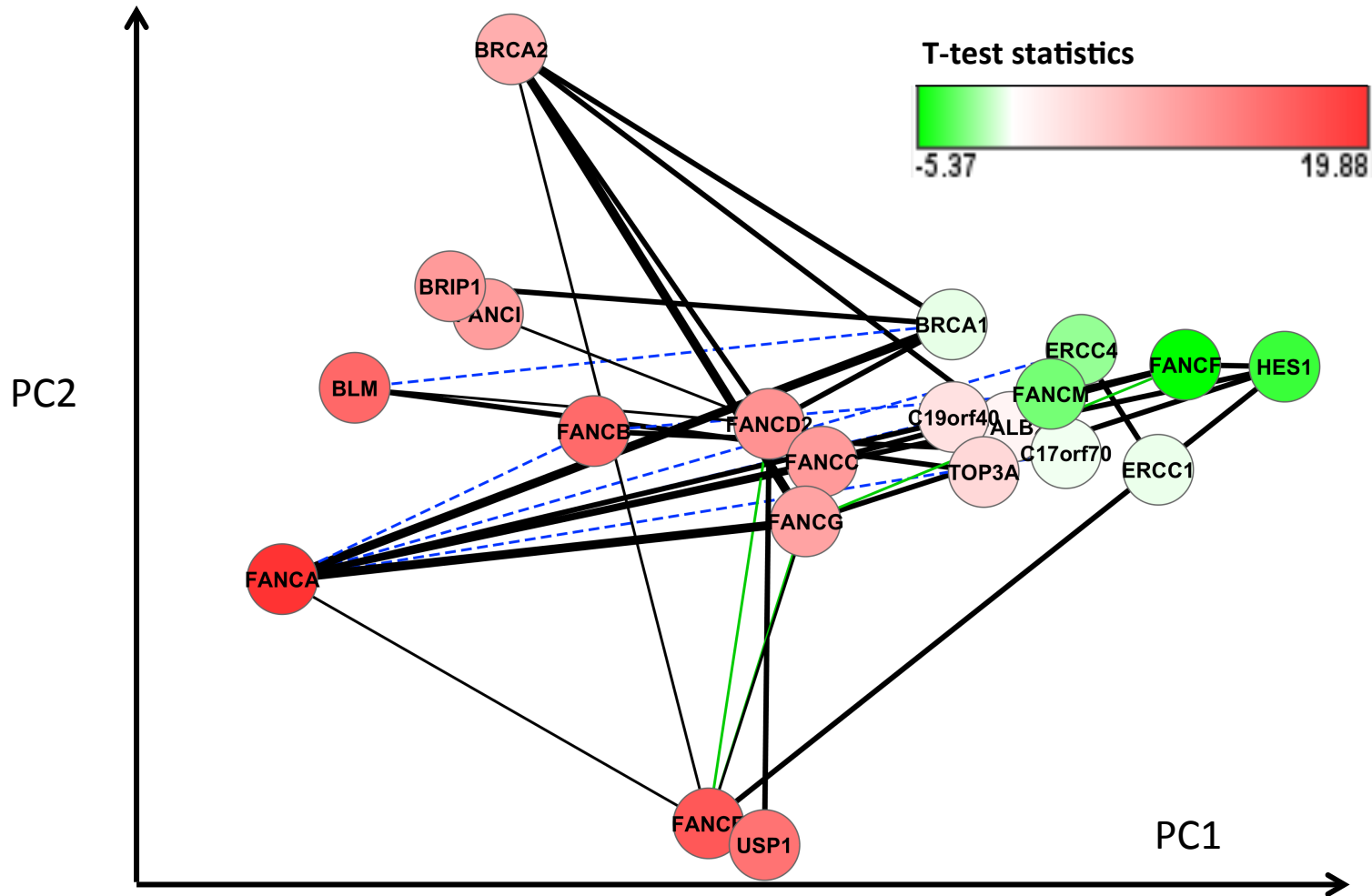
Moldovan and D'Andrea 2009
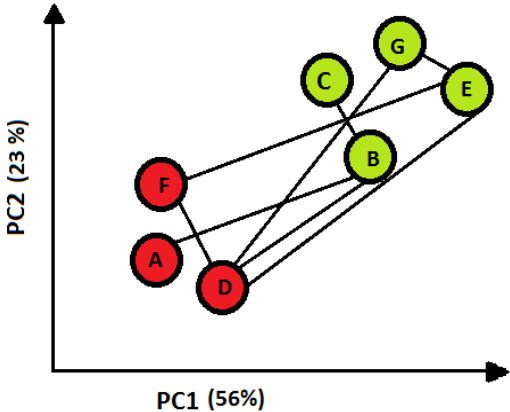Peri *et al.*, 2004
TGCA, 2012

MORPHING DATA–DRIVEN AND STRUCTURE–BASED LAYOUT

Pure network structure based layout
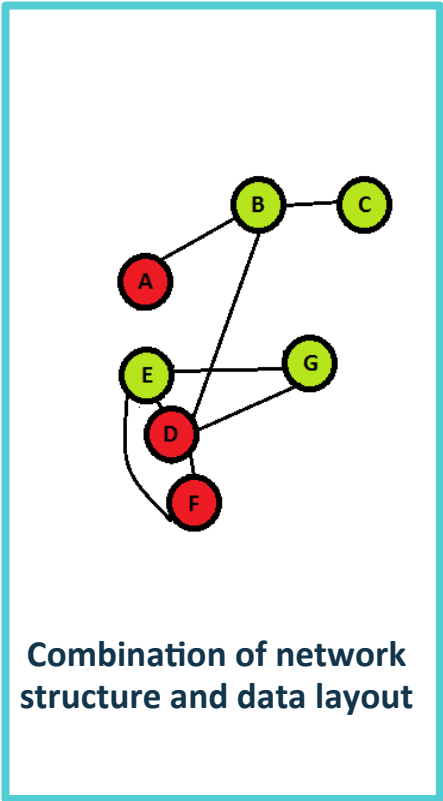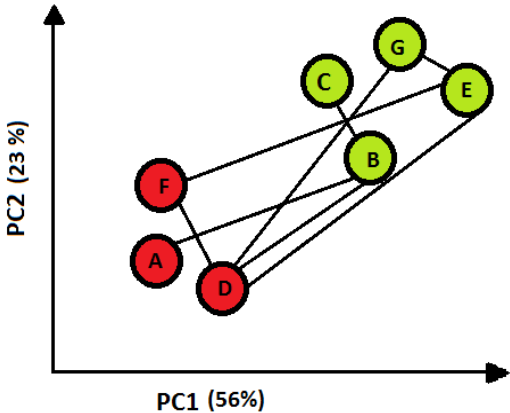
**Combination of network structure and data layout**

Purely Data Driven Layout

PCA or Elmap

MORPHING DATA-DRIVEN AND STRUCTURE-BASED LAYOUT

Pure network structure based layout

Combination of network structure and data layout

Purely Data Driven Layout

PCA or Elmap

MORPHING DATA-DRIVEN AND STRUCTURE-BASED LAYOUT

MORPHING DATA-DRIVEN AND STRUCTURE-BASED LAYOUT
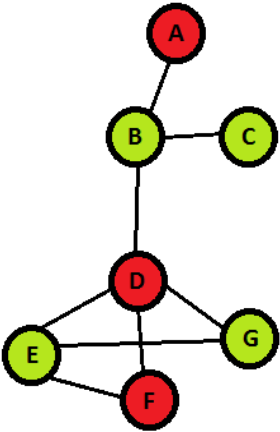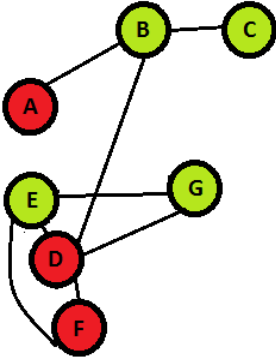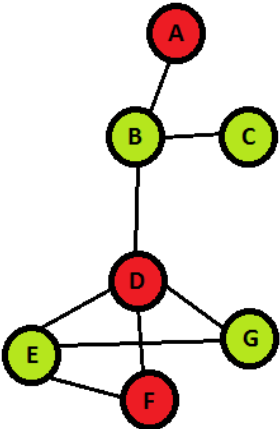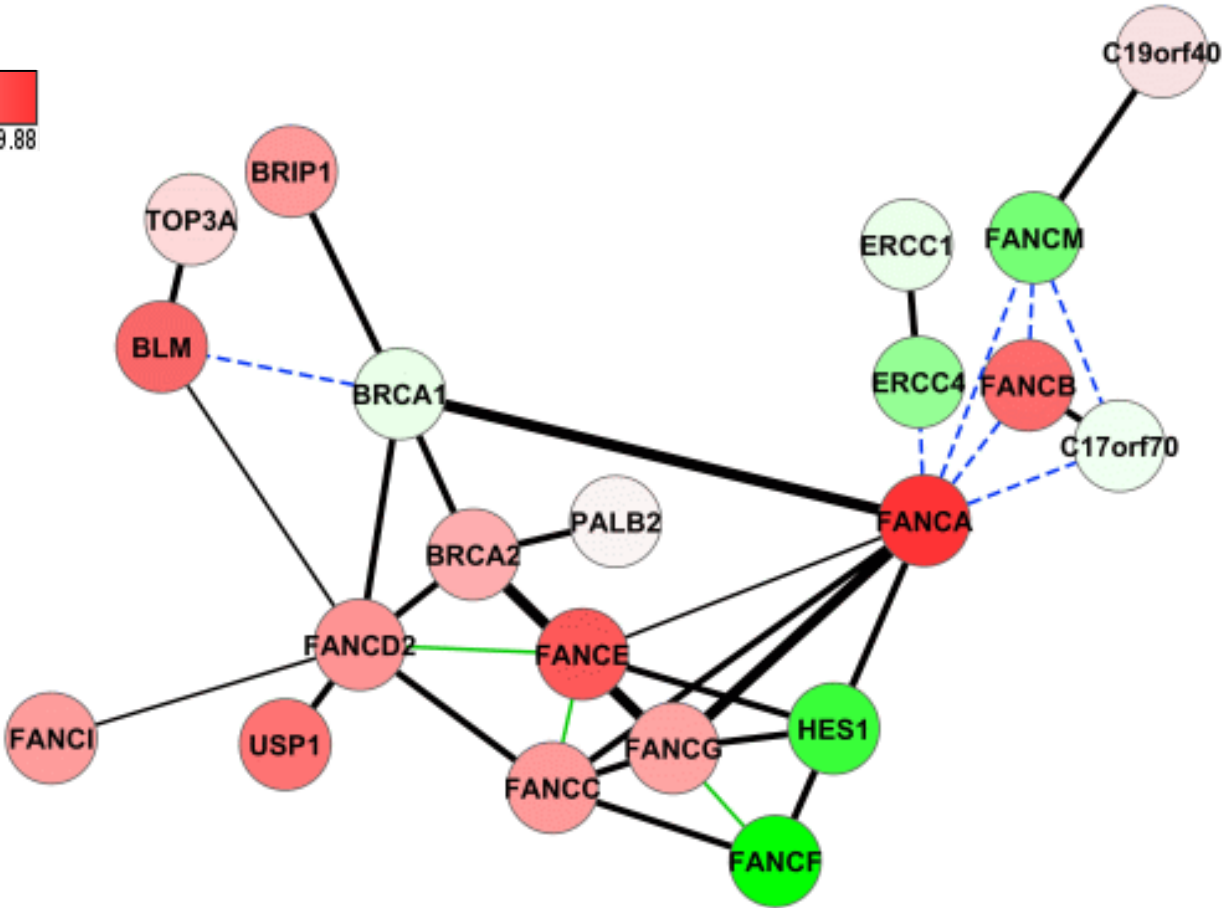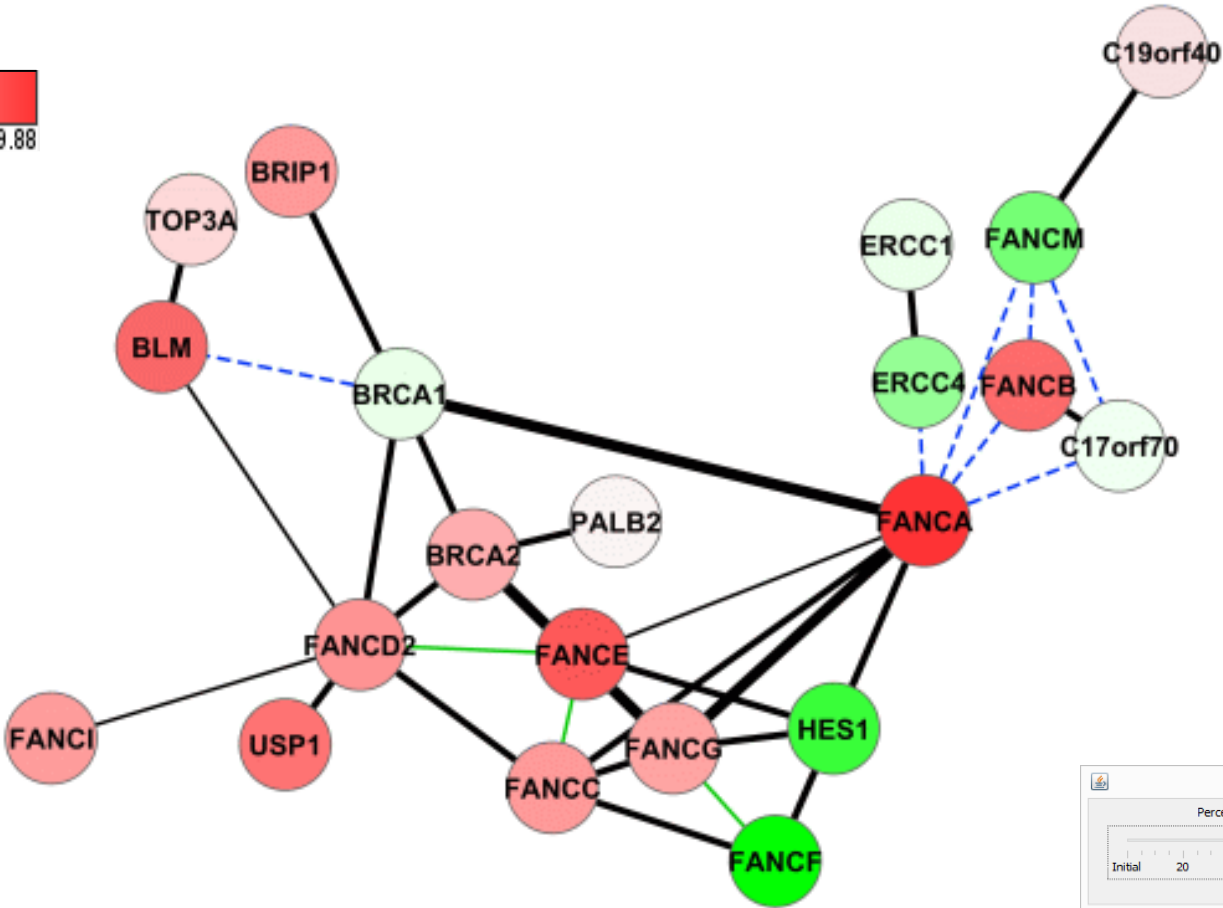
MORE ADVANCED FEATURES
OF DEDAL

ADVANCED FEATURES OF DEDAL

Pre-processing of the data:

- o Smoothing
- o Double centering
- o Quality check

Data-driven layout: PCA or nPMs
Morphing

Post-processing of the layout:

- o Alignment
- o Overlapping
- o Missing values
- o Outliers

NONLINEAR PRINCIPAL MANIFOLDS

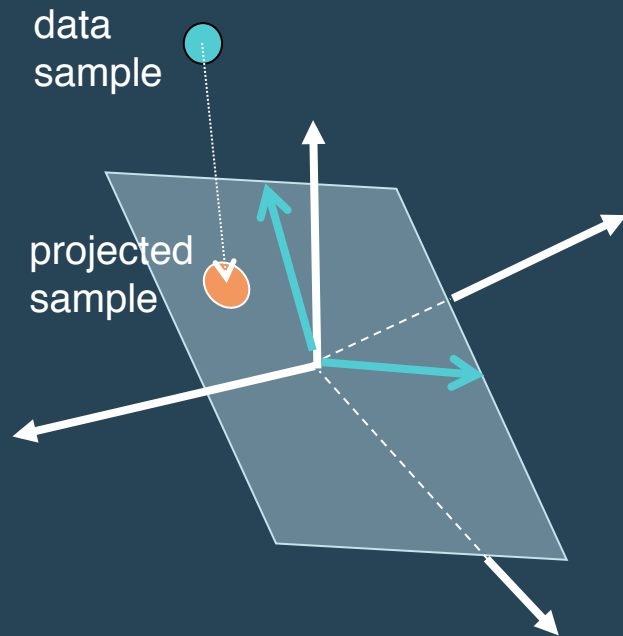**Elastic map algorithms**
Principal manifolds approximation

a) 3D PCA linear manifold.

b) ELMap2D

c) PCA2D

Linear PCA vs nonlinear Principal Manifolds for visualisation of breast cancer microarray data

Basal
LumA
LumB
ERBB2
Normal
Unclassified
Group A
Group B
ER +
ER -

a)

b) ELMAP2D

c) PCA2D

Gorban A.N., Zinovyev A. 2010

NETWORK SMOOTHING

data sample

projected sample
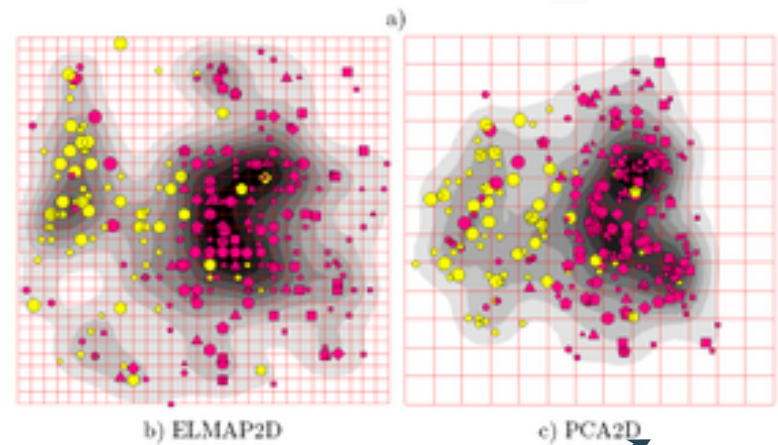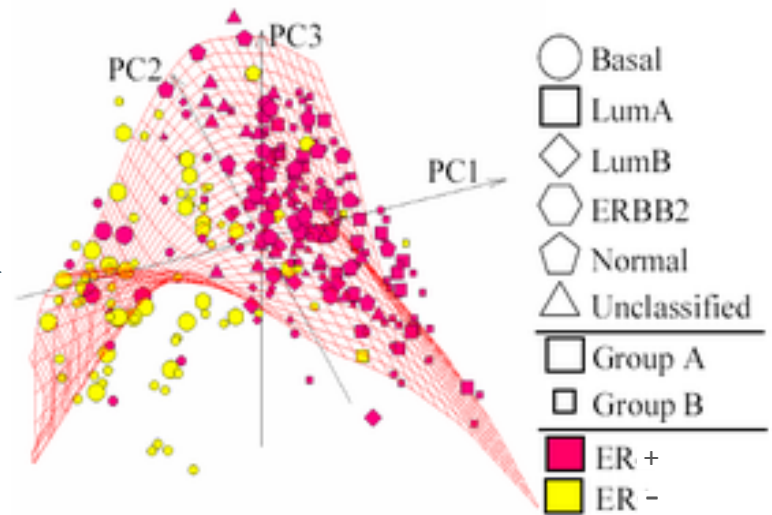
→ initial space

→ subspace of functions smooth on a gene network

basis vectors are eigenvectors of the graph Laplacian

NETWORK SMOOTHING
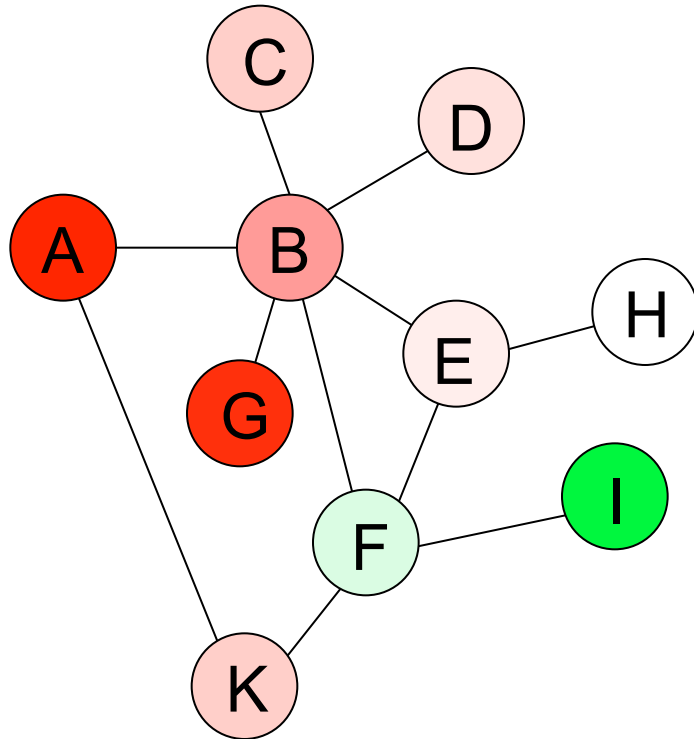
Smooth distribution,
balanced dosage
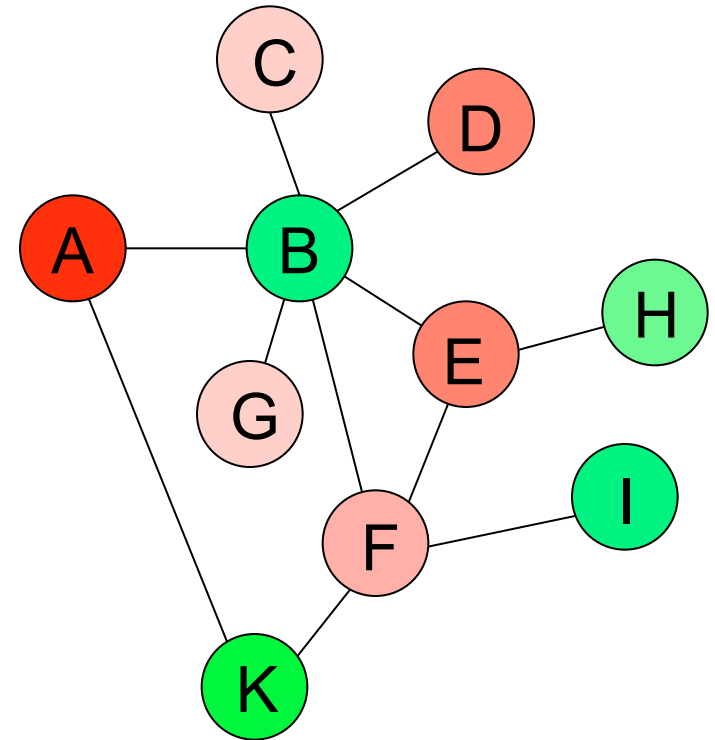
Non-smooth distribution,
unbalanced dosage

BIG NETWORK

1047 nodes
1986 edges
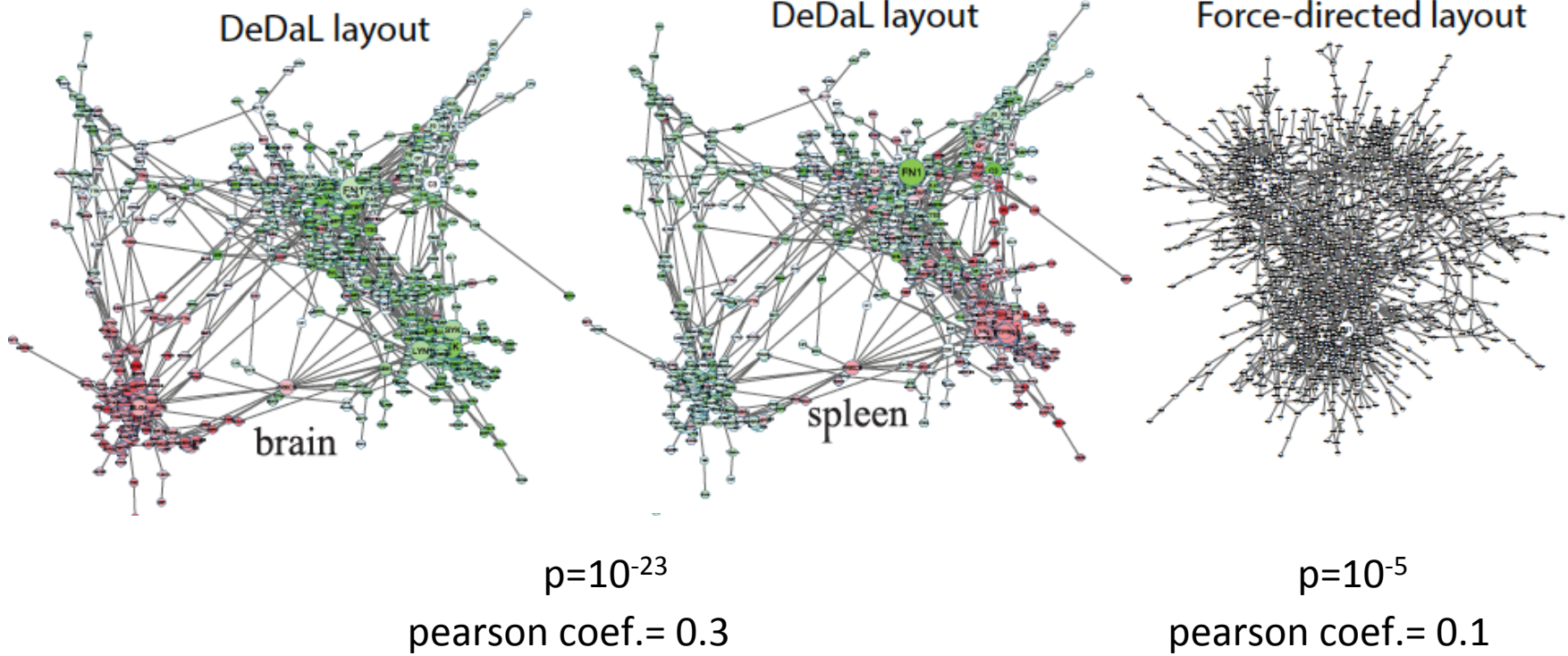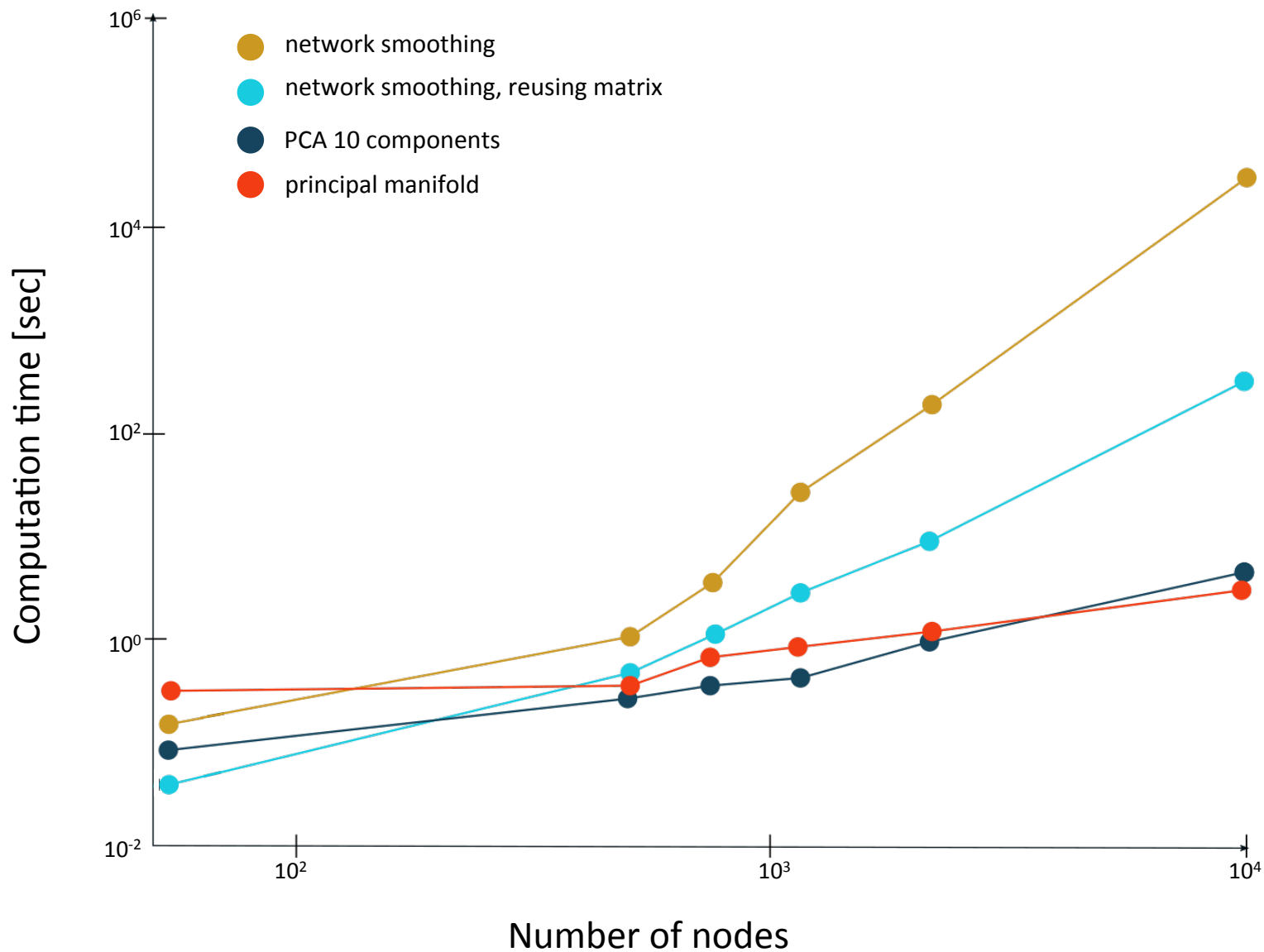
p=10^{-23}
pearson coef.= 0.3

p=10^{-5}
pearson coef.= 0.1

Czerwinska *et al.*, 2015

SCALABILITY

Legend:
- network smoothing
- network smoothing, reusing matrix
- PCA 10 components
- principal manifold

Y-axis: Computation time [sec]

X-axis: Number of nodes

Czerwinska *et al.*, 2015

TUTORIAL

bioinfo-out.curie.fr/projects/dedal/

TUTORIAL

bioinfo-out.curie.fr/projects/dedal/

SOFTWARE

DeDaL: Cytoscape 3.0 app for producing and morphing data-driven and structure-driven network layouts

Urszula Czerwinska[1,2,3], Laurence Calzone[1,2,3], Emmanuel Barillot[1,2,3] and Andrei Zinovyev[1,2,3]*

*Correspondence:
Andrei.Zinovyev@curie.fr
[1]Institut Curie, 26 rue d'Ulm, Paris, FR
Full list of author information is available at the end of the article

Abstract

Background: Visualization and analysis of molecular profiling data together with biological networks are able to provide new mechanistical insights into biological functions. Currently, high-throughput data are usually visualized on top of predefined network layouts which are not always adapted to a given data analysis task. We developed a Cytoscape app which allows to construct biological network layouts based on the data from molecular profiles imported as values of

SUMMARY

There is a need to combine networks and -omics data in biology

Network layout should be adapted to the analysis

DeDaL – Cytoscape App. performs
o different types of data-driven layouts
o morphing between strucure-based and data-driven layout
o pre-processing of data as double centering and network smoothing

THANK YOU

Urszula Czerwinska
urszula.czerwinska@curie.fr

@UlaLaParis