Systems Biology approaches for studying toxicity

Francesco Falciani

Centre for Computational Biology and Modelling

University of Liverpool

NanoMILE

Engineered nanomaterial mechanisms of interactions with living systems and the environment: a universal framework for safe nanotechnology

Coordinator: EugeniaValsami-Jones

- Project start: March 2013
- Duration: 4 years
- Number of partners: 26 (+ 2 US partners)
- Budget: 10M Euro

NanoMILE key objectives

- selection and development of an optimum MNM library to support objectives below;
- technically and computationally advanced integration of systems biology;
- balanced toxicological / ecotoxicological approaches;
- novel high throughput platforms for screening;
- feedback loops for development of safer by design MNMs;
- ***Robust framework for classification of MNMs according to their biological impacts***

UoL

BIOXM

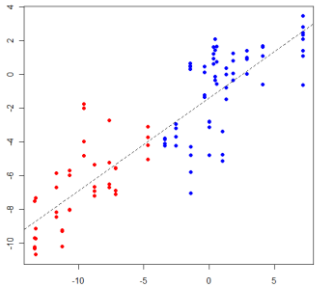
NPs

NPs Features

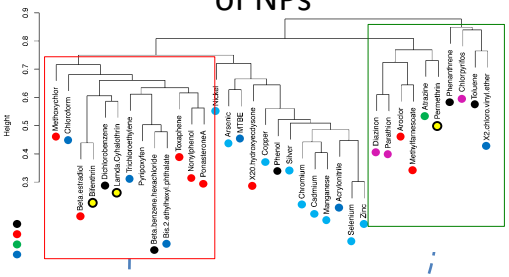
Omics Data

Physiology Toxicology

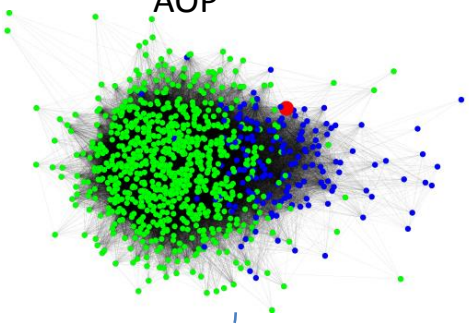
Multi-Level QSAR



Molecular Classification of NPs



AOP



NPs PCFs

Key representative NPs

AOP Modules

Identifying links between Molecular-Physiological Response and NPs PCFs

Time course Concentration response

Dynamical Models of NP Response

$$\dot{x}_i = \sum_{j=1}^{N} w_{ij} x_j + b_i x_i$$

$$E^{SQE} = \sum_{i=1}^{N} \sum_{t} \left(x_i^{measured} - x_i \right)^2$$

$$E^{Object} = \sum_{j} \sum_{t} \left(o_{ij} - w_{ij} \right)^2$$

Hypothesis and non-hypothesis driven research

- Hypothesis driven approaches focus on specific mechanisms and aim at elucidating adverse outcome pathways (AOPs) in detail.

- These are very effective but there should be sufficient evidence for supporting the original assumptions. Often this is not the case.

- Nanomaterials are a relatively novel technology and prior knowledge is scare. They are an ideal ground for the application of approaches that aim at “learning” the underlying structure of AOPs from observational data.

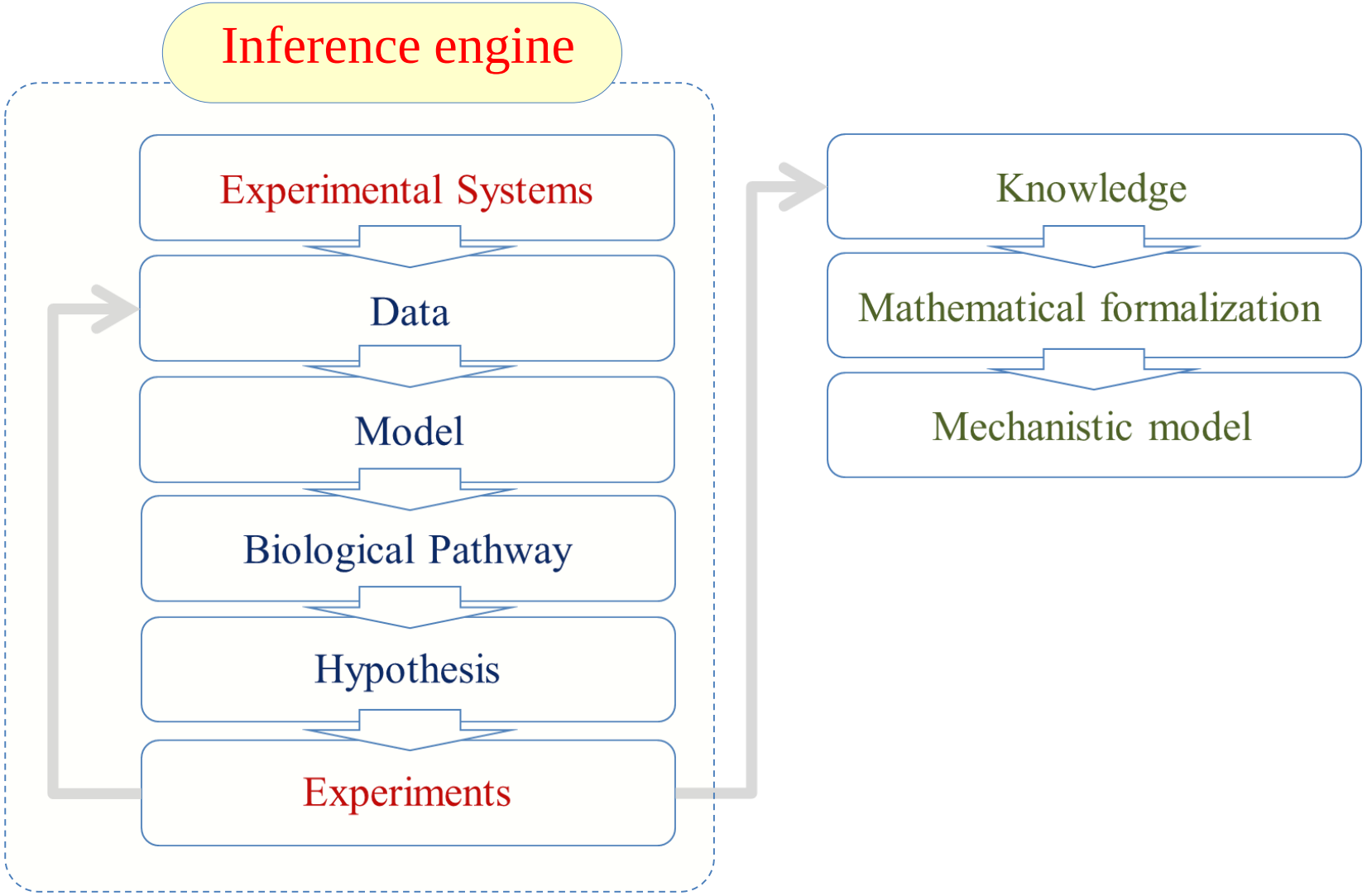
Measuring is not enough!

- Recent improvements in Functional Genomics technologies (e.g. transcriptomics, proteomics, metabolomics, physiomics) allow measuring the relative concentration of tens of thousands of cellular molecular components in single experiments.

- This generates large amounts of data but this is not sufficient to increase knowledge. On it's on this is just a large stamp collecting exercise.

- Advanced computational approaches are required to analyse and interpret large scale data.

Reverse engineering and mechanistic models



Virtual Physiological Human
network of excellence

The VPH aims to support the development of patient-specific computer models and their application in personalised and predictive healthcare.

The Virtual Physiological Human Network of Excellence aims to provide the necessary infrastructure

computational methodologies,
tools and databases

The Virtual Physiological Human
Network of Excellence
VPH NoE

Supporting sustainability and excellence
in European VPH research

Enable academic, clinical and industrial researchers to communicate, and to exchange data and technologies in a standardised way.

Synergy COPD

A **simulation environment** and a **decision-support system** aiming at enabling the deployment of systems medicine.

A knowledge base,

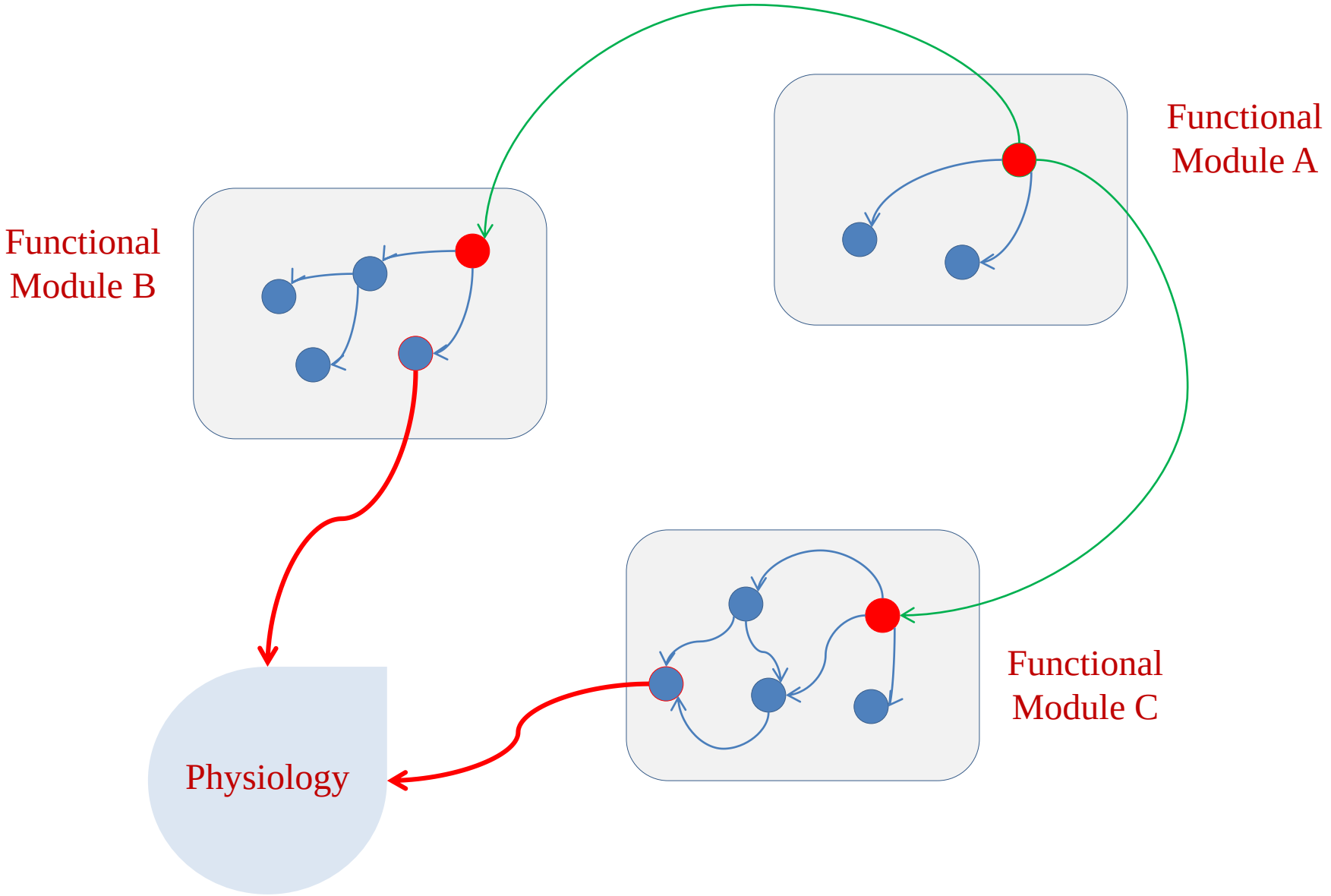
An inference engine,

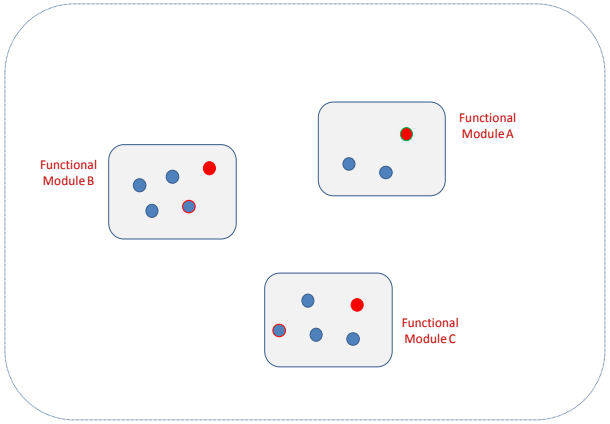
A simulation engine,

Two graphical visualisation environments.

The Synergy system focuses on patients with **COPD**, which is a major public health problem and a complex, heterogeneous and multi-component disease.

Conceptual structure of a molecular network

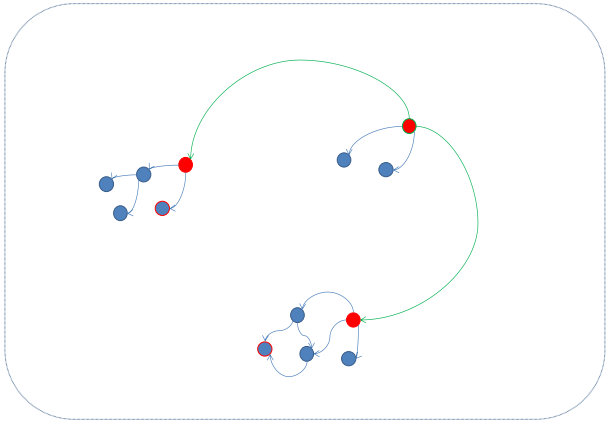




Identify the key Functional modules

Identify groups of genes which define a sub-graph within a gene regulatory network

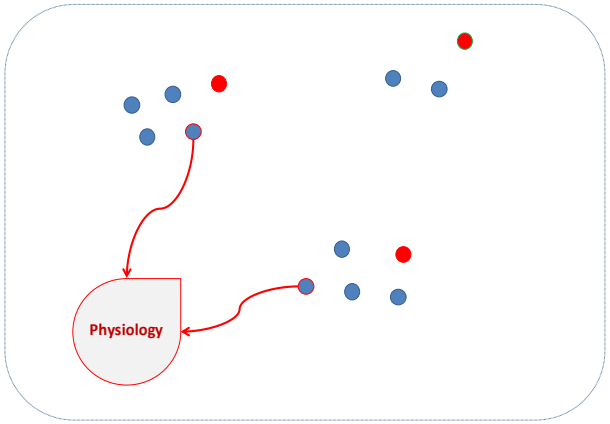
Challenge: Different definitions of functional module can lead to different results



Identify the structure of regulatory networks

Identify a network topology which is compatible with observed data (i.e. time course)

Challenge: Many possible topologies can explain the data

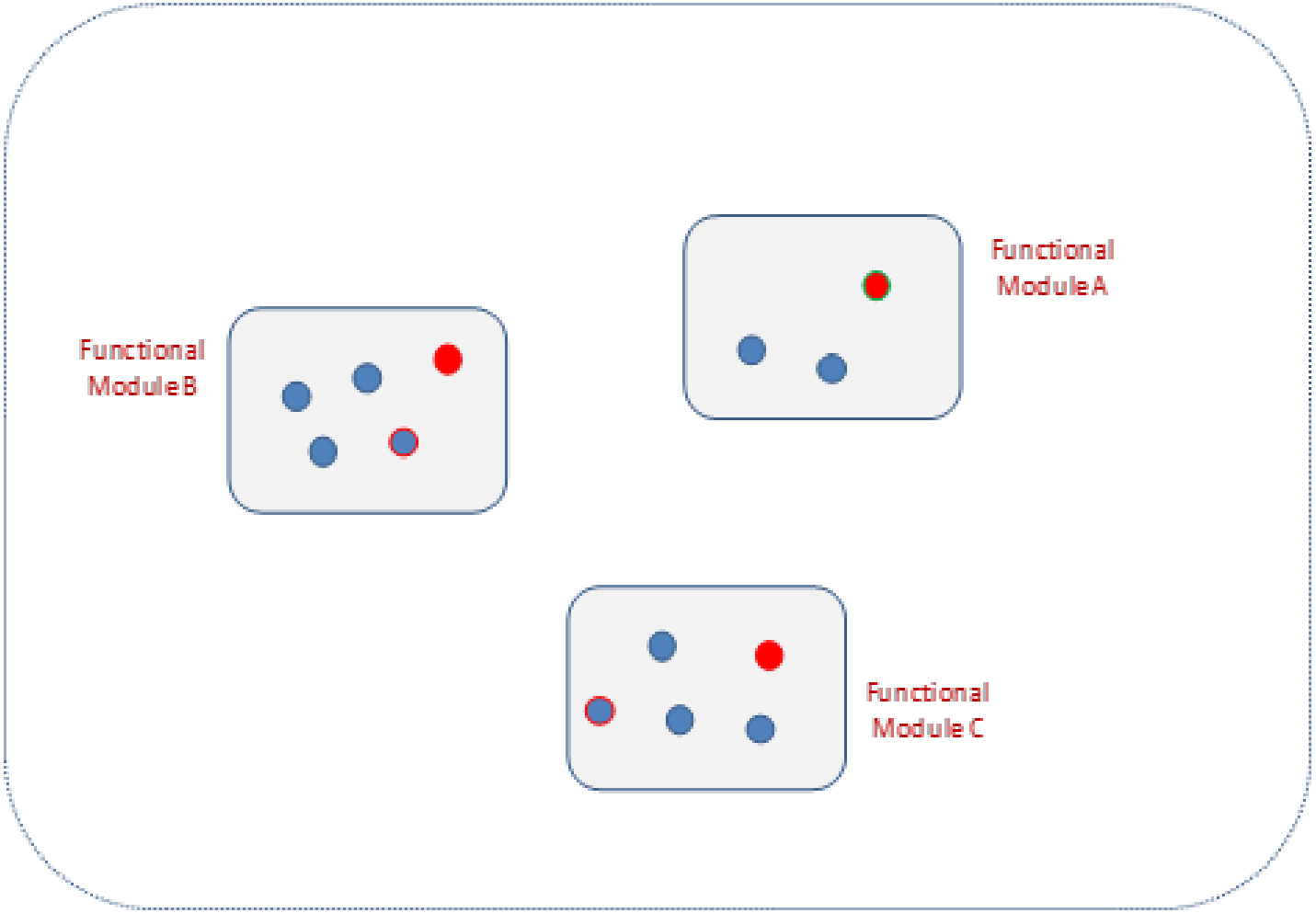


Identify molecular components controlling physiology

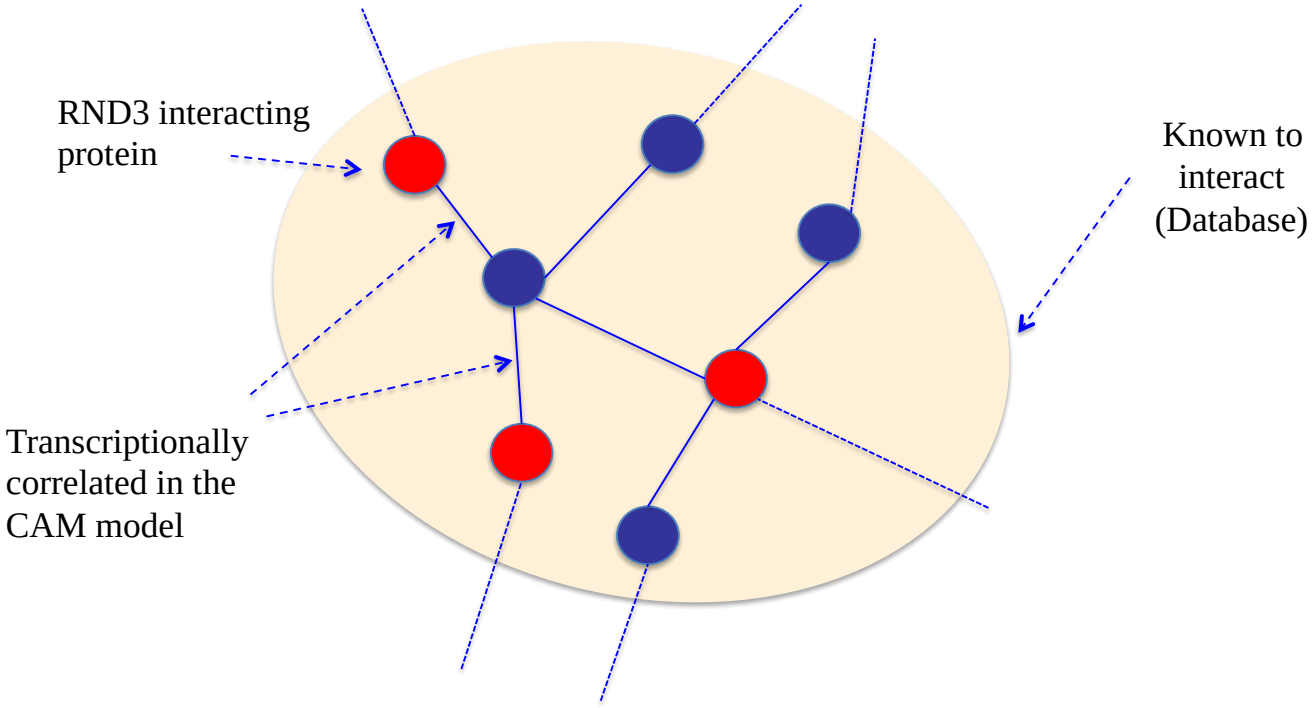
Identify the most important components whose combined activity is predictive of physiology outcome

Challenge: The methodology needs to be able to identify genes, proteins or metabolites that most directly contribute to the phenotype

Network Modularity



How can RND3 regulate cell cycle?

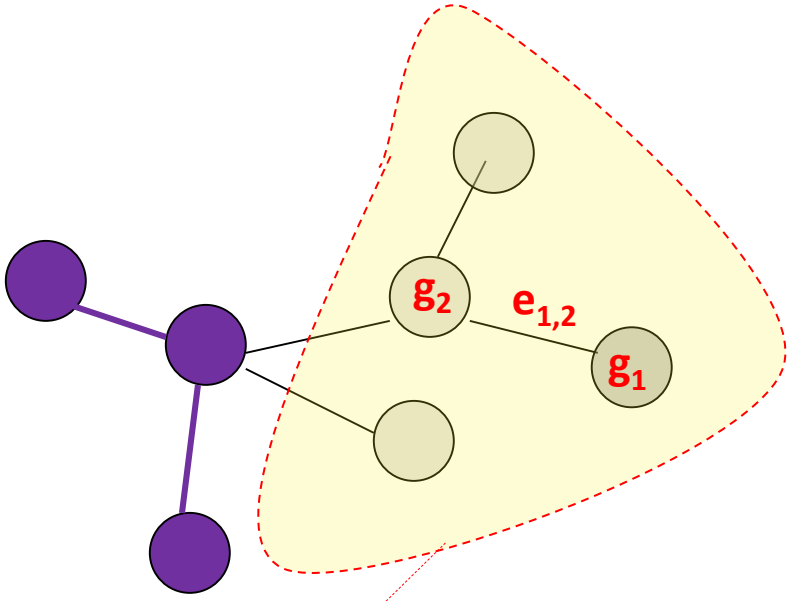


Step 1
RND3 immunoprecipitation

Step 2
MASS Spectrometry

Step 3:
Integration of the RND3 interactome with known PPI interaction databases and CAM expression profiling

A new modularization method



Nodes are scored for differential expression:

$$score_g = -\log(p)$$

Small p values-> large profits
Large p values -> small profits

Edges are scored for co-expression:

$$score_e = -w\log(I(g_i, g_j))$$

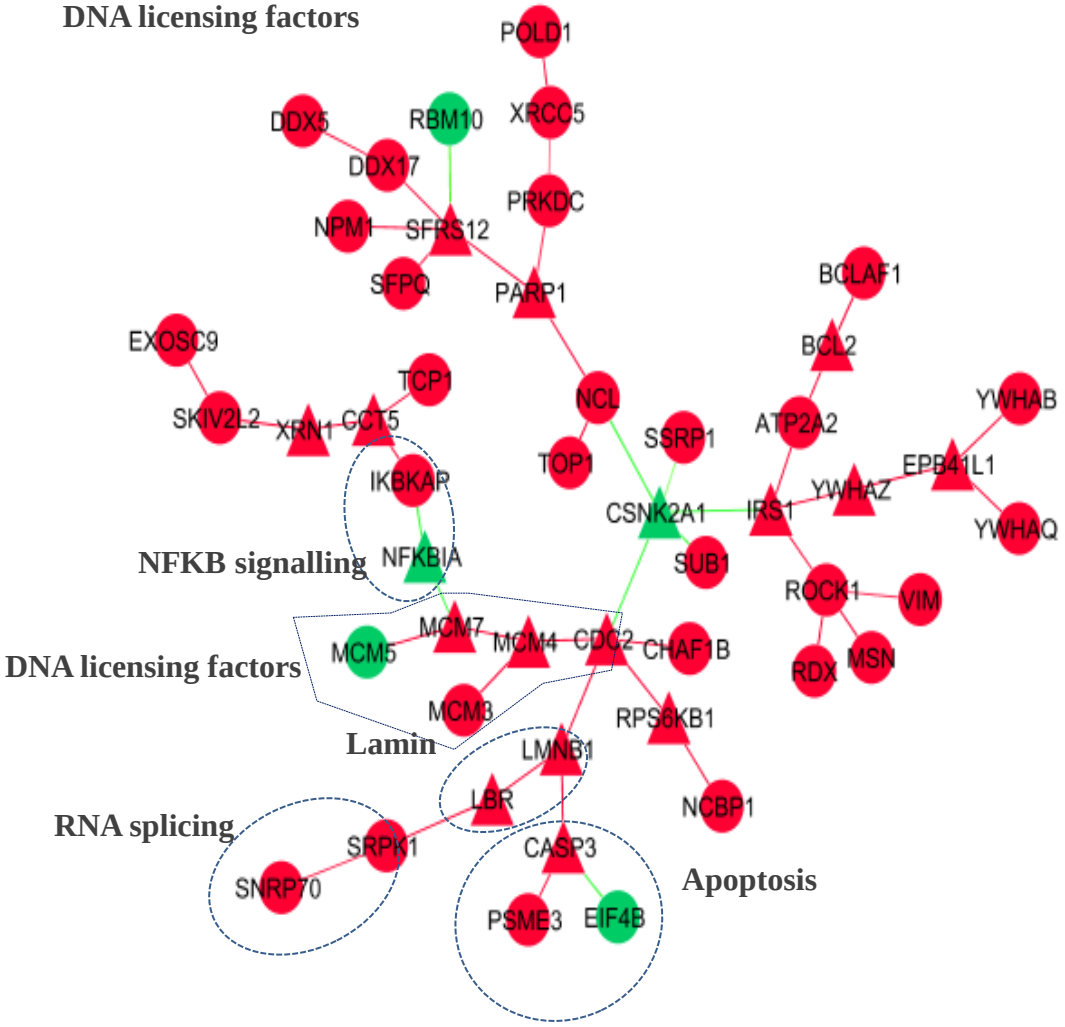
Small correlation values -> large costs
Large correlation values -> small costs

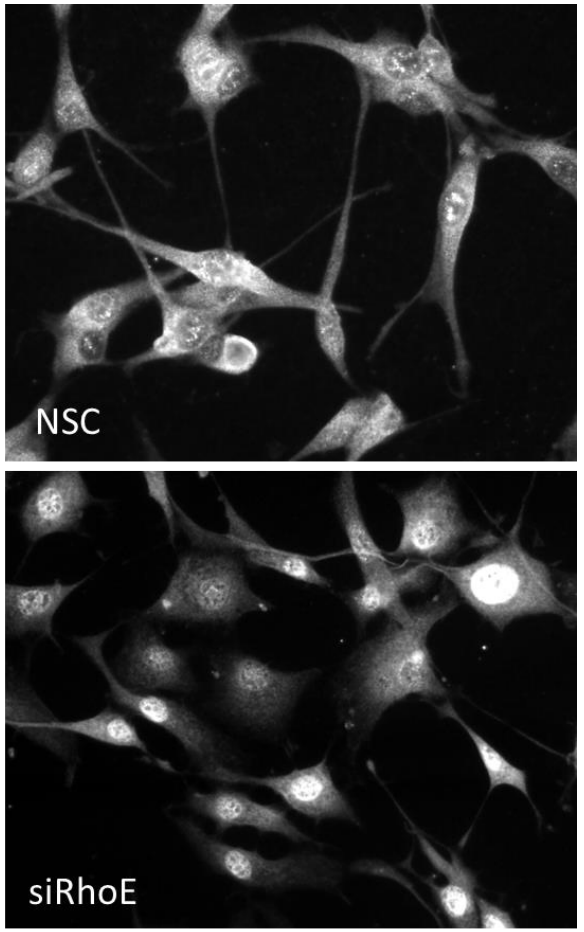
We want to maximise:

$$score_A = \sum_{g \in N} score_g - \sum_{e \in E} score_e$$

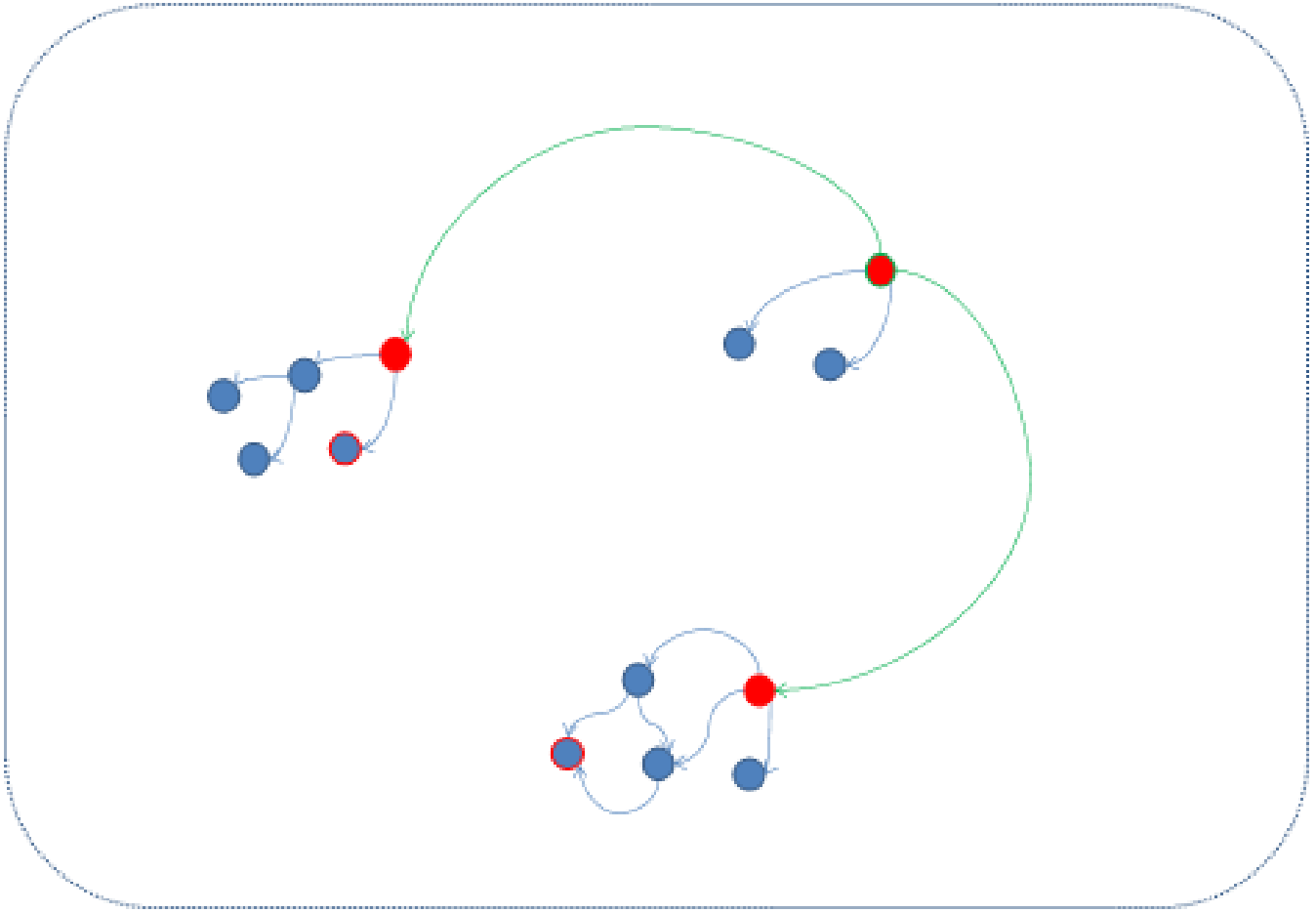
$$Mscore = \frac{\sum_{a \in A} score_a - \mu_k}{s_k}$$

RND3 is nuclear and controls MCM3 localization

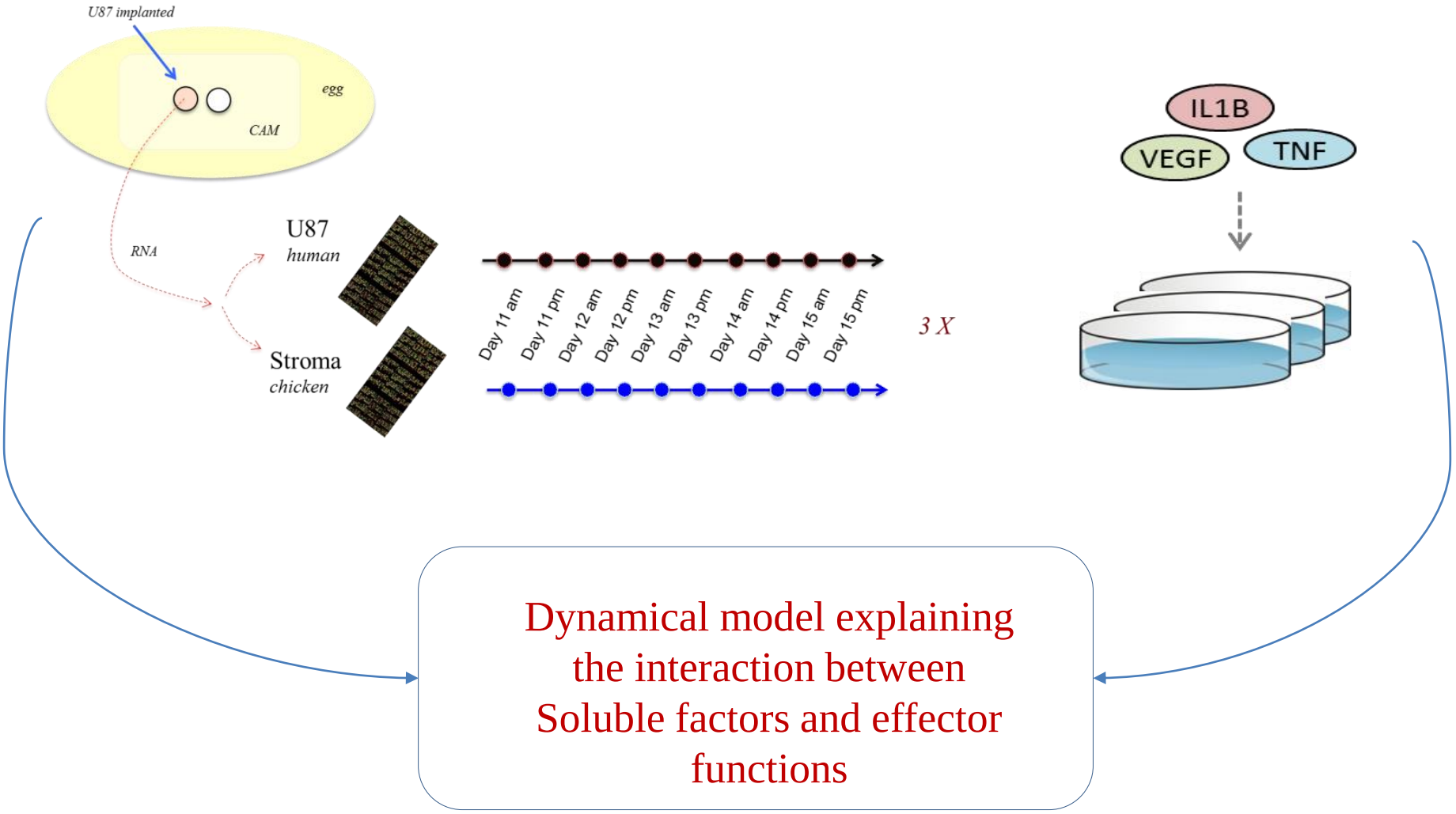




Network Inference



Learning the structure of a dynamical model by integrating multiple experiments



Gupta *et al. BMC Systems Biology* 2011, **5**:52
http://www.biomedcentral.com/1752-0509/5/52

BMC
Systems Biology

METHODOLOGY ARTICLE Open Access

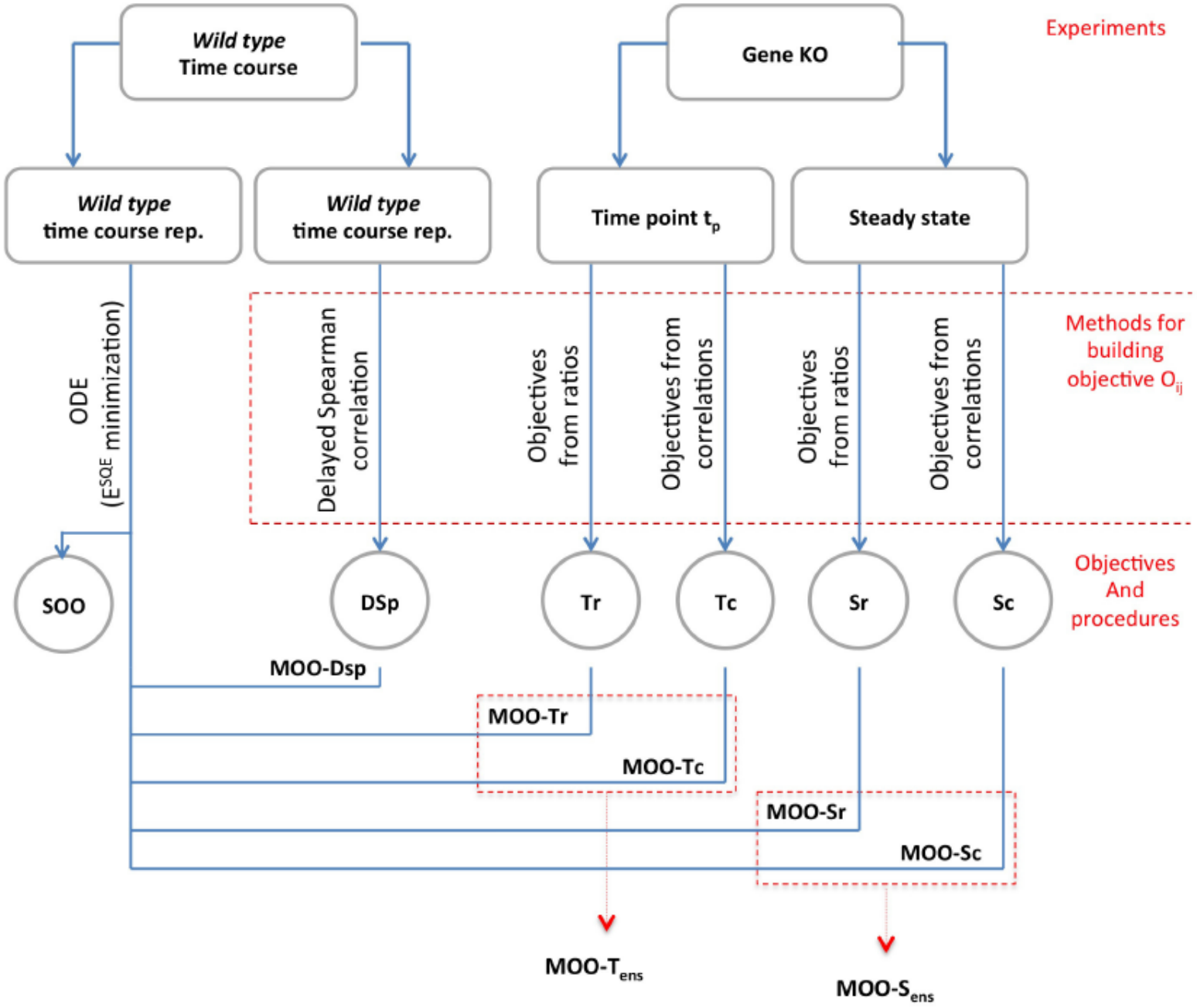
A computational framework for gene regulatory network inference that combines multiple methods and datasets

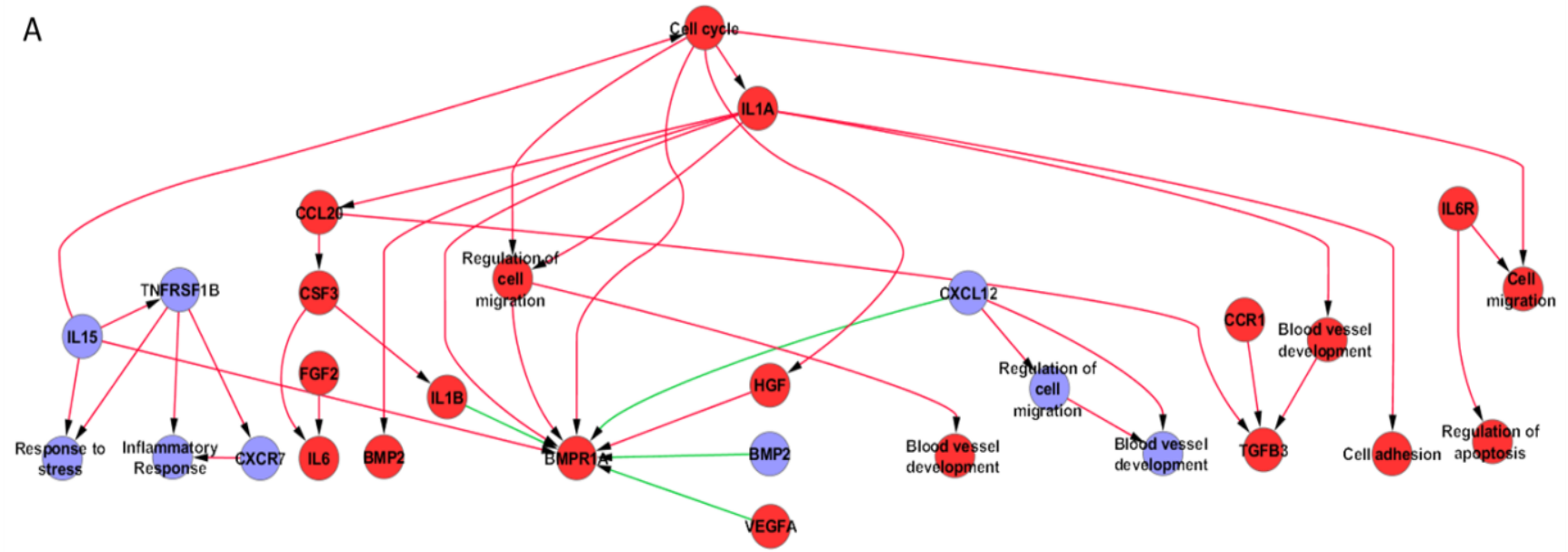
Rita Gupta[1], Anna Stincone[1], Philipp Antczak[1], Sarah Durant[2], Roy Bicknell[2], Andreas Bikfalvi[3] and Francesco Falciani[1*]

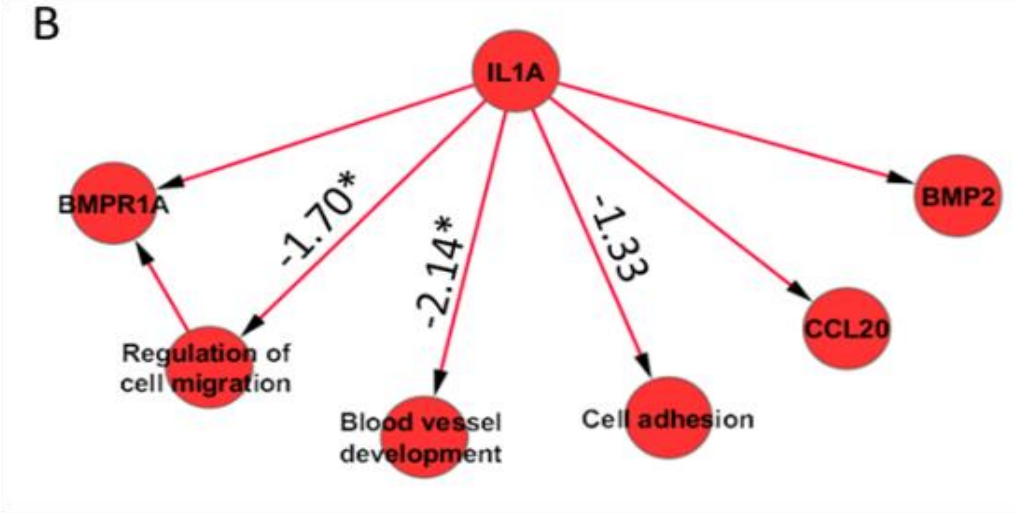
$$\dot{x}_i = \sum_{j=1}^{N} w_{ij} x_j + b_i x_i \qquad (1)$$

$$E^{SQE} = \sum_{i=1}^{N} \sum_{t} \left(x_i^{measured} - x_i \right)^2 \qquad (2)$$

$$E^{Object} = \sum_{j} \sum_{t} \left(o_{ij} - w_{ij} \right)^2 . \qquad (3)$$


Experiments
Wild type
Time course
Gene KO
Wild type
time course rep.
Wild type
time course rep.
Time point t_p
Steady state
ODE
(E^{SQE} minimization)
Delayed Spearman
correlation
Objectives
from ratios
Objectives from
correlations
Objectives
from ratios
Objectives from
correlations
Methods for
building
objective O_{ij}
SOO
DSp
Tr
Tc
Sr
Sc
Objectives
And
procedures
MOO-Dsp
MOO-Tr
MOO-Tc
MOO-Sr
MOO-Sc
MOO-T_{ens}
MOO-S_{ens}


A
Cell cycle
IL1A
CCL20
IL6R
Regulation of cell migration
Cell migration
TNFRSF1B
CSF3
CXCL12
IL15
CCR1
Blood vessel development
FGF2
Regulation of cell migration
HGF
IL1B
Response to stress
Inflammatory Response
CXCR7
IL6
BMP2
BMPR1A
BMP2
Blood vessel development
Blood vessel development
TGFB3
Cell adhesion
Regulation of apoptosis
VEGFA


B
IL1A
BMPR1A
-1.70*
-2.14*
-1.33
BMP2
CCL20
Regulation of cell migration
Blood vessel development
Cell adhesion

Systems Approaches to Nanomaterials toxicity

A Proof of Concept Study

A collaboration with Prof. Chris Vulpe

Exposures

Were done with 15-20 adult (14 day old) daphnids at 1/10 LC_{50} in 800 mL COMBO media for 24 hours.

Gene Expression using Illumina RNA sequencing

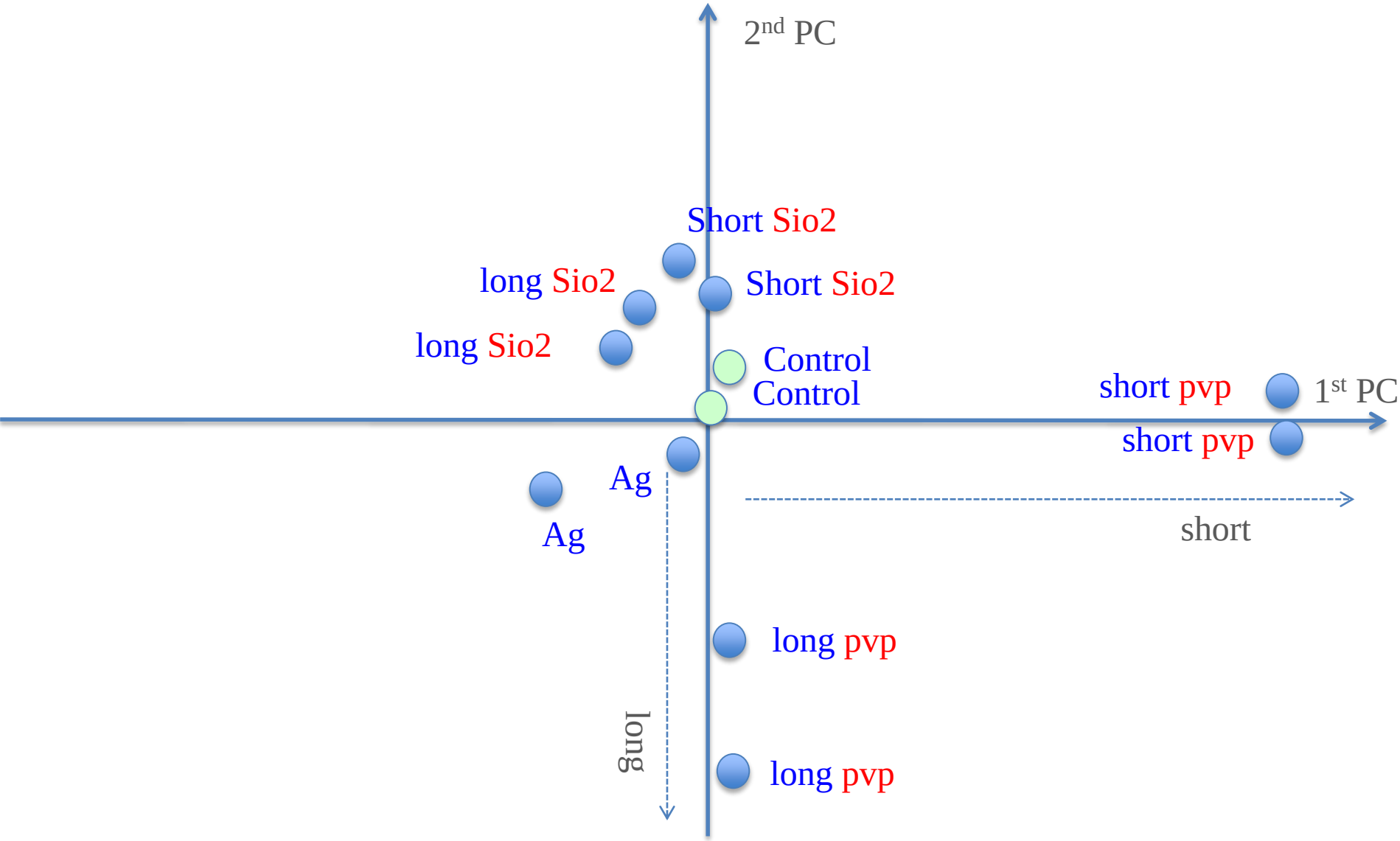
Nanowire samples			
Name	**Dimensions**	**Coating**	**% impurity nanorods**
AgNW-L-PVP	65nm x 20µm 'long'	PVP	1.6
AgNW-L-SiO2	65nm x 20µm 'long'	SiO_2	0.8
AgNW-S-PVP	30nm x 2µm 'short'	PVP	<0.1
AgNW-S-SiO2	30nm x 2µm 'short'	SiO_2	<0.1

PVP = Organic poly(vinyl pyrrolidone)

SiO2 = inorganic amorphous aluminum-doped silica

Systems Approaches to Nanomaterials toxicity

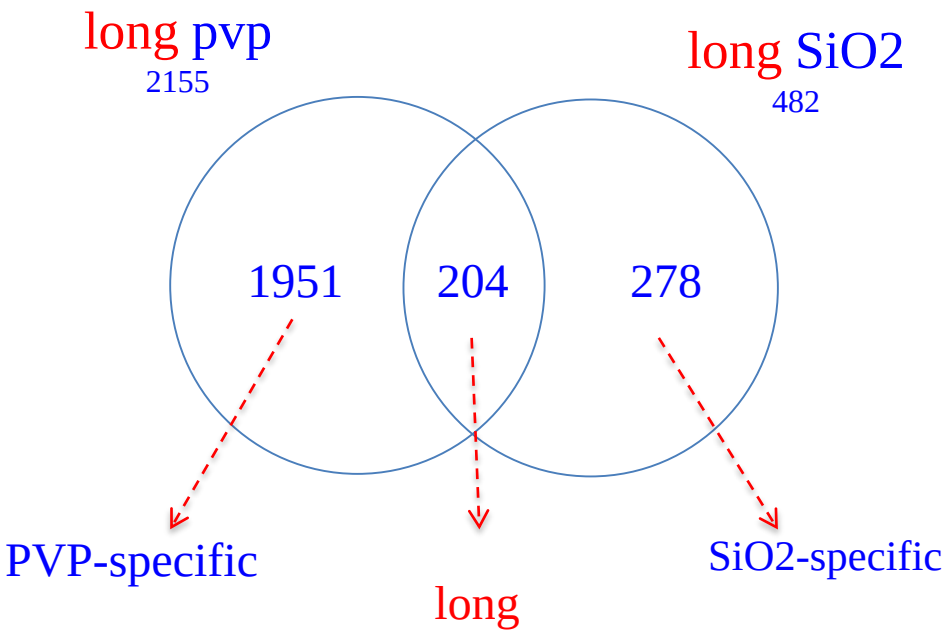
Exploratory Analysis



Systems Approaches to Nanomaterials toxicity

Coating – Functional Analysis

PVP-SiO_2

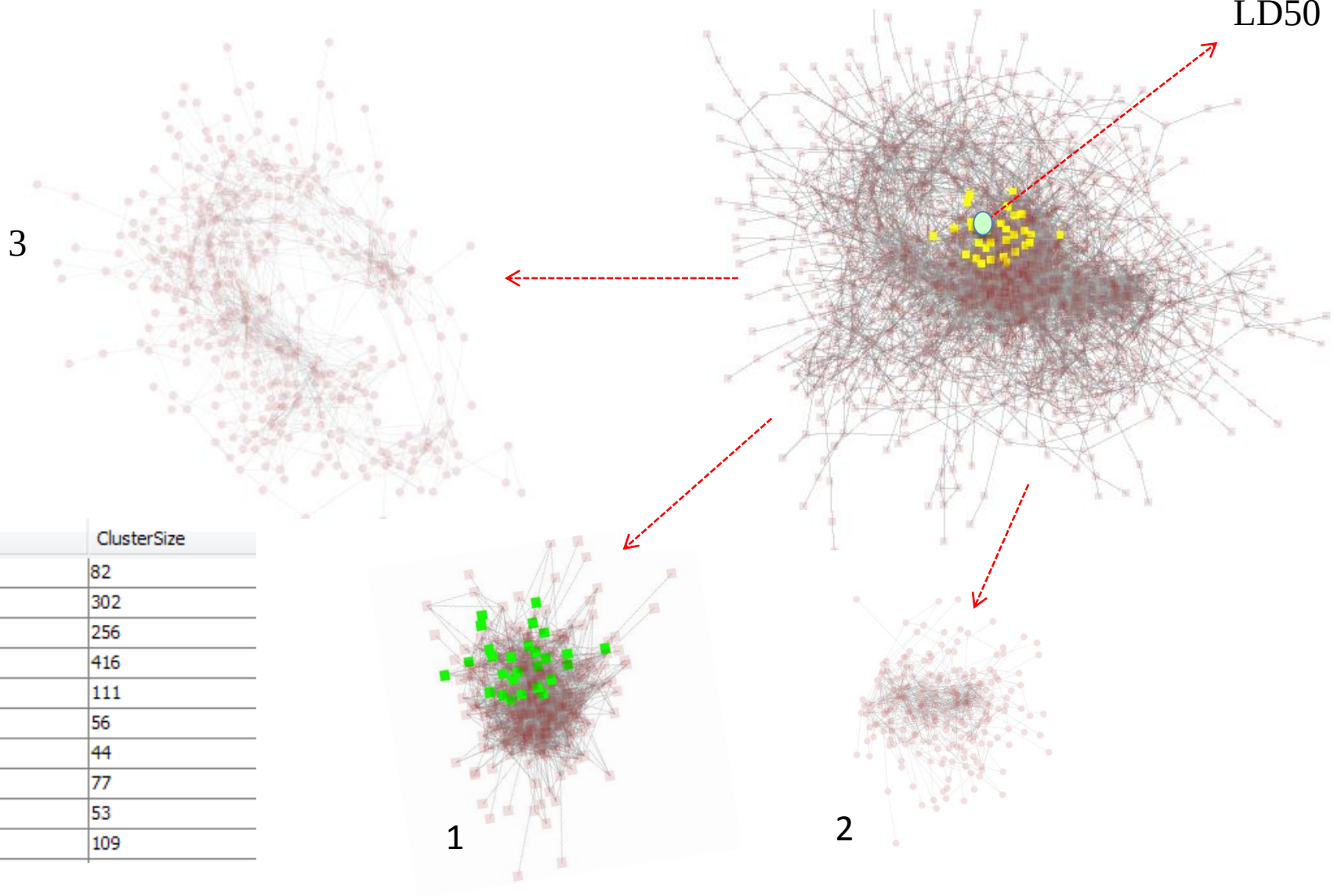


Splisosome
Ribosome
Butanoate metabolism

Peroxisome
Lysosome
Ubiquitin mediated proteolysis
Metabolism of xenobiotics by cytochrome P450

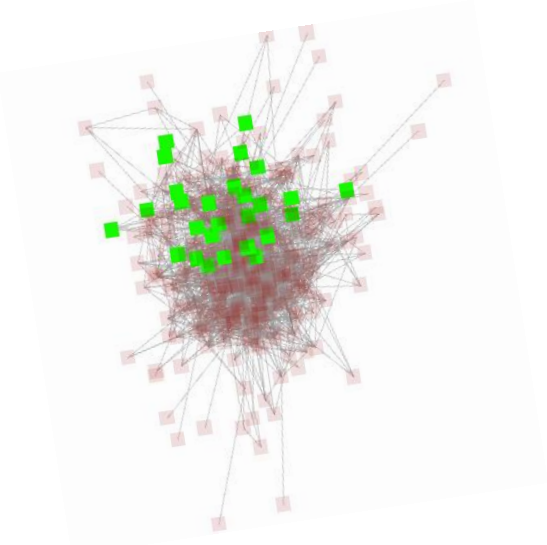
Systems Approaches to Nanomaterials toxicity

Identifying Network Modules Linking Transcriptional Response to Phenotypic endpoints



ClusterID	ClusterSize
0	82
1	302
2	256
3	416
4	111
5	56
6	44
7	77
8	53
9	109

Conclusions



Functional Enrichment

Helicase Activity (11)
DEAD and DEAH domain proteins
Snf2-related protein
Chromosome maintenance complex proteins

Intracellular Protein Transport (10)

RNA Processing (8)

Protein Catabolism (6)
Ubiquitination
Proteasome

Future Plans

1. Obtain Data of sufficient complexity to develop a multi-step QSAR model applied to specific classes of nanoparticles

2. Develop dynamical models of *Daphnia* AgNW response that include more complex phenotypic readouts

3. Use a broader range of species where mechanistic studies are possible (e.g. Zebrafish)

Acknowledgements

University of Liverpool

Francesco Falciani

Kim Clarke*

Nil Turan*

Peter Davidson*

John Herbert*

John Ankers*

Michela Bonomo*

Danilo Basili

Xiaoliang Sun

Jaanika Konberg

Philipp Antczak