Introduction to computational and systems biology

# Lecture 10: Network Controllability analysis

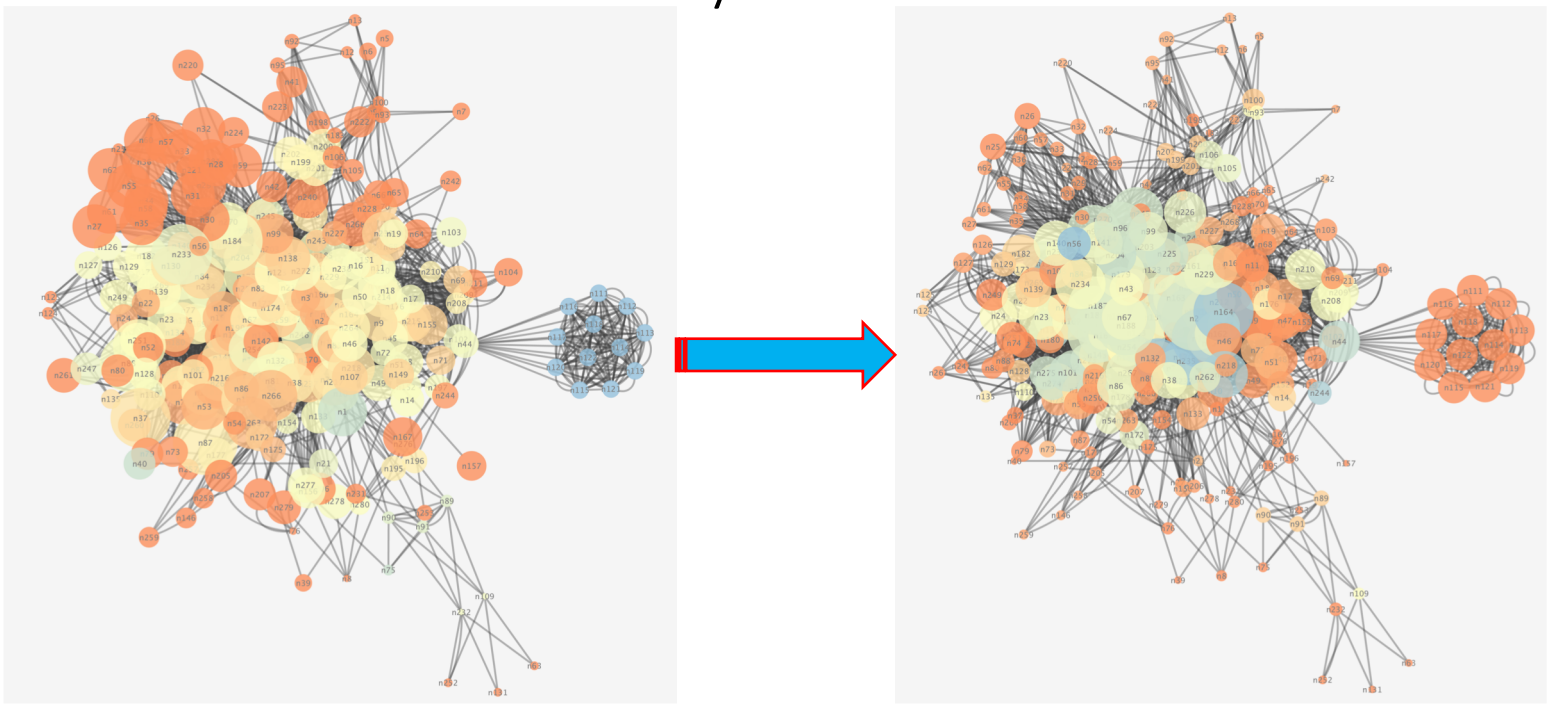
Ion Petre

Department of Mathematics and Statistics

University of Turku

Fall 2019

# What is network controllability?



The ability to change the *global behaviour* of a complex network through well-chosen *minimal local interventions*, taking advantage of the internal inter-connections in the network

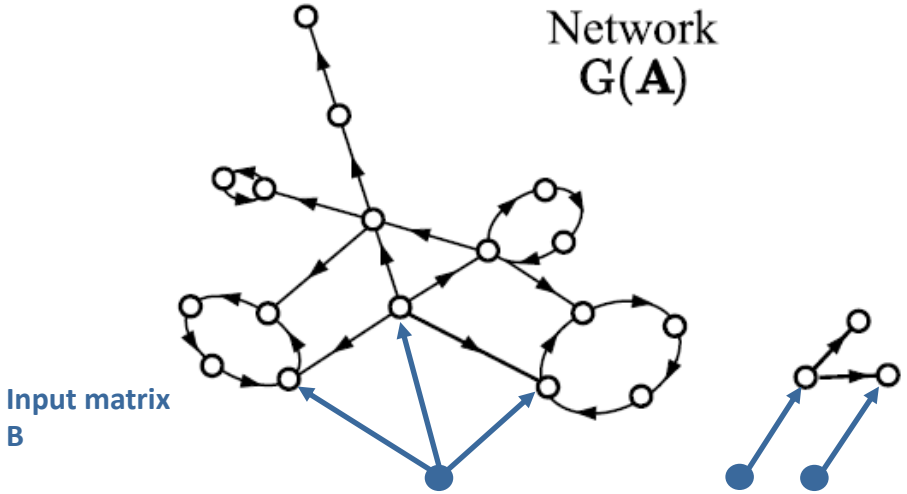
# Liu, Slotine, Barabasi (2011)

# Controllability of linear networks

- **Concept:** A linear dynamical system is *controllable* (from a given set of input nodes) if it can be "driven" from any initial state to any desired final state in finite time

- **Example:** drive a cancer cell to an apoptotic state

- $\frac{dx(t)}{dt} = Ax(t) + Bu(t)$

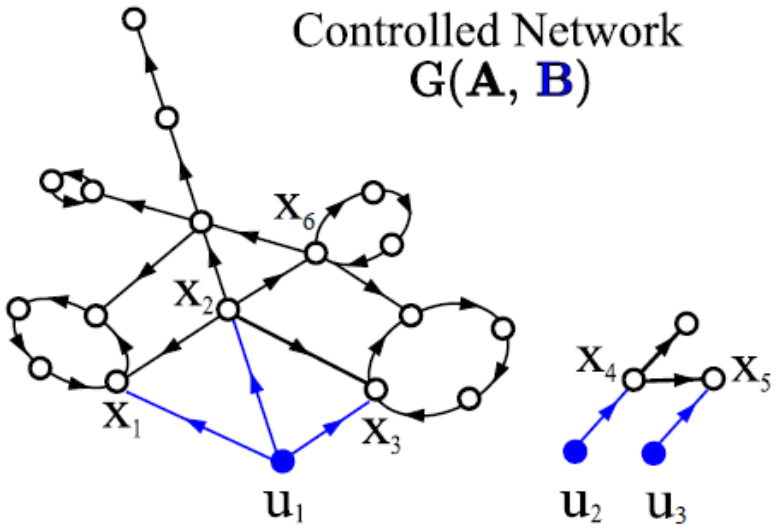


# Two questions

- **Question 1:** Is the system (A,B) controllable?

  - In other words, can we find for any initial state I and final state F, a suitable input u(t) that drives the system from I to F in finite time?

$$\frac{\boldsymbol{dx(t)}}{\boldsymbol{dt}} = \boldsymbol{Ax(t)} + \boldsymbol{Bu(t)}$$

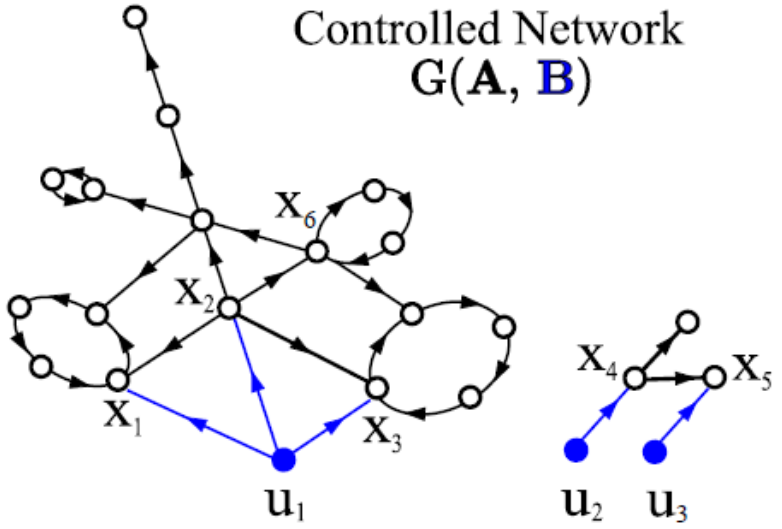



- **Question 2:** For a system A, what is the "smallest" input B such that (A,B) is controllable

# An elegant solution to Q1

- Kalman (1963): a linear model is controllable iff its controllability matrix has full rank
  - Controllability matrix: [B | AB | $A^2B$ | … | $A^{n-1}B$]
  - Note: switch from **ODEs** to **linear algebra**

- **Limitation:** all kinetic details are assumed to be known; **difficult for a large model**



# Structural controllability: elegant results for Q1 & Q2

- **Definition:** A linear model (A,B) is *structurally controllable* if there exists a suitable numeric assignment for the non-zero values of A and B such that (A',B') becomes controllable.

  - Key advantages:
    - Focus on the structure, not on the detailed kinetic setup
    - Switch from *linear algebra* to *graph theory*

  - Result (Lin 1974, Shields and Pearson 1976): if a network is structurally controllable, then it is controllable in all, except a thin set of kinetic setups

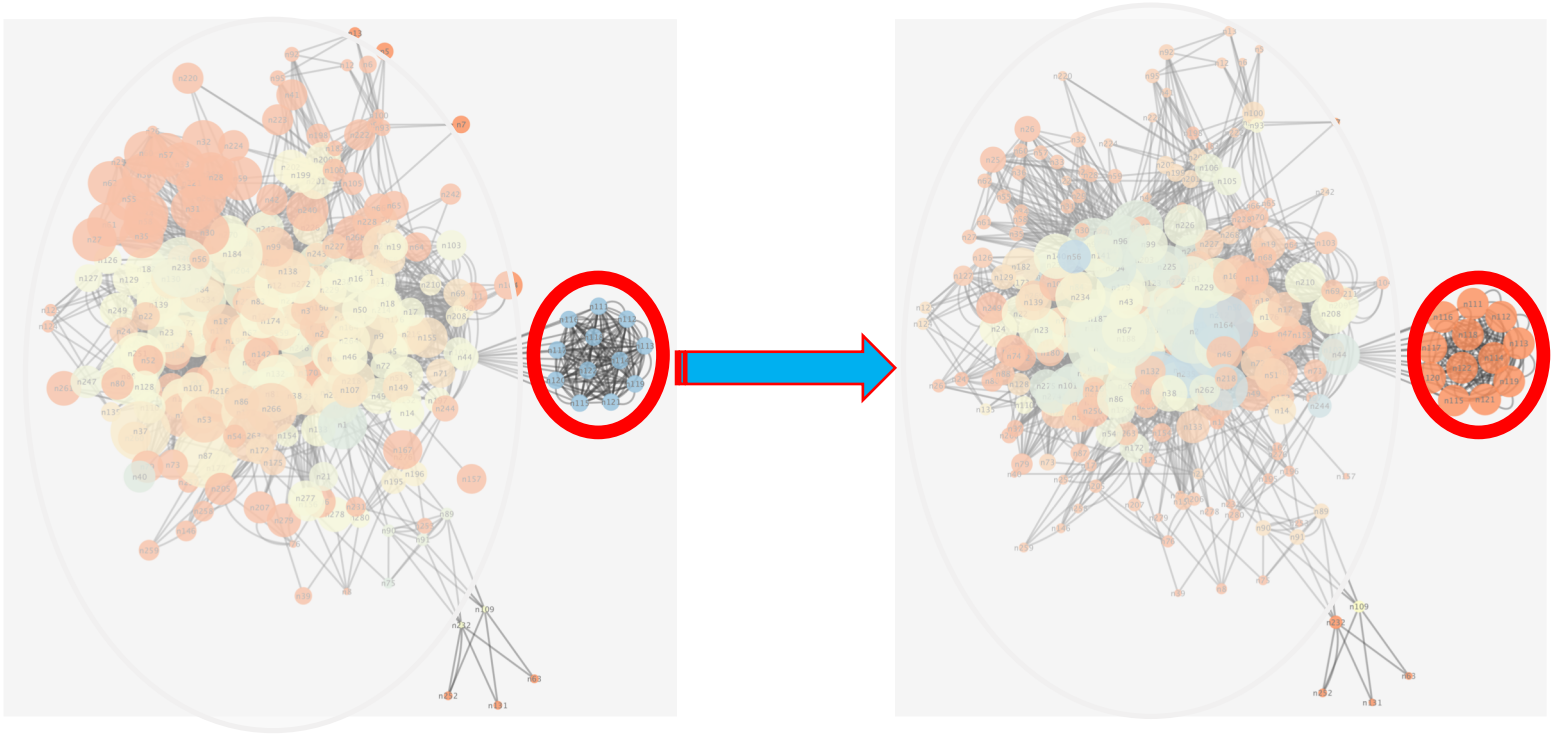

  - Result (Liu et al 2011): Elegant and efficient algorithmic solution to find the **minimum** set of driver nodes needed to control the full network.
    - Key insight: the problem is about special paths in a directed graph

# Liu, Slotine and Barabasi (2011)

**Table 1 | The characteristics of the real networks analysed in the paper**

| Type | Name | $N$ | $L$ | $n_D^{real}$ |
|---|---|---|---|---|
| Regulatory | TRN-Yeast-1 | 4,441 | 12,873 | 0.965 |
| | TRN-Yeast-2 | 688 | 1,079 | 0.821 |
| | TRN-EC-1 | 1,550 | 3,340 | 0.891 |
| | TRN-EC-2 | 418 | 519 | 0.751 |
| | Ownership-USCorp | 7,253 | 6,726 | 0.820 |
| Trust | College student | 32 | 96 | 0.188 |
| | Prison inmate | 67 | 182 | 0.134 |
| | Slashdot | 82,168 | 948,464 | 0.045 |
| | WikiVote | 7,115 | 103,689 | 0.666 |
| | Epinions | 75,888 | 508,837 | 0.549 |
| Food web | Ythan | 135 | 601 | 0.511 |
| | Little Rock | 183 | 2,494 | 0.541 |
| | Grassland | 88 | 137 | 0.523 |
| | Seagrass | 49 | 226 | 0.265 |
| Power grid | Texas | 4,889 | 5,855 | 0.325 |
| Metabolic | *Escherichia coli* | 2,275 | 5,763 | 0.382 |
| | *Saccharomyces cerevisiae* | 1,511 | 3,833 | 0.329 |
| | *Caenorhabditis elegans* | 1,173 | 2,864 | 0.302 |
| Electronic circuits | s838 | 512 | 819 | 0.232 |

# Target controllability



The ability to change the behaviour of a *target subset*

# Target controllability

- Linear dynamical systems with output
  - $\frac{dx(t)}{dt} = Ax(t) + Bu(t)$
  - $y(t) = Cx(t)$, where C is the output matrix

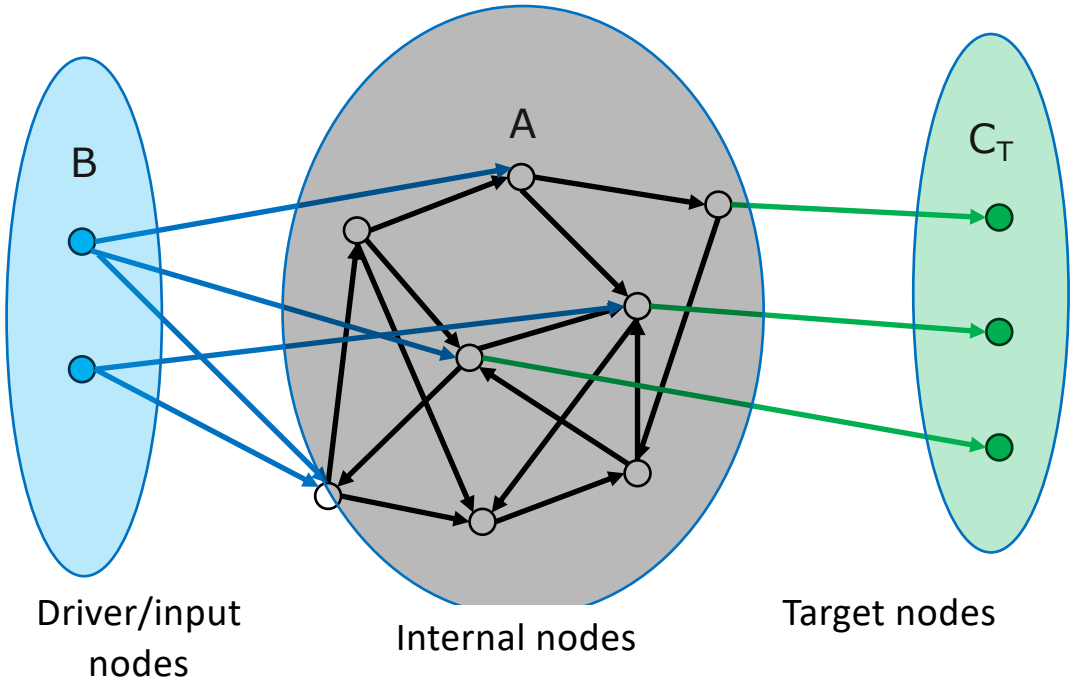
    - Our focus: $C_{t,t}$=1 for all nodes t in the target set T, and it is 0 otherwise. Also denoted as $C=C_T$

- *Structural target controllability*: the possibility to turn the system controllable by modifying the non-zero values in A, B, C
  - Result (Lin 1974, Shields and Pearson 1976): if a network is structurally target controllable, then it is target controllable in all, except a thin set of kinetic setups

# Graph-based target controllability

- Insight: *linear dynamical systems can be represented as directed weighted graphs*
  - $\frac{dx(t)}{dt} = Ax(t) + Bu(t)$
  - $y(t) = C_T x(t)$

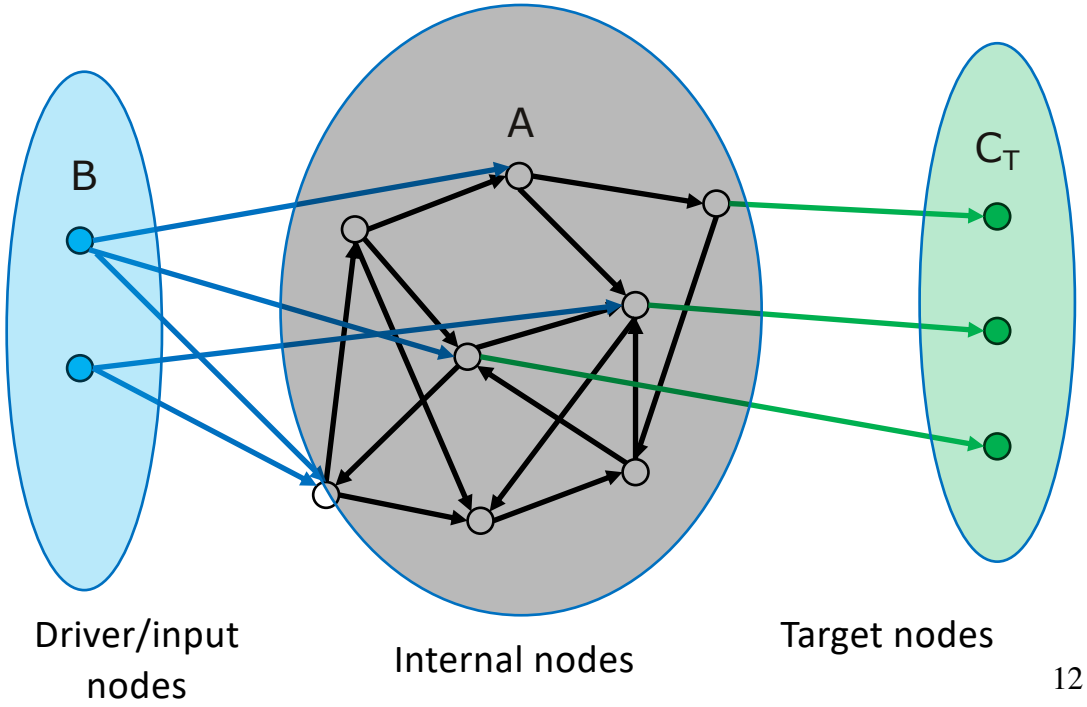
Structural target controllability:



# Graph-based structural target controllability

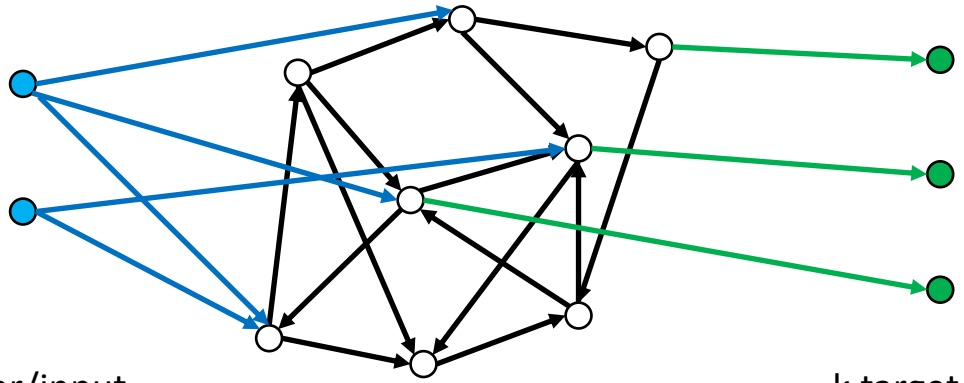
- Key result for Q1':
  - (Poljak, Murota 1990) If $(A,B,C_T)$, with $|T|=k$, is *structurally target controllable*, then
    - there is a set of k paths starting in the driver nodes and ending in each of the target nodes such that
    - no two paths intersect at the same distance from their end points

Structural target controllability:



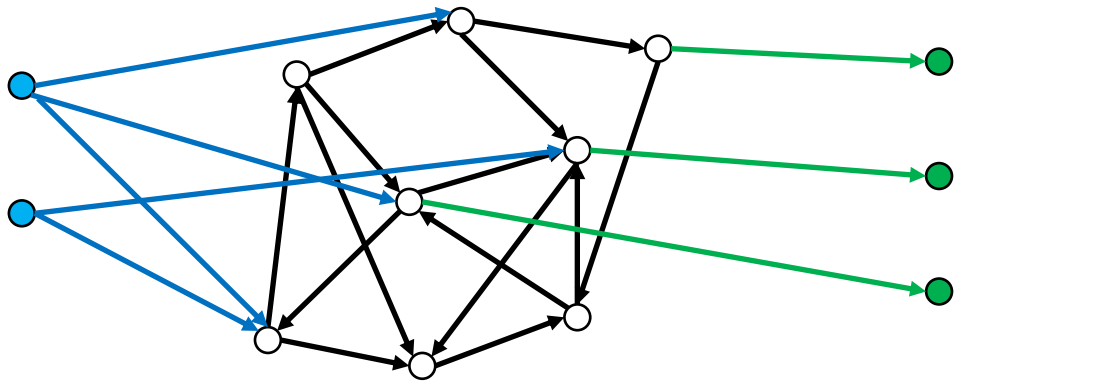
# Q2': Can a target be controlled?

- Input:
  - The internal nodes and their interconnections
  - The target nodes
- Output:
  - A (minimal?) set of driver nodes such as the linear system is structurally target controllable
- Goal (based on Poljak, Murota 1990): choose a set of driver nodes so that
  - there are k paths starting in the driver nodes and ending in each of the target nodes
  - no two paths intersect at the same distance from their end points

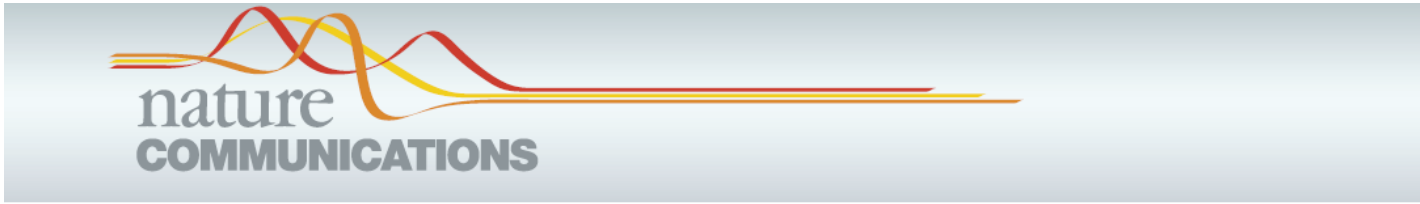


# Can a target be controlled?

- Goal (based on Poljak, Murota 1990) : choose a set of driver nodes so that
  - there are k paths starting in the driver nodes and ending in each of the target nodes
  - no two paths intersect at the same distance from their end points
- **Key insight: this is an iterated matching problem**
  - Match the targets to some internal nodes; they become the new targets
  - Match the new targets against the internal nodes;
    - Those remaining unmatched become driver nodes
    - The left side of the matching becomes the new target
  - Stop the iteration after n steps and make all remaining targets driver nodes



# Gao et al. – Target controllability: a greedy-type approximation (2014)




ARTICLE

Received 20 Feb 2014 | Accepted 30 Sep 2014 | Published 12 Nov 2014

DOI: 10.1038/ncomms6415

## Target control of complex networks

Jianxi Gao[1], Yang-Yu Liu[1,2,3], Raissa M. D'Souza[4,5] & Albert-László Barabási[1,2,3]

Controlling large natural and technological networks is an outstanding challenge. It is typically neither feasible nor necessary to control the entire network, prompting us to explore target control: the efficient control of a preselected subset of nodes. We show that the structural

# Target Controllability

- **Our results:**
  - The Target Controllability optimisation problem is **NP-hard**
  - **Fastest heuristics algorithms** and **better approximation** results
  - Control algorithms which **maximize the use of FDA-approved drugs** in controlling essential genes
  - An automated (Web-based) **software** for bio-net control
  - **Demonstrated the approach for several types of cancer**

- Controlling Directed Protein Interaction Networks in Cancer. *Scientific Reports* 2017. **Top 5% most read oncology papers in Scientific Reports 2017.**
- Structural target controlability of linear networks. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2018.
- Controlability of Linear Networks. **nVidia Best Paper Award CMSB 2016.**
- NetContrl4BioMed: A pipeline for biomedical data acquisition and analysis of network controllability. *BMC Bioinformatics* 2018.

# Target controllability in the biomedical domain

- Input controller is implemented through drugs
- Staying realistic
  - **No super-drugs**: input nodes have bounded out-degree
    - Matrix B has at most N non-zero values on every column
  - **Minimize the number of drugs** given to a patient: minimize the number of input (also called driver) nodes
  - **FDA-approved drugs**: the list of potential driven nodes may be given in the input of the problem

# Target controllability as an optimization problem

- **Target-control optimization problem**

- **Input**: Matrix A (n x n) and matrix C (k x n), with k≤n, where every row and column of C have at most one non-zero entry with value 1.

- **Output**: Matrix B (n x m) such that
  - Srank (CB | CAB | $CA^2B$ |…|$CA^{n-1}B$)=k
  - m is minimum

  - **Result: the (decision) problem is NP-hard!**

- **Target Control Optimization Problem (TCP)**
- **Input**: Matrix A (n x n) and matrix C (k x n), with k≤n, where every row and column of C have at most one non-zero entry with value 1.
- **Output**: Matrix B (n x m) such that
  - Srank (CB | CAB | $CA^2B$ |…|$CA^{n-1}B$)=k
  - m is minimum

- **N-bounded Target Control Optimization Problem (N-TCP)**
- Additional condition for the output: B contains at most N non-zero values on every column

- **Edge-bounded Target Control Optimization Problem (E-TCP)**
- Alternative optimization for the output: B contains a minimum number of non-zero rows

- **The decision versions of all these problems defined in the usual way:**
  - **TCP-d** **N-TCP-d** **E-TCP-d**

- **Result: all of these decision problems are NP-hard!**

www.nature.com/scientificreports


OPEN

# Controlling Directed Protein Interaction Networks in Cancer

Krishna Kanhaiya[1], Eugen Czeizler[1,2], Cristian Gratie[1] & Ion Petre[1]

Control theory is a well-established approach in network science, with applications in bio-medicine and cancer research. We build on recent results for structural controllability of directed networks, which identifies a set of driver nodes able to control an a-priori defined part of the network. We develop a novel and efficient approach for the (targeted) structural controllability of cancer networks and demonstrate it for the analysis of breast, pancreatic, and ovarian cancer. We build in each case a protein-protein interaction network and focus on the survivability-essential proteins specific to each cancer type. We show that these essential proteins are efficiently controllable from a relatively small computable set of driver nodes. Moreover, we adjust the method to find the driver nodes among FDA-approved drug-target nodes. We find that, while many of the drugs acting on the driver nodes are part of known cancer therapies, some of them are not used for the cancer types analyzed here; some drug-target driver nodes identified by our algorithms are not known to be used in any cancer therapy. Overall we show that a better understanding of the control dynamics of cancer through computational modelling can pave the way for new efficient therapeutic approaches and personalized medicine.

# Target control of cancer essential genes

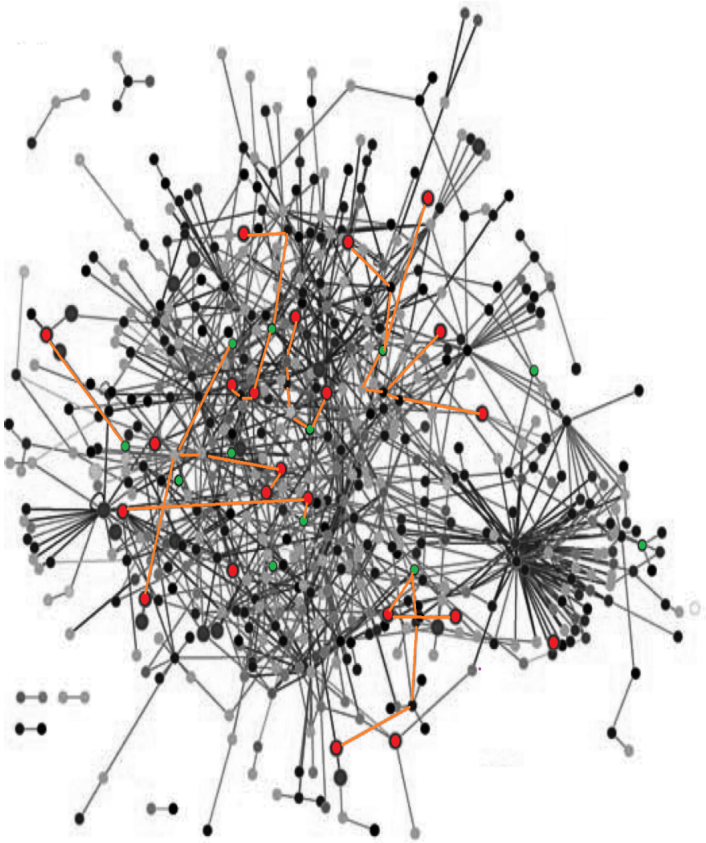
- Consider three types of cancer:
  - Breast cancer,
  - Pancreatic cancer,
  - Ovarian cancer.

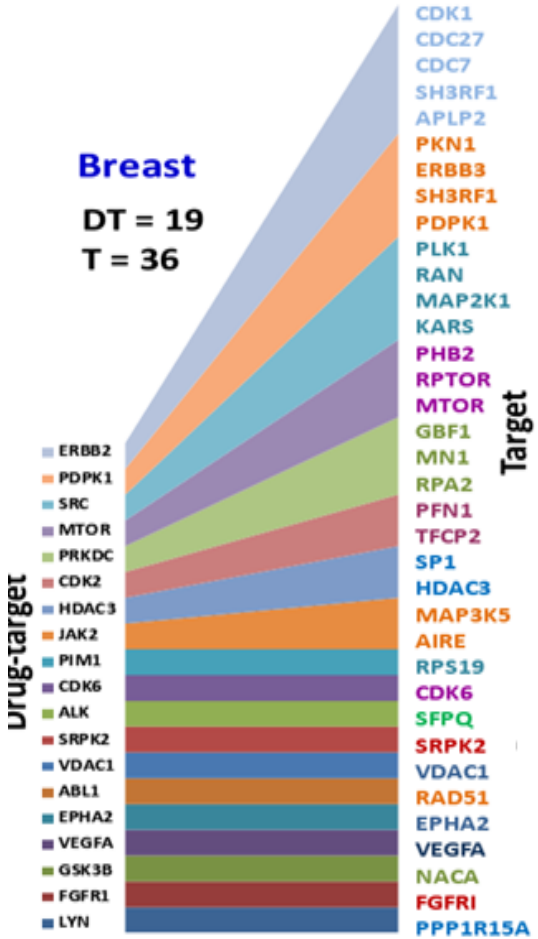
- Select the relevant set of cancer-specific genes/proteins

- Select a set of (target) essential genes

- Use known gene signalling/regulatory/interaction networks to derive the cancer-specific gene interaction network

# Essential genes, Drug targets, and Signaling networks





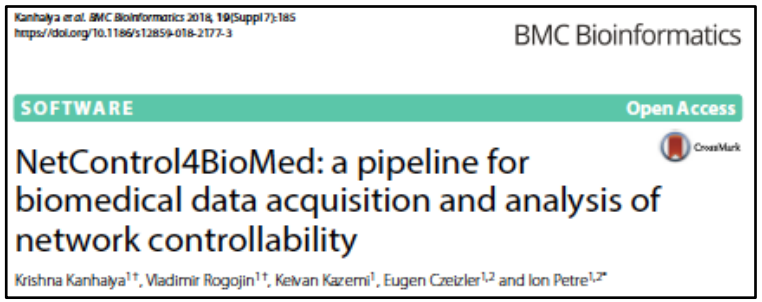
# Literature-based validation

| Drug-target | Target genes | Anti-cancer drug | Know to be used in cancer therapies |
|---|---|---|---|
| ERBB2 | CDK1, CDCH2, CDC7, SH3RF1, APLP2 | Lapatinib | Breast, Lung |
| SRC | PLK1, RAN, MAP2K1, KARS | Dasatinib, Bosutinib, Ponatinib | CML |
| PDPK1 | PNK1, ERBB3, SH3RF1, PDPK1 | None | None |
| PRKDC | GBF1, MN1, RPA2 | None | None |
| MTOR | PHB2, RPTOR, MTOR | Temsirolimus | RCC, BMC |
| JAK2 | MAP3K5, AIRE | Ruxolitinib, Erlotinib | Pancreatic |
| HDAC3 | SP1, HDAC3 | Vorinostat | CTCL |
| CDK2 | PFN1, TFCP2 | None | None |

Our predictions on drug-target proteins with highest potential impact for breast cancers.

# NetControl4BioMed

http://combio.org/

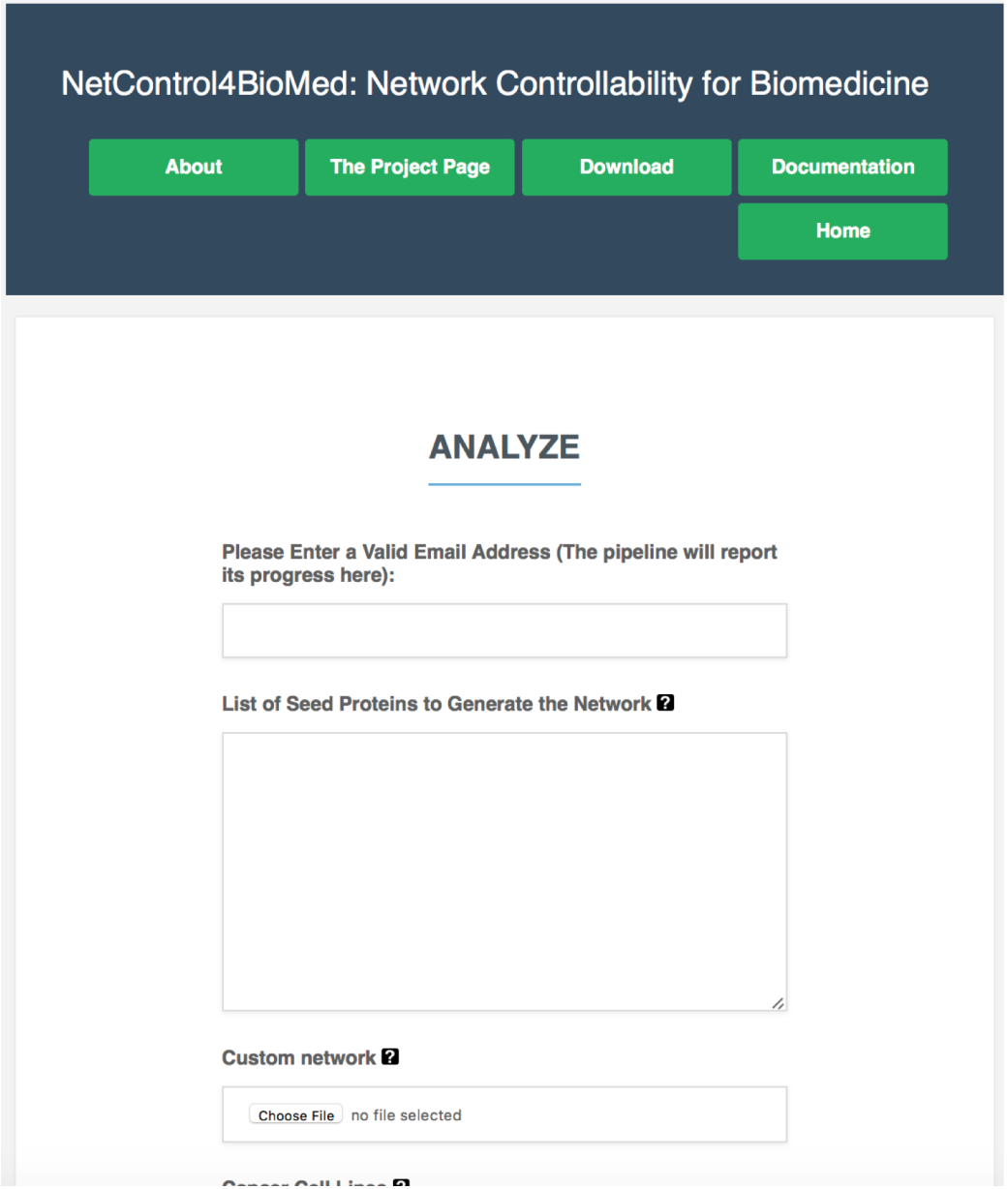

Kanhaiya et al. BMC Bioinformatics 2018, 19(Suppl 7):185
https://doi.org/10.1186/s12859-018-2177-3

BMC Bioinformatics

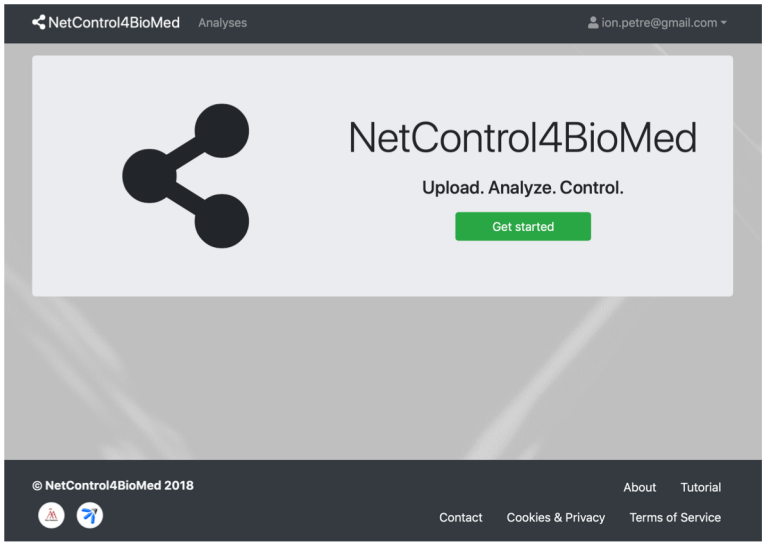
SOFTWARE Open Access

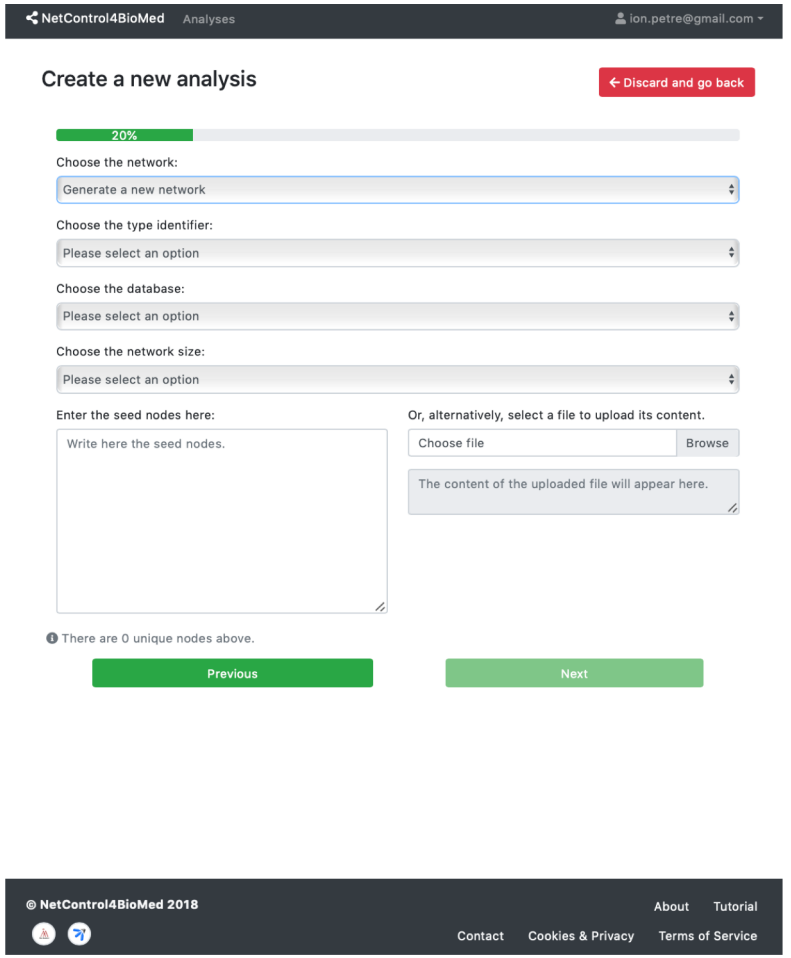
NetControl4BioMed: a pipeline for biomedical data acquisition and analysis of network controllability

Krishna Kanhaiya[1†], Vladimir Rogojin[1†], Kelvan Kazemi[1], Eugen Czeizler[1,2] and Ion Petre[1,2*]

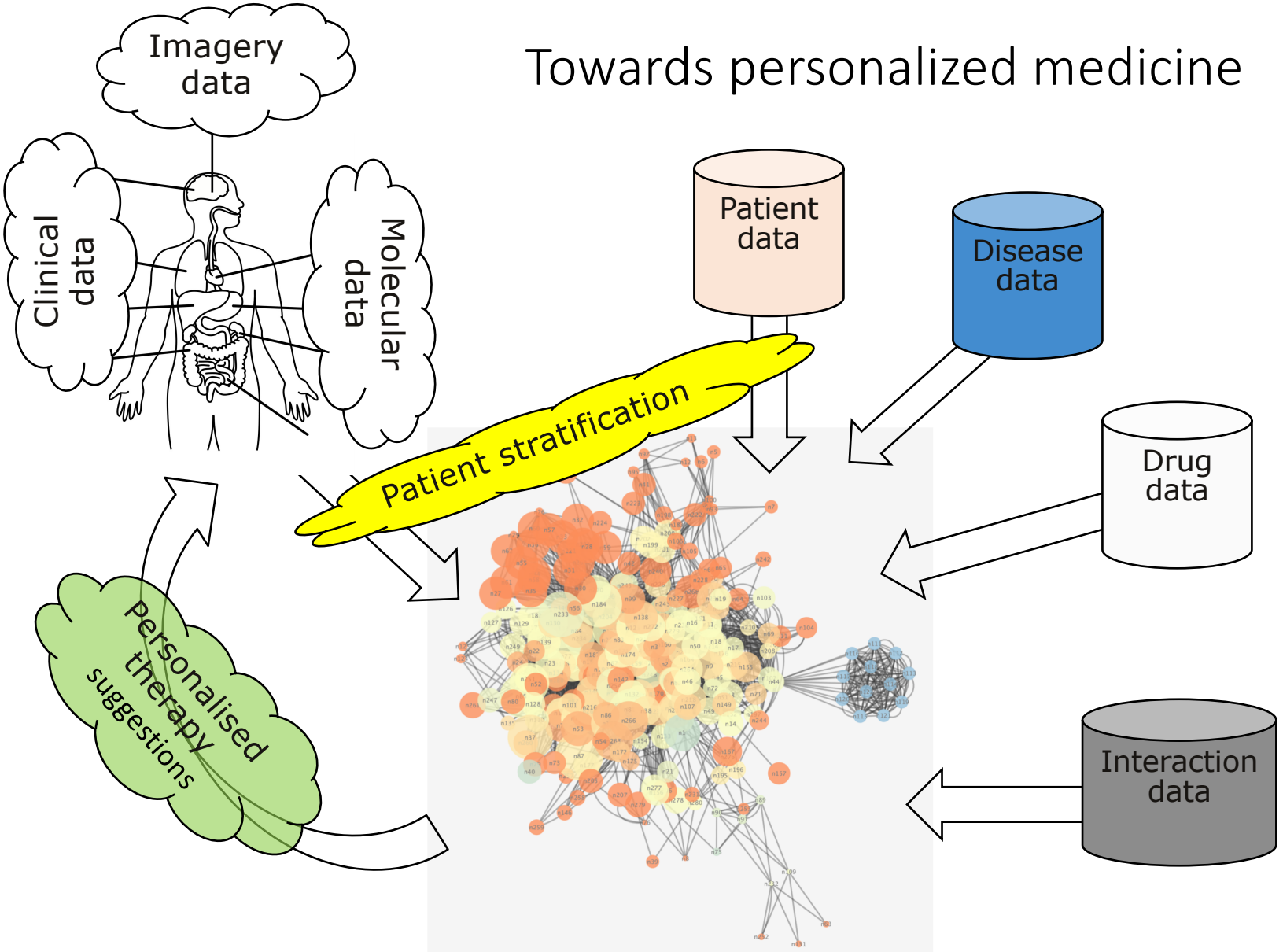


# NetControl4BioMed 2020 release



# Towards personalized medicine



# Example: multiple myeloma

- Jens Lohr et al. Widespread genetic heterogeneity in multiple myeloma: implications for targeted therapy. *Cancer Cell* 2014.
    - 203 patients
    - Genetic mutation data, copy number alterations

- Approach
    - Disease-specific analysis on this data set
        - Network around the most frequently mutated genes
        - Predict drug combinations
        - Validate against standard therapy lines
    - Patient-specific analysis
        - Network around the patient's own mutated genes
        - Predict drug combinations
        - Analyze differences wrt the standard therapy lines

# Current research lines

- Controllability in a (semi-)quantitative framework
  - Approximations over ODEs
  - Boolean networks
  - …
- Edge-labeled directed graphs
  - Activation, inhibition influences
- More complex behavior of the driver nodes
  - Side-effects
  - Synergy effects
- Approximation guarantees for our various heuristics

- Proof of concept in personalized medicine
- Multi-clone tumors
- Clinical validation

# Contributors

- E. Czeizler

K.Kanhaiya

K. Chiu Wu

V. Rogojin

C. Gratie

V. Popescu

- Andrei Andronescu (Bucharest)
- Diana Ion (Bucharest)
- Eduard Mititelu (Bucharest)
- Alex Nedea (Bucharest)

- Sadra Safa (Tehran)
- Negin Majidi (Tehran)

- Daniela Schacherer (Heidelberg)
- Jose Angel Sanchez Martin (Madrid)
- Nicoleta Balteanu (Bucharest)

- Patric Gustafsson (Turku)
- Joel Sjöblom (Turku)
- Elio Nushi (Turku)

Work funded by:

ACADEMY OF FINLAND

Challenge Finland