UNIVERSITY OF SURREY

Studying prostate cancer as a network disease by qualitative computer simulation with Petri Nets

Nicholas Stoy, Sophie Chen, Andrzej M. Kierzek*

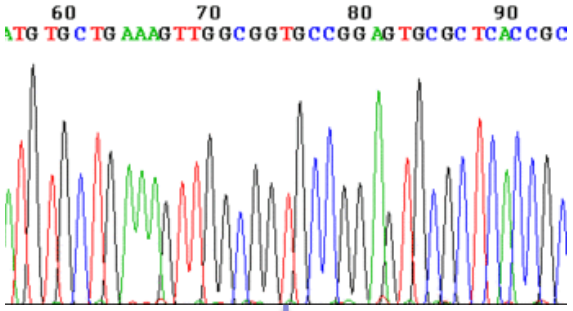
*a.kierzek@surrey.ac.uk

www.surrey.ac.uk

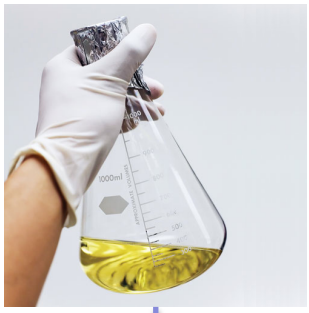
Prediction of phenotype from genotype and environmental conditions.

UNIVERSITY OF SURREY

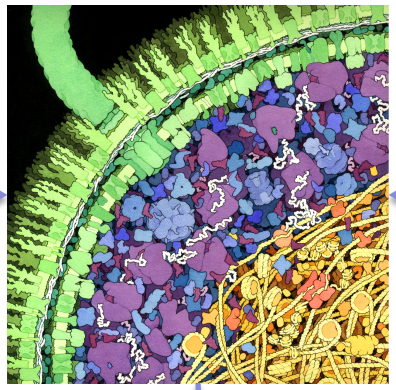
GENOTYPE

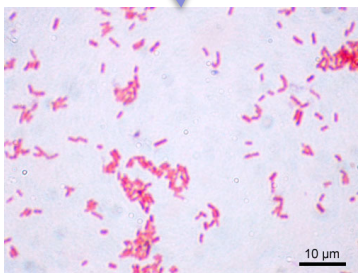


ENVIRONMENT



LIVING CELL





PHENOTYPE

We can sequence any DNA of interest, including full genome of an individual, but we are not making full use of this information yet.

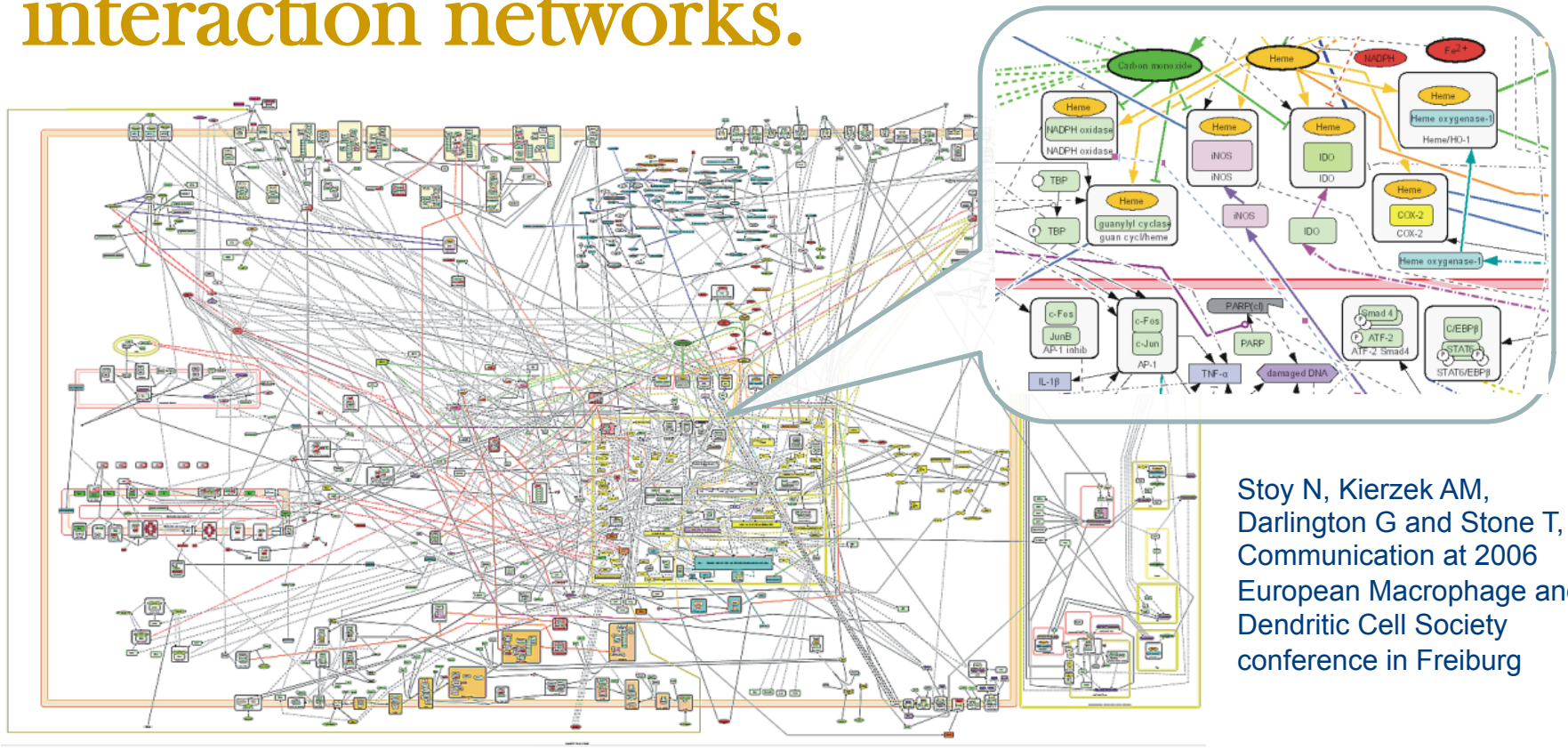
The ability to predict phenotype occurring for given genetic background and environmental conditions will revolutionise medicine and biotechnology.

Molecular Biology knowledge will be used to reverse engineer molecular machinery of the cell as a computer model and use simulation to predict cellular behaviour for particular set of genetic and environmental perturbations.

www.surrey.ac.uk

Reconstruction of molecular interaction networks.

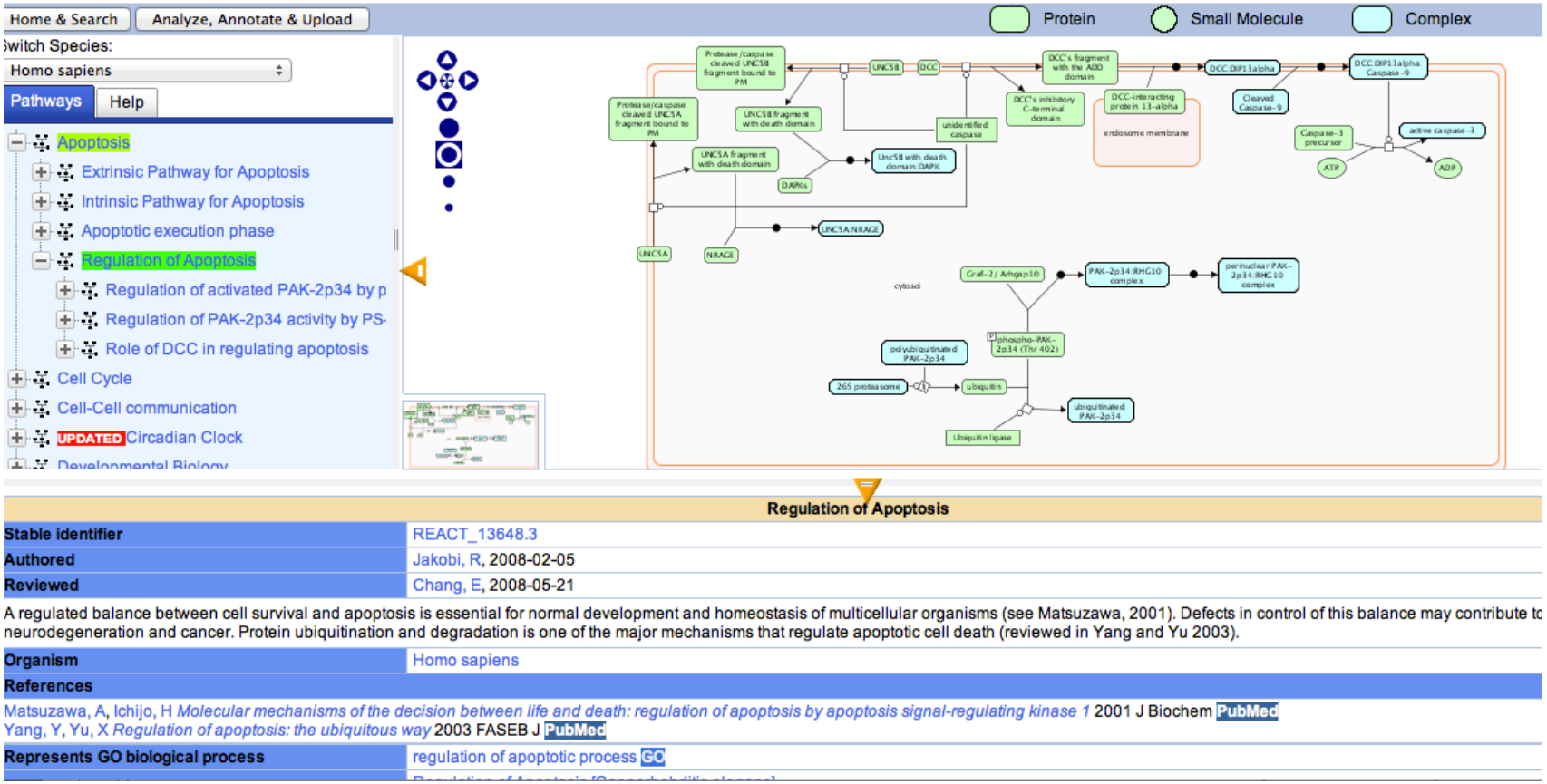
UNIVERSITY OF SURREY



Stoy N, Kierzek AM, Darlington G and Stone T, Communication at 2006 European Macrophage and Dendritic Cell Society conference in Freiburg

The model of signalling pathways in human macrophage constructed in Systems Biology Graphical Notation (SBGN). The model contains 605 molecular species and 707 interactions.

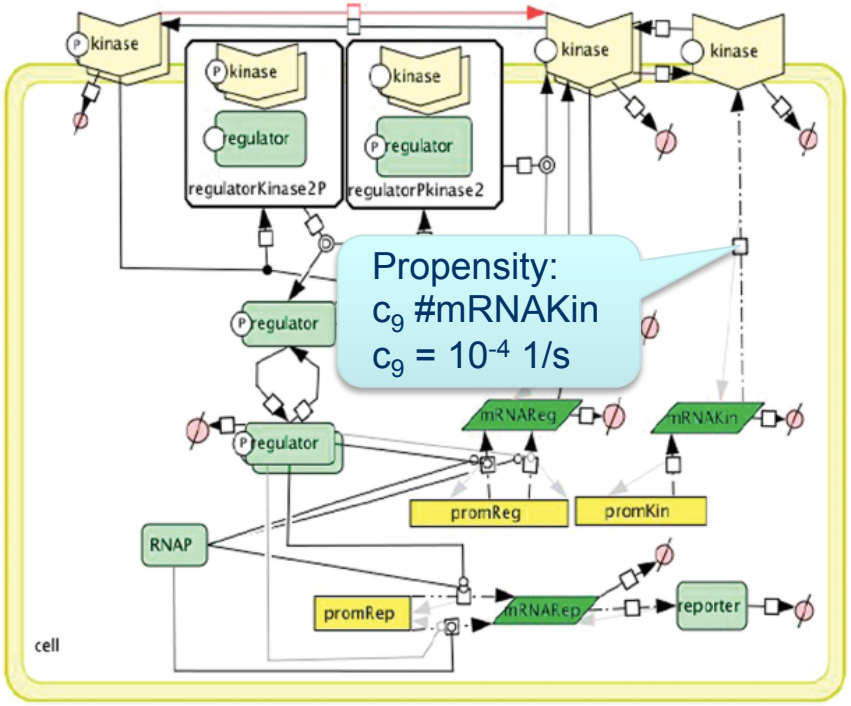
Reactome: Community based, peer-reviewed reconstruction effort.



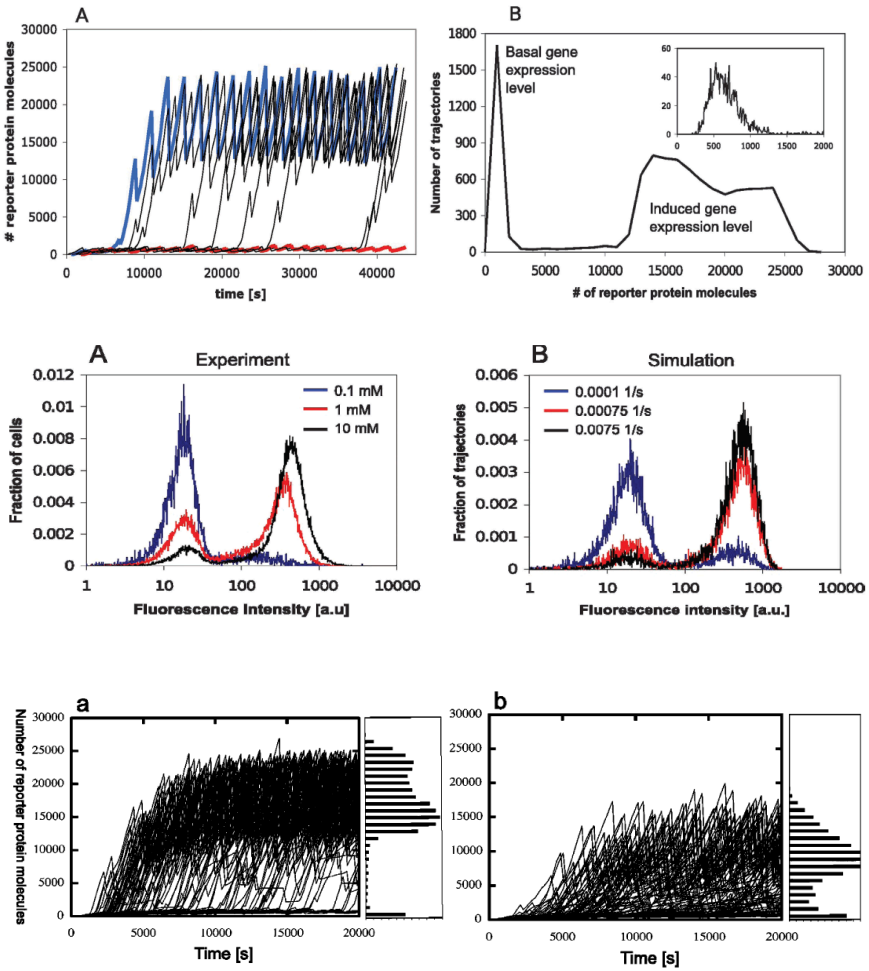
UNIVERSITY OF SURREY

How to simulate behaviour of the molecular interaction network?

In an ideal world, I would like to simulate molecular interaction network as CTMC ...



Stochastic kinetic model of two component system signalling. Kierzek, Zhou, Wanner, Molecular Biosystems, 2010

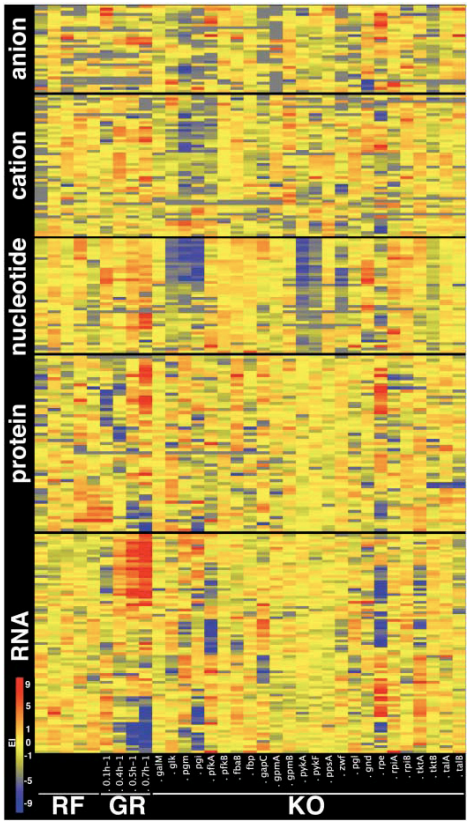


Lack of quantitative parameters is a major obstacle towards dynamic model including all genes in the genome ....

UNIVERSITY OF SURREY

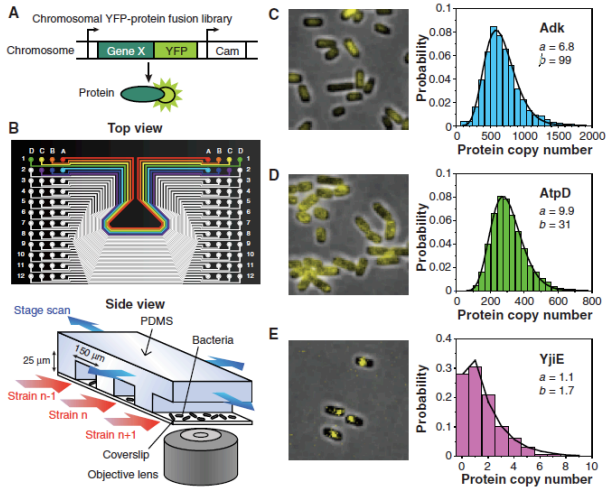
Multiple High-Throughput Analyses Monitor the Response of *E. coli* to Perturbations

Nobuyoshi Ishii,[1,2*] Kenji Nakahigashi,[1,2*] Tomoya Baba,[1,2,3*] Martin Robert,[1,2*] Tomoyoshi Soga,[1,2,6*] Akio Kanai,[1,2*] Takashi Hirasawa,[1,2*] Miki Naba,[1] Kenta Hirai,[1] Aminul Hoque,[1,2] Pei Yee Ho,[5] Yuji Kakazu,[1] Kaori Sugawara,[1] Saori Igarashi,[1] Satoshi Harada,[1] Takeshi Masuda,[1,2] Naoyuki Sugiyama,[6] Takashi Togashi,[1] Miki Hasegawa,[1] Yuki Takai,[1] Katsuyuki Yugi,[1,2] Kazuharu Arakawa,[1] Nayuta Iwata,[1,2] Yoshihiro Toya,[1,2] Yoichi Nakayama,[1,2] Takaaki Nishioka,[1,2,4] Kazuyuki Shimizu,[1,2,5] Hirotada Mori,[1,2,3] Masaru Tomita[1,2,6†]



Quantifying *E. coli* Proteome and Transcriptome with Single-Molecule Sensitivity in Single Cells

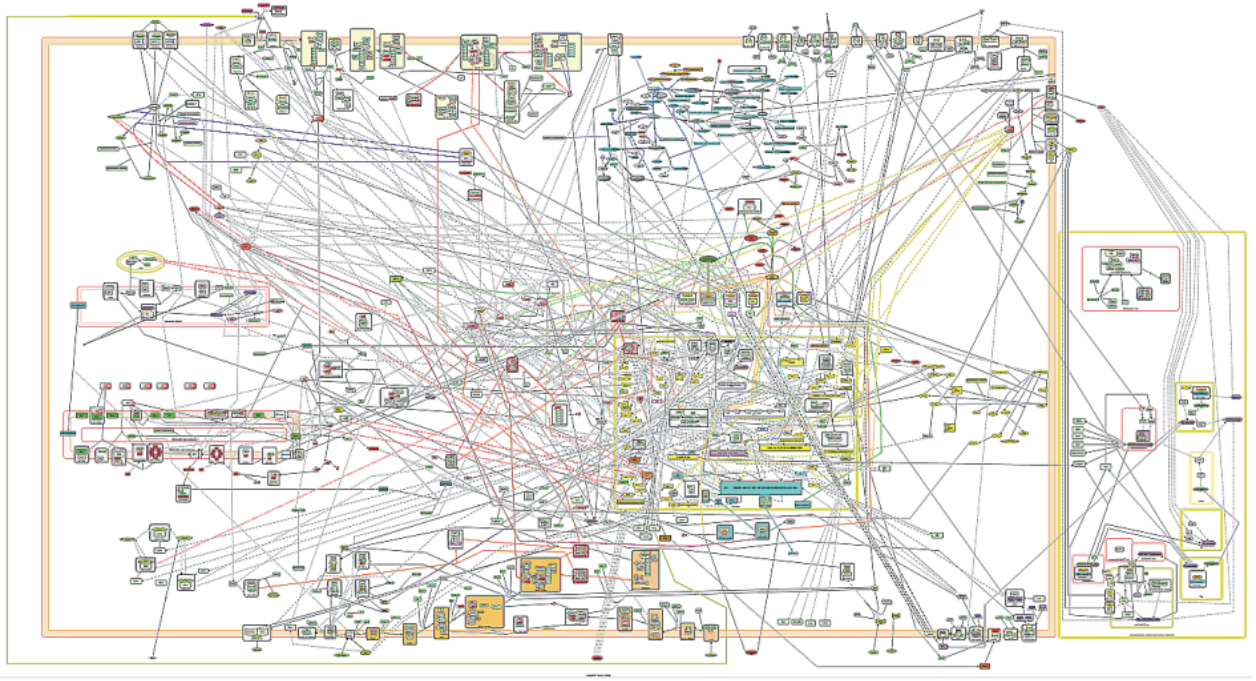
Yuichi Taniguchi,[1*] Paul J. Choi,[1*] Gene-Wei Li,[1,2*] Huiyi Chen,[1,3*] Mohan Babu,[4] Jeremy Hearn,[1] Andrew Emili,[4,5] X. Sunney Xie[1†]



.... although the progress in quantitative experimental approaches is astonishing.

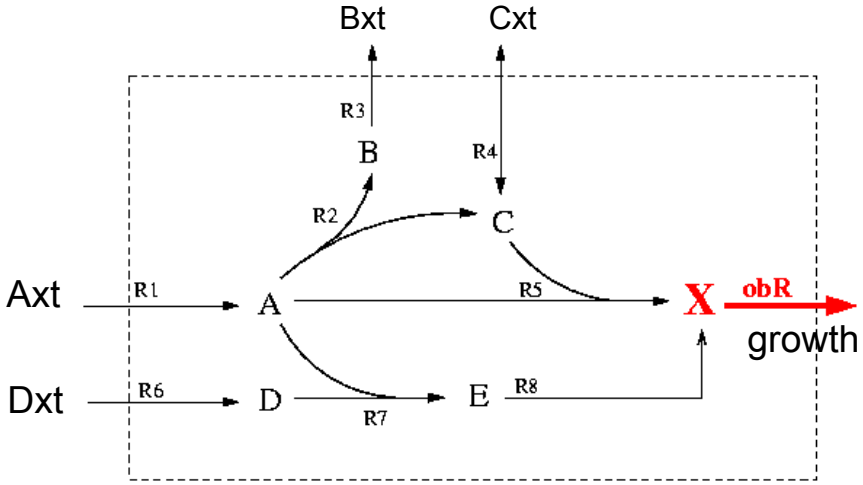
Should I wait until there is enough quantitative enough data to run dynamic simulations or should I look for approximate, qualitative methods to simulate reconstructions of molecular interaction networks with experimental data available today?

UNIVERSITY OF SURREY



The ideal solution would provide useful predictions from information about network connectivity alone while allowing gradual increase of quantitative detail by incorporation of quantitative data as they become available.

Flux Balance Analysis – good solution for metabolic network reconstructions.



Adapted from FluxAnalyzer software (Steffen Klamt,MPI Magdeburg)

The **linear programming** algorithm finds the largest possible value of dX/dt. However, there are many possible values of fluxes (F_1,..,F_8) that result in the same maximal value of objective function.

Analysis of steady state metabolic flux distributions is currently the only computer simulation method which can be used on genome scale models of molecular interaction networks of the cell.

Find **maximal** *dX/dt* if the following constraints are satisfied:

$$\frac{dX}{dt} = F_5 + F_8$$

Value to be maximised (**objective function**)

$$\frac{dAxt}{dt} = -F_1$$

$$\frac{dDxt}{dt} = -F_6$$

$$\frac{dBxt}{dt} = F_3$$

$$\frac{dCxt}{dt} = -F_4$$

Transport of extracellular (**external,** unbalanced) metabolites.

$$0 = F_1 - F_2 - F_5 - F_7$$

$$0 = F_2 - F_3$$

$$0 = F_2 + F_4 - F_5$$

$$0 = F_6 - F_7$$

$$0 = F_7 - F_8$$

Steady state (flux balance) assumption for intracellular (**internal**) metabolites.

$$0 < F_1 \leq 100$$

$$0 < F_2 \leq 100$$

$$0 < F_3 \leq 100$$

$$-100 < F_4 \leq 100$$

$$0 < F_5 \leq 100$$

$$0 < F_6 \leq 100$$

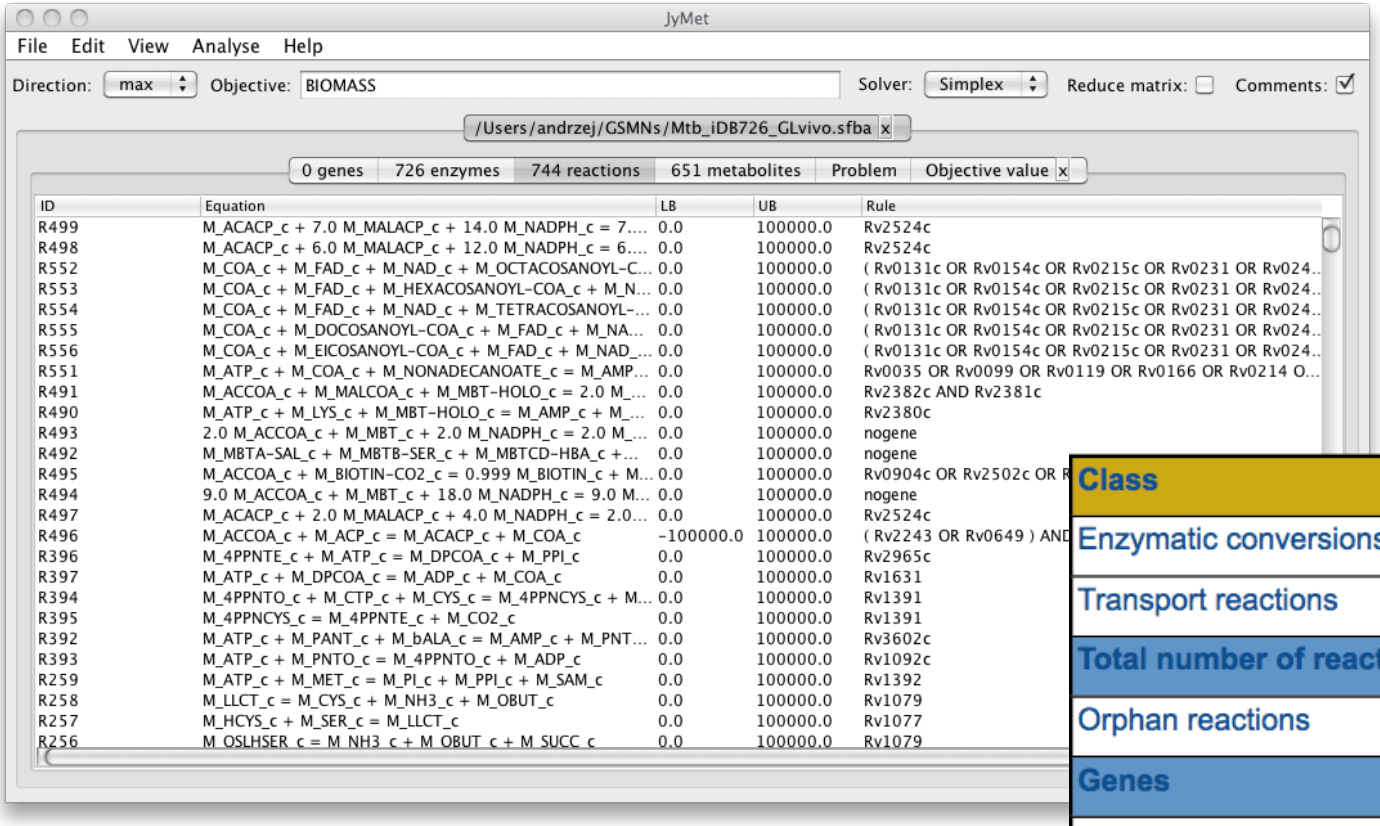
$$0 < F_7 \leq 100$$

$$0 < F_8 \leq 100$$

Minimal and maximal reaction capacities (**bounds**). R4 is the only **reversible** reaction in the system.

Flux Balance Analysis – good solution for metabolic network reconstructions.

UNIVERSITY OF SURREY



ID	Equation	LB	UB	Rule
R499	M_ACACP_c + 7.0 M_MALACP_c + 14.0 M_NADPH_c = 7....	0.0	100000.0	Rv2524c
R498	M_ACACP_c + 6.0 M_MALACP_c + 12.0 M_NADPH_c = 6....	0.0	100000.0	Rv2524c
R552	M_COA_c + M_FAD_c + M_NAD_c + M_OCTACOSANOYL-C...	0.0	100000.0	(Rv0131c OR Rv0154c OR Rv0215c OR Rv0231 OR Rv024..
R553	M_COA_c + M_FAD_c + M_HEXACOSANOYL-COA_c + M_N...	0.0	100000.0	(Rv0131c OR Rv0154c OR Rv0215c OR Rv0231 OR Rv024..
R554	M_COA_c + M_FAD_c + M_NAD_c + M_TETRACOSANOYL-...	0.0	100000.0	(Rv0131c OR Rv0154c OR Rv0215c OR Rv0231 OR Rv024..
R555	M_COA_c + M_DOCOSANOYL-COA_c + M_FAD_c + M_NA...	0.0	100000.0	(Rv0131c OR Rv0154c OR Rv0215c OR Rv0231 OR Rv024..
R556	M_COA_c + M_EICOSANOYL-COA_c + M_FAD_c + M_NAD_...	0.0	100000.0	(Rv0131c OR Rv0154c OR Rv0215c OR Rv0231 OR Rv024..
R551	M_ATP_c + M_COA_c + M_NONADECANOATE_c = M_AMP...	0.0	100000.0	Rv0035 OR Rv0099 OR Rv0119 OR Rv0166 OR Rv0214 O...
R491	M_ACCOA_c + M_MALCOA_c + M_MBT-HOLO_c = 2.0 M_...	0.0	100000.0	Rv2382c AND Rv2381c
R490	M_ATP_c + M_LYS_c + M_MBT-HOLO_c = M_AMP_c + M_...	0.0	100000.0	Rv2380c
R493	2.0 M_ACCOA_c + M_MBT_c + 2.0 M_NADPH_c = 2.0 M_...	0.0	100000.0	nogene
R492	M_MBTA-SAL_c + M_MBTB-SER_c + M_MBTCD-HBA_c +...	0.0	100000.0	nogene
R495	M_ACCOA_c + M_BIOTIN-CO2_c = 0.999 M_BIOTIN_c + M...	0.0	100000.0	Rv0904c OR Rv2502c OR R
R494	9.0 M_ACCOA_c + M_MBT_c + 18.0 M_NADPH_c = 9.0 M...	0.0	100000.0	nogene
R497	M_ACACP_c + 2.0 M_MALACP_c + 4.0 M_NADPH_c = 2.0...	0.0	100000.0	Rv2524c
R496	M_ACCOA_c + M_ACP_c = M_ACACP_c + M_COA_c	-100000.0	100000.0	(Rv2243 OR Rv0649) AND
R396	M_4PPNTE_c + M_ATP_c = M_DPCOA_c + M_PPI_c	0.0	100000.0	Rv2965c
R397	M_ATP_c + M_DPCOA_c = M_ADP_c + M_COA_c	0.0	100000.0	Rv1631
R394	M_4PPNTO_c + M_CTP_c + M_CYS_c = M_4PPNCYS_c + M...	0.0	100000.0	Rv1391
R395	M_4PPNCYS_c = M_4PPNTE_c + M_CO2_c	0.0	100000.0	Rv1391
R392	M_ATP_c + M_PANT_c + M_bALA_c = M_AMP_c + M_PNT...	0.0	100000.0	Rv3602c
R393	M_ATP_c + M_PNTO_c = M_4PPNTO_c + M_ADP_c	0.0	100000.0	Rv1092c
R259	M_ATP_c + M_MET_c = M_PI_c + M_PPI_c + M_SAM_c	0.0	100000.0	Rv1392
R258	M_LLCT_c = M_CYS_c + M_NH3_c + M_OBUT_c	0.0	100000.0	Rv1079
R257	M_HCYS_c + M_SER_c = M_LLCT_c	0.0	100000.0	Rv1077
R256	M_OSLHSER_c = M_NH3_c + M_OBUT_c + M_SUCC_c	0.0	100000.0	Rv1079

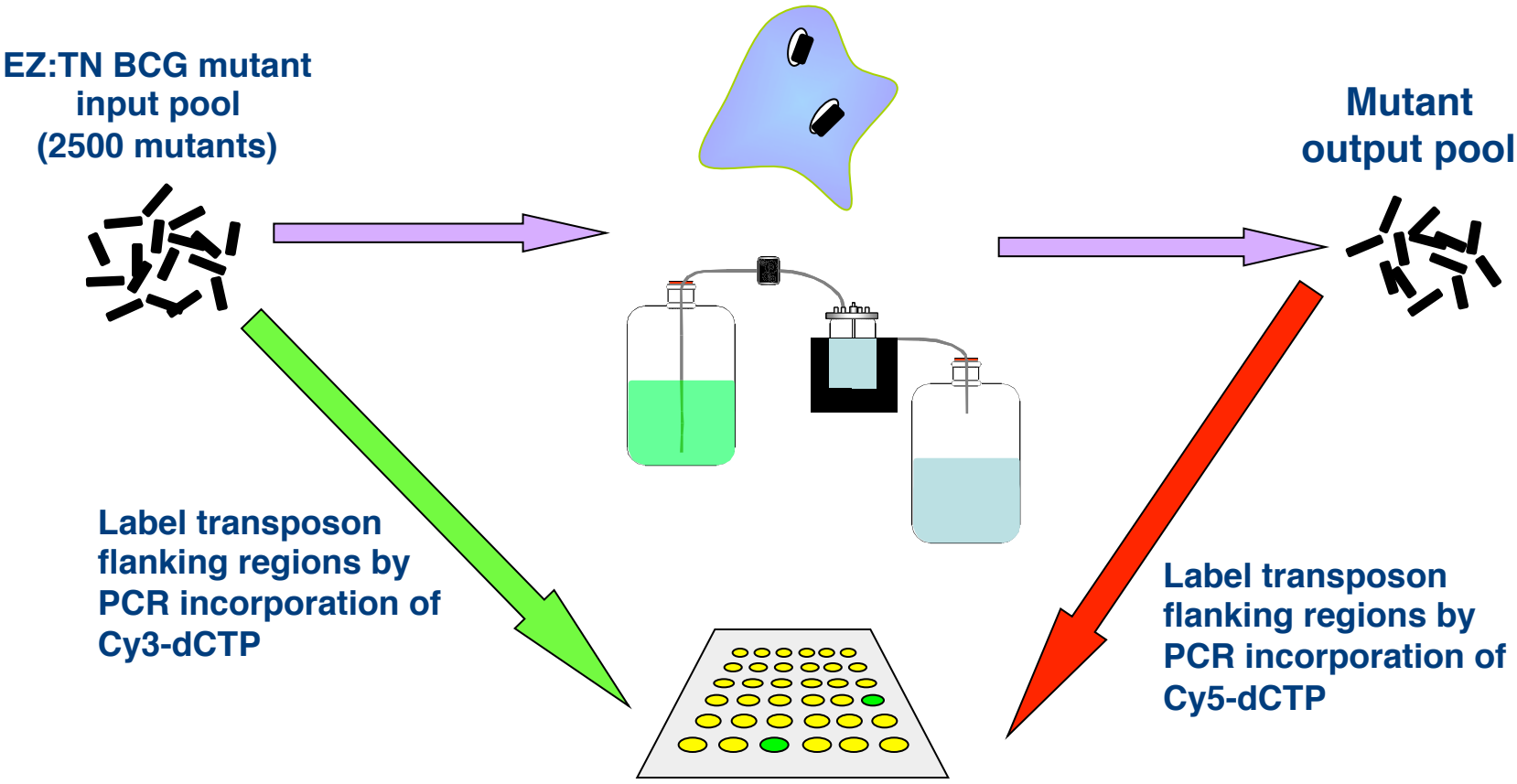
Class	Number
Enzymatic conversions	723
Transport reactions	126
Total number of reactions	**849**
Orphan reactions	210
Genes	**726**
Internal metabolites	638
External metabolites	101
Total number of metabolites	**739**

SurreyFBA software: Gevorgyan, Bushell, Avignone-Rossa, Kierzek, Bioinformatics, 2011

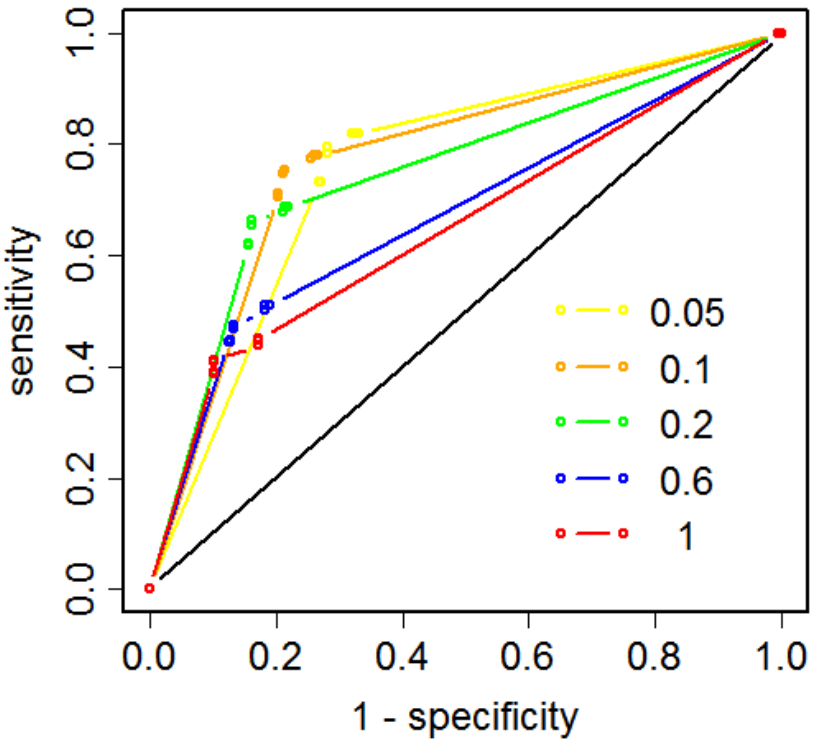
GSMN-TB model: Beste, Hooper, Stewart, Bonde, Avignone-Rossa, Bushell, Wheeler, Klamt, Kierzek, McFadden, Genome Biology 2007

Screening for essential genes by Transposon Site Hybridisation (TraSH)

UNIVERSITY OF SURREY



Receiver Operating Characteristics (ROC) of gene essentiality prediction.



Sensitivity = TP/(TP + FN)
Specificity = TN/(TN+FP)

Each ROC curve shows 100 points corresponding to sensitivity and specificity of the model predictions obtained for growth rate thresholds varying in the range from 0.0 to 0.1 (increment 0.001). The growth rate threshold has no effect on prediction accuracy.

The LP optimisation is effectively used as a qualitative test of BIOMASS producibility and it is irrelevant whether TB bacillus grows with maximal rate or not.

Different curves correspond to TraSH ratio thresholds of 0.05, 0.1, 0.2, 0.6, 1. The TraSH ratio cutoff has considerable influence on prediction accurracy.

The best ROC curve corresponds to the following prediction scores: **Sensitivity 71%, Specificity 80%, Correct predictions 78%.**

We did a lot of interesting work with Flux Balance Analysis in the areas of bacterial pathogens and biotechnology but ....

UNIVERSITY OF SURREY

BIOINFORMATICS APPLICATIONS NOTE Vol. 27 no. 3 2011, pages 433–434
doi:10.1093/bioinformatics/btq679

Systems biology

Advance Access publication December 9, 2010

SurreyFBA: a command line tool and graphics user interface for constraint-based modeling of genome-scale metabolic reaction networks

Albert Gevorgyan, Michael E. Bushell, Claudio Avignone-Rossa and Andrzej M. Kierzek*
Faculty of Health and Medical Sciences, University of Surrey, Guildford, Surrey GU2 7XH, UK

Associate Editor: Olga Troyanskaya

Metabolic Engineering 10 (2008) 227–233

Contents lists available at ScienceDirect

Metabolic Engineering

journal homepage: www.elsevier.com/locate/ymben

Selection of objective function in genome scale flux balance analysis for process feed development in antibiotic production

Chiraphan Khannapho [1], Hongjuan Zhao, Bhushan K. Bonde, Andrzej M. Kierzek, Claudio A. Avignone-Rossa, Michael E. Bushell *
Microbial Science Division, Faculty of Health and Medical Sciences, University of Surrey, Guildford, Surrey GU2 7XH, UK

OPEN ACCESS Freely available online PLoS COMPUTATIONAL BIOLOGY

Differential Producibility Analysis (DPA) of Transcriptomic Data with Metabolic Networks: Deconstructing the Metabolic Response of *M. tuberculosis*

Bhushan K. Bonde, Dany J. V. Beste, Emma Laing, Andrzej M. Kierzek, Johnjoe McFadden*
Microbial Sciences Division, Faculty of Health and Medical Sciences, University of Surrey, Guildford, United Kingdom

Mendum *et al. Genome Biology* 2011, **12**:R127
http://genomebiology.com/2012/12/12/R127

Genome Biology

RESEARCH Open Access

Interrogation of global mutagenesis data with a genome scale model of *Neisseria meningitidis* to assess gene fitness *in vitro* and in sera

Tom A Mendum, Jane Newcombe, Ahmad A Mannan, Andrzej M Kierzek and Johnjoe McFadden*

Research Open Access

GSMN-TB: a web-based genome-scale network model of *Mycobacterium tuberculosis* metabolism

Dany JV Beste[¤*], Tracy Hooper[¤*], Graham Stewart[*], Bhushan Bonde[*], Claudio Avignone-Rossa[*], Michael E Bushell[*], Paul Wheeler[†], Steffen Klamt[‡], Andrzej M Kierzek[¤*] and Johnjoe McFadden[¤*]

Addresses: *School of Biomedical and Molecular Sciences, University of Surrey, Stag Hill, Guildford, Surrey, GU2 7XH, UK. †Tuberculosis Research Group, Veterinary Laboratories Agency (Weybridge), New Haw, Addlestone KT15 3NB, UK. ‡Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstrasse, D-39106 Magdeburg, Germany.

¤ These authors contributed equally to this work.

Correspondence: Johnjoe McFadden. Email: j.mcfadden@surrey.ac.uk

Published: 23 May 2007
Genome Biology 2007, **8**:R89 (doi:10.1186/gb-2007-8-5-r89)

Received: 25 January 2007
Revised: 16 April 2007
Accepted: 23 May 2007

www.surrey.ac.uk

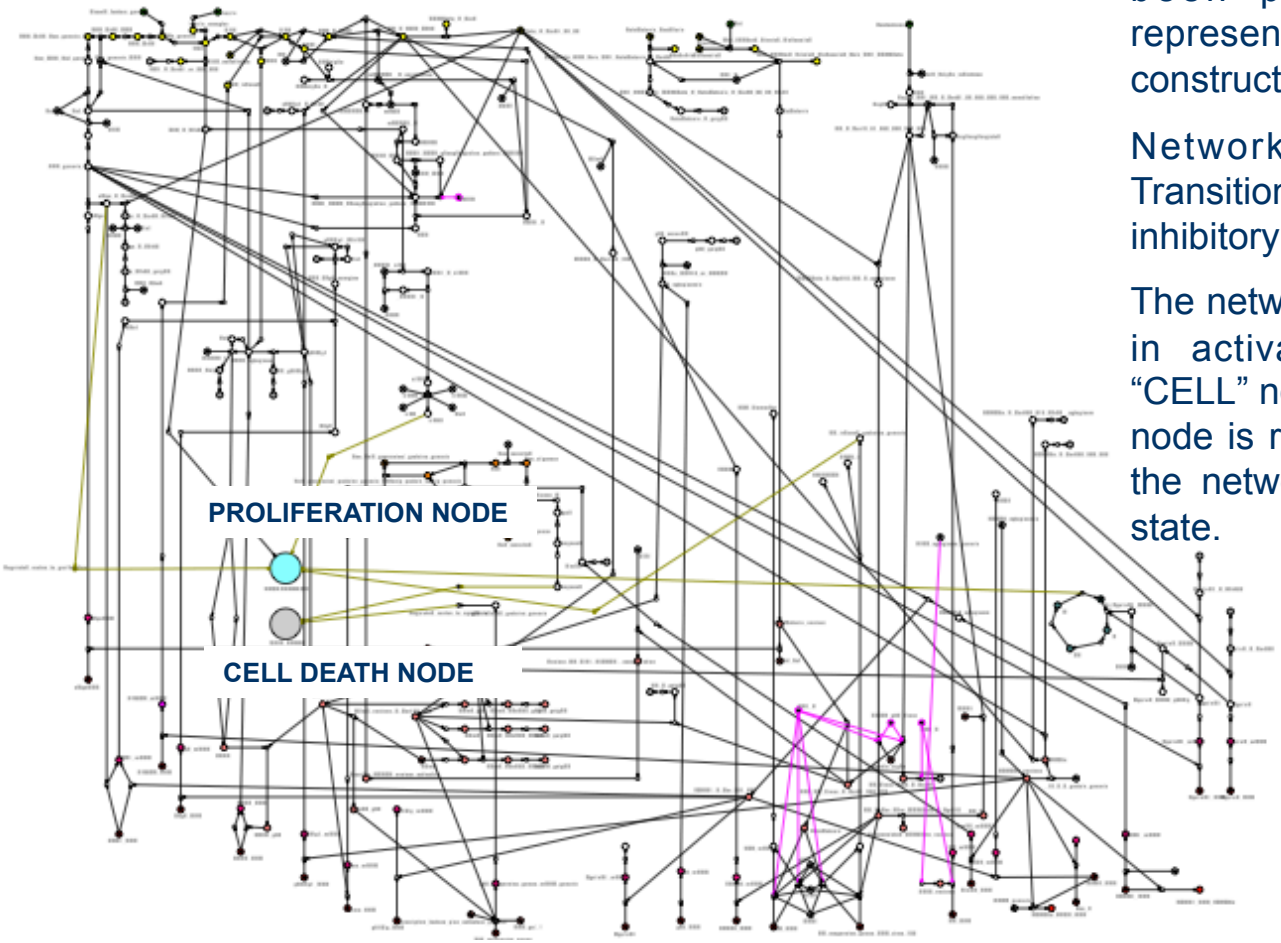
UNIVERSITY OF SURREY

... can we make qualitative predictions about genetic perturbations for general networks including dynamic regulatory processes for which steady state analysis is not useful?

Reconstruction of the signalling network involved in Prostate Cancer evolution.

UNIVERSITY OF SURREY

RECEPTORS



GENES

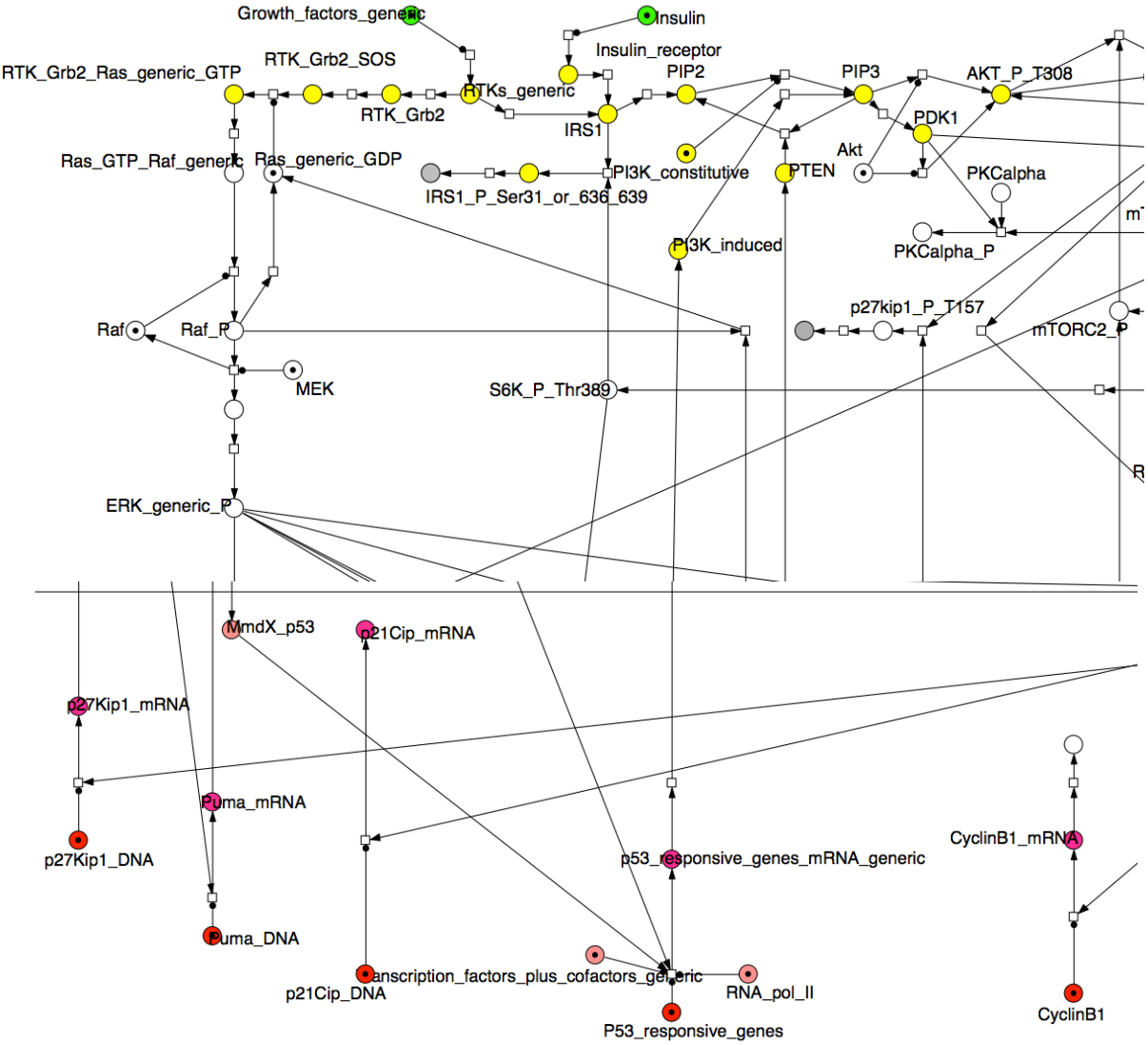
Extensive review of literature on molecular interactions involved in Prostate Cancer has been performed. Inteactions were represented as an Extended Petri Net constructed in Snoopy.

Network statistics: 251 Nodes, 195 Transitions, 420 Edges, 409 Read Edges, 0 inhibitory edges.

The network represents signal flow resulting in activation of “PROLIFERATION” or “CELL” node behaviour. If PROLIFERATION node is reached before CELL DEATH node the network is considered to enter cancer state.

Reconstruction of the signalling network involved in Prostate Cancer evolution.

UNIVERSITY OF SURREY



EXAMPLE RECEPTORS

EXAMPLE GENES

www.surrey.ac.uk

UNIVERSITY OF SURREY

Can we predict genetic and pharmacological perturbations influencing chances of proliferation before cell death without information on transition rate constants and molecular amounts?

Qualitative simulation approach.

UNIVERSITY OF SURREY

Discretise molecular activities. In all simulations the nodes were allowed to have 0, 1 or 2 tokens.

Make all transitions equally likely to fire. All transition rates were set to 1. All transitions for which pre-place nodes had more than 0 tokens were equally likely to fire.

Set initial marking of the network. Set marking of PROLIFERATION and CELL DEATH nodes to 0. Set marking of other nodes according to biological knowledge on activity of receptors and genes.

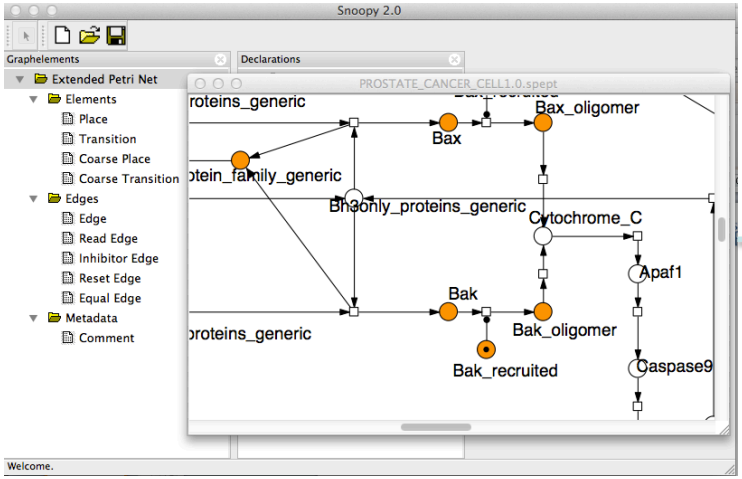
Generate ensemble of stochastic token games. Calculate F_{WT} – the fraction of trajectories in which PROLIFERATION node reached state of 1 before CELL DEATH node reached state of 1. The Gillespie algorithm simulation was used to generate token game trajectories.

Apply perturbation of interest. Gene knock-outs were simulated by setting the state of "DNA" node to 0. Increased degradation was simulated by setting transition rate to 1000.

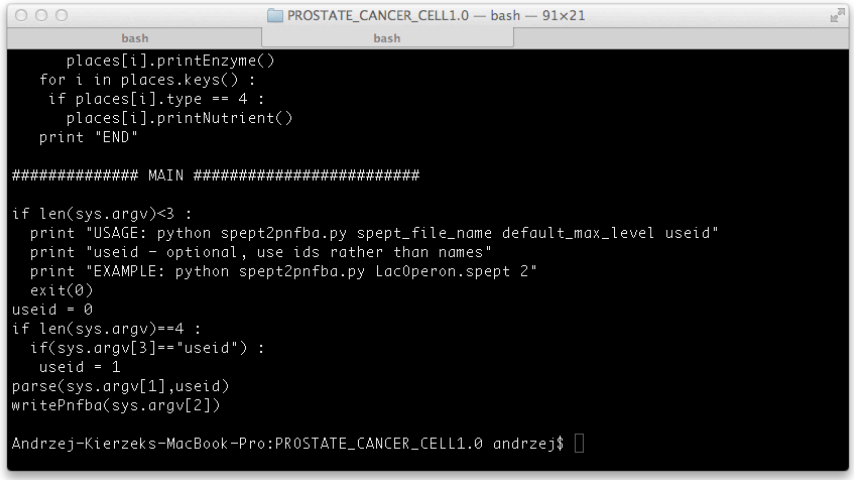
Generate ensemble of stochastic token games. Calculate F_P – the fraction of trajectories in which PROLIFERATION node reached state of 1 before CELL DEATH node reached state of 1.

Is F_P significantly different from F_{WT}? If yes, conclude that the perturbation influences the chances of prostate cancer evolution. Direction of change is meaningful, if $F_P > F_0$ the perturbation increases chances of cancer evolution.

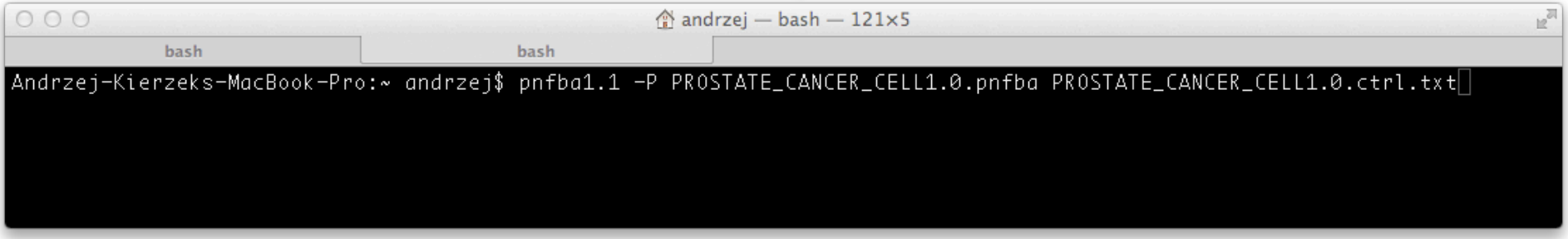
Implementation.



Snoopy

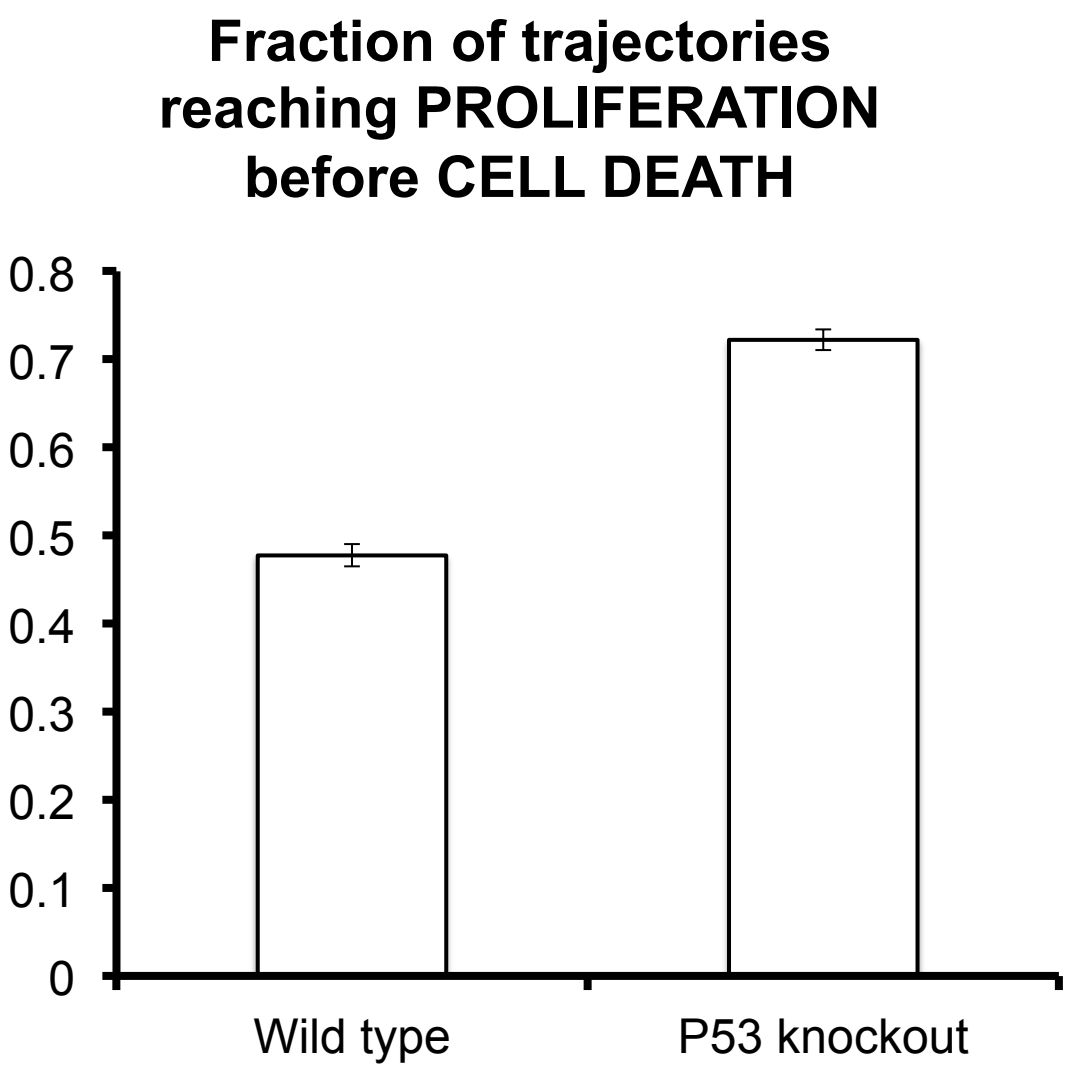


Python script reading *.spept file format.



New version of SurreyFBA software (manuscript in preparation).

Results for P53 gene knock-out.



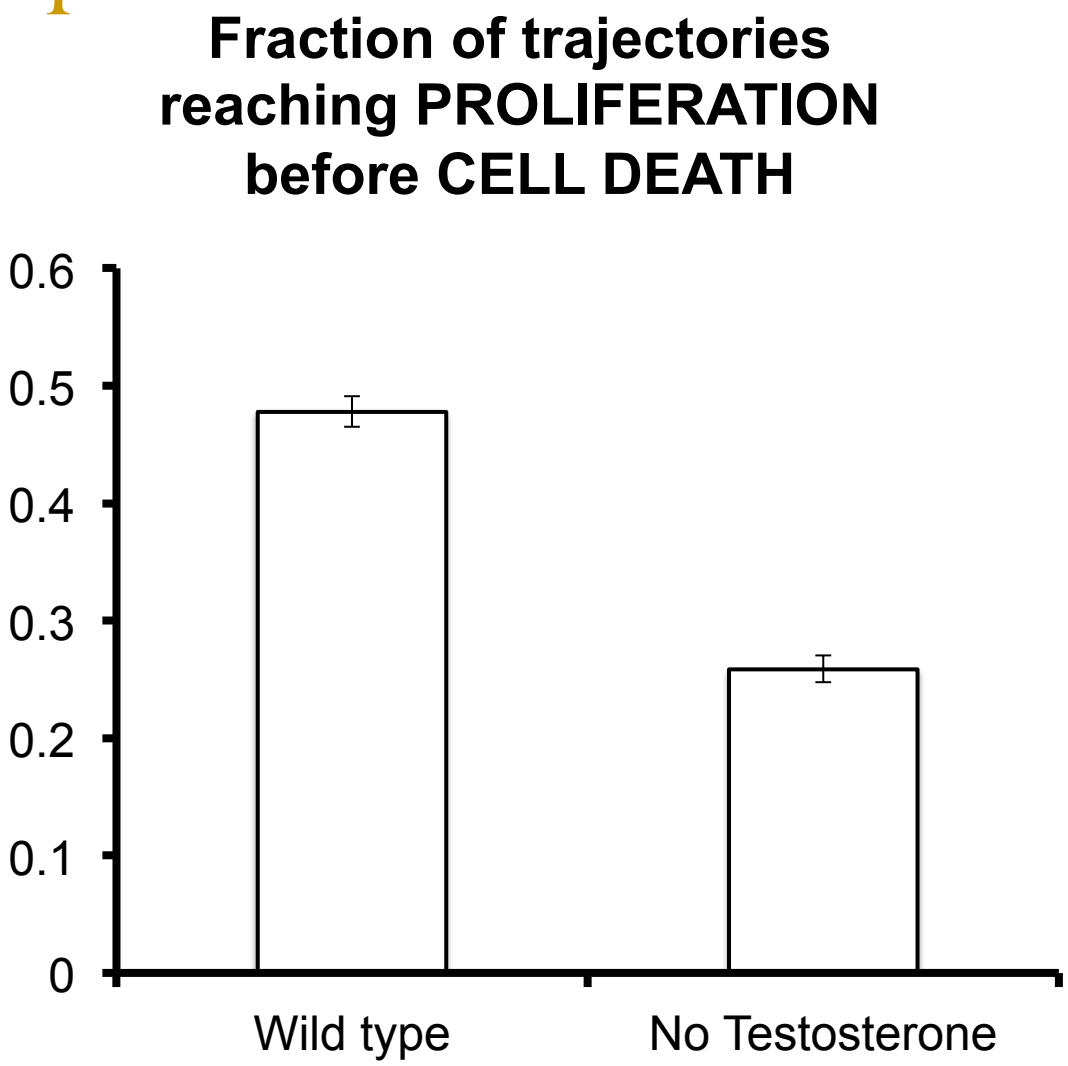
The 10,000 trajectories have been run for initial (Wild type) model. Each trajectory was run until CELL DEATH node changed state from 0 to 1 or the simulation time reached 100 arbitrary time units. The fraction of trajectories in which PROLIFERATION node changed state from 0 to 1 was calculated.

The same calculations have been performed for the model in which the state of p53_DNA node was set to 0.

The 99% binomial probability confidence intervals were calculated by binconf() function of Hmisc R package using Wilsons method.

This results is consistent with experimental data. The p53 gene is known as “guardian of the genome”. Its inactivation is associated with evolution of many types of cancer, including prostate cancer.

Results for no Testosterone input.



The 10,000 trajectories have been run for initial (Wild type) model. Each trajectory was run until CELL DEATH node changed state from 0 to 1 or the simulation time reached 100 arbitrary time units. The fraction of trajectories in which PROLIFERATION node changed state from 0 to 1 was calculated.

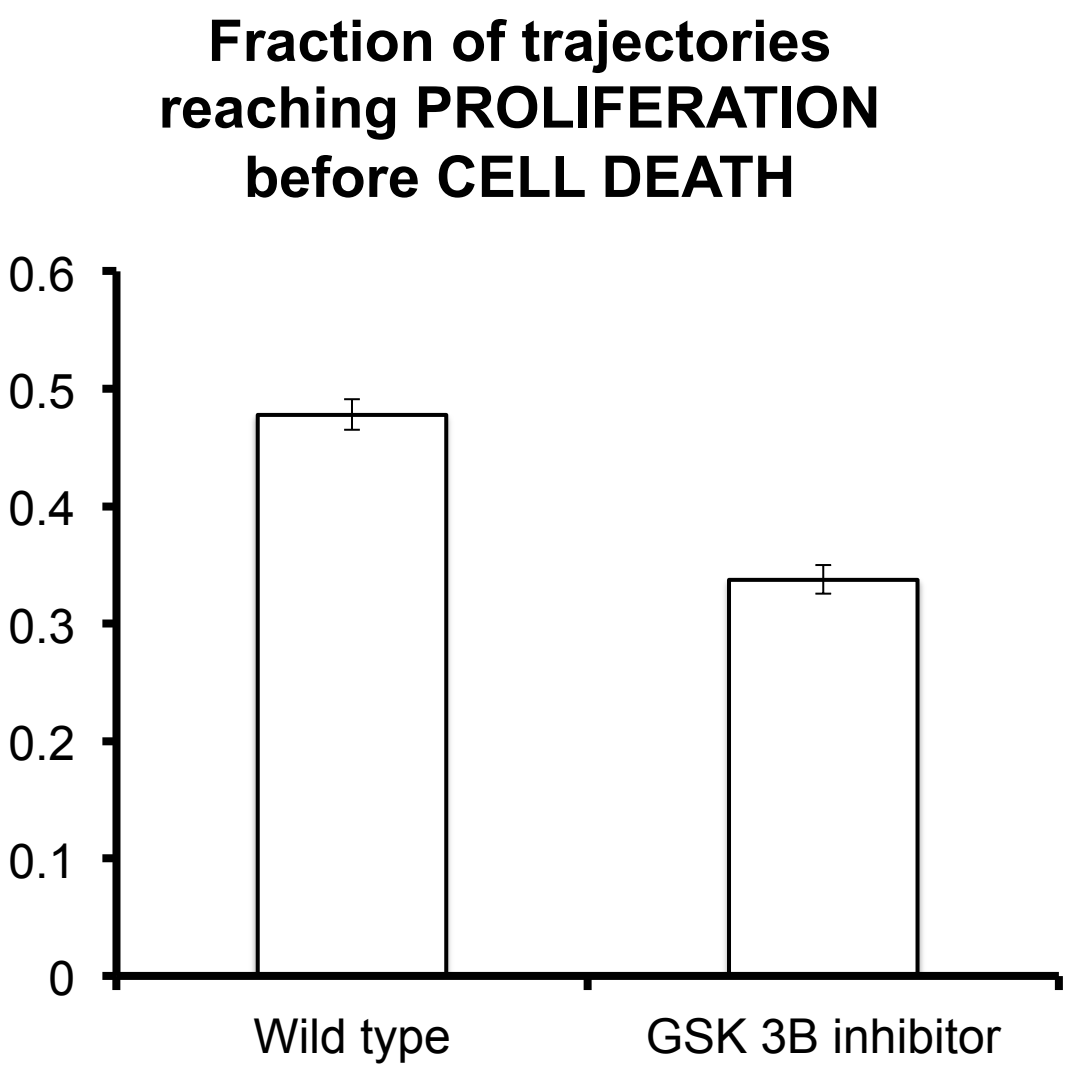
The same calculations have been performed for the model in which the state of "Testosterone" node was set to 0.

The 99% binomial probability confidence intervals were calculated by binconf() function of Hmisc R package using Wilsons method.

This results is consistent with experimental data.

Results for GSK 3B inhibitor.

UNIVERSITY OF SURREY



The 10,000 trajectories have been run for initial (Wild type) model. Each trajectory was run until CELL DEATH node changed state from 0 to 1 or the simulation time reached 100 arbitrary time units. The fraction of trajectories in which PROLIFERATION node changed state from 0 to 1 was calculated.

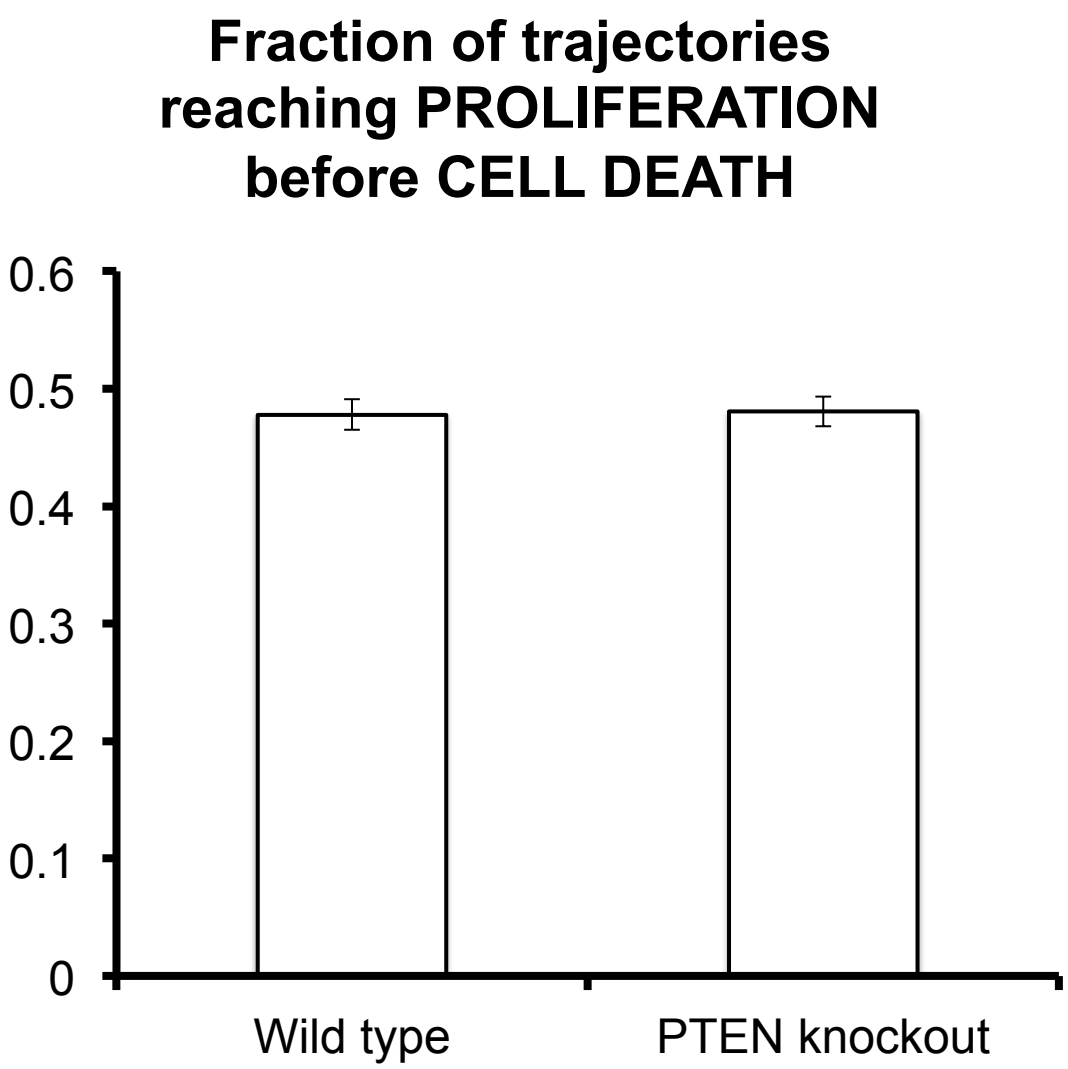
GSK 3B inhibitor is a drug that destabilises nuclear AR-GSK-3B. The rate of transition representing degradation of this molecule was set to 1000 and the same calculations as describe for "Wild type" model above were performed.

The 99% binomial probability confidence intervals were calculated by binconf() function of Hmisc R package using Wilsons method.

This results is consistent with experimental data. The GSK 3B inhibitor is a drug used in prostate cancer therapy.

Results for PTEN gene knockout.

UNIVERSITY OF SURREY



The 10,000 trajectories have been run for initial (Wild type) model. Each trajectory was run until CELL DEATH node changed state from 0 to 1 or the simulation time reached 100 arbitrary time units. The fraction of trajectories in which PROLIFERATION node changed state from 0 to 1 was calculated.

The same calculations have been performed for the model in which the state of PTEN_DNA node was set to 0.

The 99% binomial probability confidence intervals were calculated by binconf() function of Hmisc R package using Wilsons method.

This contradicts experimental data. The PTEN gene polymorphism is associated with prostate cancer evolution.

Discussion

UNIVERSITY OF SURREY

1. Application of Gillespie algorithm to generate sample of possible event sequences in qualitative Petri Net model is promising strategy for analysis of genome scale molecular interaction networks.
2. Related approach of “Signalling Petri Net” (PLoS Comput Biol 4(2): e1000005. doi:10.1371/journal.pcbi.1000005) implemented as PathwayOracle has been applied before to signalling networks. Our method is better suited for incorporation of existing qualitative knowledge about relative rates of biological processes (e.g. degradation rate increased by the drug). It is also simpler to integrate with existing model checking tools.
3. Results of this feasibility study suggest that it may be possible to gradually increase the level of detail of the molecular network reconstruction by incorporating quantitative information as it becomes available.

Acknowledgements.

OPCaRT - The Ovarian & Prostate Cancer Research Trust, Guildford, Surrey, UK
Sophie Chen

Computational systems biology group @ Surrey.
Ciaran Fisher (modelling & experiment), Ahmad Mannan, Carla Moller-Levet, Nicholas Stoy, Huihai Wu

Collaborating academics @ Surrey.
Claudio Avignone-Rossa, Dany Beste, Michael Bushell, Rebecca Hoyle, Andrzej M. Kierzek, David Lewis, Roberto La Ragione, Emma Laing, Johnjoe McFadden, Bernadette Moore, Nick Plant, Andrea Rocco, Colin P. Smith, Graham Stewart

International Collaborations.
Barry Wanner (Purdue), Hirotada Mori (Nara Institute of Technology), Kazuyuki Shimizu (Kyushu Institute of Technology), Steffen Klamt (Max Planck Institute), Olivier Neyrolles (CNRS), Ludovic Tailleux (Institute Pasteur), Maria Foti (University of Milan-Bicocca)

The work presented here and its continuation is supported by the following funding bodies.

www.surrey.ac.uk