



# Predictive biological models – an EGEE Grid application

MAX-PLANCK-INSTITUT  
FÜR MOLEKULARE GENETIK  
IHNESTRASSE 63-73

Ralf Herwig

Max Planck Institute for Molecular Genetics

Ihnestr. 73, 14195 Berlin

[herwig@molgen.mpg.de](mailto:herwig@molgen.mpg.de)

WERKSTÄTTEN  
NACHTPFORTE

# Talk structure

- 1. Introduction - the EMBRACE Network of Excellence**
- 2. Modelling human diseases**
- 3. Data on human interactions**
- 4. Modelling systems and methods**
- 5. Use case: Modelling the effect of drugs**
- 6. Conclusion & outlook**



# EMBRACE NoE

## Overall goal:

EMBRACE is a EU-sponsored Network of Excellence aimed at enabling bioinformatics research through better operability of databases, servers, and services.



# EMBRACE – project facts

Coordination: EBI Hinxton (Graham Cameron)

Duration: 01.02.2005-31.01.2010

Partners: 18

Budget: 8,000,000 EUR

Homepage:

<http://www.embracegrid.info>

Work packages:

Content integration

Tool integration

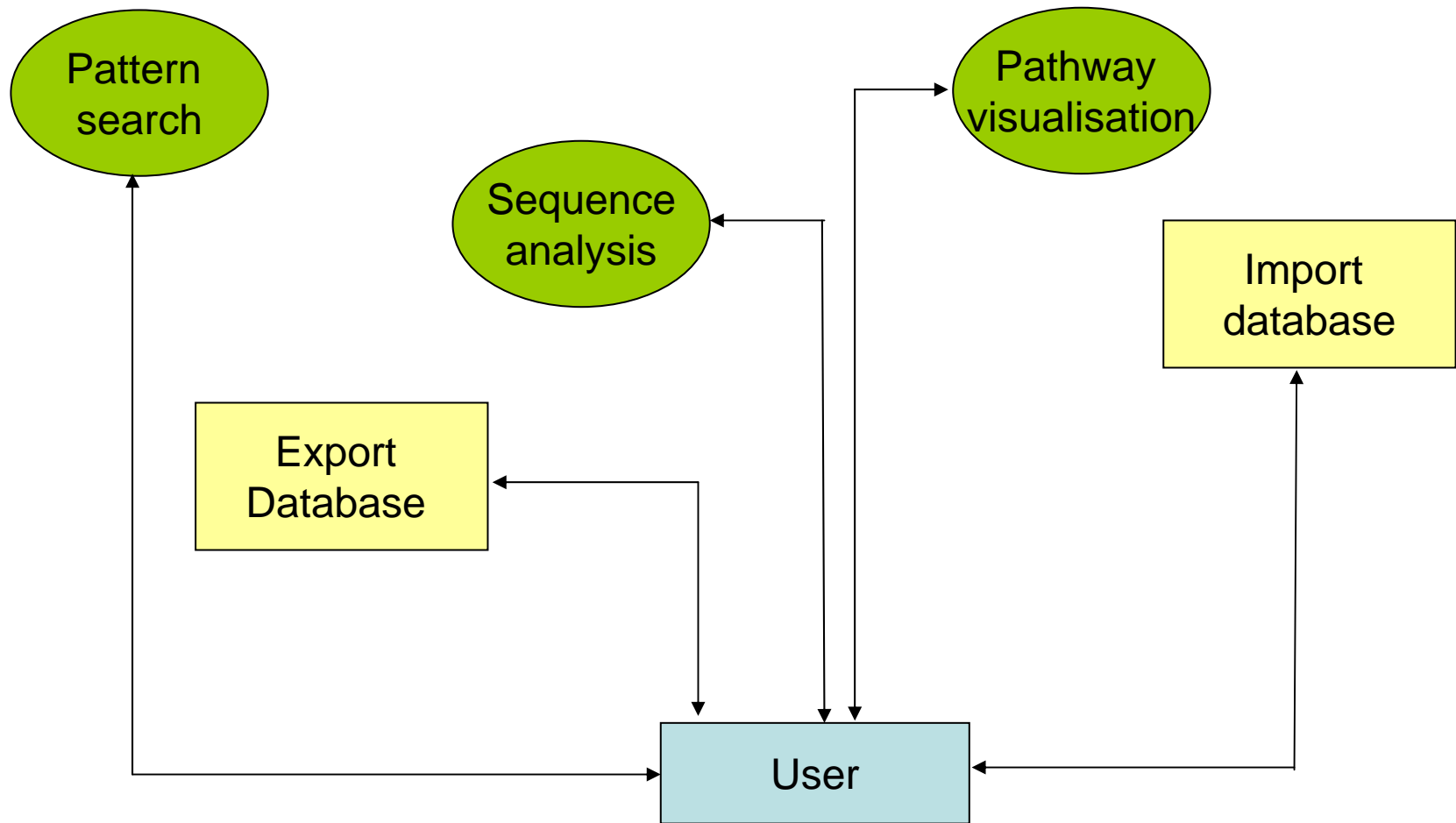
Test cases

Outreach

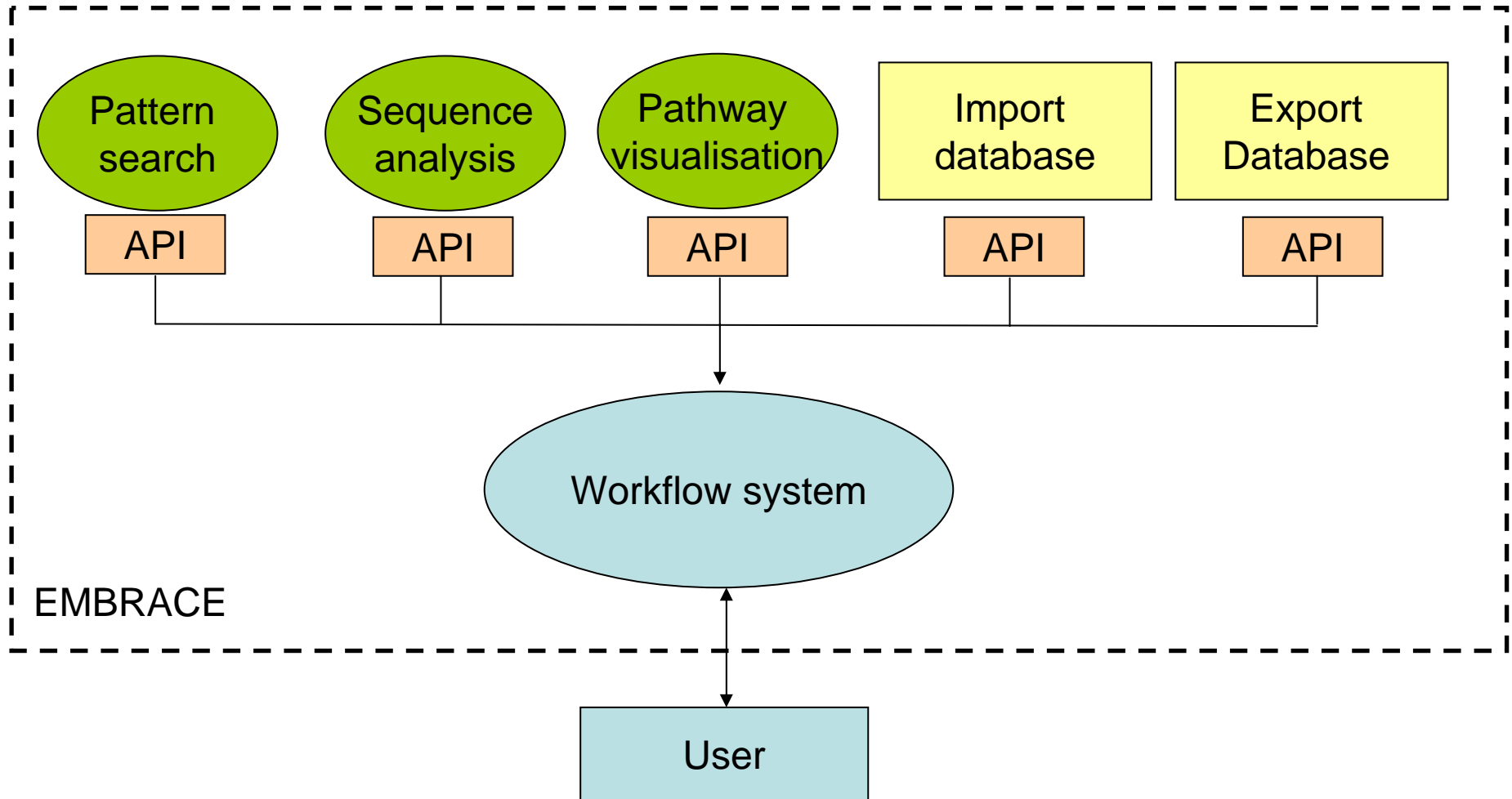
Particip. Number	Participant name	Participant short name
1	European Molecular Biology Laboratory – European Bioinformatics Institute	EMBL-EBI
2	Istituto di Tecnologie Biomediche – Sezione di Bari, CNR	ITB-BA
3	University of Manchester	UMAN
4	Swiss Institute of Bioinformatics	SIB
6	Swedish University of Agricultural Sciences Linnaeus Centre for Bioinformatics	SLU-LCB
7	Centre National de la Recherche Scientifique	CNRS
8	Technical University of Denmark	CBS-DTU
9	Consejo Superior de Investigaciones Cientificas	CSIC
10	Stockholms Universitet	SU
11	Institut National de la Recherche Agronomique	INRA
12	Max-Planck-Society, Max-Planck- Institute for Molecular Genetics	MPI-MG
13	CSC - Scientific Computing Ltd	CSC
14	University College London	UCL
15	Weizmann Institute of Science	WIS
16	Stichting Katholieke Universiteit	KUN-CMBI
17	Instituto Nacional de Tecnica Aeroespacial (Centre de Astrobiologia)	INTA-CAB
18	University of Bergen	CBU



# User interaction – current situation



# User interaction – the EMBRACE vision



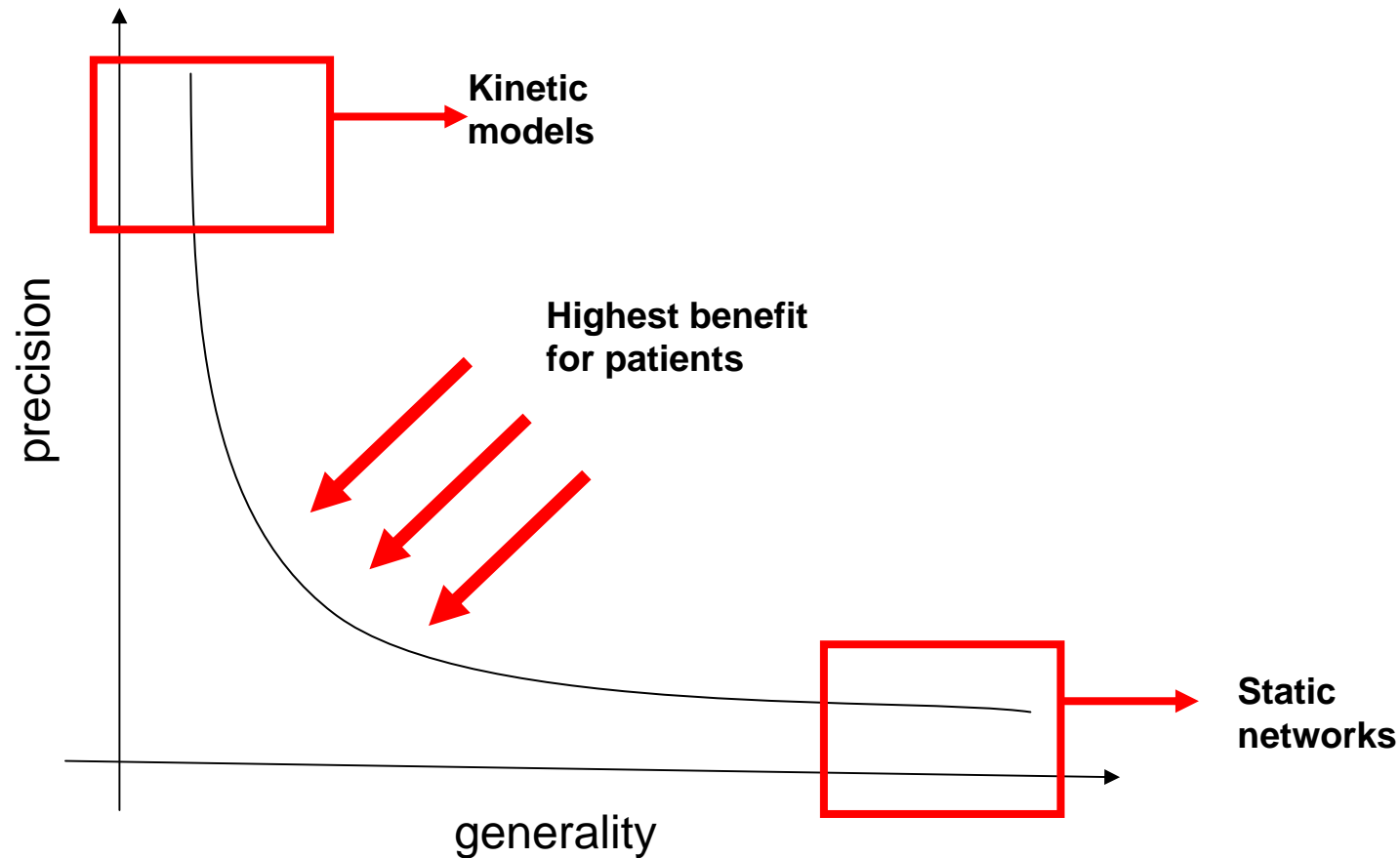
# Talk structure

1. Introduction - the EMBRACE Network of Excellence
2. **Modelling human diseases**
3. Data on human interactions
4. Modelling systems and methods
5. Use case: Modelling the effect of drugs
6. Conclusion & outlook



# Why disease models ?

Trade-off between model generality and model precision



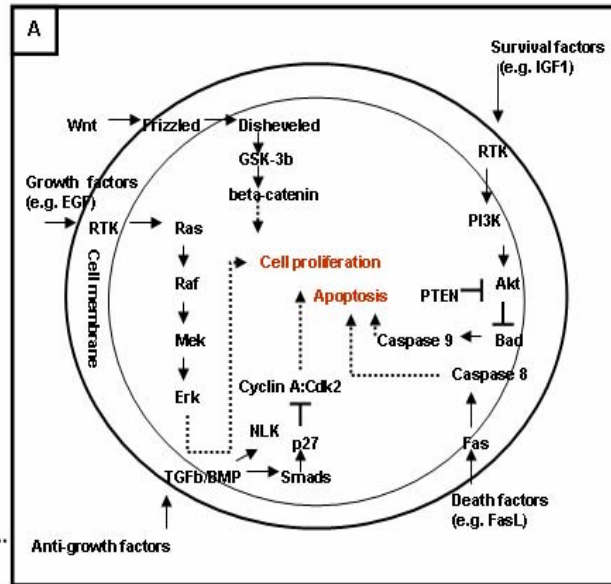


# Disease models requirements

- **Inclusion of the right model components**
- **„Minimal“ (but sufficient) coverage of the disease domain**
- **Kinetic analysis trained to phenotype**
- **Identification of the essential read-outs**



# MPIMG cancer network



**B**

Description View Reactions Network Simulation Population Analysis Export/Import

**Reactions**

Number of reactions: 1186

GR2Cancer\_361726 (SM7:SM9:4:SM9:4:TAK1:P:TAB1:TAB2-P [cytosol]) irreversible kinetic Pathways

1 NLK [cytosol] + 1 ATP [cytosol] → 1 phospho-NLK [cytosol] + 1 ADP [cytosol]

GR2Cancer\_362187 (phospho-RS-PDK [cytosol]) irreversible kinetic Pathways

1 ATP [cytosol] + 1 Phosphatidylnicotinamide 4,5-bisphosphate [plasma membrane] → 1 ADP [cytosol] + 1 Phosphatidylnicotinamide 3,4,5-trisphosphate [plasma membrane]

GR2Cancer\_362311 (PTEN [cytosol]) irreversible kinetic Pathways

1 H2O [cytosol] + 1 Phosphatidylnicotinamide 3,4,5-trisphosphate [plasma membrane] → 1 Orthophosphate [cytosol] + 1 Phosphatidylnicotinamide 4,5-bisphosphate [plasma membrane]

GR2Cancer\_361789 (phospho-NLK [nucleoplasm]) irreversible kinetic Pathways

1 TOF-4 [nucleoplasm] + 1 NTP [nucleoplasm] → 1 Phospho-TOF-4 [nucleoplasm] + 1 NTP [nucleoplasm]

GR2Cancer\_361678 (SM7:SM9:4:SM9:4:TAK1:P:TAB1:TAB2 [cytosol]) irreversible kinetic Pathways

1 SM7:SM9:4:SM9:4:TAK1:P:TAB1:TAB2 [cytosol] + 2 ATP [cytosol] → 1 SM7:SM9:4:SM9:4:TAK1:P:TAB1:TAB2-P [cytosol] + 2 ADP [cytosol]

GR2Cancer\_362221 (SHC-activated IGF-1R receptor [plasma membrane]) irreversible kinetic Pathways

1 ATP [cytosol] + 1 SHC-activated IGF-1R receptor [plasma membrane] → 1 ADP [cytosol] + 1 phospho-SHC-activated IGF-1R [cytosol]

Network components	
Reactions	1913
Pathways	20
Rb/E2F pathway	
DNA Damage Checkpoints	
DNA repair	
IGF-1 signaling	
Signaling by EGFR	
NGFR signalling	
TGFbeta signalling	
BMP signaling	
Hedgehog signaling	
MAP kinase cascade	
Extrinsic apoptosis	
Intrinsic apoptosis	
Toll Like Receptor 10 (TLR10) Cascade	
Toll Like Receptor 3 (TLR3) Cascade	
Cytokine Signaling/ JAK/STATsignaling	
Wnt signaling	
E-cadherin pathway	
PLC signalling	
Glucagon signaling	
Proteins	326
Mutated genes	8*
Druggable genes	18**
Complexes	354

Reaction	Biochemical equation	Rate law
Synthesis	$v_0 : \rightarrow S_2$	$v_0 = k_0$
Complex formation	$v_1 : S_1 + S_2 \leftrightarrow S_1 : S_2$	$v_1 = kf[S_1][S_2] - (kf/K_D)[S_1 : S_2]$ with $K_D=100$
Formation of product	$v_2 : S_1 \rightarrow S_2$	$v_2 = kI[S_1]$
Degradation	$v_3 : S_2 \rightarrow$	$v_3 = k_3 S_2$ with $k=10^{-3} (^{\circ})$

Wierling et al., under review.

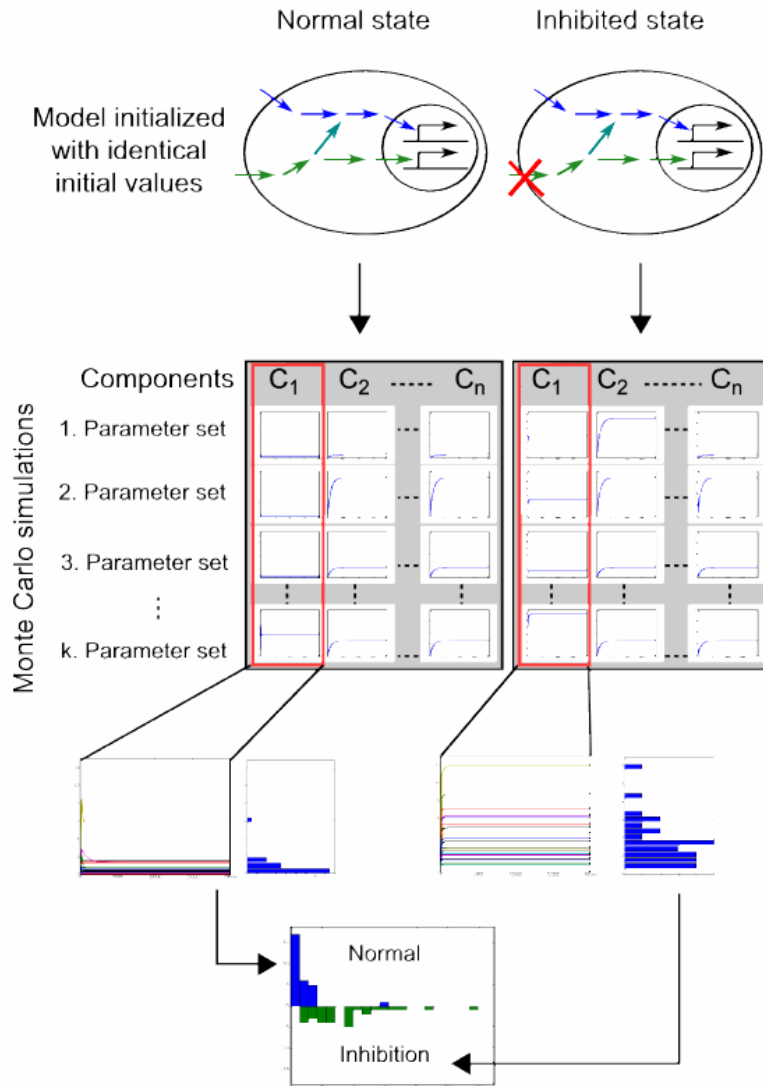
EGEE 2008 Conference, Istanbul 22.09.2008



Max-Planck-Institute for Molecular Genetics



# Modelling strategy and the grid



- annotate reaction systems
- add kinetic information
- compare two states
- random sampling of kinetic parameters
- steady state simulation
- statistical hypothesis testing of steady state ratios

# Talk structure

1. Introduction - the EMBRACE Network of Excellence
2. Modelling human diseases
3. **Data on human interactions**
4. Modelling systems and methods
5. Use case: Modelling the effect of drugs
6. Conclusion & outlook



# Sources of information on human interactions

## CPDB purpose and overview

**Problem:** Diverse and heterogeneous pathway annotation

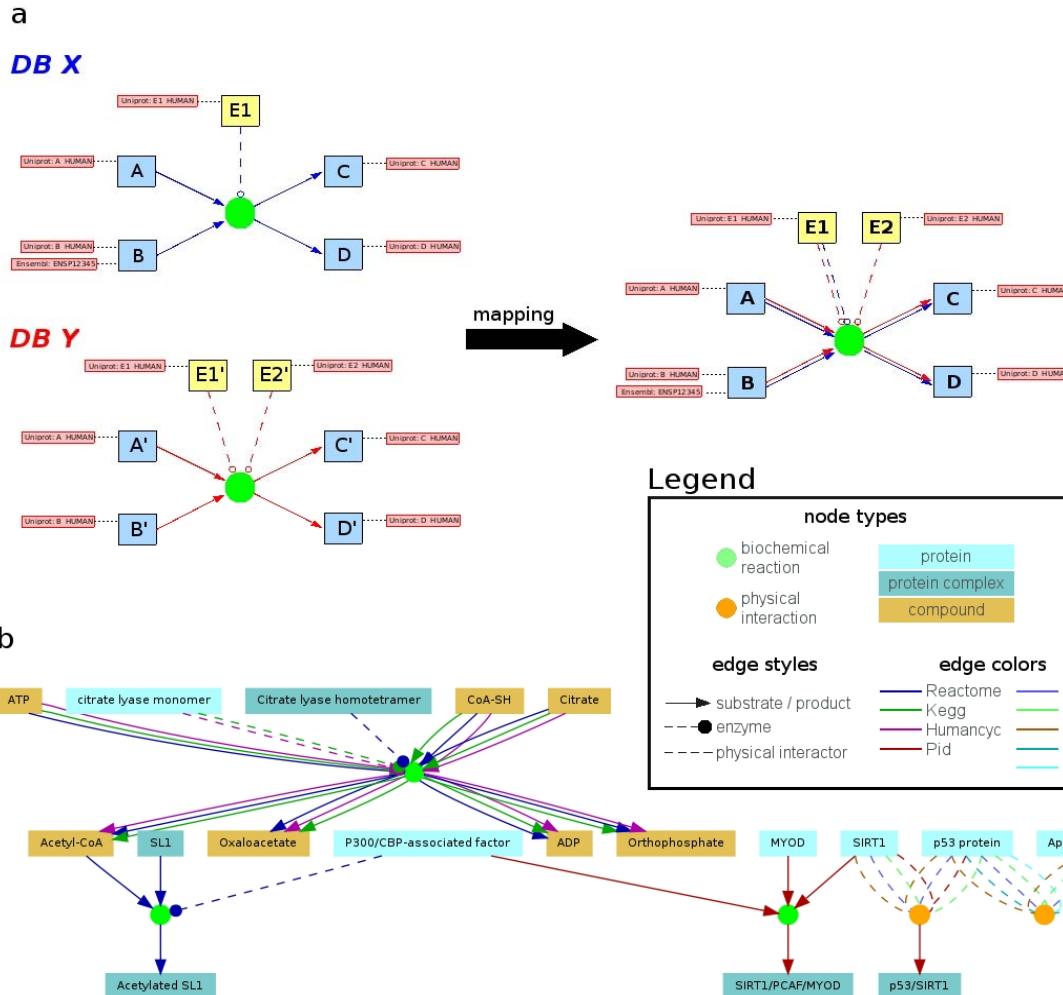
- Contains metabolic reactions, signal transduction events and gene regulatory interactions integrated in a common schema
- Data comes from different sources. Currently, 12 databases (Reactome, KEGG, HumanCyc, PID, BioCarta, IntAct, DIP, SPIKE etc.) and a small manually uploaded dataset. Detection and reduction of redundant information
- Web service with rich functionality:
  - Dynamically generated database content and coverage statistics
  - Search for specific interactions of molecules or pathways
  - Visualization of the results in **expandable** network-graphs
  - Over-representation analysis
  - Gene annotation to pathways
  - Model upload (SBML, PSI-MI, BioPAX)
  - Model download (BioPAX)

Kamburov (2007) Master Thesis, Free University Berlin.  
Kamburov et al., under review.



# ConsensusPathDB – mapping of interactions

## Integration of human functional interactions



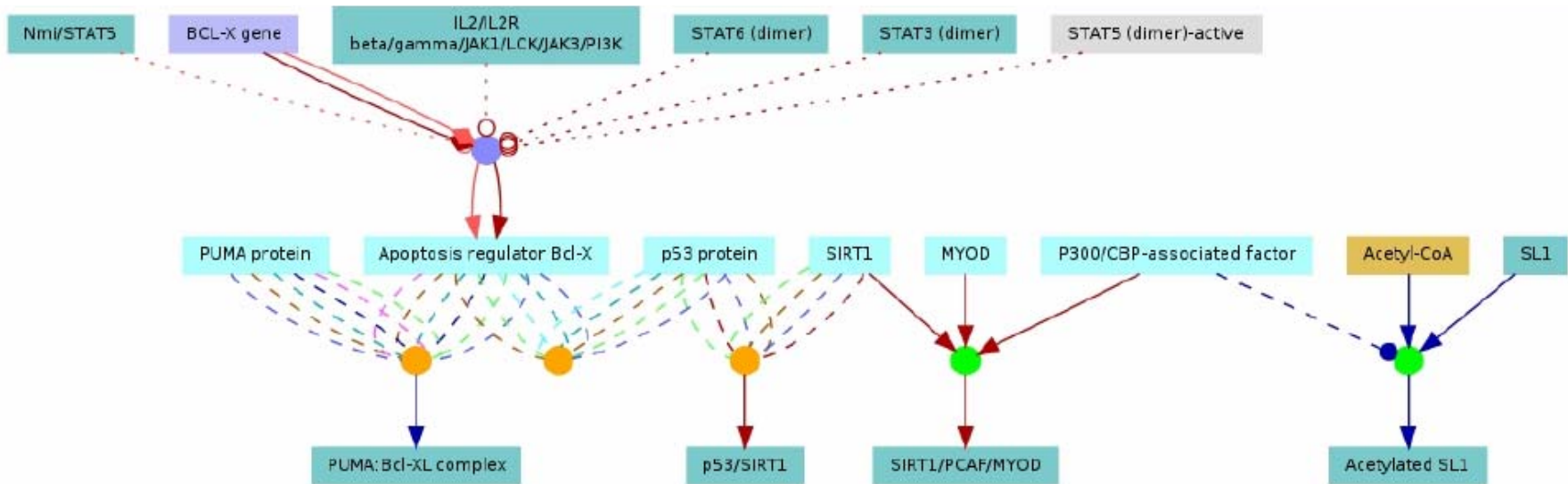
- mapping is based on identifier matching

- different layers are indicated by different node colors

- interactions are coded as edges

- origin of interaction is indicated by different edge colors

# Network visualisation



gene	— Reactome
protein	— Kegg
protein complex	— Pid
compound	— Netpath
group / unknown	— Dip
	— Hprd
	— Spike
● biochemical reaction	— Humancyc
● gene regulation	— Biocarta
● physical interaction	— Intact
	— Mint
	— Biogrid

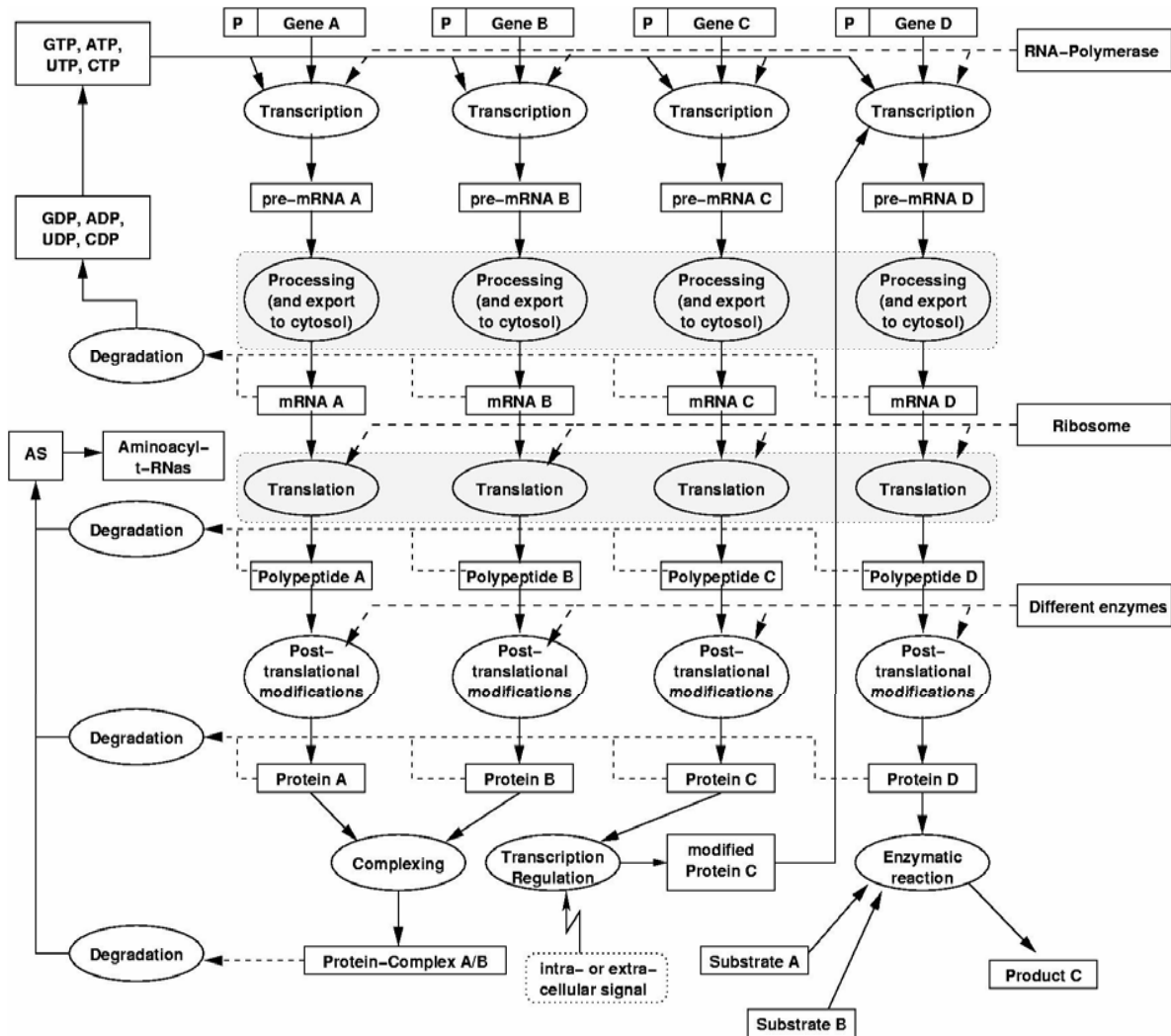
# Talk structure

1. Introduction - the EMBRACE Network of Excellence
2. Modelling human diseases
3. Data on human interactions
4. Modelling systems and methods
5. Use case: Modelling the effect of drugs
6. Conclusion & outlook





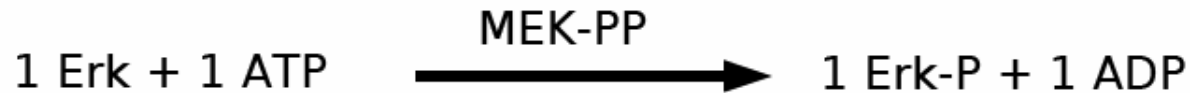
# PyBioS modelling and simulation tools



- object-oriented structure
- data structure covers main biological objects
- actions attached to objects mimic biological processes

# Reactions and kinetics

## Reaction equations



## Rate laws

$$v_1 = k_1 \times [\text{Erk}] \times [\text{ATP}] \times [\text{MEK-PP}]$$

## Kinetic parameters

$$k_1 = 1.2 \text{ mmol}^{-2} \cdot \text{min}^{-1}$$

## Initial concentrations

$$\begin{aligned} [\text{Erk}] &= 0.1 \text{ mM} & [\text{Erk-P}] &= 0.0 \text{ mM} & [\text{MEK-PP}] &= 0.05 \text{ mM} \\ [\text{ATP}] &= 3.0 \text{ mM} & [\text{ADP}] &= 0.0 \text{ mM} \end{aligned}$$

## Differential equations

$$\frac{d[\text{Erk}]}{dt} = -v_1$$

$$\frac{d[\text{Erk-P}]}{dt} = v_1$$

$$\frac{d[\text{ATP}]}{dt} = -v_1$$

$$\frac{d[\text{ADP}]}{dt} = v_1$$

# Web interface

*PyBioS* (<http://pybios.molgen.mpg.de>)

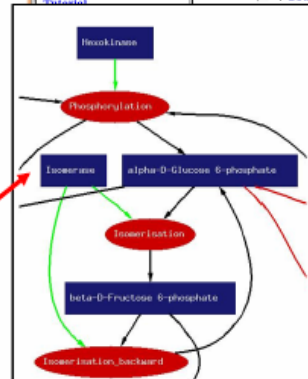
Detailed hierarchical model

Diagram of simulation results

Model repository

The screenshot shows the PyBioS web interface in a Mozilla browser. The address bar shows <http://pybios.molgen.mpg.de/biolab>. The main content area is titled "Simulation Environment at /biolab/chris2/Hyene". On the left, there is a "Model repository" showing a tree of available models: cell (Cell.cell), CytoplasmMembrane (Compartment.CytoplasmMembrane), Cytoplasm (Compartment.Cytoplasm), Cytosol (Compartment.Cytosol), alpha-D-Glucose (Compound.C00267), alpha-D-Glucose 6-phosphate (Compound.C00668), beta-D-Fructose 6-phosphate (Compound.C05345), ATP (Compound.C00002), ADP (Compound.C00008), AMP (Compound.C00020), and beta-D-Fructose 1,6-bisphosphate (Compound.C05378). On the right, a graph shows the concentration of various species over time (0 to 100). The legend indicates: alpha-D-Glucose (blue), alpha-D-Glucose 6-phosphate (red), beta-D-Fructose 6-phosphate (green), beta-D-Fructose 1,6-bisphosphate (purple), and ATP (yellow).

List of reactions



**Reactions**

**Phosphorylation** (Hexokinase): Irreversible kinetic  
 $1 \text{ alpha-D-Glucose} + 1 \text{ ATP} \rightarrow 1 \text{ alpha-D-Glucose 6-phosphate} + 1 \text{ ADP}$

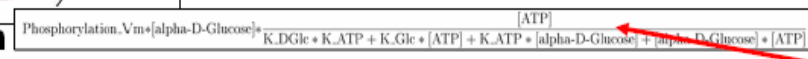
**Isomerisation backward** (Isomerase): Irreversible kinetic  
 $1 \text{ beta-D-Fructose 6-phosphate} \rightarrow 1 \text{ alpha-D-Glucose 6-phosphate}$

**Isomerisation** (Isomerase): Irreversible kinetic  
 $1 \text{ alpha-D-Glucose 6-phosphate} \rightarrow 1 \text{ beta-D-Fructose 6-phosphate}$

**Phosphorylation** (Phosphofruktokinase): Irreversible kinetic  
 $1 \text{ beta-D-Fructose 6-phosphate} + 1 \text{ ATP} \rightarrow 1 \text{ beta-D-Fructose 1,6-bisphosphate} + 1 \text{ ADP}$

Auto-generated interaction diagram of the model

Details of the kinetics



# Automated network upload

(A) Search-Interface

(B) According results

(C) Networkgraph of the selected reactions

(D) Details of selected reactions

(E) Detailed listing of the reactions dedicated for population

(A) search network reaction

Search term: glycolysis

Search accuracy: low medium high

Database: Trnspath, KEGG, **Reactome**, SPS, Kintikan

Datatype:  compound or gene name  pathway

search again

Hits from the Reactome database, that are similar to your search term:

- Glycolysis (Homo sapiens)
- Glycolysis (Mus musculus)
- Glycolysis (Rattus norvegicus)
- Glycolysis (Gallus gallus)
- Glycolysis (Fugu rubripes)
- Glycolysis (Danio rerio)

Find according reactions in the database

The following reactions will be populated: (E)

1 D-Fructose 1,6-bisphosphate [cytosol]	->	1 Dihydroxyacetone phosphate [cytosol] + 1 D-Glycerate 3-phosphate [cytosol]
actionid: Reactome_70458	enzyme/gene: fructose-bisphosphate aldolase C, class I haloenzyme	compartment: cytosol kinetic: massi <a href="#">more</a>
1 Dihydroxyacetone phosphate [cytosol]	->	1 D-Glyceraldehyde 3-phosphate [cytosol]
actionid: Reactome_70464	enzyme/gene: triosephosphate isomerase dimer [cytosol]	compartment: cytosol kinetic: massi <a href="#">more</a>
1 D-Glyceraldehyde 3-phosphate [cytosol] + 1 Orthophosphate [cytosol] + 1 NAD+ [cytosol]	->	1 3-Phospho-D-glyceroyl phosphate [cytosol] + 1 NADH [cytosol] + 1 H+ [cytosol]
actionid: Reactome_70449	enzyme/gene: glyceraldehyde-3-phosphate dehydrogenase tetramer	compartment: cytosol kinetic: massi <a href="#">more</a>

populate

(C)

```

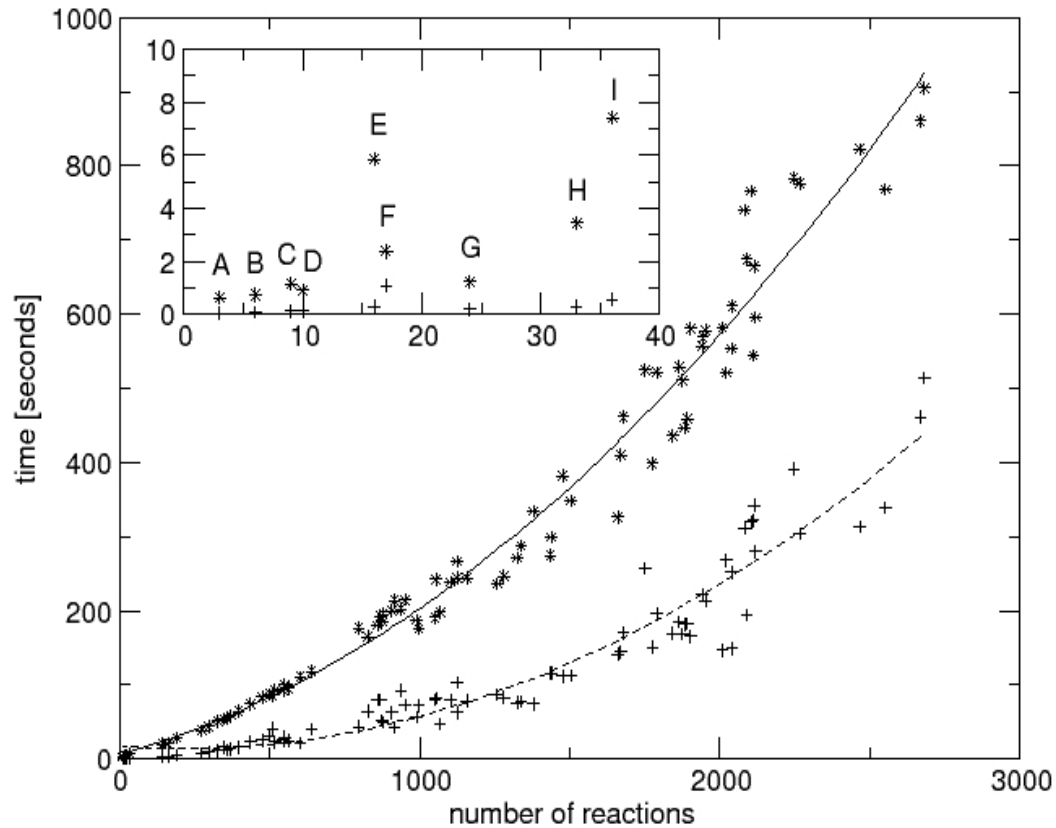
    graph TD
      A[D-Fructose 1,6-bisphosphate] --> B[Dihydroxyacetone phosphate]
      B --> C[D-Glyceraldehyde 3-phosphate]
      A --> C
    
```

(D) search network reaction

show info	populate	delete	reaction	direction	reaction
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	1 D-Fructose 1,6-bisphosphate [cytosol]	->	1 Dihydroxyacetone phosphate [cytosol] + 1 D-Glycerate 3-phosphate [cytosol]
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	1 Dihydroxyacetone phosphate [cytosol]	->	1 D-Glyceraldehyde 3-phosphate [cytosol]
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	1 D-Glyceraldehyde 3-phosphate [cytosol] + 1 Orthophosphate [cytosol] + 1 NAD+ [cytosol]	->	1 3-Phospho-D-glyceroyl phosphate [cytosol] + 1 NADH [cytosol] + 1 H+ [cytosol]

update/delete populate

# Performance of large-scale modelling



# Scripting model simulations

- generate PyBioS model
- determine kinetic constants with Monte Carlo approach
- run ODE solver for each model pair
- summarise the model pairs across all parameter vectors
- compute statistical validation of the output

## Involves several software packages:

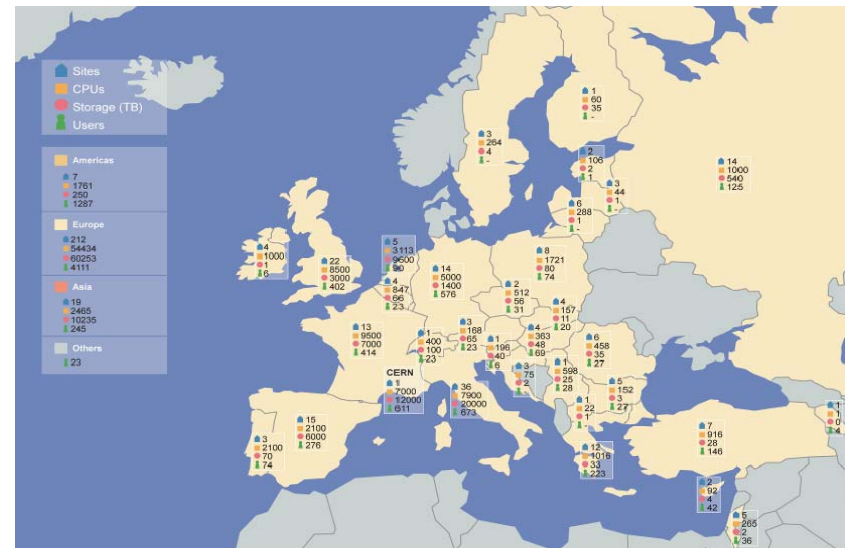
Fortran ODE solver

Python modules (numpy, scipy)

R/Bioconductor

Christophe Blanchet, CNRS, Lyon

EGEE 2008 Conference, Istanbul 22.09.2008



Max-Planck-Institute  
for Molecular Genetics



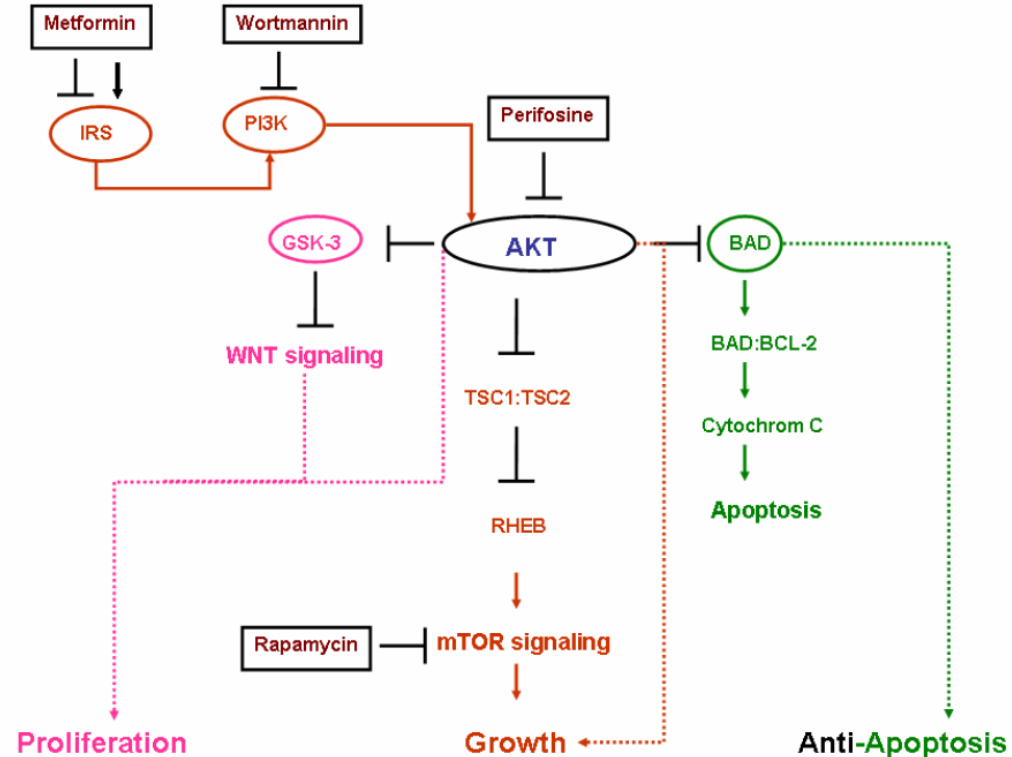
# Talk structure

1. Introduction - the EMBRACE Network of Excellence
2. Modelling human diseases
3. Data on human interactions
4. Modelling systems and methods
5. Use case: Modelling the effect of drugs
6. Conclusion & outlook



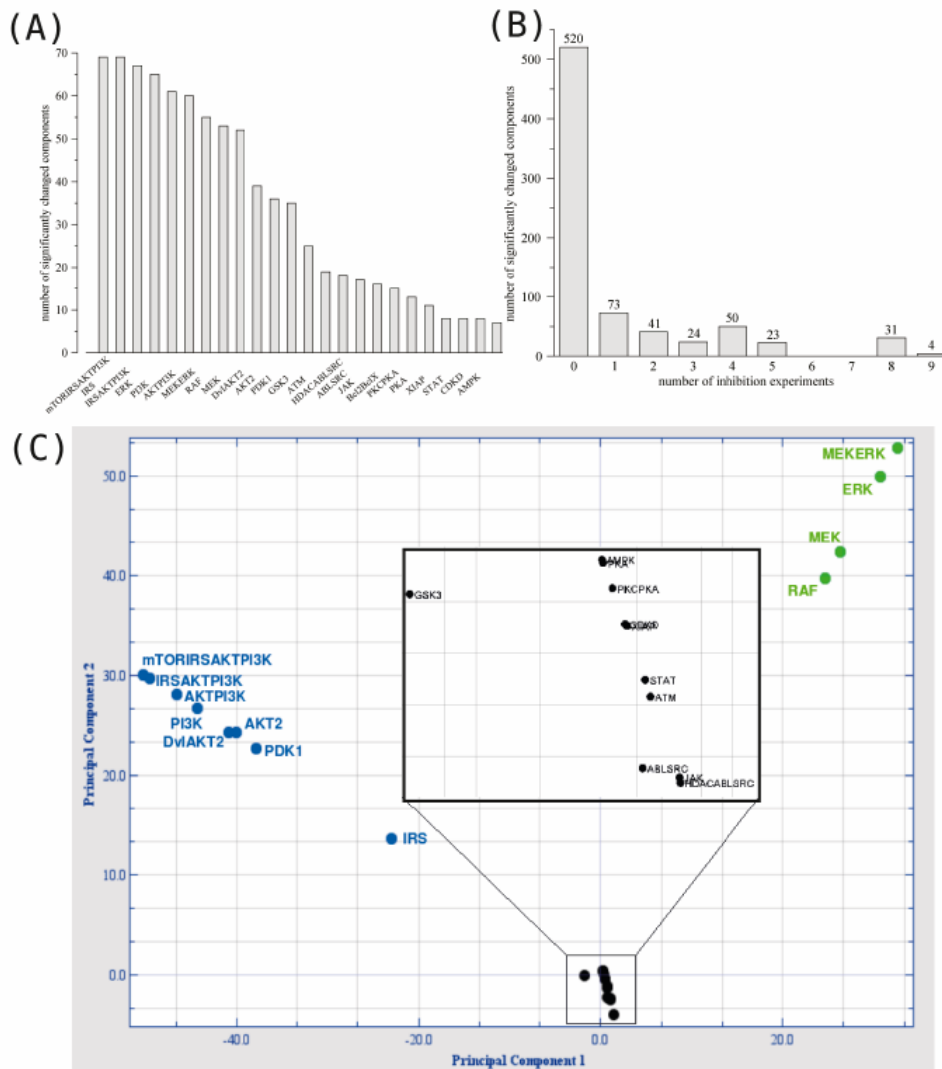
# Testing drug effects

Compound	Target protein	Target pathway	Inhibition experiments	Reference
Perifosine	RAC-beta serine/threonine protein kinase	AKT Signaling	AKT2, AKTPI3K, IRSAKTPI3K, mTORIRSAKTPI3K, DvlAKT2	Van Ummersen et al. 2004, Hennessy et al., 2005
Wortmannin analogues	Phosphatidylinositol-4,5-bisphosphate 3-kinase	PI3K Cascade	PI3K, AKTPI3K, IRSAKTPI3K, mTORIRSAKTPI3K	Ng et al. 2001, Hennessy et al., 2005
Metformin	Insulin receptor substrate-1	IRS-signaling	IRS, IRSAKTPI3K, mTORIRSAKTPI3K	Yuan et al. 2003
Indirubin-3'-oxime	AMP-activated protein kinase	Glucagon signaling	AMPK	Meijer et al., 2003
15 a.a peptide	Extraacellular regulated kinase	Raf Signaling	ERK, MEKERK	Horiuchi et al., Shen & Brown 2003
PD-325901, ARRY-142886	Dual specificity mitogen-activated protein kinase kinase 1	Raf Signaling	MEK, MEKERK	NCI, 2008, Hugnh et al. 2007
Staurosporine	Protein kinase A	Glucagon signaling	PKA, PKCPKA	Kissau 2002
Enzastaurin (LY-317615)	Protein kinase C	Protein kinase		
PD-0332991	Cyclin dependent kinase	Cell cycle		
AEG35156	X-linked Inhibitor of apoptosis	Apoptosis		
FJ9	Dishevelled	Wnt Signaling		
ATM Inhibitor (KU-0055933)	Ataxia telangiectasia mutated	Apoptosis		
UCN-01, OSU03012	3-phosphoinositide dependent protein kinase-1	AKT Signaling		
Imatinib (Glivec®), Dasatinib	Abelson murine leukemia	Cell cycle		
FR901228	Histone deacetylase	Cell cycle		
Obatoclax-Mesylate (GX15-070MS)	B-cell lymphoma 2	Apoptosis		
Obatoclax-Mesylate (GX15-070MS)	Apoptosis regulator Bcl-X	Apoptosis		
STAT-induced-STAT-inhibitor-1 (SSI-1)	Signal Transducers and Activators of Transcription	Cytokine Signa		
CP-690550	Janus kinase 1	Cytokine Signa		
CP-690550	Janus kinase 3	Cytokine Signa		
Dasatinib (Sprycel®)	Proto-oncogene tyrosine-protein kinase Src	Src Signaling		
Indirubin-3'-oxime, Aloisine A	Glycogen synthase kinase-3beta	Wnt signaling		
Everolimus (Corticin®)	Mammalian target of rapamycin	mTOR Signaling		
Diferuloglmethane	Cyclin D	Cell cycle		
Sorafenib (Naxavar®)	BRAF	Raf Signaling		
Sorafenib (Naxavar®)	c-raf	Raf Signaling		



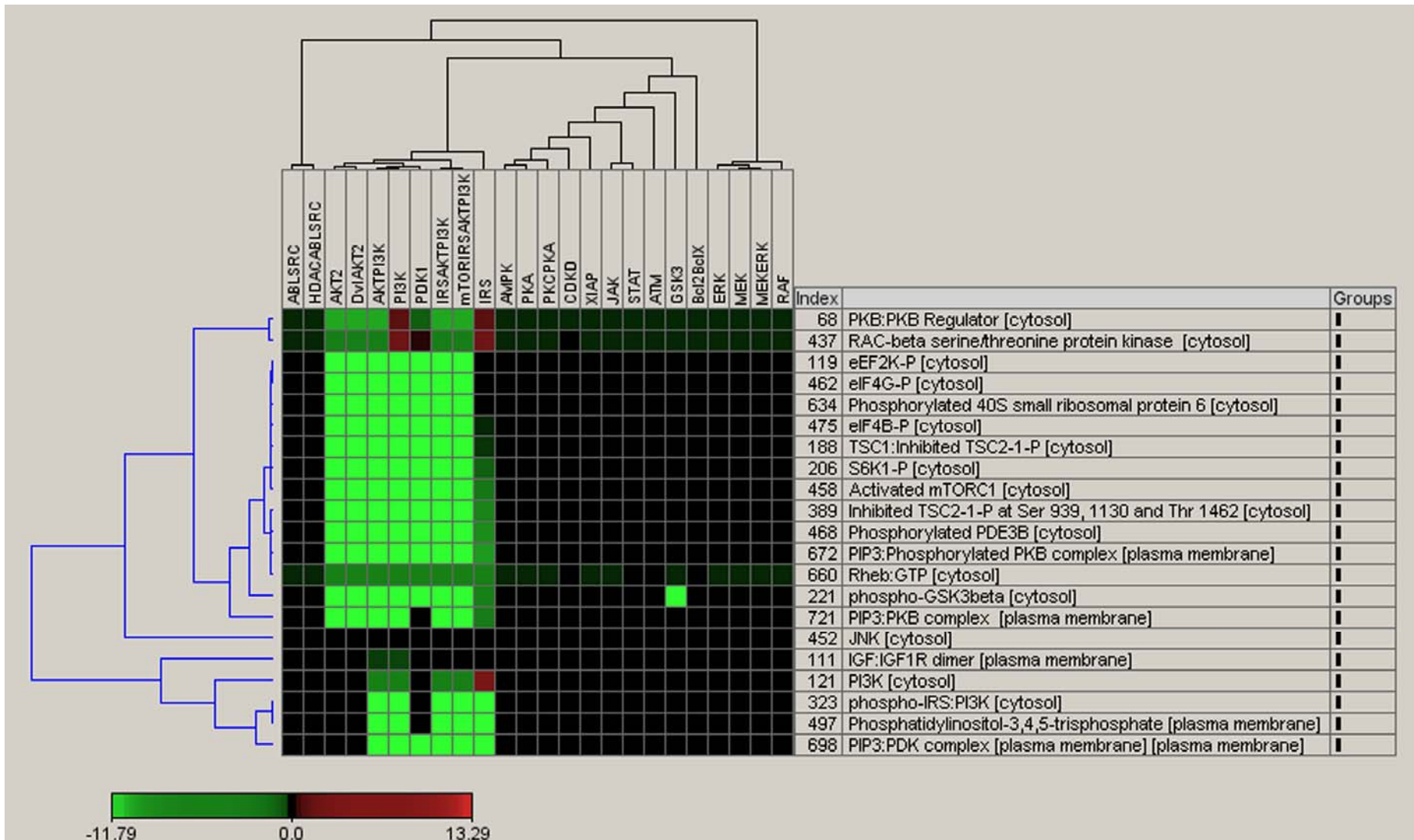


# Simulation results – global statistics

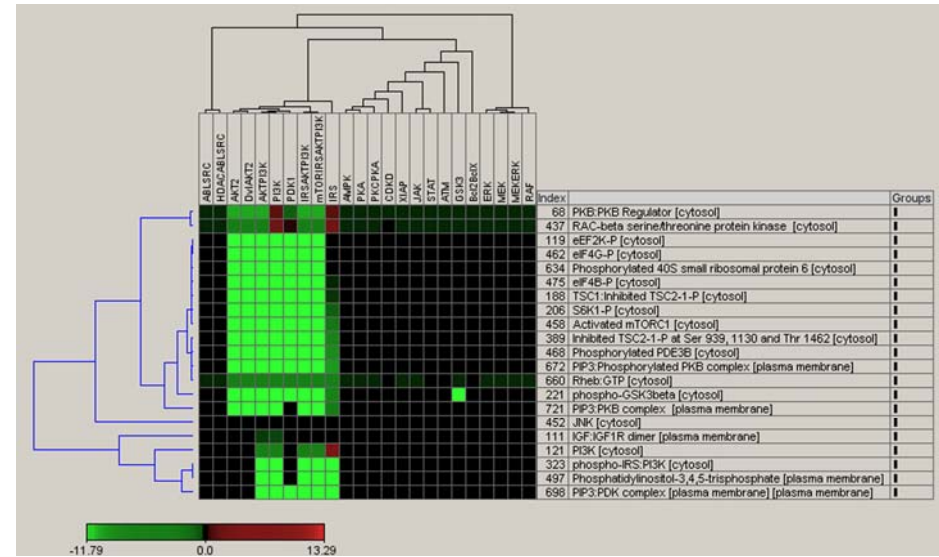
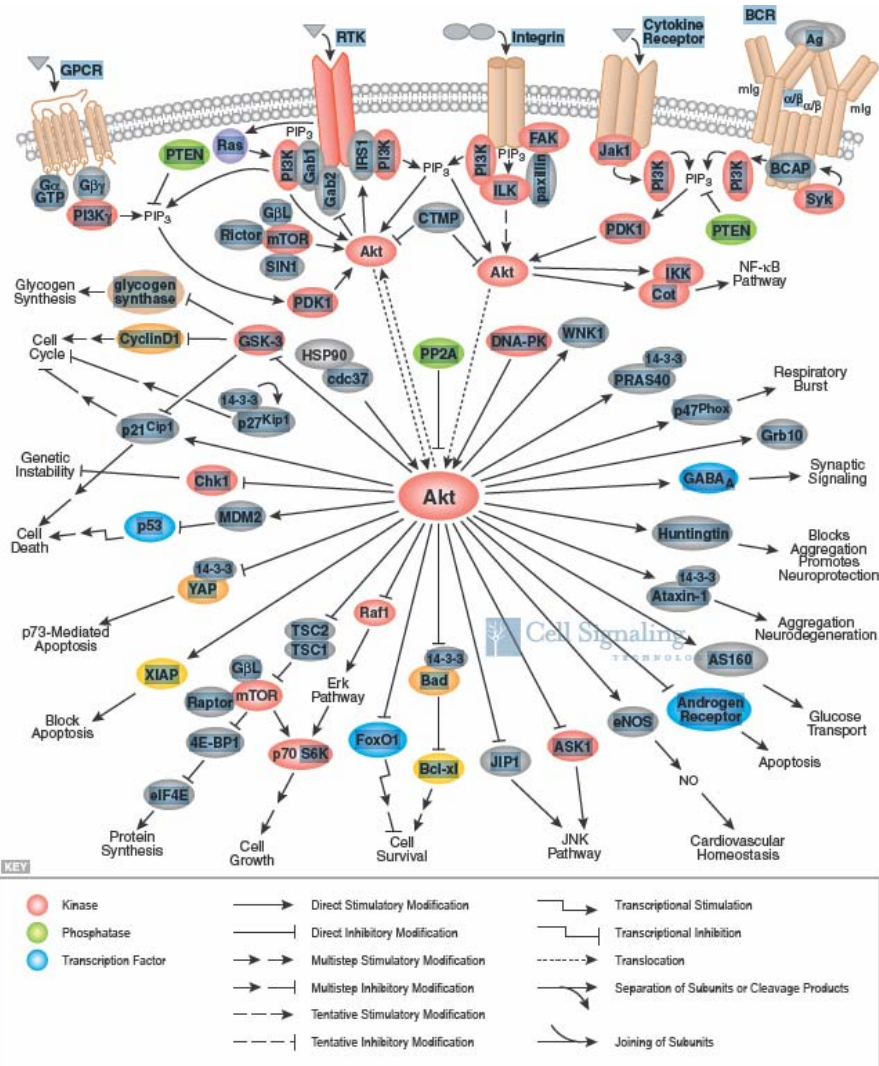


- most components not affected
- direct and indirect effects
- highest effectors with 10% of model components
- specificity of drugs different

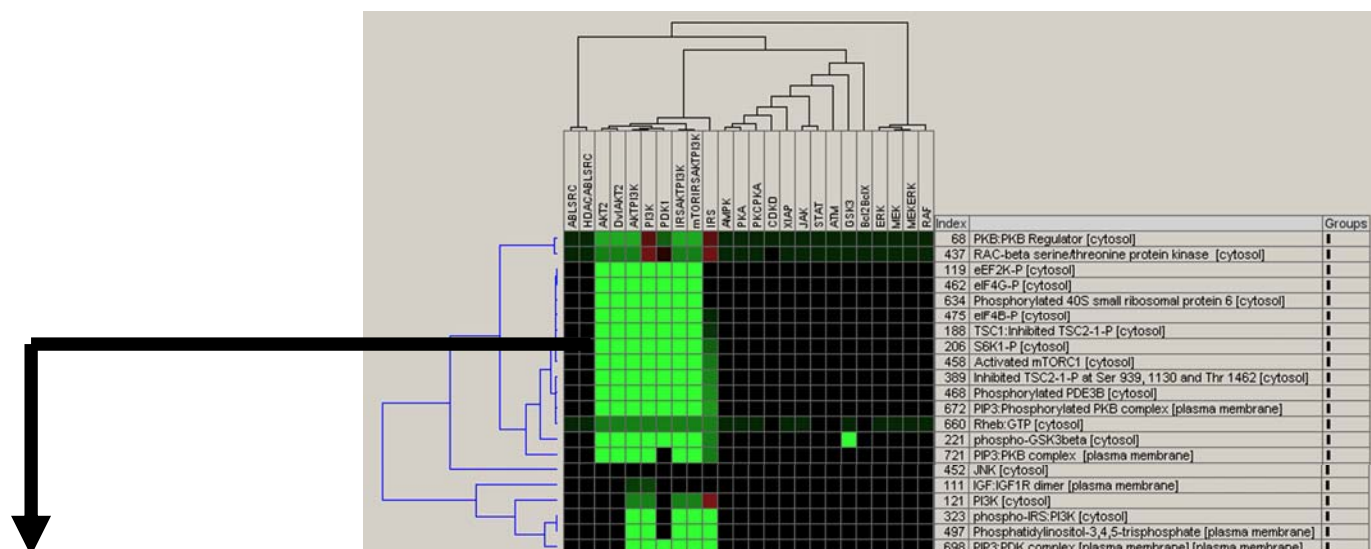
# Simulation results – clustering dependencies



# Simulation results – clustering dependencies



# Simulation results – clustering dependencies



1: [Mol Cell.](#) 2006 Oct 20;24(2):185-97.

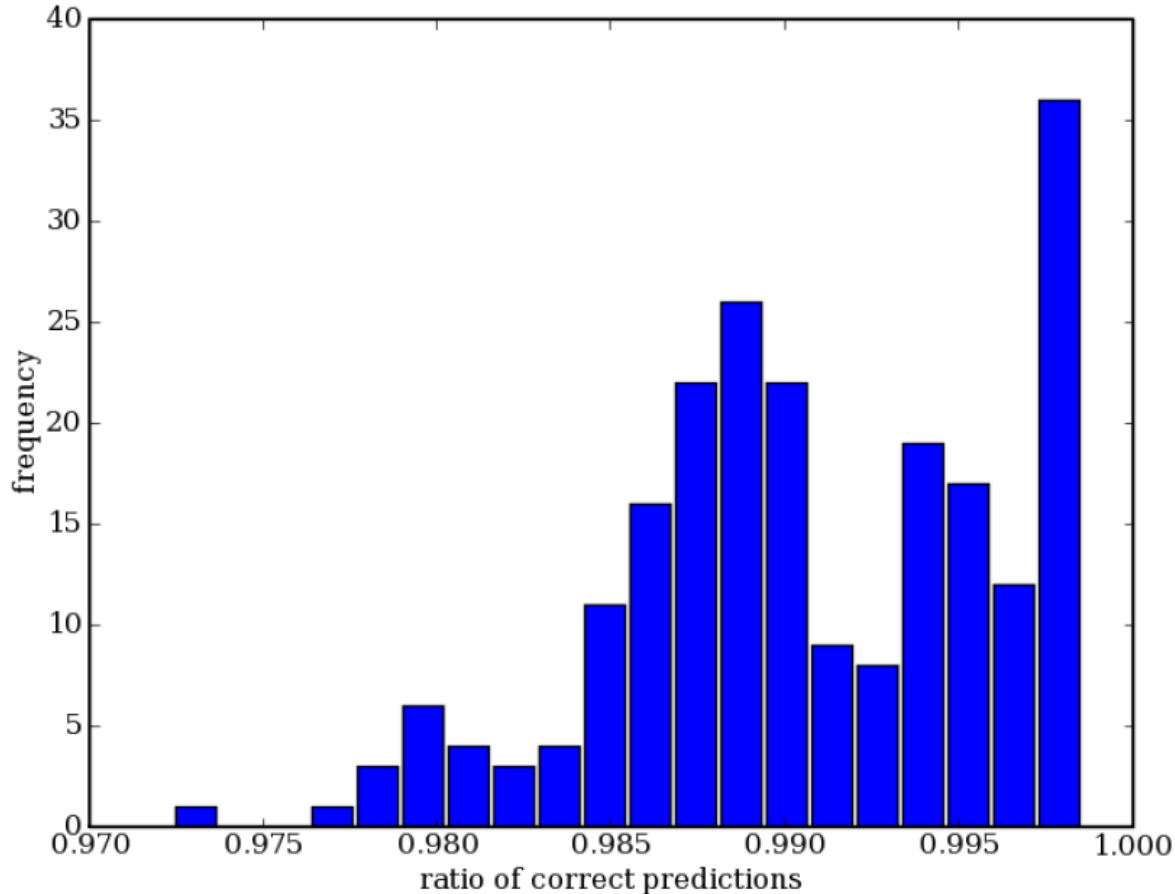
S6K1 regulates GSK3 under conditions of mTOR-dependent feedback inhibition of Akt.

[Zhang HH](#), [Lipovsky AI](#), [Dibble CC](#), [Sahin M](#), [Manning BD](#).

Department of Genetics and Complex Diseases, Harvard School of Public Health, Harvard Medical School, Boston, Massachusetts 02115, USA.

Feedback inhibition of the PI3K-Akt pathway by the mammalian target of rapamycin complex 1 (mTORC1) has emerged as an important signaling event in tumor syndromes, cancer, and insulin resistance. Cells lacking the tuberous sclerosis complex (TSC) gene products are a model for this feedback regulation. We find that, despite Akt attenuation, the Akt substrate GSK3 is constitutively phosphorylated in cells and tumors lacking TSC1 or TSC2. In these settings, GSK3 phosphorylation is sensitive to mTORC1 inhibition by rapamycin or amino acid withdrawal, and GSK3 becomes a direct target of S6K1. This aberrant phosphorylation leads to decreased GSK3 activity and phosphorylation of downstream substrates and contributes to the growth-factor-independent proliferation of TSC-deficient cells. We find that GSK3 can also be regulated downstream of mTORC1 in a HepG2 model of cellular insulin resistance. Therefore, we define conditions in which S6K1, rather than Akt, is the predominant GSK3 regulatory kinase.

# Simulation results – cross validation



- **leave-one-out cross validation**
- **„correct“ predictions against computed predictions**
- **concordance factor**

# Talk structure

1. Introduction - the EMBRACE Network of Excellence
2. Modelling human diseases
3. Data on human interactions
4. Modelling systems and methods
5. Use case: Modelling the effect of drugs
6. Conclusion & outlook



# Conclusions

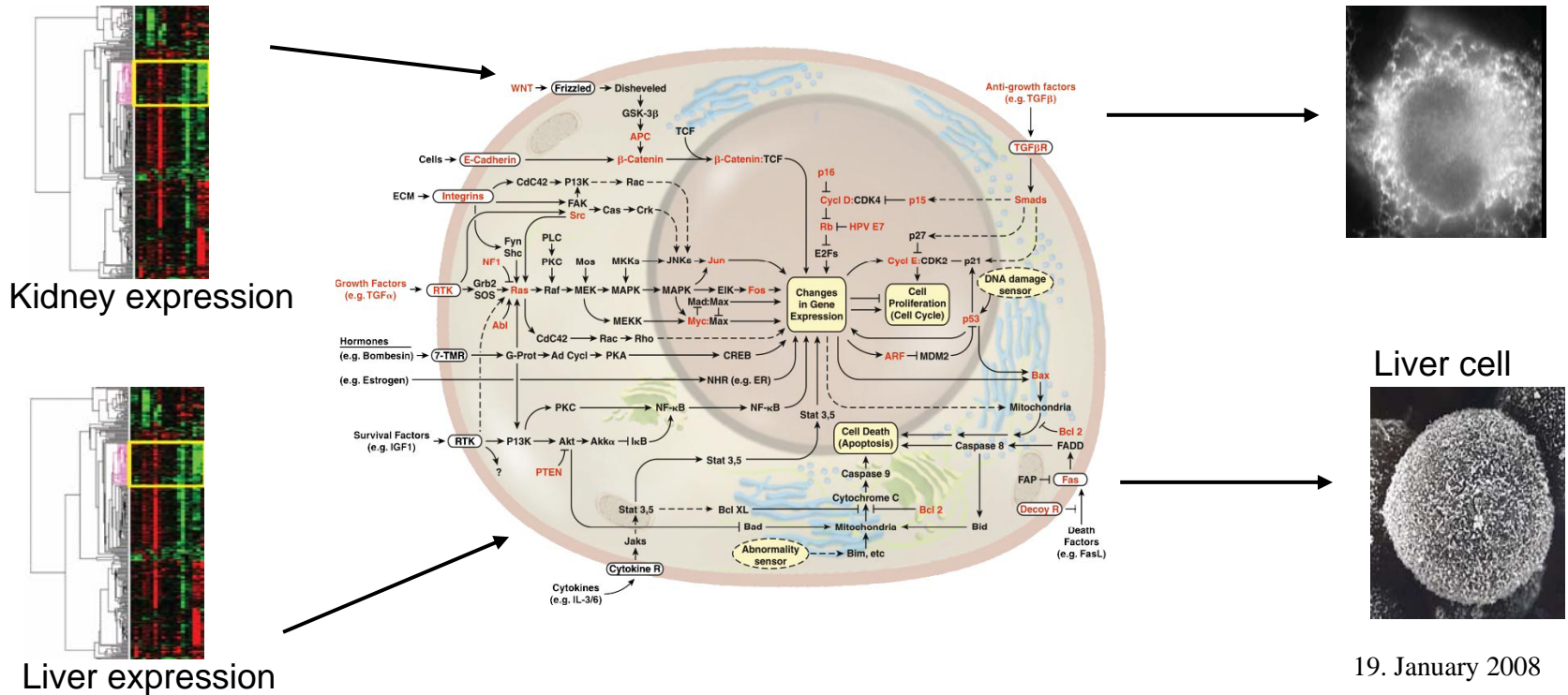
- **modelling strategy for generating predictions from medically-relevant biological networks**
- **requires massively parallel computer simulations; ideal grid application**
- **use cases have been run successfully**
- **prediction results can be verified by literature**



# Adaptation of generic models

Gene expression data can be imprinted onto the cancer cell model:

- Unexpressed genes are deleted
- Levels of proteins are adjusted to expression levels
- Proteins levels are further tuned by their half life (if known)



19. January 2008



# Acknowledgments

## MPIMG

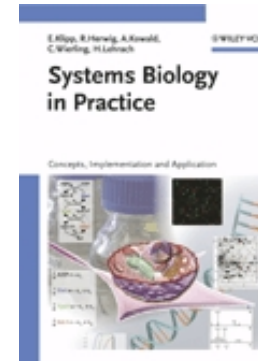
Christoph Wierling  
Andriani Daskalaki  
Elisabeth Maschke-Dutz  
Atanas Kamburov  
Hans Lehrach

CNRS, Lyon

Christophe Blanchet

## EU projects

EMBRACE  
CARCINOGENOMICS  
APO-SYS



Klipp, Herwig, Kowald, Wierling, Lehrach.  
Systems Biology in Practice, Wiley.

**Thanks for your attention !**

