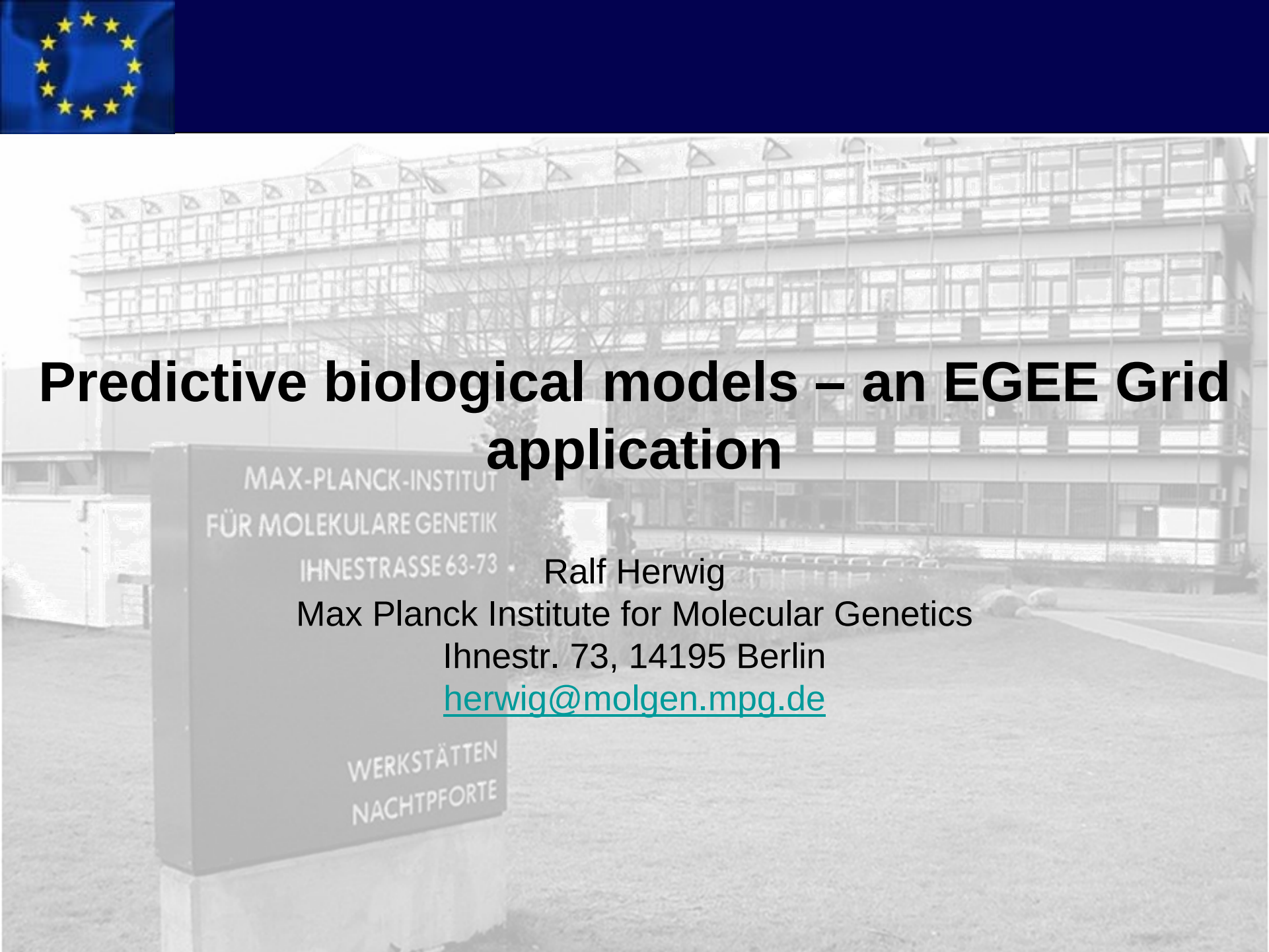



# Predictive biological models – an EGEE Grid application

Ralf Herwig
Max Planck Institute for Molecular Genetics
Ihnestr. 73, 14195 Berlin
herwig@molgen.mpg.de

# 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

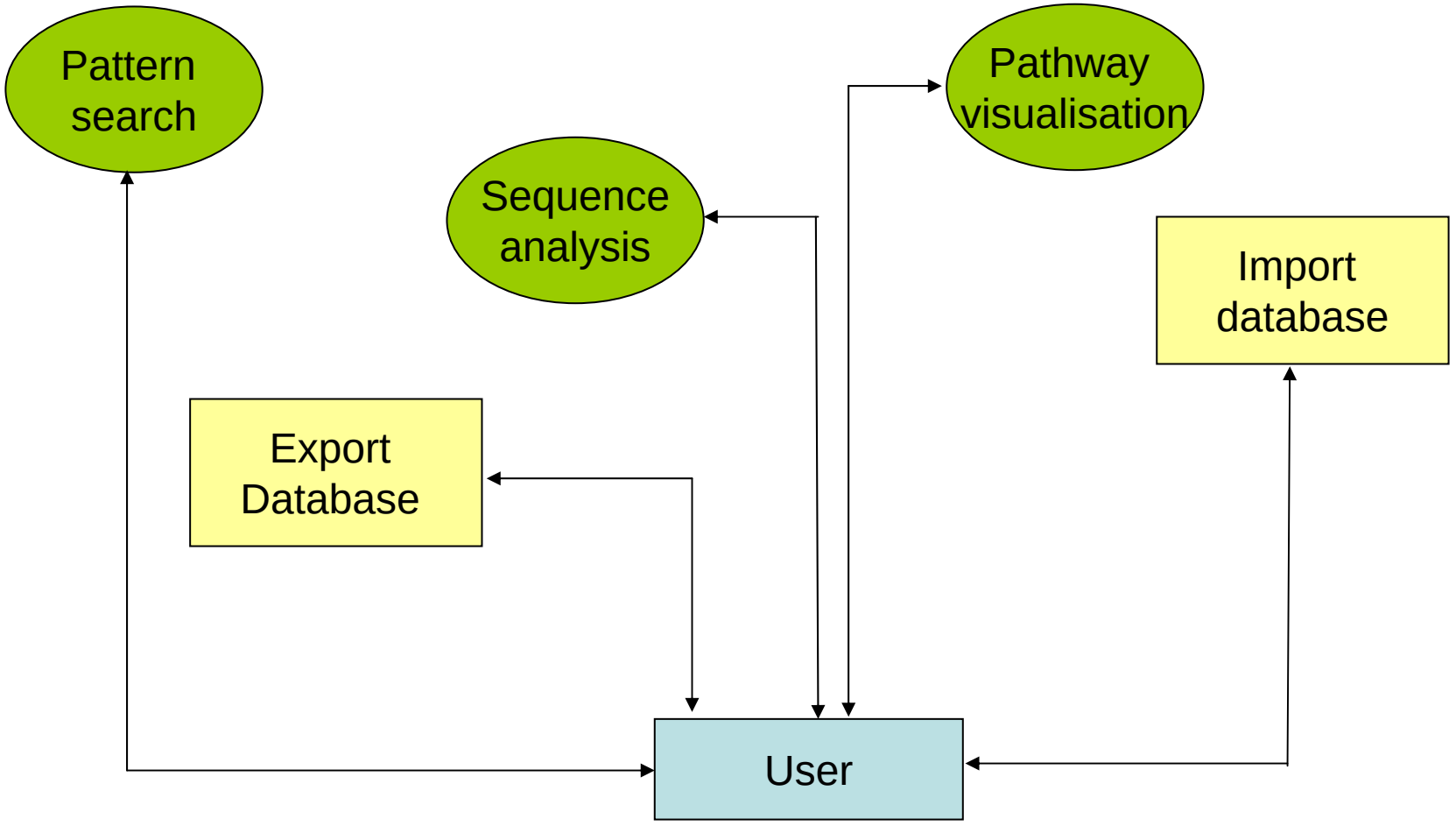
| Particip. Number | Participant name | Participant short name |
|---|---|---|
| 1 | European Molecular Biology Laboratory – European Bioinformatics Institute | EMBL-EBI |
| 2 | Instituto di Tecnologie Biomediche – Sezione di Bari, CNR | ITB-BA |
| 3 | University of Manchester | UMAN |
| 4 | Swiss Institute 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 Aerospacial (Centre de Astrobiologia) | INTA-CAB |
| 18 | University of Bergen | CBU |



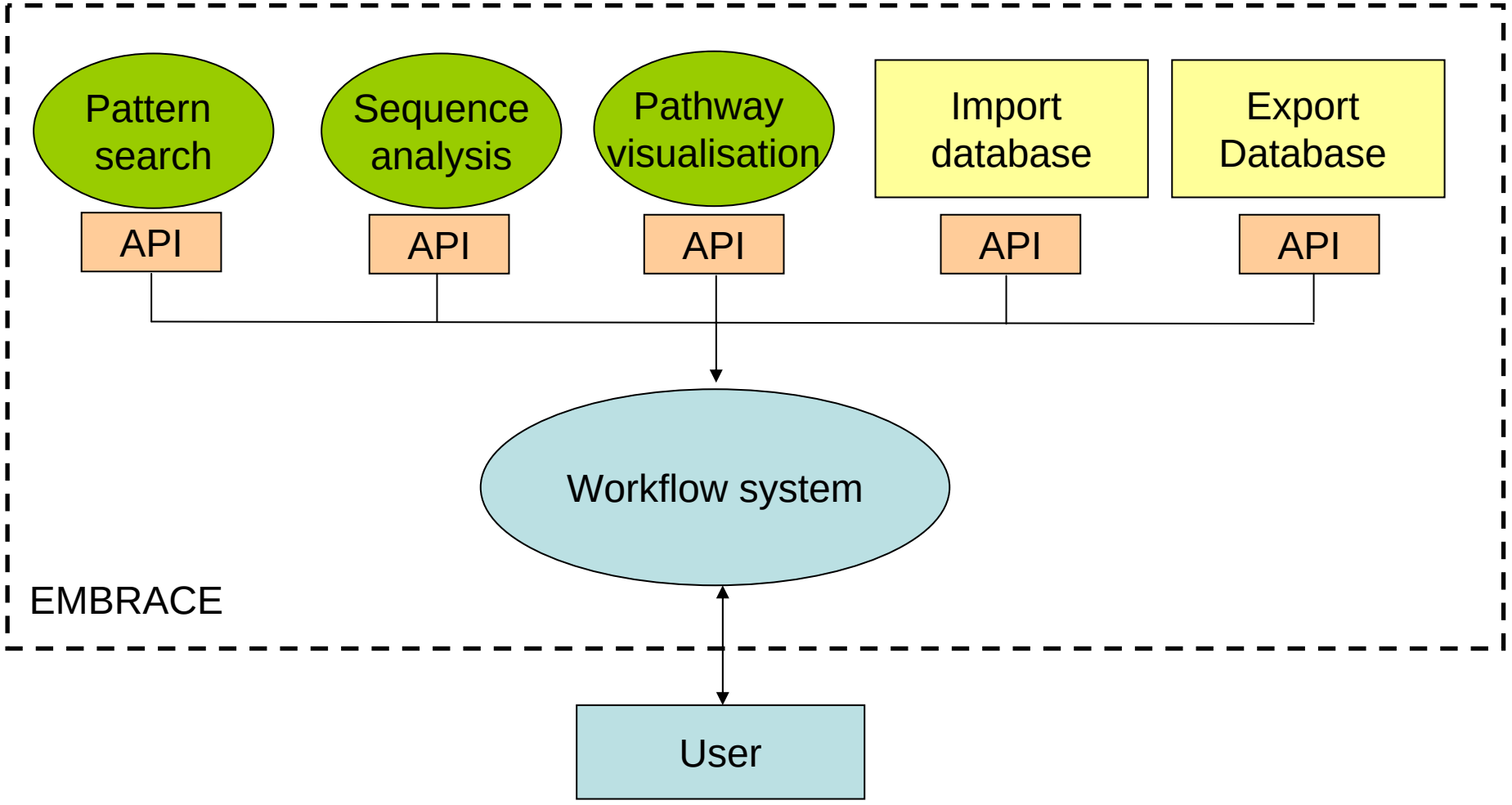
Max-Planck-Institute for Molecular Genetics

EMBRACE Grid Network of excellence

# User interaction – current situation



# User interaction – the EMBRACE vision






Max-Planck-Institute for Molecular Genetics

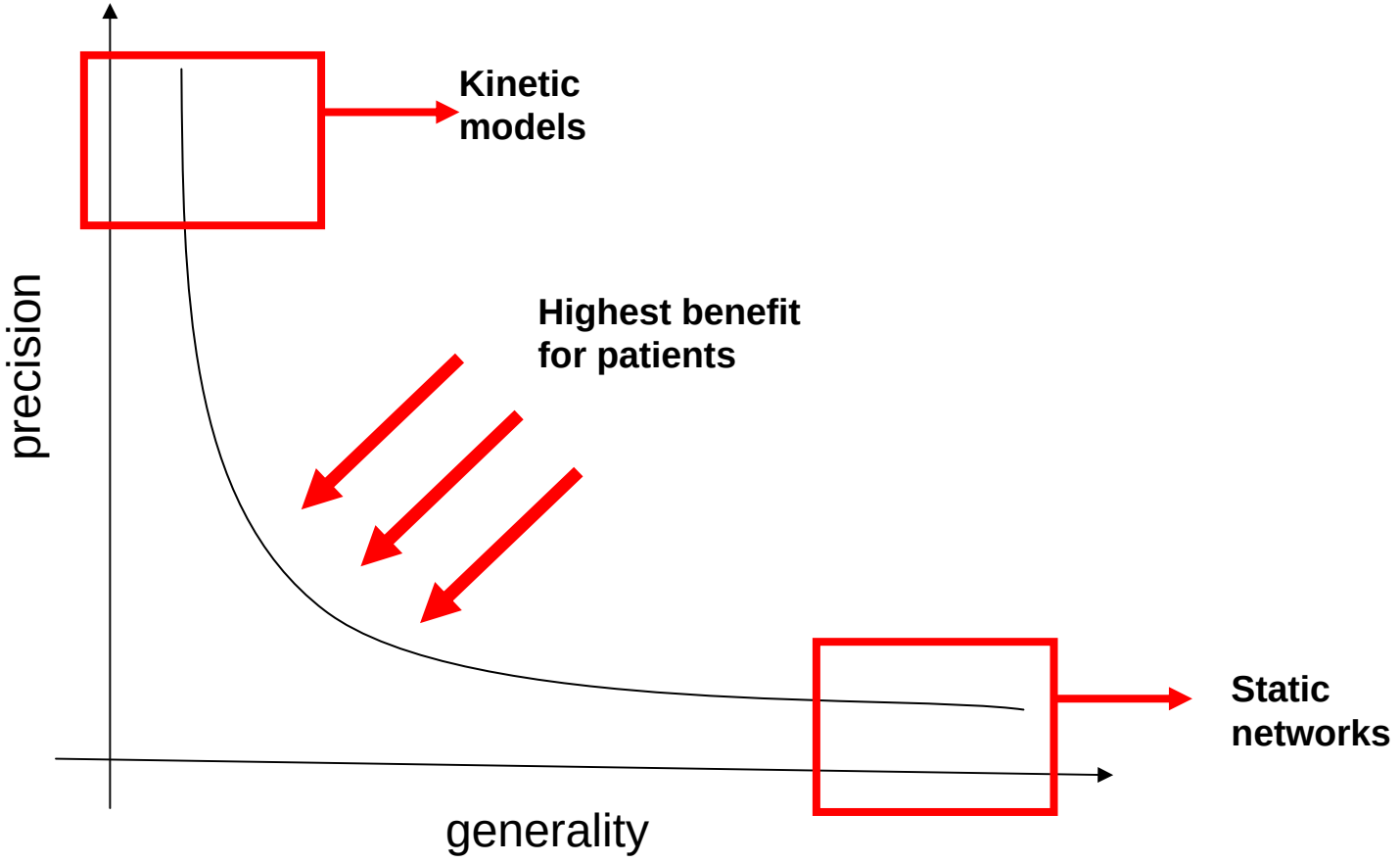
EMBRACE Grid Network of excellence

# 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 ?

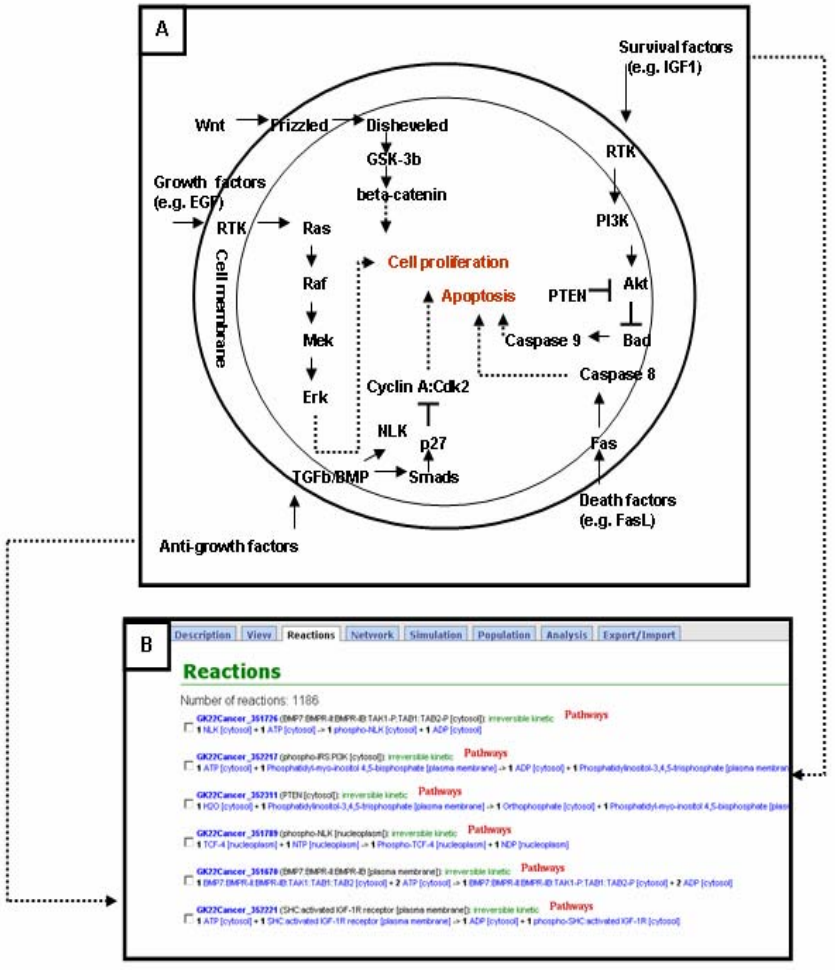
**Trade-off between model generality and model precision**



# 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




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

| Reaction | Biochemical equation | Rate law |
|---|---|---|
| Synthesis | $v_0 : \rightarrow S2$ | $v_0 = k_0$ |
| Complex formation | $v_1 : S1 + S2 \leftrightarrow S1 : S2$ | $v_1 = kf[S1][S2] - (kf / K_D)[S1 : S2]$ with $K_D$=100 |
| Formation of product | $v_2 : S1 \rightarrow S2$ | $v_2 = k1[S1]$ |
| Degradation | $v_3 : S2 \rightarrow$ | $v_3 = k_3\ S2$ with k=$10^{-3}$ (*) |

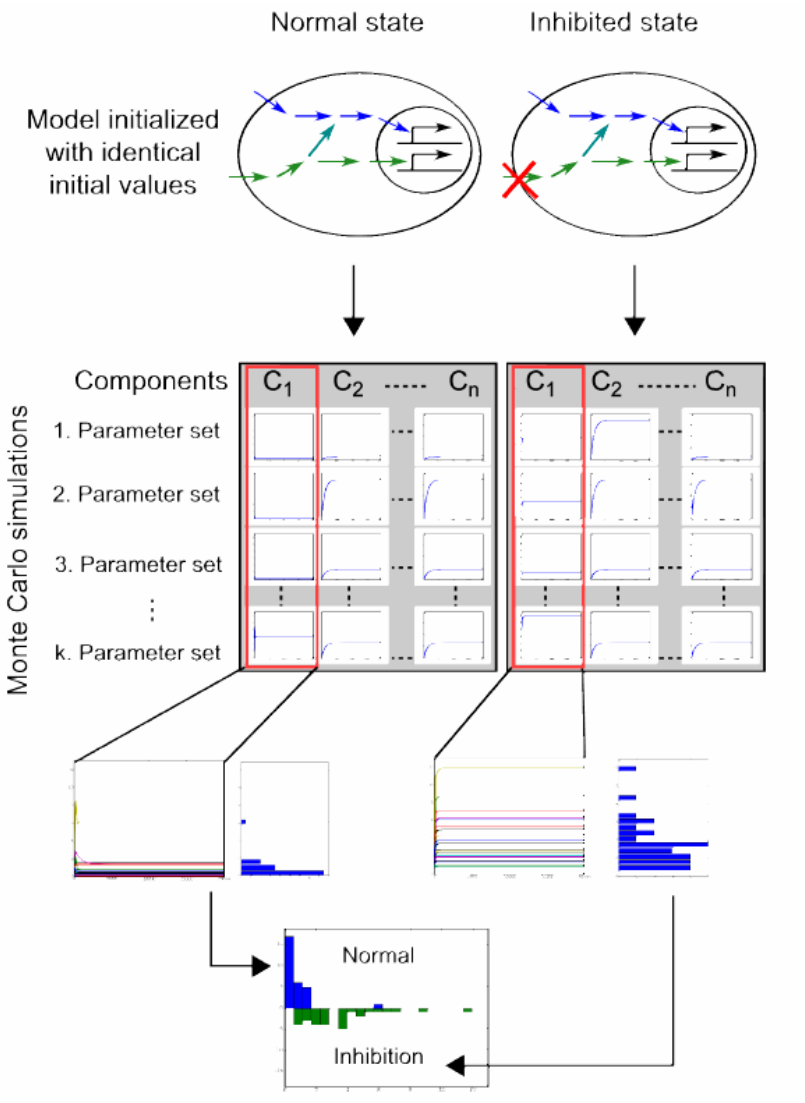
**Wierling et al., under review.**

**EGEE 2008 Conference, Istanbul 22.09.2008**

Max-Planck-Institute for Molecular Genetics

MBRACE Grid Network of excellence

# 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

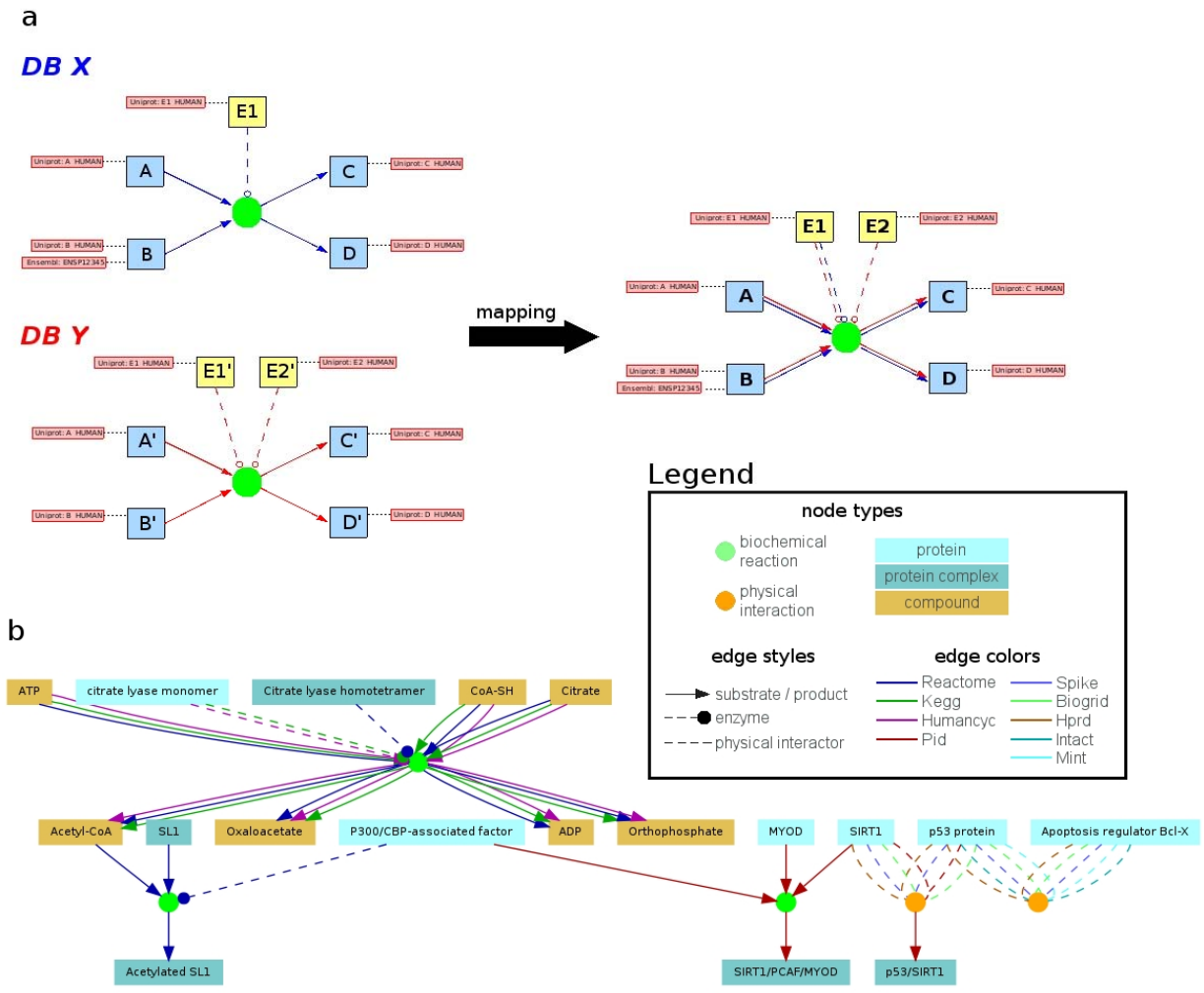
**Problem:** Diverse and heterogeneous pathway annotation

- 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.**

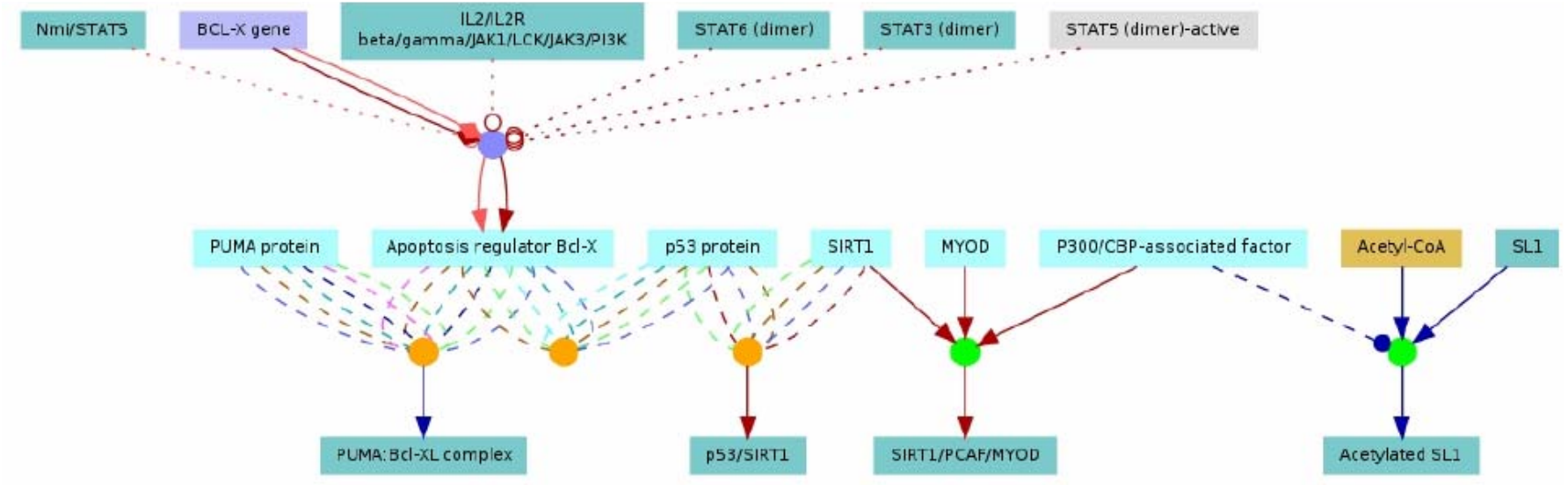
# ConsensusPathDB – mapping of interactions

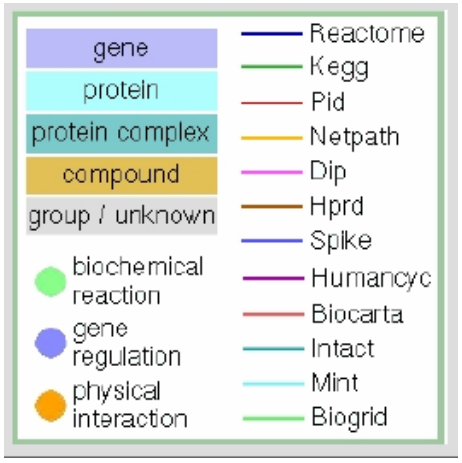
## Integration of human functional interactions




- **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





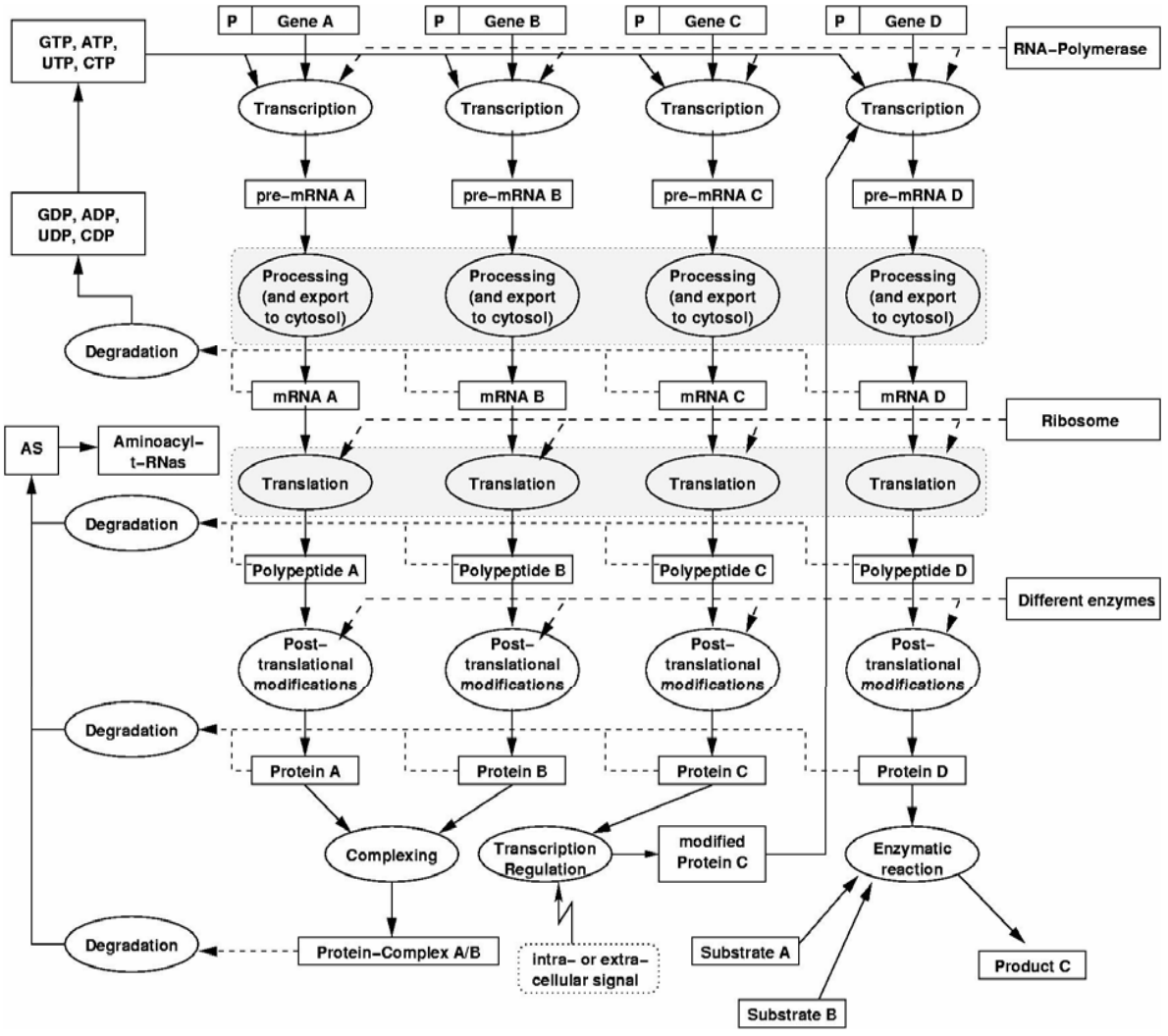

Max-Planck-Institute for Molecular Genetics

MBRACE Grid Network of excellence

# 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



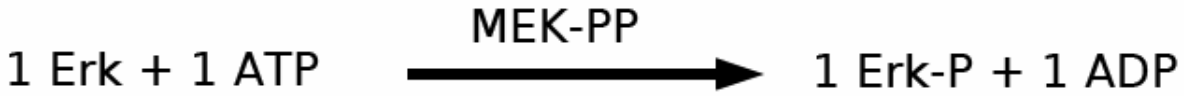

- **object-oriented structure**
- **data structure covers main biological objects**
- **actions attached to objects mimic biological processes**



Max-Planck-Institute for Molecular Genetics

# Reactions and kinetics

## Reaction equations




## Rate laws

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

## Kinetic parameters

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

## Initial concentrations

$$[Erk] = 0.1 \ \text{mM} \qquad [Erk-P] = 0.0 \ \text{mM} \qquad [MEK-PP] = 0.05 \ \text{mM}$$

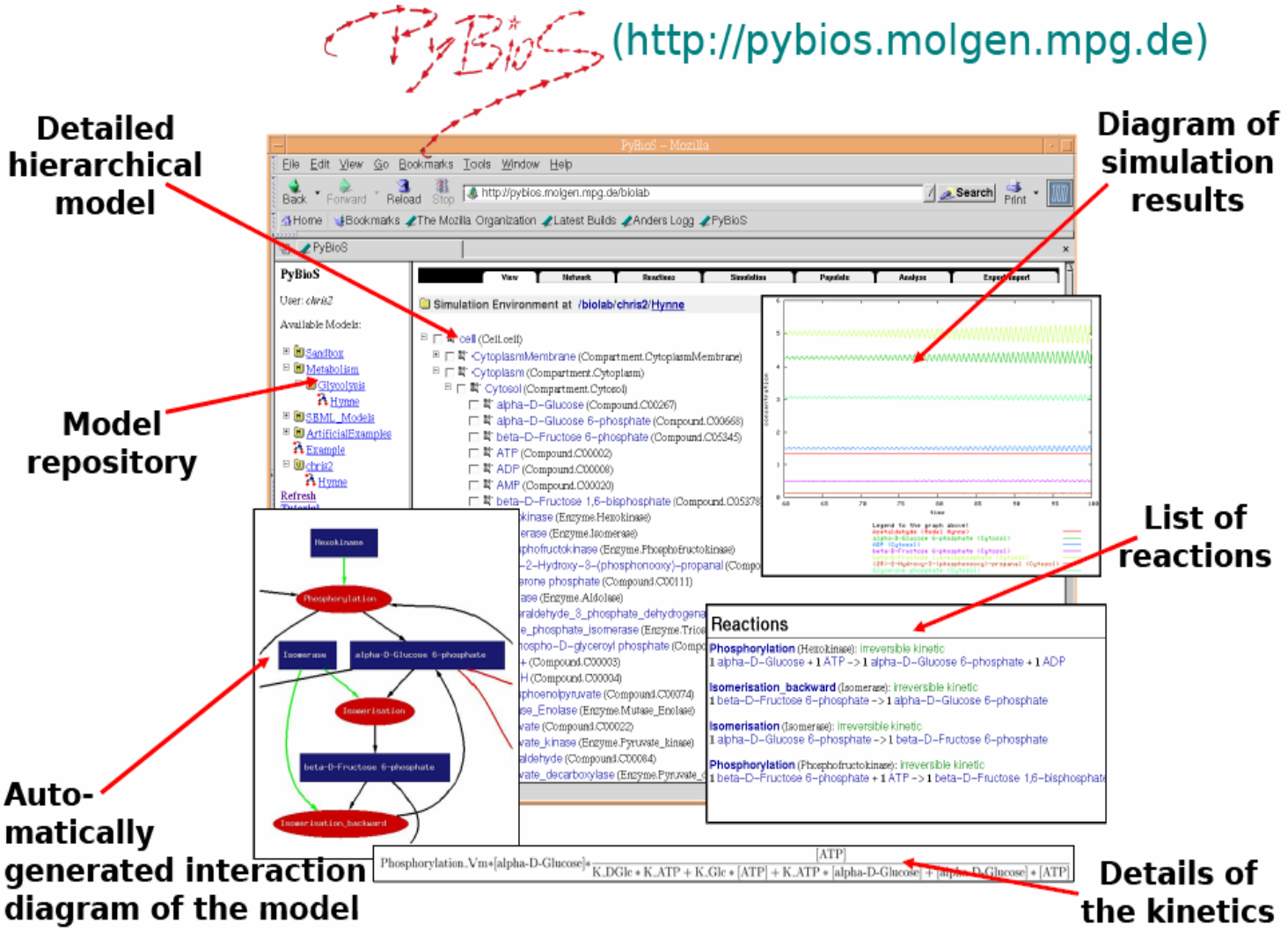
$$[ATP] = 3.0 \ \text{mM} \qquad [ADP] = 0.0 \ \text{mM}$$

## Differential equations

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

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

# Web interface

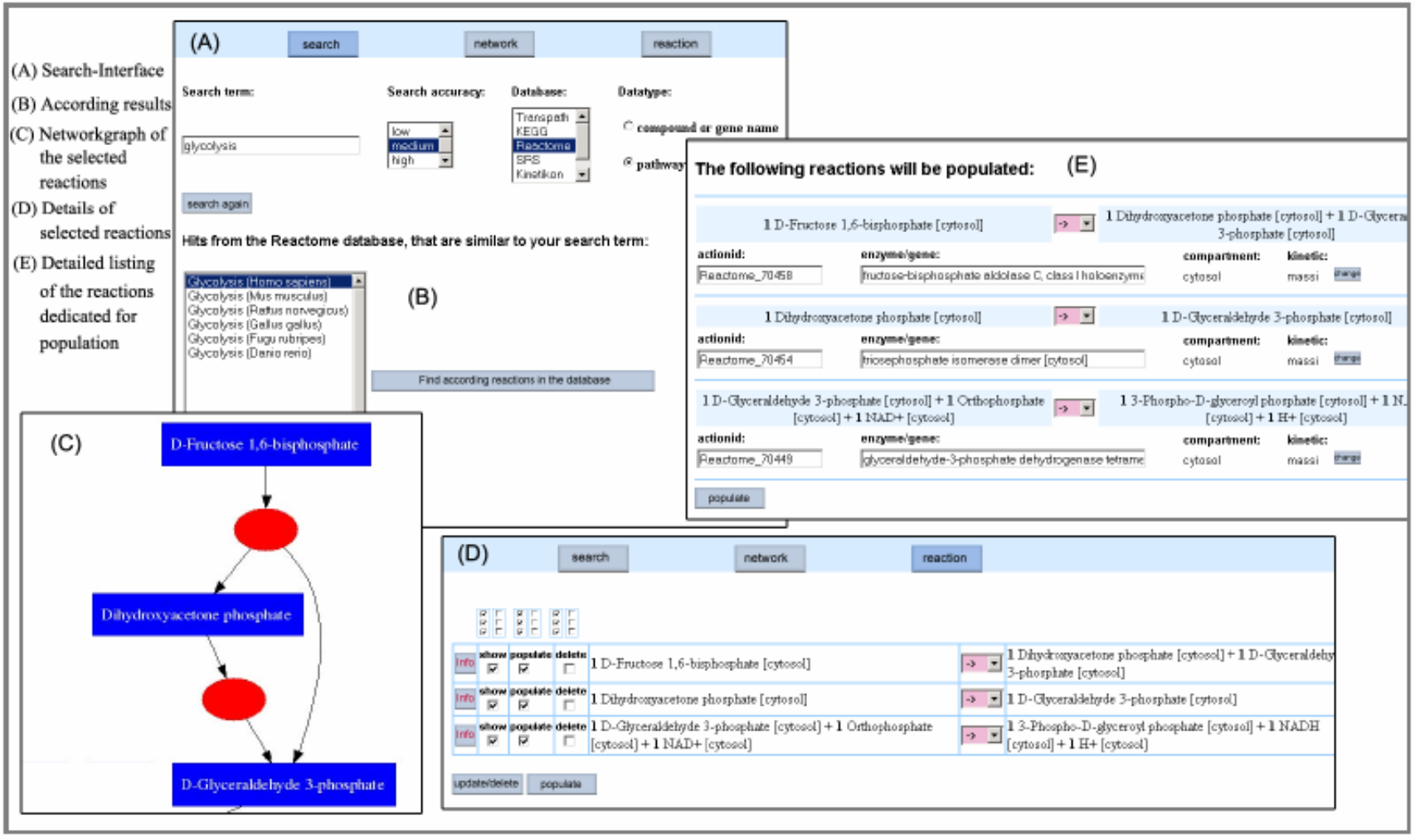




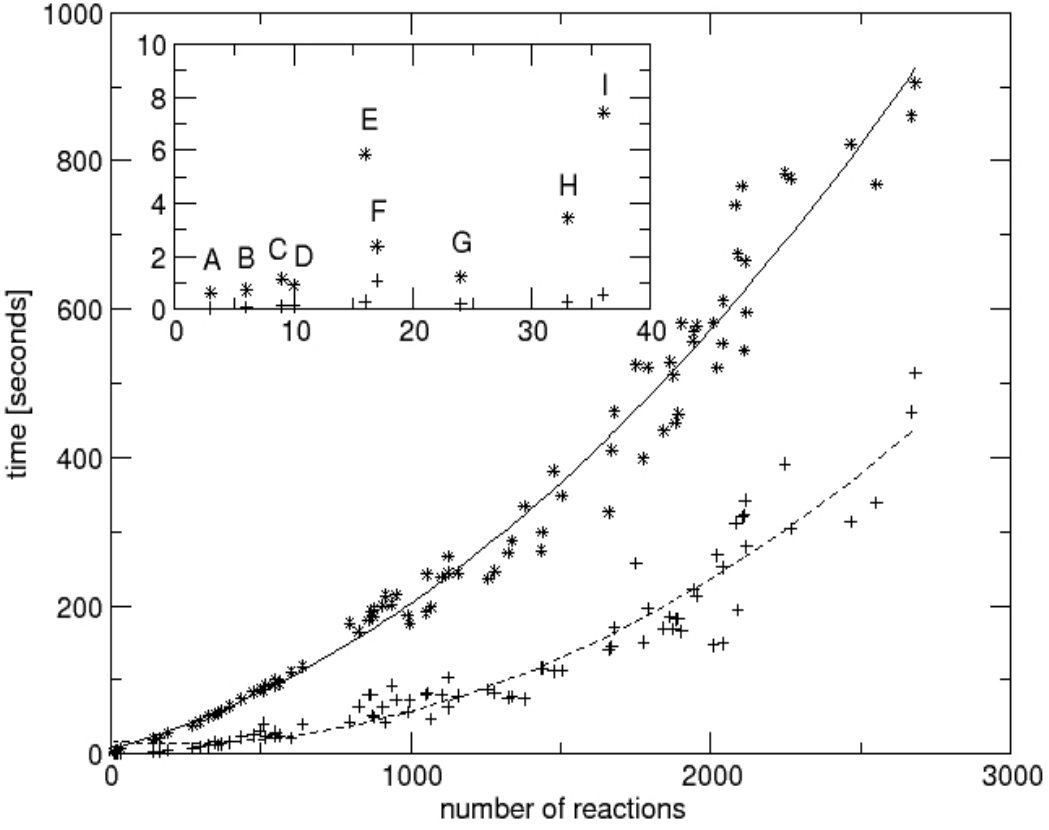

Max-Planck-Institute for Molecular Genetics

MBRACE Grid Network of excellence

# Automated network upload



# Performance of large-scale modelling



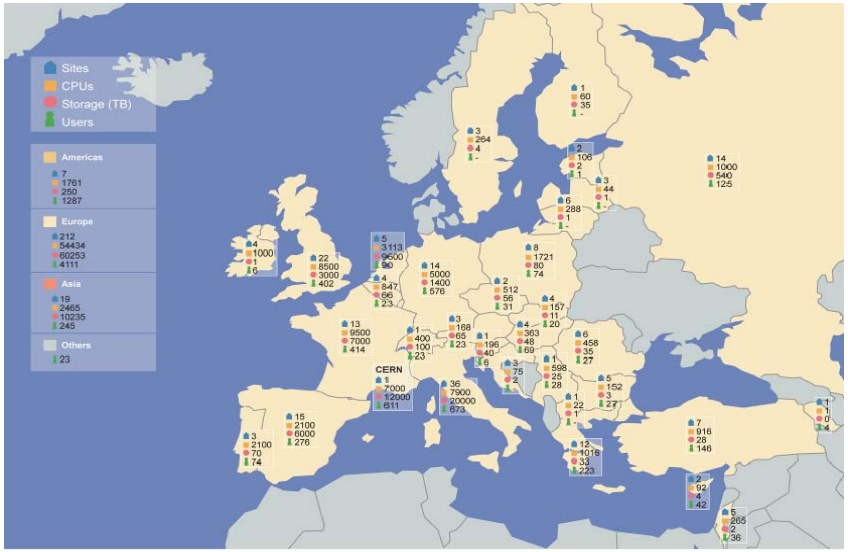



Max-Planck-Institute for Molecular Genetics

MBRACE Grid Network of excellence

# 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

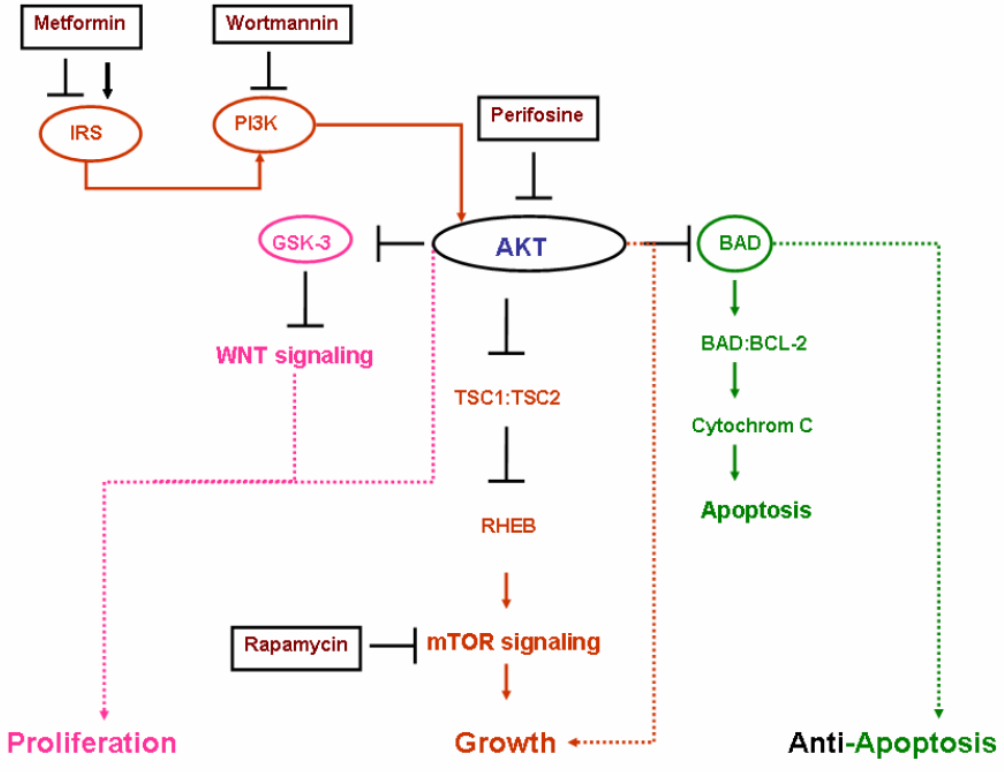
MBRACE Grid
Network of excellence

# Talk structure

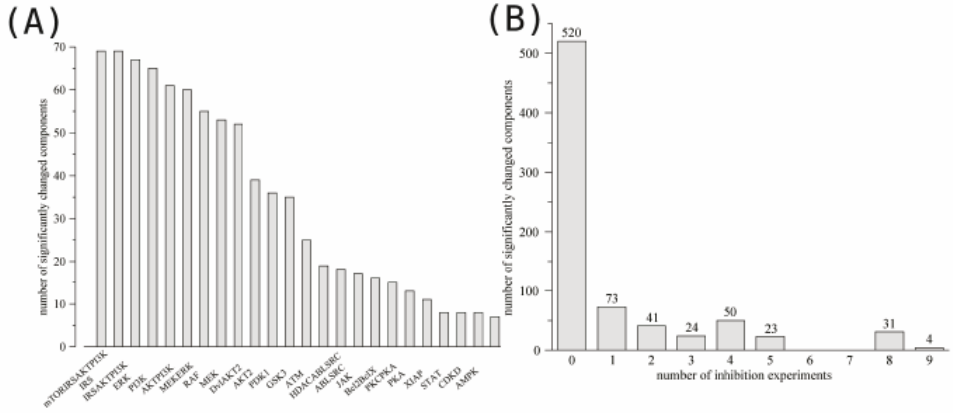
1. Introduction - the EMBRACE Network of Excellence
2. Modelling human diseases
3. Data on human interactions
4. Modelling systems and methods
5. Use case: Modelling the effect of drugs
6. Conclusion & outlook

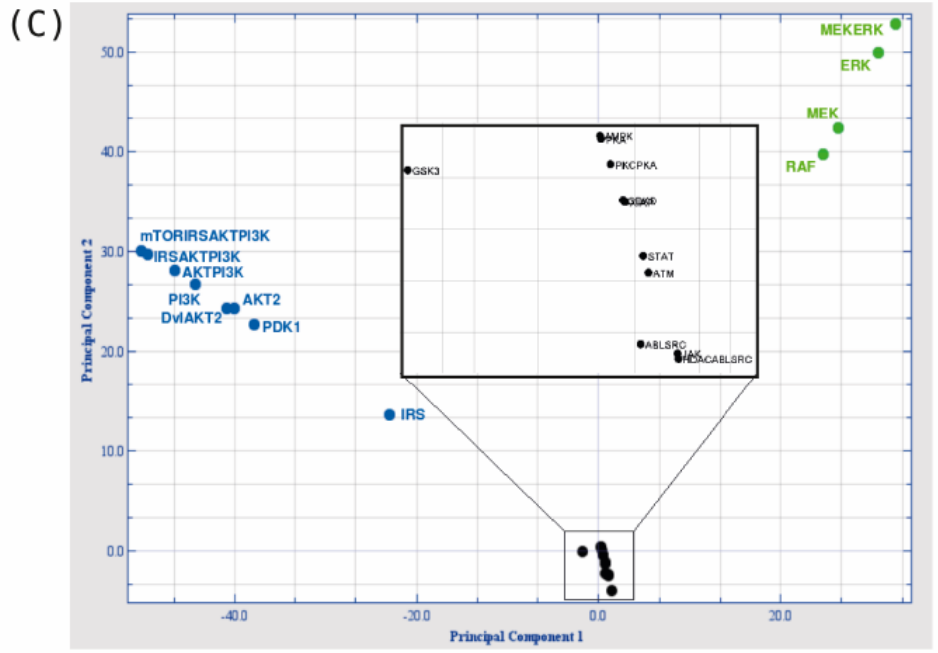
# Testing drug effects

| 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 | Extracellular 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, Huynh et al. 2007 |
| Staurosporine | Proteinkinase A | Glucagon signaling | PKA,PKCPKA | Kissau 2002 |
| Enzastaurin (LY-317615) | Proteinkinase 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(Spycel®) | Proto-oncogene tyrosine-protein kinase Src | Src Signaling | | |
| Indirubin-3'-oxime, Aloisine A | Glycogen synthase kinase-3beta | Wnt signaling | | |
| Everolimus(Certican®) | Mammalian target of rapamycin | mTOR Signalin | | |
| Diferuloylmethane | Cyclin D | Cell cycle | | |
| Sorafenib (Naxavar®) | BRAF | Raf Signaling | | |
| Sorafenib (Naxavar®) | c-raf | Raf Signaling | | |



# Simulation results – global statistics



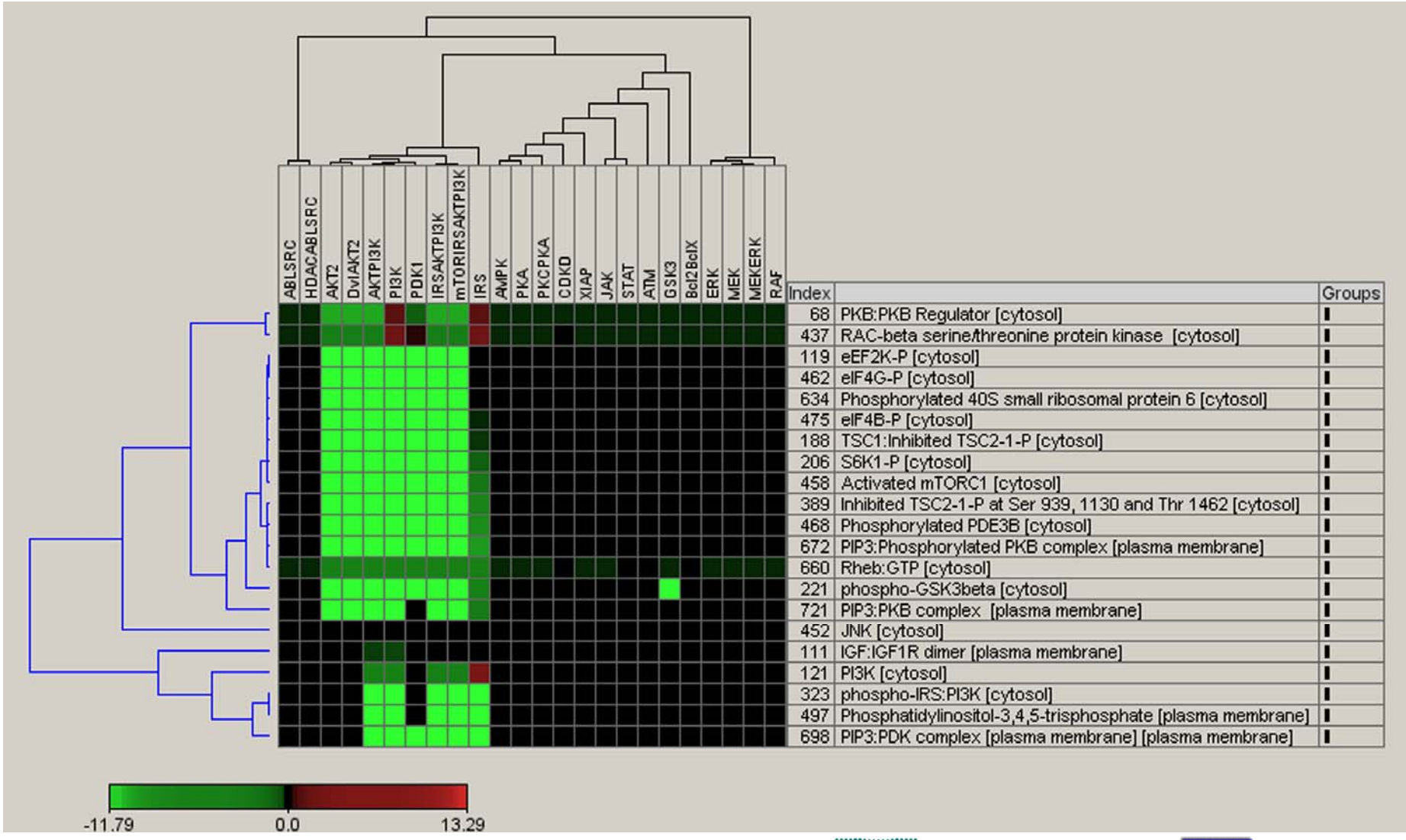

- most components not affected
- direct and indirect effects
- highest effectors with 10% of model components
- specificity of drugs different



Max-Planck-Institute for Molecular Genetics

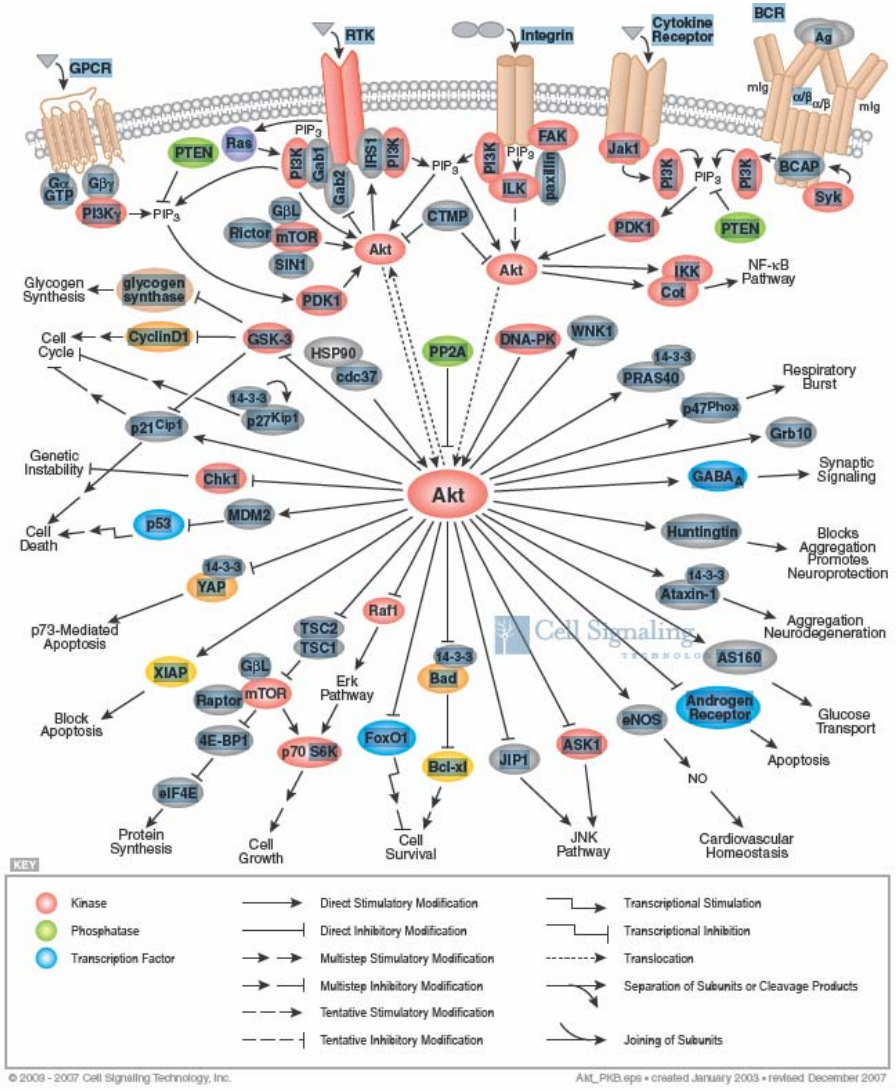
MBRACE Grid Network of excellence

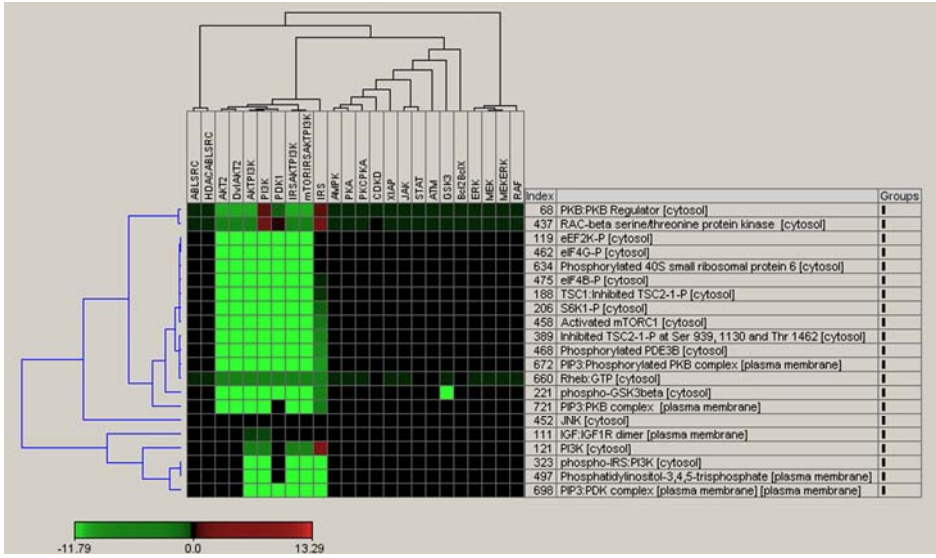
# Simulation results – clustering dependencies




| Index | | Groups |
|---|---|---|
| 68 | PKB:PKB Regulator [cytosol] | |
| 437 | RAC-beta serine/threonine protein kinase [cytosol] | |
| 119 | eEF2K-P [cytosol] | |
| 462 | eIF4G-P [cytosol] | |
| 634 | Phosphorylated 40S small ribosomal protein 6 [cytosol] | |
| 475 | eIF4B-P [cytosol] | |
| 188 | TSC1:Inhibited TSC2-1-P [cytosol] | |
| 206 | S6K1-P [cytosol] | |
| 458 | Activated mTORC1 [cytosol] | |
| 389 | Inhibited TSC2-1-P at Ser 939, 1130 and Thr 1462 [cytosol] | |
| 468 | Phosphorylated PDE3B [cytosol] | |
| 672 | PIP3:Phosphorylated PKB complex [plasma membrane] | |
| 660 | Rheb:GTP [cytosol] | |
| 221 | phospho-GSK3beta [cytosol] | |
| 721 | PIP3:PKB complex [plasma membrane] | |
| 452 | JNK [cytosol] | |
| 111 | IGF:IGF1R dimer [plasma membrane] | |
| 121 | PI3K [cytosol] | |
| 323 | phospho-IRS:PI3K [cytosol] | |
| 497 | Phosphatidylinositol-3,4,5-trisphosphate [plasma membrane] | |
| 698 | PIP3:PDK complex [plasma membrane] [plasma membrane] | |

# Simulation results – clustering dependencies



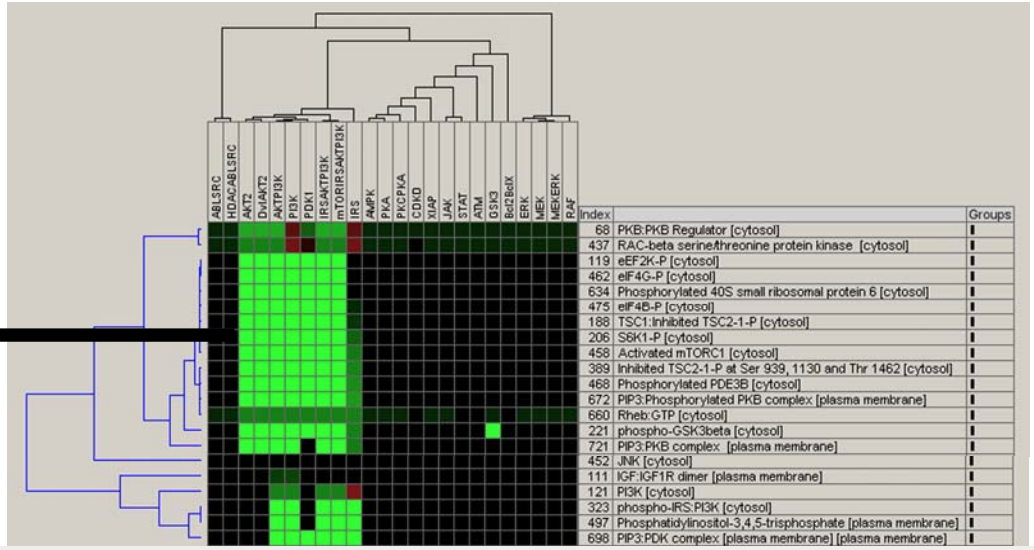




**EGEE 2008 Conference, Istanbul 22.09.2008**

Max-Planck-Institute for Molecular Genetics

MBRACE Grid Network of excellence

# 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.



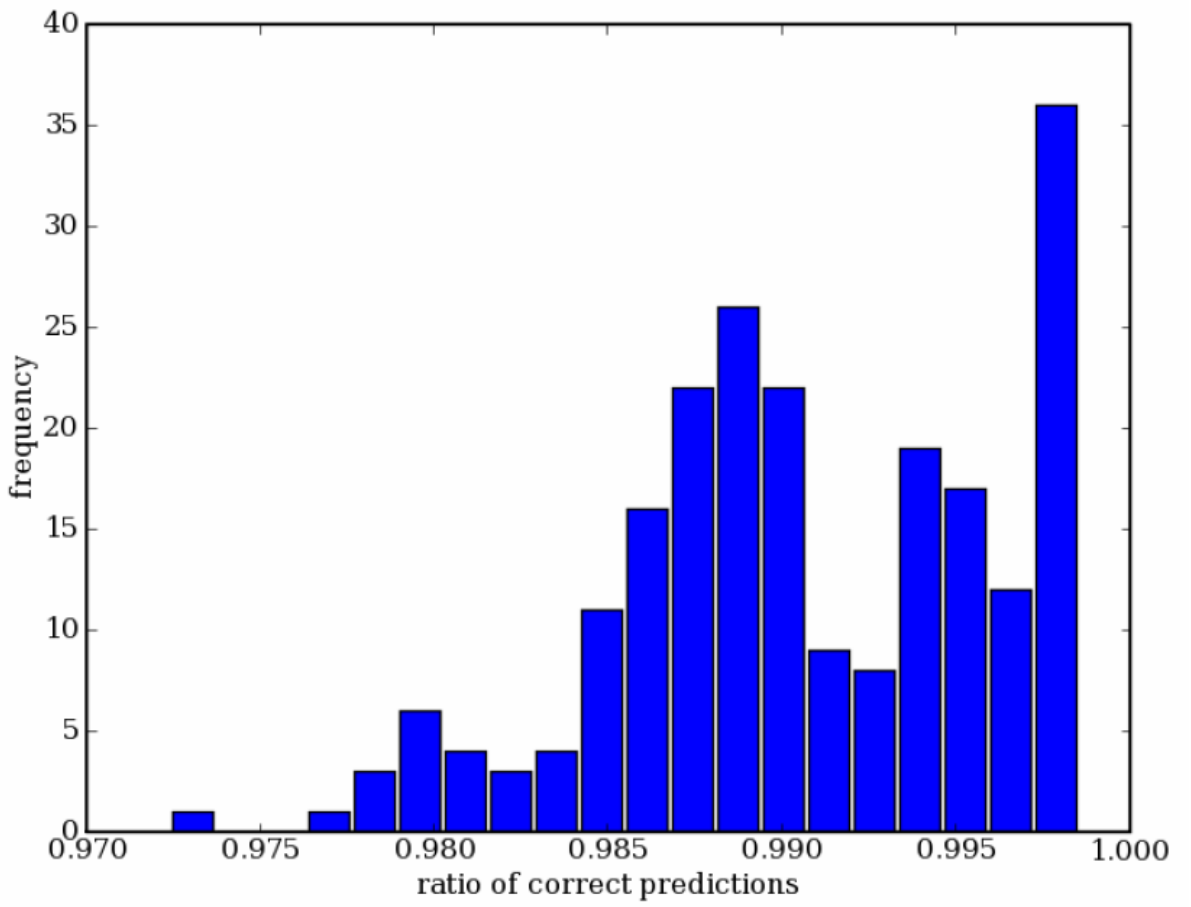
Feedback inhibition of the PI3K-Akt pathway by the mammalian target of rapamycin complex 1 (mTORC1) has emerged as an important signaling event in tumor syndromes, cancer, and insulin resistance. Cells lacking the tuberous sclerosis complex (TSC) gene products are a model for this feedback regulation. 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.



PMID: 17052453 [PubMed - indexed for MEDLINE]

PMCID: PMC1880887

# 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

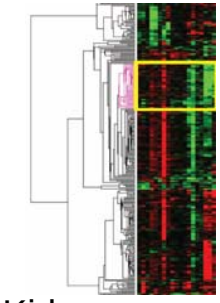
# Conclusions

- **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**

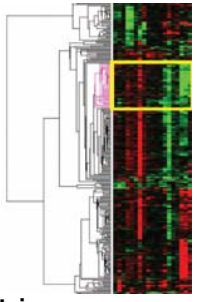
# Adaptation of generic models

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

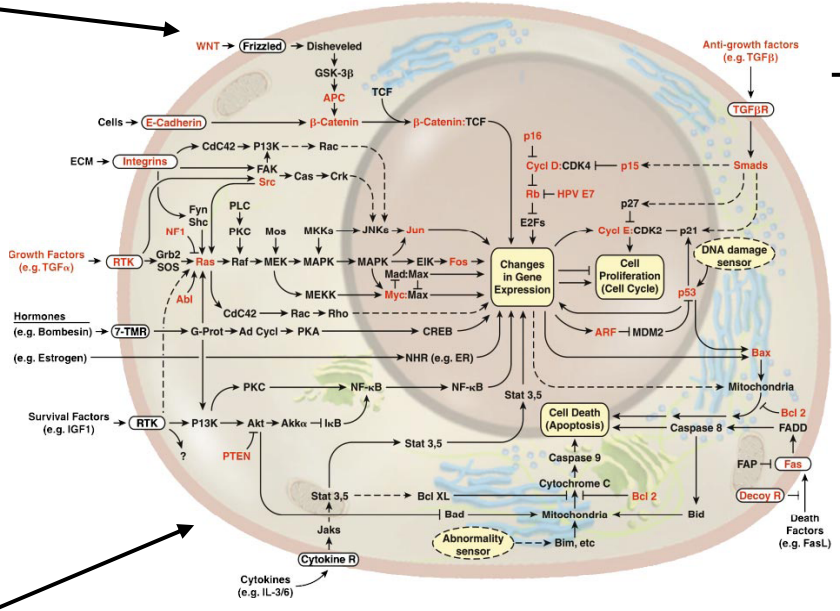
- **Unexpressed genes are deleted**
- **Levels of proteins are adjusted to expression levels**
- **Proteins levels are further tuned by their half life (if known)**



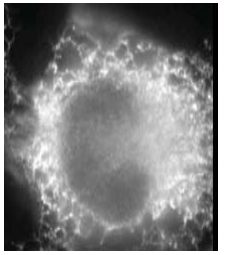
Kidney expression



Liver expression




Kidney cell



Liver cell

19. January 2008

# Acknowledgments


**MPIMG**

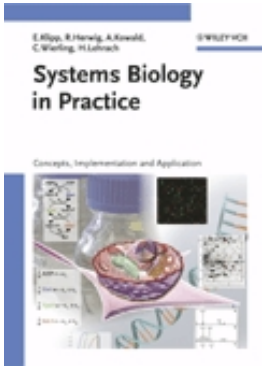
**Christoph Wierling**
**Andriani Daskalaki**
**Elisabeth Maschke-Dutz**
**Atanas Kamburov**
**Hans Lehrach**

**CNRS, Lyon**

**Christope Blanchet**

**EU projects**

**EMBRACE**
**CARCINOGENOMICS**
**APO-SYS**




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

**Thanks for your attention !**



Max-Planck-Institute
for Molecular Genetics