Systems medicine approach to model cell signaling activity uncovers disease mechanisms and predicts cancer outcome

Joaquín Dopazo

Clinical Bioinformatics research Area,
Fundación Progreso y Salud,
Functional Genomics Node, (INB-ELIXIR-es),
Bioinformatics in Rare Diseases (BiER-CIBERER),
Sevilla, Spain.

http://www.clinbioinfosspa.es
http://www. babelomics.org
@xdopazo

Máster en Bioinformática Aplicada a Medicina Personalizada y Salud, Madrid, 11 abril, 2018

Setting the problem in context:

The transition to precision medicine

Degree of personalization

Intuitive
Based on trial and error

Identification of probabilistic patterns

Decisions and actions based on knowledge

Intuitive Medicine **Empirical Medicine** **Systems Medicine**

Today

Tomorrow

Molecular biomarkers

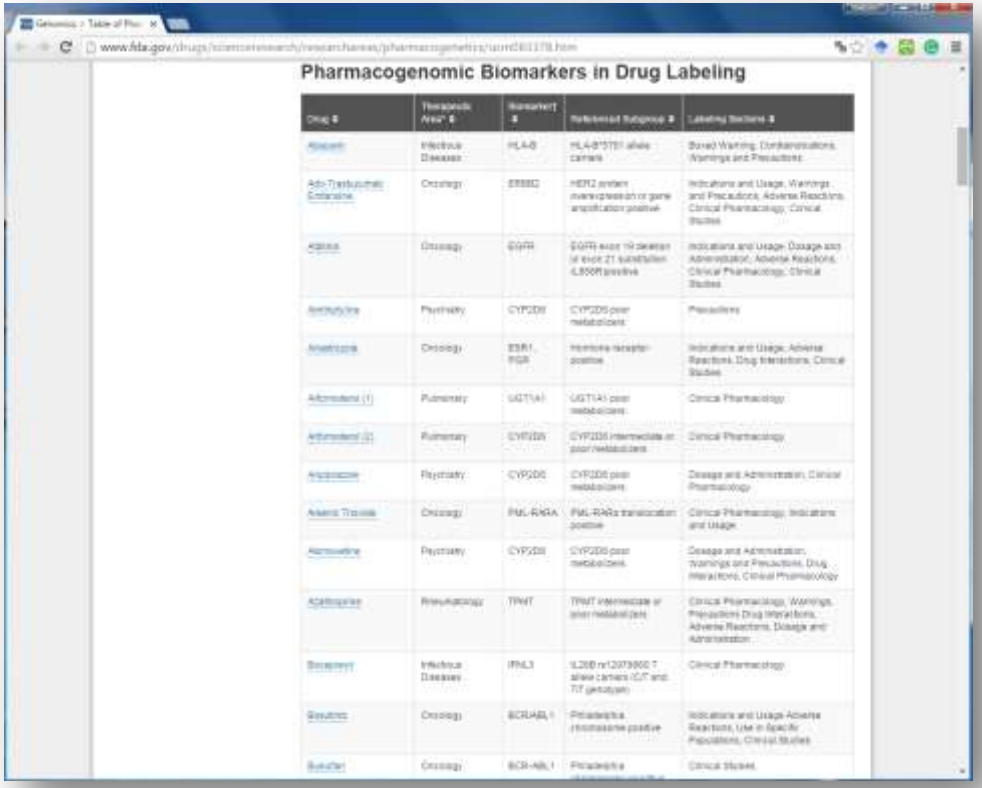
Genomic biomarkers

Mechanistic biomarkers

Precision medicine is based on a better knowledge of phenotype-genotype relationships. This ultimately involves the knowledge of **disease** and **drug action mechanisms**

Requires of a better way of defining diseases by introducing **genomic** technologies in the **diagnostic** procedures and **treatment decisions**

Single-gene biomarkers are the result of probabilistic associations



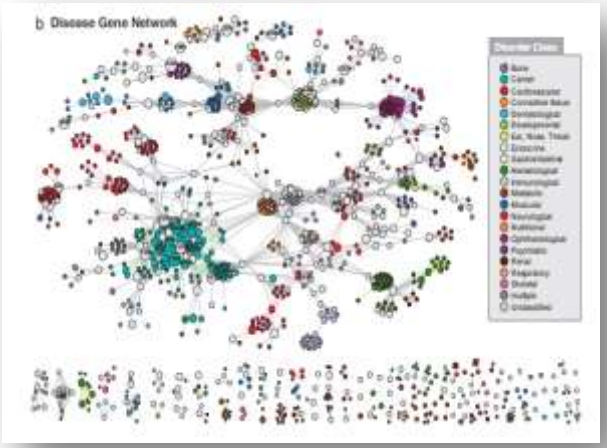
http://www.fda.gov/drugs/scienceresearch/researchareas/pharmacogenetics/ucm083378.htm

Most “personalized” therapies are based on this type of biomarkers

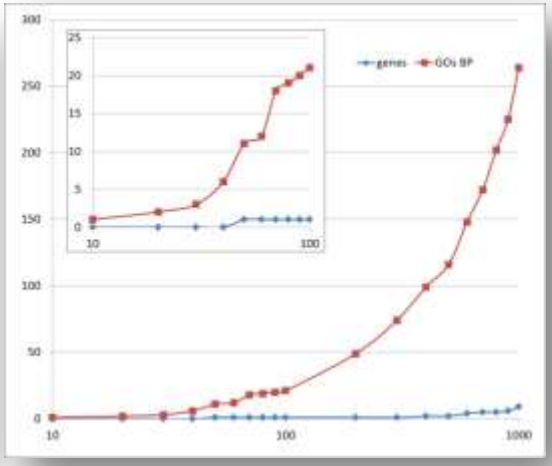
Despite most biomarkers used are single gene variants, most human genetic diseases (and almost all traits) have a modular nature

- Conventional single-gene biomarkers have a demonstrated clinical utility. However, their success is purely probabilistic, often modest and frequently lack any mechanistic anchoring to the fundamental cellular processes responsible for the disease or therapeutic response.
- **Modular nature of genetic diseases: Causative genes** for the same or phenotypically similar diseases may generally reside in the same **biological module**, either a **protein complex** (Lage et al, 2007), a **sub-network** of protein interactions (Lim et al, 2006) , or a **pathway** (Wood et al, 2007)

Disease genes are close in the interactome



Goh 2007 PNAS



Fernandez, 2013, Orphanet J Rare Dis.

Same disease in different populations is caused by different genes affecting the same functions

There are exceptions: MammaPrint, an example of successful breast cancer decision support test based on a multigenic biomarker

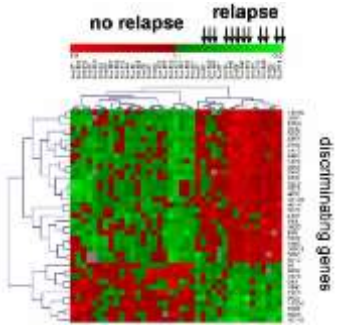
1

Finding genes

Gene expression profiling predicts clinical outcome of breast cancer

Laura J. van 't Veer*†, Hongyue Dai†‡, Marc J. van de Vijver*†, Yudong D. He‡, Augustinus A. M. Hart*, Mao Mao‡, Hans L. Peterse*, Karin van der Kooy*, Matthew J. Marton‡, Anke T. Witteveen*, George J. Schreiber‡, Ron M. Kerkhoven*, Chris Roberts‡, Peter S. Linsley‡, René Bernards* & Stephen H. Friend‡

* Divisions of Diagnostic Oncology, Radiotherapy and Molecular Carcinogenesis and Center for Biomedical Genetics, The Netherlands Cancer Institute, 121 Plesmanlaan, 1066 CX Amsterdam, The Netherlands
‡ Rosetta Inpharmatics, 12040 115th Avenue NE, Kirkland, Washington 98034, USA
† These authors contributed equally to this work



2

Assessing functions

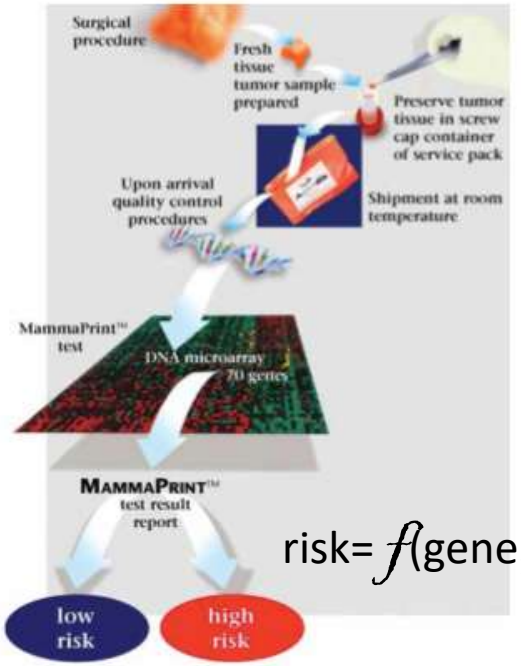
By historic reasons genes were first selected and their functionalities were assessed afterwards.

INSIGHT REVIEW

Enabling personalized cancer medicine through analysis of gene-expression patterns

Laura J. van 't Veer[1,2,3] & René Bernards[1,3,4,5]

Therapies for patients with cancer have changed gradually over the past decade, moving away from the administration of broadly acting cytotoxic drugs towards the use of more-specific therapies that are targeted to each tumour. To facilitate this shift, tests need to be developed to identify those individuals who require therapy and those who are most likely to benefit from certain therapies. In particular, tests that predict the clinical outcome for patients on the basis of the genes expressed by their tumours are likely to increasingly affect patient management, heralding a new era of personalized medicine.



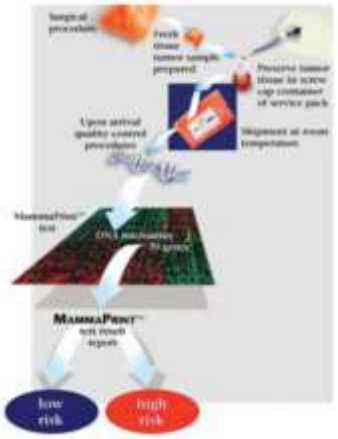
Biological Function	MammaPrint Gene Count
Metabolism	7
Cell cycle and DNA replication	12
Extracellular matrix adhesion and remodeling	5
Growth, proliferation, transformation and cell death	17
General signal transduction and intracellular transport	8
Growth factor	7
Motility or actin filament related	5
Intracellular hydrolase	1
Immune response	1
Neuropeptide	1
Predicted transmembrane protein with unknown function	2
Predicted transcriptional control or DNA binding proteins	5
Unknown function	6
Total Gene Count	70

The strength of this approach is that it is unbiased: there are no assumptions about which genes are likely to be involved in the process of interest. For example, in a **data-driven** study of the prognosis of patients with breast cancer, **little was known about the function** of 15 of the 70 genes that were found to constitute a prognostic gene-expression signature[4]. A drawback of this approach is that the **outcome relies solely on the quality of the data** (and the samples).
By contrast, using the **knowledge-driven** approach, genes that are thought to be relevant to a particular cancer trait are selected on the basis of the scientific literature.

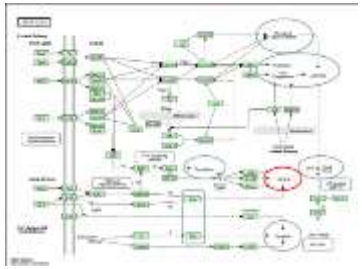
Risk is calculated as a function of the 70 gene expression levels

$$risk = f(gene_1, gene_2, \ldots gene_{70})$$

Change in the paradigm

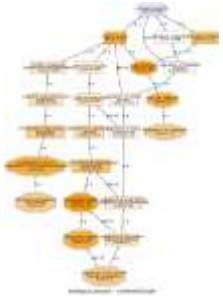


MammaPrint and other multigenic biomarkers: bottom up, from genes to functions that define one (or several) biological modules.



Models of cell functionality: top-down mechanism-based biomarkers, from biological modules to genes

Two problems: defining functional modules and modeling their behavior



Definition

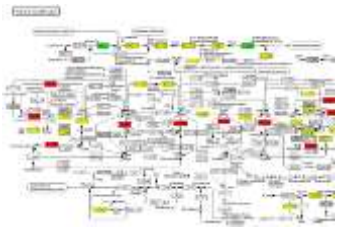
Behavior

Gene ontology: descriptive; unstructured functional labels

Enrichment methods. GO, etc. (simple statistical tests). No information on how components relate among them

Interactome: relationships among components but unknown function

Connectivity models. Protein-protein, protein-DNA and protein-small molecule interactions (tests on network properties). No information the functional roles of the components



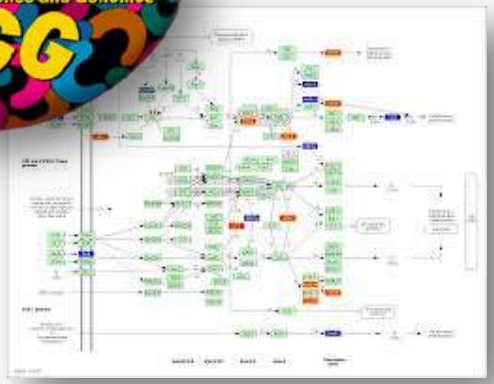
Pathways: relationships among components and their functional roles

Mathematical models. Kinetic models including stoichiometry, balancing reactions, etc.

Computational models. Models of signalling pathways, metabolic pathways, regulatory pathways, etc. (executable models)

Defining the module: Pathways: maps of cell activity (in sickness and in health)

www.genome.jp/kegg



disease-maps.org/

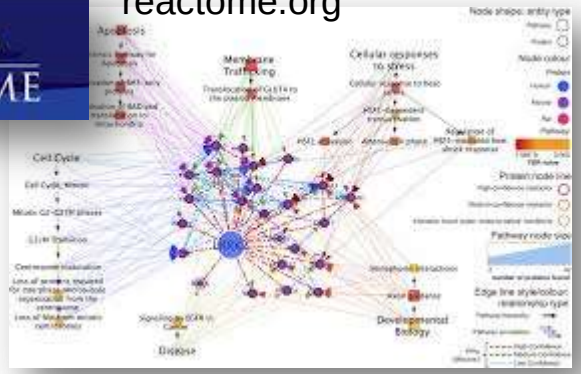
Oncogénesis

Alzheimer

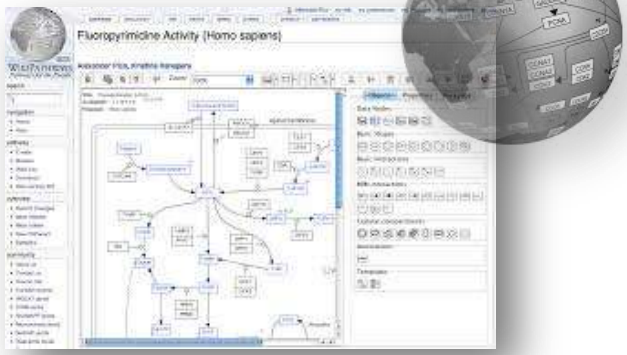
More disease maps…

Parkinson

reactome.org



www.wikipathways.org/

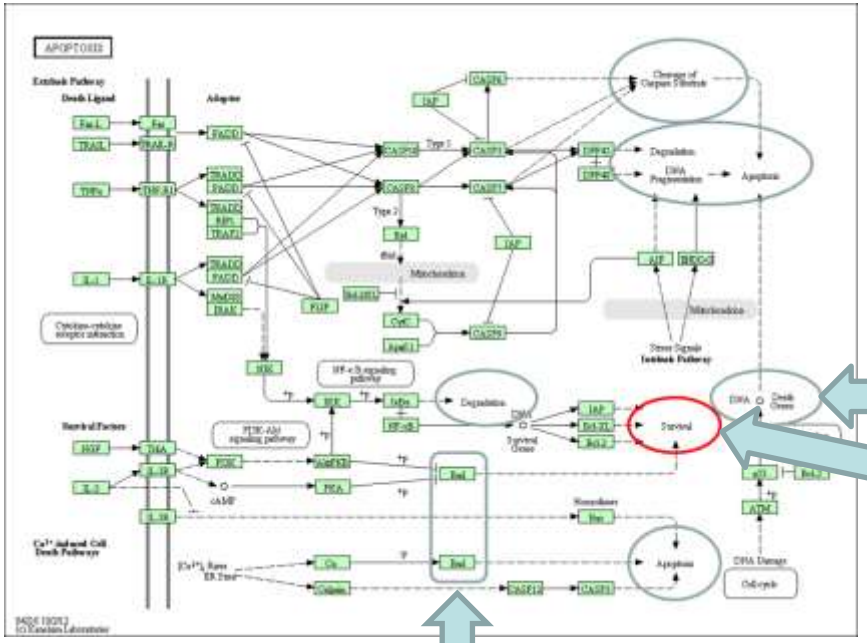


Defining pathway activity

We first need a map: pathways are defined in different repositories (**KEGG**, Reactome, Biocarta, disease maps, etc.)

What pathway level makes a real biological meaning?

Gene ⇨ sub-pathway ⇨ pathway



Enrichment methods (pathway-level): Different and often opposite cell behaviors are triggered by the same **pathway**.
E.g.: death and survival

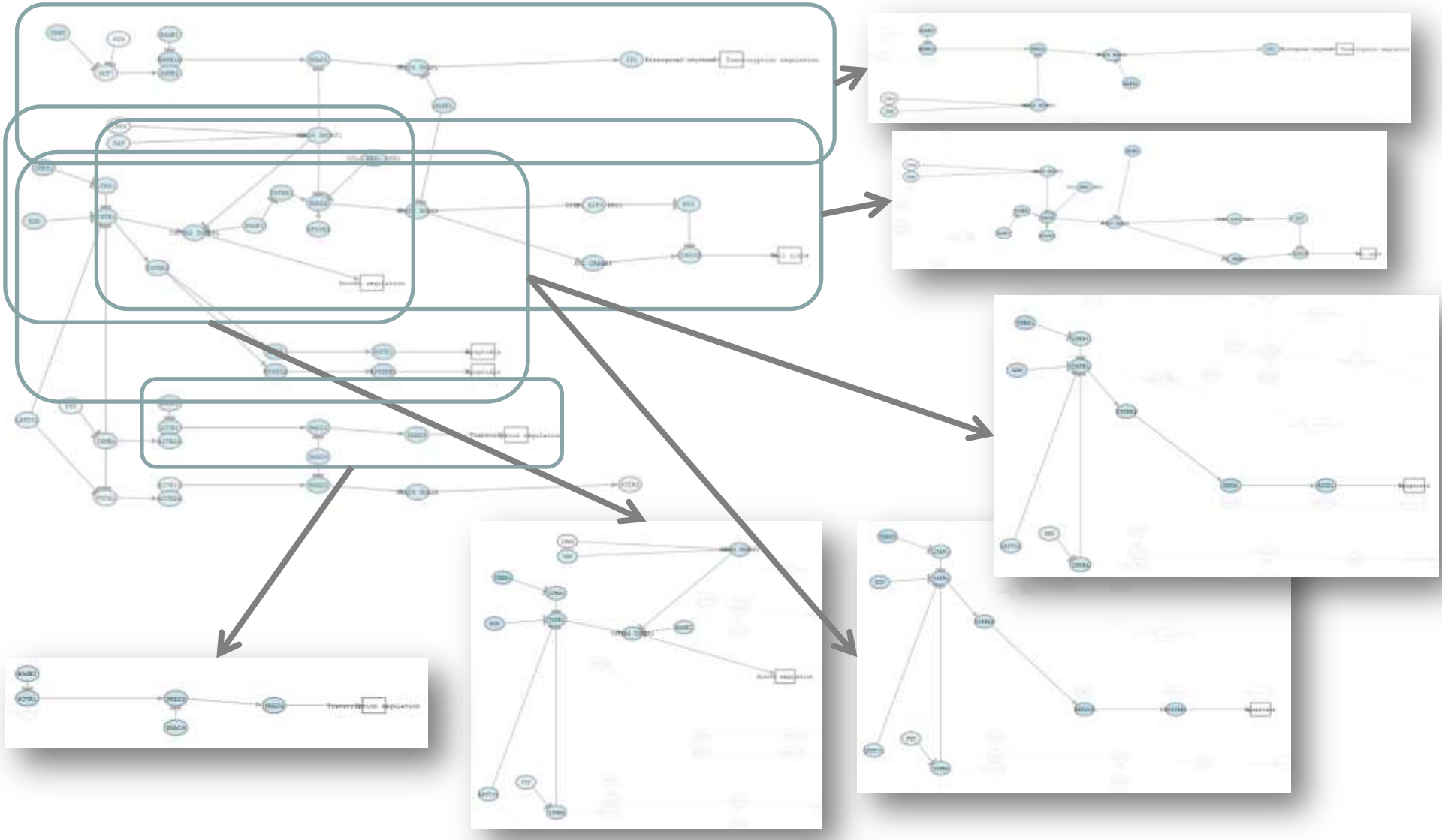
Death

Survival

Gene level: The same gene can trigger different (and often opposite) responses, depending on the stimulus

Sub-pathway (elementary circuit) connects stimulus to response

Decomposition of a pathway into their elementary circuits



How realistic are models of pathway activity?

RESEARCH ARTICLE

CANCER

Signaling pathway models as biomarkers: Patient-specific simulations of JNK activity predict the survival of neuroblastoma patients

Dirk Fey,[1] Melinda Halasz,[1] Daniel Dreidax,[2] Sean P. Kennedy,[1] Jordan F. Hastings,[3] Nora Rauch,[1] Amaya Garcia Munoz,[1] Ruth Pilkington,[1] Matthias Fischer,[4,5,6] Frank Westermann,[2] Walter Kolch,[1,7,8] Boris N. Kholodenko,[1,7,8]* David R. Croucher[1,3,9]*

Signaling pathways control cell fate decisions that ultimately determine the behavior of cancer cells. Therefore, the dynamics of pathway activity may contain prognostically relevant information different from that contained in the static nature of other types of biomarkers. To investigate this hypothesis, we characterized the network that regulated stress signaling by the c-Jun N-terminal kinase (JNK) pathway in neuroblastoma cells. We generated an experimentally calibrated and validated computational model of this network and used the model to extract prognostic information from neuroblastoma patient–specific simulations of JNK activation. Switch-like JNK activation mediates cell death by apoptosis. An inability to initiate switch-like JNK activation in the simulations was significantly associated with poor overall survival for patients with neuroblastoma with or without *MYCN* amplification, indicating that patient-specific simulations of JNK activation could stratify patients. Furthermore, our analysis demonstrated that extracting information about a signaling pathway to develop a prognostically useful model requires understanding of not only components and disease-associated changes in the abundance or activity of the components but also how those changes affect pathway dynamics.

Beyond static biomarkers—The activity of signalling networks as an alternate biomarker?

Fey et al., Sci. Signal. 8, ra130 (2015).

Inability of JNK activation (that mediates apoptosis) is associated to bad prognostic, irrespective of *MYCN* amplification status

Problem: ODE can efficiently solve only small systems

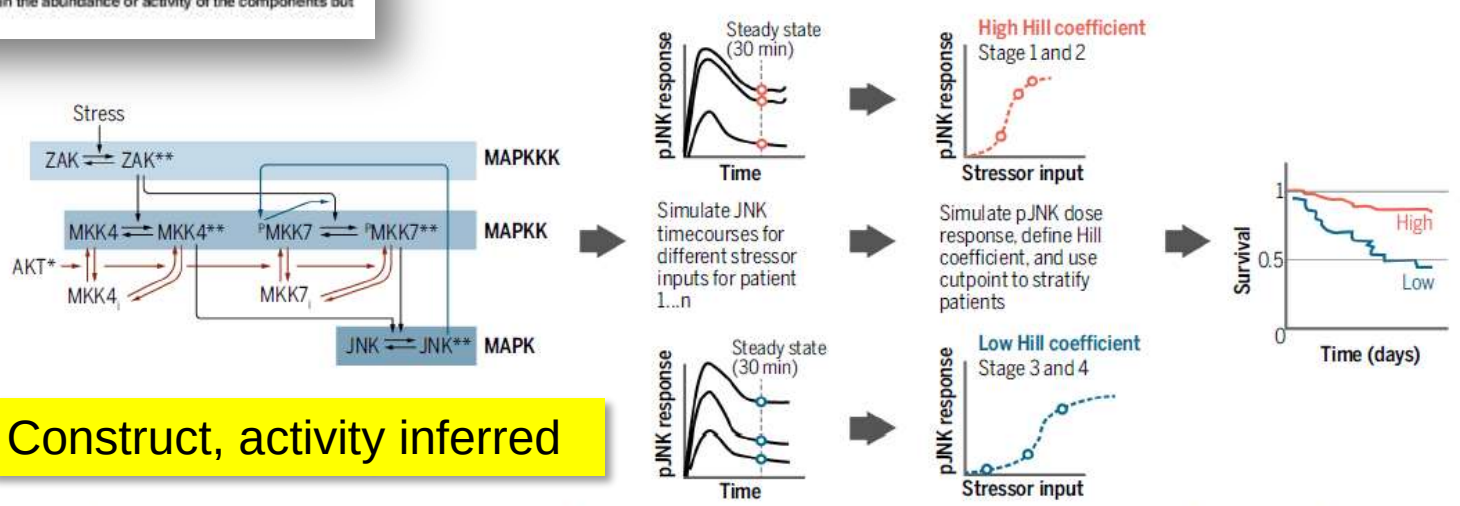
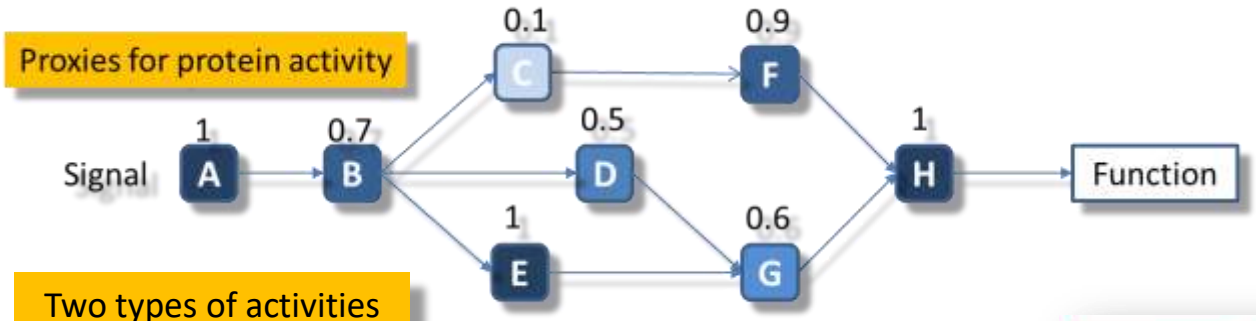
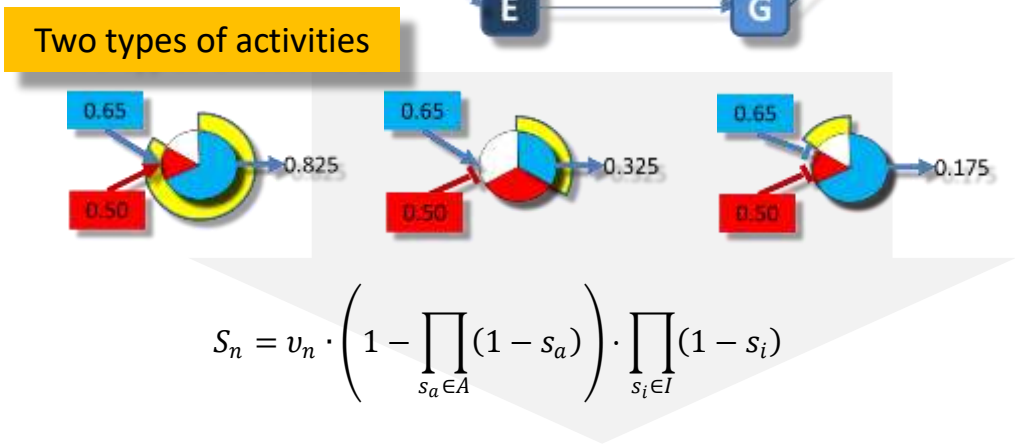


Fig. 1. Using network descriptors of signaling pathway activation potential to predict patient response. After construction of a computational model based on the validated network topology and that reproduces the signaling pathway dynamics, the model can be used to identify network descriptors, such as the Hill coefficient, that are calculated from the dynamic simulation of the activation of a signaling pathway. These in silico biomarkers cannot be directly measured.

Signal propagation models of signaling pathways



From individual gene expression profiles…




www.impactjournals.com/oncotarget/ Oncotarget, Advance Publications 2016

High throughput estimation of functional cell activities reveals disease mechanisms and predicts relevant clinical outcomes

Marta R. Hidalgo[1], Cankut Cubuk[1], Alicia Amadoz[1,2], Francisco Salavert[2,3], José Carbonell-Caballero[1], Joaquin Dopazo[1,2,3]

[1]Computational Genomics Department, Centro de Investigación Príncipe Felipe (CIPF), Valencia, 46012, Spain

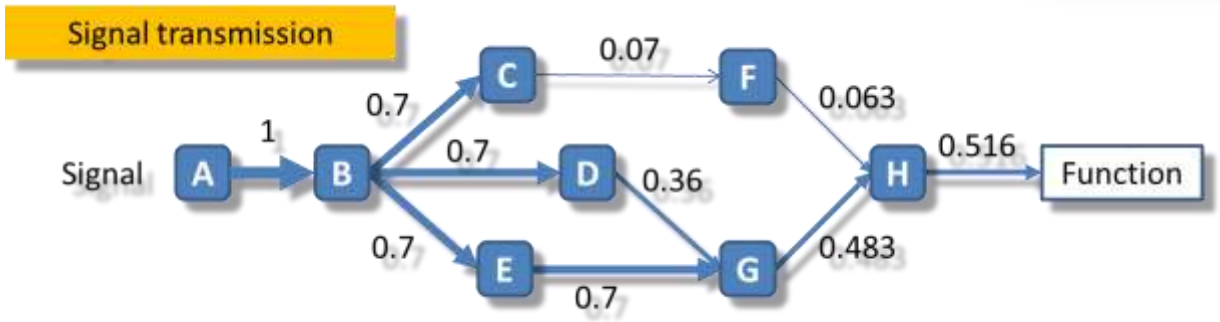
[2]Functional Genomics Node (INB-ELIXIR-es), Valencia, 46012, Spain

[3]Bioinformatics in Rare Diseases (BIER), Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Valencia, 46012, Spain

Correspondence to: Joaquin Dopazo, **email:** jdopazo@cipf.es

Keywords: signaling pathway, disease mechanism, prognostic, survival, biomarker

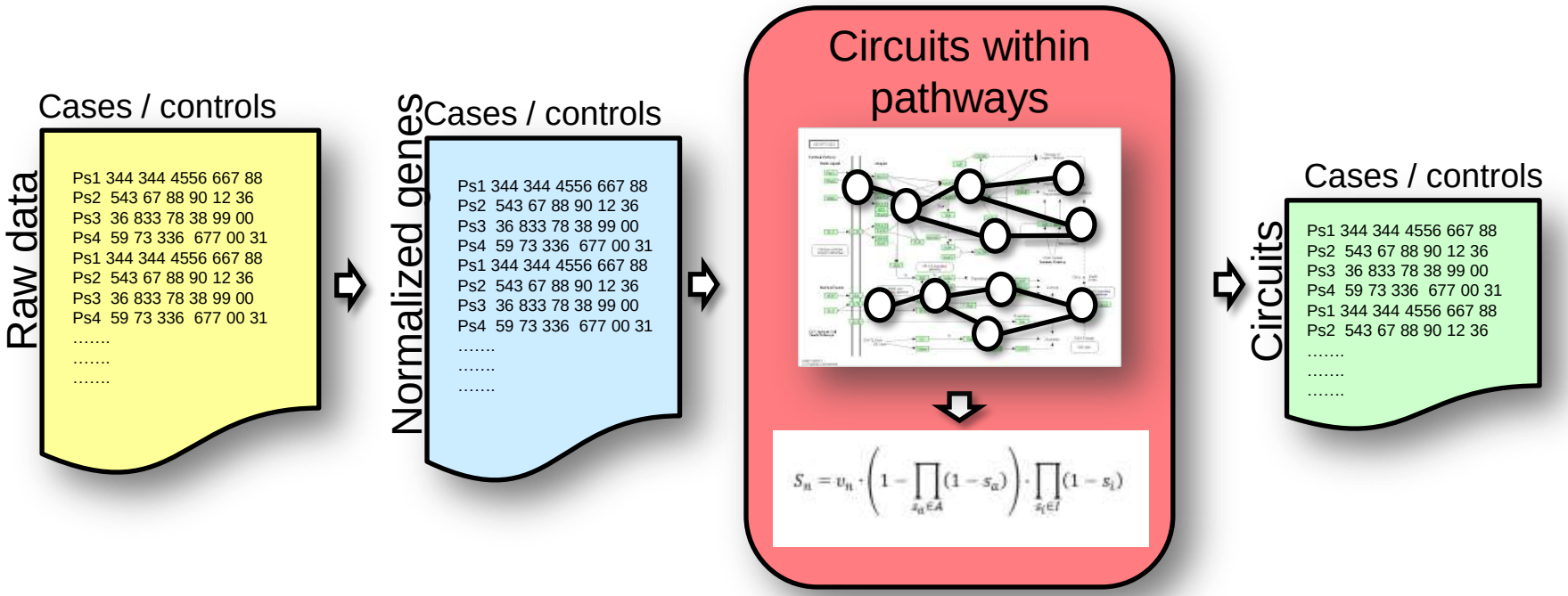
Received: September 01, 2016 **Accepted:** November 21, 2016 **Published:**



…to profiles of circuit activity (and functional activity)

Are scalable

Gene expression data are transformed into signal activity intensities

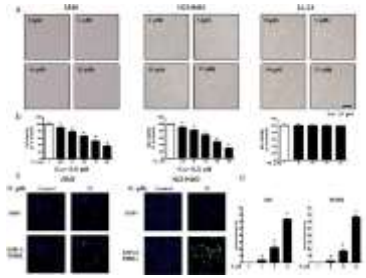


A simple transformation of raw data (normalization) and an algorithm for signal propagation results in accurate estimations of circuit activities.

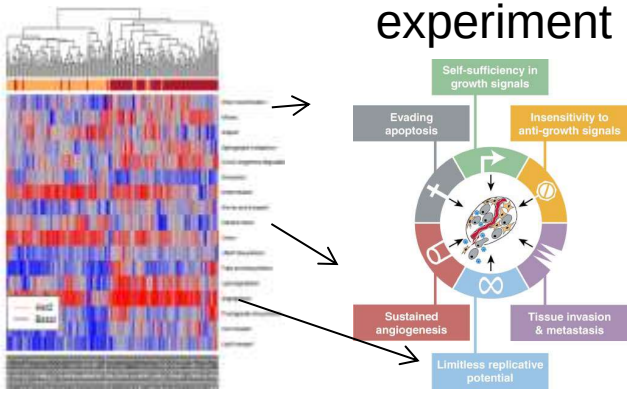
¡ The same concept that MammaPrint, risk= f(gene$_1$, gene$_2$, … gene$_{70}$) but based on biological knowledge, is used here to estimate cell functional activity !

Models of signaling activity provide high-throughput estimations of intensity activation of cell functions from gene expression measurements

Some (not all) conventional cell function can be studied, one at a time, in individual assays

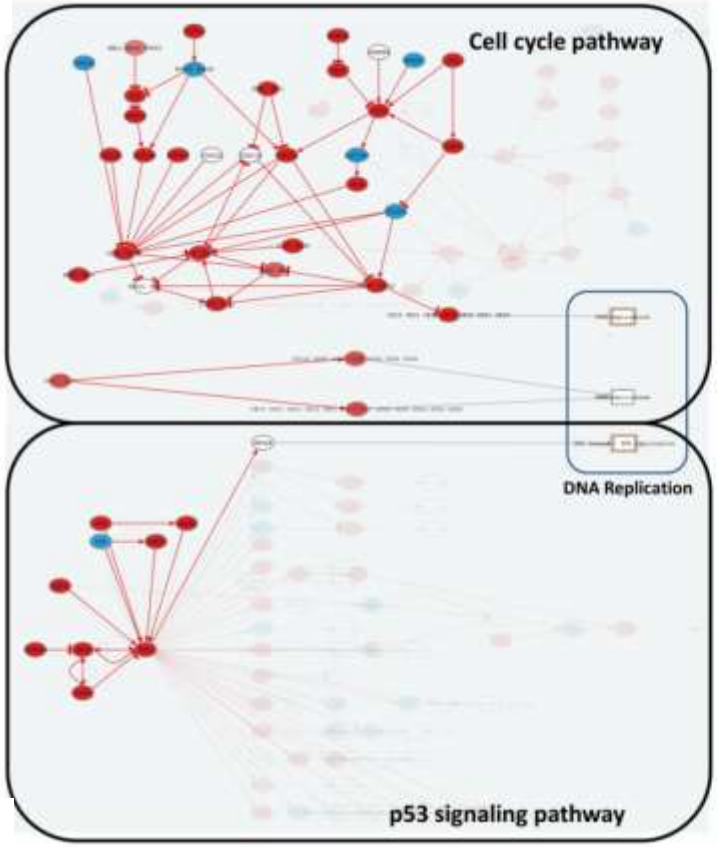


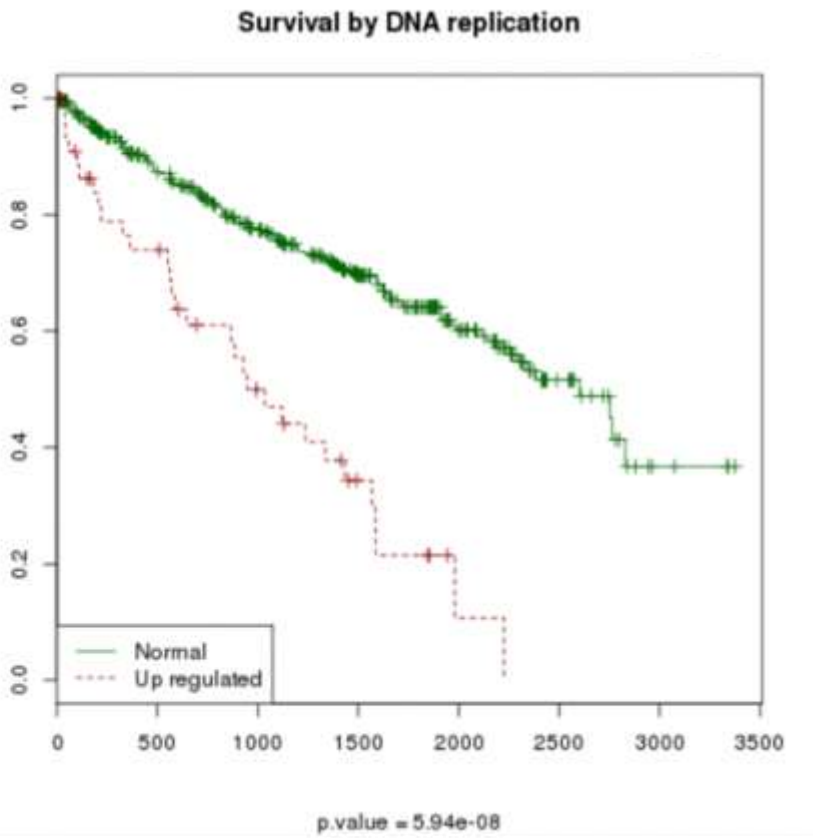
With Signaling models, the intensity at which the whole repertoire of cell functions is triggered can be measured in only one individual experiment



Hypothesis: the intensity at which functions are triggered by the signaling system of the cell is more related to phenotypes than the intensity of gene expression

Signaling activity trigger cell functions directly related to cancer progression



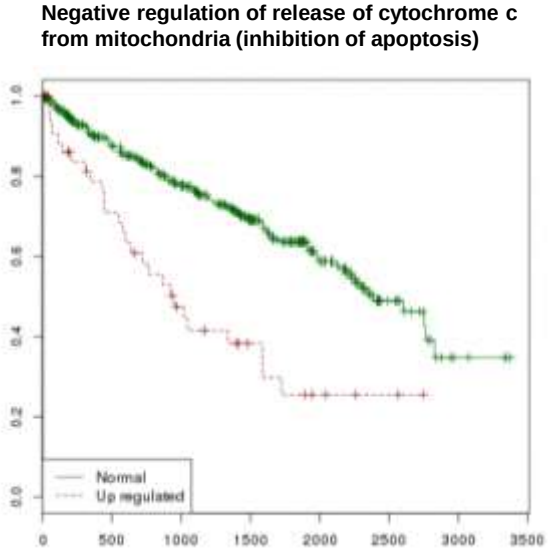


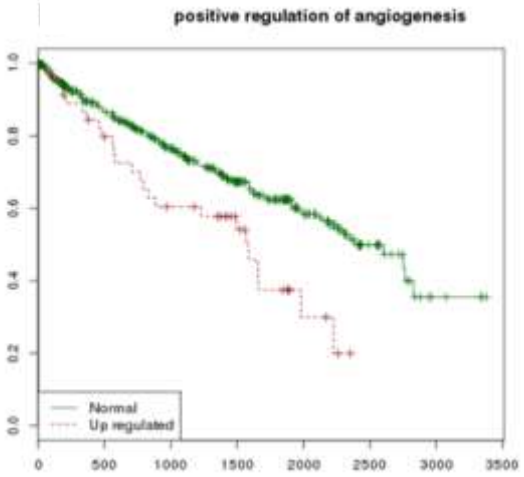
Hidalgo et al., 2017 Oncotarget

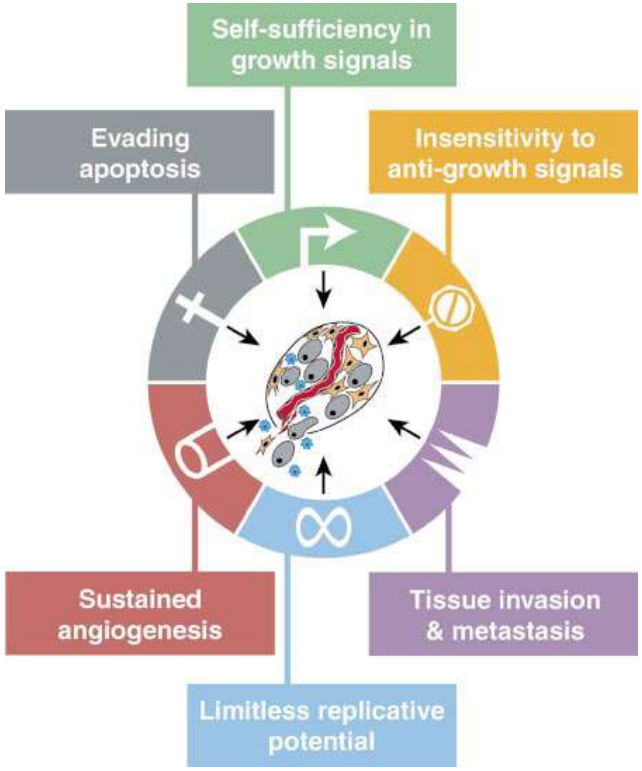
DNA replication= f(gene$_1$, gene$_2$, … gene$_n$)

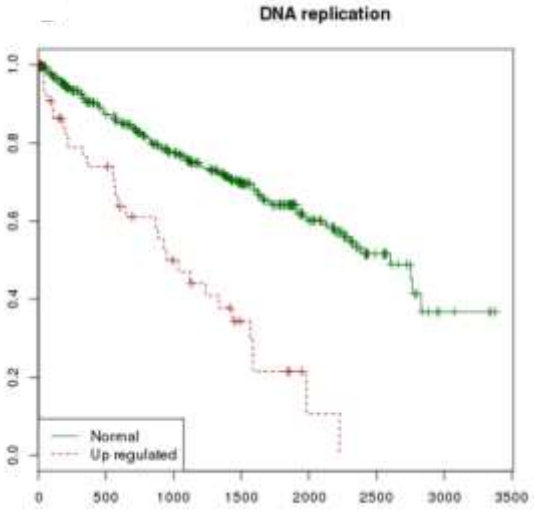
DNA replication function is a **construct**: the activity is **inferred** not measured

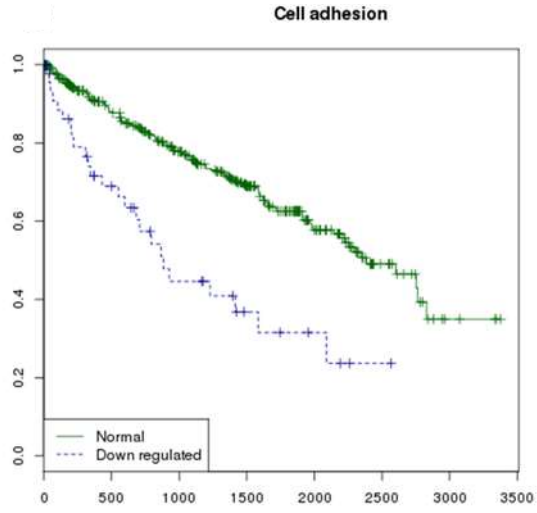
Actually, signal activity triggers all the cancer hallmarks





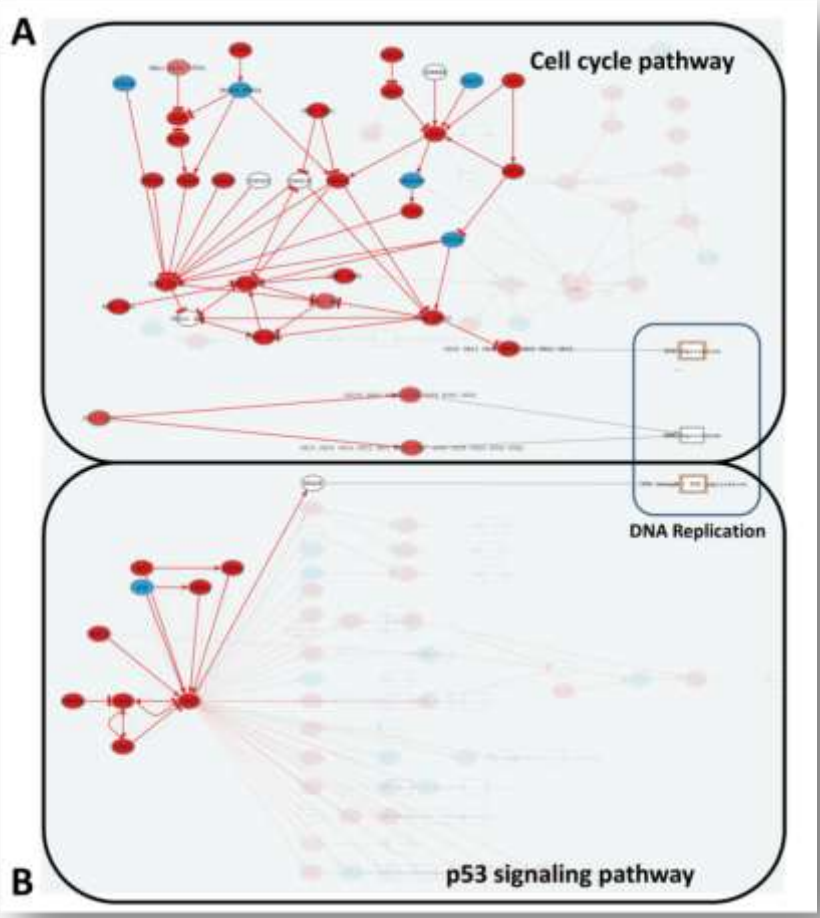


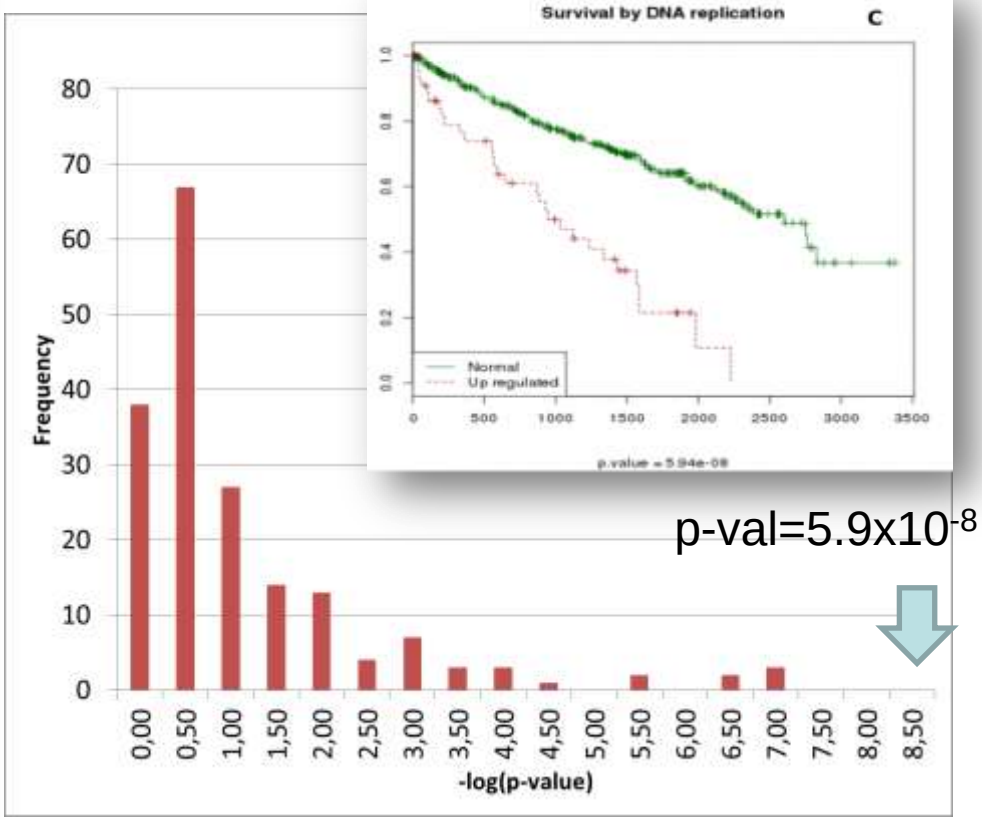




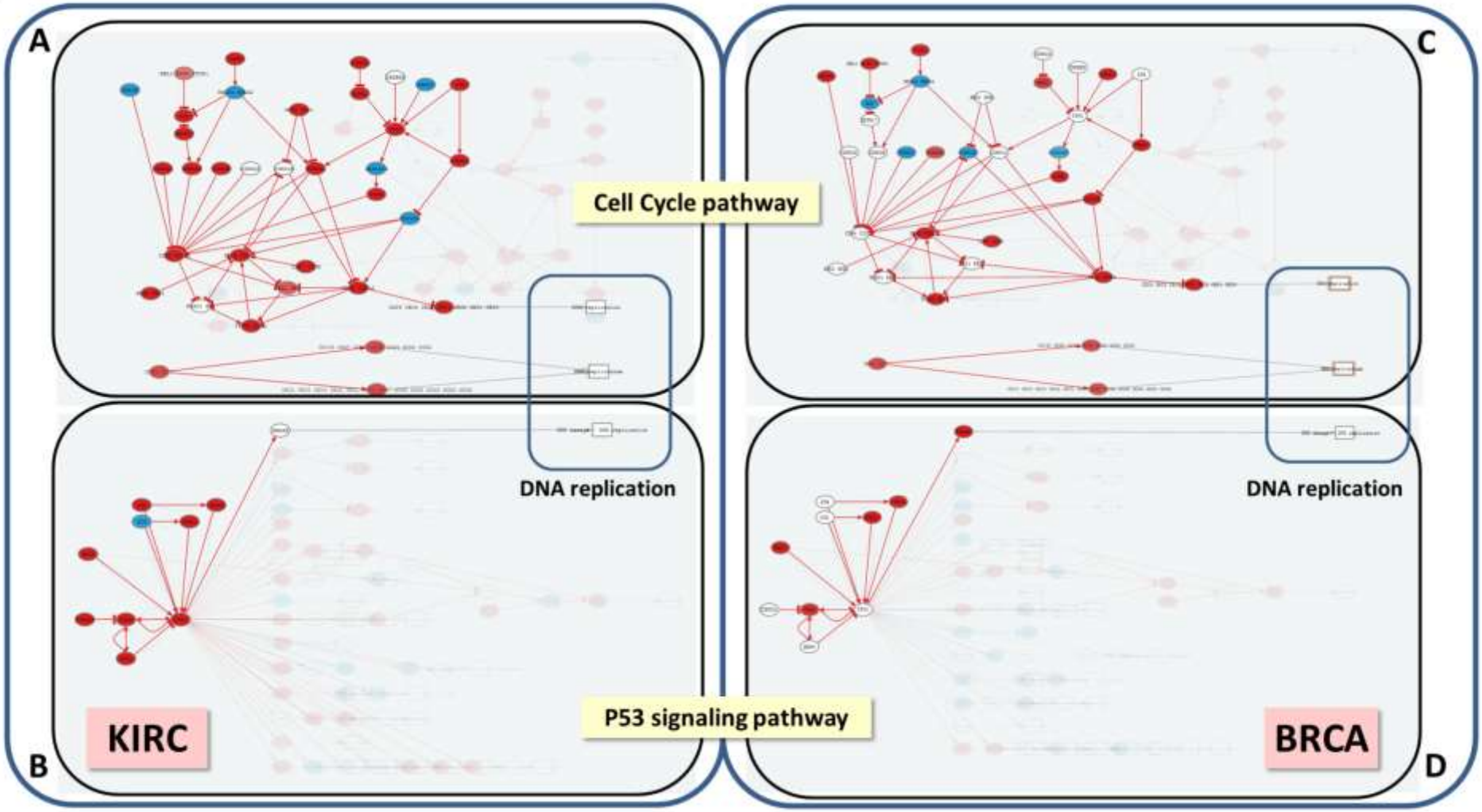
Hanahan, Weinberg, 2011
Hallmarks of cancer: the next generation. Cell 144, 646

The inferred function activity (mechanistic biomarker) is more correlated to survival than the activity of any gene (conventional biomarker) in the circuit

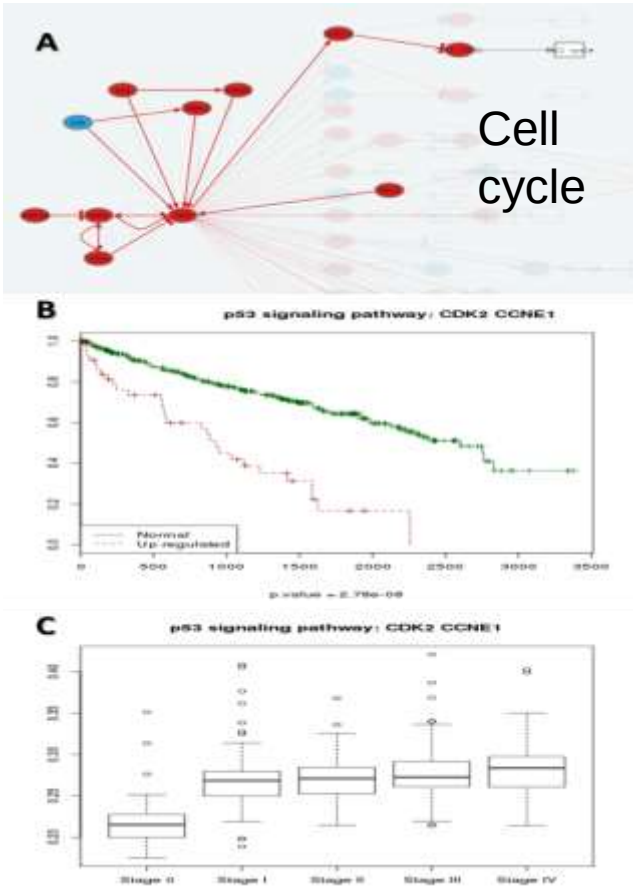




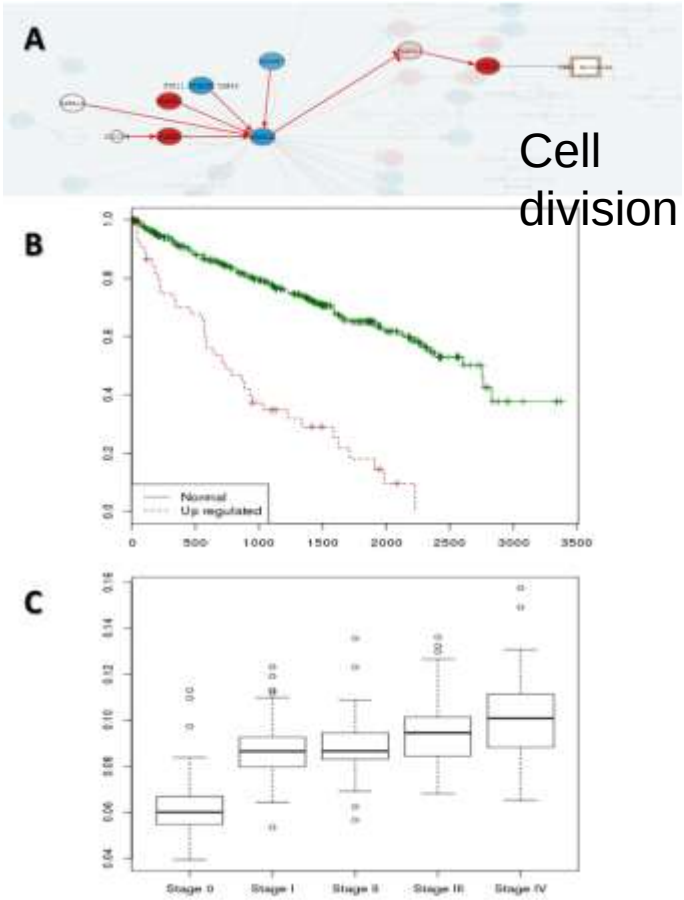
Different cancer use different gene expression programs to activate the same functions



Signal intensity over certain functions increases in the initiation of cancer while on others increases with cancer stage



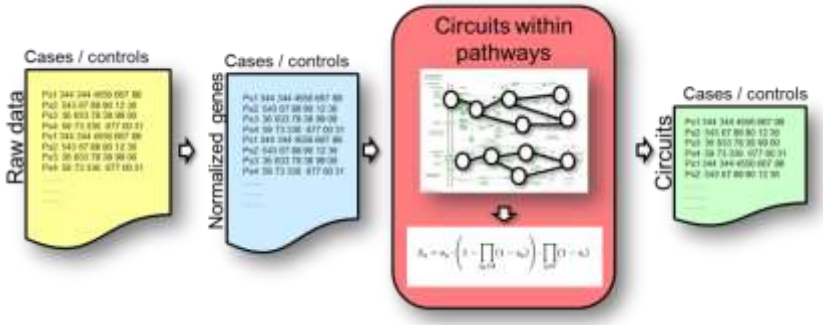
Cancer initiation



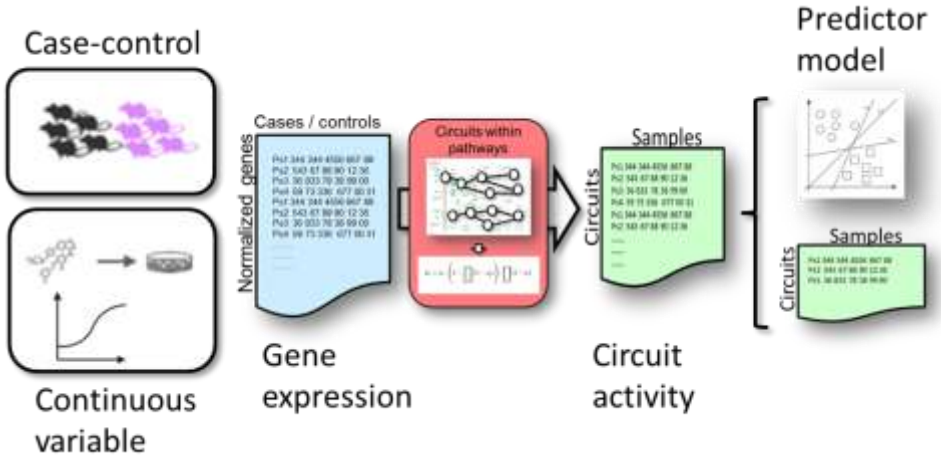
Cancer progression

Circuit activation probabilities can be used as features for predictors

1. Generation of features: signaling circuit activities

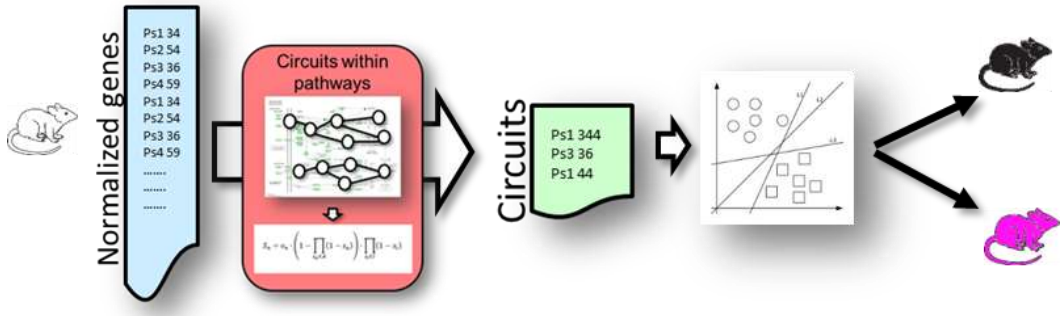


2. Training set

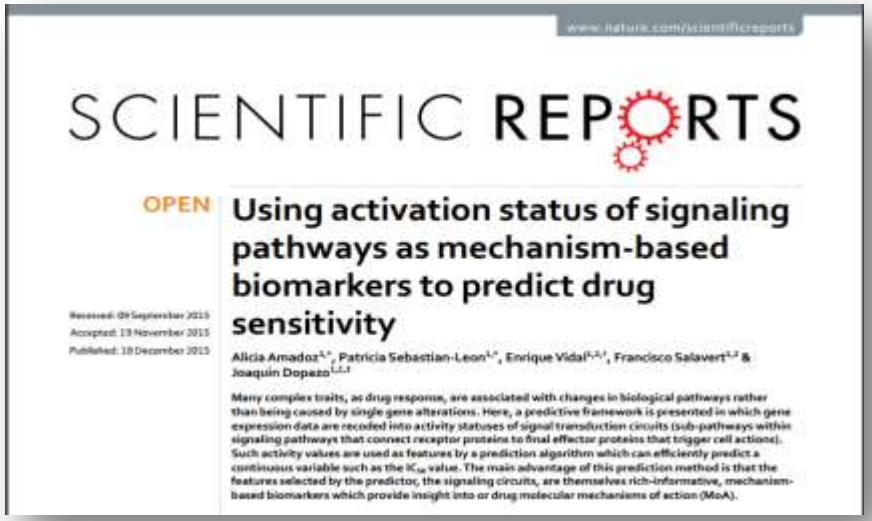


Circuit activation probabilities are mechanism-based biomarkers

3. Prediction



Prediction of IC_{50} values from the activity of signaling circuits



www.nature.com/scientificreports

SCIENTIFIC REPORTS

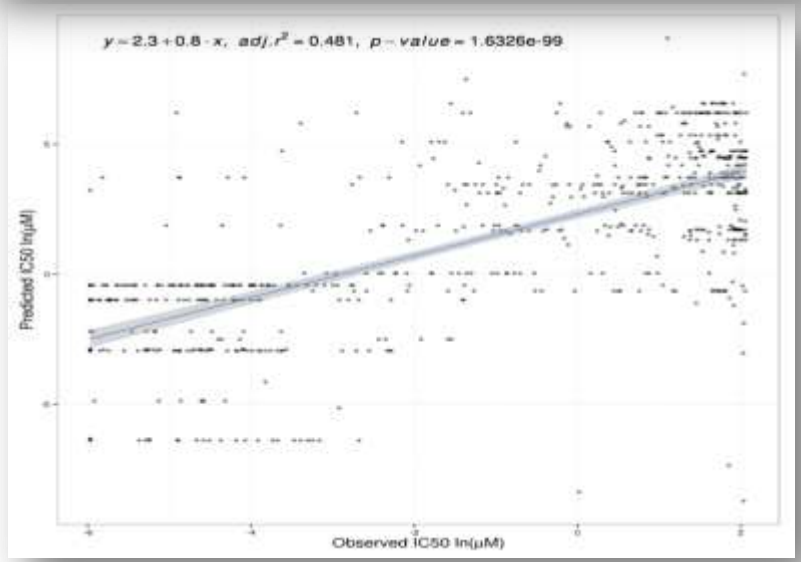
OPEN

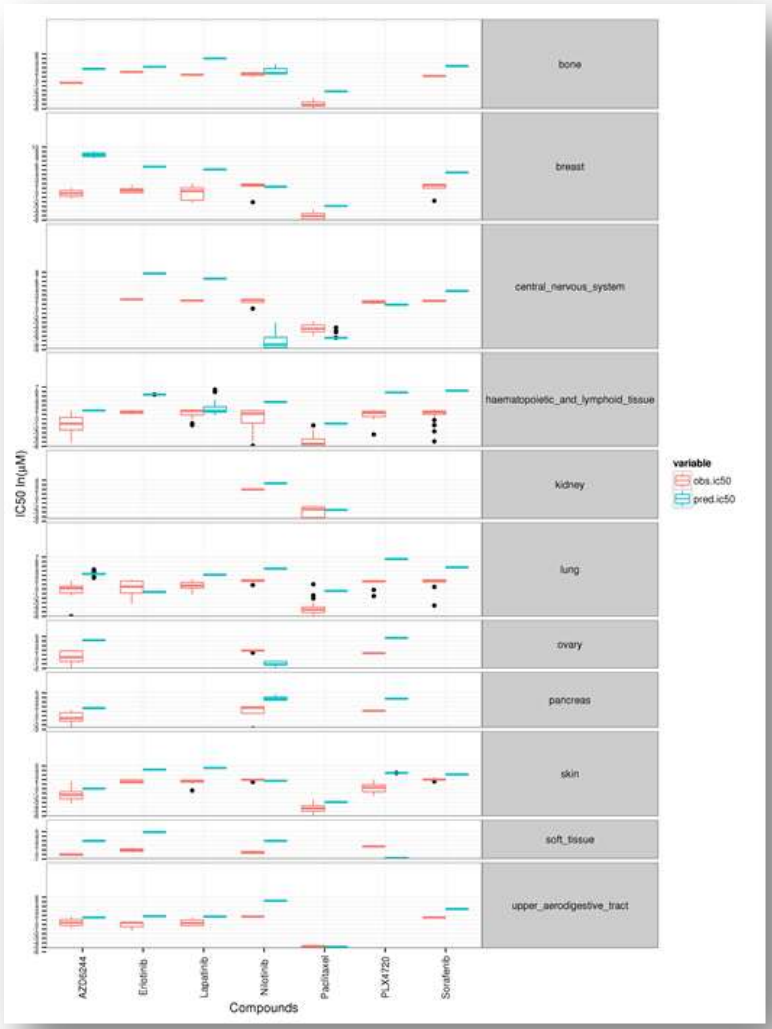
Using activation status of signaling pathways as mechanism-based biomarkers to predict drug sensitivity

Received: 09 September 2015
Accepted: 13 November 2015
Published: 18 December 2015

Alicia Amadoz[1,*], Patricia Sebastian-Leon[1,*], Enrique Vidal[1,2,†], Francisco Salavert[1,2] & Joaquin Dopazo[1,2,†]

Many complex traits, as drug response, are associated with changes in biological pathways rather than being caused by single gene alterations. Here, a predictive framework is presented in which gene expression data are recoded into activity statuses of signal transduction circuits (sub-pathways within signaling pathways that connect receptor proteins to final effector proteins that trigger cell actions). Such activity values are used as features by a prediction algorithm which can efficiently predict a continuous variable such as the IC_{50} value. The main advantage of this prediction method is that the features selected by the predictor, the signaling circuits, are themselves rich-informative, mechanism-based biomarkers which provide insight into or drug molecular mechanisms of action (MoA).





Amadoz et al., Sci. Rep. 2015

Actionable models

The real advantage of models is that, the same way they can be used to convert omics data into measurements of cell functionality that provide information on disease mechanisms and drug MoA, they can be used to test hypothesis such as "*what if I suppress (or over-express) this (these) gen(es)?*" This lead to the concept of **actionable models**.

By **simulating** changes of gene expression/activity it is easy to:

- Directly study of the consequences of induced gene over-expressions or KOs
- Carry out reverse studies of genes that need to be perturbed to change cell functionalities, such as:
  - Reverting the "normal" functional status of a cell
  - Selectively kill diseased cells without affecting normal cells
  - Enhancing or reducing cell functionalities (e.g., apoptosis or proliferation, respectively, to fight cancer)
  - Etc.

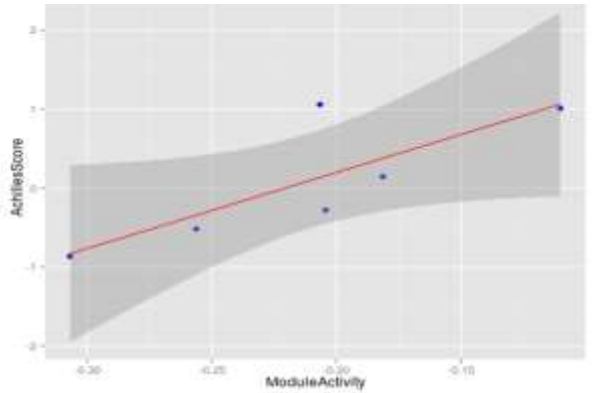
Model validation (1)

The activity of some signaling circuits is correlated with cell survival

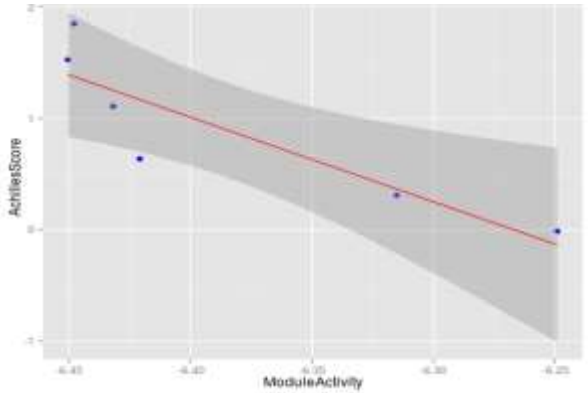
Survival data from Achilles cell line KOs (Broad Institute) can be compared to the change in circuit activities **predicted** by the model

Increased survival

Onco-circuit



Tumor suppressor circuit



Increased pathway activity

Functions:

- Repression of apoptosis
- Activation of proliferation

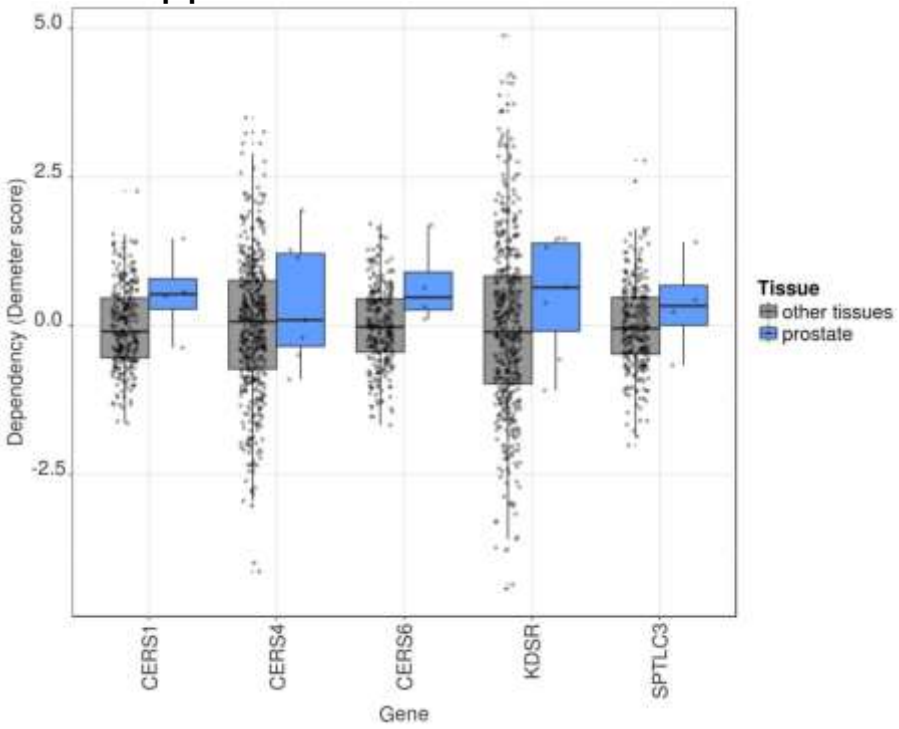
- Activation of apoptosis
- Repression of proliferation

Essential circuits: once found, other ways of deactivating these circuits can be find, opening the door to knowledge-based target discovery

Model validation (2)

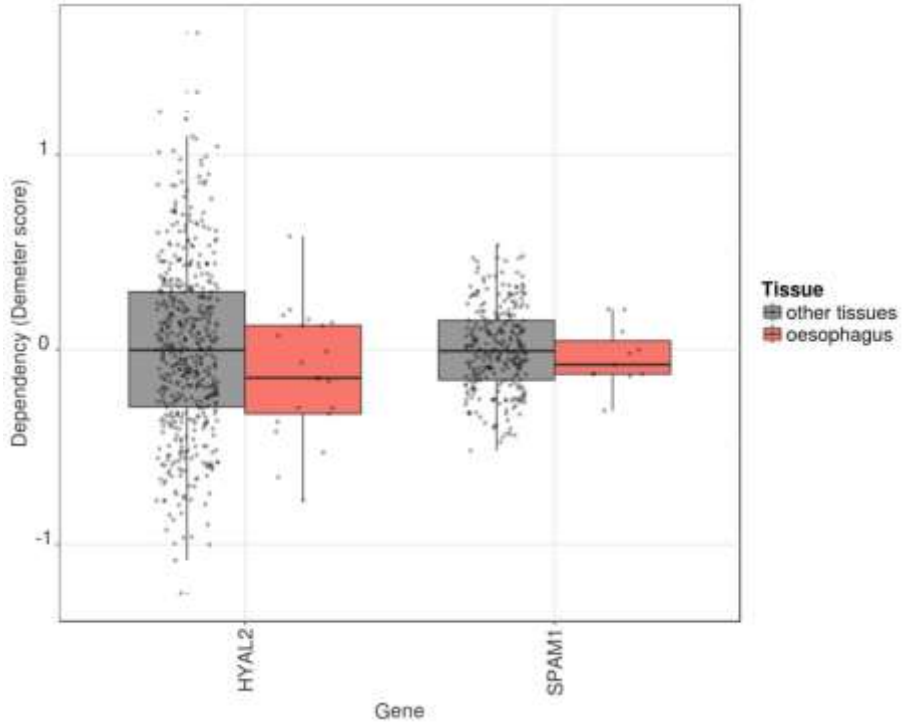
1) Prediction of other gene targets, whose inhibition (modeled KO) deactivate these circuits
2) Validation of the real KO effects with Achilles II (Tsherniak A, et al. 2017, Defining a cancer dependency map. Cell 170: 564-576)

Deactivation of tumor suppressor modules



Prostate cell line

Deactivation in onco-modules



Skin cell line

Interventions on pathways made easy

Nucleic Acids Research, 2016 **1**
doi: 10.1093/nar/gkw369

Actionable pathways: interactive discovery of therapeutic targets using signaling pathway models

Francisco Salavert[1,2], Marta R. Hidago[1], Alicia Amadoz[1], Cankut Çubuk[1], Ignacio Medina[3], Daniel Crespo[1], Jose Carbonell-Caballero[1] and Joaquín Dopazo[1,2,4,*]

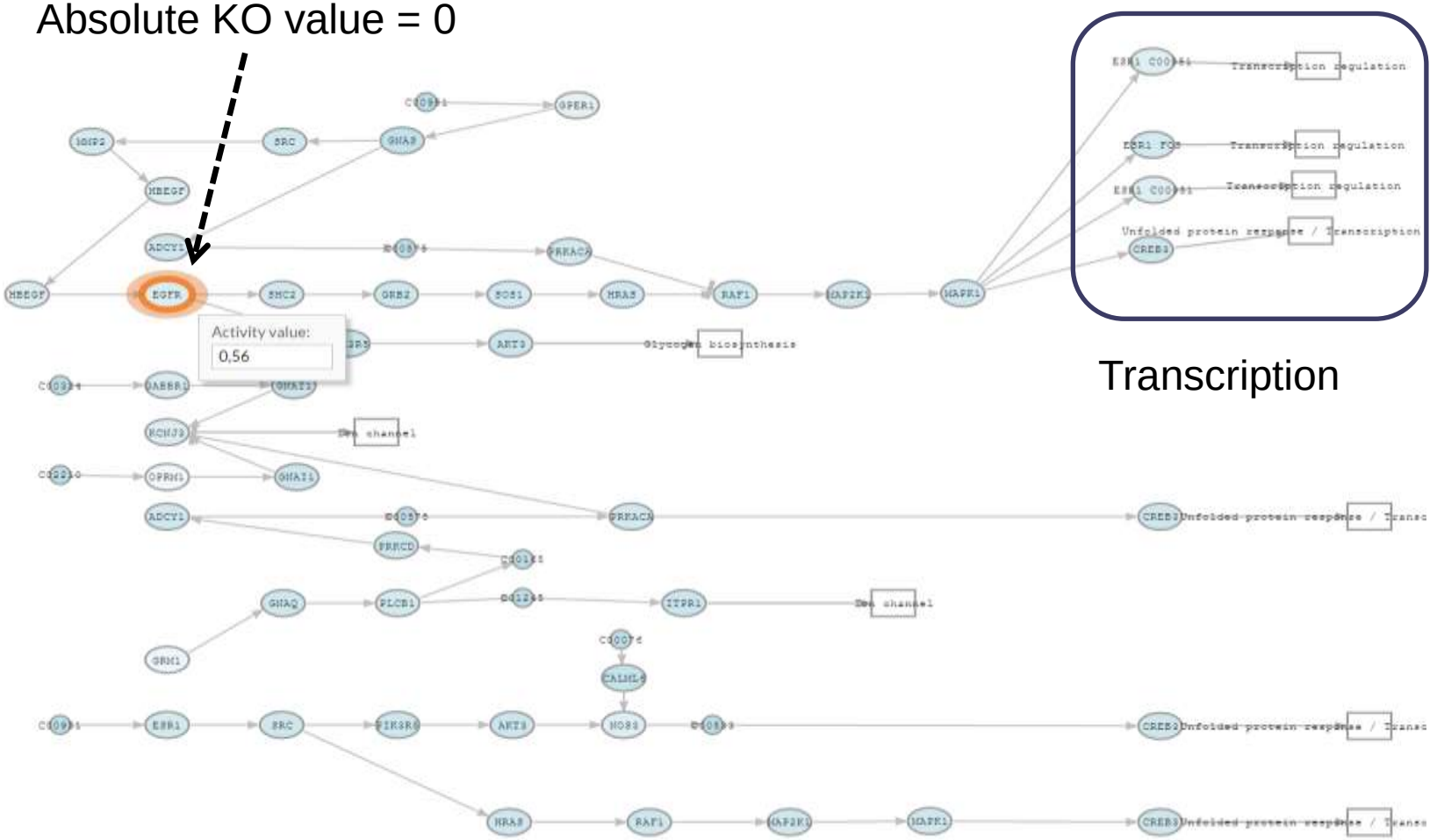
[1]Computational Genomics Department, Centro de Investigación Príncipe Felipe (CIPF), Valencia, 46012, Spain, [2]Bioinformatics in Rare Diseases (BiER), Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Valencia, 46012, Spain, [3]HPC Service, University Information Services, University of Cambridge, Cambridge, CB3 0RB, UK and [4]Functional Genomics Node, (INB, PRB2, ISCIII) at CIPF, Valencia 46012, Spain

Received February 8, 2016; Revised April 13, 2016; Accepted April 22, 2016

Freely available software PathAct
http://pathact.babelomics.org/

Actionable pathway models

We can inhibit EGFR (target of Afanatib) by reducing its activity value (0.56 in cancer).
Absolute KO value = 0

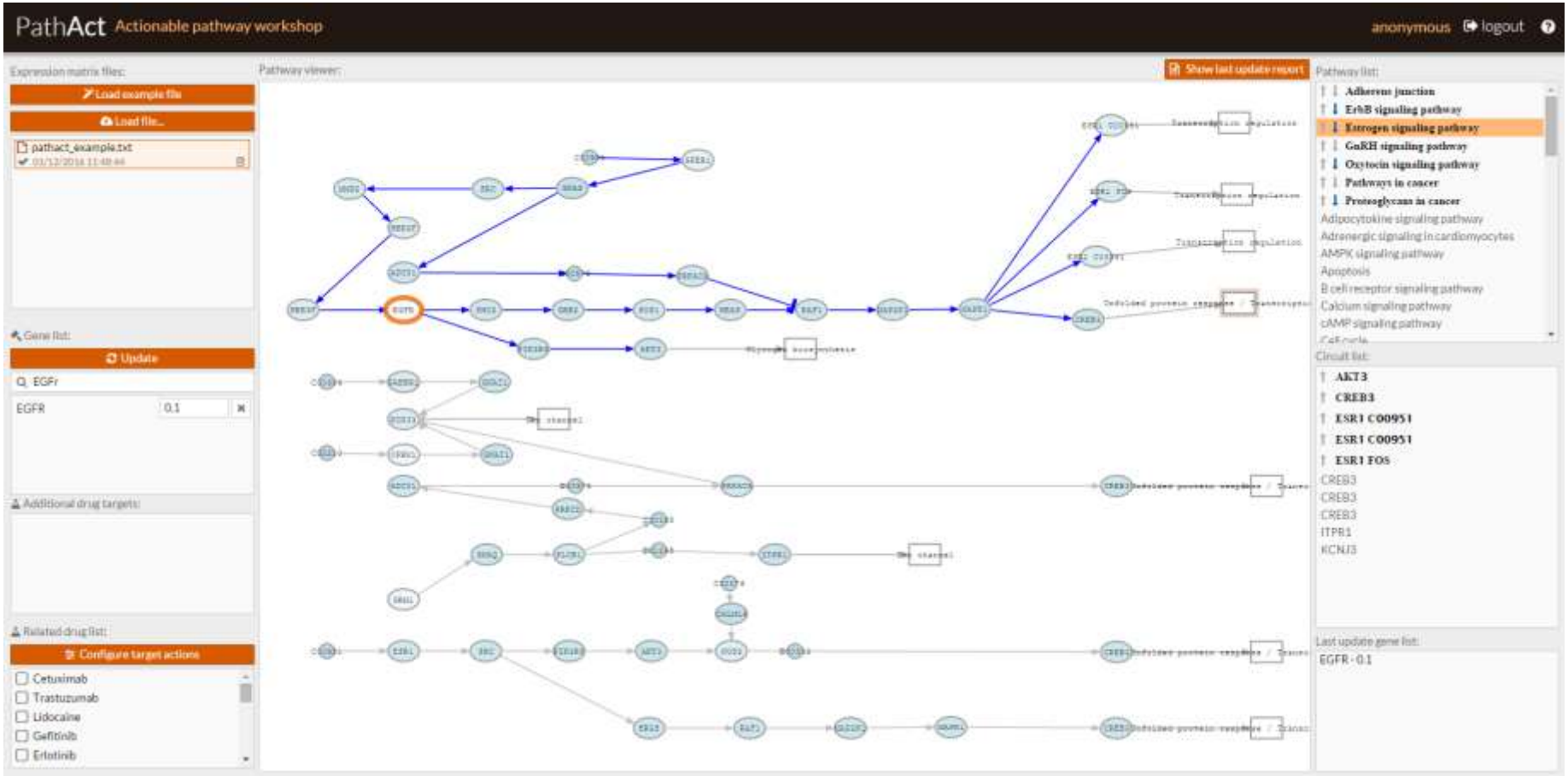


Transcription

Estrogen signaling pathway

http://pathact.babelomics.org/

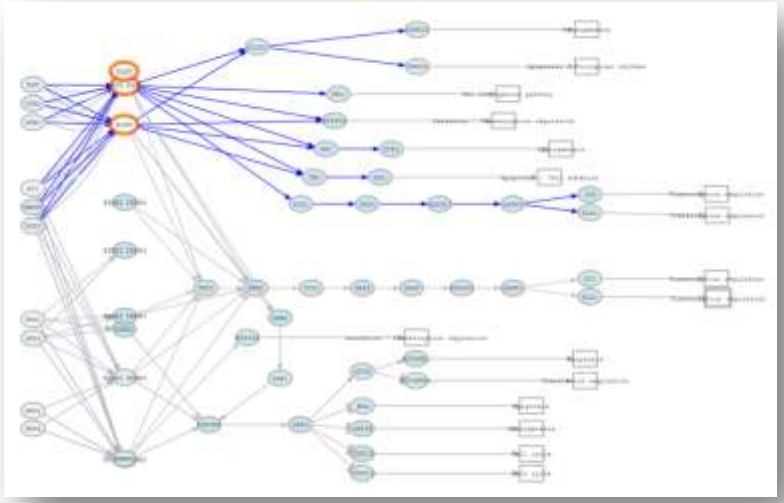
Actionable pathway models



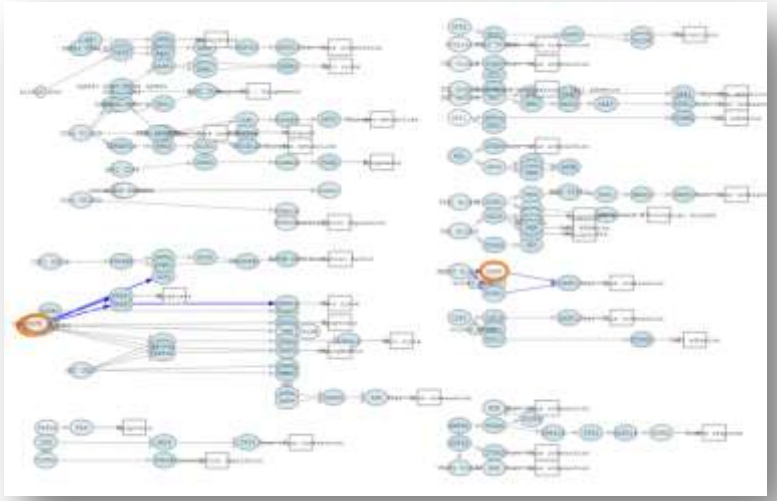
The inhibition of the transcription sought has been attained, but six more pathways have been affected in different ways

http://pathact.babelomics.org/

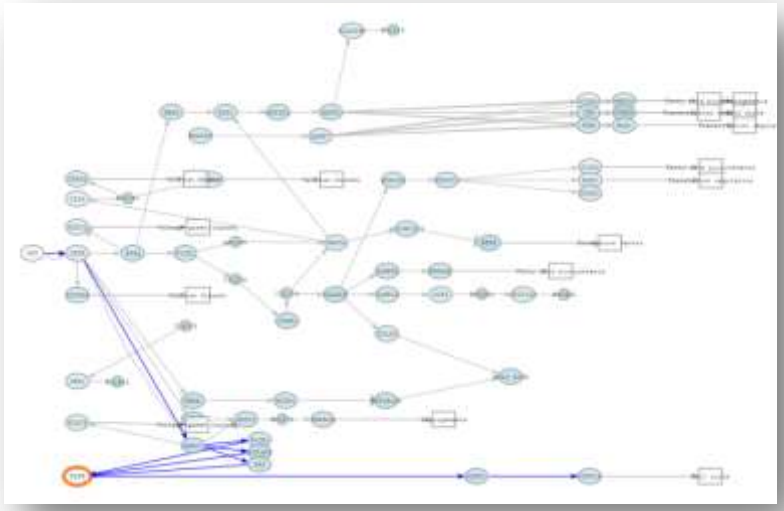
Actionable pathway models



Transcription, angiogenesis and other are inhibited in *ErbB signaling pathway*



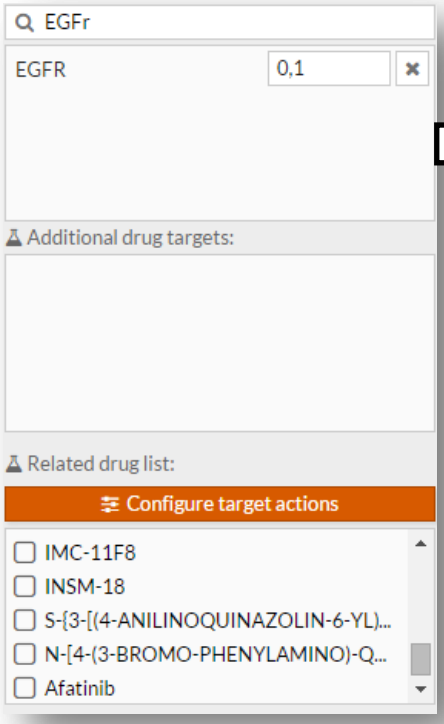
Cell cycle inhibited in *Proteoglycans in cancer pathway*

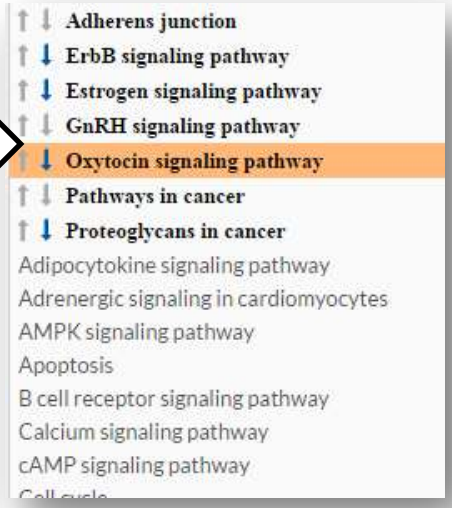


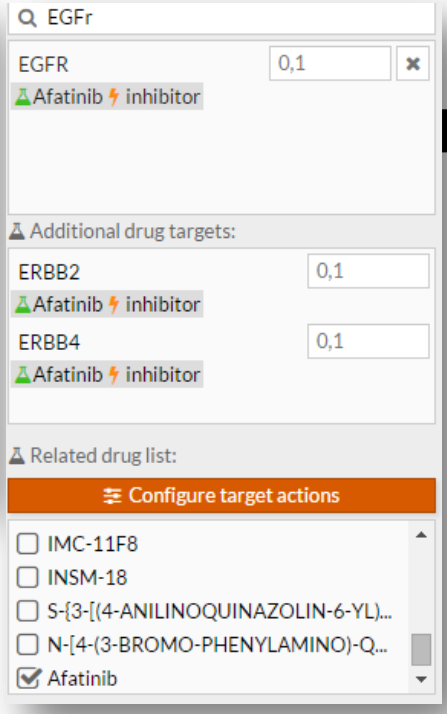
Cell cycle is inhibited in *Oxytocin signaling pathway*

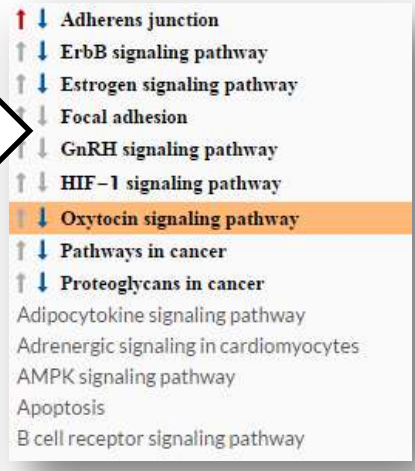
http://pathact.babelomics.org/

Simulating drug inhibition

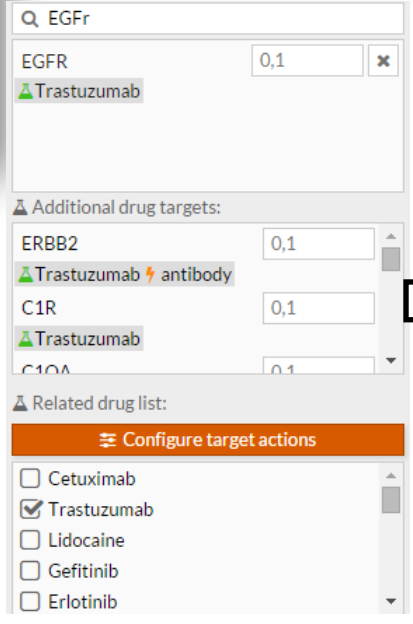


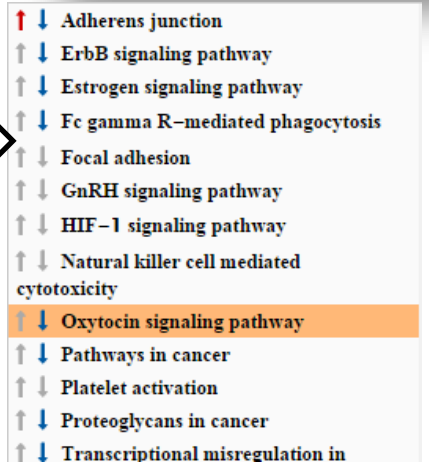






Real inhibition with Afanatib affects 11 pathways



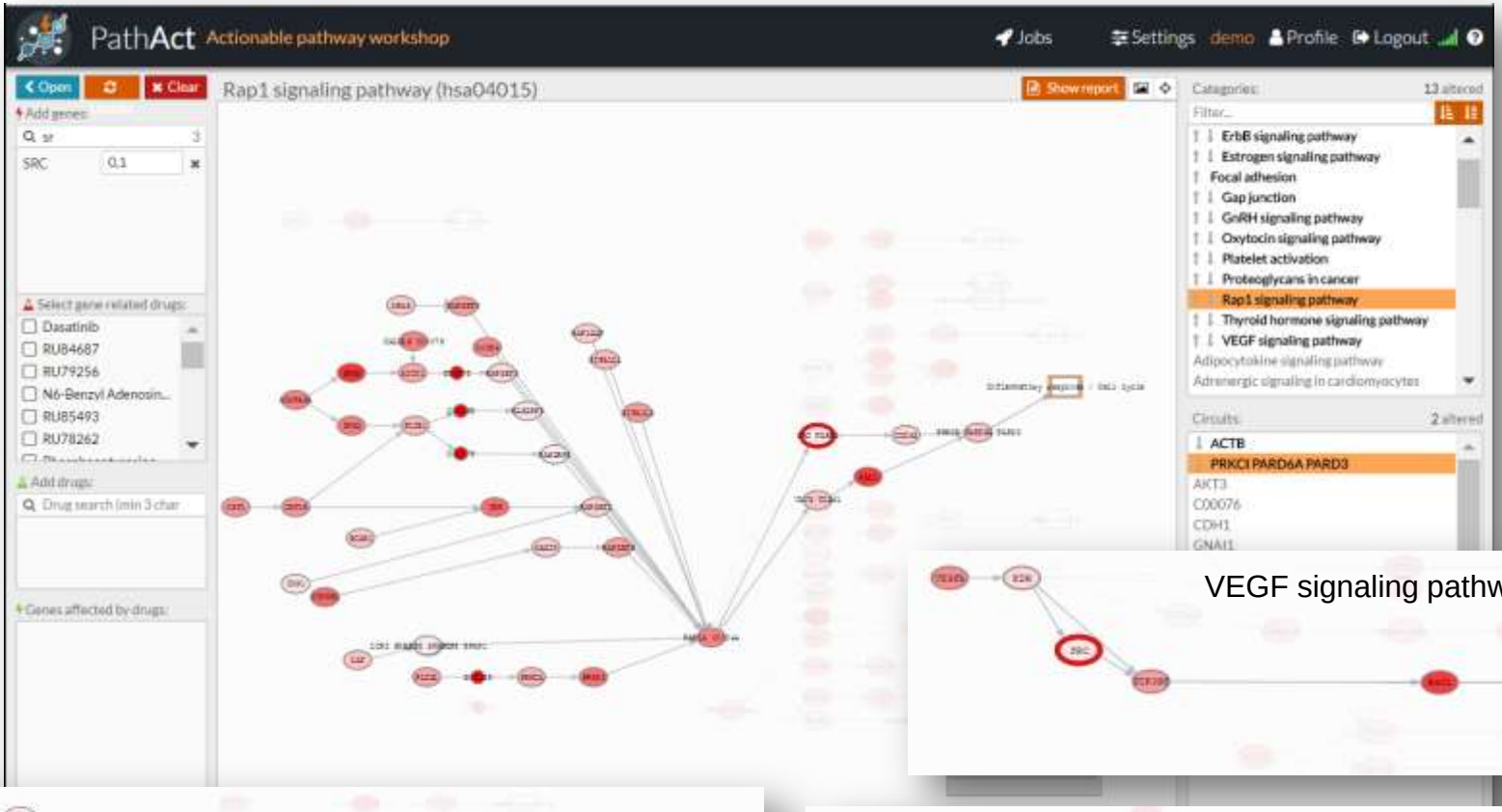


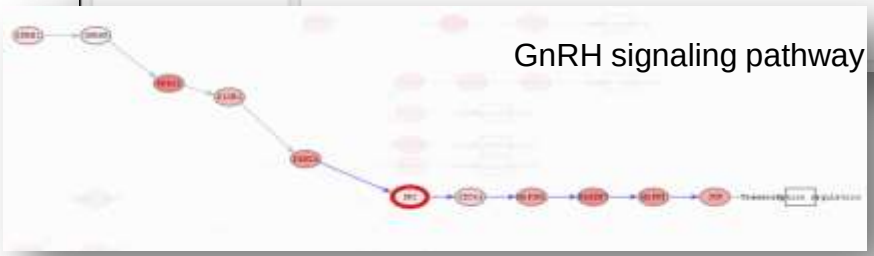
“Ideal” KO of EGFR affects 7 pathways

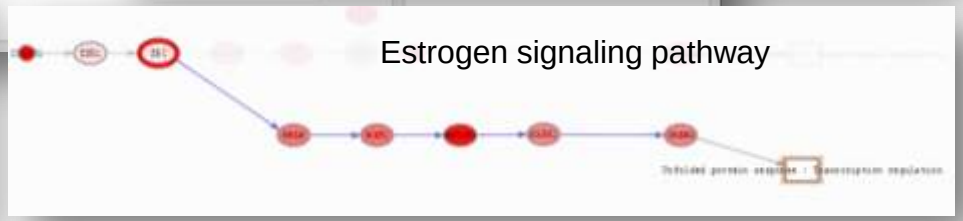
Inhibition with broader spectrum Trastuzumab affects 13 pathways

An example with *SRC* gene.

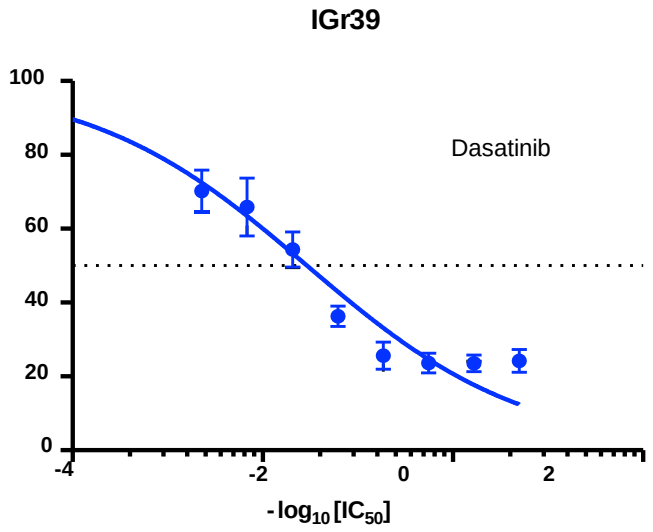
Predicted to be essential in melanoma cell lines







Prediction of essential genes that were never experimentally tested before



Dasatinib, a specific inhibitor of *SRC*, demonstrates the essentiality of *SRC* predicted because the inhibition of the gene predicts the inhibition of an onco-circuit

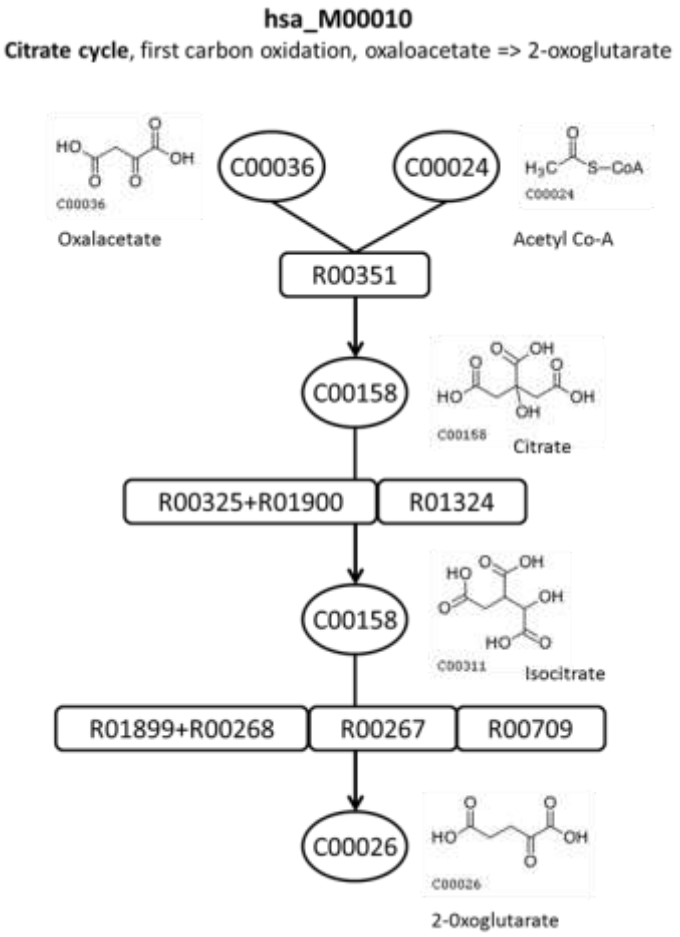
Metabolic pathways can also be modeled

There are 94 modules that recapitulate the main aspects of metabolism of carbohydrates, lipids, amino acids and nucleotides

Metabolic activity:

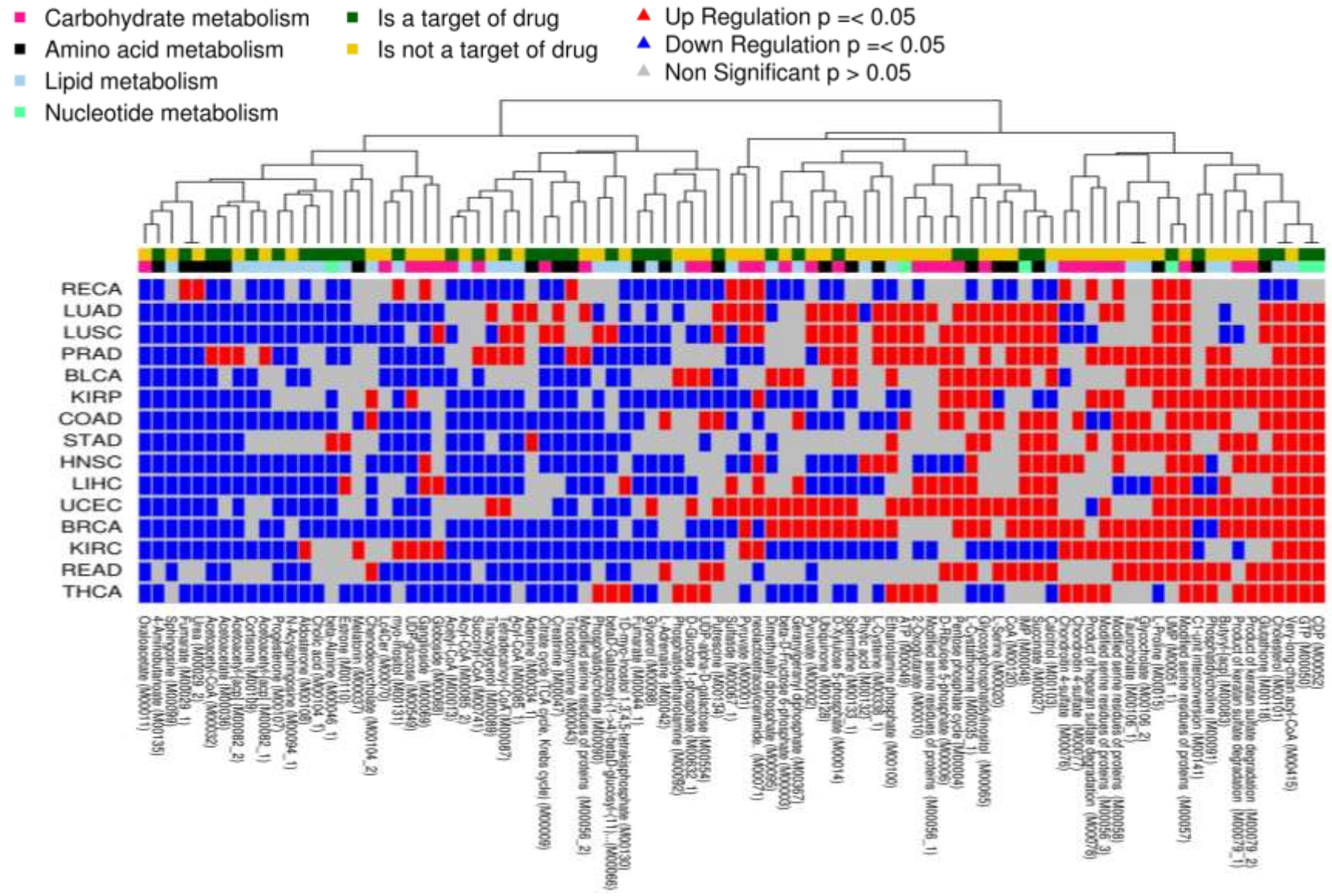
$$S_n = n_i \cdot \left(1 - \prod_{s_a \in A} (1 - n_a)\right)$$

Where n_i is the activity of the current node n, A is the total number of edges arriving to the node that account for the flux of metabolites produced in other nodes with activity values n_a.



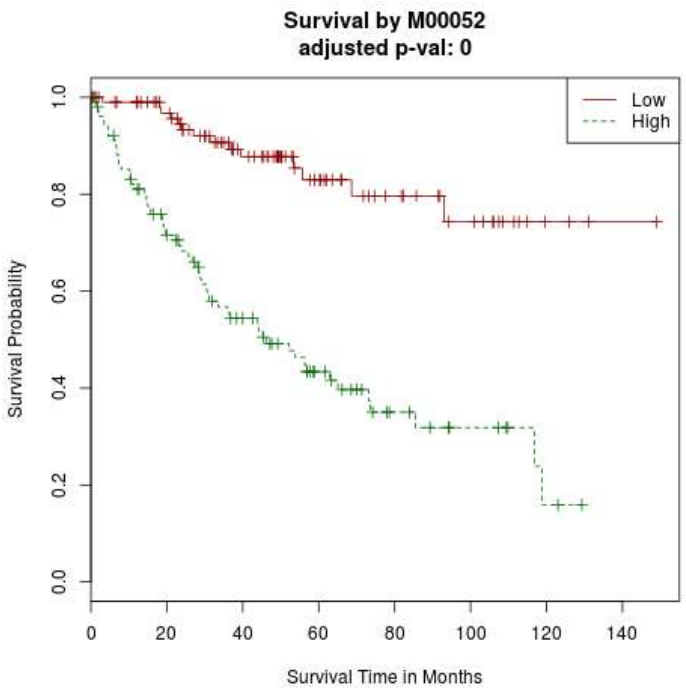
Differential metabolic activity: two conditions are compared by means of a Wilcoxon test (FDR adjusted across modules)

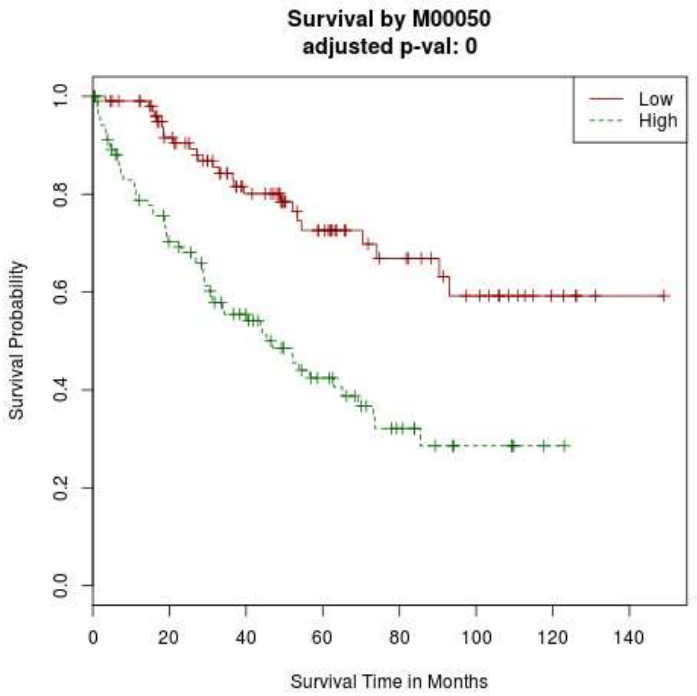
Metabolic modules capture differential metabolic activity



Metabolic modules also capture cell functionality associated to cancer prognostic

High activity of *Guanine ribonucleotide biosynthesis* and *Pyrimidine ribonucleotide biosynthesis* modules is associated to **low survival**.
These modules are target of *Mercaptopurine* and *Gemcitabine*.
The mechanism of action of these drugs involves inhibition of DNA synthesis and that leads to cell death





Prediction of gene essentiality from metabolic pathway essentiality

Pyrimidine degradation pathway was predicted to be an onco-module in gastric cancer cell lines. Predicted genes that switch the pathway off are *DPYD*, *DPYS* (confirmed in Achilles) and *UPB1*

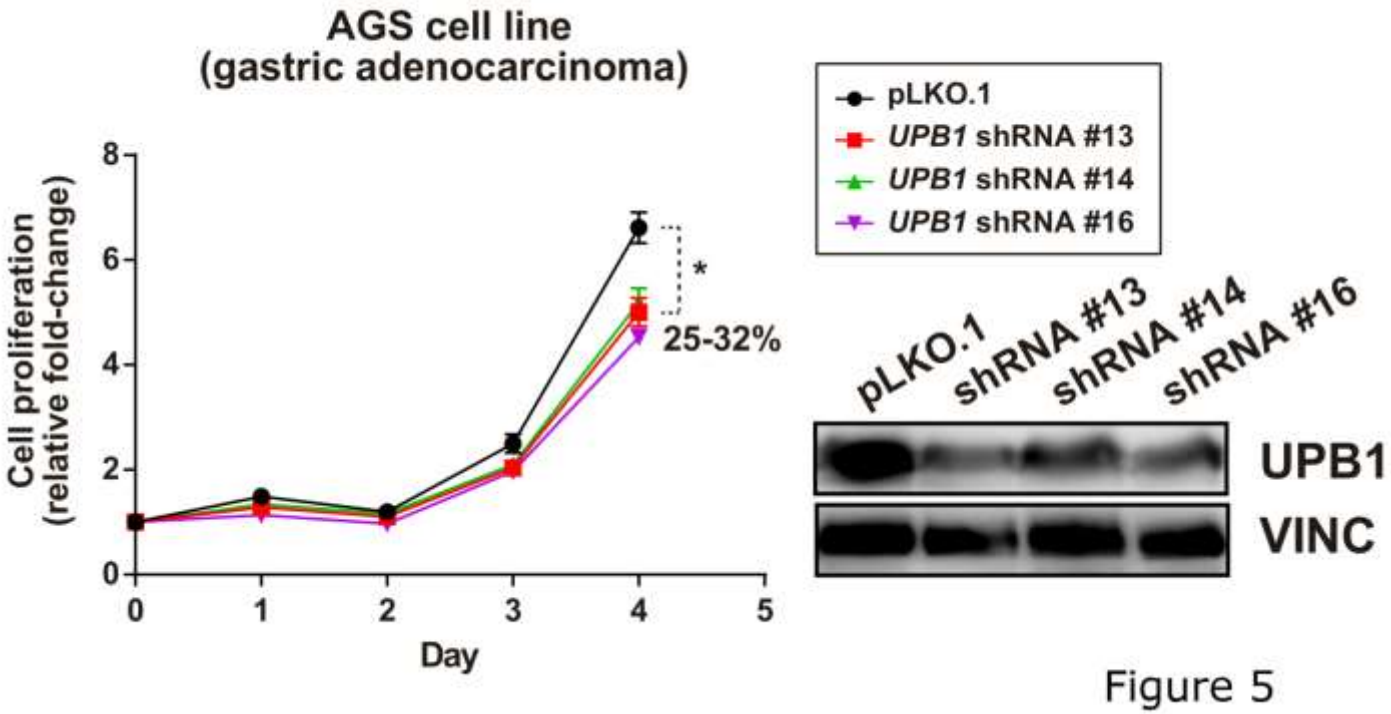
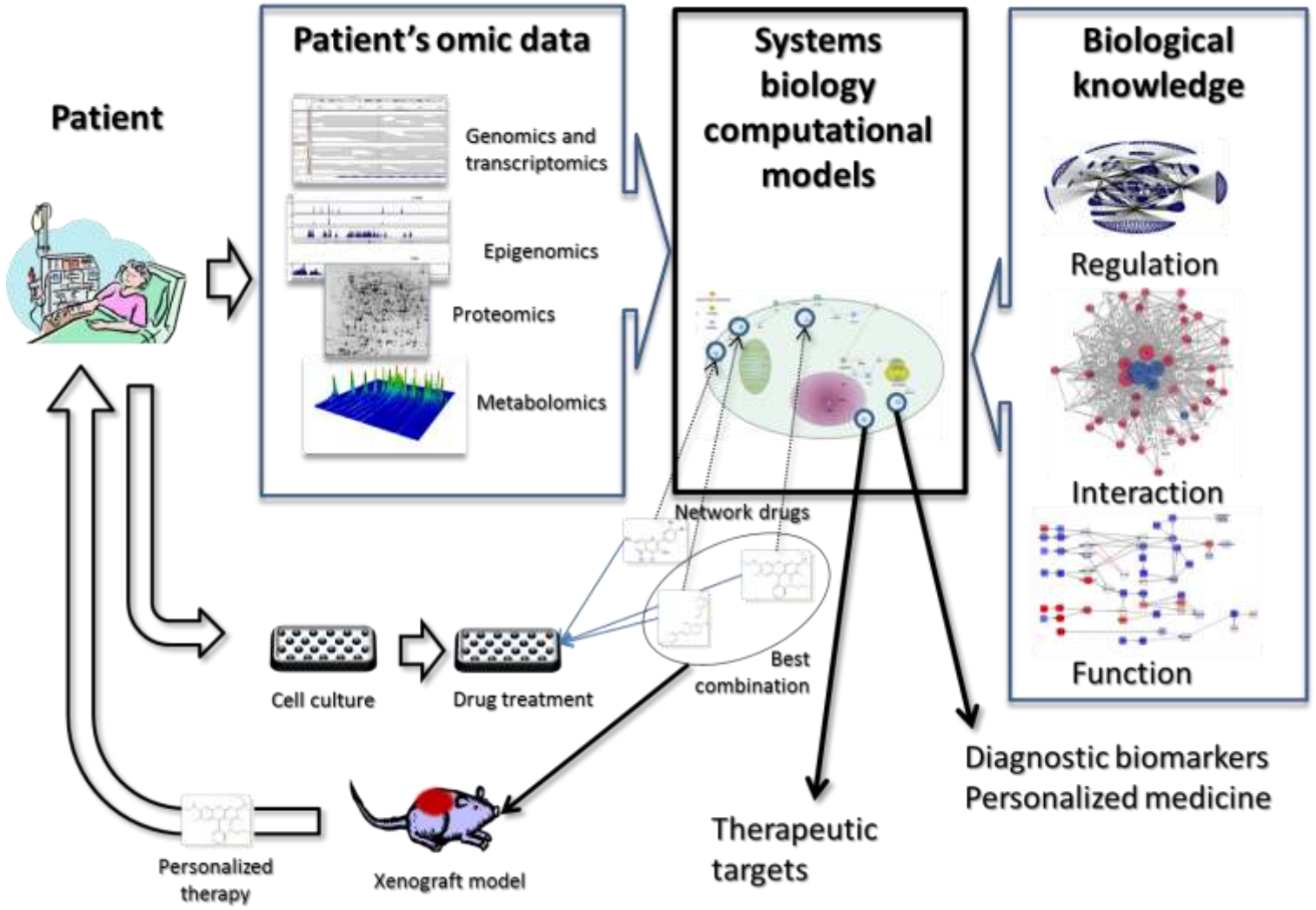


Figure 5

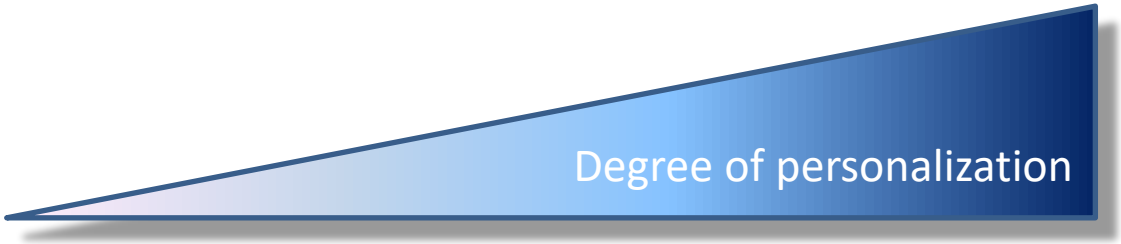
UPB1 encodes an enzyme (β-ureidopropionase) that catalyzes the last step in the pyrimidine degradation pathway, required for epithelial-mesenchymal transition

Models of cell functional activity bring the dream of precision personalized (even individualized) treatments closer



From: Dopazo, 2014, Genomics and transcriptomics in drug discovery. Drug Discovery Today

The real transition to precision medicine



Intuitive
Based on trial
and error

Identification of
probabilistic
patterns

Decisions and
actions based
on knowledge

Intuitive Medicine **Empirical Medicine** **Precision Medicine**

Today

Tomorrow

The use of new algorithms that enable the transformation of genomic measurements into cell functionality measurements that account for disease mechanisms and for drug mechanisms of action will ultimately allow the real transition from today's empirical medicine to precision medicine and provide an increasingly personalized medicine

Clinical Bioinformatics Area
Fundación Progreso y Salud, Sevilla, Spain, and…

…the INB-ELIXIR-ES, National Institute of Bioinformatics and the BiER (CIBERER Network of Centers for Research in Rare Diseases)

Follow us on twitter

@xdopazo

@ClinicalBioinfo

UNIÓN EUROPEA
Fondo Europeo de Desarrollo Regional
"Una manera de hacer Europa"

https://www.slideshare.net/xdopazo/