



Tutorial for High Throughput pathway interpretation and analysis

Kinza Rian

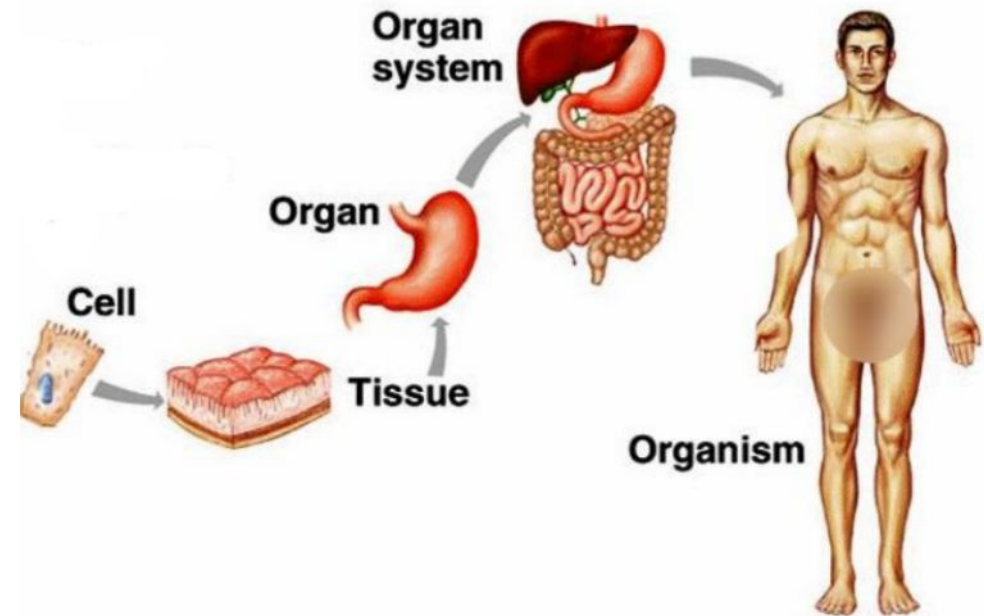
kinza.rian@juntadeandalucia.es

Clinical Bioinformatics Research Area, HVR

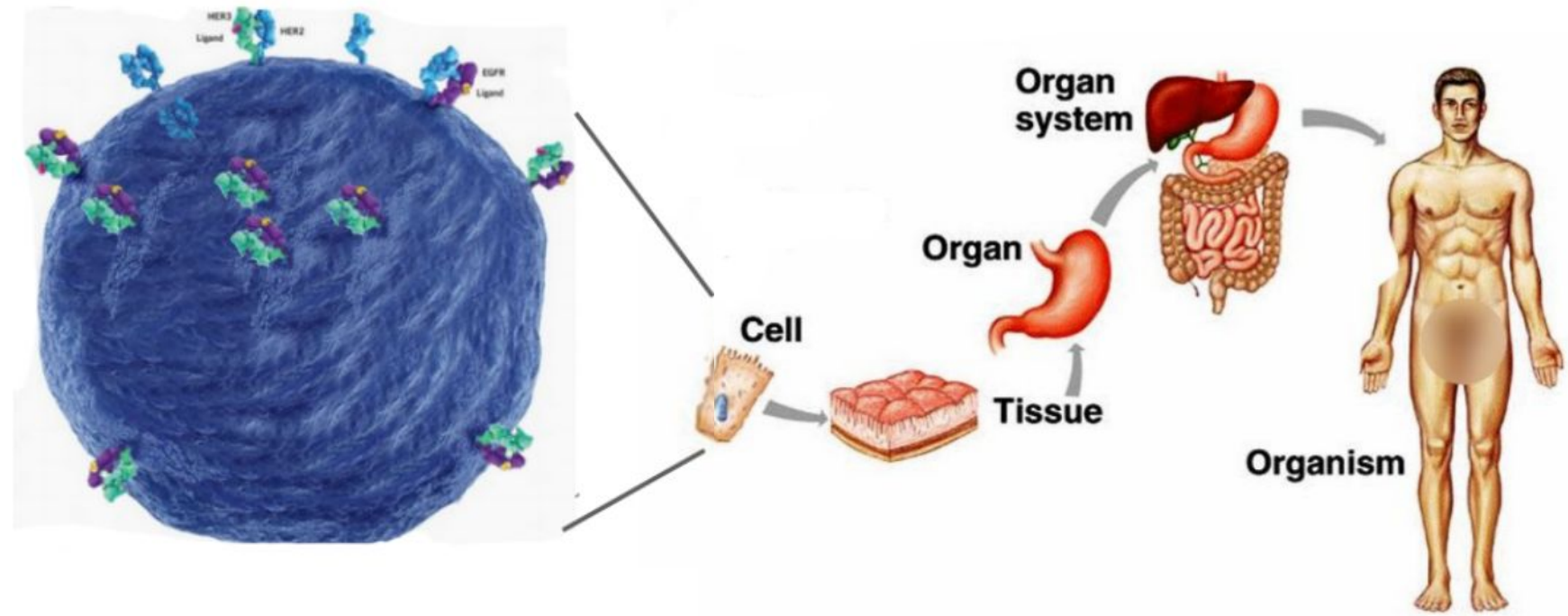
hipathia.babelomics.org



Introduction

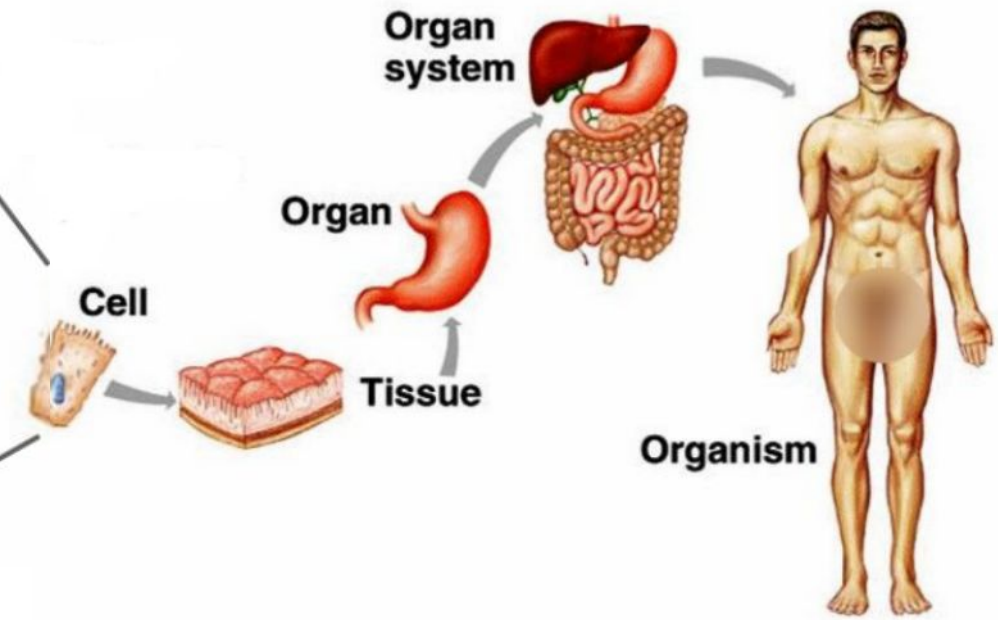
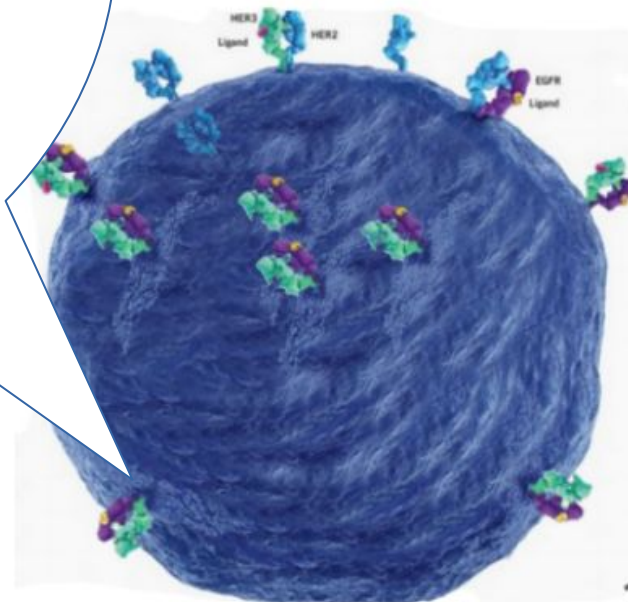
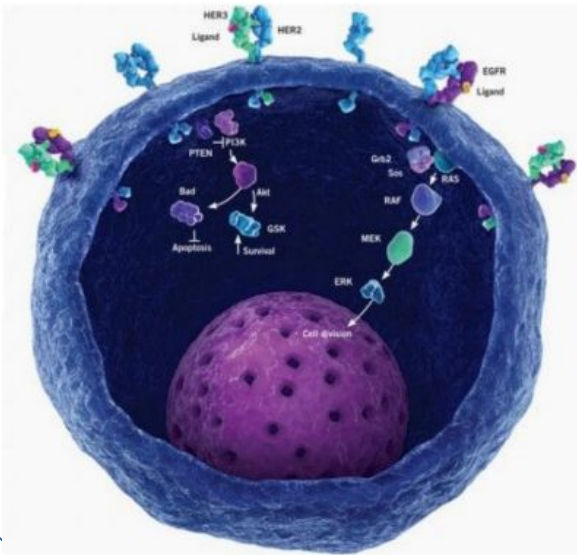


Introduction

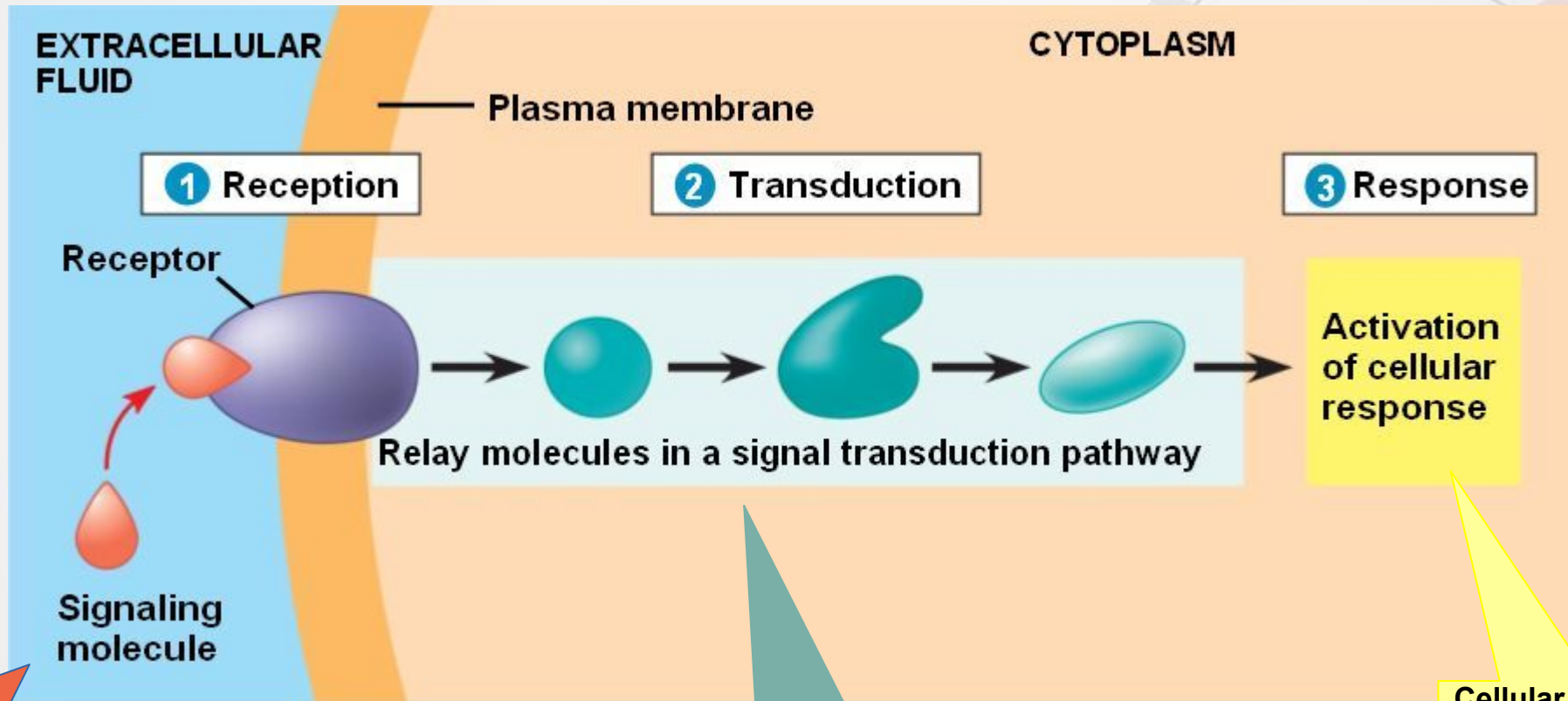


Introduction

> **Cell signaling** is part of any **communication** process that governs basic **activities of cells** and coordinates multiple-cell actions.



Signaling pathway



Chemical signals

- Hormones
- Neurotransmitters
- Growth factors
- Cytokines
- Drugs

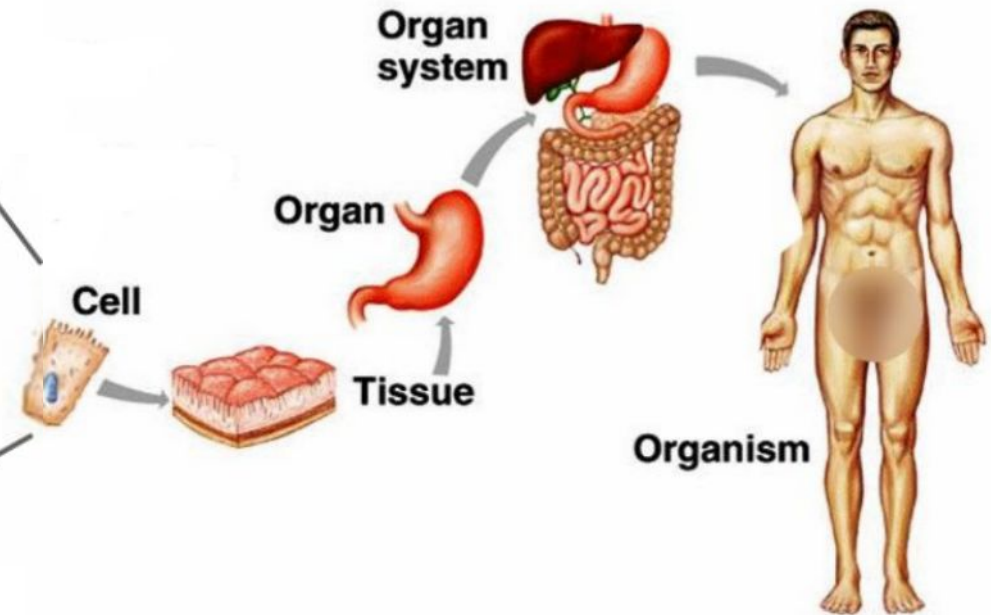
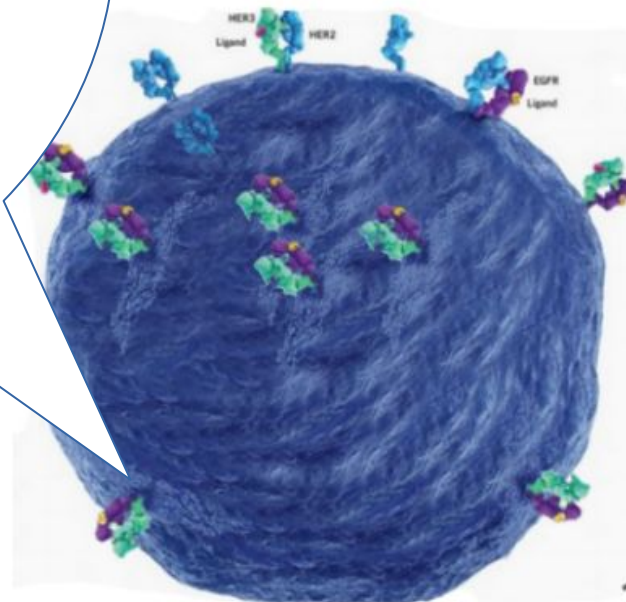
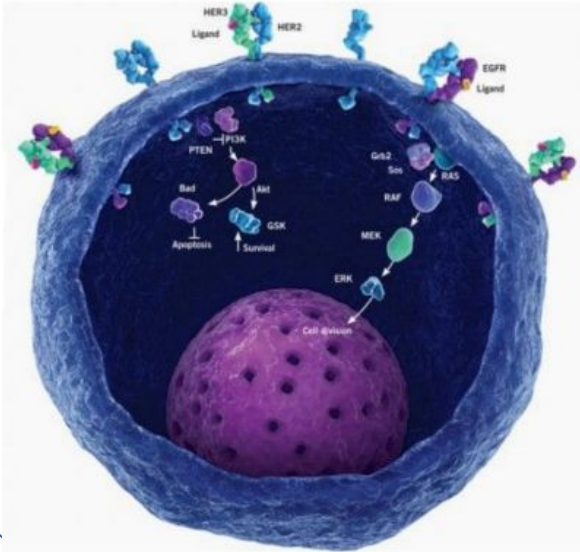
Activation & Inhibition

- Phosphorilation
- Dephosphorilation
- Glycosylation
- Ubiquitination
- Methylation

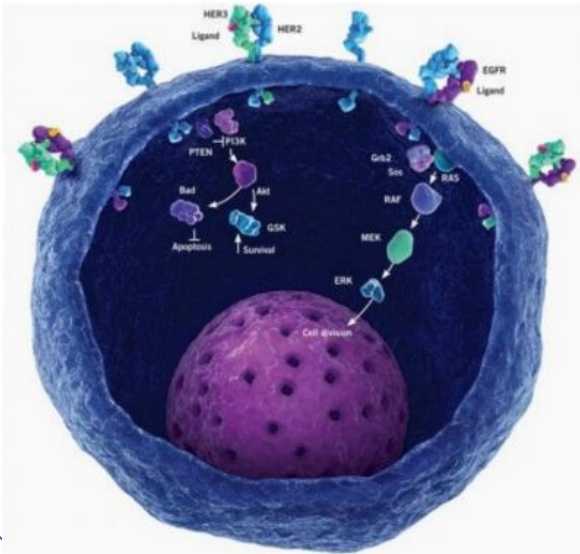
Cellular Function

- Apoptosis
- Survival
- Growth
- Migration
- Proliferation

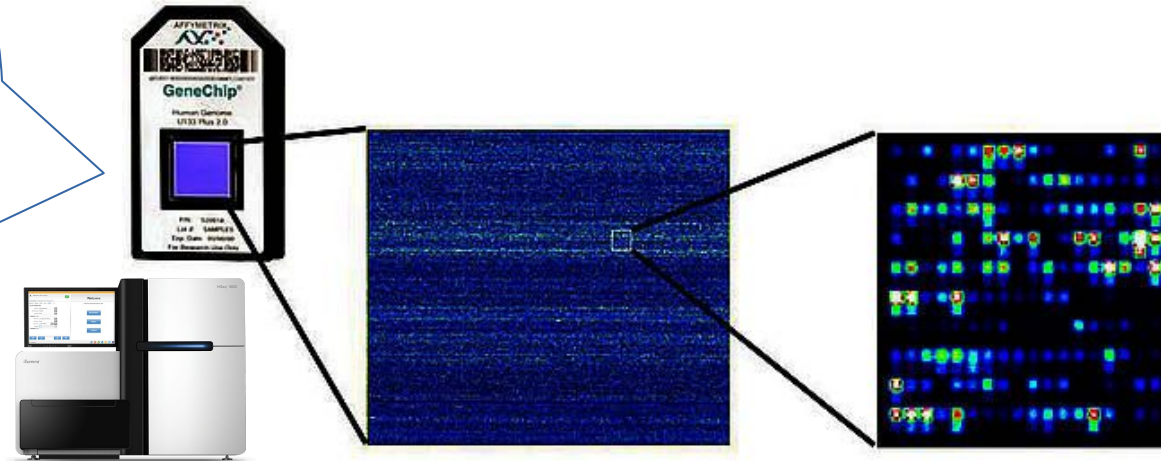
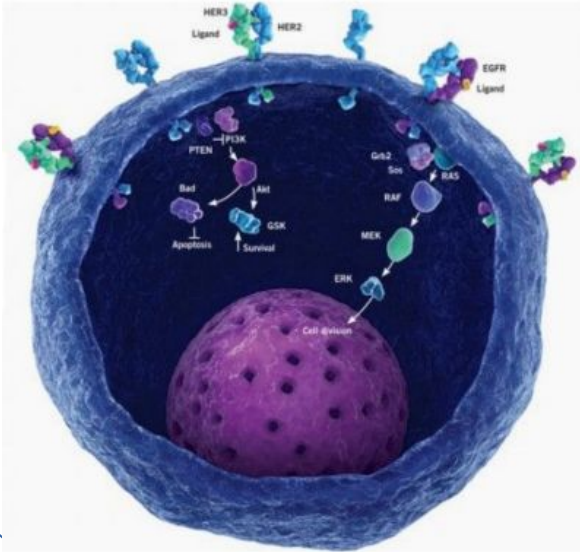
From cell to pathways



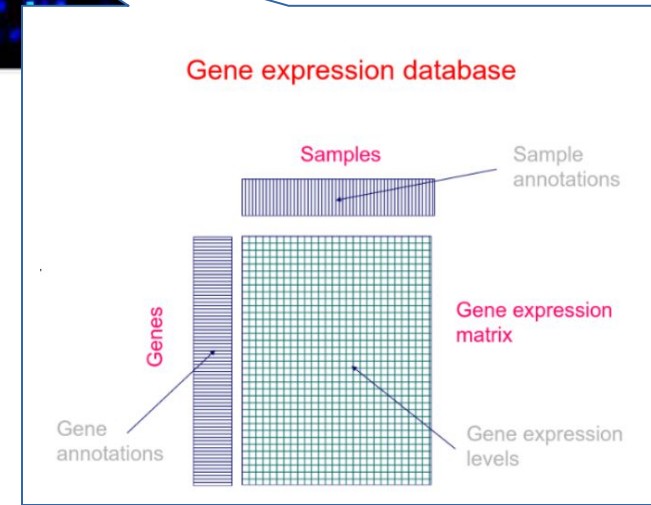
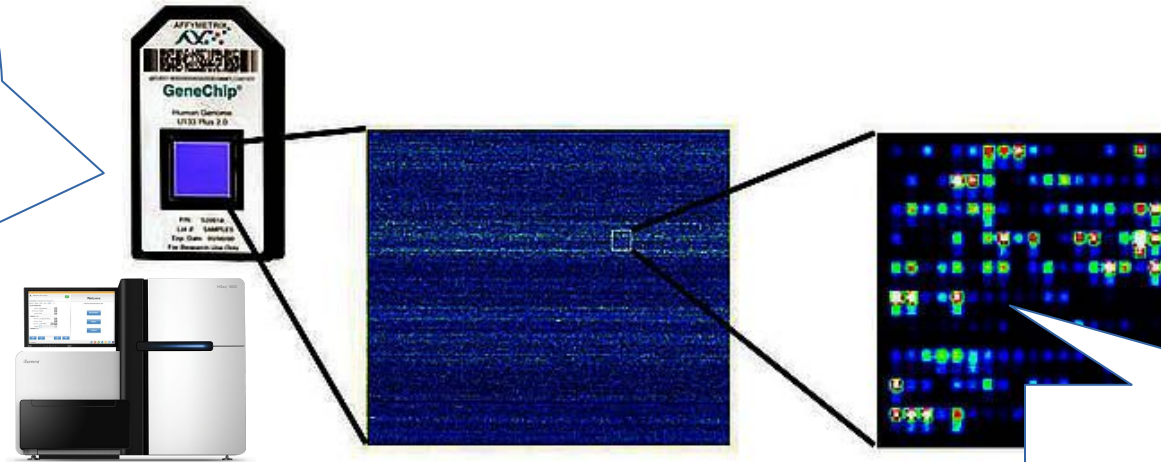
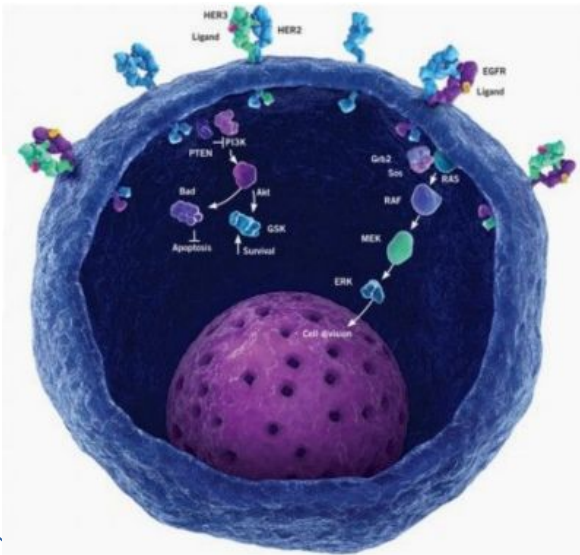
From cell to pathways



From cell to pathways



From cell to pathways



Overview / Our aim

Case Vs Controls

Case Vs Controls

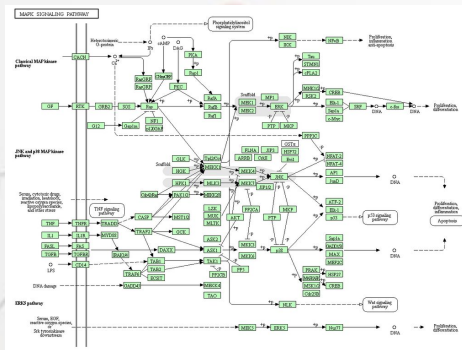
Raw Data

```
id TCGA_EW_A10V_O1A_11R_A144_07 TCGA_DB_A143_O1A_11R_A115_07  
TCGA_AR_A1AR_O1A_11R_A137_07 TCGA_AG_A129_O1A_21R_A161_07  
TCGA_AZ_A05X_O1A_12R_A086_07 TCGA_BH_A0B3_O1A_11R_A056_07  
TCGA_AN_A0AL_O1A_11R_A082_07 TCGA_BH_A1B1_O1A_11R_A129_07  
TCGA_BH_A1B3_O1A_11R_A129_07 TCGA_BH_A1E1_O1A_11R_A129_07  
TCGA_BH_A1B0_O1A_11R_A129_07 TCGA_BH_A0D0_O1A_23R_A12P_07  
TCGA_BH_A0D0_O1A_23R_A12P_07 TCGA_BH_A0D0_O1A_23R_A12P_07  
TCGA_E9_A1M6_O1A_11R_A144_07 TCGA_BH_A1E1_O1A_11R_A129_07  
TCGA_A7_A0B0_O1A_11R_A144_07 TCGA_EZ_A17_11R_A115_07  
1 0.43934783403125 0.37432818240243 0.391826953647877  
0.3925994052209 0.3889584839982 0.488062090524718  
0.381683579147373 0.425793715847912 0.372870427153651  
0.3925994052209 0.3889584839982 0.488062090524718  
0.403234658246308 0.36888730507246 0.39615832243318  
0.37579101223776 0.388882791368954 0.406672016488648  
0.34622206670989 0.4022156262319  
29974 0.385189655061793 0.2937246611168 0.33824654853348  
0.3392898969891 0.32788611975740 0.38609391256296  
0.33369898260781 0.38403211317658 0.27364411620602  
0.338637924789498 0.2933366390838 0.3860630783841  
0.327621018959854 0.2922781722352 0.29388848136702  
0.2057329308112 0.29241232531206 0.2420353797958  
0.26463861524822 0.29381831772822 0.428877879378001  
0.412381887466802 0.40273637154134 0.428877879378001  
0.43922206670989 0.412381887466802 0.38754362432187  
0.43486650769526 0.43186863933462 0.43486650769526  
0.38518972273642 0.427848612848667 0.414780838678332  
0.425237882552206 0.420885349769897 0.42095963054558  
0.426945486236763 0.416789921899738 0.426352489593857  
0.40972462432197 0.4215133318658  
2 0.58530618262448 0.545248878292849 0.545139407466753  
0.6648521524774 0.5685533864884 0.5648521524774  
0.551348543747071 0.568548873892575 0.568537611710764  
0.53181078887405 0.5811972222266 0.58678887378989  
0.581472466883485 0.683146353616241 0.597937781549334  
0.69138197208139 0.585442429697958 0.57898984610956  
CGA_AZ_A05X_O1A_12R_A086_07 0.693254979794158  
14568 0.481483537789427 0.434267988539266 0.479153487468888  
0.48986831676889 0.473713834513659 0.545139407466753
```

Normalized genes

```
id TCGA_EW_A10V_O1A_11R_A144_07 TCGA_DB_A143_O1A_11R_A115_07  
TCGA_AR_A1AR_O1A_11R_A137_07 TCGA_AG_A129_O1A_21R_A161_07  
TCGA_AZ_A05X_O1A_12R_A086_07 TCGA_BH_A0B3_O1A_11R_A056_07  
TCGA_AN_A0AL_O1A_11R_A082_07 TCGA_BH_A1B1_O1A_11R_A129_07  
TCGA_BH_A1B3_O1A_11R_A129_07 TCGA_BH_A1E1_O1A_11R_A129_07  
TCGA_BH_A1B0_O1A_11R_A129_07 TCGA_BH_A0D0_O1A_23R_A12P_07  
TCGA_BH_A0D0_O1A_23R_A12P_07 TCGA_BH_A0D0_O1A_23R_A12P_07  
TCGA_E9_A1M6_O1A_11R_A144_07 TCGA_BH_A1E1_O1A_11R_A129_07  
TCGA_A7_A0B0_O1A_11R_A144_07 TCGA_EZ_A17_11R_A115_07  
1 0.43934783403125 0.37432818240243 0.391826953647877  
0.3925994052209 0.3889584839982 0.488062090524718  
0.381683579147373 0.425793715847912 0.372870427153651  
0.3925994052209 0.3889584839982 0.488062090524718  
0.403234658246308 0.36888730507246 0.39615832243318  
0.37579101223776 0.388882791368954 0.406672016488648  
0.34622206670989 0.4022156262319  
29974 0.385189655061793 0.2937246611168 0.33824654853348  
0.3392898969891 0.32788611975740 0.38609391256296  
0.33369898260781 0.38403211317658 0.27364411620602  
0.338637924789498 0.2933366390838 0.3860630783841  
0.327621018959854 0.2922781722352 0.29388848136702  
0.2057329308112 0.29241232531206 0.2420353797958  
0.26463861524822 0.29381831772822 0.428877879378001  
0.412381887466802 0.40273637154134 0.428877879378001  
0.43922206670989 0.412381887466802 0.38754362432187  
0.43486650769526 0.43186863933462 0.43486650769526  
0.38518972273642 0.427848612848667 0.414780838678332  
0.425237882552206 0.420885349769897 0.42095963054558  
0.426945486236763 0.416789921899738 0.426352489593857  
0.40972462432197 0.4215133318658  
2 0.58530618262448 0.545248878292849 0.545139407466753  
0.6648521524774 0.5685533864884 0.5648521524774  
0.551348543747071 0.568548873892575 0.568537611710764  
0.53181078887405 0.5811972222266 0.58678887378989  
0.581472466883485 0.683146353616241 0.597937781549334  
0.69138197208139 0.585442429697958 0.57898984610956  
CGA_AZ_A05X_O1A_12R_A086_07 0.693254979794158  
14568 0.481483537789427 0.434267988539266 0.479153487468888  
0.48986831676889 0.473713834513659 0.545139407466753
```

Signaling pathways/circuits



$$S_n = v_n \cdot \left(1 - \prod_{s_a \in A} (1 - s_a) \right) \cdot \prod_{s_i \in I} (1 - s_i)$$

Activity

TCGA_EW_A10V_O1A_11R_A144_07	TCGA_DB_A143_O1A_11R_A115_07	TCGA_BH_A1B0_O1A_11R_A129_07	T
TCGA_EZ_A17_11R_A115_07	TCGA_AR_A1AR_O1A_11R_A137_07	TCGA_AG_A129_O1A_21R_A161_07	T
TCGA_AZ_A05X_O1A_12R_A086_07	TCGA_BH_A0B3_O1A_11R_A056_07	TCGA_AN_A0AL_O1A_11R_A082_07	T
TCGA_BH_A1B1_O1A_11R_A129_07	TCGA_BH_A1E1_O1A_11R_A129_07	TCGA_BH_A1B3_O1A_11R_A129_07	T
TCGA_BH_A1B0_O1A_11R_A129_07	TCGA_BH_A0D0_O1A_23R_A12P_07	TCGA_BH_A0D0_O1A_23R_A12P_07	T
TCGA_E9_A1M6_O1A_11R_A144_07	TCGA_EZ_A17_11R_A115_07	TCGA_A7_A0B0_O1A_11R_A144_07	T
0.039445913483249	0.0371829		
324628284	0.0371829	0.0371829	0.0371829
264410704	0.043216205232446	0.043216205232446	0.043216205232446
558594547	0.0420242859979419	0.054652897954109	0.0397561728292142
653014679	0.03687743181473	0.0455128562012871	0.0549544488895957
687285109			
hsa04110_68_10_11_12_13_14_9	0.56986103974761	0.6874377033341	0.578211614480503
hsa04110_46	0.030802565998871	0.031378892891783	0.55505126545452
hsa018186_439474	0.5888216025482	0.599239632166899	0.58683519731551
981366227	0.58747812238992	0.5386824215655	0.538766488524899
97731933	0.042793287958788	0.043216205232446	0.5212863
4154445			
hsa04110_96_07	0.5802443181862113	0.58071385148101	0.541334023394196
4184697	0.588792081593121	0.518111378220275	0.52485883991437
0.588844889589852	0.51589596492859	0.515561726742555	0.51178381148684
0.518859997924771	0.529145058844453	0.520540565932587	0.525675418153119
0.523837356928666	0.506449981327262	0.5291681848496468	0.51585471327926
hsa04110_48	0.483807084146882	0.51187536681908	0.43466807172819
19824862	0.58996297415117	0.567421882174162	0.599884739629708
03815836	0.46411499014705	0.6863534308735	0.581544775074276
43389246	0.512579473428481	0.46814791319854	0.51848451738826
31282827	0.515661888426823	0.468633859312094	0.538247269935679
68839815			
hsa04110_43_83_84_85	0.37811798613337	0.41954111899138	0.392681175717458
0.48546781838485	0.48788882838997	0.48238828664797	0.388953712147125
0.415676225109251	0.418071646369154	0.383384108991719	0.422887159054524
0.43888681676889	0.473713834513654	0.416681147888807	0.468528161865537

Case Vs Controls

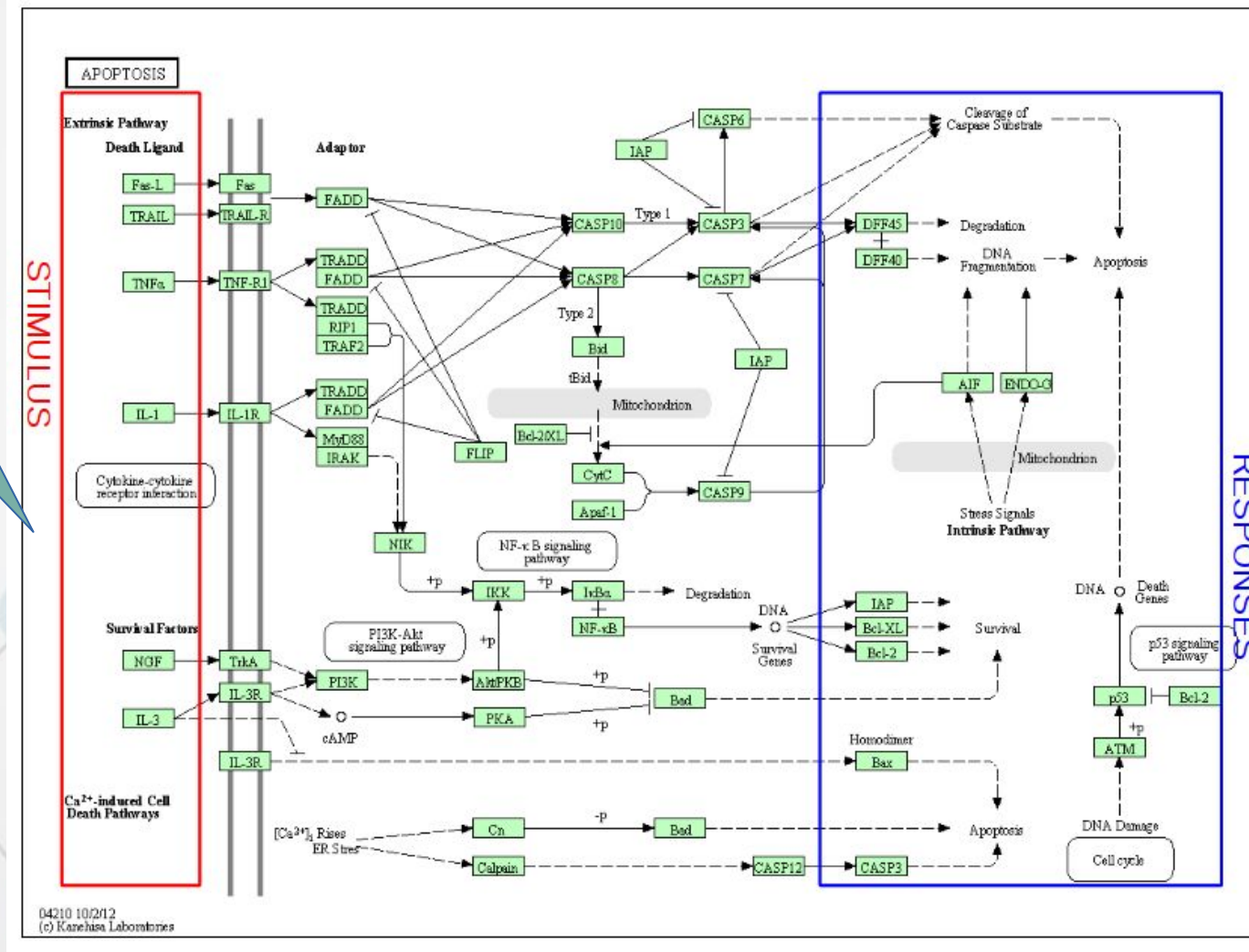
We seek for a simple transformation from individual gene expression profiles to profiles of circuit activity (and functional activity).



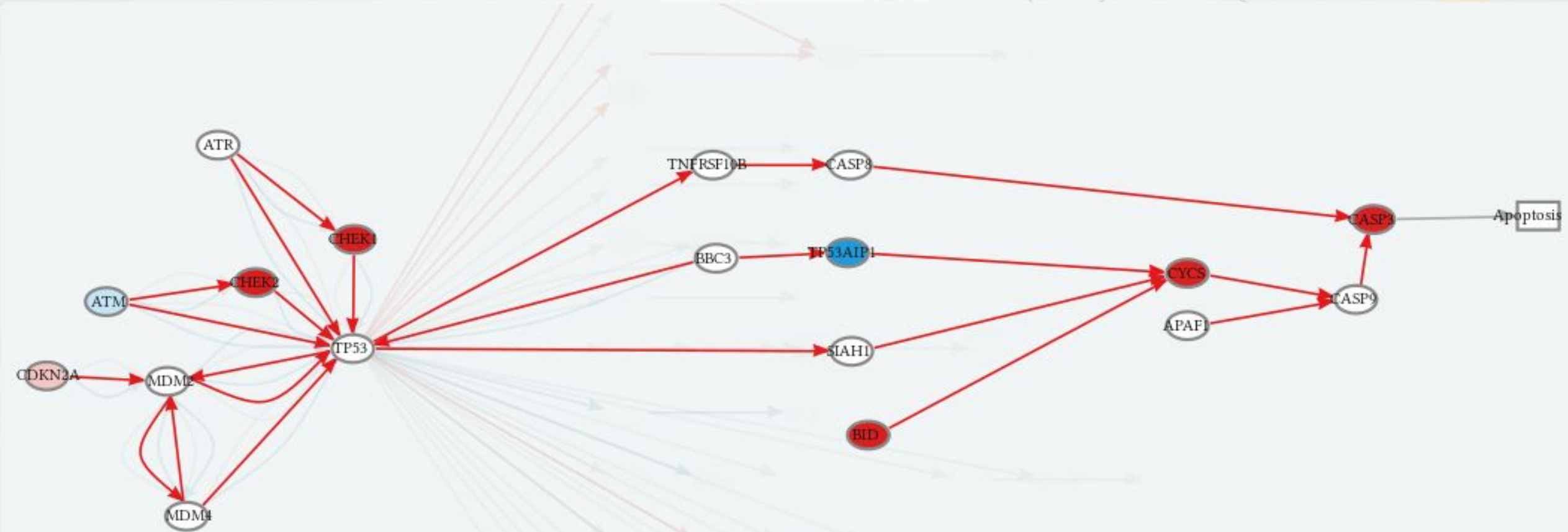
**“Torture the data, and it will confess to anything.”
– Ronald Coase**

Pathways layout

Take pathways information



Meaningful subpathways

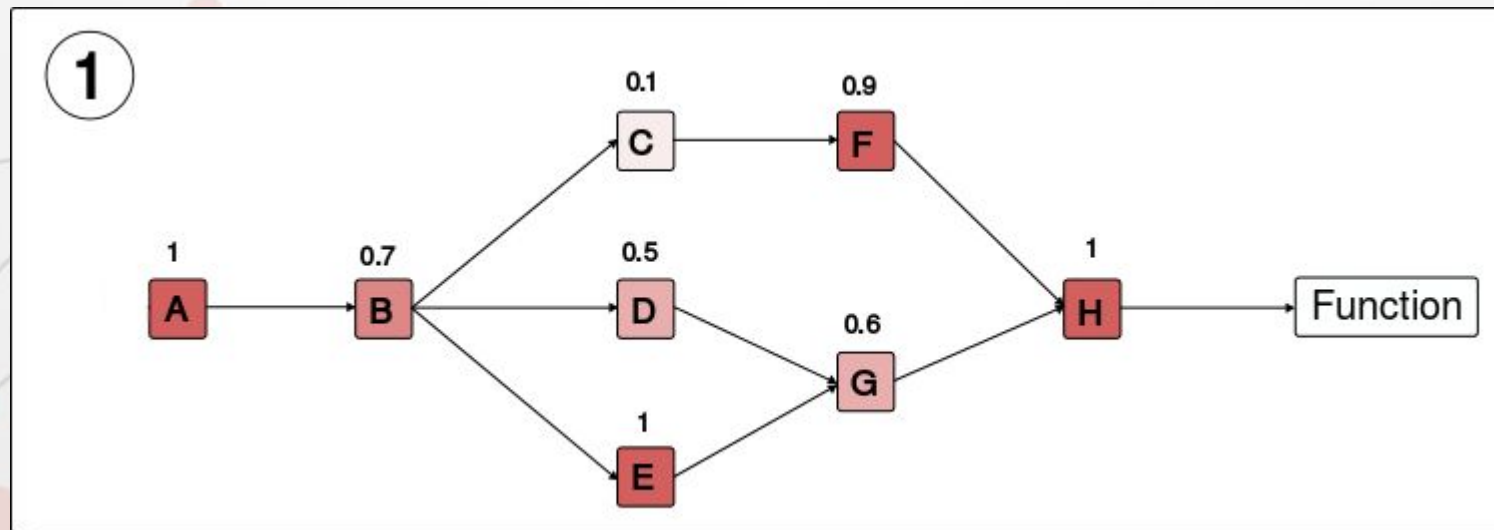


Effector subpathway

> Subpathway including any node from any **receptor** to **one effector protein**

Computing the signal

1) Compute a **node score** based on the expression.



Computing the signal

- 1) Compute a **node score** based on the expression.
- 2) Compute **signal** passing through each **node n**.

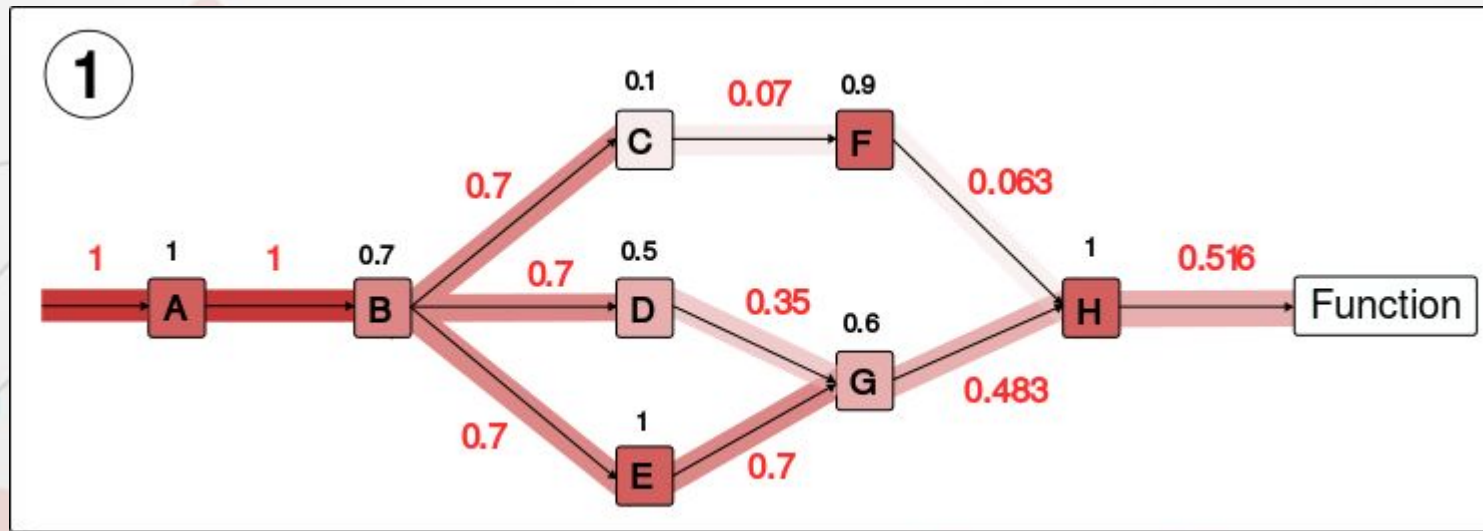
$$S_n = v_n \cdot \left(1 - \prod_{s_i \in A} (1 - s_i)\right) \cdot \prod_{s_j \in I} (1 - s_j)$$

S_n : Signal value through n

v_n : Node value

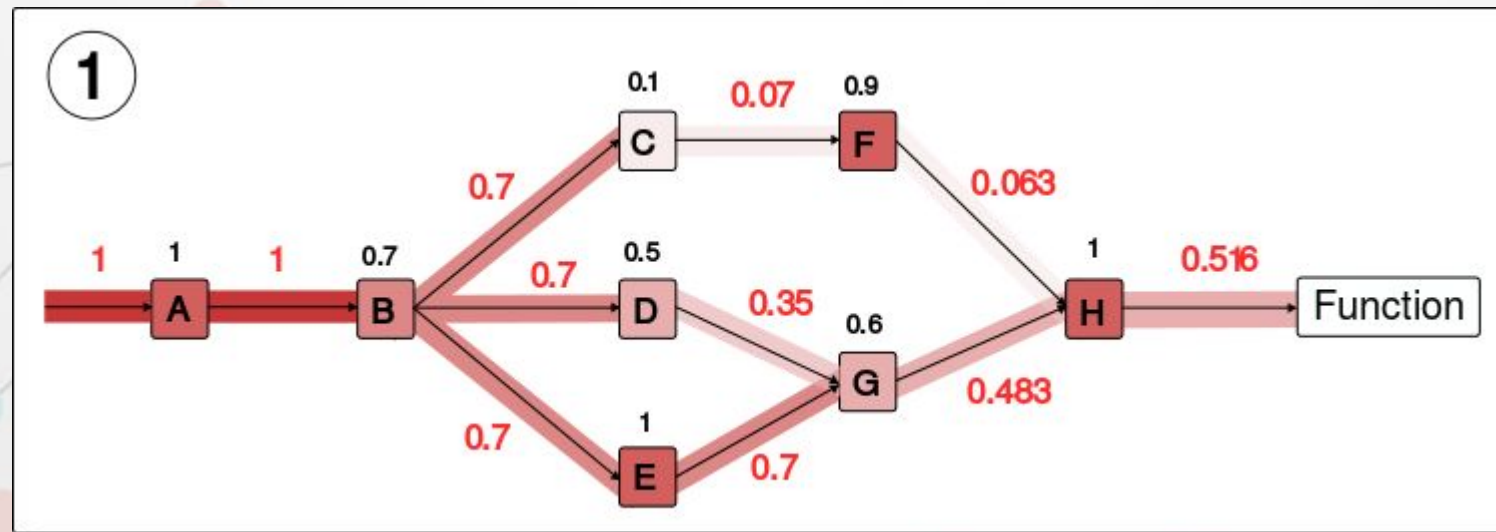
A : Activation edges

I : Inhibition edges



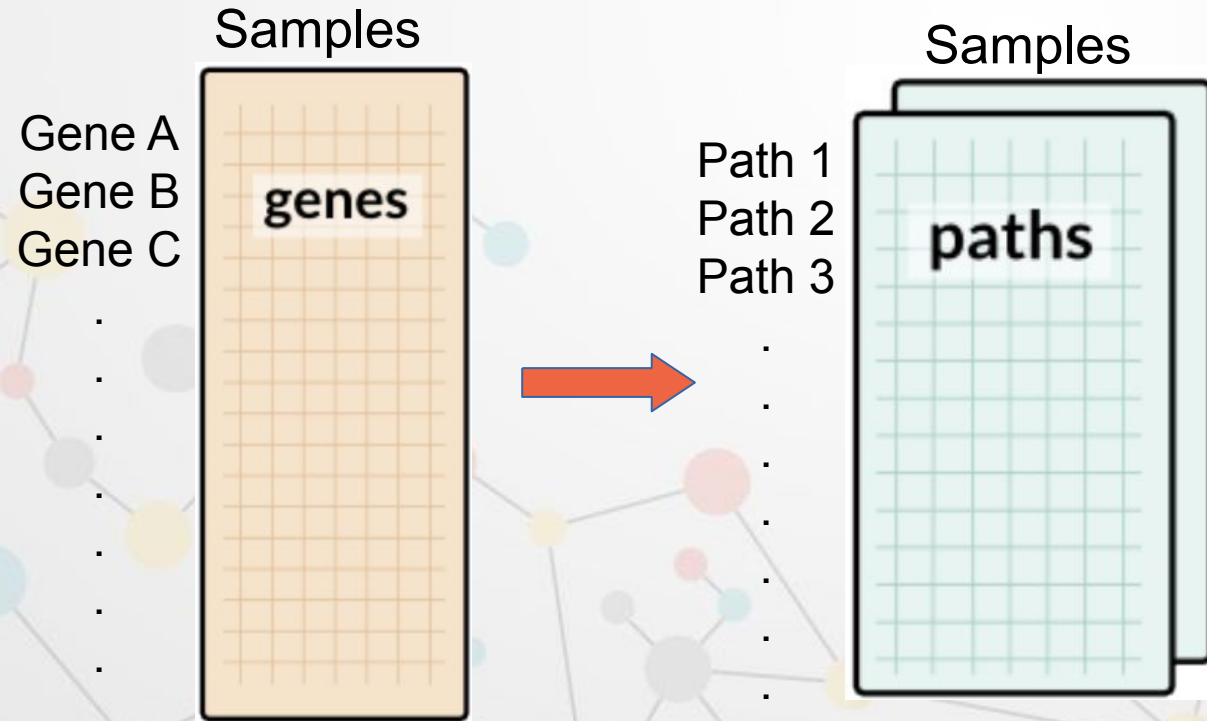
Signal values

- 1) Input signal **1** in any **input node**
- 2) Compute the signal through **each node iteratively**
- 3) **Loops** can be processed
- 4) **Subpathway signal**: last node signal



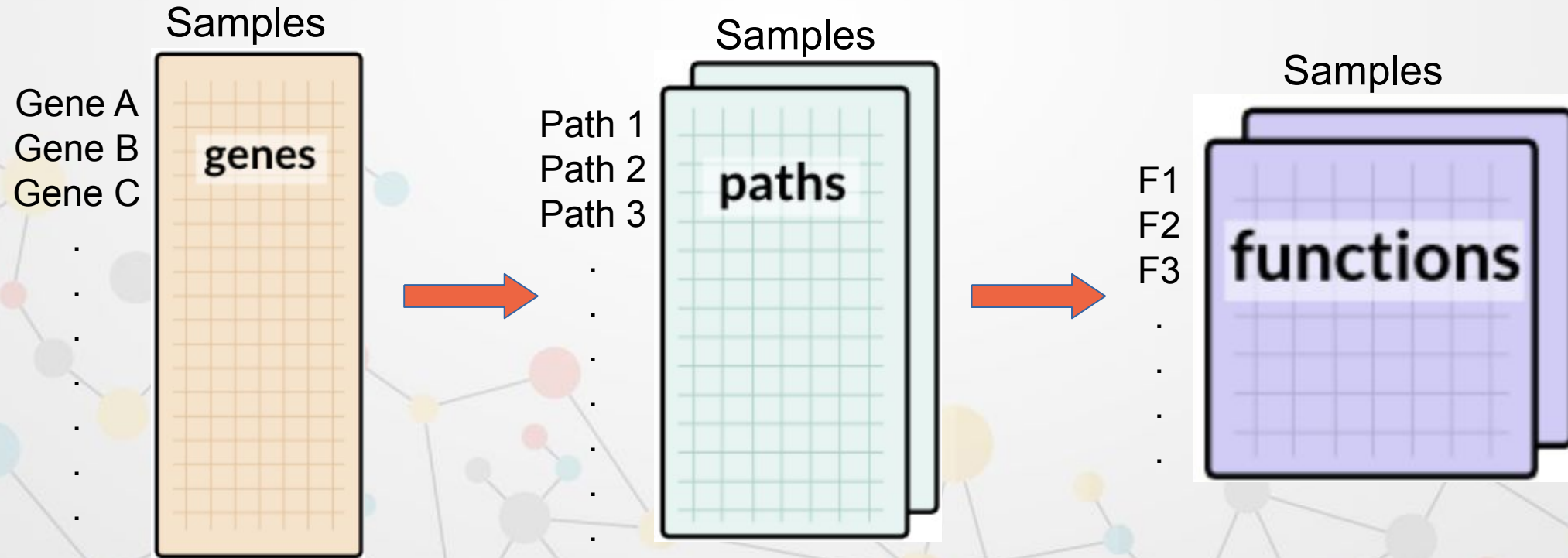
Functional annotation

1) Estimate effector proteins activation



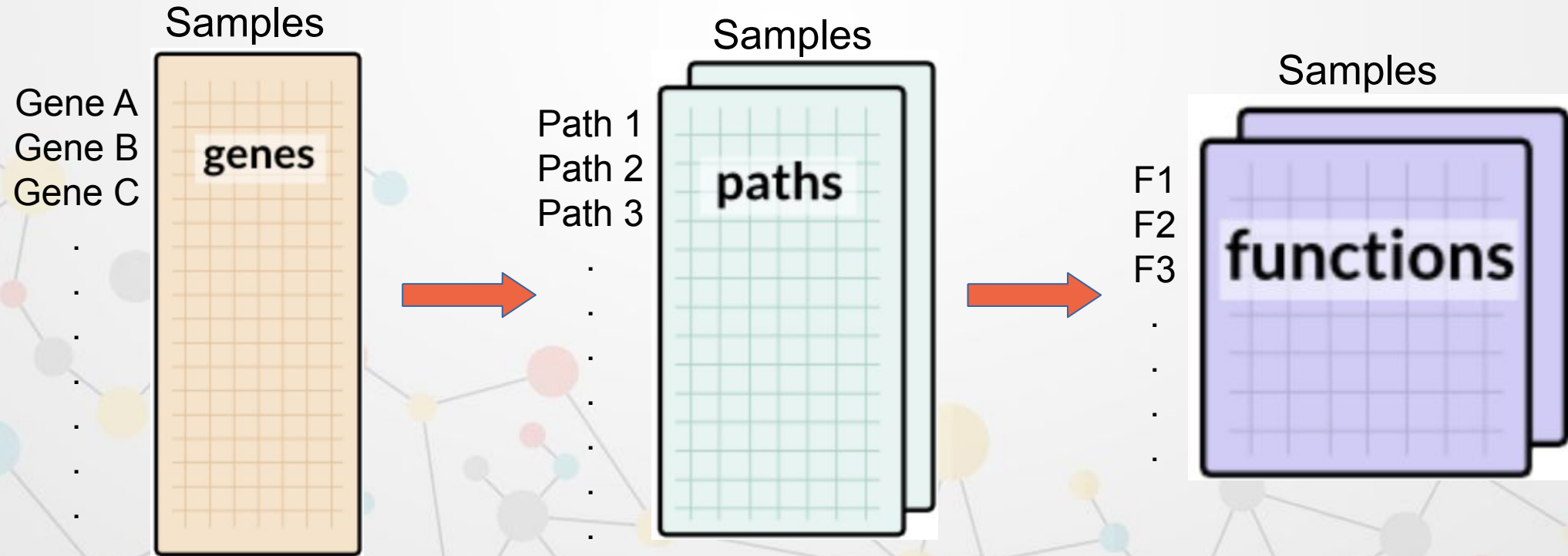
Functional annotation

- 1) Estimate effector proteins activation
- 2) Annotate effector proteins functions



Functional annotation

- 1) Estimate effector proteins activation
- 2) Annotate effector proteins functions
 - Uniprot keywords
 - GO annotation



Research Paper

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

Marta R. Hidalgo¹, Cankut Cubuk¹, Alicia Amadoz^{1,2}, Francisco Salavert^{1,3}, José Carbonell-Caballero¹, Joaquin Dopazo^{1,2,3}

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

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

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

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

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

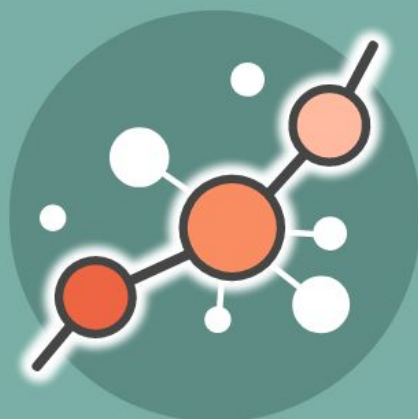
Received: September 01, 2016

Accepted: November 21, 2016

Published: December 22, 2016

ABSTRACT

Understanding the aspects of the cell functionality that account for disease or drug action mechanisms is a main challenge for precision medicine. Here we propose a new method that models cell signaling using biological knowledge on signal transduction. The method recodes individual gene expression values (and/or gene mutations) into accurate measurements of changes in the activity of signaling circuits, which ultimately constitute high-throughput estimations of cell functionalities caused by gene activity within the pathway. Moreover, such estimations can be obtained either at cohort-level, in case/control comparisons, or personalized for individual patients. The accuracy of the method is demonstrated in an extensive analysis involving 5640 patients from 12 different cancer types. Circuit activity measurements not only have a high diagnostic value but also can be related to relevant disease outcomes such as survival, and can be used to assess therapeutic interventions.



hiPathia

HIGH THROUGHPUT PATHWAY
INTERPRETATION AND ANALYSIS

Start

HiPathia is a web tool for the interpretation of the consequences of the combined changes of gene expression levels and/or genomic mutations in the context of signalling pathways. [Read more](#)

Note:

hiPathia web application makes an intensive use of the HTML5 standard and other cutting-edge web technologies such as Web Components, so only modern web browsers are fully supported.

This website is free and open to all users.



Case study: Breast Cancer

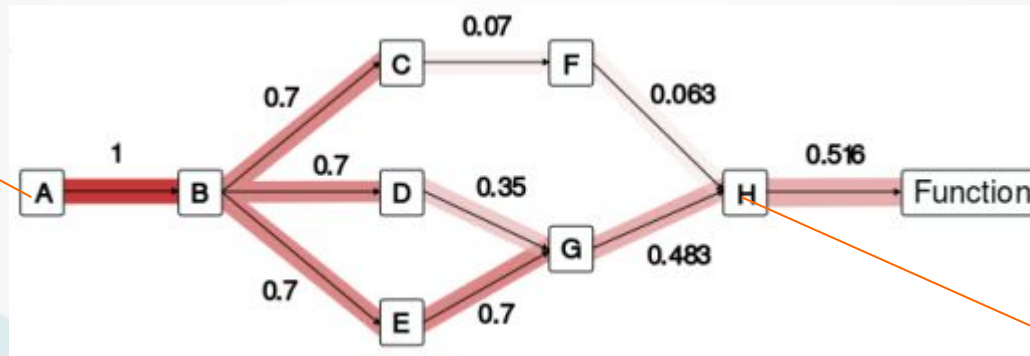
1) We will work with a **Breast Cancer** dataset from the repository
The Cancer Genome Atlas:

http://hipathia.babelomics.org/doc/doku.php?id=worked_example_differential_example

Perturbation effect tool

In our model, signal transduction is usually modulated through all nodes between a **membrane receptor** and an **effector protein**

Membrane Receptor

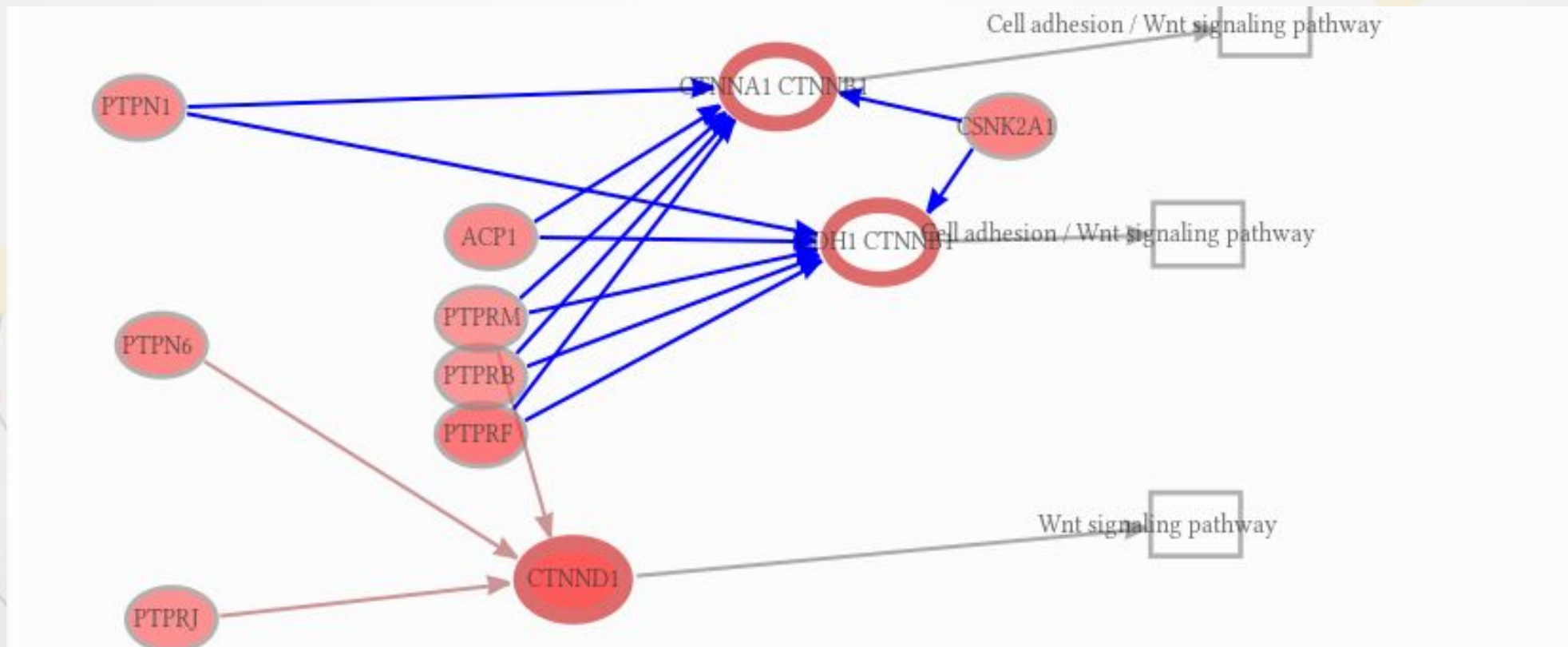


Intracellular signaling proteins

Effector protein

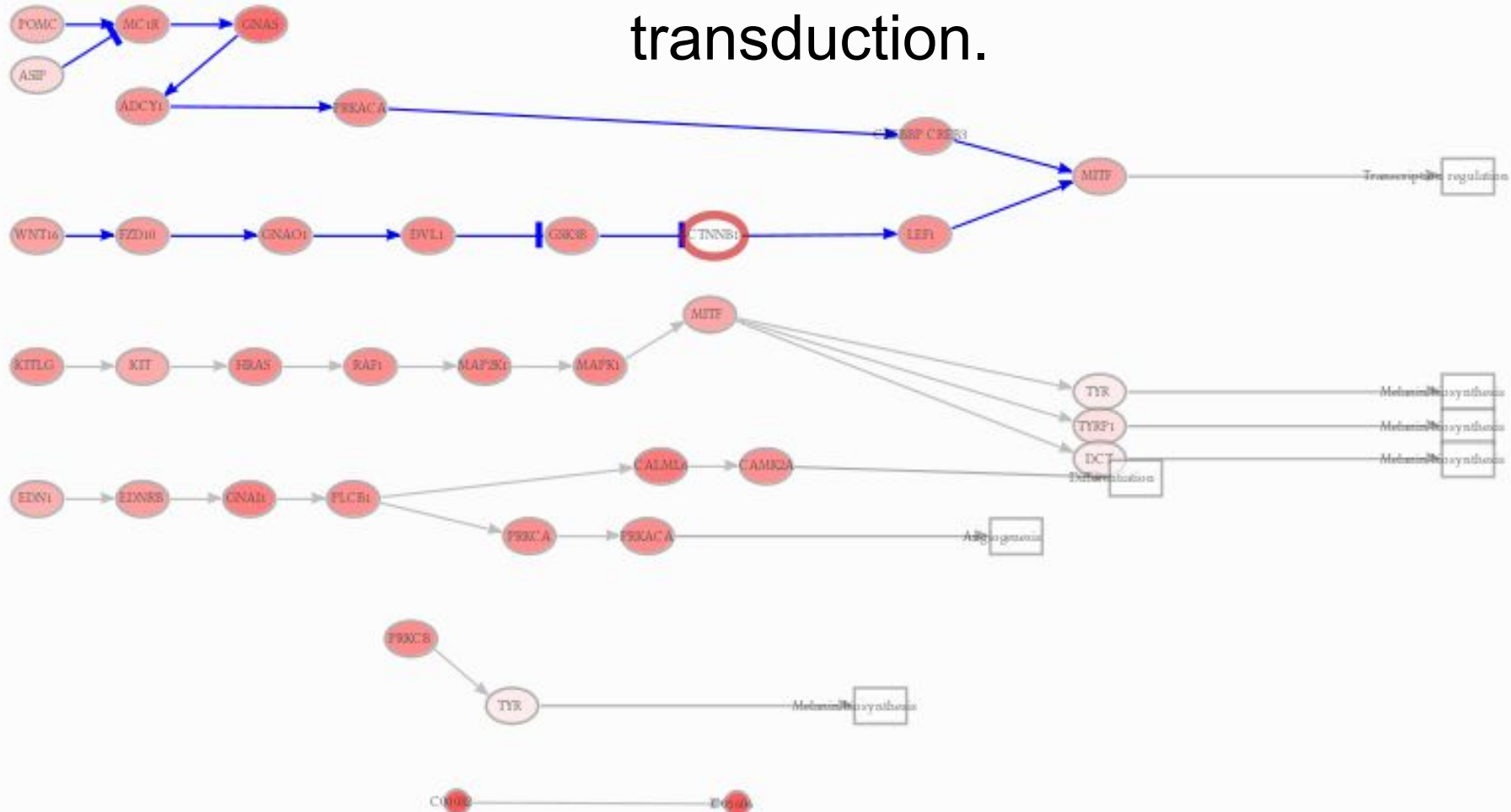
Perturbation effect tool

This approach offers a powerful framework to design targeted therapies with minimum resources.



Perturbation effect tool

We can use the same machinery to **virtually explore the consequences of knocking out** (or overexpress) a gene in signal transduction.



CyPathia

The signal propagation model has been implemented also for **Cytoscape community**



Cytoscape App Store

Submit an App ▾

Search the App Store

Kinza ▾

Want an easier way to install apps? [Click here to learn how!](#) ✕

[← Go back to home](#)



CyPathia

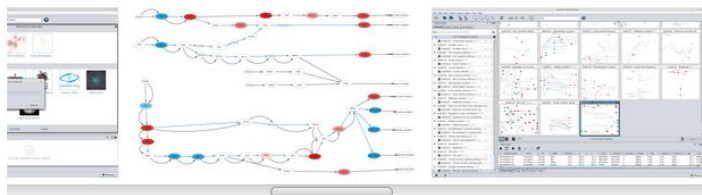
High Throughput pathway interpretation and analysis

★★★★★ (1) 198 downloads | [helpdesk](#) | [discussions](#) 

[Editor Actions ▾](#)

Details

Release History



CYTOSCAPE 3

[Download](#)

Version 1.0.0

Released 25 Sep 2019

Works with [Cytoscape 3.7](#)

[Download Stats](#) [Click here](#)

RESOURCES

- [Ask a question](#)
- [Search helpdesk](#)
- [Search BioStars](#)
- [E-mail](#)

The first version of **CyPathia** was developed for Cytoscape (3.6.1 +). It provides a user friendly and straightforward interface. The CyPathia app is based on **Hipathia Bioconductor** package, allowing the Cytoscape community for the first time the possibility of using mechanistic models. In order to call **Hipathia**, we have to invoke R from **Cytoscape**. Thus **JRI Java/R Interface** have been used (Fig.1). JRI allows to run a R session inside Java applications as a single thread. **CyPathia 1.0.0** provides two functionality, first one is the computation of signal transduction along KEGG signaling pathways from transcriptomic data. The user has to load both of the expression matrix file and experimental design file which has at least two groups. In the current version, CyPathia supports **145 networks from KEGG** for three species (**Human, Mouse and Rat**). Next version will support directed networks. **CyPathia** use **wilcoxon test** to compare the signal values -which are already calculated using Hipathia- between different conditions loaded previously in the experimental design file. The second functionality is to import generated report (sifs, natt and eatt files) from both [hipathia web](#) or [Hipathia package](#).

Case Study:



CENTRO DE DOCUMENTACIÓN VIRGEN DEL RÍO

ENTRADA HOSPITAL

KEEP CALM AND SEQ ON



@ClinicalBioinfo