Uncovering signaling differences between normal and transformed hepatocytes using cell-specific pathway models

Julio Saez-Rodriguez, Ph. D.

EMBL-EBI

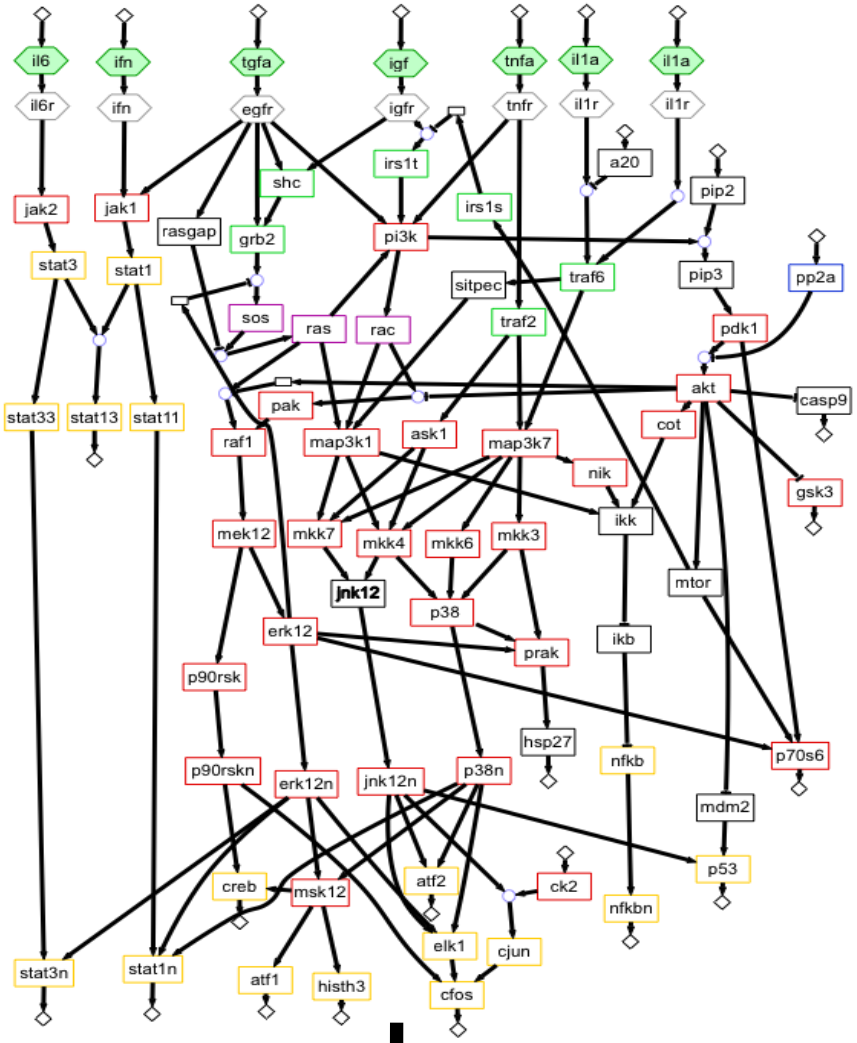
European Bioinformatics Institute
Hinxton, Cambridge - UK

Uncovering signaling differences between normal and transformed hepatocytes using cell-specific pathway models

Julio Saez-Rodriguez, Ph. D.

Department of Systems Biology, Harvard Medical School
&
Biological Engineering Department, M.I.T.

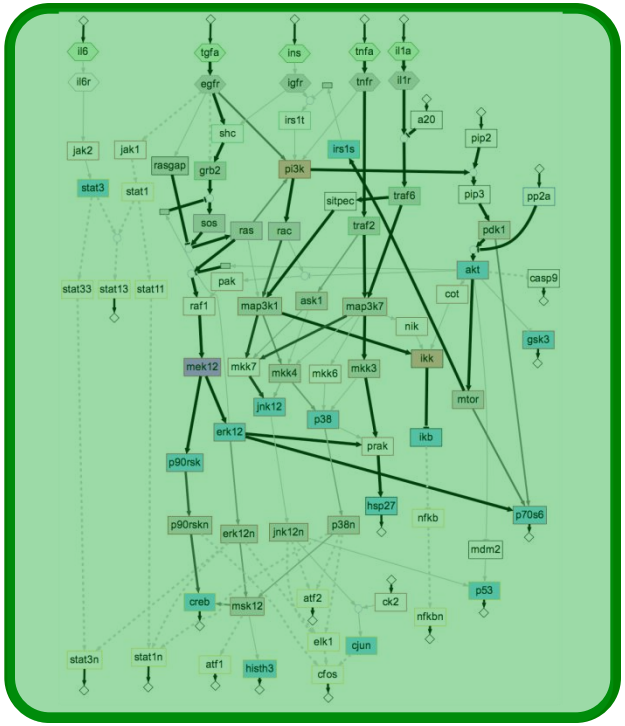
How do cells process extracellular signals?



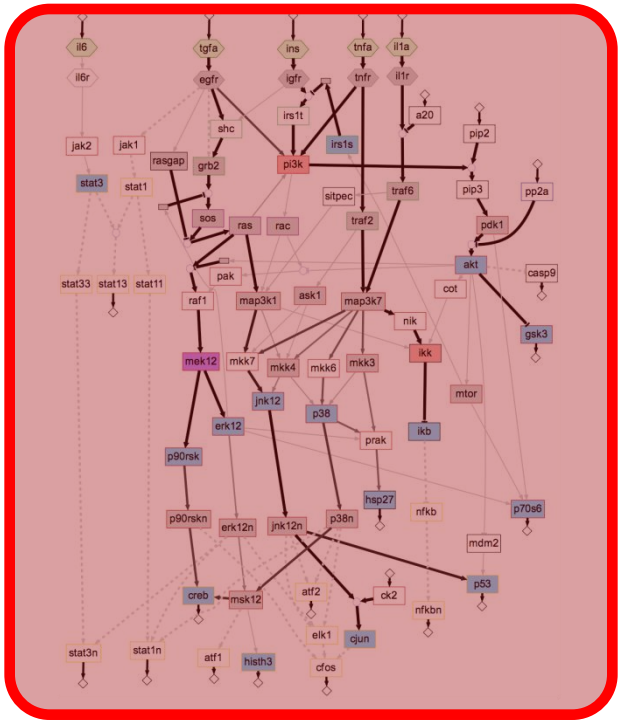
Phenotype

How is signal processing altered in disease?

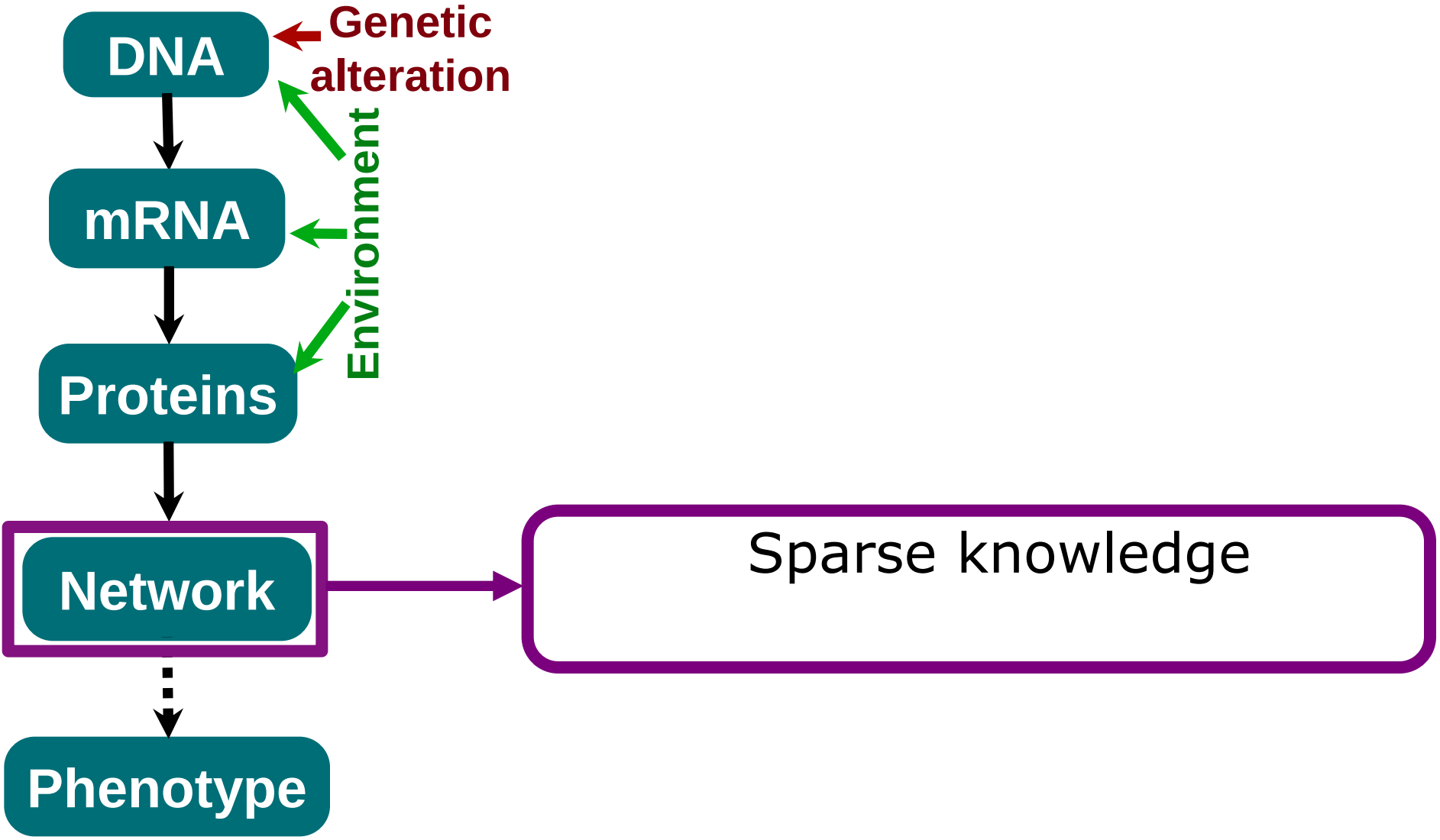
Normal



Diseased

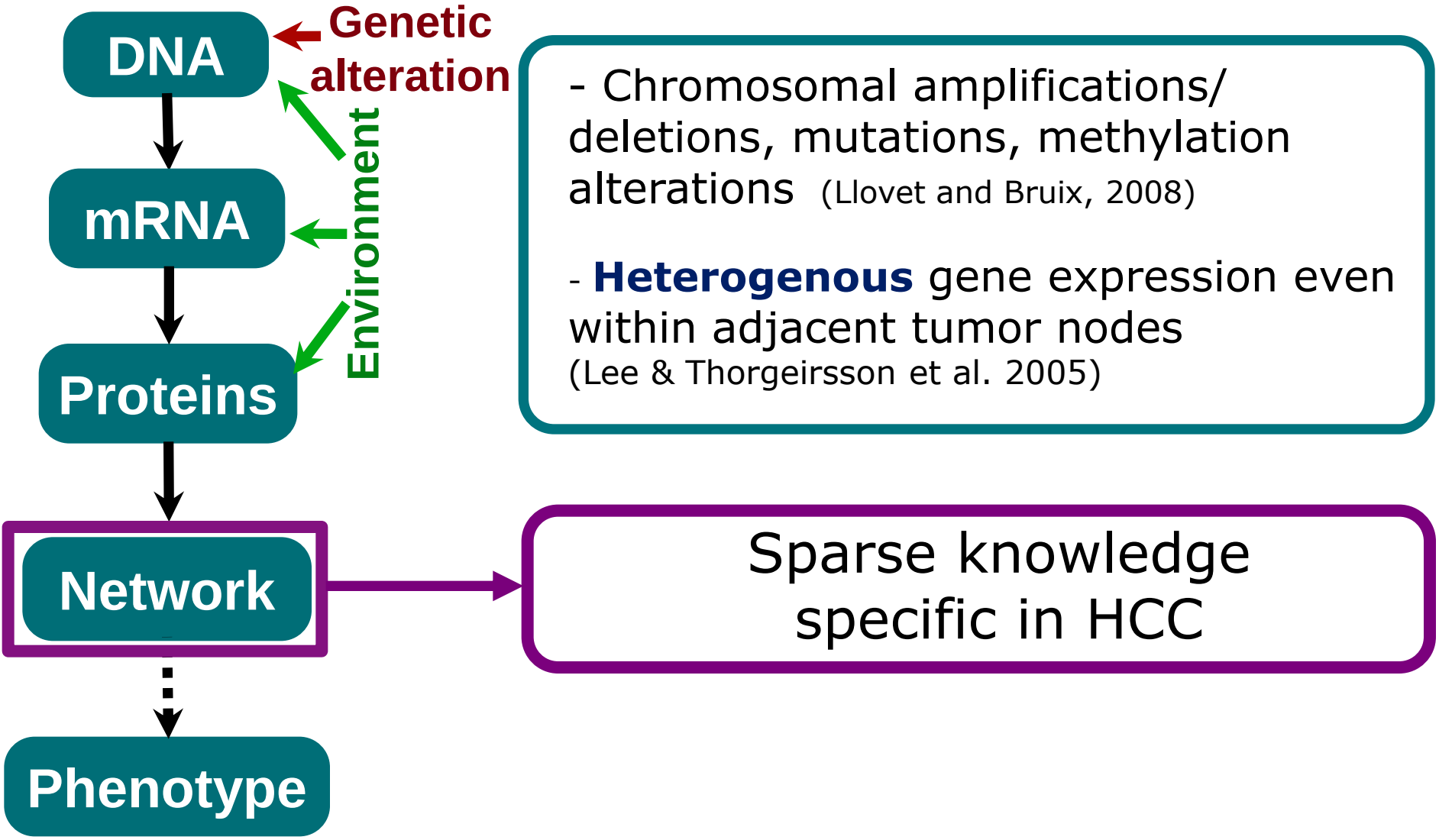


How is signal processing altered in disease?

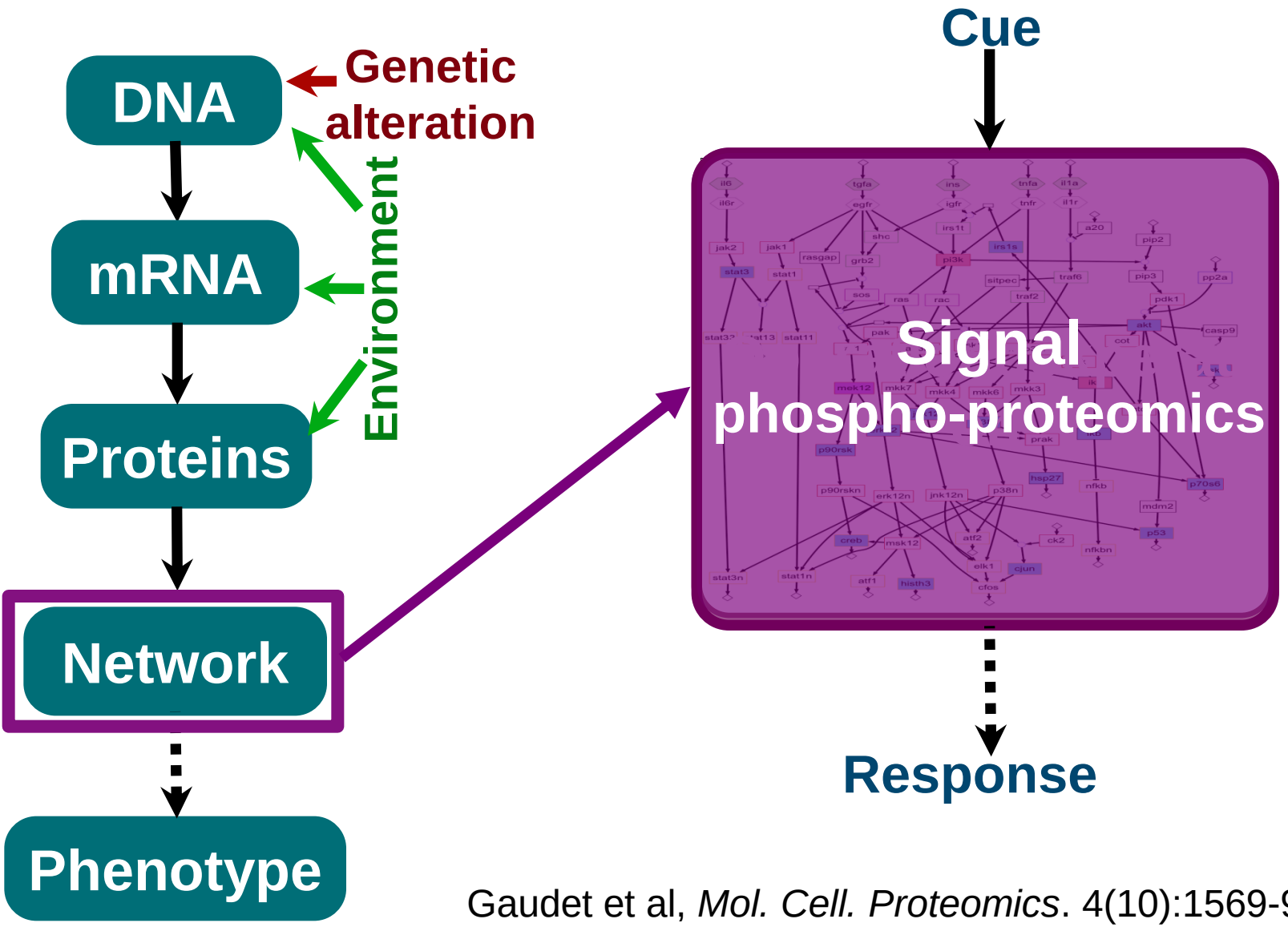


Case study: how is signaling altered in transformed vs normal hepatocytes?

Hepatocellular Carcinoma (HCC):
most frequent form of liver cancer, 3rd most lethal cancer

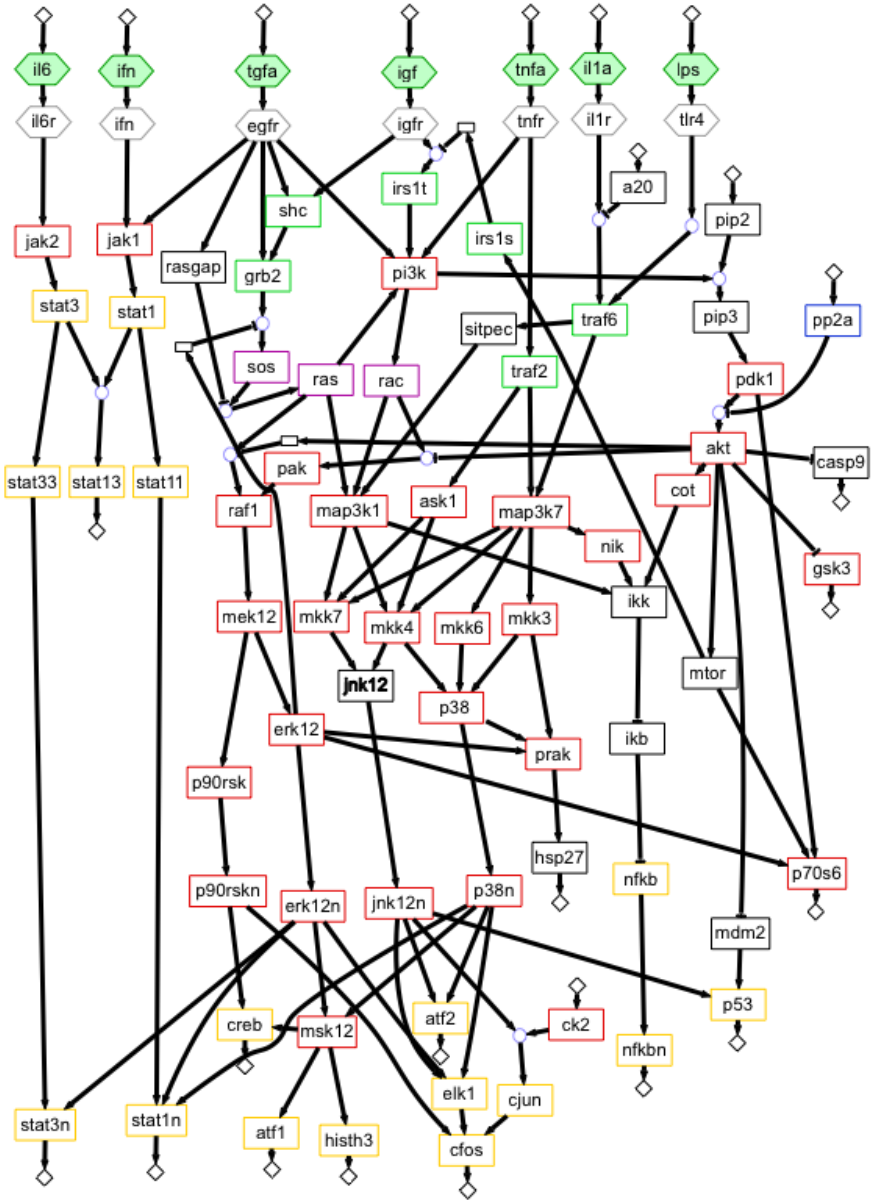


Cue-Signal-Response approach to understand operation of signal transduction pathways



Gaudet et al, *Mol. Cell. Proteomics*. 4(10):1569-90 (2005).

Design of Cue-Signal-Response experiment for HCC



Construct map of canonical pathways from *Ingenuity* (85 species)

Select

- **perturbations** (chemical inhibitors = drugs)
&
- **signals** (phosphorylations measurable with Luminex/xMAP technology)

as distributed in the network as possible

Stimulus | Perturbation

Readout | Perturb&Read

Design of Cue-Signal-Response experiment for HCC

Performed by Leonidas Alexopoulos

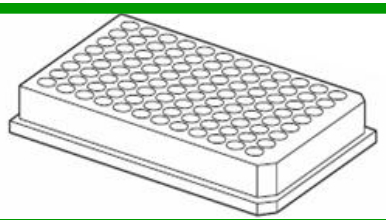
Primary human hepatocytes & HCC cell line HepG2

Cue

➜ 7 **extracellular ligands**

➜ 7 specific **chemical inhibitors** (drugs)

at different times after stimulation



Signal

➜ **Phosphorylation** of 17 key proteins (30 min, 3h)

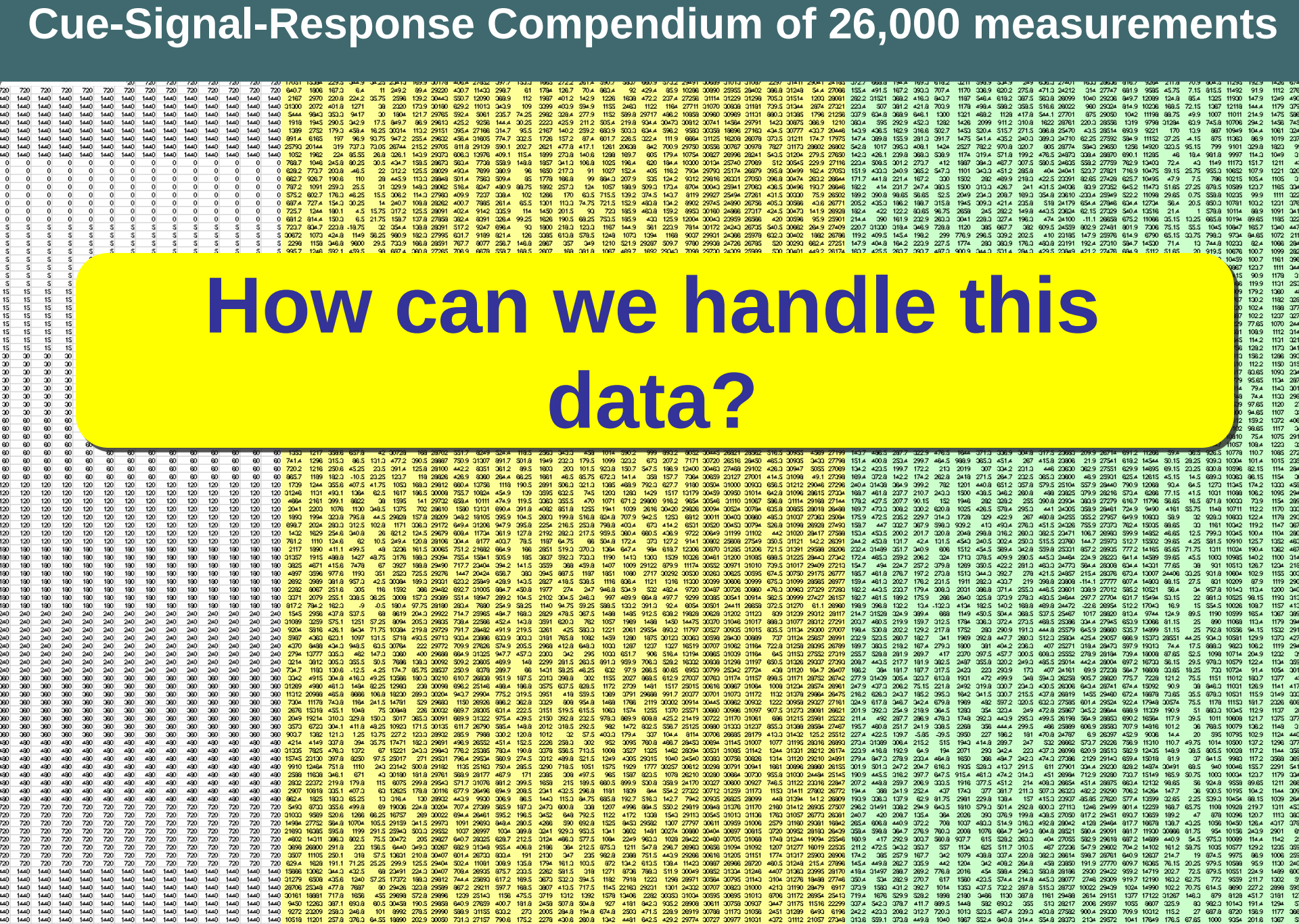
using Luminex/xMAP (bead-based ELISA)

Response

➜ **Release** of 20 cytokines (3h, 24h)

Alexopoulos L, Saez-Rodriguez J, Cosgrove B, Lauffenburger DA, Sorger PK, *Mol. Cell. Proteom, in press*

Cue-Signal-Response Compendium of 26,000 measurements



Cue-Signal-Response Compendium of 26,000 measurements

THE FAST TRACK TO MODEL TRAINING!

Flexible Informatics for Linking Experimental Data to Mathematical Models

Saez-Rodriguez J, Goldispe A, Muhlich J, Alexopoulos A, Millard B, Lauffenburger DA, Sorger PK, *Bioinformatics,* 24:6, 2008.

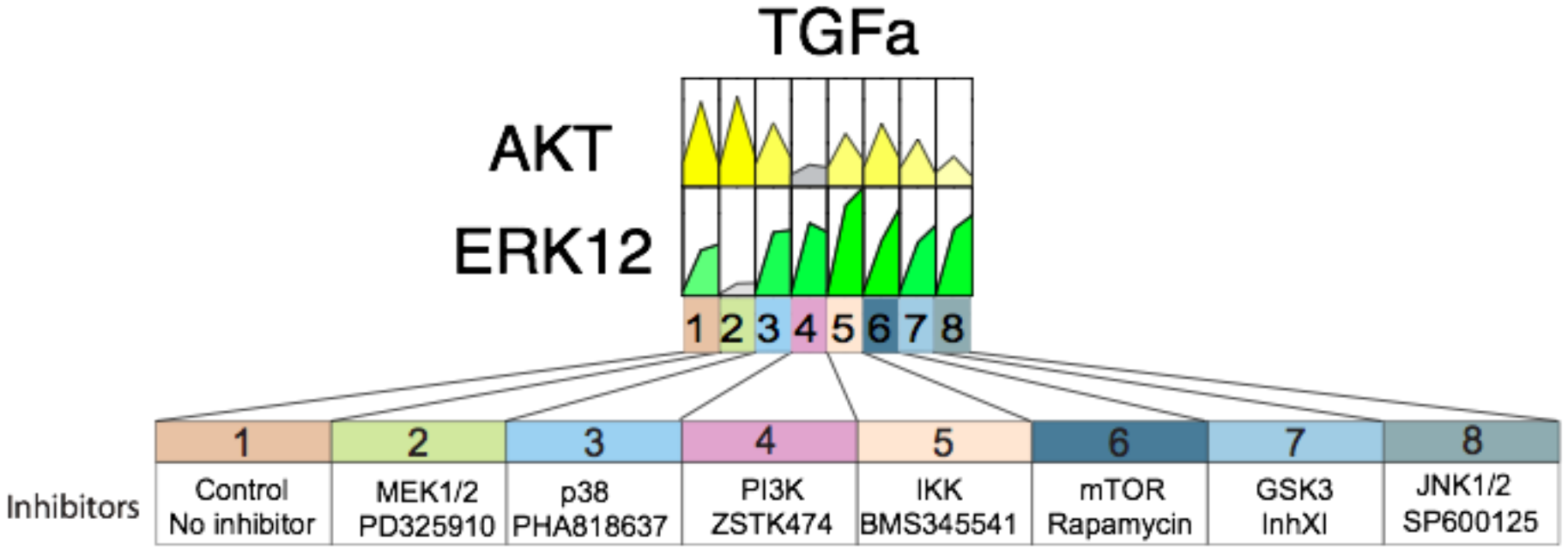
Open-source MATLAB Toolbox
Script& User Interface

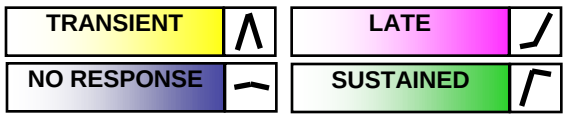
- Maintains data provenance - Modular (easy to expand)

Available at

http://code.google.com/p/sbpipeline/

Visualization of large data sets





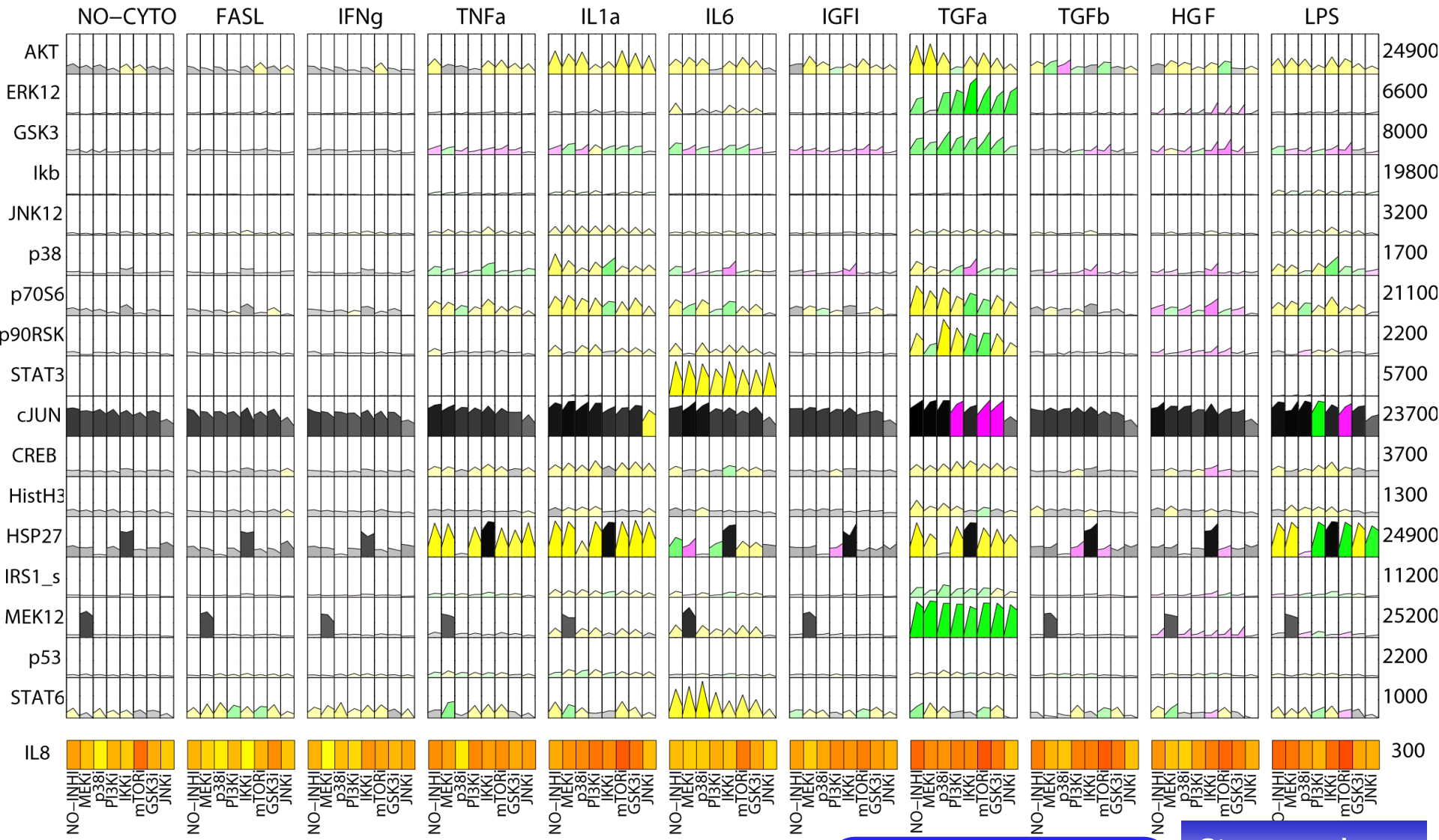
TIME POINTS:
0, 30min, 3hrs

Primary Hepatocytes

Stronger primary

Stronger HepG2

Visualization of large data sets



TRANSIENT | LATE | NO RESPONSE | SUSTAINED

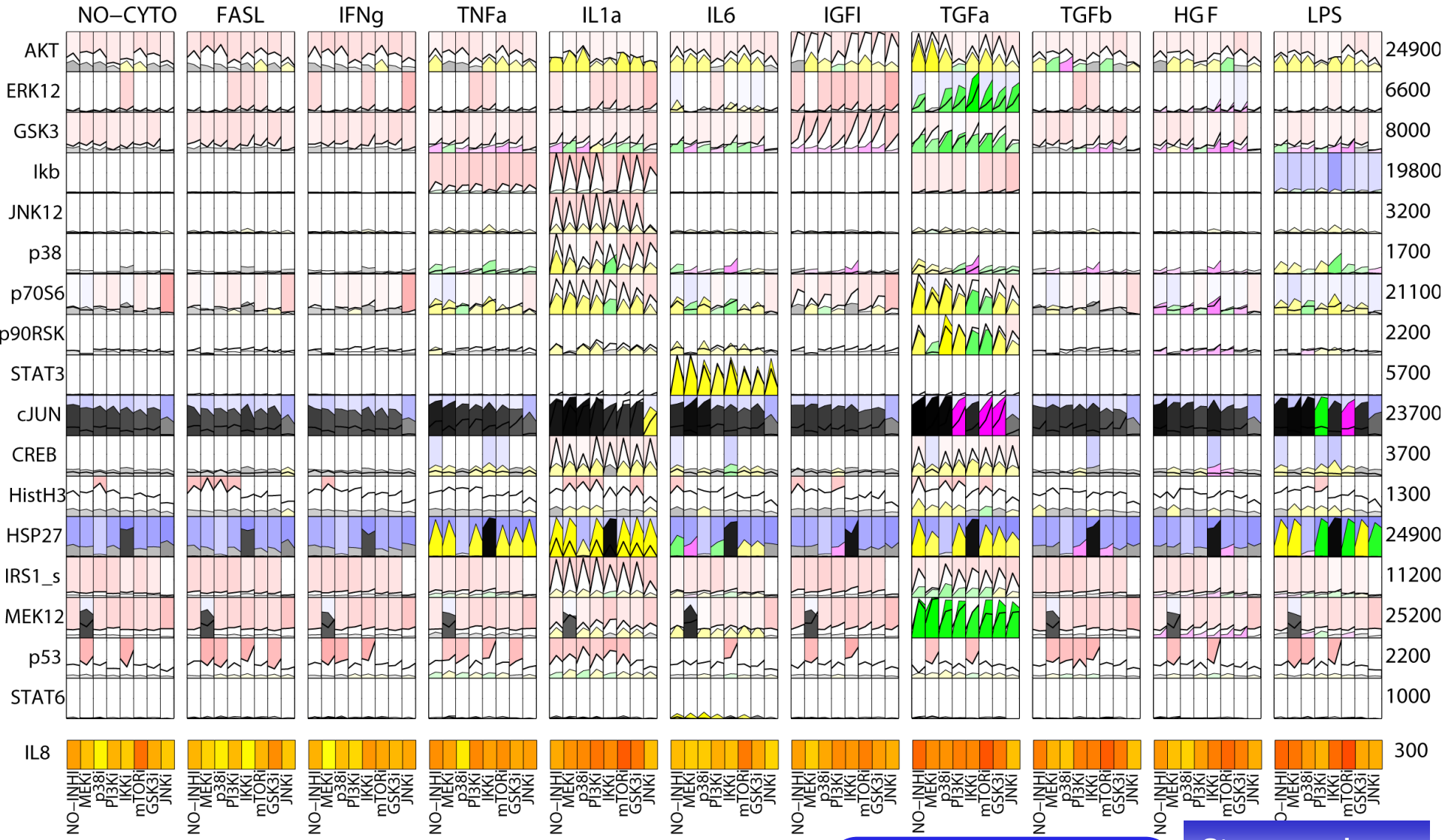
TIME POINTS:
0, 30min, 3hrs

Primary Hepatocytes

Stronger primary

Stronger HepG2

Visualization of large data sets



TRANSIENT

LATE

NO RESPONSE

SUSTAINED

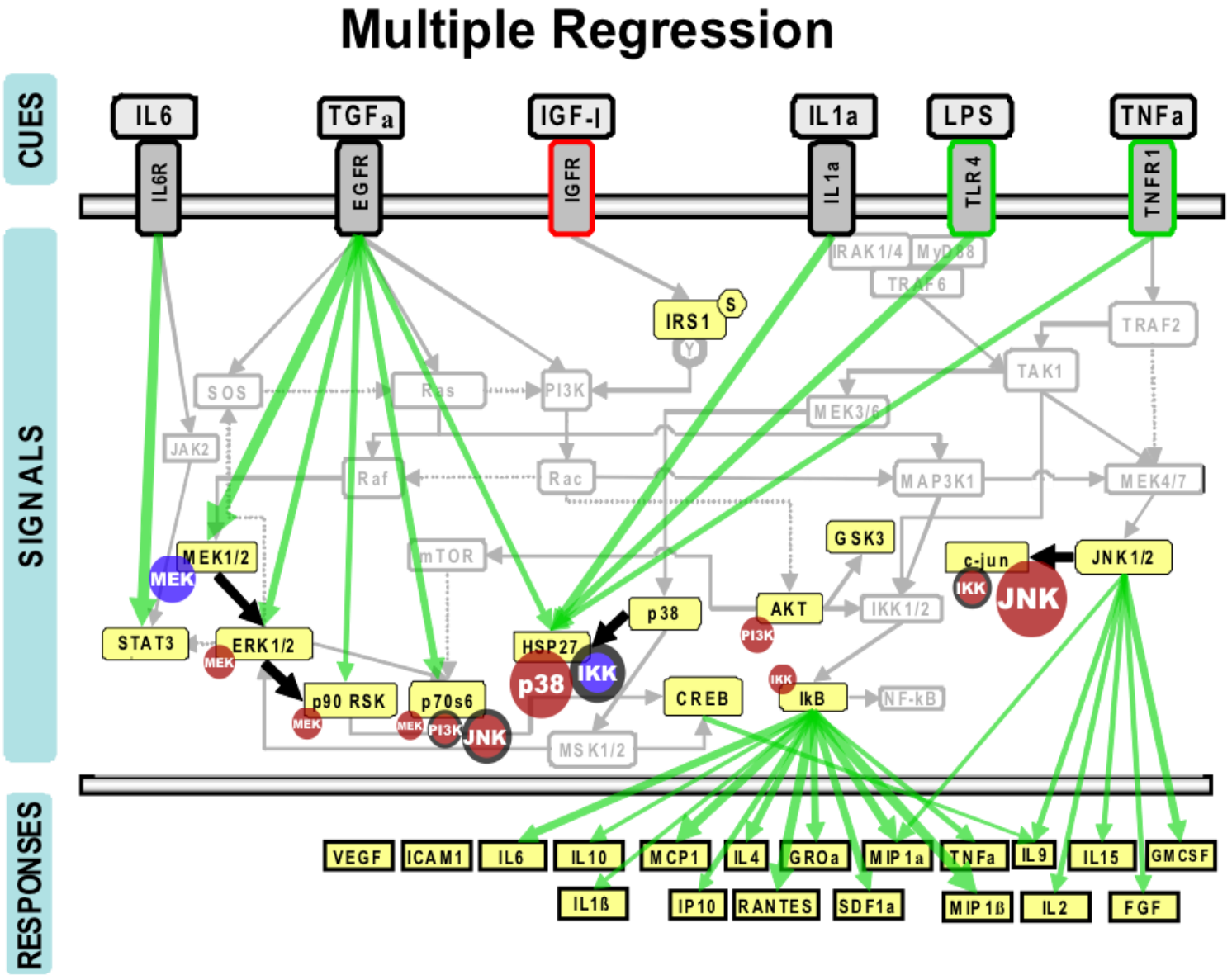
TIME POINTS:
0, 30min, 3hrs

Primary Hepatocytes

Stronger primary

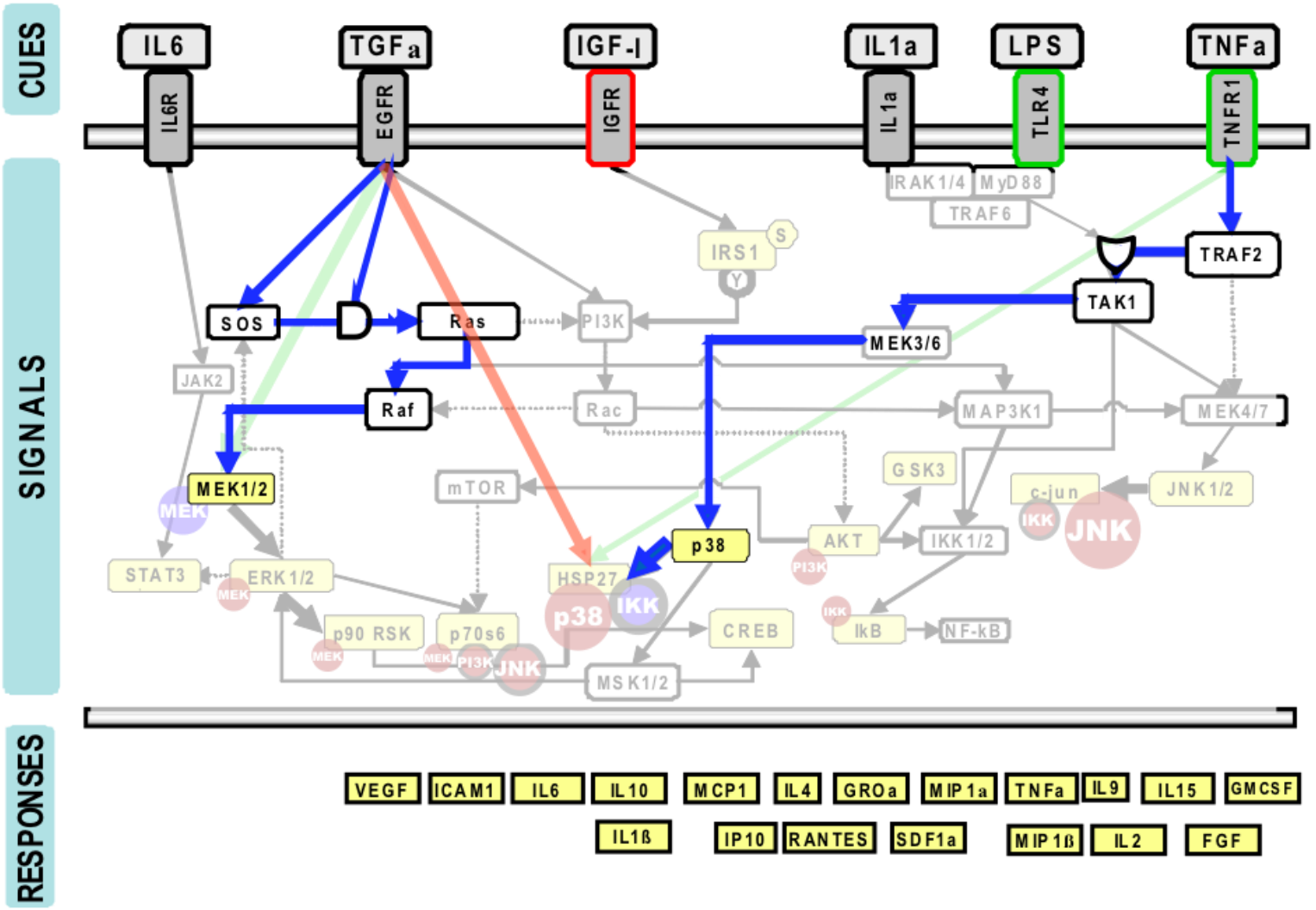
Stronger HepG2

Data-driven approaches useful but (in our case) provide limited mechanistic insight



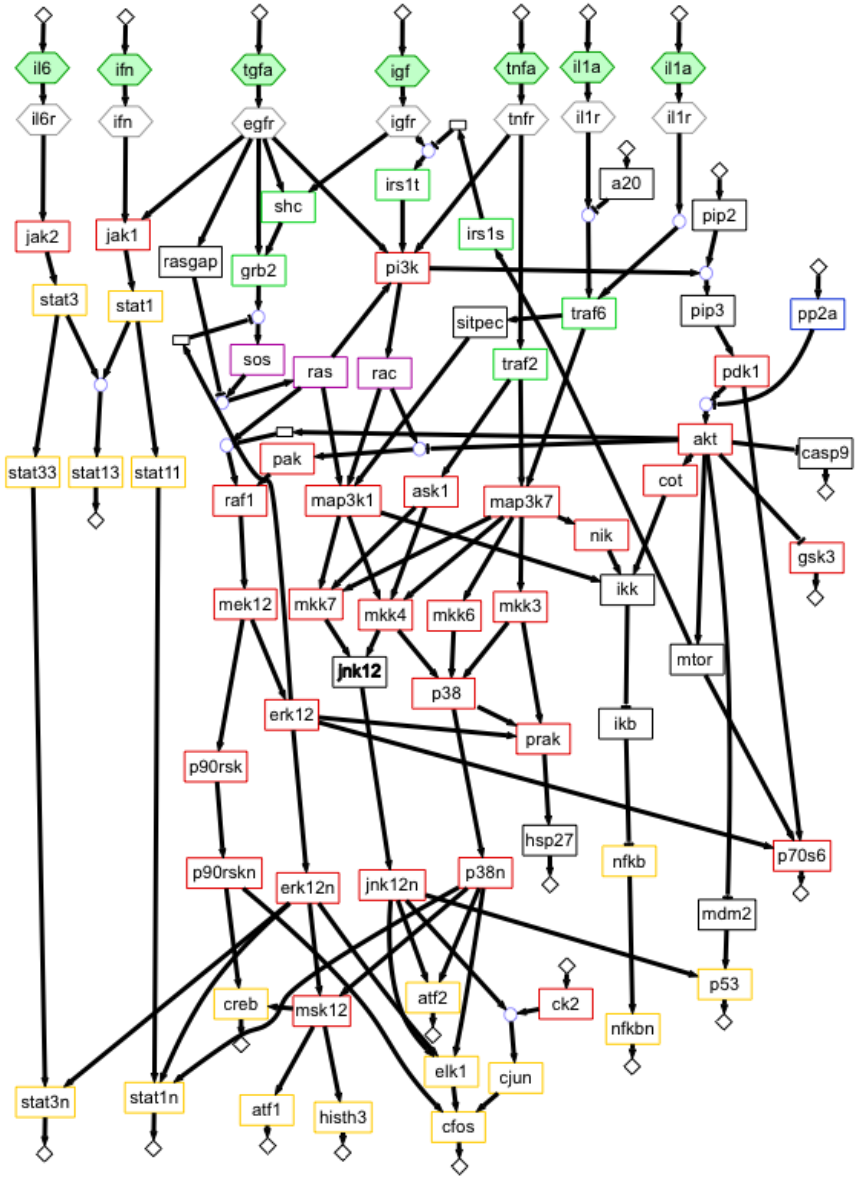
Alexopoulos L, Saez-Rodriguez J, Cosgrove B, Lauffenburger DA, Sorger PK, *Mol. Cell. Proteom, in press*

Data-driven approaches useful but (in our case) provide limited mechanistic insight



Alexopoulos L, Saez-Rodriguez J, Cosgrove B, Lauffenburger DA, Sorger PK, *Mol. Cell. Proteom, in press*

Signaling pathway maps summarize literature knowledge

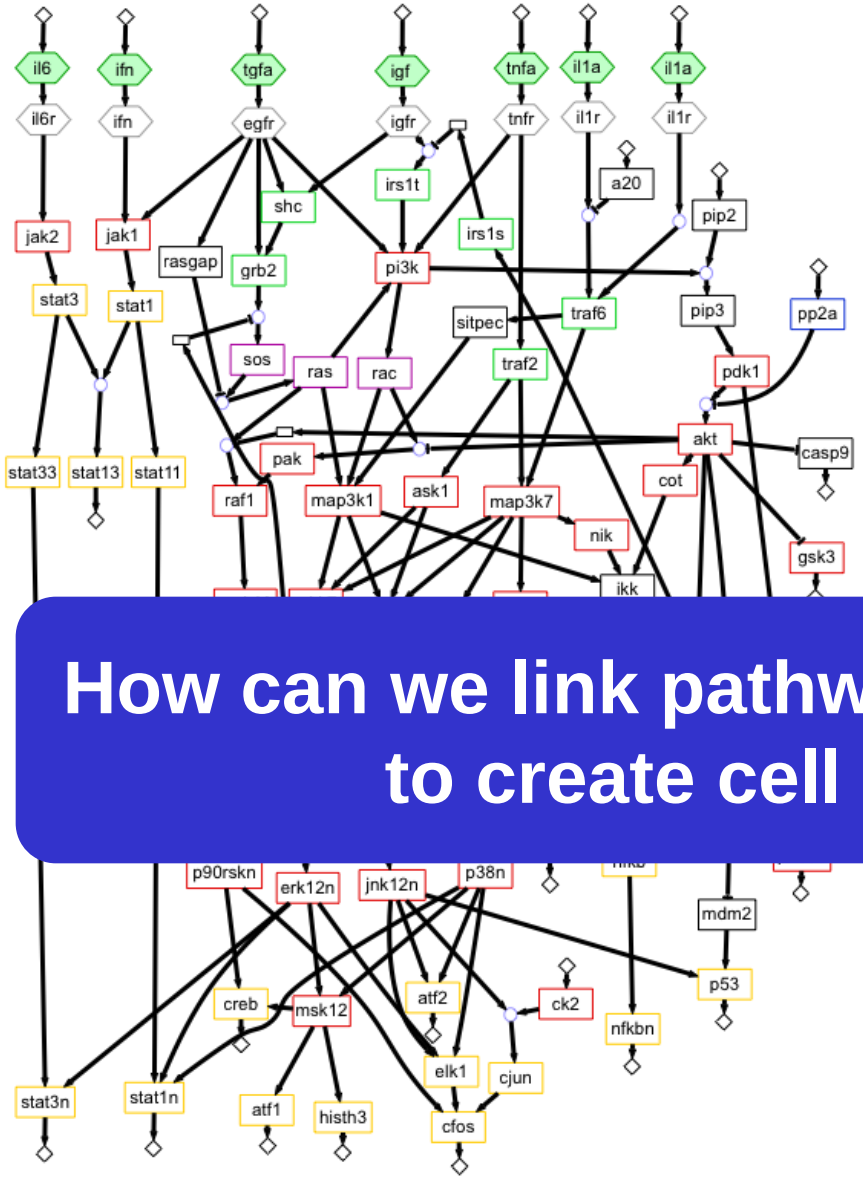


Pathway maps widespread and very **useful** but

- ***Pictures* not computable** models to study signal processing

- Not cell-type specific

Signaling pathway maps summarize literature knowledge



Pathway maps widespread and very **useful** but

- ***Pictures* not computable** models to study signal processing

- Not cell-type specific

How can we link pathway maps to signaling data to create cell specific models?

Challenges to link pathway maps to data of signal transduction

- Make **maps executable** (models) so that experiments can be simulated
- Define **metric** to **evaluate** models given the data
- Develop a framework to **explore** models & **identify best**

Saez-Rodriguez J, Alexopoulos LG, Epperlein J, Samaga R, Lauffenburger DA, Klamt S, Sorger PK, *Mol. Syst. Biol,* 5, 331, 2009.

Challenges to link pathway maps to data of signal transduction

- Make **maps executable** (models) so that experiments can be simulated

  ➪ Transform into **Boolean** (0/1) logic (AND/OR) models ✔

- Define **metric** to **evaluate** models given the data

  ➪ Balance fit to data with model size ✔

- Develop a framework to **explore** models & **identify best**

  ➪ (i) Compress map

  (ii) Construct an 'scaffold' with all possible models (all gates) compatible with map

  (iii) find model with optimal metric (train) ✔

Saez-Rodriguez J, Alexopoulos LG, Epperlein J, Samaga R, Lauffenburger DA, Klamt S, Sorger PK, *Mol. Syst. Biol,* 5, 331, 2009.

Challenges to link pathway maps to data of signal transduction

CellNetOptimizer

Matlab toolbox, script & user interface
freely available at
http://www.ebi.ac.uk/saezrodriguez/software

Different (&complementary) from standard reverse engineering:

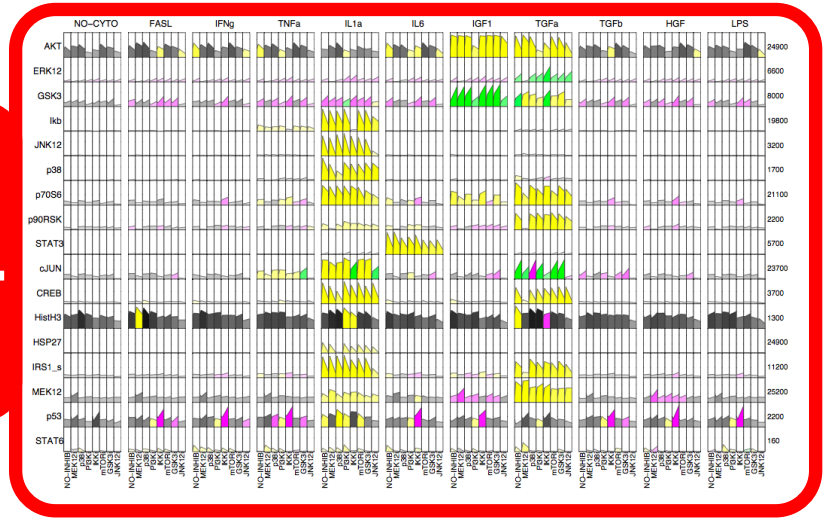
- Relies on **prior knowledge**
→ in a first step, reduces search space to plausible connections

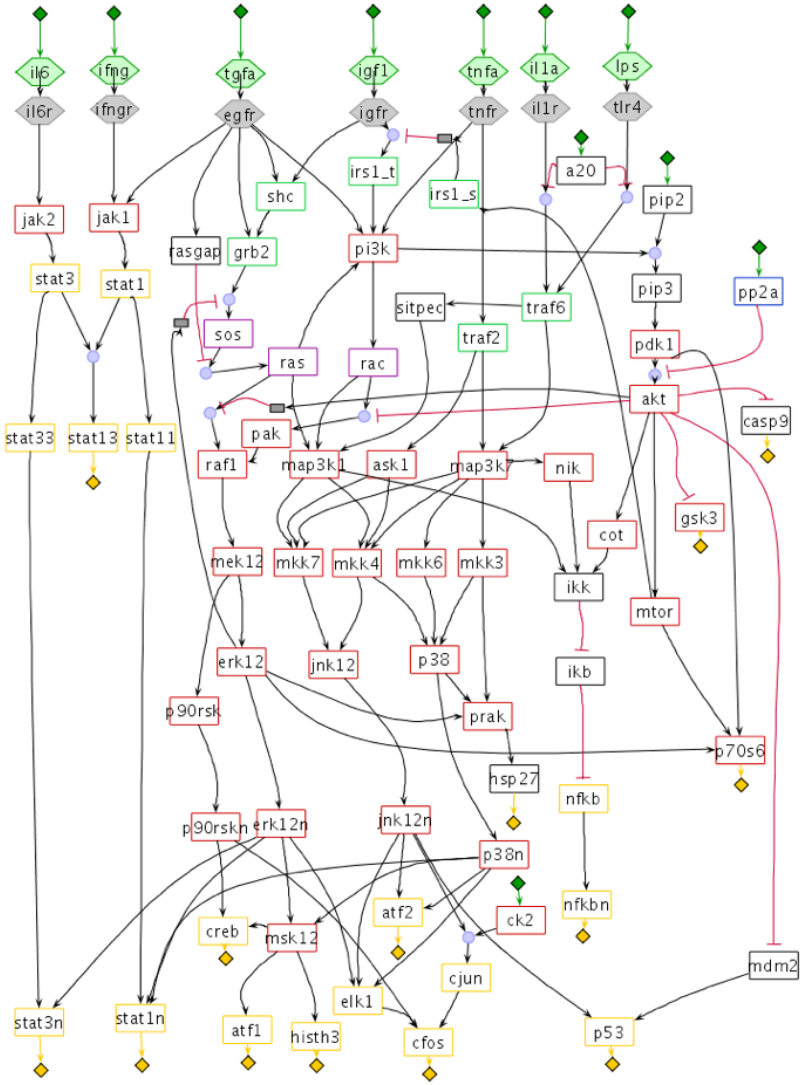
- Considers **intermediates**
→ mechanistic insight

Saez-Rodriguez J, Alexopoulos LG, Epperlein J, Samaga R, Lauffenburger DA, Klamt S, Sorger PK, *Mol. Syst. Biol,* 5, 331, 2009.

Application to signaling in hepatocytes

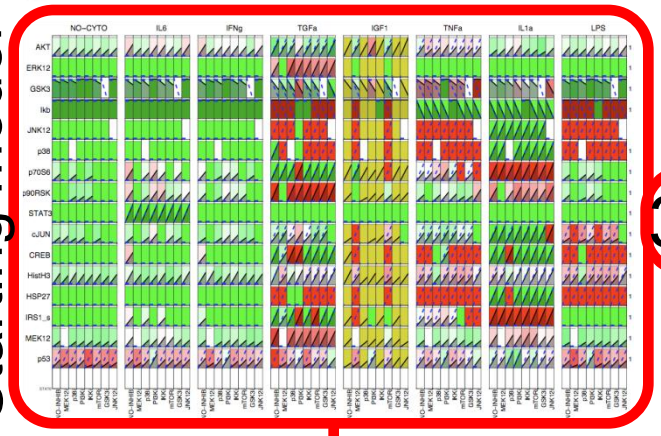
HepG2





Model trained to HepG2 data

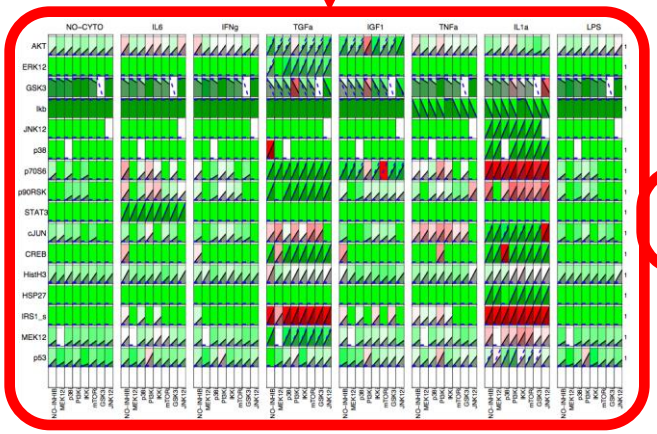
Starting model



Error

34.3%

Trained model



8.1%

Stimulus

Perturbation

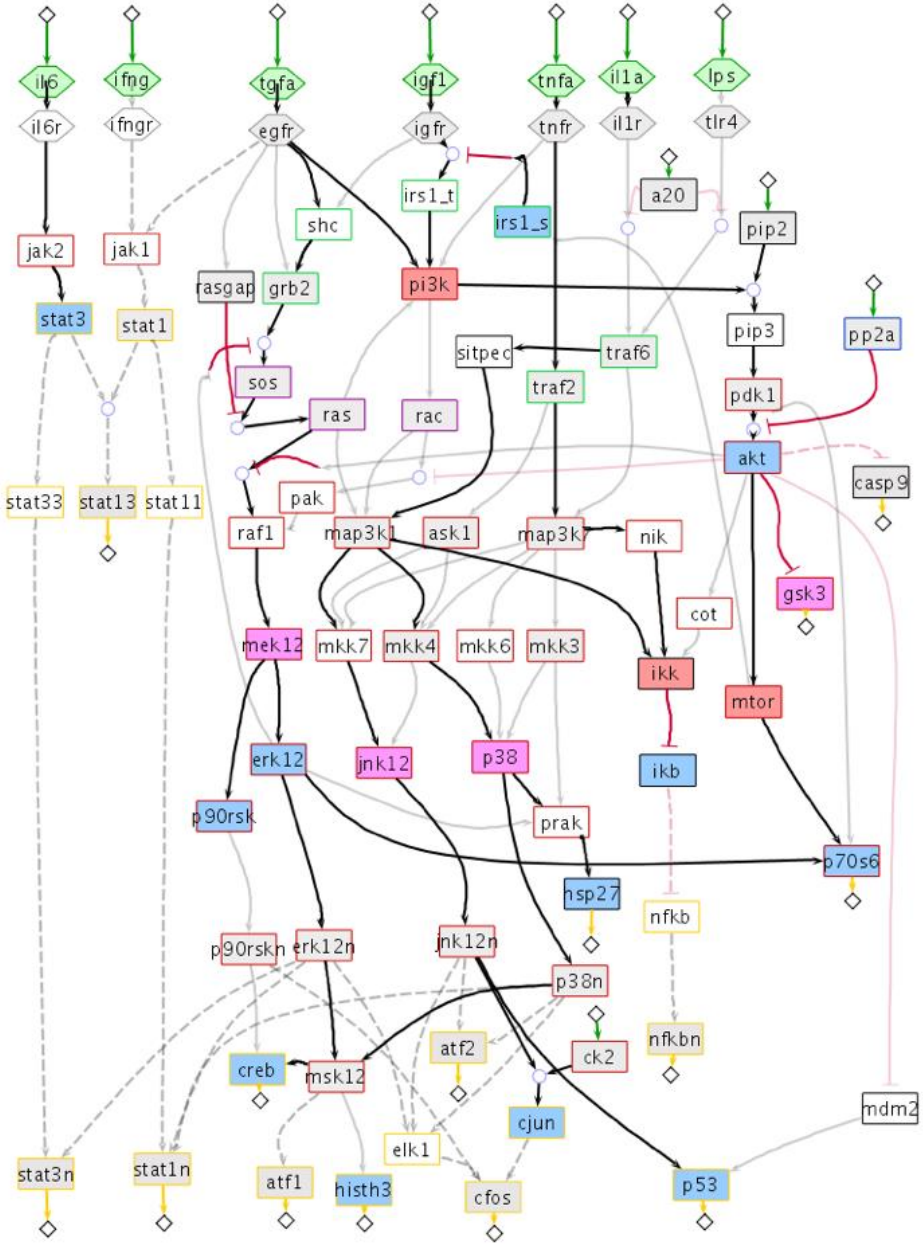
Readout

Perturb&Read

Kept

Removed

No effect

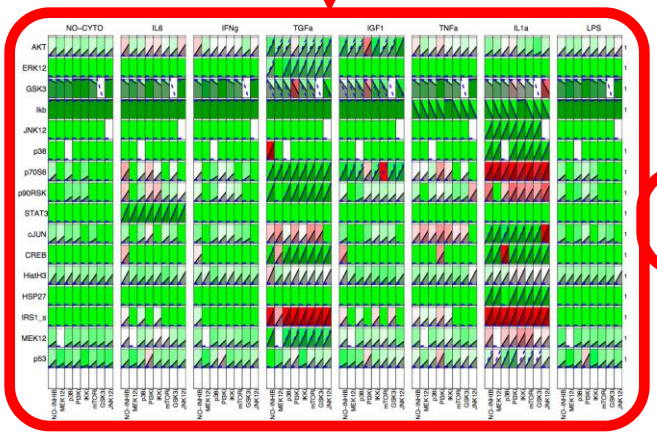


Model trained to HepG2 data

Starting model

- Model has good predictive power (cross validation + test set)
- Statistically significant better than random data & networks

Trained model



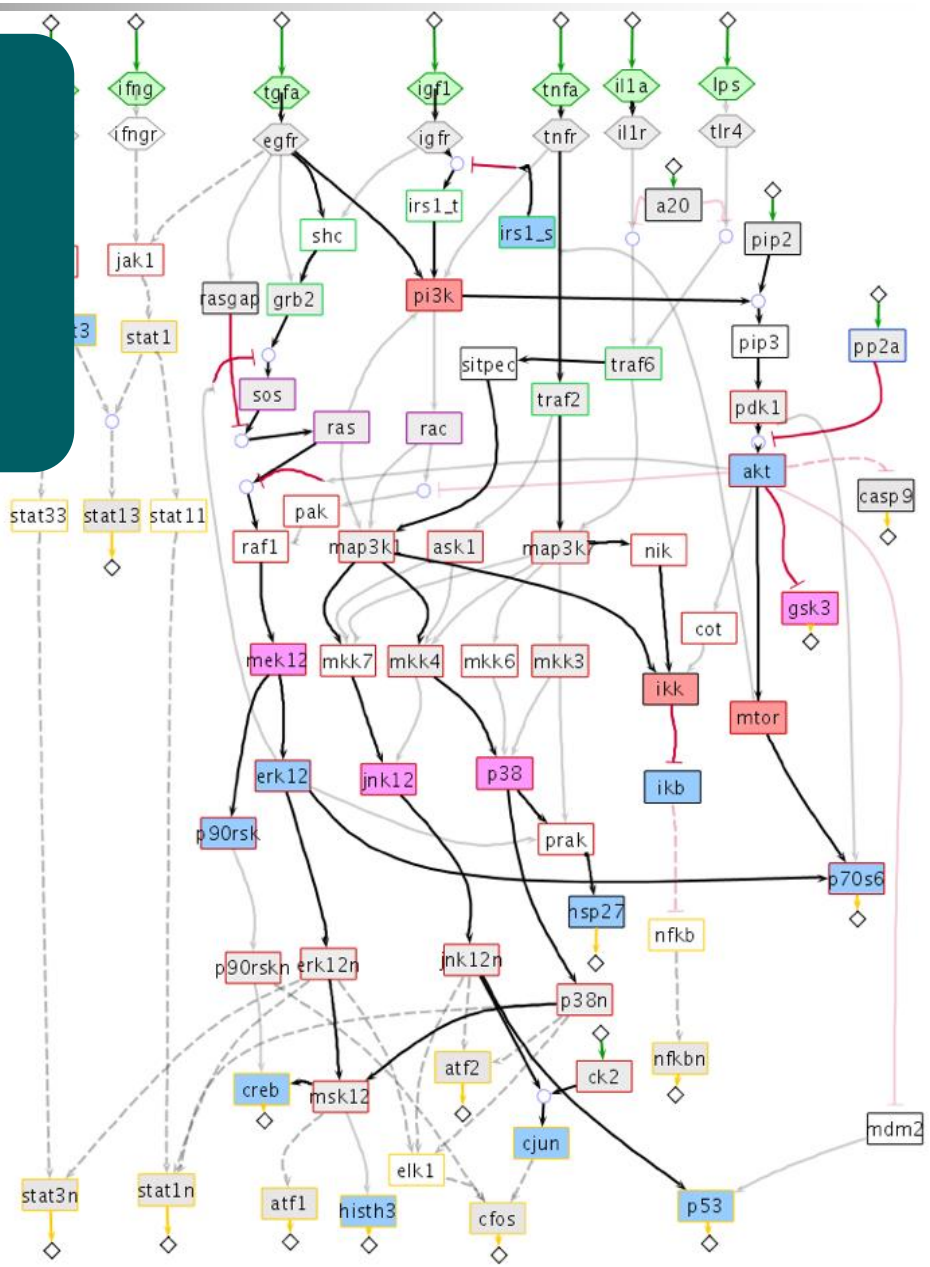
8.1%

Stimulus | Perturbation | Readout | Perturb&Read

Kept

Removed

No effect



Explanation of residual error

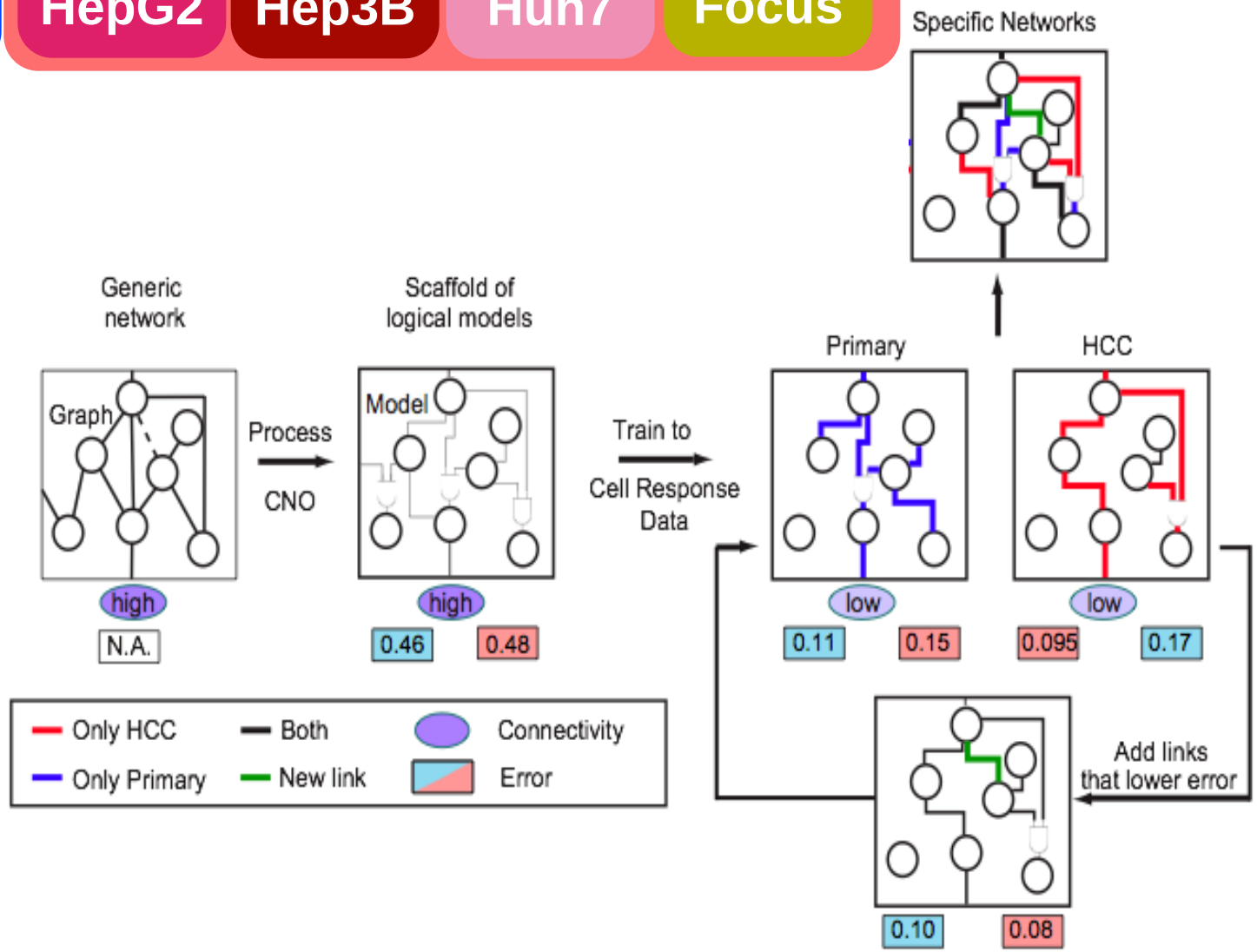
- Algorithm extended to identify **links missing** in map based on the fit of the data

➪ Identifies links that are **plausible** (supported in literature)

Saez-Rodriguez J, Alexopoulos LG, Epperlein J, Samaga R, Lauffenburger DA, Klamt S, Sorger PK, *Mol. Syst. Biol,* 5, 331, 2009.

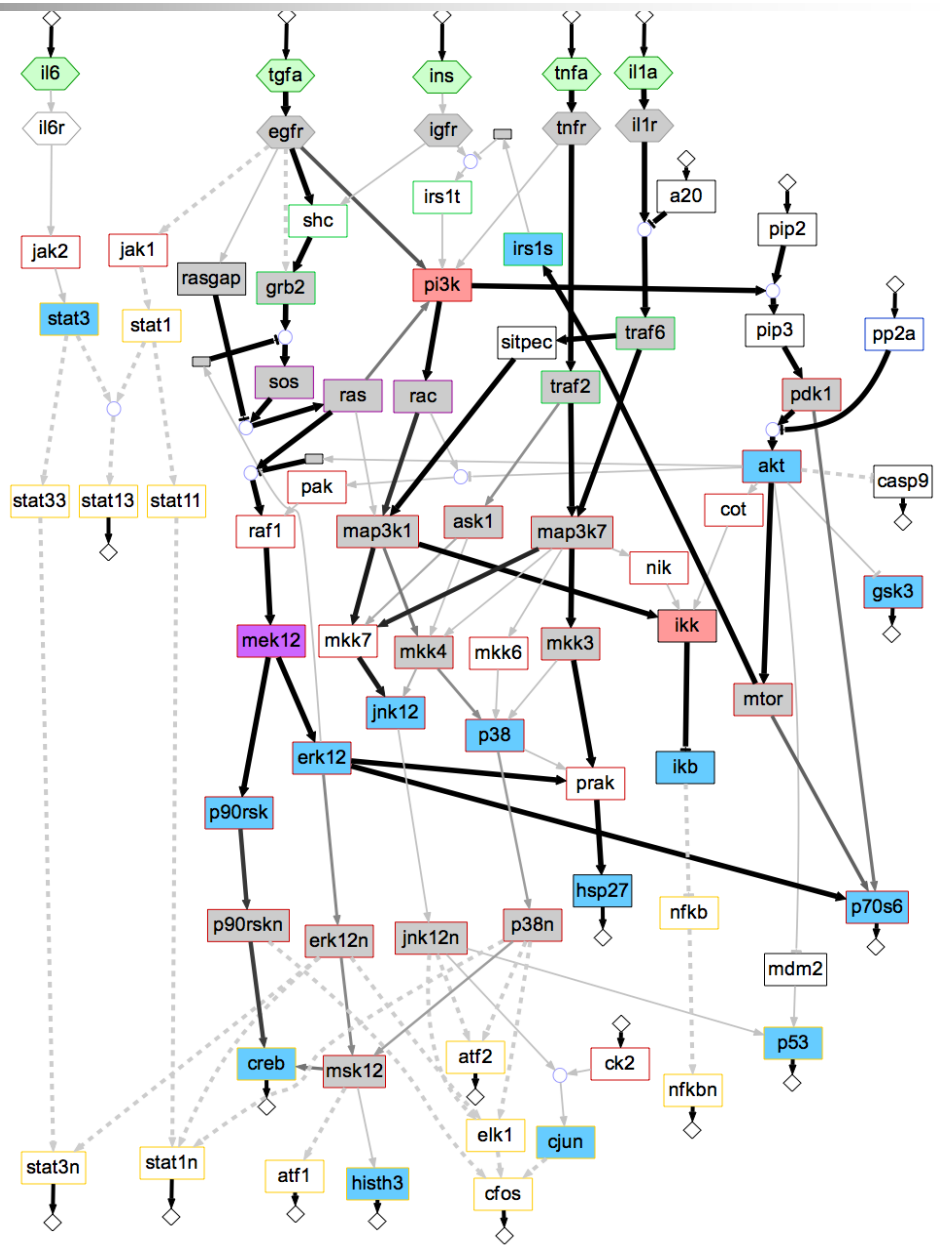
Comparison of primary hepatocytes to 4 HCC cell lines

Primary | HepG2 | Hep3B | Huh7 | Focus



Saez-Rodriguez J, Alexopoulos LG, Zheng M, Lauffenburger DA, Sorger PK, *submitted.*

Primary



Stimulus

Perturbation

Readout

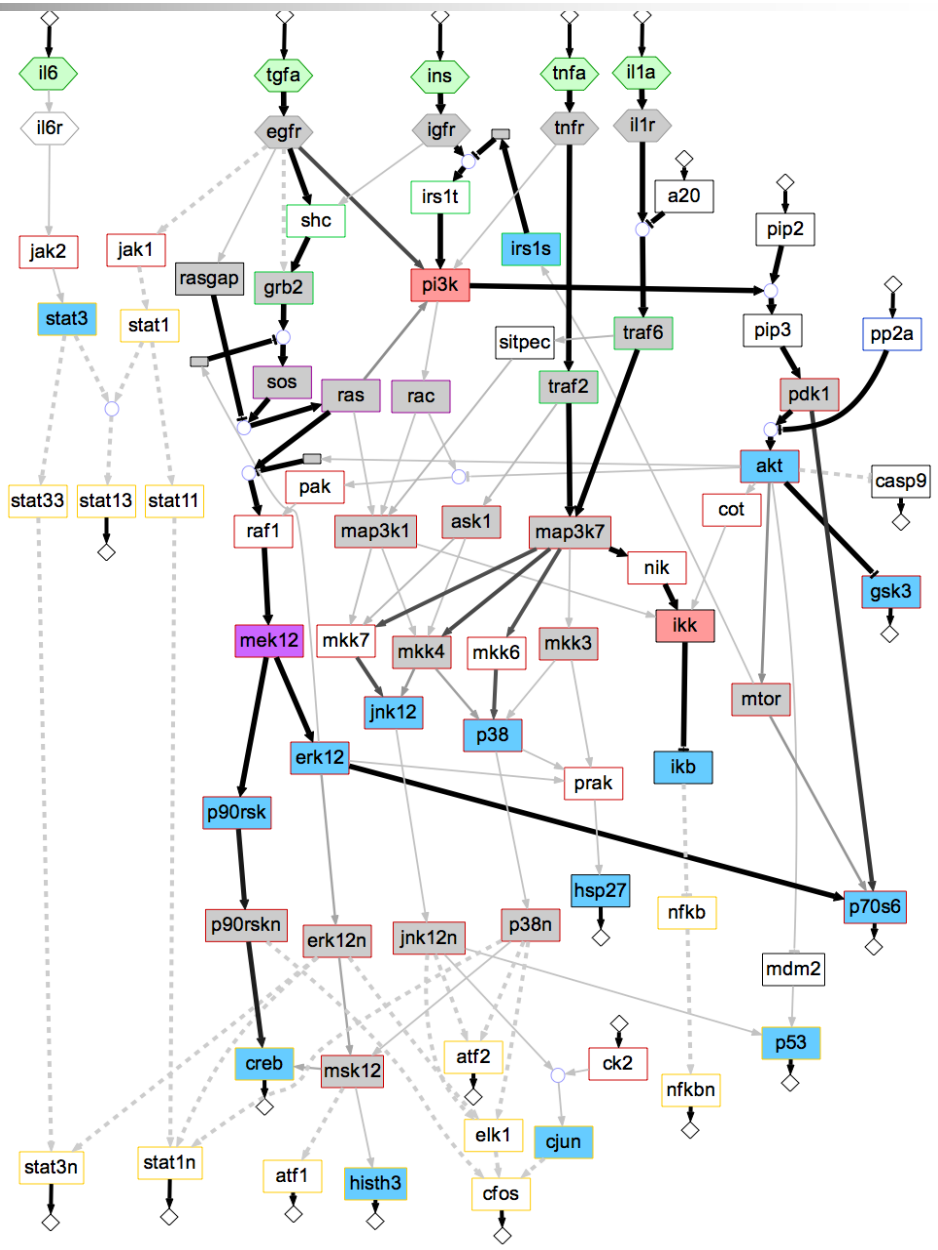
Perturb&Read

Kept

Removed

No effect

HepG2



Stimulus

Perturbation

Readout

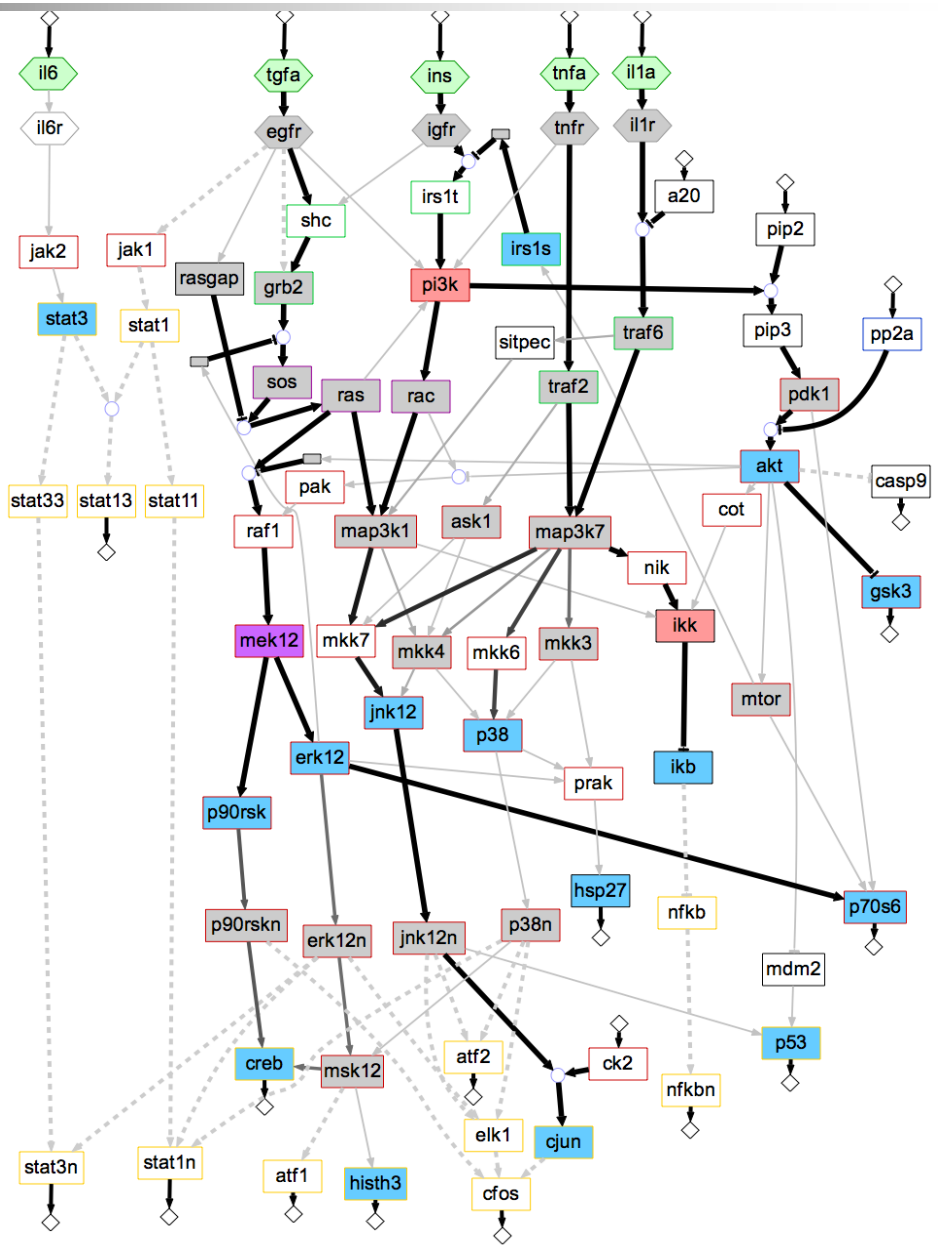
Perturb&Read

Kept

Removed

No effect

Hep3B



Stimulus

Perturbation

Readout

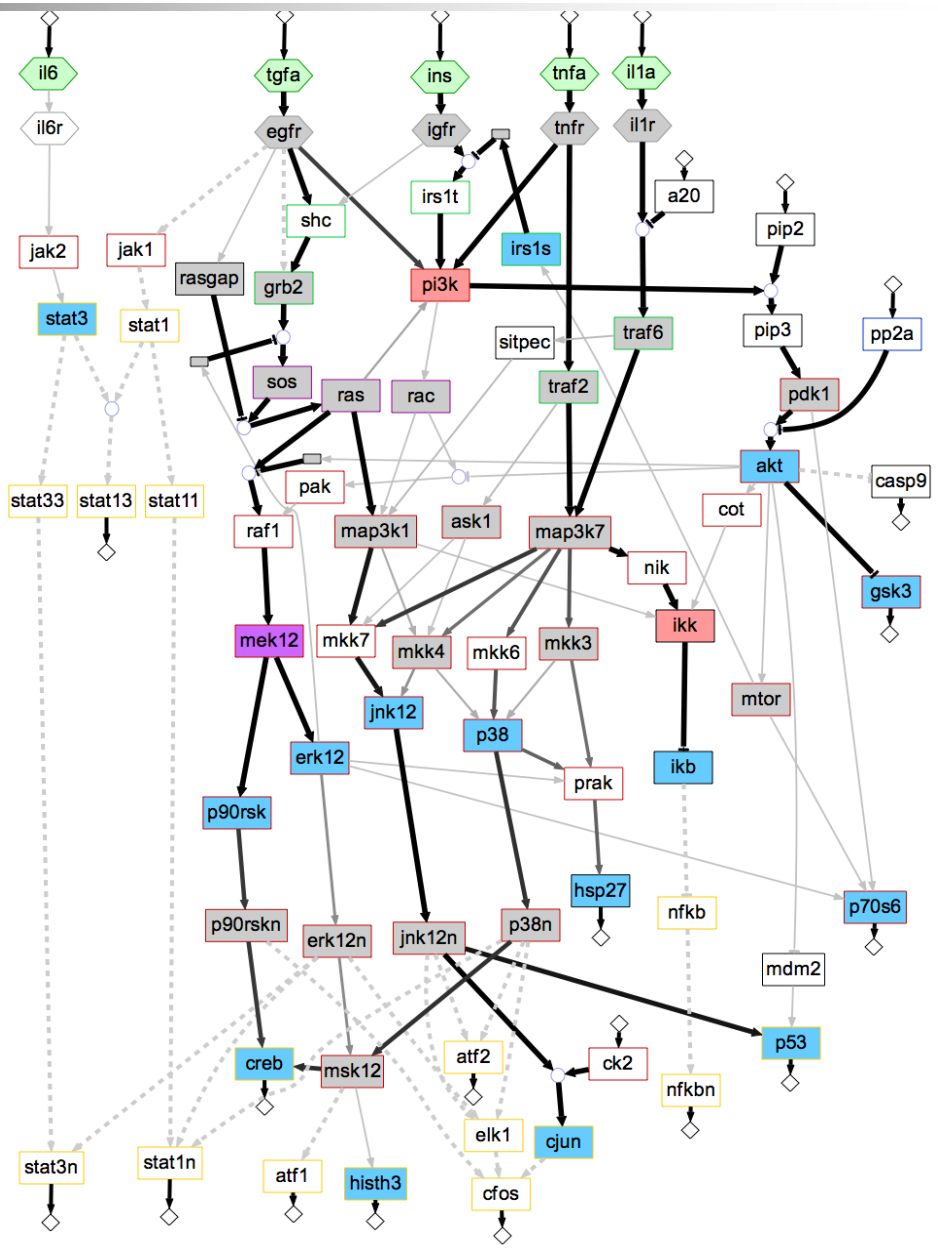
Perturb&Read

Kept

Removed

No effect

Huh7



Stimulus

Perturbation

Readout

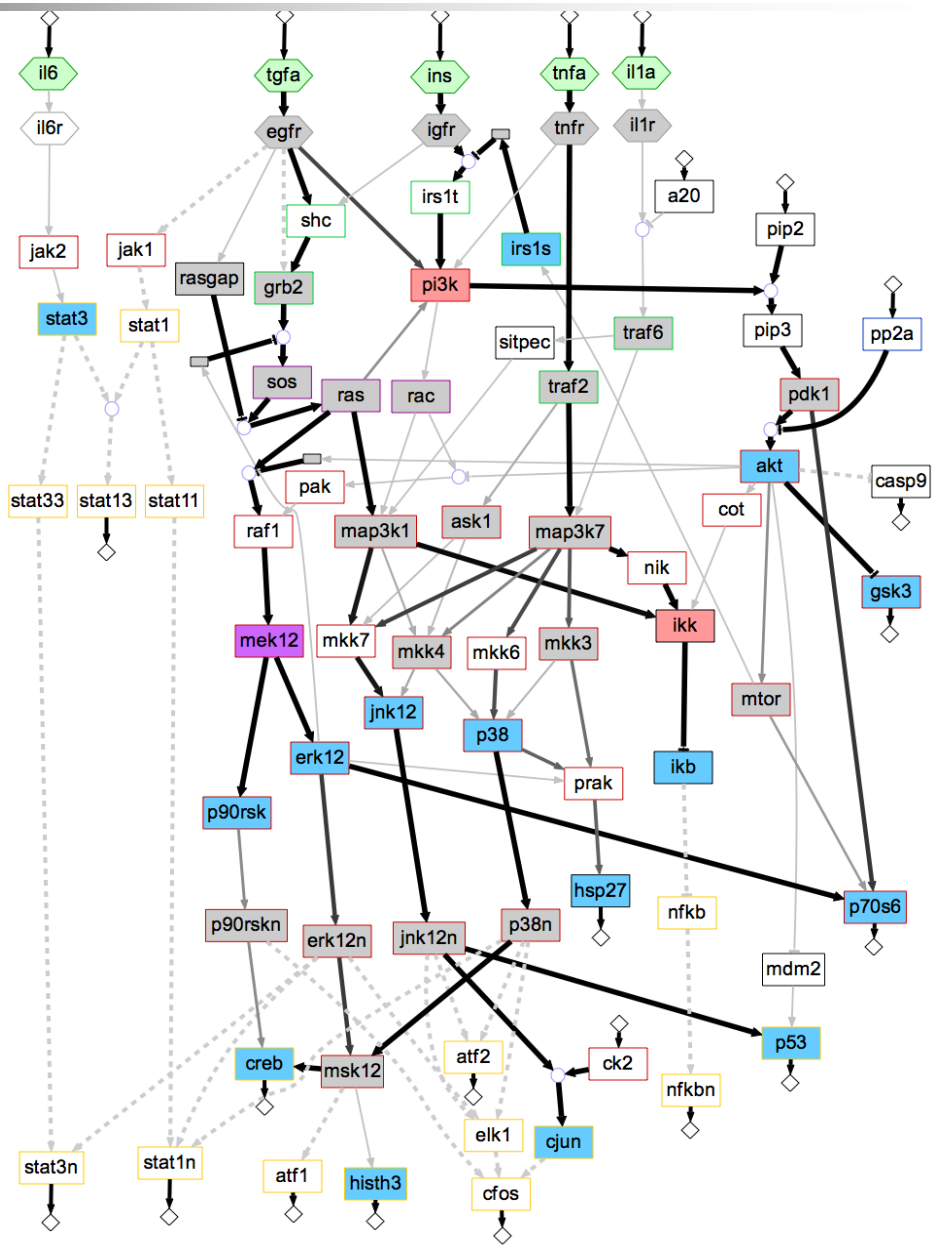
Perturb&Read

Kept

Removed

No effect

Focus



Stimulus

Perturbation

Readout

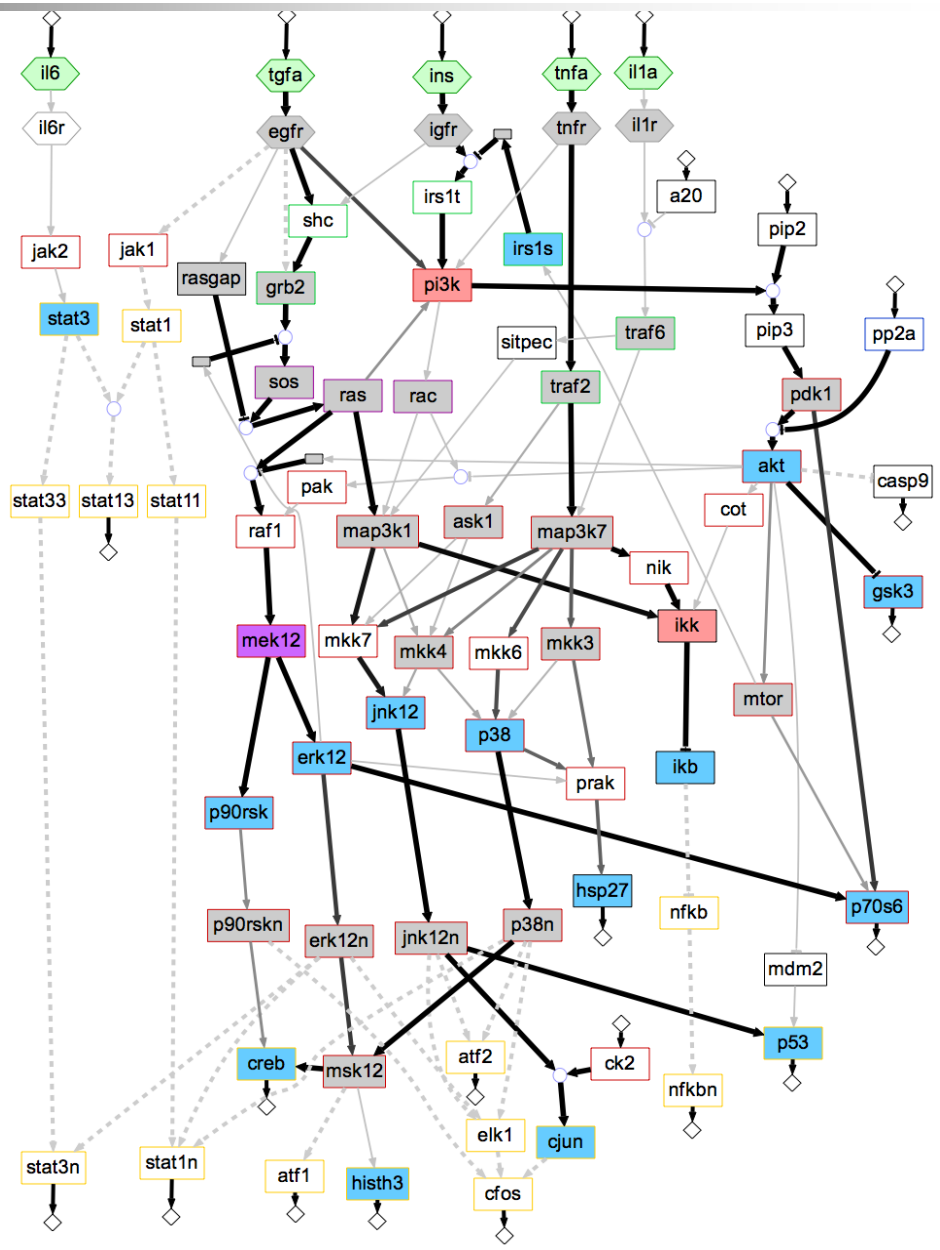
Perturb&Read

Kept

Removed

No effect

How different are HCC cell lines with respect to each other and to primary hepatocytes?



Stimulus

Perturbation

Readout

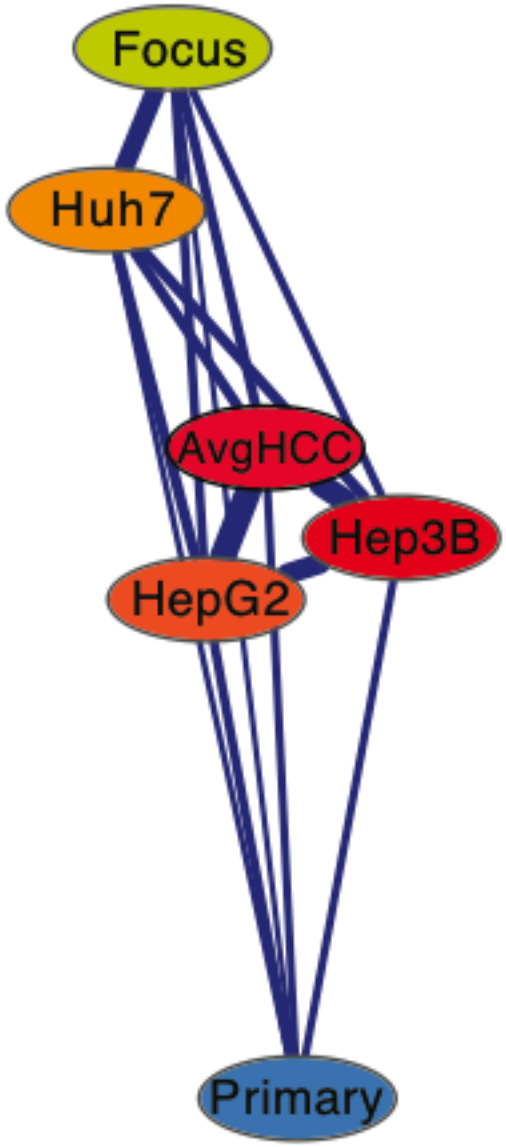
Perturb&Read

Kept

Removed

No effect

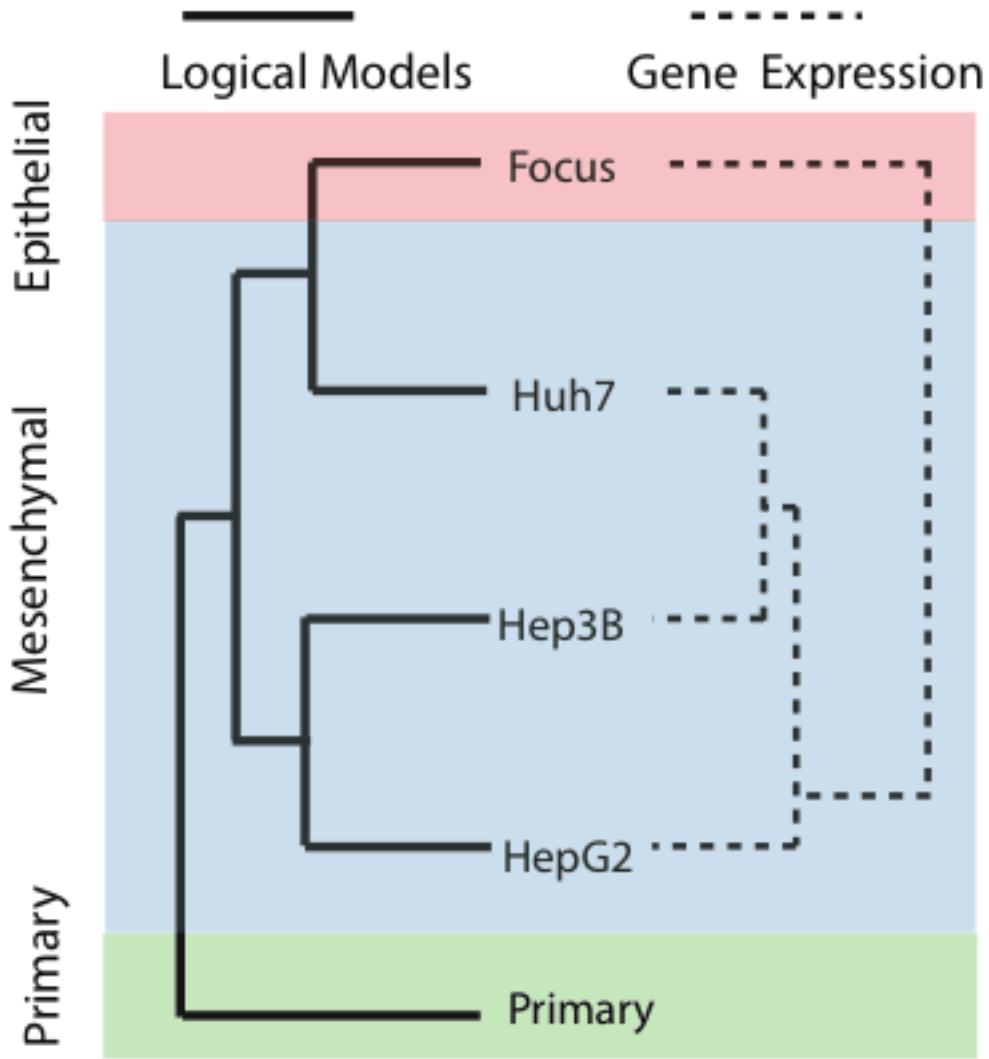
Models of all cell lines are more different to primary hepatocytes than to each other



- Define distance between cells based on differences in topology of models
- Including data set AvgHCC: mixed data from HCC cell lines

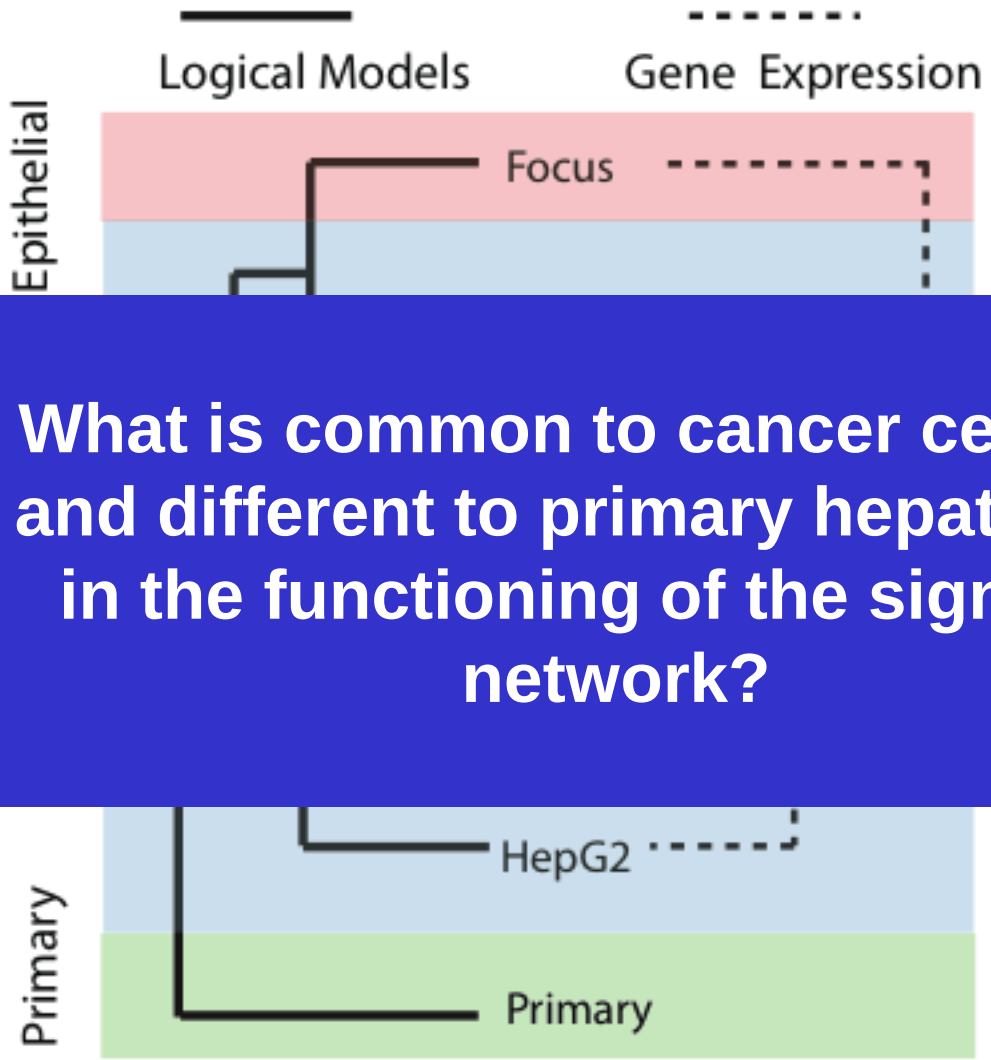
Saez-Rodriguez J, Alexopoulos LG, Zheng M, Lauffenburger DA, Sorger PK, *submitted.*

Network differences roughly correlate with gene expression & phenotypical differences



Lee and Horgeirsson, Hepatology, 35:1134, 2002.
Fuchs et al, Cancer Res., 68:2391, 2008.

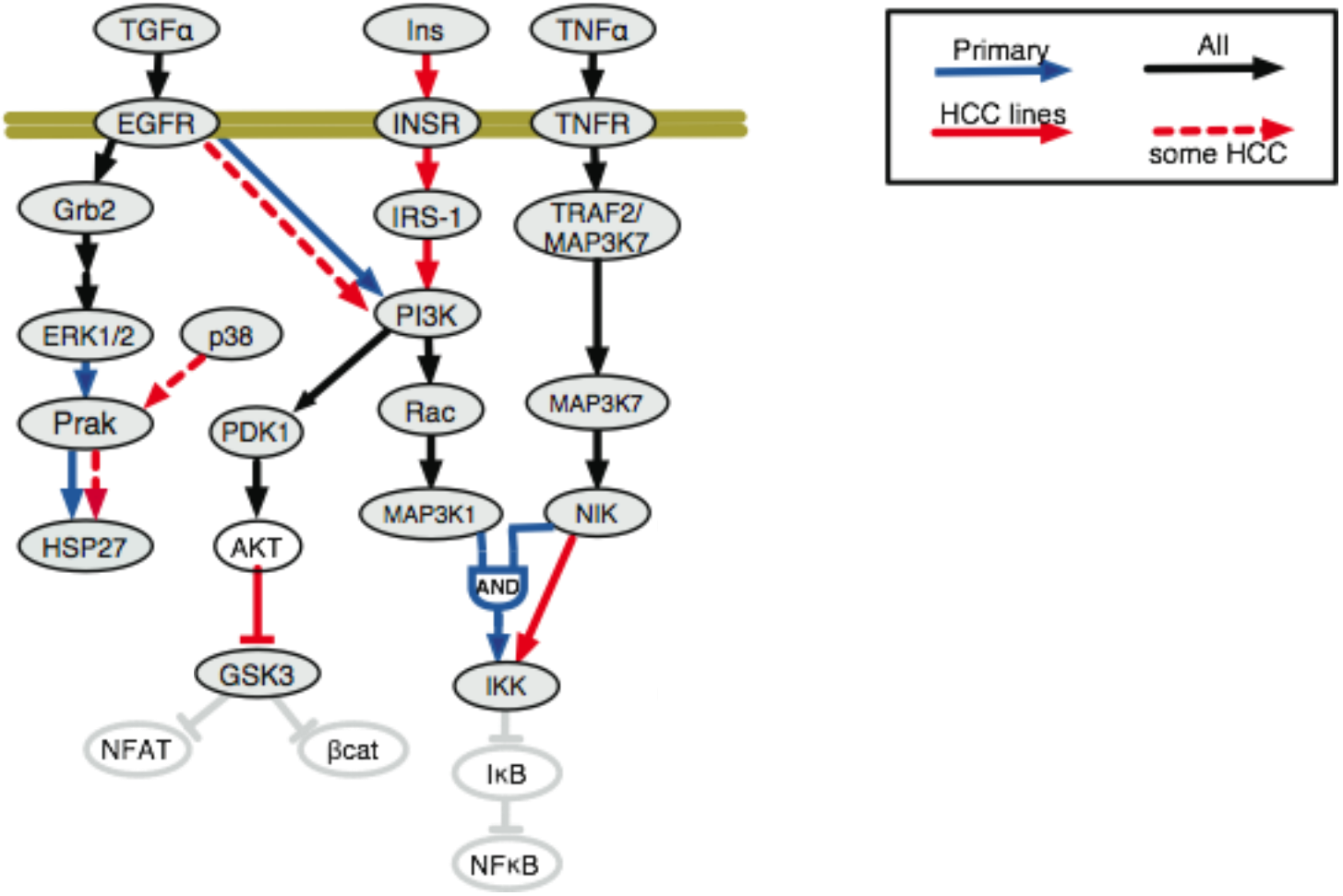
Network differences roughly correlate with gene expression & phenotypical differences



What is common to cancer cell lines and different to primary hepatocytes in the functioning of the signaling network?

Lee and Horgeirsson, Hepatology, 35:1134, 2002.
Fuchs et al, Cancer Res., 68:2391, 2008.

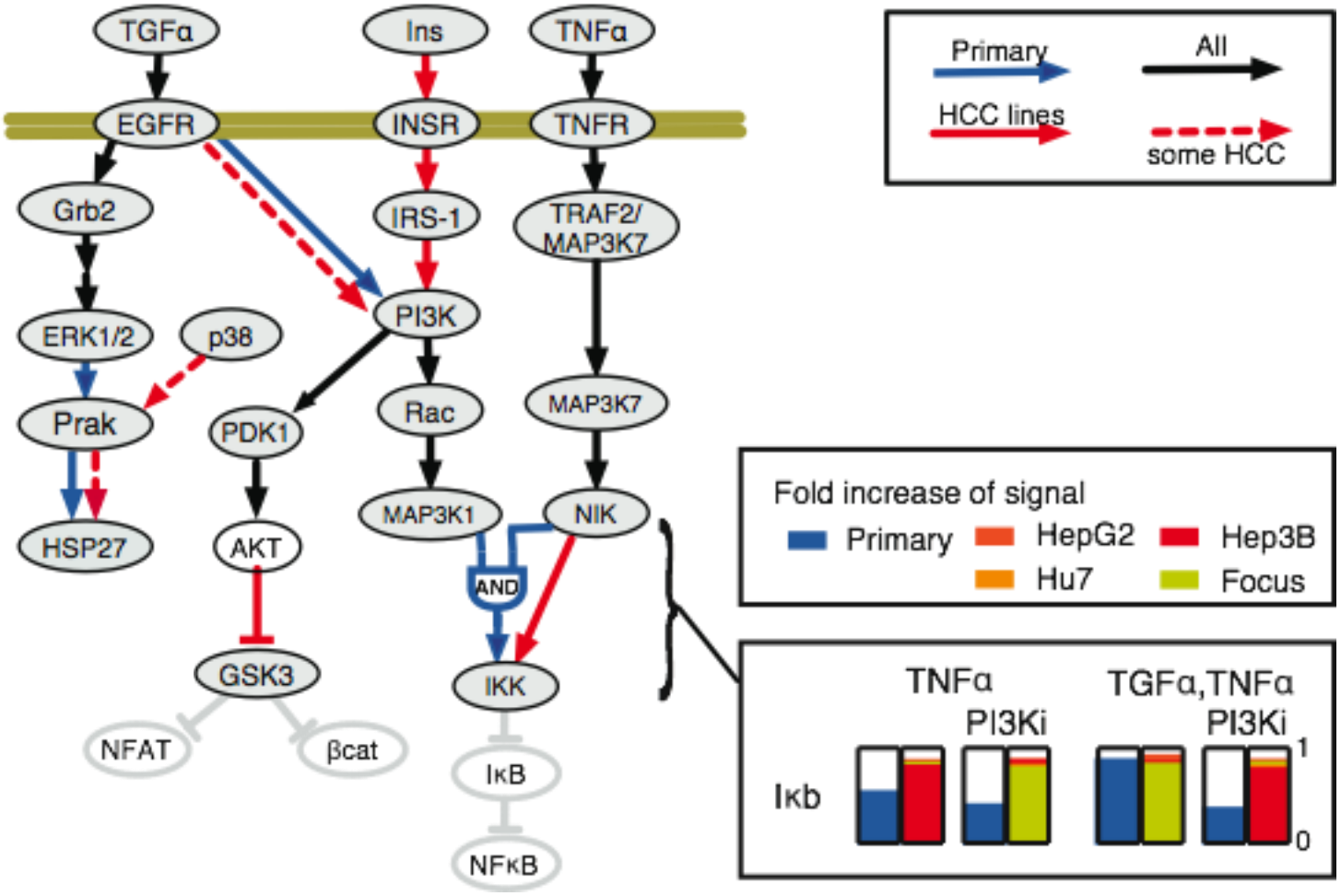
Specific pathways determine major differences



- Only active in HCC cell lines:
  Insulin→.. →AKT→GSK3
- HSP27 phosphorylation:
  ERK mediated in primary

- Difference in NFkB activation:
  TNF dependent only in HCC
  TNF+TGFα in primary

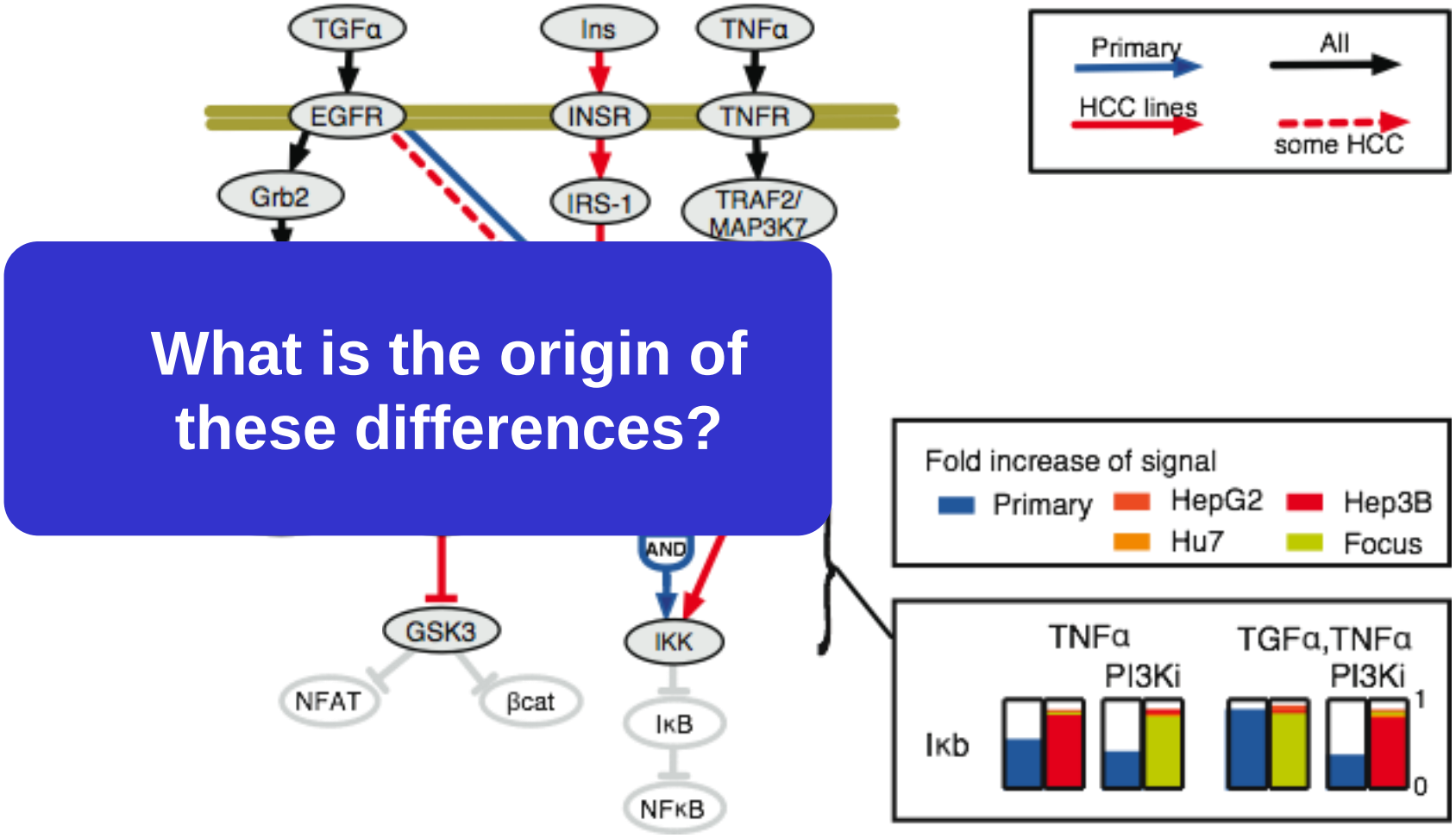
Specific pathways determine major differences



- Only active in HCC cell lines:
  Insulin→.. →AKT→GSK3
- HSP27 phosphorylation:
  ERK mediated in primary

- Difference in NFkB activation:
  TNF dependent only in HCC
  TNF+TGFα in primary

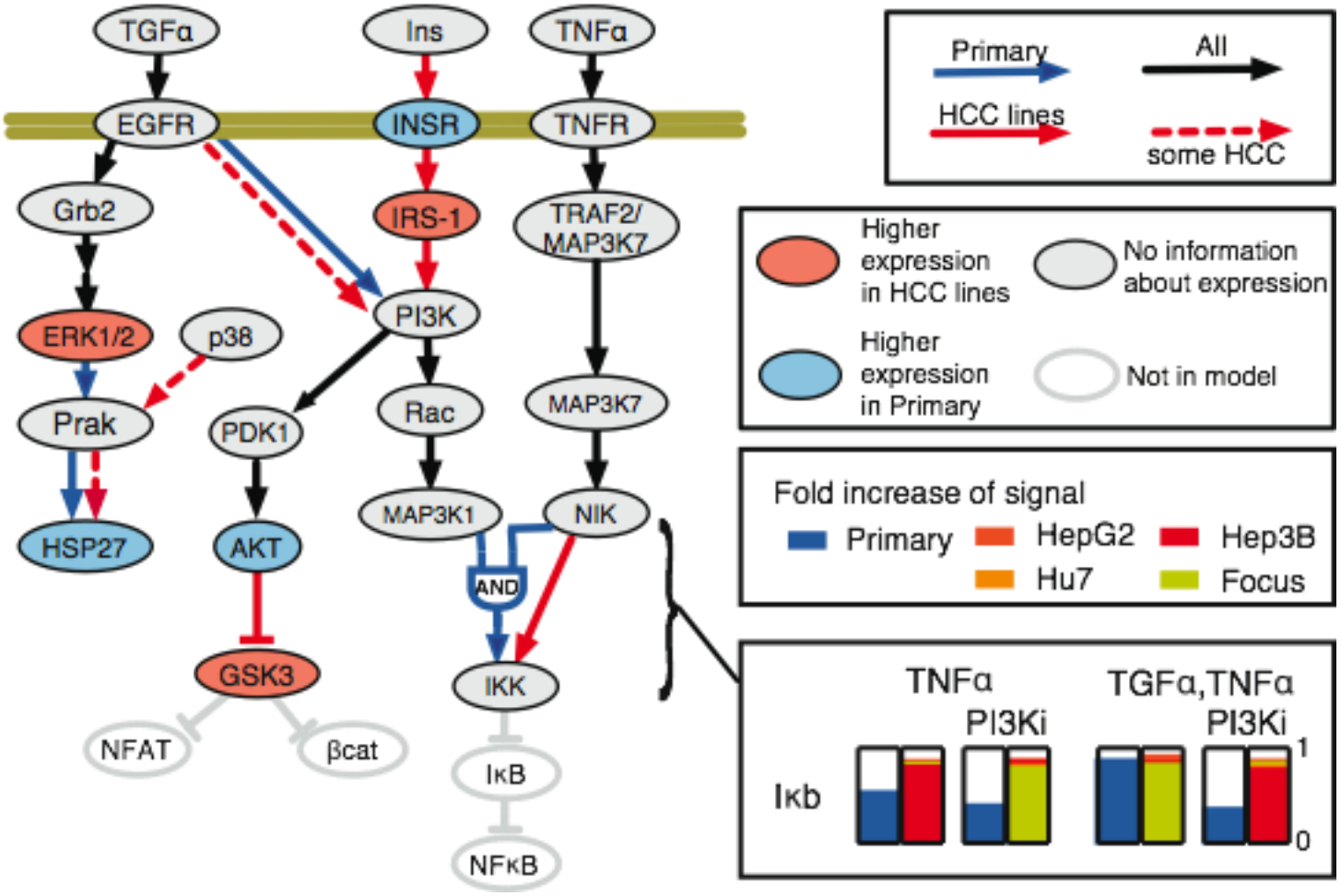
Specific pathways determine major differences



- Only active in HCC cell lines:
  Insulin→.. →AKT→GSK3
- HSP27 phosphorylation:
  ERK mediated in primary

- Difference in NFkB activation:
  TNF dependent only in HCC
  TNF+TGFα in primary

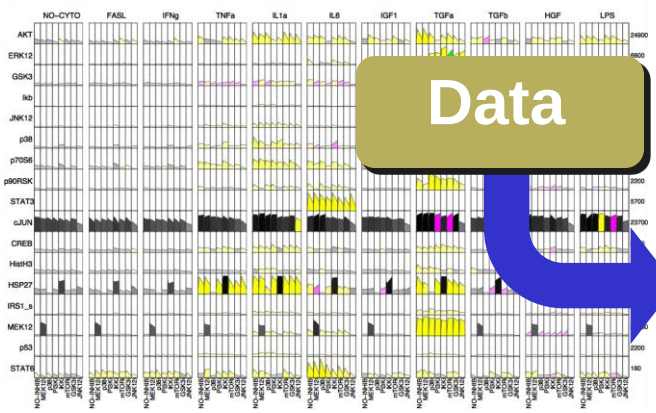
Specific pathways determine major differences



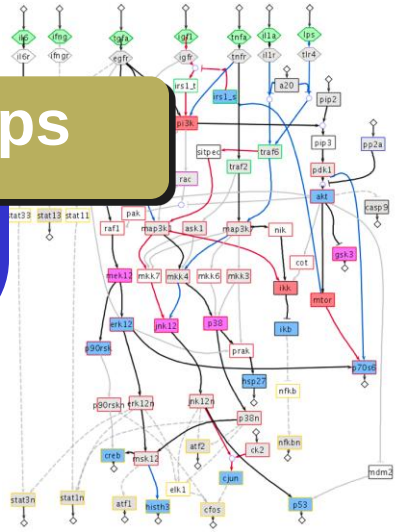
- Only active in HCC cell lines:
  Insulin→.. →AKT→GSK3
- HSP27 phosphorylation:
  ERK mediated in primary

- Difference in NFkB activation:
  TNF dependent only in HCC
  TNF+TGFα in primary

Summary

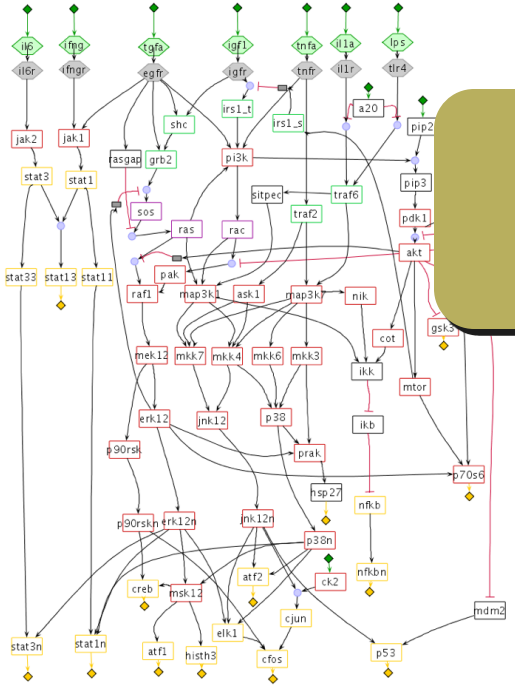


Data



Maps

CellNet Optimizer

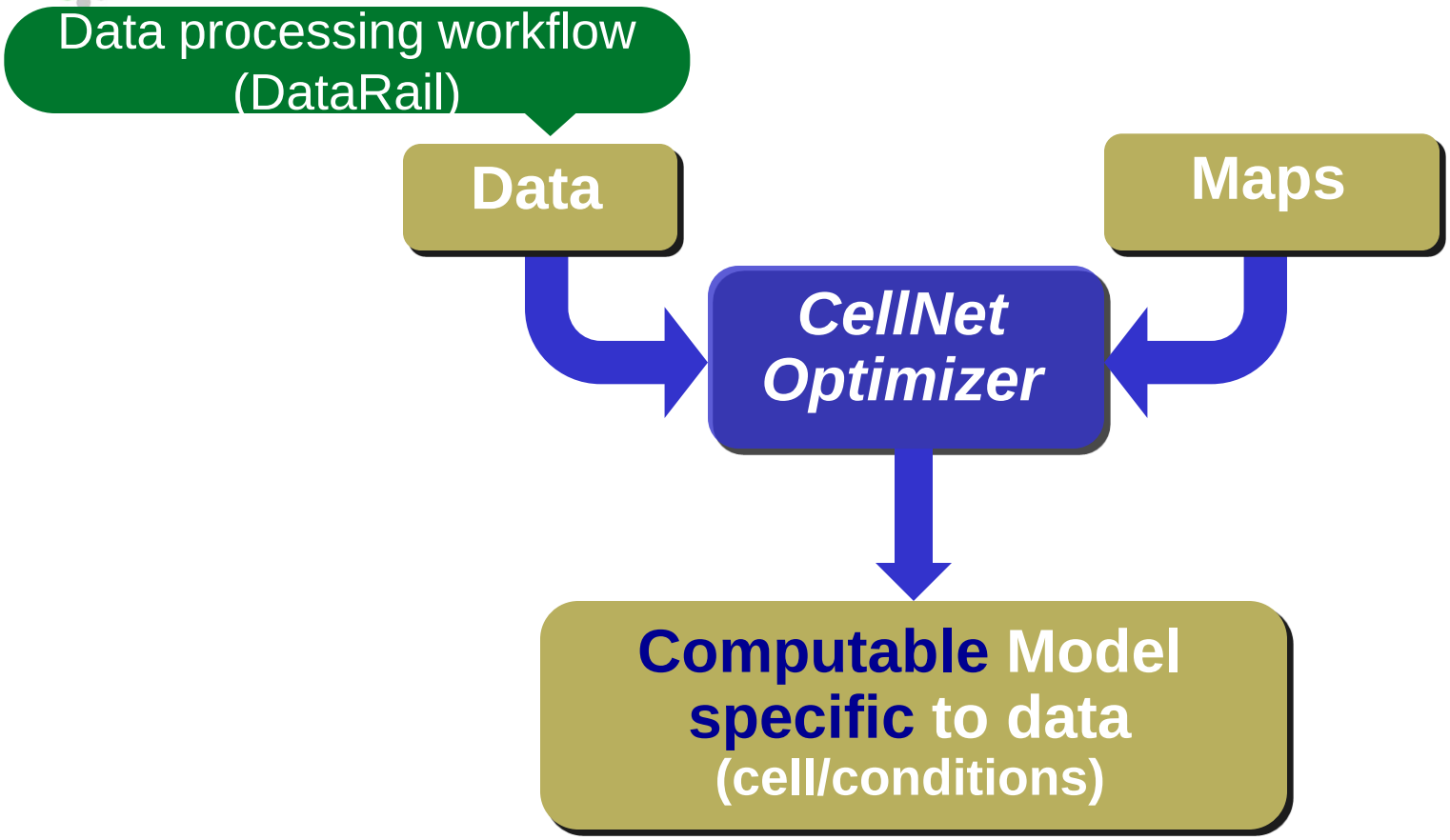


Computable Model specific to data (cell/conditions)

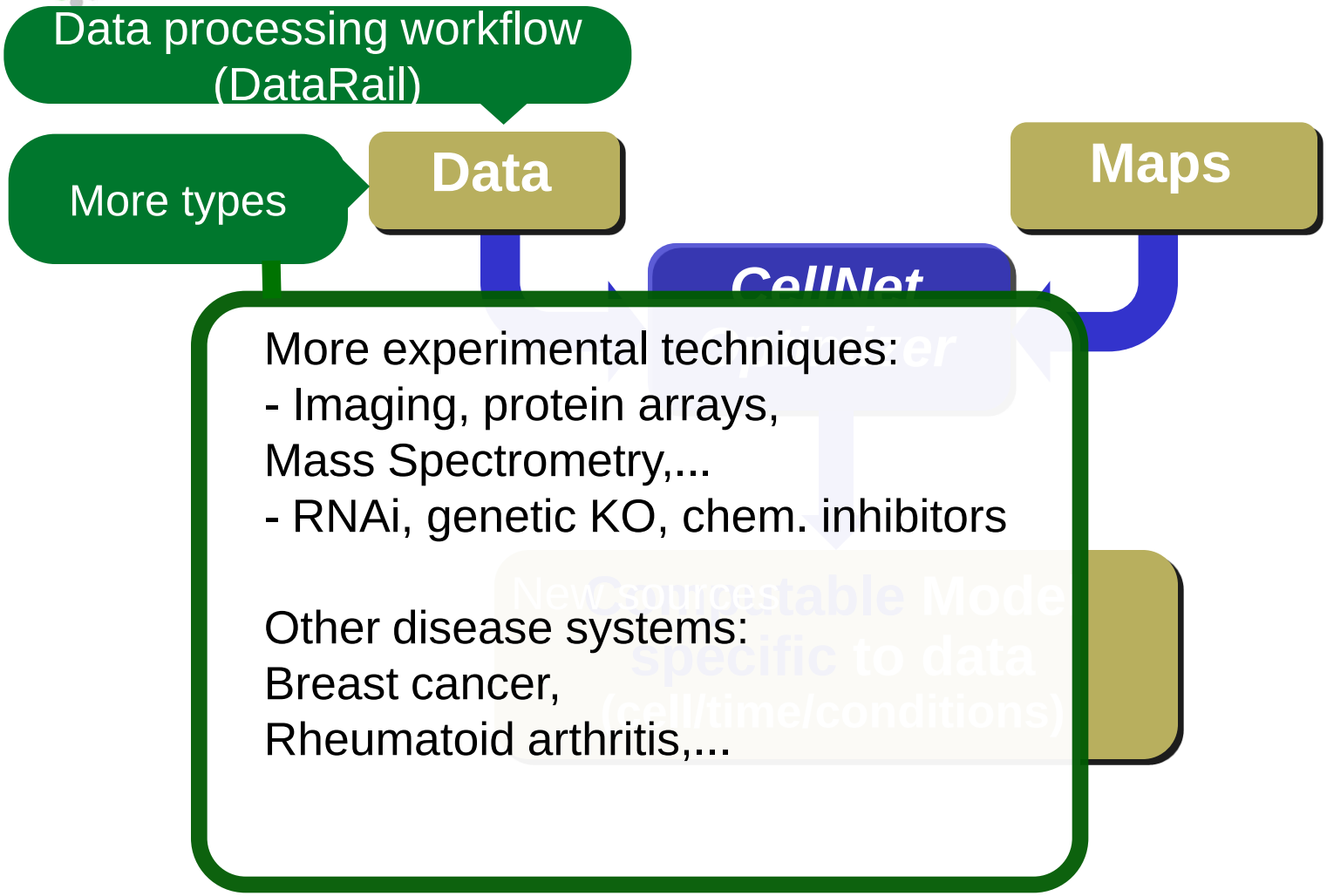
Summary

- Pathway **maps** are **not specific**
- **Models** trained to data are much **sparser** and **predictive**
- It is possible to
  - Construct models **specific** for **cell types**
  - **Cluster** cell types based on pathways models
  - Pinpoint specific **differences** between **normal** and **diseased** cells

Future and current directions



Future and current directions



Future and current directions

Data processing workflow (DataRail)

More types

Data

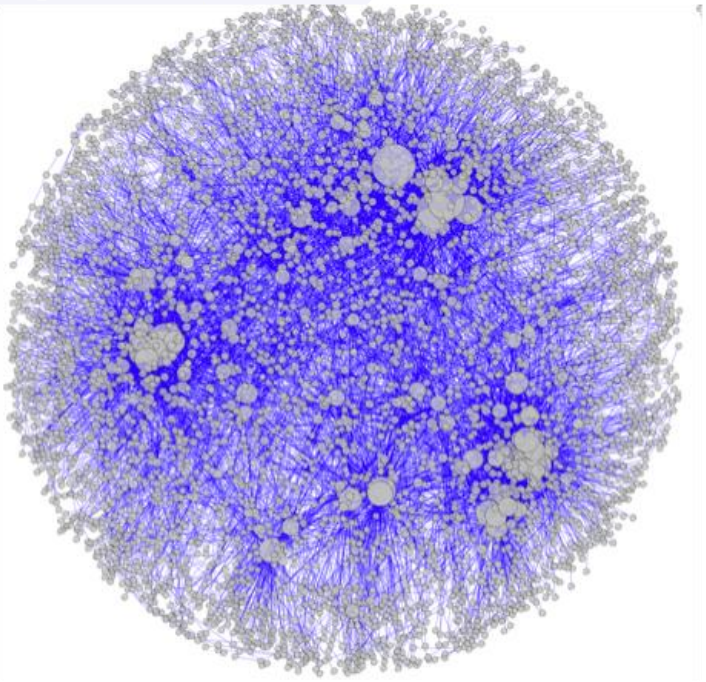
Maps

New sources

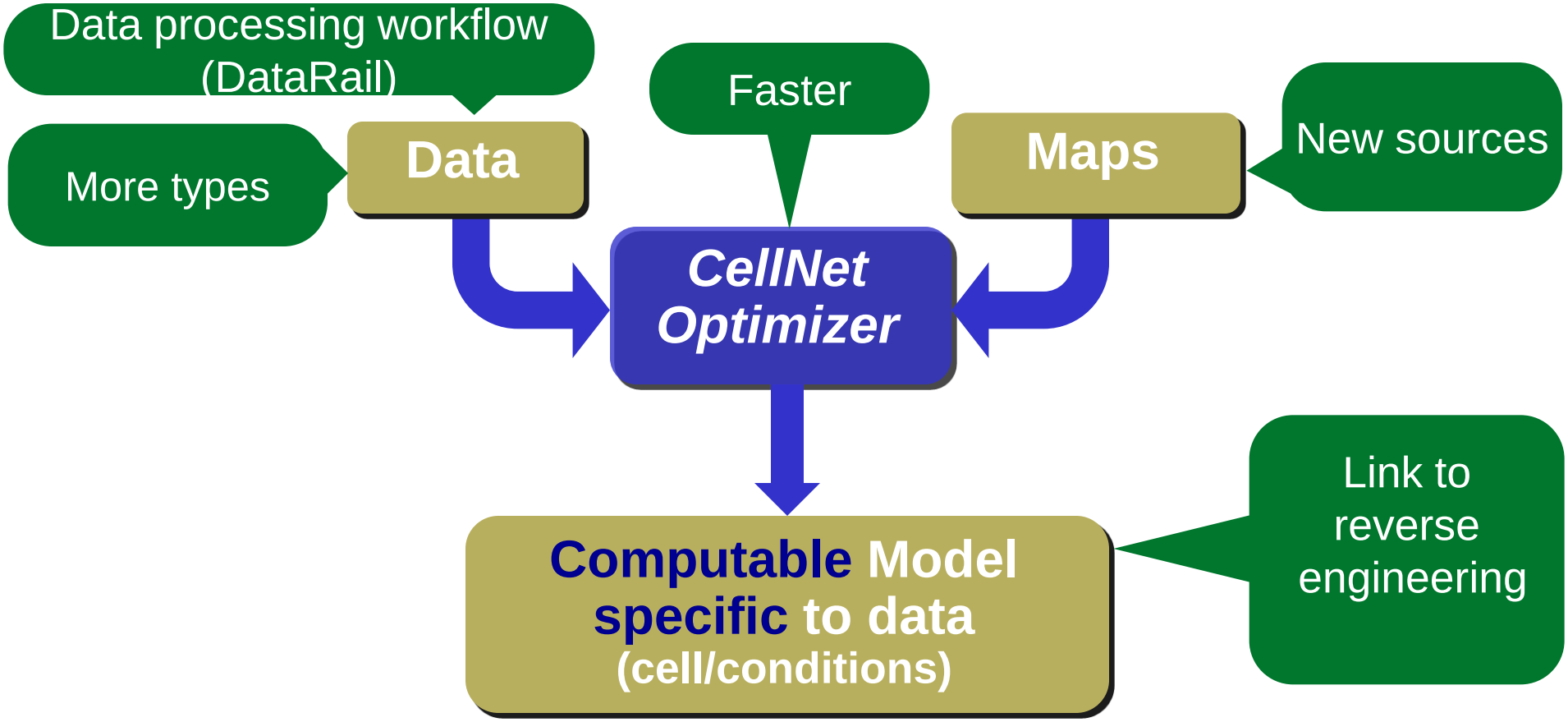
CellNet

D. Kirouac:
- Databases of curated pathways
- Protein Interaction Networks

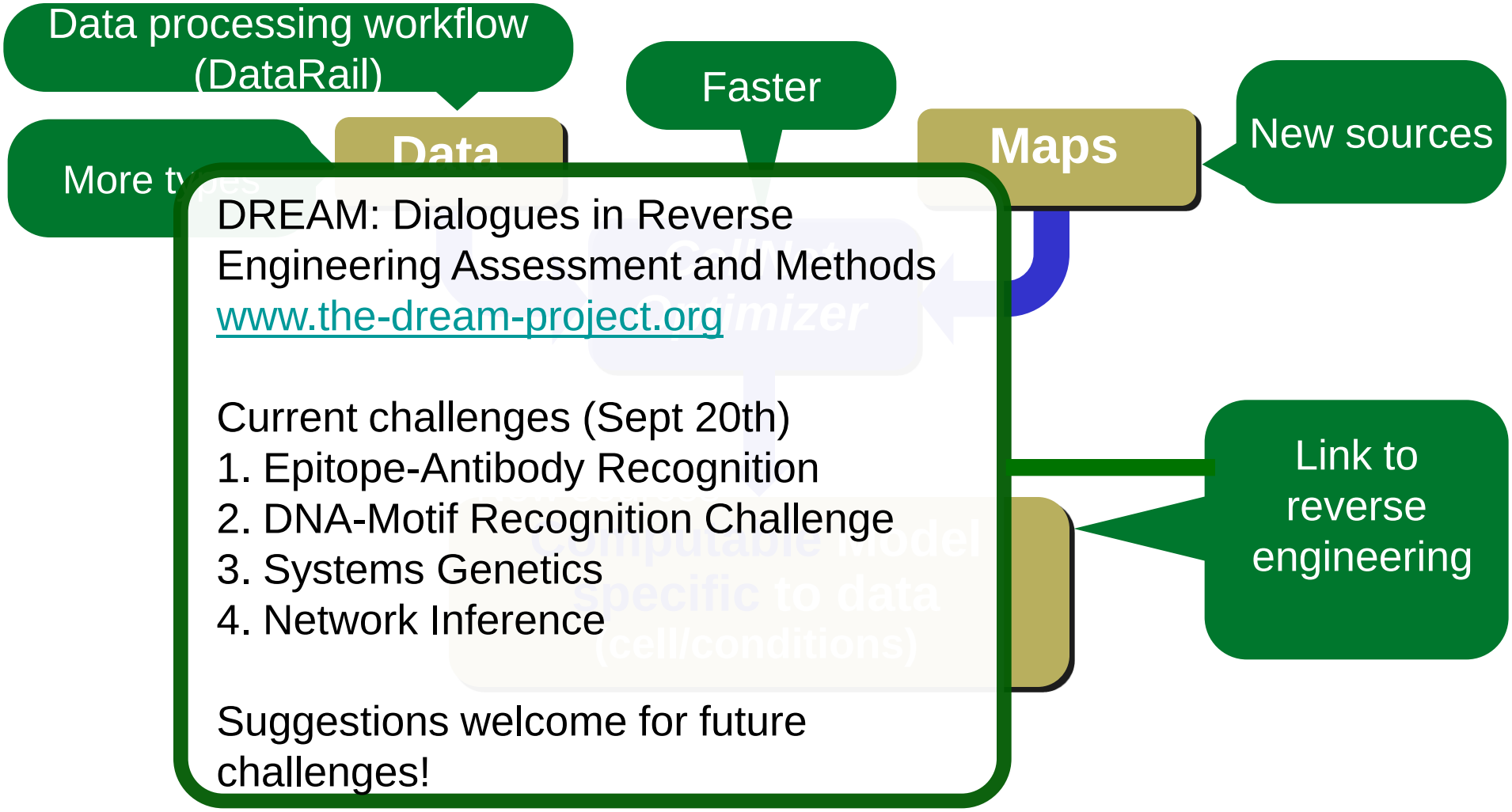
- Literature mining



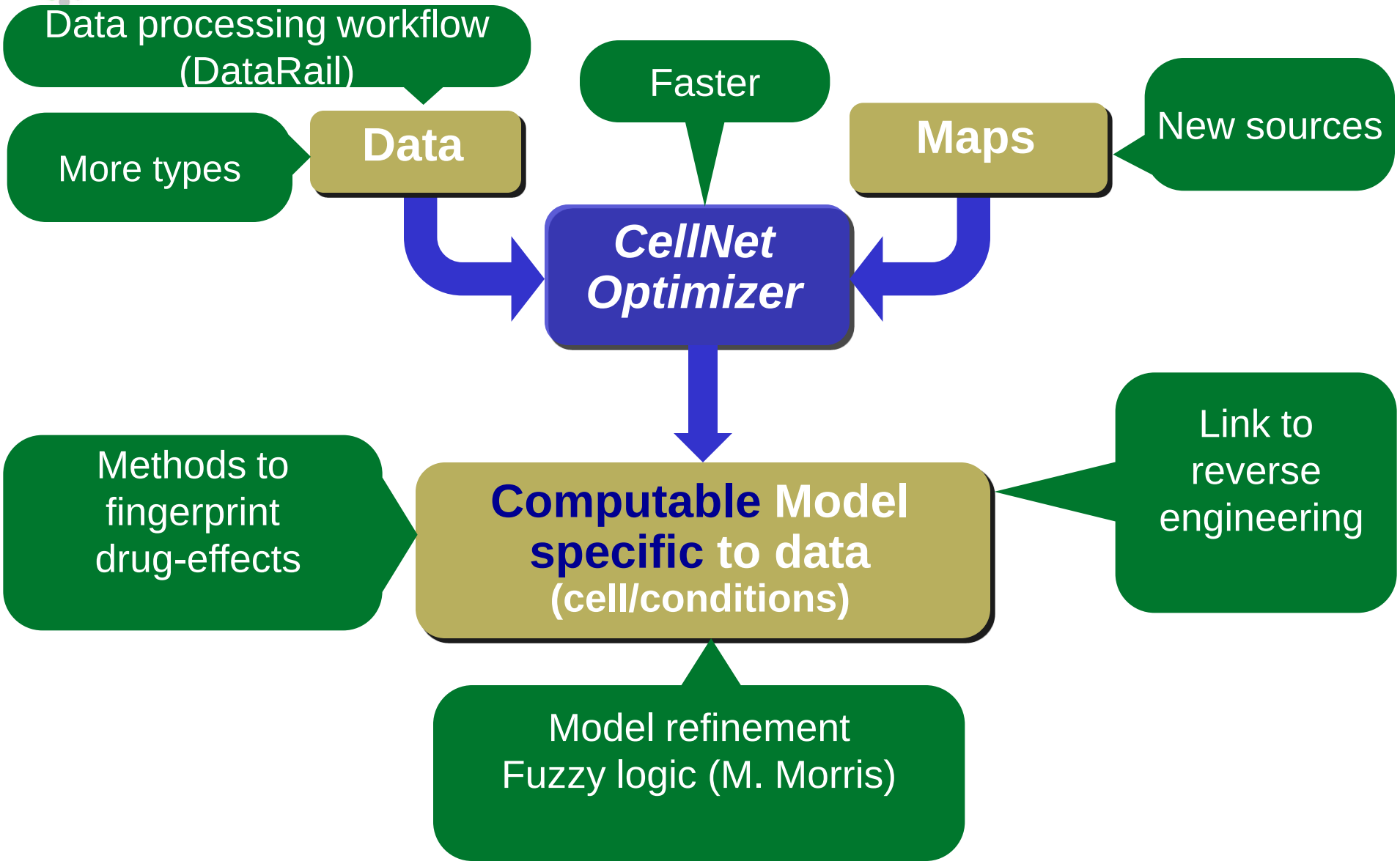
Future and current directions



Future and current directions

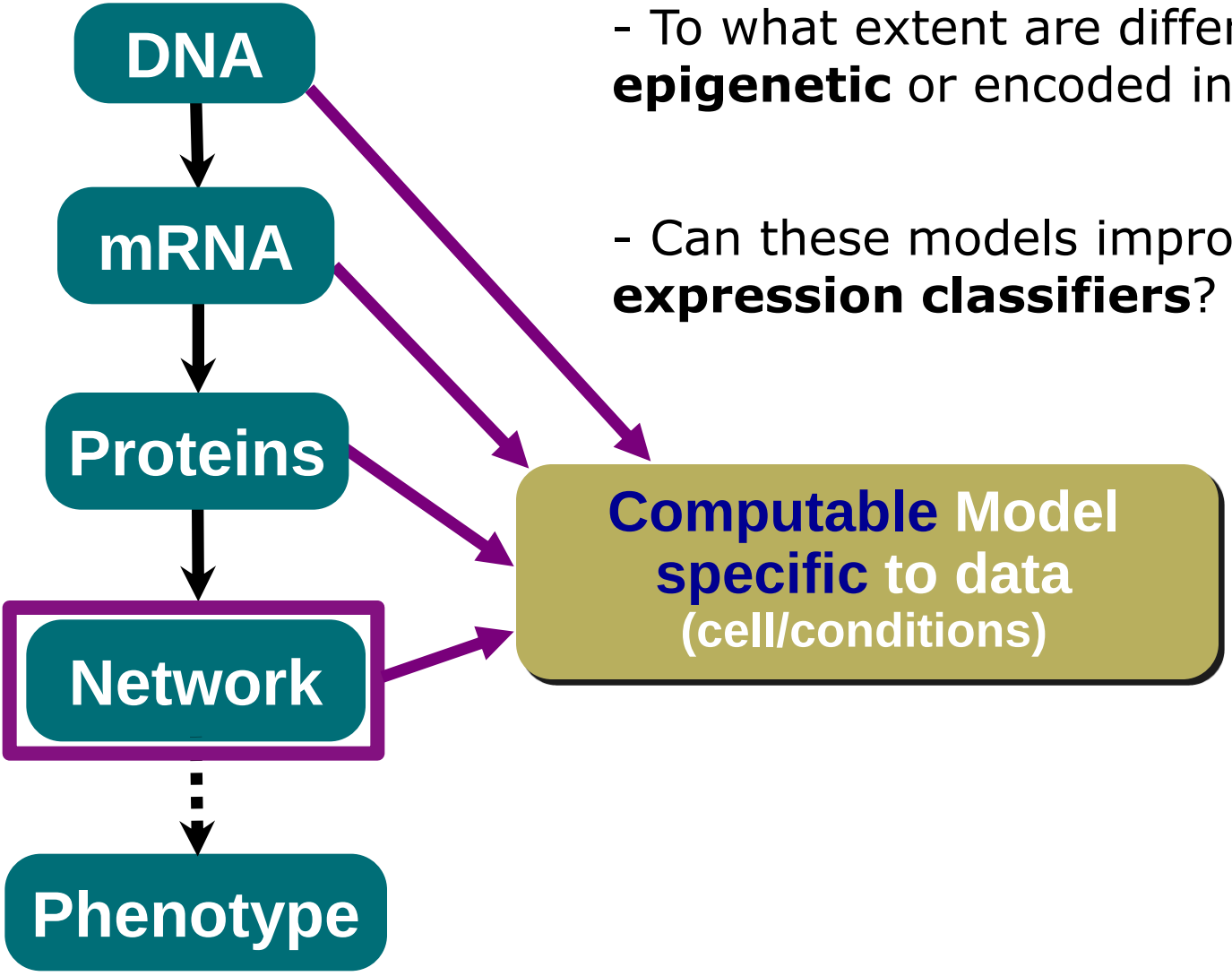


Future and current directions



Future and current directions

Linking post-translational events to genetic alterations



- To what extent are differences **epigenetic** or encoded in the genome?

- Can these models improve gene **expression classifiers**?

Acknowledgments

- Jeremy Muhlich
- Mario Niepel
- Bree Aldridge
- Sean Milton
- Jonathan Epperlein
- Laura Kleiman
- William Chen
- Peter Sorger

- Melody Morris
- Arthur Goldsipe
- Dan Kirouac
- Nickel Dittrich
- Joel Wagner
- Douglas Lauffenburger
- Alexander Mitsos (RWTH Aachen)

- Fabian Theis (Helmholtz)

- Leonidas Alexopoulos (NTUA)
- Ioannis Melas

- Bea Penalver (NW Univ)

- Eduardo Sontag (Rutgers)

- Regina Samaga
- Steffen Klamt
- Sebastian Mirschel
- Ernst Dieter Gilles

Max-Planck Institute MD

- David de Graaf (GenStruct)
- Raul Rodriguez-Esteban (Bo. In.)
- Chris Espelin (Pfizer)

Funding by Pfizer & NIH (GM68762, CA112967)