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

Julio Saez-Rodriguez, Ph. D.



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

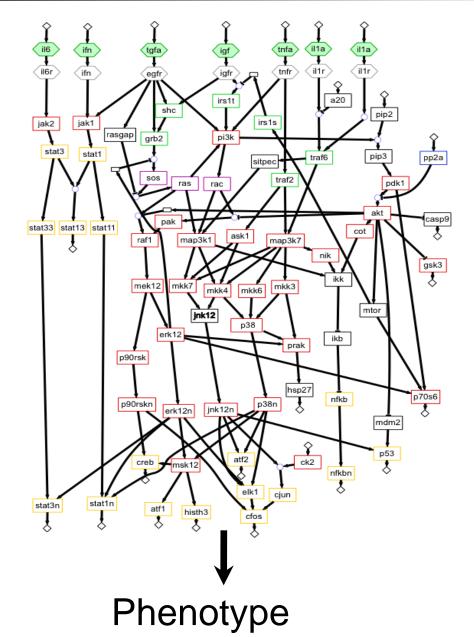
&

Biological Engineering Department, M.I.T.





### How do cells process extracellular signals?

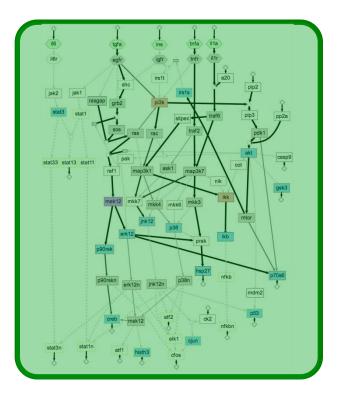


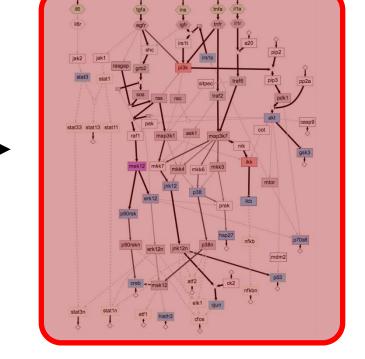


### How is signal processing altered in disease?

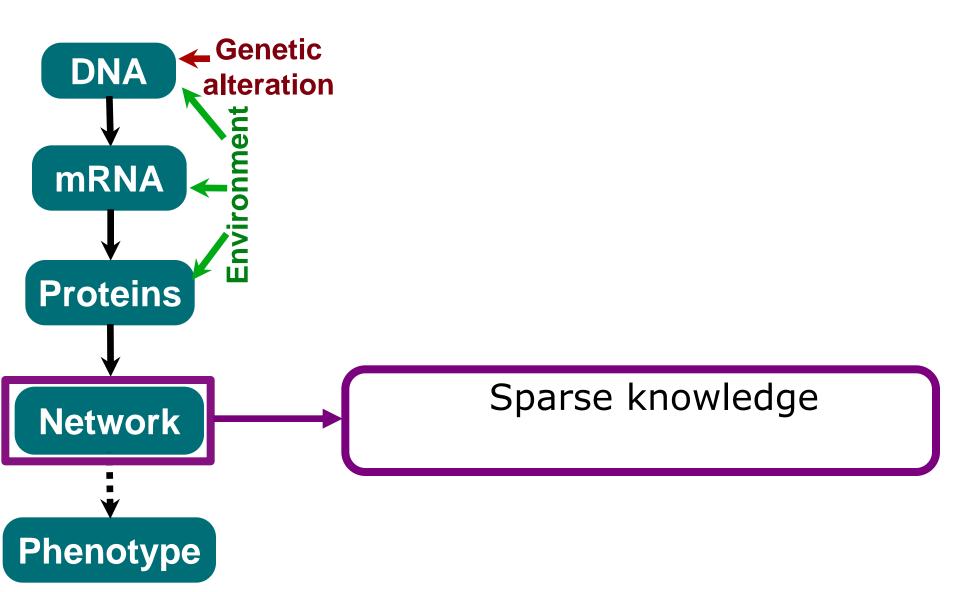
### Normal

### Diseased











DNA

**mRNA** 

**Proteins** 

Network

Phenotype

**Genetic** 

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

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

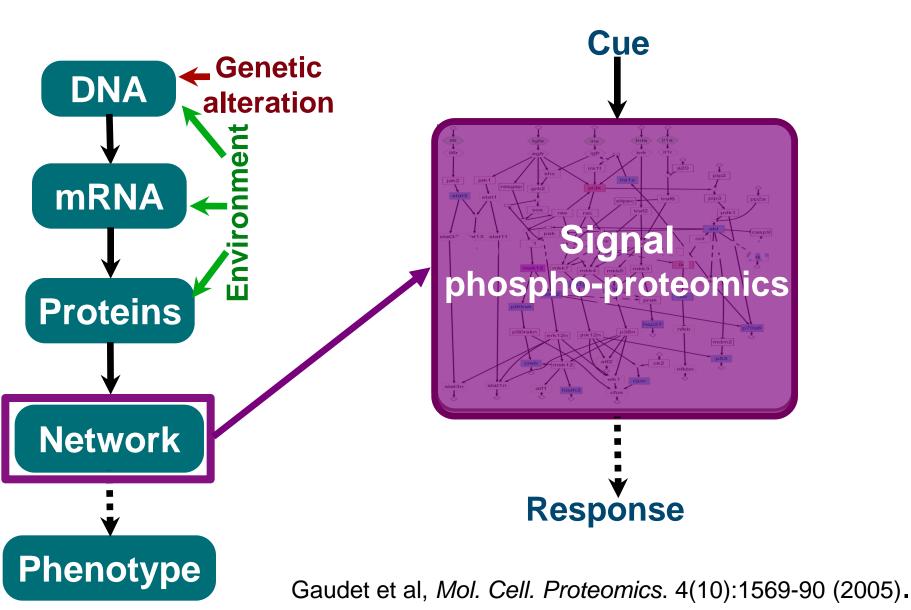
> - Chromosomal amplifications/ deletions, mutations, methylation alterations (Llovet and Bruix, 2008) - Heterogenous gene expression even

within adjacent tumor nodes (Lee & Thorgeirsson et al. 2005)

Sparse knowledge specific in HCC

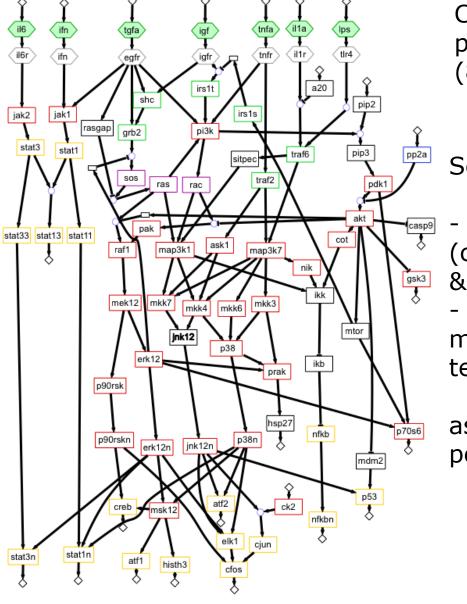


## Cue-Signal-Response approach to understand operation of signal transduction pathways





#### 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)

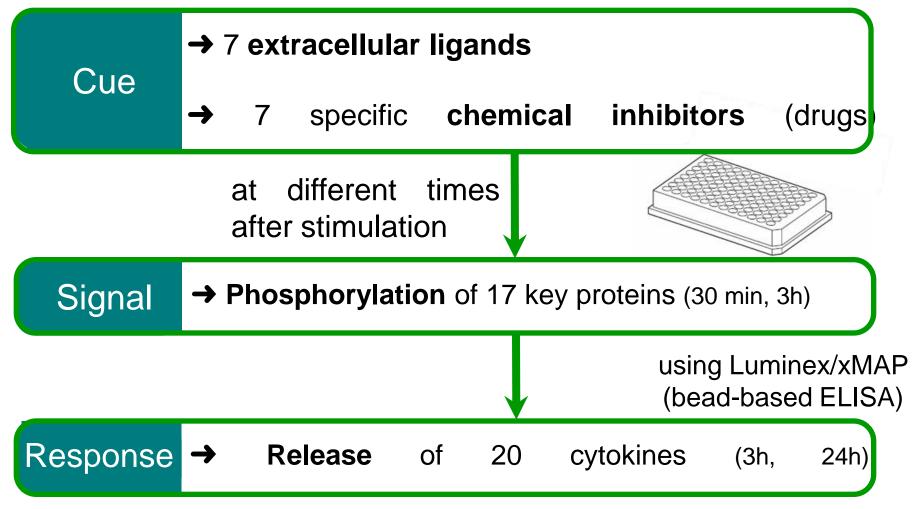
as distributed in the network as possible



#### Design of Cue-Signal-Response experiment for HCC

Performed by Leonidas Alexopoulos

Primary human hepatocytes & HCC cell line HepG2



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

#### **Cue-Signal-Response Compendium of 26,000 measurements**

92 429,4 85,9 10296 30890 25955 29402 396,8 31248

10858 30960 30969

9012 29916 26001

24728 28785

700.9 29750 30556 30767

10000 00104

705.0 01514 709.5 01044

890.0 01085 1798

\$12,00545

\$20 30293

1639 472.2 207.4 27256 01114 01229 01299 2460 1122 1164 27711 01070 00808 01181

496.2

396.2

561

223.9

54,4 27066 155,4 491,5 167,2 390,0

368.2 416.3

391.2

421.8

167.2

282.2 31521 223.4 507

542.8

1203 29091

229.9 27116

2874 27321

430.7

7827 01170 29802 26902

336.9 620.2

548.4 618.2

489.2

2758 471.0 24212

417.8 S44.1 27701 873

014 27747 681.9

29236

9686 45.75

169.5

170

323.S

849.7 12089 124.8

814.9 10236

1042 1019 11199 89.75

1256

815.5

1007

11930

11363

1120 1107 1372 94.65

1440) 1440) 1440) 1440) 1440)

720

1440 1440 1440

ريد، (مدا (مدا (مدا مدا

1440 1440 1440

167.0 6,4 11 249.2

1271

707.0

80.25 30.5

46.S

25.5

46.25

4.5

1149

9800

223.8 -18.75

35.75 2596 38 2320 30 1804

20214

404.7

329.9

28 445,9

32 354,4

1806

2970 2072 220.8 224.2 401.8 1271

9940 1945 353.3

926.7

109

804.3

1158

89.4 29220 400.7 11400 298.7

139.2 30443 550.7 12090 368.9 173.9 30180 629.2 11013 343.9

395.4

215.2 29705 811.8 29139 590.1

158.5 29670 560.4 7708 558.9

S16.4

400.7

392.4

138.8 28391 517.2 9247 696.4

9256 144.4 314.7

7560 S09,4

7237 209.4 102

205.9

9189 821.4 126

121.7 29765 592.4 5061 235.7 74.25 2992 328.4

110.0 29948 - 501.4

114.0 27980 409.9

125.5 28091 402.4

182.0 27995 601.7

29.5 700.9 188.8 28591 787.7 8077 258.7

61 1784 126.7 70.4 860.4

112

95.3 2167

202.7 2821

148,8

146.8 2967

1987 401.2 142.9 1226 3399 403.9 594.9 1155

477.9

170

613.8 357 578.5

1857 341.0

1778 166.8

1800 218.3 123.3 1167

1152

1261

1025

894.0

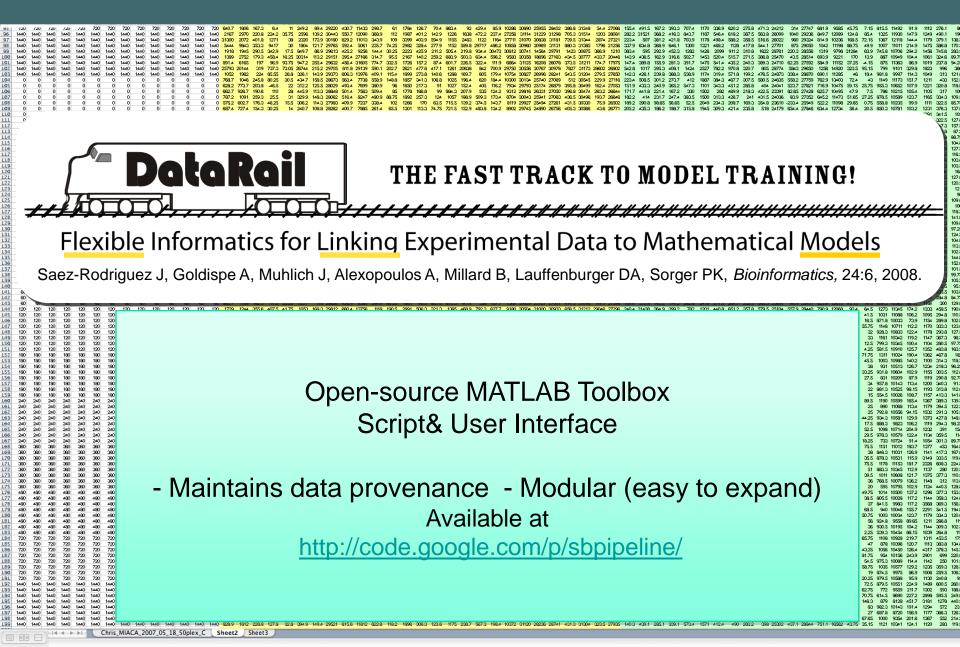
20636

1210 521.9 29287

## How can we handle this data?

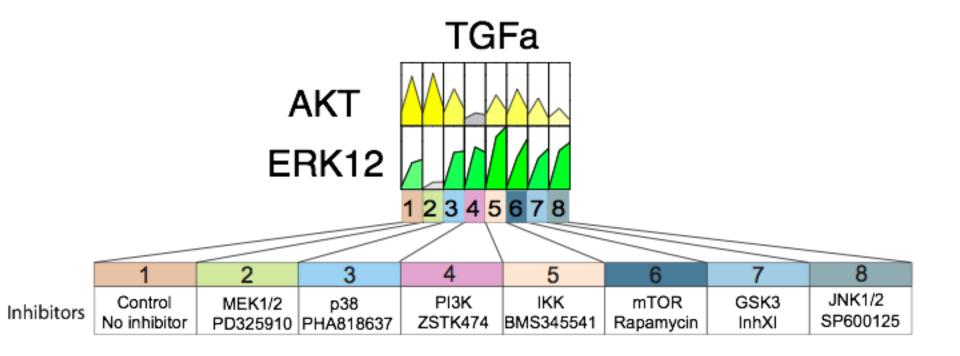
131.3 477.2 390.5 28987 750.9 31307 891.7 7171 30720 26516 29450 465.3 30935 3430 27799 151.4 1216 1189 20.5 125.8 29100 442.2 118 29826 426.9 250.6 391.A 8051 205 428.0 20947 422.5 207 304.2 120.7 -10.5 20.25 8090 66.25 672.3 157.7 625.A 1709 124 255.6 407.5 41.75 1050 109.0.29912\_090.4 10758 1118 190.5 snen 1095 827.7 790.9 12089 792.0 9180 00504 01000 00900 01212 2904B 01409 851.2 057 a 579.5 90.4 1617 186.5 30008 755.7 139 3595 602.5 1200 1280 1064 62.5 10824 1429 30960 8800 39 1595 141 29702 658.4 10111 474.9 119.5 0060 055.5 702 29610 1580 10101 690.4 091.8 4082 851.8 470 1071 671.2 29900 916.2 1941 1039 2616 30420 255 9854 00548 01110 01087 S98.9 01114 09169 07144 179.2 427.5 207.7 90.15 152 290 330.2 824.9 222.8 295.8 348.2 18105 395.9 2000 199.8 516.8 1250 175.9 229.7 014.0 267 44 S 29929 157.8 28209 104 S 6912 20011 20402 20980 485.0 01007 27080 25084 472.5 236.2 1729 329 3551 - 59 9 R10.4 282.3 1890 499.5 48 3236 161.5 20065 751.2 21682 1084 994 618.7 **PP4 9** 2851 519.0 12008 20870 21285 21208 28208 755.A 3927 7177 388 189.8 29490 189.0 29001 820.2 25949 429.9 418.5 29922 412.8 \$09.2 .12.5 290.9 8078 58.25 398.8 800 302 121.2 610.7 26808 239, 30099 188.5 607.5 1172 400.5.28008 90.9 418 1164 \$29 29680 11339 28305 631.4 19214 010.0 029 B 150.0 \$017 01022 202.5 978.1 48.25 1.25 394 10.75 7825 478.0 1372 25365 780.4 S38.5 9910 1110 179.8 01078 190 5 180.0 65.25 1440) 1440) 1440) 1440) 1440) 1440 1440 1440 1440 01287 690.8 2008 29597 1055 025.9 900.4 29000 115.2

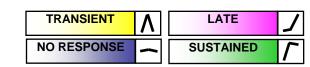
#### **Cue-Signal-Response Compendium of 26,000 measurements**











TIME POINTS: 0, 30min, 3hrs

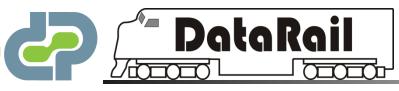
Primary Hepatocytes

Stronger primary Stronger HepG2

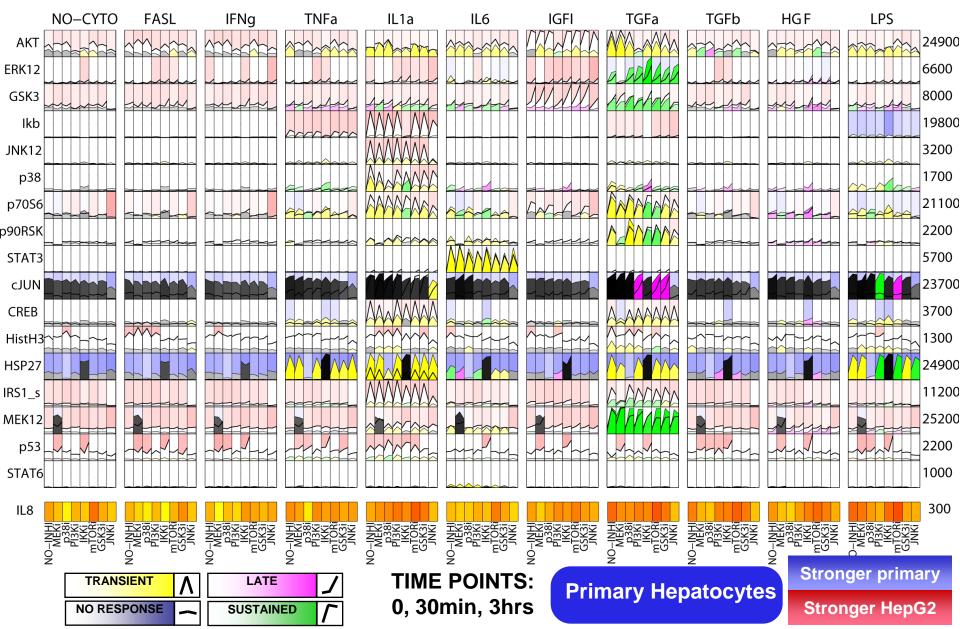


# Visualization of large data sets

	NO-CYTO	FASL	IFNg	TNFa	IL1a	IL6	IGFI	TGFa	TGFb	HGF	LPS
AKT											24900
ERK12											6600
GSK3											8000
lkb											19800
JNK12											3200
p38											1700
p70S6											21100
p90RSK											2200
STAT3											5700
cJUN											23700
CREB											3700
HistH3											1300
HSP27											24900
IRS1_s											11200
MEK12											25200
p53											2200
STAT6											1000
IL8											300
		MEKI D381 J1KKi J1CRi S5K3i J1KKi J1KKi		MEK D S S S S S S S S S S S S S S S S S S			Ĭ X S S S S S S S S S S S S S				Plassing Anti- National Nation
	MO-MEK MEK PI3K IKK MTORI GSK3i JNKi NO-INHJI	R OF FLAN			NO-INKI MEKI PI3X31 MEKI GSK31 JNKI JNKI	E C C C C C C C C C C C C C C C C C C C	MEKI PI3KI B381 IKKI GSK31 JNKi				<u>, 50</u>
-	TRANSIE		LATE			OINTS:				Strong	ger primary
	NO RESPO		SUSTAINE	ر ۲		in, 3hrs		hary Hepa	alocytes	Stron	ger HepG2



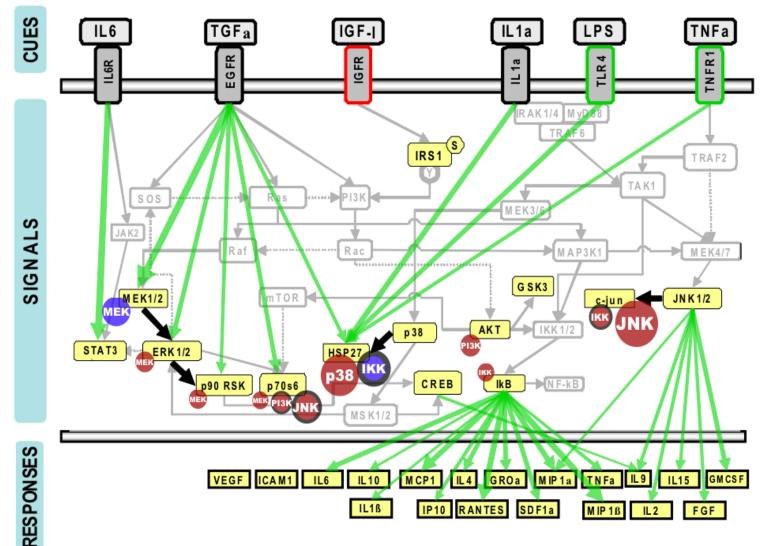
# Visualization of large data sets





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

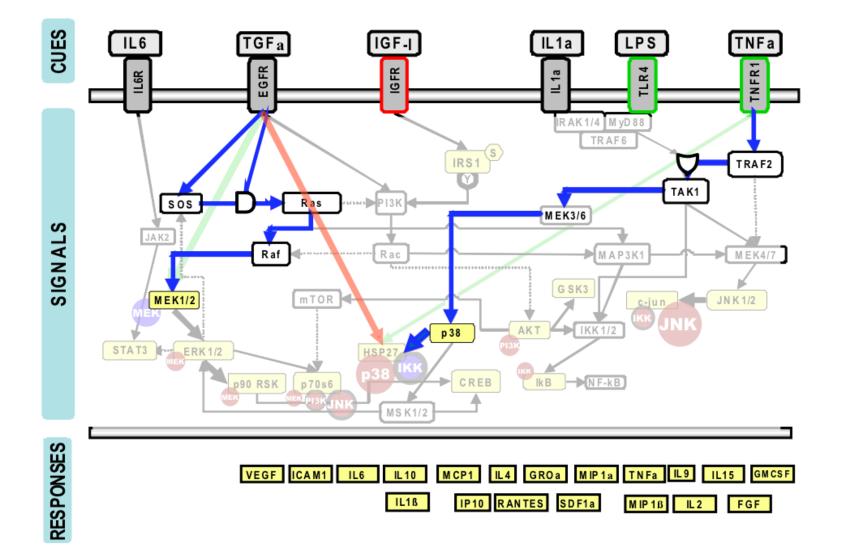
#### **Multiple Regression**



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



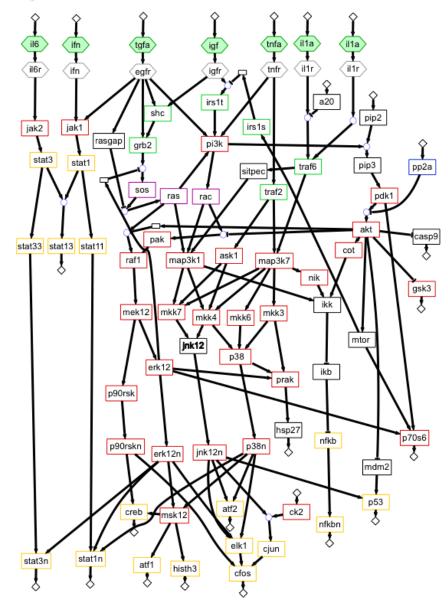
#### Data-driven approaches useful but (in our case) provide limited mechanistic insight



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



# Signaling pathway maps summarize literature knowledge

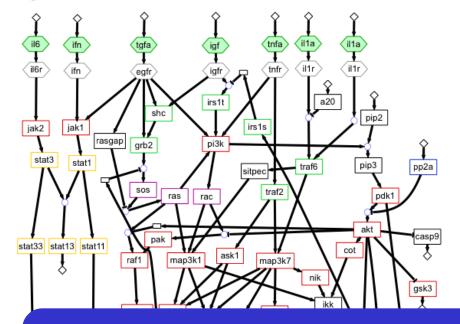


Pathway maps widespread and very useful but

- *Pictures* not computable models to study signal processing
- Not cell-type specific



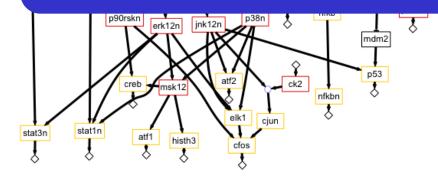
# Signaling pathway maps summarize literature knowledge



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

 $rac{>}$  Transform into **Boolean** (0/1) logic (AND/OR) models  $\checkmark$ 

Define metric to evaluate models given the data

 $\Rightarrow$  Balance fit to data with model size  $\checkmark$ 

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 <u>http://www.ebi.ac.uk/saezrodriguez/software</u>

Different (&complementary) from standard reverse engineering:

- Relies on prior knowledge

 $\rightarrow$  in a first step, reduces search space to plausible connections

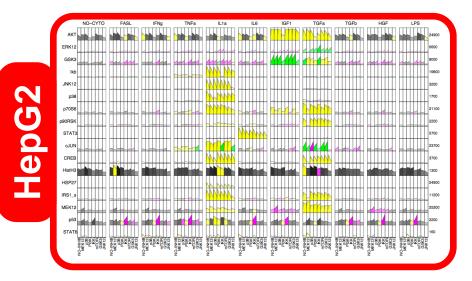
- Considers intermediates

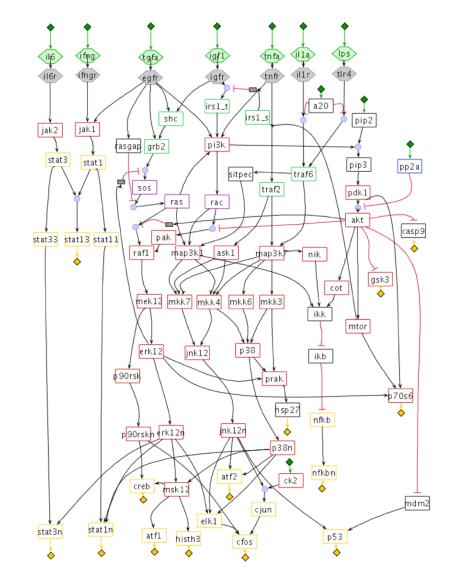
 $\rightarrow$  mechanistic insight

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



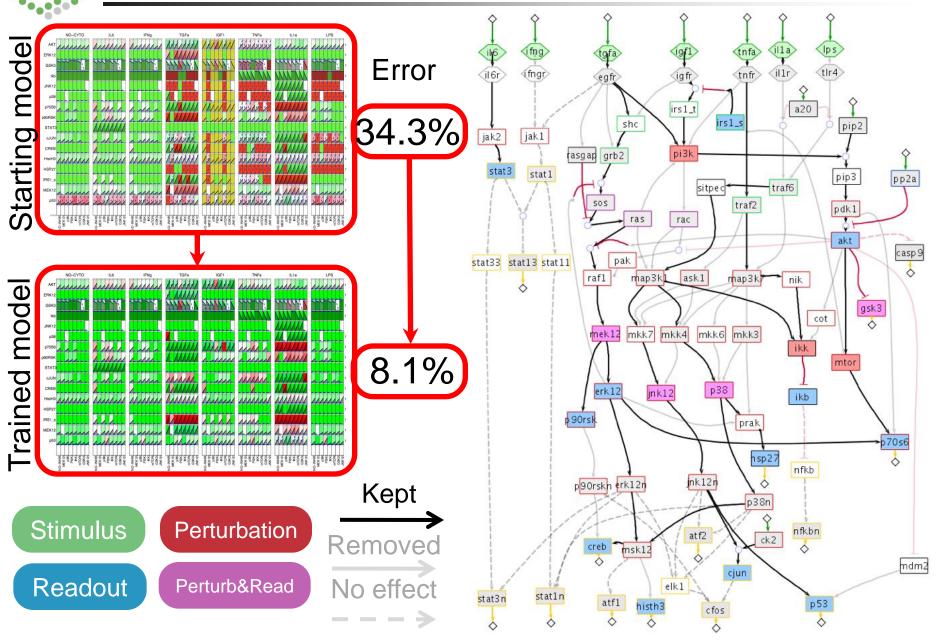
### **Application to signaling in hepatocytes**





### Model trained to HepG2

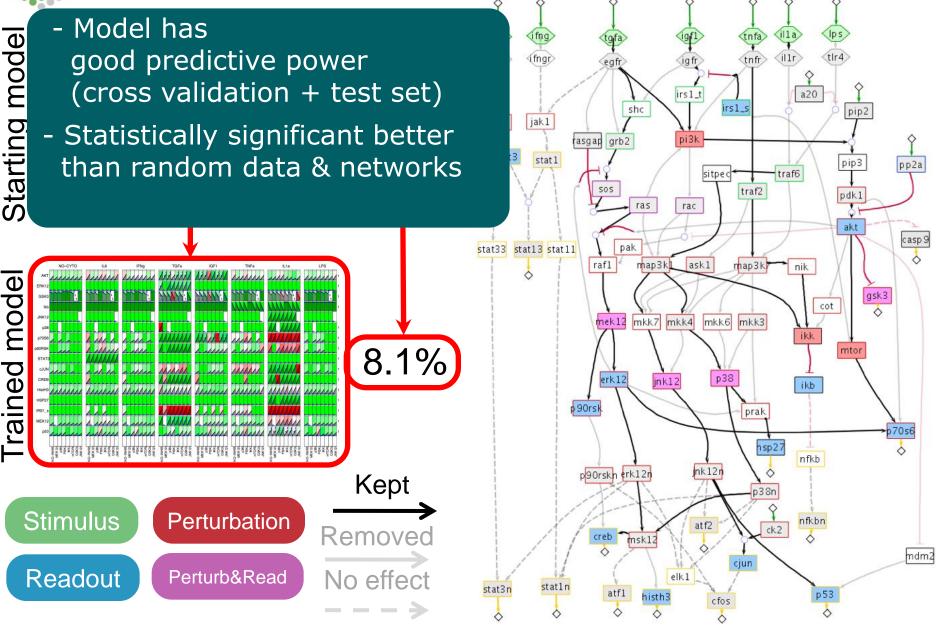
data





## Model trained to HepG2

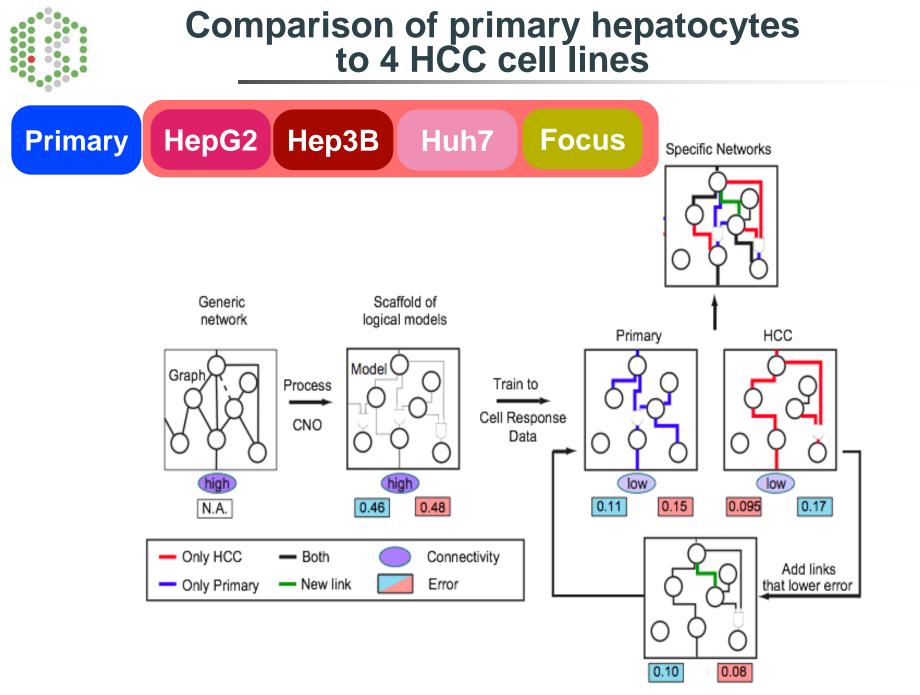
data





- 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, Zheng M, Lauffenburger DA, Sorger PK, submitted.

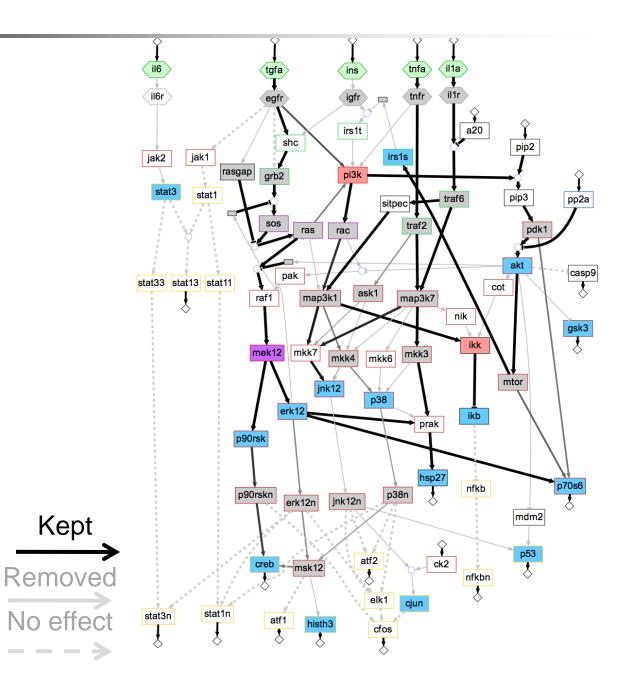


Stimulus

Readout

Perturbation

Perturb&Read

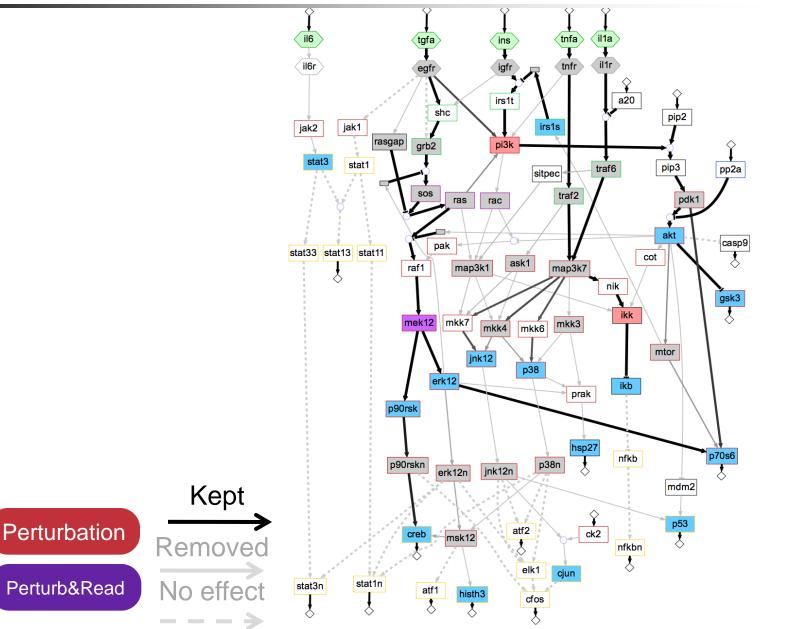


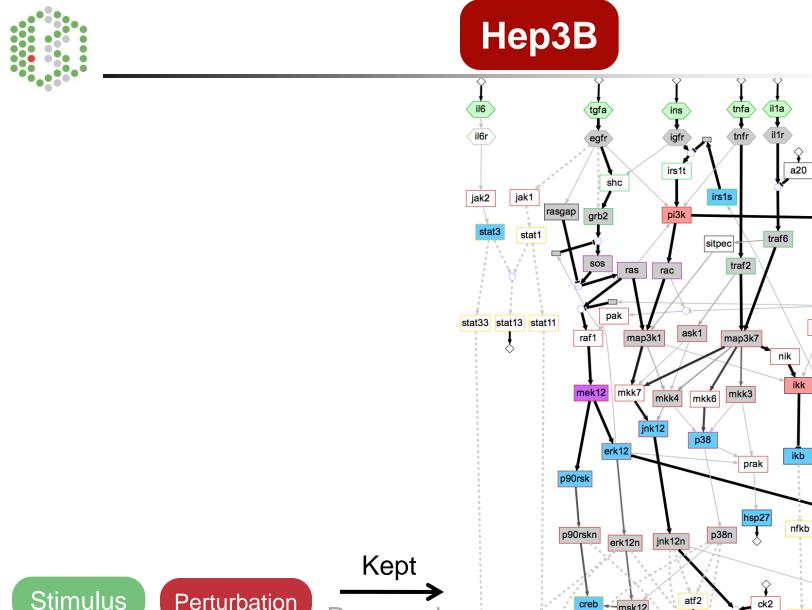


Stimulus

Readout







pip2

pip3

akt

mtor

cot

ikk

ikb

pdk1

pp2a

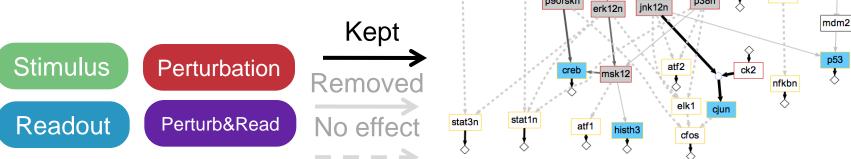
casp9

ł

gsk3

p70s6

Ŷ

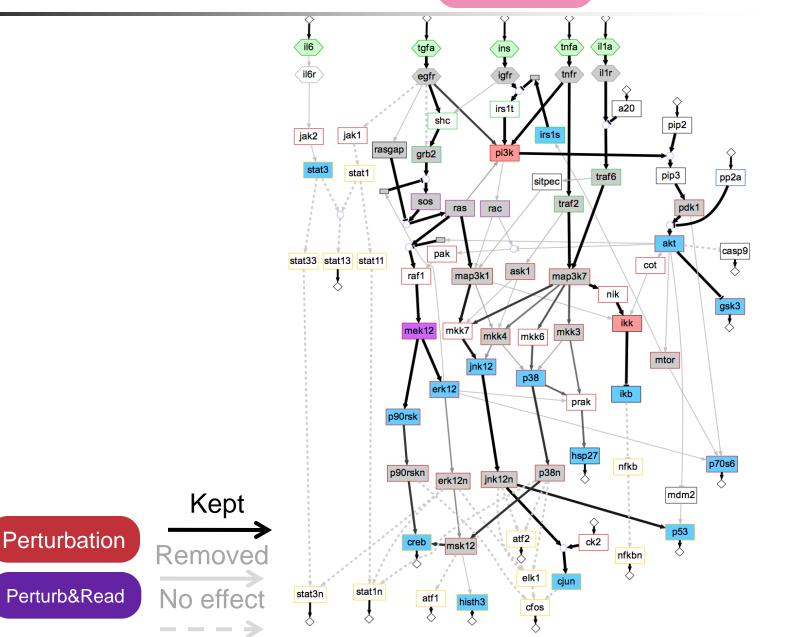




Stimulus

Readout

### Huh7

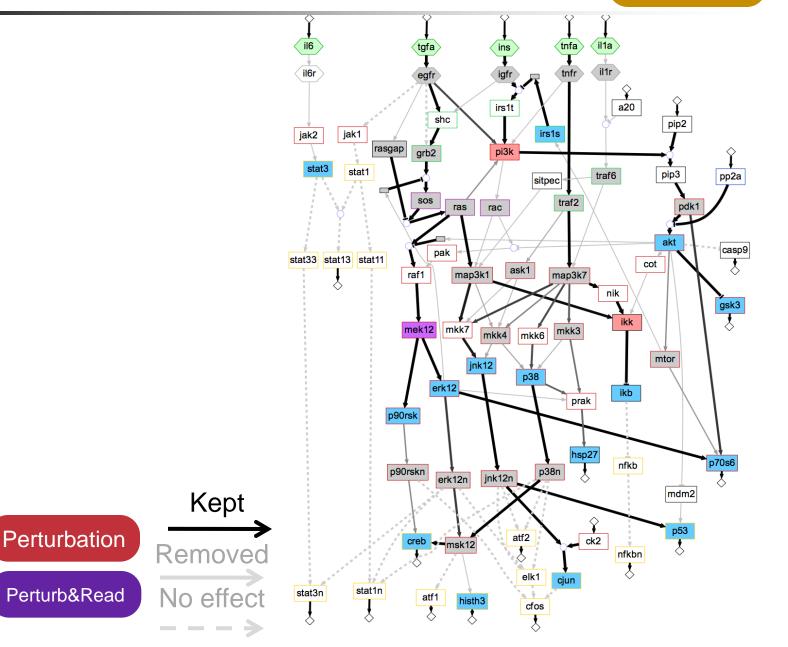




Stimulus

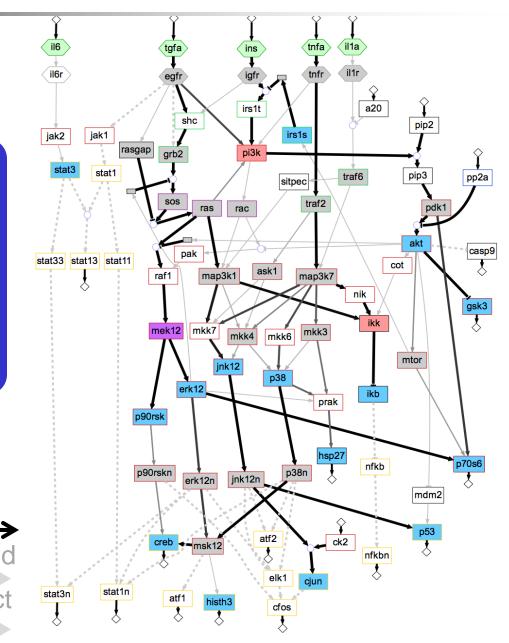
Readout







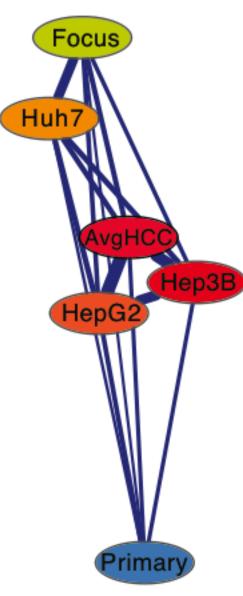
How different are HCC cell lines with respect to each other and to primary hepatocytes?







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

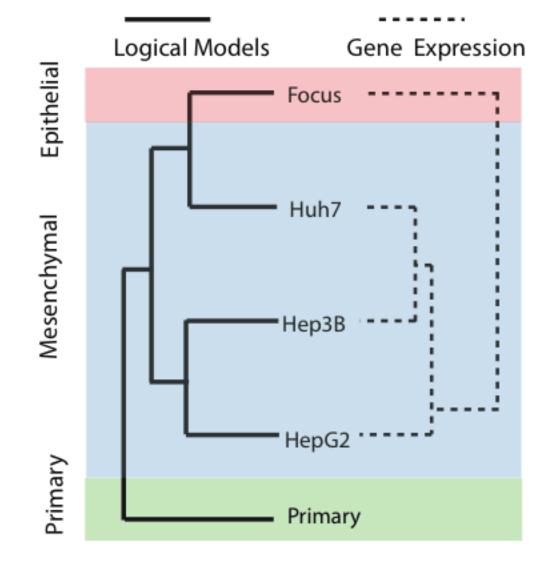


- Define distance between cells based on differences in topology of models
- Including data set <u>AvgHCC</u>: mixed data from HCC cell lines

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



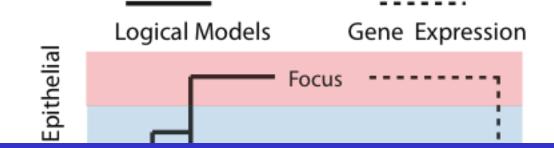
# Network differences roughly correlate with gene expression & phenotypical differences



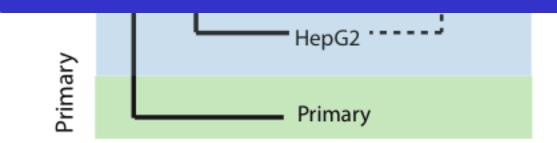
Lee and Horgeirsson, Hepatology, 35:1134, 2002. Fuchs et al, Cancer Res., 68:2391, 2008.



# Network differences roughly correlate with gene expression & phenotypical differences



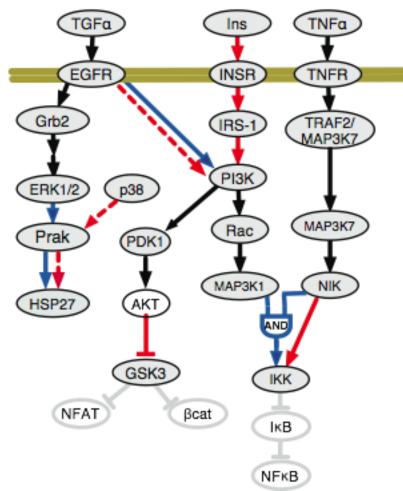
What is common to cancer cell lines and different to primary hepatocytes in the functioning of the signaling network?

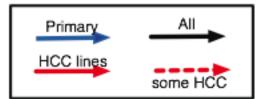


Lee and Horgeirsson, Hepatology, 35:1134, 2002. Fuchs et al, Cancer Res., 68:2391, 2008.



### Specific pathways determine major differences



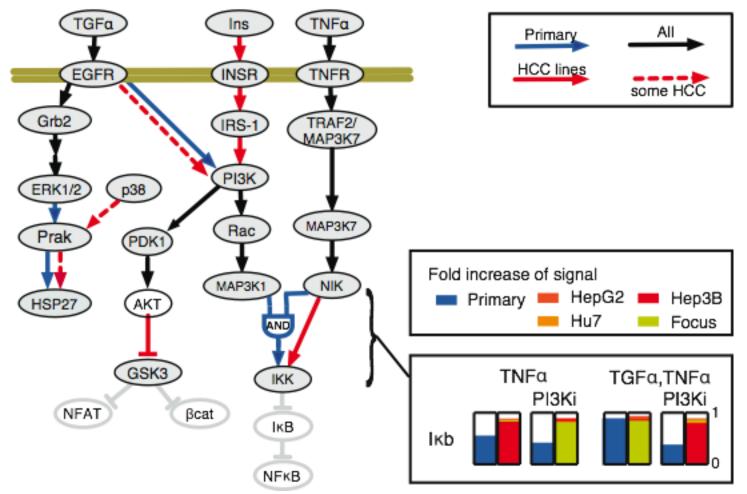


- Only active in HCC cell lines: Insulin→.. →AKT→GSK3
- HSP27 phosphorylation: ERK mediated in primary

 Difference in NFkB activation: TNF dependent only in HCC TNF+TGFa in primary



#### Specific pathways determine major differences



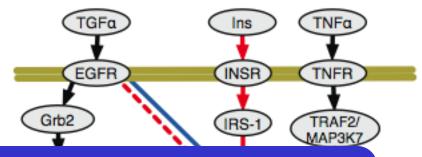
- Only active in HCC cell lines: Insulin→.. →AKT→GSK3
- HSP27 phosphorylation: ERK mediated in primary

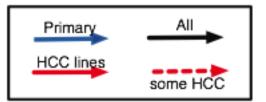
 Difference in NFkB activation: TNF dependent only in HCC TNF+TGFa in primary



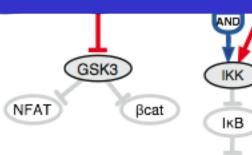
#### Specific pathways determine major differences

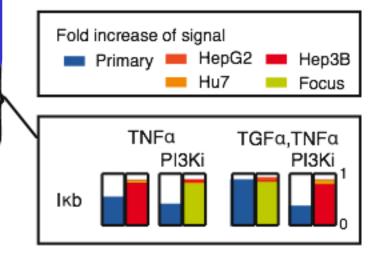
NFKB





# What is the origin of these differences?



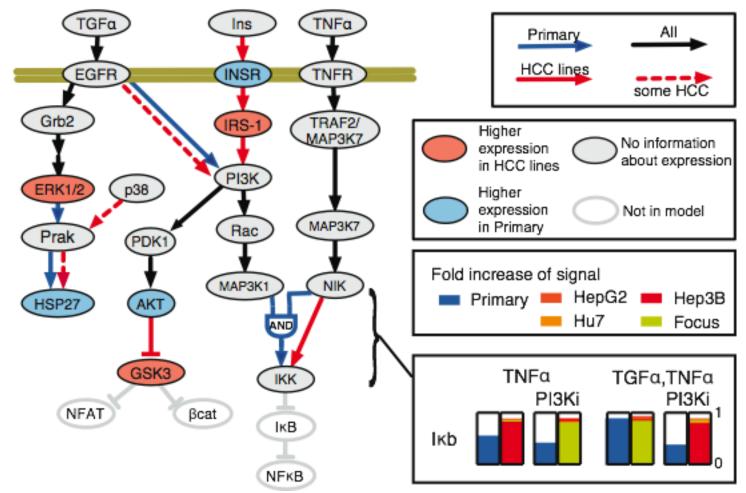


- Only active in HCC cell lines: Insulin→.. →AKT→GSK3
- HSP27 phosphorylation: ERK mediated in primary

 Difference in NFkB activation: TNF dependent only in HCC TNF+TGFa in primary



#### Specific pathways determine major differences

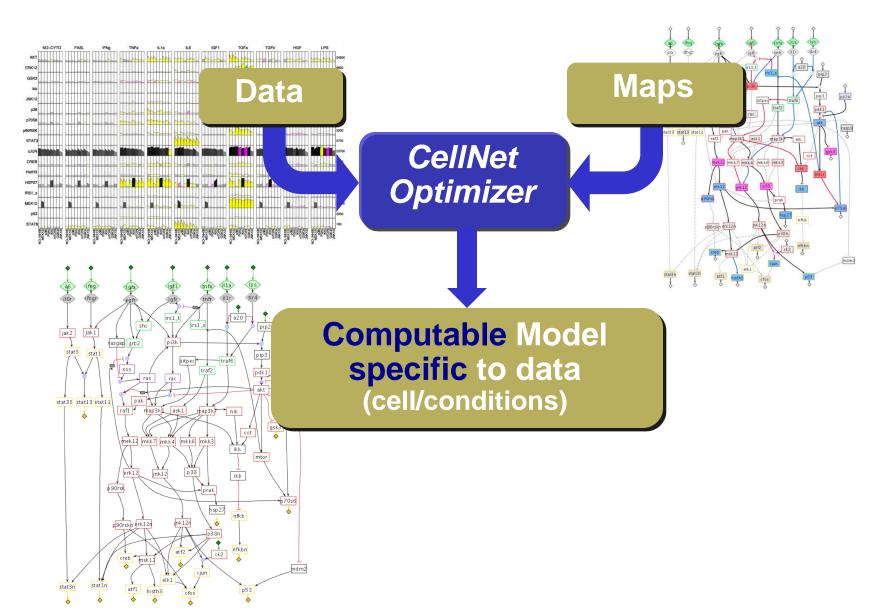


- Only active in HCC cell lines: Insulin→.. →AKT→GSK3
- HSP27 phosphorylation: ERK mediated in primary

 Difference in NFkB activation: TNF dependent only in HCC TNF+TGFa in primary



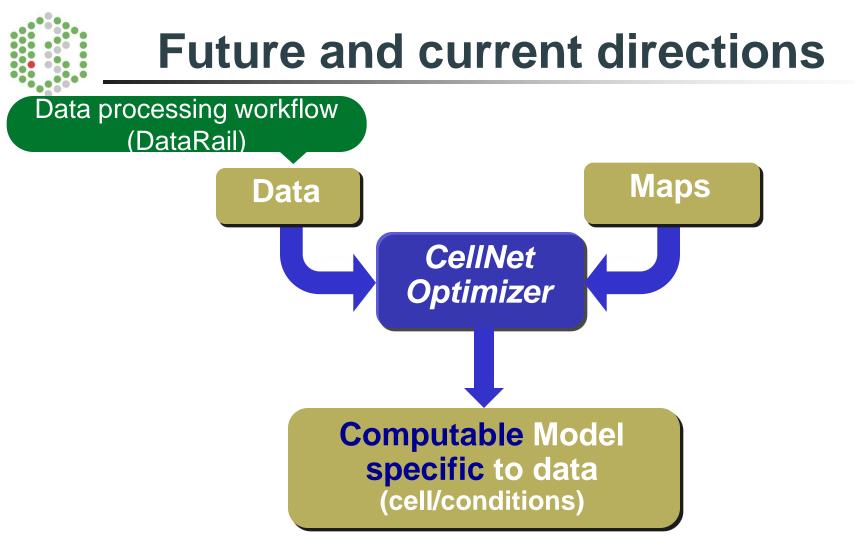
### Summary

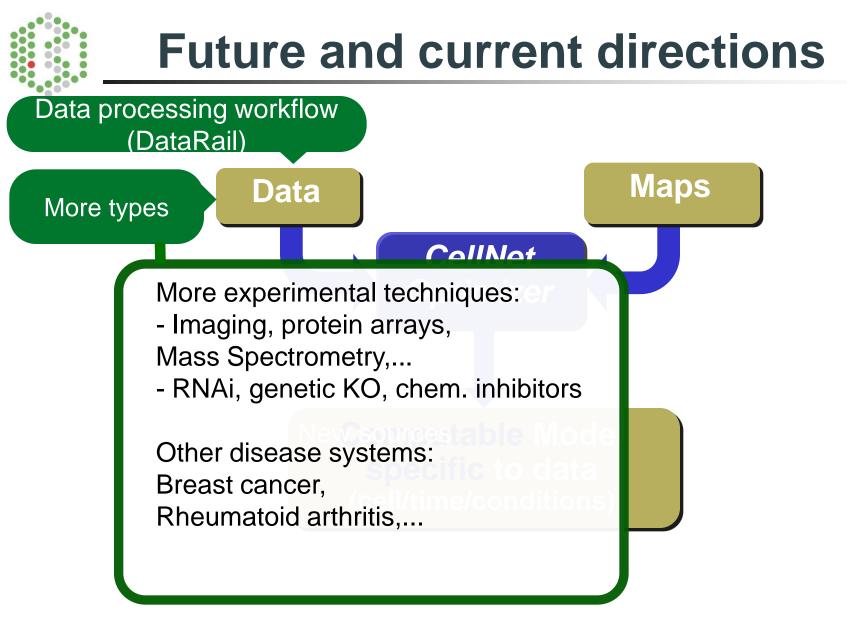


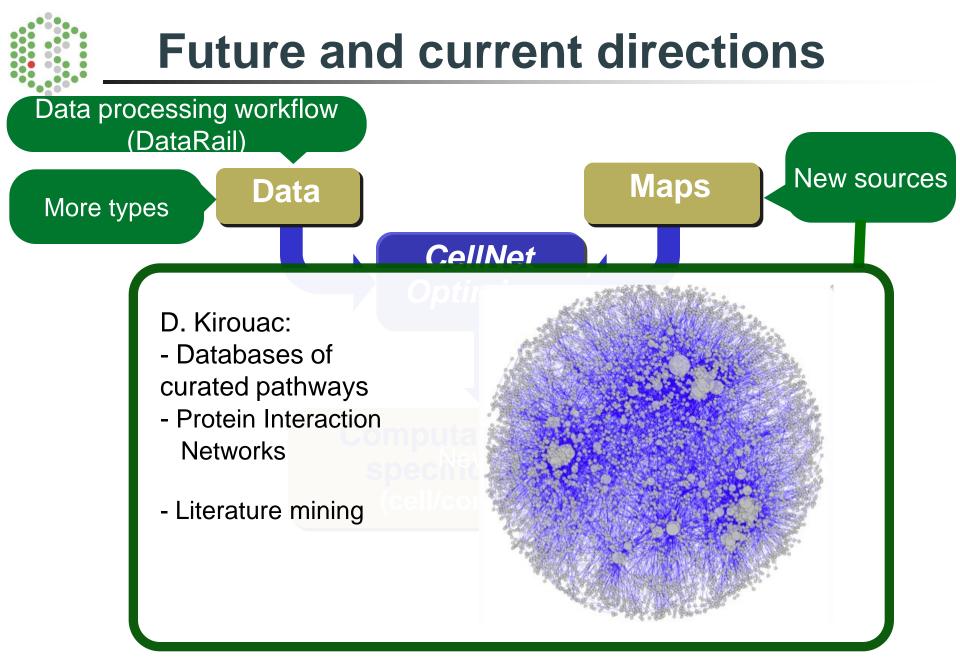


#### Summary

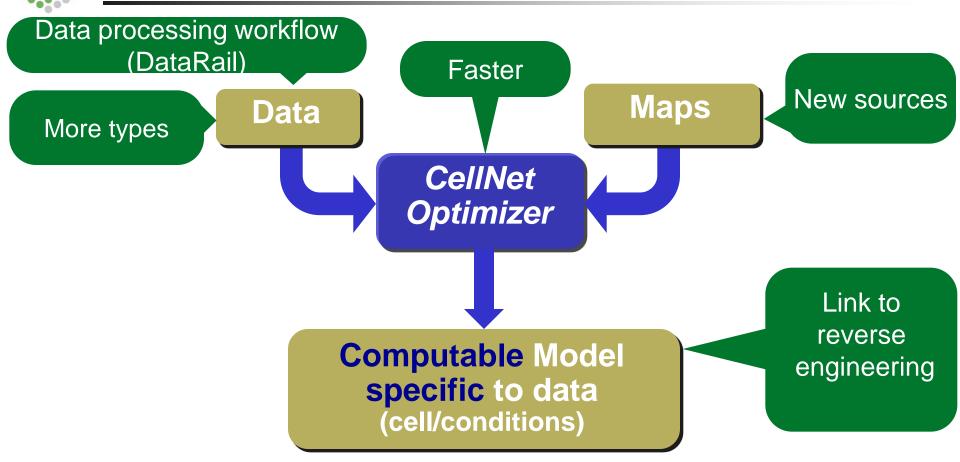
- Pathway maps are not specific
- Models trained to data are much sparser and predictive
- It is possible to
  - Construct models specific for cell types
  - Cluster cell types based on pathways models
  - Pinpoint specific differences between normal and diseased cells



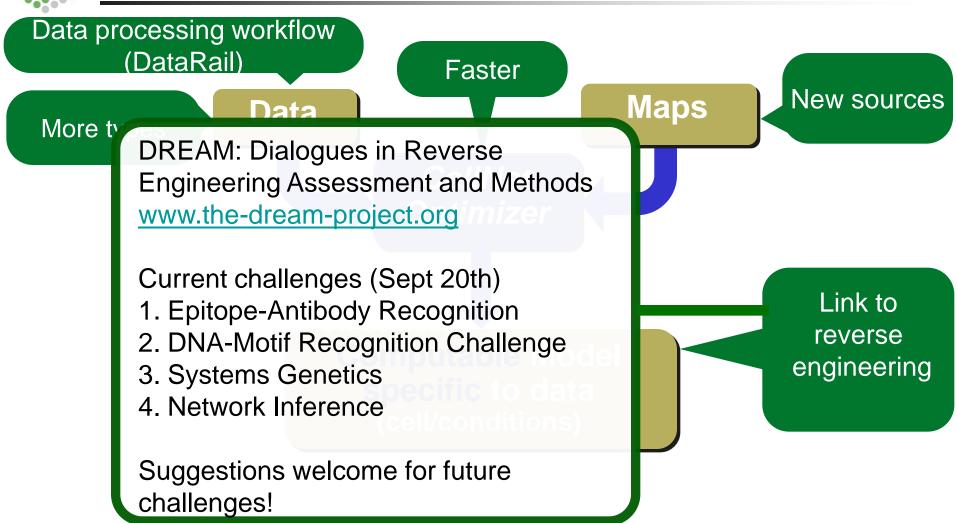


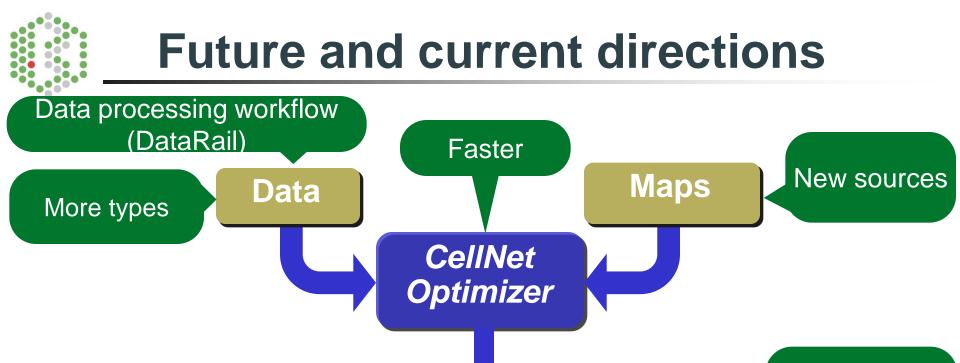


## **Future and current directions**



## Future and current directions





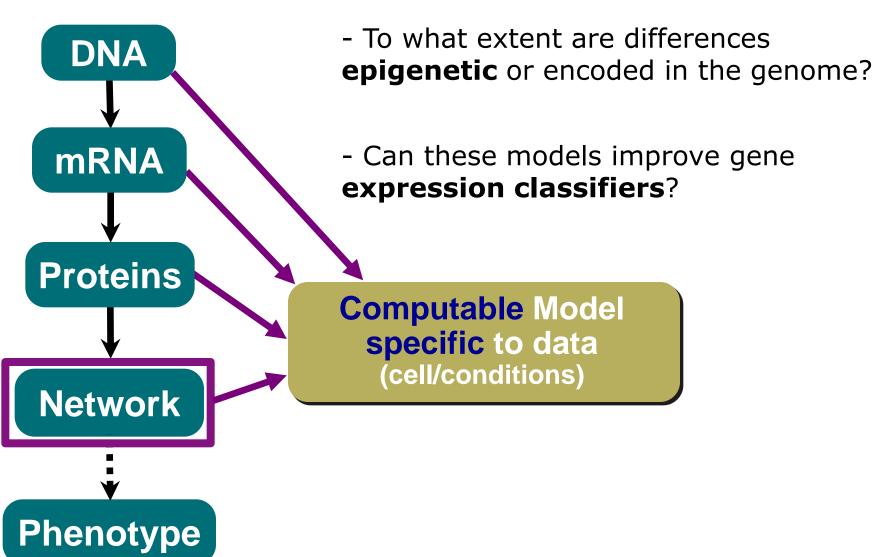
Methods to fingerprint drug-effects

Computable Model specific to data (cell/conditions) Link to reverse engineering

Model refinement Fuzzy logic (M. Morris)

## Future and current directions

#### Linking post-translational events to genetic alterations





## Acknowledgments

**Melody Morris** 

Dan Kirouac

Joel Wagner

Nickel Dittrich

Arthur Goldsipe

- Jeremy Muhlich
- Mario Niepel
- Bree Aldridge
- Sean Milton



- Jonathan Epperlein
- Laura Kleiman
- William Chen
- Peter Sorger
- Leonidas Alexopoulos (NTUA)
  Ioannis Melas
- Regina Samaga
- Max-Planck Institute

MD

- Steffen Klamt
- Sebastian Mirschel
- Ernst Dieter Gilles

Eduardo Sontag (Rutgers)

Bea Penalver (NW Univ)

Fabian Theis (Helmholtz)

**Douglas Lauffenburger** 

Alexander Mitsos (RWTH Aachen)

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

Funding by Pfizer & NIH (GM68762, CA112967)