

Using Pathway Logic
to Integrate
Signal Transduction
and
Gene Expression Data

SRI International

Linda Briesemeister

Steven Eker

Merrill Knapp

Keith Laderoute

Patrick Lincoln

Andy Poggio

Carolyn Talcott

Lawrence Berkeley
National Laboratory

Joe Gray

Laura Heiser

Paul Spellman

Symbolic Systems Biology

The *qualitative and quantitative* study of biological processes as integrated systems not just isolated parts.

Goals:

- Model causal networks of biomolecular processes and interactions in a logical framework
- Develop formal models that are as close as possible to domain expert's mental models
- Compute with and analyze these networks
 - Abstract and refine logical models
 - Simulate or use deduction to check properties
- Make predictions, experiment, update model

Pathway Logic

Pathway Logic (PL) is an approach to modeling biological entities and processes based on rewriting logic. Signal transduction processes are modeled at different levels of abstraction in the PL knowledge base. The resulting signaling networks can be queried using formal methods tools.

For example, given an initial state:

- execute--show me some signaling pathway
- search--show me all pathways leading to a specified final condition
- model-check--is there a pathway with certain given properties?

How it works

Regulation of the Raf-MEK-ERK pathway by protein phosphatase 5

Alex von Kriegsheim, Andrew Pitt, G. Joan Grindlay, Walter Kolch and Amardeep S. Dhillon

The Raf-MEK-ERK pathway couples growth factor, mitogenic and extracellular matrix signals to cell fate decisions such as growth, proliferation, migration, differentiation and survival^{1,2}. Raf-1 is a direct effector of the Ras GTPase and is the initiating kinase in this signalling cascade. Although Raf-1 activation is well studied, little is known about how Raf-1 is inactivated. Here, we used a proteomic approach to identify molecules that may inactivate Raf-1 signalling. Protein phosphatase 5 (PP5) was identified as an inactivator that associates with Raf-1 on growth factor stimulation and selectively dephosphorylates an essential activating site, Ser 338. The PP5-mediated dephosphorylation of Ser 338 inhibited Raf-1 activity and downstream signalling to MEK, an effect that was prevented by phosphomimetic substitution of Ser 338, or by ablation of PP5 catalytic function. Furthermore, depletion of endogenous PP5 increased cellular phospho-Ser 338 levels. Our results suggest that PP5 is a physiological regulator of Raf-1 signalling pathways.

Biology Research Papers



A fibroblast cell with nothing outside of the cell, EgfR in the membrane, and Src inside the cell ...

Desired initial state



Starting with NRG1 and EGF stimuli, find a pathway leading to phosphorylation of all 4 ErbB receptors

Query



```

rl[1.EgfR.on]:
  ?ErbB1L:ErbB1L
  [CellType:CellType | ct:Soup
  {CLo | clo          }
  {CLm | clm EgfR     } ]
=>
[CellType:CellType | ct:Soup
{CLo | clo ?ErbB1L:ErbB1L  }
{CLm | clm [EgfR - act]   } ] .
*** ErbB1Ls are AR EGF TGFa Btc Epr HB-EGF
*** 10076047(R) ligand binding, autophosphorylation
*** 11350724(R)
*** 11566606(R)

```

Curated Rules

```

rl[426.EgfR.xby.Dok2]:
  {CLm | clm [EgfR - act] }
  {CLi | cli              }
  {CLc | clc Dok2        }
=> ...

```

= Pathway Logic

```

[Fibroblast |
{CLo | empty
}{CLm | EgfR
}{CLi | Src
} ...

```

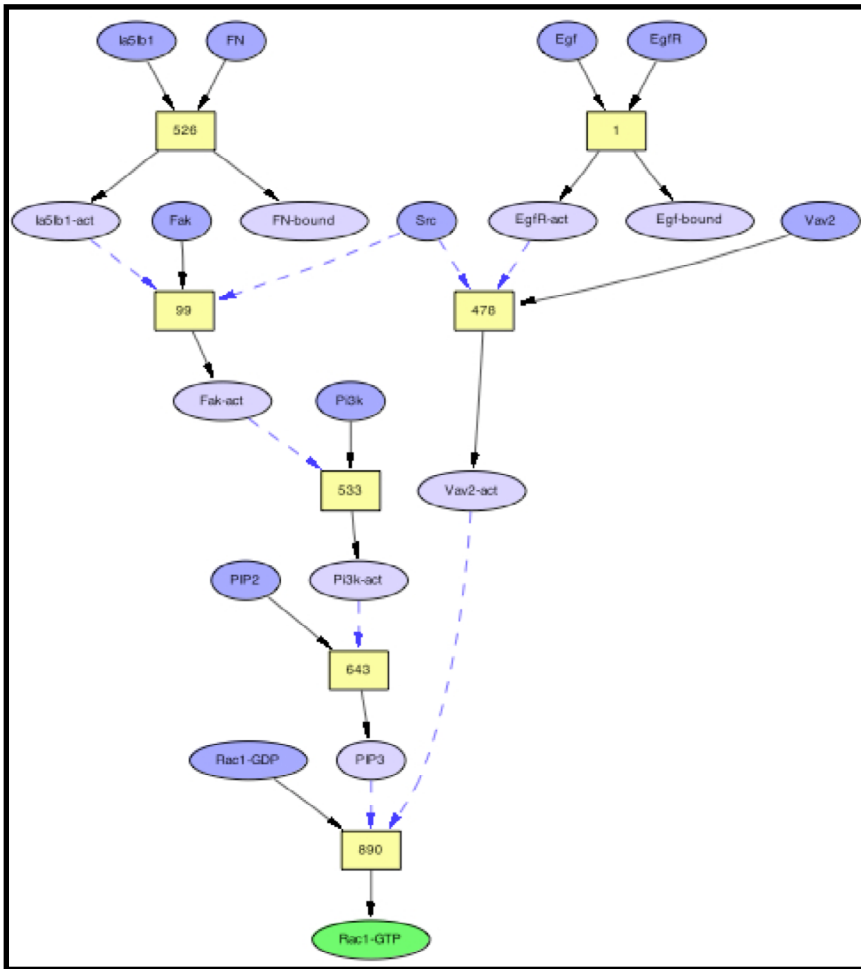
Cell state

```

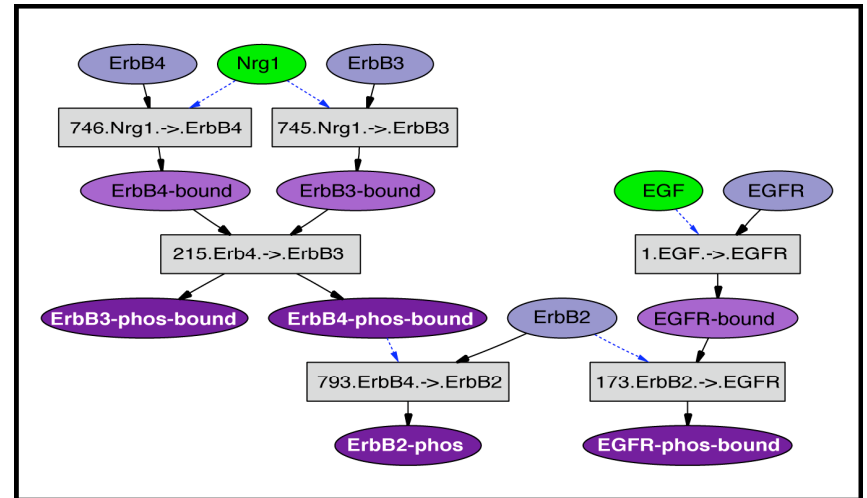
Rewrite logic
Model checking
Query and search
Graphics

```

Computer Code



Graphical representation of pathway

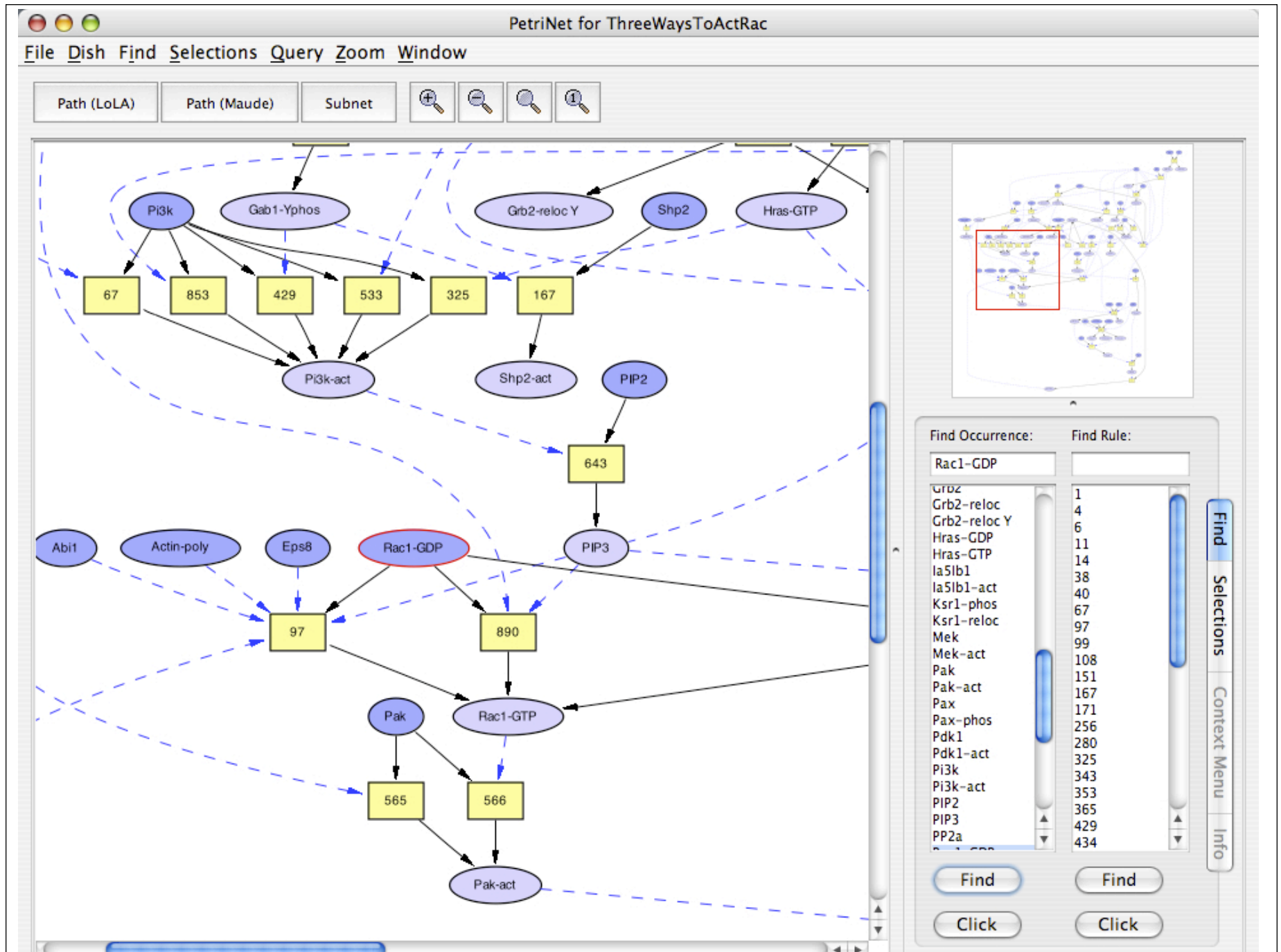


Query result

The Pathway Logic Assistant (PLA)

Provides interactive visual representation of PL models. Using PLA one can

- choose a model/initial state
- display the network of reactions for a chosen model
- formulate queries -- specify goals/avoids
- display relevant subnet or pathway
- compare two subnets and/or pathways
- compute knockouts and show effects on network
- display downstream impact of given components
- color pathway according to gene expression levels
- animate coloring of gene expression time series



Pathway Logic Basics

How is signaling represented in the PL knowledge base?

A cell and its ligands are represented as a term
ligands [cellType | locations]

Each location has the form

{ locationName | components }

A signaling rule has the form

cellStateBefore => cellStateAfter

Example Rule: Activation of PrIR

```
r1[766.PrIR.by.PrI]:  
  PrI  
  [any:CellType | ct  
    {CLO | clo}{CLM | clm PrIR}]  
=>  
  [any:CellType | ct  
    {CLO | clo [PrI - bound]}  
    {CLM | clm [PrIR - act]}] .  
-----  
*** 11566606(R) PrIR is a homodimer
```

In any cell containing the receptor PrIR in its membrane, if the ligand PrI is present in the supernatant containing the cell, then it will bind to PrIR on the outside surface of the cell, [PrI - bound], and PrIR will become activated, [PrIR - act]..

Example Rule: Phosphorylation of Cbl

```
r1[816.Cbl.by.Pr1R]:  
  {CLm | clm [Pr1R - act]}  
  {CLi | cli Fyn}{CLc | clc Cbl}  
=>  
  {CLm | clm [Pr1R - act]}  
  {CLi | cli Fyn [Cbl - Yphos]}{CLc | clc} .  
-----  
*** 9890970(D) Cbl is phosed on Y731
```

Activated Pr1R, in the presence of Fyn, causes tyrosine phosphorylation of Cbl, [Cbl - Yphos]. The specific phosphorylation site, tyrosine 731, is not represented explicitly, but kept in the annotation in case in the future making this explicit should become important.

PL models from gene expression data

mRNA expression data was used to create a putative initial state for 51 breast cancer cell lines.

Of the several hundred initial state components, most were taken to be present in all cell lines; about forty varied across the cell lines. For each of the initial states, the corresponding network of signaling rules was generated.

An unsupervised hierarchical clustering on the network components that varied across the 51 networks yielded 20 rule clusters. Some clusters were deemed not relevant. Three of the remaining clusters are shown in Figure 1: Rule Clusters.

Figure 1: Rule Clusters

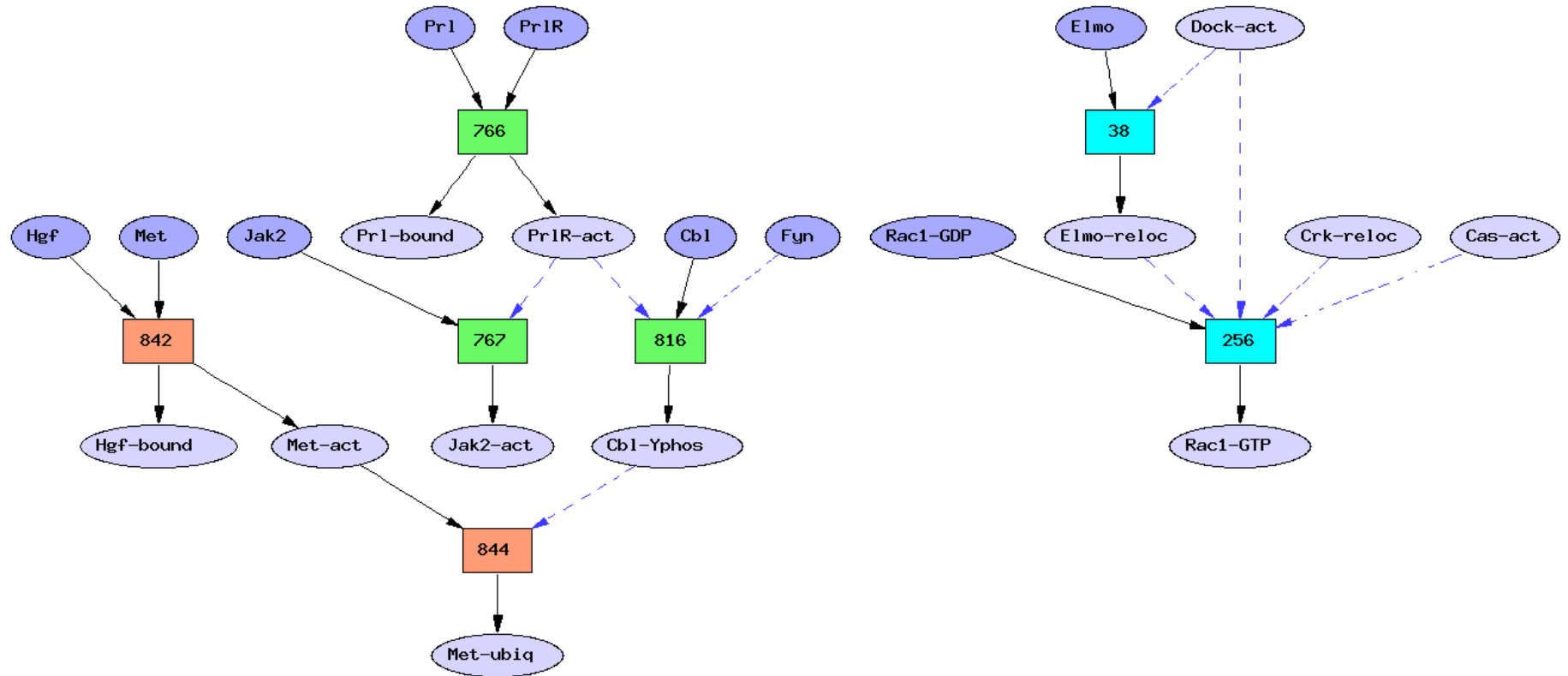


Figure 1: Caption

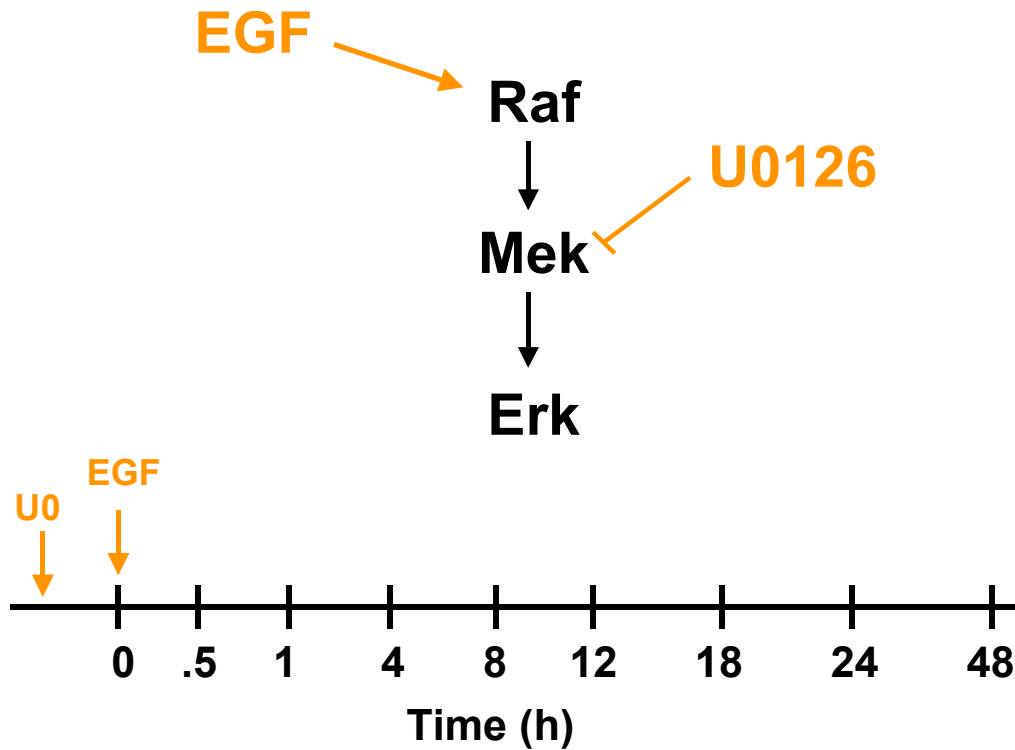
The PrlR cluster (green rules) appears in 5 cell lines (all luminal). Enhanced activity of the PrlR may be a significant risk factor for human breast cancer.

The Met cluster (orange rules) appears in many of the cell lines. Met is a potent source of signals both for the proliferation and chemotaxis of various human cancer cells, including breast cancer cells.

The Elmo cluster (cyan rules) contains rules related to Elmo and its role in activation of Rac1.

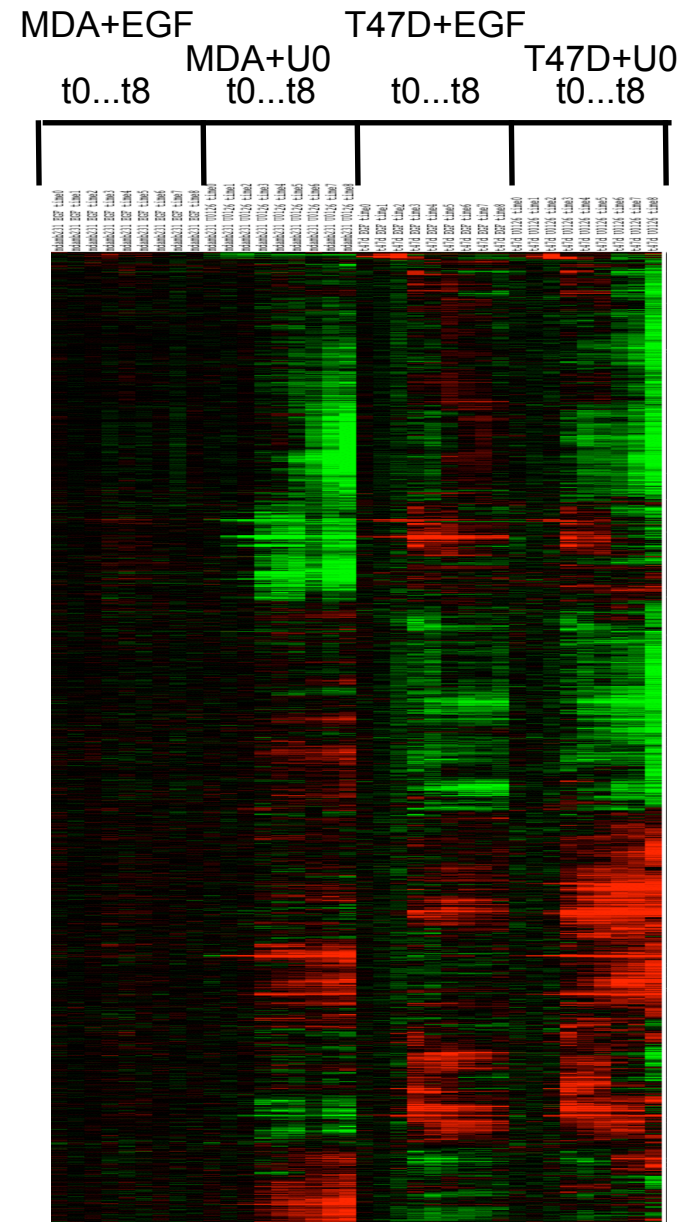
Assessment of transcriptional changes induced by EGF and the MEK inhibitor U0126

Experimental Design

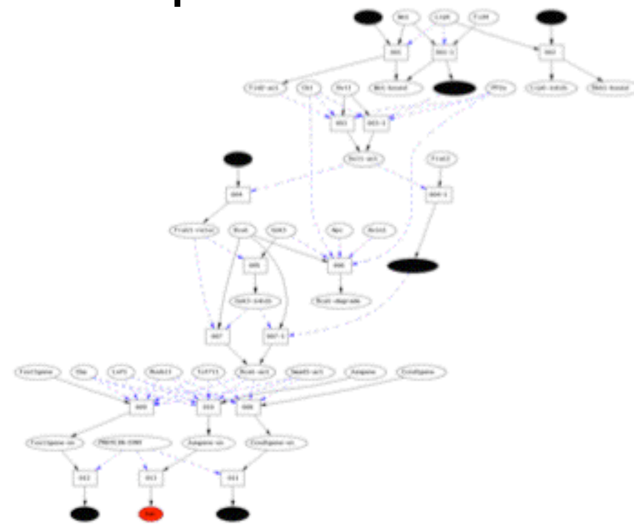


Assessment of transcriptional changes induced by EGF and the MEK inhibitor U0126 in two breast cancer cell lines: MDA-MB231 and T47D

MDA-MB231 and T47D show very different patterns of gene expression after drug treatment
 ~2100 genes show temporal variation in at least one condition



T47D + EGF timecourse Painted on the Wnt path



(a) 1 hr



(b) 4 hr



(c) 8 hr



(d) 12 hr

Red: up
Green: down
Black: no change

References

[PL web] <http://pl.csl.sri.com>

[ICBP] LBNL Integrative Cancer Biology Program <http://icbp.lbl.gov>

[Rewriting Logic] J. Meseguer. Conditional Rewriting Logic as a unified model of concurrency. *Theoretical Computer Science*, 96(1):73--155, 1992.

[PL2004] C. Talcott, S. Eker, M. Knapp, P. Lincoln, and K. Laderoute. Pathway logic modeling of protein functional domains in signal transduction. In *Proceedings of the Pacific Symposium on Biocomputing*, January 2004.

[PLA2005] Carolyn Talcott and David L. Dill. The pathway logic assistant. In Gordin Plotkin, editor, *Third International Workshop on Computational Methods in Systems Biology*, pages 228--239, 2005.

The Wnt Signaling Pathway

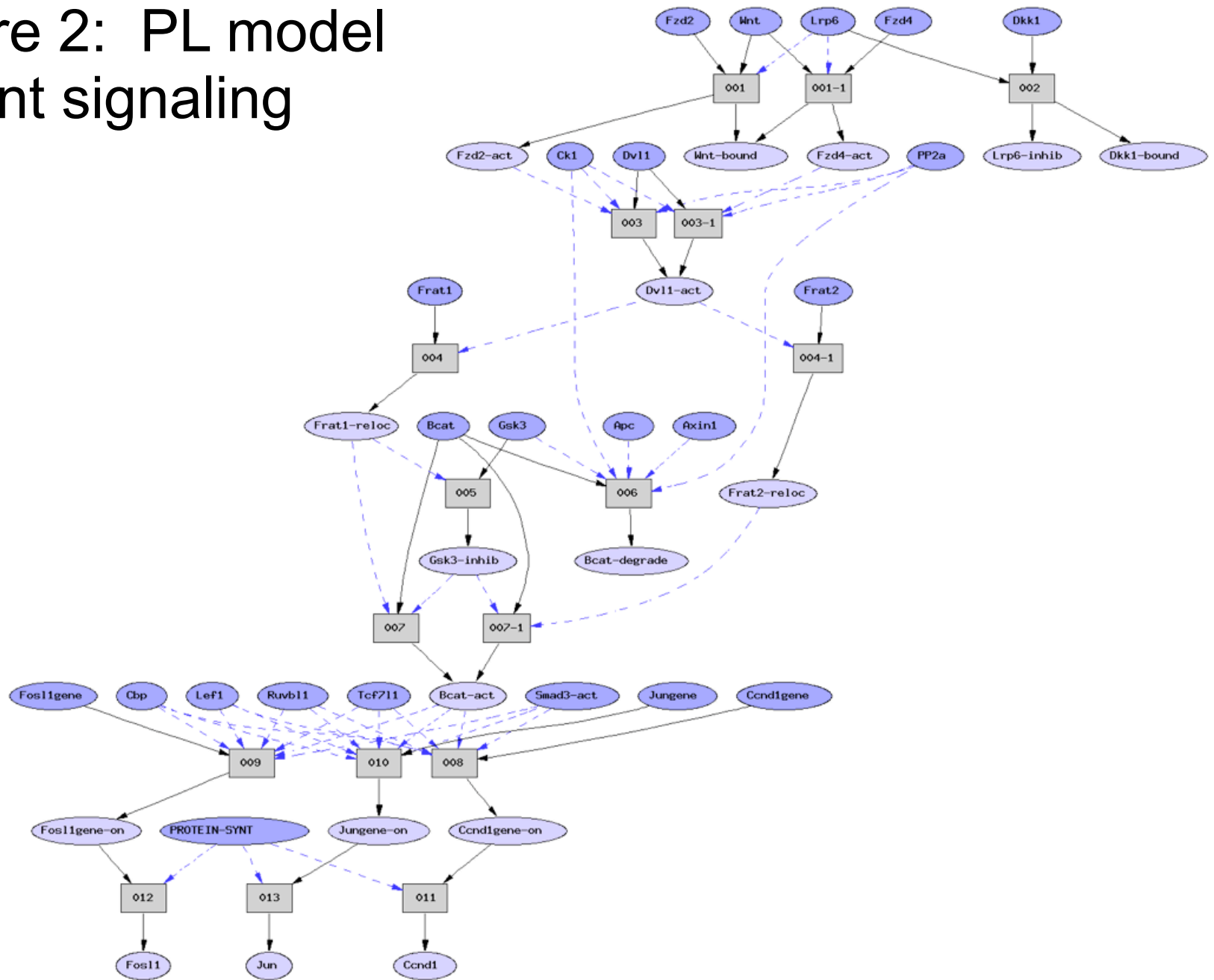
Figure 2 shows the PL model of the Wnt signaling system, which is important for both patterning of the vertebrate embryo, the maintenance of self-renewing tissues in the adult, and is implicated in the development of diverse human carcinomas.

Each oval represents a protein (or other molecule) with a specific modification and cellular location.

Rectangles represent reactions (rules).

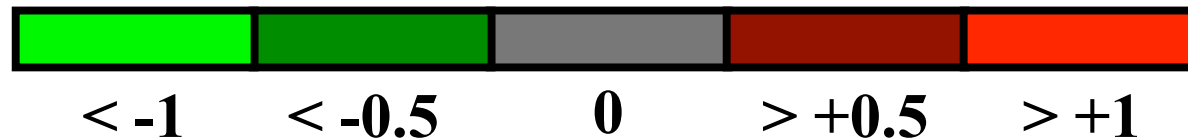
The dark ovals indicate components present in the initial state.

Figure 2: PL model of Wnt signaling



Visualizing course expression data

Figures 3-5 show the response of T47D cells to treatment of Egf at 1, 4 and 8 hours post treatment, painted on the Wnt signaling pathway. For each protein in the model, changes in expression level were mapped to one of five colors: two shades of green, two shades of red and gray. Green indicates that gene expression is down-regulated following Egf treatment, red indicates up-regulation, and gray indicates no change in expression.



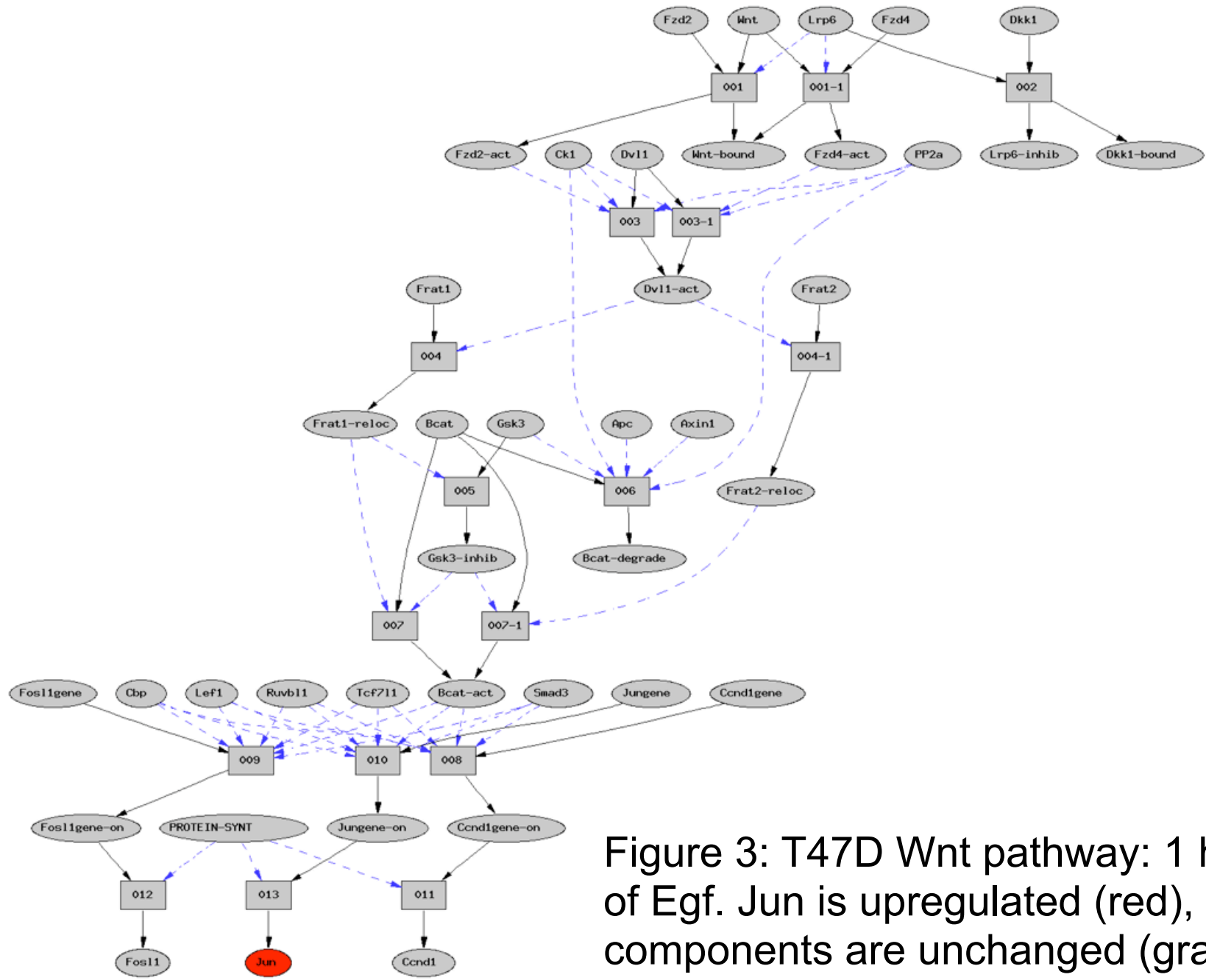


Figure 3: T47D Wnt pathway: 1 hour of Egf. Jun is upregulated (red), other components are unchanged (gray).

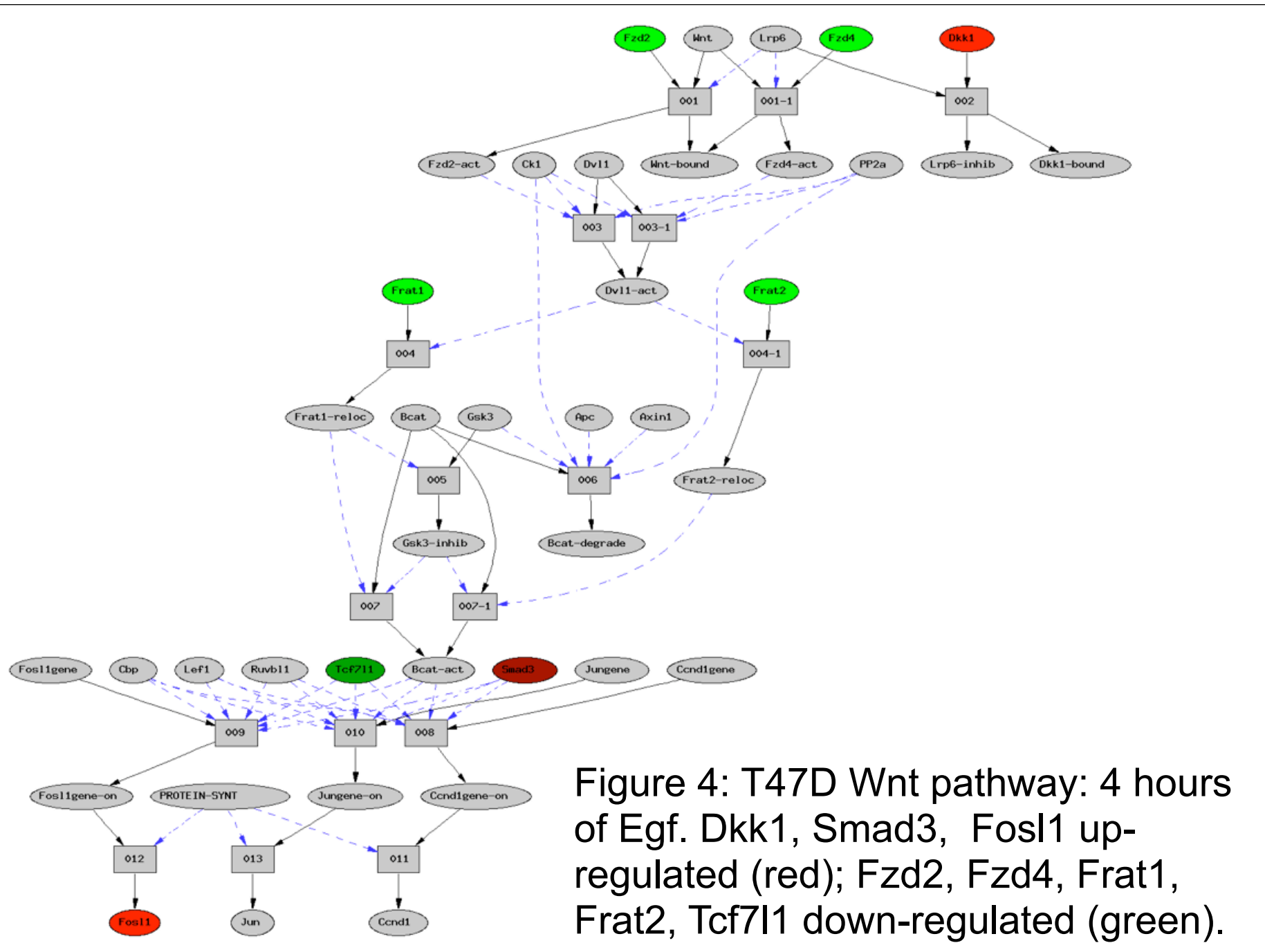


Figure 4: T47D Wnt pathway: 4 hours of Egf. Dkk1, Smad3, Fosl1 up-regulated (red); Fzd2, Fzd4, Frat1, Frat2, Tcf711 down-regulated (green).

