Using Pathway Logic to Integrate Signal Transduction and Gene Expression Data

SRI International

Linda Briesemeister
Steven Eker
Merrill Knapp
Patrick Lincoln
Andy Poggio
Carolyn Talcott
Keith Laderoute

Lawrence Berkeley
National Laboratory

Laura Heiser
Paul Spellman
Joe Gray

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.

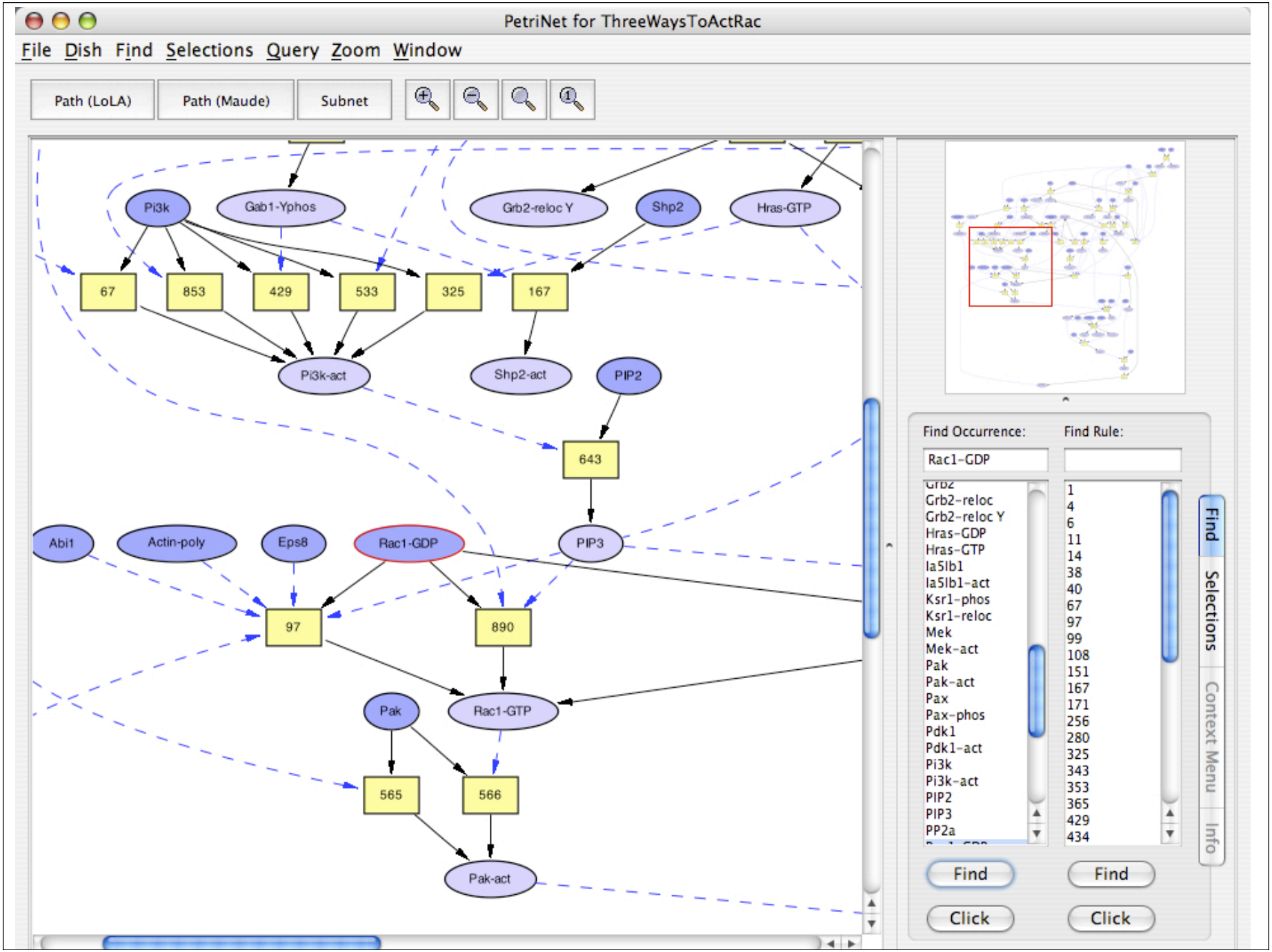
For example, given an initial state:

- 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?

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
- show knockouts
- display downstream impact of given components
- color pathway according to gene expression levels
- animate coloring of gene expression time series


PetriNet for ThreeWaysToActRac
File Dish Find Selections Query Zoom Window
Path (LoLA)
Path (Maude)
Subnet
Pi3k
Gab1-Yphos
Grb2-reloc Y
Shp2
Hras-GTP
67
853
429
533
325
167
Pi3k-act
Shp2-act
PIP2
643
Abi1
Actin-poly
Eps8
Rac1-GDP
PIP3
97
890
Pak
Rac1-GTP
565
566
Pak-act
Find Occurrence:
Find Rule:
Rac1-GDP
Grb2-reloc
Grb2-reloc Y
Hras-GDP
Hras-GTP
la5Ib1
la5Ib1-act
Ksr1-phos
Ksr1-reloc
Mek
Mek-act
Pak
Pak-act
Pax
Pax-phos
Pdk1
Pdk1-act
Pi3k
Pi3k-act
PIP2
PIP3
PP2a
1
4
6
11
14
38
40
67
97
99
108
151
167
171
256
280
325
343
353
365
429
434
Find
Selections
Context Menu
Info
Find
Find
Click
Click

Pathway Logic Basics

How is signaling representing 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 PrlR

```
rl[766.PrlR.by.Prl]:
    Prl
   [any:CellType | ct
     {CLo | clo}{CLm | clm PrlR}]
  =>
   [any:CellType | ct
     {CLo | clo [Prl - bound]}
     {CLm | clm [PrlR - act]}] .
  ------------------------------
  *** 11566606(R) PrlR is a homodimer
```

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

Example Rule: Phosphorylation of Cbl

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

Activated PrlR, 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 models each of 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

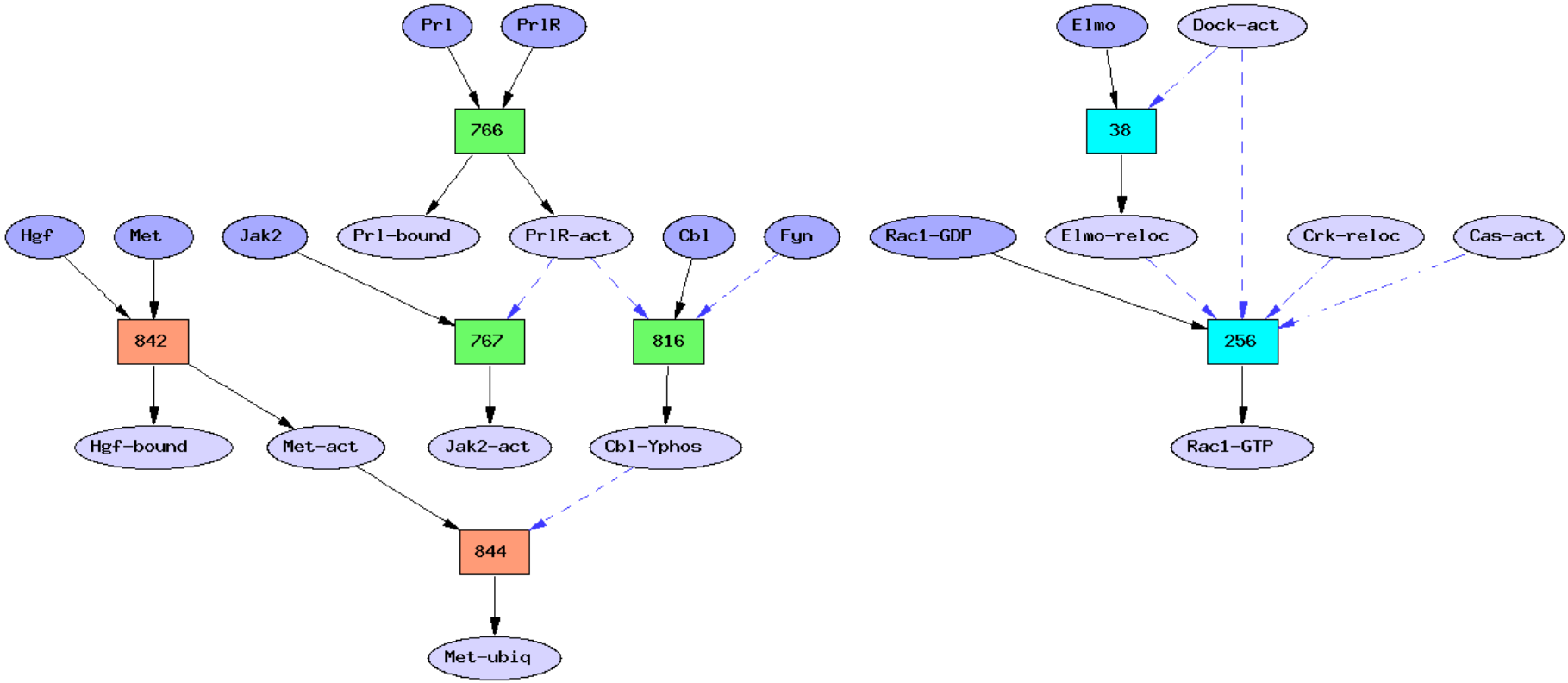


Figure 1: Caption

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.

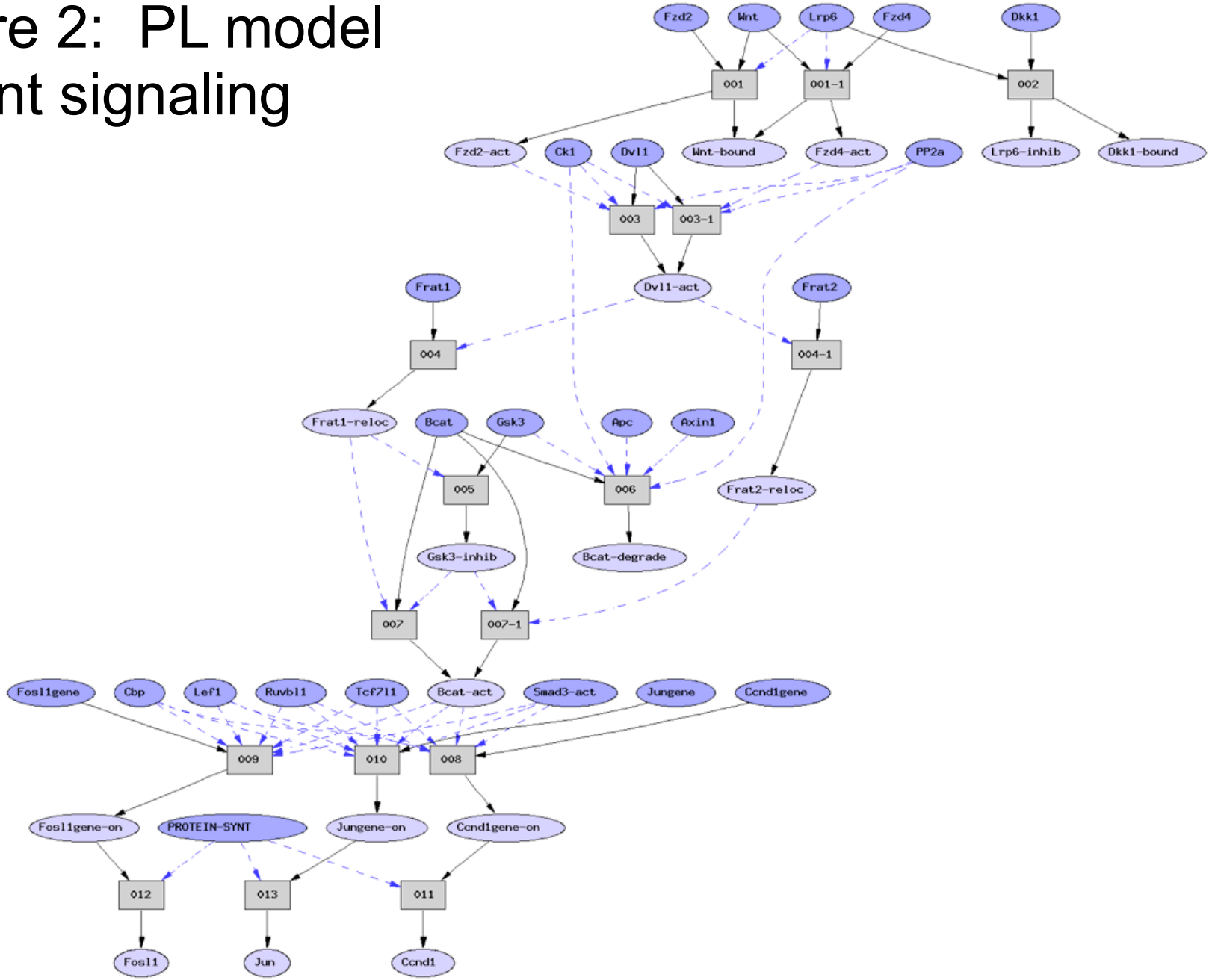
The Elmo cluster (cyan rules) contains rules related to Elmo and its role in activation of Rac1.

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.

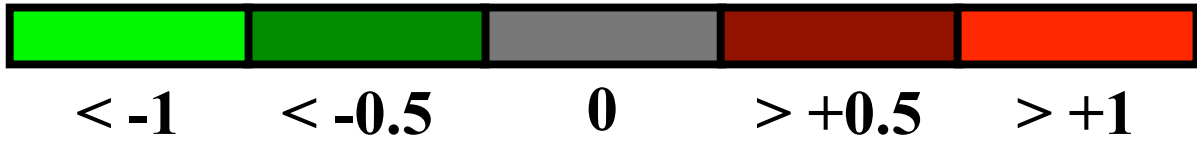
Each oval represents a protein (or other molecule) with a specific modification and cellular location.
Rectangles represent reactions (rules).
The dark ovals indicate components present in the initial state.

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.



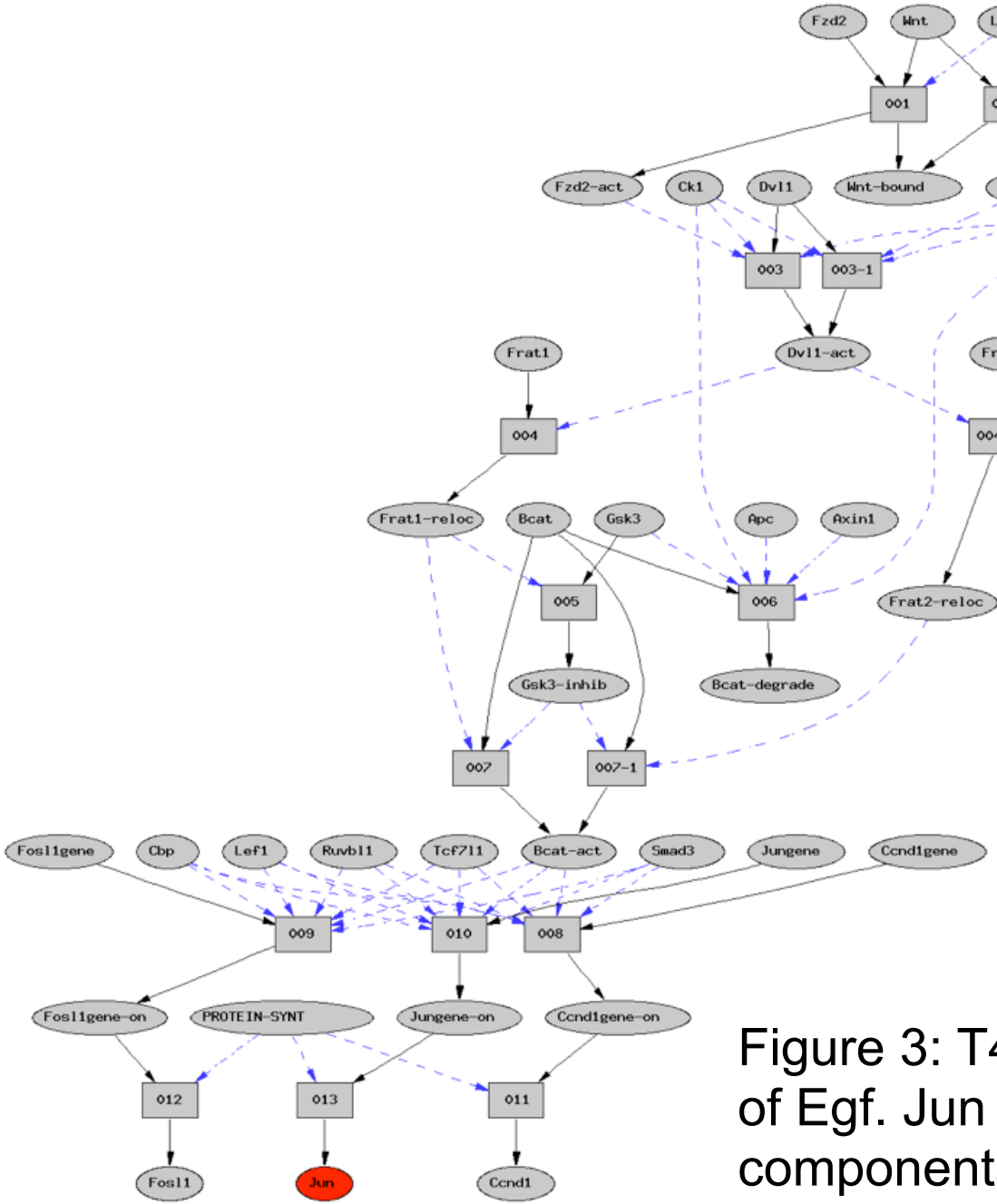


Figure 3: T47D Wnt pathway: 1 hour of Egf. Jun is upregulated (red), other components are unchanged (gray).

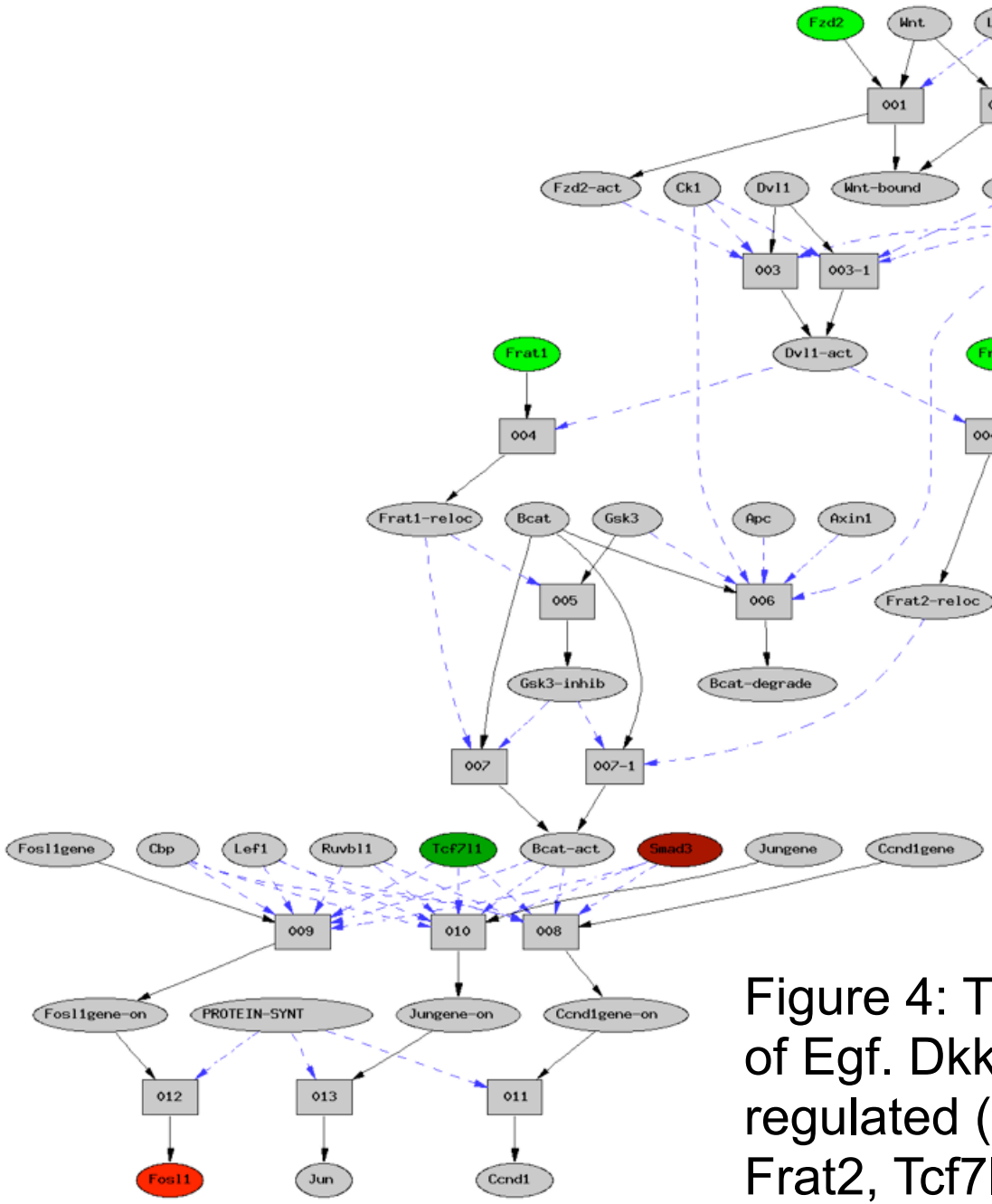


Figure 4: T47D Wnt pathway: 4 hours of Egf. Dkk1, Smad3, Fosl1 up-regulated (red); Fzd2, Fzd4, Frat1, Frat2, Tcf7l1 down-regulated (green).

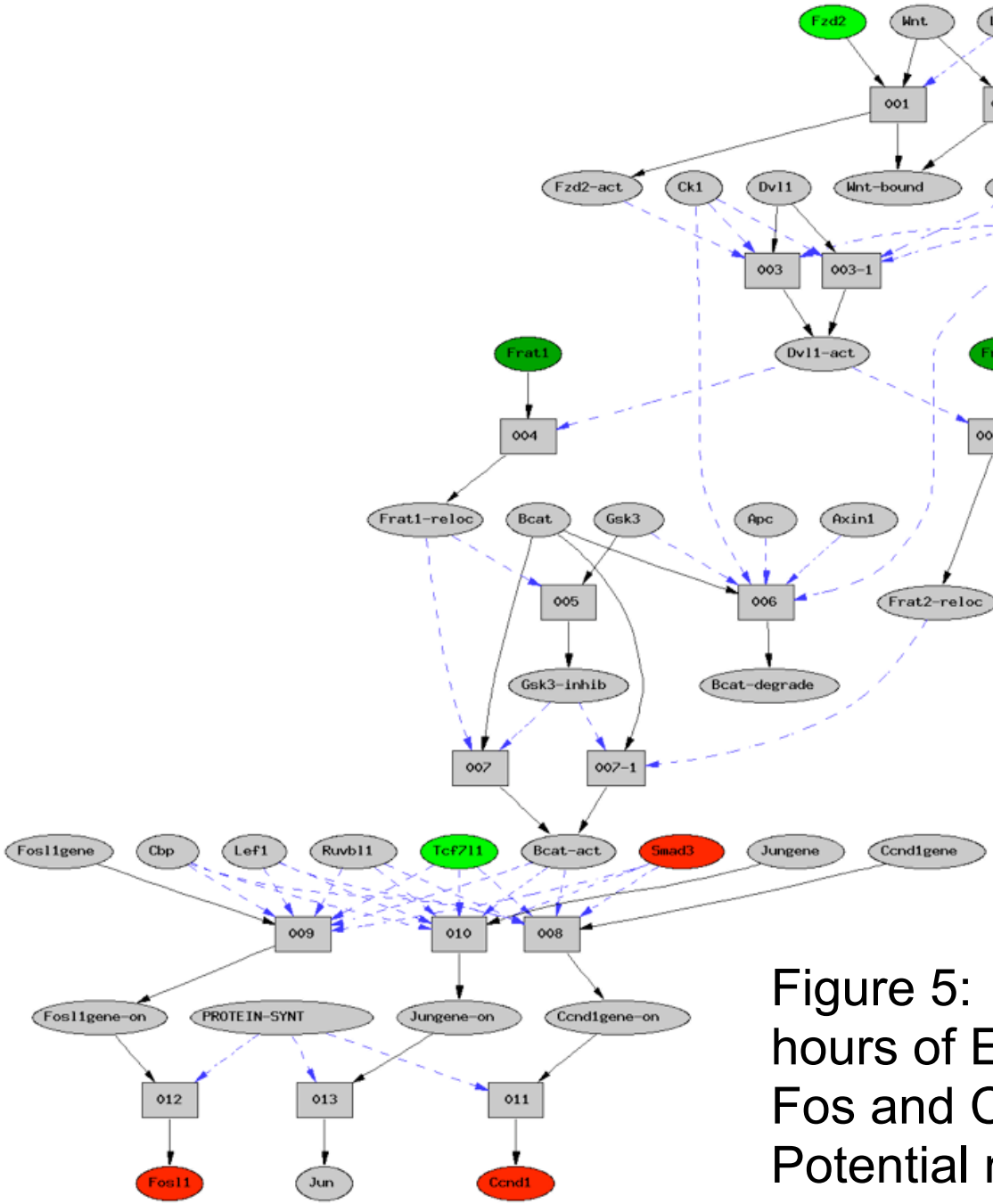


Figure 5: T47D Wnt pathway: 8 hours of Egf. Transcription factors, Fos and Ccnd1, are up-regulated. Potential new Wnt signaling effects.

References

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

[ICBP] LNBNL Integrative Cancer Biology Program http://icbp.lbl.gov

[Rewritiing 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.