Symbolic Systems Technology

Using formal systems and and reasoning tools to

- Understand how things work; why things don't work
- Design, analyze, adapt complex systems

Carolyn Talcott
ISSSB
2011 November 14

Plan

- What I mean by Symbolic Systems Biology
- Pathway Logic as an example
- Challenges & Opportunities

Elements of Symbolic Systems Biology

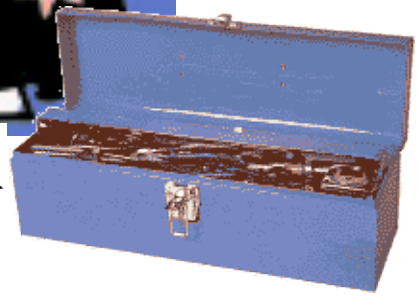
- Modeling and analysis of biological systems
- Representation in a formal system / logic
- Executable models
- Use of formal reasoning tools to answer questions

Modeling Issues

- What questions do you want the model answer?
- What can you observe/measure?
- What questions do you ***really*** want the model answer?
- What does that mean?
- Explain it to a computer!

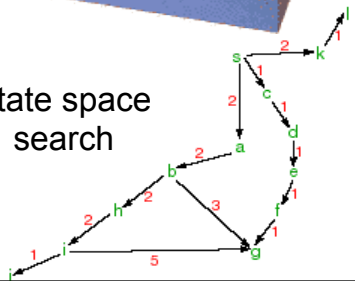
Modeling Methodology

rapid prototyping

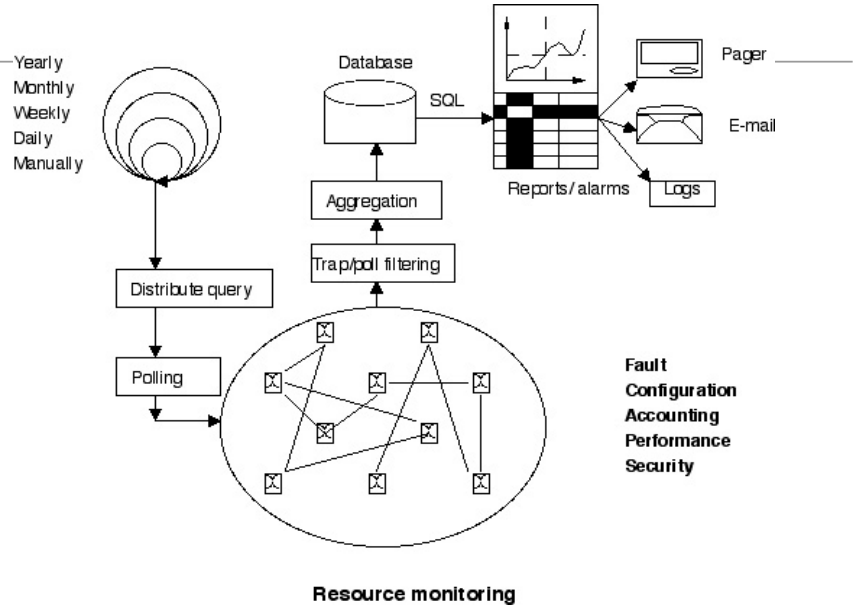


$$S \models \Phi$$

model checking



state space search



Resource monitoring

What is a formal system?

- Language: to describe things and properties
- Semantics: thing satisfies property
- Reasoning principles: proving/disproving properties of things
- [Reflection: to model and reason about models and reasoning]

Executable formal models

- Something to play with -- model train, architectural model ...
- Compter representation of
  - System state: collections of entities
    - each with name/id, location, knowledge, resources..
  - System behavior: transition rules, response to stimuli
- Execution: application of rules, possibly concurrent
  - watch it run
  - observe effects of perturbations
- Executions as objects that we can reason about
- Properties of states (P,Q) and executions (ϕ)
  - P until Q, eventually P,
  - this event depends on that event

Symbolic analysis -- answering questions

- Forward collection -- upper bound on possible states
- Backward collection -- initial states leading to states of interest
- Search -- for states of interest
- Model checking -- do all executions satisfy ϕ, find counter example
- Constraint solving -- steady state analysis

Model Transformations: new views, faster analysis

- Mapping of formal systems
  - access to tools (RWL to Petri Nets)
  - alternative semantics (qualitative, kinetic, probabilistic/stochastic)
  - richer property specification (RWL to HOL)
- Abstracting
  - quantities (points to regions)
  - qualities (phos(Y 123) vs phosY vs phos vs act/inact)
  - entities (protein vs family or homology class)
- Canonical form -- to combine/compare models

Pathway Logic (PL) http:pl.csl.sri.com

- Essence
- Examples
  - Response to Stimuli
  - Bacterial Protease Network
  - TB metabolic network
  - Sleep

In a nutshell

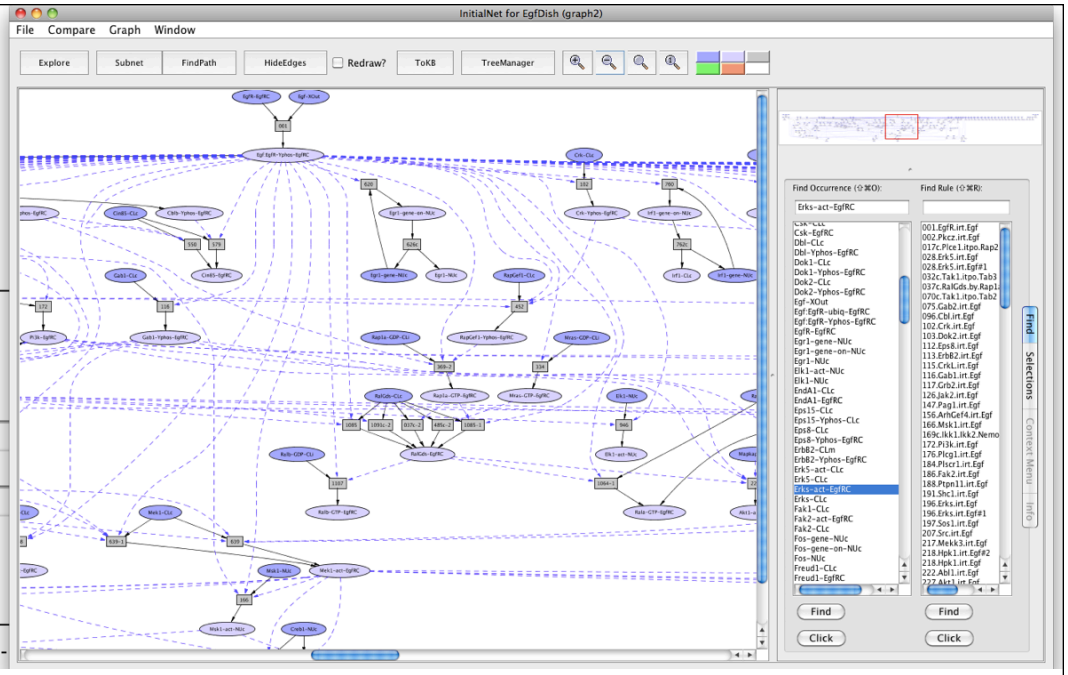
- Biological processes represented in rewriting logic to capture some detail, mapped to Petri nets for scalable model checking
- Queryable formal knowledge base of experimental findings (Datum KB)
  - subject, assay, treatment, times, cells, result, variants, source
- Algebraic signature ~ controlled vocabulary linked to standard sources (Uniprot ...).
- Terms represent reference entity with modifications, activity state, location
- Global model: network of executable rules (RKB) inferred from Datum KB
- Local Model ~ initial state (experimental setup) + relevant rules from RKB
- Pathways (sets of biomolecules that function together) and regulation effects discovered by asking questions (answered by formal reasoning tools)

Answering questions using the PL Assistant

- Given a model (dish and rules), specify goals and avoids
  - Find the relevant subnet
    - backward collection followed by forward collection
  - Find pathway
    - assert goal not achievable and ask LoLA, pathway is counter example
  - In silico knock outs
- Comparing networks
- Following connections
- Find all rule minimal pathways (separate tool, to be integrated)
  - cluster
  - ask for essential rules/occurrences, uses

Datum query

Signaling response to Egf (AMK)



Predicates

subject | entity | matches | Erk

assay | type | is a | ActivationAssay

Add a predicate Remove all predicates

50 of 407 Results

Export all results to txt or csv

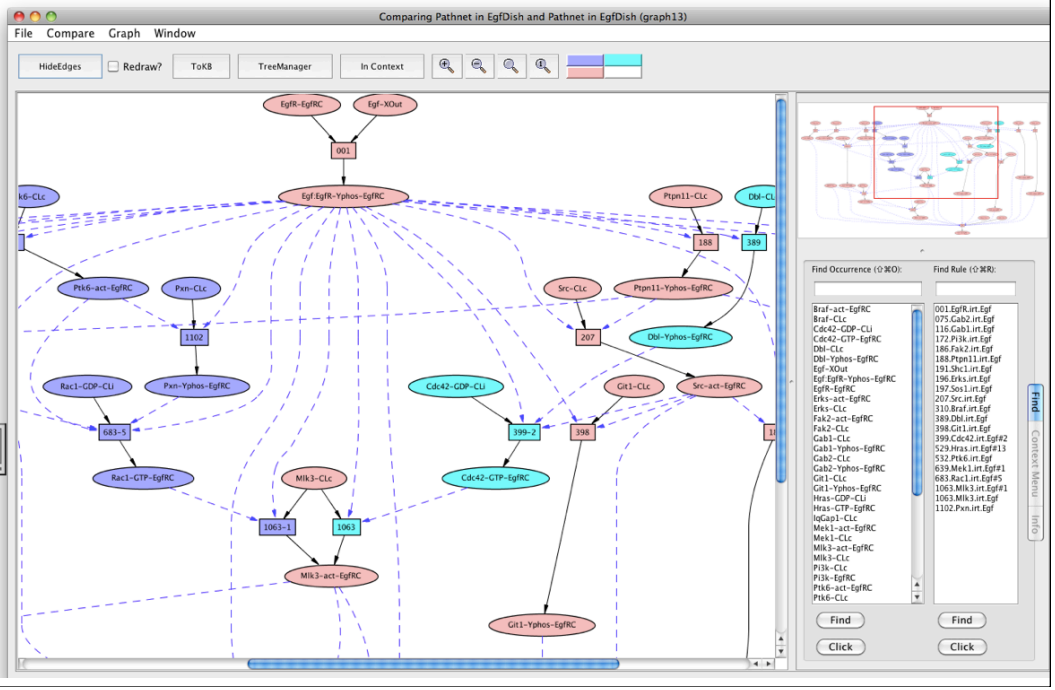
Page 1 of 9 Next

Expand All | Collapse All

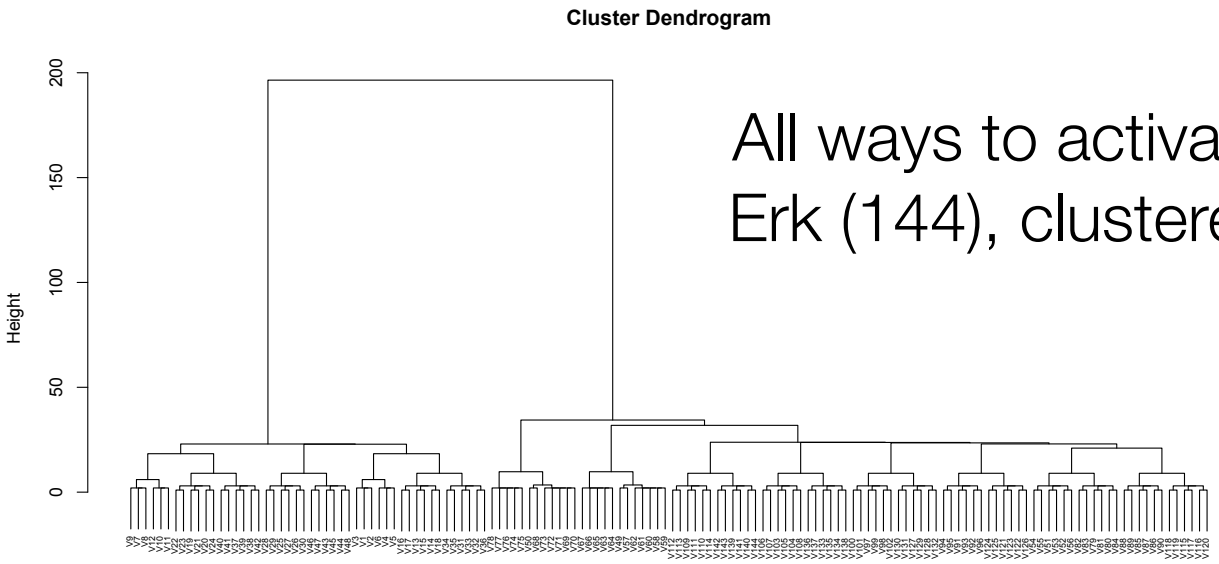
Erks[Ab]IP IVKA(MBP)[32P-ATP] is increased irt Lps (times) Expand

- cells: mPECs in BMS
- times: 0 7+ 15+ 30 60 90 min
- partially reqs: Irak1 [KO]
- source: 10754329-Fig-3c

Erks[Ab]IP IVKA(Elk1)[32P-ATP] is increased irt Lps (30 min) Expand

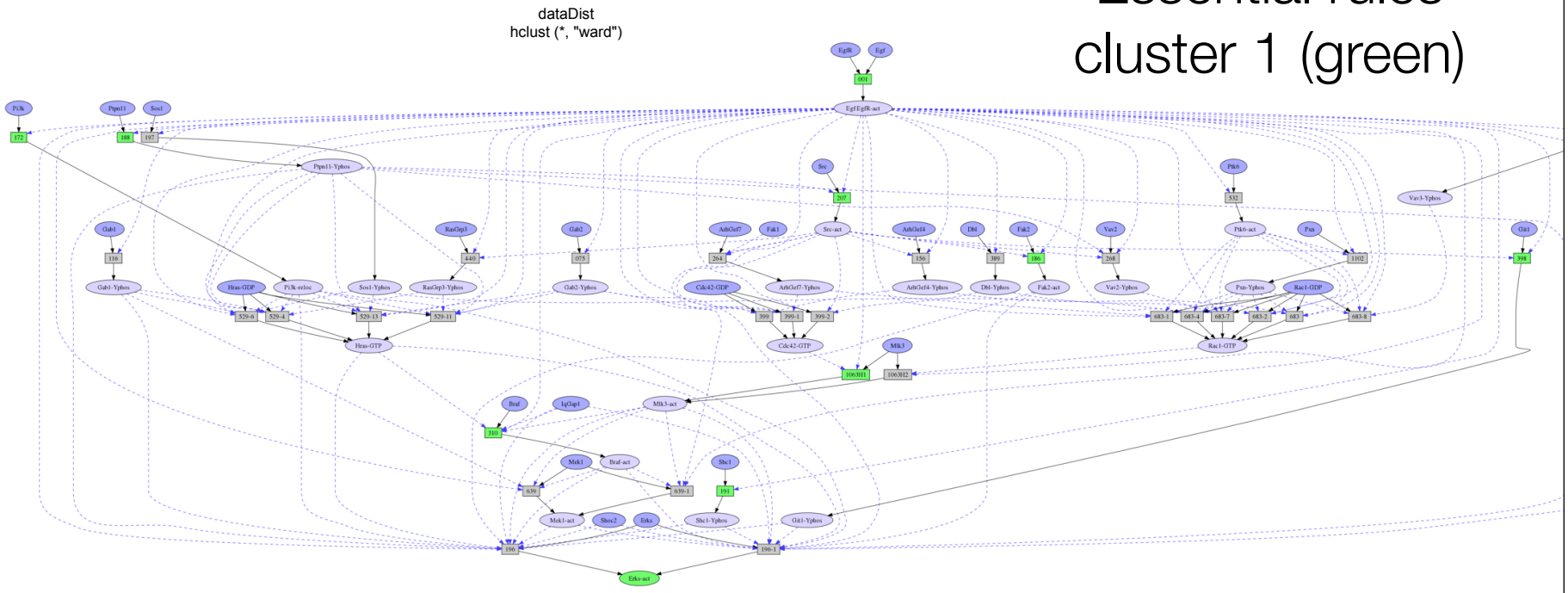


Two ways to activate Erk



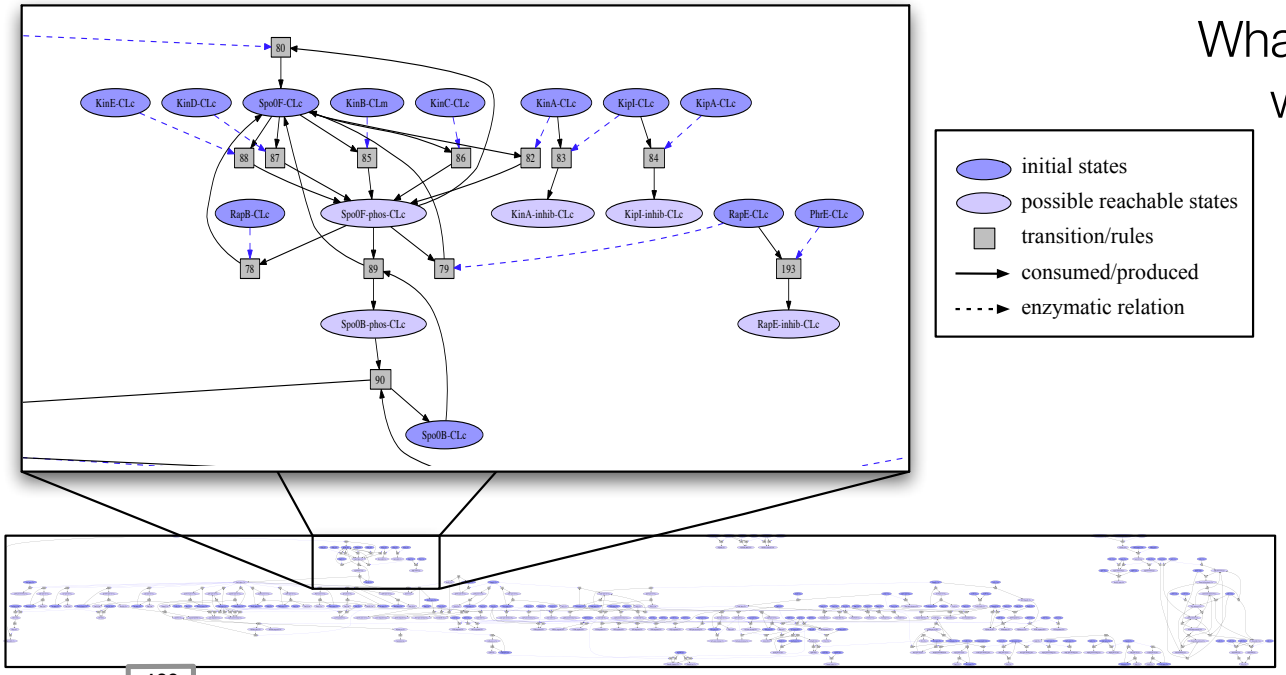
All ways to activate Erk (144), clustered

Essential rules cluster 1 (green)

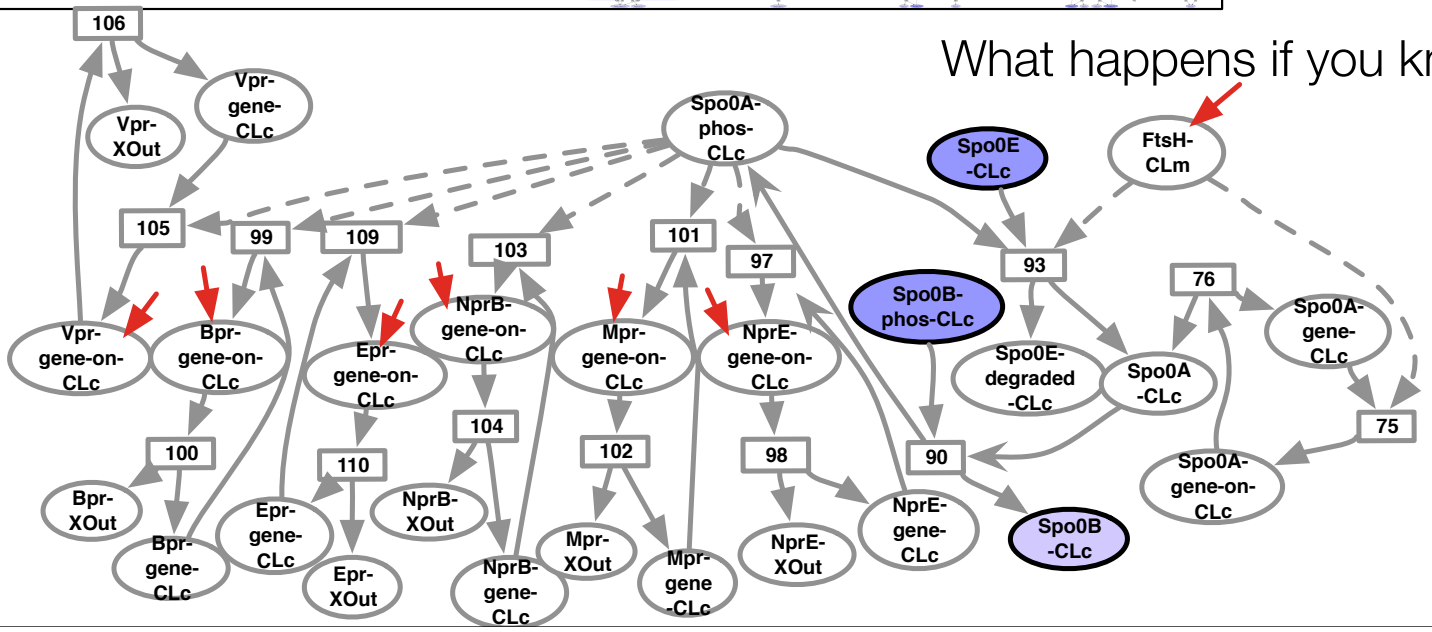


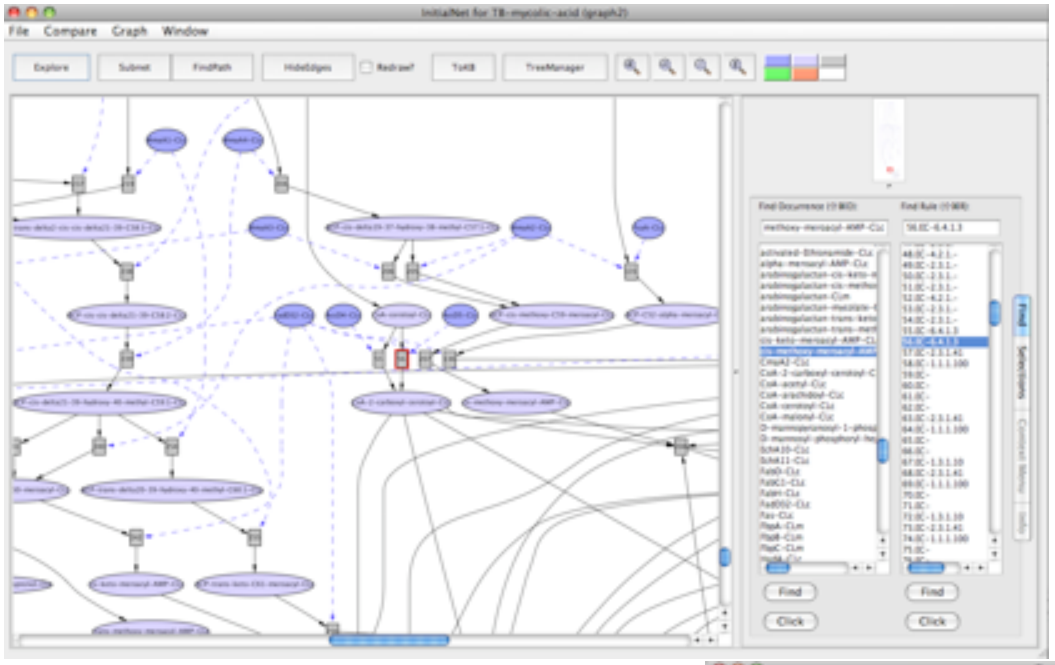
Protease interaction network (Gram Positive Bacteria) (Anupama)

What regulates a protease, what does it regulate



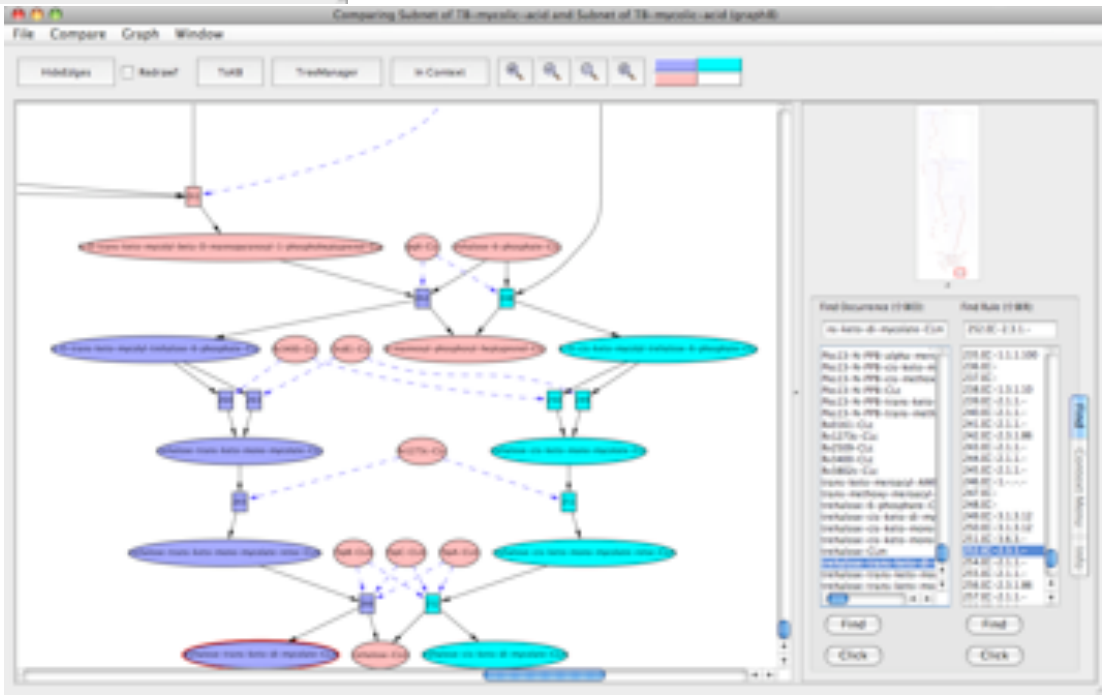
What happens if you knockout FtsH?



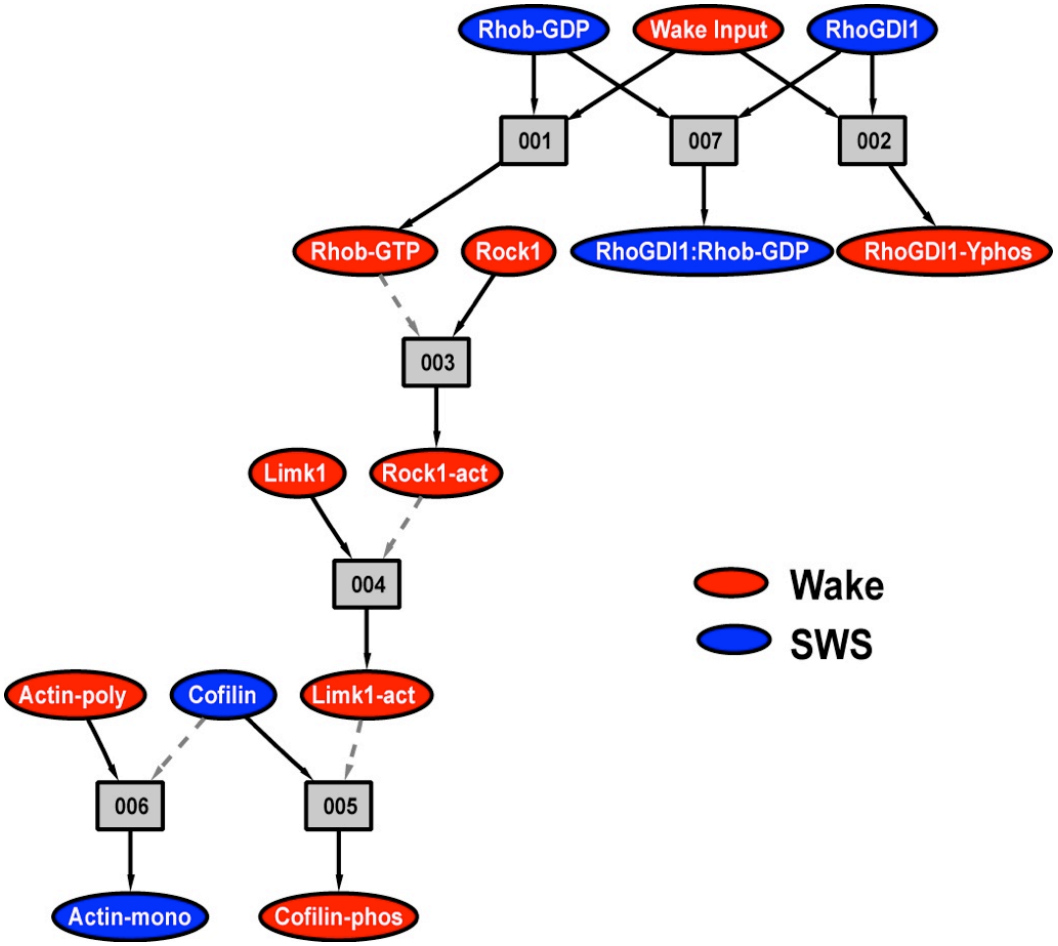


TB-mycolic-acid synthesis (Malabika)

Multiple synthesis routes



A HypotheticaL Model Pathway Relating State and Synaptic Plasticity



Wake state:
unknown signal(s)
=> phosphorylation of Rock1
=> activation of Limk1
=> phosphorylation of cofilin
=> increase in polymerized actin
(Phosphorylated cofilin is unable to depolymerize actin)

SWS:
RhoDG11 binds Rhob-GDP
(is not phosphorylated)
=> Rock1, Limk1, and cofillin would not be phosphorylated and
=> actin depolymerization
=> decrease in synaptic weight

Challenges and Opportunities

We need good models not just good analysis tools

- Alternative computational methods to check a prediction (example: proximity)
- Type checking: using basic biological knowledge
  - P has TS phosphorylation sites, Q is a tyrosine kinase
  - A rule that says Q phosphorylates P is either wrong or missing intermediate steps
- Better ways to capture experimental data.
  - Semantic data mining
  - Symbolic supplementary data
- Inferring rules from experimental data.
  - Needs representation of relevant biological knowledge
  - What do experiments tell you?
  - Roles -- adaptor, scaffold, transformer

Model Integration

- Aggregation -- different representation of the same process
    - PID EgfR signalling + Reactome EgfR signalling
- Complementation -- different parts of a response
- Making a model executable
    - Reactome pathways are disconnected -- missing rules, multitple levels of detail -- when are two entities the same?
- Mutiple computational models
    - multi scale
    - qualitative + quantitative

Managing Complexity

- Query of complex models may give complex answers (all paths)
  - How to visualize, make sense of the results
- Scaling
  - Divide and conquer
  - Property preserving abstractions
  - zooming in and out

Synergy

- Finding a shared representation of relevant biological knowledge
- Combining results from multiple analyses
    - how can our different tools work together?
- Biologist friendly scripting

The end! (The beginning?)