Modelling molecular evolution with process algebras

Marek Kwiatkowski
ETH Zürich & Eawag
marek.kwiatkowski@eawag.ch

8 June 2011
WCSB 2011, Zürich

(PhD work at the University of Edinburgh, supervised by Ian Stark)

Overview

1. Introduction and motivation
   - Some existing work
   - Towards a unifying framework
2. Modelling evolution of a signalling cascade
   - Process algebras for biology
   - The MAPK cascade and its model
   - Evolutionary setup
   - Fitness distributions and model backtracking
3. Conclusions

Some recent studies

- A. Wagner Does evolutionary plasticity evolve? Evolution **50**, 1996.
- M. Siegal and A. Bergman Waddington's canalization revisited: Developmental stability and evolution. PNAS **99**, 2002.
- A. Bergman and M. Siegal Evolutionary capacitance as a general feature of complex gene networks. Nature **424**, 2003.
- O. Soyer *et. al.* Signal transduction networks: Topology, response, and biochemical reactions. J. Theor. Biol. **238**, 2006.
- O. Soyer and S. Bonhoeffer Evolution of complexity in signalling pathways. PNAS **103**, 2006.
- L. Dematté *et. al.* Evolving BlenX programs to simulate the evolution of biological networks. Theor. Comput. Sci. **408**, 2008.
- E. Borenstein and D. Krakauer An end to endless forms: epistasis, phenotype distribution bias, and non-uniform evolution. PLoS Comp. Biol. **4**, 2008.

Common theme

models $\equiv$ genotypes, execution $\equiv$ development, results $\equiv$ phenotypes.

Towards a unifying framework

Just like systems biology has benefited from SBML, evolutionary systems biology could benefit from a standard specification and modelling format. Ideally, it should:

1. Be agent-centric, not reaction-centric,
2. Support dynamic complex formation,
3. Have deterministic primary dynamics, but
4. Admit a variety of execution modes.

In what follows we introduce and evaluate such a prototype framework.

M. Kwiatkowski A formal computational framework for the study of molecular evolution. Ph.D. thesis, The University of Edinburgh, 2010.

M. Kwiatkowski and I. Stark On executable models of molecular evolution. WCSB 2011.

Process algebra and biology

Process algebras are, loosely speaking, idealised programming languages with a focus on parallel computing. They have been used to model biochemical networks since ca. 1999.

Define:

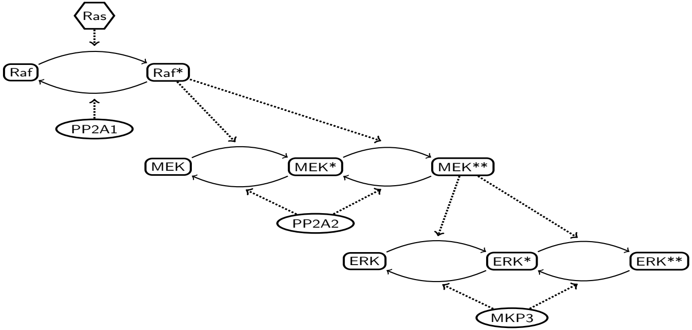
$$A \triangleq a.(A_1|A_2)$$
$$B \triangleq b.B$$

Compute: $A \mid B = a.(A_1|A_2) \mid b.B \longrightarrow A_1 \mid A_2 \mid B$

Benefits: formality, parsimony, compositionality, abstraction.

A. Regev and E. Shapiro Cellular abstractions: cells as computations. Nature **419**, 2002

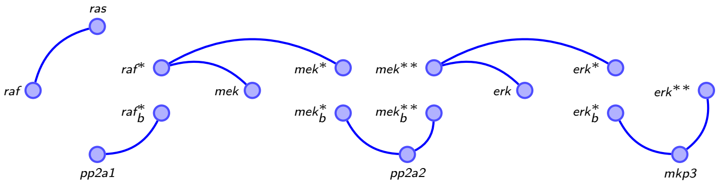
Case study: the MAPK cascade (1)



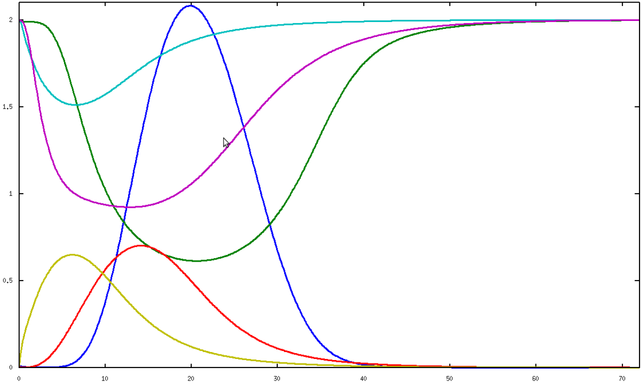
- Functionally conserved in most animals
- Crucial component of many signal transduction pathways
- Relays and amplifies the signal efficiently
- Benchmark for new modelling techniques

Case study: the MAPK cascade (2)

$$
\begin{aligned}
Ras &\triangleq (\boldsymbol{\nu} x — \overline{x})ras(x;y).(\overline{x}.Ras + y.Ras)\\
Raf &\triangleq (\boldsymbol{\nu} x — \overline{x})raf(x;y).(\overline{x}.Raf + y.Raf^*)\\
Raf^* &\triangleq (\boldsymbol{\nu} x — \overline{x})(\boldsymbol{\nu} z — \overline{z})(raf^*(x;y).(\overline{x}.Raf^* + y.Raf^*) + raf_b^*(z;y).(\overline{z}.Raf^* + y.Raf))\\
PP2A1 &\triangleq (\boldsymbol{\nu} x — \overline{x})pp2a1(x;y).(\overline{x}.PP2A1 + y.PP2A1)\\
MEK &\triangleq (\boldsymbol{\nu} x — \overline{x})mek(x;y).(\overline{x}.MEK + y.MEK^*)\\
MEK^* &\triangleq (\boldsymbol{\nu} x — \overline{x})(\boldsymbol{\nu} z — \overline{z})(mek^*(x;y).(\overline{x}.MEK^* + y.MEK^{**}) + mek_b^*(z;y).(\overline{z}.MEK^{**} + y.MEK^*))\\
MEK^{**} &\triangleq (\boldsymbol{\nu} x — \overline{x})(\boldsymbol{\nu} z — \overline{z})(mek^{**}(x;y).(\overline{x}.MEK^{**} + y.MEK^{**}) + mek_b^{**}(z;y).(\overline{z}.MEK^{**} + y.MEK^*))\\
PP2A2 &\triangleq (\boldsymbol{\nu} x — \overline{x})pp2a2(x;y).(\overline{x}.PP2A2 + y.PP2A2)\\
ERK &\triangleq (\boldsymbol{\nu} x — \overline{x})erk(x;y).(\overline{x}.ERK + y.ERK^*)\\
ERK^* &\triangleq (\boldsymbol{\nu} x — \overline{x})(\boldsymbol{\nu} z — \overline{z})(erk^*(x;y).(\overline{x}.ERK^* + y.ERK^{**}) + erk_b^*(z;y).(\overline{z}.ERK^{**} + y.ERK^*))\\
ERK^{**} &\triangleq (\boldsymbol{\nu} x — \overline{x})erk_b^{**}(x;y).(\overline{x}.ERK^{**} + y.ERK^*)\\
MKP3 &\triangleq (\boldsymbol{\nu} x — \overline{x})mkp3(x;y).(\overline{x}.MKP3 + y.MKP3)\\
\boldsymbol{\Pi} &\triangleq \mathbf{c_1 \cdot Raf \mid\mid c_2 \cdot Ras \mid\mid c_3 \cdot MEK \mid\mid c_4 \cdot ERK \mid\mid c_5 \cdot PP2A1 \mid\mid c_6 \cdot PP2A2 \mid\mid c_7 \cdot MKP3}
\end{aligned}
$$

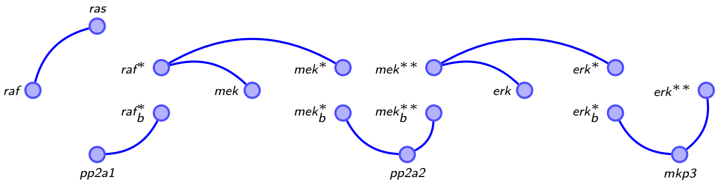


Case study: the MAPK cascade (3)



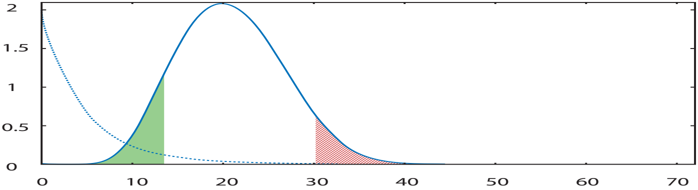
Twenty-three differential equations extracted from the $c\pi$ model and solved with Octave. Emergent Michaelis-Menten kinetics for every reaction.

Evolutionary analysis of the MAPK cascade: the plan



- Reconfigure every site in every way possible (ca. 1M variants)
- Find evolutionarily fragile and robust sites
- Compute the fitness of every variant using signal integration
- Find the distribution of mutation effects on fitness

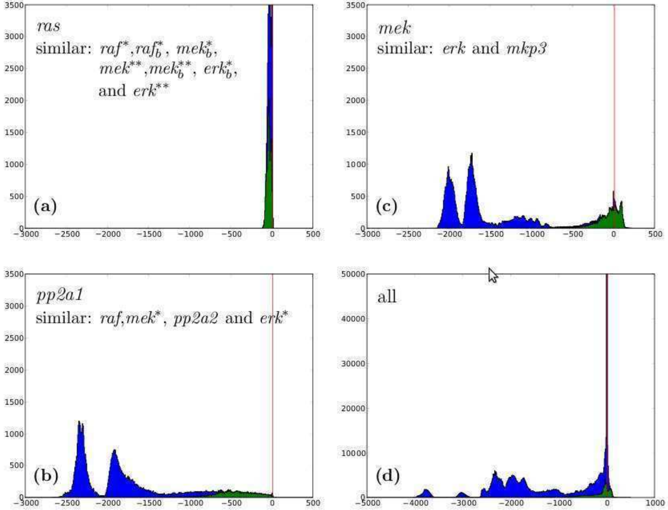
Evolutionary analysis of the MAPK cascade: fitness function



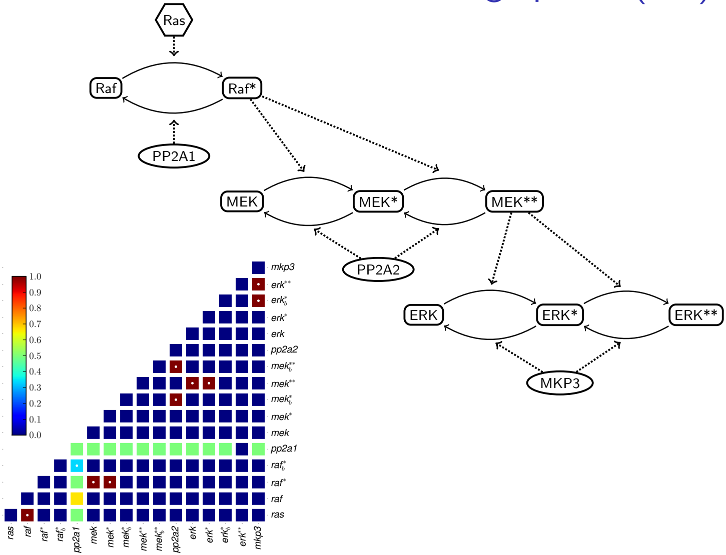
- Rewards fast and strong response (green area)
- Punishes incomplete switching-off (red area)

L. Dematté *et. al.* Evolving BlenX programs to simulate the evolution of biological networks. Theor. Comput. Sci. **408**, 2008.

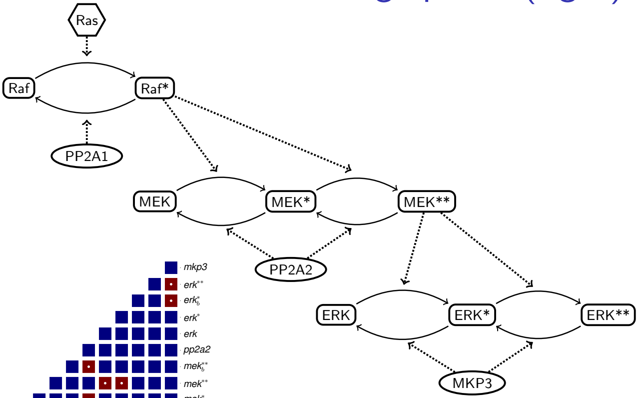
Evolutionary analysis of the MAPK cascade: fitness distributions

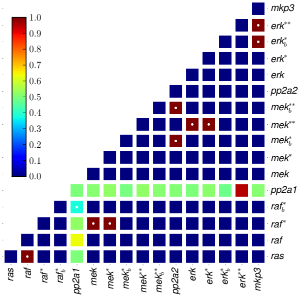


Evolutionary analysis of the MAPK cascade: two strange peaks (left)

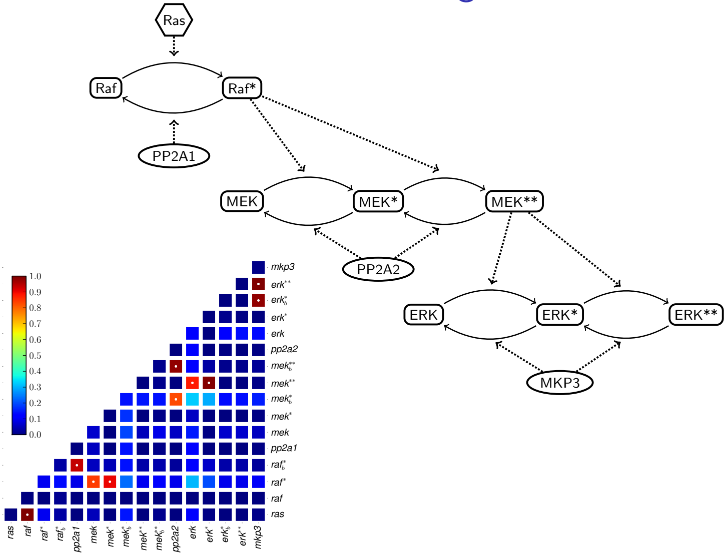


Evolutionary analysis of the MAPK cascade: two strange peaks (right)





Evolutionary analysis of the MAPK cascade: advantageous mutations



Conclusions