BMC Systems Biology

Poster presentation

Bayesian inference of the kinetic parameters of a realistic MAPK/ERK pathway

Vilda Purutçuoğlu* and Ernst Wit

Address: Department of Mathematics and Statistics, Lancaster University, Lancaster, LA1 4YF, UK

Email: Vilda Purutçuoğlu* - v.purutcuoglu@lancaster.ac.uk

* Corresponding author

from BioSysBio 2007: Systems Biology, Bioinformatics and Synthetic Biology
Manchester, UK. 11–13 January 2007

Published: 8 May 2007

BMC Systems Biology 2007, **1**(Suppl 1):P19 doi:10.1186/1752-0509-1-S1-P19

This abstract is available from: http://www.biomedcentral.com/1752-0509/1?issue=S1

Background

All cellular activations are regulated by various signal transduction pathways, which are the network of interacting proteins used to carry over signals in the cell's environment for producing associate responses. The MAPK (mitogen-activated protein kinase) or its synonymous ERK (extracellular signal regulated kinase) pathway is one of the major signal transduction systems which regulates the cellular growth control of all eukaryotes like cell proliferation or apoptosis. The complex structure of this regulatory mechanism whose main components are Ras, Raf, and MEK proteins (see Figure 1) includes a number of phosphorylations on the protein level. The functionality of these proteins is stochastic in nature and directed by positive and negative feedback loops that cause either activation or inhibition of other proteins.

Due to the importance in the cellular lifecycle, the MAPK/ERK pathway has been intensively studied, thereby a number of qualitative descriptions of this regulatory mechanism are available in the literature. However none of the sources describe the system by an explicit set of reactions. Here we combine these qualitative sources for a representation of the pathway as a list of (quasi) reactions which is used to produce a basis for stochastic simulation. For defining our reaction set we denote all components by simple notations and use multiple parametrizations to indicate different localization of the molecules in the cell and to describe the protein using different binding sites as well as various phosphorylations.

Modelling by diffusion approximation

Gene regulation is commonly modelled via ordinary differential equations (ODEs). Although ODEs are successful to represent some reactions like linear production and degradation, they cannot describe the small system variability of the actual reactions. For biochemical systems, stochastic processes are a natural choice as these kinds of dynamic formalization take into account the probabilistic manner of the different biological activations. In this study under the assumption that the probability distribution of the number of the molecules of each species at t depends on the continuous t and continuous number of molecules, we use the diffusion approximation to explain the change of state of each substrate at t. In this modelling the current state is found by a Langevin approach, where a correlated noise term describes the stochastic behaviour of the model over and above the drift term via $dY(t) = \mu(Y, \Theta)dt + \beta^{1/2}(Y, \Theta)dW(t)$ in which $dW(t)$ is a s-dimensional vector representing the change of a Brownian motion over time and s is the total number of substrates in the system. $\mu(Y, \Theta) = V'a(Y_t, \Theta)$ and $\beta(Y, \Theta) = V'diag\{a(Y_t, \Theta)\}V$ are mean, or drift, and variance, or diffusion, matrices, respectively, both depending on the state of the system Y at time t, and the parameter vector $\Theta = (\Theta_1,...,\Theta_r)'$ explicitly. $\Theta_j (j = 1,...,r)$ represents the stochastic rate constant of the j th reaction and r denotes the total number of reaction. Accordingly V is the net effect matrix and r-dimensional vector $a(Y_t, \Theta)$ describes the hazard of each reaction at time t. The algorithm computes the next state at $t + dt$ by replacing $Y(t)$ by $Y(t) + dY(t)$.

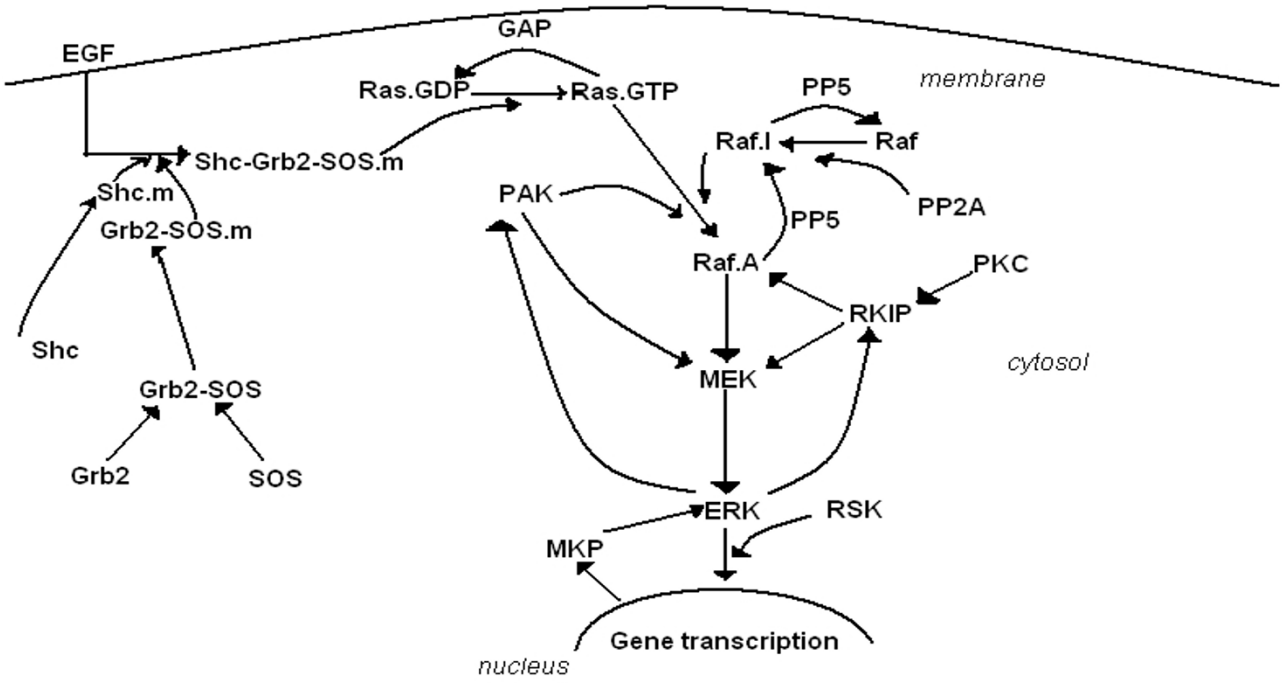

Figure 1
Simple representation of the structure of MAPK/ERK pathway.

Diffusion approximation for inference

For estimating the model parameters, i.e. the stochastic rate constants, we apply the discretized version of diffusion approximation, which is known as Euler approximation, $\Delta Y_t = \mu(Y_t, \Theta)\Delta t + \beta^{1/2}(Y_t, \Theta)\Delta W_t$, where ΔW_t shows a s-dimensional independent identically distributed $N(0, I\Delta t)$ random vector. We define our data vector as a $(n + 1) \times s$ matrix in which each column indicates a vector of $Y_i = (Y_{t_0,i}, ..., Y_{t_n,i})$ and n stands for the total number of observed time step. Finally I is the indicator of the i th substrate. Since the change in state for a given Δt has a multivariate normal distribution, the likelihood associated with this time increment is derived proportional to

$$L(Y \mid \Theta) \propto \left\{ \prod_{i=0}^{n-1} \left| \beta(Y_{t_i}, \Theta) \right|^{-1/2} \right\} \times \exp\left\{ -\frac{1}{2} \sum_{i=0}^{n-1} (\Delta Y_{t_i} - \mu(Y_{t_i}, \Theta)\Delta t_i)' \left| \beta(Y_{t_i}, \Theta)\Delta t_i \right|^{-1} (\Delta Y_{t_i} - \mu(Y_{t_i}, \Theta)\Delta t_i) \right\} \quad (1)$$

where Y_{t_i} shows the state of the i th substrate at time t and $\Delta Y_t = Y_{t+\Delta t} - Y_t$. As can be seen from equation 1, the conditional posterior density of reaction rates Θ does not have a known distribution. We compute the posterior distribution of Θ using the MCMC method. Moreover to decrease the bias causing by discretization we augment our observations by putting extra time states between given measurements. Then conditional on accepted Θ, we simulate and update the missing states by implementing the Metropolis-Hastings algorithm as one block of $\hat{Y}$ at a time. On simulated data we observe that the sampler converges well and is able to identify the dynamics of the MAPK/ERK pathway.

References

1. Golightly A, Wilkinson DJ: **Bayesian inference for stochastic kinetic models using a diffusion approximation.** *Biometrics* 2005, **61(3):**781-788.
2. Kolch W, Calder M, Gilbert D: **When kinases meet mathematics: the systems biology of MAPK signalling.** *FEBS Lett* 2005, **579:**1891-1895.