

Review of the computational models for the signaling pathway

Meijing Kong^a, Chao Wang^b and Le Zhang^{c,*}

Computer and Information Science, Southwest University, Chongqing, China

^aa369@swu.edu.cn, ^bwachest@swu.edu.cn, ^czhanglcq@swu.edu.cn

Keywords: signaling pathway; computational models ; systems biology

Abstract. In the past few years, systems biology methods have been widely used to explore the mechanism of cancer. Various signaling pathways play the important role in cancer. And a large number of computational methods are used to model the signaling pathway. Here, we reviewed many existing mechanisms and modeling techniques, which are used to model different signaling pathways. Then we introduced the systematic workflow of a system biology approach and some fundamental analysis methods. Finally we discuss some known challenges in the process of modeling signaling pathways and what we should handle in our future work.

1 Introduction

Signaling pathway is usually used to explain important internal cellular communication, the change of cell environment, integration of external or internal and to respond the change of signal and the transcription activity, metabolism, or other regulatory measures[1]. They are responsible for regulating critical cellular processes, such as autophagy, apoptosis, transcription, cell cycle progression and proliferation. Abnormal signaling pathways may cause to the diseases. For instance, protein tyrosine kinase (PTK) plays an important role in intracellular signal transduction pathways, it can affect the cell growth, proliferation and differentiation and the disturbances of the inhibitory constraints on kinase activity can lead to tumorigenic PTK signaling[2]. Moreover, the p38 mitogen-activated protein kinase (p38 MAPK) is the main signal molecule of signal transduction process, and plays an important role in the process of development and disease[3, 4].

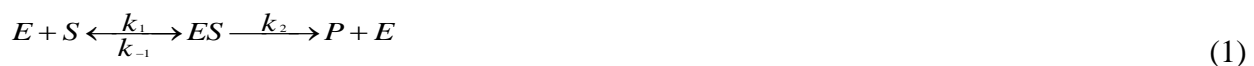
For purpose of understanding the complicated dynamic behavior of signaling pathways, researchers have put forward to models from abstract models that underline some key features of signaling pathways [5, 6] to detailed models that describe the dynamics of signaling pathways in specific organisms[7-10]. Computational models and methods are becoming an integral part of molecular biology. They are being used not only to identify cellular components, but also to determine how these components interact with one another[11].

This article will describes the modeling of signaling pathways and summarize several mathematical techniques.

2 Existing Computational models of signaling pathway

Ordinary Differential Equation (ODE) models

The enzymatic reaction products of many signal cascade consisting of a series of signal transduction process acts as the next reaction catalytic enzyme. Because the biochemical reaction mapped into concentration change with time, we usually use ODE to describe the reactions in signal transduction pathways. The reaction (Eq. 1) can be described that the enzyme E reacts with substrate S to yield Enzyme-Substrate complex ES which in turn yields product P.



Where k_1 , k_{-1} and k_2 are rate constants. The double arrow (\rightleftharpoons) means that the reaction is reversible, and the single arrow (\rightarrow) means that reaction can occur in one direction.

According to the Law of Mass Action[12], the rate of a reaction is relative to the product of the concentration of the reactants. By the Law of Mass Action, a group of nonlinear ordinary differential equations are used to describe each reactant.

$$\frac{d[S]}{dt} = k_1[E][S] + k_{-1}[ES] \quad (2)$$

$$\frac{d[E]}{dt} = -k_1[E][S] + (k_{-1} + k_2)[ES] \quad (3)$$

$$\frac{d[ES]}{dt} = k_1[E][S] - (k_{-1} + k_2)[ES] \quad (4)$$

$$\frac{d[P]}{dt} = k_2[ES] \quad (5)$$

Where [S], [E] and [ES] denote the concentration of enzyme, substrate and Enzyme-Substrate complex, respectively.

Recently, significant progress has been made in the area of modeling for better understanding of the biological behavior of the cell signaling pathways [13-16]. Several research groups have used ODE to analyze the dynamics of signaling networks and generate experimentally testable predictions [17-25]. To predict the dynamics of the IGFR network after IGF-1 stimulation in MDA-MB231 cells, Iadevaia et al.,[26] developed a mass action ODE model consisting of 65 ODEs and 161 model unknowns parameter. Chen's mass action model use 499 ODEs to describe dynamics of 828 reactions and 229 parameters[27] to reveal input-output behavior of ErbB signaling pathway. Hoffmann model consisted of 24 ODEs that described the reaction kinetics equations of the change in concentration [28]. The model also has 30 parameters, which come from experimental data, literature and estimation.

In general, the ODE models have been proved to have certain ability to predict and regulatory mechanism underlying normal and abnormal signaling [29-31]. However, it is hard to employ ODEs to describe the large signaling pathway and difficult to solve these ODEs due to the following reasons. First, the initial conditions and constants are sensitive for many mutual differential equations. Second, it may not be appropriately considered for the trivial numerical values of concentrations of species. Third, the pathway behavior can be affected by the time delays in certain cellular processes, such as transcription[32].

Hybrid Functional Petri Net (HFPN) Models

Petri net is a family of powerful discrete event models in parallel with the development of the theory of discrete event systems[33] and More formal definitions are given in[34-36]. For example, Sackmann et al., [37] proposed systematic model that used qualitative Petri net to analyzes the signaling pathways [33]. Miwa et al., [38] developed a discrete Petri net model for the IL-1 signaling pathway simulation and determined the firing frequency of each transition by applying the proposed method for a part of the Petri net model of IL-1 signaling pathway. Hybrid Functional Petri Net was firstly developed by Matsuno et.al.,[39] to represent and simulate biopathways. Since this study extended the notions of hybrid Petri net[40], functional Petri net[41] and hybrid object net [42], it is suited for biopathways modeling[39]. Moreover, Koh et al.,[citation] [11] employed HFPN to explore the Akt and MAPK pathways and their crosstalk. In conclusion, time-less discrete Petri nets combine an intuitive modelling style with well-founded analysis techniques. And structural analysis of Petri nets has indicated significant insights into mechanisms for identifying interrelated targetable node sets for interventional treatment [31].

Agent-Based Models(ABM)

According [43],agent-based modelling is defined as the combination of single entity which is individual component of a system obeying its own pre-defended rules and reacting to its environment and neighboring agents. In agent-based modeling (ABM), each entity is considered to be an agent in a multi agent system (MAS) that interacts locally with its neighbors as well as the environment [44-46]. For example, Shirazi et.al., [46] presented a multi-agent approach to model the MAPK signaling pathway, which defines each substrate in a signaling pathway as an

independent entity (agent). And Pogson et al., [43] employ ABM to demonstrate that it is a suitable to describe the cellular regulatory events such as the $NF\kappa B$ pathway.

As mentioned above, the formal agent-based modelling paradigm have demonstrated useful strategy for understanding mechanisms of biological systems. Agents provides a powerful framework for more detailed modelling of intracellular signaling pathways [43]. And the ABM has greater scope than the ODE, but it is hard to integrate with experimental data to estimate the key parameters of ABM.

Modularization Models

For the large pathway, it is difficult to estimate the values of the key unknown parameters. The common approach to obtain parameters values is from the real experimental measurements and the published literature[25]. Unfortunately, since most of the parameters reported in the literatures are dependent on experimental conditions, it is necessary to decompose the large pathway into several small and independent components for parameter estimation [47-50]. For example, Koh et al., [11] developed a decomposition approach for the parameter estimation for HFPN pathway modeling by decomposing the HFPN pathway into 6 components, each of which has no more than 25 unknown parameter. On the other hand, Peng et al.,[4] proposed a new systems biology approach by integrating quantitative LC-MS(liquid chromatography–mass spectrometry)approach and RPPA(reverse phase protein array) experimental Data into the computational pathway model and estimate the unknown parameters by modularized factor graph.

3 Method

We summarize the signaling pathway modeling procedure as three steps: (1) modeling for signaling pathway, (2) the parameter estimation, and (3) the parameter analysis. The detail description is in the following.

A. Signaling pathway modeling

Model signaling pathway can employ experimental data to put forward the biological hypotheses and use the model to show the biological complex behavior. Usually, we develop the mathematical model with respect to the structure of the experimental data.

B. Parameter estimation algorithm

Parameter estimation is a critical problem in the biological pathways modeling. It is difficult to estimate so many parameters in a large pathway for the limited experimental data. Eq.6 can be employed for parameter estimation by minimizing the fitness error between the experimental data and simulated data.

$$\theta^* = \arg \min \sum_{i=1}^M \sum_{j=1}^N w_i (x_i^{sim}(t_j, \theta) - x_i^{exp}(t_j))^2 \quad (6)$$

Where θ^* denotes the objective function of the parameter optimization, $w_i = 1/(\max x_i^{exp}(t_j))^2$, $x_i^{sim}(t_j, \theta)$ denotes the protein concentration by time series data obtained by ODE equation simulation, $x_i^{exp}(t_j)$ denotes the protein concentration by time series data obtained by experiments, i represents the protein index, j represents the time point; M represents the number of proteins and N represents the number of time points.

C. Parameter analysis

Identifiability analysis

The identifiability analysis applied the coefficient of variation to confirm whether the parameter is identifiable or not [51, 52]. The ratio of the standard deviation to the mean of the estimated values is defined as the coefficient of variation[52].If the coefficient of variation for the given parameter is smaller than 1, then it is identifiable; otherwise it is unidentifiable.

Sensitivity analysis

Parameter sensitivity analysis is a tool to quantitatively determine the effect that specific parameters on the output. To understand the relationship between system responses and variations in individual model parameter values, local parameter sensitivity analysis was performed by the following equation[53]:

$$S_i = \frac{\partial ['protein']}{\partial V_i} \bigg/ \frac{['protein']}{V_i} \approx \frac{\Delta ['protein']}{['protein']} \bigg/ \frac{\Delta V_i}{V_i} \quad (7)$$

Where S_i denotes concentration of the critical protein which directly effects the change of cell phenotype proteins, V_i denote the estimated parameters and ΔV_i denotes a small change of the corresponding parameter. Each of the parameters of the estimated value increased or decreased by 1%. So sensitivity analysis is used to examine the larger changes of a system for parameter sensitivity.

4 Discussion and Perspectives

Signaling pathway mechanisms study is important for understanding the disease and multi-drug therapies. Because of the complexity of its biological environment and behavior, it is a very difficult problem to study the signal transduction pathway. This research reviews a lot of commonly used signaling pathway modeling methods. However, since current modeling techniques can not completely present the complex behavior of biological systems, the future research work will focus on how to simulate the dynamic behavior of the signal path and look for a more effective algorithm to optimize the related parameters.

Acknowledgement

This work supported by the Natural Science Foundation of China under Grant No. 61372138, Chongqing excellent youth award and the Chinese Recruitment Program of Global Youth Experts, as well as by Fundamental Research Funds for the Central Universities No. XDJK2014B012 and NO. XDJK2016A003

References

- [1] E. Klipp and W. Liebermeister, "Mathematical modeling of intracellular signaling pathways," *Bmc Neuroscience*, vol. 7, pp. S10-S10, 2006.
- [2] P. Blume-Jensen, . and T. Hunter, . "Oncogenic kinase signalling," *Nature*, vol. 411, pp. 355-365, 2001.
- [3] M. Marderosian, A. Sharma, A. R. Funk, J. Masri, O. D. Jo, and J. F. Gera, "Tristetraprolin regulates Cyclin D1 and c-Myc mRNA stability in response to rapamycin in an Akt-dependent manner via p38 MAPK signaling," *Oncogene*, vol. 25, pp. págs. 6277-6290, 2006.
- [4] H. Peng, T. Peng, J. Wen, D. A. Engler, R. K. Matsunami, J. Su, et al., "Characterization of p38 MAPK isoforms for drug resistance study using systems biology approach," *Bioinformatics*, vol. 30, pp. 1899-1907, 2014.
- [5] R. Heinrich, B. G. Neel, Rapoport, and T. A., "Mathematical Models of Protein Kinase Signal Transduction," *Molecular Cell*, vol. 9, pp. 957-970, 2002.
- [6] J. A. Papin and B. O. Palsson, "Topological analysis of mass-balanced signaling networks: a framework to obtain network properties including crosstalk," *Journal of Theoretical Biology*, vol. 227, pp. 283-297, 2004.
- [7] V. S, M. F, Z. S, and R. M, "Signal Transduction Dynamics of the Protein Kinase-A/Phosphofructokinase-2 System in *Saccharomyces cerevisiae*," *Metabolic Engineering*, vol. 3, pp. 163-172, 2001.
- [8] B. Schoeberl, C. Eichler-Jonsson, E. D. Gilles, and G. Müller, "Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors," *Nature Biotechnology*, vol. 20, pp. 370-375, 2002.

- [9] T. M. Yi, H. Kitano, and M. I. Simon, "A quantitative characterization of the yeast heterotrimeric G protein cycle," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, pp. 10764-9, 2003.
- [10] I. Swameye, ., T. G. Muller, J. Timmer, ., O. Sandra, ., and U. Klingmuller, . "Identification of nucleocytoplasmic cycling as a remote sensor in cellular signaling by databased modeling," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, pp. 1028-1033, 2003.
- [11] G. Koh, H. F. Teong, M. V. Clément, D. Hsu, and P. S. Thiagarajan, "A compositional approach to parameter estimation in pathway modeling: a case study of the Akt and MAPK pathways and their crosstalk," *Bioinformatics*, vol. 22, pp. e271-e280(10), 2006.
- [12] S. Schnell and T. E. Turner, "Reaction kinetics in intracellular environments with macromolecular crowding: simulations and rate laws ☆," *Progress in Biophysics & Molecular Biology*, vol. 85, pp. 235-260, 2004.
- [13] V. Sharma and A. Compagnoni, "Computational and mathematical models of the JAK-STAT signal transduction pathway," in *Proceedings of the 2013 Summer Computer Simulation Conference*, 2013.
- [14] D. Y. Wang, L. Cardelli, A. Phillips, N. Piterman, and J. Fisher, "Computational modeling of the EGFR network elucidates control mechanisms regulating signal dynamics," *Bmc Systems Biology*, vol. 3, pp. 1-17, 2009.
- [15] M. L. Guerriero, A. Dudka, N. Underhill-Day, J. K. Heath, and C. Priami, "Narrative-based Computational Modelling of the Gp130/JAK/STAT Signalling Pathway," *Bmc Systems Biology*, vol. 3, pp. 1096-1098, 2009.
- [16] M. Heiner, I. Koch, and J. Will, "Model validation of biological pathways using Petri nets—demonstrated for apoptosis," *Biosystems*, vol. 75, pp. 15–28, 2004.
- [17] U. S. Bhalla and R. Iyengar, . "Emergent Properties of Networks of Biological Signaling Pathways," *Science*, vol. 283, pp. 381-387, 1999.
- [18] U. S. Bhalla, P. T. Ram, and R. Iyengar, "MAP kinase phosphatase as a locus of flexibility in a mitogen-activated protein kinase signaling network," *Science*, vol. 297, pp. 1018-1023, 2002.
- [19] M. R. Birtwistle, M. Hatakeyama, N. Yumoto, B. A. Ogunnaike, J. B. Hoek, and B. N. Kholodenko, "Ligand-dependent responses of the ErbB signaling network: experimental and modeling analyses. *Mol Syst Biol* 3:144," *Molecular Systems Biology*, vol. 3, pp. 144-144, 2007.
- [20] M. Melissa, O. Mandri, G. B. Mills, and P. T. Ram, "Network topology determines dynamics of the mammalian MAPK1,2 signaling network: bifan motif regulation of C-Raf and B-Raf isoforms by FGFR and MC1R," *Faseb Journal*, vol. 22, pp. 1393-1403, 2008.
- [21] J. J. Tyson, K. Chen, and B. Novak, "Network dynamics and cell physiology," *Nature Reviews Molecular Cell Biology*, vol. 2, pp. 908-916, 2001.
- [22] J. J. Tyson, C. N. Attila, and N. Bela, "The dynamics of cell cycle regulation," *Bioessays News & Reviews in Molecular Cellular & Developmental Biology*, vol. 24, p. 1095–1109, 2002.
- [23] J. J. Hornberg, F. J. Binder BBruggeman, B. Schoeberl, R. Heinrich, and H. V. Westerhoff, "Control of MAPK signalling: from complexity to what really matters," *Oncogene*, vol. 24, pp. 5533-5542, 2005.
- [24] S. Yamada, S. Shiono, A. Joo, and A. Yoshimura, "Control mechanism of JAK/STAT signal transduction pathway," *Febs Letters*, vol. 534, pp. 190–196, 2003.
- [25] T. M. Yi, H. Kitano, M. I. Simon, T. M. Yi, H. Kitano, and M. I. Simon, "A quantitative characterization of the yeast heterotrimeric G protein cycle," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, pp. 10764-9, 2003.
- [26] I. Sergio, L. Yiling, F. C. Morales, G. B. Mills, and P. T. Ram, "Identification of Optimal Drug Combinations Targeting Cellular Networks: Integrating Phospho-Proteomics and Computational Network Analysis," *Cancer Research*, vol. 70, pp.: 6704–6714., 2010.

- [27] W. W. Chen, B. Schoeberl, P. J. Jasper, M. Niepel, U. B. Nielsen, D. A. Lauffenburger, et al., "Input–output behavior of ErbB signaling pathways as revealed by a mass action model trained against dynamic data," *Molecular Systems Biology*, vol. 5, p.: 239., 2009.
- [28] A. Hoffmann, A. Levchenko, and M. L. Scott, "The I κ B - NF- κ B signaling module: Temporal control and selective gene activation," *Science*, vol. 298, pp. 1241–1245, 2002.
- [29] F. Dana, G. Alexey, L. Galina, S. Anatoly, M. Stuart, M. Peter, et al., "Systems Biology Reveals New Strategies for Personalizing Cancer Medicine and Confirms the Role of PTEN in Resistance to Trastuzumab," *Cancer Research*, vol. 69, pp. 6713-6720, 2009.
- [30] K. Anatoly, A. Edit, N. I. Markevich, N. M. Borisov, J. B. Hoek, and B. N. Kholodenko, "Scaffolding protein Grb2-associated binder 1 sustains epidermal growth factor-induced mitogenic and survival signaling by multiple positive feedback loops," *Journal of Biological Chemistry*, vol. 281, pp. 19925-38, 2006.
- [31] B. Behinaein, K. Rudie, and W. Sangrar, "Structural Analysis of Petri Nets for Modeling and Analyzing Signaling Pathways," in *Electrical and Computer Engineering (CCECE), 2014 IEEE 27th Canadian Conference on*, 2014, pp. 1-5.
- [32] J. A. Papin, T. Hunter, B. O. Palsson, and S. Subramaniam, "Reconstruction of cellular signalling networks and analysis of their properties," *Nature Reviews Molecular Cell Biology*, vol. 6, pp. 99-111, 2005.
- [33] J. L. Peterson, "Petri Net Theory and the Modeling of Systems," Prentice-Hall, Inc., Englewood Cliffs, N.J., 1981.
- [34] T. Murata, "Petri nets: Properties, analysis and applications," *Proceedings of the IEEE*, vol. 77, pp. 541-580, 1989.
- [35] M. Heiner, D. Gilbert, and R. Donaldson, "Petri Nets for Systems and Synthetic Biology," in *International Conference on Formal Methods for the Design of Computer, Communication*, 2008, pp. 987-992.
- [36] S. Jayasuryan, A. Bamezai, and V. Gehlot, "Petri Net Based Model Of The T Cell Receptor Signaling Pathway," in *Bioinformatics & Computational Biology*, 2006, pp. 56-62.
- [37] A. Sackmann, M. Heiner, and I. Koch, "Application of Petri net based analysis techniques to signal transduction pathways," *Bmc Bioinformatics*, vol. 7, p.: 482., 2006.
- [38] Y. Miwa, K. Hioka, Q. W. Ge, and H. Matsuno, "An Algorithm to Produce Conditional Equations for Smooth Signal Flows in the Petri Net Model of a Signaling Pathway," *Technical Report of Ieice Cst*, vol. 109, pp. 25-30, 2009.
- [39] H. Matsuno, Y. Tanaka, H. Aoshima, A. Doi, M. Matsui, and S. Miyano, "Biopathways representation and simulation on hybrid functional Petri Net," *Silico Biology*, vol. 3, pp. 389-404, 2003.
- [40] H. ALLA and R. DAVID, "CONTINUOUS AND HYBRID PETRI NETS," *Journal of Circuits Systems & Computers*, vol. 8, pp. 159-188, 2011.
- [41] R. Valk, *Self-modifying nets, a natural extension of Petri nets*: Springer Berlin Heidelberg, 1978.
- [42] R. Drath, "Hybrid object nets : an object oriented concept for modeling complex hybrid systems," in *Proc. Hybrid Dynamical Systems, 3rd International Conference on Automation of Mixed Processes, ADPM'98*, 1998, pp. 437-442.
- [43] M. Pogson, R. Smallwood, E. Qvarnstrom, and M. Holcombe, "Formal agent-based modelling of intracellular chemical interactions," *Bio Systems*, vol. 85, pp. 37–45, 2006.
- [44] C. Jacob and I. Burleigh, "Biomolecular swarms – an agent-based model of the lactose operon," *Natural Computing*, vol. 3, pp. 361-376, 2004.
- [45] C. Jacob, A. Barbasiewicz, and G. Tsui, "Swarms and Genes: Exploring λ - Switch Gene Regulation through Swarm Intelligence," in *Evolutionary Computation*, 2006. CEC 2006. IEEE Congress on, 2006, pp. 2535-2542.
- [46] A. S. Shirazi, S. Von Mammen, and C. Jacob, "Hierarchical self-organized learning in agent-based modeling of the MAPK signaling pathway," in *Evolutionary Computation (CEC), 2011 IEEE Congress on*, 2011, pp. 2245 - 2251.

- [47] E. Balsa-Canto, M. Peifer, J. R. Banga, J. Timmer, and C. Fleck, "Hybrid optimization method with general switching strategy for parameter estimation," *Bmc Systems Biology*, vol. 2, p.: 26., 2008.
- [48] C. G. Moles, P. Mendes, and J. R. Banga, "Parameter estimation in biochemical pathways: a comparison of global optimization methods," *Genome Research*, vol. 13, pp. 2467-2474, 2003.
- [49] M. Rodriguez-Fernandez, P. Mendes, and J. R. Banga, "A hybrid approach for efficient and robust parameter estimation in biochemical pathways," *Bio Systems*, vol. 83, pp. 248-265, 2006.
- [50] C. C. Wang, M. Cirit, and J. M. Haugh, "PI3K-dependent cross-talk interactions converge with Ras as quantifiable inputs integrated by Erk," *Molecular Systems Biology*, vol. 5, p.: 246., 2009.
- [51] H. Abdi, "Coefficient of variation," *Encyclopedia of research design*, pp. p. 169-171, 2010.
- [52] X. Sun, J. Su, J. Bao, T. Peng, L. Zhang, Y. Zhang, et al., "Cytokine combination therapy prediction for bone remodeling in tissue engineering based on the intracellular signaling pathway," *Biomaterials*, vol. 33, pp. 8265–8276, 2012.
- [53] P. Huiming, W. Jianguo, L. Hongwei, C. Jeff, and Z. Xiaobo, "Drug inhibition profile prediction for NFκB pathway in multiple myeloma," *Plos One*, vol. 6, pp. e14750-e14750, 2011.