OXFORD

Systems biology

NetExplore: a web server for modeling small network motifs

Dmitri Papatsenko[1,2,*] and Ihor R. Lemischka[1,2,3]

[1]Department of Regenerative and Developmental Biology, [2]Black Family Stem Cell Institute and [3]Department of Pharmacology and System Therapeutics, Icahn School of Medicine at Mount Sinai, Systems Biology Center New York, One Gustave L. Levy Place, New York, NY 10029, USA

*To whom correspondence should be addressed.
Associate Editor: Alfonso Valencia

Abstract

Motivation: Quantitative and qualitative assessment of biological data often produces small essential recurrent networks, containing 3–5 components called network motifs. In this context, model solutions for small network motifs represent very high interest.
Results: Software package NetExplore has been created in order to generate, classify and analyze solutions for network motifs including up to six network components. NetExplore allows plotting and visualization of the solution's phase spaces and bifurcation diagrams.
Availability and implementation: The current version of NetExplore has been implemented in Perl-CGI and is accessible at the following locations: http://line.bioinfolab.net/nex/NetExplore.htm and http://nex.autosome.ru/nex/NetExplore.htm.
Contact: dmitri.papatsenko@gmail.com
Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 Introduction

The analysis of biological systems using quantitative and experimental methods often produces small gene regulatory networks or network motifs which perform critical functions. Network motifs including 2–5 nodes became the major subjects of modeling and analysis (Goentoro *et al.*, 2009; Milo *et al.*, 2002; Mangan and Alon, 2003). Many types of network motifs have been identified and described (Goentoro *et al.*, 2009; Heinaniemi *et al.*, 2013; Papatsenko and Levine, 2011; Shen *et al.*, 2012; Stricker *et al.*, 2008). However, even motifs with identical topologies (order of regulatory links) may produce surprisingly rich spectrums of dynamic solutions.

Constraining network parameters given biological data is routinely used to model networks and network motifs; this approach has advantages and limitations. Major limitations include: (i) The availability of adequate biological data to fit the network parameters, (ii) overfitting problems, caused by imbalance between the amount of experimental data and the number of unknown parameters (Janssens *et al.*, 2006; Papatsenko and Levine, 2011; Papatsenko *et al.*, 2013; Warmflash *et al.*, 2012). Constraining parameters may produce many alternative solutions.

Recent progress in genomics and the accumulation of large amounts of high-throughput data allows reconstruction of Gene Regulatory Networks (GRNs) and finding essential network motifs *in silico* (Ivanova *et al.*, 2006; Lu *et al.*, 2009). Collection of data and validation of these *predicted* network motifs is difficult, but prediction of their likely behavior is desirable. Solutions may be obtained by assigning 'random' sets of parameters, centered on biologically relevant parameter values. Collection and classification of many 'emitted' random solutions predicts probable behavior of network motifs.

2 Algorithm

NetExplore includes three parts (modules): NetExplore Solver, NetExplore Browser and NetExplore Bifurcator. Each module is available online and as a standalone program for download. NetExplore Solver emits random sets of parameters for a user-defined network. NetExplore Browser explores solutions and visualizes their phase spaces. NetExplore Bifurcator carries out bifurcation analysis and displays bifurcation diagrams.

Net Explore algorithm is based on the previously developed framework for GRNs (Papatsenko *et al.*, 2011; Papatsenko *et al.*,
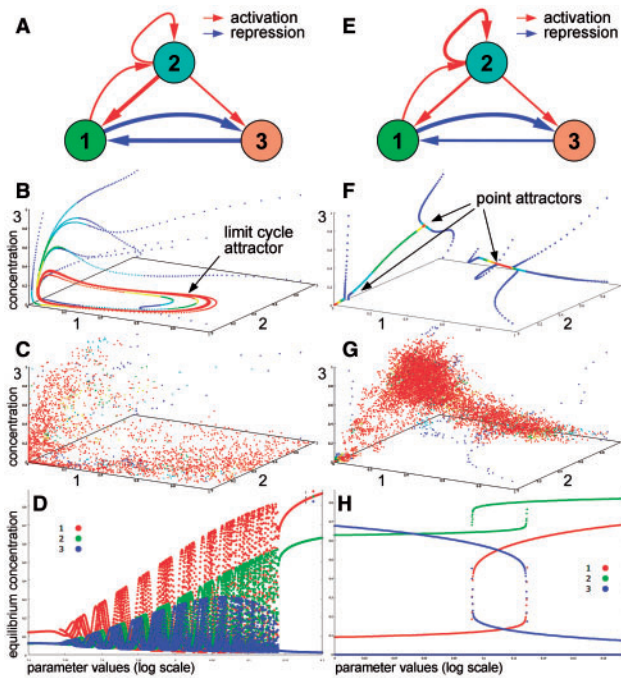
Fig. 1. Identical networks may produce different solutions. (**A**, **E**) Two networks are shown with identical topologies, but different values of K (regulatory strengths, arrow thickness). (**B**, **F**) Phase spaces are shown, (**B**) the solution is a limit cycle attractor, (**F**) the solution is a multistable system with three point attractors shown by red color. (**C**, **G**) The same phase spaces in the presence of stochastic noise, the attractors produce "clouds" of states. Colors in **B**, **C**, **F** and **G** show system evolution in time (red – finite states). (**D**, **H**) Bifurcation analysis (proportional scaling of all Ks) of motifs on **A** and **E** is shown. (**D**) Oscillator produces a characteristic pattern. (**H**) Multistable system, three dynamic point attractors are clearly visible in the transition region (see the blue line)

2013; Zinzen *et al.*, 2006). A detailed description of the framework is available in Supplementary Material. Concentration of every component in the system is calculated based on a system of ordinary differential equations (ODE) or stochastic differential equations (SDE), defined automatically by NetExplore. Synthesis of every x_i component in the system is calculated from integrated inputs (P_i, reflects the network topology), exponential decay and stochastic noise:

$$\frac{d[x_i]}{dt} = \alpha_i P_i - \beta_i [x_i] + \xi_i(0, \sigma_i)\sqrt{[x_i]} \qquad (1)$$

In Equation (1), α, β and ξ define synthesis, decay and stochastic noise rate constants. Inputs from A activators and R repressors are integrated as follows:

$$P_i = \left(1 - \prod_{j \in A} (1 - p_{ij})\right)\left(\prod_{k \in R} (1 - p_{ik})\right) \qquad (2)$$

p in Equation (2) describes probability of activation by a single input:

$$p_{ij} = \frac{(1 + C_j K_{ij}[x_j])^N - 1}{C_j + (1 + C_j K_{ij}[x_j])^N - 1}; \quad K = e^{-\frac{\Delta G_{P-DNA}}{RT}}; \quad C = e^{-\frac{\Delta G_{P-P}}{RT}} \qquad (3)$$

Equation (3) is similar to Hill function (Berg and von Hippel, 1988; Bintu *et al.*, 2005; Zinzen and Papatsenko, 2007). The model includes the following parameters: binding constant K, cooperativity C and number of binding sites N. System of ODEs describing network is automatically constructed and solved numerically by NetExplore.

3 Results and Discussion

NetExplore Solver provides options for network design. A user defines the number of network components, order and sign of regulatory links between the components. NetExplore Solver emits large number of solutions (standalone version has no limit on the number of solutions) and classifies them based on the number of dynamic attractors (monostable, bistable etc) and the presence of periodic signals (limit cycle attractors). Output of the NetExplore Solver is a list of solutions (parameter values), which may be analyzed, edited and passed to the second module, NetExplore Browser. NetExplore Browser can explore phase spaces for the selected solutions [see Fig. 1(A,B,E,F)], including analysis of the system in the presence of stochastic noise (Fig. 1C and G). 'Adjust parameters' panel of NetExplore Browser provides options for model adjustment. The 'Bifurcate' button links NetExplore Browser to NetExplore Bifurcator, a program that carries out bifurcation analysis (Fig. 1D and H) of a given model solution. The stand-alone version of NetExplore Bifurcator (see online help) can also create pictures, consecutively displaying evolution of the phase space in response to gradual adjustment of the model parameters. The frames can be combined into a video, such as the one shown here: https://www.youtube.com/watch?v=a85upWytWwA.

Collecting large a number of solutions ($>10^4$) and bifurcation analysis may require substantial time; therefore it may be desirable to use the standalone versions of Solver and Bifurcator.

References

Berg,O.G. and von Hippel,P.H. (1988) Selection of DNA binding sites by regulatory proteins. II. The binding specificity of cyclic AMP receptor protein to recognition sites. *J. Mol. Biol.*, **200**, 709–723.

Bintu,L. *et al.* (2005) Transcriptional regulation by the numbers: models. *Curr. Opin. Genet. Dev.*, **15**, 116–124.

Goentoro,L. *et al.* (2009) The incoherent feedforward loop can provide fold-change detection in gene regulation. *Mol Cell.*, **36**, 894–899.

Heinaniemi,M. *et al.* (2013) Gene-pair expression signatures reveal lineage control. *Nat. Methods*, **10**, 577–583.

Ivanova,N. *et al.* (2006) Dissecting self-renewal in stem cells with RNA interference. *Nature*, **442**, 533–538.

Janssens,H. *et al.* (2006). Quantitative and predictive model of transcriptional control of the Drosophila melanogaster even skipped gene. *Nat Genet.*, **38**, 1159–1165.

Lu,R. *et al.* (2009) Systems-level dynamic analyses of fate change in murine embryonic stem cells. *Nature*, **462**, 358–362.

Mangan,S. and Alon,U. (2003) Structure and function of the feed-forward loop network motif. *Proc. Natl Acad. Sci. USA*, **100**, 11980–11985.

Milo,R. *et al.* (2002) Network motifs: simple building blocks of complex networks. *Science.* **298**, 824–827.

Papatsenko,D. and Levine,M. (2011) The Drosophila gap gene network is composed of two parallel toggle switches. *PLoS* **6**, e21145.

Papatsenko,D. *et al*. (2011) Clusters of temporal discordances reveal distinct embryonic patterning mechanisms in Drosophila and anopheles. *PLoS* **9**, e1000584.

Papatsenko,D. *et al*. (2013). Quantitative Approaches to Model Pluripotency and Differentiation in Stem Cells. In: Sell, S. (ed.) *Stem Cell Handbook*. Springer Science, New York, pp. 59–74.

Shen,Y. *et al*. (2012) A map of the cis-regulatory sequences in the mouse genome. *Nature*, **488**, 116–120.

Stricker,J. *et al*. (2008). A fast, robust and tunable synthetic gene oscillator. *Nature*, **456**, 516–519.

Warmflash,A. *et al*. (2012) Pareto evolution of gene networks: an algorithm to optimize multiple fitness objectives. *Phys. Biol.* **9**, 1478–3975.

Zinzen,R.P. and Papatsenko,D. (2007) Enhancer responses to similarly distributed antagonistic gradients in development. *PLoS Comput. Biol.*, **3**, e84.

Zinzen,R.P. *et al*. (2006) Computational models for neurogenic gene expression in the Drosophila embryo. *Curr. Biol.* **16**, 1358–1365.