Control of apoptosis by SMAR1†

Md. Zubbair Malik,[a,d] Md. Jahoor Alam,[b] Romana Ishrat,[a] Subhash M. Agarwal,[c] and R.K. Brojen Singh*[d]

Received Xth XXXXXXXXXX 20XX, Accepted Xth XXXXXXXXX 20XX
First published on the web Xth XXXXXXXXXX 200X
DOI: 10.1039/b000000x

Supplimentary information

A. Mathematical model of the network

The $p53-SMAR1$ regulatory network is modeled by a set of variables $\{X_i\}$ $i = 1,2,...,N$ with $N = 18$ (18 molecular species) (Table 1) which undergo thirty five reaction channels ($M = 35$) (Table 2). The state vector at any instant of time t is given by $\vec{X}(t) = (X_1,\ldots,X_N)^T$, where the variables in the vector are populations of the molecular species listed in Table 1. The classical deterministic equations constructed from these reactions in the network (Fig. 1, Table 2) are given by,

$$\frac{dx_1}{dt} = -k_1x_1x_{14} + k_6 - k_8x_1x_2 + k_9x_4 - k_{12}x_1x_6 + k_{13}x_7 + k_{17}x_{10}x_{12} - k_{29}x_1x_{15} + k_{33}x_{10}x_{18} + k_{34}x_9x_{15} \quad (1)$$

$$\frac{dx_2}{dt} = k_2x_3 - k_5x_2 + k_7x_4 - k_8x_1x_2 + k_9x_4 - k_{18}x_2x_{15} - k_{20}x_2x_8 + k_{21}x_{13} - k_{32}x_2x_{11} \quad (2)$$

$$\frac{dx_3}{dt} = k_3x_1 - k_4x_3 \quad (3)$$

$$\frac{dx_4}{dt} = -k_7x_4 + k_8x_1x_2 - k_9x_4 - k_{19}x_4x_8 - k_{30}x_4x_{15} \quad (4)$$

$$\frac{dx_5}{dt} = -k_{10}x_5 + k_{11}x_6 \quad (5)$$

$$\frac{dx_6}{dt} = k_{10}x_5 - k_{11}x_6 - k_{12}x_1x_6 \quad (6)$$

$$\frac{dx_7}{dt} = k_{12}x_1x_6 - k_{13}x_7 - k_{15}x_7x_8 + k_{29}x_1x_{15} + k_{31}x_{17} \quad (7)$$

$$\frac{dx_8}{dt} = -k_{14}x_8 - k_{15}x_7x_8 - k_{19}x_4x_8 - k_{20}x_2x_8 + k_{23} - k_{35}x_8x_{15} \quad (8)$$

$$\frac{dx_9}{dt} = k_{15}x_7x_8 - k_{16}x_9 - k_{34}x_9x_{15} \quad (9)$$

$$\frac{dx_{10}}{dt} = k_{16}x_9 - k_{17}x_{10}x_{12} - k_{33}x_{10}x_{18} \quad (10)$$

$$\frac{dx_{11}}{dt} = -k_{22}x_{11} - k_{25}x_{11}x_{16} + k_{24} - k_{32}x_2x_{11} \quad (11)$$

$$\frac{dx_{12}}{dt} = -k_{17}x_{10}x_{12} + x_{25}x_{11}x_{16} \quad (12)$$

$$\frac{dx_{13}}{dt} = k_{19}x_4x_8 - k_{21}x_{13} \quad (13)$$

$$\frac{dx_{14}}{dt} = -k_1x_1x_{14} + k_{20}x_2x_8 \quad (14)$$

$$\frac{dx_{15}}{dt} = -k_{18}x_2x_{15} + k_{26} - k_{27}x_{15} - k_{29}x_1x_{15} - k_{30}x_4x_{15} \quad (15)$$

† Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/b000000x/

[a] *Centre for Interdisciplinary Research in Basic Sciences, Jamia Millia Islamia, New Delhi-110025, India.*
[b] *College of Applied Medical Sciences, University of Hail, Hail-2440, Kingdom of Saudi Arabia.*
[b] *Bioinformatics Division, Institute of Cytology and Preventive Oncology, 1-7, Sector - 39, Noida 201301, India.*
[b] *School of Computational and Integrative Sciences, Jawaharlal Nehru University, New Delhi-110067, India.; Email: brojen@jnu.ac.in*

$$\frac{dx_{16}}{dt} = k_{18}x_2x_{15} - k_{25}x_{11}x_{16} - k_{28}x_{16} + k_{31}x_{17} \quad (16)$$
$$\frac{dx_{17}}{dt} = k_{30}x_4x_{15} - k_{31}x_{17} \quad (17)$$
$$\frac{dx_{18}}{dt} = k_{32}x_2x_{11} - k_{33}x_{10}x_{18} \quad (18)$$

where $\{k_i\}$ and $\{x_i\}$, for $i = 1,2,\ldots,N$, represent the sets of rate constants of the reactions (Table 2) (see the article) and concentration of the variables corresponding to the molecular species $x_i = X_i/V$, where V is system size (Table 1) (see the article), respectively. The coupled set of non-linear differential equations can be solved numerically using the standard fourth order Runge–Kutta method of numerical integration [1] to study dynamical behavior of the system.

B. Quasi-steady state approximation

The system of reactions (Table 2) (see the article), from which the ODEs (12)–(30) were constructed, can be approximately divided into two types of elementary reactions, namely, *fast* and *slow* reactions [2]. The variables in the state vector $\vec{x}$ can be divided into fast and slow vectors given by

$$\vec{x}^s = \begin{bmatrix} x_1 \\ x_2 \\ x_8 \\ x_{11} \\ x_{15} \end{bmatrix}; \vec{x}^f = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}; \vec{x} = \begin{bmatrix} \vec{x}^s \\ \vec{x}^f \end{bmatrix};$$
$$y_1 = \begin{bmatrix} x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix}; y_2 = \begin{bmatrix} x_9 \\ x_{10} \end{bmatrix}; y_3 = \begin{bmatrix} x_{12} \\ x_{13} \\ x_{14} \end{bmatrix}; y_4 = \begin{bmatrix} x_{16} \\ x_{17} \\ x_{18} \end{bmatrix} \quad (19)$$

The fast variables are normally corresponding to complex molecular species. Generally, formation of complex molecular species due to fast reactions is followed by fast decay of these complexes, the dynamics of the fast variables reaches steady state much quickly as compared to the dynamics of slow variables [3,4]. We then use Henri–Michaelis–Menten–Briggs–Haldane approximation to assume that the time evolution of fast state vector $\vec{x}^f$ reaches equilibrium state defined by $\vec{x}^{*f}$ much faster as compared to the time evolution of slow state vector $\vec{x}^s$ [3,4]. Applying this approximation, we can reach the following steady state for fast variables,

$$\frac{d\vec{x}^f}{dt} \approx 0; \vec{x}^{*f} = \begin{bmatrix} y_1^* \\ y_2^* \\ y_3^* \\ y_4^* \end{bmatrix};$$
$$y_1^* = \begin{bmatrix} x_3^* \\ x_4^* \\ x_5^* \\ x_6^* \\ x_7^* \end{bmatrix}; y_2^* = \begin{bmatrix} x_9^* \\ x_{10}^* \end{bmatrix}; y_3^* = \begin{bmatrix} x_{12}^* \\ x_{13}^* \\ x_{14}^* \end{bmatrix}; y_4^* = \begin{bmatrix} x_{16}^* \\ x_{17}^* \\ x_{18} \end{bmatrix} \quad (20)$$

such that the dynamics of the system for sufficiently large time is governed by the dynamics of the slow variables given by

$$\frac{d\vec{x}}{dt} = \frac{d}{dt}\begin{bmatrix} \vec{x}^s \\ \vec{x}^f \end{bmatrix} \approx \frac{d\vec{x}^s}{dt} = \frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \\ x_8 \\ x_{11} \\ x_{15} \end{bmatrix} \quad (21)$$

The approximate solution of the complex model can be obtained from this reduced model using quasi-steady-state approximation.

C. Multifractal DFA approach

Multifractal detrended fluctuation analysis (MF-DFA) is a technique used to determine fractal properties, and to detect important correlations in nonstationary time series [5]. Fractal parameters, namely, Hurst exponent (H), generalized dimension (D) etc can be calculated numerically using method adopted by Kantelhardt et.al. [6] as summarized below. First, the time series signal $\{x_j\}$ of length N is taken as random walk, and can be represented by the profile, $Y(i) = \sum_{j=1}^{i}(x_j - \langle x \rangle)$, where, $\langle x \rangle$ is mean value of the signal, and $i = 1,2,...,N$. Second, profile $Y(i)$ is divided into $N_s = int(\frac{N}{s})$ equal nonoverlapping segments of size s. To consider all data points, $2N_s$ segments are considered by counting starting from both ends of the data. Third, the following variance is determined,

$$F^2(s,\nu) = \frac{1}{s}\sum_{i=1}^{s}\{Y[(\nu-1)s+i] - y_\nu(i)\}^2 \quad (22)$$

where, $\nu = N_s + 1,...,2N_s$, and $y_\nu(i)$ is the fitting polynomial in segment ν. Fourth, the qth order fluctuation function is estimated by averaging over all segments,

$$F_q(s) = \left\{\frac{1}{2N_s}\sum_{\nu=1}^{2N_s}[Y[(\nu,s)]^{q/2}\right\}^{1/q} \quad (23)$$

Fifth, the scaling behavior of the function $F_q(s)$ is represented by,

$$F_q(s) \sim s^{H_q} \quad (24)$$

where, H_q is the generalized Hurst exponent, which represents the measure of self-similarity and correlation properties of the signal. Then H_q is related to classical scaling exponent $\tau(q)$,

$$\tau(q) = qH_q - 1 \tag{25}$$

and from the definition of Holder exponent, $\alpha = \frac{d\tau}{dq}$, the singularity function $f(\alpha)$[5] is given by,

$$f(\alpha) = q\alpha - \tau(q) \tag{26}$$

Then, generalized dimension of the signal is measured by,

$$D_q = \frac{\tau(q)}{q-1} \tag{27}$$

Now, D_0, for $q = 0$, is the fractal or Hausdorff dimension, D_1 is information dimension and D_2 represents correlation dimension[5]. Multifractal signature can be observed in the system if there exists significant dependence of H_q on q in the time series due to different scaling nature of small and large fluctuations[6]. Positive dependence of H_q on q indicates large fluctuations in the time series, whereas negative dependence of H_q on q exhibits small fluctuations in the time series. Further, in multifractal time series, small and large fluctuations are characterized by large and small values of H_q.

D. Visibility graph of time series

This technique maps a time series to a network[7], where each observation in time series is translated to a node and an edge between any two nodes is introduced when the following visibility condition is satisfied i.e. two nodes corresponding to observations $x(t_a)$ and $x(t_b)$ are connected if all intermediate states $x(t_c)$ with $t_a < t_c < t_b$ satisfy,

$$\frac{x_b - x_c}{t_b - t_c} > \frac{x_b - x_a}{t_b - t_a} \tag{28}$$

These networks are undirected due to symmetry in visibility condition. Since properties of the time series are inherited to the corresponding network, the studies of this network provide useful information which can't be observed in traditional time series data.

Topological properties of networks

The following topological properties of the networks are studied to examine the important behavior of the networks.

Degree distribution: The degree k of a node indicates the number of links the node connects with other nodes in the network. Consider a network defined by a graph $G = (N, E)$, where N and E are number of nodes and edges respectively. The probability of degree distribution ($P(k)$) of the network is the probability that any chosen node will have a degree k, which is given by,

$$P(k) = \frac{n_k}{N} \tag{29}$$

where, n_k is the number nodes having degree k. $P(k)$ in random and small-world networks follow Poisson distribution, whereas, it obeys power law $P(k) \sim k^{-\gamma}$ in scale-free and hierarchical networks depending on the value of γ which indicates the importance of hubs or modules in the network[8].

Clustering co-efficient: Clustering co-efficient of a network is a measure of how strongly a node's neighborhoods are interconnected. It is the ratio of the number of triangular motifs a node has with its nearest neighbor to the maximum possible number of such motifs. For an undirected network, clustering co-efficient ($C(k_i)$) of ith node can be calculated by,

$$C(k_i) = \frac{2e_i}{k_i(k_i - 1)} \tag{30}$$

where, e_i is the number of connected pairs of nearest-neighbor of ith node, and k_i is its degree. $C(k)$ in scale free networks does not depend on k, whereas in hierarchical network it follows a power law, $C(k) \sim k^{-\alpha}$, with $\alpha \sim 1$.

Neighborhood connectivity: Neighborhood connectivity of a node of a network ($C_N(k)$) represents the average connectivities of the nearest neighbors of the node in the network[9], given by,

$$C_N(k) = \sum_q qP(q|k) \tag{31}$$

where, $P(q|k)$ is conditional probability that a link belonging to a node with connectivity k points to a node with connectivity q. The k dependence of $C_N(k)$, i.e. $C_N(k) \sim k^{-\beta}$ is a signature of hierarchical topology in the network[10]. Further, change in the sign of exponent β could be an indicator of assortivity in the network topology[11].

Betweenness centrality: Betweenness centrality of a node represents the ability to (a) extract benefits from information flows in the network[12], and (b) extent to which the node has control over the other nodes in the network through communication[13,14]. If $d_{ij}(v)$ indicates the number of geodesic paths from node i to node j passing through node v, and d_{ij} represents number of geodesic paths from node i to j, then betweenness centrality ($C_B(v)$) of a node v can be measured by,

$$C_B(v) = \sum_{i,j;i\neq j\neq k} \frac{d_{ij}(v)}{d_{ij}} \tag{32}$$

If M denotes the number of node pairs excluding v, then normalized betweenness centrality is given by, $C_{NB}(v) = \frac{1}{M} C_B(v)$.

Closeness centrality: Closeness centrality (C_C) estimates how fast information is spread from a node to other nodes reachable from it in the network[15]. C_C of a node i is the reciprocal of the mean geodesic distance between the node and all other nodes connected to it in the network, and is given by,

$$C_N(k) = \frac{n}{\sum_j d_{ij}} \tag{33}$$

where, d_{ij} is geodesic path length between nodes i and j, and n is the total number of nodes in the network connected to node i.

Eigenvector centrality: Eigenvector centrality of a node i ($C_E(i)$) in a network is proportional to the sum of i's neighbor centralities [16], and it is measured by,

$$C_E(i) = \frac{1}{\lambda} \sum_{j=nn(i)} v_j \tag{34}$$

where, $nn(i)$ indicates nearest neighbors of node i in the network. λ is the eigenvalue of the eigenvector v_i given by, $Av_i = \lambda v_i$, where, A is the adjacency matrix of the network. The principal eigenvector of A, which corresponds to maximum eigenvalue λ_{max}, is taken to have positive eigenvector centrality score [17]. Since node's eigenvector centrality function smoothly varies over the network and depends on its neighbors, node with high eigenvector centrality is embedded in the locality of nodes of high eigenvector centralities, and chance of having isolated nodes in and around the locality is very low [16]. Hence, eigenvector centrality can be used as an indicator of node's spreading power in the network.

E. Permutation entropy: a measure of complexity

The complex information contained in a system is inherited in the time series of the constituting variables of the system, and can be measured by calculating permutation entropy of the time series [18,19]. Permutation entropy H of a time series of a dynamical variable $x(t)$ of a system can be calculated as follows. The time series $x(t)$ can be mapped onto a symbolic sequence of length N: $x(t) = \{x_1, x_2, ..., x_N\}$. This sequence is then partitioned into M number of short sequences of equal length U i.e. $x(t) = \{w_1, w_2, ..., w_M\}$, where ith window is given by $w_i = \{x_{i+1}, x_{i+2}, ..., x_{i+U}\}$. This window is allowed to slide along $x(t)$ with maximum overlap. Permutation entropy of a window w_i can be calculated by defining a short sequence of embedded dimension r, $S_i = \{x_{i+1}, x_{i+2}, ..., x_{i+r}\}$ in r-dimensional space, finding all possible inequalities of dimension r and mapping the inequalities along the ascendingly arranged elements of w_i to find the probabilities of occurrence of each inequality in w_i. Since q out of $r!$ permutations are distinct, one can define a normalized permutation entropy as $H_i = -\frac{1}{ln(r!)} \sum_{j=1}^{q} p_j ln(p_j)$, where $0 \leq H_i(r) \leq 1$. The mapped permutation entropy spectrum of time series $x(t)$ is represented by $H = \{H_1, H_2, ..., H_M\}$. In self-organized state one has $H \to 0$.

References

1 W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, *Numerical Recipe in Fortran*, Cambridge University Press, 1992, **(1.ISBN)**, 9780521430647.

2 J.D. Murray, *Mathematical Biology I and II*, Springer-Verlag 3rd Edition, 2003, **ISBN**, 9780387224374.

3 M. Schauer, R. Heinrich, *Math. Biosc.*, 1983, **198365(2)**, 155-170.

4 M.G. Pedersen, A.M. Bersani and E. Bersani, *J. Math. Chem.*, 2007, **43(4)**, 1318.

5 T.C. Halsey, M.H. Jensen, L.P. Kadanoff, I. Procaccia and B.I. Shraiman, *Phys. Rev.*, 1986, **A 33**, 1141.

6 J.W. Kantelhardt, *et al*. *Physica*, 2002, **A 316**, 87-114.

7 L. Lacasa, B. Luque, F. Ballesteros, J. Luque and J.C. Nuno, *Pro. Natl. Acad. Sc*, 2008, **105**, 4972–4975.

8 R. Albert and A. L. Barabasi, *Rev. Mod. Phys.*, 2002, **74**, 47.

9 S. Maslov and K. Sneppen, *Science*, 2002, **296**, 910–913.

10 P.S. Romualdo, V. Alexei and V. Alessandro, *Phys. Rev. Lett.*, 2001, **87**, 258701.

11 A. Barrat, M. Barthelemy, R. Pastor-Satorras and A. Vespignani, *PNAS*, 2004, **101**, 3747-3752.

12 S.P. Borgatti and M.G.Everett, *Social Networks*, 2006, **28**, 466–484.

13 U.A. Brandes, *J. Math. Sociol.*, 2001, **25**, 163–177.

14 O. Mason and M. Verwoerd, *IET Syst. Biol.*, 2007, **1**, 89–119.

15 G. Canright and E.M. Kenth, *Sci. Comput. Program.*, 2004, **53**, 195–214.

16 G. Canright, and E.M. Kenth, *Complexus*, 2006, **3**, 131–146.

17 P. Bonacich, *Am. J. Sociol.*, 1987, **92**, 1170–1182.

18 C. Bandt and B. Pompe, *Phys. Rev. Lett*. 2002, **88(17)**, 174102.

19 Y. Cao, W. Tung, J.B. Gao, V.A. Protopopescu and L.M. Hively, *P hys Rev.*, 2004, **E70**, 046217.