

Analysis of the Effect of Branch Position on the Trajectory Stability of Gene Regulatory Networks

Yunjun Zhang

Department of Natural Sciences, Peking University Health Science Center, Beijing
Email: yunjun.zhang@pku.edu.cn

Received: Feb. 6th, 2015; accepted: Feb. 19th, 2015; published: Feb. 25th, 2015

Copyright © 2015 by author and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Revealing the relationship between the dynamics and structure of gene regulatory networks is a challenging issue in system biology. This study focuses on the trajectory stability which is a typical dynamic characteristic of regulatory networks. Based on the ITC network model which forms the main structure in a regulatory network, we analyzed how the positions of branches affect the trajectory stability of the network. In theoretical analysis, it has been proven that the branch's position will affect the trajectory stability in two ways. On one side, if inserting an independent branch or a branch at the first node of the ITC model, the trajectory stability of the network will decrease slightly; on the other side, if inserting a branch at the end node of the ITC model, the trajectory stability will considerably increase. In simulation, it has also been shown that if inserting a branch at middle position, the trajectory stability will be affected by a combination of these two effects. These findings can not only help to reveal the topological origin of trajectory stability in regulatory networks, but to provide guidance in designing robust artificial network in synthetic biology.

Keywords

Attractor, Boolean Model, Chain-Like Network

分支链路对基因调控网络路径稳定性的影响分析

张云俊

北京大学医学部医用理学系, 北京
Email: yunjun.zhang@pku.edu.cn

收稿日期: 2015年2月6日; 录用日期: 2015年2月19日; 发布日期: 2015年2月25日

摘要

揭示基因调控网络的动力学特征与网络结构的关系是系统生物学的研究热点, 本文针对典型的动力学特征——路径稳定性, 以调控网络的主干结构ITC模型为基础, 分析了在ITC模型的不同位置接入分支链路后, 网络路径稳定性的变化情况。从理论上证明了在ITC模型的不同位置接入分支链路对路径稳定性具有不同的影响。一方面, 在ITC模型的起点中加入分支链路或者在ITC模型之外加入独立分支链路, 路径稳定性降低; 另一方面, 在ITC模型的末端加入分支链路会提高路径稳定性。此外, 实验模拟的结果表明, 在ITC模型的中间节点加入分支链路, 网络的路径稳定性会受到上述两种作用的综合影响。本研究结果不仅能够帮助揭示调控网络中路径稳定性的形成机制, 也能够合成生物学中指导设计稳定的人工生物网络, 具有理论和实践两方面的意义。

关键词

吸引子, 布尔模型, 链式网络模型

1. 引言

基因调控网络是由一系列相互作用的生物分子组成的动态演化系统[1]。每一个网络节点代表一种生物分子, 而网络边则表示分子间的相互作用(包括抑制作用和激活作用)。调控网络各节点的表达状态受相互作用的影响而不断变化, 这种顺序变化的状态序列是控制细胞内各种生命活动的控制信号。在实际演化环境中, 调控网络能够克服生物化学过程中的各种随机干扰, 稳定地保持状态演化的稳定性。因此, 研究调控网络的状态演化规律及其影响因素是系统生物学的重要议题[2]-[5]。

前期研究表明, 调控网络的拓扑结构与网络的稳定性特征有密切联系。Li 等人[2]运用经典的布尔模型对调控网络的状态演化过程进行建模, 首次发现了调控网络的状态稳定性和路径稳定性, 同时, 该结果也指示了网络拓扑结构是产生这些稳定性特征的决定性因素。Wang 等人[5]和 Zhang 等人[6]进一步研究了如何利用网络的稳定性特征来提高逆向学习网络结构的效率。Wu 等人[7]提出了调控熵来描述调控网络的结构特征。调控熵是指网络内部节点之间的调控关系的一致性。通过模拟数据和实际网络验证发现, 调控网络的稳定性特征与其调控熵大小有显著的相关性。

为了准确定位产生网络稳定性的特征结构模块, Zhang 等人[8]从链式网络模型出发, 建立了调控网络的主干结构模型——理想传递链(Ideal Transmission Chain, ITC)模型, 并首次从理论上证明了 ITC 模型是产生状态稳定性的充分和必要条件。此外, 真实调控网络的实例研究也表明, ITC 模型对其状态稳定性具有决定性作用。

本文在文献[8]的基础进一步讨论影响网络路径稳定性的结构因素。真实的调控网络是以 ITC 模型为基本主干, 同时包含多个分支链路的树状结构, 而且, 分支链路的结构基本相同, 只是接入 ITC 模型的位置不同。因此, 本文研究了三种典型接入位置, 即在 ITC 模型之外, 在 ITC 模型的起点以及在 ITC 模型的末端接入分支链路。针对这三种情况, 分别证明了加入分支链路后路径稳定性的变化情况。最后,

用实验模拟的方式研究了在其他位置加入分支链路后路径稳定性的变化情况。

2. 方法

2.1. 布尔网络模型简介

本研究使用布尔网络模型来描述基因调控网络的状态演化过程。布尔网络模型(Boolean network model)是一种常用的网络建模方法[2] [5] [9]，它用 0~1 取值的布尔变量来表示网络节点的状态。网络节点的状态取 0，对应实际情况该种基因不表达或低表达；节点取值为 1 表示该基因高表达。节点之间的作用边也被简化成抑制和激活两种类型，分别取值+1 和- λ 两种情况。 λ 体现了抑制作用对激活作用的相对权重： $\lambda = 1$ 表示抑制作用和激活作用等权；另一种情况是 $\lambda = \infty$ ，表示抑制作用具有决定性作用。后一种模型也成为一票否决模型。在本文的研究中，我们主要考虑一票否决模型。

在布尔模型下，各节点的状态演化是同步进行的，具体规则如下：对于包含 n 个节点的网络，假定在 t 时刻节点 i 的状态是 $s_i(t)$ ，那么，在 $t+1$ 时刻，有

$$s_i(t+1) = \begin{cases} 0 & \sum_{j=1}^n a_{ji} s_j(t) < 0 \\ 1 & \sum_{j=1}^n a_{ji} s_j(t) > 0 \\ s_i(t) & \sum_{j=1}^n a_{ji} s_j(t) = 0 \end{cases}$$

其中， a_{ji} 是节点 j 到节点 i 的作用边权重：如果是激活边， $a_{ji} = 1$ ；是抑制边的话， $a_{ji} = -\infty$ 。我们将同一时刻(如 t 时刻)网络各个节点的状态组合成 n 维的二值向量，即 $S^n(t) = (s_1(t), s_2(t), \dots, s_n(t))'$ ，该向量就表示网络在该时刻的状态。

在布尔网络模型的框架下，调控网络内部的各种生物分子之间的调控关系被简化为相应网络节点之间的连接关系(用连接关系矩阵 $(a_{ji})_{n \times n}$ 表示)。这种连接关系可以用网络拓扑图表示(如下面的图 1 所示)。布尔网络简单而有效，能够较完整地刻画调控网络的整体动力学行为特征[2]，因此，在布尔模型的基础上，调控网络的研究重点就是网络的拓扑关系对调控网络动力学性质的影响。

2.2. 路径稳定性简介

路径稳定性表示网络状态演化路径的重叠性。文献[2]引入了 W 值来表示路径稳定性。具体计算思路如下：假设网络从状态 $S_{i_0}^n$ 开始，经 L 步到达演化终点 $S_{i_L}^n$ ，每一步路径上的路径通量(定义为通过该路径的所有网络状态的数目)记作 T_{i_k} ，

$$\bar{S}_{i_0}^n \xrightarrow{T_{i_0}} \bar{S}_{i_1}^n \xrightarrow{T_{i_1}} \dots \xrightarrow{T_{i_{L-1}}} \bar{S}_{i_L}^n$$

那么，对于状态点 $S_{i_0}^n$ 的 w_{i_0} 值为

$$w_{i_0} = \sum_{k=i_0}^{i_{L-1}} T_{i_k} / L_i$$

进一步，网络整体的 W 值是各个状态 w 值的平均， $W = \langle w_{i_0} \rangle$ 。 W 值体现了网络中各个状态路径的平均重合度。但是，这种对路径稳定性的定义方法容易受到网络规模的影响，因为节点数目多的网络拥有更多的网络状态数，因此，其状态空间中每一条演化路径上的路径通量会偏大。从而导致节点数目多的网络 W 值越大。

为了消除由于网络规模不同导致的网络状态数目的变化及其对路径稳定性的影响，本研究中采用相对重合度来表示路径稳定性，即定义 W 值为

$$W = \frac{\langle w_{i_0} \rangle}{2^n}$$

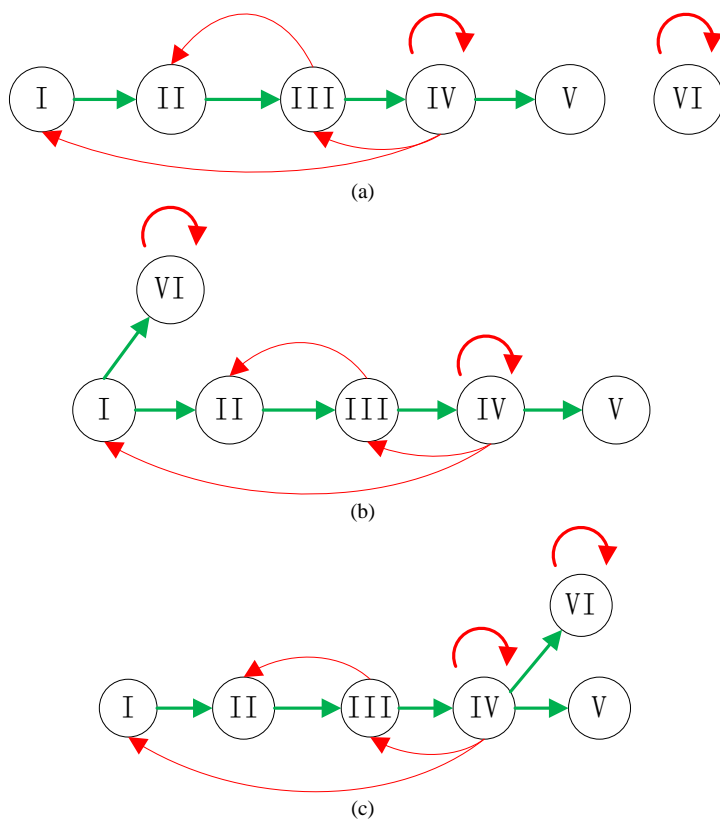


Figure 1. Examples of branches in ITC model; (a) Independent branch; (b) Branch at the start node; (c) Branch at the end node
图 1. 分支链路位置示意图; (a) 独立分支链路; (b) 起点分支链路; (c) 末端分支链路

3. 结果

前期研究表明[8], 真实的调控网络是以 ITC 模型为基本主干, 同时包含多个分支链路的树状结构。在本节中, 我们将通过严格的论证来分析不同位置的分支结构对网络的路径稳定性的影响情况, 揭示在真实调控网络中路径稳定性的形成机制。

结合典型的生物学特征, 我们主要考虑了三种情况的分支链路, 即独立链路(分支链路 with ITC 模块没有激活关系连接), 起点分支链路(分支链路被调控网络的起始节点激活)和末端分支链路(分支链路被调控网络的最后一个传递节点激活)。在真实的调控网络中, 这三种情况都有典型的例子: 例如, Fission yeast 调控网络中的 SK 节点就是独立链路; Budding yeast 网络中的 SBF 节点和 Cln1,2 节点是起点分支链路; 而 Budding yeast 网络中的 Swi5 节点就是末端分支链路。这三种分支链路的具体示意图如图 1 所示。对于在网络中其他位置插入分支链路, 网络路径稳定性的变化情况将在模拟实验中给出。

3.1. 独立链路对 ITC 路径稳定性的影响

独立链路(如图 1(a)所示)与 ITC 模块之间没有激活作用相联系, 因此, 该节点会在自身抑制作用的控制下被关闭。图 2 中对比了加入独立链路前后的网络状态演化路径的变化情况。

对于原网络(即加入链路之前)任意一条起点为 $S_{i_0}^n$, 长度为 L_i 的演化路径, 在加入独立链路之后, 该路径上的每一个状态点都会派生出两个状态点, 即 $(S_{i_k}^n, 0)$ 和 $(S_{i_k}^n, 1)$ 。它们分别对应着独立链路上的节点取 0 或者 1。对于状态点 $(S_{i_k}^n, 0)$, 它们会保持原有的演化路径。而 $(S_{i_k}^n, 1)$ 状态点会在推迟一个单位时间

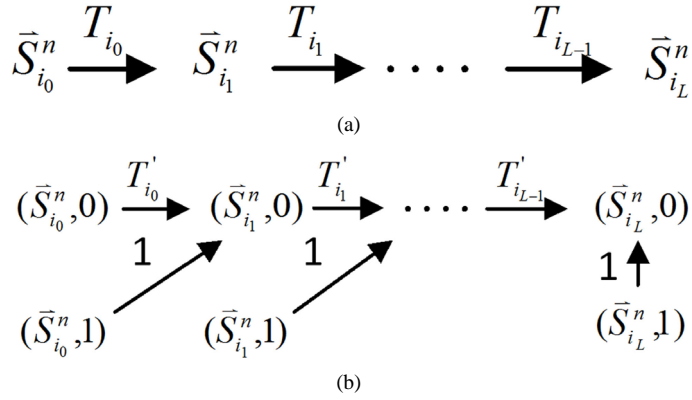


Figure 2. Analysis of the trajectory stability after inserting an independent branch; (a) Before inserting the branch, the trajectory of state evolution and the flow of each link; (b) After inserting the branch, the trajectory of state evolution and the flow of each link
图 2. 独立链路对 ITC 模型的路径稳定性影响分析; (a) 加入独立链路之前, 状态演化路径及各路径通量情况; (b) 加入独立链路之后, 状态演化路径及各路径通量情况

汇入原有的演化路径(如图 2 所示)。对于这种情况, 有如下的结论:

结论 1: 加入独立链路之后, ITC 网络的 W 值降低。

证明: 记 T_k 表示原有演化路径上第 k 步的路径通量, 那么, 加入独立链路之后该通量变为 T'_k

$$T'_k = 2T_k - 1 < 2T_k$$

进一步, 原有状态点 $S_{i_0}^n$ 以及新的状态点 $(S_{i_0}^n, 0)$ 和 $(S_{i_0}^n, 1)$ 所对应的 w 值分别为

$$w_{i_0} = \sum_{k=i_0}^{i_{L-1}} T_k / L_i$$

$$w_{i_0,0} = w_{i_0,1} = \left(\sum_{k=i_0}^{i_{L-1}} T'_k \right) / L_i \quad (1)$$

将公式(1)带入, 即可以推导出 $\frac{w_{i_0,0} + w_{i_0,1}}{2} < 2w_{i_0}$ 。

最后, 加入独立链路口之后, 网络的路径稳定性指标变成

$$W' = \frac{\left\langle \frac{w_{i_0,0} + w_{i_0,1}}{2} \right\rangle}{2^{n+1}} < \frac{\langle w_{i_0} \rangle}{2^n} = W \quad (2)$$

这表明加入独立链路之后, 网络的 W 值下降了。

3.2. 起点分支链路对 ITC 路径稳定性的影响

接下来, 考虑起点处的分支链路(如图 1(b)所示), 分析方法与独立链路的情况相同。

在加入分支链路之前, 依据起始节点的 0~1 状态, 所有的网络状态可以分成两类, 即 $(1, S_{i_k}^{n-1})$ 和 $(0, S_{i_k}^{n-1})$ 。这两类点的一般演化途径及各路径通量由图 3(a)所示。加入分支链路之后, 原有的两类状态派生出四类状态点(依据分支链路节点的 0~1 状态), 即 $(0, S_{i_k}^{n-1}, 0)$ 、 $(0, S_{i_k}^{n-1}, 1)$ 、 $(1, S_{i_k}^{n-1}, 0)$ 和 $(1, S_{i_k}^{n-1}, 1)$ 。这四类状态的演化路径及其通量由图 3(b)所示。

从图中可以看出, 加入分支链路之后, 两条从状态点 $(0, S_{i_k}^{n-1}, 0)$ 和状态点 $(1, S_{i_k}^{n-1}, 1)$ 起始的演化路径替代了原来的主干路径, 而其他的状态点(如 $(0, S_{i_k}^{n-1}, 1)$ 和 $(1, S_{i_k}^{n-1}, 0)$)都会延迟一个单位时间汇入干路径。

而且,这两条主体路径在第 $k+1$ 步(比原来推迟了一个单位时间)发生汇聚。由此分析,可以得出以下结论:

结论 2: 加入起点分支链路之后,网络的路径稳定性指标值会降低。

证明:具体的证明过程类似于独立链路的情形,在此省略。

3.3. 末端分支链路对 ITC 路径稳定性的影响

本小节考虑最后一种情况,即分支链路联接在 ITC 网络的最后一个传递节点上(如图 1(c)所示)。

在没有加入分支链路之前,对于任意一条收敛路径(如图 4(a)所示),其终点状态就是网络的全局吸引子。也就是说, $S_{i_L}^n$ 中仅有最后一个节点表达(取值为 1),而其他所有的网络节点都被关闭(取值为 0)。

在加入分支链路之后,原路径上的每一个状态点都会对应成两个状态,分别为 $(S_{i_k}^n, e_{i_k})$ 和 $(S_{i_k}^n, 1-e_{i_k})$ 。其中, e_{i_k} (取值 0 或者 1)表示主路径上每一个时间点链路节点的状态。它们的演化路径如图 4(b)所示。之所以这么标记是因为加入末端分支链路之后,原有的演化路径发生了变化,无法确定是由主路径上状态点中分支节点的具体取值。但是,可以肯定,无论 e_{i_k} 取值如何,其对应的状态 $(S_{i_k}^n, 1-e_{i_k})$ 会延迟 1 个单位时间汇入主路径。

另一方面,在状态 $S_{i_{L-1}}^n$ 中,网络的最后一个传递节点(也就是分支链路接入的节点)必然是打开状态,这使得 $(S_{i_{L-1}}^n, e_{i_{L-1}})$ 不会直接变成 $(S_{i_L}^n, 0)$,而是先经过 $(S_{i_L}^n, 1)$ 再到 $(S_{i_L}^n, 0)$ 。这一现象说明,原来长度为 L_i 步的演化路径被增加了一个步长(L_i+1 步)。这能够增加单个状态的 W ,从而提高网络整体的 W 值。

结论 3: 加入末端分支链路之后,网络的路径稳定性指标值会增加。

证明:加入末端分支链路之后, $(S_{i_0}^n, e_{i_0})$ 和 $(S_{i_0}^n, 1-e_{i_0})$ 分别对应的 W 值为

$$w_{i_0,0} = w_{i_0,1} = \left(\sum_{k=i_0}^{i_{L-1}} T'_k + T'_{L-1} + 1 \right) / (L_i + 1) \quad (3)$$

所以,对该状态点而言,加入分支链路后 W 值的增加量为

$$\begin{aligned} \frac{1}{2} w_{i_0,0} - w_{i_0} &= \frac{\left(\sum_{k=i_0}^{i_{L-1}} T'_k + T'_{L-1} + 1 \right)}{2 \times (L_i + 1)} - \frac{\sum_{k=i_0}^{i_{L-1}} T_k}{L_i} \\ &= \frac{L_i \times T'_{L-1} + L_i - \sum_{k=i_0}^{i_{L-1}} T_k}{2 \times (L_i + 1) \times L_i} \\ &= \frac{L_i \times T'_{L-1} + L_i + L_i \times T'_{L-1} - \sum_{k=i_0}^{i_{L-1}} T_k}{2 \times (L_i + 1) \times L_i} \end{aligned}$$

由于 $T_{L-1} > 1$, 而且 $\left(L_i \times T'_{L-1} - \sum_{k=i_0}^{i_{L-1}} T_k \right) > L_i - 1$, 故

$$\frac{L_i \times T'_{L-1} + L_i + L_i \times T'_{L-1} - \sum_{k=i_0}^{i_{L-1}} T_k}{2 \times (L_i + 1) \times L_i} > \frac{4 \times L_i - 1}{2 \times (L_i + 1) \times L_i} > \frac{1}{L_i}.$$

从而网络整体的 W 值也相应增加了。

3.4. 模拟实验结果

为了验证上述理论分析的结果,我们进行了随机模拟实验。对给定网络规模(即网络节点数目)的 ITC 模块,在不同的位置加入分支链路,并计算加入链路后的网络路径稳定性(W 值)。典型的调控网络都包含 10 个左右的网络节点,我们的模拟实验将网络主干的节点数定在 6~12 个。给定节点数目,随机选择一个 ITC 链路构成网络主干,然后在指定位置上加入分支链路并计算 W 值。重复上述过程 10,000 次,计

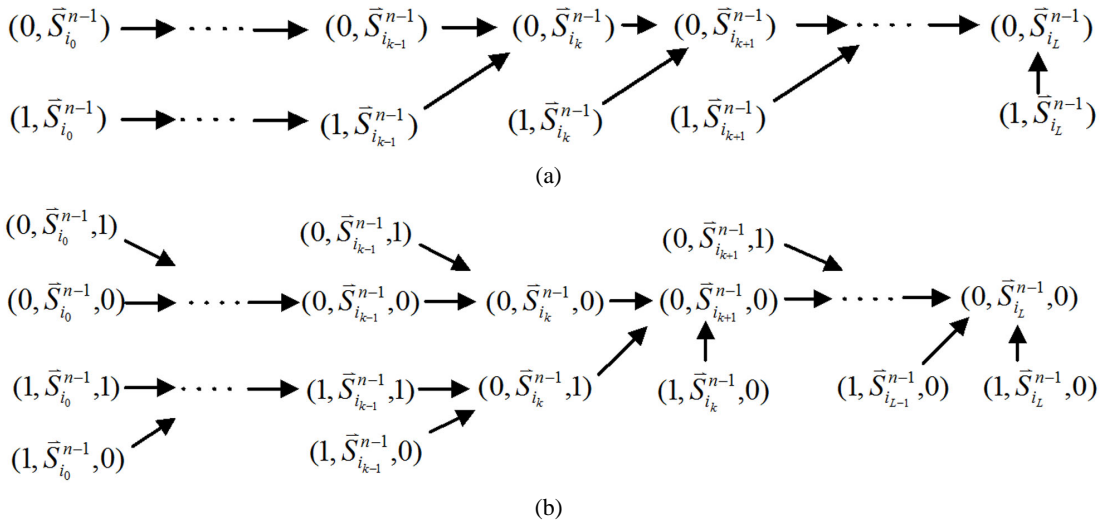


Figure 3. Analysis of the trajectory stability after inserting a branch at the start node; (a) Before inserting the branch, the trajectory of state evolution and the flow of each link; (b) After inserting the branch, the trajectory of state evolution and the flow of each link
图 3. 起点分支链路对 ITC 模型的路径稳定性影响分析; (a) 加入起点分支链路之前, 状态演化路径及各路径通量情况; (b) 加入起点分支链路之后, 状态演化路径及各路径通量情况

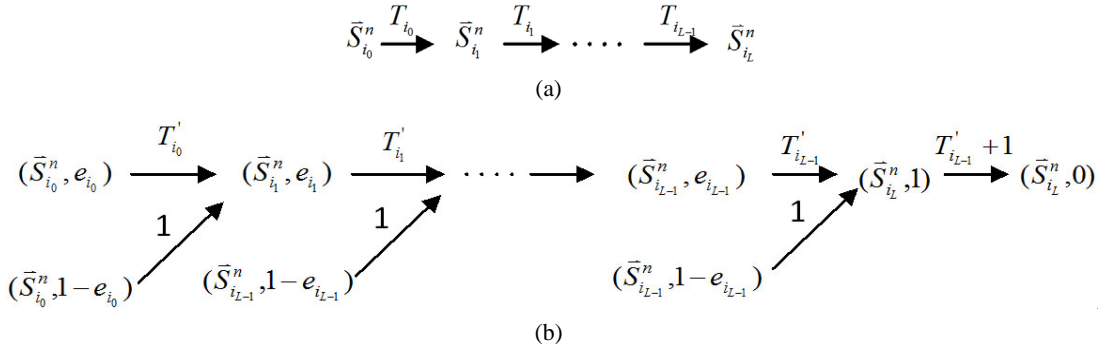


Figure 4. Analysis of the trajectory stability after inserting a branch at the end node; (a) Before inserting the branch, the trajectory of state evolution and the flow of each link; (b) After inserting the branch, the trajectory of state evolution and the flow of each link
图 4. 末端分支链路对 ITC 模型的路径稳定性影响分析; (a) 加入末端分支链路之前, 状态演化路径及各路径通量情况; (b) 加入末端分支链路之后, 状态演化路径及各路径通量情况

算出指定节点数目和指定分支位置下的平均 W 值。总的计算结果如图 5 所示。

横坐标表示接入分支链路的位置, -1 表示未接入分支链路, 0 表示独立链路, 1 表示起始节点, 以此类推。对于 n 个节点的网络, 横坐标为 $n-1$ 就是末端分支链路。从图 5 中可以看出, 随着接入位置的后移, 网络 W 值先降低然后增加。这反映了接入分支链路后的两种作用, 接入位置靠前, 减小路径重合度, 从而降低 W 值; 而接入位置靠后, 演化路径会被延长, 从而导致 W 值会显著增加, 这一点与理论分析的结果完全吻合。

4. 讨论

本文以 ITC 网络模块为基础, 研究了调控网络中分支链路的接入位置对路径稳定性影响情况, 通过理论推导表明, 一方面, 在 ITC 模型的起点或者 ITC 模型之外接入分支链路, 状态演化路径的汇聚时间点被推迟, 从而导致路径重叠度降低, 进而是 W 值减少; 另一方面, 在 ITC 模型的末端接入分支链路,

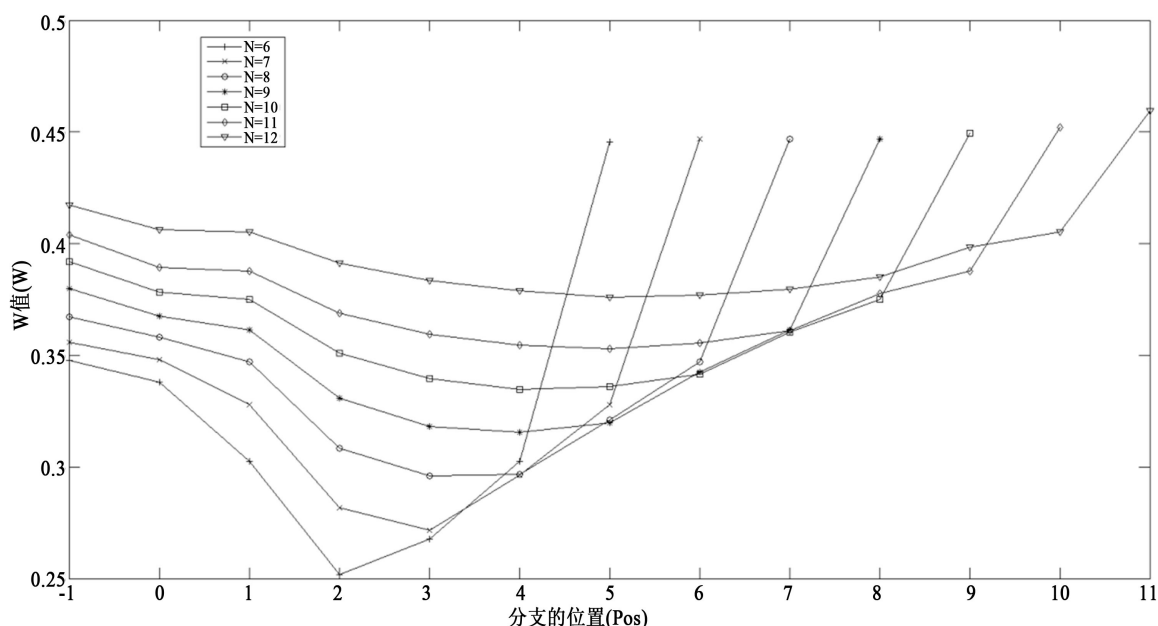


Figure 5. Simulation results of the effect of branch's position on the trajectory stability. Pos = -1 means there is no branch. Pos = 0 means the independent branch. Pos = 1 - n - 1 indicates that inserting branch from the start node to the end node in succession

图 5. 分支链路对路径稳定性的影响模拟分析图。位置为 -1 表示没有接入分支，位置为 0 表示接入独立分支，位置为 1~n-1 表示从网络起点开始依次接入分支

状态演化路径会被延长，从而提高了网络的 W 值。此外，实验模拟的结果也进一步表明在 ITC 模型的中间节点接入分支链路，其对网络路径稳定性的影响是这两种作用的综合。而且，越靠近起点， W 值减少的效果明显，越靠近末端， W 值增加的效果明显。本研究结论不仅为解析路径稳定性的结构成因奠定基础，还能够指导合成生物学研究中设计出具有较高路径稳定性的网络结构。

参考文献 (References)

- [1] Karlebach, G. and Shamir, R. (2008) Modeling and analysis of gene regulatory networks. *Nature Reviews Molecular Cell Biology*, **9**, 770-780.
- [2] Li, F., Long, T., Lu, Y., et al. (2004) The yeast cell-cycle network is robustly designed. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 4781-4786.
- [3] Kitano, H. (2004) Biological robustness. *Nature Reviews Genetics*, **5**, 826-837.
- [4] 欧阳颀 (2012) 系统生物学中的物理问题. *物理*, **1**, 25-30.
- [5] Wang, G., Du, C., Chen, H., et al. (2010) Process-based network decomposition reveals backbone motif structure. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 10478-10483.
- [6] Zhang, Y.J., Ouyang, Q. and Geng, Z. (2014) Refining network reconstruction based on functional reliability. *Journal of Theoretical Biology*, **353**, 170-178.
- [7] Wu, Y., Zhang, X., Yu, J., et al. (2009) Identification of a topological characteristic responsible for the biological robustness of regulatory networks. *PLoS Computational Biology*, **5**, e1000442.
- [8] Zhang, Y.J., Ouyang, Q. and Geng, Z. (2014) Topological origin of global attractors in gene regulatory networks. *Science China Physics, Mechanics & Astronomy*, **58**, 1-8.
- [9] Davidich, M.I. and Bornholdt, S. (2008) Boolean network model predicts cell cycle sequence of fission yeast. *PLoS One*, **3**, e1672.

汉斯出版社为全球科研工作者搭建开放的网络学术中文交流平台。自2011年创办以来，汉斯一直保持着稳健快速发展。随着国内外知名高校学者的陆续加入，汉斯电子期刊已被450多所大中华地区高校图书馆的电子资源采用，并被中国知网全文收录，被学术界广为认同。

汉斯出版社是国内开源（Open Access）电子期刊模式的先行者，其创办的所有期刊全部开放阅读，即读者可以通过互联网免费获取期刊内容，在非商业性使用的前提下，读者不支付任何费用就可引用、复制、传播期刊的部分或全部内容。

