由基因调控网络数据分析揭示振荡斑图的功能结构

张朝阳, 黄旭辉, 郑志刚, 胡岗*

北京师范大学物理学系, 北京 100875

*联系人, E-mail: ganghu@bnu.edu.cn

收稿日期: 2014-04-23; 接受日期: 2014-07-14

国家自然科学基金资助项目(批准号: 11174034, 11075016, 11135001)

摘要 自然和社会系统的许多实际问题都可以用复杂网络动力学来描述, 特别是生物网络和社会网络的应用最为广泛. 人们已经积累了各种网络行为的大量数据, 如何提出合理的物理思想和有效的数学方法来分析这些数据、从中提取有用的信息; 从这些信息出发, 理解在许多实际过程中由复杂的网络拓扑结构产生的简单动力学关系, 并辨识支持各种网络动力学行为的自组织结构, 是一个重要的课题. 本文综述了其中一种研究方法. 以生物学基因调控网络(Gene Regulatory Network, GRN)为研究对象, 提出了一种功能权重(Functional Weight, FW)计算方法. 通过 FW 方法定量计算在 GRN 的振荡动力学中不同调控作用(称为作用边)的重要性. 进一步研究发现在各种 GRN 中这种 FW 的分布普遍极不均匀, 绝大部分作用边具有很小权重(弱作用边), 只有少数边具有大的权重(强作用边). 而这些强作用边构成的子网络比原网络简单得不可比拟, 但它们揭示了网络节点之间的相互作用关系并控制着整个网络的动力学行为. 由这些强作用边构成的网络, 我们可以辨识 GRN 的功能骨架和振荡核心子网络. 振荡核心起到网络振荡源的作用, 而振荡骨架则成为网络中信号传播的主要路径. FW 方法可以广泛应用到不同尺寸、不同调控规则(AND, OR 等形式)和不同动力学行为(周期振荡、混沌动力学和暂态过程)的 GRNs, 并且由 FW 分析出发我们可以对 GRN 动力学进行有效调控.

关键词 动力学, 功能权重, 基因调控网络, 周期振荡, 混沌

PACS: 89.75.Fb, 05.10.-a, 05.65.+b

doi: 10.1360/SSPMA2014-00123

1 引言

近年来复杂网络斑图动力学的研究引起了极大重视, 其原因不仅来源于这一问题在理论上的重要

意义, 还源于问题在各个领域, 特别是在社会与生物等交叉领域的广泛应用^[1,2], 其中基因调控网络是一种极为重要的实例. 在实际问题中往往保存有

引用格式: 张朝阳, 黄旭辉, 郑志刚, 等. 由基因调控网络数据分析揭示振荡斑图的功能结构. 中国科学: 物理学 力学 天文学, 2014, 44: 1319–1333
Zhang Z Y, Huang X H, Zheng Z G, et al. Exploring cores and skeletons in oscillatory gene regulatory networks by a functional weight approach (in Chinese). Sci Sin-Phys Mech Astron, 2014, 44: 1319–1333, doi: 10.1360/SSPMA2014-00123

大量数据, 而所涉及的复杂网络的拓扑结构又极为复杂^[3,4]. 如何充分利用这些信息从复杂的网络结构中找出尽可能简单而易于理解的功能结构, 用以提高对复杂网络斑图动力学的认识、预言及调控能力成为研究复杂网络具有实际意义的重要问题. 本文针对基因调控网络建议了一类有用的分析方法.

振荡是许多实际网络的一类普遍动力学行为. 基因调控网络(Gene Regulatory Networks, GRNs)的振荡经常控制着一些重要的生物节律, 对不同的生物学功能发挥着不可或缺的重要作用^[5-9]. 数学上, GRN 的动力学常用耦合常微分方程来表示^[10-12], 网络中的单个节点由方程中的一个变量来描述, 而基因之间的相互调控则由网络节点间的相互作用来体现. GRN 通常有大量节点和极多且复杂的相互作用, 使得 GRN 的斑图动力学行为显得很复杂、似乎杂乱. 这些大量相互作用在实际动力学过程中常常发挥非常不同的调控作用^[13-18], 有的作用可能非常关键, 而另一些作用则可能不那么重要, 而这些重要性的区别则常常隐藏在看似复杂的动力学数据中, 不被人们所轻易观察到. 本文综述了一种计算各种相互作用重要性的方法, 称为功能权重(Functional Weight, FW)计算方法. 运用这种方法可以从已有振荡网络的动力学数据出发定量计算每一条相互作用边的重要性程度, 并依据权重大小从大量相互作用中找出起关键驱动功能的少部分作用边. 简化的网络揭示了产生给定振荡斑图的关键功能结构, 加深对这些简化结构的认识有助于理解网络动力学的起因以及信号在网络中的主要传播路径, 并在此基础上可根据需要对网络动力学进行有效调控.

2 GRN 模型及功能权重计算原理和方法

我们首先从系统生物学中常用的 GRN 模型出发, 这里 GRN 可被描述为耦合的常微分方程组^[19-22] (Ordinary Differential Equations, ODEs).

$$\frac{dp_i}{dt} = f_i(\mathbf{p}) - p_i, [\mathbf{p}] = (p_1, p_2, \dots, p_N), \quad (1)$$

$$f_i(\mathbf{p}) = \begin{cases} A_i(\mathbf{p}), & \text{只有正调控,} \\ R_i(\mathbf{p}), & \text{只有负调控,} \\ A_i(\mathbf{p})R_i(\mathbf{p}), & \text{两种形式联合调控,} \end{cases}$$

$$A_i(\mathbf{p}) = \frac{act_i^{h_i}}{act_i^{h_i} + K_i^{h_i}}, R_i(\mathbf{p}) = \frac{K_i^{h_i}}{rep_i^{h_i} + K_i^{h_i}},$$

$$act_i = \sum_{j=1}^N \alpha_{ij} p_j, rep_i = \sum_{j=1}^N \beta_{ij} p_j, i, j = 1, 2, \dots, N,$$

这里 p_i 是基因 i 表达的蛋白质浓度, 通过适当标度化其范围限定为 $0 \leq p_i \leq 1$, N 是系统大小即调控网络中的基因个数. (1)式中忽略了漏转录速率的影响, 并假定所有蛋白质的降解速率相同. 矩阵 α_{ij} , β_{ij} 表示调控关系(拓扑耦合强度), 如果 j 正调控 i (促进作用)有 $0 < \alpha_{ij} < 1$; 如果 j 负调控 i (抑制作用)有 $0 < \beta_{ij} < 1$; 而在 j 不参与 i 的调控时有 $\alpha_{ij} = \beta_{ij} = 0$. act_i 和 rep_i 分别表示基因 i 所受到的正调控和负调控转录因子作用的总和. 基因调控形式用希尔函数(Hill Function)来表达, 其中希尔指数(Hill Coefficient)取值范围为 $1 < h_i < 10$, 而活化/阻抑系数(Activation/Repression Coefficient) $0 < K_i < 1$. 所有这些参数区间处于生物学中常用的参数范围之内.

方程(1)的一个最基本的特点是: 如果没有相互作用, 单个变量自身是不会有振荡并且总会单调地趋近于固定浓度值, 任何节点运动的复杂性均来源于其他节点对该节点的相互作用与该节点自身动力学的耦合, 而这些相互作用都包含在场 $f_i(\mathbf{p})$ 之中. 进一步研究可知, 任何变量产生振荡的变化 $\frac{dp_i}{dt}$ 来源于与 i 相互作用的相关节点的浓度变化, 可用变分形式表示为

$$\dot{f}_i^c(\mathbf{p}) = \sum_{j=1, j \neq i}^N \dot{f}_{ij} = \sum_{j=1, j \neq i}^N \frac{\partial f_i}{\partial p_j} \frac{dp_j}{dt}, \quad (2)$$

这里 $\dot{f}_i^c$ 表示由其他节点浓度变化引起 f_i 的总变化率. (2)式的一个非常有用的特点是节点 i 的总变化率 $\dot{f}_i^c$ 表达为各个邻居贡献作用之和, 其中每个邻居的贡献都可以分别计算. 在计算节点 i 振荡中, 第 j 个节点所作贡献的权重可计为

$$w_{ij}(t) = \frac{|\dot{f}_{ij}|}{\sum_{j=1, j \neq i}^N |\dot{f}_{ij}|} = \frac{\left| \frac{\partial f_i}{\partial p_j} \frac{dp_j}{dt} \right|}{\sum_{j=1, j \neq i}^N \left| \frac{\partial f_i}{\partial p_j} \frac{dp_j}{dt} \right|}, \quad (3)$$

即为 t 时刻雅克比矩阵乘以权重 $\frac{dp_j}{dt}$ 之后归一化权值. 这个权重是对 t 时刻计算的, 而在一个周期中节

点 j 对节点 i 振荡贡献的整体权重可用平均值计算为

$$w_{ij} = \frac{1}{T} \int_0^T w_{ij}(t) dt, \quad (4)$$

其中 T 为振荡周期; 对于混沌运动 T 则需要取更长时间来求平均. w_{ij} 代表在节点 i 运动中节点 j 所作贡献的功能权重(Functional Weight, FW). 所谓功能权重是对动力学过程而言, 它代表在某个动力学斑图中某条边所涉及的权重. 对同一个拓扑网络不同的动力学过程可计算出不同的功能权重.

w_{ij} 具有很好的半正定性 ($0 \leq w_{ij}$) 和归一性 ($\sum_{j=1, j \neq i}^N w_{ij} = 1$). $w_{ij}=0$ 或 $w_{ij} \ll 1$ 代表 j 对 i 运动的作用较弱, 而 w_{ij} 接近或等于 1 则表示在所研究的振荡态中 j 对 i 运动起决定性的控制和驱动作用.

3 GRN 振荡态的运动骨架(Skeletons)和振荡核心(Cores)

有了上述的计算方法, 我们就可以来实际计算相互作用的 FW 值. 计算之中还有两个重要问题有待回答. 其一是, 由(3)和(4)式计算出的 w_{ij} 是否真的能够量度 j 对 i 作用重要性; 其二是, 即使其一为真, w_{ij} 的分布是均匀(即如 i 受其他 k 个基因的调控, 而 k 个邻居的作用权重都接近)还是不均匀. 如 w_{ij} 分布均匀, 这一 FW 分析方法对认识和调控 GRN 振荡的作用就不大. 而对第二个问题的回答与计算方法的正确与否无关, 完全由所研究的复杂网络自身的动力学性质决定.

图 1(a)中一个 6 节点($N=6$)15 条边($I=15$)的 GRN 网络用来作为 FW 方法研究的实例. 连线粗细正比于耦合强度 α_{ij} , β_{ij} , 此例中所有耦合强度相同; 实线表示正调控, 虚线代表负调控. 在 $\alpha_{ij}=\beta_{ij}=1$, $h_i=h=3$, $K_i=K=0.10$ 时网络自持续振荡. 在给定的参数条件下, 该调控网络具有唯一的周期振荡态吸引子. 把各节点的振荡数据代入(3)和(4)式, 我们可以对该网络中所有相互作用画出 FW 权值图(图 1(b)). 图中连线粗细正比于权重大小, 图中一个非常有趣且实际上非常有用的特点是该网络 FW 的分布极不均匀. 大部分相互作用权值接近零或在计算精度内等于零, 代表这些相互作用对所作用的基因调控作用非常微弱, 而少数的相互作用具有接近于 1 或在精度内等于 1 的权值, 代表该相互作用对作用节点的关键调控和对该

图 1 (网络版彩图)利用公式(3)和(4)对一个振荡 GRN(a)计算功能权重(FW)的实例(b), 可以得到较强相互作用边构成的振荡骨架((c), skeleton)和不可约化的核心结构((d), core). 核心子网络是原网络的振荡源, 振荡信号主要通过骨架网络传播. 约化子网络产生的动力学轨道很好地重复了原复杂 GRN 的轨道(e), (f). 引自文献[23]

Figure 1 (Color online) An example of drawing functional weight (FW) map (b) computed by eqs. (3) and (4) for an oscillatory GRN (a). Interaction skeleton (c) and irreducible core structure (d) can be obtained from (b), which serves as the oscillation source generating waves propagating mainly through the paths of the skeleton. The agreements between the orbits of the full GRN and that of the reduced subnetworks are satisfactory (e), (f). Quoted from ref. [23].

节点振荡的决定性驱动作用. 本节开始时提出的第 2 个问题在图 1 振荡网络中得到极为满意的回答.

由图 1(b)的 FW 分布, 我们可以来进行有效的网络复杂性约化(Complexity Reduction). 将图 1(b)中所有 $w_{ij} \leq 0.30$ 的相互作用除去, 在剔除这些较弱的相互作用时, 将其平均值 $\left\langle p_j \right\rangle = \frac{1}{T} \int_0^T p_j(t) dt$ 保留在相互作用中以避免改变网络的参数条件. 在进行上述操作并保留较强相互作用边之后原网络图 1(a)约化为子网络图 1(c). 该网络是原网络的子网络, 它揭示了该 GRN 振荡态中基因调控之间的主要驱动关系及

原网络中信号传播路径, 称为振荡斑图的功能骨架 (Skeleton). 图 1(c) 的约化只针对相互作用, 图中保留了原网络的所有节点, 而这些节点在功能网络中发挥的作用也互不相同, 有些节点只受其他节点驱动, 而它们不在骨架中驱动别的节点, 显然只具有被动作用. 将图 1(c) 中只有输入没有输出的节点逐个剔除最后达到一个不可约化的子网络, 该子网络就是图 1(d). 图 1(d) 的子网络中每个节点都具有输入和输出, 它们对网络的全局振荡起着决定性的作用, 称为振荡 GRN 的核心结构(Core). 这个核心结构就是原网络(图 1(a))的振荡源. 整个 GRN 的振荡斑图现在可以理解: 核心结构(图 1(d))产生了振荡, 而振荡信号主要以骨架(图 1(c))通道控制了整个网络的振荡.

图 1(b)–(d) 回答了上述第 2 个问题, 而第 1 个问题的回答还需要进行检验. (3) 和 (4) 式 w_{ij} 计算有效性的检验则要通过确认用 FW 计算出来的重要(不重要)相互作用是否真的重要(不重要)来实现. 而最好的检验计算是看图 1(c) 和 (d) 中由重要相互作用边和节点构成的网络产生的动力学能否重复原复杂网络图 1(a) 的主要动力学特征. 我们在图 1(e) p_1 – p_2 二维相平面中比较了图 1(d) 的核心子网络(3 节点 3 条边)与图 1(a) 的原网络产生的运动轨道, 两者很好地重叠. 图 1(f) 在 p_4 – p_6 相平面中比较骨架图 1(c) 与原网络图 1(a) 的轨道, 二者重要特征也很好吻合. 就图 1(a) 的振荡网络而言, 前面提出的两个问题都得到了肯定回答.

图 1 仅讨论了一个 6 节点的小网络. 真正的复杂性出现在具有多个节点和大量相互作用边的网络动力学之中, 这使复杂性约化问题成为真正有意义并且具有挑战性的问题. 图 2(a) 是一个 50 节点 250 条边的 GRN 网络, 其连边随机选取, 每个节点有 5 条随机输入边(正负调控等概率随机分配), 参数按方程(1)所列范围进行随机选取. 其网络拓扑结构看起来很复杂, 节点相位分布也很杂乱. 用图 1 同样的方法, 我们可以用网络的振荡数据由(3)和(4)式计算所有节点连线(即相互作用)的功能权重, 在图 2(b)中 w_{ij} 的大小正比于作用线的粗细. 图中可以看出粗线的数目比起总的拓扑连线数要少得多. 图 2(c) 保留了图 2(b) 中所有 $0.30 < w_{ij}$ 的作用边得到了振荡骨架子网络, 振荡信号的传播路径已经清楚显示在图中. 对图 2(c) 进行如图 1(c) 同样方法的约化, 可以得到图 2(a) 振荡网络的核心结构——振荡核心(图 2(d)). 非常有趣的现象是尽管骨架(图 2(c))和核心结构(图 2(d))的子网络比原

图 2 (网络版彩图)(a)–(f)与图 1(a)–(f)完全相同, 只是采用了更复杂的 GRN. 我们看到了通过 FW 计算(b), 可以得到极为简化的骨架((c))和核心子网络(d), 并且这些简化网络可以如(e)和(f)那样再现原复杂 GRN 的动力学. 引自文献[23]

Figure 2 (Color online) (a)–(f): the same as Figures 1(a)–(f), respectively, with the oscillatory GRN ($N=50$ and $I=250$) in (a) considered. It is remarkable that the greatly simplified compact core and skeleton structures (which can be obtained by using the simple and unified computation rules of eqs. (3), (4)) can well reproduce the oscillatory dynamics of a very complicated network (a). Quoted from ref. [23].

网络(图 2(a))简单得不可比拟, 但它们很好地重复了原网络的斑图动力学行为(见图 2(e)和(f)).

4 振荡 GRNs 中各种不同特征的骨架和核心子网络

上节的分析指出权重计算方法能够实际度量网络中连边作用的重要性程度, 并且在 GRN 振荡斑图中权重值 w_{ij} 的分布极不均匀, 这为我们提供了一种从振荡数据分析揭示 GRN 中产生和传播振荡信号的

关键子网络(骨架和核心)的途径. 在骨架和核心的分析中另一个重要问题是在 GRN 振荡斑图中有哪些不同类型的核心结构以及它们是怎样决定了 GRN 的不同运动形态.

图 3 给出了另一个随机构建的 $N=50, I=250$ 的振荡 GRN, 该网络具有周期振荡吸引子, 对振荡数据进行分析后得到图 3(a)的运动骨架子网络. 有趣的现象是在骨架中 50 个节点分成了两个互不相联的集团(互不相联是指没有功能权重大于 0.3 的连边, 而在网络拓扑上两个集团骨架则是紧密地相互连接的), 而每个集团都被自己独立的核心子网络(图 3(b))所控制. 这样的集团结构到底是否真实存在, 可以通过轨道比较的方法来验证. 图 3(c)和 3(d)分别用两个核心子网络独立产生的轨道与原网络整体运行产生的轨道相比较, 核心子网络的轨道与原整体网络轨道具有高度一致性, 验证了在原 GRN 振荡中相对独立的振荡集团(Clusters)的存在.

图 4 给出了另一个有趣的振荡态的例子($N=50, I=250$). 研究该网络不同节点的振荡运动发现, 网络中节点的运动性质可区分成两类: 一类节点具有简单的周期运动如图 4(a)所示, 而另一类运动则明显复杂得多, 具有多峰结构如图 4(b)所示. 在图 4(c)的 GRN 中我们画出了这两类节点, 它们在网络中的分

图 3 (网络版彩图)(a)–(d)与图 2(c)–(f)的作法完全相同. 该振荡网络中存在两个独立的核心子网络, 支配了两个独立节点集团(clusters)的振荡运动. 从(c)和(d)可以看出两个“独立”的核心子网络的确独立地复制了原整体复杂网络相应节点的动力学行为. **Figure 3** (Color online) The same as Figures 2(c)–(f). Two independent cores appear in the network dominating the oscillatory dynamics of two independent clusters. The cores can well reproduce the oscillatory dynamics of a very complicated network.

布似乎杂乱无章. 用 FW 计算(3)和(4)式的方法算出 w_{ij} , 并用图 1–3 相同的方法构建约化的骨架(图 4(d))和核心(图 4(e))子网络. 在子网络中两类节点非常清楚的区分开来. 图 4(e)的核心网络具有多层的环(Loop)结构. 其中四节点主动环(图 4(e)中四节点负环 $3 \cdots 45 \rightarrow 40 \cdots 43 \cdots 3$)在骨架中只驱动其他节点, 而不受环外其他节点的作用. 这个独立负环(所谓负环是指环中负调控个数为奇数, 人们已经证明 GRN 的所有振荡模式必须来源于负环结构^[24])会产生简单振荡行为, 所有受这一主动环直接控制的下游会呈现简单周期运动. 而在图 4(e)核心结构中主动环下游的节点构成新的负环结构, 又会产生新的振荡模式. 主动环和下游环产生振荡模式的叠加就形成了多模的复杂运动形式, 所有复杂运动形式的节点在骨架中都处在多环控制的下游, 受多模核心振荡的驱动. 这里仅有节点 21 是个例外, 在骨架中它被主动环控制, 但具有复杂运动模式. 仔细研究发现节点 21 还受到来自节点 10 的权重 $w=0.28$ 的负调控, 这一调控是节点 21 复杂运动的根源, 但由于它的权重稍低于阈值在振荡骨架构建中被剔除. 图 4(f)和(g)的轨道比较显示, 极为简化的骨架和核心子网络能成功区分并同时复制原复杂 GRN 产生的明显具有不同复杂性特点的运动轨道.

网络拓扑及运动斑图的复杂性随节点数 N 和相互作用 I 的增加而急剧增加. 实际网络系统常常具有很大的尺寸, 用复杂得多的大网络系统来检验 FW 权重分析方法的有效性就显得很重要. 图 5 分别对两个 $N=400, I=2000$ 的大得多的 GRN 振荡例子进行了研究. 用与图 1–4 同样的方法我们仍然可以找到相当简单的核心结构(图 5(a)和(c)), 并发现这些与原网络相比简单得不可比拟的子网络仍然可产生与原网络定性一致、定量具有很好近似的节点的运动轨道行为.

图 3–5 对不同集团结构、运动形态和复杂程度的 GRN 振荡行为进行了 FW 分析和适用性检验, 其结果说明有效的数学和物理方法可以从大量的动力学数据中找到有用的信息, 并从看似复杂的网络拓扑结构中得到一些易于理解的简化功能结构.

5 GRN 振荡态 FW 分析方法的统计研究

第 2 和 3 节中列举各种例子显示了 FW 方法对分析 GRN 振荡行为的应用, 但若干孤立的例证难以说明

图 4 (网络版彩图) $N=50$, $I=250$ 的 GRN 振荡态的另一个例子

网络中呈现两类运动性质的节点, 一类具有简单周期运动(a), 仅被单个主动负环驱动, 另一类具有复杂的多峰周期运动(b), 被多个负环控制, 具有多模的复杂振荡形式. 极为简化的骨架和核心子网络能同时复制原复杂 GRN 的简单和复杂运动

Figure 4 (Color online) There are two types of oscillations in the network: one with simple period driven by single negative feedback loop and the other with multiple-peak controlled by different oscillatory modes. It is remarkable that the greatly simplified compact core and skeleton structures can well reproduce all the motions of the original network (a).

该方法的成功应用. 要在没有严格数学证明的情况下可信地验证 FW 方法的有效性必须使用各种统计方法. 本文将用系统的统计来分析讨论 FW 方法对 GRN 振荡斑图应用的有效性.

首先我们在图 6 中分析振荡 GRN 中核心结构的统计数据. 图 6(a)和(b)是 $N=20$, $I=100$ 的 GRN 中用图 1 相同的方法得到的核心子网络所含的不同节点数 N_c 和相互作用边数 I_c 出现的概率. 由图可见绝大多数核心子网络集中在 3–6 的少节点数并且连边较少的区段(注意网络总边数 $I=100$). 图 6(c)统计了核心网

图 5 (网络版彩图)两个 $N=400$, $I=2000$ 的自持续振荡 GRN 的例子

(a)和(c)分别为两振荡态的不可约核心子网络. (b)和(d)核心子网络中节点轨道与相应原网络轨道的比较. 尽管子网络与原网络相比有了不可比拟的简化, 但它们仍能较好地复制原网络相应节点的轨道. 这验证了 FW 分析方法在大网络中应用的有效性

Figure 5 (Color online) Two examples of large GRNs. (a), (c) Irreducible core structures. (b), (d) Comparisons of a dynamic orbit of the original GRN with that of core in a 2D phase plane. The agreements between the orbits of the full GRN and that of the greatly reduced subnetworks are satisfactory. Thus, FW method can be effectively applied to large GRNs.

络中所含有的环数(J), 绝大多数集中在 $J \leq 4$ 的范围内. 有趣的是由图 6(a)–(c)可见三节点单环核心在所有各种不同核心中具有最大比例. 图 6(d)给出最有意义的结果: 原网络中长度为 L 的环出现概率由方块表示, 在核心结构中长度为 L 的环概率由圆点代表. 核心环长 $L \leq 9$ 出现的概率大于 99.5%, 而原网络中这些短环出现的概率小于 0.5%. 即 GRN 振荡非常鲁棒地并自组织地选择了网络中以极小概率出现的短环作为自己的振荡源.

为了定量地计算骨架和核心网络再现原网络动力学的有效性, 我们提出了一个标度化的误差率 (Scaled Error, SE) 的量. 其计算方法如下: 首先用相位 ($\phi = 2\pi t/T$) 来统一描述各种振荡变量. 分别定义 $\Delta p_i = p_i - \bar{p}_i$, $\Delta \hat{p}_i = \hat{p}_i - \bar{\hat{p}}_i$ 为原网络和约化子网络第 i 个节点的浓度量的振荡变量 ($\bar{p}_i$, $\bar{\hat{p}}_i$ 分别为 p_i ,

图 6 (网络版彩图) $N=20, I=400$ 的振荡 GRN 中核心子网络所含有的节点、边和环的数目及环长的统计

5000 个振荡 GRN 样本用于所有概率统计. 网络振荡动力学自组织到以短环为振荡源的振荡态中. 引自文献[23]

Figure 6 (Color online) Some statistics of core structures (nodes, links and loops) with $N=20$ and $I=100$. 5000 oscillatory samples are applied in the statistics. These core loops with small lengths are self-functionally selected to dominate the pattern dynamics of GRNs. Quoted from ref. [23].

$\hat{p}_i$ 的平均值). 对所有振荡变量 $\phi=0$ 都取在其最大值上. 这样可以对网络所有节点计算平均绝对误差 (Average Absolute Error, AAE)

$$AAE = \frac{1}{2\pi N} \sum_{i=1}^N \int_0^{2\pi} |\Delta p_i - \Delta \hat{p}_i| d\phi, \quad (5)$$

AAE 代表了网络中绝对误差总和. 这一误差可用总的振荡强度 (Total Oscillation Strength, TOS)

$$TOS = \frac{1}{2\pi N} \sum_{i=1}^N \int_0^{2\pi} (|\Delta p_i| + |\Delta \hat{p}_i|) d\phi \quad (6)$$

来重新标度为标度化的误差 (SE)

$$SE = \frac{AAE}{TOS}. \quad (7)$$

SE 的取值在 $[0,1]$ 区间, 量度约化子网络与原网络动力学行为的整体偏差. $SE=0$ 代表子网络与原网络动力学的完全一致, 而 $SE=1$ 代表完全不符 (原网络振荡而约化子网络不振荡).

图 7(a) 计算了 $N=20, I=100$ 阈值 $w_{th}=0.30$ 时, 用骨架网络计算的 SE 的概率分布图 (黑方块). 在计算中采用了 5000 个对拓扑、参数、初值随机取值的振荡样本. 这里在骨架构建中平均剔除了 76.1% 的相互作用边. 与此同时我们选择了两个对比组, 分别在同样网络中随机剔除了 30.0% (圆块) 和 60.0% (三角) 的连边. 对比发现, 尽管在 FW 计算中骨架剔除的连边对比组更多, 骨架组的 SE 分布都在 SE 较小的区段,

且很少有振荡死亡 ($SE=1$ 在 20% 以下). 而对比组在去边数远小于骨架组的情况下出现大比例的振荡死亡 (对 60% 随机去边组, $SE=1$ 在 95% 以上). SE 的统计分布显示了 FW 分析约化方法的有效性.

在 FW 分析中另一个问题是阈值 w_{th} 的选择. 在这之前的讨论中都采用了 $w_{th}=0.30$, 那么 FW 分析方法对 w_{th} 的选择是否具有鲁棒性, 并且什么样的阈值选择最为优化? 图 7(b) 画出了对不同阈值 $w_{th}=0.20, 0.25, 0.30$ 条件下, $N=(20, 50, 100), I=(100, 250, 500)$ 的 GRN 振荡网络的骨架的平均边数占总边数的比例. 图中充分显示了权重 w_{th} 分布的极其不均匀性. 在各组数据中 $w_{th} \leq 0.10$ 弱权重边数占 70% 以上的概率, 而 $w_{th} \geq 0.40$ 边数占的比例仅有 20% 左右 (即平均每个节点一条连边). 有趣的是这一不均匀性的特点在各种尺寸的 GRN 中都存在. 而且不均匀性随 $[N, I]$ 的增大而增加. 这为我们在 $w \in [0.1, 0.4]$ 之间鲁棒地选取 w_{th} 阈值, 有效地构建既比原网络极大约化又能理想地重复原网络动力学, 并帮助人们透彻地理解运动机制的骨架和核心子网络提供了坚实的基础. 图 7(c) 分别对 $w_{th}=0.20, 0.25, 0.30$ 的不同阈值作了骨架子网络的 SE 概率分布图. 三种分布非常接近, 并且它们都根本区别于图 7(a) 中的随机去边网络的 SE 分布. 我们在图 7(d) 中画出了功能权重 w_{ij} 的概率分布. 权重分布的极度不均匀性很好地解释了图 7(b) 和 (c) 的行

图 7 (网络版彩图)(a)–(c)用(7)式对不同约化子网络计算误差 SE 分布概率的统计结果. (d)–(e) FW 分布的统计结果. 引自文献[23]

Figure 7 (Color online) (a)–(c) Statistical results of scaled error (SE) defined in (7) for different reduced subnetworks. (d)–(e) Probability distributions of functional weight. Quoted from ref. [23].

为. 图 7(e)画出了每条边的功能权重 w_{ij} 与它对应的拓扑耦合强度 α_{ij} , β_{ij} 的关联. 图中显示的相对均匀的分布说明功能权重的强烈不均匀分布是由动力学自组织形成的而不是拓扑耦合强弱所产生的结果.

定义 $SE \leq 0.50$ 作为子网络能否正确重现原网络动力学标准, 我们可以计算各种不同方法构造子网络的“正确率”. 表 1 对 $N=20, 50, 100, l=5N$ 三种尺寸 GRNs, 分别统计 $w_{th}=0.20, 0.25, 0.30$ 三种骨架子网络(Su_S)和随机去边(Su_R)30%和 60%子网络的正确率. 每组数据都由 5000 个随机抽取的不同样本(原网络均为自持续振荡态)计算得出. 从表 1 看出尽管 FW 分析构建的骨架平均去边数远大于随机去边数, 但骨架的正确率远大于随机去边子网络. 而且有意义的现象是随(N, l)的增大在所有骨架子网络的计算中留边率单调减少, 但正确率反而单调增加. 而反观随

表 1 方程(1)不同子网络重现原网络动力学正确率的统计值. 引自文献[23]

Table 1 Probabilities for different subnetworks to success-fully reproduce the oscillatory dynamics of the full GRNs of eq. (1). Quoted from ref. [23]

Network	$Su_S(0.20)$	$Su_S(0.25)$	$Su_S(0.30)$	$Su_R(30\%)$	$Su_R(60\%)$
$N=20$	87.2%	84.5%	80.6%	24.2%	3.7%
$N=50$	88.5%	85.0%	81.5%	18.6%	2.0%
$N=100$	90.2%	87.3%	85.4%	15.5%	1.8%

机去边样本, 在去边留边率不变的情况下, 正确率随(N, l)增加而快速减少. (N, l)越大 FW 分析的优势越加明显.

方程(1)在促进和抑制的联合调控中采用了 AND 模式, 这种模式的确存在于一些实际基因调控网络中. 同时也有许多基因采取不同的调控模式, 如以下的 OR 调控模式也被广泛使用^[5,7].

$$\frac{dp_i}{dt} = \gamma_i + \sum_{j=1}^N f_{ij} - p_i, \quad i = 1, 2, \dots, N, \quad (8)$$

$$f_{ij} = \begin{cases} 0.0, & j \text{ 对 } i \text{ 没有作用} \\ (1.0 - p_i) \frac{p_j^{h_{ij}}}{p_j^{h_{ij}} + K_{ij}^{h_{ij}}}, & j \text{ 正调控 } i \\ -p_i \frac{p_j^{h_{ij}}}{p_j^{h_{ij}} + K_{ij}^{h_{ij}}}, & j \text{ 负调控 } i \end{cases}$$

其中 $0 \leq p_i \leq 1$ 代表第 i 个基因表达的蛋白质浓度范围, 参数取 $1 < h_i < 10$, $0 < K < 1$ 为实际常用生物 Hill 指数和活化/阻抑系数范围. 漏调控转录率 (Leakage Transcription Rates) 取 $0 < \gamma_i < 0.1$, 所有 p_i 的降解率都标度化为 1.

图 8(a)–(d) 以 (8) 式为模型作了图 6(a)–(d) 对模型 (1) 所作的完全相同的统计计算. 图 8 在不同模型中完全再现了图 6 中显现的 GRN 网络 ($N=20$, $I=100$) 中核心结构 (振荡源) 的统计行为.

表 2 中对 (8) 式模型作了与表 1 完全相同的子网络“正确率” S_u 的计算. 用 FW 构建的骨架子网络的成功率甚至比表 1 的成功率还高, 比随机去边子网络的优势更有相当的扩大. 特别是 FW 方法的有效性随 (N , I) 的增加而逐步提高的结论仍然被各类测试所普遍验证.

图 8 (网络版彩图)(a)–(d) 采用了与图 6(a)–(d) 完全一样的计算, 只是这里采用了方程 (8) 的节点调控动力学方程

本图中观察到的现象和得到的结论与图 6 完全一致, 证实了 FW 方法对不同 GRN 模型的普遍适用性

Figure 8 (Color online) The same as Figures 6(a)–(d) by applying the regulatory dynamics of eq. (8). The FW method can be applied to different models of GRNs with the same conclusion as Figure 6.

6 FW 方法对 GRN 混沌振荡的应用

到上节为止, 本综述一直在讨论 GRN 的周期振荡斑图动力学问题. FW 方法不仅可应用于周期振荡网络, 也可应用于混沌基因调控网络. 在非线性动力系统中混沌的存在极为普遍, 例如三变量自治非线性系统就很容易在一些参数条件下发现混沌. 以此为参考, 人们期待在基因调控网络和其他生物系统中应该广泛观察到混沌行为. 然而由于实验数据有限混沌的判断并不容易, 但是在一些清楚的前兆行为的观察条件下混沌行为还是可以相当可信的确定. 例如在病态的心肌组织中会观察到电信号或离子信号中的倍周期分岔行为 (称为 Altanans), 而在这种分岔之后出现的具有混乱特点的心肌活动能够比较可信地断定为心肌动力系统的混沌或缺陷湍流, 而这种行为对心脏正常功能有致命的影响. 另外, 基因调控和许多生物学过程可以很好地用数学模型来描述, 例如本文中的方程 (1) 和 (8). 在这些模型方程中准确揭示混沌现象的规律, 对实际系统混沌动力学的理解、预言和调控就具有启示或指导意义.

在 GRN 研究中人们常常使用模体 (Motifs) 的概念, 模体所指的是具有某种生物功能而在生物网络中以远大于在随机网络中出现的概率存在的小的子

表 2 方程(8)不同子网络重现原网络动力学正确率的统计值. 引自文献[23]

Table 2 Probabilities for different subnetworks to successfully reproduce the oscillatory dynamics of the full GRNs of eqs. (8). Quoted from ref. [23]

Network	$Su_S(0.20)$	$Su_S(0.25)$	$Su_S(0.30)$	$Su_R(30\%)$	$Su_R(60\%)$
$N=20$	91.4%	89.5%	86.3%	16.5%	1.4%
$N=50$	92.0%	90.9%	89.7%	14.8%	1.1%
$N=100$	93.5%	92.2%	90.8%	12.1%	0.9%

图 9 (网络版彩图)节点数小于等于 5 的全部振荡模体
它们都是节点数少于等于 5 的不可约的单负环结构. 引自文献 [25]

Figure 9 (Color online) All the oscillatory motifs with $N \leq 5$. They are the structures with single irreducible negative feedback loops. Quoted from ref. [25]

网络. 振荡 GRN 中也存在振荡功能模体, 例如在图 1–5 的核心子网络中我们就看到了起振荡源作用的各种功能振荡模体. 所有振荡模体都应以各种负反馈环存在, 图 9 中展示了节点数少于等于 5 的所有振荡模体. 这些模体以单独或结合的形式高概率地出现在图 1–5 的核心子网络中.

相比图 9 的振荡模体, 混沌模体应有更复杂的结构, 因为混沌运动可以粗略地理解为不同(至少两种)振荡模式竞争的产物. 把混沌模体(Chaotic Motifs)定义为能产生混沌运动的最小不可约的单元, 那么该单元一定要满足一些网络拓扑的要求. 例如, (i) 混沌模体必须包含两个环结构以支持不同振荡模式的竞争. (ii) 在两个环中必须至少有一个负环. 因为只有负环才能独立地产生振荡, 而正环只可能在负环的驱动下产生影响振荡的模式. 在文献[18]中我们通

过定性分析和数值验证提出了混沌模体的另外条件. (iii) 如混沌模体中包含一个正环和一个负环, 那么正环一定不能被负环完全控制, 即正环中一定要包含独立于负环的节点, 以使被动的正环能产生与主动的负环相当强度的竞争以产生混沌. 图 10 展现了 4 节点及以下以高频次出现的若干混沌模体结构. 有趣的是所有不可约的混沌模体都具有 N 个节点和 $N+1$ 条边的最简单的结构, 其中只有一个节点受两个邻居驱动, 而其他节点都只受到一条边驱动. 发生在这个受到联合调控的节点上的动力学代表了它所处的两个环的运动模式的竞争, 这是产生混沌的关键.

现在我们对若干混沌 GRN 例证进行分析. 图 11(a)给出了一个三节点混沌 GRN(采用方程(1)的动力学), 而(b)展示了在改变参数 K 时网络动力学从周期到混沌的分叉序列, (c)为在 $K=0.102$ 时混沌吸引子形态. 对该混沌运动数据进行(3)和(4)式的 FW 计算, 其中(4)式的平均时间 T 为 100 个平均振荡周期. 计算得到的各个作用边的权重标记在(a)上, 由图可见调控线 $3 \rightarrow 1$ 具有较弱的权重值, 而节点 2 受到的两个调控作用权值相近, 反映了两个负环 $2 \cdots 1 \cdots 3 \cdots 2$ 和 $2 \cdots 1 \rightarrow 2$ 之间的高强度的竞争是产生混沌运动的主要原因. 核心子网络(d)正是图 10 中的混沌模体(11), 它可以用图 1 同样的方法从混沌数据的 FW 分析中得到. 由图 11(d)子网络产生的混沌分叉序列和混沌吸

图 10 (网络版彩图)一部分具有较高出现概率的节点数少于等于 4 的混沌模体

它们都是能够产生混沌运动的最小网络结构单元. 引自文献[18]

Figure 10 (Color online) Some significant chaotic motifs with $N \leq 4$ are shown. Chaotic motifs are the minimal and irreducible building blocks for chaotic motions in GRNs. Quoted from ref. [18].

图 11 FW 分析方法对混沌 GRN 的应用

分岔图与吸引子用来比较原网络与混沌模体的动力学性质. 引自文献[18]

Figure 11 (Color online) FW method can be applied to chaotic GRN. The bifurcation diagrams and chaotic attractors are compared between the original network and chaotic motif. Quoted from ref. [18].

引子分别在图 11(e)和(f)给出, 它们与图 11(b)和图 11(c)的行为非常相似. 在图 11(e)和(f)的计算中保留了文献[18]的操作方法, 未将剔除边的平均值保留.

图 12 对另一 4 节点的混沌 GRN 进行讨论, 其中图 12(a)–(f)进行与图 11(a)–(f)完全同样的计算和分析. 由图可见 FW 方法可成功辨识对混沌运动起关键作用的结构(图 12(d), 即图 10 的混沌模体(j)). 节点 1 所受到的权值相近的联合调控是两振荡模式竞争导致混沌运动的关键.

当网络尺寸(N , I)增大时, FW 方法对混沌 GRN 的分析效果和显著性就有所下降. 由于混沌运动对小的微扰特别敏感, 按作用边权重值对网络约化产生的误差会引起混沌行为的较大偏离. 但即使在这种情况下, FW 分析仍然能提供许多有用信息. 在大多数情况下, FW 方法仍能成功地揭示复杂的 GRN 中决定混沌运动的核心子网络, 并且这些子网络能够很好地复制原网络的整个分岔序列和吸引子特征, 但产生分岔序列和不同吸引子的参数条件则有了相

图 12 (网络版彩图)与图 11 相同, FW 方法分析了另一个 4 节点的混沌 GRN. 引自文献[18]

Figure 12 (Color online) The same as Figure 11 FW method can be also applied to another 4-node chaotic GRN. Quoted from ref. [18].

当变化(即使图 11, 12 的少节点混沌也是如此). 但也有许多例子中, 混沌斑图的核心子网络本身就很大, 约化空间远不如周期振荡 GRN 那么大, 并且产生的分岔序列和吸引子行为也与原网络差别较大. 这些都与混沌运动的敏感性和不同振荡模式的竞争相关. 即便如此, 核心子网络产生的动力学斑图仍然保持了原网络分岔和吸引子的主要特征. 如何依据混沌运动的特点加以改进 FW 方法以便对混沌网络提出更好的分析和描述手段, 仍然是一个重要的研究任务.

7 GRN 振荡行为的调控

理解斑图动力学过程的一个重要目的是为了对其动力学过程进行影响和控制. FW 方法可以告诉我们复杂网络中的哪些节点和相互作用对某种动力学行为(振荡或混沌)最为敏感, 这为我们提供了高效率调控这种运动的手段.

对振荡 GRN 网络进行有效调控, 一种最简单的

图 13 (网络版彩图)用不同的去边方法抑制振荡. 控制核心网络可以更有效地抑制振荡

Figure 13 (Color online) Suppressing oscillations of GRNs by different strategies. It is found that much more significant effect on network oscillations can be observed in manipulations of links in cores than in manipulations of the links randomly chosen in GRNs.

检测方法是如何最有效地抑制振荡, 即只改变网络中最少的边和节点来消除网络整体的振荡状态. 以图 1 为例. 直接从图 1(a)很难看出如何以最少的边和节点来抑制振荡. 但通过读取图 1(b)权重分布图(FW map)及约化的高权重子网络(图 1(c)和(d))可以清楚地看到剔除在核心子网络中的 $1 \cdots 3$, $3 \cdots 2$, $2 \cdots 1$ 三边中任何一条, 就有很大的可能消除振荡源而抑制整个网络振荡. 这一预言的确被证实(图 13(a)的阻尼振荡曲线). 而剔除随意选取的其他边则对振荡影响不大, 如图 13(a)振荡曲线所示. 读取图 2(c)和(d)的 FW 图可见, 如剔除 $17 \rightarrow 42$, $17 \rightarrow 48$ 两边或单独剔除 $16 \rightarrow 17$ 一条边, 就可以完全破坏图 2(a)振荡核心中的环结构, 则可成功抑制振荡. 而调控核心子网络外的随机选取的两条边, 如 $42 \rightarrow 49$, $17 \rightarrow 27$ 则不会重大改变振荡态. 这些预言都被图 13(b)所验证.

为了对以上调控过程进行统计研究, 我们对振荡 GRN($N=20$, $I=100$)的 5000 样本进行了两种不同的去边操作. 一种是对每个样本在核心子网络中任意剔除 J 条边(J 为样本核心子网络中环的个数), 另一种是在整个网络中任意剔除 J 条边. 其结果在图 13(c)

中给出. 对核心子网络去边后的余网络中振荡死亡率近 80%, 重现原网络动力学的正确率 S_u 低于 1.5%, 即原网络振荡受到实质性破坏. 原网络随机去边的余网络振荡死亡率仅 9.1%, 重现原网络动力学的正确率则高达 87.5%, 即网络振荡运动并未受到重大影响. 尽管两种方法去除了完全同样数目的边, 但它们对 GRN 振荡的影响不可同日而语. 这里清楚地显示了调控核心子网络对控制整个网络动力学所发挥的关键作用.

除以上抑制振荡的调控外, FW 方法还可以对其他各种专门性的调控提供信息. 例如读取 FW 图数据可以告诉如何调控振荡频率. 在从图 2(d)可预言仅剔除作用边 $17 \rightarrow 42$ 不能完全抑制振荡(因为核心子网中还有 4 节点负环存在), 但这一调控可以明显地降低频率, 这是因为核心子网络中短的 3 节点负环振荡源被破坏, 而长的 4 点负环产生的振荡会有更长的周期(图 14(a)). 另外有时实际需要抑制一部分节点的振荡而不影响其他节点的振荡. 图 2(c)中剔除相互作用 $42 \cdots 11$ 可重大影响骨架中 11 下游的所有节点振荡行为但不会对网络其他节点行为产生重大影响(见图

图 14 (网络版彩图)用不同去边方法进行不同目的的调控

核心网络中最短环被破坏后,长环建立新的振荡源产生周期较长的振荡.剔除上游作用边可以严重影响下游节点的振荡幅度,而不会对其他分支的节点的振荡幅度产生重大影响

Figure 14 (Color online) To regulate the network dynamics by different strategies. Some long loops play roles of source loops for network oscillation with long periods after some nodes on the original short core loops are discarded. Moreover, discarding an upstream node can seriously affect the motions of its down stream nodes, but not considerably change the motions of nodes in different branches.

14(b));而在图 3 中剔除(b)中左上 3 节点环的任何一边可对(a)中左上集团的整个节点运动产生严重抑制而不显著影响右下集团的振荡行为(见图 14(c)).如此等等,FW 图会给我们提供丰富的信息帮助我们为目的地对 GRN 斑图动力学进行调控,实现所需要的动力学(功能)行为.类似的调控行为也可以应用于混沌网络,包括调控极少的边或节点来抑制混沌和改变混沌的特性等等,这里不再一一详述.

8 总结

本文综述了对基因调控网络振荡问题的研究,讨论了如何从复杂网络的大量动力学数据中提取有用信息,和如何利用这些信息从复杂的网络拓扑结构中理解其中相对简单的动力学和功能结构.具体地说本文综述了一种所谓的权重(FW)分析方法,用以定量计算网络中各个作用边对所作用对象节点驱动的重要性,并通过观察高权重边构成的骨架子

网络和核心子网络来辨识网络中产生振荡的源结构及信号在动力学网络中从振荡源到整个网络的传播路径,以此理解在整个基因调控网络中振荡的产生及传播过程.这种 FW 分析方法不仅可应用于周期振荡 GRN,还可应用于分析混沌的 GRN 系统;不仅可用于理解网络中振荡形成及表现的机制,而且可用来对 GRN 振荡作有效调控.

本文综述的方法可应用于不同的基因调控模型,我们期待它在基因调控网络动力学行为研究中有实质性的突破和应用.特别是这种方法指出,在大量基因参与的生物过程(从模型中表现为动力学过程)中,真正起决定作用的基因可能只是极少数,而这些少数基因所构成的子功能网所决定的运动模式通过相互作用影响到大量基因的运动行为.从看似复杂的大量基因的集体行为中深入寻找被掩盖的关键基因和核心相互作用,并找到生物学过程的机制及调控要点,可能是研究复杂基因调控过程的一个重要任务.另外本文综述的方法集中讨论了基因调控网络

动力学问题, 但作为一个动力学分析的数学物理方法, 其应用范围并不仅限于生物系统, 但必须考虑方法所要求的动力学条件, 其具体应用范围目前还不是很清楚. 我们估计 FW 方法很可能普遍应用于一大类复杂网络系统, 其局域动力学具有被动特点(不受外界驱动时, 处于稳定定态), 而相互调控作用具有单调性(正调控: 被调控变量的增加率随调控变量的增加而增加; 负调控: 被调控变量的增加率随调控变量的增加而减少). 在实际中具有这些性质的复杂网络非常普遍, 如一些蛋白质网络、生态捕食网络以及许多社会经济网络等等. 我们希望 FW 方法能在更广阔的领域中有更多的应用.

然而本方法也存在很大局限性, 期待能够在进一步发展中解决. 最大限制是 FW 分析方法要在大量已

知信息的基础上进行操作. 为了计算权重值, 我们必须知道全部的网络动力学运动数据; 知道网络节点方程及其中的参数; 又要知道相互作用的所有拓扑作用边, 而在实际生物和其他系统中这些信息是很难全部掌握的. 在实际问题中, 以及在数据库中往往知道其中很多但是并不完全的信息, 这样在部分信息缺失的条件下进行近似但又有效的计算, 就成为很重要的要求. 另外许多数据和参数掺杂了噪声, 所以在有误差的信息下进行可靠地计算也是一个重要要求. 在以往的研究中已经有一系列在缺失部分信息下进行计算的方法^[26–28], 同时从动力学数据反推网络拓扑连接的研究在近年也有很大发展^[29,30]. 将这些已有工具与 FW 分析方法相结合发展强有力的分析复杂网络动力学机制的工具, 是一个非常具有前景的课题.

参考文献

- 1 Watts D J, Strogatz S H. Collective dynamics of “small-world” networks. *Nature*, 1998, 393: 440–442
- 2 Barabasi A L, Albert R. Emergence of scaling in random networks. *Science*, 1999, 286: 509–512
- 3 Albert R. Scale-free networks in cell biology. *J Cell Sci*, 2005, 118: 4947–4957
- 4 Barabasi A L. Scale-free networks: A decade and beyond. *Science*, 2009, 325: 412–413
- 5 Tsai T Y C, Choi Y S, Ma W, et al. Robust, tunable biological oscillations from interlinked positive and negative feedback loops. *Science*, 2008, 321: 126–129
- 6 Novak B, Tyson J J. Design principles of biochemical oscillators. *Nat Rev Mol Cell Biol*, 2008, 9: 981–991
- 7 Tyson J J, Novk B. Functional motifs in biochemical reaction networks. *Annu Rev Phys Chem*, 2010, 61: 219–240
- 8 Goldbeter A. Computational approaches to cellular rhythms. *Nature*, 2002, 420: 238–245
- 9 Ferrell Jr J, Tsai T C, Yang Q. Modeling the cell cycle: Why do certain circuits oscillate? *Cell*, 2011, 144: 874–885
- 10 Glass L, Mackey M C. *From Clocks to Chaos: The Rhythms of Life*. Princeton: Princeton University Press, 1988
- 11 Goldbeter A. *Biochemical Oscillations and Cellular Rhythms*. Cambridge: Cambridge University Press, 1996
- 12 Alon U. *Introduction to Systems Biology: Design Principles of Biological Networks*. Boca Raton: Chemical Rubber Company Press, 2007
- 13 Albert R, Jeong H, Barabasi A L. Error and attack tolerance of complex networks. *Nature*, 2000, 406: 378–382
- 14 Callaway D S, Newman M E J, Strogatz S H, et al. Network robustness and fragility: Percolation on random graphs. *Phys Rev Lett*, 2000, 85: 5468–5471
- 15 Kitano H. Biological robustness. *Nat Rev Genet*, 2004, 5: 826–837
- 16 Qian Y, Huang X, Hu G, et al. Structure and control of self-sustained target waves in excitable small-world networks. *Phys Rev E*, 2010, 81: 036101
- 17 Liao X, Xia Q, Qian Y, et al. Pattern formation in oscillatory complex networks consisting of excitable nodes. *Phys Rev E*, 2011, 83: 056204
- 18 Zhang Z, Ye W, Qian Y, et al. Chaotic motifs in gene regulatory networks. *PLoS ONE*, 2012, 7: e39355
- 19 Albert R, Othmer H G. The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in *drosophila melanogaster*. *J Theor Biology*, 2003, 223: 1–18
- 20 Ishihara S, Fujimoto K, Shibata T. Cross talking of network motifs in gene regulation that generates temporal pulses and spatial stripes. *Genes Cells*, 2005, 10: 1025–1038
- 21 Tian X J, Zhang X P, Liu F, et al. Interlinking positive and negative feedback loops creates a tunable motif in gene regulatory networks. *Phys Rev E*, 2009, 80: 011926
- 22 Kobayashi Y, Shibata T, Kuramoto Y, et al. Robust network clocks: Design of genetic oscillators as a complex combinatorial optimization problem. *Phys Rev E*, 2011, 83: 060901
- 23 Zhang Z, Li Z, Hu G, et al. Exploring cores and skeletons in oscillatory gene regulatory networks by a functional weight approach. *Eur Phys*

- Lett, 2014, 105: 18003
- 24 Goh K I, Kahng B, Cho K H. Sustained oscillations in extended genetic oscillatory systems. *Biophys J*, 2008, 94: 4270–4276
 - 25 Ye W, Zhang Z, Lv B, et al. Oscillatory and anti-oscillatory motifs in genetic regulatory networks. *Chin Phys B*, 2012, 21(6): 60203
 - 26 Klemm K, Bornholdt S. Topology of biological networks and reliability of information processing. *Proc Nat Acad Sci USA*, 2005, 102: 18414–18419
 - 27 Pigolotti S, Krishna S, Jensen M H. Symbolic dynamics of biological feedback networks. *Phys Rev Lett*, 2009, 102: 088701
 - 28 Pigolotti S, Krishna S, Jensen M H. Oscillation patterns in negative feedback loops. *Proc Nat Acad Sci USA*, 2007, 104: 6533–6537
 - 29 Wang W X, Chen Q, Huang L, et al. Scaling of noisy fluctuations in complex networks and applications to network prediction. *Phys Rev E*, 2009, 80: 016116
 - 30 Ren J, Wang W X, Li B, et al. Noise bridges dynamical correlation and topology in coupled oscillator networks. *Phys Rev Lett*, 2010, 104: 058701

Exploring cores and skeletons in oscillatory gene regulatory networks by a functional weight approach

ZHANG ZhaoYang, HUANG XuHui, ZHENG ZhiGang & HU Gang^{*}

Department of Physics, Beijing Normal University, Beijing 100875, China

The topological structures of complex networks are often very complicated and there are huge amount of data in their dynamics. It is important to extract useful information from the available data and to explore some simple reduced structures (if they exist) which play key roles determining different functions of the networks. In this paper, models of gene regulatory networks (GRNs) are studied and a method of functional weight analysis is proposed to quantitatively measure the importance of all topological interactions from the data of oscillation. This method allows us to obtain some simple skeleton interactions and core structures. The latter play the role of oscillation sources while the former reveal the main signal propagation paths from the sources throughout the networks. Different control strategies and detailed statistical results demonstrate the validity, significance and efficiency of the analysis.

dynamics, functional weight, gene regulatory network, oscillation, chaos

PACS: 89.75.Fb, 05.10.-a, 05.65.+b

doi: 10.1360/SSPMA2014-00123