

酵母蛋白质网络的动力学性质*

李方廷^{1,2} 吕莹^{1,2} 龙涛^{1,2} 欧阳颀^{1,2,†} 汤超^{1,3}

(1 北京大学理论生物中心 北京 100871)

(2 北京大学物理学院 北京 100871)

(3 NEC Laboratories America, 4 Independence Way, Princeton, NJ 08540, USA)

摘要 蛋白质-蛋白质、蛋白质-DNA 相互作用网络决定了细胞中各种关键功能的执行. 基于芽殖酵母(*budding yeast saccharomyces cerevisiae*)的蛋白质-蛋白质相互作用网络数据和相关的实验文献,我们建立了调控细胞周期和生命周期(*cell cycle and life cycle*)的蛋白质网络,并利用离散模型研究了该网络的动力学性质. 研究表明:细胞周期网络的动力学性质具有很强的稳定性,约 94% 的蛋白质初态将演化到对应于生物学 G_1 基态的稳定态,使其成为唯一的全局吸引点;同时,绝大多数的初态的演化路径都通过由 G_1 激发态到 G_1 基态的细胞周期演化路径,使细胞周期路径成为全局性的“吸引”路径.

关键词 蛋白质网络动力学 细胞周期过程 芽殖酵母

Dynamical properties of protein networks in yeast

LI Fang-Ting^{1,2} LU Ying^{1,2} LONG Tao^{1,2} OUYANG Qi^{1,2,†} TANG Chao^{1,3}

(1 Centre for Theoretical Biology, Peking University, Beijing 100871, China)

(2 School of Physics, Peking University, Beijing 100871, China)

(3 NEC Laboratories America, 4 Independence Way, Princeton, NJ 08540, USA)

Abstract Protein-protein and protein-DNA interactions in living cells constitute molecular dynamic networks that govern various cellular functions. Based on recent studies of budding yeast and other related literature, we have established the protein network governing the cell cycle and life cycle processes. Here we report our dynamical analysis of the cell-cycle network of budding yeast using a discrete two-state model. Our study reveals important properties of the network: (1) almost all (~94%) initial protein states evolve to the stationary state G_1 , making it a global attractor; (2) the dynamical paths to G_1 for these initial states are through the cell-cycle sequence, making it a globally attracting trajectory.

Key words dynamical properties of protein networks, cell cycle process, budding yeast

芽殖酵母(*budding yeast saccharomyces cerevisiae*)是生物学研究中广泛应用的单细胞真核模式生物. 酵母在细胞周期调控的研究中有着极其重要的作用. 1996 年,作为第一个真核生物,芽殖酵母的全基因组测序工作完成并公布^[1]. 近年来,芽殖酵母的蛋白质相互作用的数据迅速增加积累^[2-4],这些蛋白质-蛋白质相互作用网络的数据和相关的生物学研究进展,为进一步全面系统地研究蛋白质网络的性质提供了可能. Barabási 小组^[5]及 Maslov 等^[6]对芽殖酵母蛋白质网络的拓扑性质进行了研究, Ih-

mels 等对基因调控网络中的基本调控单元进行了研究^[7]. 相对于较为稳定的基因组,由蛋白质网络组成的蛋白质组,对不同的环境信号,通过蛋白质的状态不断变化产生反应,即通过动力学过程来完成生物学功能的. 所以,蛋白质网络动力学的研究成为生物学家和生物物理学家共同关心的重要问题^[8]. 虽然,蛋白质网络动力学的研究^[9]有了一定程度的进展,但其全局性质并不清楚.

* 2003-07-29 收到初稿, 2003-08-15 修回

† 通讯联系人. E-mail: qiqi@pku.edu.cn

基于最近的芽殖酵母的蛋白质 - 蛋白质相互作用网络数据^[2-4]和相关的实验文献^[10],我们建立了调控芽殖酵母细胞周期和生命周期的蛋白质网络,并利用离散模型对该网络的动力学性质进行了研究.研究表明:基于网络相互作用规则,细胞周期网络的动力学性质具有很强的稳定性,约 94% 的蛋白质初态将演化到对应于生物学 G_1 基态的稳定态,使其成为惟一的全局吸引点,同时,绝大多数的初态的演化路径都通过由 G_1 激发态到 G_1 基态的细胞周期演化路径,使细胞周期路径成为全局性的“吸引”路径.

芽殖酵母具有简单的生命周期,能够以单倍体和双倍体形式存在.在营养丰富的条件下,单倍体和双倍体的芽殖酵母细胞都能够以正常的细胞分裂周期进行繁殖. CDK 蛋白激酶(cyclin - dependent kinase)的基因表达和活性最终调控了整个细胞周期过程^[10].在营养缺乏条件下,触发孢子形成信号(sporulation signal)双倍体细胞能够通过减数分裂产生孢子,形成单倍体细胞来适应恶劣的外界条件,减数分裂主要是由 Ime1 蛋白的表达和活化来调控的^[11].受到信息素(pheromone)刺激,两个单倍体细胞将融和成为一个新的双倍体细胞^[12].

首先,我们研究了生物学中最为清楚的细胞周期(cell - cycle)调控网络.该网络是我们基于以前的动力学模型^[13]并通过大量的文献调研(省略)和对蛋白质数据库(<http://mips.gsf.de/>)的分析建立起来的.简化的细胞周期网络如图 1 中插图所示.该网络中的蛋白质可分为以下三大类:第一类为 cyclin/Cdc28 复合物,包括 Cln3/Cdc28 复合物(图中简称为 Cln3),Cln1/Cdc28 与 Cln2/Cdc28 复合物(简称为 Cln1,2),Clb5/Cdc28 与 Clb6/Cdc28 复合物(简称为 Clb5,6),Clb1/Cdc28 与 Clb2/Cdc28 复合物(简称为 Clb1,2);第二类为转录因子,包括 MBF, SBF, Mcm1/SFF 和 Swi5;第三类为 cyclin/Cdc28 复合物的抑制蛋白与降解蛋白,包括 Sic1, Hct1, Cdc20/APC.网络中的正相互作用表示蛋白质间的激发或活化作用,负相互作用表示蛋白质间的抑制或去活化作用,自指的箭头表示蛋白质的自降解作用.

细胞周期过程可以简述如下:在营养丰富的条件下,当双倍体或单倍体的酵母细胞长得足够大时, Cln3/Cdc28 蛋白复合物将被活化,促使细胞进入“激发的” G_1 态,这时细胞的 Sic1 浓度较高, Hct1 处于活化状态.活化的 Cln3/Cdc28 复合物将活化转录因子 MBF 和 SBF,活化的 MBF 和 SBF 与 DNA 结合

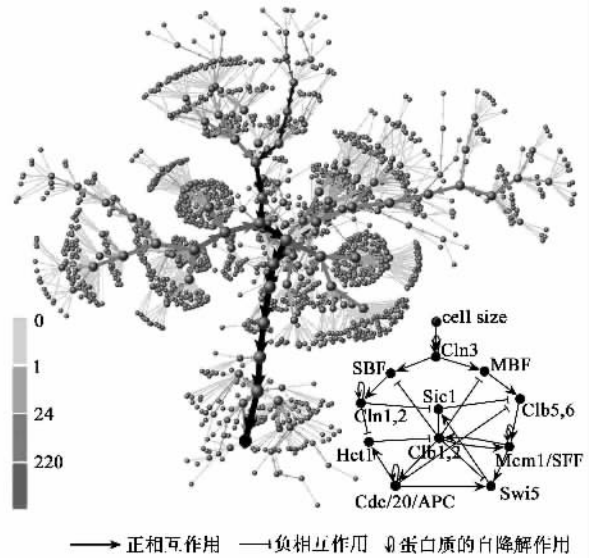


图 1 简化的芽殖酵母细胞周期蛋白质调控网络(插图)与其所有可能初始态的演化路径

[蛋白质网络中的细胞大小(cell size)节点表示外界信号,其他节点表示某类蛋白质或蛋白质复合物,正相互作用表示蛋白质间的激发或活化作用,负相互作用表示蛋白质间的抑制或去活化作用,自指的箭头表示蛋白质的自降解作用.演化路径中的每个节点表示一个蛋白质状态,箭头表示由一个蛋白质状态向另一个蛋白质状态演化.图中每个点的大小和每条线的宽度正比于 $\ln(2+m)$, m 为蛋白质初态经过该点(边)的数目]

后,转录相应的 mRNA,然后翻译形成 Cln1, Cln2, Clb5, Clb6 蛋白,上述蛋白抑制了 Sic1 和 Hct1 的作用,并控制着 G_1 后期基因的表达.在 S 期(S phase)细胞复制自己的 DNA.通过 G_2 期, Clb1 和 Clb2 活化,细胞进入有丝分裂期(M phase).有丝分裂使得复制的 DNA 等量地分配到细胞相对的两极,然后一个细胞分裂产生两个子细胞,在该过程中 Cdc20/APC 和 Swi5 被活化,导致 Sic1 浓度升高, Hct1 活化,并对 cyclin/Cdc28 复合物产生抑制作用.最后,细胞又回到细胞周期的静息(rested) G_1 态, G_1 基态,等待下一次分裂信号.总的来说,细胞周期过程起始于“激发” G_1 态——该状态 Cln3/Cdc28 复合物处于活化状态,通过一系列大家熟知的细胞周期过程,最后回到 Cln3/Cdc28 复合物未活化的 G_1 基态.

为了研究蛋白质网络的全局动力学性质,我们选择以下简单的离散动力学模型:每类蛋白质只有两种状态 0 与 1,分别表示该蛋白质处于活化与未活化状态.下一个时刻蛋白质的状态是由当前时刻的蛋白质状态按照以下规则决定的:

$$S_i(t+1) = \begin{cases} 1, & \sum_j a_{ij} S_j(t) > 0 \\ 0, & \sum_j a_{ij} S_j(t) < 0 \\ S_i(t), & \sum_j a_{ij} S_j(t) = 0 \end{cases}, \quad (1)$$

其中 a_{ij} 是第 j 类蛋白质对第 i 类蛋白质的作用系数。模型中的时间步长是逻辑步长,而非实际意义上的时间。我们选择 a_{ij} 取值为 1 与 $-\infty$ (负无穷),分别表示正相互作用和负相互作用。自降解作用具有时间延迟的性质:一个具有自降解作用的蛋白质,若在 t 时刻被活化 [$S_i(t) = 1$],而且在 $t+1$ 到 $t=t+t_d$ 时间内一直没有其他的正负输入,那么它将在 $t=t+t_d$ 时刻降解 [$S_i(t+t_d) = 0$]。在模型中选择 $t_d = 9$ 。采用简单的离散动力学模型的优势,在于能够分析网络动力学状态全空间的性质,从而得到网络的全局动力学性质。

利用以上的离散动力学模型,我们研究了细胞周期调控网络(11 结点网络,不包括 cell size 信号)。首先,把激发 G_1 态作为初始态(Cln3, Sic1 和 Het1 的状态为 1,其余蛋白质的状态为 0),我们观察到系统逐步演化到 G_1 基态(Sic1 和 Het1 的状态为 1,其余蛋白质的状态为 0),蛋白质状态的时间演化过程(略去)与生物学实验观察相符合^[14,15]。我们进一步研究了网络的动力学吸引点,遍历所有可能的 2048 个初始态(11 结点网络),该蛋白质网络最后演化到 12 个稳定的状态,其中 1925 个初始态(约 94%)演化到静息的 G_1 态——细胞周期的生物学稳定态——而且该状态是惟一的全局吸引点。为了验证是否所有网络都具有类似的性质,我们在相同演化规则下,研究了具有相同结点数目和相同连接数目的网络,随机产生网络中结点间的相互作用。随机网络的动力学性质的研究(略去)表明,随机网络平均来说具有更多的吸引点,而且出现类似细胞周期网络全局吸引点的概率极小。细胞周期网络的特殊结构使得其能够具有较好的动力学稳定性,具有吸引约 94% 初始态的全局吸引点—— G_1 基态。

那么,11 结点的细胞周期网络中的所有初始状态是怎样一步一步地演化到最后的吸引点呢?图 1 中央自上向下的黑色路径为生物学路径——细胞周期路径,其末端最大的黑色节点为 G_1 基态,演化路径图中每个点的大小和每条线的宽度正比于 $\ln(2+m)$, m 为蛋白质初态经过该点(边)的数目。我们发现细胞周期路径是最可几路径:大部分的网络初始状态首先被吸引到细胞周期路径上来,然后沿着

细胞周期路径逐步到达稳定态—— G_1 基态。这意味着不仅 G_1 基态是一个全局性的稳定点,而且从 G_1 激发态到 G_1 基态的细胞周期路径同样是一个全局性的稳定的动力学路径,即蛋白质网络通过长期的进化,其动力学性质具有双重稳定性。我们发现在裂殖酵母(fission yeast)和蛙卵细胞(frog egg)的细胞周期网络中,也有类似的动力学性质。

最后,我们研究了调控芽殖酵母中调控细胞周期和生命周期(life-cycle)过程的蛋白质网络的动力学性质,简化的网络如图 2 中插图所示。为了保证相关蛋白处于正常的生物学状态,我们在网络中引进了 Housekeeping 节点,该节点一直属于激活的 1 状态。当 Mat a1/alpha2 蛋白质(双倍体)与孢子形成信号(sporulation signal)必须同时处于 1 时,蛋白 Ime1 才能活化,其他动力学演化规则与细胞周期网络相同。遍历所有的初始状态,可以得到演化路径,其中单倍体的 G_1 基态与双倍体的 G_1 基态形成 2 个全局吸引点(演化路径中最大的两个黑色节点),分别吸引了 64.2% 和 29.0% 的全部初始态。图 2 中画出了其中随机选取的 9000 个初态的演化路径,相应的生物学路径同时是两条吸引路径。

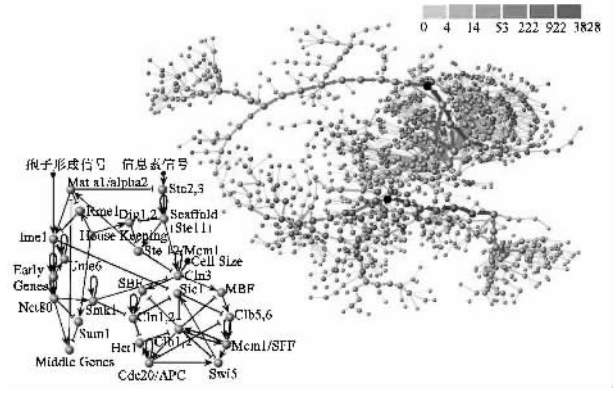


图 2 简化的芽殖酵母细胞周期和生命周期过程蛋白质调控网络(插图)及其中 9000 个随机初始态的演化路径 [演化路径中每个点的大小和每条线的宽度正比于 $\ln(2+m)$, m 为蛋白质初态经过该点(边)的数目]

总之,我们的结果表明,蛋白质网络不仅拓扑性质与随机网络不同,其动力学性质也与随机网络完全不同,而且该性质与蛋白质网络的生物学功能紧密相关。简单的离散的动力学模型能够作为研究生物网络的微分方程模型的补充,并且对于大规模的生物网络动力学研究可能较为有效。

参考文献

[1] Goffeau A et al. Science, 1996, 274 :546
 [2] Mering C et al. Nature, 2002, 417 :399

[3] Uetz P *et al.* Nature 2000 403 :623
 [4] Schwikowski B *et al.* Nat. Biotechnol. 2000 18 :1257
 [5] Albert R , Jeong H , Barabási A - L. Nature ,2000 ,406 :378
 [6] Maslov S , Sneppen K. Science ,2002 ,296 :910
 [7] Ihmels J *et al.* Nature Genetics. 2002 ,31 :370
 [8] Alliance for cellular signalling. Nature ,2002 ,420 :703
 [9] Tyson J J , Chen K , Novak B. Nature Review Molecular Biology ,2001 ,2 :908

[10] Mendenhall M D , Hodge A E. Microbiol Mol Biol Rev. , 1998 ,62 :1191
 [11] Vershon A K , Pierce M. Curr. Opin. Cell Bio. ,2000 ,12 34
 [12] Gustin M C *et al.* Microbiol Mol Biol Rev. ,1998 ,62 :1264
 [13] Chen K *et al.* Mol. Biol. Cell ,2000 ,11 369
 [14] Spellman P T *et al.* Mol. Biol. Cell ,1998 ,9 :3273
 [15] Cross F R *et al.* Mol. Biol. Cell ,2002 ,13 52

· 物理新闻与动态 ·

超强的光丝

激光束在开阔的空气中传播时显示出透明的特征,它们可以继续不断地在天空中穿行;但若碰到小水滴层时,它们将会大量地吸收激光。近来在法国里昂的 Claude Bernard 大学的一群以 Courvoisier 教授为首的物理学家们利用超强(约为 10^{14} W/cm^2)超短时(约为 120fs)的激光脉冲来创造出一种超强光丝,其宽度约为 $150 \mu\text{m}$,长度为 100m 左右。这种光丝的特点是能在实验室的模拟水滴层中传播而能量损失极小。模拟水滴层的条件非常接近于真实的大气状态。因此这项研究工作对于激光的应用具有很重要的意义。特别是在自由空间内进行的激光通讯、对大气污染的监测以及某些测距方面的应用等。

对于这种光丝的形成机理,现在初步认为是由两种非线性光学效应所促成的,一是光学中的“克尔效应”,它能使高强度的光在传播介质中(例如这里的空气和水滴等)改变折射率,从而导致了光的自聚焦现象,另一个是散焦等离子体效应。由于在实验室条件下研究的成功,法国科学家们准备将这项研究推广到在一定控制条件下的实际大气层中。

(云中客 摘自 Applied Physics Letters , 14 July 2003)

蛋白质的飞秒 X 射线结晶学

利用 X 射线结晶学可以对成千个蛋白质分子的冻结过程拍摄下快照。但却无法记录下单个蛋白质分子的全部运动过程。这是由于过去使用的 X 射线拍摄技术在时间上是纳秒量级的,这相对于反映蛋白质分子的变化过程来说是显得太慢了。现在位于法国的欧洲同步加速器与辐射研究所(ESRF)的一群以 P. Anfinrud 博士为首的科学家们展现了他们的飞秒 X 射线结晶摄影技术。他们对肌浆珠蛋白(myoglobin)分子的突变体是如何排出 CO 有毒分子的过程进行了飞秒量级的记录。肌浆珠蛋白是一种能在肌肉组织内贮存氧气的蛋白质分子。研究者们选择肌浆珠蛋白的突变体作为对象是因为当蛋白质分子内的部分原子结构发生变形时,它会比正常的肌浆珠蛋白分子更快地排除 CO 分子。

为了能记录下这个过程,科学家们首先对蛋白质分子施加一飞秒时间的激光脉冲以便于排除 CO 分子,接着迅速地从 ESRF 同步加速器中发出的 150fs 的 X 射线脉冲照射蛋白质,其中最关键的一点是要在同步加速器中离析出单一 X

射线的能力。与此同时再用一台 CCD 摄像机拍摄下 X 射线脉冲连续通过蛋白质分子时,分子发生的各种形态变化。最终摄制成的影片显示出了 CO 分子在蛋白质分子内不同位置处的移动与转移过程,也清晰地显示了肌浆珠蛋白分子是如何改变自身的形态以适应于 CO 分子逐步排除的过程。这个飞秒 X 射线技术不仅可以研究蛋白质中的许多分子的迁移过程,而且还可以拍摄下这个时间尺度下的动态影片来与动力学模拟过程作对比,从而可促进实验结果与理论模型之间的比较与配合。他们的工作在 7 月 26—31 日在美国 Cincinnati 召开的美国结晶学年会上进行演示。

(云中客 摘自 Science , 20 June 2003)

研究活体肿瘤的新技术

美国普渡大学与位于英国伦敦的帝国研究院的 Ping Yu 和 D. Nolte 博士利用红外线及特殊的半导体全息照相术拍摄下了老鼠的肿瘤组织。这种肿瘤组织过去都是通过切片后,利用离子辐射才能看到它的结构。这项新的技术是利用光相干成像法(optical coherence imaging ,OCI)来拍摄下贯穿整个肿瘤组织的图象,这项技术是与过去的光相干层析 X 射线摄影(optical coherence tomography ,OCT)技术有关。OCT 的操作过程包括如下步骤,先是通过激光束对样品进行扫描,随后再逐点地收集信息,最后将它们装配成完整的图像。但 OCI 与此不同,它是收集活体组织整个薄片的完全图像,然后直接记录到摄像机上。

OCI 的关键技术是动力学全息摄影,它是将不相干的散射光背景进行过滤,只保留了相干部分的全息图像。活体组织可对红外光线快速地反射成像,当然它也能对普通光线发生强烈地散射,如果不过滤去散射光,那么它将会淹没相干光形成的图像。适当地调节在 OCI 系统图像干涉仪上成像光束与参照光束间的位相差,科学家们就能有控制地拍摄到活体样品不同深度的图像,从而能对肿瘤活体组织进行一片一片地拍摄,这样拍摄下的图像就是反映活体肿瘤内部结构的一部电影。

将 OCI 技术应用在经过培育的老鼠肿瘤上,拍摄下的图像揭示出了组织结构内新的死亡部位的范围和钙化区的特征,这些特征非常类似于人体内的肿瘤发展。

总之,科学家们相信,全息 OCI 技术能够代替 X 射线和微切片方法来研究活体组织新的破坏与死亡状况。

(云中客 摘自 Applied Physics Letters , 21 July 2003)