

基于 Petri 网的基因调控网络建模与分析 *

黄佳良¹⁾ 郭红¹⁾ 孟春^{2)**}

(¹⁾福州大学数学与计算机科学学院, 福州 350002; ²⁾福州大学生物科学与工程学院, 福州 350002)

摘要 如何有效描述与分析复杂的基因调控网络(GRN)是生物学家研究基因表达调控机制的关键步骤. 现有大部分方法在建模过程中忽略了生物中广泛存在的协同作用, 模型预测结果与实际生物行为之间存在误差. 基于混合函数 Petri 网(HFPN)理论提出了一种对基因调控网络进行定量分析的新方法. 首先简要介绍 GRN 与 HFPN 的基础理论, 然后为 HFPN 引入两类新元素: 逻辑库所与逻辑变迁, 描述基因调控网络的逻辑规则以及转录因子间的协同作用, 最后构建海胆 *endo16* 基因调控网络的 Petri 网模型, 并预测模型在不同位点发生突变时的基因表达水平变化. 分析结果与文献实验数据相一致, 验证了方法的正确性.

关键词 混合函数 Petri 网, 基因调控网络, 海胆 *endo16* 基因, 协同作用, 逻辑规则

学科分类号 Q811.4, TP391.9

人类基因组计划的基本完成, 标志着生命科学研究进入后基因组时代. 研究工作逐渐从基因测序转向基因功能^[1], 开始以系统角度研究基因表达调控机制. 如何有效描述与分析基因调控网络(GRN)成为其中的关键步骤. 传统上基于反应动力学理论采用一组微分方程可以有效地描述 GRN 的动态行为. 但由于参数的相关性与生物系统自身的鲁棒性, 准确估计复杂生物系统的所有参数依然相当困难, 同时该方法缺乏直观性. Vohradsky 等^[2]采用人工神经网络描述 GRN, 以基因作为网络结点, 弧权值描述基因间的信号放大级数. Cherry^[3]考虑 GRN 调控机制与逻辑电路之间的相似性, 采用 Boolean 网络描述 GRN, 在很大程度上简化了调控网络, 但与连续方法相比缺乏对于动态行为描述力度.

由于 Petri 网不仅具有直观易懂的图形表示和严格数学定义, 而且已有大量相关算法和分析工具, 近年来逐渐成为描述和分析各种复杂生化网络的有效工具^[4~6]. Steggles 等^[7]以 Petri 网的两个互补库所表示一个基因结点, 将基因调控网络的逻辑规则转化为 Petri 网的变迁规则, 但网络随之变得更加复杂. Chen^[8]基于混合 Petri 网构建一些 GRN 模型. Doi 等^[9]在混合 Petri 网与混合动态 Petri 网的基础上针对生化网络将 Petri 网元素进一步细分, 并增加一些新元素, 扩展得到混合函数 Petri 网

(Hybrid Functional Petri Net, HFPN). 利用 HFPN 中的离散和连续元素可以恰当地描述生物体内各种不同生理机制的复杂生物系统. 例如, 利用离散元素可以描述基因表达的开关系统^[1], 利用连续元素可以描述转录、转化、酶催化、代谢反应等连续改变的事件^[9]. 此时, 类似米氏方程的生物反应公式作为速率函数几乎可以直接加入模型中^[8].

然而, 这些方法忽略了生物系统中广泛存在的协同效应, 例如多个转录因子的协同作用往往并不等于这些转录因子单独作用的简单累加^[10]. 模型预测结果与实际生物行为之间存在误差. 鉴于此, 本文通过对混合函数 Petri 网进行扩展, 增加了逻辑库所与逻辑变迁, 描述基因调控网络的逻辑规则以及转录因子之间的协同作用. 基于 Davison 等建立的海胆 *endo16* 基因调控网络构建 Petri 网模型, 并验证了该方法的正确性.

1 理论与方法

1.1 协同作用及其动力学理论

真核基因的表达调控是一个很复杂的过程, 涉

* 福建省自然科学基金资助项目(C0510006).

** 通讯联系人.

Tel: 13509369204, E-mail: mengchun@fzu.edu.cn

收稿日期: 2007-08-03, 接受日期: 2007-11-07

及很多转录因子(transcription factors, TFs)与 DNA 上的调控元件(位点)之间的相互作用^[1]. 表达调控过程具有协同性的特征, 表现为多个转录调控位点的共同作用效果大于每个位点单独作用的简单累加^[10]. 转录因子通过三种方式对协同性转录激活行为产生影响, 即多个转录因子之间的相互作用、转录因子与各 DNA 位点的协同结合以及转录因子与转录机器的协同作用^[10]. 本文暂不考虑转录因子作为蛋白质存在的相互作用, 并假定不同转录因子与位点的结合过程相互独立, 主要考虑第三种方式, 即多个转录因子对转录过程的协同作用.

基因调控网络中, 通常认为转录因子数量远超过 DNA 位点, 自由转录因子的浓度并不因其与位点结合而显著减少^[12]. 假设转录因子 T_A 结合到基因 G 上的结构域 DNA 位点, 对于可能出现的两种结合状态(DNA、DNA· T_A)基因表达以不同的速率启动. 利用复合物 DNA· T_A 的离解平衡方程, 以转录因子 T_A 浓度 $[T_A]$ 与离解常数 K_A 描述 T_A 与 DNA 位点的结合状态.

$$\text{DNA} \cdot T_A \leftrightarrow \text{DNA} + T_A \quad (1)$$

其中, $K_A = [\text{DNA}][T_A]/[\text{DNA} \cdot T_A]$ 表示离解平衡常数, $\alpha = [\text{DNA} \cdot T_A]/([\text{DNA} \cdot T_A] + [\text{DNA}])$ 表示基因 G 上 T_A 位点的结合饱和度, 满足 $0 \leq \alpha \leq 1$. 转录因子 T_A 对于基因 G 表达水平的影响可描述为^[12]

$$\varphi = w_0 + w_A \alpha \quad (2)$$

其中, w_0 、 w_A 表示不同结合状态的复合物(DNA、DNA· T_A)对于基因表达的影响权值($0 \leq w_i \leq 1$).

对于两个转录因子 T_A 与 T_B 结合到基因 G 上的结构域 DNA 位点, 方程(2)可改写为^[12]

$$\varphi = w_0 + w_A \alpha + w_B \beta + w_{AB} \alpha \beta \quad (3)$$

考虑以下三种情况: 若转录过程只能由复合物 DNA· T_A · T_B 启动(即转录因子 T_A 与 T_B 均为启动转录过程的必要条件), $w_0 = w_A = w_B = 0$, $w_{AB} = 1$, 则转录因子的作用可描述为 T_A 与 T_B 的逻辑与($T_A \wedge T_B$); 若复合物 DNA· T_A 、DNA· T_B 与 DNA· T_A · T_B 在相同水平启动基因表达过程, $w_0 = 0$, $w_A = w_B = w_{AB} = 1$, 则转录因子的作用可描述为 T_A 与 T_B 的逻辑或($T_A \vee T_B$); 如果 DNA· T_B 促进转录, 但 T_A 结合后转录过程被完全抑制, $w_0 = w_A = w_{AB} = 0$, $w_B = 1$, 转录因子的作用可等价于逻辑表达式 $\neg T_A \wedge T_B$. 三种情况的代数表达式描述如下:

$$\varphi_{(\alpha \& \beta)} = \alpha \beta \quad (4)$$

$$\varphi_{(\alpha \vee \beta)} = \alpha + \beta - \alpha \beta \quad (5)$$

$$\varphi_{(\neg \alpha \& \beta)} = \beta - \alpha \beta \quad (6)$$

这里, α 、 β 表示位点的结合饱和度 ($0 \leq \alpha$, $\beta \leq 1$), 因此这些操作不仅仅适用于 Boolean 类型数据(0, 1), 同样也适用于连续值变量. 实际上, 使用集合操作符 $\cap$ 、 $\cup$ 更适合描述这些连续值变量的逻辑行为. 上述方程可以扩展描述 n 个转录因子的协同作用.

1.2 扩展混合函数 Petri 网

本文针对基因调控网络增加两类新元素(逻辑库所与逻辑变迁), 将混合函数 Petri 网(HFPN)进一步扩展为 HFPNe, 如图 1 所示. 逻辑库所是对连续库所进行扩展: 为每个库所增加一个时间函数 $s(t)$ 描述该库所的输入值(input)随时间 t 发生变化. 逻辑变迁是对离散变迁进行扩展: 为每个变迁增加一个布尔函数 f , 其自变量为前集库所的状态 M , 用于选择关联库所的更新模式. 变迁发生时, 首先由前集库所状态计算布尔函数 f , 根据计算结果选择不同的更新方式. 对于修饰弧(激活弧、抑制弧)关联的库所状态保持不变, 对于输入输出弧关联的库所需要更新其托肯值及时间函数 $s(t)$. 逻辑变迁的授权发生条件为前集库所均已更新. 变迁发生满足同步机制^[13], 即某一个时刻 t , 不同变迁引用相同时间函数 $s(t)$ 时, 变量 t 的取值必须一致. 只有当所有授权逻辑变迁均发生后, 全局时钟 t 才更新为 $t + \Delta t$.

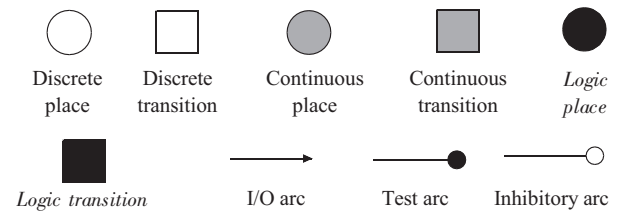


Fig. 1 The basic elements of the extend of HFPN (new in *italic*)

1.3 HFPNe 与 GRN

基因调控网络中, 转录因子直接或间接与基因调控区域(*cis-regulatory regions*)上的位点结合. 结合的转录因子激励或抑制基因启动子(gene promoter)的装配过程, 从而影响基因表达水平. 转录因子作为输入信号(input)结合到位点上, 基因调控模块像一个条件逻辑装置, 输出的结果(output)随着输入信号不同而不同. 通过这种方式, 每个基因都可以对整个生命循环中遇到的所有周围环境发生应答. 根据转录因子活性在生命周期中的变化, 输入信号分为两类^[14,15]: 时间依赖输入(time-varying inputs)与非时间依赖输入(time invariant inputs).

基于 HFPNe 构建基因调控网络时, 逻辑库所表示时间依赖输入的转录因子, 其时间函数 $s(t)$ 描述转录因子对于基因表达水平的影响(即转录因子的活性). 连续库所表示非时间依赖输入的转录因子. 这两类库所的托肯值描述 DNA 相关位点的结合饱和度 α ($0 < \alpha < 1$). 采用逻辑变迁描述基因调控网络中的一次逻辑运算, 即根据当前转录因子的结合状态选择基因表达的调控行为(如增强或减弱). 连续库所与逻辑变迁之间以修饰弧(激活弧、抑制弧)连接, 逻辑库所与逻辑变迁之间采用输入/输出弧连接. 此外, 以连续库所存放逻辑运算中间结果, 并采用输入输出弧与逻辑变迁连接. 逻辑变迁发生(即完成一次逻辑运算)时, 根据连接弧类型更新转录因子库所及中间结果库所状态: 对于输入输出弧, 需要根据逻辑运算结果更新关联库所中的时间函数 $s(t)$ 值和托肯值, 对于激活弧与抑制弧则无需更新.

1.4 实验平台

目前已有大量 Petri 网工具用于处理生化模型, 比如基于混合函数 Petri 网的 Cell Illustrator^[8]以及基于时间 Petri 网的 Snoopy^[4]. 但由于需要增加部分逻辑元素, 这些工具显然不能满足我们的实验要求. 另一方面, 如果要重新开发一个 Petri 网工具包又将需要耗费大量的时间和精力. 考虑到本课题目前的研究重点是验证文中所提出基于 Petri 网的生化网络建模与分析方法的正确性与有效性, 而不仅仅开发某一 Petri 网工具本身. 最终我们选择基于开源项目 Petri 网工具包(PNK)及系统生物学软件包(SBW)搭建一个基于 Petri 网的系统生物学实验平台, 如图 2 所示. 这样可以大大节省搭建实验平台的工作量, 从而将更多的精力与时间用于实验部分.

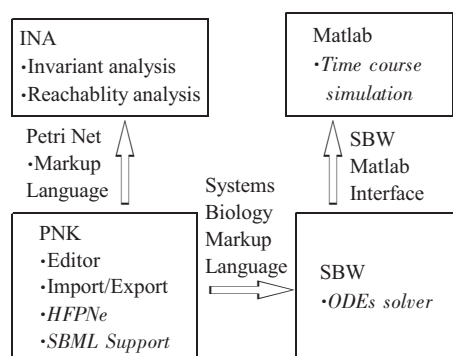


Fig. 2 Architecture of Petri Net Workbench (new feature in italic)

该 Petri 网实验平台具体包括以下方面的扩展:

- Extend of Hybrid Functional Petri Net(混合函数 Petri 网扩展模块). 定义 HFPNe, 允许以可视化方式编辑该类型 Petri 网, 还支持以 Petri 网标记语言(PNML)格式将模型导出到 Petri 网分析工具(INA)进行分析.
- ODEs solver(微分方程求解模块). 实现与系统生物学工作平台(SBW)连接, 允许对 Petri 网模型进行基于 ODEs 的仿真.
- SBML Support(SBML 支持模块). 支持 Petri 网模型以系统生物学标记语言(SBML)导入或导出系统. 转化过程中, SBML 中的生化反应对应 Petri 网中的变迁(Transition), 而反应物对应于库所(Place).
- Time Course Simulation(时间过程仿真模块). 利用 SBW 的 Matlab 接口将实验仿真数据以图形方式输出. 此前, 我们已经基于该平台构建一些生化网络并验证了平台中 HFPN 模块的正确性^[16]. 本文主要基于该实验平台对 HFPN 进一步扩展: 针对基因调控网络增加逻辑库所与逻辑变迁两类基本元素.

2 结果与讨论

2.1 海胆 endo16 基因调控网络

由于海胆胚胎发育形式简单, 在进化上低级, 在现代实验胚胎学中研究比较清楚, 因此非常适于作为分析基因调控网络(GRN)的发育系统. 关于海胆胚胎内中胚层(endomesoderm)特化的基因调控网络(GRN)模型其实是一个临时理论. 它的目的是阐明使这些细胞系中的细胞接受特定命运和命运进一步分化的基因组编码机制. 目前, 海胆内中胚层的基因调控网络包括大约 50 个基因, 大多是调控基因. 本文以海胆 endo16 基因调控网络作为实验模型. 被誉为“破解发育之谜的真正的先驱之一”的 Davidson 长期以海胆为材料研究早期发育的调控机制. 他们领导的实验室自 20 世纪 80 年代开始致力于基因的转录调控机制研究. 他们对海胆 endo16 基因顺式调控序列的研究到目前为止仍是人们对基因调控机制了解最清楚的基因之一. Davidson^[14]的工作表明, 海胆 endo16 基因调控网络可以类比为复杂的生物计算机, 负责解读发育过程中千变万化的各种信息, 最终决定每个细胞的命运. 海胆 endo16 基因表达是由每个细胞特异的“调控状态”, 即每个细胞中特异表达转录因子的组合状态所决定的. 根据转录因子活性在生命周期中的变化, 可分为两类^[14, 15]: 时间依赖输入(time-varying input)与非时间依赖输入(time invariant input). 第一

类转录因子是指在胚胎发育周期中该转录因子的活性(即对基因表达水平的影响程度)随着时间发生变化. 通常这一类转录因子的活性可视为时间驱动函数, 作为基因调控主要的输入信号; 第二类转录因子与生长时间并无直接联系, 但当关联的时间依赖转录因子的活性变化超过一定阈值时, 该转录因子的活性作为调控因子将发生变化, 从而影响基因表达调控模式.

2.2 Petri 网模型构建

本文基于 Davison 等建立的海胆 *endo16* 基因调控网络^[14, 15]构建了 Petri 网模型(图 3). 该 Petri 网模型由 27 个库所和 12 个变迁组成. 其中, 包括 3 个表示时间依赖输入转录因子的逻辑库所, 9 个表示非时间依赖输入转录因子的连续库所, 3 个表示上游抑制因子的离散库所, 12 个存放中间计算结果的连续库所, 12 个表示 GRN 逻辑调控规则的逻辑变迁.

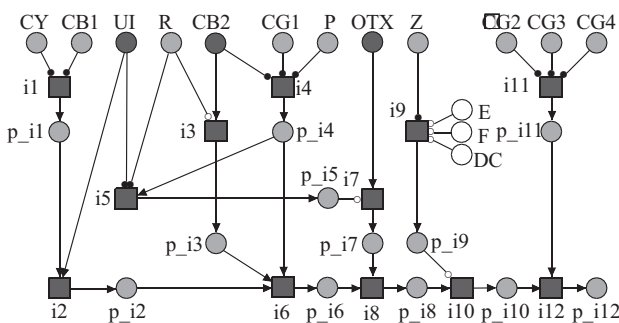


Fig. 3 Petri net model of the genetic regulatory network of the sea urchin *endo16* gene^[14, 15]

海胆 *endo16* 基因调控网络包括两个特定的调控模块: 模块 A 和模块 B. 调控模块 A 包括 6 个位点(CG1、P、OTX、CG2、CG3、CG4), 负责胚胎发育早期的基因表达调控; 调控模块 B 包括 6 个位点(CY、CB1、UI、R、CB2)负责胚胎发育后期的基因表达调控. 该 GRN 模型包括 4 个布尔逻辑子系统. a. CY 与 CB1 位点控制的输入增强子系统(i1, i2)负责当这 2 个位点均有转录因子结合时增强 UI 输入; b. CB2、P 与 CG1 位点控制的 BA 连接子系统(i3, i4, i6)负责将模块 B 的调控输出传递给模块 A 并将其放大; c. R 位点控制的模块 BA 内部输入开关(i5, i7)负责当 UI 输入超出阈值时(通常在囊胚后期出现), 切断模块 A 的 OTX 输入, 模块 B 成为主要的调控方式; d. Z 位点控制的关闭开关(i9)负责当上游抑制因子 F、E、DC 中的某一个结合到位点时关闭基因表达.

2.3 文献实验数据

Davidson 等采用多种扰乱基因表达的方法, 进行了大量干扰分析. 转录因子结合位点发生突变或者转录因子被抑制或发生缺失, 将导致海胆 *endo16* 基因的表达被阻碍或者改变, 采用定量 PCR 方法测量海胆胚胎细胞中特异表达转录因子的组合状态对于海胆 *endo16* 基因表达的影响. 在这些干扰实验中, 针对不同的转录因子结合位点发生突变(mutation), Davidson 等^[14, 15]做了几组实验, 观察海胆 *endo16* 基因的表达水平变化情况. a. 野生型(即 BA 型)经过 20 h 发育后, 基因表达活性开始上升, 在 30 h 左右活性下降, 在 60 h 左右迅速增强, 基因表达出现晚期上升现象. b. 当 CY 位点被破坏后得到 B(CYm)A 型, 基因表达仍出现晚期上升特征, 只是程度上相对弱一点. 实验还发现 B(CYm)A 型、B(CB1m)A 型的基因表达与 B(CYmCB1m)A 型一样. c. 当 CB2 位点发生变异后得到 B(CB2m)A 型, 晚期表达活性上升的特征几乎全部消失. 实验还发现 BA(Pm)型、BA(CG1m)型的基因表达均与 B(CB2m)A 型相同. d. 位点 R 变异后得到 B(Rm)A 型, 基因表达与野生型 BA 相比在后期显著上升.

2.4 Petri 网模型预测结果

通过改变 Petri 网模型中转录因子库所的初始值参数, 根据 HFPNe 中的变迁发生规则, 对该 Petri 网模型进行动态行为模拟. 预测海胆 *endo16* 基因在不同位点发生突变(或转录因子缺失)时表达水平的变化. 如图 4 所示, 横坐标表示海胆胚胎发育周期, 纵坐标表示海胆 *endo16* 基因的表达水平. 每条曲线表示对于某一变异类型, 海胆 *endo16* 基因在海胆胚胎发育不同阶段中的表达水平变化.

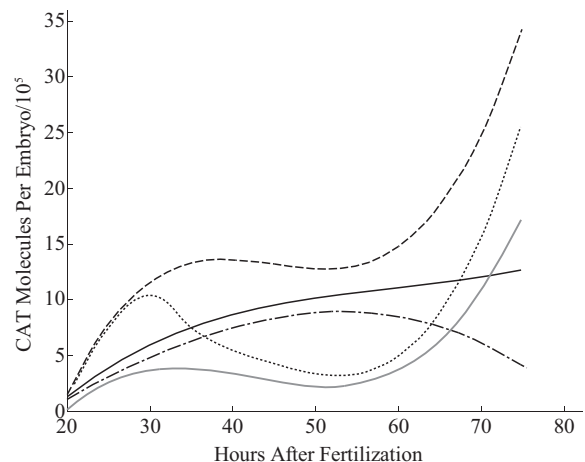


Fig. 4 Different outputs for different mutations

..... BA; —: B(Ulm)A; ---: B(CB2m)A; -.-: B(Rm)A; —: B(CYm)A.

根据该 Petri 网模型预测结果, 结合海胆 *endo16* 基因调控逻辑, 以基因组调控编码的形式解释 Davidson 等在实验中观察到基因表达水平变化现象. a. 图 4 中的 BA 曲线表明对于野生型(BA 型)经过 20 h 发育后, 海胆 *endo16* 基因表达活性开始上升, 在 30 h 左右由于 UI 输入超过阈值, 系统切断模块 A 的 OTX 输入, 活性下降, 在 60 h 左右 UI 输入迅速增强, 基因表达出现晚期上升现象. b. 当 CY 位点被破坏后, B(CYm)A 型输出仍满足晚期上升特征, 只是程度上相对弱一点. 预测结果中还发现 B(CYm)A 型、B(CB1m)A 型的输出与 B(CYmCB1m)A 一样, 表明这两个位点均为增强 UI 输入的必要条件, 满足逻辑与关系. c. 当 CB2 位点发生变异, B(CB2m) 型从功能上将模块 B 与模块 A 分开, 晚期表达活性上升的特征几乎全部消失, 表明 CB2 位点是模块 B 输出传入模块 A 并对其放大的必要条件. 实际上, 预测结果还表明模块 A 中的 P 位点和 CG1 位点也是该调控行为的必要条件, 即 P、CG1 与 CB2 位点上的转录因子满足逻辑与关系. d. 位点 R 变异后, 导致 OTX 去抑制现象, 即使 UI 输入超出阈值, 模块 A 的 OTX 输入仍未被切断, 因此 B(Rm)A 型的活性与野生型 BA 相比在后期显著上升.

这些预测结果均与参考文献[14, 15]的实验结果相一致. 此外, 模型还产生许多额外的预测, 均可通过实验进行验证. 例如, 模型预测得出结论: 当上游抑制因子 F、E、DC 中的任意一个结合到位点时均会关闭基因表达, 即这 3 个转录因子作用满足逻辑或关系.

3 结 论

本文主要基于混合函数 Petri 网, 通过引入逻辑库所与逻辑变迁来描述基因调控网络逻辑规则以及转录因子间的协同作用, 并构建海胆 *endo16* 基因调控网络验证了该方法的正确性. 值得注意的是, 本文对基因调控网络的分析主要集中在模型动态行为预测上. 实际上, 利用 Petri 网的分析理论(如不变量分析、可达性分析等)进行拓扑结构分析往往有助于生物学家对复杂生物系统的理解, 参见我们的前期工作^[16].

基于混合函数 Petri 网构建生化网络模型, 也存在着一些不足. 由于 Petri 网中变迁发生规则的限制, 可达性等方面分析的算法复杂度往往随着网络规模增大而迅速增加. 到目前为止, 大部分研究

集中于不超过 100 个结点的生化网络. 实际上, 如何处理大规模生化网络也是系统生物学目前的一个研究热点. 此外, 虽然基于 Petri 网的建模理论在代谢网络、信号传导途径、基因调控网络均已有关报道, 但是如何将各种类型生化网络集成到一个生化模型, 使其理论预测能够更真实地反映出生物系统的实际行为, 仍然需要大量的研究工作.

参 考 文 献

- Chen M, Hofstaedt R. Quantitative petri net model of gene regulated metabolic networks in the cell. In *Silico Biology*, 2003, **3** (29): 347~365
- Vohradsky J. Neural network model of gene expression. *Faseb J*, 2001, **15**: 846~854
- Cherry J L, Adler F R. How to make a biological switch, *J Theor Biol*, 2000, **203**: 117~133
- Sackmann A, Heiner M, Koch I. Application of petri net based analysis techniques to signal transduction pathways. *BMC Bioinformatics*, 2006, **17**(482): 482~500
- Peleg M, Rubin D, Altman B. Using petri net tools to study properties and dynamics of biological systems. *Journal of the American Medical Informatics Association*, 2005, **12**(2): 181~199
- 罗若愚, 廖莎, 曾绍群, 等. FluxExplorer: 一个通用的基于化学计量矩阵的代谢网络建模和分析平台. *科学通报*, 2006, **51**(6): 665~671
- Luo R Y, Liao S, Zeng S Q, *et al.* Chinese Science Bulletin (in Chinese), 2006, **51**(6): 665~671
- Steggless L J, Banks R, Wipat A. Modelling and Analysing Genetic Networks From Boolean Networks to Petri Net. Heidelberg: Springer Berlin, 2006. 127~141
- Doi A, Fujita S, Matsuno H, *et al.* Constructing biological pathway models with hybrid functional petri nets. In *Silico Biol*, 2004, **14** (23): 271~291
- Gilbert D, Heiner M. From Petri nets to differential equations—an integrative approach for biochemical network analysis. The 27th International Conference on Applications and Theory of Petri Nets, Turku, Finland, 2006
- 薛文, 王进, 黄启来, 等. 真核基因转录激活的多位点协同调控. *生物化学与生物物理进展*, 2002, **29**(4): 510~513
- Xue W, Wang J, Huang Q L, *et al.* *Prog Biochem Biophys*, 2002, **29** (4): 510~513
- 张昆林, 张静, 罗静初. 酵母基因上游与内含子可能存在的转录协同作用. *生物化学与生物物理进展*, 2005, **32**(1): 46~52
- Zhang K L, Zhang J, Luo J C. Potential transcriptional synergy between upstream regions and introns of yeast genes. *Prog Biochem Biophys*, 2005, **32**(1): 46~52
- Maria J, Schilstra A, Bolouri H. Modelling the regulation of gene expression in genetic regulatory networks. <http://strc.herts.ac.uk/bio/maria/Apostrophe/index.html>, 2001-11-03
- 袁崇义. Petri 网原理与应用. 北京: 电子工业出版社, 2005
- Yuan C Y. Principle and Application of Petri Nets. Beijing:

- Publishing House of Electronics Industry, 2005
- 14 Yuh C H, Boluri H, Davidson E H. Cis regulatory logic in the endo16 gene switching from a specification to a differentiation mode of control. *Development*, 2001, **128**: 617~629
- 15 Davidson E H, Rast J P, Yuh C H. A genomic regulatory network for development. *Science*, 2002, **295**(5560): 1669~1678
- 16 Huang J L, Guo H, Meng C. Application of petri net based analysis techniques to biochemical networks. 2007 International Conference on Bio-inspired Computing: Theories and Applications, Zhengzhou, China, 2007

Modeling and Analysis of Genetic Regulatory Networks Based on Petri Net

HUANG Jia-Liang¹⁾, GUO Hong¹⁾, MENG Chun^{2)*}

¹⁾ College of Mathematics and Computer Science, Fuzhou University, Fuzhou 350002, China;

²⁾ College of Biological Science and Technology, Fuzhou University, Fuzhou 350002, China)

Abstract With the success of the human genomic project, the amount of experimental data is increasing rapidly. In the post-genomic era, the focus is now shifting to the so called "from the sequence to the function", i.e., in addition to completing genome sequences, it is possible to learn about gene expression patterns and protein interactions on the genomic scale. Description and analysis of complex Genetic Regulatory Networks (GRN) is a critical problem for biologists to understand genetic regulatory mechanism. Most of existing methods ignore the synergistic effects observed widely in biologic systems. Thus there should be some errors between the predictions of the model and the actual biologic behaviors. To address this problem, a new quantitative analysis approach for genetic regulatory networks was proposed in the context of Hybrid Functional Petri Net (HFPN). Firstly, basic theory of GRN and HFPN was presented briefly. Petri-net-based models have been widely adopted for studying biological systems since Petri net has graphical modeling representation and strict mathematic background. In GRN, the combined effect of transcription factors to induce or repress gene transcription is usually different from the simple sum of their individual effects. This is so called Synergistic effect. Two kinds of new elements, logic places and logic transitions were introduced to describe the logic rules in GRN and the synergy between transcription factors. It was decided to extend existing tools instead of writing new software in order to have more time for the experimental part of this project. A Petri Net Workbench was developed based on the Open Source project the Petri Net Kernel (PNK) and the Systems Biology Workbench (SBW) during this project. The Petri Net Workbench was extended by the following features: the extended definition of Hybrid Functional Petri Net, the Ordinary Differential Equations (ODEs) solver, Systems Biology Markup Language (SBML) support and time course simulation. Finally, a Petri net model for the genetic regulatory network of the sea urchin endo16 gene was developed base on the GRN model published in literature and a list of quantitative outputs for different mutations was predicted. The analysis result corresponded to the experimental data published in literature properly and demonstrated the correctness of the Petri net model. It demonstrated how Hybrid Functional Petri net analysis techniques can be applied to Genetic Regulatory Networks. The new concepts of logic places and logic transitions had been proven to be useful for describe the logic rules in GRN and the synergy between transcription factors.

Key words hybrid functional Petri net, genetic regulatory networks, sea urchin endo16 gene, synergy, logic rules

*This work was supported by The Natural Science Foundation of Fujian Province of China (C0510006).

**Corresponding author .

Tel: 13509369204, E-mail: mengchun@fzu.edu.cn

Received: August 3, 2007 Accepted: November 7, 2007