

电子细胞模型 Analog-Cell 中前体 mRNA 剪接过程的模拟与研究*

欧阳丹彤^{1,2}, 王 珏^{1,2}, 韩霄松^{1,2}, 卢欣华^{1,2+}

¹(吉林大学 计算机科学与技术学院, 吉林 长春 130012)

²(吉林大学 符号计算与知识工程教育部重点实验室, 吉林 长春 130012)

Simulation and Research on Pre-mRNA Splicing in Analog-Cell

OUYANG Dan-Tong^{1,2}, WANG Jue^{1,2}, HAN Xiao-Song^{1,2}, LU Xin-Hua^{1,2+}

¹(College of Computer Science and Technology, Jilin University, Changchun 130012, China)

²(Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education, Jilin University, Changchun 130012, China)

+ Corresponding author: E-mail: luxh@jlu.edu.cn

Ouyang DT, Wang J, Han XS, Lu XH. Simulation and research on pre-mRNA splicing in Analog-Cell. *Journal of Software*, 2012, 23(9): 2273-2284 (in Chinese). <http://www.jos.org.cn/1000-9825/4121.htm>

Abstract: Pre-mRNA splicing is a key process in the gene expression of eukaryotic cells, which includes two major procedures: the cutting of introns and the union of exons. This paper proposes an algorithm to simulate the Pre-mRNA splicing process in the Analog-Cell, which was developed independently. Using the algorithm, the correct results are consistent with the biological principles. Through simulation of the process, the study has found that this algorithm has the capacity to make the introns to form the lariat structure efficiently and accurately, meanwhile to unit the exons selectively. In this way, the Analog-Cell can finish the formation of the mature mRNA.

Key words: electronic cell; Analog-Cell; alternative splicing; pre-mRNA splicing

摘 要: 真核细胞前体 mRNA 的剪接加工包含内含子剪切和外显子拼接两个过程,是真核细胞基因表达过程中的一个重要环节.针对这一环节,提出了一种模拟真核细胞前体 mRNA 内含子剪切及其选择性剪接的算法,并在自主研发的电子细胞模型 Analog-Cell 中实现了该算法,且获得了符合生物学原理的模拟结果.模拟结果表明,该算法可使剪接体高效、准确地模拟内含子的剪接过程,并形成套索结构,同时有选择地将外显子片段拼接起来,最终形成成熟的 mRNA.

关键词: 电子细胞; Analog-Cell; 选择性剪接; 前体 mRNA 的剪接

中图法分类号: TP301 **文献标识码:** A

近年来,随着生命科学领域研究的深入,使用计算机构建细胞模型,即应用精确的数理计算模型模拟复杂的细胞生命活动过程,逐渐成为计算生物学的研究热点之一.日本和美国的一些大学相继研制出一种称为电子细

* 基金项目: 国家自然科学基金(60973089, 61133011, 61170092, 61003101); 吉林省科技发展计划(20101501, 20100185); 浙江省自然科学基金(Y1100191); 吉林大学符号计算与知识工程教育部重点实验室开放项目(93K-17-2009-K05); 欧盟合作项目(155776-EM-1-2009-1-IT-ERAMUNDUS-ECW-L12)

收稿时间: 2011-04-01; 定稿时间: 2011-09-01

胞(electronic cell,简称 E-Cell)的细胞建模软件^[1,2].这种软件基于人们对细胞掌握的知识,应用现代计算机科学技术,构建出一种虚拟现实的软件环境,在分子层面上模拟细胞的一系列生命活动,如 DNA 解旋、RNA 翻译、膜转运等等.生物学家使用 E-Cell 可以在模拟的过程中,寻找细胞生命活动中潜在的、未被发觉的规律.

2007年,本项目组构建了国内第一个电子细胞模型 Analog-Cell^[3].Analog-Cell 以真核细胞基因表达为模拟对象,采用状态控制机制,在核苷酸分子层面上模拟基因通过解旋、转录、mRNA 加工和翻译,表达成蛋白质的过程,并模拟了基因点突变以及一系列基因表达调控反应.然而,Analog-Cell 中模拟前体 mRNA 加工,仅仅模拟了前体 mRNA 的加帽反应,而内含子剪切及外显子选择性拼接等真核细胞中涉及到基因表达和调控的重要生命活动并没有进行模拟.国际上较为成熟电子细胞,如美国的 Virtual Cell、日本的 E-Cell,研究重点多是在细胞模型的建立,重点探讨的是细胞模型中的数值计算方法^[4]、图形表达^[5]以及生物编程环境^[6],而对细胞模型的实现部分的研究却不是很多.美国康涅狄格大学的 Virtual Cell 采用 J2EE 构建的客户端-服务器系统结构,生物工作者可以使用 Virtual Cell 提供的生物环境编程接口来定制各种不同的细胞及各种细胞生命活动^[7].日本庆应大学的 E-Cell 是世界上最早的电子细胞模型,该细胞模型曾取得多项科研成果,例如对疾病基因的预测、探讨细胞生命活动的未知规律等等^[8].

为了精确地模拟真核细胞的生命活动,本文在已有 Analog-Cell 成果的基础上,提出了剪接体在细胞核中对前体 mRNA 进行剪接加工的算法,完善了 Analog-Cell 中的前体 mRNA 的加工过程,即增加了对前体 mRNA 的内含子剪切和外显子选择性表达这两个过程的模拟.

本文第 1 节介绍前体 mRNA 剪接过程的生物学原理.第 2 节给出计算机模拟细胞基因剪接过程算法.第 3 节将应用这一算法在 Analog-Cell 中得到的模拟结果.

1 前体 mRNA 的加工和修饰

真核生物细胞的 DNA 经过转录后,形成前体 mRNA.前体 mRNA 不能立即进行翻译,需要对其进行一系列的加工,以形成成熟的 mRNA^[9].这一过程首先是在前体 mRNA 的 5'端加一个鸟苷帽子,在 3'端添加一段腺嘌呤序列,然后在剪接体的催化下进行基因剪接过程.即由剪接体将前体 mRNA 链上不参与基因表达的内含子部分剪掉,同时,剪接体还将需要表达的外显子片段拼接起来.其中,在前体 mRNA 中内含子进行剪切的位置,或者外显子进行拼接的位置,称为剪接位点^[10].最终,成熟的 mRNA 游离出细胞核进行翻译过程.这一过程如图 1 所示.

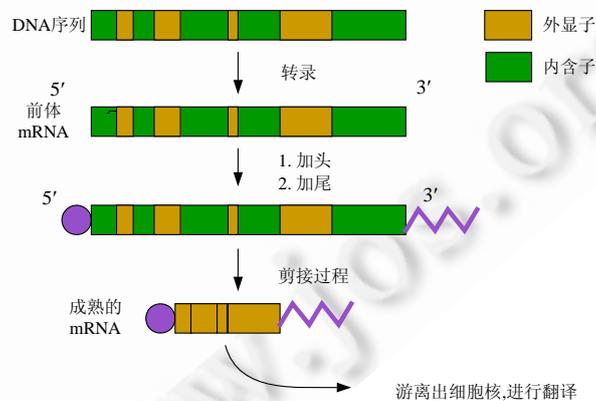


Fig.1 Eukaryotic mRNA processing

图 1 真核细胞中 mRNA 的加工过程

1.1 内含子剪切过程

真核细胞中的内含子剪切过程需要进行两次转酯反应^[11].第 1 步是靠近内含子 3'端内含子分支位点的 A

残基的 2'-OH 对 5'端的磷酸基团进行亲核攻击,形成 2'-5'磷酸二酯键连接的套环结构;在第 2 步转酯反应中,被剪接的外显子 3'核苷酸分子的-OH 进攻 3'内含子末端的磷酸基团,最后在 3'内含子外显子剪接点处断开,套索结构释放,两个外显子建立连接.这一过程如图 2^[11]所示.

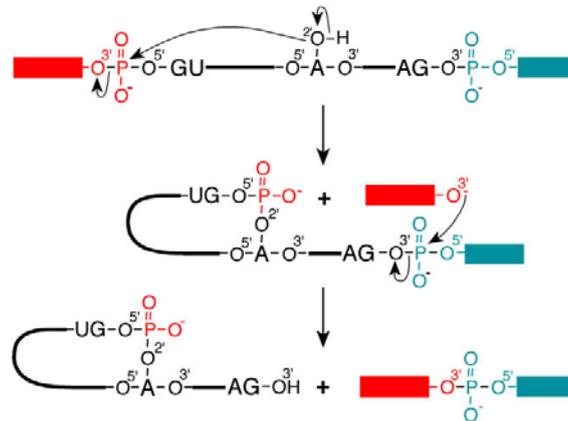


Fig.2 Pre-mRNA splicing process

图 2 内含子剪切的过程

1.2 剪接体

剪接过程中剪接体起到催化剪接反应的功能,实际上是由多种因子组装的复合体,包括杂性核 RNA 结合蛋白(hnRNP)、小型核 RNA 结合蛋白(snRNP)等等.Reed 等人通过蛋白质组学方法证明,与前体 mRNA 剪接相关的蛋白质因子大约有 145 种,它们参与 RNA 剪接过程中的不同环节^[9,10].图 3 是组成剪接体最主要的几个剪接因子参与内含子剪切的组装过程^[9].

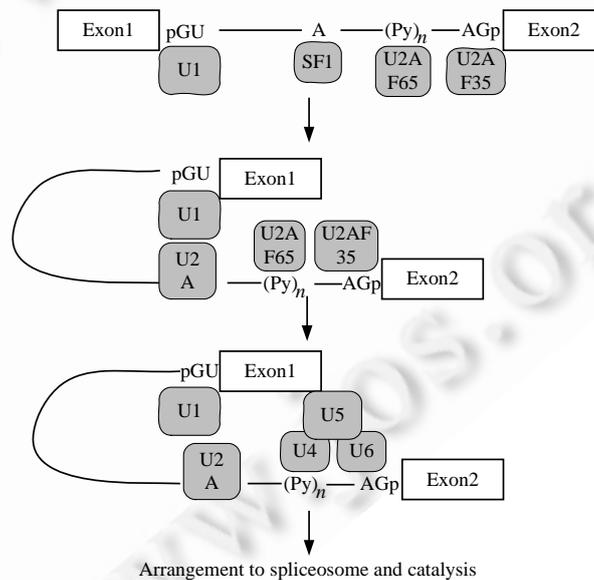


Fig.3 Process of spliceosome assembly

图 3 剪接体的组装过程

当前,生物学界对这些蛋白因子的研究尚不完整,所以目前不对这些小型蛋白进行模拟,而是在 Analog-Cell 的细胞核中添加一种新物质代表剪接体,并且在反应规则集中为这个新物质添加相应的反应规则.

1.3 选择性剪接

前体 mRNA 的选择性剪接就是有选择性地需要将进行基因表达的外显子拼接成为成熟的 mRNA,即在同一段基因中,其剪接位点和拼接方式可以改变^[11].选择性剪接在细胞分化的基因调控中起到重要的作用,使得同一段基因可以产生多种不同的蛋白,以应对外界环境变化和行使不同的生理功能.对个体分化发育和一些关键的细胞生理过程如凋亡、细胞兴奋等的精确调控有重要意义^[14,15].其过程如图 4 所示.

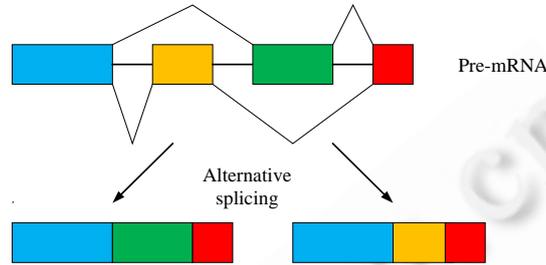


Fig.4 Elementary alternative splicing events

图 4 选择性剪接示例

以上我们对真核生物细胞中前体 mRNA 剪接过程的生物学原理进行了简要描述.可见,真核细胞剪接过程的主要过程是:从前体 mRNA 上将内含子的剪掉,以及选择性地需要将进行基因表达的外显子部分拼接起来.

2 Analog-Cell 中前体 mRNA 剪接过程的模拟方法

我们在 Analog-Cell 中提出一种新算法,模拟本文第 1 节叙述的前体 mRNA 的剪接过程.

算法的开始条件是,当游离的剪接体移动到前体 mRNA 的 5'端附近时,剪接体会发生反应并附着在前体 mRNA 的 5'端,然后在基因链上移动,当移动到剪接位点时进行剪接反应.由于目前剪接位点预测技术的精度不是很理想,所以当前使用剪接位点预测算法不能满足实验的精确性,因此使用在核酸序列数据库 GenBank 中经实验验证过的剪接位点信息.

我们在前体 mRNA 基因链上为每一个核苷酸分子添加一个递增的 ID 值属性,代表此核苷酸分子在前体 mRNA 链中的位置.同时,也为剪接体添加一个 ID 值(初始值为 0),标识剪接体当前在前体 mRNA 上的位置.主要有两个作用:

- 第 1 是剪接体通过判别 ID 值的大小,使其在基因链上沿着从 5'端到 3'端的方向移动;
- 第 2 是结合查询 GeneBank 获得的剪切位点信息,使剪接体准确地在前体 mRNA 剪接位点处进行剪接反应.

2.1 模拟剪接体移动的反应规则与算法

在 Analog-Cell 中,反应底物的状态可以在反应过程中发生改变,如果几种反应使底物的状态改变形成循环,就能够实现几种反应的交替进行,形成一种类似于自动机的反应循环机制.例如,剪接体在前体 mRNA 上的移动就是由两个反应(Reaction2 和 Reaction3)交替进行的^[16].剪接体在基因链上移动的具体反应规则如图 5 所示.

图 5 中,标识为 J 的球代表剪接体,实线代表两个核苷酸分子之间的氢键连接,粗线表明正在形成的氢键,虚线表示正在断开的氢键.反应条件(reaction condition)中记录着触发此反应所需要具备的条件,反应条件之间是“并且”关系,反应之后(after reaction)中记录的是反应的结果.所有的反应都由 3 种底物参与,其中,Cell1 和 Cell2 代表具备反应条件的底物.算法中通过修改剪接体的状态值,控制剪接体进行不同的反应.参照图 5 中的反应规则,利用状态控制模拟剪接体在前体 mRNA 上移动的算法见算法 1.

算法 1. 剪接体移动算法.

1. 如果某段前体 mRNA 未经过剪接,并且前体 mRNA 的 5'端的反应范围存在剪接酶,则进行 Reaction 1:

- 1) 剪接体与前体 mRNA 5' 端的两个核苷酸分子建立氢键连接;
 - 2) 修改参与反应的核苷酸分子的状态;
 - 3) 修改剪接体的状态,进行算法第 2 步;
2. 循环,直到剪接体移动到剪接位点:
- 1) 比较与剪接体连接着的两个反应物的 ID 值,进行 Reaction 2:
 - a) 剪接体断裂与 ID 值较小的反应物的氢键连接;
 - b) 修改剪接体的状态,进行循环第 2 步;
 - 2) 剪接体判断当前 ID 值是否等于内含子剪切位点:
 - 如果为假,进入循环第 3 步;
 - 如果为真,进入算法第 3 步;
 - 3) 剪接体当前 ID 值不是内含子的剪切位点,进行 Reaction 3:
 - a) 剪接体与下一个核苷酸分子进行连接反应;
 - b) 剪接体的 ID 值加 1;
 - c) 修改剪接体状态,跳回至循环第 1 步;
3. 剪接体处于内含子的剪切位点:
修改剪接体的状态,进入剪切过程.

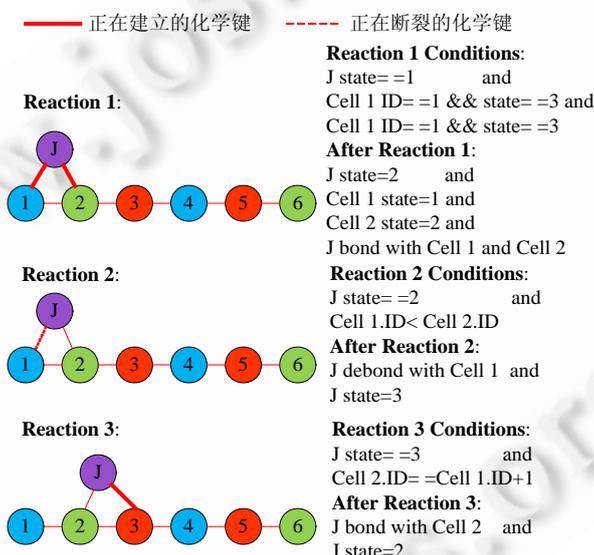


Fig.5 Reactions of the spliceosome movement

图 5 剪接体移动反应

2.2 模拟内含子剪切过程的反应规则与算法

当剪接体移动到内含子的 5' 端的剪接位点时,剪接体触发剪接反应,进入内含子剪切过程.在内含子剪切过程中,剪接体主要进行两次剪切反应和一次连接反应:分别是在内含子的 5' 端和 3' 端位点各进行一次剪切,并催化内含子的头节点与分支位点进行连接.

剪切过程中,第 1 次剪切发生在内含子的 5' 端剪接位点,剪接体催化基因链断裂,之后,剪接体保持与基因链两端核苷酸分子的连接,并继续在内含子上移动.当剪接体在内含子上遇到分支位点时,剪接体催化分支位点处的核苷酸分子与内含子的起始位点处的核苷酸分子建立氢键连接,形成内含子套索结构.接着,剪接体的状态跳

回到在前体 mRNA 上的移动循环,直至遇到内含子的 3'端剪接位点,触发第 2 次剪切反应.这次剪切会从前体 mRNA 上剪掉套索结构的内含子.

在使用内含子剪切功能之前,先要通过 GeneBank 获取所要模拟的前体 mRNA 的基因信息,包括内含子的端点位置、内含子的分支位点等.接着,将这些数据输入 Analog-Cell 中储存基因信息的 STL 容器,程序中使用的 STL 容器中每一个容器内都保存有 5'端剪接位点,内含子的分支位点以及 3'端剪接位点这 3 个数据.每当执行完一次内含子剪切后,该容器指针指向下一个容器,剪接体可以根据这里的数据,把基因链上内含子从前体 mRNA 上精确地剪掉,并形成符合生物学原理的套索结构.

为了描述模拟内含子剪切过程所涉及到的反应,假设一段长度为 13 的前体 mRNA,其中的内含子部分是从 5~11,内含子的分支位点为 9.这一过程中所涉及到的反应如图 6、图 7 所示.

结合图 6 和图 7 中的反应规则,利用状态控制模拟内含子剪切过程的算法如下:

算法 2. 内含子剪切算法.

1. 剪接体在前体 mRNA 上移动到内含子的 5'端位点,进行 Reaction 4:
剪接体断裂此位点处的氢键连接;
2. 剪接体进行 Reaction 5:
 - 1) 剪接体与 ID 值较大一端的核苷酸分子的下一个核苷酸分子建立氢键连接;
 - 2) 剪接体的 ID 值加 1;
3. 循环,直至剪接体移动到分支位点:
 - 1) 剪接体进行 Reaction 6:
 - a) 剪接体与 ID 值较大一端的核苷酸分子的下一个核苷酸分子建立氢键连接;
 - b) 剪接体的 ID 值加 1;
 - 2) 剪接体判断自身当前 ID 值是否为内含子上的分支位点:
如果为假,进入循环第 3 步;
如果为真,进入算法第 4 步;
 - 3) 如果不是分支位点,修改剪接体状态,进行 Reaction 7:
 - a) 剪接体断裂与当前所连接的 ID 值第 2 大的核苷酸分子之间的连接;
 - b) 修改剪接体状态,跳至循环第 1 步;
4. 如果是分支位点,修改剪接体状态,进行 Reaction 8:
 - 1) 剪接体将所连接的内含子头部节点与分支位点建立连接;
 - 2) 剪接体断裂与内含子头节点的连接;
 - 3) 修改剪接体状态,跳至剪接体在前体 mRNA 上移动的循环,直到剪接体移动到内含子的末端位点时,进行算法第 5 步;
5. 剪接体移动到内含子末端位点,进行 Reaction 9:
 - 1) 剪接体与内含子末端节点下一位核苷酸分子建立连接;
 - 2) 剪接体将末端位点与其下一位核苷酸分子断裂连接;
 - 3) 剪接体断裂与末端位点的连接;
 - 4) 完成内含子末端位点剪切,进行算法第 6 步;
6. 剪接体进行 Reaction 10:
将连接着的两段 DNA 链连接;
7. 完成一次内含子剪切过程,修改剪接体状态,剪接体进行 Reaction 2,继续在前体 mRNA 上进行移动.

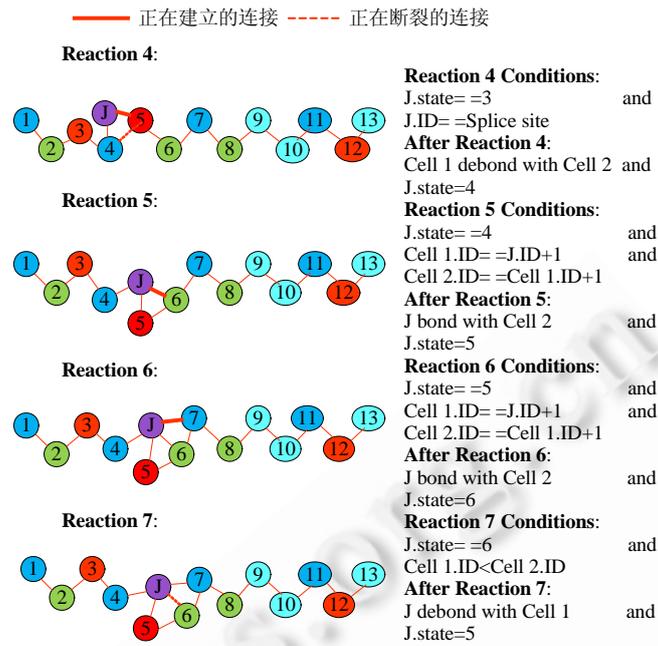


Fig.6 Reactions of spliceosome's first splicing and moving on intron
 图 6 剪接体进行第 1 次剪切及在内含子上移动

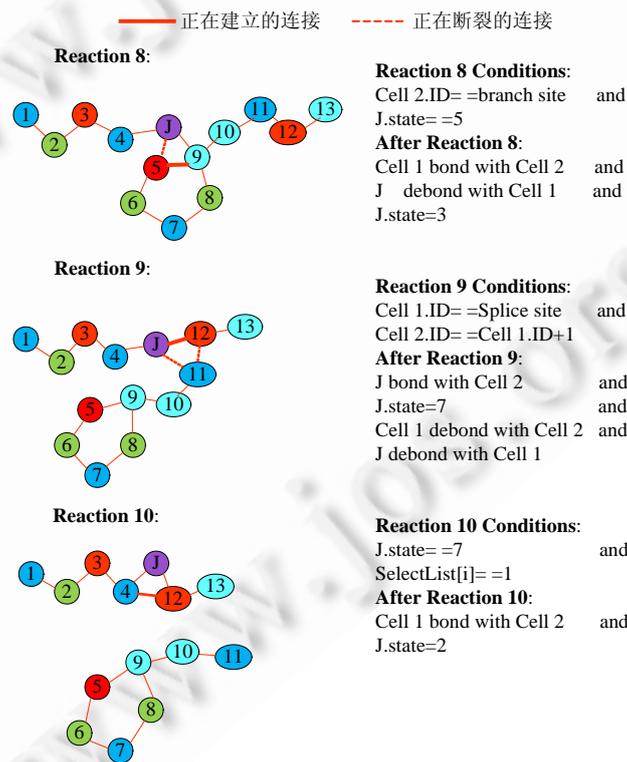


Fig.7 Reactions of spliceosome's second splicing and formation of the lariat structure
 图 7 剪接体第 2 次剪切和内含子套索结构的形成

2.3 模拟外显子选择性表达过程的反应规则与算法

目前,生物学界对细胞的基因选择性表达的研究大多数是为了揭示某个基因的多种剪接表达形式,目前依然没有对于选择性剪接一般性的结论^[18],即现在还无法使用算法准确预测外显子的选择性拼接.为了提高实验的模拟精确度,我们采用外显子选择性向量的方法模拟外显子的选择性剪接.

我们定义的外显子选择性向量是一个包含若干个 1 和 0 构成的数字向量,其中:

- 0 代表对应的外显子片段不表达,即下一段外显子不在成熟的 mRNA 中;
- 1 代表该外显子表达,将被添加到成熟的 mRNA 中,外显子选择性向量的长度根据外显子个数而定.实验时,可根据需要定义这个向量的长度.

例如,

SelectList=(010110...).

这种方法可以让实验者根据需要,自定义外显子选择性向量来模拟不同的基因表达.我们可以将细胞的多个基因的多种外显子选择性向量添加到数据库,这样,细胞可以根据自身目前所处环境和细胞生理的需求,采用不同的外显子选择性向量,进而产生不同的蛋白,这样可以构建细胞随环境改变而采取的自我调节机制.

当剪接体进行算法 1 中的第 1 步,或者进行算法 2 中的第 6 步时,会查找当前外显子选择性向量,如果为 1,则继续进行算法 2;如果为 0,则按照算法 3 进行外显子的剪切.算法 3 及所涉及的反应见算法 3(如图 8 所示).

算法 3. 选择性剪接算法.

1. 当剪接体剪切掉一个内含子片段时,或者当剪接体刚附着在前体 mRNA 的头部时,判断当前选择性向量中的值,判断结束后,向量指针自增,并更新反应队列;
 - 如果为 1,进入算法第 2 步;
 - 如果为 0,进入算法第 3 步;
2. 如果当前外显子向量中的值为 1:
 - 剪接体进行 Reaction 10:将左右两端的氨基酸链连接上;
3. 如果当前外显子向量中的值为 0,剪接体进行 Reaction 11:
 - 剪接体与下一位核苷酸分子建立连接;
4. 循环,直到剪接体移动到前体 mRNA 的 3'端或者外显子的 3'端:
 - 1) 剪接体进行 Reaction 12:
 - 剪接体断裂当前与其连接着的 ID 值第 2 大的核苷酸分子之间的连接;
 - 2) 剪接体进行 Reaction 13:
 - 剪接体与下一位核苷酸分子建立连接;
 - 3) 剪接体判断自己的 ID 值是否等于下一段内含子的前端位点:
 - a) 如果为假,进入循环第 1 步;
 - b) 如果为真,进入算法第 4 步;
 - c) 如果剪接体的 ID 值等于前体 mRNA 的长度值,则进行算法第 5 步;
5. 如果剪接体的 ID 值等于下一段内含子前端位点,则进行 Reaction 14:
 - 1) 剪接体与外显子末端位点下一位核苷酸分子建立连接;
 - 2) 剪接体断裂与外显子末端位点之间的连接;
 - 3) 剪接体断裂外显子的末端位点和其下一位核苷酸分子之间的连接;
 - 4) 修改剪接体的状态,进行内含子剪切反应;
6. 剪接体已经移动到前体 mRNA 的尾部,剪接体脱离前体 mRNA,算法结束.

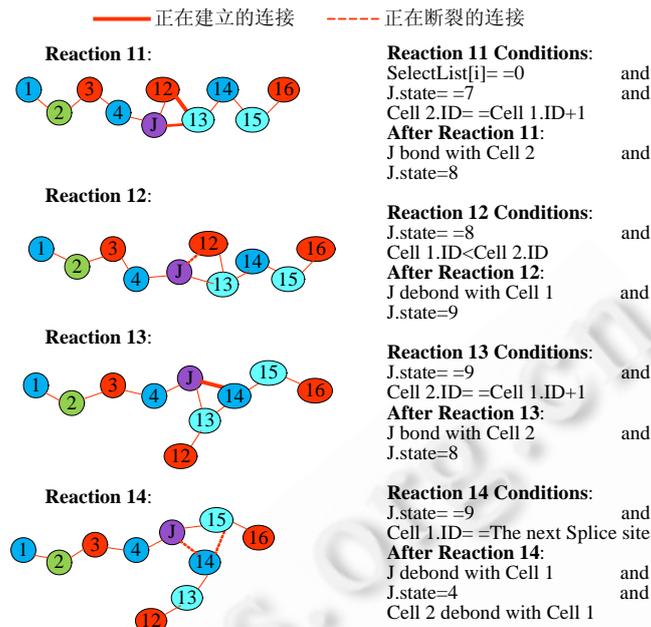


Fig.8 Reactions of spliceosome in alternative splice process

图 8 剪接体参与选择性剪接过程的反应

3 Analog-Cell 中前体 mRNA 剪接过程的模拟结果

我们使用 Analog-Cell 构建了一个特殊的细胞模型,并使用上文中描述的算法模拟真核细胞的基因剪接过程.在这个细胞模型中,有一条长度为 25 的前体 mRNA 链和若干个游离的剪接体,该前体 mRNA 的基因结构如图 9 所示.

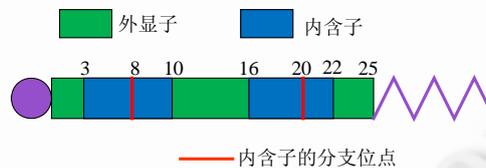


Fig.9 Structure of pre-mRNA in experiment

图 9 实验中的前体 mRNA 的结构

在 Analog-Cell 中输入基因序列后,再将上文中提到的反应规则全部输入进程序.运行程序,可以看到剪接体在前体 mRNA 上移动并进行基因剪接的全部过程.剪接体会在基因链上边移动边判断当前位置是否是剪接位点,如果是,则进入剪接过程,如图 10(a)所示.图 10(b)~图 10(d)中是内含子剪切过程,剪接体依次进行上文中所提到的反应,将内含子以套索形状从前体 mRNA 基因链上剪掉.图 10(e)、图 10(f)是剪接体将不表达的外显子部分从基因链上剪掉.进行了剪接反应过程之后,剪接体会在前体 mRNA 上一直移动下去,直到在基因链的末端脱离基因链.此时,基因剪接过程结束,形成成熟的 mRNA.图 10(a)~图 10(f)是当外显子的选择性序列为(101)时的程序运行截图.

图 10 中,方框中标识的 J 代表剪接体,第 2 个数字是剪接体此时的状态值,最后一位的数字是剪接体的当前 ID 值.图 10(a)是当一个剪接体游离到前体 mRNA 的头部时,剪接体附着在此段前体 mRNA 的头部,而另外一个剪接体状态没有发生改变.图 10(b)是剪接体移动到了第 1 个内含子的 5'端的剪接位点,并开始进行内含子剪切.

图 10(c)是剪接体引导内含子的头部核苷酸分子与该内含子的分支位点建立连接,并形成内含子套索结构.图 10(d)是剪接体从前体 mRNA 上剪掉了一段内含子.图 10(e)中由于选择性向量的第 2 位是 0,所以该段外显子被从前体 mRNA 上剪切下来.图 10(f)是剪接体将此段外显子剪掉,并开始进行第 2 段内含子的剪切.3 个典型的模拟结果如图 11 所示.

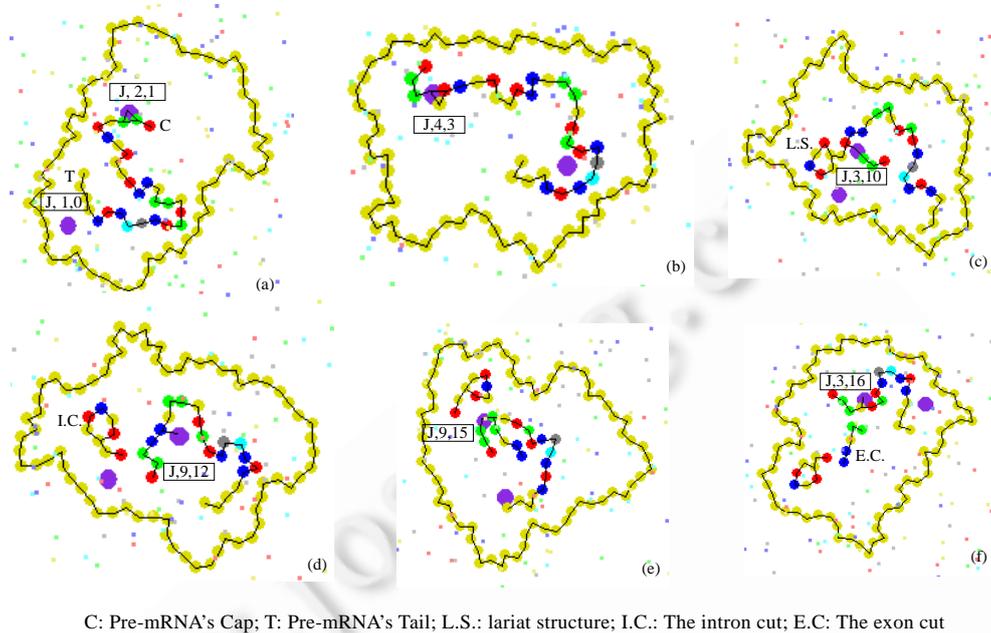


Fig.10 Process of pre-mRNA splicing in experiment

图 10 实验中的前体 mRNA 的剪接过程

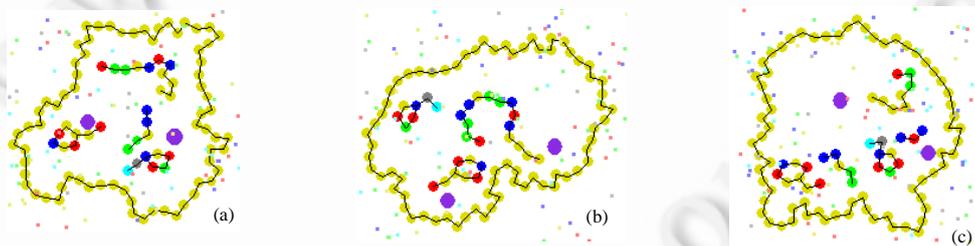


Fig.11 Three typical examples of pre-mRNA splicing result in experiment

图 11 前体 mRNA 剪接模拟结果的 3 个典型示例

图 11(a)是选择性向量为(101)时运行的结果,中间的外显子被剪掉.图 11(b)是选择性向量为(111)时运行的结果,3 个外显子全部保留在成熟的 mRNA 中,两个内含子被剪掉.图 11(c)是选择性剪接向量为(100)时运行的结果,后两个外显子和两个内含子被剪掉.

4 相关工作

由于目前电子细胞的研究比较重视细胞建模以及计算效率,并且生物学界在对细胞剪接过程的研究也不是很彻底,到目前为止还没有出现关于细胞基因剪接的相关模拟算法.在一些电子细胞中,比如 Smart Cell 和 VICE Cell 的模拟还停留在原核生物的层面上,没有对剪接过程的模拟.而在一些较为成熟的电子细胞中,例如 Virtual Cell 和 E-Cell,它们可以对真核生物细胞进行模拟,但出于提高计算效率和模拟速度的目的,它们没有使

用算法动态随机地模拟剪接过程.目前,支持随机动态模拟基因剪接过程的只有 E-Cell,但是在 E-Cell 中没有图形化界面模拟剪接过程.

在 Analog-Cell 中应用本文提出的算法,增强了细胞模拟的深度,为了模拟更为复杂的真核生物细胞打下基础.并且,借助 Analog-Cell 的图形化显示以及动态随机性等模拟特性,使得前体 mRNA 剪接的模拟相比较其他的电子细胞具有一定的优势.

对于剪接过程动态模拟的研究,还与其他电子细胞相比,几个主要的电子细胞^[18-22]之间的对比见表 1.

Table 1 Contrast between several major electronic cells

表 1 几种主流电子细胞之间的比较

E-Cells	Simulate mRNA splice process	Simulate eukaryotic cell	Graphical interface	Nation & birth time	Application
Virtual cell	No	Yes	No	America, 2004	A software environment for computational cell biology
E-Cell	Yes	Yes	No	Japan, 1998	A computational system for constructing whole cell models
Cyber cell	No	Yes	Yes	Canada, 2003	Simulator for complex metabolic and chemical pathways
VICE cell	No	No	No	Italy, 2005	Success to simulate prokaryotic cell to survive in normal environment
Smart cell	No	No	No	Germany, 2004	A framework to simulate cellular processes that combines stochastic approximation with diffusion and localisation
Analog-Cell	Yes	Yes	Yes	China, 2007	Permit user to observe the stochastic activities of cell in molecular level

5 结 论

本文提出了一种新的基于电子细胞模型 Analog-Cell 模拟真核细胞前体 mRNA 剪接过程的算法.该算法具有两个特点:(1) 依据细胞学研究成果,模拟出功能一致,结构类似的剪接体及由剪接过程所形成的套索结构内含子;(2) 模拟了真核细胞的选择性剪接,为 Analog-Cell 下一步模拟基于环境变化的基因调控打下基础.

致谢 在此,我们向对本文工作给予支持和建议的同行,尤其是吉林大学计算机科学与技术学院及智能信息处理实验室讨论班上的同学们和老师表示感谢.

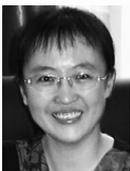
References:

- [1] Zhao MS, Shang T, Sun DY, Jiang JH, Tang J, Wu YS. The current state and prospect of E-Cell. *Acta Electronica Sinica*, 2001,29(12A):1740-1743 (in Chinese with English abstract).
- [2] Loew LM, Shaff JC. The virtual cell: A software environment for computational cell biology. *TRENDS in Cell Biology*, 2001, 19(10):401-406. [doi: 10.1016/S0167-7799(01)01740-1]
- [3] Lu XH, Sun JG. Analog-Cell: A new pattern model of e-cell. *Acta Electronica Sinica*, 2007,35(1):49-53 (in Chinese with English abstract).
- [4] Csikasz-Nagy A. Computational systems biology of the cell cycle. *Briefings in Bioinformatics*, 2009,10(4):424-434. [doi: 10.1093/bib/bbp005]
- [5] Cheong R, Paliwal S, Levchenko A. Models at the single cell level. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 2010,2(1):34-48. [doi: 10.1002/wsbm.49]
- [6] Hucka M, Finney A, Sauro HM, Bolouri H, Doyle JC, Kitano H. The system biology markup language: A medium for representation and exchange of biochemical network models. *Bioinformatics*, 2003,19(4):524-531. [doi: 10.1093/bioinformatics/btg015]
- [7] Li W. Research on chemical reaction quantitative calculation method and algorithm in virtual cell [MS Thesis]. Harbin: Harbin Institute of Technology, 2006 (in Chinese).
- [8] Zhang EH. Research on semi-automatic constructs virtual cells ontology [MS Thesis]. Harbin: Harbin Institute of Technology, 2009 (in Chinese).

- [9] Black DL. Mechanisms of alternative pre-messenger RNA splicing. Annual Review of Biochemistry, 2003,72:291–336. [doi: 10.1146/annurev.biochem.72.121801.161720]
- [10] Wu Y, He RQ. Structural analyses of RNA splicing factors: progress and perspective, 2003,30(4):503–508 (in Chinese with English abstract).
- [11] Schellenberg MJ, Ritche DB, MacMilan AM. Pre-mRNA splicing: A complex picture in higher definition. Trends in Biochemical Science, 2008,33(6):243–246. [doi: 10.1016/j.tibs.2008.04.004]
- [12] Sperling J, Azubel M, Sperling R. Structure and function of the pre-mRNA splicing machine. Structure, 2008,16(11):1605–1615. [doi: 10.1016/j.str.2008.08.011]
- [13] Staley JP, Jr. Woolford JL. Assembly of ribosomes and spliceosome: Complex ribonucleoprotein machines. Current Opinion in Cell Biology, 2009,21(1):109–118. [doi: 10.1016/j.ceb.2009.01.003]
- [14] Matlin AJ, Clark F, Smith CWJ. Understanding alternative splicing: Towards a cellular code. Nature Reviews, 2005,6(5):386–398. [doi:10.1038/nrm1645]
- [15] Chen M, Manley JL. Mechanisms of alternative splicing regulation: Insights from molecular and genomics approaches. Nature Reviews Molecular Cell Biology, 2009,10(11):741–754. [doi:10.1038/nrm2777]
- [16] Hutton TJ. Evolvable self-replicating molecules in an artificial chemistry. Artificial Life, 2002,8(4):341–356. [doi:10.1162/106454602321202417]
- [17] Nilsen TW, Graveley BR. Expansion of the eukaryotic proteome by alternative splicing. Nature, 2010,463:457–463. [doi:10.1038/nature08909]
- [18] Loew LM, Schaff JC. The virtual cell: A software environment for computational cell biology. Trends in Biotechnology, 2001, 19(10):401–406. [doi: 10.1016/S0167-7799(01)01740-1]
- [19] Tomita M, Hashimoto K, Takahashi K, Shimizu TS, Matsuzaki Y, Miyoshi F, Saito K, Tanida S, Yugi K, Venter JC, Hutchison CA. E-Cell: Software environment for whole-cell simulation. Bioinformatics, 1999,15(1):72–84. [doi: 10.1093/bioinformatics/15.1.72]
- [20] Broderick G, Ru'aini M, Chan E, Ellison MJ. A life-like virtual cell membrane using discrete automata. In Silico Biology, 2005,5(2):163–178.
- [21] Chiarugi D, Cuti M, Chiarugi D, Marangoni R, LoBrutto G. VICE: A virtual cell. Computational Methods in System Biology, 2005,3082:207–220. [doi: 10.1007/978-3-540-25974-9_16]
- [22] Ander M, Beltrao P, Ventura BD, Ferkinghoff-Borg J, Foglierini M, Kaplan A, Lemerle C, Tomas-Oliveira I Serrano L. SmartCell, a framework to simulate cellular processes that combines stochastic approximation with diffusion and localization: Analysis of simple networks. System Biology, 2004,1(1):129–138. [doi: 10.1049/sb:20045017]

附中文参考文献:

- [1] 赵明生,尚彤,孙冬泳等.电子细胞的研究现状与展望.电子学报,2001,29(12A):1740–1743.
- [3] 卢欣华,孙吉贵.Analog-Cell:一种新的电子细胞图形模型.电子学报,2007,35(1):49–53.
- [7] 李巍.虚拟细胞中化学反应定量计算的方法与算法的研究[硕士学位论文].哈尔滨:哈尔滨工业大学,2006.
- [8] 张恩慧.虚拟细胞本体半自动构建方法研究[硕士学位论文].哈尔滨:哈尔滨工业大学,2009.
- [10] 吴瑛,赫荣乔.RNA 剪接因子结构与功能研究进展.生物化学与生物物理进展,2003,30(4):503–508.



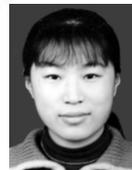
欧阳丹彤(1968—),女,吉林长春人,博士,教授,博士生导师,CCF 高级会员,主要研究领域为基于模型的诊断,自动推理.



韩霄松(1984—),男,博士生,主要研究领域为计算机模拟与建模,生物信息学.



王珏(1985—),男,硕士,主要研究领域为电子细胞模型,计算机模拟与建模.



卢欣华(1977—),女,博士,讲师,主要研究领域为电子细胞模型,计算机模拟与建模.