

# The Evolution and Regulatory Function of Transcription in Intron

Jun Hu, Jianhong Yang

School of Science and Information Engineering, Yunnan Agricultural University, Kunming  
Email: hjlyj8302@sina.com

Received: Dec. 17<sup>th</sup>, 2013; revised: Dec. 20<sup>th</sup>, 2013; accepted: Dec. 23<sup>rd</sup>, 2013

Copyright © 2013 Jun Hu, Jianhong Yang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. In accordance of the Creative Commons Attribution License all Copyrights © 2013 are reserved for Hans and the owner of the intellectual property Jun Hu, Jianhong Yang. All Copyright © 2013 are guarded by law and by Hans as a guardian.

**Abstract:** There are two opposing theories about intron originated and evolved. The introns-early theory believed that introns had already existed at the beginning of life. However, the introns-late theory believed that the introns were inserted into genes during the course of evolution. Some cis-elements in introns play a key role in transcriptional regulatory. The statistical analysis of transcriptional regulatory function showed that there are some potential patterns of transcriptional synergy between 5'-UTR and introns.

**Keywords:** Intron; Evolution; Transcriptional Regulatory; Statistical Analysis; Transcriptional Synergy

## 内含子的进化及其转录调控功能的统计分析

胡俊, 杨建红

云南农业大学基础与信息工程学院, 昆明  
Email: hjlyj8302@sina.com

收稿日期: 2013年12月17日; 修回日期: 2013年12月20日; 录用日期: 2013年12月23日

**摘要:** 关于内含子的起源和进化有内含子早现(introns-early)和内含子晚现(introns-late)两种假说。内含子早现理论认为内含子是基因原来就有的结构; 内含子晚现理论认为内含子是在进化过程中插入基因中的。除了提供多种剪切方式外, 内含子中还有多种顺式作用元件参与转录调控, 对酵母基因内含子转录调控功能的统计分析显示酵母基因上游与内含子之间可能存在的一些转录协同作用模式。

**关键词:** 内含子; 进化; 转录调控; 统计分析; 转录协同作用

### 1. 内含子的进化和存在的意义

#### 1.1. 内含子的起源和进化

内含子是基因中编码区之间插着的非编码 DNA 序列(见图 1)。内含子可以转录成 mRNA, 但翻译成蛋白质前被剪切掉<sup>[1]</sup>。它不出现在成熟的 mRNA 中。在原核生物基因组中也发现极少数的内含子, 它们可能起着某种调节作用<sup>[2]</sup>。而几乎所有的真核生物都有

内含子, 只是在低等真核生物中较少见, 酵母基因组 6000 个基因中仅包含 239 个内含子, 哺乳动物细胞中多数基因都包含 50 个以上的内含子<sup>[3]</sup>。内含子分为 GU-AG 内含子, AU-AC 内含子, I 类、II 类、III 类内含子, 孪生内含子, 前-tRNA 内含子, 古细菌内含子<sup>[3]</sup>。有关内含子起源和进化的争论主要围绕 GU-AG 内含子进行的。主要存在两种相对的假说: 一种是由 Walter Gillber 提出的内含子早现(introns-early)理论,

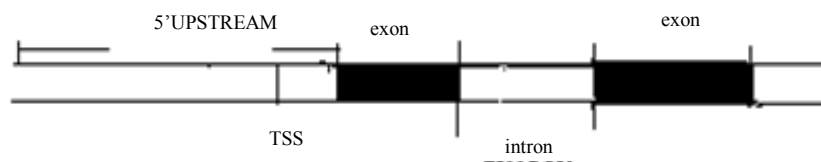


Figure 1. Simple illustration of introns  
图1. 内含子的简单图例

认为内含子和它位于其上的基因一样古老，在装配第一个这样的基因时，内含子就已存在<sup>[4,5]</sup>，在进化过程中原核生物细胞中的内含子丢失了，而真核生物细胞中的内含子部分保留了下来。另一种内含子晚现 (introns-late) 假说认为内含子不是基因原有的，而是在进化的某一过程中通过转座作用插入到连续基因中去的<sup>[6,7]</sup>。目前，吴加金等对肌动蛋白家族内含子插入位置分布的演化做了分析后得出内含子可能是在蛋白异化过程中获得的<sup>[8]</sup>。关于内含子起源和进化的争论已经持续了 20 多年，目前获得的相关证据都不能推翻任何一种假说<sup>[3]</sup>。

## 1.2. 内含子存在的意义和功能

内含子普遍存在于真核生物的基因中，它不编码蛋白质，但在长期的进化过程中为它又能保留下来，它存在的意义一直是生物进化中的问题。内含子的存在提供了多种剪切方式。基因转录后通过选择性剪切，相同的 DNA 序列可产生不同的蛋白质产物；可变剪切是基因表达的重要调控机制，由于可变剪切存在多样性，一个基因能在转录后通过 hnRNA 的剪切产生两个或多个蛋白质<sup>[9]</sup>。除此之外，内含子还具有其它的一些功能。例如，真核生物的内含子可编码蛋白质和酶<sup>[10]</sup>；内含子中还可包含其他 snRNA 或小核仁 RNA 的转录模板；已经证实，如果在开放阅读框 (ORF) 中保留了部分内含子，其中的终止密码子能够被识别并将剪接错误的 mRNA 降解，使之不能进入细胞质内<sup>[9]</sup>。

## 2. 内含子转录调控功能的统计分析

内含子的存在不仅提供多种剪切方式，而且内含子中有多种顺式作用元件参与转录调控<sup>[9]</sup>，其中的一些序列起着启动子 (promoter) 和增强子 (enhancer) 的作用，有的则具有抑制子 (repressor) 的功能<sup>[11-14]</sup>。这些结果大多数是基于实验分析获得的。张静等通过比较分

析的方法和期望频率法对酵母基因组中转录效率较高 (>30 mRNAs/hr) 和转录效率较低 ( $\leq 10$  mRNAs/hr) 的基因内含子序列特征进行统计分析<sup>[15,16]</sup>，结果表明一些寡核苷酸在这两组转录效率不同的基因内含子中出现的频率有显著差异，而且高效转录基因内含子具有有利于基因转录的序列特征。同时也提示高效转录基因内含子中含有较多潜在的转录因子作用位点。对这些潜在位点在内含子中的分布特征用 R-scan 统计分析法进行分析，进一步揭示出高效转录基因内含子中不仅含有较多潜在的转录因子结合位点，而且从这些位点的分布来看，在高效内含子组中的显著聚类明显多于低效内含子组。这样的分布特征同样有利于提高基因的转录效率。

从内含子在转录效率较高的酵母基因中的实际位置来看，由于多数基因的第一个外显子 (exon) 较短 (图 1)，因此内含子比较靠近基因的 5' 上游 (5' upstream)，有的内含子甚至位于 5' UTR 区。张静等<sup>[17]</sup>通过研究酵母内含子在基因序列中的分布对基因转录效率的影响后，发现高效转录基因内含子可能参与了转录调控。

真核生物转录调控的一个重要特征是基因受多个转录因子的联合调控，在联合调控过程中采用了协同性和协同作用的原则<sup>[18]</sup>，多数真核基因有两个或两个以上的调控位点<sup>[19]</sup>，只需启用少数调控蛋白就可以完成多种调控。对真核基因组上游转录激活协同性的研究表明在转录调控过程中转录因子与各 DNA 位点协同结合，结合方式可能是单个转录因子与多个 DNA 位点结合，也可能是几个转录因子一起与 DNA 位点协同结合。张昆林等<sup>[20]</sup>仍使用频率分析法抽提出高效转录基因上游区域和内含子序列中的转录调控元件，并定义一个位于上游，另一个位于内含子的两个寡核苷酸为寡核苷酸对，寡核苷酸对在一个特定基因上的尾-头距离为最近距离。对最近距离在 84 bp 内的寡核苷酸对进行统计分析得到了酵母基因上游与内含

子之间可能存在的一些转录协同作用模式。

### 3. 结语和展望

对内含子的起源和进化及内含子转录调控功能的统计分析有助于完善对内含子功能的认识；有助于提高对其它真核生物基因转录调控机制的认识；为其它真核生物基因组相应问题的研究提供理论分析方法的基础。

### 参考文献 (References)

- [1] 罗静初 (2002) 生物信息学概论. 北京大学出版社, 北京, 255.
- [2] 诸葛健, 王正祥 (2003) 现代英汉生物工程词典. 科学出版社, 北京.
- [3] 袁建刚, 周严, 强伯勤, 等, 译 (2002) 基因组. 科学出版社, 北京, 238.
- [4] Gilbert, W. (1978) Why genes in pieces? *Nature*, **271**, 501.
- [5] Darnell, J.E. and Doolittle, W.F. (1986) Speculations on the early course of evolution. *Proceedings of the National Academy of Sciences USA*, **83**, 1271-1275.
- [6] Cavalier-Smith, T. (1991) Intron phylogeny: A new hypothesis. *Trends in Genetics*, **7**, 145-148.
- [7] Rzhetsky, A., Ayala, F.J., Hsu, L.C., et al. (1997) Exon/intron structure of aldehyde dehydrogenase genes supports the "intronless" theory. *Proceedings of the National Academy of Sciences USA*, **94**, 6820-6825.
- [8] 吴加金, 刘涛, 贺福初, 吴晓霞 (1998) 动物肌动蛋白基因中内含子的来源及存在意义的探讨. *遗传学报*, **5**, 409-415.
- [9] 沈羽珩, 方德福 (1997) 真核基因表达调控. 高等教育出版社, 施普林格出版社, 北京, 39-47.
- [10] 刘次全, 白春礼, 张静 (1997) 结构分子生物学. 高等教育出版社, 北京, 159-163.
- [11] Katharina, H.S., Cox, T.C. and May, B.K. (1998) Identification and characterization of a conserved erythroid-specific enhancer-located in intron 8 of the Human 5-Aminolevulinate Synthase 2 Gene. *The Journal of Biological Chemistry*, **273**, 16798-16809.
- [12] Bhattacharyya, N. and Banerjee, D. (1999) Transcriptional regulatory sequences within the first intron of the chicken apolipoprotein AI(apoAI) gene. *Gene*, **234**, 371-380.
- [13] Brinste, R.L., Allen, J.M., Bbehlinger, R.R., et al. (1988) Introns increase transcriptional efficiency in transgenic mice. *Proceedings of the National Academy of Sciences USA*, **85**, 836-840.
- [14] Clement, J.Q. and Wilkinson, M.F. (2000) Rapid induction of nuclear transcripts and inhibition of intron decay in response to the polymerase II inhibitor DRB. *Journal of Molecular Biology*, **299**, 1179-1191.
- [15] 张静, 石秀凡 (2003) 酵母基因中转录正调控内含子序列特征的统计分析. *生物化学与生物物理进展*, **2**, 231-238.
- [16] Zhang, J., Hu, J., Shi, X.-F., Cao, H. and Liu, W.-B. (2003) Detection of potential positive regulatory motifs of transcription in yeast introns by comparative analysis of oligonucleotide frequencies. *Computational Biology and Chemistry*, **27**, 497-506.
- [17] 张静, 石秀凡, 杨恒芬 (2003) 酵母内含子在基因序列中的分布对基因转录效率的影响. *生物化学与生物物理进展*, **6**, 945-949.
- [18] Carey, M. and Smale, S.T. (2002) Transcriptional regulation in eukaryotes concepts strategies and techniques. Tsinghua University Press, Beijing.
- [19] 薛文, 王进, 黄启来, 郑伟娟, 华子春 (2002) 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*, **4**, 510-513.
- [20] 张昆林, 张静, 罗静初 (2005) 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*, **10**, 953-958.