

34-39

RNA, 剪接, 剪接调控

7965(7)

*(生物工程进展)*1996, Vol. 16, No. 5

RNA 剪接和剪接调控

桂建芳

(中国科学院水生生物研究所 武汉 430072)

Q52202

基因是遗传的功能单位, 蛋白是由基因表达的行使生理功能的主体。70 年代以前, 人们对基因和蛋白的了解主要是来自于对细菌研究的认识, 认为基因、mRNA 和蛋白之间是一种线性关系, 即基因上的脱氧核苷顺序全部转录成 mRNA 上的核苷顺序, mRNA 上的核苷顺序再全部翻译成蛋白上的氨基酸顺序。然而, 70 年代后期至今的一系列发现已彻底变革了传统的基因表达模式, 特别是近年来众多转录因子和剪接因子的鉴定和克隆以及它们的定位和互作研究已开始显露出真核生物基因调控理论的端倪。这里, 特就 RNA 剪接及其调控作一概述。

一、基因割裂现象的发现和 RNA 剪接研究领域的兴起

RNA 剪接研究起始于对基因割裂现象的发现。1977 年, 当美国麻省理工学院的 Phillip Sharp 和冷泉港实验室的 Richard Roberts 分别领导的两个研究小组研究真核细胞病毒腺病毒 (adenovirus) 基因组和其编码衣壳蛋白的 mRNA 之间的关系时, 他们创造了这一 mRNA 链和其互补基因组 DNA 链结合的杂种分子。当获得这一杂种分子的电镜照片时, 两个研究小组都能看出基因组的那一部分产生了成熟的 mRNA 分子, 但奇异的是 mRNA 并不是自始至终直线排列在其互补的 DNA 链上, 有些区段显示出巨大的未杂交的 DNA 环。当这些照片在 1977 年 6 月的冷泉港会议上展示后, 生物学家们立即认识到他们对基因结构的理解已经变革。会后不久, 很快在其它真核基因中也得到了广泛证实^[1-3]。

这一创造性的发现意味着真核细胞的基因

结构和 mRNA 合成要比细菌中复杂得多, 其基因在表达过程中发生了剪接 (splicing), 一些区段为基因中的内含子 (intron), 在剪接后丢失了, 一些区段为基因的外显子 (exon), 在剪接后重新拼接起来, 形成成熟的 mRNA。也就是说, 真核基因结构实质上是割裂 (split) 的, 即一个基因由多个外显子和内含子间断相接而成。这样, 一个新的称为 RNA 剪接的研究领域随着这一创造性的发现从此诞生了, 其发现的主要领导者 Phillip Sharp 和 Richard Roberts 也由此分享了 1993 年度的诺贝尔生理和医学奖^[4, 5]。

二、体外剪接反应和剪接体

RNA 剪接研究发展的关键步骤之一是其体外剪接反应 (in vitro splicing reaction) 系统的引入。在这种由可溶性核提取液组成的系统中, 加入放射性标记的 pre-mRNA 底物, 就可在体外分析和再现剪接反应的生化过程^[6]。

采用体外剪接反应, 很快发现 RNA 剪接在一复合体中进行, 且该复合体可通过甘油梯度沉降鉴定出来, 其沉降系数为 60S。这一复合体被命名为剪接体 (Spliceosome)^[7], 它与核糖体很相似, 含有底物 RNA 和大量稳定的 RNA 蛋白成分。

剪接体是复合的核糖核蛋白 (ribonucleo-protein) 颗粒, 它由 5 种核内小 RNA (即 U1, U2, U4, U5, 和 U6snRNAs) 和许多蛋白组成^[8]。

剪接体的组装是一复杂多步骤的过程。首先, 前体 RNA 上的 5' 剪接点、分支点和 3' 剪接点必须被剪接机器所识别, 在这识别过程中, 主

要决定子是 5' 剪接点与 U1 snRNA 之间的互补。接着, 几种蛋白结合于 3' 剪接点的附近, 并将 U2 snRNP 推向于分支点, 使 U2 snRNA 与分支点发生碱基配对。然后, 这种含 U1、U2 snRNP 和许多蛋白的前剪接体 (pre-spliceosome) 再与 U4、U5 和 U6 snRNP 的复合体结合, 形成最终可催化剪接反应的剪接体^[8]。就现今已了解的情况来看, 其组装过程中的 RNA 与 RNA 间的相互作用以及其催化步骤相当复杂^[9], 这里就不详述。

三、选择性剪接及其生物学意义

RNA 剪接研究中重大的发现之一是选择性剪接 (alternative splicing), 即在同一个基因中, 其剪接位点和形式可以改变, 从而导致一个基因能产生多个具有明显差异的蛋白产物。因而, 选择剪接在调控发育和分化期间的基因表达中起了重要作用^[8]。

粘连蛋白 (fibronectin) 基因是选择性剪接的典型例子之一。人的粘连蛋白基因由三种类型的外显子组成, 它们分别编码蛋白上的特定区段。由粘连蛋白基因转录的 pre-RNA 可被有选择地剪接成编码 20 种不同蛋白的 mRNA。这些 mRNA 翻译的蛋白产物具有不同的功能, 且与其剪接的结构变化有关。特别是在肝细胞中, 其 mRNA 与其它细胞的明显不同, 有两个外显子被剪接掉了, 因而由肝细胞分泌的粘连蛋白更易在血流中循环, 因为它比其它类型的粘连蛋白拥有较少的被细胞表面受体识别的区段。由其它细胞分泌的粘连蛋白更多地存在于固体组织的细胞内基质中找到。此外, 由于另三个外显子剪接变化所产生的 5 种不同的 mRNA 所翻译的蛋白产物也表现出明显不同的生理作用, 依赖于其剪接的差异, 有些粘连蛋白可与淋巴细胞上的受体结合, 有些则不能。因此, 通过有选择性的剪接, 同一个基因能被用作不同的生理功能^[3]。

抗体的形成也是 RNA 有选择性剪接的一个著名例子。研究表明, B 淋巴细胞分泌免疫球蛋白 IgM 和 IgD 的开关是由 RNA 剪接导致

的。在 B 淋巴细胞中仅生成一种免疫球蛋白的 per-mRNA, 它包含有编码结合抗原的可变区和编码 IgM 恒定区以及 IgD 恒定区的外显子, 且这些外显子被不同的内含子所隔开。IgM 和 IgD 的产生依赖于两种不同的选择剪接机制。如果编码 IgD 恒定区的外显子被剪接了, 细胞将分泌 IgM; 如果编码 IgM 恒定区的外显子被剪接了, 则生成 IgD。在两种剪接机制都启动的过渡期间, IgM 和 IgD 两类免疫球蛋白都可产生^[10]。

果蝇的性别决定也与 RNA 的有选择剪接紧密相关。一般来说, 果蝇的性别表型是由常染色体和性染色体的比例决定的。当其比例为 1 (即二倍体细胞中有两个 X 染色体) 时, 胚胎发育为雌性个体; 当其比例为 0.5 时, 该胚胎发育为雄性。调控这一过程的关键基因之一是一种称为转化子 (transformer, 简称 Tra) 的基因。这一基因是决定雌性的必需品, 它的丧失将导致雄蝇形成。在幼虫期, 该基因的转录产物或者加工成雌性和雄性都共有的普通 mRNA, 或者加工成雌性个体特异的 mRNA。仅仅雌性个体中含有这一有选择剪接的信使, 且仅仅是它才具有功能作用。发现于雌性和雄性中的共有普通 mRNA 在第二个外显子的前面部分含有一个终止密码, 由这一 mRNA 产生的小蛋白没有功能作用, 因而与性别决定无关。然而, 在特异于雌性的信使中, 这一终止密码位于内含子之中而在 mRNA 形成期间被剪接掉了, 因而没有干扰信使的翻译。也就是说, 雌性的转录产物才是这一基因的功能产物。有趣的是, 由类似这些调控性别的基因编码的蛋白产物大多是剪接因子 SR-蛋白超级家簇的成员^[11, 12], 因下文还将提到, 这里就不详述。

四、SR 蛋白和 SR 蛋白相关多肽家簇

RNA 剪接是如何进行的呢? 其调控机制怎样呢? 最近, 一组含有 SR (S-丝氨酸, R-精氨酸) 二肽重复序列区段的剪接因子 SR 蛋白的鉴定和克隆以及它们在细胞内的定位已燃起了这一研究的热点。SR 重复区段最先是在遗传

鉴定的果蝇剪接调控因子 *su(W^o)*、*Tra* 和 *Tra-2* 中描述的,后来在与 U1 snRNP 结合的 70 kD 的蛋白以及由生化鉴定的剪接因子 SF2/ASF 和 SC₃₅ 都发现了这一 SR 重复区段。据统计,现已鉴定出 20 多种 SR 蛋白和 SR 蛋白相关的蛋白(表 1)。SR 蛋白有 5 个共同点:即(1)都含有能被单克隆抗体 mAb104 识别的磷酸抗原决定部位,(2)能用盐两步沉淀法共纯化,(3)其

结构基本类似,都至少在 N-端含有 1 个 RNA 识别区段、在 C-端含有一个 SR 重复区段(4)它们在 SDS-PAGE 胶上的大小在进化上是保守的,(5)都有补偿 RNA 剪接的作用。SR 蛋白相关蛋白也含有 SR 重复区段,是具有多种作用的剪接调节因子,与 SR 蛋白一起被称为富含 SR 二肽的剪接因子超级家簇^[13-15]。

表 1 已鉴定的含有 SR 重复区段的 pre-mRNA 剪接因子超级家簇的成员
(引自 Fu, 1993[15])

第一类:SR 蛋白	第二类:SR 蛋白相关多肽
来源于人的 SR 蛋白	来源于人的 SR 相关多肽
SRp20	U1 70 k
SRp30a/SF2/ASF	U2 AF65
SRp30b/SC35	U2 AF35
SRp30c	HCCL1 ~ U2AF 65
SRp40	HRH1 ~ Pep22
SRp55	Clk-1
SRp75	Clk-2-Clk-1
p54	Clk-3-Clk-1
9G8	
来源于其它生物的 SR 蛋白	来源于其它生物的 SR 相关多肽
X16(小鼠) = SRp20	<i>Tra</i> (果蝇)
PR264(鸡) = SC35	<i>Tra-2</i> (果蝇)
HRS(大鼠) = SRp40	<i>su</i> (又称 SWAP,果蝇)
B52(果蝇) = SRp55	xU1 70 k(瓜瓣) = U1 70 k
RBP-1(果蝇) = SRp20	dU1 70 k(果蝇) = U1 70 k
SR-1(植物) ~ SF2/ASF	mU2AF65(小鼠) = U2AF65
Np13(酵母)	U2afbp-rs1(小鼠) ~ U2AF35
	U2afbp-rs2(小鼠) ~ U2AF35
	dU2AF50(果蝇) = U2AF65
	dU2AF38(果蝇) = U2AF65
	Prp2(酵母) ~ U2AF65
	YCL11c(酵母) ~ U2AF65
	MUD2(酵母) ~ U2AF65
	mClk-1 \ Sty(小鼠) = Clk-1

注: =, 同源基因; ~, 结构相似基因。

五、SR 蛋白的活性和作用

SR 蛋白是剪接的基本要素。在缺少 SR 蛋白的细胞抽提液例如 S100 中,不能检测到特异的剪接复合体^[14,15],表明 SR 蛋白参与了剪接体的早期组装。进而在无剪接活性的细胞抽提液中,加入 SR 蛋白可恢复剪接活

性^[14,16,17];同时用 SR 蛋白的抗体如抗-SC35^[18]或抗-9G8^[19]处理核抽提液可抑制剪接体组装的早期步骤。这些早期观察与最近有关 SR 蛋白相互作用的发现^[20-22]意味着 SR 蛋白可能是核抽提液中普通蛋白各复合体的部分成员。

然而,最近的进一步研究显示,SR-蛋白的

每一个成员对 mRNA 前体表现有不同的底物特异性范围。例如,人免疫缺陷病毒(HIV)tat pre-mRNA 的剪接似乎是依赖于 SF2/ASF,而不是 SC35^[23];在另一个例子中,是 SC35 而不是 SF2/ASF 能够代替 9G8 抗原恢复缺失了 9G8 核抽提液的剪接活性^[19]。不同的 SR 蛋白也以不同的效率和不同的机制影响剪接位置的选择^[14,24]。此外,不同的剪接强化成分可通过与不同的 SR 蛋白成员相互作用而行使功能。例如,果蝇的双性(*dsx*) mRNA 前体的雌性特异性剪接是由两个均含有 SR 区段的转化子 *tra* 和 *tra-2* 将 SR-蛋白调遣到雌性特异性外显子中的调节位置,并因此激活雌性特异性剪接而完成的^[25];牛生长激素(bGH)mRNA 前体最后一个内含子的无效剪接是因为它有一个微弱的 5'剪接位点,其剪接与否取决于 SF2/ASF 是否特异性地结合于其内含子下游外显子中一富含嘌呤的区段(命名为剪接强化子)上^[26]。这也是说,选择性剪接位点的使用至少部分地是由 SR-蛋白结合于剪接位置介导的,或者是由特异性选择剪接调节子将 SR-蛋白调遣到剪接位置介导的^[27]。与这些发现一致的是, B52/SRp55 已证实对果蝇的发育是必需的^[28,29],然而,在这一剪接因子缺失的幼虫中,仍旧能剪接 5 种内源性的 pre-mRNA^[26],表明 B52/SRp55 对体内和所有剪接不是必需的。因此,SR 蛋白的作用可能不是重复的,有些可能还是独特的^[29,30]。

六、剪接因子的定位及剪接发生的位置

剪接因子的亚核定位已通过抗原和抗体特异性结合和第二抗体荧光标记的方法在光镜和电镜上进行了广泛的观察研究,其最引人注目的特征是发现这些剪接因子在核内聚集成 20-50 个核斑(nuclear speckles)^[31]。这一特征最初是用于特异于 U2, U4—U6 和 U5snRNPs 的抗体进行亚核定位时观察到的,同时也观察到这些 snRNPs 也集中分布于 1—5 个称为螺旋体(coiledbodies)的核内焦点(foci)中。随后,用 SR-蛋白的特异性抗体与 snRNPs 的特异性抗

体进行的共定位研究进一步表明非 snRNPs 的剪接蛋白也集中分布在这些核斑区域。有趣的是,存在于活性剪接体中的这些 SR-蛋白和其 snRNPs 可能与一种特殊的核基质依附在一起。这种核基质在核经 DNA 酶 I 和高盐连续抽提几乎移去了所有的染色质蛋白和 DNA 之后,其结构仍然保持完整,但对 RNA 酶敏感。研究已经表明,剪接体与核基质是相联系的,因为从这核基质抽提物中可获得脉冲标记的内源性 RNA 前体^[3]。最近,有人还从核基质中分离出剪接因子 SR 蛋白,并发现这些 SR 蛋白还特异地与外显子相结合。

高分辨的原位杂交方法和脉冲标记研究已将新转录的 RNA 定位在核斑区域以及从基因直到核孔的核通道中^[32,33],在这通道中,其 RNA 前体离其基因的所在位置愈远,其内含子顺序与其外显子顺序相比的相对浓度愈低。这些研究者还发现,核斑实际上由核心和周边两部分组成,核心主要是剪接成分,周边由新转录的 RNA 围绕,并与活性基因紧密相连。他们已检验了 10 个活性基因的定位,发现 7 个基因与其 mRNA 一起位于这些周边位置。因而,这批研究人员认为,一些基因的转录和剪接发生在由剪接因子聚集核斑核心的周边区域。

核斑经电镜观察分析包含有两类不同的结构,即染色质间颗粒簇(interchromatin granule clusters)和染色质周边纤维(perichromatin fibrils)。美国冷泉港实验室的 Spector 教授认为,染色质周边纤维是转录位置,当基因转录开始时,剪接因子从其染色质间颗粒簇的贮存位置调遣到染色质周边纤维转录位置,剪接就发生在这一位置或其附近。也就是说,转录和剪接是紧密协调的。最近,Spector 实验室报道了一个十分有趣的发现;当一个含有内含子的基因被转移到培养细胞时,剪接因子被定位到基因被转录位置的概率是 90%;而当一个没有内含子的基因被转移到细胞时,剪接因子不被簇拥到转录位置的概率是 85%。因为含有内含子的 RNA 必须被剪接,剪接因子才被调遣出来,可见在细胞核内涉及剪接的 RNA 及其因子存

在有一种十分有序的组织结构^[34]。

七、特异于 SR 蛋白激酶的鉴定及其调控作用

从对剪接因子聚集的核斑点进行研究中观察到的一个最有趣的现象是其在细胞周期中发生组装和去组装,即它们在细胞分裂间期,成典型的斑点状结构;随着细胞进入有丝分裂,斑点逐渐分散,到有丝分裂中期时,斑点消失;当分裂完成后,斑点又在两个子细胞中重新出现。这一现象表明,剪接因子聚集的核内斑点与那些在细胞周期中发生了组装和去组装的细胞器如染色体和核膜类似,可能也存在类似的调控机制。由此出发,终于鉴定,纯化并克隆出了一种新的特异于 SR-蛋白的激酶-SRPK1^[35-36]。关于 SRPK1 的发现经过,因已有一篇综述专门介绍^[37],这里就不再重复。

SRPK1 的发现既揭示了 SRPK1 特异性地催化 SR 蛋白磷酸化在调节剪接因子在细胞周期中定位的生理意义,又解释了 SR 蛋白的磷酸化和去磷酸化在调节 RNA 剪接和剪接位置选择中的功能作用^[38-43]。最近,另一个以 SR 蛋白为磷酸化底物称为 CLK/STY 的激酶也被鉴定出来了^[44],与 SRPK1 相比,这一激酶还含有类似于 SR 的区段,该区段对于识别特异底物似乎是重要的^[44]。现有的这些结果表明,细胞中可能存在多种 SR 蛋白和其它剪接因子的激酶,它们对不同的 SR 蛋白具有不同的底物特异性和亲和性,且其活性的调节可能还具有细胞和组织的特异性。这些激酶的鉴定正在将剪接调节和信号传导通道连接起来^[15]。

参考文献

- [1] Berget, S. M., Moore, C. & Sharp, P. A., Proc. Natl. Acad. Sci. USA, 1977, 74:3171-3175
- [2] Chow, L. T., Gelinas, R. E., Broker, T. R. & Roberts, R. J., Cell, 1977, 12:1-8
- [3] Sharp, P. A., Cell, 1994, 77:805-815
- [4] Travis, J., Science, 1993, 262:506
- [5] Cohen, J., Science, 1995, 268:1706-1708

- [6] Padgett, R. A., Hardy, S. F. & Sharp, P. A., Proc. Natl. Acad. Sci. USA, 1983, 80:5230-5234
- [7] Brody, E. & Abelson, J., Science, 1985, 228:963-967
- [8] Horowitz, D. S., & Krainer, A. R., Trends in Genetics, 1994, 10:100-106
- [9] Nilsen, T. W., Cell, 1994, 78:1-4
- [10] Maki, R., Roeder, W., Traunecker, A., Sidman, C., Wabl, M., Rasjke, W. & Tontgawa, S., Cell, 1981, 24:353-365
- [11] Roth, M. B., Zahler, A. M., & Stolk, J. A., J. Cell Biol. 1991, 115:587-596
- [12] Tian, M. and Maniatis, T., Science, 1992, 256:237-240
- [13] Zahler, A. M., Lane, W. S., Staik, J. A., & Roth, M. B., Genes & Dev., 1992, 6:837-847
- [14] Zahler, A. M., Neugebauer, K. M., Lane, W. S., & Roth, M. B., Science, 1993, 260:219-222
- [15] Fu, X. T., RNA, 1995, 1:663-680
- [16] Krainer, A. R., Conway, G. C., & Kozak, D., Cell, 1990, 62:35-42
- [17] Fu, X. D., & Maniatis, T., Science, 1992, 256:535-538
- [18] Fu, X. D., & Maniatis, T., Nature, 1990, 343:437-441
- [19] Cavaloc, Y., Popielarz, M., Fuchs, J. P., Gattoni, R. & Stevenin, J., EMBO J., 1994, 13:2639-2649
- [20] Wu, J., & Maniatis, T., Cell, 1993, 75:1061-1070
- [21] Amrein, H., Hedley, M. & Maniatis, T. Cell, 1994, 76:735-746
- [22] Kohtz, J. D., Jamison, S. F. et al. Nature, 1994, 368:119-124
- [23] Fu, X. D. Nature, 1993, 365:82-85
- [24] Zahler, A. M. & Roth, M. B. Proc. Natl. Acad. Sci. USA, 1995, 92:2642-2626
- [25] Tian, M., Maniatis, T. Cell, 1993, 74:105-114
- [26] Sun, Q., Meyeda, A., Hampson, et al. Genes & Dev. 1993, 7:2598-2608,
- [27] Lavigueur, A., La Branche H., et al. Genes & Dev., 1993, 7:2405-2417
- [28] Ring H. Z., & Lis J. T., Mol. Cell Biol, 1994, 14:7499-7506

- [29] Peng X. & Mouut, S. M., *Mol. Cell Biol.* 1995, 15:
- [30] Rana, S. Singh J., Zahler A. M., et al, *Mol. Cell Biol.*, 1995, 15:4898-4907
- [31] Spector, D. L., *Annu. Rev. Cell Biol.* 1993, 9:265-315
- [32] Carter K. C., Bowman, D. et al. *Science*, 1993, 259: 1330-1334
- [33] Xing Y., Johnson, C. V., et al. *Science*, 1993, 259: 1326-1330
- [34] Baskin Y., *Science*, 1995, 268:1564-1565
- [35] Gui J. F., Lane, W. S. et al. *Nature*, 1994a, 369:678-682
- [36] Gui, J. F., Tronchere, H., et al *Proc Natl. Acad Sci USA.*, 1994b, 91:10824-10828
- [37] 桂建芳, 细胞生物学杂志, 1996, 18(2):49-54
- [38] Tazi J., Kornstadt, U., Rossi, F., et al. *Nature* 1993, 363:283-286
- [39] Mermod J. E., Cohen, P. T. W., et al *EMBO J.*, 1994, 13:5679-5688
- [40] Cardinal E., Cohen, P., et al. *FEBS Lett.* 1994, 352: 276-280
- [41] Spector, D. L., *Nature*, 1994, 369:604
- [42] Wansink, D. G., van Driel, R. et al. *Molecular Biology Reports*, 1995, 20:45-55
- [43] Hendzel, M. J. & Bazett-Jones, D. P., *Chromosoma*, 1995, 103:509-516
- [44] Colwill, K., Pawson, T., et al. *EMBO J.*, 1995, 14: 6011-6023

RNA Splicing and Regulation of Splicing

Gui Jianfang

(Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan 430072)

Abstract

The discovery of the split gene resulted in the advances of RNA splicing research field. Significant progress has been made recently in identifying and cloning factors involvde in RNA splicing. Genetic and biochemical evidence suggests that the SR proteins might play important roles in the regulation of RNA splicing. The identification of SR protein-specific kinases may ultimately connect splicing to other biological processes, such as signal transduction pathways, transcriptional regulation, and the cell cycle control. Gene regulation is central to all biological phenomena, and RNA splicing is important in the regulation of genes, so a lot of molecular biologists are further attempting to elucidate and understand the mechanism and the cell biology of RNA splicing and regulation.

Key Words RNA splicing, kinases, cell cycle control, SR proteins, gene regulation