

户的检索词检索 GenBank, 并把检出的序列记录通过 e-mail 返回给用户; BLAST server 进行序列对 GenBank 的类似性检索. ②FTP 实现文件在用户和服务器之间的传送. 据此用户可免费获得一个完整的最新版本 GenBank 或最近一周释放的 GenBank 新记录 (GenBank update). ③Gopher、WAIS 和 WWW 服务器. 它们与用户计算机上安装的 client 软件组成相应的 server-client 软件系统, 进行数据库检索或者实现文件传送. Internet 的电子公告牌, 例如生物学家感兴趣的 BIOSCI 公告牌为分子生物学家进行专题学术讨论、寻求别人帮助以及与数据库人员的交流提供了极大的方便.

致谢 本项工作得到本院计算中心周士波高级工程师的帮助. 作者的 e-mail 地址是: zhousb@bepc2. ihp. ac. cn.

参 考 文 献

- 1 王槐春编著. 蛋白质与核酸序列分析基础. 北京: 人民军医出版社, 1994: 1
- 2 Henikoff S. Trends Biochem Sci, 1993; 18: 267
- 3 Gribskov M, Devereux J. Sequence Analysis Primer. New York: Stockton Press, 1991: 1
- 4 Parker M. Trends Biochem Sci, 1993; 18: 485
- 5 Apple R D, Bairoch A, Hochstrasser D F. Trends Biochem Sci, 1994; 19: 258
- 6 Bleasby A, Griffiths P, Hines D *et al.* Trends Biochem Sci, 1993; 18: 310

Electronic Communication and Sequence Analysis. Wang Huaichun (*Institute of Medical Information, Beijing 100850, China*).

Abstract Electronic communication through Internet has provided on-line computer users an effective way for information exchange. For molecular biologists, they can not only use the electronic mail (e-mail) system to send and receive message as fast as fax, but also more importantly, can access large numbers of new molecular biology databases and softwares. They can perform various sorts of sequence analysis tasks, including homology search against databases, gene coding region identification and protein secondary structure analysis, etc. A database, such as GeneBank, can be accessed through the following ways: (1) e-mail file servers, (2) file transfer protocol (FTP), and (3) Gopher, wide area information server (WAIS) or world-wide web (WWW). The BIOSCI bulletin board, specially for the interest of molecular biologists, gives great convenience for molecular biologists to hold discussions of scientific topics, ask for others' help and communicate with database staffs.

Key words electronic communication, Internet, sequence analysis, database

内含子对提高转基因动物基因表达效率的影响

卢一凡 邓继先 肖成祖 马清钧

(军事医学科学院生物工程研究所, 北京 100071)

摘要 综述了内含子在转基因动物基因表达中的作用, 对内源性内含子和异源内含子对转基因表达的影响进行分析, 讨论了内含子提高转基因动物基因表达效率的三种可能机制, 并指出在表达载体构建中应考虑到内含子的重要性.

关键词 内含子, 转基因动物, 基因表达

转基因动物在发展的短短10多年中, 几乎已经涉及到医学研究的所有前沿课题. 无论在发病机理、疾病模型建立、药理学抗药性研究, 还是在基因治疗、药用蛋白质生产方面都已展示出良好前景, 已取得了许多令人瞩目的成就. 但由于生物机体的复杂性以及对发育机制所知甚少, 许多基础理论问题仍在探索之中. 在转基因动物的建立过程中, 所遇到的一个主要问题是导入的外源基因多数表达水平过低. 为解决这一难题各国学者一直在进行不懈的努力. 尽管影响外源基因表达水平的因素很多, 但现有的文献表明仍需要在元件构建上深入研究. 从理论上讲, 将在细胞特异表达的任一基因调控区与任一编码蛋白质基因融合, 在动物细胞中表达都应是不可能的. 这一思想也已被广泛地应用于基因表达与调控研究之中. 然而在转基因动物中, 并非所有构建的表达载体都能正常表达外源基因, 遇到的两个最为普遍的问题是表达方式不适合和不能达到足够的表达水平. Brinster 等^[1]发现内含子在提高外源基因表达上起重要作用. 以后这方面的研究逐渐增多, 它对外源基因在转基因动物中表达提供了许多有益的帮助.

1 内含子在细胞表达系统中研究概况

转基因动物的建立极为复杂繁琐, 花费的时间长、投资大. 而将所构建的表达载体预先在哺乳动物细胞中进行表达则是验证元件正确与否的有效手段. 因此有关内含子对表达效率的影响已在细胞表达系统做了众多研究. Hamer 等^[2]研究了所构建的一系列 SV₄₀病毒载体, 这些载体上含有多个位点可与鼠珠蛋白基因相连, 当该基因插入其中一个克隆位点时, 就可以产生稳定的 mRNA. 当用突变方法使 SV₄₀缺失最后一个内含子时, 则不能产生稳定的转录产物, 但通过添加异源内含子则能使之得以改善. 这表明内含子的剪接对 mRNA 在细胞内的积累是必要的. 因此80年代初的 cDNA

表达载体除启动子和 polyA 序列外, 通常含有异源内含子^[3]. 但随后又发现, 当删除内含子时并不总是对 mRNA 的产生有所影响. 例如一些病毒基因, 包括 E₁A 蛋白^[4]、多瘤病毒大 T 抗原基因^[5]都不含内含子却同样具有功能. 同时菜豆基因^[6]、酵母肌动蛋白基因^[7]和鸡胸苷激酶基因^[8]等不管有无内含子都能产生相同的 mRNA. 这些试验证实, 80年代中后期, cDNA 表达载体尽管没有内含子, 但表达水平仍可达到一定高度. 这表明仅在一些特殊情况下内含子对 mRNA 的积累是必要的.

将编码基因受控于调控序列之下, 并在转基因动物体内各种类型细胞中进行表达的方法已用于许多基础研究和应用研究. 由于 cDNA 易于获取, 并与基因组序列相比, cDNA 并非很大而又便于操作, 因而深受人们青睐. 许多 cDNA 表达载体已不断构建出来. 尽管很多在哺乳动物细胞表达系统试验良好, 但绝大多数在转基因鼠上试验却不易表达或者表达水平过低. 而内含子却可以提高转基因动物基因表达的效率^[1,9]. 因而内含子是否为转基因动物基因表达所必须存在的元件之一成为人们关注的焦点.

2 内源性内含子的影响

在研究内含子对转基因动物基因表达效率的影响时, 首先考虑到的问题是内源性内含子的作用. 因为现有的一些观点认为, 内含子含有某些调控元件, 特别是增强子. Brinster 等^[1]、Palmiter 等^[9]研究了内源性内含子对表达效率的影响. 他们将鼠的生长激素 (rGH) 基因组序列 (含有 A、B、C、D 4 个内含子) 与鼠金属硫蛋白基因 (rMT-1) 启动子连接, 注入鼠受精卵中, 在出生的转基因鼠肝脏中检测到的 mRNA 水平是 cDNA 编码基因表达的10倍. 当更换不同的启动子时, 含有内含子的构件, 其转录效率明显高于 cDNA 构件. 同时在产生转基因阳性鼠中, 检测到的表达个体数目也大为

增加,有的组别可高达80%^[1]. 由于这些试验所使用的样本数比较大,因而结果颇具代表性. 这一令人兴奋的结果又促使他们进一步考虑在 rGH 表达过程中, 4个内含子对提高表达效率是否具有相同的作用, 试验结果表明, 当只含有内含子 A 时, 其表达效率为 rGH 基因组序列表达 mRNA 水平的 $\frac{1}{2}$ (为 cDNA 构件表达水平的5倍). 而含有内含子 C、D 的构件则对表达略有提高 (为 cDNA 构件表达的3倍). 在含有 A、C、D 3个内含子构件中, 表达水平与只含有内含子 A 的表达水平相近. 令人惊奇的是内含子 B 有抑制内含子 A 的作用, 因为在含有 A、B 内含子的构件中, 其表达的 mRNA 水平比 cDNA 构件还要低得多. 内含子 B 的这一负向影响可通过内含子 C、D 的加入而得以克服^[9]. 这一结果说明内含子 A 对表达水平有较大影响. 然而 Whitelaw 等^[10]在羊的 β -乳球蛋白基因 (BLG) 中保留第一内含子, 将该构件注入鼠受精卵中, 11只阳性鼠仅有3只表达, 与 BLG 基因组构件相比, 表达水平也低许多. 当将 BLG 基因启动子与保留有第一内含子的人 α_1 -抗胰蛋白酶 (α_1 -AT) 基因相连接, 在获得的 8只阳性鼠中也仅有1只检测到 α_1 -AT 的 mRNA. 相比之下, 采用 α_1 -AT 基因组序列, 则 15只转基因鼠中, 10只在乳腺定位表达 α_1 -AT 蛋白质, 表达水平高达 (7~8) g/L. 这表明第一内含子的作用尚需要进一步证实研究. 但采用基因组序列能提高表达水平已有众多报道, 已为人们所接受^[11,12].

3 异源内含子的影响

内源性内含子虽然对提高转基因动物基因表达有较大影响, 但由于内含子在 mRNA 剪接上的特殊作用, 以及在研究外源基因时, 特别在采用 cDNA 构件时, 研究异源内含子对转基因表达的影响则更具有意义. Palmiter 等^[9]试验了将几种异源基因内含子分别置于已删除内含子的 rGH 基因的 5' 端、3' 端以及中间位置, 探讨其对表达水平的影响. 将鼠的胰岛素基因 II (rIns-II) 内含子 A 置于 rGH cDNA 5' 端, 与

cDNA 构件相比可以提高 mRNA 转录近7倍. 另一个试验中将 rIns-II 基因内含子 A 置于人的神经生长因子 (hNGF) cDNA 5' 端, 提高表达约75倍. 需指出的是, 将内含子放入 5' 端时应仔细处理好内含子与第一外显子之间的序列, 防止改变阅读框架而阻止 rGH 的翻译.

尽管将异源内含子放于 cDNA 5' 端可增强转基因的表达, 但当将两个相同的异源内含子相排列共同置于 5' 端时, 却几乎不表现出内含子的作用, 表达水平与 cDNA 表达相近. 将几种异源内含子置于 cDNA 3' 端对表达也无影响, 置于下游也未能援救表达. 这些结果证实, 插入内含子的位置对表达是相当重要的.

上述试验采用的是 MT 启动子, 改用其它启动子, 如将鼠的弹性蛋白酶 I 启动子 (rE-I) 分别与 rGH 和 hAF (人血管形成因子) 基因连接, 将 rIns-II 内含子置于 hAF 5' 端, 结果提高表达81倍, 而对 rGH 却无影响. 这进一步说明提高表达水平与所用的组织特异性启动子和异源内含子的位置有密切关系. 另一些报告也都表明, 在 cDNA 中添加内含子可使转基因阳性鼠中获得表达的个体数目大为增加^[1,10,13].

4 内含子作用机制

内含子可能是通过几种不同的机制来提高转基因动物基因表达效率的. 一是某些内含子包含有增强子或其它顺式调控元件, 它们同某些蛋白质结合影响其转录的起始和延伸^[14]; 二是内含子的剪接增加了 mRNA 在核内的稳定性, 导致在细胞质中积累更多的成熟 mRNA^[15]; 另一种可能性是内含子含有这样一些序列, 即能开放染色体的功能域, 也许通过影响核质成分、位置等提高转基因的表达^[16].

目前这些机制尚不甚明了. 有功能的内含子位置依赖性表明其对剪接能起一定控制作用, 或在其内部存在有极强的增强子元件^[17], 它们对决定表达都极为重要. 弱增强子或位置依赖性顺式作用元件也可能包含于内含子内^[18,19]. 这些内含子天然地靠近于启动子位

置. rGH 和 rIns-II 基因内含子的位置依赖性也是由于这些机制^[9]. 已有报道指出 rGH 内含子含有 DNase I 超敏感位点, 它是甲状腺素受体结合位点^[20], 这证实了有效内含子必须结合一些转录因子. 目前关于这一点的普遍看法是, 启动子附近的序列已经在以这样一种方式进化, 即有利于转录因子的聚集, 这些序列明显含有转录因子结合位点^[9,10]. 而 DNA 序列对有关重要启动子元件定位于核小体有着精巧的影响. 蛋白质与 DNA 的相互作用也能改变核小体所处位置. 当把构建的基因连接到 5' 未翻译区时, 则这些序列能以最大可能性去使之适合于功能区构象. 进一步说, 当第一个内含子靠近转录起始点时, 启动子发挥功能的序列可以扩展或影响到内含子, 因而这一内含子具有相当的重要性^[9].

Brinster 等^[1]推测在 MT 启动子与 rGH 连接的基因构件中, 内含子也许含有重要调控元件. 在人 β -珠蛋白基因研究中, 发现在内含子 B/外显子 3 以及结构基因 3' 端均有重要增强子元件表现^[21]. 但关于是否内含子内存在增强子尚存在争议. 针对 rGH, 有人指出其内含子并非含有增强子. 第一, 采用 MT 启动子与 rGH 基因相连接, 已使 rGH 基因的表达建立于适合的组织特异性上; 第二, rGH 基因在肝细胞和胰腺细胞中并没有正常表达, 由于在细胞表达系统内含子具有一定的正向作用, 因而似乎其中不可能含有组织特异性增强子; 第三, 如果增强子样元件存在于内含子的话, 则可以预料内含子应对转染细胞后的表达有所影响. 但某些细胞系试验结果并不支持这一点^[13]; 第四, 当将内含子置启动子上游^[13]或置于 cDNA 下游并不增强表达活性^[9]. Brinster^[1]指出, 也许在内含子内存在另外一种促进转录的机制, 即存在一种独立的、普遍性的增强子.

真核细胞的内含子剪接能以组织特异性、发育特殊阶段或特定性别等方式进行^[19]. 但这种剪接作用是否对转基因动物基因表达效率有重要影响, 为证实这一点, Choi 等^[13]构建了一个内含子拼接体 (即在内含子内插入一段序

列), 将其置于 CAT 基因 5' 端及中间位置, 使它们受控于人组蛋白基因启动子. 发现不仅这些构件比对照在阳性鼠中表达 CAT 基因的个体数目大大增加, 而且发现几乎在检测的所有组织中都获得了表达, 与 cDNA 构件相比, 每种组织提高表达 5~250 倍. 这进一步说明内含子的剪接作用能较大地增加细胞质的 mRNA 水平. 一些研究表明, 这主要表现在内含子对 mRNA 的加工和其稳定性有重要影响^[22]. 核 RNA 前体的稳定性极差, 这一前体与“基因的拼接体”相连后将导致核稳定性增强或更为有效地将 mRNA 运输到细胞质中. 这表明内含子对转录的速度并无大的影响, 但对 mRNA 的积累起关键作用. 另一方面, Brinster 等^[1]指出, mRNA 的丰富度是与转录的相对速度成正比的, 在对转基因鼠的分析中, 内含子主要影响是转录而不是 RNA 的加工^[2,22]. Choi 等^[13]发现在转基因鼠中, 内含子拼接体的剪接增加 CAT 的活性, 指出部分原因是多聚腺苷酸化的作用以及运输到细胞质速率造成的. 这在细胞表达系统已被证实. 但在转基因鼠, 对核进行分离时由于大量 RNase 释放出来, 使其评价有一定困难.

通过在有关位置插入异源内含子来提高表达效率的试验中, 发现内含子的许多作用是通过改变转录因子的聚集或核小体的相位得以发挥的^[9]. 尽管内含子的作用早已在细胞表达系统做了深入研究^[18,19,22,23], 但问题是, 内含子对转基因动物基因表达有所影响, 却不影响转染的细胞. 例如内含子对牛 GH 基因在细胞表达系统就并非重要^[24]. 在细胞培养中, 选择作用主要是针对已建立的稳定的克隆, 而在转基因动物中却无法进行. 有无内含子在转基因动物基因表达上的明显差别清楚地表明, 位置独立性 (即不依赖于整合位点) 转基因表达依赖于除 cDNA 以外的其它序列, 它可以有助于建立开放染色体功能域^[25], 也许这些调控元件、增强子元件和相关序列都包含在内含子内, 它改变核小体与 DNA 相互作用, 并决定是否功能性转录复合物能聚集在启动子上. 换句话说, 内

含子所含有的序列对于有关重要启动子元件所处的核质状态相当重要。内含子的位置是达到这一状态的一个必要条件, 这与具有的增强子样特性有所区别。对于为何内含子对转染细胞后的表达几乎没有影响的一个可能的解释是, 在发育过程中核质所处的状态的差别, 即在培养细胞中核质所处的状态不如转基因动物体内严格^[1]。

5 结 论

内含子能较明显地提高转基因动物的基因表达效率, 它可能通过这样一些机制来起作用, 即含有增强子样作用元件, 或通过剪接增加 mRNA 的稳定性, 或通过改变核质状态开放染色体功能区来提高转基因表达效率。其存在于胚胎发育中所识别的序列中, 这也是内含子在转基因动物和细胞表达系统的差别之一。尽管内含子发挥作用的这些机制仍需进一步研究, 但内含子对转基因动物基因表达的重要影响已毋庸置疑。从实践的观点来看, 它给我们的启示是, 在表达载体构建策略中, 使用基因组序列一般有可能获得表达, 而且在绝大多数情况下可能获得高效表达。如果采用 cDNA, 那么插入异源内含子, 或将 cDNA 插入基因第一外显子 5' 端, 也可能改善表达水平。总之, 考虑到内含子的一般需要, 能构建出有效的表达构件。然而由于对机制尚缺乏充分了解, 因而还不能对构件进行直接的选择, 同时还需研究更为有效的评价方法。

参 考 文 献

- 1 Brinster R L, Allen J A, Behringer R R *et al.* Proc Natl Acad Sci USA, 1988; **85**: 836
- 2 Hamer D, Leder P. Cell, 1979; **18**: 1299
- 3 Mulligan R C, Berg P. Science, 1980; **209**: 1422
- 4 Carlock L, Jones N C. Nature, 1981; **294**: 572
- 5 Treisman R, Novak L, Favaloro J *et al.* Nature, 1981; **292**: 595
- 6 Chee P P, Klassy R C, Slighton J L. Gene, 1986; **41**: 47
- 7 Ng R, Domdey H, Larson G *et al.* Nature, 1985; **314**: 183
- 8 Gross M K, Kainz M S, Merrill C F. Mol Cell Biol,

1987; **7**: 4576

- 9 Palmiter R D, Sandgren E P, Avarbock M R *et al.* Proc Natl Acad Sci USA, 1991; **88**: 478
- 10 Whitelaw C B A, Archibald A L, Harris S *et al.* Transgenic Res, 1991; **1**: 3
- 11 Meade H, Gates L, Lacy E *et al.* Biotechnology, 1990; **8**: 443
- 12 Simons J P, McClenaghan M, Clark A J. Nature, 1987; **328**: 539
- 13 Choi T, Huang M T F, Gorman C. Mol Cell Biol, 1991; **6**: 3070
- 14 Mitchell P J, Tjian R. Science, 1989; **244**: 371
- 15 Dreyfuss G, Swanson M S, Pinol-Roma S. Trends Biochem Sci, 1988; **13**: 86
- 16 Sveren J, Chalkley R. Trends Genet, 1990; **6**: 52
- 17 Hammer R E, Swift C H, Ornitz D W *et al.* Mol Cell Biol, 1987; **7**: 2956
- 18 Chung S, Perry R P. Mol Cell Biol, 1989; **9**: 2075
- 19 Aronow B, Lattier D, Sibiger R *et al.* Genes Dev, 1989; **3**: 1384
- 20 Nyborg J K, Spindler S R. J Biol Chem, 1986; **26**: 5685
- 21 Behringer R R, Hammer R E, Brinster R L *et al.* Proc Natl Acad Sci USA, 1987; **84**: 7056
- 22 Buchman A R, Berg P. Mol Cell Biol, 1988; **8**: 4395
- 23 Callis J, Fromm M, Walbot V. Genes Dev, 1987; **57**: 47
- 24 Palseau F, Leung F, Kopchick J J. Gene, 1987; **57**: 47
- 25 Townes T M, Behringer R R. Trends Genet, 1990; **6**: 219

Effect of Intron on Increasing Efficiency of Gene Expression in Transgenic Animal. Lu Yi-fan, Deng Jixian, Xiao Chengzu, Ma Qingjun (Biotechnology institute, Military Medicine Academy, Beijing 100071, China).

Abstract The effect of intron on gene expression in transgenic animal is reviewed. The difference between natural and heterologous introns in increasing gene expression efficiency is analyzed and three possible mechanisms which introns promote gene expression in transgenic animal are discussed. It is necessary to consider intron effects on constructing expression vectors of producing transgenic animal.

Key words intron, transgenic animal, gene expression