

核基质结合序列(MAR)与基因表达调控

张可伟 王健美 郑成超*

(山东农业大学生命科学学院,泰安 271018)

摘 要 核基质结合序列(MAR)是能在体外与核基质特异结合的 DNA 序列,广泛存在于染色质 Loop 结构的边界序列中。随着研究的深入,发现 MAR 序列不仅在染色质折叠中起到重要作用,影响邻近内源基因表达,而且将 MAR 序列构建到外源基因表达盒两侧转化动植物时,也影响外源基因的表达。因此,MAR 序列是基因组中一种重要的基因间边界序列,为阐明非编码序列在基因表达中的作用和构建真核生物高效表达载体提供了新途径。

关键词 染色质,核基质结合序列(MAR),基因表达,转基因沉默

中图分类号 Q78 文献标识码 A 文章编号 1000-3061(2004)01-0006-04

真核生物染色质在 11nm 纤维、30nm 纤维和 300nm 纤维三个水平上进行包装,其中 30nm 纤维通过与核基质结合形成一系列大小不一的 Loop 环,构成 300nm 纤维^[1]。在 Loop 环的基部序列中,发现 MAR(matrix attachment region)序列有规律地存在,推测 MAR 序列在 Loop 环的形成和内源基因表达中起重要作用^[2,3]。来自酵母、植物、动物中的 MAR 序列被构建到真核生物表达载体外源基因表达盒两侧,能影响外源基因在转基因动植物中表达。MAR 的这一特性可应用于高效表达载体的构建^[4]。近年来,随着研究的深入,MAR 序列的作用及其作用机制日益清晰。本文综述了 MAR 序列在体内的存在形式、MAR 序列的作用及作用机制,同时对 MAR 序列的研究前景进行了展望。

1 核基质与核基质结合序列

核基质是存在于细胞核中由蛋白质纤维构成的网状结构,在完整的细胞核中,染色质纤维频繁地与核基质结合,形成染色质的高级结构,利于染色质行使其复制、分配等遗传功能^[5]。从生化角度来说,核基质就是选择性地去除组蛋白和 DNA 后,剩余的主要由蛋白质、少量 DNA 和结构 RNA 组成的复合体。早期的制备方法是运用高盐抽提和 DNase I 处理。后来,采用更接近于生理条件的低浓度 3,5-二碘水杨酸锂(LIS)来选择性的抽提,并用限制酶代替 DNase I 切割降解核基质不结合 DNA,得到结合有少量 DNA 和结构 RNA 的蛋白网状结构,即核基质。残留在核基质上的 DNA 约占 DNA 总量的 3%~6%,其中包含能与核基质紧密结合的 MAR 序列。鉴定一段 DNA 序列是否为 MAR 序列,主要根据它能否在体外与制备的核基质发生特异性结合来判断^[6]。

2 染色质状态与 Loop 结构模型

染色质是由 DNA、组蛋白、非组蛋白及少量 RNA 组成的线性复合体结构,按功能状态不同可以分为活性染色质和非活性染色质。基因的表达不仅需要 RNA 聚合酶和转录因子的存在,而且需要能接纳 RNA 聚合酶和转录因子的活性染色质状态的存在。当染色质处于活性状态时,DNA 序列上的顺式作用元件才能与反式作用因子相互作用,这样细胞就能够利用调控特异性转录因子的多少或活性来对基因转录实行精细调控。测定染色质活性状态通常通过测定其对 DNase I 的敏感性,具有转录潜能或正在转录的区域都对 DNase I 敏感。一般的 DNA 敏感性区域分布在基因两侧的几千个碱基对内,所以推测真核生物基因组由许多独立限定的区域组成,形成相对开放的染色质结构。当 DNA 敏感性区域的边缘序列被分离后,发现在这些区域中存在多段 MAR 序列^[2]。后来验证 MAR 序列在体外也能与核基质特异结合,所以有理由推测在体内,染色质纤维很可能通过 MAR 序列与核基质结合形成一系列的 Loop 结构。每一个 Loop 结构形成一个拓扑上独立的区域,染色质的凝缩与 DNA 的转录在此区域可以不受干扰的独立进行。这样,一些 Loop 结构即使位于常染色体内也可能永久性地处于固缩和失活状态,另一些 Loop 结构在细胞的合适分化时期就能够形成转录的结构,基因的转录表达不受 Loop 结构所处的染色质状态和周围其它基因的影响^[7]。

最近研究表明,在玉米、高粱和水稻中,某些同源基因的上下游序列尽管缺少序列的同源性,但 MAR 序列的数目与分布趋向于保守^[1,3],在动物中的小鼠和兔子中也发现了同

收稿日期 2003-06-30,修回日期 2003-10-16。

基金项目:国家十五“863”计划(No.2002AA224101)、国家自然科学基金(No.30270145)和国家转基因植物研究与产业化专项(No.J99-A-038)基金资助。

* 通讯作者。Tel 86-538-8242894,Fax:86-538-8226399,E-mail:cczheng@sdau.edu.cn

© 中国科学院微生物研究所期刊联合编辑部 <http://journals.im.ac.cn>

样的现象^[8]。这些结果表明在生物进化中,尽管物种间序列发生变化,但是基因所处的染色质的局域结构是趋向于保守的,也从另一个方面说明了基因两侧的 MAR 序列对于维持染色质的结构和基因功能的重要性。

3 MAR 对外源基因表达的影响

3.1 MAR 对外源基因表达的影响

早期研究认为,外源基因表达的差异主要是由于插入位点不同造成,外源基因整合到凝缩染色质或其附近就会降低其基因表达强度直至产生基因沉默。现在认为内源调节元件可能影响外源基因的表达,比如外源基因插入到抑制子附近,受到插入序列边缘抑制子序列的影响而降低基因表达^[9]。Loop 结构模型暗示 MAR 序列可能作为一个边界元件阻止染色体的凝缩区域或外源基因附近的调控元件对外源基因表达的影响。

目前已从许多种真核生物中分离出 MAR 序列,部分已被连接到外源基因表达盒两侧通过转基因鉴定其在体内对外源基因表达的影响,结果表明 MAR 序列能改善外源基因的表达状况,但普遍不能降低转基因个体间表达的差异性^[6,10,11]。我们分离的 MAR 序列中,有的尽管能显著增强外源基因稳定表达但对消除个体间的表达差异性影响不大^[4]。原因可能是外源基因插入位置不同,MAR 序列在不同染色质环境中参与染色质的包装方式也不尽相同,所以在不同的转基因个体中 MAR 序列虽然普遍改善了外源基因的表达环境,但改变的程度仍与外源基因的整合位置有关,外源基因在不同的个体中仍有表达水平上的差异。

Brouwer 等发现来自玉米乙醇脱氢酶 5' 端的一段 MAR 序列,在玉米愈伤中能增强外源基因表达,但在转基因再生玉米植株中却抑制外源基因在根中组成型表达^[1]。说明 MAR 序列在体内起的作用不仅仅是单独与核基质结合,而且可能接受组织特异的调控蛋白的调节,调整其在染色质折叠中的作用,影响外源基因的表达。Fujiwara 也分离到了在不同细胞时期含量相差很大的 MAR 序列结合蛋白^[12],直接说明了在细胞的不同时期,MAR 序列的结合蛋白可能不一样。最近,Mankin 等发现对于特定的 MAR 序列,其对外源基因表达的作用依赖于外源基因的启动子,MAR 序列对因组织特异性不表达的启动子没有增强表达的作用而对表达的启动子有增强表达的作用,验证了 MAR 序列对染色体状态的改变作用与外源基因顺式元件的调节作用是决定启动子活性并直接影响基因表达的两个关键因素^[13],两者之间是否存在一种平衡机制有待于进一步的实验证实。

3.2 MAR 序列对外源基因沉默影响

Allen 等将酵母 MAR 序列构建到 *Gus* 报告基因表达盒两侧转化烟草悬浮系,最早提出 MAR 序列能降低同源基因依赖基因沉默的假说^[14]。他们在研究中虽然没有发现拷贝数依赖的基因表达,但是发现当拷贝数高于某一阈值时,拷贝数的增加降低基因表达水平。有趣的是在外源基因两侧加 MAR 序列的转基因植物中这个阈值增大。当外源基因的

拷贝数超过某一阈值时,基因的表达量降低,这种现象可能由于同源抑制引起,而不是经典的位置效应引起。MAR 序列提高了这个阈值,说明 MAR 序列在一定程度上减少了外源基因沉默。

Anandalakshmi 等研究发现,两侧携带 MAR 的外源基因的转基因植株发生基因沉默主要在转录后水平^[15]。当转 MAR-*Gus* 基因的沉默株系与表达一个转录后基因沉默抑制子(PIHC-Pro)的株系杂交时,在杂交后代中,沉默抑制子消除了转录后基因沉默从而使后代中的 *Gus* 基因被激活表达;但是,单纯转 *Gus* 基因的沉默株系与 PIHC-Pro 株系杂交,后代中的 *Gus* 基因没被激活。表明 MAR 序列对基因沉默的消除作用发生在转录水平,但不能消除转录后水平基因沉默。

3.3 MAR 的核基质体外结合能力与体内功能的关系

假如 MAR 序列仅仅通过与核基质结合在体内起作用,那么具有强核基质结合活性的 MAR 序列应该具有更强的增强作用。Allen 等发现自主复制序列 ARS(autonomous replicating sequence)能使外源基因在悬浮细胞的表达增强 12 倍,而一个与核基质结合能力更强的 MAR 序列,能增强基因表达 60 倍^[14,16]。Brouwer 等发现,体外核基质结合能力强的 *Mha* MAR 序列在玉米愈伤中对外源基因表达的影响小于核基质结合能力相对较弱的 *Adh* MAR 序列的增强作用^[1]。我们的研究结果也表明体外结合能力较强的 TM3 增强外源基因表达的能力不及体外结合能力较弱的 TM2^[4]。说明 MAR 序列的体外结合能力虽然是必要的,但也不能直接决定其在体内对外源基因的增强功能。MAR 序列的体内作用不仅仅是简单的参与核基质结合决定染色质状态,还可能存在着一种与调控蛋白结合进而调控启动子活性的机制。MAR 序列与核基质的结合和 MAR 序列与调控蛋白的结合也可能存在一个相互竞争的过程。MAR 序列与核基质的结合可能阻碍了其调控蛋白的结合能力,既可能削弱调控蛋白的增强作用,也可能削弱调控蛋白的阻遏作用,MAR 序列在体内结合与体外结合之间的关系可能是多种因素平衡的结果。因此,我们认为 MAR 序列与核基质的体外结合仅仅反映了 MAR 序列与蛋白的结合能力,并不能决定其体内功能。

4 MAR 作用的分子机制

迄今,关于 MAR 序列对基因表达的作用的报道几乎都把 MAR 序列当作一个具有多种功能的 DNA 元件^[6]。由于目前仍然不能在活体内直接检测到 MAR 序列与核基质的结合情况,因此,难以确定 MAR 序列对转基因的影响是依赖于与核基质的结合还是依赖于其它功能。人们根据所得的实验结果推测出 MAR 序列作用的 4 种模型,可分别用来解释 MAR 序列的不同作用。

4.1 MAR 的终止子模型(Terminator model)

MAR 序列可以被想象为一个强的终止子。弱的转录终止作用能极大的降低基因表达的水平,当 RNA 聚合酶转录到转录单元的周围区域时,转录干涉现象就有可能发生^[17]。弱的终止可能会产生能诱导转录后基因沉默的异常 RNA 产

物。此外,转录通读能够在一个多拷贝位点产生多个含启动子序列的转录本。另一实验证明细胞核中含启动子的转录物的积累能导致基因沉默现象的发生^[18]。所以假如 MAR 序列能提高终止效率,就可望解决或至少能部分解决转录终止效率不高的问题,从而提高基因表达效率。

4.2 染色质开放模型(Chromatin opening model)

染色质开放模型最早由 Laemmli 等提出^[19]。长久以来人们就知道 H1 偏爱与富含 AT 的 DNA 序列结合。Laemmli 的工作进一步阐明了 H1 与 MAR 序列协作,排布 H1 蛋白分子并压缩成染色质。寡聚肽偏端霉素能选择性地竞争 H1 与 MAR 序列结合,使分离出来的 H1 再分布,染色质变得对拓扑异构酶 II 和限制性内切酶更为敏感,这种状态的染色质为活性染色质。他们发现非组蛋白 HMGI/Y 类蛋白能通过凹槽以类似偏端霉素的方式与富含 AT 的 DNA 序列结合,使染色质状态变得疏松。所以,富含 AT 的 MAR 序列与类似 HMGI/Y 的蛋白质结合,可能是驱动染色质开放调节染色质状态的关键因子。Allen 等证明 MAR 序列在悬浮细胞中能平均提高外源基因表达 60 倍,但在再生烟草植株中只能增强外源基因表达 2 到 3 倍,蛋白分析表明悬浮细胞中含有更多的 HMGI/Y 类蛋白^[14,16],说明 HMGI/Y 蛋白很可能在体内与 MAR 序列作用影响染色质状态进而影响基因功能。

4.3 促外源基因整合模型(Integration enhancing model)

目前,对转基因时 DNA 是如何整合到染色质中还不清楚。然而,在动物和酵母的非同源重组中发现 MAR 序列的存在,在拟南芥和水稻中发现在外源基因的整合位点也存在 MAR 序列,说明 MAR 可能以某种方式产生整合热点,提供外源基因插入^[20]。因此有理由推测在外源基因两侧构建上 MAR 序列可能以某种方式增强转基因的整合效率。这种推测已被实验证实,Petersen 等发现 P1-MAR 序列和 TBS-MAR 序列能提高外源基因在大麦愈伤中的整合效率^[21];Han 等也发现 MAR 序列能提高转基因杨树抗性芽的生成^[22]。其原因可能是由于染色体蛋白如组蛋白 H1、HMGI/Y 等能选择性的与 MAR 序列结合,对外源 DNA 有保护作用,刺激整合过程,甚至帮助外源基因定位到基因组的活性区域^[6]。

4.4 抑制同源配对模型(Pairing interference model)

多数实验表明,MAR 序列能抑制整合到相同位点的多个外源基因拷贝之间的配对作用。一般来说,在植物系统和动物系统中,当多个外源基因拷贝整合到同一位点时,由于外源基因相距太近而极易配对,抑制了外源基因正常转录而导致基因沉默。Allen 等的研究结果表明当外源基因两侧插入 MAR 序列后,多拷贝的外源基因往往比单拷贝的基因更易于表达^[14]。说明 MAR 序列可能通过与核基质的结合使每个拷贝的外源基因形成一个独立的 Loop 结构进而抑制了外源基因各拷贝间的配对作用,促进了外源基因的转录。

然而,当外源基因的拷贝数高于一个阈值时,抑制配对模型变得更加复杂,MAR 序列就不能抑制基因沉默^[12,14]。这里可能存在两种机制:一是多拷贝外源基因的存在可能饱和了核基质上的 MAR 序列结合位点,一些外源基因拷贝尽

管两侧有 MAR 序列但不能与核基质有效结合,这些未与核基质结合的拷贝将与已结合或未结合的拷贝配对产生同源抑制引起的基因沉默;二是两侧带有 MAR 序列的转基因其转录活性较强,使转录出来的 mRNA 量与拷贝数成正比,当 mRNA 的量高于某一阈值时,将诱发转录后基因沉默。这两种推测还有待于今后实验证实。

5 展 望

目前,MAR 序列结合蛋白的分离与鉴定,MAR 序列在体内与结合蛋白如何结合的直接证据成了研究的热点和难点。这些问题的解决可以为以上的作用模型提供直接的实验证据。将 GFP 与染色质蛋白或 MAR 序列结合蛋白融合转化动植物,将有可能直接在活体细胞中观察到 MAR 与蛋白的体内作用,利用 FISH 技术对转基因个体基因整合位点分析则可望直接观察到 MAR 对染色质状态的消除作用。MAR 序列是基因边界序列的一部分,是研究较早的一类基因间 DNA 序列。在动植物已大规模测序后的今天,了解这些序列在染色质折叠和基因表达调控方面的作用,对于功能基因组的研究、生物基因组进化和真核高效表达载体的构建等均有重要意义。相信随着 MAR 序列研究的深入,MAR 序列及其它一些基因间序列的作用将会更加清楚,最终可以协助破译复杂的基因组测序后的天文数据密码,使转基因技术成为高效的物种改良手段。

REFERENCES(参考文献)

- [1] Brouwer C, Bruce W, Maddock S *et al.* Suppression of transgene silencing by matrix attachment regions in maize: a dual role for the maize 5'ADH1 matrix attachment region. *Plant Cell*, 2002, **14**: 2251 - 2264
- [2] Pikaard CS. Chromosome topology-organizing genes by loops and bounds. *Plant Cell*, 1998, **10**: 1229 - 1232
- [3] Tikhonov AP, Bennetzen JL, Avramova ZV. Structural domains and matrix attachment regions along collinear chromosomal segments of maize and sorghum. *Plant Cell*, 2000, **12**: 249 - 264
- [4] Zhang KW(张可伟), Wang JM(王健美), Yang GX(杨国栋) *et al.* Isolation of a strong matrix attachment region(MAR) and identification of its function *in vitro* and *in vivo*. *Chin Sci Bulletin(科学通报)* 2002, **47**(20): 1572 - 1577
- [5] Mesner LD, Hamlin JL, Dijkwel PA. The matrix attachment region in the Chinese hamster dihydrofolate reductase origin of replication may be required for local chromatid separation. *Proc Natl Acad Sci USA*, 2003, **100**(6): 3281 - 3296
- [6] Allen GC, Spiker S, Thompson WF. Use of matrix attachment regions(MARs) to minimize transgene silencing. *Plant Mol Biol*, 2000, **43**: 361 - 376
- [7] Xin L, Liu DP, Ling CC. A hypothesis for chromatin domain opening. *Bioessays*, 2003, **25**(5): 507 - 514
- [8] Millot B, Montolieu L, Fontaine ML *et al.* Hormone induced modifications of the chromatin structure surrounding 5 upstream regulatory regions conserved between the mouse and rabbit WAP genes. *Biochem J*, 2003, **372**(1): 41 - 52
- [9] Campisi L, Yang YZ, Yi Y *et al.* Generation of enhancer trap lines

- rescence. *Plant J*, 1999, **17**: 699 – 707
- [10] Li XC(李旭刚), Zhu Z(朱桢), Xu JW(徐军望) *et al.* Isolation of pea matrix attachment region and study on its function in transgenic tobaccos. *Science in China Series C(中国科学 C 辑)*, 2001, **31** (3): 230 – 237
- [11] Wu ZY(吴震宇), Zhao MJ(赵慕钧), Li ZP(李载平). Regulation of gene expression by SAR of silkworm *Attacus ricini* rRNA gene. *Acta Biochem Biophys Sin(生物化学与生物物理学报)*, 2001, **33** (1): 59 – 64
- [12] Fujiwara S, Matsuda N, Sato T *et al.* Molecular properties of a matrix attachment region-binding protein located in the nucleoli of tobacco cells. *Plant Cell Physiol*, 2002, **43** (12): 1558 – 1567
- [13] Mankin SL, Allen GC, Phelan T *et al.* Elevation of transgene expression level by flanking matrix attachment regions (MAR) is promoter dependent : a study of the interactions of six promoters with the RB7 3' MAR. *Transgenic Res*, 2003, **12** (1): 3 – 12
- [14] Allen GC, Hall GJ, Spiker S *et al.* Scaffold attachment regions increase reporter gene expression in stably transformed plant Cells. *Plant cell*, 1993, **5**: 603 – 613
- [15] Anandalakshmi R, Pruss GJ, Ge X *et al.* A viral suppressor of gene silencing in plants. *Proc Natl Acad Sci USA*, 1998, **95**: 13079 – 13084
- [16] Allen GC, Hall GJ, Michalowski S *et al.* High-level transgene expression in plant cells : Effects of a strong scaffold attachment region from tobacco. *Plant Cell*, 1996, **8**: 899 – 913
- [17] Greger IH, Proudfoot NJ. Poly(A) signals control both transcriptional termination and initiation between the tandem GAL10 and GAL7 genes of *Saccharomyces cerevisiae*. *EMBO J*, 1998, **17**: 4771 – 4779
- [18] Al-Kaff NS, Covey SN, Kreike MM *et al.* Transcriptional and post-transcriptional plant gene silencing in response to a pathogen. *Science*, 1998, **279**: 2113 – 2115
- [19] Käs E, Poljak L, Adachi Y *et al.* A model for chromatin opening : stimulation of topoisomerase-II and restriction enzyme cleavage of chromatin by distamycin. *EMBO J*, 1993, **12**: 115 – 126
- [20] Dietz-Pfeilstetter A, Arndt N, Kay V *et al.* Molecular structure and regulatory potential of a T-DNA integration site in petunia. *Transgenic Res*, 2003, **12** (1): 83 – 99
- [21] Petersen K, Leah R, Knudsen S *et al.* Matrix attachment regions (MARs) enhance transformation frequencies and reduce variance of transgene expression in barley. *Plant Mol Biol*, 2002, **49**: 45 – 58
- [22] Han KH, Ma CP, Strauss SH. Matrix attachment regions (MARs) enhance transformation frequency and transgene expression in poplar. *Transgenic Res*, 1997, **6**: 415 – 420

The Potential Role of Nuclear Matrix Attachment Regions (MARs) in Regulation of Gene Expression

ZHANG Ke-Wei WANG Jian-Mei ZHENG Cheng-Chao*

(College of Life Sciences, Shandong Agricultural University, Taian 271018, China)

Abstract Gene transfer technology is being used to enhance agronomic performance or improve quality traits in a wide variety of crop species. However, it is sometimes severely handicapped by difficulty in obtaining material in which transgene expression is predictable and stable over many generations. Because integration seemed to occur randomly in the plant genome, it was thought that some transgenes would be integrated in a relatively uncondensed, transcriptionally active chromatin environment, while others in a condensed, transcriptionally inert chromatin structure. Nuclear matrix attachment regions (MARs) are defined as DNA sequences that bind preferentially to the proteins of the nuclear matrix. They typically are localized at the borders of gene domains, implicating them in the formation of individual loops of higher order chromatin structure and transcription regulation. When MARs are positioned on either side of a transgene their presence usually results in higher and more stable expression in transgenic plants, most likely by minimizing gene silencing. In this review, we focus mainly on novel findings and our observations concerning the function of MARs in transcription regulation. Our objective is not only to summarize the current data and present several possible models to explain MAR effects on the transcription regulation, but also to point out some open questions involving the utilization of MARs in constructing high efficient expression vectors.

Key words chromatin structure, matrix attachment regions (MAR), gene expression, transgene silencing

Received : 06-30-2003

This work was supported by grants from The National Natural Sciences Foundation of China (No. 30270145), The State " Ten-five " 863 High Technology R&D Project of China (No. 2002AA224101) and The National Special Program for Research and Industrialization of Transgenic Plants of China (No. J99-A-038).

* Corresponding author. Tel : 86-538-8242894 ; Fax : 86-538-8226399 ; E-mail : cczheng@sdau.edu.cn ©: 中国科学院微生物研究所期刊联合编辑部 http://journals.im.ac.cn