

文章编号: 1005-0906(2007)04-0037-04

玉米基因组重复序列的研究进展

韩永华^{1,2}, 宋运淳³, 金危危¹

(1. 中国农业大学国家玉米改良中心, 北京 100094; 2. 广西师范大学生命科学学院, 广西 桂林 541004;

3. 武汉大学植物发育生物学教育部重点实验室, 武汉 430072)

摘要: 玉米基因组中大部分是重复序列。重复序列可分为两种类型: 一种为串联重复序列, 另一种为散在重复序列。串联重复序列主要分布在染色质组、核仁组织区、着丝粒和端粒区域, 还有一些短的串联重复序列随机分布于基因组中。散在重复序列主要是可转移因子, 可转移因子分为转座子和返座子两种类型。文中综述了玉米基因组中各种重复序列的研究进展。

关键词: 玉米基因组; 重复序列; 研究进展

中图分类号: S513

文献标识码: A

Research Progress of Repetitive Sequence in Maize Genome

HAN Yong-hua^{1,2}, SONG Yun-chun³, JIN Wei-wei¹(1. *National Maize Improvement Center of China, China Agricultural University, Beijing 100094;**2. College of Life Sciences, Guangxi Normal University, Guilin 541004;**3. The Key Laboratory of MOE for Plant Developmental Biology, Wuhan University, Wuhan 430072, China)*

Abstract: The maize genome is made up largely of repetitive sequence which consists of tandem repetitive sequence and dispersed repetitive sequence. Tandem repetitive sequence mainly locates on knob, nucleolar organizer region, centromere and telomere of chromosome. Short tandem repeat randomly distributes in whole genome. Dispersed repetitive sequence is mainly transposable elements which are composed of transposon and retroposon. This paper reviewed progress of repetitive sequence in maize genome.

Key words: Maize genome; Repetitive sequence; Research progress

重复序列在真核生物基因组中占有很大比重, 绝大部分存在于间隔序列和调控序列等非编码序列中, 也分布在有些结构基因的序列中, 后者多为低度重复序列。植物基因组重复序列的比例与基因组大小有关, 基因组较大的植物其重复序列所占的比重也较大^[1]。如植物中拟南芥基因组最小, 仅为 120 Mb, 重复序列大约占 25%; 玉米基因组较大, 为 2 500 Mb, 其中 60%~80%是重复序列^[2]。根据重复序列在基因组中组织方式的不同, 重复序列可分为两种类

型: 一种为串联重复序列(tandem repetitive sequence), 许多个重复单位按照串联形式排列在一起; 另一种为散在重复序列(dispersed repetitive sequence), 主要指可转移因子(transposable elements), 其重复单位与其它无关重复序列或单拷贝序列相间排列。重复序列对维持染色体的空间结构、基因的表达调控、遗传重组等都具有重要作用。重复序列单元、拷贝数及其在基因组中分布位置、存在状态都具有多样性。通过对重复序列进行研究, 可以揭示基因的结构、功能机制、基因组的多态性、标记基因等^[3]。

1 玉米中的串联重复序列

1.1 染色质组中的串联重复序列

染色组(knob)是玉米染色体上明显可见的异染色质区, 在不同的品系中, 组的数目、大小和分布位置不同^[4]。Peacock 等^[5]证实玉米染色组是由不编码的高度重复的卫星 DNA 组成, 卫星 DNA 的重复单位

收稿日期: 2006-12-05; 修回日期: 2007-04-05

基金项目: 国家自然科学基金(39870423)、广西博士点学科建设项目

作者简介: 韩永华(1971-), 女(蒙古族), 吉林双阳人, 副教授, 博士, 从事植物分子细胞遗传学研究。Tel: 010-62734249

E-mail: hanyh329@163.com

金危危为本文通讯作者。Tel: 010-62734249

E-mail: jinweiwei_71@hotmail.com

是 180bp。进一步用原位杂交方法证实这种卫星 DNA 只存在于异染色质组中,在着丝粒异染色质、NOR 异染色质及 B 染色体的异染色质中都不含有这种卫星 DNA 成分。Ananiev 等^[6]发现玉米染色组的分子结构极复杂,除 180bp 的重复序列外,还含有一个 350bp 的重复序列(TR-1)以及各种返座子。这两种重复序列分布在同一个组中、不同的组中以及染色体上没有明显的染色组的位置。

1.2 核糖体 DNA

核糖体 DNA 即 rDNA,是细胞中编码核糖体 RNA(rRNA)的基因。rDNA 有 2 种,一种为 45S rDNA,另一种为 5S rDNA。其中 45S rDNA 编码细胞质中 18S、5.8S 和 28S(或 25S、26S)rRNA 是真核生物中一类高度重复的串联重复序列,集中分布在一个或数个位点,大都与核仁组织区(NOR)相联系。植物中 45S rDNA 的重复单位长度为 7.8~18.5 kb,单倍体基因组中的拷贝数为 600~8 500。5S rDNA 编码细胞质中 5S rRNA 是真核生物中另一类高度保守的串联重复序列,重复单位长度为 200~900 bp,单倍体基因组中的拷贝数为 1 000~50 000。真核生物基因组中,5S rDNA 与 45S rDNA 在染色体上的分布彼此独立。除小麦、节节麦和黑麦等少数物种中两者处于相同的染色体上外,大多数物种中 5S rDNA 并不位于核仁组织区所在染色体上^[7]。玉米中 45S rDNA 和 5S rDNA 重复单位长度分别为 9.1 kb 和 322 bp^[8],45S rDNA 位于第 6 染色体短臂上,5S rDNA 位于第 2 染色体长臂近末端^[9]。

1.3 着丝粒中的串联重复序列

着丝粒是真核生物有丝分裂和减数分裂染色体的极向运动所必需的染色体位点。植物的着丝粒主要由卫星 DNA 序列和多种还原转座子构成。卫星 DNA 是 140~180 bp 的串联重复单位,它构成植物着丝粒的中心域,而还原转座子序列插入到串联重复的卫星 DNA 阵列中或位于中心域的两侧。玉米着丝粒区串联重复的卫星 DNA 序列称为 CentC,重复单位长度为 156 bp^[10]。FISH 分析显示,CentC 仅分布于玉米染色体的着丝粒区,而且每条染色体上 CentC 的拷贝数存在显著差异,但在第 4 染色体上 CentC 的杂交信号最弱。Page 等^[11]从玉米第 4 染色体的着丝粒区分离出一个重复 DNA 序列 Cent4, FISH 分析显示 Cent4 仅存在于第 4 染色体的着丝粒区,是第 4 染色体专一的着丝粒序列。在玉米 B 染色体上存在 B 染色体着丝粒特有的重复序列 pZmBs^[12]。pZmBs 与 CentC 无明显同源性,但它与

Cent4 的部分序列具有同源性。这提示玉米第 4 染色体与 B 染色体有进化上的关系。

1.4 端粒中的串联重复序列

高等植物中端粒重复序列是高度保守的,大多数植物的端粒中都含有拟南芥形式的端粒序列 5'-TTTAGGG-3'。但是,植物中也有非拟南芥形式的端粒序列^[13]。拟南芥的端粒重复序列能与玉米的端粒杂交,说明玉米的端粒中含有拟南芥形式的端粒序列^[14]。

1.5 MR68

重复单位长度为 410 bp,在玉米基因组中的拷贝数约为 3 000。FISH 分析显示 MR68 序列位于玉米第 3 和第 6 染色体长臂的亚端粒区以及第 6 染色体短臂的随体上^[15]。

1.6 短串联重复序列

短串联重复序列(short tandem repeats,STR),又称简单重复序列(simple sequence repeats,SSR),是由重复单元为 1~5 个碱基对的 DNA 序列所构成的串联重复的 DNA 家族,它有以下特点:①序列简单,重复单元为 1~5 bp,但是具有多种组合方式;②高度重复多拷贝,从几十到几百乃至千万拷贝数;③散状分布,大多数 SSR 以随机方式存在于单拷贝或低拷贝基因间,因此可以用于标记相邻编码基因,此类重复片段一般较长。还有一些 SSR 存在于基因内,可用于评价等位基因的遗传多样性和杂种鉴定等方面的研究;④共显性分子标记,等位的 DNA 区段多态性易保留在基因组内,因为大多数 SSR 是不编码的,不易被自然选择和人工选择而淘汰;⑤SSR 的多态性经 PCR 扩增技术就可呈现出来,技术简便易行。玉米基因组中富含 CT/GA、CCT/GGA、CCG/GGC、CAG/GTC、TAG/ATC 和 GAA/CTT 等 2、3 个碱基的重复,还有其它形式多样的 4 个碱基和 5 个碱基的重复^[16]。玉米不同自交系之间的 SSR 具有明显的多态性,因此它在玉米的遗传学研究方面具有广泛的应用价值,如用于遗传连锁图的构建、种质资源的评估以及辅助田间育种等^[3]。

2 玉米中的散在重复序列

在植物基因组中,大部分重复序列是散在重复序列,主要是可转移因子。根据转座机制的不同,可转移因子分为两种类型:类型 I 称为转座子(transposon),其遗传信息的流动是从 DNA 到 DNA。转座子成分又可分为两大类:一类叫自主成分,它们具备使自身切除或转座的能力;另一类叫非自主成分,它们

只有在同一组自主成分存在时才能切除或转座。转座子成分在基因组中的含量较低,通常单倍体基因组中一种转座子成分的拷贝数少于 100。但有一类特殊类型的非自主成分,称为微型倒转重复转移因子 MITEs (miniature inverted repeat transposable elements),在单倍体基因组中的拷贝数非常高,约为 500 ~ 15 000,但这类转座子的长度较短(<500 bp)^[17]。类型 II 称为返座子(retroposon),其遗传信息的流动是从 DNA 到 RNA 再到 DNA。根据结构特征的不同,返座子又可分成两个大家族(superfamily),一个是病毒大家族,另一个是非病毒大家族。病毒大家族包括所有的还原病毒和一类结构与还原病毒的原病毒相似的转座成分,称为还原转座子(retrotransposon)。非病毒大家族包括长散布重复序列 LINEs (long interspersed repeated elements) 和短散布重复序列 SINEs(short interspersed repeated elements)等^[18]。

2.1 玉米基因组中的返座子

在玉米基因组中,50%以上是可转移因子,其中大部分是还原转座子中的 Ty1-copia 和 Ty3-gypsy 成分,但没有发现 LINEs 和 SINEs 返座子。玉米基因组中已识别的丰富的还原转座子为 Huck, Clinful, Ji, Opie^[19], Prem-2^[20], Zeon^[21], Grande(GenBank accession no. X82087), Prem-1^[22] 和 ANLI^[23]。其中 Ji, Opie 和 Prem-2 是 Ty1-copia 型的还原转座子成分, Huck, Clinful, Grande, ANLI 和 Prem-1 是 Ty3-gypsy 型的还原转座子成分。在玉米基因组中还有一些含量较少的还原转座子,它们是 Tekay(GenBank accession no. AF050455A), Hopscotch^[24], Stoner^[25], Magellan, Reina, Fourf, Kake, Victim, 和 Milt^[26]。已有研究表明玉米中含量较多的还原转座子通常分布在异染色质区,处于甲基化状态,因而不表达的。如 Huck, Grande, Prem-2 和 Opie 广泛分布在着丝粒串联重复的卫星 DNA 阵列中或其两侧^[27]。Huck, Grande, Opie, Zeon-1, ANLI, Prem-1 和 Prem-2 分布在染色体的串联重复阵列中,其中 Prem-1 和 Prem-2 在染色体中分布最广。而含量较少的还原转座子如 Tekay, Reina, Fourf, Victim 和 Hopscotch 通常分布在去甲基化的基因丰富区,是能够表达的。

2.2 玉米基因组中的转座子

玉米基因组中已识别的转座子包括 Ac/Ds, Spm/En 和 Mu 等系统,已识别的 MITEs 家族有 B2^[28], mPIF^[29], Hb2^[30] 和 Hbr^[31]。这些转座子系统和 MITEs 家族都是优先分布在去甲基化的基因丰富区^[32], 而它们常被用作从玉米基因组中寻找基因的标签。

参考文献:

- [1] Haberer G, Young S, Bharti A K, Gundlach H, Raymond C, et al. Structure and architecture of the maize genome[J]. Plant Physiol, 2005, 139(4): 1612-1624.
- [2] 铁双贵,郑用琰. 玉米基因组的简单重复序列遗传研究进展[J]. 生物工程进展, 2001, 21(5): 59-62.
- [3] McClintock B, Kato T A, Blumenschein A. Chromosome constitution of races of maize[A]. Colegio de Postgraduados Chapingo, Mexico. 1981.
- [4] Peacock W J, Dennis E S, Phoades M M, Pryor A J. High repeated DNA sequence limited to knob heterochromatin in maize[J]. Proc Natl Acad Sci. USA, 1981, 78: 4490-4494.
- [5] Ananiev E V, Phillips R L, Rines H W. A knob-associated tandem repeat in maize capable of forming fold-back DNA segments: are chromosome knobs megatransposons[J]. Proc Natl Acad Sci USA, 1998, 95: 10785-10790.
- [6] Lapitan N L V. Organization and evolution of higher plant nuclear genomes[J]. Genome, 1992, 35: 171-181.
- [7] Zimmer E A, Jupe E R, Walbot V. Ribosomal gene structure, variation and inheritance in maize and its ancestors[J]. Genetics, 1988, 120: 1125-1136.
- [8] Li L J, Arumuganthan K. Physical mapping of 45S and 5S rDNA on maize metaphase and sorted chromosomes by FISH[J]. Hereditas, 2001, 134: 141-145.
- [9] Ananiev E V, Phillips R L, Rines H W. Chromosome-specific molecular organization of maize (*Zea mays* L.) centromeric regions[J]. Proc Natl Acad Sci. USA, 1998, 95(22): 13073-13078.
- [10] Page B T, Wanous M K, Birchler J A. Characterization of a maize chromosome 4 centromeric sequence: evidence for an evolutionary relationship with the B chromosome centromere[J]. Genetics, 2001, 159: 291-302.
- [11] Alfenito M R, Birchler J A. Molecular characterization of a maize B chromosome centric sequence[J]. Genetics, 1993, 135(2): 589-597.
- [12] Adams S P, Leitch I J, Bennett M D, Leitch A R. Aloe L.-a second plant family without (TTTAGGG)_n telomeres[J]. Chromosoma, 2000, 109: 201-205.
- [13] 柴守诚,员海燕. 高等植物 DNA 重复序列的主要类型和特点[J]. 西北植物学报, 1999, 19(3): 555-563.
- [14] Chen C X, Yan H H, Zhai W X, Zhu L H, Sun J S. Identification and chromosomal location of a new tandemly repeated DNA in maize[J]. Genome, 2000, 43: 181-184.
- [15] Ananiev E V, Chamberlin M A, Klaiber J, Svtashev S. Microsatellite megatracts in the maize (*Zea mays* L.) genome[J]. Genome, 2005, 48 (6): 1061-1069.
- [16] Jiang N, Wessler S R. Insertion preference of maize and rice miniature inverted repeat transposable elements as revealed by the analysis of nested elements[J]. Plant Cell, 2001, 13 (11): 2553-2564.
- [17] 孙乃恩,孙东旭,朱德煦. 分子遗传学[M]. 南京: 南京大学出版社, 1990.
- [18] SanMiguel P, Tikhonov A, Jin Y K, et al. Nested retrotransposons in the intergenic regions of the maize genome [J]. Science, 1996, 274: 765-768.

- [19] Hu W, Das O P, Messing J. Zeon-1, A member of a new maize retrotransposon family[J]. *Mol Gen Genet*, 1995, 248: 471-480.
- [20] Tureich M P, Mascarenhas J P. PREM-1, A putative maize retroelement has LTR (long terminal repeat) sequences that are preferentially transcribed in pollen[J]. *Sex Plant Reprod*, 1994, 7: 2-11.
- [21] Blake C. Meyers, Scott V. Tingey, michele morgante. mbundance, mistribution, and mtranscriptional mctivity of repetitive elements in the maize genome[J]. *Genome Research*, 2001, 11(10): 1660-1676.
- [22] White S E, Habera L F, Wessler S R. Retrotransposons in the flanking regions of normal plant genes: A role for copia-like elements in the evolution of gene structure and expression[J]. *Proc Natl Acad Sci.*, 1994, 91: 11792-11796.
- [23] Marillonnet S, Wessler S R. Extreme structural heterogeneity among the members of a maize retrotransposon family[J]. *Genetics*, 1998, 150: 1245-1256.
- [24] Tikhonov A P, SanMiguel P J, Nakajima Y, et al. Colinearity and its exceptions in orthologous adh regions of maize and sorghum[J]. *Proc Natl Acad Sci., USA*, 1999, 96: 7409-7414.
- [25] Theuri J, Phelps-Durr T, Mathews S, Birchler J. A comparative study of retrotransposons in the centromeric regions of A and B chromosomes of maize[J]. *Cytogenet Genome Res.*, 2005, 110(1-4): 203-208.
- [26] Bureau T E, Wessler S R. Tourist: A large family of small inverted repeat elements frequently associated with maize genes[J]. *Plant Cell*, 1992, 4: 1283-1294.
- [27] Zhang X Y, C é dric Feschotte, Zhang Q, et al. P instability factor: An active maize transposon system associated with the amplification of Tourist-like MITEs and a new superfamily of transposases[J]. *Proc Natl Acad Sci., USA*, 2001, 98(22): 12572-12577.
- [28] Spell M, Baran G, Wessler S R. An RFLP adjacent to the maize waxy gene has the structure of a transposable element[J]. *Mol Gen Genet*, 1988, 211: 364-366.
- [29] Zhang Q, Arbuckle J, Wessler S R. Recent, extensive, and preferential insertion of members of the miniature inverted-repeat transposable element family Heartbreaker into genic regions of maize[J]. *Proc Natl Acad Sci., USA*, 2000, 97(3): 1160-1165.
- [30] Brunner S, Fengler K, Morgante M, et al. Evolution of DNA sequence nonhomologies among maize inbreds[J]. *Plant Cell*, 2005, 17(2): 343-360.
- [31] Cowperthwaite M, Park W, Xu Z N, et al. Use of the transposon Ac as a Gene-searching engine in the maize genome[J]. *Plant Cell*, 2002, 14 (3): 713-726.
- [32] Jiang N, Wessler S R. Insertion preference of maize and rice miniature inverted repeat transposable elements as revealed by the analysis of nested elements[J]. *Plant Cell*, 2001, 13 (11): 2553-2564.

(责任编辑:李万良)