

番茄 RAPD 分析有效引物 3' 端序列的偏倚分布及其与基因组序列结构的关系

余诞年

(海南师范学院生物系, 海口 571158)

摘要 对番茄 RAPD 分析的有效引物与 Operon 引物系列(OP 系列)的 3' 端序列作了统计比较。研究结果指出, 番茄有效引物系列不是 OP 系列的随机样本, 而是 3' 端序列的种类和分布上有偏倚性的选择性样本。虽然两个系列的引物都具有 3' 端同序现象, 但对其分布频率正交比较结果差异极显著, 此类同序性 3' 端在番茄基因组的 RAPD 座位中是普遍分布的。RAPD 引物 3' 端序列的非随机分布, 反映了番茄基因组内的多态性区段侧翼具有序列的特异性, 这种特异性结构应是番茄基因组序列的一种属性, 这种属性为番茄属基因组所共有。

关键词 番茄; RAPD 引物; 3' 端序列; 基因组序列

中图分类号: Q75 文献标识码: A 文章编号: 0253-977X(2001)02-0138-05

A Partial Distribution of the 3' Sequences of Effectual Primers in RAPD Analysis for *Lycopersicon* sp. and with Relation to Sequential Feature of Tomato Genome

YU Dan-nian

(Dept. of Biology, Hainan Normal College, Haikou, Hainan 571158, China)

Abstract The study of compared effectual primers 3' end sequences between RAPD analysis used in tomato and Operon system. It is suggested that 3' end sequences of primers for tomato RAPD are not the random samples of OP system, their 3' end sequences are partial distribution in kinds and frequency. Though there have homologous sequences at 3' end in both systems. But their distribution of frequency shows significant difference in orthogonal comparison. The 3' ending homologous sequences in RAPD loci are found all over in tomato genome. Therefore the nonrandom distribution of 3' end sequences of effectual primers in tomato RAPD reflected the sequence-specific structure may occur in tomato genome at flank of polymorphic regions. It is probably an attribute of tomato genomic sequence and is common in *Lycopersicon*.

Key words tomato; RAPD primer; 3' end sequence; genomic sequence

1991 年, G. B. Martin 首次报道应用 RAPD 技术分析番茄的 DNA 多态性以来, 已有不少涉及番茄 RAPD 的研究报道, 其中大部分选用的是 Operon 公司生产的引物系列(下文简称 OP 系列)^[1~26]。鉴于 RAPD 扩增的关键序列是引物的 3' 端序列, 作者曾

对 OP 系列的 520 种引物 3' 端序列进行统计, 在三核苷酸水平上, OP 系列 3' 端序列的分布不是三核苷酸组合的随机样品^[27], 因此欲知番茄引物 3' 端序列分布是否有其特异性, 还必需以 OP 系列作为参照系进行比较。作者曾自有关番茄 RAPD 文献

收稿日期: 2000-05-12; 修回日期: 2000-09-30

基金项目: 本文获海南省教育厅科研基金项目及省自然科学基金项目资助

作者简介: 余诞年(1938-), 男, 浙江镇海籍, 北京农业大学园艺系 59 届毕业, 教授, 专业: 遗传学及植物育种。电话: 0898-5882246, E-mail: ydn@hainnu.edu.cn

发表的有效引物中,主要选择已知序列的十核苷酸的引物组成样本(下简称番茄系列),将其 3'端序列与 OP 系列引物进行比较。经比较发现,番茄系列的 3'端序列在三核苷酸水平上亦非 OP 系列的随机样本,而是序列种类有偏倚分布的选择性样本,并指出这种番茄引物 3'端序列分布的偏倚性,可能反映了番茄基因组序列结构的某些特征^[27]。通过特异性简单序列的频率分布探讨基因组序列特征的研究,目前尚未见到其他作物上有类似报道。

本文将在四核苷酸水平上进一步进行比较。因为四核苷酸序列已占引物全序列的 40%,其频率分布状况的比较更具有说服力。而且基因组内编码序列侧翼的某些寡核苷酸序列(包括四核苷酸序列),已知是与基因的活性调控、复制、转座、内切(可能包含重组)等功能有关的结构域。同时还发现在卫星 DNA 中也不乏四核苷酸的重复序列,也有一定的遗传功能。虽然目前对此类序列的功能知之尚少,但

如果某些序列种类较之其他序列种类更频繁或更集中地存在于基因组内或基因组的某些特定区域或部位,那么至少在某种程度上反映出这些核苷酸序列在基因组内存在的意义。

1 RAPD 引物 3'端序列的分布和比较

1.1 OP 系列和番茄引物 3'端序列的分布

分别将 OP 系列及番茄 RAPD 系列引物 3'端按四核苷酸不同序列(下简称 3'端序列)分组统计。OP 系列 A-Z 组共 520 种引物,3'端序列中并未出现全部可能的四碱基组合,只有其中的 149 类组合,而且是非均匀分布的(图 1)。番茄引物作为一个样本,来自不同的番茄 RAPD 研究报道,选择其中已在 DNA 多态性分析中应用,并已知序列者,除大部分来自 OP 系列外,其中还有约 25%的引物不属于该系列,共 153 种,但 3'端序列只有 90 类,其分布也是不均匀的(图 2)。

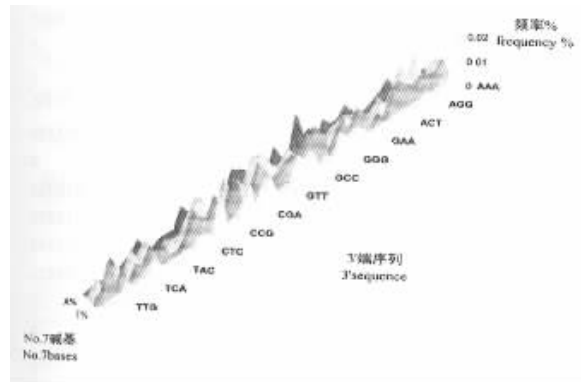


图 1 OP 系列 3'端序列频率
Fig.1 The frequency distribution of 3' end sequences in OP

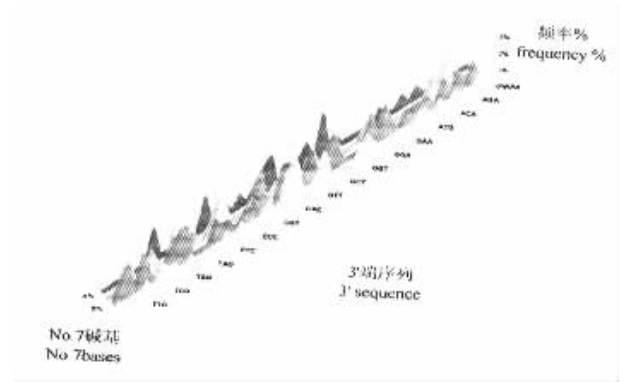


图 2 番茄引物 3'端序列频率
Fig.2 The frequency distribution of 3' end sequences in tomato

图注 3'端序列由右至左依次为

Note 3' end sequences from right to left are

AAA ,AAG ,AAC ,AGA ,AGG ,AGT ,ACA ,ACC ,ACT ,ATG ,ATC ,ATT ,GAA ,GAG ,GAC ,
GGA ,GGG ,GGC ,GGT ,GCA ,GCC ,GCT ,GTG ,GTC ,GTT ,CAA ,CAG ,CAC ,CGA ,CGG ,CGT ,
CCA ,CCG ,CCC ,CCT ,CTG ,CTC ,CTT ,TAA , ,TAG ,TAC ,TGA ,TGG ,TGT ,TCA ,TCG ,TCC ,TCT ,TTG ,TTC ,TTT .

No.7 碱基由上至下依次为(No.7 bases from up to down are)A% , G% , C% , T% .

以每组引物中不同序列所占的比例作为该组各序列的出现频率,并转换为 $\sin^{-1}\sqrt{\%}$ 值。将番茄样本的频率与 OP 系列的频率分布曲线作配合测验(表 1), $\chi^2 = 327.88$ ($df = 148$),估算该自由度下的 $\chi^2_{(0.001)} = 209.44$ ^[28],两个分布间差异极显著。

1.2 引物 3'端序列的同源性现象

OP 系列的 520 种引物只有 149 类末端的四聚序

列,意味着至少平均每 3~4 种引物有一个共同的末端序列,作者拟将此称为引物的“同源性序列”或“同序末端”。OP 系列中有两个以上(含两个)引物共有的同序末端序列共 119 类,合计有 490 种引物,占 3'端序列种类数的 79.9%,引物数的 94.2%。而在番茄样本中,同源性序列共 38 类,101 种引物,分别相当于 OP 系列同序末端的 31.9%和 20.6%,为番茄样本 3'

端序列种类的 42.2% ,引物数的 66.0% ,其中包括五碱基序列相同的 12 类 26 种 ,更有 2 种引物具有六碱基同源性 3' 端序列。作者曾在三核苷酸水平上 ,对番茄样本中同源性序列的分布与 OP 系列同类序列的分布进行过正交比较^[27,29] ,以相同的方法 ,在四核苷酸

水平上对同源性序列的分布进行比较 ,结果同样在二者间存在显著差异(表 2) ,说明其同源性序列的分布也不是 OP 系列的随机样本 ,相比之下其同源性序列在种类上更为集中 ,这种分布的偏倚性 ,显然取决于番茄基因组的序列结构。

表 1 番茄系列 OP 系列引物 3' 端序列频率的成对比较 (频率值已作 $\text{Sin}^{-1} \sqrt{\%}$ 的变换)

Table 1 The pairing differences of frequency of 3' sequences compared between primers of tomato and OP systems (arc-sin transformation).

No. 8 ~ 10 碱基 No. 8 ~ 10bases	No. 7 碱基 No. 7base				No. 8 ~ 10 碱基 No. 8 ~ 10bases	No. 7 碱基 No. 7base			
	A%	G%	C%	T%		A%	G%	C%	T%
AAA	0	0	2.513392	0	CAG	0	0.868697	0.575006	2.403485
AAG	0	5.031637	0.156272	0.01813	CAC	1.810007	0.868697	0.119945	1.042518
AAC	2.513392	4.356122	4.356122	2.513392	CGA	0.156272	4.356122	4.356122	0
AGA	0	0	1.120187	2.513392	CGG	0.379272	5.627356	0	0.615786
AGG	2.513392	0.667009	5.031637	0.328978	CGT	0	0.328978	1.047923	0.328978
AGT	0	0.030928	1.120187	0	CCA	0.328978	7.125016	0.379272	2.513392
ACA	0	0.043998	5.031637	0	CCC	0.328978	3.555615	6.662684	0.868697
ACC	0.030928	5.031637	0.379272	2.513392	CCT	1.794526	0.025774	0.01813	2.513392
ACT	3.555615	0	0.17424	1.794526	CTG	4.356122	0.288648	0.030928	0.328978
ATG	1.794526	4.356122	0	1.794526	CTC	1.042518	0.030928	0.001429	3.555615
ATC	0	2.513392	0.328978	0	CTT	2.513392	2.513392	3.555615	0
GAA	0.328978	4.356122	0.030928	2.513392	TAA	2.513392	1.794526	0	0
GAG	5.031637	4.356122	0.575006	4.356122	TAG	1.794526	3.555615	0	0
GAC	3.555615	0.030928	0.328978	5.031637	TAC	0	0	3.131408	1.794526
GGA	3.555615	0.01813	0.575006	0.01813	TGA	0	3.131408	3.131408	0
GGG	5.031637	4.356122	4.356122	0.156272	TGG	5.031637	5.627356	0.156272	0.030928
GGT	0.030928	2.513392	6.166452	3.555615	TGT	5.679665	3.555615	0.328978	0.328978
GCA	4.356122	0.156272	0.868697	0	TCA	2.513392	0.156272	4.356122	2.547272
GCC	0.402962	0	0.288648	5.627356	TCC	5.031637	5.031637	1.810007	0.01813
GCT	0	4.356122	1.120187	5.031637	TCT	0	2.547272	2.513392	0
GTG	2.547272	0.156272	5.627356	0.575006	TTG	2.513392	2.513392	0.119945	0
GTC	0.17424	0.328978	1.810007	1.794526	TTC	0	5.031637	2.547272	2.513392
GTT	0	4.356122	0.01813	2.513392	TTT	0	0	2.513392	0
CAA	0	0.379272	3.555615	1.794526					

$$\chi^2 = 327.8755, df = 148, P < 0.001$$

注 :表中数值为 (the values is) $\chi^2 = (\text{TOM} - \text{OP})^2 / \text{OP}$,其中 TOM 为番茄变换值 ,OP 为 OP 系列变换值 (TOM :tomato ,OP :OP system)

2 番茄 RAPD 标记在染色体上的分布

归纳文献中已定位的标记 ,将其中属于 OP 系列的引物进行分析 ,这类引物共 84 种 ,并散布于番茄的全部染色体上 ,但分布不均匀 ,标记最多的是染色体 9 ,达 22 个 ;最少的是染色体 8 和 12 各只有 2 个 ,二者竟相差 11 倍。有 10 种引物可同时标记两条不同的染色体 ,其中有 1 种引物 ,甚至可用于标记 3 条染色体 ,说明不同染色体上存在着可与同一种引物配对的同源序列 ,此外还有 6 种属于 3' 端同源的引

物。除染色体 12 外 ,在俩俩染色体间 ,都可找到包括上述两种同源性系列 (相同引物和同源性 3' 端引物) 在内的 ,同源 3' 端序列 (引物) 的标记座位 ,这类座位共有 44 个 ,包含 39 种引物。其分布也并不均匀 ,同一染色体上最多的可达 11 个 ,最少也有一个 ,平均占标记总数的 46.4% ,占 RAPD 标记座位总数的 46.32% ,覆盖了除 12 号染色体外的所有染色体 (表 3) ,这也说明同源性序列并非个别染色体的序列结构特点 ,而是基本上覆盖了整个番茄基因组的一种普遍性序列结构 ,其种类和分布也不是随机的。

表 2 番茄和 OP 系列 3'端同源末端序列分布正交比较

Table 2 The orthogonal comparison of frequency of primer 3' end homologous sequences between tomato and OP systems

方差分析 analysis of variance :	<i>df</i>	<i>SS_q</i>	<i>MS_q</i>	<i>F</i>	
序列间 sequences	19	96.68487	5.088677		
处理(样本)间 treatment(samples)	4	43.6218	10.90545	6.31793 * *	
剩余 error	76	131.1845	1.726111		
总计 total	99	271.4911			
比较 comparisons	番茄系列 tomato 番茄(OP) OP in tomato	OP 系列 OP system	随机样本 1 sample1	随机样本 2 sample2	
总和 sum	155.6212	160.0432	123.766	137.5991	138.1459
对应的% corresponding%	0.372549	0.421053	0.236538	0.3	0.3
$\lambda_1 =$	3	3	-2	-2	-2
$\lambda_2 =$	1	-1	0	0	0
$\lambda_3 =$	0	0	2	-1	-1
$\lambda_4 =$	0	0	0	1	-1
析因分析 Factorial analysis :	<i>df</i>	<i>SS_q</i>	<i>MS_q</i>	<i>F</i>	
番茄与 OP, tomato vs. OP	1	36.49235	36.49235	12.60502 * *	
番茄内 within tomato	1	0.488861	0.488861		
OP 与样本 OP vs. sample	1	6.633114	6.633114		
样本内 within sample	1	0.007475	0.007475		
处理合计, sum of treatment	4	43.6218	10.90545		
误差, error	95	275.0312	2.895065		
总计, total	99	318.653			

表 3 RAPD 标记在番茄染色体上的分布

Table 3 The distribution of RAPD markers on tomato chromosomes

染色体 No. Chromosome No.	1	2	3	4	5	6	7	8	9	10	11	12
RAPD 标记数 Number of RAPD markers	10	13	8	6	5	13	6	2	22	4	4	2
具同源末端的标记数 Number of homologous 3' end	3	4	4	3	3	7	2	1	11	3	3	0
占标记数的% ratios %	30.0	30.8	50.0	50.0	60.0	53.8	33.3	50.0	50.0	75.0	75.0	0.0

3 讨论与结论

番茄属内以及番茄与马铃薯分子作图的比较中发现,它们都具有相同的分子标记座位的排列顺序,但番茄与马铃薯间有 5 个整臂倒位,在番茄种间也发现有小区段的基因座倒排现象,而且基因座之间的图距有变化^[9,30,31],作者认为番茄种间杂种的遗

传分析中,基因座之间图距的相对集中,有可能是倒位抑制效应所致。从番茄种间染色体同源性较高这一点来看,番茄种间的多态性更可能表现为小片段的倒位方面,因此不能排除这类片段的侧翼存在上述引物末端序列的可能性。

番茄 RAPD 引物 3'端序列是一组非随机的选择性样本。引物 3'端序列的非随机分布是番茄基因组

内多态性区段侧翼有特异性结构的反映。

番茄 RAPD 引物同源性末端的存在,并基本覆盖番茄整个基因组,可以认为特异性结构是番茄基因组序列结构的一种属性。这种属性也是番茄属基因组所共有的特性。

参考文献 (References):

- [1] 余诞年. 番茄基因的分子标记与遗传作图—文献综述 [J]. 园艺学报, 1998, 25 (4) : 361 ~ 366.
- [2] 恩和巴雅尔等. 随机引物在番茄 DNA 扩增及多态性分析中的应用 [J]. 内蒙古师大学报 (自然科学, 汉文版), 1998, 27 (3) : 223 ~ 226.
- [3] 恩和巴雅尔等. 几种番茄品种基因组 DNA 的相似指数初步分析 [J]. 内蒙古师大学报 (自然科学, 汉文版), 1998, 27 (4) : 307 ~ 309.
- [4] Martin G B, et al. Rapid identification of markers linked to a *Pseudomonas* resistance gene in tomato by using random primers and near-isogenic lines [J]. Proc Natl Acad Sci USA, 1991, 88 : 2336 ~ 2340.
- [5] Klein-Lankhorst R M, et al. Isolation of molecular markers for tomato (*L. esculentum*) using random amplified polymorphic DNA [J]. TAG, 1991, 83 : 108 ~ 114.
- [6] Williams C E, St Clair D A. Phenetic relationships and levels of variability detected by restriction fragment length polymorphism and random amplified polymorphic DNA analysis of cultivated and wild accessions of *Lycopersicon esculentum* [J]. Genome, 1993, 36 : 619 ~ 630.
- [7] Foolad M R, et al. RAPD markers for constructing intraspecific tomato genetic map [J]. Pl Cell Rept, 1993, 12 : 293 ~ 297.
- [8] Carland F M, Staskiewicz B J. Genetic characterization of Pto locus of tomato: Semi-dominance and cosegregation of resistance to *Pseudomonas syringae* pathovar tomato and sensitivity to the insecticide Fenitrothion [J]. Mol Gen Genet, 1993, 239 : 17 ~ 27.
- [9] Weide R, et al. Integration of the classical and molecular linkage map of tomato chromosome 6 [J]. Genet, 1993, 135 : 1175 ~ 1186.
- [10] Dickinson M J, et al. Close linkage between the *Cf-2/Cf-5* and *Mi* resistance loci in tomato [J]. MPMI, 1993, 3 (3) : 341 ~ 347.
- [11] Dax E, et al. A random amplified polymorphic DNA (RAPD) molecular marker for the *Tm-2^a* gene in tomato [J]. Euphytica, 1994, 74 : 159 ~ 163.
- [12] Wing R A, et al. Map-based cloning in crop plants. Tomato as a model system: Genetic and physical mapping of jointless [J]. Mol Gen Genet, 1994, 242 : 681 ~ 688.
- [13] Hossain M, et al. Production of somatic hybrids between tomato (*Lycopersicon esculentum*) and night shade (*Solanum lycopersicoides*) by electrofusion [J]. Breeding Sci, 1994, 44 : 405 ~ 412.
- [14] van der Beek J G, et al. Resistance to powdery mildew (*Oidium lycopersicum*) in *Lycopersicon hirsutum* controlled by an incompletely-dominant gene *Ol-1* on chromosome 6 [J]. TAG, 1994, 89 : 467 ~ 473.
- [15] van Wordragen M F, et al. Genetic and molecular organization of the short arm and pericentromeric region of tomato chromosome 6 [J]. Euphytica, 1994, 79 : 169 ~ 174.
- [16] Chunwongse J, et al. Chromosomal localization and molecular-marker tagging of powdery mildew resistance gene (*Lv*) in tomato [J]. TAG, 1994, 89 : 76 ~ 79.
- [17] Yaghoobi J, et al. Mapping a new nematode resistance locus in *Lycopersicon peruvianum* [J]. TAG, 1995, 91 : 457 ~ 464.
- [18] Stevens M R, et al. Mapping the *Sw-5* locus for tomato spotted wilt virus resistance in tomatoes using RAPD and RFLP analysis [J]. TAG, 1995, 90 : 451 ~ 456.
- [19] Ohmori T, et al. Identification of RAPD markers linked to the *Tm-2* locus in tomato [J]. TAG, 1995, 90 : 307 ~ 311.
- [20] Ohmori T, et al. Molecular characterization of RAPD and SCAR markers linked to the *Tm-1* locus in tomato [J]. TAG, 1996, 92 : 151 ~ 156.
- [21] Grandillo S, Tanksley S D. Genetic analysis of RFLPs GATA microsatellites and RAPDs in across between *L. esculentum* and *L. pimpinellifolium* [J]. TAG, 1996, 92 : 957 ~ 965.
- [22] Pillen K, et al. Construction of a high-resolution genetic map and YAC-contig in the tomato *Tm-2^a* region [J]. TAG, 1996, 93 : 228 ~ 233.
- [23] Chague V, et al. Identification and mapping on chromosome 9 of RAPD markers linked to *Sw-5* in tomato by bulked segregant analysis. [J] TAG, 1996, 92 : 1045 ~ 1051.
- [24] Samoylov V M, Sink K C. The role of irradiation dose and DNA content of somatic hybrid calli in producing asymmetric plants between an interspecific tomato hybrid and eggplant [J]. TAG, 1996, 92 : 850 ~ 857.
- [25] Kobayashi R S, et al. Somatic hybridization between *Solanum ochroanthum* and *Lycopersicon esculentum* [J]. Pl Cell Tiss & Organ Cult, 1996, 45 : 73 ~ 78.
- [26] Bogani P, et al. Genome flux in tomato cell clones cultured in vitro in different physiological equilibria. II. A RAPD analysis of variability [J]. Genom, 1996, 39 : 846 ~ 853.
- [27] 余诞年. 番茄 (*Lycopersicon sp.*) RAPD 分析引物序列的特异性 [J]. 海南师范学院学报, 1998, 11 (2) : 88 ~ 95.
- [28] 杨纪珂. 数理统计方法—在医学科学中的应用 [M]. 上海: 上海科技出版社, 1964, 124.
- [29] G. W. 斯奈迪格等 (杨纪珂, 汪安琦译). 应用于农学和生物学实验的数理统计方法 [M]. 北京: 科学出版社, 1963, 368 ~ 372.
- [30] 余诞年, 吴定华, 陈竹君. 番茄遗传学 [M]. 长沙: 湖南科技出版社, 1999, 133 ~ 150.
- [31] Tanksly S D, et al. High density molecular linkage maps of the tomato and potato genomes [J]. Genet, 1992, 132 : 1141 ~ 1160.