

水稻黑条矮缩病毒基因组片段 S7 的 cDNA 克隆 及全序列分析*

张恒木^{1, 2} 陈剑平^{1, **} 雷娟利¹ 程 晔¹ 薛庆中²

(¹ 浙江省农业科学院病毒学实验室 杭州 310021)

(² 浙江大学农业与生物技术学院 杭州 310029)

摘 要 :应用 RT-PCR 技术克隆了 2 个水稻黑条矮缩病毒(rice black-streaked dwarf virus ,RBS-DV)中国分离物 ,即浙江分离物和河北分离物的基因组片段 S7 ,并测定了他们的全序列。结果表明 :RBSDV 浙江分离物(RBSDV-Zj)基因组片段 S7 全长 2193nts(EMBL 登录号为 AJ297427) ,RBSDV 河北分离物基因组片段 S7 全长 2190nts(EMBL 登录号为 AJ297428) ,二者均含有两个开放阅读框(open reading frame ,ORF) ,分别编码约 41kD 和 36kD 多肽。2 个中国分离物核苷酸同源性高达 99% ,相应的 ORF 编码的多肽同源性分别为 100% 和 94.4% ,与日本 RBSDV 基因组片段 S7 核苷酸同源性为 93.4% 和 93.8% ,相应 ORF 编码的多肽同源性分别为 98.1%(ORF1) 96.5% 和 97.8%(ORF2) ,与意大利 MRDV S6 核苷酸同源性为 85.1% 和 85.3% ,相应多肽同源性分别为 92.3%(ORF1) 85.5% 和 86.8%(ORF2)。

关键词 :水稻黑条矮缩病毒 ,基因组片段 S7 ,序列分析

中图分类号 S432.41 文献标识码 :A 文章编号 0001-6209(2002)02-0200-08

水稻黑条矮缩病毒(rice black-streaked dwarf virus ,RBSDV)是一种特殊的多组分双链 RNA(double RNA ,dsRNA)病毒 ,隶属呼肠孤病毒科斐济病毒属 ,主要分布于中国、朝鲜、韩国和日本等东亚国家和地区 ,对水稻、大麦、小麦、玉米等禾谷类作物造成危害。自 80 年代以来 ,本实验室对其流行学和生物学特性进行了深入的调查和研究^[1~4]。近期的研究还表明我国北方发生严重的玉米粗缩病和小麦绿矮病都是由 RBSDV 引起的^[5]。RBSDV 病毒粒子呈正二十面体 ,双层外壳均含有突起(spike) ,内含 10 条 dsRNA 基因组片段 ,按其在凝胶上的迁移率由慢到快依次称为 S1 ~ S10。此病在水稻上表现的主要症状是病株明显矮缩 ,叶色浓绿 ,叶脉上出现肿瘤状突起。侵染不同的作物 ,其症状略有不同。而且该病毒不仅能在植物体内增殖 ,而且能在介体昆虫中增殖 ,但介体昆虫并不表现症状^[6]。

迄今国内尚未发现抗黑条矮缩病毒的种质资源 ,主要是通过喷洒杀虫剂来防治 ,然而利用病毒自身的基因抗病毒已有许多成功的先例。因此 ,加快 RBSDV 的基因组研究 ,不仅有利于有效控制水稻黑条矮缩病的发生和流行 ,并且有利于进一步研究病毒基因组的功能、致病性、转录、复制、组装、以及病毒-介体-宿主三者之间基因互作机制。本文旨在通过对 2 个 RBSDV 中国分离物基因组片段 S7 的比较研究来了解其基因组成及编码特性 ,为水稻抗病毒育种及其功能分析奠定分子基础。

* 浙江省青年科技人才专项基金(RC9604)和浙江省重点实验室专项资金资助

** 通讯作者 ,E-mail :jpchen@mail.hz.zj.cn 电话 0571-86404242

作者简介 张恒木(1970 -)男 ,安徽滁州人 ,浙江大学博士研究生 ,从事分子生物学研究。

收稿日期 2001-02-20 ,修回日期 2001-04-04

1 材料和方法

1.1 材料

2 个 RBSDV 中国分离物分别采自浙江省兰溪市感病的水稻(1998 年 7 月)和河北保定市感病的小麦(1999 年 5 月),采集的病株保存在 - 80℃ 备用。

1.2 方法

1.2.1 病毒基因组提纯 :参照 Uyeda 等^[7]方法分别提取 2 个 RBSDV 分离物的 dsRNA 总基因组 ,并利用低熔点琼脂糖凝胶法分离基因组片段 S7。

1.2.2 dsRNA 的变性 :分别取纯化的 2 个病毒分离物的基因组片段 S7 约 2μg 溶于 25μL 无菌水 ,加入 25μL DMSO(dimethyl sulfoxide)中 ,95℃ 加热 5min ,然后冰浴 10min ,再用 3 倍体积乙醇和 1/10 体积乙酸钠(3mol/L ,pH5.2)沉淀 ,备用。

1.2.3 基因组片段 S7 全序列克隆 (1)中间序列的扩增 :首先根据已报道的 RBSDV 日本分离物的序列设计引物 S7p-1 和 S7p-2(表 1),并利用这些引物合成 cDNA ,37℃ 保温 60min ,cDNA 合成后 ,加 1μL RNaseH(2U/μL),再通过 PCR 纯化 Kit(QIAGEN)去除引物和小片段 cDNA ,并以此纯化的 cDNA 为模板进行 PCR 反应扩增基因组片段 S7 的中间序列 ,PCR 扩增反应条件均为 94℃ 变性 30s ,55℃ 退火 30s ,72℃ 延伸 1min ,经 30 次循环后 72℃ 再延伸 10min。(2)两端序列扩增 :根据自己测定的中间序列进一步设计引物 S7p-3 和 S7p-4 (表 1),为了获得完整的末端序列 ,首先在变性的 S7 片段的两条 RNA 链的两个 3'-端连接上引物 zhm-1(表 1)。为了防止引物 zhm-1 在随后的 dsRNA/DNA 连接反应中自身环化 ,引物 zhm-1 的 3'-端被氨基化。20μL 的连接反应体系中包括 :10U T4RNA ligase(Biolab) 2μg dsRNA 200 pmol 引物 zhm-1 50mmol/L Tris-HCl ,pH7.8 ,10mmol/L MgCl₂ ,10mmol/L dithiothreitol ,1mmol/L ATP ,10μg/mL BSA 25% PEG 8000 和 10U RNase 抑制剂(Gibico) ,然后 37℃ 温育 15min ,RNA/zhm-1 连接产物用 QIAGEN RNeasy plant minikit 纯化除去未连接的引物 zhm-1 分子。然后以 zhm-2(表 1 ,与引物 zhm-1 序列互补)作引物 ,用 Superscript Revers Transcriptase(MBI)将 RNA/zhm-1 连接产物逆转录成 cDNA(方法按照商家的操作手册)。42℃ 温育 60min 后 ,加热到 70℃ 15min ,终止反应 ,然后置于冰上 3min ,加 1μL RNase H (2U/μL) ,37℃ 温育 20min。再用 PCR Gel Extraction Kit (QIAGEN)纯化获得的 cDNA 用作 PCR 扩增模板 ,引物 S7p-3 和 zhm-2、S7p-4 和 zhm-2 分别用来扩增 5'-端和 3'-端序列。PCR 扩增均为 94℃ 变性 30s ,60℃ 退火 30s ,72℃ 延伸 1min ,经 30 次循环后 72℃ 延伸 10min。

表 1 用于扩增 RBSDV S7 的引物

Table 1 Primers used for amplification of RBSDV genome segments S7

Primer	Position	Sequence (5' ~ 3')	Tm/℃	Direction
S7p-1	nt922-941	GAG CTC TTC TAG TTA TTG CG	58	+
S7p-2	nt1468-1450	TGT CAC ACC ACT CTT CTC C	58	-
S7p-3	nt1135-1116	CAA GAT TAA GCA GAA GGA GA	56	-
S7p-4	nt1183-1204	ATG AAT TAC ACT TTA GGT GAT C	58	+
zhm-1		PO ₄ -CTC TTC CCC TCC CTC CTC-NH ₂	60	
zhm-2		GAG GAG GGA GGG GAA GAG	60	

1.2.4 PCR 扩增产物的克隆和测序: PCR 扩增的预期片段通过 PCR Gel Extraction Kit (QIAGEN) 纯化, 并与 pGEM T-vector (Promega) 进行连接, 连接产物转化 TG1 感受态细胞, 利用氨苄抗生素、 α -互补和电泳法筛选重组子, 重组质粒再经 PCR 扩增印证。用 QIAprep spin mini prep kit (QIAGEN) 提纯质粒, 用 BigDye™ Terminator Cycle Sequencing Ready Reaction Kit (PERKIN ELMER) 试剂以及引物 T7 和 SP6 在 ABI PRISM™ 377 DNA Sequencer 上进行序列测定, 方法均按照商家说明。应用 GCG 程序软件进行序列分析^[8]。

2 结果

2.1 RBSDV S7 中间序列 cDNA 的合成和 PCR 扩增

根据发表的 RBSDV 日本分离物 S7 的全序列设计了中间序列的两个引物, 分别以纯化的 2 个 RBSDV 分离物的 dsRNA 基因组片段 S7 变性后的 ssRNA 为模板, 与这两个引物退火, 合成 cDNA, 再按所设计的反应条件进行 PCR 扩增, 扩增产物经 1.2% 琼脂糖凝胶电泳检测均为单一条带, 约 0.55kb, 与预期的片段大小一致(图 1)。



图 1 2 个中国 RBSDV 分离物基因组片段 S7 RT-PCR 扩增产物电泳结果

Fig.1 RT-PCR products of genome segment S7 of two Chinese

olates examined on a 1.2% agarose gel stained with ethidium bromide
M. Marker DNA; (from up to down: 2kb, 1kb, 0.75kb, 0.5kb, 0.25kb, 0.1kb)

- 1: Internal region of RBSDV-Zj genome segment S7;
- 2: 5' terminal region of RBSDV-Zj genome segment S7;
- 3: 3' terminal region of RBSDV-Zj genome segment S7;
- 4: Internal region of RBSDV-Heb genome segment S7;
- 5: 5' terminal region of RBSDV-Heb genome segment S7;
- 6: 3' terminal region of RBSDV-Heb genome segment S7.

2.2 RBSDV S7 两端序列 cDNA 的合成和 PCR 扩增

为了得到 S7 的全长, 首先将其变性, 在其正义链和反义链的 3'端分别连上寡核苷酸 zhm-1, 然后利用与 zhm-1 互补的寡核苷酸 zhm-2 作引物引导第一链 cDNA 的合成, 以此 cDNA 为模板, 再按设计的 PCR 反应条件, 利用引物 S7p-3 和 zhm-2、S7p-4 和 zhm-2 分别扩增基因组 S7 的 5'和 3'端的序列, 扩增产物经 1.2% 琼脂糖凝胶电泳检测均出现单一条带, 分别约 1.15kb 和 1.03kb, 与预期的片段大小一致(图 1)。

2.3 扩增产物的克隆和序列分析

测定的 2 个中国 RBSDV 分离物基因组片段 S7 序列均已登录 EMBL 数据库, 登录号分别为 AJ297427(RBSDV-Zj S7)和 AJ297428(RBSDV-Heb S7)。RBSDV 浙江分离物基因组片段 S7 全长 2193ns, 而 RBSDV 河北分离物基因组片段 S7 相对于浙江分离物 S7 而言, 在 nt523-525 缺失了 3 个碱基, 全长 2190ns。测序结果表明: 2

个中国分离物的保守末端高度一致, 紧接着寡核苷酸 zhm-2/zhm-1 的分别是 5'端保守序列 "AAGTTTTTTT——" 和 3'端保守序列 "——GTC", RBSDV-Heb 存在一个完整的倒转重复序列(inverted repeats)紧挨着末段保守序列, 分别位于 nt9-17(TTTCGACCT)和 nt2169-2177

FAELLDQEI~~GE~~EWCDTSNFM~~DL~~RIVEDENDFEFVSSHITRHL~~LIV~~LSNPNVLTSTCLLAKLSLIQHVN~~NF~~DVINYWEA
.....D...EIH...A...N...E.....F.....I.....E.....
MNR~~WE~~ITDELKMGFVFR~~AF~~NLKS~~NQ~~FEVVTKLLSDSLFC~~PG~~ISVIGKLSM~~IP~~MLTVHSIPEYLDHWFTDD~~FS~~SDNEL
.....L.....L.....I.....G.....II.....IM.....G.....LY.....N.....V.....F.....A.....QRCK..
SFIRFGEITVPK~~WK~~KVVVQFYLRQVFSRV~~RT~~KVLIANTD~~VD~~YWYSLFM~~Q~~TLIFESML~~RT~~TKTLIKKILNS RBSDV-Zj S7 ORF2
.....R.....K.....R.....K.....N.....V.....RBSDV-Heb S7 ORF2
.....R.....V.....K.....S.....NM.....RBSDV-Jap S7 ORF2
.....MRDV S6 ORF2

图 2 4 个病毒分离物相应基因组片段全序列及其编码的多肽同源性分析

Fig.2 Comparison of complete nucleotide sequences and deduced amino acid sequences among the corresponding genome segments of four virus isolates
a.Comparison of complete nucleotide sequences;b.Comparison of amino acids sequences encoded by ORF1;c.Comparison of amino acids sequences encoded by ORF2.

表 2 4 个不同来源的病毒分离物相应基因组片段的同源性比较

Table 2 Homologies (% identical nucleotides,with % identical amino acids of the open reading frames in brackets) among four isolates of RBSDV and MRDV

S7	Hebei	Japan	MRDV-S6
Zhejiang	99.0(100.0,98.4)	93.4(98.1,97.8)	85.1(92.3,85.5)
Hebei		93.8(98.1,96.5)	85.3(92.3,86.8)
Japan			85.1(91.2,86.1)

表 3 4 个不同来源的病毒分离物相应基因组片段的组织结构

Table 3 Organization of the corresponding genome segments of four virus isolates

	Total size nts	G + C content %	5'-UTR nts	ORF1		Intergenic region nts	ORF2		3'-UTR nts
				nts	kD		nts	kD	
Zhejiang S7	2193	34.3	41	1089	41.0	52	930	36.3	81
Hebei S7	2190	34.4	41	1086	40.9	52	930	36.3	81
Japanese S7	2193	34.0	41	1089	41.1	52	930	36.4	81
MRDV S6	2193	34.7	41	1089	41.0	52	930	36.3	81

Notes:Sequence accession numbers were:Zhejiang S7 (AJ297427),Hebei S7 (AJ297428),Japanese S7 (S63917),MRDV S6 (X55701).

3 讨论

研究病毒基因组末端序列能为病毒复制及其基因的表达提供重要信息,由于 RBSDV 基因组的 3'-端没有 poly(A),不能按经典的方法用 oligo(dT)引导 cDNA 的合成以及克隆末端序列,加 poly(A/T/G/C)的效率较低又不能一次性有效地确定末端的碱基,作者利用了 RBSDV 的基因组是双链 RNA 性质,采取直接加特异性寡核苷酸的方法,一次连接即确定 RBSDV 基因组片段的两个末端的序列,方法简便有效。研究还表明 2 个 RBSDV 分离物 S7 的末端序列以及靠近末端的倒转重复序列和 Azuhata 等^[9]用 RBSDV-日本分离物所得结果基本一致,它们的保守性暗示 RBSDV 基因组末端序列具有重要的作用,在已知的斐济病毒中都含有类似的结构,且绝对保守,这说明在功能上起重要作用的残基在进化上是高度保守的,体现了结构和功能的一致性。有意义的是,在 RBSDV-Zj S7 除了 nt9-17 和

nt2172-2180 互为倒转重复序列外,在靠近末端处还存在另外 1 对倒转重复序列 nt17-22 (TGTCTTG) 和 nt2159-2164 (CAGACA),这 2 对末端倒转重复序列,在其他的 3 个病毒分离物中也存在,且均位于非编码区,二级结构预测表明它们正好形成发夹结构(资料未显示),暗示了两个末端在病毒基因组的转录、复制、表达过程中存在着复杂的相互作用。

2 个中国分离物与 RBSDV-日本分离物 S7、意大利 MRDV S6 的全序列相比,在碱基点突变中,其中大多数是同类碱基之间的转换(即 A/G, C/T)(图 2a),这一特性可能与 dsRNA 复制有关,因为依赖 RNA 的 RNA 聚合酶(RNA-dependent RNA polymerase, RdRp)没有 3'—5'的外切酶活性,便不能在复制过程中校正错误配对的碱基,而在 RNA 分子中,除了 A/U, G/C 之间的正常配对外, G/U 之间也可能形成较弱的氢键配对,这样在转录和复制过程中发生错误配对的碱基对中 G/U 占多数,从而导致了 G/A, C/U 的突变较多,这种突变究竟对病毒的进化有何意义尚待探讨。另外我们发现较多的突变发生在密码子的第 3 位碱基上,因而多数情况下不引起氨基酸序列的变化,使病毒在进化过程中存在保持相对稳定性。

对 RBSDV-Zj S7 的两个 ORF 作密码子偏爱性分析,发现在第一个 ORF 中密码子 XYA/XYT 的使用频率高达 75%,第二个 ORF 中,密码子 XYT/XYG 的使用频率也高达 73%,而 RBSDV-Zj S7 的 A/T 含量为 66%,在 RBSDV-Heb S7 中也发现类似的密码子使用特性。而在其他水稻病毒中,如 RDV(rice dwarf virus)基因组的 A + T 含量(58%)以及利用密码子 XYA + XYT 的频率(57%)基本持平^[10],这体现了两种水稻病毒在分子水平上的差异。

通过 4 个不同来源的病毒基因组片段序列比较可以看出,它们在基因组结构上非常接近,其核苷酸和氨基酸水平上的同源性也较高,根据 RBSDV S7 两个 ORF 及 MRDV、OSDV、NLRV 相应的同源基因组片段的序列做进化树分析可以看出(资料未显示),3 个 RBSDV 分离物和意大利 MRDV 应该是同一种病毒的不同地理小种^[11]。

RBSDV S7 第一个 ORF 编码的蛋白质已证实为病毒的非结构蛋白^[12],第二个 ORF 编码的蛋白质在病株和感染的介体昆虫体内尚未检测到它的存在,对它们的功能也一无所知。在 RBSDV 侵染的病株中, Northern 杂交仅检测到全长的 S7 mRNA,也就是说病株中不存在亚基因组 S7 mRNA(subgenomic RNA),第二个 ORF 究竟是如何表达的,尚需要进一步研究,但是 ORF2 在 4 个不同来源的病毒分离物中是保守的,暗示了它具有重要的功能,没有检测到可能是由于其表达量极低以至于难以检测的缘故。根据真核细胞翻译双顺反子 mRNA(dicistronic mRNA)规律,第二个 ORF 可能通过重新启动机制(reinitiation mechanism)^[13]或内部核糖体进入模式(the internal ribosome entry model)^[14,15]。而无论采用哪一种翻译方式,双顺反子的第二个 ORF 表达量都是极低的,尤其在植物中^[16,17]。从起始密码子的上下游序列来看,2 个中国分离物和 RBSDV-日本分离物基因组 S7 以及意大利 MRDV S6 的第一个 ORF 的 -3 和 +4 位置均分别是 A 和 G,与真核细胞核糖体翻译起始的最适序列一致,表明了病毒对寄主的适应性;而第二个 ORF 的起始密码子上下游序列不符合真核细胞核糖体翻译起始的最适序列,可能也是其表达量极低而难以检测到的原因之一。

参 考 文 献

- [2] 阮义理 陈声祥 林瑞芬,等. 浙江农业科学,1984 **4** :185 ~ 192.
- [3] 陈声祥. 浙江农业科学,1996 **1** :41 ~ 42.
- [4] 陈声祥 余舰斌 秦文胜,等. 中国病毒学,1993 **8** (4) 374 ~ 378.
- [5] Zhang H M ,Lei J L ,Chen J P. A dwarf disease on rice ,wheat and maize in different region of China is caused by rice black-streaked dwarf virus (RBSDV) In Zhou G H ,Li H F ed. Proceedings of the First Asian Conference on Plant Pathology . Beijing :China Agricultural Sciencetech Press ,2000. 155.
- [6] Boccardo G ,Milne R G. Plant reovirus group. In :Harrison B D ,Murrent A F ed. CMI/AAB ,Descriptions of Plant Viruses. Wellesbourne ,Warwick :Association of Applied Biologists ,1984. 294.
- [7] Uyeda I ,Lee B C ,Ando Y. Reovirus isolation and RNA extraction. In :Foster G D ,Taylor S C ed. Methods in Molecular Biology ,Plant virology protocols :From virus isolation to transgenic resistance. Totowa ,New Jersey :Humana Press ,1998 ,**81** :65 ~ 75.
- [8] Genetics Computer Group Wisconsin Package version 10.0.575 Science Drive ,Madison ,Wisconsin ,USA 53711. 1998.
- [9] Azuhata F ,Uyeda I ,Kimura I ,et al . *J Gen Virol* ,1993 **74** :1227 ~ 1232.
- [10] 肖 锦 李 毅 张 净,等. 微生物学报. 1998 **38** (5) 348 ~ 358.
- [11] Marzachi C ,Boccardo G ,Milne R ,et al . *Sem Virol* ,1995 **6** :103 ~ 108.
- [12] Isogai M ,Uyeda I ,Lee B C. *J Gen Virol* ,1998 **79** :1487 ~ 1494.
- [13] Kozak M. *Molecular and Cellular Biology* ,1987 **7** :3438 ~ 3435.
- [14] Jang S K ,Krausslich H G ,Nicklin M ,et al . *Journal of Virology* ,1988 **62** :2636 ~ 2643.
- [15] Pelletier J ,Sonenberg N. *Nature* .1988 **334** :320 ~ 325.
- [16] Kaufman R ,Murtha P ,Davies M. *EMBO Journal* ,1987 **6** :187 ~ 193.
- [17] Angenon G ,Votile J ,Kurkela S A ,et al . *Molecular and Cellular Biology* ,1989 **9** :5676 ~ 5684.

cDNA Cloning and Sequence Analysis of Genome Segment S7 of Rice Black-streaked Dwarf Virus

Zhang Hengmu^{1 2} Chen Jianping¹ Lei Juanli¹ Cheng Ye¹ Xue Qingzhong²

(¹ College of Agriculture and Biotechnology ,Zhejiang University ,Hangzhou 310029 ,China)

(² Virology Laboratory ,Zhejiang Academy of Agricultural Sciences ,Hangzhou 310021 ,China)

Abstract : Genome segments 7 of zhejiang and Hebei isolates of rice black-streaked dwarf virus (RBSDV) were amplified and sequenced. Segment 7 of Zhejiang isolate was consisted of 2193 nts (EMBL accession no. AJ297427) in length and that of Hebei isolate was 2190 nts (AJ297428). Both segments contained two non-overlapping open reading frames (ORFs), which encoded two polypeptides with molecular weights of 41 kD and 36 kD. These two segments shared 99% nucleotide identity , 100% and 94.4% amino acid identities of ORF1 and ORF2 , shared 93.5% and 93.8% identities at nucleotide level , 98.1% (ORF1) and 96.5%/97.8% (ORF2) at amino acid level with S7 of Japanese RBSDV , and shared 85.1% and 85.3% identities at nucleotide level , 92.3% (ORF1) , 85.5%/86.8% (ORF2) at amino acid level with S6 of Italian MRDV.

Key words : RBSDV , Genome segment 7 , Sequence analysis