GENOMIC ORGANIZATION OF ODONTOGLOSSUM
RINGSPOT VIRUS (Cy-1 STRAIN) RNA AND
COMPARISON WITH THAT
OF KOREAN STRAIN

Masato IKEGAMI* and Narinobu INOUEY

The complete nucleotide sequence of the genomic RNA of odontoglossum ringspot virus Cy-1 strain (ORSV Cy-1) was determined using cloned cDNA. This sequence is 6611 nucleotides long containing four open reading frames, which correspond to 126 K, 183 K, 31 K and 18 K proteins. The 5' non-coding region of ORSV Cy-1 is 62 nucleotides. The ORFs encoded a 126 K polypeptide and a 183 K read-through product in which helicase-sequence and polymerase-sequence motifs are found. The 5' non-coding region, which extends from bases 1 to 62 has 2G residues and the ribosome binding site (AUU). The 3' non-coding region of ORSV Cy-1 comprises 414 nucleotides in length. The genomic organization of ORSV Cy-1 is nearly identical to that of ORSV Korean strain (ORSV-K). However, the ORF encoding 183 K protein overlaps the ORF encoding 31 K protein in ORSV Cy-1, but not in ORSV-K. The 183 K read-through product of ORSV Cy-1 is 16 amino acids longer than that of ORSV-K. The homology of the nucleotide sequences of ORSV Cy-1 and ORSV-K is 96%.

Key words: Tobamovirus, Odontoglossum ringspot virus, Nucleotide sequence, Genome organization

INTRODUCTION

Odontoglossum ringspot virus (ORSV) was first isolated and characterized from Odontoglossum grande orchids showing ringspots in the leaves (Jensen and Gold 1951). This virus also causes diamond mottle, mosaic and flower colour breaking in Cymbidium, and flower colour breaking in Cattleya

Research Institute for Bioresources, Okayama University, Kurashiki 710, Japan
Received December 18, 1995
*NODAI Research Institute, Tokyo University of Agriculture, Tokyo 156, Japan
(Inouye 1966a and b, Isomura et al. 1991). The particles of ORSV are rigid rod and contain single-stranded RNA molecules, approximately 6 kb in length, which is a typical morphology of the tobamoviruses (Jensen and Gold 1951). ORSV in easily distinguishable from tobacco mosaic virus (TMV) and other tobamoviruses based on host range, serology, and the nucleotide sequence of 3'-terminal region (Inouye 1966a, Isomura et al. 1990, Isomura et al. 1991).

We present here the complete nucleotide sequence of the genomic RNA of ORSV (Cy-1 strain) isolated in Japan and compare it to that of the Korean strain of ORSV (ORSV-K).

MATERIALS AND METHODS

Virus, Virus purification, and RNA isolation

ORSV Cy-1 was originally from Cymbidium sp. (Inouye 1966a) and was propagated in Cymbidium plants in a greenhouse. The virus was purified from Cymbidium plants as described by Inouye (1966a), and RNA was extracted by incubation with 1% SDS, followed by phenol/chloroform extraction and subsequent ethanol precipitation (Ikegami et al. 1987).

Complementary DNA synthesis and molecular cloning

ORSV Cy-1 RNA was polyadenylated using poly (A) polymerase (Takara Shuzo, Kyoto, Japan) as described by Sippel (1973). DNA complementary to the RNA was synthesized using a cDNA kit (purchased from Pharmacia, Uppsala, Sweden) according to the manufacturer's specifications. Synthesized double-stranded cDNA was inserted into the EcoRI site of pUC119 using EcoRI linkers and transformed into Eschericia coli JM109 made competent. Transformants carrying plasmids with ORSV Cy-1 cDNA inserts were identified by colony hybridization using single-stranded cDNA synthesized from ORSV Cy-1 RNA.

Nucleotide sequence determination and analysis

Nucleotide (nt) sequence analysis was carried out by the enzymic dideoxynucleotide chain termination sequencing method with a DNA sequencer (model 370A, Applied Biosystems, Foster, USA).

ORSV Cy-1 RNA was decapped with NaIO₄ and aniline (Dasgupta et al. 1976), the 5' end was labeled with [γ-³²P] ATP by polynucleotide kinase (PNK), and subjected to direct enzymic sequencing (García-Arenal et al. 1987). The sequence of the 5' end of the RNA was further established by the deoxynucleotide chain termination method on the RNA (Gould and Symons 1989) by priming with a 22-base oligonucleotide complementary to nt 81 to 102.
Analysis and comparison of the ORSV Cy-1 RNA sequence, and of the deduced amino acid sequences of the encoded proteins, were carried out using the program DNASIS.

RESULTS AND DISCUSSION

Fig. 1 shows the complete sequence of ORSV Cy-1 RNA as determined from the cDNA clones. The genomic organization of ORSV Cy-1 was compared to that of ORSV-K (GenBANK, accession number; X82130). The genomic RNA of ORSV Cy-1 is 6611 nt long, 7 nt shorter than that of ORSV-K. The homology of the nucleotide sequences of ORSV Cy-1 and ORSV-K is 96%. Computer analysis of the complete sequence of ORSV Cy-1 reveals four open reading frames (ORF) in the (+) strand. The first AUG initiation codon was found at residues 63 to 65, which starts an ORF encoding a protein composed of 1111 amino acids (126 K). This ORF terminates at an amber codon, UAG, positioned at residues 3399 to 3401. The ORF encoding the read-through protein composed of 1611 amino acids (183 K) terminates at residues 3999 to 4901. The terminal 23 bases of the ORF encoding the 183 K protein overlap the ORF encoding the 31 K protein, which terminates with UGA. The ORF encoding the 18 K protein initiates 2 bases upstream from the terminal codon for the 31 K protein. This terminal ORF is the coat protein gene, as determined by expression in E. coli (Isomura et al. 1991). The coat protein gene terminates at residues 6195 to 6197. The coat protein is 157 amino acids. Such genomic organization is similar to that of ORSV-K (Fig. 2). However, the ORF encoding 183 K protein does not overlap the ORF encoding 31 K protein in ORSV-K. The 5' non-coding region of ORSV Cy-1 has 62 nt, 7 nt shorter than that of ORSV-K.

The nucleotide sequence of 5' non-coding region of ORSV Cy-1 has 85 to 90% similarity to that of ORSV-K. The 5' non-coding regions of both strains of ORSV, which extends from base 1 to 62, have 2 G residues, at positions 6 and 10, but the 5' non-coding region of TMV-V is characteristically free of the G residue (Mandeles 1968). The 5' non-coding regions of TMV-L (Ohno et al. 1984), pepper mild mottle tobamovirus (PMMV) (Avilá-Rincon et al. 1989), and CGMMV (Ugaki et al. 1991) are devoid of the G residue. The deletion of the 5'-terminal 8 nt (GUAAUUUU) of TMV-L caused the complete loss of infectivity and the other sequence of the region seemed not to be crucial, although a large deletion had a large effect (Tomenius et al. 1987). This region is also conserved in ORSV Cy-1 and ORSV-K, although it has a G residue at position 6. However, the role of the sequence of 8 nt is not known. The ribosome-binding site is thought to be AUU (Tyc et al.
ORSV genome sequence

Fig. 1-1.
ORSV genome sequence

Fig. 1. Nucleotide sequence of ORSV Cy-1 RNA and encoded amino acid sequences for each ORF. ATG initiation codons are underlined, and approximate molecular weight of the translation product is indicated. Asterisks indicate termination codons.

```
5' ORSV Cy - 1  (1111)  (1611)  (157)  (278)  6611 nt

3' ORSV - K  (1111)  (1595)  (157)  (278)  6618 nt
```

Fig. 2. Comparison of ORFs of ORSV Cy-1 and ORSV-K deduced from the nucleotide sequences. Figures in parentheses present the number of amino acid residues. Data from GenBank, accession number: X82130 (ORSV-K).

142 1996
1984), which is present from base 20 to 22 in both strains.

We previously reported on the 3' non-coding region of ORSV Cy-1 (Isomura et al. 1990). The 3' non-coding region of ORSV Cy-1 RNA is different from those of the other tobamoviruses. The 3' non-coding region of ORSV Cy-1 RNA is composed of 414 nt, but the 3' non-coding regions of TMV-V, TMV-L, TMGMV, and CGMMV are 204, 202, 210, and 176 nt, respectively (Goelet et al. 1982, Ohno et al. 1984, Solis and Garcia-Arenal 1990, Ugaki et al. 1991). ORSV Cy-1 has 3 repeated domains, each of 35 nt, which were very similar. Such repeated sequences are not present in the other tobamoviruses. It is unknown if all 3 copies within the ORSV Cy-1 sequence are required for replication of ORSV Cy-1 RNA. ORSV-K has also 3 repeated domains, each of 35 nucleotides, which were very similar (Ryu et al. 1995).

The 31 K cell-to-cell movement protein of ORSV Cy-1 is the same number of amino acid residues as that of ORSV-K (Table 1). The ORSV Cy-1 31 K protein has 98% amino acid sequence similarity to that of ORSV-K. We previously reported the sequences of 31 K and 18 K protein genes (Isomura et al. 1990, Isomura et al. 1991). The putative 126 K and 183 K proteins correspond to the 130 K and 180 K proteins of TMV-V and TMV-L, respectively (Goelet et al. 1982, Ohno et al. 1984). The former protein of ORSV Cy-1 has 1111 amino acids, which is the same number as that of ORSV-K. The 183 K read-through product of ORSV Cy-1 is 16 amino acids longer than that of ORSV-K. The amino acid sequence identity of the ORSV Cy-1 183 K and 126 K proteins is approximately 94% and 94% with the corresponding 183 K and 126 K proteins of ORSV-K.

| Table 1. Genome organization and sequence homology between ORSV Cy-1 and K |
|-----------------|-----------------|-----------------|-----------------|-----------------|
|                  | ORSV Cy-1       | ORSV K          | Total           | Alignment       |
| Total genome size| 6611nt          | 6618nt          | 96              | 96              |
| 5' non-coding region | 1-62            | 1-69            | 85              | 90              |
|                  | (62nt)          | (69nt)          |                 |                 |
| 126 K protein cistron | 63-3401         | 70-3408         | 97              | 97              |
| (No. of amino acid residues) | (1111a.a.)      | (1111a.a.)      | (94)            | (94)            |
| 183 K protein cistron | 63-4901         | 70-4858         | 94              | 97              |
| (No. of amino acid residues) | (1611a.a.)      | (1595a.a.)      | (93)            | (94)            |
| 31 K protein cistron | 4879-5718       | 4886-5725       | 98              | 98              |
| (No. of amino acid residues) | (278a.a.)       | (278a.a.)       | (98)            | (98)            |
| 18 K protein cistron | 5721-6197       | 5728-6204       | 97              | 97              |
| (No. of amino acid residues) | (157a.a.)       | (157a.a.)       | (95)            | (95)            |
| 3' non-coding region | 6198-6611       | 6205-6618       | 97              | 97              |
|                  | (414nt)         | (414nt)         |                 |                 |
ORSV genome sequence

(a)

ORSV Cy-1
VDGVPGCOKTKEILETVNFDEDLILVPGEACKMIKRNKSHVGRATKDNVRTVDSFLM
ORSV-K
VDGVLCGOKTKEILETVNFDEELILVPGEACKMIKRNKSHVGRATKDNVRTVDSFLM

824
883

884
943

HLKPKTYKNKLIDEGLMMLHTGCVNFIALSHCREAMVFDTEQIPFINRVANFPYPKHF
HLKPKTYKNKLIDEGLMMLHTGCVNFIALSHCREAMVFDTEQIPFINRVANFPYPKHF

944
1003

TLVYDHERVRLSLRCPADVTTHFMNSKYDGVKLCTNDVIRSVDAEVRGKGVFPNPSKPL
HTCLHRREVRLSLRCPADVTTHFMNSKYDGVKLCTNDVIRSVDAEVRGKGVFPNPSKPL

1004
1063

KGKIIIIFTQDSKAELKERGYYEVESTFGEINTVHEIQGETFEDSVVRLTPTPLELISKSS
KGKIIIIFTQDSKAELKERGYYEVESTFGEINTVHEIQGETFEDSVVRLTPTALELISKSS

1064
1080

PHHLVALTRHTKSFYY

(b)

ORSV Cy-1
LELDISKYDKSONEFHCAVEYLIWEKLGLNQFEELWKQGHRKTSLKDYTAKITCLWYG
ORSV-K
LVLDISKYDKSONEFHCAVEYFIWEKLGLNQFEELWKQGHRKTSLKDYTAKITCLWYG

1376
1435

1436
1495

RKSVDVTFTIIGNTVIIAACLASMI PMDKVI KAAFGDCDSMLYI PKGLDLPDIQSGANLMW
RKSVDVTFTIIGNTVIIAACLASMI PMDKVI KAAFGDCDSMLDILDI PKGLDLPDIQSGANLMW

1496
1515

NFEAKLYRKYGYFCGRYII
NFEAKLYRKYGYFCARYII

Fig. 3. Comparison of partial amino acid sequences of the 183 K protein of ORSV Cy-1 with those of the 183 K protein of ORSV-K (Goelet et al., 1982). These sequences contain motifs indicated by Habili and Symons (1989). (a) The region contains nucleic acid helicase motifs; (b) the region contains RNA polymerase motifs. Motifs are indicated by solid lines.
Habili and Symons (1989) showed that positive-strand viruses could be grouped based on amino acid sequence motifs of nucleic acid helicases and RNA polymerases, and proposed a new luteovirus supergroup. Within the 126 K protein of ORSV Cy-1 there are six areas of sequential amino acids associated with proteins having helicase activity (Fig. 3). These regions show a high level of sequence conservation within ORSV strains, Regions II, III, IV, V, and VI have complete sequence conservation between ORSV Cy-1 and ORSV-K, whereas region I have 1 conservative amino acid difference. Within the 183 K protein gene region there are four areas of conserved amino acids associated with putative viral replicase proteins (Fig. 3). These four regions show a high level of sequence conservation within ORSV strains. Region II has complete sequence conservation between ORSV Cy-1 and ORSV-K, whereas regions I, III and IV have 1, 2, and 1 conservative amino acid difference, respectively.

REFERENCES

Isomura, Y., Matumoto, Y., Murayama, A., Chatani, M., Inouye, N. and Ikegami, M. 1990. Nucleotide
ORSV genome sequence


オドントログサムリングスポットウイルス Cy-1 株
RNA のゲノム構成および韓国株との比較

池上 正人・井上 成信

摘 要

オドントログサムリングスポットウイルス Cy-1 株（ORSV Cy-1）の cDNA クローンを
用いて完全長ゲノム RNA の塩基配列を決定した。この配列の長さは6611塩基であり、126
K、183K、31 K および18 K のタンパク質に対応する 4 個のオープンリーディングフレーム
（ORF）を含んでいた。ORSV Cy-1 の 5 ' 非翻訳領域は 62 塩基であった。126 K および、その
183 K リードスルー産物をコードする ORF 内にはヘリカーゼとポリメラーゼの塩基配列モ
チーフが含まれていた。5 ' 非翻訳領域である 1 から 62 塩基目の領域には 2 個の G 残基とリポ
ソーム結合領域（AUU）を保有していた。ORSV Cy-1 の 3 ' 非翻訳領域は 414 塩基長であっ
た。ORSV Cy-1 の遺伝子構成は、韓国株（ORSV-K）と殆ど同じであった。しかしながら、
ORSV Cy-1 の 183 K タンパク質をコードする ORF は 31 K タンパク質をコードしている
ORF と一部重なっていたが、ORSV-K では重なっていなかった。ORSV Cy-1 および ORSV
K の 183 K タンパク質のアミノ酸数を比較すると、ORSV Cy-1の方が 16 アミノ酸長かった。
ORSV Cy-1 と ORSV-K の塩基配列の相似性は 96% であった。

キーワード：コパウイルス，オドントログサムリングスポットウイルス，
塩基配列，ゲノム構造

Vol. 4

147