岡大資生研報 (Bull. Res. Inst. Bioresour. Okayama Univ.) 4 : 137-147 (1996)

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

Masato IKEGAMI* and Narinobu INOUYE

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 composes 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 overlapes 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 *Eco*RI site of pUC119 using *Eco*RI linkers and transformed into *Eschericia coli* JM109 made competent. Transformants carrying plasmids with ORSV Cy-1 cDNA inserts were identifed 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 $NaIO_4$ and aniline (Dasgupta *et al.* 1976), the 5' end was labeled with [γ-^{32}P] 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) (Avílá-Rincon *et al.* 1989), and CGMMV (Ugaki *et al.* 1991) are devoid of the G residue. The deletion of the 5'-terminal 8 nt (GUAUUUUU) 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.*

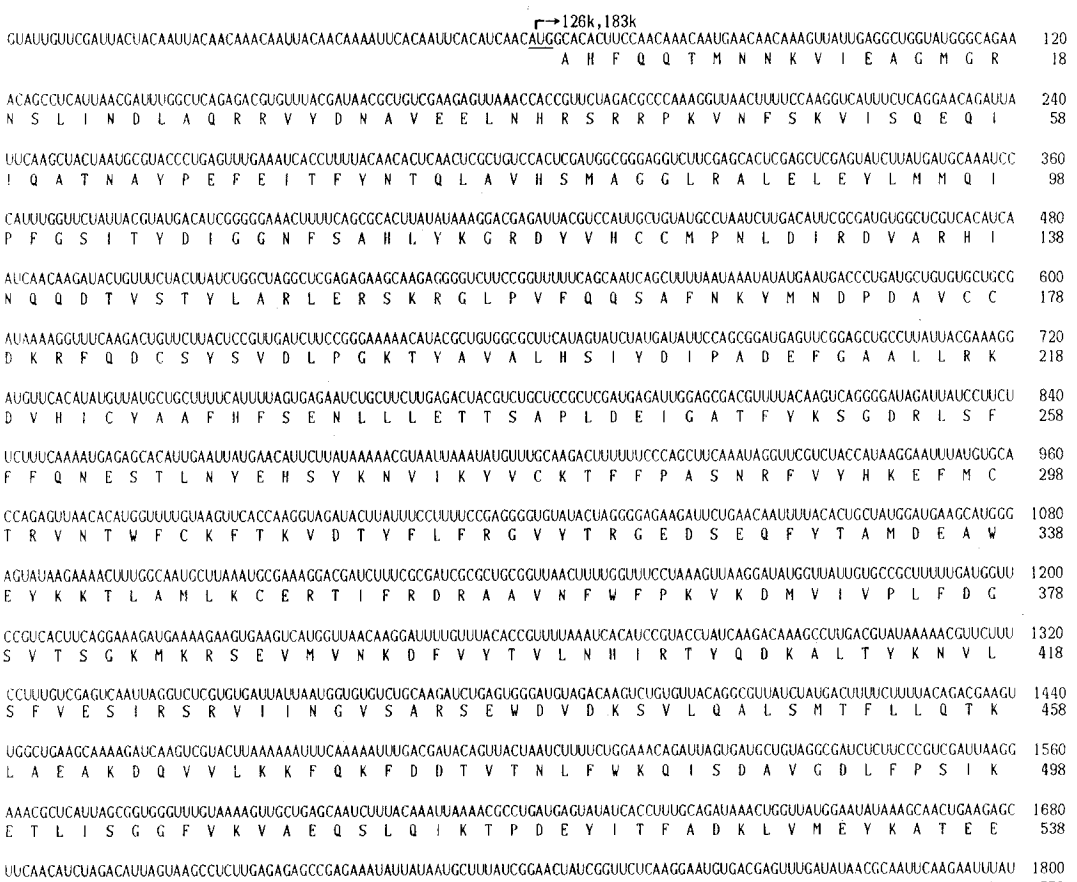

Fig. 1-1.

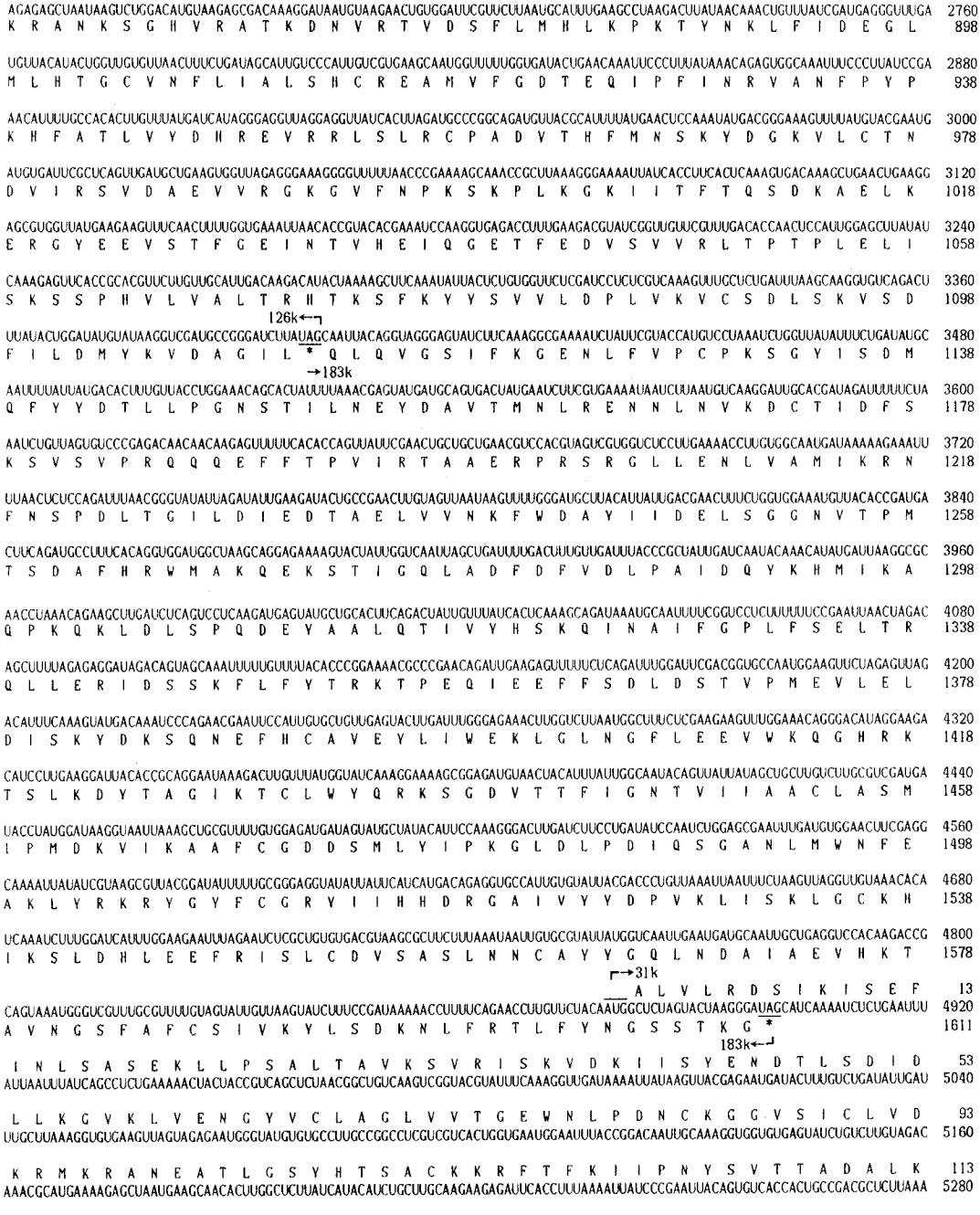

Fig. 1-2.

```
 G  I  W  Q  V  M  T  N  I  R  G  V  E  M  E  K  G  F  C  P  L  S  L  E  F  V  S  I  C  V  V  V  L  N  N  I  K  L  G  L   153
GGAAUAUGGCAAGUUAUGACAAAUAUUAGAGGUGUGGAAAUGGAAAAGGGUUUUUGCCCAUUAUCUUUGGAGUUUGUUUCUAUUUGCGUGGUCUAUUUAAAUAAUAUAAAACUCGGACUG  5400

 R  E  K  I  L  N  V  T  E  G  G  P  T  E  L  T  E  A  V  V  D  E  F  V  E  K  V  P  M  A  A  R  L  K  S  F  R  S  V  N   193
AGAGAGAAGAUUUUGAACGUAACAGAAGGAGGGCCCACCGAACUUACUGAAGCAGUUGUUGAUGAGUUCGUGGAGAAAGUUCCUAUGGCUGCUAGGCUCAAGUCUUUUCGCUCUGUCAAC  5520

 K  K  K  P  S  N  S  S  K  F  V  N  G  K  S  R  L  N  S  R  N  K  L  N  Y  E  N  G  D  S  D  V  G  I  S  V  V  D  D  I   233
AAAAAGAAACCUAGUAACAGUUCUAAGUUUGUUAAUGGAAAAUCUAGACUUAAUAGUCGUAAUAAAUUAAAUUAUGAAAAUGGUGAUAGUGAUGUUGGUAUUAGUGUUGUAGAUGAUAUU  5640
                                                                                                  31k←┐
 V  V  G  N  G  V  S  D  I  R  I  D  D  D  C  E  S  F  D  A  Q  S  D  S  Y  *                                              258
GUGGUUGGUAAUGGUGUUAGUGAUAUUCGUAUUGAUGAUGAUUGUGAGUCAUUUGACGCACAAUCUGAUUCGUAUUGAAUAUGUCUUACACUAUUACAGACCCGUCUAAGCUGGCUUAUU  5760
                                                                                                  S  Y  T  I  T  D  P  S  K  L  A  Y    12
                                                                                                  └→18k
UAAGCUCGGCUUGGGCUGACCCCAAUUCACUGAUCAACCUUUGUACCAAUUCUCUGGGUAAUCAGUUCCAAACACAACAAGCUCGAACAACUGUUCAACAGCAGUUUGCUGAUGUUUGGC  5880
L  S  S  A  W  A  D  P  N  S  L  I  N  L  C  T  N  S  L  G  N  Q  F  Q  T  Q  Q  A  R  T  T  V  Q  Q  Q  F  A  D  V  W    52

AGCCGGUUCCUACUUUGACCAGUAGGUUCCCUGCAGGCGCUGGUUACUUCAGAGUUUAUCGCUAUGAUCCUAUAUUAGAUCCUUUAAUAACUUUCUUAAUGGGUACUUUUGAUACUCGUA  6000
Q  P  V  P  T  L  T  S  R  F  P  A  G  A  G  Y  F  R  V  Y  R  Y  D  P  I  L  D  P  L  I  T  F  L  M  G  T  F  D  T  R    92

AUAGAAUAAUCGAGGUAGAAAAUCCGCAGAAUCCGACAACUACGGAAACAUUAGAUGCAACUCGUAGAGUUGAUGAUGCAACUGUAGCAAUAAGAUCUGCAAUAAAUAAUCUAUUAAAUG  6120
N  R  I  I  E  V  E  N  P  Q  N  P  T  T  T  E  T  L  D  A  T  R  R  V  D  D  A  T  V  A  I  R  S  A  I  N  N  L  L  N   132

AGUUAGUUAGGGGAACUGGUAUGUACAAUCAAGUCUCAUUUGAGACGAUAUCUGGACUUACUUGGACCUCUUCCUAAUCAUAUUUAGGAAAAUAACGUUGAUAGUGUUGAACUAUCCGUG  6240
E  L  V  R  G  T  G  M  Y  N  Q  V  S  F  E  T  I  S  G  L  T  W  T  S  S  *                                          157
                                                                                  18k←┘
GUGCAUACGAUAAUGCAUAGUGGUUAUCCCUCCACUUAAAUCGAAGGGUUUUUCACUGCAGAUAUGUAGGUUUCCUACGGUGAAUAUAAAACUUAUAUCCCGUUGUGUACACGAUAGUAC  6360

AUAGUGUUUAUCCCUCCACUUAAAUCGAAGGGUUUUGUGUCAGACGCGUGUAAGGAGUGGUCAACCUUACGACACAUUUAAAAUAAUGCGUCCGUGGUGCAUACGAUAAUGCAUAGUGUU  6480

UGUCCCUCCACUUAAAUCGAAGGGUUGUGUAUAUGGAUCAUGCGGAUAAAGUUAUACUGGUGCAGUAUAACCCGUUAUACACAUAAAAUUAUGAGGGAUUCGAAUUCCCCCUUACCUCGG  6600

GUAGAGGCCCA                                                                                                              6611
```

Fig. 1-3.

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.

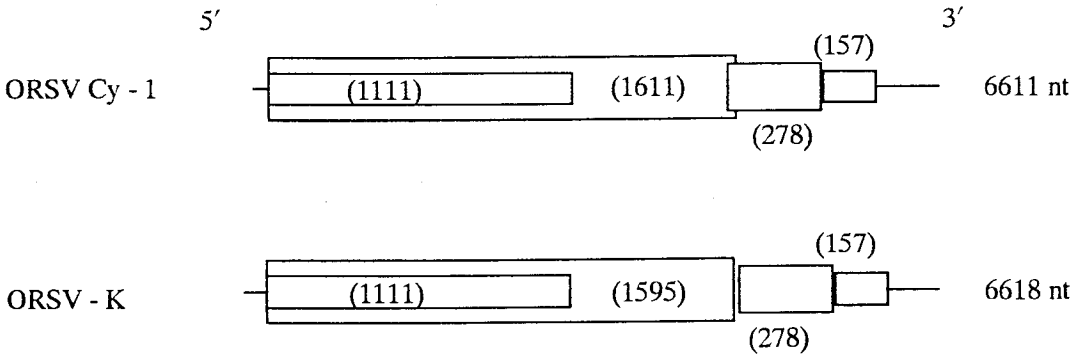

Fig. 2. Comparison of ORFs of ORSV Cy-1 and ORSV-K deduced from the nucleotide sequences. Figures in parentheses present the number of amino acid residues. Data from GenBANK, accession number : X82130 (ORSV-K).

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 García-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

			Sequence homology (%)	
	ORSV Cy-1	ORSV K	Total	Alignment
Total genome size	6611nt	6618nt	96	96
5' non-coding region	1-62 (62nt)	1-69 (69nt)	85	90
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 (414nt)	6205-6618 (414nt)	97	97

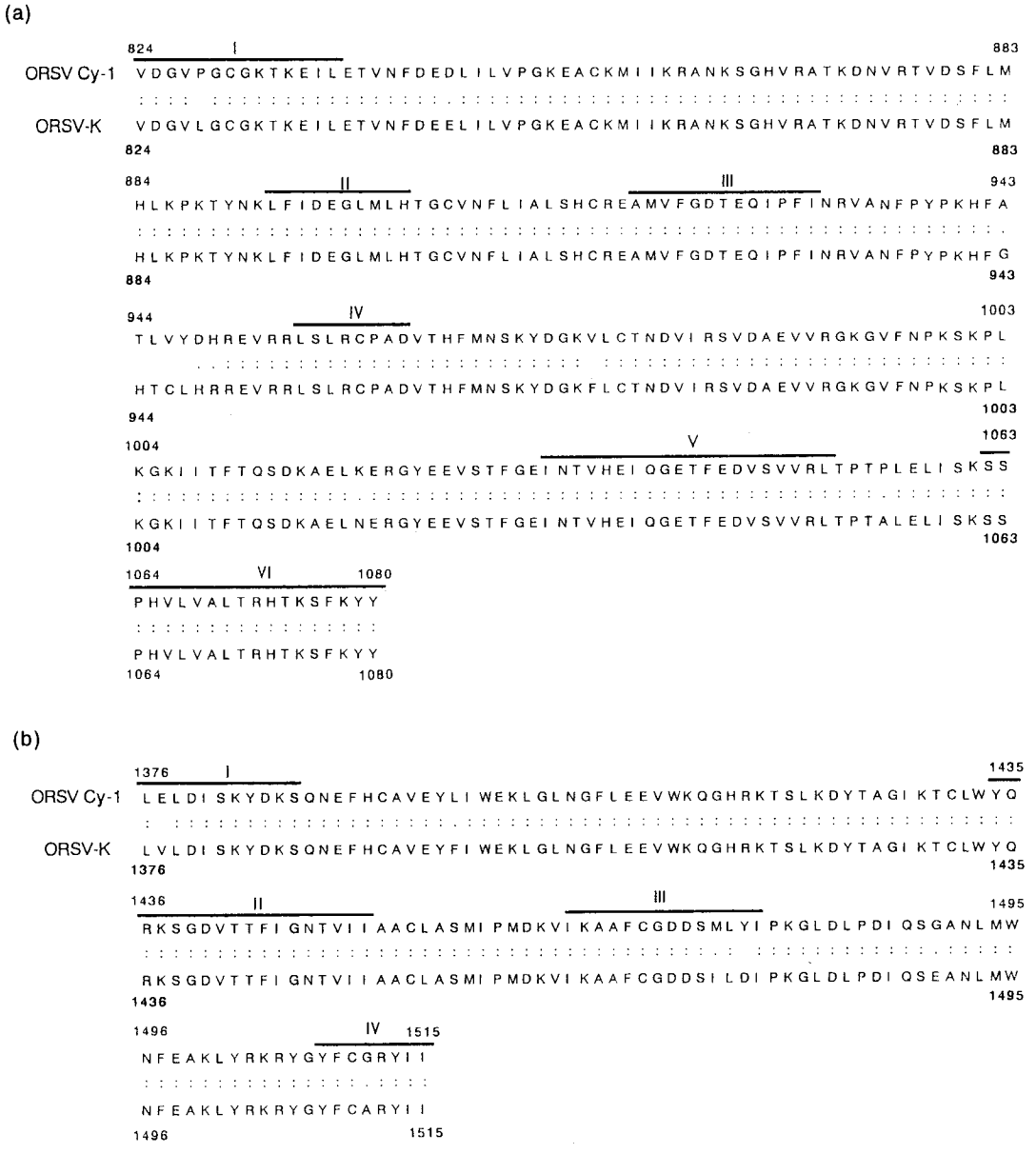

Fig. 3. Comparison of partial amino acid sequences of the 183 K protein of ORSV Cy-1 with those of the 183 K protein of ORSV-K (Goelet *et al.* 1982). These sequences contain motifs indicated by Habili and Symons (1989). (a) The region contains nucleic acid helicase motifs; (b) the region contains RNA 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

Avílá-Rincon, M. J., Ferrera, M. L., Alonso, E., Garcia-Luque, I. and Diaz-Ruiz, J. R. 1989. Nucleotide sequences of 5' and 3' non-coding regions of pepper mild mottle virus strain S RNA. J. Gen. Virol. 70 : 3025-3031.

Dasgupta, R., Harada, F. and Kaesberg, P. 1976. Blocked 5' termini brome mosaic virus RNA. J. Virol. 18 : 260-267.

García-Arenal, F., Zaitlin, M. and Palukaitis, P. 1987. Nucleotide sequence analysis of six satellite RNAs of cucumber mosaic virus : Primary sequence and secondary structure alterations do not correlate with differences in pathology. Virology 158 : 339-347

Goelet, P., Lomonossoff, G. P., Butler, P. J. G., Akam, M. E., Gait, M. J. and Karn, J. 1982. Nucleotide sequence of tobacco mosaic virus RNA. Proc. Natl. Acad. Sci. U. S. A. 79 : 5818-5822.

Gould, A. R. and Symons, R. H. 1982. Cucumber mosaic virus RNA3 : Determination of the nucleotide sequence provides the amino acid sequences of protein 3A and viral coat protein. Eur. J. Biochem. 126 : 217-226.

Habili, N. and Symons, R. H. 1989. Evolutionary relationship between luteoviruses and other RNA plant viruses based on sequence motifs in their putative RNA polymerases and nucleic acid helicases. Nucleic Acids Res. 17 : 9543-9555.

Ikegami, M., Osaki, T. and Inouye, T. 1987. Single-stranded DNA in tobacco leaf curl virus. Ann. Phytopathol. Soc. Japan. 53 : 269-273.

Inouye, N. 1966a. A virus disease of *Cymbidium* caused by odontoglossum ringspot virus in Japan. Ber. Ohara Inst. landw. Biol. Okayama Univ. 13 : 149-159.

Inouye, N. 1966b. Virus diseases of orchids. IV. Symptoms of viruses in *Cattleya*. Japan Orchid Soc. Bull. 12 : 2-5 (in Japanese).

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

sequence of cell-to-cell transport protein gene of odontoglossum ringspot virus. Nucleic Acids Res. 18 : 7448.

Isomura, Y., Matumoto, Y., Murayama, A., Chatani, M., Inouye, N. and Ikegami, M. 1991. Molecular cloning, sequencing and expression in *Escherichia coli* of the odontoglossum ringspot virus coat protein gene. J. Gen. Virol. 72 : 2247-2249.

Jensen, D. D., and Gold, H. A. 1951. A virus ring spot of *Odontoglossum* orchid : symptoms, transmission and electron microscopy. Phytopathology 41 : 648-653.

Mandeles, S. 1968. Localization of unique sequence in tobacco mosaic virus ribonucleic acid. J. Biol. Chem. 243 : 3671-3674.

Ohno, T., Aoyagi, M., Yamanashi, Y., Saito, H., Ikawa, S., Meshi, T. and Okada, Y. 1984. Nucleotide sequence of the tobacco mosaic virus (tomato strain) genome and comparison with the common strain genome. J. Biochem. 96 : 1915-1923.

Ryu, K. H., Choi, J. K. and Park, W. M. 1995. Analysis of the coat protein gene of odontoglossum ringspot virus-Cy (Korean isolate). Mol. Cells 5 : 181-186.

Sippel, A. E. 1973. Purification and characterization of adenosine triphosphate : Ribonucleic acid adenyltransferase from *Escherichia coli*. Eur. J. Biochem. 37 : 31-40.

Solis, I. and Garcia-Arenal, F. 1990. The complete nucleotide sequence of the genomic RNA of the tobamovirus tobacco mild green mosaic virus. Virology 177 : 553-558.

Tomenius, K., Clapham, D. and Meshi, T. 1987. Localization by immunogold cytochemistry of the virus-coded 30 K protein in plasmodesmata of leaves infected with tobacco mosaic virus. Virology 160 : 363-371.

Tyc, K., Kornaska, M., Gross, H. J. and Filipowicz, W. 1984. Multiple ribosome binding to the 5'-terminal leader sequence of tobacco mosaic virus RNA, assembly of the 80S robosome mRNA complex at the AUU codon. Eur. J. Biochem. 140 : 503-511.

Ugaki, M., Tomiyama, M., Kakutani, T., Hidaka, S., Kiguchi, T., Nagata, R., Sato, T., Motoyoshi, F. and Nishiguchi, M. 1991. The complete nucleotide sequence of cucumber green mottle mosaic virus (SH strain) genomic RNA. J. Gen. Virol. 72 : 1487-1495.

オドントグロッ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

オドントグロッ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，126K，183K，31Kおよび18K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．126Kおよび，その183K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残基とリボソーム結合領域（AUU）を保持していた．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183Kタンパク質をコードするORFは31K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183K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の塩基配列の相同性は96%であった．

キーワード：トバモウイルス，オドントグロッサムリンスポットウイルス，塩基配列，ゲノム構造