Short Communication

Characterization of *Cestrum yellow leaf curling virus*: a new member of the family *Caulimoviridae*

Livia Stavolone,[1]† Antonio Ragozzino[2] and Thomas Hohn[1]‡

[1]Friedrich Miescher Institute, Basel, Switzerland

[2]Università di Napoli, via Università, 80055 Portici (NA), Italy

Correspondence
Thomas Hohn
thomas.hohn@fmi.ch

Received 10 June 2003
Accepted 12 September 2003

Cestrum yellow leaf curling virus (CmYLCV) has been characterized as the aetiological agent of the *Cestrum parqui* mosaic disease. The virus genome was cloned and the clone was proven to be infectious to *C. parqui*. The presence of typical viroplasms in virus-infected plant tissue and the information obtained from the complete genomic sequence confirmed CmYLCV as a member of the *Caulimoviridae* family. All characteristic domains conserved in plant pararetroviruses were found in CmYLCV. Its genome is 8253 bp long and contains seven open reading frames (ORFs). Phylogenetic analysis of the relationships with other members of the *Caulimoviridae* revealed that CmYLCV is closely related to the *Soybean chlorotic mottle virus* (SbCMV)-like genus and particularly to SbCMV. However, in contrast to the other members of this genus, the primer-binding site is located in the intercistronic region following ORF Ib rather than within this ORF, and an ORF corresponding to ORF VII is missing.

The *Caulimoviridae* are plant pararetroviruses and replicate their genome through a process of reverse transcription similar to that of retroviruses (Rothnie *et al.*, 1994). Viral genomes form supercoiled minichromosomes in the host cell nucleus that are transcribed into terminally redundant polyadenylated RNA. This migrates to the cytoplasm where it is used both as polycistronic mRNA to produce viral proteins (Hohn & Fütterer, 1992) and as a replication template. In the latter case, the full-length RNA, primed by the host cytosolic initiator methionine tRNA ($tRNA^{Met}$), is reverse transcribed by the viral reverse transcriptase (RT)/RNase H into genomic dsDNA, which may re-enter the nucleus for amplification or be encapsidated into a virion. Virions accumulate in characteristic inclusion bodies in the cytoplasm of infected cells.

Caulimovirus genomes are single, circular, double-stranded DNA molecules ranging from 7000 to 8300 bp with between one and eight closely spaced open reading frames (ORFs). The functions of most genes are known, particularly in *Cauliflower mosaic virus* (CaMV), the type member of the family. ORF I encodes the cell-to-cell movement protein (MP) (Thomas & Maule, 1995) containing an RNA-binding domain conserved in the MP of most plant viruses and in the family I transport protein group (Thomas *et al.*, 1993; Koonin *et al.*, 1991). The ORF II product is the aphid transmission factor (ATF) (Woolston *et al.*, 1983), whereas that of ORF III is a virion-associated protein (VAP) (Leclerc *et al.*, 2001). VAP has been found to be an indispensable co-factor of ATF (Leh *et al.*, 1999).

The product of ORF IV is the coat protein (CP) precursor (Kirchherr *et al.*, 1988). It is processed by an aspartic proteinase (Torruella *et al.*, 1989) included in the poly-functional protein (Pol) encoded by ORF V. RT and RNase H are the other known activities of Pol (Takatsuji *et al.*, 2003). In CaMV, expression of genes from the polycistronic pregenomic RNA is efficiently transactivated by TAV, the product of ORF VI, which is translated from a subgenomic monocistronic viral mRNA (Bonneville *et al.*, 1989). TAV is also the major constituent of the virus inclusion body matrix, and influences virus host range and symptom development (Schoelz & Shepherd, 1988). Despite the diversity of caulimovirus genome organization, most of the structural and functional protein domains described for CaMV are conserved in all members of the family, albeit not always in the same order (Calvert *et al.*, 1995; Hohn & Fütterer, 1997).

A virus disease of *Cestrum parqui* was first described in the Campania region of Italy (Ragozzino, 1974). *C. parqui* is a member of the Solanaceae family, and is ubiquitous in the southern part of Italy, where it is used mainly for surrounding cultivated fields because it repels goats and has strong resistance to bacterial and cryptogrammic diseases. Infected plants show leaf mosaic, curling and stunting (Fig. 1a). Isometric virus particles of about 50 nm in diameter prevalently located in cytoplasmic inclusion bodies and very similar to those typical of caulimoviruses have been found in infected tissues (Fig. 1b).

†Present address: Istituto di Virologia Vegetale CNR, via Amendola 165/A, 70126 Bari, Italy.

‡Present address: University of Basel, Department of Botany, Plant Health Unit, Schönbeinstr. 6, CH-4056 Basel, Switzerland.

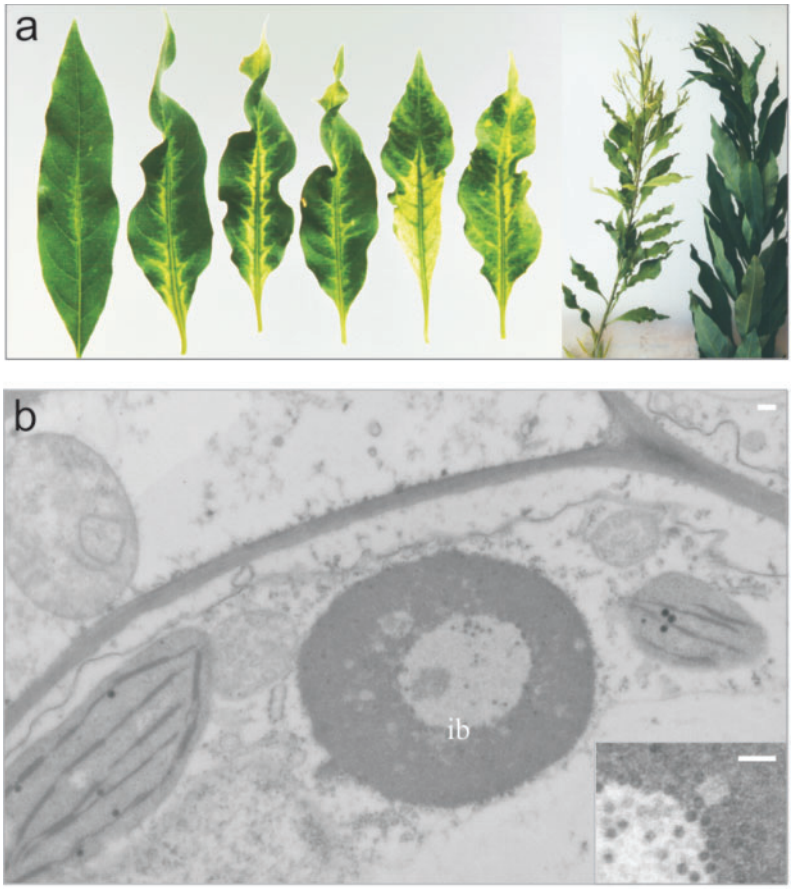


Fig. 1. Pathological alteration induced by CmYLCV on *Cestrum parqui*. (a) In infected plants symptoms develop from leaf yellow mosaic and curling (left panel) to shoot stunting (right panel). (b) Electron micrograph of ultrathin section of infected *C. parqui* leaf showing virus inclusion body (ib). In the lower right corner a magnified image shows single virions within the inclusion body. White bars in the upper right corners of the figures represent 200 nm.

The isolated infectious virus (formerly *Cestrum virus*) was named *Cestrum yellow leaf curling virus* (CmYLCV) and its biological and serological properties were studied. CmYLCV was mechanically inoculated by leaf rubbing (55 plants) and grafting (7 plants) to test and crop species (Table 1). Symptoms were observed only on *Cestrum elegans* and *C. parqui* (4 weeks after inoculation, 5 to 6 weeks after grafting) and *N. clevelandii* (6 to 7 weeks after grafting), while the other plants remained symptom-free. CmYLCV infection and non-infection was confirmed by PCR analysis with virus-specific oligonucleotides and by back-inoculation to *C. parqui* plants.

Serological affinity among members of the *Caulimoviridae* family is commonly used for classification (van Regenmortel *et al.*, 2000). We tested any potential CmYLCV serological relationship with CaMV, *Dahlia mosaic virus* (DaMV) and *Carnation etched ring virus* (CERV) by heterologous DAS-ELISA of infected *C. parqui* leaf extracts with specific antibodies (Agdia Inc.) against the three viruses. The absence of any reaction indicates distinct serological properties of CmYLCV.

CmYLCV was partially purified from infected *C. parqui* leaves and DNA was extracted according to a conventional procedure (Gardner & Shepherd, 1980). Electrophoretic separation in an agarose gel revealed two DNA components, corresponding to the linear and circular forms of caulimoviral DNAs (Menissier *et al.*, 1983). pCmYLCV, a clone of the entire CmYLCV genome flanked by 141 bp terminal repeats (Fig. 2a), was constructed in the pBluescript vector (Stratagene). To determine whether this sequence was infectious, *C. parqui* plants were mechanically inoculated with viral DNA excised from pCmYLCV at the unique *Kas*I restriction site present in the terminal repeats. Six weeks after inoculation, the first systemic symptoms appeared on plant leaves. The presence of complete CmYLCV genome molecules in non-inoculated symptomatic leaves was proven by PCR amplification using different pairs of virus-sequence-specific oligonucleotides spanning the entire sequence. Back-inoculation of healthy *C. parqui* also resulted in infection. The sequence of both strands of the cloned viral DNA was determined (GenBank acc. no. AF364175). Several clones of distinct fragments spanning the whole virus genome were also obtained and their sequences aligned with that of pCmYLCV. DNA sequencing reactions were performed using BigDye Terminators with a GeneAmp PCR system 9700 thermocycler, and analysed on an ABI PRISM 3700 DNA Analyser (all from Applied Biosystems). SeqWeb, the web-based version of the Wisconsin Package version 10.2 [Genetics Computer Group (GCG), Madison, WI, USA] and the VectorNTI package from Informax Inc. (Frederick, MD, USA) were used for all DNA and deduced amino acid sequence analyses.

Table 1. List of test and crop plants inoculated with CmYLCV

Plant species are grouped by family. Plants susceptible to CmYLCV infection are indicated in bold; plants indicated by * were also grafted with CmYLCV-infected bark tissue.

Amaranthaceae	*Brassica ruvo*	*Nicotiana benthamiana*
Gomphrena globosa	*Eruca sativa*	*Nicotiana bigelovii*
Caryophyllaceae	*Raphanus sativus*	***Nicotiana clevelandii****
Dianthus caryophyllus	Cucurbitaceae	*Nicotiana debneyi*
Chenopodiaceae	*Cucumis sativus*	*Nicotiana edwardsonii*
Beta vulgaris	*Cucurbita pepo*	*Nicotiana glauca*
Chenopodium amaranticolor	Lamiaceae	*Nicotiana glutinosa*
Chenopodium foetidum	*Ocimum basilicum*	*Nicotiana langsdorfii*
Chenopodium murale	Leguminosae	*Nicotiana occidentalis*
Chenopodium quinoa	*Phaseolus vulgaris*	*Nicotiana rustica*
Compositae	*Pisum sativum*	*Nicotiana samsum*
Calendula officinalis	*Vicia faba*	*Nicotiana sylvestris*
*Chrysanthemum leucanthemum**	*Vigna unguiculata*	*Nicotiana tabacum*
Dahlia pinnata	Nyctaginaceae	*Nicotiana xanthii**
Lactuca sativa	*Mirabilis jalapa*	*Petunia hybrida*
Sonchus officinalis	Solanaceae	*Physalis floridana*
Tagetes sp.	*Capsicum annuum**	*Physalis oxycarpa*
Zinnia elegans	***Cestrum elegans***	*Physalis peruviana*
Cruciferae	***Cestrum parqui****	*Solanum melongena**
Brassica campestris cv. Just Right	*Datura methel*	*Solanum tuberosum*
Brassica oleracea var. *Capitata*	*Datura stramonium*	Tetragoniaceae
Brassica oleracea sp.	*Lycopersicon esculentum**	*Tetragonia expansa*

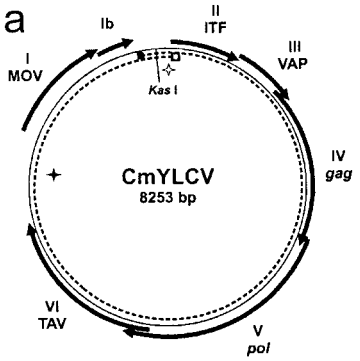
The viral genome comprises 8253 nucleotides and has a GC content of 37 %. According to the convention for caulimovirus sequences, numbering begins at the 5′-end of the putative minus-strand primer-binding site. This sequence is highly conserved in all members of the *Caulimoviridae* family and, in the case of CmYLCV, has 12 nucleotides complementary to the consensus $tRNA^{Met}$ sequence (Fig. 2c). The virus shows the highest overall nucleotide sequence similarity with members of the *Soybean chlorotic mottle virus* (SbCMV)-like genus, e.g. SbCMV (47·8 %) and *Peanut chlorotic streak virus* (PCSV, 45·5 %); 45 % to CaMV and less than 41 % to all other genus members.

Two non-coding regions are found in the viral DNA (Fig. 2a). The larger (6085–6755) contains a TATA box at nt 6395–6402 (TATAAATA) that is part of the promoter directing transcription of the (pre)genomic RNA (L. Stavolone and others, unpublished) followed by a polyadenylation signal (AATAAA; nt 6687–6693). As in other caulimoviruses, a long pregenomic RNA 5′-leader, folding into a typical stem–loop structure, precedes the first viral ORF (Pooggin *et al.*, 1999). Unusually, in CmYLCV, the predicted stem–loop structure starts immediately after the TATA-box and constitutes almost the entire intergenic region. This particular position of the stem–loop, limiting the space necessary to accommodate a ribosome on the RNA molecule, and the absence of a short ORF proximal to the RNA 5′-end (Ryabova *et al.*, 2000), suggest a mechanism of translation distinct from the 'ribosome shunting' typical of other members of the *Caulimoviridae* family (Ryabova & Hohn, 2000). The second non-coding region is 326 bp long (7961–0033) and includes the primer ($tRNA^{Met}$) binding site.

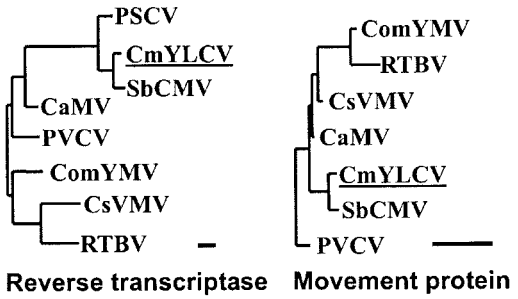
Computer analysis predicts seven putative ORFs encoding proteins of more than 10 kDa on the positive strand of the virus genome sequence (Fig. 2a) and none in the minus strand. All motifs characteristic of the *Caulimoviridae* and those shared with retroid elements are conserved in the CmYLCV sequence (Hohn & Fütterer, 1997). Genome organization (number and distribution of the ORFs) of CmYLCV resembles that of SbCMV-like genus members (Hasegawa *et al.*, 1989). Indeed, the SbCMV-like genus is characterized by the presence of three ORFs between ORF I and ORF IV instead of the two typical of CaMV. The additional ORF ('Ib' in SbCMV and 'A' in PCSV) following gene I contains the $tRNA^{Met}$ binding site but, excluding this domain, it is dispensable for virus infectivity (Mushegian *et al.*, 1995; Takemoto & Hibi, 2001). In CmYLCV, ORF I terminates the larger intergenic region and is followed by a small ORF Ib (Fig. 2a) separated by a single ATGA start/stop codon characteristic of the SbCMV-like genus. However, CmYLCV ORF Ib is smaller and followed by an intergenic region (bp 7961–0033) including the $tRNA^{Met}$ binding site.

ORF I shows significant homology with caulimovirus MPs, particularly in the conserved RNA-binding domain (Fig. 2b). It has been suggested for CaMV that MPs mediate virus movement by modifying plasmodesmata, forming

a



d



b

Protease

```
CmYLCV ORFV      23 AYVDTGATICL
SoyCMV ORFV      33 AYIDTGATLCF
CaMV   ORFV      42 CFVDTGASLCI
CVMV   ORFIII    23 GLFDTGANICI
PVCV   ORFI    1223 AFYDTGAAYSI
CoYMV  ORFIII  1217 AIVDTGATACL
RTBV   ORFIII   984 ALIDSGSTHNI
                    D G
```

Transactivator/Viroplasmin

```
CmYLCV ORFVI  101 LFPGADPVFAYTAYLMGYVKQIIIFEEF
SoyCMV ORFVI  235 FFIGADPHLLSSFYQSGLISYIYLQEDE
CaMV   ORFVI  293 FRKNANPQMVREAYAAGLIKTIYPSNNL
CVMV   ORFIV  134 ADIYTDRNLVISAYNYGLLQVLYIEHPS
                                  Y   G
```

Reverse transcriptase

```
CmYLCV  ORFV     291 WFSTLDAKSGYWQLRL<19>WNVLPFGLKQAPGIYQN<11>FCLAYIDDILVFT
SoyCMV  ORFV     291 WFSSLDAKSGYYQLRL<19>WNVLSFGLKQAPCIYQR<11>IYLAYIDDILIFT
CaMV    ORFV     336 IFSSFDCKSGFWQVLL<18>WNVVPFGLKQAPSIFQR<11>FCCVYVDDILVFS
CVMV    ORFIII   296 IISKFDCKSGFYHIKL<18>WKVLPFGYHNSPSIFQQ<11>FIIVYIDDILVFS
PVCV    ORFI    1474 LFSKFDLKSGFWQLGI<18>WKVMPFGLKTAPSLFQK<11>SALVYIDDILLFS
CoYMV   ORFIII  1498 IYSKFDLKSGFWQVAM<18>WLVMPFGLKNAPAIFQR<11>FIAVYIDDILVFS
RTBV    ORFIII  1274 IFSKFDLKAGFHHMKL<18>WNVCPFGIANAPCAFQR<10>FALLYIDDILIAS
                     FS  D K G          W V  FG    P   Q         Y DDIL
```

Ribonuclease H

```
CmYLCV   ORFV     512  IIETDASGDTWA<19>ERLCKYTSGIFS<77>EYIKGEKNSLADTL
SoyCMV   ORFV     508  IVETDASQHSWS<63>LLLCKYVSGTFT<72>ELIKSENNPFEIRL
CaMV     ORFV     547  IIETDASDDYWG<14>ELLCRYASGSFK<72>EHIKGTDNHFADFL
CVMV     ORFIII   512  IIEVDASNNAYG<12>EYLCRYNSGTFK<73>QLIDGKTNIIADYL
PVCV     ORFI    2612  ILQTDASDQYWS<10>RKICGFASGKFK<61>KHLKGKDNILADFL
CoYMV    ORFIII  1711  IIETDGCMTGWG<15>ERICAYASGSFN<72>EHIDGKHNGLADAL
RTBV     ORFIII  1486  IIETDASEEGWG<15>EKIAGYASGNFG<72>EHIKGNKNFLPNFL
                       I   D                  SG F           N     L
```

Movement Protein

```
CmYLCV   ORFI     50  NIFDRKNVIFY<36>IHISTIQIVLKSTFLKGLDYPISLAITDERI
SoyCMV   ORFI     46  NVFSRDNILKF<45>VHISTLQVLIKSTFLKGLDTPLELTLRDNRL
CaMV     ORFI     71  KAFSPKNEILY<45>VHLGAVKILLKAQFRNGIDTPIKIALIDDRI
CVMV     ORFI    943  KLFGKKEIEYY<48>IHLAAVEIVVKAYFREGIDTPFEIILCDDRI
PVCV     ORFI    115  SAFAKRSVTPW<56>IHFGAVKIALTYHGRKGQPVVARLALLDTRY
CoYMV    ORFIII  102  RTFHGSVNHRV<37>IHIGVMLVRIQILHRKFAGTMALIVFRDTRW
RTBV     ORFIII   44  ENFDWKKLSGI<56>YHIGMMAIGVKGLHRRKIGTKVMIMFYDDSF
                       F              H                          D
```

c

tRNA^{met}-binding site

1- **TGGTATCAGAGCA**TTCAA

ORF II: Putative coiled-coil domain

129- RTISNLEINRTDIKSDLEKQISLNKENLKGL<23>IANLTQKVKEMQTFLNDKL

ORF III: Coiled-coil domain

1- MGLEDSVKQLQGAIEELKGLI<3>VEEMKEIVKRESED<3>QEYLRSFLEPLAK

ORF IV: Zinc-finger

431- **CKCWIC**HEEGHYANE**C**

tubules that can accommodate virus particles in the channel (Linstead *et al.*, 1998). Indeed, in the course of our electron microscopy investigations we observed such tubules containing CmYLCV particles in *C. parqui*-infected leaf tissue. This suggests that CmYLCV could employ a similar mechanism for cell-to-cell movement.

The amino acid sequence of ORF II shows 41 % similarity

Fig. 2. CmYLCV genome and predicted proteins properties. (a) Predicted genome organization of CmYLCV based upon the nucleotide sequence of cloned viral DNA. The circle represents the double-stranded DNA genome. The outer arrows depict the position of the ORFs (I through VI) in the three reading frames. The black and white stars indicate the position of the potential full-length transcript promoter TATA-box (TATAAATA) and the $tRNA^{Met}$ binding site, respectively. The dashed arc represents the virus DNA fragment cloned in pCmYLCV. The position of the *Kas*I restriction enzyme is marked. (b) Alignment of biologically significant homologous regions of proteins encoded by one member per each genus of the *Caulimoviridae* family. The virus acronym, ORF designation and the number of residues separating the protein N termini from the aligned segment are indicated in front of each sequence. Identical amino acids are shown in italics below the alignments. The spacing between amino acid blocks is given in angle brackets. (c) CmYLCV domains of biological significance conserved in the *Caulimoviridae* family. Numbers indicate the distance from the N termini of the proteins. Position *a* and *d* of the coiled–coil heptad repeats, prevalently occupied by hydrophobic amino acids that form the helix interface, are underlined. (d) Phylogenetic trees. Phylogeny was inferred from reverse transcriptase (RT) and movement protein (MP) domains as aligned above (second block in the MP alignment, first block in RT). Genetic distances represented by the branches of the trees are indicated by bars corresponding to 10 (RT tree) and 100 (MP tree) substitutions per 100 residues.

with SbCMV ORF II and contains a putative double α-helix domain (Fig. 2c). A similar motif in CaMV ORFII is responsible for ATF trimerization (Hebrard *et al.*, 2001) and interaction with the product of ORFIII (Leh *et al.*, 1999). Although no aphid (or other insect) transmission test was performed, we observed that healthy *C. parqui* transplanted under field conditions in the vicinity of a plant naturally infected with CmYLCV acquired the virus and displayed symptoms during the spring season. This suggests possible insect-mediated CmYLCV transmission.

CmYLCV ORF III also contains a coiled–coil domain (Fig. 2c) responsible for tetramerization, a conserved feature of caulimovirus VAPs (Stavolone *et al.*, 2001).

ORF IV encodes the viral coat protein precursor as suggested by the presence of an RNA-binding domain (zinc finger; Fig. 2c) arranged as $CX_2CX_4HX_4C$ and preceded by a lysine-rich domain (aa 380–431; 44 % lysine content) (Covey, 1986). An acidic domain within its C-terminal part resembles the PEST-related degradation signal of CaMV responsible for protein destabilization (Karsies *et al.*, 2001).

The putative protein encoded by ORF V contains all motifs conserved in caulimovirus replicases: aspartic protease, RT and RNase H (Fig. 2b). All domains conserved in ORFs IV and V follow the typical '*gag–pol*' arrangement of most retro- and pararetroviruses (Rothnie *et al.*, 1994) and show the highest homology with SbCMV, confirming that CmYLCV is most appropriately classified with this genus.

The product of CmYLCV ORF VI conserves the key domains of the CaMV TAV (Fig. 2b) and hence probably also protein function. Although there is some indication for the absence of a shunting mechanism in SbCMV (Takemoto & Hibi, 2001) and CmYLCV, a possible alternative strategy of gene expression involving TAV-mediated reinitiation of translation (Park *et al.*, 2001) is not excluded.

Predicted relationships among representative members of the *Caulimoviridae* family were determined using SeqWeb. One member per each genus of the family was included in the phylogenetic analyses. These were based on conserved amino acid regions in MP and RT, aligned as in Fig. 2(b). A matrix of the pairwise evolutionary distances between aligned sequences, corrected using the Kimura method, was generated. Unrooted phylogenetic trees were reconstructed from the generated matrixes using the neighbour-joining method, and displayed as phylograms. Both trees identify CmYLCV as member of the SbCMV-like genus of the *Caulimoviridae* family with the closest homology to SbCMV (Fig. 2d).

However, CmYLCV differs from the two other members of the SbCMV-like genus in the absence of the nonessential ORF VII (Mushegian *et al.*, 1995; Takemoto & Hibi, 2001). The lack of ORF VII and the exclusion of the $tRNA^{Met}$ binding site from ORF Ib, the other dispensable gene in the genome, could represent an evolutionary strategy of CmYLCV to increase efficiency.

Analysis of the CmYLCV sequence, and comparison with the other available sequences of members of the *Caulimoviridae*, confirms that despite the many conserved features within this virus family, single virus properties can differ significantly, even among members of the same genus. The most intriguing question deriving from this study is whether ribosome shunting does indeed not occur in CmYLCV and what kind of translation mechanism is used instead.

Finally, further investigations on the activity of the potential full-length transcript promoter could uncover a new biotechnological tool for expression of exogenous genes in plants.

ACKNOWLEDGEMENTS

We thank Maciej Pietrzak and Herbert Angliker for DNA sequencing and Andreas Gisel for valuable help with bioinformatic analysis. We are also grateful to Antonio Peluso and Sjoerd van Eeden for expert technical assistance and to Helen Rothnie for critical reading of the manuscript. L. S. was partially supported by an EMBO short-term fellowship.

REFERENCES

Bonneville, J. M., Sanfaçon, H., Fütterer, J. & Hohn, T. (1989). Posttranscriptional trans-activation in cauliflower mosaic virus. *Cell* **59**, 1135–1143.

Calvert, L. A., Ospina, M. D. & Shepherd, R. J. (1995). Characterization of cassava vein mosaic virus: a distinct plant pararetrovirus. *J Gen Virol* **76**, 1271–1278.

Covey, S. N. (1986). Amino acid sequence homology in *gag* region of reverse transcribing elements and the coat protein gene of cauliflower mosaic virus. *Nucleic Acids Res* **14**, 623–633.

Gardner, R. C. & Shepherd, R. J. (1980). A procedure for rapid isolation and analysis of cauliflower mosaic virus DNA. *Virology* **106**, 159–161.

Hasegawa, A., Verver, J., Shimada, A., Saito, M., Goldbach, R., Van Kammen, A., Miki, K., Kameya-Iwaki, M. & Hibi, T. (1989). The complete sequence of soybean chlorotic mottle virus DNA and the identification of a novel promoter. *Nucleic Acids Res* **17**, 9993–10013.

Hebrard, E., Drucker, M., Leclerc, D. & 9 other authors (2001). Biochemical characterization of the helper component of cauliflower mosaic virus. *J Virol* **75**, 8538–8546.

Hohn, T. & Fütterer, J. (1992). Transcriptional and translational control of gene expression in cauliflower mosaic virus. *Curr Opin Genet Dev* **2**, 90–96.

Hohn, T. & Fütterer, J. (1997). The proteins and functions of plant pararetroviruses: knowns and unknowns. *Crit Rev Plant Sci* **16**, 133–161.

Karsies, A., Hohn, T. & Leclerc, D. (2001). Degradation signals within both terminal domains of the cauliflower mosaic virus capsid protein precursor. *Plant J* **27**, 335–343.

Kirchherr, D., Albrecht, H., Mesnard, J. M. & Lebeurier, G. (1988). Expression of the cauliflower mosaic virus capsid gene *in vivo*. *Plant Mol Biol* **11**, 271–276.

Koonin, E. V., Mushegian, A. R., Ryabov, E. V. & Dolja, V. V. (1991). Diverse groups of plant RNA and DNA viruses share related movement proteins that may possess chaperone-like activity. *J Gen Virol* **72**, 2895–2903.

Leclerc, D., Stavolone, L., Meier, E., Guerra-Peraza, O., Herzog, E. & Hohn, T. (2001). The product of ORF III in cauliflower mosaic virus interacts with the viral coat protein through its C-terminal proline rich domain. *Virus Genes* **22**, 159–165.

Leh, V., Jacquot, E., Geldreich, A., Hermann, T., Leclerc, D., Cerutti, M., Yot, P., Keller, M. & Blanc, S. (1999). Aphid transmission of cauliflower mosaic virus requires the viral PIII protein. *EMBO J* **18**, 7077–7085.

Linstead, P. J., Hills, G. J., Plaskitt, K. A., Wilson, I. G., Harker, C. L. & Maule, A. J. (1988). The subcellular localization of the gene I product of cauliflower mosaic virus is consistent with a function associated with virus spread. *J Gen Virol* **69**, 1809–1818.

Menissier, J., De Murcia, G., Lebeurier, G. & Hirth, L. (1983). Electron microscopic studies of the different topological forms of the cauliflower mosaic virus DNA: knotted encapsidated DNA and nuclear minichromosome. *EMBO J* **2**, 1067–1071.

Mushegian, A. R., Wolff, J. A., Richins, R. D. & Shepherd, R. J. (1995). Molecular analysis of the essential and nonessential genetic elements in the genome of peanut chlorotic streak caulimovirus. *Virology* **206**, 823–834.

Park, H. S., Himmelbach, A., Browning, K. S., Hohn, T. & Ryabova, L. A. (2001). A plant viral 'reinitiation' factor interacts with the host translational machinery. *Cell* **106**, 723–733.

Pooggin, M. M., Futterer, J., Skryabin, K. G. & Hohn, T. (1999). A short open reading frame terminating in front of a stable hairpin is the conserved feature in pregenomic RNA leaders of plant pararetroviruses. *J Gen Virol* **80**, 2217–2228.

Ragozzino, A. (1974). Una virosi del *Cestrum parqui* L'Erit (Fam. Solanaceae) in Campania. *Ann Fac Sci Agrar Univ Stud Napoli Portici IV* **8**, 249.

Rothnie, H. M., Chapdelaine, Y. & Hohn, T. (1994). Pararetroviruses and retroviruses: a comparative review of viral structure and gene expression strategies. *Adv Virus Res* **44**, 1–67.

Ryabova, L. A. & Hohn, T. (2000). Ribosome shunting in the cauliflower mosaic virus 35S RNA leader is a special case of reinitiation of translation functioning in plant and animal systems. *Genes Dev* **14**, 817–829.

Ryabova, L. A., Pooggin, M. M., Dominguez, D. I. & Hohn, T. (2000). Continuous and discontinuous ribosome scanning on the cauliflower mosaic virus 35 S RNA leader is controlled by short open reading frames. *J Biol Chem* **275**, 37278–37284.

Schoelz, J. E. & Shepherd, R. J. (1988). Host range control of cauliflower mosaic virus. *Virology* **162**, 30–37.

Stavolone, L., Herzog, E., Leclerc, D. & Hohn, T. (2001). Tetramerization is a conserved feature of the virion-associated protein in plant pararetroviruses. *J Virol* **75**, 7739–7743.

Takatsuji, H., Hirochika, H., Fukushi, T. & Ikeda, J. E. (2003). Expression of cauliflower mosaic virus reverse transcriptase in yeast. *Nature* **319**, 240–243.

Takemoto, Y. & Hibi, T. (2001). Genes Ia, II, III, IV and V of *Soybean chlorotic mottle virus* are essential but the gene Ib product is non-essential for systemic infection. *J Gen Virol* **82**, 1481–1489.

Thomas, C. L. & Maule, A. J. (1995). Identification of the cauliflower mosaic virus movement protein RNA-binding domain. *Virology* **206**, 1145–1149.

Thomas, C. L., Perbal, C. & Maule, A. J. (1993). A mutation of cauliflower mosaic virus gene I interferes with virus movement but not virus replication. *Virology* **192**, 415–421.

Torruella, M., Gordon, K. & Hohn, T. (1989). Cauliflower mosaic virus produces an aspartic proteinase to cleave its polyproteins. *EMBO J* **8**, 2819–2825.

van Regenmortel, M. H. V., Fauquet, C. M. Bishop, D. H. L. Carstens, E. B. Estes, M. K. Lemon, S. M. Maniloff, J. Mayo, M. A. McGeoch, D. J. Pringle C. R. & Wickner, R. B. (editors) (2000). *Virus Taxonomy. Seventh Report of the International Committee on Taxonomy of Viruses.* San Diego: Academic Press.

Woolston, C. J., Covey, S. N., Penswick, J. R. & Davies, J. W. (1983). Aphid transmission and a polypeptide are specified by a defined region of the cauliflower mosaic virus genome. *Gene* **23**, 15–23.