

# Characterization of *cis*-acting elements affecting strength and phloem specificity of the coconut foliar decay virus promoter

Alain Hehnt and Wolfgang Rohde

MPI für Züchtungsforschung, Carl-von-Linné-Weg 10, D-50829 Köln, Germany

During replication in its host plant, coconut foliar decay virus (CFDV) remains restricted to the phloem tissue. Previous *in vivo* studies on subgenomic CFDV DNA had provided evidence for the phloem specificity of the CFDV promoter. Here, new promoter constructs are described which are distinguished by the presence or absence of various *cis*-acting signals and which gave rise to a 16-fold higher reporter gene ( $\beta$ -glucuronidase) activity (reaching 30% of the cauliflower mosaic virus 35S promoter) in tobacco protoplasts, while the phloem specificity in transgenic tobacco plants was conserved. Surprisingly, the CFDV stem-loop structure dramatically influenced transcriptional efficiency. From these studies and sequence comparisons with other phloem-specific promoters, *cis*-signals involved in CFDV promoter strength and tissue specificity were identified.

Foliar decay disease affects coconut (*Cocos nucifera* L.) palms introduced into Vanuatu, while the local coconut ecotypes are tolerant to the disease (Calves *et al.*, 1980). It is caused by coconut foliar decay virus (CFDV), which is transmitted by the plant hopper *Mundus taffini* (Julia *et al.*, 1985). The 20 nm CFDV particles (Randles & Hanold, 1989) contain a unique, circular covalently closed, single-stranded DNA molecule of 1291 nucleotides (nt) with the potential to encode at least six viral proteins of molecular mass greater than 5 kDa (Rohde *et al.*, 1990). Based on these properties CFDV, together with banana bunchy top virus (Harding *et al.*, 1993; Wu *et al.*, 1994), subtterranean clover stunt virus (Boevink *et al.*, 1995) and faba bean necrotic yellow virus (Katul *et al.*, 1995), is a member of a new plant virus group.

As pointed out previously (Rohde *et al.*, 1990, 1994), CFDV shares several structural and sequence homologies with other

viruses. A potential stem-loop (SL) located at position 941–962 (Fig. 1A) is highly homologous to the SL structure characteristic of geminiviruses and involved in geminivirus replication (Orozco & Hanley-Bowdoin, 1996). A further loop

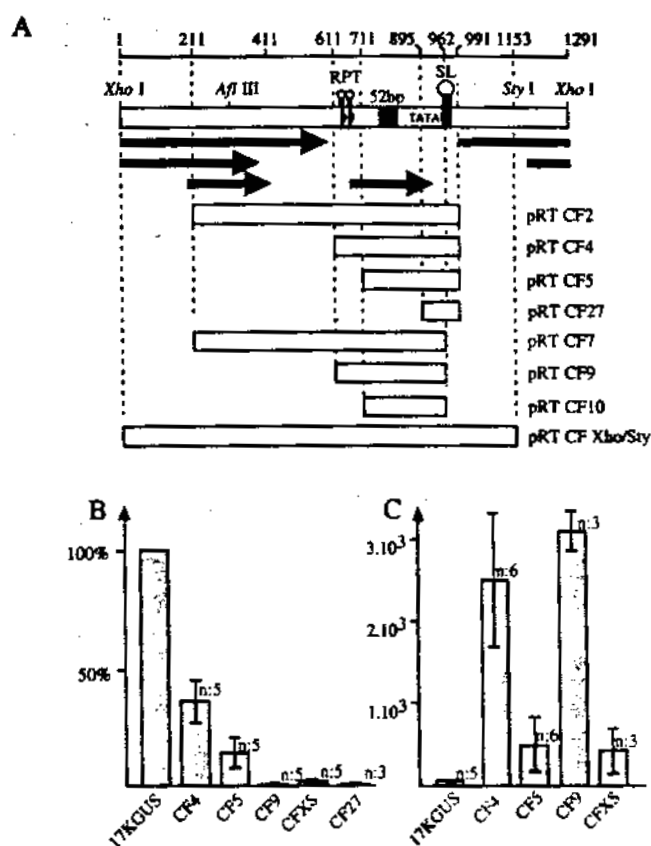


Fig. 1. CFDV promoter constructs and their *in vivo* activities. (A) Schematic representation of genomic and subgenomic CFDV fragments used for CFDV-GUS constructs on the basis of *Xho*I-linearized double-stranded CFDV DNA. Large black arrows represent ORFs 1–4 on the positive-sense strand. RPT, repeated sequence; 52bp, DNA stretch homologous to CoYMV; SL, stem-loop homologous to geminivirus stem-loop. (B) Transient expression experiments in *N. tabacum* protoplasts. GUS activity was calculated relative to the CaMV 35S 17K-GUS control. The boxes represent the mean of *n* experiments and the error bars correspond to the standard deviation; *n* is the number of independent experiments (for each experiment, three transformations were performed). (C) GUS activity of CFDV-GUS constructs in *E. coli* JM109. Results are represented as the absolute value in fluorimetric units/ $\mu$ l bacterial culture. The error bars correspond to the standard deviation for *n* colonies tested.

Author for correspondence: Wolfgang Rohde.

Fax +49 221 5062 413. e-mail rohde@mpiz-koeln.mpg.de

† Current address: BMP-CNRS UPR 406, 28 rue Goethe, F-67000 Strasbourg, France.

sequence, CGAAG, and its arrangement in two directly repeated stem-loop sequences (nt 658–676 and nt 685–701; RPT in Fig. 1A) display sequence and structural homologies with the *as-1* sequence of the cauliflower mosaic virus (CaMV) 35S promoter (Lam *et al.*, 1989). Finally, a stretch of 52 nt (position 734–785) shares 68% identity with a sequence (nt 7277–7226) of the phloem-specific *Commelina* yellow mottle badnavirus (CoYMV; Medberry *et al.*, 1992). Full-length and subgenomic fragments of double-stranded, *Xho*I-linearized CFDV DNA in fusion with the *uidA* ( $\beta$ -glucuronidase, GUS) reporter gene displayed phloem-specific promoter activity in stably transformed tobacco plants and measurable, but low, activity as compared to the CaMV 35S promoter in transient expression experiments (Rohde *et al.*, 1994). Here, we describe further promoter studies in an effort to localize *cis*-acting signals important for phloem specificity.

For these studies, two sets of subgenomic CFDV DNAs differing by the absence or presence of the SL structure were constructed by PCR using appropriate primers and introducing additional restriction sites (5' end, *Hind*III; 3' end, *Nco*I; Fig. 1A). The 3' primer which conserved the SL structure was GGA TCC ATG GGG TTA GCA GAG GAG GTT GG (the *Nco*I site is indicated in bold, the CFDV-homologous sequence is underlined); the 3' primer resulting in deletion of the SL structure was GGA TCC ATG GTA ATA CTA GCC CCC GCG. The subgenomic fragments CF2 and CF7 start at position 211 (5' primer GCC AAG CTT ATC CGA GAG CCA GCC GA, *Hind*III site in bold), CF4 and CF9 start at position 611 (5' primer GCC AAG CTT ATA CCC GCC CGC CCA CG), and CF5 and CF10 start at position 711 (5' primer GCC AAG CTT AAT CGA GTT ATG GGC GGG). For CF27 (start at position 895), three nucleotides of the wild-type CFDV sequence were mutated in order to create a *Hind*III site (CGG ACG GCT GAG TTA AGC TTG CGC CAA AAA CC). The PCR products were digested with *Hind*III and *Nco*I and introduced into the pRT2-syn-GUS vector derived from pRT2-syn-LUC (Turner *et al.*, 1994) by removing the luciferase gene by *Nco*I/*Bam*HI digestion and replacing it by the GUS reporter gene. The *Hind*III site located downstream of the CaMV 35S terminator was filled-in with Klenow polymerase and thus converted into an *Nhe*I site upon recircularization.

For the resulting pRTCF constructs, three to five independent transient expression experiments were done in *Nicotiana tabacum* SR1 protoplasts as described by Tacke *et al.* (1990), with a CaMV 35S-driven 17K-GUS construct as the positive control. GUS activity was determined in protein extracts prepared 16 h post-transfection and using 4-methylumbelliferyl glucuronide (4-MUG) as the substrate (Jefferson *et al.*, 1987). Neither pRTCF9 (Fig. 1B) nor the other two constructs lacking the SL structure (pRTCF7, pRTCF10) displayed significant GUS activity under the assay conditions (1.2% and 0.4%, respectively; data not shown). In contrast, the pRTCF constructs containing the SL structure (except for pRTCF27) showed GUS activity with pRTCF4 (30.6%  $\pm$  7.8 of

17K-GUS) and pRTCF5 (14.1%  $\pm$  6.5 of 17K-GUS), being much more active than the previously reported pRTCF<sub>Xho</sub>/Styl construct (1.9%  $\pm$  0.7 in comparison with 17K-GUS; Rohde *et al.*, 1994). Deletion of the RPT sequences (pRTCF4 versus pRTCF5) resulted in a 2-fold decrease of promoter efficiency. Finally, significant promoter activity was not detected with pRTCF27. These results limit the region for *cis* signals important for efficient transient expression to between nt 611–991, resulting in an active CFDV promoter of some 380 nt.

The CFDV fragments CF2, CF4, CF5 and CF7 in fusion with the GUS gene and the CaMV 35S terminator were cloned into the binary vector pBIN19 (Bevan, 1984), introduced into *Agrobacterium tumefaciens* strain LBA 4404 by electroporation and used for transformation of *N. tabacum* SR1. For each transformation event, regenerated plants from several independent lines (ten for CF2 and CF4, five for CF5 and CF7) were tested for GUS activity. For this, transverse thin-sections were stained with 5-bromo-4-chloro-3-indolyl  $\beta$ -D-glucuronide (X-gluc) at 37 °C and observed by light microscopy (Fig. 2). All constructs tested, including pBINCF7 which was completely inactive in protoplasts, displayed phloem-specific activity as shown for construct CF5 in Fig. 2(A, C). The strength of expression, however, differed in a manner analogous to that in the transient expression experiments (Fig. 1B), with CF2-, CF4- and CF5-transgenic tissue completely stained after 2 h, while CF7-transgenic material needed overnight incubation for colour development. For all transgenic plants, expression in the phloem tissue was apparently stronger in the adaxial (adp) than in the abaxial phloem (abp; Fig. 2A; data not shown). No expression was detected in the xylem (Fig. 2C) as compared to the 35S promoter control (Fig. 2D). GUS expression was also tested at the callus stage for all of the transgenic lines with results similar to the transient expression experiments in protoplasts (high GUS activity for CF2, CF4 and CF5, poor staining for CF7; data not shown). Thus, the CFDV promoter is active in differentiated (plant) and undifferentiated cells (callus, protoplasts).

Finally, the expression level of the pRTCF constructs was examined in transformed *E. coli* JM109 cells. For each construct, three or six single colonies were grown overnight at 37 °C in 3 ml of LB medium. The OD<sub>600</sub> of the bacterial cultures was then adjusted to 1 and 10  $\mu$ l of each was incubated for 1 h at 37 °C in the presence of 3 mM 4-MUG in GUS extraction buffer (50 mM NaH<sub>2</sub>PO<sub>4</sub>-Na<sub>2</sub>HPO<sub>4</sub>, pH 7, 10 mM EDTA, 0.1% Triton X-100). In all experiments, high GUS activity of the CFDV constructs was observed in comparison to the CaMV 35S 17K-GUS control (Fig. 1C). pRTCF4 and pRTCF9 were the most efficient constructs with an activity as high as 257 141 units/ $\mu$ l and 317 733 units/ $\mu$ l, respectively. These results also further highlighted the importance of the SL structure for efficient expression: the pRTCF5 construct reached 19.6% and the CaMV 35S 17K-GUS only 2.4% of pRTCF4 activity.

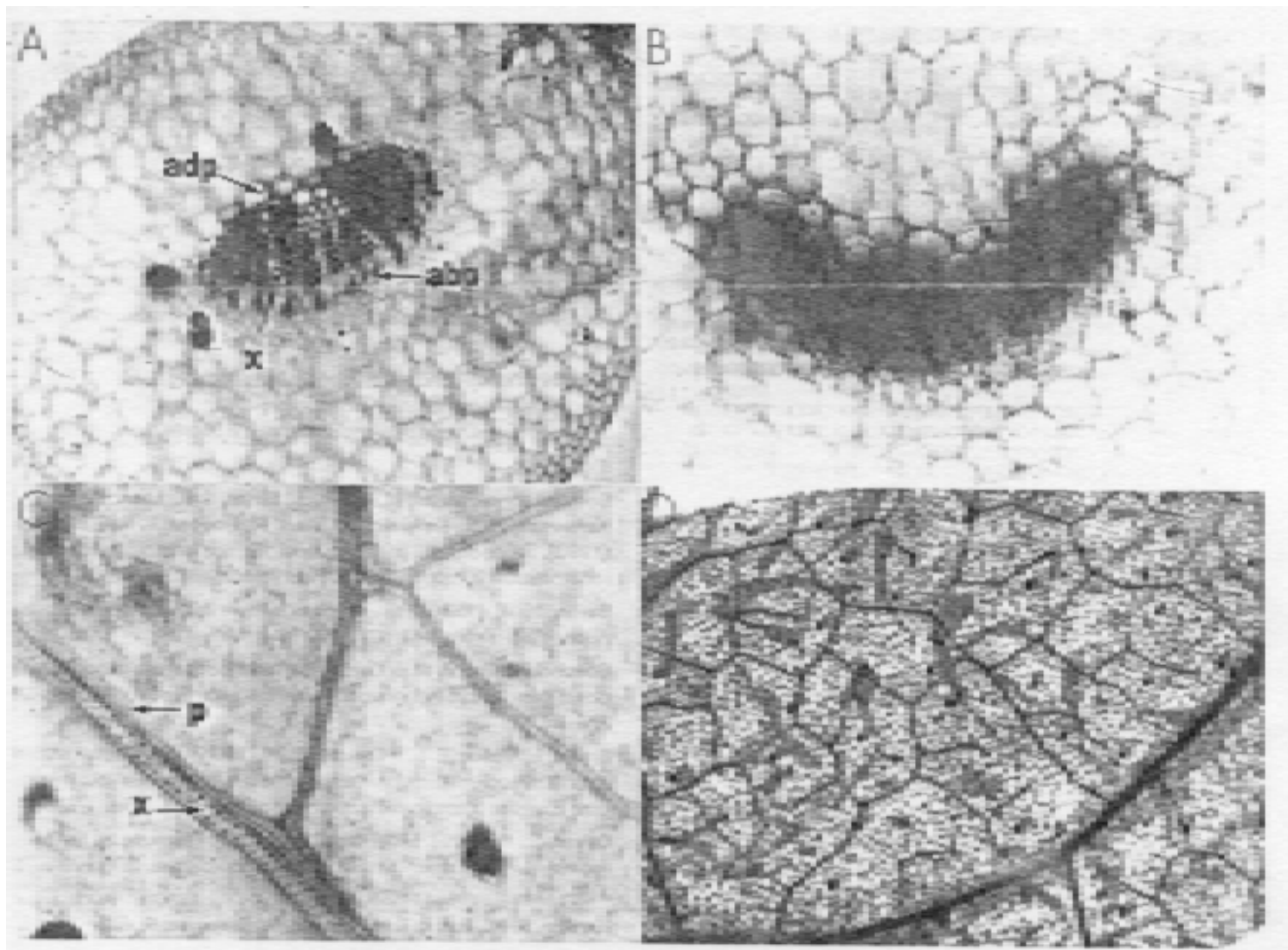


Fig. 2. Tissue specificity of the CFDV promoter in transgenic tobacco. Sections and thin-sections of leaves from transgenic tobacco plants transgenic for the CF5 CFDV-GUS construct (A, C) and the CaMV 35S-GUS construct (B, D) stained with X-gluc. abp, abaxial phloem; adp, adaxial phloem; p, phloem; x, xylem.

From the results obtained, the *cis*-acting signals for the phloem specificity of the CFDV promoter are located on a fragment of approximately 250 nt (position 711–962; Fig. 1A) which is devoid of the double repeat sequence RPT and the SL structure. The only prominent, previously described signal is a stretch of 52 nt (Rohde *et al.*, 1994) which is present in all constructs except for CF27. A sequence comparison of CFDV in this region with other phloem-specific promoters identified a highly conserved motif of 13 nt to the centre of the 52 nt stretch (Fig. 3). For rice tungro bacilliform virus (RTBV), this consensus sequence is apparently not essential for phloem specificity, since the RTBV core promoter lacks half of this box (Yin *et al.*, 1997). However, it is interesting to note that – except for CoYMV – this motif is localized upstream of and in proximity to the TATA box (169 nt for CFDV, 129 nt for RTBV, 123 nt for *rolC* and 255 nt for *GS3A*; Fig. 3). Further potential *cis*-elements of the CFDV promoter may be involved in binding factors necessary for phloem-specific promoter activity. One such candidate signal contributing to phloem specificity may be the repetitive CCA motif immediately

downstream of the 52 nt stretch. Indeed, the last CCA is near a TGG and Y. Yin and co-workers showed that CCA(N)<sub>n</sub>TGG is a recognition site for RF2a transcriptional activator to the RTBV promoter (Yin & Beachy, 1995; Yin *et al.*, 1997). Finally, the TGCTAAGT repeat localized directly upstream of the TATA box is unique to CFDV (Fig. 3). Thus, with CFDV the promoter activity in specialized, differentiated tissue (phloem) as well as in undifferentiated or dedifferentiated cells (callus, protoplasts) supports the assumption that more than one transcriptional activator protein might interact with CFDV-specific motifs.

Other strong viral and phloem-specific promoters have been described. One example is the CoYMV promoter which in transient expression experiments reached 30% of the activity of a double CaMV 35S promoter (Medberry *et al.*, 1992). The RTBV promoter also directs phloem-limited GUS activity (Bhattacharyya-Pakrasi *et al.*, 1993; Yin & Beachy, 1995; Yin *et al.*, 1997). For CFDV, the activity of pRTCF4 in tobacco cells represented 30.8% of the CaMV 35S 17K-GUS activity and, therefore, 33.9% activity of the 35S promoter, since 17K-GUS



Fig. 3. Sequence comparison of the CFDV promoter region determining phloem specificity with promoter sequences of other phloem-specific promoters. Sequence comparison of CFDV CF5 to the promoter regions of the rice tungro bacilliform virus (RTBV) promoter (Yin *et al.*, 1997), *Commelina* yellow mottle virus (Medberry *et al.*, 1993), the ro1C gene of *Agrobacterium rhizogenes* (Kiyokawa *et al.*, 1994) and the pea glutamine synthase GS3A gene (Brears *et al.*, 1991). The 52 bp region (Fig. 1A) of CFDV is underlined. Homologies between all promoters are shaded (grey box). The dotted boxes highlight the repeated sequences CCA or TGCTAAGT. The TATA boxes are shown as black boxes.

is 1.1 times more efficient than a simple 35S promoter (Tacke *et al.*, 1990). A further property of the CFDV promoter is noteworthy, namely its high activity in prokaryotic cells. As compared to CaMV 35S promoter activity in bacteria (Assaad & Signer, 1990; Fig. 1C), pRTCF4 as the most active CFDV promoter construct is more than 40 times more efficient. The *Chlorella* virus adenine methyltransferase gene promoter has also been demonstrated to be a strong promoter in plants and bacteria (Mitra *et al.*, 1994; Mitra & Higgins, 1994). In contrast to CFDV, however, its activity in plants is not limited to the phloem, but can be detected in leaves, stems and flowers.

Regulation of gene expression in different kinds of organisms at the level of transcription is of particular interest in gene technology. Testing the expression ability of constructs in bacteria prior to their use in transient expression experiments or stable transformation will not only save time, but provides

an opportunity to efficiently test every stage of cloning or DNA amplification and, thereby, prevent any deletion or knock-out mutation in the sequence of interest. However, few shuttle promoters have been described up to now. In this respect, the CFDV CF4 construct has definitely interesting and original properties and may be of particular interest for application in gene technology.

The technical help of Ms A. Kaufmann is gratefully acknowledged. A.H. was the recipient of an EC TMR fellowship.

**References**

Assaad, F. F. & Signer, E. R. (1990). Cauliflower mosaic virus P35S promoter activity in *Escherichia coli*. *Molecular & General Genetics* **223**, 517-520.

Bhattacharyya-Pakrasi, M., Peng, J., Elmer, J. S., Laco, G., Shen, P., Kaniewska, M. B., Konowicz, H., Wen, F., Hodges, T. K. & Beachy, R. (1993). Specificity of a promoter from the rice tungro bacilliform virus for expression in phloem tissues. *Plant Journal* **4**, 71-79.

Bevan, M. W. (1984). *Agrobacterium* vectors for plant transformation. *Nucleic Acids Research* **12**, 8711-8721.

Boevink, P., Chu, P. W. & Keese, P. (1995). Sequence of subterranean clover stunt virus DNA: affinities with the geminiviruses. *Virology* **207**, 354-361.

Brears, T., Walker, E. L. & Coruzzi, G. M. (1991). A promoter sequence involved in cell-specific expression of the pea glutamine synthase GS3A gene in organs of transgenic tobacco and alfalfa. *Plant Journal* **1**, 235-244.

Calves, C., Renard, J. L. & Marty, G. (1980). Tolerance of the hybrid coconut local x Rennel to New Hebrides disease. *Oléagineux* **35**, 443-451.

Harding, R. M., Burns, T. M., Hafner, G., Dietzgen, R. G. & Dale, J. L. (1993). Nucleotide sequence of one component of the banana bunchy top virus genome contains a putative replicase gene. *Journal of General Virology* **74**, 323-328.

Jefferson, R. A., Kavanagh, T. A. & Bevan, M. W. (1987). GUS fusions:  $\beta$ -glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO Journal* **6**, 3901-3907.

Julia, J. F., Dollet, M., Randias, J. & Calves, C. (1985). Foliar decay of coconut by *Myndus taffini* (FDMT): new results. *Oléagineux* **40**, 19-27.

Katul, L., Maiss, E. & Vetter, H. J. (1995). Sequence analysis of a faba bean necrotic yellows virus DNA component containing a putative replicase gene. *Journal of General Virology* **76**, 475-479.

Kiyokawa, S., Kobayashi, K., Kikuchi, Y., Kamada, H. & Harada, H. (1994). Root inducing region of mikomipine type Ri plasmid pRi1724. *Plant Physiology* **104**, 801-802.

Lam, E., Benfey, P. N., Gilmartin, P. M., Fang, R. & Chua, N. H. (1989). Site-specific mutations alter *in vitro* factor binding and change promoter expression pattern in transgenic plants. *Proceedings of the National Academy of Sciences USA* **86**, 7890-7894.

Medberry, S. L., Lockhart, B. E. L. & Olszewski, N. E. (1992). The *Commelina* yellow mottle virus promoter is a strong promoter in vascular and reproductive tissues. *Plant Cell* **4**, 185-192.

Mitra, A. & Higgins, D. W. (1994). The *Chlorella* virus adenine methyltransferase gene promoter is a strong promoter in plants. *Plant Molecular Biology* **26**, 85-93.

Mitra, A., Higgins, D. W. & Rohe, N. J. (1994). A *Chlorella* virus gene promoter functions as a strong promoter both in plants and bacteria. *Biochemical and Biophysical Research Communications* **204**, 187-194.

- Brozco, B. M. & Hanley-Bowdoin, L. (1996). A DNA structure is required for geminivirus replication origin function. *Journal of Virology* **70**, 148–158.
- Randles, J. W. & Hanold, D. (1989). Coconut foliar decay virus particles are 20-nm icosahedra. *Intervirology* **30**, 177–180.
- Rohde, W., Randles, J. W., Langridge, P. & Hanold, D. (1990). Nucleotide sequence of a circular single-stranded DNA associated with coconut foliar decay virus. *Virology* **176**, 648–651.
- Rohde, W., Becker, D. & Randles, J. W. (1994). The promoter of coconut foliar decay-associated circular single-stranded DNA directs phloem-specific reporter gene expression in transgenic tobacco. *Plant Molecular Biology* **27**, 623–628.
- Jacke, E., Prüfer, D., Salamini, F. & Rohde, W. (1990). Characterization of a potato leafroll luteovirus subgenomic RNA: differential expression by internal initiation and UAG suppression. *Journal of General Virology* **71**, 2265–2272.
- Turner, R., Bate, N., Twell, D. & Goster, G. D. (1994). *In vivo* characterization of a translational enhancer upstream from the coat protein open reading frame of potato virus S. *Archives of Virology* **137**, 123–132.
- Wu, R. Y., You, L. R. & Soong, T. S. (1994). Nucleotide sequences of two circular single-stranded DNAs associated with banana bunchy top virus. *Phytopathology* **84**, 952–958.
- Yin, Y. & Beachy, R. (1995). The regulatory regions of the rice tungro bacilliform virus promoter and interacting nuclear factors in rice (*Oryza sativa*). *Plant Journal* **7**, 969–980.
- Yin, Y., Zhu, Q., Dai, S., Lamb, C. & Beachy, R. (1997). RF2a, a bzip transcriptional activator of the phloem-specific rice tungro bacilliform virus promoter, functions in vascular development. *EMBO Journal* **16**, 5247–5259.

---

Received 28 November 1997; Accepted 16 February 1998