Cherry leaf roll virus (CLRV): genome organisation of the RNA1

S. von Bargen[1], J. Langer[1], A. Rumbou[1], J. Gentkow[2], C. Büttner[1]

[1] Humboldt-Universität zu Berlin, Institute of Horticultural Sciences, Section Phytomedicine, Lentzeallee 55/57, 14195 Berlin, Germany, susanne.von.bargen@agrar.hu-berlin.de

[2] present address: Leibniz-Institute of Plant Biochemistry, Weinberg 3, 06120 Halle, Germany

Introduction

The complete organisation of the *Cherry leaf roll virus* genome, a virus which affects many fruit trees and other woody hosts, has not been determined to date. However, partial sequence information of the bipartite virus which is available of the 3´ proximal portion including the complete 3´ non-coding region (NCR) of the genomic RNA1 and RNA2 has led to the classification as a subgroup c nepovirus.

Methods and Results

Sequences of the RNA1 of two CLRV isolates from different host plants (CLRV-E395 originating from *Rheum rhabarbarum* and CLRV-E326 from *Juglans regia*) were obtained. The genomic structure of the CLRV-RNA1 coding for a polyprotein (PP1) was compared with other established subgroup c nepoviruses like *Tomato ringspot virus* (ToRSV), *Blackcurrant reversion virus* (BRV) and *Peach rosette mosaic virus* (PRMV). The polyprotein of the rhubarb isolate (ORF$_{12\text{-}6350\ nt}$; 2112 amino acids) contains a N-terminal protease cofactor (**PCo**), adjacent is a nucleotide-binding protein-domain (**NTBp**), followed by the sequences coding for the genome-linked viral protein (**VPg**), a protease (**Pro**) and the viral replicase (**RdRp**). Putative protein functions were predicted by identification of characteristic sequence motifs. The region coding for the CLRV-VPg protein was identified with the computer programs NetPicoRNA V1.0 and NetCorona V1.0., and exhibited highest similarities with the corresponding ToRSV-VPg. Predicted specific protease recognition sequences in the CLRV isolates (Q_{1121}/S_{1122} and Q_{1150}/S_{1151}) also corresponded to ToRSV.

Conclusion

The genome organisation of the RNA1 of CLRV strains E395 from rhubarb and E326 from the woody host walnut corresponds with other subgroup c nepoviruses.

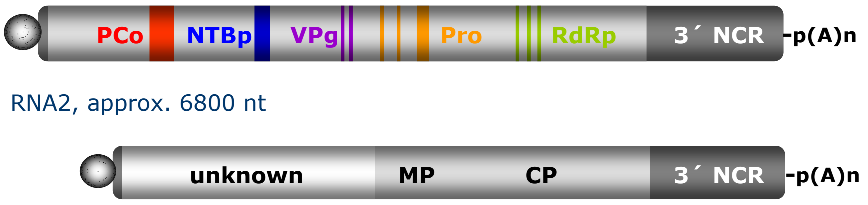
References

Argos (1988) NAR 16, 9909-9919; **Gorbalenya** et al. (1989a) J. Mol. Evol. 24, 256-268; **Gorbalenya** et al. (1989b) FEBS letters 243, 103-114; **Rott** et al. (1995) JGV 76, 465-473; **Wang** et al. (1999) JGV 80, 799-809.

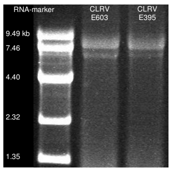
Genome organisation of CLRV-E395

RNA1, 7918 nt (ORF nucleotides: 12-6350; polyprotein 1: 2112 amino acids)

RNA2, approx. 6800 nt



viral RNAs



Viral RNAs of CLRV strains E603 and E395 separated by denaturating agarose gel electrophoresis.

protease-cofactor (PCo) motif

```
                    490       500       510       520       530       540       550       560
               ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
CLRV-E395      FVRALRDEFSNSVGKYLNAVSEYKAQIENFWAWAVRWSNNLMNKVDVSLRALQGSAFFAAALVIVGGIVYLIENLLPPAA 467
CLRV-E326      FVRALRDEFSNSVGKYLNAVSEYKAQIENFWAWAVRWSNNLMNKVDVSLRTLQGSAFFAAALVIIGGIVYLIENLLPPAT 467
BRV-NC003509   FFQTLKSKFFEFLKPYIQHAIYASAEIEKYWAFIHGWATKMWNNVGVELQALGDAAWWAIGITMVCGIVTLVEKLLVYLG 463
ToRSV-NC003840 FVKLLREHFDDTIGKWIPKLLGATQKIEELWRWSLEWAQNMSKKLDVSLRVLRGSALVGVGLLLVSGILYFAEQLLRSFG 553
PRMV-AF06626   ILDKCKYCFMSTFSPFLASLQSAKAEIEKFWQNCMSWARNLWSKAHLALQALGLYAIWALVLTILCGIVYLLESLFITAG 455
Consensus      :.   :  * .  .  ::      . :**: *     *: :: .:  : *:.*    *   .  : :: **: : *.*:
                       FxxxxxxxxxxxxxxxxxxxxxxxxxxxWxxxxxxxxxxxLxxxxxxxxxxxxxxxxxxxxxLxE
```

F-x$_{27}$-W-x$_{11}$-L-x$_{21}$-L-x$_{1}$-E (Rott et al., 1995)

NTB protein (NTBp) motif

```
                    810       820       830       840       850       860       870       880
               ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
CLRV-E395      YFFGKARTGKTIFANNFKNTMVQYLGTTSENIFYKNARDQFWPKYRQQAIVIVDDLSAVENPPS--LESEFIQLMSTMPY 778
CLRV-E326      YFFGKARTGKTIFANNFKNTMVQYLGTTSENIFYKNARDQFWPKYRQQAIVIVDDLSAVENPPS--LESEFIQLMSTMPY 778
BRV-NC003509   YIYGKSHCGKSLFMEDVSRALLKENGHAPNDIYAKNARDSFWSGYLQHACVQVDDLSACVTRPS--LESEFLQLVGSKMY 773
ToRSV-NC003840 YLYGGPRCGKSLFAQSFMNAAVDFMGTTVDNCYFKNARDDFWSGYRQEAICCVDDLSSCETQPS--IESEFIQLITTMRY 870
PRMV-AF06626   YIFGASQCGKSTLANYLCPLLLAHMGWDAHDVYSKDPTEGYWSGYYQQKCLKMDDLSAVVPKQVSPLEQQLIPLISTEEK 772
Consensus      *::* .: **: : : .     :   *   .: : *:. : :*. * *.     :****:        :*.::: *: :
                  GxxGxGKS/TxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxDD
```

Gxx(G)xGKS/T-x$_n$-D/ED (Gorbalenya et al., 1989a)

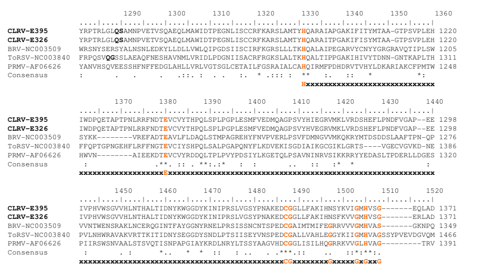
putative VPg (M: 3.28 kDa) with flanking 3C-like protease (Pro) recognition sites

```
                    1240      1250      1260      1270      1280      1290      1300      1310
               ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
CLRV-E395      RQIFCGGVG-AGSAGAMMKMSVQS-------TIPSGSDVQSYASRNLRRVYRPTRLGLQSAMNPVETVSQAEQLMAWIDT 1171
CLRV-E326      RQIFCGGVG-AGSAGAMMKMSVQS-------TVPSGSDVQSYASRNLRRVYRPTRLGLQSAMNPVETVSQAEQLMAWIDT 1171
BRV-NC003509   QTFVAGTSSALVSAGMITQLSANAGSVCTSASNPSGAASYVSSNIPIHHRWRSNYSERSYALNSNLEDKYLLDLLVWLQI 1155
ToRSV-NC003840 QAIFCGAAGGTVSMAAVGKMTVQS-------TIPSGSYADVYNARNMTRVFRPQSVQGSSLAEAQFNESHAVNMLVRIDL 1262
PRMV-AF06626   KTFLSGSVGFGSALALKNQLDVHSSVASSGSIATQSYARSIPIVWAKVARYANVHSQVEESSHFNFFEDGLAHLLVRLVG 1198
Consensus      : :..*  .   : .   :: .::          ...           :       .   .     .  .::. :  274
```

predicted by NetPicoRNA V1.0 (0.546) NetCorona V1.0 (0.545)
NetCorona V1.0 (0.677)

Q/E ✂ G/S (Wang et al., 1999)

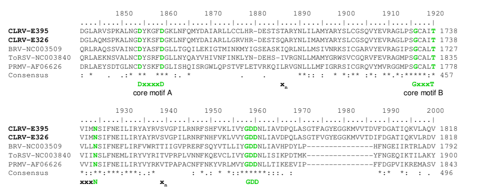
viral cysteine protease (Pro) motif



H-x$_{30\text{-}46}$-E-x$_{69\text{-}101}$-CG-x$_{6\text{-}8}$-G-x$_{3\text{-}5}$-G-x-H/LxxG
(Gorbalenya et al., 1989b)

H-x$_{47}$-E-x$_{104}$-CG-x$_{6\text{-}8}$-G-x$_{3\text{-}5}$-N-x-H/LxxG CLRV, nepoviruses subgroup C

viral replicase (RdRp) core motif



DxxxD-x$_n$-DxxxTxxxN-x$_n$-GDD
(Argos, 1988)

Acknowledgements

Research was funded by the DFG projects Bu890/8-1 and Bu890/8-2.