Virus Research 163 (2012) 678–683

Contents lists available at SciVerse ScienceDirect

Virus Research

journal homepage: www.elsevier.com/locate/virusres

Short communication

# Complete nucleotide sequence of *Cherry leaf roll virus* (CLRV), a subgroup C nepovirus

Susanne von Bargen*, Juliane Langer, Jenny Robel, Artemis Rumbou[1], Carmen Büttner

*Humboldt-Universität zu Berlin, Department of Crop and Animal Sciences, Division Phytomedicine, Lentzeallee 55/57, D-14195 Berlin, Germany*

ARTICLE INFO

*Article history:*
Received 28 October 2011
Received in revised form
22 December 2011
Accepted 26 December 2011
Available online 31 December 2011

*Keywords:*
CLRV rhubarb isolate
Genome organization
Protein domains
Terminal sequence duplication

ABSTRACT

The complete nucleotide sequence of both genomic (+)ss RNAs of a rhubarb isolate of *Cherry leaf roll virus* (CLRV) was determined. The larger RNA1 is 7918 nucleotides and the shorter RNA2 6360 nucleotides in size, each genome component comprising a single open reading frame (ORF). The RNA1-encoded polyprotein (P1) is 2112 amino acids long (235.6 kDa) containing domains characteristic for a proteinase-cofactor (PCo), nucleotide-binding helicase (Hel), genome-linked protein (VPg), proteinase (Pro), and an RNA-dependent RNA polymerase (Pol). The RNA2-encoded polyprotein (P2) has a molecular mass of 174.9 kDa (1589 aa) encoding the putative movement protein (MP) and the coat protein (CP) of CLRV. The genome region upstream of the MP has a coding capacity of 77 kDa, however processing of P2 by the putative virus-encoded proteinase and protein-function encoded by this region is unknown. Furthermore, it could be demonstrated that the 5′-termini including the N-terminal region (208 aa) of P1 and P2 of the rhubarb isolate of CLRV are nearly identical among the two genome segments.

The taxonomic position of CLRV as member of the genus *Nepovirus* was confirmed by phylogenetic analyses employing the amino acid sequences of the conserved Pro-Pol region of RNA1, the complete P2, and the CP. However, clustering of *Nepovirus*-species according to allocated subgroups was inconsistent and depended on the compared genome fragment.

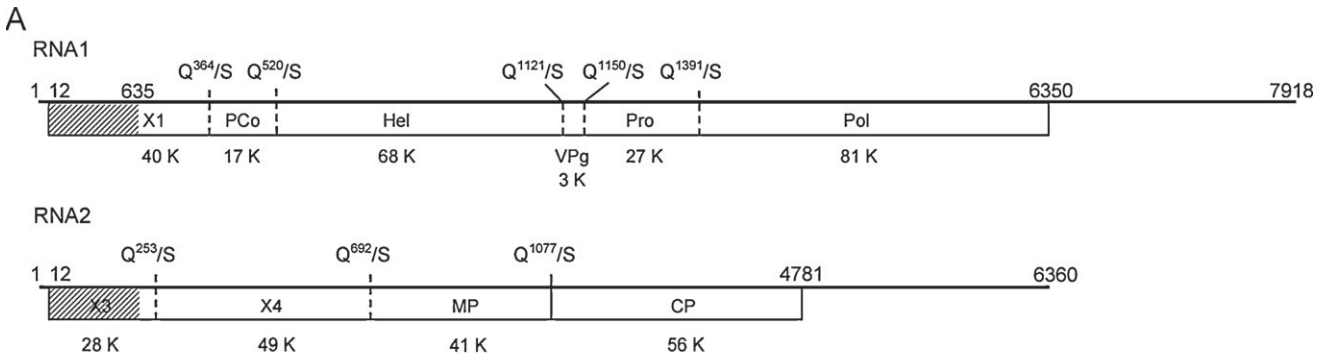


*Cherry leaf roll virus* (CLRV) was reported for the first time in the 1930s in English walnut (*Juglans regia* L.) and sweet cherry (*Prunus avium* L.). Since then, numerous CLRV-infected plant species have been recorded worldwide revealing its wide natural host range which comprises at least 18 genera of broad-leaved trees and shrubs as well as a variety of herbaceous plants (summarized in Büttner et al., 2011). The virus has been reported to infect rhubarb (*Rheum rhabarbarum* L.) in the United Kingdom (Scott et al., 1992; Tomlinson and Walkey, 1967) and has been repeatedly detected in plants in Germany exhibiting chloroses and ringspots on leaves (Rebenstorf, 2005). Available partial sequences of certain CLRV isolates mostly representing 3′-proximal regions of the genomic RNA1 or RNA2 together with morphological and biological characteristics of the virus have led to the classification as a nepovirus, subgroup C within the newly established family *Secoviridae* (Sanfacon et al., 2009) of the order *Picornavirales* (Le Gall et al., 2008). Until now, 3 genomes of members of the subgroup C nepoviruses have been completely sequenced, i.e. *Tomato ringspot virus* (ToRSV, Rott et al., 1991a, 1995), *Blackcurrant reversion virus* (BRV, Latvala-Kilby and Lehto, 1999; Pacot-Hiriart et al., 2001), and *Grapevine Bulgarian latent virus* (GBLV, Elbeaino et al., 2011) showing that the available molecular information of this subgroup is still limited.

Here, we report the complete sequence of a rhubarb isolate of *Cherry leaf roll virus*. The CLRV isolate E395 was originally obtained in 1987 from rhubarb (*Rheum rhabarbarum*) collected in Bornheim (Germany) by transmission to biotest plants. The virus isolate was propagated in *Chenopodium quinoa* (L.), which was also used as source material for the purification of virus particles by adaption of a protocol of Wood (1998). Purified viral particles were used for construction of a random primed cDNA-library followed by PCR-amplification according to Froussard (1992). PCR-products were cloned applying standard protocols (Sambrook et al., 1989). Plasmids containing cDNA-inserts above 1000 bp were selected and sequenced using a "BigDye® Terminator v1.1 Ready Reaction Cycle Sequencing Kit" applying standard protocols and an ABI PRISM® 310 Genetic Analyzer from Applied Biosystems (USA). Received sequences were subjected to a nucleotide database search using Nucleotide BLAST (Altschul et al., 1990). cDNA-clones exhibiting significant identities to partial sequences of other CLRV isolates or related nepoviruses in the database were completely sequenced by primer walking. Gaps between non-overlapping cDNA-library clones were closed by RT-PCR using new primers based on received sequence data and information of available CLRV isolates in the NCBI database. In order to sequence the 3′ non-coding regions

* Corresponding author. Tel.: +49 30 2093 46447; fax: +49 30 2093 46450.
*E-mail address:* susanne.von.bargen@agrar.hu-berlin.de (S. von Bargen).
[1] Permanent address: Plant Protection Institute, National Agricultural Research Foundation, P.O. Box 1303, 38011 Volos, Greece.

0168-1702/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.virusres.2011.12.018




B CLRV polyprotein 1

| Aa | clve | surface | peptide | cleavage between putative proteins |
|---|---|---|---|---|
| $Q^{66}$ | 0.175 | 0.504 | IVCRQ/SVES | |
| $Q^{364}$ | 0.242 | 0.530 | EGARQ/SAID | X1/PCo |
| $Q^{520}$ | 0.283 | 0.506 | NKA**V**Q/SGNM | PCo/Hel |
| $Q^{1121}$ | 0.538 | 0.546 | KMS**V**Q/STIP | Hel/VPg |
| $Q^{1131}$ | 0.252 | 0.492 | GSD**V**Q/SYAS | |
| $Q^{1150}$ | 0.209 | 0.567 | RLG**L**Q/SAMN | VPg/Pro |
| $Q^{1391}$ | 0.387 | 0.537 | FMS**A**Q/SGLR | Pro/Pol |
| $Q^{1929}$ | 0.402 | 0.576 | RDFHQ/SQYS | |

polyprotein2

| Aa | clve | surface | peptide | cleavage between putative proteins |
|---|---|---|---|---|
| $Q^{66}$ | 0.175 | 0.504 | IVCRQ/SVES | |
| $Q^{253}$ | 0.242 | 0.618 | RMT**L**Q/SAQT | X3/X4 |
| $Q^{496}$ | 0.201 | 0.665 | ALW**A**Q/SQVS | |
| $Q^{692}$ | 0.318 | 0.606 | YLP**A**Q/SGIG | X4/MP |
| $Q^{1077}$ | 0.329 | 0.552 | NMP**L**Q/SGIL | MP/CP |

**Fig. 1.** (A) Schematic representation of the genome organization of *Cherry leaf roll virus* (CLRV). Numbers above the genome segments of the virus indicate nucleotide positions including beginning and end of the open reading frames. Nearly identical N-terminal regions among the first 635 nucleotides between both genome components are indicated by a striped box. Boxes represent RNA1- and RNA2-encoded polyproteins with predicted cleavage sites (Q/S) at respective positions including putative function and molecular mass of mature proteins. (B) Putative cleavage sites of the polyproteins 1 and 2 by the RNA1-encoded 3C-like proteinase of the rhubarb isolate of CLRV as predicted by NetPicoRNA 1.0. Calculated cleavage (clve) and surface probabilities (surface) of the respective peptides are indicated as well as cleavage between the conserved amino acids glutamine (Q) and serine (S). Small, hydrophobic amino acids conserved in position -2 are given in bold.

(NCRs) specific for viral RNA1 and RNA2, respectively, the universal CLRV primer RW1 (Werner et al., 1997) was applied in cDNA-synthesis. Segment specific 3′ NCRs were amplified by combination of RW1 with a forward primer located within the Pol-coding region of RNA1 or the CP-coding region of RNA2, respectively. Either the 5′ RACE-PCR system (Roche Diagnostics, Germany) or inverse PCR (Ochman et al., 1990) was used for determination of the 5′-termini of viral RNA1 and RNA2. At least a 2 fold-coverage of the CLRV genome was achieved by independently amplified fragments. In case of sequence inconsistencies which might be attributed to PCR or sequencing artifacts, a third fragment was generated and sequenced. Complete nucleotide and protein sequences were assembled in BioEdit 7.0.9.0 (Hall, 1999) and aligned with ClustalX 2.0 (Larkin et al., 2007) applying default parameters.

The genome of CLRV is composed of two (+) single stranded RNA species as confirmed by sequence analysis. RNA1 of CLRV-E395 is 7918 nucleotides in length (EMBL ID: FR851461), RNA2 is 6360 nucleotides long (EMBL ID: FR851462). Sequences reported here exclude the last two nucleotides of the 3′ non-coding region as well as the polyA-tail. However, the presence of the ultimate 3′-terminus of the CLRV genomic RNAs was confirmed for the isolate. The 3′-terminus including the polyA-tail was amplified by RT-PCR using M4T and M4-primer (Chen et al., 2001) in combination with RW2 (Werner et al., 1997), and subsequently sequenced (data not shown). As products amplified with these primers are not specific for one of the genomic RNAs of CLRV, the ultimate 3′-terminal nucleotides including the polyA-tail were not included in the sequence annotation.

RNA1 of CLRV encodes a single ORF in frame 3 with a length of 6339 nucleotides (Fig. 1). The first potential initiation codon of the sequence at position 12–14 is the only in frame start codon in a heptamer AAAAUGG context, which is optimal for translation initiation in dicotyledonous plants (Pesole et al., 2000), and known to enhance translation in eukaryotic cells when containing an A in position −3, and a G in position +4 (Kozak, 2005). Assuming that translation begins at this first AUG, the translated product of CLRV-RNA1 is a polyprotein precursor comprising 2112 amino acids with a calculated Mr of 235,600. RNA2 of CLRV encodes one large ORF in frame 3 with a length of 4770 nucleotides. Corresponding to viral RNA1 we assume that translation of RNA2 is initiated at the first AUG, because of sequence identity of the 5′-terminal regions of both genome components. The P2 encoded by CLRV-RNA2 is a polypeptide of 1589 amino acids with a calculated Mr of 174,900. Putative cleavage sites within polyprotein 1 (P1) and polyprotein 2 (P2) by the virus-3C-like proteinase were predicted by alignment of CLRV and ToRSV, the latter containing experimentally confirmed cleavage sites (Carrier et al., 2001; Chisholm et al., 2001; Wang et al., 1999; Wang and Sanfacon, 2000). Identified sites were further examined by the program NetPicoRNA1.0 (Blom et al., 1996) using the settings for the 3C-like proteinase of the genus *Aphthovirus* affiliated to the family *Picornaviridae*. The proteinase recognition sites confirmed for ToRSV are correlating with the analysis of the cleavage sites predicted for the putative proteinase encoded by the rhubarb isolate of CLRV. For P1 they were found to be positioned in good accordance with experimentally confirmed proteinase cleavage sites reported for ToRSV. On the contrary, cleavage sites predicted in P2 of CLRV are less corresponding with the equivalent polypeptide sequence of ToRSV. This is partly due to the different size of P2 which is 175 kDa in case of CLRV and was reported to be 207 kDa for ToRSV (Rott et al., 1991a). However, the Q/S cleavage site between the RNA2-encoded MP and CP was experimentally confirmed by Scott et al. (1993) for another CLRV isolate derived from birch. It may be worth mentioning that with the exception of the putative cleavage site between X1 and PCo valine, leucine, or alanine is found in position -2 adjacent to all cleavage sites of P1 and P2 of CLRV postulated here (Fig. 1B). Such small hydrophobic amino acids were found to be critical for the recognition of the cleavage site by the ToRSV-encoded proteinase (Carrier et al., 1999). However, additional investigations are required to clarify the processing of the polyproteins by the CLRV-encoded 3C-like proteinase.

P1 of CLRV contains a proteinase cofactor domain (PCo, Rott et al., 1995), adjacent is a putative helicase (Hel) containing a nucleoside triphosphate-binding protein-domain (NTB, Gorbalenya et al., 1989a), followed by the genome-linked viral protein (VPg, Hellen and Cooper, 1987), a proteinase (Pro, Gorbalenya et al., 1989b) and the viral replicase (Pol, Koonin and Dolja, 1993) as displayed in Fig. 1A. The N-terminal 395 aa upstream of the PCo domain of CLRV P1 do not contain a sequence motif recognized by blastp or affiliated programs. However, it displays some similarities to the 5′-proximal-coding regions of the other subgroup C nepoviruses ToRSV and BRV, respectively. The N-terminal amino acid stretch of CLRV P1 is rich in alanine residues (54 out of 395 amino acids, 13.7%). This result also corresponds to observations reported from the other subgroup C nepoviruses ToRSV (Wang and Sanfacon, 2000) and BRV (Latvala-Kilby and Lehto, 1999).

The characteristic amino acid motif of the putative viral proteinase cofactor (PCo) previously identified among other nepoviruses (Rott et al., 1995 and references therein) is located between positions $F^{396}$ and $E^{460}$ of P1 of CLRV. It was demonstrated by Zhang and Sanfacon (2006) that the corresponding region (X2) of ToRSV is highly hydrophobic and contains several transmembrane helices targeting the protein to the endoplasmic reticulum (ER). Considerable amounts (35.3%) of hydrophobic amino acids are also present in the putative PCo ($S^{365}$–$Q^{520}$, 17 kDa) of CLRV-E395 and are predominantly located near the C-terminal part of the protein. The Hel domain present in CLRV P1 between position $G^{704}$ and $N^{802}$ indicates that the putative protein ($S^{521}$–$Q^{1121}$, 68 kDa) belongs to helicase superfamily III (Gorbalenya et al., 1990; Koonin and Dolja, 1993). The protein is involved in NTP-binding and hydrolysis (Le Gall et al., 2008). Specific 3C-like proteinase recognition sequences $Q^{1121}$/S and $Q^{1150}$/S within P1 of CLRV predicted by NetPicoRNA correspond to cleavage sites confirmed for ToRSV flanking the VPg region. The consensus sequence for nepoviral VPg-proteins postulated by Mayo and Fritsch (1994) is also present within the putative VPg-coding region of the *Cherry leaf roll virus* RNA1. However, with the exception of the characteristic tyrosine residue, the motif is poorly conserved within the putative VPgs of investigated subgroup C nepoviruses CLRV, ToRSV, BRV, GBLV, *Peach rosette mosaic virus* (PRMV), and *Artichocke yellow ringspot virus* (AYRSV). Downstream of the $Y^{1133}$ included in the putative VPg of CLRV-E395 several arginine residues are alternating with hydrophobic amino acids supporting the observation of Thompson et al. (2002) that basic amino acids which are surrounded by bulky hydrophobic residues are often found in nepovirus-encoded VPgs. The putative VPg of the rhubarb isolate of CLRV consists of 29 amino acids with a calculated molecular mass of 3.3 kDa which is in good accordance with the molecular mass of 3.5 kDa of the VPg of a birch isolate of CLRV estimated by denaturating PAGE by Hellen and Cooper (1987). Notably, the VPg-coding region of CLRV-E395 isolate shared 90.8% identity on nucleotide level with the VPg of a recently sequenced CLRV fragment covering over 4000 nucleotides of the 3′ terminal region of RNA1 originating from nasturtium (accession no. GU167974, Marais et al., 2010) and is perfectly conserved in the amino acid sequence. The putative 3C-like proteinase of CLRV is located downstream of the VPg. It is most probable released from the P1 precursor protein between positions $Q^{1150}$/S and $Q^{1391}$/S yielding a mature protein of 27 kDa. The catalytic triad which is part of the conserved domain of viral cysteine proteases described by Gorbalenya et al. (1989b) is present in the putative Pro of the CLRV-E395 isolate ($H^{1192}$, $E^{1240}$, and $C^{1345}$). A second histidine in the Pro domain has been shown to be essential for substrate specificity of the ToRSV proteinase (Hans and Sanfacon, 1995). It is also conserved within the RNA1-encoded putative viral proteinase of the rhubarb isolate of CLRV ($H^{1363}$). However, the Pro domain of this isolate is imperfectly conserved, because of a G to N substitution at position 1355, which applies also for the nasturtium isolate of CLRV. Thus, this residue is probably not essential for enzymatic activity or specific substrate recognition of 3C-like proteinases of picorna-like viruses (Le Gall et al., 2008). Further, the Pro-coding region of CLRV isolates exhibit high sequence conservation of 90.4% on nucleotide and 97.0% identity on protein level. The Pol-coding regions of CLRV originating from rhubarb and nasturtium are less conserved than the Pro-coding regions. The 2166 nucleotide long sequence of the rhubarb isolate of CLRV, most probably yielding a mature protein of 81 kDa after cleavage at position $Q^{1391}$/S, shows 87.6% nucleotide identity to the corresponding region of the virus isolated from nasturtium resulting in 95.5% identity at the protein level. Conserved regions spanning motifs I to VIII of the RdRp, as defined by Koonin and Dolja (1993), are present in the RNA-encoded P1 of isolate E395 and are well conserved between both CLRV-variants. In the rhubarb isolate the domains are located between positions $E^{1593}$ and $S^{1868}$. The conserved peptide KxE in motif I, the core sequence B comprised in motif V, and the core sequence YGDD in a strong hydrophobic context in motif VI confirm that the putative Pol-coding region of CLRV contains a RNA-dependent-RNA polymerase (RdRP) of the supergroup I characteristic for members belonging to the order *Picornavirales* (Le Gall et al., 2008 and references within).

Protein blast of the N-terminal region of P2 of CLRV (253 aa) reveals a low similarity to the N-terminus of the P1 polyprotein encoded by the RNA1 of BRV (score 43.5, *E* value 0.01). The amino acid sequence between $M^{359}$ and $E^{419}$ downstream of the N-terminus of CLRV P2 produced 46% amino acid identity in blastp to the repeat region in the X4-protein of the polyprotein encoded by ToRSV-RNA2. Sequence alignment of the X4-protein of three different ToRSV isolates with the equivalent region of the CLRV-RNA2-encoded polyprotein confirmed that a short stretch of 30 amino acids ($L^{390}$ to $E^{419}$) is conserved in all four sequences. Notably, this conserved stretch is part of a repetitive amino acid sequence found in the ToRSV-X4 protein (Jafarpour and Sanfacon, 2009), but it is not duplicated within the CLRV P2 sequence. The motif described for MPs of nepoviruses by Mushegian (1994) containing a conserved proline in a hydrophobic context is also present in the putative MP of the rhubarb-isolate of CLRV ($P^{915}$). The equivalent region of CLRV P2 ($S^{693}$–$Q^{1077}$, 41 kDa) shows closest relationship to the putative MP of ToRSV exhibiting 58.1% amino acid identity. Therefore, we conclude that this polypeptide represents the MP of CLRV; however, its biological function as well as the exact size of the mature MP of CLRV needs experimental confirmation. The only region recognized by the blastp affiliated CDD-database (Marchler-Bauer et al., 2009) is the coat protein of nepoviruses located at the C-terminus consisting of three domains forming a characteristic pseudo T = 3 icosahedral capsid structure (Le Gall et al., 2008). In accordance with previous studies carried out by Pallas et al. (1991) and Zhou et al. (1998) who investigated several walnut strains of CLRV, the CP-coding region of the rhubarb isolate of CLRV is 1536 nucleotides long encoding 512 aa. The resulting Mr of 56,340 of the capsid protein of the rhubarb isolate correlates well with the molecular mass of 56 kDa previously reported for the other CLRV isolates.

Sequence conservation of the 5′-proximal region as well as the 3′-terminus of the genomic RNAs of CLRV-isolate E395 was investigated by northern blot analysis using E395-specific digoxigenin-labeled probes targeting the terminal regions. In parallel, an RNA1- as well as an RNA2-specific probe was developed. Hybridization of probes to total nucleic acid preparations from CLRV-infected *C. quinoa* plants under stringent conditions confirmed the sequence duplication between both genomic RNAs of the rhubarb isolate of CLRV (Fig. 2). As expected, the probes targeting the genomic region between nucleotides 579 and 814 of RNA1 and RNA2, respectively, were specific for one genomic RNA of CLRV. These findings correspond to the sequence analysis of both

A

| probe | position within CLRV genome [nt] | size [bp] |
|---|---|---|
| 5′ region | 8-216 (RNA2) | 209 |
| RNA1 | 599-814 | 216 |
| RNA2 | 579-770 | 192 |
| 3′ region | 7507-7918 (RNA1), 5949-6360 (RNA2) | 412 |

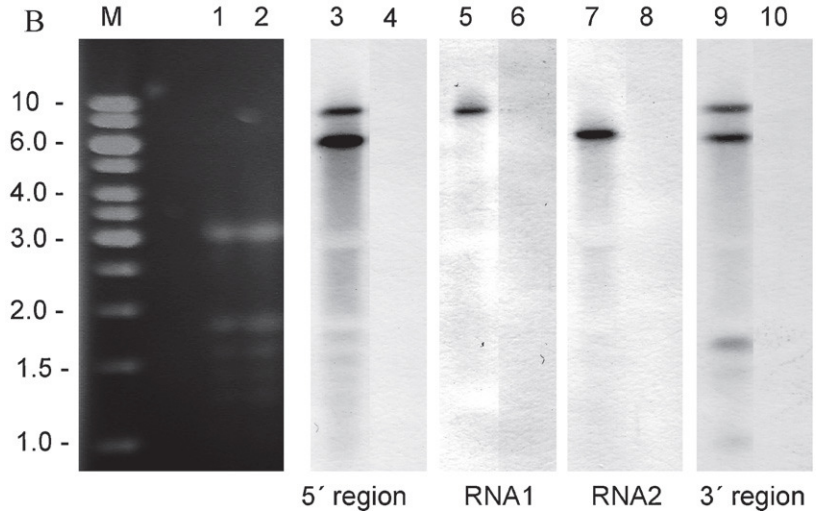



**Fig. 2.** Determination of sequence duplication between the 5′- and 3′-terminal regions of CLRV genome segments by northern blots. (A) Overview of CLRV-specific DNA-probes including their position and size. (B) Northern blot of total RNA extracted from CLRV-E395 infected (lanes 3, 5, 7, 9) and healthy (lanes 4, 6, 8, 10) *Chenopodium quinoa* leaves. Loading control of total RNA obtained from CLRV-infected (lane 1) and healthy (lane 2) plants are displayed in the agarose gel prior to blotting using the 1 kb ladder (M, Fermentas, St. Leon-Rot, Germany) as size marker. Genomic regions targeted by the probes are indicated below the blots.

genomic RNAs of CLRV-E395. The first 635 nucleotides of the 5′-termini of CLRV-RNA1 and -RNA2 are nearly identical exhibiting 84.7% sequence identity on nucleotide level. The stretch includes the 11 nucleotides of the 5′ NCRs which are perfectly conserved among both genomic RNAs. Notably, the short 5′ NCRs of CLRV also contain the complete consensus sequence reported for nepoviruses by Fuchs et al. (1989). The coding region of the ORFs exhibits 88.9% identity among the first 208 amino acids of RNA1-encoded P1 and

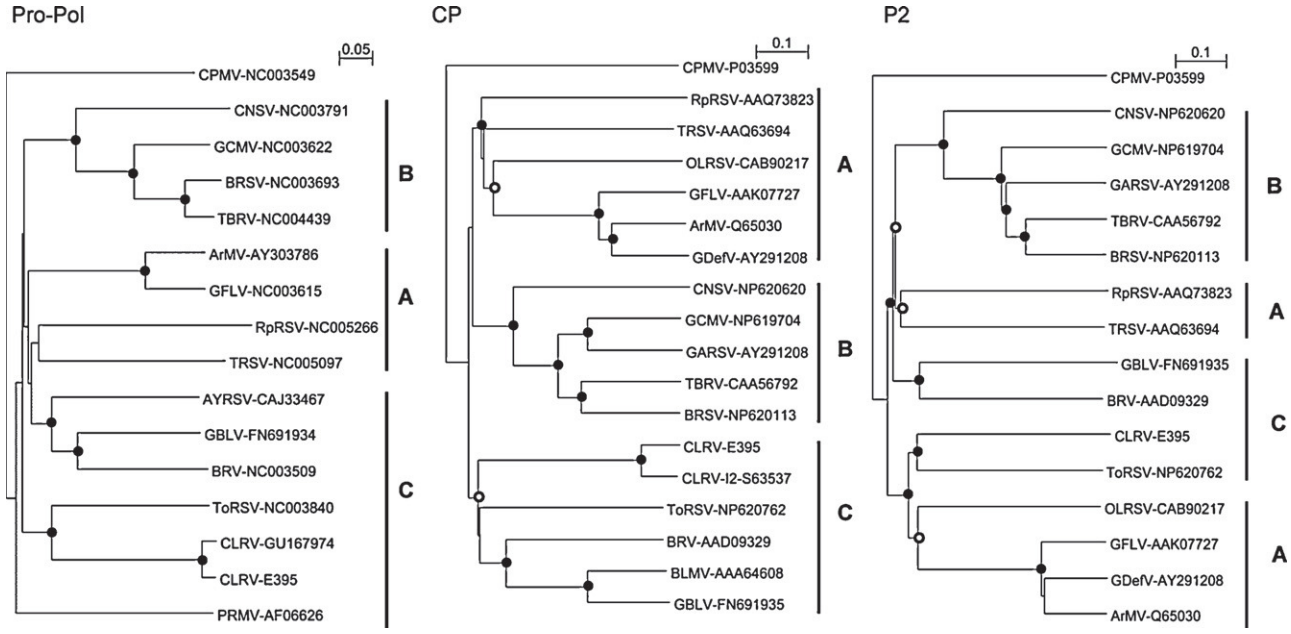



**Fig. 3.** Phylogenetic trees were inferred from amino acid sequence alignments of the Pro-Pol region starting at the domain containing the CG-motif of the proteinase and ending with domain VI including the GDD motif of the polymerase as proposed by Le Gall et al. (2008) (left), the coat-protein (CP) region covering the three conserved N-terminal, central and C-terminal domains (middle), and the full length polyprotein (P2) encoded by RNA2 of members of the genus nepovirus. Neighbor-joining trees were generated with ClustalX 2.0 including bootstrap analysis with 1000 repetitions. Nodes supported by bootstrap values above 80% are shown as black dots; values between 70% and 79% are represented by open circles. Nepovirus subgroups A, B, and C are indicated on the right side of each tree. The scale bar of 0.05 represents 5 amino acid substitutions per 100 amino acids of the aligned sequences, and 0.1 indicates 10 substitutions per 100 amino acids respectively. Virus acronyms for *Arabis mosaic virus* (ArMV), *Artichocke yellow ringspot virus* (AYRSV), *Beet ringspot virus* (BRSV), *Blackcurrant reversion virus* (BRV), *Blueberry leaf mottle virus* (BLMV), *Cycas necrotic stunt virus* (CNSV), *Grapevine Anatolian ringspot virus* (GARSV), *Grapevine Bulgarian latent virus* (GBLV), *Grapevine chrome mosaic virus* (GCMV), *Grapevine deformation virus* (GDefV), *Grapevine fanleaf virus* (GFLV), *Olive latent ringspot virus* (OLRSV), *Peach rosette mosaic virus* (PRMV), *Raspberry ringspot virus* (RpRSV), *Tobacco black ring virus* (TBRV), *Tobacco ringspot virus* (TRSV), *Tomato ringspot virus* (ToRSV) are shown with their respective GenBank accession numbers. Cowpea mosaic comovirus (CPMV) was used as outgroup.

RNA2-encoded P2 of CLRV. Such sequence duplication between the N-termini of P1 and P2 has also been described for the subgroup C nepovirus ToRSV by Rott et al. (1991b), but was neither found in the completed genomes of BRV nor GBLV. An additional approx. 1.6 kb fragment detected by the probe targeting the 3′ NCR of CLRV was already described by Brooks and Bruening (1995) as subgenomic RNA2 containing the complete 3′ non-coding region of the virus.

Comparing the 3′ non-coding regions of the rhubarb isolate of CLRV, the RNA1-3′ NCR (1568 nt) being slightly shorter than the 3′ NCR of RNA2 (1579 nt), also revealed a very high degree of sequence conservation of 98.6%. This is in accordance with findings reported by Scott et al. (1992) who showed that both genomic 3′ NCRs of a CLRV-birch isolate ($I_2$) were nearly identical (98%) and also exhibited considerable sequence conservation with another CLRV isolate obtained from rhubarb (R25). Additionally, Borja et al. (1995) analyzed the 3′ NCRs of a CLRV isolate obtained from walnut (W8) and also showed that the 3′-proximal untranslated regions are of similar length and highly conserved exhibiting 99.5% sequence identity.

Taxonomic position of *Cherry leaf roll virus* within the genus *Nepovirus* was investigated addressing the conserved Pro-Pol region of RNA1 as the most suitable region to determine taxonomic positions between Picorna-like viruses as suggested by Le Gall et al. (2008). Additionally, the more variable P2 including the CP cistron encoded by RNA2 was used for comparison (Fig. 3). Phylogenetic analyses were carried out using the neighbor-joining algorithm included in the program ClustalX 2.0. The neighbor-joining phylogenetic tree constructed from the conserved domains of the CP among members of the genus *Nepovirus* confirmed the current taxonomic position of CLRV within subgroup C. Both CLRV-isolates obtained from rhubarb and birch, respectively, clustered as a sister group of ToRSV, BRV, *Blueberry leaf mottle virus* (BLMV), and GBLV. All CP coding regions of species included in subgroup C exhibit low amino acid identities between 22.6% and 25.0%. This clade was clearly separated from the other members of the nepoviruses representing subgroups A and B, respectively being supported by bootstrap values above 70%. However, the CP-derived tree and the phylogram based on the alignment of the RNA1-encoded P1 between the "CG" domain of the 3C-like proteinase and the "GDD" containing domain VI of the Pol showed incongruent results. While CLRV-sequences were again grouped with ToRSV (48.8–49.3% aa identity between the two virus species), BRV, GBLV, and *Artichoke yellow ringspot virus* (AYRSV) clustered in a distinct group together with four nepovirus species assorted to subgroup A (GFLV, ArMV, TRSV, RpRSV), and *Peach rosette mosaic virus* (PRMV, subgroup C). Only subgroup B species formed a distinct lineage supported by significant bootstrap values. This was confirmed by comparison of the complete P2 encoded by RNA2 showing that the phylogenetic cluster of subgroup C nepoviruses was no longer maintained. Thus the most conserved Pro-Pol sequence as well as the more variable complete P2 emphasized the diversity among members of the genus. The incongruent clustering among species belonging to the nepoviruses have been previously reported by Sanfacon et al. (2009) leading to the conclusion that this genus needs revision and may be split up into separate genera.

## Acknowledgements

We thank PD Dr. Joachim Hamacher for providing the CLRV isolate, the Deutsche Forschungsgemeinschaft (DFG grant BU 890/14-1) for financial support, and Jana Gentkow and Renate Junge for assistance in the amplification and cloning of certain genomic regions of the investigated CLRV isolate.

## References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.
Blom, N., Hansen, J., Blaas, D., Brunak, S., 1996. Cleavage site analysis in picornaviral polyproteins: discovering cellular targets by neural networks. Protein Sci. 5, 2203–2216.
Borja, M., Sanchez, F., Rowhani, A., Bruening, G., Ponz, F., 1995. Long, nearly identical untranslated sequences at the 3′ terminal regions of the genomic RNAs of cherry leafroll virus (walnut strain). Virus Genes 10, 245–252.
Brooks, M., Bruening, G., 1995. A subgenomic RNA associated with Cherry leafroll virus infections. Virology 211, 33–41.
Büttner, C., von Bargen, S., Bandte, M., Myrta, A., 2011. Cherry leaf roll virus. In: Hadidi, A., Barba, M., Candresse, T., Jelkmann, W. (Eds.), Virus and Virus-Like Diseases of Pome and Stone Fruits. APS Press, St. Paul, USA, pp. 119–125.
Carrier, K., Hans, F., Sanfacon, H., 1999. Mutagenesis of amino acids at two tomato ringspot nepovirus cleavage sites: effect on proteolytic processing *in cis* and *in trans* by the 3C-like protease. Virology 258, 161–175.
Carrier, K., Xiang, Y., Sanfacon, H., 2001. Genomic organization of RNA2 of Tomato ringspot virus: processing at a third cleavage site in the N-terminal region of the polyprotein in vitro. J. Gen. Virol. 82, 1785–1790.
Chen, J., Chen, J., Adams, M.J., 2001. A universal PCR primer to detect members of the Potyviridae and its use to examine the taxonomic status of several members of the family. Arch. Virol. 146, 757–766.
Chisholm, J., Wieczorek, A., Sanfacon, H., 2001. Expression and partial purification of recombinant tomato ringspot nepovirus 3C-like proteinase: comparison of the activity of the mature proteinase and the VPg-proteinase precursor. Virus Res. 79, 153–164.
Elbeaino, T., Digiaro, M., Fallanaj, F., Kuzmanovic, S., Martelli, G.P., 2011. Complete nucleotide sequence and genome organisation of grapevine Bulgarian latent virus. Arch. Virol. 156, 875–879.
Froussard, P., 1992. A random-PCR method (rPCR) to construct whole cDNA library from low amounts of RNA. Nucleic Acids Res. 20, 2900.
Fuchs, M., Pinck, M., Serghini, M.A., Ravelonandro, M., Walter, B., Pinck, L., 1989. The nucleotide sequence of satellite RNA in grapevine fanleaf virus, strain F13. J. Gen. Virol. 70, 955–962.
Gorbalenya, A.E., Blinov, V.M., Donchenko, A.P., Koonin, E.V., 1989a. A NTP-binding motif is the most conserved sequence in a highly diverged monophyletic group of proteins involved in positive strand RNA viral replication. J. Mol. Evol. 28, 256–268.
Gorbalenya, A.E., Donchenko, A.P., Blinov, V.M., Koonin, E.V., 1989b. Cysteine proteases of positive strand RNA viruses and chymotrypsin-like serine proteases. FEBS Lett. 243, 103–114.
Gorbalenya, A.E., Koonin, E.V., Wolf, Y.I., 1990. A new superfamily of putative NTP-binding domains encoded by genomes of small DNA and RNA viruses. FEBS Lett. 262, 145–148.
Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95–98.
Hans, F., Sanfacon, H., 1995. Tomato ringspot nepovirus protease: characterization and cleavage site specificity. J. Gen. Virol. 76, 917–927.
Hellen, C.U.T., Cooper, J.I., 1987. The genome-linked protein of cherry leaf roll virus. J. Gen. Virol. 68, 2913–2917.
Jafarpour, B., Sanfacon, H., 2009. Insertion of large amino acid repeats and point mutations contribute to a high degree of sequence diversity in the X4 protein of tomato ringspot virus (genus Nepovirus). Arch. Virol. 154, 1713–1717.
Koonin, E.V., Dolja, V.V., 1993. Evolution and taxonomy of positive strand RNA viruses: implications of comparative analysis of amino acid sequences. Crit. Rev. Biochem. Mol. Biol. 28, 375–430.
Kozak, M., 2005. Regulation of translation via mRNA structure in prokaryotes and eukaryotes. Gene 361, 13–37.
Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins, D.G., 2007. Clustal W and Clustal X version 20. Bioinformatics 23, 2947–2948.
Latvala-Kilby, S., Lehto, K., 1999. The complete nucleotide sequence of RNA2 of blackcurrant reversion nepovirus. Virus Res. 65, 87–92.
Le Gall, O., Christian, P., Fauquet, C.M., King, A.M.Q., Knowles, N.J., Nakashima, N., Stanway, G., Gorbalenya, A.E., 2008. Picornavirales, a proposed order of positive-sense single-stranded RNA viruses with a pseudo-T = 3 virion architecture. Arch. Virol. 153, 715–727.
Marais, A., Faure, C., Candresse, T., Hulle, M., 2010. First report of nasturtium as a natural host of Cherry leaf roll virus on Amsterdam Island. Plant Dis. 94, 477.
Marchler-Bauer, A., Anderson, J.B., Chitsaz, F., et al., 2009. CDD: specific functional annotation with the Conserved Domain Database. Nucleic Acids Res. 37, 205–210.
Mayo, M.A., Fritsch, C., 1994. A possible consensus sequence for VPg of viruses in the family Comoviridae. FEBS Lett. 354, 129–130.
Mushegian, A.R., 1994. The putative movement domain encoded by nepovirus RNA-2 is conserved in all sequenced nepoviruses. Arch. Virol. 135, 437–441.
Ochman, H., Ajioka, J.W., Garza, D., Hartl, D.L., 1990. Inverse polymerase chain reaction. Biotechnology 8, 759–760.
Pacot-Hiriart, C., Latvala-Kilby, S., Lehto, K., 2001. Nucleotide sequence of black currant reversion associated nepovirus RNA1. Virus Res. 79, 145–152.
Pallas, V., Rowhani, A., Daubert, S., Bruening, G., 1991. Properties and translation studies of cherry leaf roll virus (CLRV). Acta Hortic. 309, 45–53.

Pesole, G., Gissi, C., Grillo, G., Licciulli, F., Liuni, S., Saccone, C., 2000. Analysis of oligonucleotide AUG start codon context in eukaryotic mRNAs. Gene 261, 85–91.
Rebenstorf, K., 2005. Untersuchungen zur Epidemiologie des Cherry leaf roll virus (CLRV) – genetische und serologische Diversität in Abhängigkeit von der Wirtspflanzenart und der geographischen Herkunft. PhD thesis, Humboldt-Universität zu Berlin, Landwirtschaftlich-Gärtnerische Fakultät, 127 pp.
Rott, M.E., Tremaine, J.H., Rochon, D.M., 1991a. Nucleotide sequence of tomato ringspot virus RNA-2. J. Gen. Virol. 72, 1505–1514.
Rott, M.E., Tremaine, J.H., Rochon, D.M., 1991b. Comparison of the 5′ and 3′ termini of Tomato ringspot virus RNA1 and RNA2: evidence for RNA recombination. Virology 185, 468–472.
Rott, M.E., Gilchrist, A., Lee, L., Rochon, D.M., 1995. Nucleotide sequence of tomato ringspot virus RNA1. J. Gen. Virol. 76, 465–473.
Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning – A Laboratory Manual. Cold Spring Harbour Laboratory, Cold Spring Harbour, New York.
Sanfacon, H., Wellink, J., Le Gall, O., Karasev, A., van der Vlugt, R., Wetzel, T., 2009. Secoviridae: a proposed family of plant viruses within the order Picornavirales that combines the families Sequiviridae and Comoviridae, the unassigned genera Cheravirus and Sadwavirus, and the proposed genus Torradovirus. Arch. Virol. 154, 899–907.
Scott, N.W., Cooper, J.I., Liu, Y.Y., Hellen, C.U.T., 1992. A 1.5 kb sequence homology in 3′-terminal regions of RNA-1 and RNA-2 of a birch isolate of cherry leaf roll nepovirus is also present, in part, in a rhubarb isolate. J. Gen. Virol. 73, 481–485.
Scott, N.W., Cooper, J.I., Edwards, J.L., 1993. The identification, cloning and sequence analysis of the coat protein coding region of a birch isolate (I2) of cherry leaf roll nepovirus. Arch. Virol. 131, 209–215.
Thompson, J.R., Leone, G., Lindner, J.L., Jelkmann, W., Schoen, C.D., 2002. Characterization and complete nucleotide sequence of Strawberry mottle virus: a tentative member of a new family of bipartite plant picorna-like viruses. J. Gen. Virol. 83, 229–239.
Tomlinson, J.A., Walkey, D.G.A., 1967. The isolation and identification of rhubarb viruses occurring in Britain. Ann. Appl. Biol. 59, 415–427.
Wang, A., Sanfacon, H., 2000. Proteolytic processing at a novel cleavage site in the N-terminal region of the tomato ringspot nepovirus RNA-1-encoded polyprotein in vitro. J. Gen. Virol. 81, 2771–2781.
Wang, A., Carrier, K., Chisholm, J., Wieczorek, A., Huguenot, C., Sanfacon, H., 1999. Proteolytic processing of tomato ringspot nepovirus 3C-like protease precursors: definition of the domains for the VPg, protease and putative RNA-dependent RNA polymerase. J. Gen. Virol. 80, 799–809.
Werner, R., Muehlbach, H.P., Büttner, C., 1997. Detection of cherry leaf roll nepovirus (CLRV) in birch, beech and petunia by immuno-capture RT-PCR using a conserved primer pair. Eur. J. Forest Pathol. 27, 309–318.
Wood, K.R., 1998. Nepovirus isolation and RNA extraction. In: Foster, G.D., Taylor, S.C. (Eds.), Methods in Molecular Biology, Vol. 81: Pant Virology Protocols: from Virus Isolation to Transgenic Resistance. Humana Press Inc., Totowa, NJ, USA, pp. 197–204.
Zhang, G., Sanfacon, H., 2006. Characterization of membrane association domains within the tomato ringspot nepovirus X2 protein, an endoplasmic reticulum-targeted polytopic membrane protein. J. Virol. 80, 10847–10857.
Zhou, Y., Civerolo, E.L., Melvin, J.A., 1998. Comparative nucleotide sequences of the coat protein-encoding region of the genomes of cherry leafroll virus strains infecting walnut. Acta Hortic. 472, 321–324.