COMPLETE NUCLEOTIDE SEQUENCE AND MOLECULAR PROBING OF POTATO VIRUS S GENOME

J. MATOUŠEK[1*], J. SCHUBERT[2], J. PTÁČEK[3], P. KOZLOVÁ[1,4], P. DĚDIČ[2]

[1]Department of Molecular Genetics, Institute of Plant Molecular Biology Czech Academy of Sciences, Branišovská 31, 370 05 České Budějovice, Czech Republic; [2]Federal Center for Breeding Research on Cultivated Plants, Institute for Resistance Research and Pathogen Diagnostics, Aschersleben, Germany; [3]Institute for Potato Research, Havlíčkův Brod, Czech Republic; [4]Faculty of Biological Sciences, University of South Bohemia, České Budějovice, Czech Republic

Received January 17, 2005; accepted June 14, 2005

Summary. – Complete genomes of three isolates of Potato virus S (PVS) were cloned and sequenced. The PVS ORF-1 was characterized for the first time. It encodes a putative replication protein (RPT) that shares the highest homology (about 52%) with that of Blueberry scorch virus (BlScV). ORF-1 motifs, characteristic for carlaviruses were found for methyltransferase (MTR), helicase (HEL) and RNA-dependent RNA polymerase (RdRp). The complete sequence of PVS genome enabled to develop an immunocapture RT-PCR probing of the PVS genome. Using this system, the sequence variability of 11 genome zones was examined for 34 PVS isolates including 15 PVS-CS variants that caused a systemic infection in *Chenopodium quinoa*. A broad variability between PVS isolates and diverse sequence variants was found. cDNA fragments covering the coat protein (CP) leader and CP-coding region (approx. 420 bp) were pooled for PVS-O and *Chenopodium*-systemic PVS isolates (PVS-CS) and corresponding cDNA libraries were screened for sequence variants. Both cDNA pools differred mainly in the 5'-end of the CP gene. Methionine at the position 17 in combination with serine at the position 34 were frequently associated with the CS character of PVS. In general, hydrophobic and polar amino acids were characteristic for the positions 17 and 34, respectively in PVS-CS isolates. Genome probing and evolutionary distances suggested that the PVS-CS isolates analyzed were close to the ordinary European isolates of ordinary strain of PVS (PVS-O) but distant to the original Andean strain of PVS (PVS-A).

Key words: genome sequence; heteroduplex analysis; Potato virus S; RT-PCR; temperature gradient gel electrophoresis; Solanum tuberosum L.; virus genome analysis

Introduction

PVS, a member of the *Carlavirus* genus belongs to the recently created family of *Flexiviridae* (Adams *et al*., 2004). The virus has 610–700 nm long flexuous filamentous particles and is transmitted by aphids in a non-persistent manner to plant species of the *Solanaceae* and *Chenopodiaceae* families. The capped single-stranded genomic RNA is polyadenylated and encapsidated in a CP of approximately 34 K (Mackenzie *et al.*, 1989; Foster, 1992; Monis *et al.*, 1987). The 3'-part of the virus genome has been characterized in detail earlier (Mackenzie *et al.*, 1989; Matoušek *et al*., 2000b); it contains a characteristic block of three ORFs encoding 25 K, 12 K and 7 K proteins) (Morozov *et al.*, 1989). The virus forms minor amounts of two subgenomic RNAs (2.5 kb and 1.5 kb), the smaller of which codes for CP and 11 K protein (Foster and Mills, 1992). It has been proposed that the smaller subgenomic RNA encoding the coat protein has a highly active translation enhancing leader sequence (Turner and Foster, 1997). To date, only the sequence of the 3'-part of the PVS genome

*E-mail: jmat@umbr.cas.cz; fax: +42038-5300356.

Abbreviations: AcoLV-D = Aconitum latent virus; BlScV = Blueberry scorch virus; CP = coat protein; DAS-ELISA = double-antibody sandwich ELISA; GarLV = Garlic latent virus; HEL = helicase; HpLV = Hop latent virus; LSR = low similarity region; LSV = Lily symptomless virus; MTR = methyltransferase; p.i. = post inoculation; PVM = Potato virus M; PVS = Potato virus S; PVS-A = PVS Andean strain; PVS-O = PVS ordinary strain; PVS-Y = Potato virus Y; RPT = putative PVS replication protein; RdRp = RNA-dependent RNA polymerase; TGGE = temperature gradient gel electrophoresis

(Mackenzie *et al.*, 1989; Matoušek *et al.*, 2000b) and restriction analysis of complete 7.5 kb PVS genome (Monis and de Zoeten, 1990) have been published. Information about the 5'-part of the PVS genome including a detailed description of ORF-1 is not available.

PVS is represented by PVS-O and highly virulent PVS-A (Mackenzie *et al.*, 1989). Although the disease, caused by PVS-O (Foster and Mills, 1992), may be symptomless on leaves and tubers of infected potato, the disease incidence may reach 100% with yield losses of 15%. Whereas PVS-O only causes local lesions on inoculated *Ch. quinoa* leaves, PVS-A does systemic infection. In previous work, a broad natural sequence variability among closely related Central European PVS-O isolates, has been reported (Matoušek *et al.*, 2000b). This study reports sequencing of complete PVS genome and differences between the PVS-O (ordinary) and *Chenopodium*-systemic PVS-A (Andean) isolates found by molecular probing, as described previously for Potato virus Y (PVY) (Ptáček *et al.,* 2002).

Materials and Methods

PVS isolates and plant inoculation. Nineteen PVS isolates originating from different potato cultivars were collected and maintained *in vitro*. The presence of PVS was assessed by ELISA. The isolates included those from *in vitro* germplasm collections as well as new isolates from breeding potato lines. Potato plants were maintained *in vitro* on the Murashige and Skoog (1962) medium with 0.7% agar solidified at 10°C and with an 8 hrs/day photoperiod. For the experiments new cuttings were prepared and maintained at 20°C and a 16 hrs/day photoperiod. These were maintained on *Lycopersicum esculentum* cv. Nevski and kept in climatic chambers at a similar temperature and photoperiod. For a PVS bioassay, the isolates were first propagated on tomato plants *L. esculentum* cv. Nevski. Indicator plants (*Ch. quinoa* and *L. esculentum* cv. Nevski) with fully developed third leaves were then mechanically inoculated with a crude sap containing each virus isolate in 50 mmol/l phosphate buffer pH 7.4. Three plants were used per isolate and maintained in a greenhouse at natural temperature and photoperiod. Inoculated plants were scored repeatedly for visual symptoms. Three PVS isolates, namely Czech isolates Kobra and Vltava and German PVS isolate Leona were subjected to genome sequencing. All the isolates were originally collected from homonymous potato cultivars and maintained in germplasm since nineties.

Double-antibody sandwich ELISA (DAS-ELISA) was used to verify the infection at 3 and 5 weeks post inoculation (p.i.) (Clark and Adams, 1977). A broad-spectrum of PVS antibodies from Bioreba or PRI Wageningen (the Netherlands) was employed.

RNA extraction and full length cDNA cloning. A total RNA was isolated from purified preparations of all three virus isolates by digestion with 1 U of proteinase K (Sigma) at 42°C for 30 mins in 50 mmol/l Tris-HCl buffer pH 7.0 containing 0.2% SDS. The released RNA was extracted with phenol/chlorophorm and ethanol-precipitated. cDNAs were synthesized using the forward primer 5'-GTTTCAGTGGCACATGCTTGAG-3' (nt 4973–4952) and reverse transcriptase (SuperScript II, Gibco, USA) according to the manufacturer's instructions. Double stranded cDNAs were generated using DNA Polymerase I using 1 U DNA Polymerase I (Gibco) in the buffer provided by the manufacturer. The resulting double-stranded cDNAs were cloned into pBluescript II/SK+ plasmid (Stratagene) and resulting recombinant plasmids were used for transformation of *Escherichia coli* DH5αc. The clones containing inserts of at least 5 kb were used for nucleotide sequencing.

As by the method described above the 5'-end of the RNA could not be cloned it was obtained by the 5'/3'-RACE technology using a correspoding kit from Roche Pharma. For PCR, the primers race1 (5'-GAACCATTCGAAGTTATCCTTCTCA-3', nt 191–166) and race2 (5'-TCGAAAAGTGTTAACACTTCCTCGA-3', nt 109–87) were employed. The PCR-products resulting from the 5'/3'-RACE technology were cloned into pGEM-T plasmid (Promega) and resulting recombinant plasmids were used for transformation of *E. coli* DH5αc. Four clones were chosen for each isolate and sequenced. This cloning step provided sequence information about the 5'-terminus of PVS from which the primer PVS5-*Aat*II was derived. Full length cDNAs for the isolate Leona was obtained from corresponding RNA using the primer PVS3-Sgr (5'-CTCACCGGT GTTTTTTTTTTTATGCTAAAATATTTTAAAAACAT ATT-3', nt 8489–8453) containing a unique *Sgr*I site (underlined) and Superscript II reverse transcriptase (Invitrogen). The reaction ran in a volume of 25 µl for 2 hrs at 42°C according to the manufacturer's instructions. For the long range PCR the forward primer PVS3-Sgr and the reverse primer PVS5-*Aat*II (5'-TAAGACGTCGATAAA CACTCCCGAAAATAATTTGAC-3', nt 1–27) containing a unique *Aat*II site (underlined) was used in combination with. The reaction mixture (50 µl) contained 5 µl of cDNA, 200 µmol/l dNTPs, 1 µmole of each primer, 1 U of TripleMaster *Taq*-Polymerase in corresponding reaction buffer provided by the supplier of the *Taq*-polymerse. The reaction consisted of initial denaturation for 3 mins at 98°C, addition of the polymerase and 12 cycles of denaturation at 93°C for 30 secs, annealing at 66°C for 30 secs, primer extension at 68°C for 9 mins. In each cycle the annealing temperature was lowered by 0.4°C and the extension time was prolonged by 20 secs. The next 25 cycles had the following conditions: a denaturation at 96°C for 45 secs, annealing at 62°C for 30 secs, and extension at 68°C. In each of this cycles initial extension time of 10 mins was prolonged by 20 secs. The reaction was completed by a final extension at 72°C for 10 mins. After agarose gel electrophoresis of the PCR fragment a band of the expected size was cut out, purified using the NucleoSpin Extract Kit (MacheryNagel) and cloned into pGEM-T easy plasmid (Promega).

The resulting recombinant plasmid was used for transformation of ultracompetent *E. coli* XL2-Blue cells (Stratagene). From this cloning procedure only one complete clone was obtained and this was sequenced in an Alf-Express sequencer (Amersham-Pharmacia). The obtained data were analyzed using the DNASIS software (Hitachi).

Immunocapture RT-PCR probing of PVS genome. PCR and RT-PCR primers were designed from the full length PVS genome sequence. The latter was divided into 11 zones (A-K). Each zone was subdivided into three parts (I-III) according to a strategy for the genome of PVY described earlier (Ptáček *et al*., 2002) In this way three fragments ranging from 237 to 492 bp were obtained. Such a fragments can be easily analyzed by thermodynamic methods (Matoušek *et al*., 2000a). In total, 33 primer pairs (Table 1)

Table 1 Characteristics of primers

Primer designation and sequence	Position	Gene region	Annealing temp. (°C)
Zone A			
A-I-1: 5′TCACTATAGGGCGAATTGGG3′	3-22	RPT	57.0
A-I-2: 5′CTCTTGAAAGCTGCTGGTGG3′	255-236	RPT	57.0
A-II-1: 5′ATCCCTAATTTCCAACGTCG3′	216-235	RPT	56.0
A-II-2: 5′TCGGCACTGCTCACATACC3′	539-521	RPT	55.6
A-III-1: 5′CGGTGTGCAAGACATTGGA3′	371-389	RPT	56.3
A-III-2: 5′CCAACGAATATCCCCGTTG3′	788-770	RPT	57.0
Zone B			
B-I-1: 5′TTGTGCACCATTGTGTACCC3′	751-770	RPT	55.6
B-I-2: 5′GGTGTGATCAAATCGCCCT3′	1022-1004	RPT	56.4
B-II-1: 5′GCAAATTCGCCCATCATCT3′	971-989	RPT	57.1
B-II-2: 5′ATCATGACTCTCAAAGTGCCAG3′	1286-1265	RPT	55.8
B-III-1: 5′CGACAAGCAGTCTGCAATGG3′	1164-1183	RPT	58.5
B-III-2: 5′CACCCACCTCCTATCCATAGC3′	1533-1513	RPT	57.8
Zone C			
C-I-1: 5′GACATGATTGAAGCTATGGATAG3′	1501-1523	RPT	53.4
C-I-2: 5′CGAGAAGCGAGCAACTTG3′	1759-1742	RPT	53.3
C-II-1: 5′CCTACGTGGCTCAGATCGC3′	1721-1739	RPT	57.5
C-II-2: 5′CCCACTCCGTGAAAGCTGTA3′	2033-2014	RPT	57.4
C-III-1: 5′CCGATTGGGCAAGGTACAA3′	1922-1940	RPT	58.0
C-III-2: 5′AGTACCATGCAGCACGCTTG3′	2284-2265	RPT	58.4
Zone D			
D-I-1: 5′AAGGGGGTTCCAGGCAA3′	2251-2267	RPT	56.6
D-I-2: 5′GACTGACCTCCAACCTCAGC3′	2510-2491	RPT	55.2
D-II-1: 5′TGATGGAGGCCAGTTCATCC3′	2451-2470	RPT	59.0
D-II-2: 5′CTCATTGGGCTTGGAACTCC3′	2751-2732	RPT	58.4
D-III-1: 5′CTGCCACAACGCCAGGTG3′	2711-2728	RPT	59.4
D-III-2: 5′GCACTCAGCACCACGATCCT3′	3050-3031	RPT	59.2
Zone E			
E-I-1: 5′CGCAACAAGTAAGCAATACCG3′	3006-3026	RPT	58.4
E-I-2: 5′CTTCACCCTTCAGCACATCCT3′	3268-3248	RPT	58.1
E-II-1: 5′CGCGCAGTTGGGTAATGAG3′	3222-3240	RPT	58.9
E-II-2: 5′GTTGTACGGTAGCAACGTGC3′	3528-3508	RPT	58.9
E-III-1: 5′GTGCATGATGCTAGCTGTGA3′	3451-3471	RPT	55.4
E-III-2: 5′CAAGCTCATTGAAAGAATCT3′	3799-3778	RPT	56.5
Zone F			
F-I-1: 5′ATCTACGTATCCCCAAGGAAG3′	3751-3771	RPT	54.8
F-I-2: 5′TCGCTTTGGCACGGATC3′	4013-3997	RPT	56.3
F-II-1: 5′CTCTATGTTGGCACCAGCG3′	3954-3972	RPT	55.9
F-II-2: 5′AAGAGCTCGGGAAGTAAGTCT3′	4300-4280	RPT	54.0
F-III-1: 5′CACAATCATTGAGGGCATTG3′	4188- 4207	RPT	55.6
F-III-2: 5′GGCAACAACATCTTGAGGTC3′	4550-4531	RPT	53.7
Zone G			
G-I-1: 5′GCTCAAAACAGCAAAGATTGAT3′	4509-4530	RPT	56.2
G-I-2: 5′CTTTTCAACCCACCTTGCC3′	4782-4764	RPT	56.7
G-II-1: 5′ATGGTTTCGCACACATTTGC3′	4719-4738	RPT	58.1
G-II-2: 5′GAAATGGCCCATACATCTC3′	5029-5011	RPT	58.6
G-III-1: 5′AAGATTGAGGTTTTCGGACC3′	4959-4978	RPT	55.0
G-III-2: 5′GCGGCATAGGACGCTATG3′	5298-5281	RPT	55.7
Zone H			
H-I-1: 5′GCAGCACAAACCATTGTCTG3′	5254-5273	RPT	55.7
H-I-2: 5′TCAATGCTAACTCAAACGCC3′	5500-5481	RPT	54.1
H-II-1: 5′GACTATGAAGCCTTTGACGC3′	5440-5459	RPT	54.6
H-II-2: 5′TGAATAGAGCCCTGTTGGC3′	5731-5713	RPT	56.1
H-III-1: 5′CTTGTTCAACACAATGGCCA3′	5619-5638	RPT	56.1
H-III-2: 5′TCTCGGACTTCAGCAAATGC3′	6049-6030	RPT	57.5
Zone I			
I-I-1: 5′TTGTGTGAGGGTGATTATTAAG3′	6003-6024	RPT-25K	51.8
I-I-2: 5′ACTTAAGCAACTCCCGAATT3′	6240-6221	RPT-25K	53.0
I-II-1: 5′AGCTGGTAAAAGTTCTGCAA3′	6203-6222	25K	52.1

I-II-2: 5´AATCTTGACCTTCTGCCTGT3´	6531-6512	25K	52.5
I-III-1: 5´AGTACCGCTGTGAGCCCAT3´	6411-6429	25K-12K	56.4
I-III-2: 5´GTGTAGTTAGGCGGCGGT3´	6800-6783	25K-12K	54.7
Zone J			
J-I-1: 5´CTGGTTGATAAGGCTGCTGC3´	6702-6720	25K-12K	57.4
J-I-2: 5´GCTTTAGTGCCGTCGCGA3´	6938-6921	25K-12K	59.3
J-II-1: 5´ACTGCCAATTGTTGGGGAT3´	6878-6888	12K-7K	56.1
J-II-2: 5´TCCCCAGTAATGAGTAGGACG3´	7180-7160	12K-7K	56.0
J-III-1: 5´TAAGTTGGTATGCGCTCAGG3´	7123-7142	7K-CPL-CP	55.6
J-III-2: 5´GGGTCGGCTCAAGCGA3´	7544-7529	7K-CPL-CP	55.8
Zone K			
K-I-1: 5´GGTAGGCCCTCGCTTGAG3´	7520-7537	CP	56.6
K-I-2: 5´CACAGTCCCTGCTGGATCTAG3´	7783-7763	CP	56.0
K-II-1: 5´GAAAGTGGTGATCATGTGTGC3´	7720-7740	CP	54.3
K-II-2: 5´TGAGCTATTGCTTCCTCAGGT3´	8055-8035	CP	56.0
K-III-1: 5´CGTAGAGGGGCTCATACGC3´	8008-8026	CP-11K	57.1
K-III-2: 5´CTCTGACTTTGCACCATGGG3´	8499-8480	CP-11K	57.3

were designed. The thermodynamic properties and structural features of individual primer pairs were calculated using the Gene Runner, version 3.02 (Hastings Sofware, Inc.) to obtain minimal differences in annealing temperatures and to avoid complicated secondary structures as much as possible.

Immunocapturing was performed on ELISA microplates (Nunc Maxisorp) using polyclonal antibodies (PRI, Wageningen). First strand cDNAs were synthesized with the Enhanced Avian HS RT-PCR Kit from Sigma using random DNA nonamers. The RT ran at 42°C for 60 mins according to the manufacturer's instructions. In PCR (final volume of 50 µl), 5 µl of the RT reaction mixture was amplified using the abovementioned kit and primers described in Table 1 in the PTC-100 Thermocycler (MJ Research). The cycling conditions included a denaturation at 94°C for 3 mins, 35 cycles of 94°C/30 secs, 52–57°C/45 secs, and 72°C/1 min, and final extension at 72°C for 5 mins. The PCR products were separated in 1% agarose gel stained with ethidium bromide and visualized using the EDAS 290 (Kodak) photo-documentation system. The electrophoreograms were quantified by means of the ImageQuaNT software (Molecular Dynamics, USA). Band intensities were expressed in percents.

Analysis of PCR fragments by temperature gradient gel electrophoresis (TGGE) and cDNA library screening by the method of heteroduplexes. TGGE was performed in 5% denaturing acrylamide gels as described by Matoušek *et al*. (2000b). Screening of cDNA libraries for heteroduplexes was performed according to the method described by Matoušek *et al*. (2000a). Briefly, a sample cDNA was hybridized with a standard DNA (clones O5 or CS7 for O and CS variants, respectively). If a sample cDNA is identical with standard DNA, homoduplexes are formed; but if a sample cDNA is different from standard DNA (it is a sequence variant), heteroduplexes are formed. After hybridization, the hybrids were electrophoresed at 47°C a temperature estimated from TGGE profiles and favorable for partial denaturation of strands and formation of transient structures. These transient DNA structures differ in their electrophoretic mobilities from more stable homoduplexes. As a result of such an electrophoretic analysis and silver nitrate staining (Schumacher *et al*., 1986), some heteroduplexes can be identified as extra bands in the gel.

Database sequences and computer analysis. The EMBL and GenBank Nucleotide Sequence Databases were used as a source of full length replicase sequences. For comparison, the sequences of following carlaviruses were used: Aconitum latent virus (AcoLV-D, Genbank NC_002795), Blueberry scorch virus (BlScV, Genbank NC_003499), Garlic latent virus (GarLV, Genbank NC_003557), Hop latent virus (HpLV, Genbank NC_002552), Lily symptomless virus (LSV, Genbank NC_005138), and Potato virus M (PVM, D14449). For rooting the phylogenetic tree of the replicase sequences the recently described sequence of Sugarcane striate mosaic-associated virus (SCSMaV, Genbank NC_003870), characterized as positioned between the genera Foveavirus and Carlavirus (Thompson and Randles, 2001), was used as the outgroup. In some comparisons other sequences were used: PVS-A (PVS-AN/D00461); PVS-O (PVS-O/S45593); PVS strain SE (PVS-SE/U74375); PVS strain RB (PVS-RB/U74376); sequences of the PVS 3'-region (PVSKOBRA/Y15625). Sequence data analysis was carried out using the programs DNASIS, version 2.5 and DNASIS MAX (Hitachi). Protein sequence alignments were performed using the CLUSTAL W program with the amino acid classes described by Bork *et al.* (1996). The prediction of cDNA melting temperatures in TGGE gels was performed using the POLAND program (Poland, 1974), kindly provided by Dr. G. Steger, Institute Physical Biology, Heinrich-Heine University, Düsseldorf, Germany.

Results and Discussion

Full length genomes of three PVS isolates, Leona, Vltava and Kobra were cloned and genome sequence was obtained directly for only one of them, Leona. For the remaining two isolates complete genome sequences were reconstituted from shorter cDNA fragments. The genome sequences for the isolates Leona, Vltava and Kobra were deposited in the EMBL nucleotide database under Acc. Nos. AJ863509, AJ863510 and Y15625, respectively.

In comparing complete genome sequences an approx. 8% difference among the three isolates was found. The genome

of the isolate Leona consisted of 8,478 bp; ORF-1 of 5,919 bp encoded a 223 K RPT (1,973 aa) with a predicted pI of 6.97. Phylogenetic analysis (Fig. 1) showed that the putative RPT shared the highest similarity with those of the isolates BlScV and Lily symptomless virus (LSV), forming a common branch on the phylogenetic tree. The highest similarity of PVS to BlScV was also observed for CP (Matoušek *et al*., 2000b) and other parts of the 3'-region, such as the translational enhancer upstream of CP (Turner *et al.*, 1999). A more detailed comparison between whole RPTs of PVS and BlScV showed an amino acid identity of 51%, while a 71% identity was found for groups of similar amino acids. In addition, the amino acid identity was considerably higher for N- and C-terminal parts of RPT, reaching 65% and 80%, respectively. A relatively low (37%) amino acid identity was observed for the central part of RPT (aa 392–1489), forming a low similarity region (LSR). A comparable similarity was observed for PVS as compared with other carlaviruses listed in Fig. 1 (data not shown).

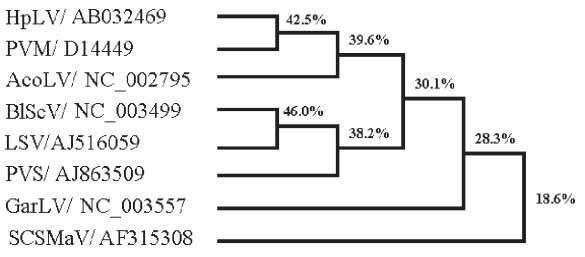

Fig. 1

Phylogenetic tree of various carlaviruses based on deduced amino acid sequence of RPT

The tree was generated using the Clustal W program. SCSMaV was chosen as the outgroup member. Amino acid identity (%) is indicated. AcoLV-D = Aconitum latent virus; BlScV = Blueberry scorch virus; GarLV = Garlic latent virus; HpLV = Hop latent virus; LSV = Lily symptomless virus; PVM = Potato virus M; PVS = Potato virus S.

A domain arrangement, characteristic for the carlavirus ORF-1 protein, i.e. the motifs for MTR, HEL and RdRp were identified in the PVS RPT (Fig. 2). A very high similarity was observed within these motifs. When comparing PVS and BlScV, the level of amino acid identity for various domains ranged from 52% to 100% and for similar amino acids from 70% to 100% (Table 2). In general, the highest similarity was observed for RdRp motifs and the lowest one for HEL domains localized in LSR.

A broad sequence variability of PVS genomes, based on detailed characterization of 3'-parts of genomes of several Central European PVS isolates has been reported by Matoušek *et al*. (2000b). This variability suggested the existence of potentially divergent variants within the Central European population of PVS. In the present study, a more complex analysis of PVS was attempted in order to obtain some molecular characteristics of the "Central European" variant of PVS (PVS-CS) that infects *Ch. quinoa* systemically. PVS-CS could be closely related to PVS-A that also infects *Ch. quinoa* systemically but causes more severe pathogenic reactions on the potato (Mackenzie *et al*., 1989). To investigate this possibility, a more complex collection of PVS isolates was obtained from various potato cultivars and inoculated to indicator plants of *Ch. quinoa* and *L. esculentum.* In total, 34 isolates were collected, of which 19 were PVS-O and 15 were PVS-CS variants (Table 3) of undetermined molecular relationship to the PVS-A, described by Mackenzie *et al*. (1989).

In order to characterize the PVS collection in more detail, a system of immunocapture RT-PCR probing of PVS genome was employed. This method has previously been used for PVY for quick molecular comparison of various isolates and their characterization including prediction of mutations in individual genome regions (Ptáček *et al*., 2002). Complete PVS genome sequence enabled designing a similar probing system for PVS using the complete genomic sequence as a template (Fig. 3). PVS genomes were probed within 11 zones (A-K), each approached with three pairs of balanced primers to amplify three short subfragments. Thus, a specific fingerprint based on 33 primer pairs was obtained for each isolate (Table 3). A broad variability of PVS was found throughout the whole PVS genome, with characteristic patterns for individual isolates, suggesting numerous changes at the nucleotide level. This finding is consistent with previous ones of broad nucleotide variability in PVS (Matoušek *et al*., 2000b). Furthermore, numerous differences among isolates

Table 2. Homology of PVS and BlScV protein domains encoded by ORF1

Amino acid sequence homology/identity (%)																		
Protein domains																		
RPT	MTR			HEL							RdRP							
	I	II	III	I	IA	II	III	IV	V	VI	I	II	III	IV	V	VI	VII	VIII
52/ 71	93/ 100	97/ 100	64/ 81	63/ 77	66/ 80	75/ 87	70/ 78	67/ 70	64/ 77	77/ 85	100/ 100	93/ 97	92/ 100	100/ 100	93/ 93	100/ 100	60/ 70	90/ 90

Maximum homology for whole RPT sequence was pairwise calculated.

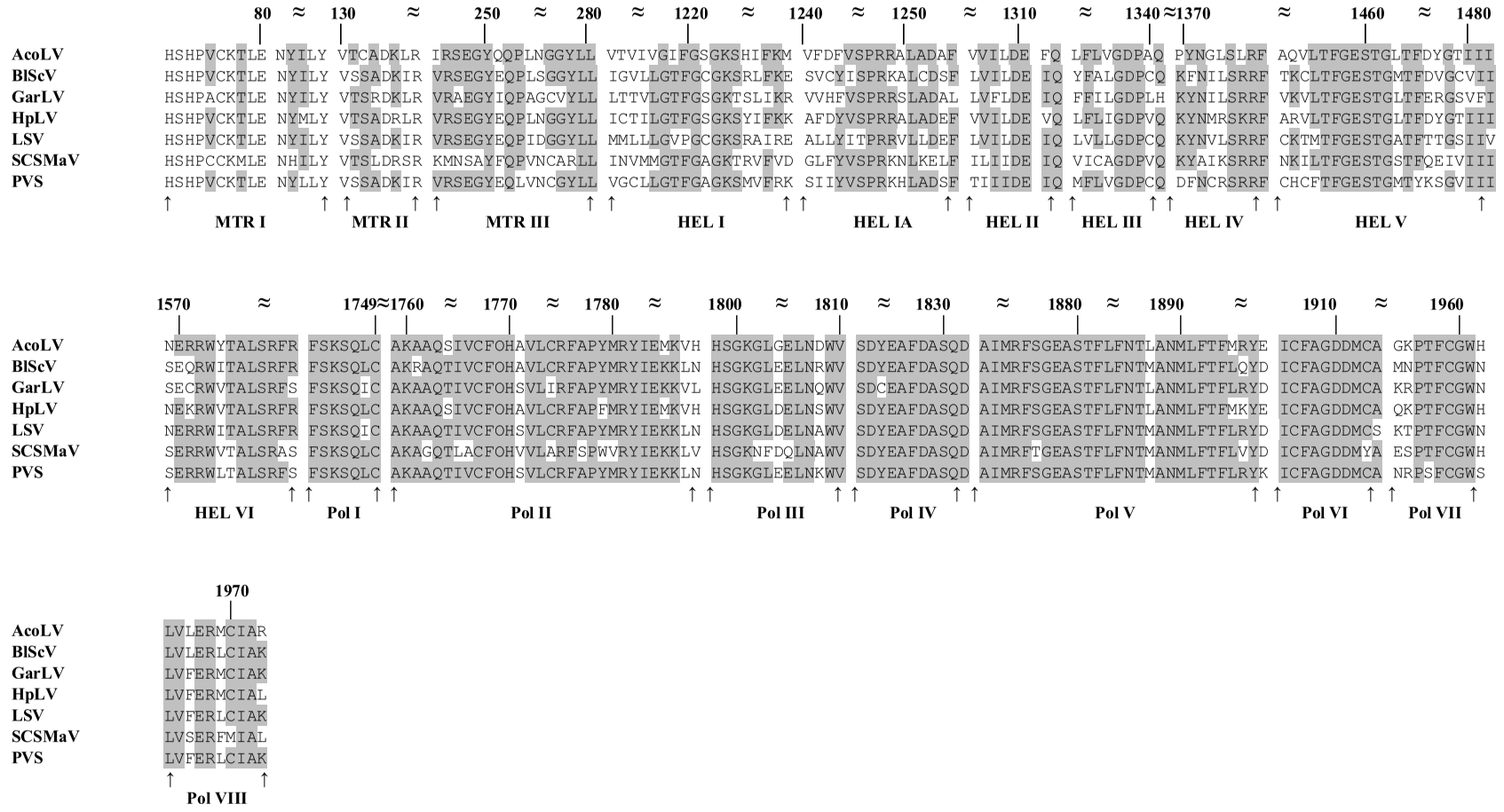

Fig. 2

Multiple alignment of amino acid sequences of MTR, HEL, and RdRp motifs in RPT of various carlaviruses

Sequences identical in at least four carlaviruses are shaded in gray. Three type 1 MTR motifs, seven HEL motifs of superfamily 1 and eight RdRp motifs of supergroup 3 are shown (Koonin and Dolja, 1994).

Table 3 Immunocapture RT-PCR probing of PVS isolates

Isolates	strain	Primer zones																																	
		A–I	A–II	A–III	B–I	B–II	B–III	C–I	C–II	C–III	D–I	D–II	D–III	E–I	E–II	E–III	F–I	F–II	F–III	G–I	G–II	G–III	H–I	H–II	H–III	I–I	I–II	I–III	J–I	J–II	J–III	K–I	K–II	K–III	
Alka	o	–	+	–	+	+	–	–	+	(+	–	–	–	+	+	+	+	(+	+	+	+	+	+	(+)	–	–	(+)	+	+	+	+	+	–	+	
Arran Banner	o	+	–	–	–	–	–	–	+	–	–	–	–	+	(+	–	+	–	+	–	(+	–	+	–	–	–	+	–	–	–	+	(+	–	(+	
Karin	o	–	+	+	+	+	–	–	+	+	+	+	+	+	+	+	+	+	+	+	+	+	(+	–	+	–	+	+	+	+	+	+	+	+	
Kobra	o	–	–	–	+	+	–	–	+	+	+	–	–	+	+	+	+	+	(+	–	+	+	+	+	–	–	+	+	+	+	+	+	+	+	
Kordoba	o	+	+	(+)	+	+	–	–	+	+	+	(+	(+	+	+	+	+	+	+	+	+	+	+	+	–	–	+	+	–	–	+	+	+	(+	
Nordchip 54	o	+	+	–	+	+	–	+	+	+	+	–	–	+	+	+	+	+	+	–	+	+	+	+	–	–	–	–	+	–	+	+	+	+	
Oddesky	o	+	+	–	+	+	–	–	–	–	–	–	–	+	+	+	+	(+	+	+	+	+	+	–	–	–	+	–	+	+	+	–	+	+	
Amarilla del Centro	o	+	+	+	+	+	+	–	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	–	–	–	–	+	–	+	+	+	+	+	
Aquila	o	+	+	(+	+	+	+	–	+	+	(+	(+	–	–	+	+	+	+	+	+	+	+	+	+	–	+	–	+	+	+	+	+	+	+	
Bograital	o	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	(+	–	–	–	–	–	–	+	–	–	+	–	–	–	
Centinial Russet	o	–	+	–	+	+	+	–	+	+	+	+	–	+	+	+	+	+	+	–	+	+	+	–	–	–	–	+	–	+	+	+	+	–	
Fransen	o	(+	+	+	+	+	+	–	+	+	+	+	–	+	+	+	+	–	–	+	+	+	+	–	–	–	–	+	–	+	+	+	+	+	
Heideniese	o	(+	+	+	+	+	+	–	+	+	+	+	–	+	+	+	+	+	+	+	+	+	+	+	–	–	–	+	–	+	+	+	+	+	
Linzer Rose	o	(+	+	+	+	+	+	–	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	–	–	–	+	+	–	+	+	+	+	
Long Blue	o	–	+	–	+	+	+	–	+	+	+	+	–	+	+	+	+	+	+	+	+	+	–	–	–	+	–	+	+	+	+	+	+	+	
Purple and White	o	+	+	+	+	+	+	–	+	+	+	+	–	+	+	–	+	+	+	+	+	+	–	+	–	–	–	–	+	+	+	+	+	+	
Reichskanzler	o	+	+	+	+	+	+	–	+	+	+	+	(+	+	+	+	+	+	+	+	+	+	+	–	–	+	–	+	+	+	+	+	+	+	
VAC 67	o	+	+	–	+	–	+	–	+	+	+	(+	+	+	+	+	+	+	+	+	+	+	+	–	–	+	–	–	+	+	+	+	+	+	
Wohltmann	o	+	+	–	+	+	+	–	+	+	+	(+	–	+	+	+	+	+	+	+	+	+	–	–	–	+	–	–	+	+	–	+	+	+	
ASS	cs	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	(+)	–	–	–	–	–	–	–	–	–	–	–	–	–	
Dutch	cs	–	+	–	+	+	–	–	+	+	(+	–	–	+	+	–	+	–	–	+	+	(+	+	–	–	–	–	–	–	–	+	(+)	–	+	
Peru	cs	(+	–	–	–	–	–	–	–	–	–	–	–	(+)	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	
Germ	cs	–	+	–	+	+	–	–	+	+	+	–	–	+	+	–	+	–	–	+	+	+	+	–	–	–	–	+	–	–	+	+	(+)	+	
I/47	cs	+	–	–	+	+	–	+	–	+	(+	–	–	+	+	+	+	(+	+	+	+	+	+	–	–	–	(+)	(+)	–	–	+	+	+	+	
I/84	cs	+	+	–	+	+	–	–	+	+	+	+	(+)	+	+	(+	+	+	+	+	+	+	+	+	–	–	+	–	(+	(+	+	+	+	+	
II/24	cs	(+	+	–	+	+	–	–	+	+	+	–	–	+	+	+	+	+	(+	+	+	+	(+)	+	+	–	+	(+)	+	+	+	–	+	+	
II/31	cs	–	+	+	+	+	–	–	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	–	+	–	+	+	+	+	+	+	+	+	
10 II/35	cs	+	+	–	+	+	–	–	+	+	+	–	–	+	+	+	+	+	+	+	+	+	+	+	+	–	+	–	+	+	+	(+	+	+	
Dunbar Rover	cs	+	+	(+)	+	+	–	–	+	+	(+)	+	(+)	+	+	+	+	(+	+	+	+	+	+	–	–	–	–	–	+	(+	–	–	–	+	
Tarpan	cs	+	+	–	+	+	–	–	+	+	+	–	–	+	+	+	–	+	+	+	+	+	(+	–	–	–	+	+	+	+	(+	–	+	+	
Blaue Mandel	cs	+	+	+	+	–	+	–	+	+	+	+	(+	+	+	+	+	+	+	–	+	+	+	+	–	+	–	–	+	+	+	+	+	+	
Clivia	cs	+	+	+	+	+	+	–	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	–	–	–	+	+	+	+	+	+	+	
Peruanita	cs	+	–	–	–	–	+	–	–	–	–	–	–	+	–	–	(+	–	–	–	+	–	–	–	–	–	–	+	–	–	–	–	–	–	
Yuguima	cs	+	–	–	+	–	+	–	–	–	–	–	–	+	–	+	+	–	–	–	+	–	–	+	–	–	–	+	–	–	–	–	–	–	
Desiree VF		–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	

+, 100–70%; (+, 69–40%; (+), 39–10%. The percentage corresponds to the band intensity.

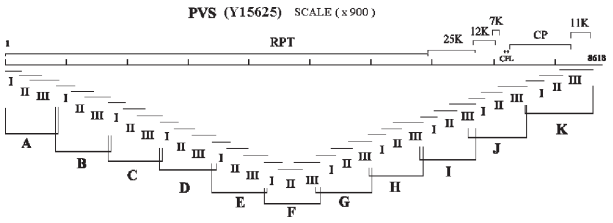

Fig. 3

Analysis of cDNA fragments from PVS Kobra isolate using selected primers

Except for RPT, CP and CPL, PVS open reading frames are designated according to the protein sizes as described earlier (e.g. Matoušek *et al.*, 2000b). RPT, gene for putative PVS replication protein; CP, coat protein gene; CPL, coat protein leader.

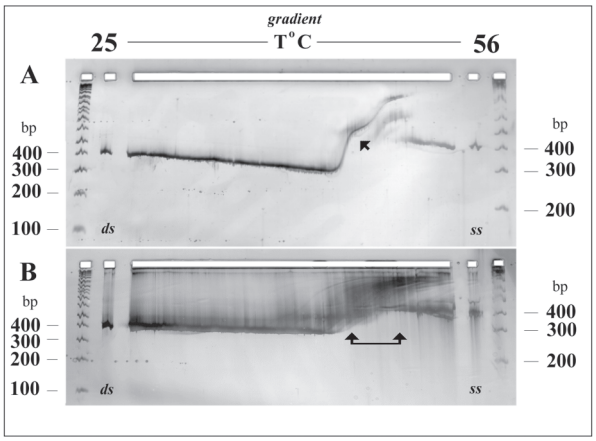

Fig. 4

Analysis of J-III (7KCPLCP) cDNA fragments of PVS by TGGE

A linear temperature gradient from 25 to 56°C was used. Melting point positions are indicated by arrows. (A): Re-amplification from clone O-7. (B): cDNA re-amplified from the cDNA library of PVS-CS variants. DNA size marker (100 bp ladder) is shown on both sides. ss = single-stranded. ds = double-stranded.

A

```
                    10         20         30         40         50
O-19      1 MPPKPDPTSS GETPQTVPLV PPPRNVEEHR VGPNQGHGQN EEAVLEQRFI  50
O-38      1 .......... .....AL... .......... ...I...... ...M....L.  50
O-40      1 .......... .....A...A .......... .......... ......R.L.  50
O-42      1 .......... .......... .......... .......... ........L.  50
O-5       1 .......... .....A...A .......... .......... ......R.L.  50
O-52      1 .......... .....AL... .......... ...I...... ...M....L.  50
O-53      1 .......... .......... ...W...... .......... ........L.  50
O-57      1 .......... .......... .......... .......... ........L.  50
O-58      1 .......... .......... .......... .......... ........L.  50
O-59      1 .......... .......... .......... .......... ........L.  50
O-6       1 .......... .....A...A .......... A......... ...M....L.  50
O-61      1 .......... .....A...A .......... A......... ...M....L.  50
O-68      1 .......... .......... .......... .......... ..........  50
O-7       1 .......... .....AL... .......... ...I...... ...M....L.  50
O-74      1 .......... .....A...A .......... A......... ...M....L.  50
O-75      1 .......... .......... .......... .......... ........L.  50
O-8       1 .......... .......... ...W...... .......... ........L.  50
CS-10     1 .......... .....AM... .......... ...S...... ...M....L.  50
CS-11     1 .......... ..A..AI..A .......... ...RE.PR.. ...M....L.  50
CS-14     1 .......... ..A..AM..A .........K ...SE.PR.. ...M....L.  50
CS-26     1 .......... ..A..AM..A ......G..K ...SE.PR.. ...M....L.  50
CS-27     1 .......I.. ..A..AM..A .........K ...SE.PR.. ...M....L.  50
CS-28     1 .......... .....AM..A .......... ...SE.PR.. ...M....L.  50
CS-29     1 .......... ..A..AM..A .........K ...SE.PR.. ...M....LV  50
CS-31     1 .......... .....AM... .......... ...S...... ...M....L.  50
CS-33     1 .......I.. ..A..AM..A .........K ...SE.PR.. ........L.  50
CS-34     1 .......... ..A..AM..A .........K ...SE.PR.. ...M....L.  50
CS-35     1 .......... ..A..AMR.A .........K ...SE.PR.. ...M....L.  50
CS-36     1 .......I.. ..A..AM..A .........K ...SE.PR.. ...M....L.  50
CS-39     1 .......... .....AM... .......... .......... ...M....L.  50
CS-40     1 .......... ..A..AM..A .........K ...SEEPR.. ...M....L.  50
CS-41     1 .........P ..A..AI..A .......... ...RE.PR.. ...M....L.  50
CS-42     1 .......... .....AM..A .........K .A.SE.PR.. ...M....L.  50
CS-5      1 .......... ..A..AM..A .........K ...SE.PR.. ...M....L.  50
CS-7      1 .......I.. ..A..AM..A .........K ...SE.PR.. ...M....L.  50
```

B

```
          (17)                (34)
O  MPPKPDPTSSGETPQslPLsPPPRNVEEHRsGP.QGHGQNEEAhLEQRhIRLIELMASKRHNSTLSNISFEIGRPSLEPT
CS MPPKPDPhSSGEsPQAhPLsPPPRNVEEH+VGPppG.tQNEEAMLEQRLIRLIELMASKRHNSTLSNISFEIGRPSLEPT
```

Fig. 5

Multiple alignment of amino acid sequences of N-termini of CP of cloned J-III (7KCPLCP) fragments selected from PVS-O and PVS-CS pools (A). Consensus sequences of PVS-O and PVS-CS pools (B)

Dominant amino acid differences between the consensus sequences are marked by asterisks.

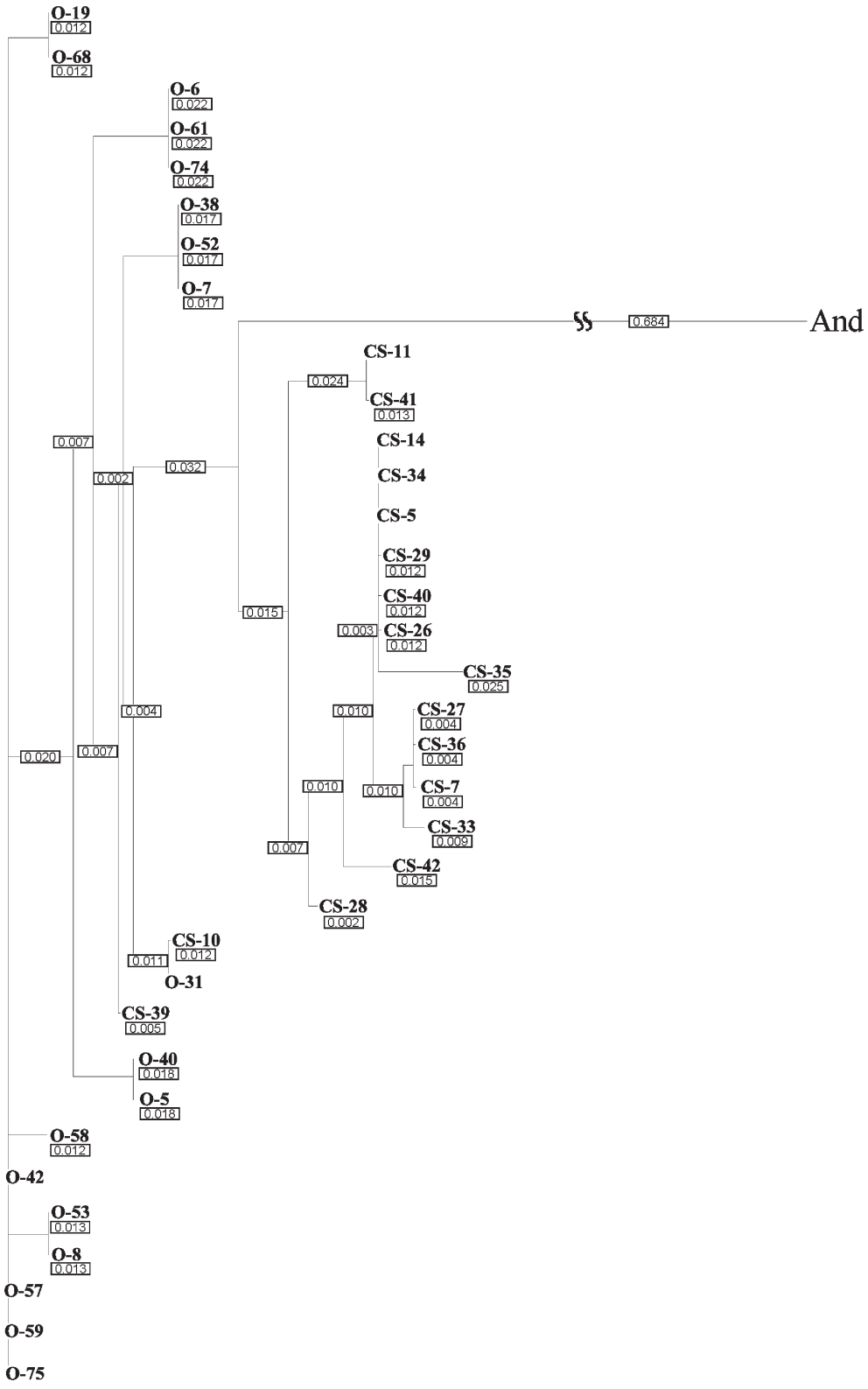

Fig. 6

Cluster analysis of amino acid J-III (7KCPLCP) sequences selected from PVS-O and PVS-CS pools and that of PVS-A

Genetic distances are given in boxes, zero distances on individual branches are not shown.

were observed, also within ORF-1 covered with primers from zone A-I to zone I-I (Fig. 3). Also some nucleotide changes in ORF1 were observed among the isolates Leona, Vltava and Kobra in the high similarity 5'- and 3'-parts of RPT coding sequence as well as in LSR (data not shown). However, these changes did not include functional protein domains shown in Fig. 2. No specific differences in the probing patterns were observed among the isolates PVS-O and PVS-CS, suggesting that the sequence determining the "CS character" of PVS must be restricted to a short sequence(s) and protein domain(s). Although most of PVS-CS isolates showed high similarity with the PVS-O ones, four (ASS, Peru, Peruanita and Yuguima) exhibited a high divergence of probing patterns from the rest of PVS isolates. The same was true for the isolate Bograital that, according to the bioassay,was a PVS-O. The lowest probing pattern similarity was observed for the isolate ASS, with which a weak PCR band was observed for the zone G-II, only. The primers of this zone covered the HEL motif V (Fig. 2, Table 2) at the 5'-end of the fragment and a stretch of conserved nucleotides at its 3'-end. A nucleotide probing system, developed for the 3'-portion of PVS-A (Mackenzie *et al*., 1989), also showed that this strain was very divergent from other PVS isolates tested (data not shown).

In previous studies on a limited number of isolates (Foster, 1991; Matoušek *et al*., 2000b), it was proposed that major differences between PVS-A and PVS-O genomes concerned the region of genes for 7 K and 11 K protein. For detailed investigations by thermodynamic analysis and sequencing the fragment J-III (now designated 7KCPLCP) of approximately 420 bp that covered the region from the C-terminal part of the 7 K protein to the N-terminal part of CP, including the CP translational enhancer (CP leader, CPL; Turner *et al*., 1999) was chosen. 7KCPCPL fragments from positive PVS-O and PVS-CS variants were pooled separately to form PVS-O- and PVS-CS-specific fragment pools. These pools containing sequence populations with melting points in the range from 42°C to 50°C, as shown on TGGE patterns (Fig. 4) were then analyzed by the cDNA heteroduplexes method and some variants were selected for sequencing. In total, 17 and 18 distinct sequence variants were identified in the PVS-O and PVS-CS pools, respectively. It was found that both cDNA pools differed in distinct nucleotide positions especially within the 5'-end of the CP gene, suggesting some divergence of these two groups of sequences. For instance, G or C at the position 232 of the 7KCPLCP fragment in the PVS-O group was replaced by A in all the PVS-CS variants. Also, distinct mutations accumulated at high frequency in other positions and in total 23 such positions were detected (data not shown). Some specific amino acid combinations were found in these two groups (Fig. 5A). E.g., methionine at the position 17 in combination with serine at the position 34 at the N-terminal part of CP was frequently associated with the CS character of PVS isolates. In addition, hydrophobic and polar amino acids were characteristic for the positions 17 and 34, respectively, in CS variants (Fig. 5B). Evolutionary distances presented in Fig. 6 showed that PVS-CS variants, positive by the RT-PCR probing (Table 3), were closely related to the European PVS-O variants and were more distantly related to the PVS-A 7KCPLCP sequence published by Mackenzie *et al*. (1989). According to previous work (Matoušek *et al*., 2000b), the 3'-portion of genome of PVS-A differed from that of PVS-O in 582 loci, 568 single nucleotide changes and 14 deletions/insertions.

Acknowledgements. This work was supported by the grants Nos. AV0Z50510513 from the Academy of Sciences of the CR GAČR 521/03/H160 and NAZV QF 3109 from the Ministry of Agriculture of the CR. This work was a part of a co-operative project No. 8/04 including Federal Center for Breeding Research on Cultivated Plants, Institute for Resistance Research and Pathogen Diagnostics, Aschersleben, Germany, Institute of Plant Molecular Biology, Academy of Sciences of Czech Republic, České Budějovice, Czech Republic. The authors thank Mmes M. Nielitz and H. Matoušková from the Institutes in Aschersleben and České Budějovice (for full addresses see the front page), respectively, for excellent technical assistance, and Drs C. Jeffries and V. Mulholland from Scottish Agricultural Science Agency, Edinburgh, UK for help in preparing the manuscript.

References

Adams MJ, Antoniw JF, Bar-Joseph M, Brunt AA, Candresse T, Foster GD, Martelli GP, Milne RG, Fauquet CM (2004): The new plant virus family *Flexiviridae* and assessment of molecular criteria for species demarcation. *Arch. Virol.* **149**, 1045–1060.

Bork P, Brown NP, Hegyi H, Schultz J (1996): The protein phosphatase2C (PP2C) superfamily: Detection of bacterial homologues. *Protein Sci*. **5**, 1421–1425.

Choi SA, Ryu KH (2003): The complete nucleotide sequence of the genome RNA of Lily symptomless virus and its comparison with that of other carlaviruses. *Arch. Virol.* **148**, 1943–1955.

Clark MF, Adams AN (1977): Characteristic of the microplate method of enzyme-linked immunosorbent assay for the detection of plant viruses. *J. Gen. Virol.* **34**, 475–483.

Foster GD (1991): Molecular variation between ordinary and Andean strains of potato virus S. *Res. Virol.* **142**, 413–416.

Foster GD (1992): The structure and expression of the genome of Carlaviruses. *Res. Virol.* **143**, 103–112.

Foster GD, Mills PR (1992): The 3´-nucleotide sequence of an ordinary strain of potato virus S. *Virus Genes* **6**, 213–220.

Koonin EV, Dolja VV (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.

Mackenzie DJ, Tremaine JH, Stace-Smith R (1989): Organization and interviral homologies of the 3'-terminal portion of potato virus S RNA. *J. Gen. Virol.* **70**, 1053–1063.

Matoušek J, Ptáček J, Dědič P, Schubert J (2000a): Analysis of variability of PVY^N P1 region using TGGE and electrophoresis of heteroduplexes. *Acta Virol.* **44**, 40–46.

Matoušek J, Schubert J, Dědič P, Ptáček J (2000b): A broad variability of potato virus S (PVS) revealed by analysis of virus sequences amplified by reverse transcriptase-polymerase chain reaction. *Can. J. Plant Pathol.* **22**, 29–37.

Monis J, Daniels S, de Zoeten GA, Slack SA (1987): Characterization of potato virus S (PVS) genomic RNA. *Phytopathology* **77**, 1742.

Monis J, de Zoeten GA (1990): Molecular cloning and physical mapping of potato virus S complementary DNA phytopathology. *Phytopathology* **80**, 446–450.

Morozov SY, Dolja VV, Atabekov JG (1989): Probable reassortment of genomic elements among elongated RNA-containing plant viruses. *J. Mol. Evol.* **29**, 52–62.

Murashige T, Skoog F (1962): A revised medium for rapid growth and bioassay with tobacco tissue cultures. *Physiol. Plant.* **14**, 473–497.

Poland D (1974): Recursion relation generation of probability profiles for specific-sequences macromolecules with long-range correlation. *Biopolymers* **13**, 1859–1871.

Ptáček J, Škopek J, Dědič P, Matoušek J (2002): Immunocapture RT-PCR probing of potato virus Y isolates. *Acta Virol.* **46**, 63–68.

Schumacher J, Meyer N, Riesner D, Weidemann HL (1986): Diagnostic procedure for detection of viroids and viruses with circular RNAs by "return"-gel electrophoresis. *J. Phytopathol.* **115**, 332–343.

Thompson N, Randles JW (2001): The genome organisation and taxonomy of Sugarcane striate mosaic associated virus. *Arch. Virol.* **461**, 1441–1451.

Turner R, Foster GD (1997): Deletion analysis of a translational enhancer upstream from the coat protein open reading frame of potato virus S. *Arch. Virol*. **142**, 167–175.

Turner RL, Glynn M, Taylor SC, Cheung M-K, Spurr C, Twell D, Foster GD (1999): Analysis of a translational enhancer present within the 5'-terminal sequence of the genomic RNA of potato virus S. *Arch. Virol.* **144**, 1451–1461.