MOLECULAR VARIABILITY OF THE COAT PROTEIN GENE OF POTATO VIRUS Y FROM TOBACCO IN CHINA

N. LI[1,2], X. WANG[1*], G. ZHOU[1], J. DONG[2]

[1]State Key Laboratory for Biology of Plant Diseases and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing, 100094, P.R. China; [2]Hebei Agricultural University, Baoding, P.R. China

Received December 9, 2005; accepted April 7, 2006

Summary. – Thirty-three tobacco samples showing typical symptoms of Potato virus Y (PVY) infection were obtained from tobacco fields in various regions of China. The results of indirect ELISA confirmed the infection with PVY. All the isolates had a capacity to infect tobacco systemically in greenhouse, causing either of two main symptoms: veinal necrosis and mosaic. The nucleotide and amino acid sequences of the coat protein (CP) gene and protein, respectively, of the isolates were determined. Comparison of the isolates revealed a high conservation of the CP gene with an identity of 83.2%. A phylogenetic tree of 41 Chinese isolates of PVY, based on complete CP gene, showed 3 groups corresponding to the strains PVY^{NTN} (A group), PVY^{o} (C group) and a putative new strain similar to PVY^{N} (B group). The amino acid sequences of complete CP protein of the isolates showed an identity of 87.6%. The highest identity was observed in the C-terminal half of the CP protein, where only 11 amino acid differences could be observed, in contrast to the N-terminal half with 22 differences.

Key words: Potato virus Y; Chinese isolates; coat protein; multiple sequence alignment; phylogenetic tree; tobacco; nucleotide sequencing

Introduction

PVY (the species *Potato virus Y*, the genus *Potyvirus*) contains a positive-sense genomic RNA encoding in a single ORF a polyprotein, which is processed into functional proteins by internal proteases P1, HC-Pro and NIa (Reichmann *et al*., 1992). PVY is naturally transmitted by aphids in a non-persistent manner with a great efficiency, causing epidemics and significant crop losses in potato, tomato, pepper, tobacco and other solanaceous plants worldwide (Shukla *et al*., 1994). The incidence and severity of the tobacco disease caused by PVY has increased rapidly over past years in central-eastern China along the Yellow River and Huai River as well as in Northeast China.

Several PVY strains are recognized on the basis of symptoms induced in infected potato (*Solanum tuberosum*) and tobacco (*Nicotiana tabacum*) (Kerlan *et al*., 1999). The ordinary or common strain (PVY^{O}) induces mild to severe mosaic and leaf drop in potato, and systemic mottle in tobacco. The tobacco veinal necrosis strain (PVY^{N}) induces very mild mottling in most potato cultivars with occasional necrotic leaves in some cultivars (Kerlan *et al*., 1999; Chachulska *et al*., 1997), but it induces severe systemic necrosis of leaf veins and petioles in tobacco. The potato tuber necrosis strain (PVY^{NTN}) induces chlorotic mottle to mosaic symptoms in potato plants and superficial to deeply sunken necrotic rings in tubers, but veinal necrosis symptoms in tobacco (Le Romancer *et al*., 1994).

Classical biological approaches to identify individual potyviruses and their strains on the basis of the host range,

*Corresponding author. E-mail: xfwang@caas.net.cn; fax: +86010-62896114.

Abbreviations: CP = coat protein; HC-Pro = helper component proteinase; MMLV = Moloney murine leukemia virus; NIa = nuclear inclusion protein a; 5'-NTR, 3'-NTR = 5'- and 3'-non-translated regions; PVY = Potato virus Y; PVY^{O} = PVY ordinary strain; PVY^{N} = PVY tobacco veinal necrosis strain; PVY^{NTN} = PVY tuber necrosis strain

symptomatology and serology, have so far failed. The sequences of the CP gene and 3'-non-translated region (3'-NTR) have been widely accepted as a molecular basis to distinguish individual potyviruses (Van der Vlugt *et al.*, 1993) or other viruses (Lin *et al.*, 2003; Handley *et al.*, 1998).

The potential for genetic variation in RNA viruses is high, because their RNA-dependent RNA polymerases lacks a proof-reading ability (Domingo and Holland, 1997; Roossinck, 1997). Several regions of the PVY genome, namely the 5'-non-translated region (5'-NTR) (Tordo *et al.*, 1995), P1 gene (Tordo *et al.*, 1995), CP gene (Čeřovská *et al.*, 2001; Shukla *et al.*, 1988a,b; Shukla and Ward, 1988; Sudarsono *et al.*, 1993), 3'-NTR, as well as the whole genome (Glais *et al.*, 1998, 2002) have been thoroughly studied by RFLP (Glais *et al.*, 1996, 1998; Blanco-Urgoiti *et al.*, 1996) and sequencing (Van der Vlugt *et al.*, 1993; Marie-Jeanne Tordo *et al.*, 1995; Aleman-Verdaguer *et al.*, 1997). The studies dealing with geographical distribution and molecular variation of closteroviruses or their isolates have provided useful information (Rubio *et al.*, 1999, 2001). Sequence comparison and phylogenetic analysis of 5'-NTR and P1 gene of 11 isolates of PVY^{N} and PVY^{NTN} indicated that potato PVY^{N} isolates from Europe and North America form separate groups corresponding to original geographic locations of collection: European (EU) $PVY^{N/NT}$ and North American (NA) $PVY^{N/NTN}$ (Nie and Singh, 2002a). Although the CP gene of several Chinese PVY isolates from tobacco and potato has been sequenced, there is still limited information about their molecular variability.

This study reports on molecular variability of the CP gene of 33 Chinese tobacco PVY isolates in relation to their strain assignment and geographical distribution.

Materials and Methods

Plant material. Samples of tobacco plant with typical disease symptoms were collected in tobacco fields in various regions of China (Table 1).

PVY isolates were maintained by mechanical inoculation of tobacco (*Nicotiana tabacum* cv. White Burley) kept in conditioned insect-proof greenhouse.

Total RNA was extracted from plant samples using Trizol reagent (Invitrogen) according to the manufacturer's instructions. It was finally dissolved in RNase-free distilled water and stored at -80°C.

RT-PCR. cDNA to the CP gene was synthesized using total RNA as template, MMLV reverse transcriptase and CP1 oligonucleotide primer (5'-TCACATGTTCTTGACTCCAAG-3', nt 738–804) in a reaction mixture and under conditions described earlier (Du *et al.*, 2004). The cDNA was amplified by using the primers CP1 (reverse) and CP2 (5'-TCAAGCAAATGACACAATTG-3', nt 1–16, forward). The amplicon of 804 bp corresponded to complete PVY CP gene (nt 1–804).

Cloning and sequencing. The PCR products were resolved by agarose gel electrophoresis. The 804 bp target bands were excised and purified using the TaKaRa Agarose Gel DNA Purification Kit Version 2.0 (Takara Co, P.R. China). The purified DNA fragments were cloned by using the pGEM T-Easy Vector System 1 (Promega). The obtained clones were sequenced by the BOYA Biotech Co. (P.R. China). Each sequence was determined in both directions.

Sequence comparison and phylogenetic analysis. The determined nucleotide and deduced amino acid sequences of the CP gene of the PVY isolates were compared to those deposited in the GenBank database using BLAST tools. The sequences were assembled and analyzed with the aid of VectorNTI Version 9.0 software (Invitrogen). Multiple sequence alignment and phylogenetic analysis were carried out using the CLUSTAL W Version 5.00 program of the DNAStar software. The Pepper severe mosaic virus was used as an out-group.

Results

Disease symptoms in tobacco

All 33 PVY isolates obtained in this study were capable to systemically infect *N. tabacum* (cv. White Burley), but the symptoms varied (Table 1). Twenty-two isolates induced typical veinal necrosis, often combined with leaf distortion, vein clearing, and/or systemic necrosis. The disease severity varied, so that some isolates induced systemic. Nine isolates induced systemic mottle, sometimes combined with necrotic spots, leaf distortion and/or systemic necrosis. Two isolates induced no typical mottle symptoms except necrotic spots. Based on the symptoms on tobacco, all the isolates could be divided into two groups, the vein necrosis and mosaic groups.

CP gene and protein sequences

Analysis of the nucleotide sequences of the CP gene of all thy isolates revealed an uninterrupted 801 nts long ORF encoding a protein of 267 amino acids, beginning with Ala, as reported for most PVY strains (Van der Vlugt *et al.*, 1993). The DAG motif, required for aphid transmission, was found in all sequences at the positions 6–8. The obtained nucleotide sequences were deposited in the GenBank (Table 1).

Comparison of the isolates based on CP gene and protein sequences

The nucleotide sequence of complete CP gene showed a 83.2% identity among all the PVY isolates, while pairwise comparisons among them gave 1–98 % of polymorphic positions.

The deduced amino acid sequences of complete CP protein showed a 87.6% identity among all the PVY isolates,

Table 1. Tobacco PVY isolates examined in this study

Group	Isolate	Place of origin	Symptoms on tobacco	Acc. No.
	AFY1	Fengyang	VCN, SN	AY841257
	AFY4-1	Fengyang	VCN, SN	AY742717
	BJ	Beijing	VCN, NS, LD, N	AY742714
	BJ0-1	Beijing	VCN	AY742723
A	BJ0-3	Beijing	VCN	AY742724
	BJ2-1	Beijing	VCN	AY742722
	HMDJ11	Mudanjiang	VCN, SN	AY742728
	HXCH30	Xuchang	VN, VB, NS	AY841260
	HXCH44	Xuchang	VCN, LD	AY841268
	AFY10	Fengyang	VCN, SN	AY742720
	AFY6	Fengyang	VCN, SN	AY742718
	AFY8-2	Fengyang	VCN, SN	AY742729
	HMDJ13	Mudanjiang	VCN, SN	AY742732
	HMDJ8-2	Mudanjiang	VCN, SN	AY742727
B	HMDJ8-3	Mudanjiang	VCN, CS,LD, N, VN	AY742731
	AFY8-1	Fengyang	VCN, SN	AY742719
	HQZH	Qingzhou	VCN	AY742716
	HXCH24	Xuchang	NS, VN, LD	AY841258
	HXCH47	Xuchang	M, LD, VN	AY841267
	LSHY	Shenyang	VCN	AY742715
	AFY12-3	Fengyang	VCN, SN	AY742721
	HXCH43	Xuchang	VCN, LD	AY841266
	GFG1	Fengyang	M	AY742725
	GFG5	Fengyang	NS	AY742730
	GJSH	Jinsha	M, LD, NS	AY742726
	HXCH2	Xuchang	M, VN	AY742733
C	HXCH25	Xuchang	NS, NS, LD	AY841259
	HXCH31	Xuchang	M	AY841261
	HXCH35	Xuchang	M, NS	AY841262
	HXCH36	Xuchang	M, NS	AY841264
	HXCH38	Xuchang	M, NS	AY841263
	HXCH39	Xuchang	M, NS	AY841265
	HXCH46	Xuchang	M, LD, NS	AY841269

VN = vein necrosis; LD = leaf distortion; VC = vein clearing; M = mosaic; SN = systemic necrosis; VCN = vein clearing and necrosis; NS = necrotic spots.

while pairwise comparisons gave 1–17% polymorphic positions.

The amino acid sequence identities within the isolates of the mosaic and veinal necrosis groups were 97% and 92%, respectively, but the identity between these two groups was 91%.

In total, 33 positions varied within the 268 amino acids long protein: 22 in the N-terminal and 11 in the C-terminal halves of the protein. The CP core region was highly homologous, with most differences restricted to only a few positions in two main regions in the N terminal half: 8–36 (Fig. 1) and 98–142 (Fig. 2), indicating "hot-spots" of mutation (Van der Vlugt *et al*., 1993). The positions 8–36 had the highest mutation rate. There was no apparent association of the sequences with the geographic origin of the respective isolates.

Phylogenetic analysis of the isolates

A phylogenetic tree was constructed based on the nucleotide sequence of complete CP gene of 41 Chinese tobacco PVY isolates (33 recent isolates and 8 isolates published earlier) and 6 standard "non-Chinese" tobacco PVY isolates (Fig 3).

The isolates obtained in this study clustered in 3 main groups, A, B and C (Table 1). As the BJ isolate of the A group had with the isolate NN-UK-NTN (a standard PVY^{NTN}) the highest nucleotide identity (99.0%), it was concluded that the A group isolates may belong to the PVY^{NTN}. For the LSHY isolate, a representative of the B group there was found no standard PVY strain with at least 97% nucleotide identity. The closest was the SCRI-N isolate (PVY^{N}) with a 94% identity. Therefore it was concluded that, as the

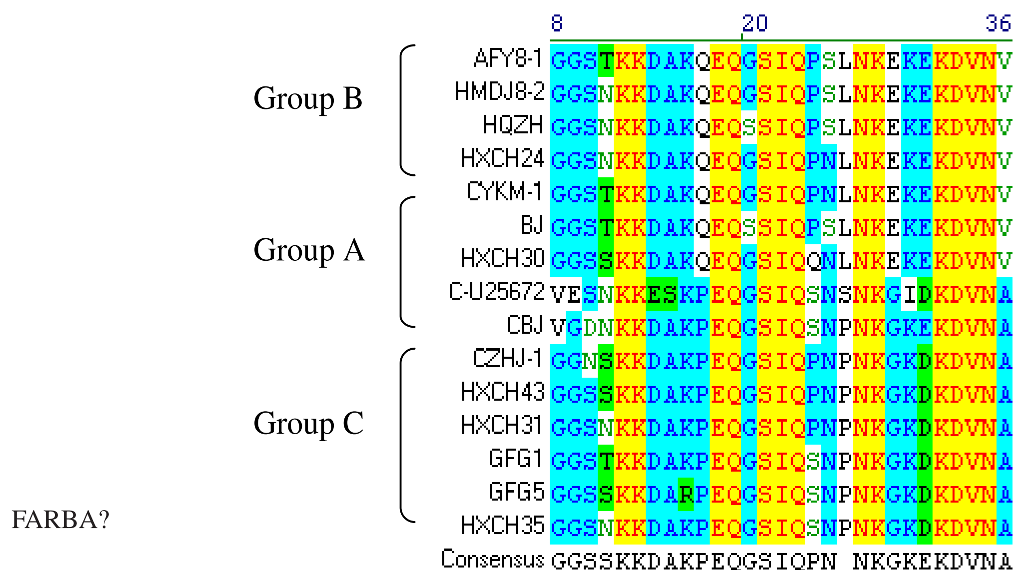

Fig. 1

Amino acid sequences (positions 8–36) of CP gene of selected Chinese PVY isolates

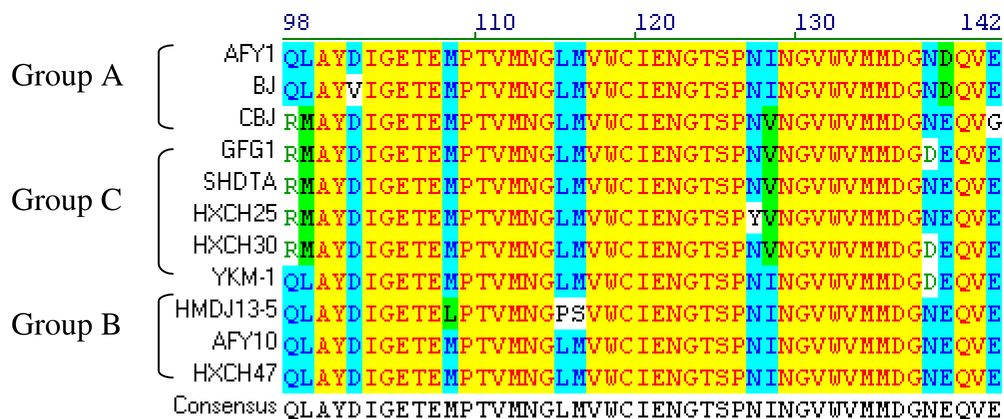

Fig. 2

Amino acid sequences (positions 98–142) of CP gene of selected Chinese PVY isolates

isolates of this group cannot be assigned to any standard PVY strain they may represent a new strain or a new PVY^{N} variant. The nucleotide identities between the GFG1 isolate, a representative of the C group, and standard isolates of PVY^{O} were all over 98%. Therefore the C group isolates were assigned to PVY^{O}

The symptoms on tobacco infected with the isolates corresponded to their group type. The 20 isolates in the groups A (PVY$^{N/NTN}$) and B (most similar PVY^{N}) produced a typical veinal or vein necrosis and clearing, diagnostic symptoms for PVY^{N}. Most of the 13 isolates of group C produced a typical mosaic symptom like PVY^{O}, except for four group C isolates (AFY12-3, HXCH43, GFG5, XCH25) which produced veinal and spot necrosis. There was no correlation between the groups and geographic origin of the respective isolates.

Discussion

The tobacco PVY isolates obtained in this study originated from 10 different regions of China. Therefore,

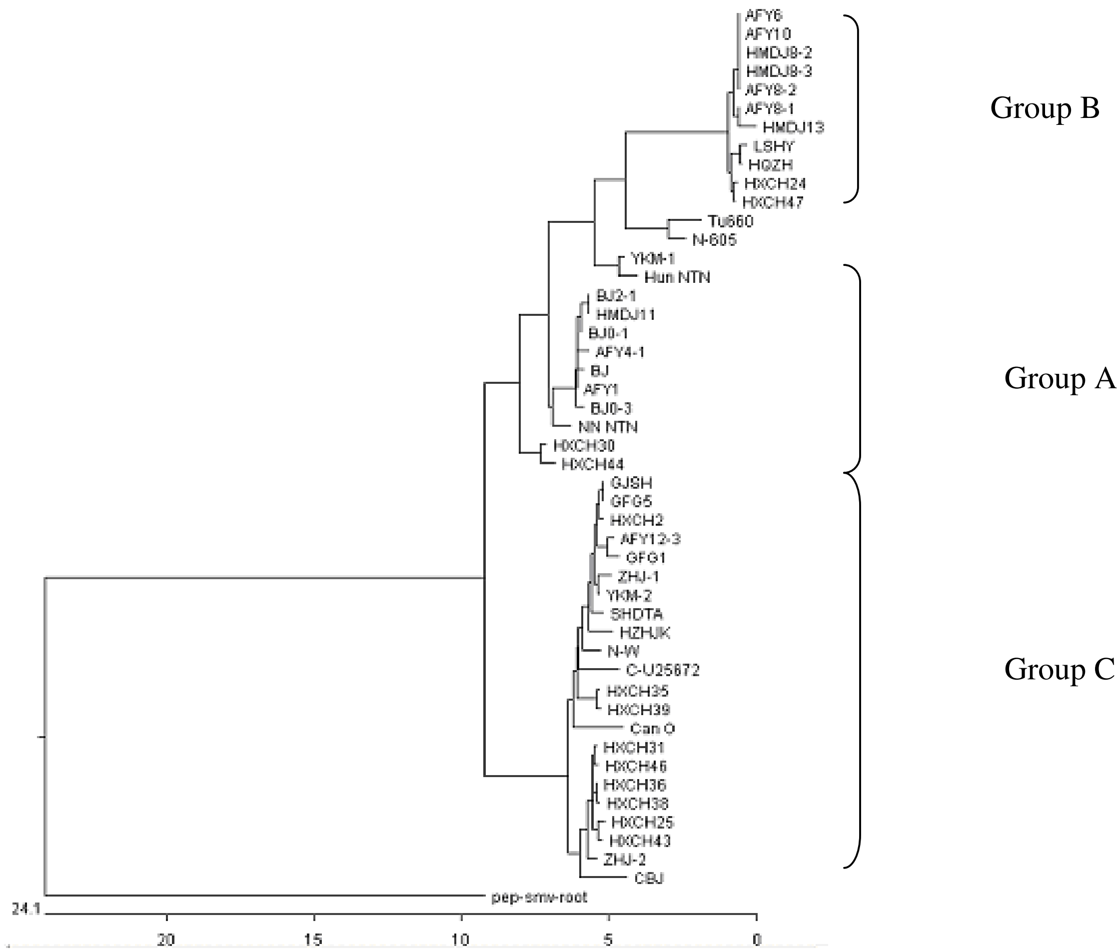

Fig. 3

Phylogenetic tree of Chinese PVY isolates and strains based on CP gene

Forty-one Chinese tobacco isolates (33 recent isolates and 8 isolates published earlier) and 6 standard "non-Chinese" tobacco isolates were included. The "standard" isolates from other countries that represent the N, NTN, N-W, NA-N, and O strains/variants are: X97895 (N605), M95491 (Hun NTN), AJ390296 (NN-UK-NTN), Z70238 (N-W), AY166866 (Tu660), U09509 (Can O).

these isolates may be regarded as representative of most of the PVY strains infecting tobacco in this country. The tobacco isolates collected and sequenced in this study did not show correspondence between geographic origin and phylogenetic clustering. Based on the symptoms and phylogenetic similarity, PVY strains can be divided into two main groups, PVY^{O} (group C) and $PVY^{N/NTN}$ (groups A and B), the latter including a major PVY^{NTN} branch, (group A) (Chachulska *et al.*, 1997; Glais *et al.*, 1998; Tordo *et al.*, 1995). The PVY^{N} and PVY^{NTN} isolates from a given geographical region were more closely related than the PVY^{N} and PVY^{NTN} isolates from another continent, suggesting that the PVY^{NTN} isolates might have evolved from locally predominant PVY^{N} strains (Nie and Singh, 2002a). However, both groups of PVY^{N} serotype have now been reported in both Europe and Northern America (J. Lorenzen, personal communication). Furthermore, the isolates in groupsA and B did not cluster closely with the isolates from other continents but formed their own separate group. However, the group C isolates clustered closely with PVY^{O} isolates from other countries. So the isolates in group C collected from 11 different regions in China (inclusive of 8 sequences from GenBank) and the PVY^{O} isolates were related more closely to each other and to PVY^{O} isolates from other countries than to the $PVY^{N/NTN}$ isolates of closer geographic proximity. The group C members that produced VN on tobacco are probably similar to recombinant PVY^{N}-W (Chachulska *et al.*, 1997; Glais *et al.*, 2002) or $PVY^{N:O}$ (Nie

and Singh, 2003; Piche *et al*., 2004). However, the PVY^{N}-W isolates do not appear as abundant as it was reported for Europe (Kerlan *et al*., 1999).

PVY is a heterogeneous virus, as the efforts to logically assign individual PVY isolates to strains or variants has proven to be challenging. It has been attempted to characterize PVY isolates by RT-PCR targeting specific areas of the genome (Boonham *et al*., 2002; Nie and Singh, 2002a,b), especially the P1 gene and 3'-NTR (Marie-Jeanne Tordo *et al*., 1995; Van der Vlugt *et al*., 1993). In this way different PVY strains could be distinguished (Van der Vlugt *et al*., 1993).

We can conclude that this study, based on the phylogenetic analysis of the CP gene in 33 Chinese tobacco PVY isolates, (i) the isolates of the A group belong to PVY^{NTN}, (ii) the isolates of the C group belong to PVY^{O}, and (iii) the isolates of the B group may represent a new PVY strain or new variant of PVY^{N}. Nevertheless, the latter possibility should be further investigated.

Acknowledgements. The authors thank Dr. J. Lorenzen, University of Idaho, Moscow, USA, and Dr. Z. Fan, China Agriculture University, Beijing, P.R. China, for critical reading of the manuscript. A financial support was provided by the project No. 2002-1870 (Tobacco China).

References

Aleman-Verdaguer ME, Goudou-Urbino C, Dubern J (1997): Analysis of the sequence diversity of the P1, HC, P3, NIb and CP genomic regions of several yam mosaic potyvirus isolates. *J. Gen. Virol.* **78**, 1253–1264.

Boonham N, Walsh K, Preston S, North J, Smith P, Barker I (2002): The detection of tuber necrotic isolates of *potato virus Y,* and the accurate discrimination of PVY^{o}, PVY^{N}, and PVY^{C} strains using RT-PCR. *J. Gen. Virol.* **102**, 103–112.

Blanco-Urgoiti B, Sanchez F, Dopazo J (1996): A strain-type clustering of potatovirus Y based on the genetic distance between isolates calculated by RFLP analysis of the amplified coat protein gene. *Arch. Virol.* **141**, 2425–2442.

Čeřovská N, Moravec T, Filigarová M (2001): Nucleotide sequences of 5'-terminal parts of coat protein genes of various isolates of NTN strain of *potato virus Y*. *Acta Virol.* **45**, 55–59.

Chachulska AM, Chrzanowska M, Robadlia C (1997): Tobacco veinal necrosis determinants are unlikely to be located within the 5' and 3'-terminal sequences of the potato virus Y genome. *Arch. Virol*. **142**, 765–779.

Domingo E, Holland JJ (1997): RNA virus mutations and fitness for survival. *Annu. Rev. Micorbiol.* **51**, 151–178.

Du G, Wang X, Zhou G (2004): Digoxigenin-labelled cDNA probes for the detection of *Tobacco mosaic virus, Cucumber mosaic virus* and *Potato virus Y. Acta Phytopatohol. Sinica.* **34**, 75–79.

Glais L, Kerlan C, Tribodet M (1996): Molecular characterization of potato virus Y^{N} isolates by PCR-RFLP. *Eur. J. Plant Pathol.* **102**, 655–662.

Glais L, Tribodet M, Kerlan C (2002): Genomic variability in Potato potyvirus Y (PVY)-evidence that PVY^{NW} and PVY^{NTN} variants are single to multiple recombinants between PVY^{o} and PVY^{N} isolates. *Arch. Virol.* **147**, 363–378.

Glais L, Tribodet M, Gauthier JP (1998): RFLP mapping of the whole genome of ten viral isolates representative of different biological groups of potato virus Y. *Arch. Virol.* **143**, 2077–2091.

Handley JA, Smith GR, Dale1 JL (1998): Harding1 Sequence diversity in the coat protein coding region of twelve sugarcane mosaic potyvirus isolates from Australia, USA and South Africa. *Arch. Virol.* **143**, 1145–1153.

Kerlan C, Tribodet M, Glais L (1999): Variability of *potato virus Y* in potato crops in France. *Phytopathology* **147**, 643–651.

Le Romancer M, Kerlan C, Nedellec M (1994): Biological characterization of various geographical isolates of *potato virus Y* inducing superficial necrosis on potato tubers. *Plant Pathol*. **43**, 138–144.

Lin HX, Rubio L, Smythe A (2003): Genetic diversity and biological variation among California isolates of Cucumber mosaic virus. *J. Gen. Virol.* **84**, 249–258.

Marie-Jeanne Tordo V, Chachulska A M, Fakhfakh H, Le Romancer M, Robaglia C, Astier-Manifacier S (1995): Sequence polymorphism in the 5'NTR and P1 coding region of genomic RNA. *J. Virol. Methods* **76**, 939–949.

Nie XZ, Singh RP (2002a): Probable geographical grouping of PVY^{N} and PVY^{NTN} based on sequence variation in P1 and methods for differentiating North American PVY^{NTN}. *J. Virol. Methods* **103**, 145–156.

Nie XZ, Singh RP (2002b): A new approach for the simultaneous differentiation of biological and geographical strains of *potato virus Y* by uniplex and multiplex RT-PCR. *J. Virol. Methods* **104**, 40–54.

Nie XZ, Singh RP (2003): Specific differentiation of recombinant $PVY^{N:O}$ and PVY^{NTN} isolates by multiplex RT-PCR. *J. Virol. Methods* **113**, 69–77.

Piche LM, Singh RP, Nie X, Gudmestad NC (2004) Diversity among PVY field isolates obtained from potatoes grown in the United States. *Phytopathology* **94**, 1368-1375.

Riechmann JL, Lain S (1992): Highlights and prospects of potyvirus molecular biology. *J. Gen. Virol.* **73**, 1–16.

Roossinck MJ (1997): Mechanisms of plant virus evolution. *Annu. Rev. Phytopathol.* **35**, 1953–1965.

Rubio L, Soong J, Kao J (1999): Geographic distribution and molecular variation of isolates of three whitefly-borne closteroviruses of cucurbits. *Phytopathology* **89**, 707–711.

Rubio L, Abou-Jawdah Y, Lin HX (2001): Geographically distant isolates of the Crinivirus cucurbit yellow stunting disorder virus show very low genetic diversity in the coat protein gene. *J. Gen. Virol.* **82**, 929–933.

Shukla DD, McKern NM, Ward CW (1988a): Coat proteins of potyviruses 5. Symptomatology, serology and coat protein sequences of three strains of passionfruit woodiness virus. *Arch. Virol.* **102**, 221–232.

Shukla DD, Thomas JE, Mckern NM, Tracy SL (1998): Coat protein of Potyviruses 4. Comparison of biological properties, serological relationships and coat protein amino acid sequences of four strains of potato virus Y. *Arch. Virol.* **102**, 207–219.

Shukla DD, Ward CW (1988b): Amino acid sequence homology of coat proteins as a basis for identification and classification of the potyvirus group. *J. Gen. Virol.* **69**, 2703–2710.

Shukla DD, Ward CW (1989): Identification and classification of potyvirus on the basis of coat protein sequence data and serology. *Arch. Virol.* **106**, 171–200.

Shukla DD, Ward C W, Brunt AA (1994): *The Potyviridae Wallingford*. CAB International.

Sudarsono, Woloshuk SL, Xiong Z (1993): Nucleotide sequence of the capsid protein cistrons from six potato virus Y (PVY) isolates infecting tobacco. *Arch. Virol.* **132**, 161–170.

Tordo VM, Chachulska AM, Fakhfakh H (1995): Sequence polymorphism in the 5'NTR and in the P1 coding region of potato virus Y genomic RNA. *J. Gen. Virol.* **76**, 939–49.

Van der Vlugt RAA, Leunissen J, Goldbach R (1993): Taxonomic relationships between distinct potato irus Y isolates based on detailed comparisons of the viral coat proteins and 3'-nontranslated regions. *Arch. Virol.* **131**, 361–375.