Genetic Diversity of *Potato virus Y* Infecting Tobacco Crops in China

Y. P. Tian, J. L. Liu, C. L. Zhang, Y. Y. Liu, B. Wang, X.-D. Li, Z. K. Guo, and J. P. T. Valkonen

First, third, fourth, fifth, and sixth authors: Laboratory of Plant Virology, Department of Plant Pathology, Shandong Agricultural University, Tai'an, Shandong 271018, P. R. China; second author: College of Plant Sciences, Jilin University, Changchun, Jilin 130062, P. R. China; first and eighth authors: Department of Plant Biology and Forest Genetics, Genetics Center, Swedish University of Agricultural Sciences, PO Box 7080, SE-75007, Uppsala, Sweden, and Department of Agricultural Sciences, PO Box 27, FIN-00014, University of Helsinki, Finland; seventh author: Heilongjiang Tobacco Institute, Mudanjiang, Heilongjiang 157011, P. R. China.
Accepted for publication 20 October 2010.

ABSTRACT

Tian, Y. P., Liu, J. L., Zhang, C. L., Liu, Y. Y., Wang, B., Li, X.-D., Guo, Z. K., and Valkonen, J. P. T. 2011. Genetic diversity of *Potato virus Y* infecting tobacco crops in China. Phytopathology 101:377-387.

Genetic variability of *Potato virus Y* (PVY) isolates infecting potato has been characterized but little is known about genetic diversity of PVY isolates infecting tobacco crops. In this study, PVY isolates were collected from major tobacco-growing areas in China and single-lesion isolates were produced by serial inoculation on *Chenopodium amaranticolor*. Most isolates (88%) caused systemic veinal necrosis symptoms in tobacco. Of these, 16 isolates contained a PVY^{O}-like coat protein (CP) and PVY^{N}-like helper component proteinase (HC-pro) and, in this respect, were similar to the $PVY^{N\text{-}Wi}$, $PVY^{N:O}$, and PVY-HN2 isolates characterized from potato in Europe, the United States, and China, respectively; two isolates contained a PVY^{O}-like HC-pro and a PVY^{N}-like CP; another two isolates had recombination junctions in the CP-encoding region. Both the HC-pro and CP of PVY were under negative selection as a whole; however, seven amino acids in HC-pro and six amino acids in CP were under positive selection. Selection pressures differed between the subpopulations of PVY distinguished by phylogenetic analysis of HC-pro and CP sequences. When PVY isolates from potato were included, no host-specific clustering of the PVY isolates was observed in phylogenetic and nucleotide diversity analyses, suggesting frequent spread of PVY isolates between potato and tobacco crops in the field.

Potato virus Y (PVY, genus *Potyvirus,* family *Potyviridae*) is the type species of the largest genus of plant-infecting RNA viruses (13). It has a positive-sense single stranded RNA genome of ≈9,700 nucleotides (nt) (24,36) and contains an open reading frame (ORF) translated into a large polyprotein, which is subsequently cleaved into 10 mature proteins by three viral-encoded proteinases (Fig. 1). Additionally, a small ORF (pipo, 75 codons) created by frame-shifting was recently identified in the P3 protein-encoding region (10).

PVY was first described in potato (*Solanum tuberosum* L.; family *Solanaceae*) as the aphid-transmissible component of a virus complex which also included the non-aphid-transmissible *Potato virus X* (*Potexvirus*) (55). PVY mainly infects solanaceous crop plants and causes great yield losses in them (50,54). In tobacco (*Nicotiana tabacum* L.), PVY is found in many parts of the world (17,18,28,29,35,38,41,57). The isolates of PVY which cause systemic vein-banding symptoms reduce yields significantly (51,52,57) but infection with isolates that cause systemic veinal necrosis (PVY^{N}) (11) results in the greatest yield losses in tobacco cultivation (29,35,41).

Induction of systemic veinal necrosis in tobacco leaves by PVY is associated with certain amino acids such as K-400 and E-419 at the C-proximal part (60) and D-205 in the central part (25) of the viral helper component proteinase (HC-pro). However, it is not ascertained that these amino acids would be the only determinants of veinal necrosis in PVY (25,48). The host determinants for the veinal necrosis phenotype are unknown. Some PVY isolates which normally induce vein-banding symptoms in tobacco can induce necrotic symptoms in cultivars carrying the nematode resistance gene *Rk* (17,18). The replicase protein (NIb) of PVY may determine the necrotic response in the presence of *Rk* but mutational analysis of NIb is needed to identify the significant amino acids (14).

Isolates of the PVY^{N} strain group are not harmful only to tobacco but are also problematic in potato, in which they overcome the genes *Ny* and *Nc* conferring hypersensitive resistance to PVY isolates belonging to the strain groups PVY^{O} and PVY^{C}, respectively (54). Despite attempts to restrict their spread (53), PVY^{N} isolates are now globally distributed and of great concern because of frequent appraisal of new variants via mutation and recombination with other PVY isolates (24,25, 42,54). Some of them cause necrotic symptoms in potato tubers and are designated to the strain PVY^{NTN}, while isolates belonging to the PVY^{N} strain group usually cause mild or no obvious symptoms in potato foliage (30,54). The viral determinants for the potato tuber necrosis phenotype seem to be different from those needed for systemic veinal necrosis in tobacco and remain to be identified (8,25).

Systemic veinal necrosis in tobacco is another criterion to classify an isolate to strain group PVY^{N}, in addition to response in potato indicators. Detection of PVY^{N} with monoclonal antibodies (MAbs) specific for PVY^{N} coat protein (CP) and by phylogenetic analysis of the CP-encoding sequences will provide some information but with limited value (7,39,54). Recombination is common in *Potyvirus* spp. (46) and many types of recombinants have been detected in PVY (8,16,24,48). Recombination between isolates of PVY^{N} and PVY^{O} or PVY^{C} may generate novel isolates in which detection of the CP with MAbs or phylogenetic grouping based on the CP-encoding sequence would fail to predict

Corresponding author: X.-D. Li;
E-mail address: xdongli@sdau.edu.cn; zhaiyiqing@126.com

*The *e*-Xtra logo stands for "electronic extra" and indicates that the online version contains a supplementary table.

doi:10.1094/PHYTO-02-10-0055

induction of veinal necrosis by the isolate. Indeed, isolates first found in Poland and designated to the strain PVY^{N-Wi} (considered sometimes synonymous to $PVY^{N:O}$) contain a PVY^{O}-type CP and, hence, are not detected with PVY^{N}-specific MAbs; however, they carry a PVY^{N}-type HC-Pro and cause veinal necrosis in tobacco (9,16,54). Some isolates reacting with certain PVY^{N}-specific MAbs actually belong to strain group PVY^{O} (26). In some cases, mutation of a single amino acid in the CP may alter the serotype (7).

Multiple studies have been carried out on isolates of PVY^{N} and PVY^{NTN} recently but these studies were focused on PVY isolates infecting potato in the field (23,24,42,54). However, no comprehensive study is available on molecular genetic diversity of PVY populations from tobacco crops (31). In China, tobacco is an important crop in which PVY causes economically significant losses (22,32,64,65). Some studies suggest that, as in potato crops in other parts of the world (44,48,61), tobacco veinal necrosis-inducing isolates of PVY have become more common in tobacco crops in China in the recent past. For example, whereas, in the early 1990s, ≤60% of the tobacco plants tested in Shandong province were infected with PVY isolates which did not induce veinal necrosis symptoms (65), ≤70% of the isolates detected in tobacco crops a few years later induced systemic veinal necrosis (63).

The aim of this study was to collect and characterize PVY isolates from tobacco plants in several tobacco-growing areas in China; determine the strain groups of the isolates using biological, serological, and molecular assays; and analyze the HC-pro- and CP-encoding sequences of the isolates for recombination, selection pressure, and phylogenetic relationship.

MATERIALS AND METHODS

Virus isolates and detection. Leaves displaying symptoms of mottle, mosaic, or veinal necrosis were collected from tobacco plants from five major tobacco-growing areas in China, including Shandong, Henan, Anhui, Gansu, and Heilongjiang Provinces. They were tested by plate-trapped antigen enzyme-linked immunosorbent assay (ELISA) (33) using PVY polyclonal antibodies prepared in the Laboratory of Plant Virology, Department of Plant Pathology, Shandong Agricultural University (21). Because plants infected in the field may contain mixed infections with virus strains, the original ELISA-positive tobacco samples were subjected to three successive cycles of single-lesion isolation on *Chenopodium amaranticolor* plants to reduce likelihood of mixed infections in the samples subjected to the various analyses. This was done being aware that some PVY isolates (i.e., belonging to PVY^{N} and PVY^{NTN} strains) may not induce local lesions on *C. amaranticolor* (3,30,38). Twenty-five isolates were obtained and maintained in plants of *N. tabacum* cv. Samsun in a greenhouse under natural daylight at 20 to 30°C. Symptoms were recorded. Serogroups of the isolates were determined by triple-antibody sandwich ELISA using PVY strain-specific MAbs following the supplier's instructions (Neogen Europe Ltd., Ayr, Scotland, UK). MAb1129 is specific to PVY^{O} and MAb1128 is specific to PVY^{N}. Four additional PVY-positive leaf samples were directly collected from the fields and included in sequence analysis without passages on *C. amaranticolor*.

RNA extraction, reverse-transcriptase polymerase chain reaction, and sequencing. Total RNA was extracted from the single-lesion PVY isolates or naturally infected tobacco leaves following the Qiagen Kit manufacturer's instructions (Qiagen, Hilden, Germany). Two primers, Poty-R (5′-GGTCGACTGC AGGATCCAAGC$(T)_{15}$-3′) and HC-R (5′-ACCAACTCTATAATG TTTTATATC-3′), were used to synthesize the cDNAs of CP- and HC-pro-encoding sequences separately (12,58). RevertAid *Moloney murine leukemia virus* reverse transcriptase was used for the cDNA synthesis according to the product protocol (MBI Fermentas, St. Leon-Rot, Germany).

The 3′-proximal part of the PVY genome was amplified by polymerase chain reaction (PCR) using Poty-R and degenerate forward primer PVY-CP-F (5′-GBAAAYGAHACAATYGATGC-3′), designed according to the first nucleotides of PVY CP-encoding sequences.

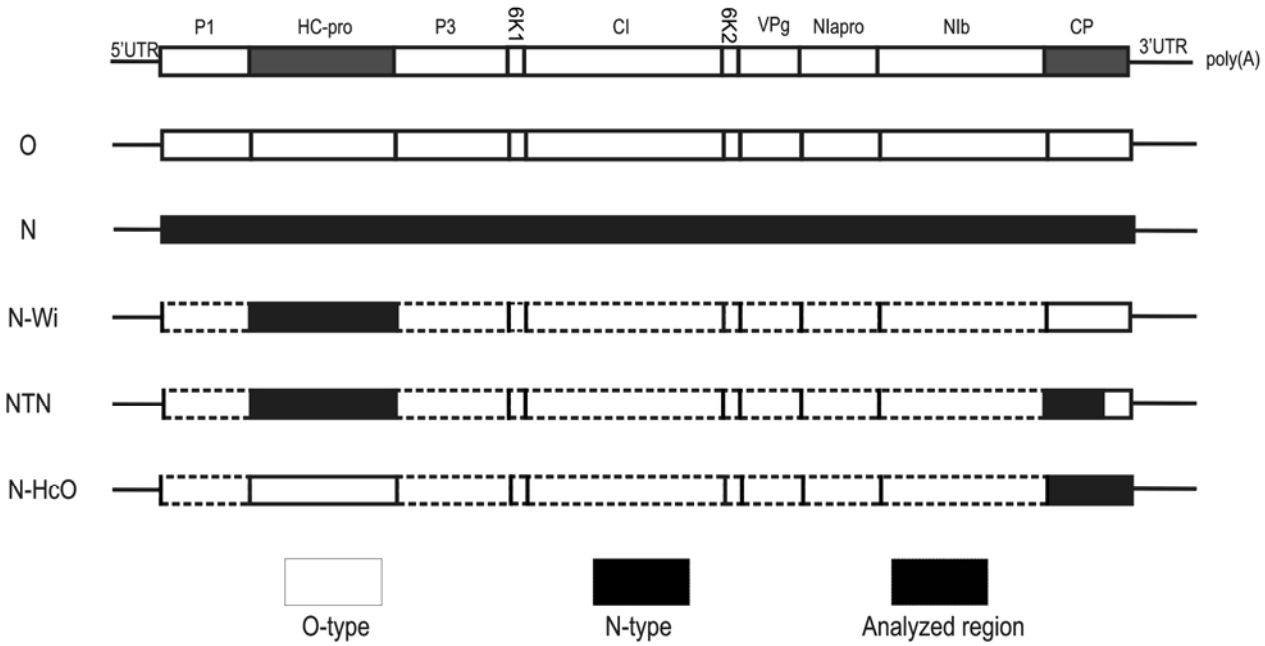

Fig. 1. Schematic presentation of the genome structure of *Potato virus Y* (PVY) and the structures of various recombinants. Genomic structures of the recombinant strains N-Wi, NTN, and N-HcO predicted to result from recombination between the O and N strains are schematically presented. Genomic regions corresponding to helper component proteinase (HC-pro) and coat protein (CP) were analyzed. Dashed line indicates that the respective part of the genome was not characterized in this study. The single-stranded positive-sense RNA genome contains an untranslated region (UTR) at the 5′ and 3′ end (5′-UTR and 3′-UTR, respectively) and is polyadenylated. Depicted parts of the polyprotein correspond to P1, a serine proteinase; P3, the third protein; 6K1 and 6K2, 6-kDa proteins; CI, cylindrical inclusion protein; VPg, viral genome-linked protein; NIa-Pro, the main viral proteinase; and NIb, the replicase.

Primers HC-R and HC-F (5′-GGCCTATTCATAGTGCGTGG-3′, corresponding to the 3′ end of the PVY P1 protein-encoding sequence) were used to amplify the HC-pro-encoding sequences according to the procedure used in previous study (58).

PCR products were subjected to 1.0% (wt/vol) agarose gel electrophoresis. The fragments of ≈1.2 kb (3′-proximal genomic region) and 1.4 kb (HC-pro-encoding region) were purified separately using the QIA-quick PCR purification kit (Qiagen). Cloning and sequencing strategies were the same as in previous studies (34,58). All sequences were deposited to the National Center for Biotechnology Information database (Supplementary Table 1).

Recombination detection and phylogenetic analysis. In addition to the sequences determined in this study, additional PVY sequences were obtained from the DPVweb database (1) (Supplementary Table 1). All nucleotide sequences were aligned using ClustalX 1.83 and adjusted manually. Nucleotide and amino acid sequence identities were calculated using Clustal W assembled in the MegAlign program of DNASTAR 6.0 software package (DNASTAR Inc., Madison, WI).

The software package Recombination Detection Program (RDP, version 3.27) (37) was used to screen HC-pro- and CP-encoding sequences for possible recombination events. The RDP package includes programs RDP, GENECONV, BOOTSCAN, MaxChi, CHIMAERA, SiSCAN, and 3SEQ. For each putative recombination break point, a Bonferroni correction P value (with a cut-off of $P < 0.05$) was calculated. To avoid false positive events, only those supported by at least four kinds of software or two methods (phylogenetic, substitution or distance comparison method) were considered to be "clear" recombinants, whereas those detected with fewer kinds of software and methods were regarded as "tentative" recombinants (43,59,62).

To determine whether the recombination events detected using the RDP software package were associated with major changes in tree topology, we conducted phylogenetic analysis on either side of the putative recombination break point using maximum likelihood (ML) methods (6). Nonrecombinant isolates were included for comparison.

Only HC-pro and CP sequences which did not contain predictable recombination junctions were analyzed for phylogenetic relatedness, including the CP- and HC-pro-encoding sequences of 30 PVY isolates with complete genomic sequences available and the sequences of Chinese PVY isolates. Due to the limited number of HC-pro sequences, additional PVY HC-pro sequences were included from the DPVweb database. Phylogenetic analyses were conducted using the neighbor-joining method assembled in Molecular Evolutionary Genetics Analysis software package (MEGA, version 4.0) (27) with the default parameters.

Selection pressure on CP and HC-pro. The ML codon substitution model implemented in CODEML program of Phylogenetic Analysis by Maximum Likelihood (PAML) package version 4.0 (66) was used to calculate the values of ω (the ratio between nonsynonymous and synonymous substitutions, d_N/d_S) and estimate the selection pressures on CP and HC-pro proteins. When $\omega > 1$, it implied that positive (or diversifying) selection existed, whereas $\omega < 1$ and $\omega = 1$ indicated negative (purifying) and neutral selection, respectively. To test diversifying selection at individual sites, different models (M0, M1a, M2a, M3, M7, and M8), which allow for heterogeneous ω ratios among sites, were compared. Likelihood ratio tests (LRTs) were used to examine whether allowing for sites with $\omega > 1$ significantly improves the fitting to data. M1a versus M2a and M7 versus M8 were compared with verify the positive selection hypothesis with degree of freedom (df) = 2. M0–M3 comparison was used to check the heterogeneity of selective constrains among sites with df = 4. Bayes empirical Bayes method, which accounts for sampling errors in the ML estimates of parameters, was used to calculate the posterior probabilities that each codon is from the site class of positive selection under models M2a and M8 (68). Selection pressures on subpopulations of PVY from different hosts (potato and tobacco) were also analyzed. The branch-site model was used to detect positive selection that affects only a few sites on prespecified lineages (67). In this analysis, the branches under analysis for positive selection are called foreground branches while other branches are background branches.

Sequences diversity and population demography analyses. DnaSP version 4.10 (47) was used to calculate Tajima's D test (56), Fu and Li's (15) D and F tests, haplotype diversity, and nucleotide diversity. Tajima's D and Fu and Li's D and F tests hypothesize that all mutations are selectively neutral. Tajima's D test depends on the differences between the numbers of segregating sites and the average number of nucleotide differences. Fu and Li's D test considers the differences between the number of singletons (mutations appearing only once among the sequences) and the total numbers of mutations, whereas Fu and Li's F test is based on the differences between the numbers of singletons and the average number of nucleotide differences among all pairs of sequences. Haplotype diversity refers to the frequency and number of haplotypes in the population. Nucleotide diversity estimates the average pairwise differences among sequences. The nucleotide diversities were calculated within and between groups. Mismatch distributions of all populations were estimated on all pairs of haplotypes present in a population.

RESULTS

Symptoms, serogroups, and sequence identities. In total, 25 isolates of PVY were obtained from tobacco plants sampled from five provinces of China and passaged through single lesions in *C. amaranticolor*. Of these 25 isolates, 22 caused systemic veinal necrosis symptoms in tobacco, whereas three isolates caused symptoms of mosaic, mottle, or vein banding (Table 1). In all, 3 isolates causing no necrotic symptoms as well as 16 isolates which induced veinal necrosis were detected with MAb1129 that is specific to PVY^O according to the information provided by the supplier. These isolates were not detected with the MAb1128 specific to PVY^N, whereas the remaining six isolates that caused veinal necrosis in tobacco were detected with MAb1128 but not MAb1129 (Table 1). Four additional isolates were collected directly from the fields without single-lesion isolation on *C. amaranticolor* or serological detection, and their CP-encoding sequences but not HC-pro-encoding sequences were determined.

The size of the 5′-proximal genomic region amplified from the 25 PVY isolates was 1,461 nt, of which the HC-pro-encoding sequence constituted 1,368 nt (sequence accessions AM236829 to AM236853) (Supplementary Table 1). The length of the 3′-proximal part of the PVY genome characterized from 29 isolates (including the 25 single-lesion isolates and 4 field samples) was 1,130 to 1,132 nt (sequence accessions AM236790 to AM236818) (Supplementary Table 1). The CP-encoding region was of the same size (804 nt) in all isolates but the 3′-untranslated region (3′-UTR) varied in length (328 nt in most isolates; 327 nt in isolates Anhui33, Feixian8, and Mengyin60; and 326 nt in isolate Henan8).

The HC-pro sequences were 81.6 to 99.7 and 89.4 to 100% identical at the nucleotide and amino acid levels, respectively, in the 25 PVY isolates characterized. The CP-encoding sequences of the 29 characterized isolates were somewhat more conserved, having identities of 87.2 to 100 and 91.3 to 100% at the nucleotide and amino acid levels, respectively.

Recombination within HC-pro- and CP-encoding sequences. Other than the 25 HC-pro-encoding sequences determined in this study, there were only 6 additional PVY HC-pro sequences available from China (Supplementary Table 1). In only one single isolate, Henan10, a recombination junction was detected within the HC-pro-encoding sequence (Table 2).

In addition to the 29 CP-encoding sequences determined in this study, CP sequences of 35 additional PVY isolates characterized from tobacco crops in China are available and included for analysis (Supplementary Table 1). Thirteen of these sequences (Table 2) showed clear evidence for recombination. According to the break-point position, these recombinants could be divided to four recombinant types: BJ, Taoxu, HXCH44, and Anhui (Table 2). In addition, phylogenetic analysis was carried out independently using two segments of the CP-encoding region (nucleotides 22 to 245 and 599 to 774). Results supported those obtained by recombination analysis and provided evidence for recombination between isolates of PVY^N and PVY^O in the aforementioned isolates (Table 1; data not shown).

Phylogenetic relationships based on HC-pro- and CP-encoding sequences. The HC-pro- and CP-encoding sequences of the isolates characterized in this study were subjected to

TABLE 1. Strain designations of *Potato virus Y* (PVY) strains characterized in this study

Isolate	Symptoms[a]	MAb[b]	HC-pro[c]	CP[d]			Designation
				22–783	22–245	599–774	
Anhui5	VB	O	O	O	O	O	O
Anhui33	VN	N	N	N	N	N	N
Anqiu1	VN	N	N	nd	N	O	NTN
Feixian8	VN	N	O	N	N	N	N-HcO
Gansu1	VN	O	N	O	O	O	N-Wi
Gansu5	VN	O	N	O	O	O	N-Wi
Gansu8	VN	O	N	O	O	O	N-Wi
Gansu10	VN	O	N	O	O	O	N-Wi
Gansu14	VN	O	N	O	O	O	N-Wi
Gansu15	VN	O	N	O	O	O	N-Wi
Heilongjiang11	VN	O	N	O	O	O	N-Wi
Heilongjiang18	VN	O	N	O	O	O	N-Wi
Heilongjiang20	VN	O	N	O	O	O	N-Wi
Henan10	M	O	nd	O	O	O	O
Henan43	VN	O	N	O	O	O	N-Wi
Henan6	VN	O	N	O	O	O	N-Wi
Henan8	VN	N	N	N	N	N	N
Mengyin55	M, Mo	O	O	O	O	O	O
Mengyin59	VN	O	N	O	O	O	N-Wi
Mengyin60	VN	N	O	N	N	N	N-HcO
Taoxu91	VN	N	N	nd	N	O	NTN
Zhucheng37	VN	O	N	O	O	O	N-Wi
Zhucheng39	VN	O	N	O	O	O	N-Wi
Zibo1	VN	O	N	O	O	O	N-Wi
Zibo3	VN	O	N	O	O	O	N-Wi

[a] Symptoms in *Nicotiana tabacum*. VB, vein banding; VN, veinal necrosis; M, mosaic; Mo, mottle.
[b] Detection with monoclonal antibody (MAb).
[c] Complete helper component proteinase (HC-pro)-encoding sequence was used for phylogenetic analyses using the neighbor-joining method, nd, not determined because the sequence contains a recombination junction.
[d] Different parts of the coat protein (CP)-encoding sequence were used for phylogenetic analysis using the neighbor-joining method; nd, not determined because the sequence contains a recombination junction.

TABLE 2. Recombination analysis on *Potato virus Y* (PVY) helper component proteinase (HC-pro)- and coat protein (CP)-encoding regions

Recombinant	Recombination region (nt)[a]	Parent-like isolates		Recombination detection	
		Major	Minor	Methods[b]	*P* value
HC-Pro					
Henan10	776–1308	O/Feixian8	O/Anhui5	B, M, **S**, 3Q	1.40×10^{-4}
CP					
BJ-type					
AFY1	?–384	O/U25672	N/AF126258	R, G, B, M, **S**, 3Q	6.40×10^{-11}
AFY4-1	?–384	O/U25672	N/AF126258	G, M, **S**, 3Q	6.40×10^{-11}
BJ	?–384	O/U25672	N/AF126258	R, G, M, **S**, 3Q	7.10×10^{-10}
BJ0-1	?–384	O/U25672	N/AF126258	R, G, M, **S**, 3Q	6.40×10^{-11}
BJ0-3	?–384	O/U25672	N/AF126258	R, G, M, **S**, 3Q	6.40×10^{-11}
BJ2-1	?–384	O/U25672	N/AF126258	R, G, M, **S**, 3Q	3.28×10^{-11}
HMDJ11	?–384	O/U25672	N/AF126258	R, G, M, **S**, 3Q	6.40×10^{-11}
Taoxu-type					
Anqiu1	599–774	N/AF126258	O/X68222	G, B, M, **S**, 3Q	1.29×10^{-7}
Taoxu91	599–774	N/AF126258	O/X68222	G, B, M, **S**, 3Q	1.29×10^{-7}
Yunan1	599–774	N/AF126258	O/X68222	G, B, M, **S**, 3Q	1.29×10^{-7}
HXCH44-type					
HXCH30	?–245	O/U25672	N/AF126258	R, G, B, M, S, **3Q**	5.82×10^{-13}
HXCH44	?–245	O/U25672	N/AF126258	R, G, B, M, S, **3Q**	1.05×10^{-14}
Anhui potato-type					
Anhui/potato	?–491	O/U25672	N/AF126258	R, G, M, C, S, **3Q**	5.53×10^{-13}

[a] Numbers refer to the nucleotides (nt) of the respective protein-encoding region of the isolate; ? indicates that one recombination junction lies at an unknown position outside the sequenced part of the viral genome.
[b] Methods which support recombinant events are indicated. R, RDP; G, GENECONV; B, BOOTSCAN; M, MAXIMUM CHI-SQUARE; C, CHIMAERA; S, SiSCAN; 3Q, 3seq. *P* value corresponds to the method marked in bold.

phylogenetic analysis with the corresponding sequences of additional PVY isolates from China (Supplementary Table 1). The recombinant sequences mentioned above were excluded.

In the phylogenetic tree of HC-pro sequences, isolates were clustered into two main groups (CHHCN and CHHCO) (Fig. 2A) supported by high bootstrap values. CHHCN contained 22 isolates, most of which were isolates characterized in this study, and was divided to two subgroups, CHHCN1 and CHHCN2. The CHHCO group contained eight isolates.

Phylogenetic analysis of the CP sequences revealed two main groups designated as CPO and CPN. CPO contained 51 isolates, including 24 characterized in this study. CPN contained 19 isolates, which could be further divided into three subgroups (CPN1, CPN2, and CPN3) supported by high bootstrap values (Fig. 2B).

Subsequently, the PVY isolates from China were compared with isolates from other countries (Supplementary Table 1). In the phylogenetic tree based on HC-pro sequences, the isolates of PVY were placed in three main groups corresponding to N, O, and C&NP (non-potato) (Fig. 3A). The group N consisted of 66 isolates, and could be further divided into two subgroups, I and II (Fig. 3A). Subgroup I contained 48 isolates, including 14 isolates from China. Many isolates in this group were reported to belong to the strain PVY^{N-Wi} in previous studies (Supplementary Table 1). Subgroup II contained 18 isolates, including 8 isolates characterized in this study (Fig. 3A). Group O of HC-pro contained 15 isolates, including 8 isolates from China. Four of the Chinese isolates were characterized in this study: Anhui5 and Mengyin55 that caused non-necrotic symptoms in tobacco and were detected with the MAb for PVY^{O}, and Feixian8 and Mengyin60 that caused veinal necrosis in tobacco leaves and reacted with the MAb for PVY^{N} (Fig. 3A; Table 1). The third group, C&NP, contained 14 isolates which had been characterized as belonging

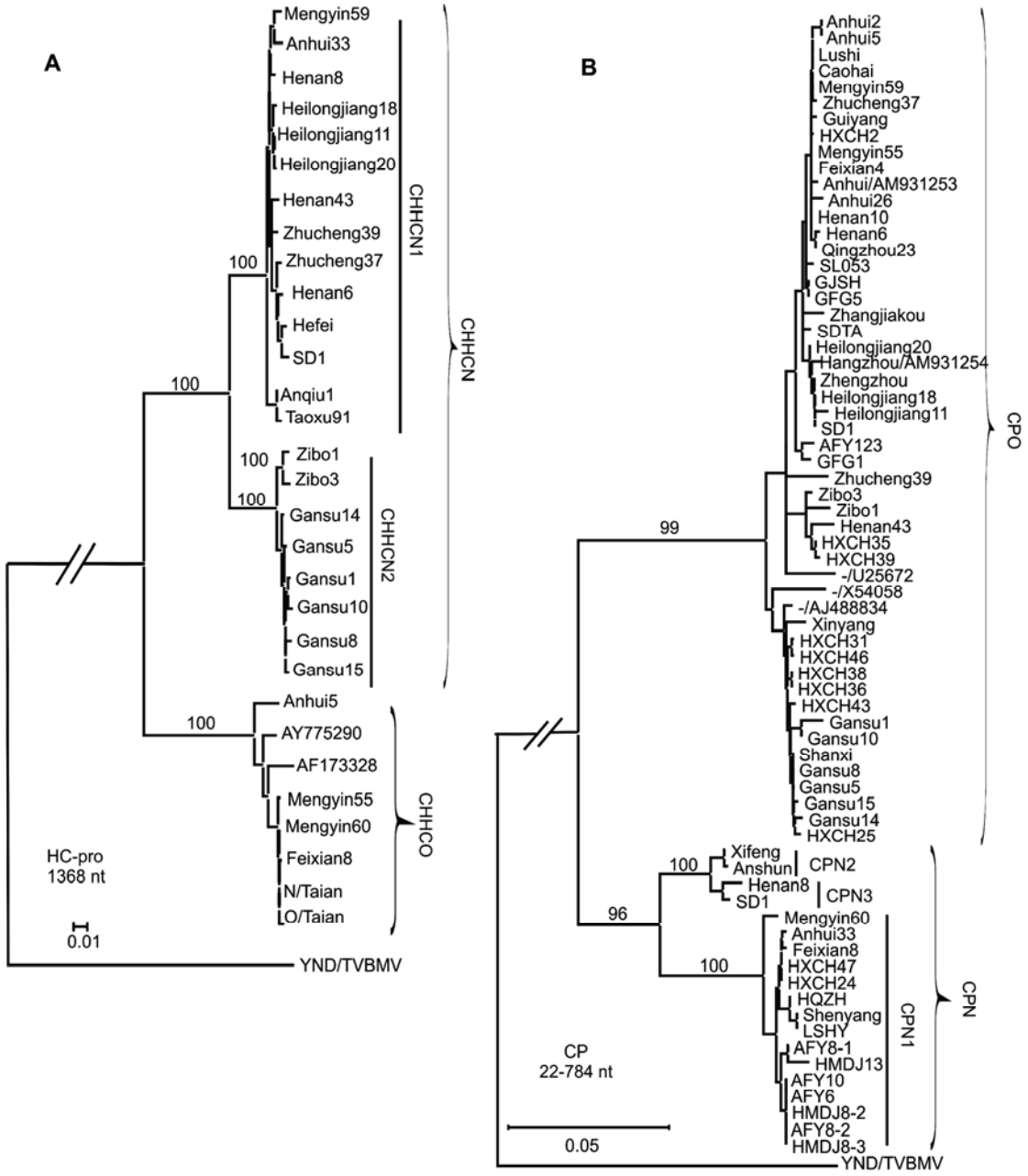

Fig. 2. Neighbor-joining trees based on the helper component proteinase (HC-pro)- and coat protein (CP)-encoding sequences of Chinese *Potato virus Y* (PVY) isolates. All available Chinese PVY isolates were included except the recombinant sequences and sequences which contain degenerate nucleotides. **A,** Tree based on HC-pro-encoding sequences (n = 30). **B,** Tree based on CP-encoding sequences (n = 70). Bars represented Kimura nucleotide units; // indicates not drawn to scale. Corresponding sequences of *Tobacco vein banding mosaic virus* (TVBMV; accession no. EF219408) were used as outgroup.

to the strain group PVY^{C} or being of non-potato (NP) origin (Supplementary Table 1). No isolate from China was found in this cluster.

The HC-pro sequences of 20 PVY isolates which caused veinal necrosis symptoms in tobacco leaves in this study (Table 1) contained the amino acid residues K-400 and E-419 (60) and D-205 (25), as expected based on previous studies. However, isolates Feixian8 and Mengyin60, whose HC-pro sequences were placed to group O, contained residues D-205, R-400, and D-419 and also induced veinal necrosis in tobacco leaves (Table 1).

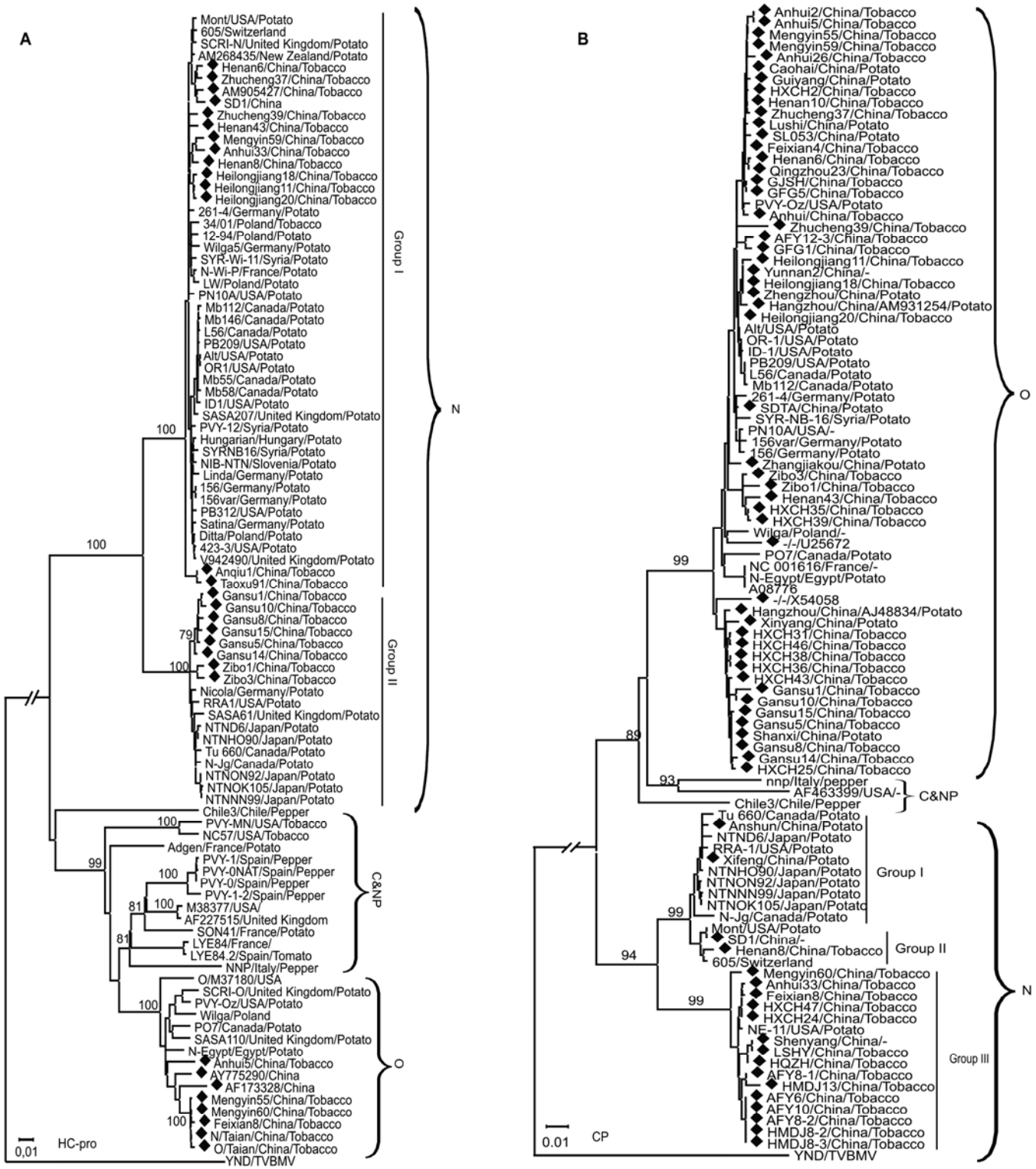

Fig. 3. Neighbor-joining trees based on the helper component proteinase (HC-pro)- and coat protein (CP)-encoding sequences of the isolates characterized from China and selected isolates from other countries. **A,** Tree based on 96 complete HC-pro-encoding sequences, including all the HC-pro-encoding sequences which were available in the databases and which were not recombinants or did not contain degenerate nucleotides. **B,** Tree based on 101 CP-encoding sequences. Nucleotides 22 to 783 of the CP-encoding sequence were used for analysis from all *Potato virus Y* (PVY) isolates available from China and, additionally, from 30 PVY isolates whose complete genomic sequences were available. No recombinant sequences or sequences containing degenerate nucleotides were included. Chinese isolates are marked with ◆; // indicates not drawn to scale. Corresponding sequences of *Tobacco vein banding mosaic virus* (TVBMV; accession no. EF219408) were used as outgroup.

Analysis of CP-encoding sequences placed these PVY isolates in three clusters (Fig. 3B). Cluster O contained 68 isolates, including all isolates detected with the MAb for PVY^O in this study. In all, 16 isolates from China (two from potato, one of unknown origin, and 13 from tobacco) seemed to form a subcluster in cluster O but the subcluster was not supported with a high bootstrap value. Cluster N contained 30 isolates, including four PVY isolates characterized in this study (Anhui33, Feixian8, Henan8, and Mengyin60), which correspond to PVY^N for all the properties analyzed. It could be further divided into three subgroups, I to III (Fig. 3B). Isolates Xifeng and Anshun from China in subgroup I showed at least 98.5% nucleotide identities to isolates which have been reported as "North American type" (NA-NTN) isolates of PVY^{NTN} (42,54). Subgroup II contained two isolates from China (SD1 and Henan8) and also isolates Mont and N605 that can be considered as type isolates of PVY^N (>99.2% nucleotide identity among the four isolates). In subgroup III, all isolates except NE-11 were from China and shared nucleotide identities of >98.1%. This subgroup contained most of the Chinese veinal necrosis isolates and was supported with a high bootstrap values. The third cluster, designated as C&NP, was small and contained only three isolates, of which none was from China.

Direction of the selection and sites under positive selection. In the site model, M0 model yielded estimated ω values of 0.170 and 0.090 for CP- and HC-pro-encoding sequences, respectively (Tables 3 and 4), indicating strong purifying selection on the evolution of these two proteins. Analysis of the selection constraints on different codons in CP- and HC-pro-encoding sequences using LRTs of M0-M3, M1a-M2a, and M7-M8 indicated that M3 was a significantly better fit to CP and HC-pro data than M0, and M8 was a significantly better fit to CP data than M7. In the CP-encoding region of all 70 PVY isolates characterized from China, 2.3% of the sites (five amino acid sites) were under strong positive selection with ω = 3.176 (Table 3). When the 52 and 13 CP sequences from known hosts (tobacco and potato, respectively) were analyzed separately, most amino acid sites in the CP were under strong purifying selection regardless of the host. In the PVY isolates from tobacco, eight amino acid sites of the CP were under positive selection (Table 3) whereas, in the potato isolates, three sites were under positive selection.

In the HC-pro-encoding region, only one site (amino acid 71) was under positive selection (ω = 2.108) when all 30 characterized isolates were included. Analysis of the HC-pro sequences from 27 tobacco isolates using the M8 model showed also that amino acid site 71 was under strong positive selection (ω = 3.117) (Table 4).

The branches with high bootstrap values in the phylogenetic trees of Chinese PVY HC-pro and CP were tested with a branch-site model (Table 4) in an attempt to detect positive selection that

TABLE 3. Selection pressures on different genomic regions of *Potato virus Y* (PVY): site models[a]

Sequence	*n*	ω	M3 vs. M0	M2a vs. M1a	M8 vs. M7	Parameter estimates under M8 (β & ω)	Positively selected sites[b]
CP							
All	70	0.170	81.300**	5.630	10.843*	p_1 = 0.023, ω = 3.176, p_0 = 0.977, β (0.242, 1.564)	10, 11, 16, **24**, **26**, 161
Potato	13	0.093	8.041	0	0.142	p_1 = 0.070, ω = 1.000, p_0 = 0.930, β (0.012, 0.204)	*11*, 16, 138
Tobacco	52	0.146	55.007**	6.240*	7.965*	p_1 = 0.004, ω = 8.321, p_0 = 0.996, β (0.133, 0.781)	11, 16, 17, 24, 26, 115, 127, **161**
HC-pro							
All	30	0.090	29.484**	1.740	3.166	p_1 = 0.004, ω = 2.108, p_0 = 0.996, β (0.519, 5.148)	16, 27, 71, 91, *99*, *163*, 346
Tobacco	27	0.081	36.824**	0	4.653	p_1 = 0.003, ω = 3.117, p_0 = 0.997, β (0.444, 5.168)	16, 71, **163**, 238

[a] Symbols: ω, d_N/d_S; p_1, proportion of sites under positive selection; p_0, proportion of sites under negative selection. Parameters p and q for the β distribution β (p, q) were given under model M8. For the coat protein (CP), the first 21 and the last 18 nucleotides were omitted; *, **, and ***: significant at 5, 1, and 0.1% risk levels, respectively.

[b] Sites potentially under positive selection identified under model M8 were numbered according to the complete helper component proteinase (HC-pro) and CP amino acid sequences. Probabilities: >0.9, underlined; 0.8 to 0.9, bold; 0.7 to 0.8, italics; 0.5 to 0.7 in plain text.

TABLE 4. Selection pressures on different genomic regions of *Potato virus Y* (PVY): branch-site models[a]

Branch	*n*	A1 vs. A		Parameters[b]	Positively selected sites[c]
CP					
CPN	19	A1	0	$\omega_{0-B} = \omega_{0-F}$ = 0.055 (P = 0.874), $\omega_{1-B} = \omega_{1-F}$ = 1.000 (P = 0.126)	Not allowed
		A		$\omega_{0-B} = \omega_{0-F}$ = 0.055 (P = 0.874), $\omega_{1-B} = \omega_{1-F}$ = 1.000 (P = 0.126)	None
CPN1	15	A1	0.113	$\omega_{0-B} = \omega_{0-F}$ = 0.046 (P = 0.823), $\omega_{1-B} = \omega_{1-F}$ = 1.000 (P = 0.114)	
				ω_{2a-B} = 0.046, ω_{2a-F} = 1.000 (P = 0.055), $\omega_{2b-B} = \omega_{2b-F}$ = 1.000 (P = 0.008)	Not allowed
		A		$\omega_{0-B} = \omega_{0-F}$ = 0.047 (P = 0.842), $\omega_{1-B} = \omega_{1-F}$ = 1.000 (P = 0.116)	
				ω_{2a-B} = 0.047, ω_{2a-B} = 1.498 (P = 0.037), ω_{2b-F} = 1.000, ω_{2b-F} = 1.498 (P = 0.005)	*25*, **116**
CPO	51	A1	7.699**	$\omega_{0-B} = \omega_{0-F}$ = 0.045 (P = 0.874), $\omega_{1-B} = \omega_{1-F}$ = 1.000 (0.084)	
				ω_{2a-B} = 0.045, ω_{2a-B} = 1.000 (P = 0.078), $\omega_{2b-F} = \omega_{2b-F}$ = 1.000 (P = 0.008)	Not allowed
		A		$\omega_{0-B} = \omega_{0-F}$ = 0.056 (P = 0.874), $\omega_{1-B} = \omega_{1-F}$ = 1.000 (P = 0.093d)	
				ω_{2a-B} = 0.056, ω_{2a-B} = 3.470 (P = 0.031), ω_{2b-F} = 1.000, ω_{2b-F} = 3.470 (P = 0.003)	9, *11*, 15, 16, 32, 161
HC-pro					
CHHCN1	14	A1	0	$\omega_{0-B} = \omega_{0-F}$ = 0.051 (P = 0.854), $\omega_{1-B} = \omega_{1-F}$ = 1.000 (P = 0.030)	
				ω_{2a-B} = 0.051, ω_{2a-B} = 1.000 (P = 0.113), $\omega_{2b-F} = \omega_{2b-F}$ = 1.000 (P = 0.004)	Not allowed
		A		$\omega_{0-B} = \omega_{0-F}$ = 0.051 (P = 0.854), $\omega_{1-B} = \omega_{1-F}$ = 1.000 (P = 0.030)	
				$\omega_{2a-B} = \omega_{2a-B}$ = 1.000 (P = 0.113), $\omega_{2b-F} = \omega_{2b-F}$ = 1.000 (P = 0.004)	131, 153, 164, 178, 185, 319, 414
CHHCN2	8	A1	–1.734	$\omega_{0-B} = \omega_{0-F}$ = 0.069 (P = 0.965), $\omega_{1-B} = \omega_{1-F}$ = 1.000 (P = 0.035)	Not allowed
		A		$\omega_{0-B} = \omega_{0-F}$ = 0.069 (P = 0.965), $\omega_{1-B} = \omega_{1-F}$ = 1.000 (P = 0.035)	None
CHHCO	8	A1	0.427	$\omega_{0-B} = \omega_{0-F}$ = 0.068 (P = 0.961), $\omega_{1-B} = \omega_{1-F}$ = 1.000 (P = 0.028)	
				ω_{2a-B} = 0.068, ω_{2a-B} = 1.000 (P = 0.011), $\omega_{2b-F} = \omega_{2b-F}$ = 1.000 (P = 0.000)	Not allowed
		A		$\omega_{0-B} = \omega_{0-F}$ = 0.069 (P = 0.965) $\omega_{1-B} = \omega_{1-F}$ = 1.000 (P = 0.027)	
				ω_{2a-B} = 0.069, ω_{2a-B} = 1.693 (P = 0.008), ω_{2b-F} = 1.000, ω_{2b-F} = 1.693 (P = 0.000)	**91**, 238, 252, 295, **346**, 358

[a] For the coat protein (CP), the first 21 and the last 18 nucleotides were omitted; *, **, and ***: significant at 5, 1, and 0.1% risk levels, respectively.

[b] The ω of different site class (0, 1, 2a, and 2b). B, background; F, foreground.

[c] Sites potentially under positive selection identified under model M8 were numbered according to the complete helper component proteinase (HC-pro) and CP amino acid sequences. Probabilities: >0.9, underlined; 0.8 to 0.9, bold; 0.7 to 0.8, italics; 0.5 to 0.7 in plain text.

affects only a few sites among a few lineages (Fig. 2). In branch CPN (n = 19), 87.4% of the sites were under strong purifying selection (ω = 0.055). However, subbranch CPN1 (n = 15), which was located within branch CPN and had four isolates less, contained two sites under positive selection. In branch CPO (51 isolates), model A was a significantly better fit to the data than null model A1. Six sites (3.1%) were under strong positive selection in this branch (ω = 3.470) (Table 4).

In the phylogenetic tree constructed using HC-pro-encoding sequences of PVY from China, three branches were defined (Fig. 2A). Seven sites were found under positive selection in the branch CHHCN1. In branch CHHCO (eight isolates), three sites were under weak positive selection (ω = 1.693) (Table 4).

Sequence diversity and population demography. The PVY isolates from China whose CP-encoding sequence were subjected to phylogenetic analysis in Figure 2B were divided into two groups based on their original host, potato (13 isolates) and tobacco (52 isolates). The nucleotide diversity of the tobacco isolates (0.0640 ± 0.0062) was higher than that of the potato isolates (0.0407 ± 0.0141) (Table 5). According to the phylogenetic tree (Fig. 2B), 37 tobacco isolates were clustered to strain group O and 15 isolates to strain group N. The nucleotide diversity between the tobacco O subgroup and tobacco N subgroup gave a high value (0.1228 ± 0.0106); it also showed that the tobacco N subgroup (0.0291 ± 0.0146) is more variable than tobacco O subgroup (0.0162 ± 0.0014). Comparison of the nucleotide diversity showed that the difference between tobacco O and potato O subgroups was not significant (0.0145 ± 0.0017).

Nucleotide diversity among groups was also determined regardless of the host factor. The isolates were grouped based on the results of Figure 2 supported by high bootstrap values. However, CPN2 and CPN3 were not included because numbers of isolates were insufficient. The nucleotide diversity within CPN (0.0260 ± 0.0062) was higher than that with CPO (0.0164 ± 0.0013) (Table 6). Nucleotide diversity within CHHCN (0.0474 ± 0.0042) was higher than that within CHHCO (0.0218 ± 0.0062). Although both CHHCN1 and CHHCN2 belong to the CHHCN group, there was an obvious difference between the two groups (0.0817 ± 0.0105). Similar to the result above, there was a significant difference between groups CHHCN and CHHCO (0.2027 ± 0.0207) (Table 6). Most of the groups gave negative values for the Tajima's D and Fu and Li's D and F tests, which suggested that the population size was in a state of expansion; however, for CPO only, the values were statistically significant (Table 7).

Comparison between the groups CPN and CPO showed that CPN had lower haplotype diversity (h) but higher nucleotide diversity (π) than CPO; however, comparing with CPO, the subgroup CPN1 of CPN had smaller h and π, which could be considered as evidence of recent population bottleneck or founder event (19) (Table 7).

Mismatch distribution was also analyzed on all the groups. The CPN1 group had a star-like distribution while other groups had multiple peaks (data not shown), which indicated that subpopulation CPN1 was emergent while other subpopulations were long-existing ones.

DISCUSSION

This study analyzed the genetic diversity of PVY isolates infecting tobacco crops, which has gained little attention. Collection of the PVY isolates from five major tobacco-growing provinces in China and production of single-lesion isolates on *C. amaran-*

TABLE 5. Neutrality tests and haplotype and nucleotide diversity analyses within and between subpopulations of *Potato virus Y* isolates characterized from China: coat protein–host[a]

Group	Tobacco (all) (n = 52)	Tobacco O (n = 37)	Tobacco N (n = 15)	Potato (all) (n = 13)	Potato O (n = 11)
Tobacco (all)	0.0640 ± 0.0062	–	–	–	–
Tobacco O	–	0.0162 ± 0.0014	–	–	–
Tobacco N	–	0.1228 ± 0.0106	0.0291 ± 0.0146	–	–
Potato (all)	0.0518 ± 0.0555	–	–	0.0407 ± 0.0141	–
Potato O	–	0.0145 ± 0.0017	–	–	0.0126 ± 0.0022

[a] Symbols: n, number of isolates within the population; –, not determined.

TABLE 6. Neutrality tests and haplotype and nucleotide diversity analyses within and between subpopulations of *Potato virus Y* isolates characterized from China: coat protein (CP) and helper component proteinase (HC-pro) subpopulations[a]

	CP			HC-pro			
Subgroup	CPN (n = 19)	CPN1 (n = 15)	CPO (n = 51)	CHHCN (n = 22)	CHHCN1 (n = 14)	CHHCN2 (n = 8)	CHHCO (n = 8)
CPN	0.0260 ± 0.0062	…	…	…	…	…	…
CPN1	–	0.0069 ± 0.0014	…	…	…	…	…
CPO	0.1284 ± 0.0060	0.1320 ± 0.0067	0.0164 ± 0.0013	…	…	…	…
CHHCN	...	…	…	0.0474 ± 0.0042	…	…	…
CHHCN1	…	…	…	–	0.0159 ± 0.0014	…	…
CHHCN2	…	…	…	–	0.0817 ± 0.0105	0.0123 ± 0.0023	…
CHHCO	…	…	…	0.2027 ± 0.0207	0.2004 ± 0.0255	0.2066 ± 0.0342	0.0218 ± 0.0062

[a] Symbols: n, number of isolates within the population; –, not determined.

TABLE 7. Neutrality tests and haplotype and nucleotide diversity analyses within subpopulations of *Potato virus Y* isolates characterized from China[a]

Population	n	Tajima's D	Fu and Li's D	Fu and Li's F	h	π
CPN	19	–0.0654	0.1762	0.0815	0.930 ± 0.047	0.0251 ± 0.0060
CPN1	15	–1.0963	–1.1302	–1.2894	0.886 ± 0.069	0.0068 ± 0.0014
CPO	51	–1.4764	–3.1638*	–3.0828*	0.976 ± 0.012	0.0162 ± 0.0013
CHHCN	22	0.2994	–0.5258	–0.4017	1.000 ± 0.014	0.0452 ± 0.0039
CHHCN1	14	–1.2886	–1.4067	–1.6199	1.000 ± 0.027	0.0157 ± 0.0013
CHHCN2	8	–0.9915	–0.9358	–1.0065	1.000 ± 0.063	0.0122 ± 0.0020
CHHCO	8	–1.4750	–1.3655	–1.5236	1.000 ± 0.063	0.0214 ± 0.0060

[a] Symbols: n, number of isolates within the population; h, haplotype diversity; π, nucleotide diversity. *, $P < 0.05$.

ticolor enabled reliable biological, serological, and molecular characterization of the isolates. Some PVY^{N} and PVY^{NTN} isolates do not cause local lesions on *C. amaranticolor* (3,30,38), and isolates similar to them may have been missed in the process of producing single-lesion isolates of PVY. Most isolates (88%) characterized in this research caused systemic veinal necrosis symptoms in tobacco. Sixteen of these isolates were detected with an MAb for PVY^{O}, and the CP-encoding sequences of these isolates were related to PVY^{O} according to phylogenetic analysis; however, they induced systemic veinal necrosis in tobacco cv. Samsun, indicating that they belonged to the strain group PVY^{N} (11,54). Their HC-pro-encoding sequences were phylogenetically related to isolates of PVY^{N}, which suggested that they were recombinants. Isolates with a similar combination of properties have been detected previously in potato and designed to $PVY^{N\text{-}Wi}$, $PVY^{N:O}$, or PVY-HN2 in previous studies (9,16,23,54).

Two isolates, Feixian8 and Mengyin60, were unique in that they induced veinal necrosis in tobacco leaves but contained PVY^{O}-like HC-pro and PVY^{N}-like CP. Because their complete genome sequences remain to be determined, the genomic composition of these isolates (tentatively designated as $PVY^{N\text{-}HcO}$) cannot the ascertained at this time. The data suggest that their genomic structure might be different from the previously described veinal-necrosis-inducing PVY recombinants (24). The amino acid residues D-205, K-400, and E-419 of HC-pro have been reported as determinants of the tobacco veinal necrosis phenotype (8,25,60). The two $PVY^{N\text{-}HcO}$ isolates detected in this study contained amino acid residues D-205, R-400, and D-419 in HC-pro but also caused veinal necrosis in tobacco leaves. This result is consistent with that of Hu et al. (25) but not with that of Tribodet et al. (60). However, PVY isolate SASA110, which contains these three amino acid residues in the corresponding positions of HC-pro, induces mosaic symptom in plants of *N. tabacum* (3). Therefore, other HC-pro amino acids or other PVY proteins regulating necrosis induction in tobacco must also exist (8,25,48).

Thirteen Chinese PVY isolates analyzed in this study contained a predictable recombination junction in the CP-encoding region. They were detected in all provinces studied and could be placed in four groups according to the different predicted positions of recombination junctions. Isolates Anqiu1 and Taoxu91 were similar to recombinant isolates of the PVY^{NTN} strain described from elsewhere (4,8,54). The only valid criterion to designate a PVY isolate to strain PVY^{NTN} is the tuber necrosis symptoms induced on potato tubers, which vary according to potato cultivar and environmental conditions (54). Some isolates that cause mosaic symptoms in tobacco can induce potato tuber necrosis (8,25). Reliable, reproducible, and sensitive biological assays for producing the tuber necrosis symptoms with PVY^{NTN} isolates appear to be urgent to establish. Unfortunately, the goal is challenging to achieve, as shown by the ring tests carried out by eight European laboratories (5).

The HC-pro- and CP-encoding regions in isolates analyzed in this study were mainly under purifying selection but seven amino acids in HC-pro and six amino acids in CP were found to be under positive selection. Except for the amino acid at position 161, the other five positively selected residues in CP were located at the CP N terminus, which is predicted to be exposed on the virion surface (2,49). Diversification of these codons might enhance interactions of CP with other viral or host proteins and, thereby, increase the fitness of this multifunctional protein. The amino acids of HC-pro that were under positive selection were located in domain 1, which is responsible for aphid transmission, RNA binding, and genome amplification (45). Diversification of this region might increase adaptation of PVY to aphid transmission and also the host. Furthermore, this study showed that the selection pressure on the isolates belonging to the phylogenetic clusters CHHCO (HC-pro-encoding region) and CPO (CP) was higher than that on the clusters CHHCN (HC-pro) and CPN (CP), indicating that selection pressures on the same protein might differ between PVY subpopulations. Previously, positive selection has been reported on the CP and 6K2 protein (40) and the P1 protein (42) of PVY.

PVY isolates can spread frequently between potato and tobacco crops (54), which is consistent with no host-specific clustering of the PVY isolates observed in the phylogenetic and nucleotide diversity analyses of this study. Because PVY is not seedborne whereas tobacco is propagated via seed, infections of tobacco crops with PVY occur via transmission of the virus from other hosts by aphids during the growing season. Potato can, indeed, function as an efficient source of PVY inoculum for tobacco crops because the incidence of PVY in the nearby potato crop correlates positively with the final disease incidence, rate of disease progress, and the magnitude of radial dispersion of PVY into the tobacco crops (20). Other inoculum sources of PVY for tobacco crops can be tomato, pepper, and *Physalis* spp. (28). The frequent spread of PVY isolates between other hosts and tobacco may be one of the driving forces of PVY evolution, formation of viral subpopulations, and generation of novel recombinants detected by analysis of PVY isolates from tobacco crops in this study.

ACKNOWLEDGMENTS

Y. P. Tian, J. L. Liu, and C. L. Zhang contributed equally to this research. This study was supported by the Program for New Century Excellent Talents in University (NCET-07-0520), Natural Scientific Foundations of China (NSFC, 31011130031, 30471138, 30971895) and Shandong province (Z2007D04), Specialized Research Fund for the Doctoral Program of Higher Education (SRFDP, 20080434006), Swedish-Asian Research Link (348-2003-5011), Shandong Tobacco Company, Center of International Mobility (CIMO, Finland), and the Academy of Finland (grants 1118766 and 1134759).

LITERATURE CITED

1. Adams, M. J., and Antoniw, J. F. 2006. DPVweb: A comprehensive database of plant and fungal virus genes and genomes. Nucleic Acids Res. 34:D382-D385.
2. Baratova, L. A., Efimov, A. V., Dobrov, E. N., Fedorova, N. V., Hunt, B., Badun, G. A., Ksenofontov, A. L., Torrance, L., and Järvekülg, L. 2001. In situ spatial organization of potato virus A coat protein subunits as assessed by tritium bombardment. J. Virol. 75:9696-9702.
3. Barker, H., McGeachy, K. D., Toplak, N., Gruden, K., Zel, J., and Browning, I. 2009. Comparison of genome sequence of PVY isolates with biological properties. Am. J. Pot. Res. 86:227-238.
4. Boonham, N., Walsh, K., Hims, M., Preston, S., North, J., and Barker, I. 2002. Biological and sequence comparisons of *Potato virus Y* isolates associated with potato tuber necrotic ringspot disease. Plant Pathol. 51:117-126.
5. Browning, L., Charlet, K., Chrzanowska, M., Dedic, P., Kerlan, C., Kryszczuk, A., Schubert, J., Varveri, C., Werkman, A., and Wolf, I. 2004. What is PVY^{NTN}? The reaction of potato cultivars to inoculation with a range of PVY isolates. Pages 48-50 in: Proc. 12th EAPR Virol. Sect. Meet. Rennes, France.
6. Chare, E. R., and Holmes, E. C. 2006. A phylogenetic survey of recombination frequency in plant RNA viruses. Arch. Virol. 151:933-946.
7. Chikh Ali, M., Maoka, T., and Natsuaki, K. T. 2007. A point mutation changes the serotype of a *Potato virus Y* isolate; genomic determination of the serotype of PVY strains. Virus Genes 35:359-367.
8. Chikh Ali, M., Maoka, T., and Natsuaki, K. T. 2008. Whole genome sequence and characterization of a novel isolate of PVY inducing tuber necrotic ringspot in potato and leaf mosaic in tobacco. J. Phytopathol. 155:409-415.
9. Chrzanowska, M. 1991. New isolates of the necrotic strain of potato virus Y (PVY^{N}) found recently in Poland. Potato Res. 34:179-182.
10. Chung, B. Y. W., Miller, W. A., Atkins, J. F., and Firth, A. E. 2008. An overlapping essential gene in the *Potyviridae*. Proc. Natl. Acad. Sci. USA 105:5897-5902.
11. De Bokx, J. A. 1961. Waardplanten van het aardappel-YN-virus [Host plants of the potato virus YN (Tobacco veinal necrosis virus)]. Eur. J. Plant Pathol. 67:273-277.

12. Fan, Z., Chen, H., Cai, S., Deng, C., Wang, W., Liang, X., and Li, H. 2003. Molecular characterization of a distinct potyvirus from whitegrass in China. Arch. Virol. 148:1219-1224.
13. Fauquet, C. M., Mayo, M. A., Maniloff, J., Desselberger, U., and Ball, L. A., eds. 2005. Virus Taxonomy. Eighth Report of the International Committee on Taxonomy of Viruses. Elsevier Academic Press, San Diego, CA.
14. Fellers, J. P., Tremblay, D., Handest, M. F., and Lommel, S. A. 2002. The *Potato virus Y* M^SN^R NIb-replicase is the elicitor of a veinal necrosis-hypersensitive response in root knot nematode resistant tobacco. Mol. Plant Pathol. 3:145-152.
15. Fu, Y. X., and Li, W. H. 1993. Statistical tests of neutrality of mutations. Genetics 133:693-709.
16. Glais, L., Tribodet, M., and 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.
17. Gooding, G. V. 1985. Relationship between strains of potato virus Y and breeding for resistance, cross protection, and interference. Tobacco Sci. 29:99-104.
18. Gooding, G. V., and Tolin, S. A. 1973. Strains of potato virus Y affecting flue-cured tobacco in the southeastern United States. Plant Dis. Rep. 57:200-204.
19. Grant, W. S., and Bowen, B. W. 1998. Shallow population histories in deep evolutionary lineages of marine fishes: insights from sardines and anchovies and lessons for conservation. J. Hered. 89:415-426.
20. Gray, S. M., and Lampert, E. P. 1988. Relationship between inoculum density and vector phenology on the incidence of potato virus Y in tobacco. Ann. Appl. Biol. 112:303-312.
21. Guo, X. Q., Zhu, H. C., Lu, S. E., Wang, Z. F., and Chen, R. T. 1998. Studies on tobacco PVY detection using indirect ELISA. Acta Tabacaria Sin. 4:38-42. (In Chinese)
22. Han, X. D., Chen, R. T., Jiang, R. J., Shi, J. K., Zhang, W. L., Chen, Z. Y., Zheng, Q. X., Shen, J. Y., and Zhu, B. M.. 1983. The tobacco vein necrosis disease caused by multiple infections with CMV and PVY. Acta Biochim. Biophys. Sin. 15:279-282. (In Chinese)
23. Hu, X., He, C., Xiao, Y., Xiong X., and Nie, X. 2009. Molecular characterization and detection of recombinant isolates of *Potato virus Y* from China. Arch. Virol. 154:1303-1312.
24. Hu, X., Karasev, A. V., Brown, C. J., and Lorenzen, J. H. 2009. Sequence characteristics of *Potato virus Y* recombinants. J. Gen. Virol. 90:3033-3041.
25. Hu, X., Meacham, T., Ewing, L., Gray, S. M., and Karasev, A. V. 2009. A novel recombinant strain of *Potato virus Y* suggests a new viral genetic determinant of vein necrosis in tobacco. Virus Res. 143:68-76.
26. Karasev, A. V., Nikolaeva, O. V., Hu, X., Sielaff, Z., Whitworth, J., Lorenzen, J. H., and Gray S. M. 2010. Serological properties of ordinary and necrotic isolates of *Potato virus Y*: A case study of PVY^N misidentification. Am. J. Pot. Res. 87:1-9.
27. Kumar, S., Tamura, K., and Nei, M. 2004. MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief. Bioinf. 5:150-163.
28. Lapp, N. A., and Gooding, G. V. 1976. Occurrence and sources of inoculums of potato virus Y in tobacco in North Carolina. Plant Dis. Rep. 60:1014-1016.
29. Latorre, B. A., Peñaloza, E., and Escaffi, O. 1982. A severe outbreak of potato virus Y in Chilean tobacco. Plant Dis. 66:893-895.
30. Le Romancer, M., Kerlan, C., and Nedellec, M. 1994. Biological characterization of various geographical isolates of potato virus Y inducing superficial necrosis on potato tubers. Plant Pathol. 43:138-144.
31. Li, N., Wang, X., Zhou, G., and Dong, J. 2006. Molecular variability of the coat protein gene of *Potato virus Y* from tobacco in China. Acta Virol. 50:107-113.
32. Li, X.-D., Li, Y.-Q., Gao, B.-C., Geng, J.-G., Wang, H.-G., Wang, W.-C., and Zheng, Q.-L. 2001. An investigation of the epidemics of tobacco viral disease in Henan province and the strategy of its control. Acta Tabacia Sin. 7:23-26. (In Chinese)
33. Ling, K. S., Zhu, H. Y., Petrovic, N., and Gonsalves, D. 2007. Serological detection of *Grapevine leafroll virus* 2 using an antiserum developed against the recombinant coat protein. J. Phytopathol. 155:65-69.
34. Liu, Y. Y., Yu, S. L., Lan, Y. F., Zhang, C. L., Hou, S. S., Li, X.-D., Chen, X. Z., and Zhu, X. P. 2009. Molecular variability of five *Cucumber mosaic virus* isolates from China. Acta Virol. 53:89-97.
35. Lockhart, B. E. L., and Fischer, H. U. 1976. A disease of tobacco in morocco caused by a veinal necrosis isolate of Potato virus Y. Plant Dis. Rep. 60:114-116.
36. Lorenzen, J. H., Meacham, T., Berger, P. H., Shiel, P. J., Crosslin, J. M., Hamm, P. B., and Kopp, H. 2006. Whole genome characterization of *Potato virus Y* isolates collected in the western USA and their comparison to isolates from Europe and Canada. Arch. Virol. 151:1055-1074.
37. Martin, D. P., Posada, D., and William, C. 2005. RDP2: Recombination detection and analysis from sequence alignments. Bioinformatics 21:260-262.
38. McDonald, J. G., and Kristjansson, G. T. 1993. New diseases and epidemics properties of strains of potato virus YN in North America. Plant Dis. 77:87-89.
39. Mih, A. M., Atiri, G. I., and Rossel, H. W. 1995. Strain-typing of potato virus Y isolates from potato in Nigeria by infectivity tests and ELISA. Afr. Crop Sci. J. 3:99-104.
40. Moury, B., Morel, C., Johansen, E., and Jacquemond, M. 2002. Evidence for diversifying selection in *Potato virus Y* and in the coat protein of other potyviruses. J. Gen. Virol. 83:2563-2573.
41. Noguchi, S., Tajima, T., Yamamoto, Y., Ohno, T., and Kubo, T. 1999. Deletion of a large genomic segment in tobacco varieties that are resistant to potato virus Y (PVY). Mol. Gen. Genet. 262:822-829.
42. Ogawa, T., Tomitaka, Y., Nakagawa, A., and Ohshima, K. 2008. Genetic structure of a population of *Potato virus Y* inducing potato tuber necrotic ringspot disease in Japan; comparison with North American and European populations. Virus Res. 131:199-212.
43. Ohshima, K., Yamaguchi, Y., Hirota, R., Hamamoto, T., Tomimura, K., Tan, Z., Sano, T., Azuhata, F., Walsh, J. A., Fletcher, J., Chen, J., Gera, A., and Gibbs, A. 2002. Molecular evolution of *Turnip mosaic virus*: evidence of host adaptation, genetic recombination and geographical spread. J. Gen. Virol. 83:1511-1521.
44. Piche, L. M., Singh, R. P., Nie, X., and Gudmestad, N. C. 2004. Diversity among *Potato virus Y* isolates obtained from potatoes grown in the United States. Phytopathology 94:1368-1375.
45. Plisson, C., Drucker, M., Blanc, S., German-Retana, S., Le Gall, O., Thomas, D., and Bron, P. 2003. Structural characterization of HC-Pro, a plant virus multifunctional protein. J. Biol. Chem. 278:23753-23761
46. Revers, F., Le Gall, O., Candresse, T., Le Romancer, M., and Dunez, J. 1996. Frequent occurrence of recombinant potyvirus isolates. J. Gen. Virol. 77:1953-1965.
47. Rozas, J., Sanchez-DelBarrio, J. C., Messeguer, X., and Rozas, R. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19:2496-2497.
48. Schubert, J., Fomitcheva, V., and Sztangret-Wisniewska, J. 2007. Differentiation of *Potato virus Y* strains using improved sets of diagnostic PCR-primers. J. Virol. Methods 140:66-74.
49. Shukla, D. D., Strike, P. M., Tracy, S. L., Gough, K. H., and Ward, C. W. 1988. The N and C termini of the coat protein of potyviruses are surface-located and N terminus contains the major virus-specific epitopes. J. Gen. Virol. 69:1497-1508.
50. Shukla, D. D., Ward, C. W., and Brunt, A. A. 1994. The Potyviridae. Cambridge University Press, Cambridge.
51. Sievert, R. C. 1971. Effect of time of inoculation with potato virus Y on yield and quality of burley tobacco. Phytopathology 61:588-589.
52. Sievert, R. C. 1978. Effect of potato virus Y on cultivars and hybrids of burley tobacco. Phytopathology 68:974-978.
53. Singh, R. P. 1992. Incidence of the tobacco veinal necrotic strain of potato virus Y (PVY^N) in Canada in 1990 and 1991 and scientific basis for eradication of the disease. Can. Plant Dis. Surv. 72:113-119.
54. Singh, R. P., Valkonen, J. P. T., Gray, S. M., Boonham, N., Jones, R. A. C., Kerlan, C., and Schubert, J. 2008. The naming of *Potato virus Y* strains infecting potato. Arch. Virol. 153:1-13.
55. Smith, K. M. 1931. On the composite nature of certain potato virus diseases of the mosaic group as revealed by the sue of plant indicators and selective methods of transmission. Proc. Royal Soc. B 109:251-267.
56. Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585-595.
57. Thomson, A. D., and Wright, D. S. C. 1966. Incidence and some effects of potato virus Y on New Zealand flue-cured tobacco. N. Z. J. Agric. Res. 9:886-893.
58. Tian, Y. P., Liu, J. L., Yu, X. Q., Lei, L. P., Zhu, X. P., Valkonen, J. P. T., and Li, X.-D. 2007. Molecular diversity of *Tobacco vein banding mosaic virus*. Arch. Virol. 152:1911-1915.
59. Tomitaka, Y., and Ohshima, K. 2006. A phylogeographical study of the *Turnip mosaic virus* population in East Asia reveals an 'emergent' lineage in Japan. Mol. Ecol. 13:4437-4457.
60. Tribodet, M., Glais, L., Kerlan, C., and Jacquot, E. 2005. Characterization of *Potato virus Y* (PVY) molecular determinants involved in the vein necrosis symptom induced by PVY^N isolates in infected *Nicotiana tabacum* cv. Xanthi. J. Gen. Virol. 86:2101-2105.
61. Verbeek, M., Piron, P. G. M., Dullemans, A. M., Cuperus, C., and van der Vlugt, R. A. A. 2010. Determination of aphid transmission efficiencies for N, NTN and Wilga strains of *Potato virus Y*. Ann. Appl. Biol. 156: 39-49.
62. Wang, H. Y., Liu, J. L., Gao, R., Chen, J., Shao, Y. H., and Li, X.-D. 2009. Complete genomic sequence analyses of *Turnip mosaic virus* basal-BR isolates from China. Virus Genes 38:421-428.
63. Wang, J., Wang, F., Qian, Y., Shi, J., Wang, C., Yang, D., Hu, M., Zhang,

Y., and Liu, C. 1998. Identification of tobacco virus strains in Shandong. Acta Tabacia Sin. 4:24-32. (In Chinese)
64. Wang, Z. F., Yan, D. Y., Zhang, G. M., Liu, K. Q., and Chen, R. T. 1986. Identification of tobacco virus type in Shandong. J. Shandong Agric. Univ. 17:75-82. (In Chinese)
65. Xie, C. S., Liu, Y. R., Wang, Z. F., and Zhang, X. G. 1993. Detection of tobacco virus disease by serological method. Chin. Tobacco 4:1-5. (In Chinese)
66. Yang, Z. 2007. PAML 4: A program package for phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24:1586-1591.
67. Yang, Z., and R. Nielsen. 2002. Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. Mol. Biol. Evol. 19:908-917
68. Yang, Z., Wong, W. S. W., and Nielsen, R. 2005. Bayes empirical Bayes inference of amino acid sites under positive selection. Mol. Biol. Evol. 22:1107-1118.