

Complete sequence analysis reveals two distinct poleroviruses infecting cucurbits in China

Hai-ying Xiang · Qiao-xia Shang · Cheng-gui Han ·
Da-wei Li · Jia-lin Yu

Received: 3 October 2007 / Accepted: 3 March 2008 / Published online: 15 April 2008
© Springer-Verlag 2008

Abstract The complete RNA genomes of a Chinese isolate of cucurbit aphid-borne yellows virus (CABYV-CHN) and a new polerovirus tentatively referred to as melon aphid-borne yellows virus (MABYV) were determined. The entire genome of CABYV-CHN shared 89.0% nucleotide sequence identity with the French CABYV isolate. In contrast, nucleotide sequence identities between MABYV and CABYV and other poleroviruses were in the range of 50.7–74.2%, with amino acid sequence identities ranging from 24.8 to 82.9% for individual gene products. We propose that CABYV-CHN is a strain of CABYV and that MABYV is a member of a tentative distinct species within the genus *Polerovirus*.

Cucurbit aphid-borne yellows virus (CABYV) belongs to the genus *Polerovirus* (family *Luteoviridae*). An isolate of this virus (CABYV-FRA) was first described in France, where it was the first polerovirus reported to infect cultivated cucurbits naturally and to cause a severe disease [10]. Later, the virus was found in Italy, Lebanon, Spain, Tunisia, and USA [1, 8, 11, 12, 16]. Thus far, only the CABYV-FRA genome has been completely sequenced, whereas partial sequences are available for Spanish and Italian isolates [4, 12, 16].

In China, only limited research has been done on a CABYV-like virus that Gu et al. [5] briefly mentioned. The identity of the virus could not be elucidated by routine serological tests because of strong serological cross-reactions among poleroviruses [3]. In 2006, we reported for the first time that CABYV occurs widely in mainland China, infecting nine different cucurbitaceous species in ten provinces surveyed [15, 18]. In the current study, we have extended our findings to present the complete genomic sequence of a Chinese CABYV isolate and have identified a new distinct polerovirus co-infecting cucurbits in China.

In a survey for CABYV in Beijing from July to October in 2006, samples showing yellowing symptoms were collected from nine different cucurbit crops: bitter melon (*Momordica charantia*), calabash gourd (*Lagenaria siceraria*), cucumber (*Cucumis sativus*), cushaw (*Cucurbita moschata*), muskmelon (*Cucumis melo*), squash (*Cucurbita pepo*), suakwa vegetable sponge (*Luffa cylindrical*), watermelon (*Citrullus lanatus*) and wax gourd (*Benincasa hispida*). These samples were stored at 4°C for several days before analysis, or at –20°C for long-term storage. Virus particles were purified from infected leaves as described previously [14]. Total RNA from infected plants or purified viral RNA was prepared and used for first-strand cDNA synthesis as described by Han et al. [6].

Hai-ying Xiang and Qiao-xia Shang contributed equally.

H. Xiang · Q. Shang · C. Han (✉) · D. Li · J. Yu
Department of Plant Pathology and State Key Laboratory
for Agrobiotechnology, China Agricultural University,
100094 Beijing, China
e-mail: hanchenggui@cau.edu.cn

H. Xiang
e-mail: xhy236@yahoo.com.cn

D. Li
e-mail: lidw@cau.edu.cn

J. Yu
e-mail: bnyvv@public.bta.net.cn

Q. Shang
Department of Plant Science and Technology,
Beijing University of Agriculture, 102206 Beijing, China
e-mail: shangjiaoxia@bac.edu.cn

Universal primers (PococpR/PoconF) similar to those described previously for polerovirus detection [13] were designed by multiple alignment of known polerovirus sequences available from GenBank. Primers for amplification and sequencing of the genomic sequences were derived from the sequences of CABYV-FRA and other poleroviruses (see legend to Fig. 1 for accession numbers) (Table 1).

Synthesis of cDNA was done with the primer PococpR, which in combination with the universal primer PoconF yielded an RT-PCR product of the expected size (~1.4 kb), encompassing the 3' third of ORF2, the intergenic NCR and the complete ORF3. Subsequently, the resulting RT-PCR products were cloned, and at least two independent PCR clones were sequenced. NCBI Blast searches revealed that there were two different types of sequences. One sequence type (accession numbers: EU091148, EF063707, and EU000535) shared 93.7–94.9% nucleotide sequence identity with CABYV-FRA, and all the Chinese isolates within this type shared 93.8–99.1% identity. In contrast, the other sequence type (accession numbers: EU091149–EU091151) shared only 77.2–79.2% identity with CABYV-FRA, and all the Chinese isolates of this type were very similar to each other (99.3–99.8%). Based on the molecular criterion (10% sequence difference) for polerovirus species demarcation [2], these data suggested that the former isolates can be assigned to CABYV (and are hence referred to as CABYV-CHN) and the latter to a new polerovirus species tentatively referred to here as Melon aphid-borne yellows virus (MABYV) due to its first detection in muskmelon and wax gourd in Beijing.

To verify this hypothesis, primers CA3414F and MA3566F were designed for specific detection of CABYV and MABYV, respectively, by an RT-PCR method similar to that described by Hauser et al. [7]. In total, 39 of 79 plants were positive when primers permitting detection of both CABYV and MABYV were used, while 31 and 34 of 79 plants tested positive, respectively, when CABYV- and MABYV-specific primer combinations were used. Interestingly, the results also showed that the vast majority of plants analysed appear to be infected with both CABYV and MABYV in cushaw, squash, suakwa vegetable sponge, and wax gourd. However, single infections with CABYV were observed only for bitter melon and cucumber and single infections with MABYV only for muskmelon and calabash gourd (Table 2).

In order to obtain the complete sequence of a CABYV-CHN isolate from diseased leaves of cushaw, four overlapping fragments of the entire genome of CABYV-CHN were amplified using four primer pairs (CA001F/CA784R, CA613F/CA2886R, CA2715F/CA4117R, and CA3488F/CA5682R; Table 1), cloned and sequenced. Thus, the

complete genomic sequence of CABYV-CHN was determined, apart from short regions where the primers annealed at the 5' and 3' termini (GenBank accession number EU000535). The genomic sequence of CABYV-CHN is 5,682 nt in length, and computer analysis revealed a genomic organization and features very similar to those of CABYV-FRA [4]. The first ORF (ORF0) begins with an AUG at nt 21–23 and terminates with a UAA at positions 738–740 to encode a putative protein (P0) of 27.6 kDa. The second ORF (ORF1) begins with an AUG at nt 142–144 and ends with an UGA at nt 2035–2037 to encode a putative protein (P1) of 69.2 kDa. The third ORF (ORF2) overlaps with ORF1 and begins at nt 142–144 and terminates with an UGA at nt positions 3309–3311 to produce a putative P1–P2 fusion protein of 119.2 kDa. An intergenic NCR of 199 nt (3312–3510) is present between the ORF2 and ORF3. The fourth ORF (ORF3) begins with AUG at nt 3511–3513 and terminates with an UAG at nt 4108–4110 and can be translated into a major coat protein (P3) of 22.1 kDa. The fifth ORF (ORF4) overlaps with ORF3 in a different reading frame, beginning with AUG at nt 3539–3541 and ending with an UGA at nt 4112–4114 to code a movement protein (P4) of 20.7 kDa. The last ORF (ORF5), immediately adjacent to ORF3 and in the same reading frame, ends with an UAA at positions 5515–5517. ORF3 and ORF5 potentially produce a minor structural protein of 74.5 kDa by a translational readthrough. The 3'-noncoding region is 165 nt in length. We did not define any function of the ORFs identified but assume their function is analogous to those of other poleroviruses [2].

Further sequence comparisons showed that the full sequence of CABYV-CHN shared 89.0% nucleotide sequence identity with CABYV-FRA. The nucleotide sequence identities for the individual ORFs of CABYV-CHN and CABYV-FRA were 87.5% (ORF0), 88.3% (ORF1), 89.9% (ORF1–ORF2), 92.5% (intergenic NCR), 94.7% (ORF3), 94.8% (ORF4), 85.7% (ORF5) and 74.8% (3' NCR). At the amino acid sequence level, the ORF identities between CABYV-CHN and CABYV-FRA were 80.3% (P0), 87.8% (P1), 90.5% (P1–P2), 94.0% (P3), 90.1% (P4), and 90.8% (P5). In addition, the CABYV-CHN sequences were also compared with those of other members of the genus *Polerovirus*. This revealed that the complete nucleotide sequence identity of CABYV-CHN with other poleroviruses ranged from 50.8 to 68.5%, and the amino acid sequence identities for the individual gene products ranged from 23.7 to 71.5%. Based on the more than 10% differences in the amino acid sequences of P0 and P1 between CABYV-CHN and CABYV-FRA, we identified CABYV-CHN as a strain of CABYV [2].

Similarly, the complete genomic sequence of a wax gourd isolate of MABYV was also determined. Four overlapping fragments of the MABYV genome were

Fig. 1 Phylogenetic trees generated from peptide sequence alignments of poleroviruses. The bootstrap analysis was performed by the unrooted NJ method and drawn using the MEGA 3.1 program. **a** P1–P2 fusion protein, **b** P3 coat protein, **c** coat protein of different isolates of CABYV and MABYV. The CABYV-CHN sequence has been assigned the accession number EU000535 and MABYV EU000534. GenBank database accession numbers of the other sequences used here are as follows: beet chlorosis virus (*BChV*, NC002766), beet mild yellowing virus (*BMV*, NC003491), beet western yellows virus (*BWYV*-USA, NC004756), cereal yellow dwarf virus-RPS (*CYDV-RPS*, NC002198), cereal yellow dwarf virus-RPV (*CYDV-RPV*, NC004751), cucurbit aphid-borne yellows virus (*CABYV-FRA*, X76931), chickpea chlorotic stunt virus (*CpCSV*, AY956384), potato leafroll virus (*PLRV*, NC001747), sugarcane yellow leaf virus (*ScYLV*, NC000874), and turnip yellows virus (*TuYV-FLI*, NC003743). The accession numbers of the coat protein sequences of different CABYV and MABYV isolates from China, France, Italy and Spain are given in **c**

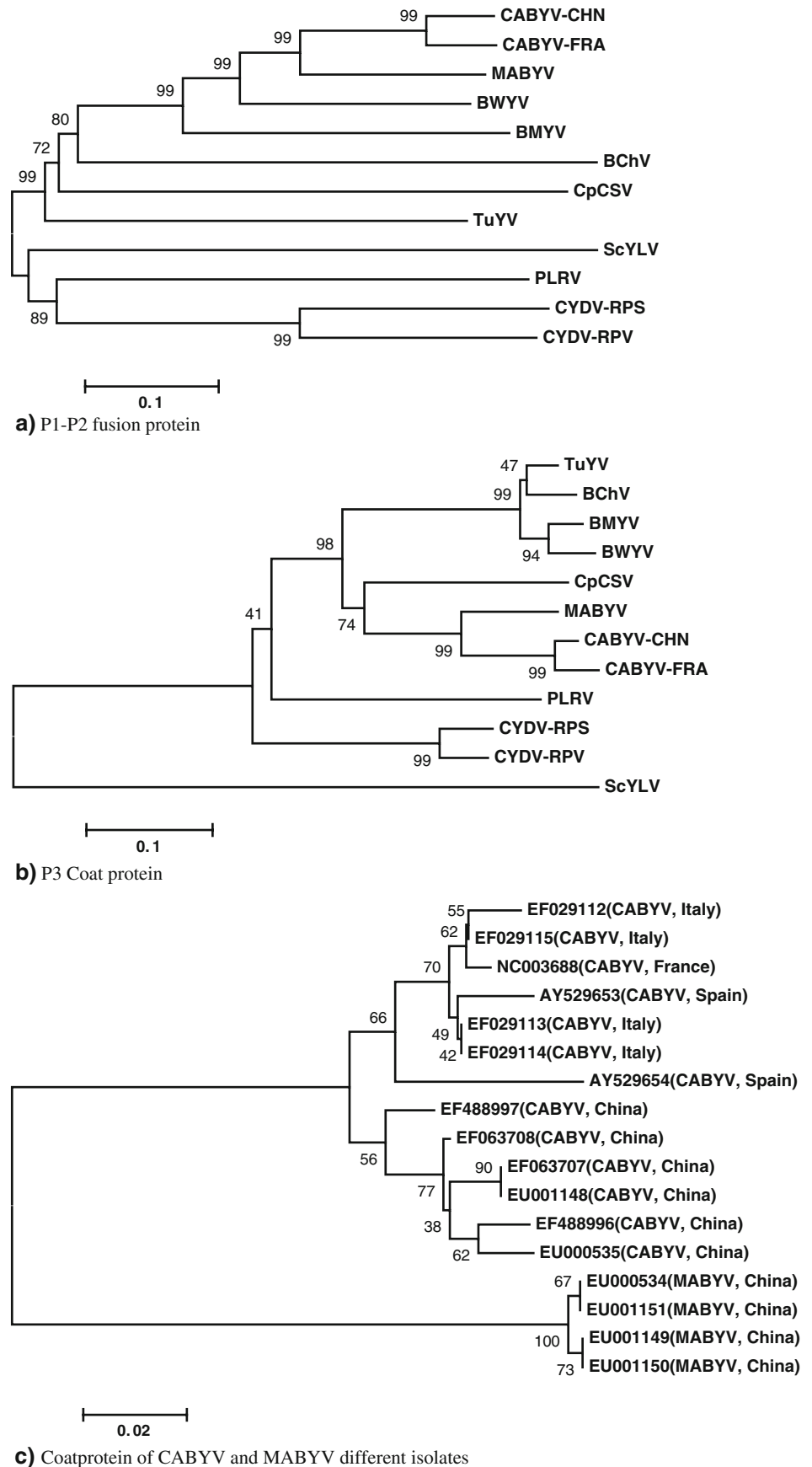


Table 1 Oligonucleotide primers used for RT-PCR and sequencing

Primer ^a	Primer sequence	Positions
CA5682R	5'-ACACCGAAACGCCAGGGGGAATC-3'	CA5660-5682
Pocon3R	5'-ACACCGAA-3'	CA5675-5682, MA5667-5674
CA5637R	5'-AACRYTCGTTGCTATCC-3'	CA5620-5637, MA5612-5629
MA5276F	5'-CGTGTGGATGAACAC-3'	MA5276-5291
MA5324F	5'-GAGCTTGGGAGGATC-3'	MA5324-5338
MA5436F	5'-ATCCGGAACCAACTC-3'	MA5436-5450
CA4077F	5'-CACCATCAAGTGCCAGGTCC-3'	CA4077-4096
CA4117R	5'-CGTCTACCTATTTCTGGG-3'	CA4100-4117, MA4105-4119
MA3566F	5'-GAACCGTCGACGCACTTCAAAG-3'	MA3566-3587
CA3488F	5'-GCACACGTGCGTTCAATTGTTAAT-3'	CA3488-3513
CA3414F	5'-GGATTTTCTAGCGGTCTA-3'	CA3414-3431
CA2886R	5'-GCGCTAATAATGTTCCGTCAG-3'	CA2866-2886
MA2837R	5'-AGCTGATCTTAAGCGCTTGG-3'	MA2818-2837
CA2715F	5'-GACTGCTCCGGTTTTGACTGG-3'	CA2715-2735, MA2724-2744
CA2070R	5'-CCAGCCAACGCGCGGCTTG-3'	CA2052-2070
MA2063R	5'-TTGTAGCCGCAAGGATTTTCAG-3'	MA2043-2063
CA784R	5'-GGCTCTGTGGTATTTTGAATG-3'	CA764-784
MA713R	5'-CGCGAGATACCGCTCGCTGC-3'	MA694-713
CA613F	5'-ACCGTATWTATGGTGAAGA-3'	CA613-631, MA619-637
CA001F	5'-ACAAAAGATACGAGCGGGTGATGCAAATTGA-3'	CA001-031, MA001-031
Pocon5F	5'-ACAAAAGAN-3'	CA001-009, MA001-009
PococpR	5'-CGTCTACCTATTTSGGRTTN-3'	Universal reverse primer
PoconF	5'-TGYTCYGGTTTTGACTGG-3'	Universal forward primer

CA CABYV, MA MABYV; W A or T, S G or C, R A or G, N A or T or C or G, Y C or T

^a F matching, or R reverse complementary, to virus sequence

Table 2 RT-PCR detection of CABYV and MABYV infections in various field-grown host species

Host	CABYV and MABYV ^a	CABYV only ^b	MABYV only ^c
Cushaw	10/21	10/21	10/21
Squash	3/4	3/4	3/4
Wax gourd	9/10	9/10	9/10
Suakwa vegetable sponge	3/8	3/8	3/8
Muskmelon	4/4	0/4	4/4
Watermelon	1/10	1/10	1/10
Calabash gourd	4/8	0/8	4/8
Bitter melon	2/3	2/3	0/3
Cucumber	3/11	3/11	0/11
Total plants	39/79	31/79	34/79

Numerator: number of plants testing positive, denominator: number of plants tested

^a Primers CA4117R and CA2715F were used for detection of both CABYV and MABYV (c.1.4 kbp)

^b Primers CA4117R and CA3414F were used for detection of CABYV only (704 bp)

^c Primers CA4117R and MA3566F were used for detection of MABYV only (554 bp)

amplified using four primer pairs (CA001F/CA784R, CA613F/MA2837R, CA2715F/CA4117R, CA3488F/CA5637R; Table 1). The 3'-terminal fragment of 240 bp was generated by a nested PCR approach: the primers Pocon3R and MA5276F were used for the first PCR, Pocon3R and MA5324F for the second, and Pocon3R and MA5436F for the third. The 5' terminal region was also amplified using a similar strategy (the primers CA784R and Pocon5F were used for the first PCR, and MA713R and Pocon5F for the second). All these PCR products were cloned and sequenced. The genomic sequence of MABYV was 5,674 nt long, apart from short regions where the primers annealed to the 5' and 3' termini, which contain eight nucleotides that are highly conserved in poliovirus genomes (GenBank accession number EU000534). The genome is also predicted to contain six large ORFs, whose functions we assume to be analogous to those of other polioviruses [2]. The first ORF (ORF0) begins at the first AUG (nt 21–23) and terminates with an UGA at nt 744–746 to encode a putative protein (P0) of 27.9 kDa. The second ORF (ORF1) begins at the second AUG (nt 151–153) and terminates with an UGA at nt 2044–2406, to produce a P1 of 68.6 kDa that is predicted to be translated by leaky scanning. ORF2 begins at nt 1479 and ends at

nt 3320 with a UAA termination codon, followed by an intergenic NCR of 195 nts. The 5' end of ORF2 overlaps ORF1, and expression of P2 is predicted to be generated by a -1 frameshift at a shifty hepta-nucleotide sequence GGGAAAC (nt 1476–1482). This putative translation product results in a 117.9 kDa P1–P2 fusion protein thought to be involved in virus replication. ORF3 begins at nt 3516 and terminates at nt 4115 with an UAG and potentially codes for the major coat protein (22.0 kDa). ORF4 (nt 3544–4119), containing an UAG stop codon, nearly completely overlaps ORF3 but is in a different reading frame, and is thought to encode the 21.2 kDa movement protein (P4). The last ORF (ORF5) is immediately adjacent to ORF3 and ends at a UAA (nt 5505–5507). Translation of ORF3–ORF5 via an in-frame translational read-through of the ORF3 stop codon, is predicted to produce a read-through fusion protein of 73.3 kDa. The 3' non-coding region is 167 nt in length. These data reveal that MABYV has a genome organization very similar to that of CABYV and other poleroviruses [2, 4].

The complete nucleotide sequence identities between MABYV and other poleroviruses including CABYV ranged from 50.7–74.2%. The nt sequence identities for the individual genome parts of MABYV and other poleroviruses were 51.8–82.1% for ORF0, 54.0–82.7% for ORF1, 60.0–76.0% for ORF1–ORF2, 54.0–71.4% for the intergenic NCR, 58.8–82.9% for ORF3, 59.5–82.5% for ORF4, 51.1–65.3% for ORF5 and 43.1–84.7% for the 3' NCR. At the amino acid sequence level, the most conserved protein P3 (CP) shared identities ranging from 41.5 to 82.9% with other poleroviruses; P0 had 25.1–73.2%, P1 had 32.0–63.9%, P1–P2 had 47.6–73.7%, P4 had 36.5–67.4%, and P5 had 25.6–57.0%. Interestingly, MABYV and CABYV shared highest similarities in all the proteins and non-coding regions. However, the differences in amino acid sequences of all the gene products were greater than 10% between MABYV and other poleroviruses including CABYV.

In order to understand the relationships of CABYV-CHN and MABYV with other poleroviruses, phylogenetic trees were generated by ClustalX (Vers. 1.83) [17] and visualized with MEGA (Vers. 3.1) [9]. The phylogenetic trees of all the proteins further confirmed that MABYV is most closely related to CABYV, indicating that they may have a common origin. Phylogenetic analysis also showed that the P0, P1 and P2 proteins of CABYV and MABYV are closely related to the corresponding proteins of beet western yellows virus (BWYV) and beet mild yellowing virus (BMV), whilst they are more similar to chickpea chlorotic stunt virus (CpCSV) in the P3 protein (Fig. 1a, b).

In addition, the alignment of the complete CP sequences of different CABYV and MABYV isolates from Europe and China showed that the poleroviruses infecting cucurbit plants could be classed into two groups, the MABYV

group and the CABYV group, and that the latter could be further divided into two subgroups, a Chinese and a European subgroup, suggesting that there is geographically associated variation among CABYV isolates (Fig. 1c).

The complete genomic sequences of CABYV and MABYV provide critical information needed to resolve their taxonomic status. Based on the 10% amino acid sequence difference criterion [2], and the fact that CABYV and MABYV often co-exist in infected plants, thus not providing cross-protection against each other (Table 2), it is evident that two distinct cucurbit poleroviruses exist in China, and we propose that MABYV be a new polerovirus infecting cucurbits. However, relatively little is known about the pathogenicity, serological specificity, transmission specificity and geographical distribution of MABYV.

Acknowledgments We would like to thank Professor Andrew O. Jackson for helpful comments on the manuscript. This work was supported by the National Key Basic Research Program of China (2006CB101903), the Natural Science Foundation of Beijing (6082006) and the Science and Technology Nova Program of Beijing (2007-B-032).

References

1. Abou-Jawdah Y, Sobh H, Fayyad A, Lecoq H (1997) First report of cucurbit aphid-borne yellows luteovirus in Lebanon. *Plant Dis* 81:1331
2. D'Arcy CJ, Domier LL (2005) *Luteoviridae*. In: Fauquet CM, Mayo MA, Maniloff J, Desselberger U, Ball LA (eds) *Virus taxonomy*, VIIIth Report of the ICTV. Elsevier/Academic Press, London, pp 891–900
3. D'Arcy CJ, Torrance L, Martin RR (1989) Discrimination among luteoviruses and their strains by monoclonal antibodies and identification of common epitopes. *Phytopathology* 79:869–873
4. Guilley H, Wipf-Scheibel C, Richards K, Lecoq H, Jonard G (1994) Nucleotide sequence of cucurbit aphid-borne luteovirus. *Virology* 202:1012–1017
5. Gu Q (2002) Progress on management of main viral diseases on cucurbitaceous crops (in Chinese). *China Watermelon Muskmelon* 2:27–30
6. Han C, Li D, Xing Y, Zhu K, Tian Z, Cai Z, Yu J, Liu Y (2000) Wheat yellow mosaic virus widely occurring in wheat in China. *Plant Dis* 84:627–630
7. Hauser S, Weber C, Vetter G, Stevens M, Beuve M, Lemaire O (2000) Improved detection and differentiation of poleroviruses infecting beet or rape by multiplex RT-PCR. *J Virol Methods* 89:11–21
8. Juarez M (2004) First report of cucurbit aphid-borne yellows virus in Spain. *Plant Dis* 88:907
9. Kumar S, Tamura K, Nei M (2004) MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Bioinformatics* 5:150–163
10. Lecoq H, Bourdin D, Wipf-Scheibel C, Bon M, Lot H, Lemaire O, Herrbach E (1992) A new yellowing disease of cucurbits caused by a luteovirus, cucurbit aphid-borne yellows virus. *Plant Pathol* 41:749–761
11. Lemaire O, Gubler WD, Valencia J, Lecoq H, Falk BW (1993) First report of cucurbit aphid-borne yellows virus in the United States. *Plant Dis* 77:1169

12. Mnari M (2005) First report of cucurbit aphid-borne yellows virus in Tunisia causing yellows on five cucurbitaceous species. *Plant Dis* 89:776
13. Robertson NL, French R, Gray SM (1991) Use of group-specific primers and the polymerase chain reaction for the detection and identification of luteoviruses. *J Gen Virol* 72:1473–1477
14. Rowhani A, Smith RS (1979) Purification and characterization of potato leafroll virus. *Virology* 98:45
15. Shang Q, Xiang H, Han C, Li D, Yu J (2008) Partial sequence analysis of two isolates of cucurbit aphid-borne yellows virus from Hubei and Yunnan in China. *Acta Phytopathol Sin* 38:64–68
16. Tomassoli L, Meneghini M (2007) First report of cucurbit aphid-borne yellows virus in Italy. *Plant Pathol* 56:720
17. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673–4680
18. Xiang H, Shang Q, Han C, Li D, Yu J (2007) First report on the occurrence of cucurbit aphid-borne yellows virus on nine cucurbitaceous species in China. *New Dis Rep* 15. <http://www.bspp.org.uk/ndr>