Arch Virol (2004) 149: 435–445
DOI 10.1007/s00705-003-0234-2

Archives of Virology
Printed in Austria

Detection and classification of allexiviruses from garlic in China

J. Chen[1], **H.-Y. Zheng**[1], **J. F. Antoniw**[2], **M. J. Adams**[2], **J.-P. Chen**[1], and **L. Lin**[1]

[1]Virology and Biotechnology Department, Zhejiang Academy of Agricultural Sciences, Hangzhou, P.R. China
[2]Plant Pathogen Interactions Division, Rothamsted Research, Harpenden, Herts, U.K.

Received April 9, 2003; accepted September 23, 2003
Published online November 13, 2003 © Springer-Verlag 2003

Summary. Degenerate primers for RT-PCR were designed and used to amplify genome fragments (*c.* 750 nt in the coat protein-ORF6 region) of allexiviruses from a total of 28 garlic samples from 24 provinces in China. Many samples contained more than one distinct sequence. A total of 60 different sequences were obtained. Phylogenetic analysis and two-way comparisons were used to assess the status of the sequences and to re-examine the criteria for distinguishing species within the genus. Most of the sequences could be allocated to either *Garlic virus D* or *Garlic virus X* on the basis of sequence similarity but some appeared to be intermediate between existing species. There were no sequences of *Garlic virus C* or *Shallot virus X*. A comparison with the related genera *Carlavirus*, *Foveavirus* and *Potexvirus* suggests that the published allexivirus species demarcation criteria may have been drawn too tightly and should be re-examined.

Introduction

Virus diseases of garlic (*Allium sativum* L.) are widespread throughout the world, causing serious damage to crop yields and quality. Whatever their primary means of transmission, the vegetative propagation of the crop favours their dissemination and accumulation in bulbs. Various filamentous viruses associated with mosaic symptoms have been reported from garlic, often in complex mixtures. These include at least two members each of the genera *Potyvirus* (*Onion yellow dwarf virus*, OYDV and *Leek yellow stripe virus*, LYSV) and *Carlavirus* (*Garlic common latent virus*, GarCLV and *Garlic latent virus*, GarLV = *Shallot latent virus*, SLV [13]). Several members of the recently established genus *Allexivirus* have also been reported from such plants [3, 4, 6, 7, 9, 11–14, 16, 17]. Viruses of the genus

Allexivirus are thought to be mite-transmitted and have a genome organisation similar to that of carlaviruses, with a single stranded polyadenylated RNA genome encoding a large (220 kDa in carlaviruses; 170–195 kDa in allexiviruses) alpha-like replicase (containing methyl transferase, helicase and RNA-dependent RNA polymerase motifs) and 5 smaller ORFs. In carlaviruses, ORFs 2–4 (approx 26, 12 and 7 kDa) slightly overlap one another in different reading frames and constitute the triple gene block, the proteins of which are implicated in cell-to-cell movement. ORF5 is the coat protein (32–36 kDa) and ORF6 is a nucleic-acid binding protein (11–16 kDa). Allexiviruses are similar but the third triple gene block ORF lacks a classical initiation codon and there is an ORF4 encoding a 40 kDa serine-rich protein with no known homology to other reported proteins and the coat protein (28 kDa) is a little smaller. The genera *Potexvirus* and *Foveavirus* are similar to carlaviruses in their genome organisation but lack ORF6 [15] (Fig. 1).

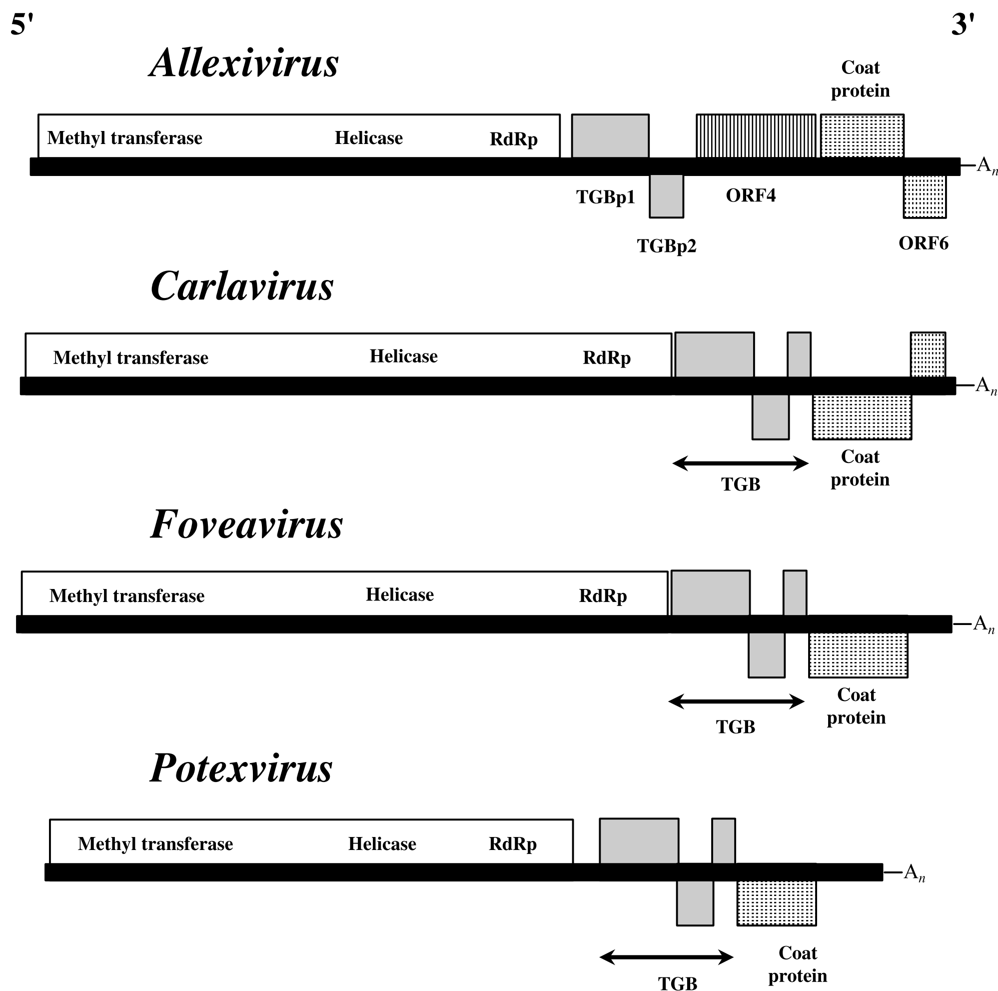

Fig. 1. Diagram showing the genome organisation of viruses in the genera *Allexivirus*, *Carlavirus*, *Foveavirus* and *Potexvirus*. RdRp, RNA-dependent RNA polymerase; TGB, triple gene block proteins

Traditional phytopathological methods based on serology, host range and symptoms have limited ability to differentiate *Allium* viruses because of multiple infection by viruses with similarly restricted host ranges. However, recent advances in molecular biology have provided new tools for identification and classification of viruses. Genome sequences and organisation have proved invaluable for distinguishing virus strains and for determining the relationships between genera, species and subspecies of distinct viruses. Largely based on such methods, a number of species have been described in the genus *Allexivirus*. All these viruses were isolated from *Allium* spp. There are complete sequences of five viruses (*Garlic virus A* (GarV-A), *Garlic virus C* (GarV-C), *Garlic virus E* (GarV-E), *Garlic virus X* (GarV-X) and *Shallot virus X* (ShV-X)) and partial sequences of three others (*Garlic mite-borne filamentous virus* (GarMbFV), *Garlic virus B* (GarV-B) and *Garlic virus D* (GarV-D)) [3, 6–10].

Garlic is an economically important crop in China and mosaic disease causes severe loss of yield and poor quality. Recently, the nucleotide sequences of many Chinese garlic potyvirus and carlavirus isolates were reported [2–4]. Two allexiviruses were also reported from Yuhang, Zhejiang province [3] but there have been no other reports of the distribution and molecular characterisation of Chinese allexivirus isolates.

In this paper, we report the use of degenerate primers to detect allexiviruses in garlic plants collected from 28 sites distributed in 24 provinces in China. The amplified cDNAs were sequenced and an analysis to determine the correct names and classification of these viruses is presented.

Materials and methods

Dormant garlic bulbs were collected from 28 commercial sites between April 2000 and December 2002 (Table 1) and planted in the greenhouse at the Virology Laboratory, Zhejiang Academy of Agricultural Sciences, Hangzhou. Leaves were collected and stored at −80 °C until used. Plant total RNAs were extracted using the RNeasy Plant Mini Kit (QIAGEN). First-strand cDNA was synthesised using M-MLV Reverse transcriptase (Life Technologies Ltd) according to the manufacturer's instructions and with M4T (5′-GTT TTC CCA GTC ACG AC $(T)_{15}$-3′) as the initial primer. The PCR reaction used a LATaq PCR system (TaKaRa) according to the manufacturer's protocols. Fragments were separated by electrophoresis through 1% (w/v) agarose gels and purified using the QIAGEN Gel Extraction Kit (QIAGEN). Two degenerate primers were designed from an alignment of published sequences of members of the genus *Allexivirus*: Allex-CP(+): 5′-TGG RCX TGC TAC CAC AAY GG-3′ and Allex-NABP(−): 5′-CCY TTC AGC ATA TAG CTT AGC-3′ (X = A, T, C, G; Y = T, C; R = A, G). These were predicted to amplify a fragment of *c.* 750 bp, covering the 3′-half of the coat protein (*c.* 440 of 750 nts) and most of ORF6 (*c.* 310 of 390 nts). PCR fragments were cloned into the pGEM-T vector (Promega) and were auto-sequenced in both directions by the ABI PRISM™ 377 DNA Sequencer, using the BigDye™ Terminator Cycle Sequencing Ready Reaction Kit (PERKIN ELMER). Sequence analysis used programs from the Wisconsin (GCG) package version 10.3 [1] and, in particular, pairwise comparisons were done with GAP (using a gap creation penalty of 50 and a gap extension penalty of 3 for nucleotide comparisons and values of 8 and 2 respectively for amino acids). To assist with large numbers of pairwise comparisons, software was written to generate batch files that were run in GCG

Table 1. Origins of the garlic bulbs and the allexivirus sequences that were obtained from them. Each entry represents a sequence and shows the number of clones that had identical sequences

Abbr	Site	GarV-B	GarV-D	GarV-E	GarV-A	Group 1	Group 2	Group 3	GarV-X	Group 4
AB	Aba Tibetan and Qiang autonomous prefecture, Sichuan province		2						9 (AB1) 1 (AB2)	
BD	Baoding city, Hebei province								12	
BT	Baotuo city, Inner Mongolia Autonomous Region						6		6	
CFH	Hongshan county, Chifeng city, Inner Mongolia Autonomous Region (red skin cultivar)		3	4				5		
CFL	Linxi county, Chifeng city, Inner Mongolia Autonomous Region (white skin cultivar)		1				11			
CQ	Chongqing city, Sichuan province								12	
CS1	Changsha city, Hunan province (purple skin cultivar)						4		3 (CS1–1)	1 (CS1–2)
CS2	Changsha city, Hunan province (white skin cultivar)			1				2	4 (CS2–1)	4 (CS2–2)
DL	Dali city, Yunnan province		6						6	
GY	Guiyang city, Guizhou province								13	
HK	Haikou city, Hainan province		10		4					
JX	Jinxiang city, Shandong province								14	
LA	Liu’an city, Anhui province		8			3			1	
LZ	Liuzhou city, Guangxi Zhuang Autonomous Region		9			1			2	

NJ	Nanjing city, Jiangsu province						13
SS	Sanshui city, Guangdong province		12			2	
TM	Tumen city, Jilin province		5				6
TP	Turpan prefecture, Xinjiang Uygur Autonomous Region						9
TS1	Tianshui city, Gansu province (multiple bulb cultivar)	4 (TS1–1) 1 (TS1–2)		5			4
TS2	Tianshui city, Gansu province (single bulb cultivar)	1		6			5
WH	Wuhan city, Hubei province		9		1		
XA	Xi'an city, Shanxi province		3				9
XX	Xixia city, Henan province		1				11
YC	Yinchuan city, Ningxia Hui Autonomous Region	1		5			6
YY	Yiyang city, Jiangxi province		4 (YY1) 1 (YY2)		2 (YY1)	5 (YY2)	
ZhaZ	Zhangzhou city, Fujian province		1				11
ZheZ	Zhengzhou city, Henan province						12

and also to extract and summarise data from the output files. Phylogenetic analyses were done using programs in PHYLIP version 3.6 [5]. Phylogenetic trees were constructed by DNADIST (with 3 categories of substitution rate estimated from TREEPUZZLE) and NEIGHBOR using the original data set and 100 bootstrap data sets generated by the program SEQBOOT from the original set. Trees with the same topology were obtained using both FITCH and DNAML.

The nucleotide sequences obtained were deposited with the DDBJ/GenBank/EMBL databases with the accession numbers AJ551467–AJ551526.

Results and discussion

Genome fragments were successfully amplified from a total of 28 garlic samples from 24 provinces using the allexivirus degenerate primers. Most samples contained 2 or 3 distinct sequences and therefore, to detect these multiple infections, about 12 clones were sequenced from each sample. Only 6 samples had the same sequence in all clones and a total of 60 different sequences (from 335 clones) were obtained and were provisionally allocated to species or groups on the basis of sequence similarity. Most sequences were allocated to either GarV-D or GarV-X and there were no sequences closely related to the published GarV-C or ShV-X sequences. Although samples from the same region sometimes had similar virus isolates (e.g. the GarV-B isolates from Gansu and Ningxia in north-central China and the two sequences from Hunan province labelled group 4 in Table 1) there was little overall relationship between sequence and geographical origin, suggesting that the viruses have mostly been distributed with the vegetatively propagated planting material. The only virus to be found on its own was GarV-X, while GarV-D could be found in association with each of the other sequence types.

For phylogenetic analysis, the sequences were aligned with all the publicly available sequences of allexiviruses covering the same region, to give an alignment of 748 nt with 74 unique sequences. In addition, a complete set of pairwise comparisons was made between the nucleotide and predicted amino acid sequences.

The frequency distribution of the 2850 pairwise nucleotide comparisons of allexivirus sequences (Fig. 2) indicates several levels of grouping. There was a major peak corresponding to the lowest level of sequence identity (62–72%) and this separated the new sequences into two major groups. One group included the GarV-E, GarV-D and GarV-A sequences, while the other included GarV-B and GarV-X. This also corresponds with an organisational difference as sequences of GarV-B, GarV-C and GarV-X with the 2 new sequences in group 4 have a short (9–17 nt) non-translated region between the two genes, whereas in the other sequences, the final nucleotide of the coat protein stop codon (TAA or TGA) is the first nucleotide of the ORF6 start codon (ATG). The distribution shows a second peak at 74–84% nucleotide identity that could represent either closely related species or distantly-related strains of the same virus. It would therefore be possible in principle to use either about 73% or 86% nucleotide identity to discriminate between species. These values would separate the existing species into a total of either 4 or 11 species as shown on the phylogenetic tree (Fig. 3) which indicates either that some existing species would be amalgamated (GarV-E,

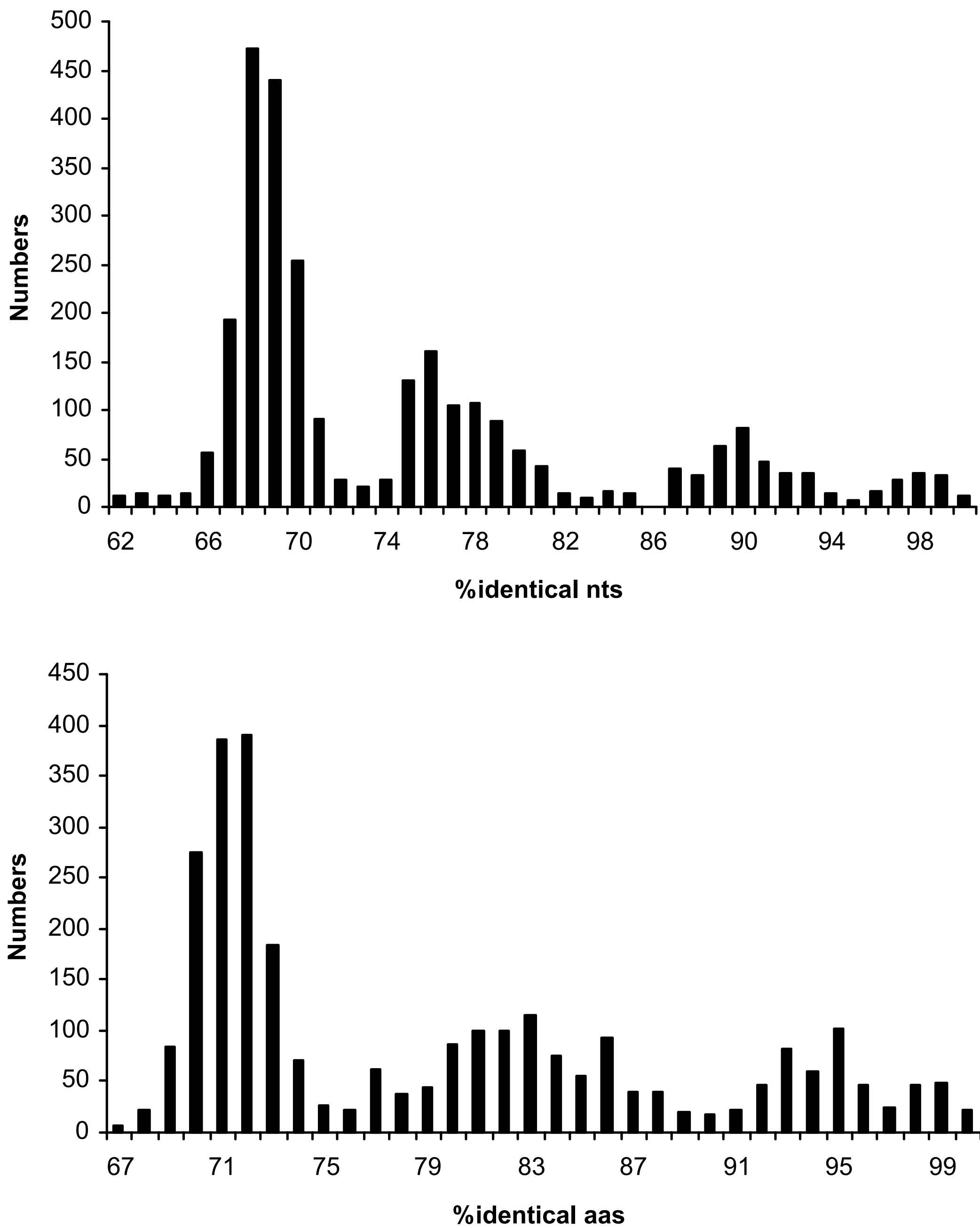

Fig. 2. Histograms showing the distribution of nucleotide and amino acid identity amongst pairwise comparisons of all allexivirus sequence fragments

GarV-D and GarV-A as one species and GarV-B and GarV-X as a second) or that four additional species are represented amongst the new sequences. Pairwise comparisons of the predicted amino acid sequences gave a similar pattern with either about 76% or 90% identical amino acids as the demarcation between species

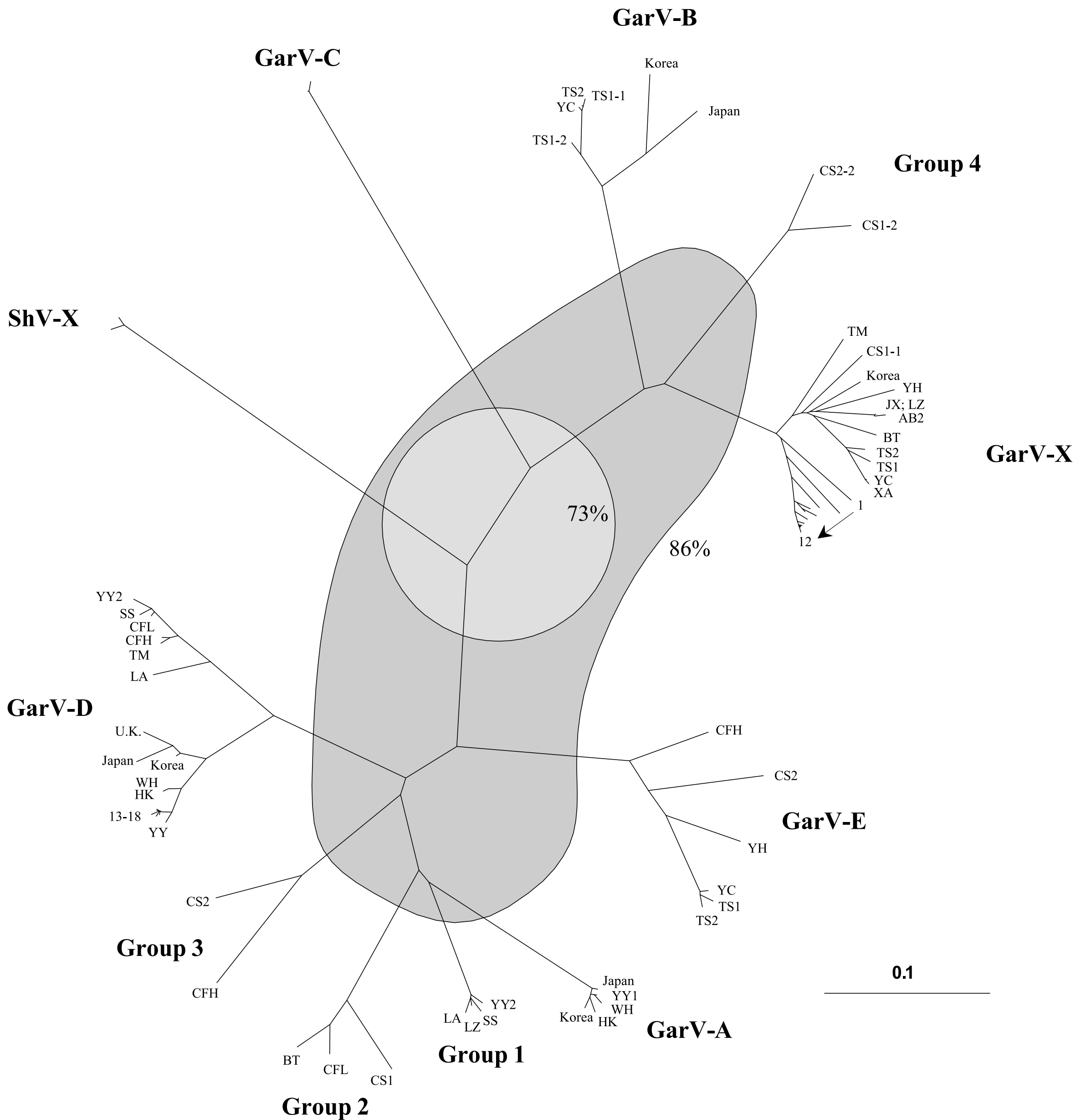

Fig. 3. Unrooted phylogenetic tree of the nucleotide sequences of the allexivirus sequence fragments using NEIGHBOR. The scale bar shows the number of substitutions per base. Possible positions for species discrimination using nucleotide identities of 73% or 86% are indicated. All major branches were supported by bootstrap values >88% (based on 100 replicates). Sequence codes are as in Table 1. GarV-X sequences 1 to 12 are respectively GY, CS2, CQ, ZhaZ, AB1, ZheZ, NJ, BD, TP, XX, DL, LA and GarV-D sequences 13 to 18 are XX, LZ, XA, AB, DL, ZhaZ. Sequences in addition to those determined in this study are: GarV-A from Japan (AB010300) and Korea (AF478197), GarV-B from Japan (AB010301) and Korea (AF543829), GarV-C (AB010302, D49443), GarV-D from Japan (AB010303), Korea (AF519572) and the UK (L38892), GarV-E from Yuhang, Zhejiang province (YH, AJ292230), GarV-X from Yuhang, Zhejiang province (YH, AJ292229) and ShV-X (L76292, M97264)

(Fig. 2). The higher number is consistent with a coat protein amino acid sequence identity of <90%, one of the species demarcation criteria for the genus *Allexivirus* in the 7th ICTV report [15].

There are proposals to establish a new virus family that would include the genus *Allexivirus* with related viruses, including the genera *Carlavirus*, *Foveavirus* and *Potexvirus*. To help decide the appropriate sequence difference for species discrimination within the allexiviruses, an attempt was made to establish robust criteria for species discrimination across these related genera. The variability within and between viruses of the other three genera was therefore examined. Pairwise comparisons were made between all completely sequenced genomes and also for the nucleotide and amino acid sequences of each completely sequenced gene of all available sequences. The data are summarised for the genomes, polymerase, coat protein and ORF6 in Table 2, which shows that different species within these genera usually have less than about 75% identical nucleotides and their coat proteins less than about 80% amino acid identity. Most individual genes fit with this pattern except for the very small ones where the sequence is probably too short to give a reliable statistic.

Table 2. Percentage nucleotide and amino acid identity from pairwise comparisons within and between species (as currently defined) of the genera *Allexivirus*, *Carlavirus*, *Foveavirus* and *Potexvirus*

	%nt identity			%aa identity		
	Between species		Within species	Between species		Within species
	min	max	min	min	max	min
Complete genomes						
Allexivirus	60.9	65.2	90.2			
Carlavirus	48.8	58.3	76.6			
Foveavirus	45.4	67.2	75.9			
Potexvirus	45.3	70.1	77.1			
Polymerase						
Allexivirus	63.0	69.5	89.9	66.9	81.7	93.4
Carlavirus	49.4	57.4	75.5	40.3	56.4	85.6
Foveavirus	45.6	66.8	75.6	36.4	71.2	87.8
Potexvirus	45.2	70.1	76.3	40.4	73.3	88.8
Coat protein						
Allexivirus	60.7	81.1	90.4	61.3	91.3	93.0
Carlavirus	43.3	71.1	74.7	29.7	76.1	86.8
Foveavirus	38.9	71.4	80.8	26.0	76.7	87.4
Potexvirus	37.3	75.0	77.3	20.5	79.8	74.4
ORF6						
Allexivirus	55.8	78.6	92.4	50.8	79.7	91.4
Carlavirus	34.2	63.2	68.1	11.5	61.9	69.9

To be consistent with the other genera, it therefore seems preferable to use the lower values (73% rather than 86% identical nucleotides, Fig. 3; 76% rather than 90% identical amino acids) as the guide for species discrimination in the genus *Allexivirus* and to amalgamate some of the species as described above. However, the large polymerase genes of allexiviruses are less closely related than the 3′-parts of the genome, and if such criteria were adopted using only the coat protein gene, the GarV-A and GarV-E sequences that would be allocated to the same species have only 64.8% nucleotides identical over their entire genomes and 69.0% identical amino acids between their polymerases. This inconsistency does not seem to occur to the same extent in the other three genera. A re-assessment of the criteria used for species demarcation in the genus *Allexivirus* is clearly necessary, preferably including biological data. This ambiguity suggests that it is best not to allocate new species names to the uncertain sequences described here but provisionally to allocate them to the nearest currently-recognised species (Groups 1–3 as GarV-A and Group 4 as GarV-X).

The degenerate primers Allex-CP(+) and Allex-NABP(−) have proved useful for detecting members of the genus *Allexivirus*. By sequencing the RT-PCR products, additional useful genome information has been readily obtained and this demonstrates the complexity of the pattern of variation amongst viruses in the genus.

Acknowledgements

The project was supported by the Zhejiang Provincial Natural Science Foundation of China (ZA0207), the 10th five year research programme of Zhejiang province (011102181), the National Natural Science Foundation of China for Distinguished Scholars (30225031) and for Young Scientists (30200008). Rothamsted Research receives grant-aided support from the Biotechnology and Biological Sciences Research Council of the United Kingdom.

References

1. Anon (2001) Wisconsin Package version 10.3. Accelrys Inc., San Diego, CA, USA
2. Chen J, Chen JP, Adams MJ (2001) A universal PCR primer to detect members of the *Potyviridae* and its use to examine the taxonomic status of several members of the family. Arch Virol 146: 757–766
3. Chen J, Chen JP, Adams MJ (2001) Molecular characterisation of a complex mixture of viruses in garlic with mosaic symptoms in China. Arch Virol 146: 1841–1853
4. Chen J, Chen JP, Adams MJ (2002) Characterisation of some carla- and potyviruses from bulb crops in China. Arch Virol 147: 419–428
5. Felsenstein J (1993) PHYLIP (Phylogeny Inference Package) version 3.6. Distributed by the author. Department of Genetics, University of Washington, Seattle
6. Kanyuka KV, Vishnichenko VK, Levay KE, Kondrikov DY, Ryabov EV, Zavriev SK (1992) Nucleotide sequence of shallot virus X RNA reveals a 5′-proximal cistron closely related to those of potexviruses and a unique arrangement of the 3′-proximal cistrons. J Gen Virol 73: 2553–2560
7. Ryabov EV, Generozov EV, Vetten HJ, Zavriev SK (1996) Analysis of the 3′-region of the mite born filamentous virus genome testifies its relation to the shallot virus X group. Mol Biol (Mosk) 30: 103–110

8. Song SL, Song JT, Kim CH, Lee JS, Choi YD (1998) Molecular characterization of the garlic virus X genome. J Gen Virol 79: 155–159
9. Sumi S, Matsumi T, Tsuneyoshi T (1999) Complete nucleotide sequences of garlic viruses A and C, members of the newly ratified genus *Allexivirus*. Arch Virol 144: 1819–1826
10. Sumi S, Tsuneyoshi T, Furutani H (1993) Novel rod-shaped viruses isolated from garlic, *Allium sativum*, possessing a unique genome organization. J Gen Virol 74: 1879–1885
11. Takaichi M, Nagakubo T, Oeda K (2001) Mixed virus infections of garlic determined by a multivalent polyclonal antiserum and virus effects on disease symptoms. Plant Dis 85: 71–75
12. Tsuneyoshi T, Matsumi T, Deng TC, Sako I, Sumi S (1998) Differentiation of *Allium* carlaviruses isolated from different parts of the world based on the viral coat protein sequence. Arch Virol 143: 1093–1107
13. Tsuneyoshi T, Matsumi T, Natsuaka KT, Sumi S (1998) Nucleotide sequence analysis of virus isolates indicates the presence of three potyvirus species in *Allium* plants. Arch Virol 143: 97–113
14. van Dijk P, van der Vlugt RA (1994) New mite-borne virus isolates from Rakkyo, shallot and wild leek species. Eur J Plant Path 100: 269–277
15. van Regenmortel MHV, Fauquet CM, Bishop DHL, Carstens EB, Estes MK, Lemon SM, Maniloff J, Mayo MA, McGeoch DJ, Pringle CR, Wickner RB (eds) (2000) Virus taxonomy: Seventh Report of the International Committee on Taxonomy of Viruses. Academic Press London, San Diego
16. Vishnichenko VK, Konareva TN, Zavriev SK (1993) A new filamentous virus in shallot. Plant Path 42: 121–126
17. Yamashita K, Sakai J, Hanada K (1996) Characterization of a new virus from garlic (*Allium sativum* L.), garlic mite-borne virus. Ann Phytopathol Soc Jpn 62: 483–489

Author's address: Prof. Jiong Chen, Virology and Biotechnology Department, Zhejiang Academy of Agricultural Sciences, Hangzhou, 310021, People's Republic of China; e-mail: jchen1@netease.com