ADVANCED TOPICS IN SCIENCE AND TECHNOLOGY IN CHINA

Jishuang Chen

Experimental Plant Virology

ZHEJIANG UNIVERSITY PRESS
浙江大学出版社

Springer

ADVANCED TOPICS
IN SCIENCE AND TECHNOLOGY IN CHINA

ADVANCED TOPICS
IN SCIENCE AND TECHNOLOGY IN CHINA

Zhejiang University is one of the leading universities in China. In Advanced Topics in Science and Technology in China, Zhejiang University Press and Springer jointly publish monographs by Chinese scholars and professors, as well as invited authors and editors from abroad who are outstanding experts and scholars in their fields. This series will be of interest to researchers, lecturers, and graduate students alike.

Advanced Topics in Science and Technology in China aims to present the latest and most cutting-edge theories, techniques, and methodologies in various research areas in China. It covers all disciplines in the fields of natural science and technology, including but not limited to, computer science, materials science, life sciences, engineering, environmental sciences, mathematics, and physics.

Jishuang Chen

Experimental Plant Virology

With 126 figures

ZHEJIANG UNIVERSITY PRESS
浙江大学出版社

Springer

Author
Dr. Jishuang Chen
Institute of Bioengineering
Zhejiang Sci-Tech University
Hangzhou 310029, China
E-mail: chenjs@zstu.edu.cn

ISSN 1995-6819 e-ISSN 1995-6827
Advanced Topics in Science and Technology in China

ISBN 978-7-308-07369-1
Zhejiang University Press, Hangzhou

ISBN 978-3-642-14118-8 e-ISBN 978-3-642-14119-5
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2010929389

Cover design: Frido Steinen-Broo, EStudio Calamar, Spain

Printed on acid-free paper

Springer is a part of Springer Science+Business Media (www.springer.com)

Preface

Plant Virology, the study of plant viruses and their diseases, is an important branch of science comprised of challenges and opportunities in China and developing countries, where agriculture is of key importance to date. The discovery of principles and the development of new techniques are helpful in feeding the population more safely and in providing more healthy systems of food production, along with many other aspects of social life such as ornamentals and natural medicine.

It is easy to show people fish, no matter how they are obtained, but it is more difficult to tell others how to fish, no matter how simple it becomes after years of practice. Based on the progress of biological studies over the last decades, molecular biology techniques have established a good foundation for realizing the investigation of viral genomes, which is the "stem" for plant virus exploration. Using *Cucumber mosaic virus* (CMV) with its satellite RNA and several other plant viruses with single stranded RNA genome as examples, some newly-obtained principles and the progress of the research are shown in this book, which is composed of six chapters. Chapters 1 to 5 mainly involve topics of genomic characterization, detection and quantitative techniques, especially the infection clone systems of CMV. Host responses to virus infection through plant microRNAs have also been demonstrated as groundwork. In Chapter 6, several plant cryptic viruses with double stranded virus genomes have been described for the first time, treating an understanding of plant viruses as a kind of bio-resource.

I hope this book will provide some practical clues and insight for people who work in this field and in related areas.

I warmly thank all the contributors who have worked with me in the same laboratory, during the year 2000 to 2009, including Ph.D candidates, Junli Feng, Zhiyou Du, Qiansheng Liao, Liqiang Li, Shaoning Chen and Qiulei Lang, with Master Students Liang Cheng, Yanfei chen, Rong Zeng, Jianguang Zhang, Zuodong Qin, Liping Zhu, Qinghua Tian, Hong Guo, Shijie Yan and Susu Shentu.

Dr. Jishuang
Hangzhou, China
December, 2009

Contents

1

Gene Cloning of *Cucumber Mosaic Virus* and Some Related Viral Agents

1.1 Introduction

Cucumber mosaic virus (CMV) is a typical member of the genus *Cucumovirus*. It has infected more than 1,000 species of monocots and dicots, including many economically important crops (Palukaitis and García-Arenal, 2003; Palukaitis et al., 1992). In China, CMV is commonly detected as the principal virus infecting field crops in the families *Solanaceae* (including tobacco, tomato, potato, pepper, etc.), *Brassicaceae* (including brassicas, radish, turnip, etc.) and *Fabaceae* (including soybean, cowpea, etc.). As shown in Fig. 1.1, CMV strain containing a satellite RNA co-infected with *Tomato mosaic virus* (ToMV), brought fruit necrosis and killed off the whole plant when the temperature was high. CMV infection in early spring used to bring a major lost of radishes and other cruciferous crops.

Fig. 1.1. Symptoms caused by infection of *Cucumber mosaic virus* in the field
(a) Field tomato plant complexly infected by *Cucumber mosaic virus* with a satellite RNA and a strain of *Tomato mosaic virus*; (b) Field radish plant infected by CMV

CMV can easily transmit mechanically to a wide range of plants. This characteristic makes it easier to do more research for virus-host interactions, and also for virus-virus interactions. Typical symptoms induced by CMV are supplied in Fig. 1.2. The rapid replication and high accumulation in leaf tissues of systemic hosts provide another advantage for genomic and quantitative studies. As shown in Fig. 1.3, CMV particles reach a high accumulation condition within four days of inoculation in cells of the inoculated leaf. And the potyvirus infection also shows distinguished characteristics.

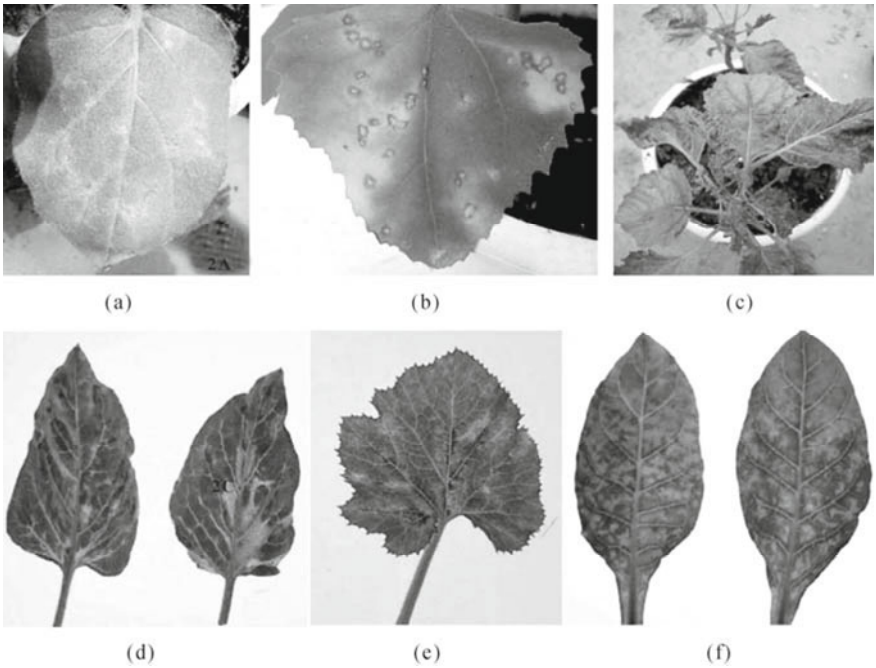


Fig. 1.2. Typical symptoms caused by inoculation of *Cucumber mosaic virus*
(a) *Nicotiana glutinosa*, presenting irregular yellowing spots on the inoculated leaves; (b) *Chenopodiu amaranticolor*, presenting local lesions on the inoculated leaves; (c) *Nicandra physalodes*, presenting systemic mosaic and distortion the lower leaves inoculated dropped; (d) *N. glutinosa*, presenting systemic mosaic; (e) *Lagenaria siceraia*, presenting systemic mosaic; (f) *N. tobacum* (cv. HuangMiaoYu), presenting systemic mosaic

Major differences are also characterized for CMV and the co-infection potyvirus, for CMV is more a hot time virus occurring in seasons with higher temperature, whilst potyvirus such as TuMV is more likely to occur in cool seasons (Table 1.1).

In the same ecological position and transmitted via similar methods (both by aphids and by mechanical transmission), CMV and potyvirus are found to have infected the same crops and express typical mosaic symptoms. It could be considered that the two kinds of viruses can be evaluated together or with similar mechanisms. The morphological characteristics are presented in Fig. 1.4.

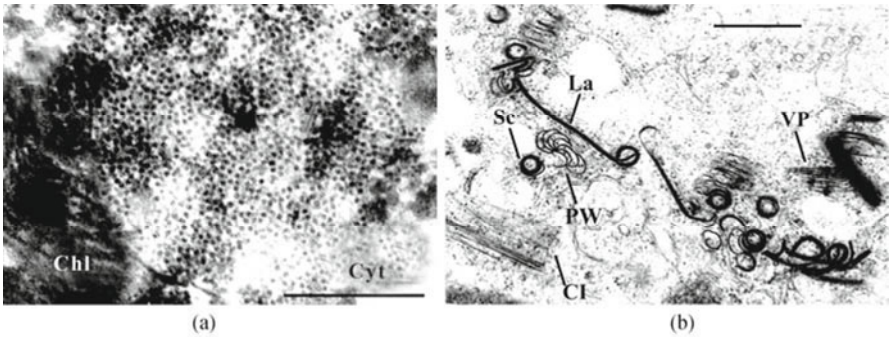


Fig. 1.3. Cytoplasmic alteration of the host tissues by infection of *Cucumber mosaic virus* and *Turnip mosaic virus* (TuMV)
(a) Cytoplast structure of *N. glutinosa* (inoculated leaf 4 days post inoculation) infected by *Cucumber mosaic virus*, presenting numerous spherical virus particles in the cytoplasm and the remaining chloroplast layers (Chl: chloroplast layer; Cyt: cytoplasm), bar = 750 nm; (b) Cell structure of *Brassica compestris* ssp. chinensis infected with TuMV, presenting different structures of cylindrical inclusion body and aggregated filamentous virus particles (PW: pinwheel structure; CI: cylindrical structure; Sc: scroll-like structure; La: laminated aggregates; VP: virus particle), bar = 600 nm

Table 1.1 Seasonal occurrence in frequencies of principal viruses infecting cruciferous crops

Seasons	CMV	TuMV	CMV +TuMV	CMV (%)	TuMV (%)	CMV/TuMV	Total isolate obtained
Spring (March–May)	12	22	5	44.0	69.2	0.63	39
Summer (June–August)	5	5	0	50.0	50.0	1.00	10
Autumn (September–November)	24	9	5	72.5	35.0	2.10	40
Winter (December–February)	42	51	7	43.3	51.3	0.85	113

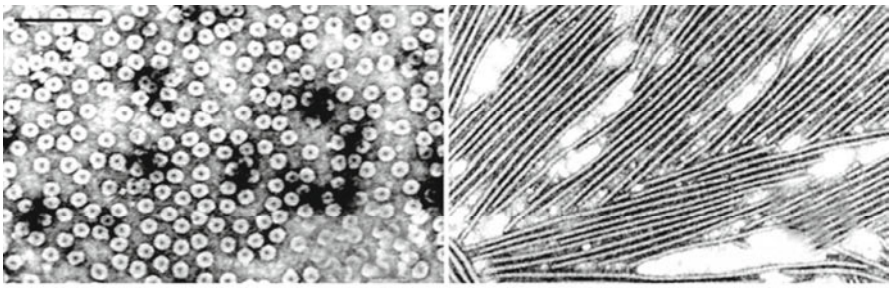


Fig. 1.4. Morphological characteristics of *Cucumber mosaic virus* and potyvirus infecting crucifers
(a) Virus particles of *Cucumber mosaic virus*, bar = 150 nm; (b) Virus particles of Turnip mosaic virus, bar = 350 nm

The survey was done during 1997 and 2001, in Zhejiang Province, eastern China.

In comparison with conventional techniques, viral genome sequence analysis is the most direct and one of the most valid methods to date. As a tripartite RNA virus, CMV contains three capped single-stranded positive-sense genomic RNAs named RNA1, RNA2 and RNA3. RNA1 and RNA2 encode 1a protein and 2a protein respectively, which are involved in virus replication (Hayes and Buck, 1990). RNA3 encodes 3a protein (MP) and coat protein (CP). 3a protein is responsible for virus movement (Ding et al., 1995; Li et al., 2001). CP is translated via a subgenomic RNA4 and involved in virus movement and aphid-mediated transmission (Kaplan et al., 1998; Perry et al., 1994). In addition, 2b protein encoded by subgenomic RNA4A via RNA2 functions in long-distance movement and as a post-transcriptional gene silencing suppressor (Brigneti et al., 1998; Ding et al., 1994; 1995).

A major advantage in using CMV as model ssRNA virus is that it has high copies of double stranded RNAs of the full-length genome. The dsRNA segments are easily extracted and analyzed, and can be regarded as replication forms, as shown in Fig. 1.5.

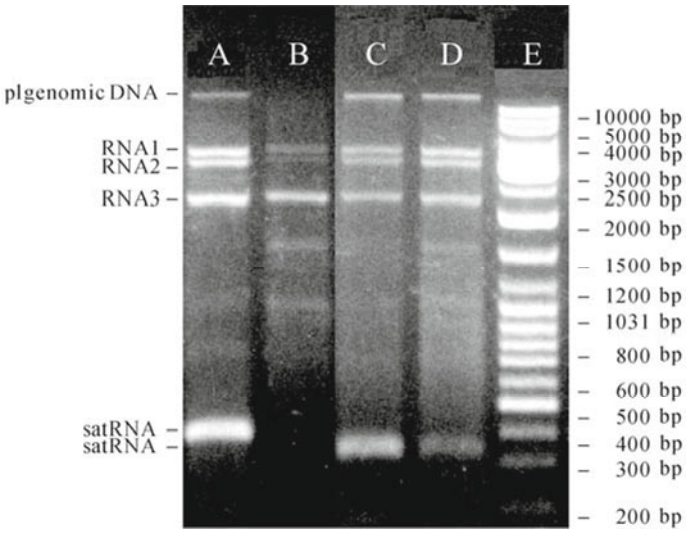


Fig. 1.5. Patterns of dsRNA analysis of several *Cucumber mosaic virus* isolates
The dsRNA pattern shows relative copy numbers of carried satellite RNA and the genome RNAs

When determining the symptom development by host, coat protein is involved in quite a few functions. It is highly expressed via subgenomic RNA4. To outline the fine functional domain and related sequence of CMV RNA3 subgenomic promoter region (SgPr) as a region for regulation between 3a protein and CP sequence, the sequences for SgPr are determined and compared to the reported sequences for CMV RNA3 with different origins. Among the CMV isolates compared, SgPrs are found to consist of 284–323 nt, varying among the isolates.

The SgPr sequence for subgroup I is obviously different from that of subgroup II because of a sequence similarity of <70% between them. This indicates that the SgPr region may have additional biological significance and that the SgPr may have no direct relationship with the presence or absence of a satRNA in CMV (Fig. 1.6).

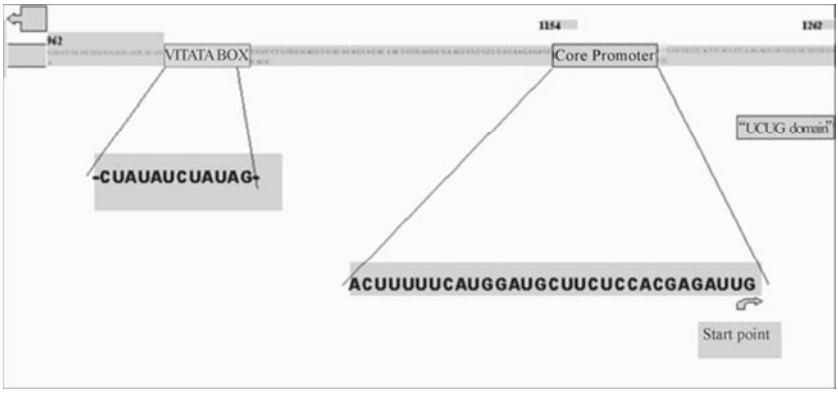


Fig. 1.6. Fine structure and functional motifs of RNA3 subgenomic core promoter region of *Cucumber mosaic virus*

Introduced in this chapter, the evolution mechanism between CMV and the co-infecting potyvirus and the evolution mechanism between CMV and its satellite RNAs (satRNAs) could be different from each other. CMV, as a multipartite ssRNA virus, pseudo-recombination is considered frequently, while potyvirus infecting new host plants seems to utilize a combination among and within gene motifs. For example, the new CMV strain infecting the tomato is discovered as a natural reassortment between the strain of the same virus, but to fit the infection to *Pinellia ternata* from the *Araceae* family, a potyvirus jointly coming from *Soybean mosaic virus* (SMV) and *Watermelon mosaic virus* (WMV). At the same time, an unusual strain of CMV infects *P. ternata* after a period of fitness to vegetative propagation of the host plant displays site mutation and deletion (insertion) mechanisms, with some independent evolution at their UTR terminus (Figs. 1.3 and 1.4). It is believed that the host and geographical environment had an impact on evolutionary types of this virus.

As for satRNA of CMV, the secondary structures are essential for replication and stability. Changes at a single base could influence survival or interaction.

1.2 A Tomato Strain of *Cucumber Mosaic Virus*, a Natural Reassortant Between Subgroups IA and II

According to serological relationships and nucleic acid identities, CMV isolates have been classified into two main subgroups, namely subgroup I and subgroup II (Palukaitis et al., 1992). The analysis of a larger numbers of CP genes and 5′

non-translated regions of CMV isolates' RNA3 has led to a further division of subgroup I into subgroups IA and IB (Roossinck et al., 1999). The nucleotide sequence identity between CMV subgroups I and II strains ranges from 69% to 77%, while it is above 90% within a subgroup (Palukaitis et al., 1992). CMV strains of subgroups IA and II have been reported from most parts of the world, while subgroup IB strains are considered to be mainly restricted to Asia (Roossinck, 2002). Considering the tripartite nature of the CMV genome, reassortment is one of the mechanisms for genetic variation and new strain generation of multipartite RNA viruses (Chao, 1997). Reassortment of multipartite RNA viruses has been displayed for many animal viruses and plant viruses, such as the influenza virus (McCullers et al., 1999) and tobravirus (Robinson et al., 1987). Among cucumoviruses, an interspecific reassortant, composed of CMV RNA3 and *Peanut stunt virus* (PSV) RNAs 1 and 2, and an intraspecific reassortant of PSV, have been discovered (Hu and Ghabrial, 1998; White et al., 1995). Studies of natural CMV populations have showed that mixed infection by different CMV strains is frequent and genetic exchange by reassortment occurred (Bonnet et al., 2005; Fraile et al., 1997).

However, natural reassortants between CMV subgroups and strains should survive against selection and could not become established as dominating populations before a favorite condition appears. Furthermore, reassortment does not occur randomly. The fraction of reassortants between CMV subgroups IA and IB is found to be larger than that of reassortants between subgroups I and II. Before, only one naturally occurring reassortant between CMV subgroups I and II strains was found by Bonnet et al. (2005).

A CMV strain, represented as an isolate, namely CMV-Tsh has been detected for its wide distribution in a tomato field in Shanghai, emerging in spring 2005. This isolate is found to be a natural reassortant between subgroups I and II based on sequence analysis.

Based on biological inoculation, virus isolation, serological identification and, especially, double stranded RNA analysis, the existence of CMV with a satRNA is found to co-exist with ToMV to cause severe systemic mosaic and necrosis synergy. After gene cloning with full-length cDNA amplified with primer pairs against all the subgroups I and II strains, the genomic sequences of this CMV are obtained. The full length RNA1 is obtained by cloning two RT-PCR products respectively. The primers used for amplification CMV genomic RNAs are listed in Table 1.2. The full length sequences of CMV-Tsh RNA1, RNA2 and RNA3 have been submitted to GenBank under the accession number EF202595, EF202596 and EF202597, respectively. CMV-Tsh RNA1 is found to consist of 3,394 nucleotides (nt), encoding 1a protein of 994 amino acids from 96 to 3,077 nt. RNA2 is consisted of 3,047 nt, containing two partially overlapped ORFs 2a and 2b. The 2a ORF encoding 2a protein of 858 amino acids extends from 86 to 2659 nt, and the 2b ORF is positioned at the sequence from 2,418 to 2,750 nt, encoding 2b protein of 111 amino acids. RNA3 contains 2,206 nt, encoding 3a protein of 280 amino acids and CP of 219 amino acids, corresponding to the sequences from 97 to 936 nt, and 1,229 to 1,885 nt respectively.

Table 1.2 Primers for RT-PCR amplification of CMV genomic RNAs

Primer[a]	Positions[b]	Nucleotide sequence[c]	Enzyme site
RNA1-F	5′ end of RNA1	5′-AATCGGATCC**TAATACGACTCACTATA G**GTTTTATTTACAAGAGCGTA-3′	*Bam*HI
RNA1-R	3′ end of gRNAs	5′-AATTGTCGACTGGTCTCCTT-3′	*Sal*I
RNA2-F	5′ end of RNA2	5′-AATCGGATCC**TAATACGACTCACTATA G**GTTTATTYWCAAGAGCGTA-3′	*Bam*HI
RNA3-F	5′ end of RNA3	5′-AATCGGATCC**TAATACGACTCACTATA G**GTAATCTTACCACT-3′	*Bam*HI
RNA23-R	3′ end of gRNAs	5′-AATTCTGCAGTGGTCTCCTT-3′	*Pst*I
RNA1-1750-R	1607–1623 nt	5′-AATTGTCGACGATGATATCACGTCCCA-3′	*Sal*I
RNA1-1600-F	1607–1623 nt	5′-AATCGGATCCTGGGACGTGATATCATC-3′	*Bam*HI

[a]: "F" represent forward primer, "R" represent reverse primers. All primers are useful against all CMV subgroups; [b]: combining area (for against CMV-Fny), "gRNAs" represent all genomic RNA; [c]: Underlined are sites for restriction enzymes, blocked are T7 promoter sequences, Y= C or T, W= A or T

The sequence comparison results between CMV-Tsh and other strains from subgroups I and II are shown in Table 1.3. RNA1 and 1a ORF of CMV-Tsh show 97.5% nucleotide sequence identity with those of the strain Q and less than 78% with those of the two strains CMV-Fny and CMV-Sd, which are obtained from Shandong, China. These results revealed that CMV-Tsh RNA1 is derived from a CMV subgroup II strain. Comparisons of the nucleotide sequences of CMV-Tsh 2a, 2b and RNA2 to those of CMV-Fny, CMV-Sd and Q strain revealed that CMV-Tsh is more closely related to the CMV-Fny (with over 96% sequence identity) than to CMV-Sd (less than 91% identity) and Q strain (less than 73% identity). So, it is most likely that CMV-Tsh RNA2 is derived from a CMV subgroup IA strain. The nucleotide sequences of CMV-Tsh RNA3, 3a and CP ORF show over 98% sequence identity with those of the strain Q and less than 80% sequence identity with those of the CMV-Fny and CMV-Sd. These results indicated that CMV-Tsh RNA3 is derived from a CMV subgroup II strain. The same classification of CMV-Tsh RNAs 1, 2 and 3 are observed by comparing the deduced amino acid sequences of five ORFs.

The results of phylogenetic analysis for five ORFs between CMV-Tsh and another 15 CMV strains are shown in Fig. 1.7. For *1a*, *3a* and *CP* genes, CMV-Tsh forms an independent clade with subgroup II strains with supporting values of 100%, while 2a and 2b ORFs of CMV-Tsh form a clade with subgroup IA strains, supported by >98% bootstrap values. These results are similar to those obtained from the sequence comparisons.

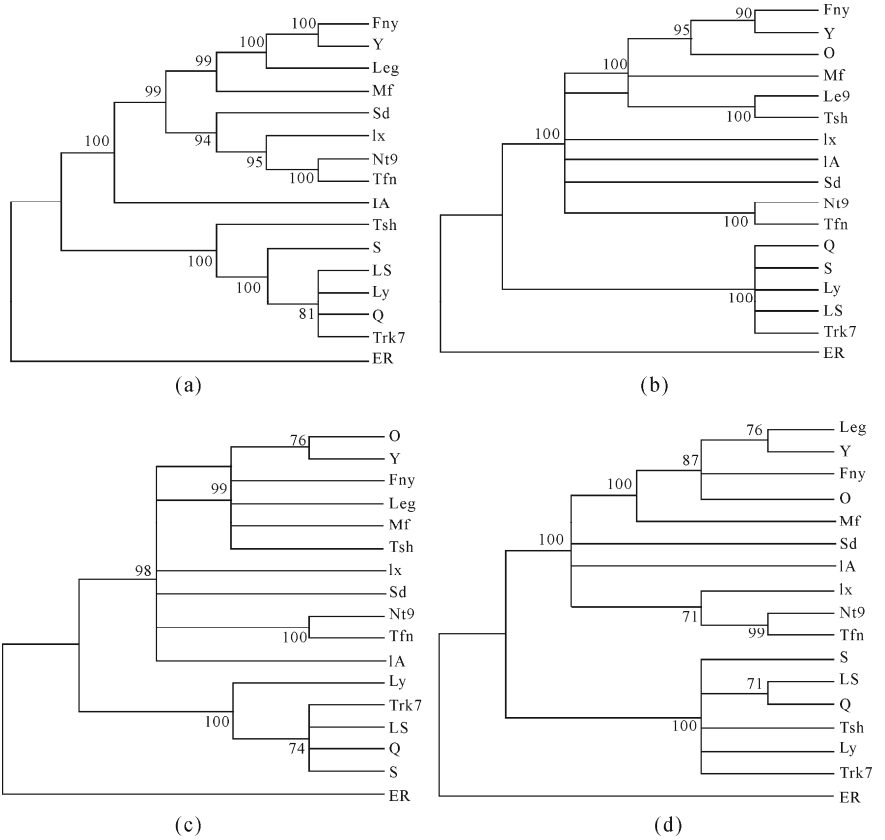
Using a restriction enzyme analysis, firstly described by Rizos et al. (see Chapter 2), RT-PCR products obtained from 15 field tomato samples collected across 2006–2008 in Shanghai has shown the same results of occurrence for

reassortant between CMV subgroups IA and II strains. It is thus judged that this new strain of CMV has become stable in certain tomato varieties in this area.

Table 1.3 Sequence comparisons of CMV-Tsh genomic RNAs and deduced amino acids sequences with those of subgroups I and II strains

RNAs	Identity of nucleotide sequence between CMV-Tsh and other CMV strains (%)			Identity of deduced amino acid sequences between CMV-Tsh and other CMV strains (%)		
	CMV-Fny[a]	CMV-Sd	Q strain	CMV-Fny	CMV-Sd	Q strain
1a	77.9	77.8	97.5	85.5	85.4	97.6
2a	96.1	91.2	72.4	97.1	92.9	72.7
2b	96.1	86.9	64.0	96.4	81.1	48.6
3a	78.9	78.9	98.5	83.9	83.6	98.7
CP	76.9	77.8	99.2	82.6	82.6	100
RNA1	76.9	76.6	97.5			
RNA2	96.0	90.4	71.2			
RNA3	74.0	74.3	98.3			

[a]CMV-Fny, CMV-Sd and Q strains are used as representative for subgroups IA, IB and II respectively. The GenBank accession numbers are given in Fig. 1.7



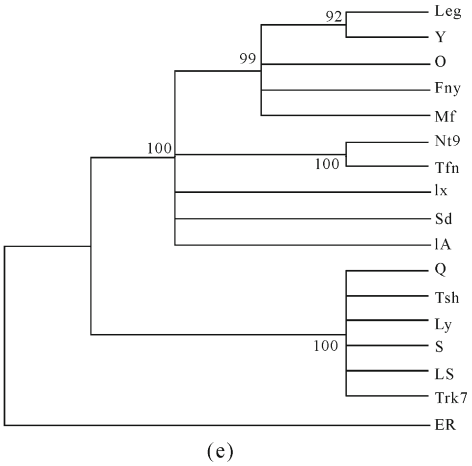


Fig. 1.7. Bootstrap majority rule consensus trees of the five ORFs of selected CMV strains ORFs are listed as (a) 1a; (b) 2a; (c) 2b; (d) 3a; (e) CP. Bootstrap percentage values are placed at major nodes. Fny, Leg, Mf, Y and O are subgroup IA strains; Nt9, Tfn, Ix, Sd and IA are subgroup IB strains; Q, Ly, S, LS and Trk7 are subgroup II strains. PSV ER strain is used as an outgroup. The GenBank accession numbers of nucleotide sequences used here are as follows: CMV-Fny (D00356, D00355, D10538); Leg (D16403, D16406, D16405); Mf (AJ276479, AJ276480, AJ276481); Y (D12537, D12538, D12499); O (*, D10209, D00385); Nt9 (D28778, D28779, D28780); Tfn (Y16924, Y16925, Y16926); Ix (U20220, U20218, U20219); Sd (AF071551, D86330, AB008777); IA (AB042292, AB042293, AB042294); Q (X02733, X00985, M21464); Ly (AF198101, AF198102, AF198103); S (Y10884, Y10885, U37227 and AF063610); LS (AF416899, AF416900, AF4127976); Trk7 (AJ007933, AJ007934, L15336); ER (U15728, U15729, U15730)

Reassortment has been proposed as an important mechanism in the evolution of the RNA virus with divided genomes. Reassortment would offset the fitness losses induced by deleterious mutations of nucleotides and recombination of viral genes (Henderson et al., 1995). CMV has been successful in adapting to different hosts and environments, leading to an extremely large host range and a worldwide distribution. Bonnet et al. (2005) analyzed the role of recombination and reassortment in the evolution of CMV with 159 filed Spanish CMV isolates collected from 1989 to 2002. According to their results, only 5% of isolates were reassortants between CMV subgroups. Amongst them, only one reassortant between CMV subgroups I and II was detected. This suggested that the occurrence of natural reassortants between strains of CMV subgroups IA and IB is more frequent than those between subgroups I and II. This phenomenon can be interpreted by the following reasons. First, CMV subgroup I strains have a higher incidence (Lin et al., 2003) and more rapid viral accumulation (Wang et al., 2002) compared with subgroup II strains. Thus, CMV subgroup I strains have competitive advantages in mixed infection. Second, the thermal optima of CMV subgroup I strains are much higher than those of subgroup II strains (Fraile et al., 1997). The resultant possibility of mixed infections between subgroups IA and IB is much higher than that between subgroups I and II. In

addition, the high nucleotide sequence divergence between CMV subgroups I and II strains is one of the important influencing factors that induce a low possibility of forming viable natural reassortants between subgroups I and II.

Although some studies have revealed that the reassortment between CMV subgroups is a rare event, it does not mean that reassortment is not important in CMV evolution. The phylogenetic analysis of 15 CMV strains have shown that reassortment had led to the high genetic diversity and evolutionary success of CMV (Roossinck, 2002). CMV-Tsh is found to be a natural reassortant between CMV subgroups IA and II strains. The sequence analysis and restriction enzyme analysis of CMV-Tsh genomic RNAs demonstrate that RNAs 1 and 3 of CMV-Tsh are derived from one or two subgroup II strain(s), while RNA2 is derived from a subgroup I strain. Furthermore, the restriction pattern of the CMV-Tsh-infected tomato plant is the same as the other five tomato plants sampled from the same planting area. It is suggested that the infection of CMV-Tsh occurred frequently in this planting area. It could be hypothesized that CMV-Tsh might be derived from a mixed infection by CMV subgroups IA and II strains, and it is likely that the subgroup II RNA2 has been washed out by the subgroup IA RNA2 because of its low efficiency in inhibiting host responses during the mixed infection of parental viruses of CMV-Tsh. In addition, a ToMV isolate and a CMV satRNA are found to co-exist with CMV-Tsh in the diseased tomato plants from Shanghai, China. Some studies have found that other plant viruses and CMV satRNAs may change the accumulation levels of different genomic RNAs of CMV in the co-infected plants (Palukaitis et al., 1992; Poolpol, 1986). Wang et al. found that Fny-CMV (subgroup IA strain) and LS-CMV (subgroup II strain) showed obviously different changes in accumulation profiles of viral RNAs, while each virus co-infected with *Zucchini yellow mosaic virus* (ZYMV) in zucchini squash and similar results are to be introduced in Chapter 3 in this book. It is also possible that ToMV and CMV satRNA co-infected with CMV-Tsh give different selection pressures on the genomic RNAs of parental strains of CMV-Tsh and bring on the occurrence of CMV-Tsh.

As a widely distributed plant virus with tripartite ssRNA genome, CMV is chosen as a good model virus for studying its mutation, detection, qualification and interaction with host plants.

1.3 The *Araceae* Strain of *Cucumber Mosaic Virus* Infecting *Pinellia ternate* Suggested to be a Novel Class Unit Under Subgroup I

As a kind of traditional Chinese medicinal plant, *Pinellia ternata* has been used for thousands of years. It has been cultivated since the end of the 1970s, and it is propagated mostly in a vegetative manner. This cultivated plant is found to be commonly infected by viral diseases which have been found recently with a new

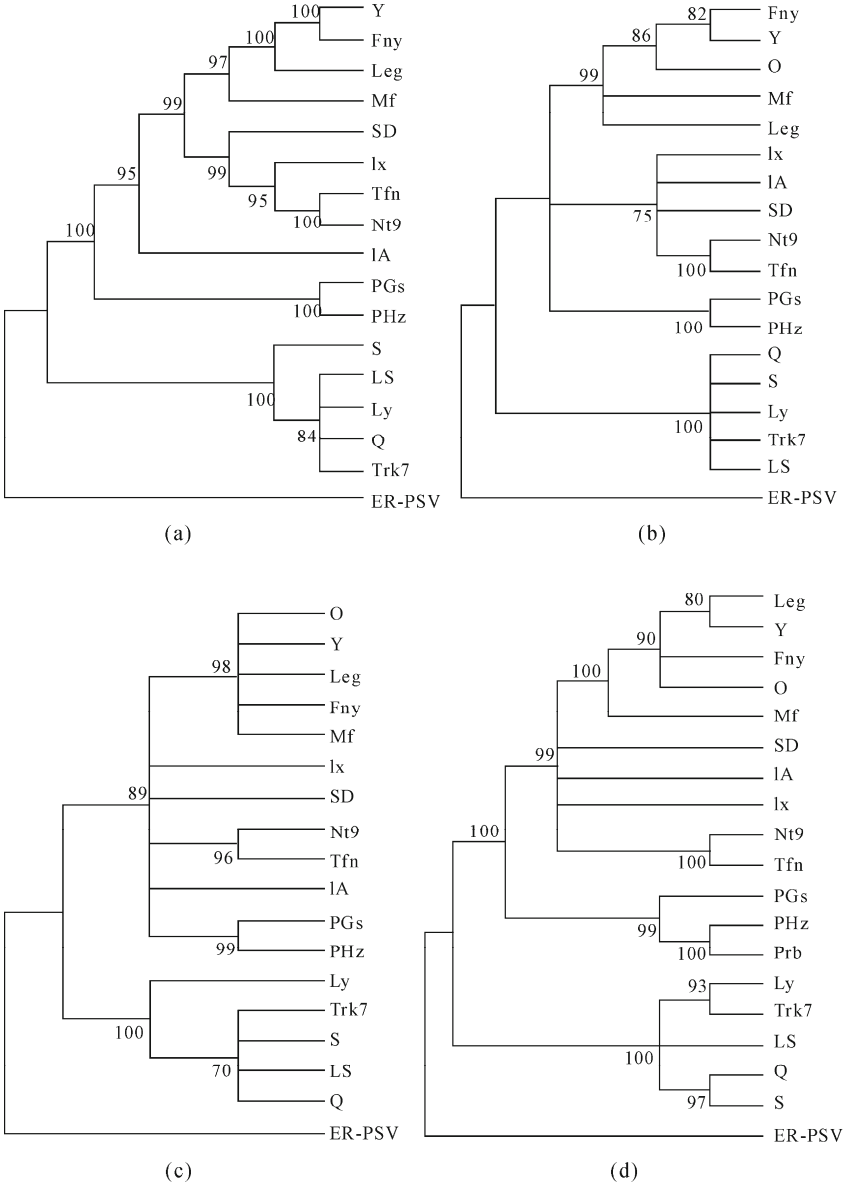
strain of *Cucumber mosaic virus* and a potyvirus. At the same time, the above viruses are seldom found in wild plants. Depending on dsRNA analysis and RT-PCR amplification, combined with full-length cDNA cloning by modified single-primer amplification technique (SPAT, see description in Chapter 6), full genomes of the newly discovered CMV strain are determined. Based on the serological and biological characteristics, the new strain of CMV limited to *Araceae* but not transported mechanically to *Solaniouse* species is considered as subgroup IC strain. Two isolates, namely CMV-PHz isolated from *P. ternata* grown in Hangzhou, eastern China, and CMV-PGs isolated from *P. ternata* grown in Gansu Province, northwest China, are sequenced for genomic phylogenetice and sequence divergence analysis with known CMV strains. The two isolates were obtained during 2005-2008. Partial sequence of other isolates, CMV-PNb isolated from cultivated *P. ternata* grown in Ningbo, eastern China, in 2003 has also been obtained and compared. The accession numbers of full length sequences for CMV-PHz RNA1, RNA2 and RNA3 are EU723568, EU723570 and EU723569, respectively. The accession numbers of the full length sequences of CMV-PGs RNA1, RNA2 and RNA3 are DQ399548, DQ399549 and DQ399550, respectively. The sizes of genomic RNAs and positions for each ORF of consensus CMV I, consensus CMV II, CMV-PGs, PHz and CMV-PNb are shown in Table 1.4.

Table 1.4 Size of genomic RNAs and position of each ORF in the CMV (CMV I, CMV II, CMV-PGs, CMV-PHz and CMV-PNb)

	CMV I	CMV II	CMV-PGs	CMV-PHz	CMV-PNb
RNA1	3357—3365	3389—3391	3336	3346	
1a ORF	95—98 to 3076—3079	96—98 to 3073—3078	86—3067	95—3076	
RNA2	3036—3060	3038—3053	3037	3037	
2a ORF	78—79 to 2652—2673	93 to 2612—2615	76—2652	76—2652	
2b ORF	2414—2432 to 2746—2836	2409—2413 to 2712—2715	2411—2743	2411—2743	
RNA3	2213—2220	2197—2209	2179	2180	2179
3a ORF	120—123 to 959—973	96—97 to 935—936	98—937	96—935	96—935
CP ORF	1255—1263 to 1911—1918	1220—1232 to 1876—1888	1232—1888	1237—1893	1237—1893

Phylogenetice and sequence divergence analysis of 1a ORF: The products of 1a ORF have been found to be related for determining the host range of CMV strains. Phylogenetice analysis shows that CMV-PHz and CMV-PGs could be clustered into a single clade with 100% supporting values apart from other strains of subgroup I, with the rest forming their own clusters (Fig. 1.8(a)). The *1a* gene of CMV-PHz and CMV-PGs shows 12.0%–13.2% and 11.2%–12.8% nucleotide divergence with subgroup IA strains, respectively. They have 13.2% to 14.1% and 13.3%–14.1% nucleotide divergence with subgroup IB strains respectively. In

addition, they are 26.2%–27.0% and 26.8%–27.5% divergent when they are compared pairwise with subgroup II strains (Table 1.5). The pairwise comparison results between subgroups of 1a divergence with remarkable difference are a/b/c models for both CMV-PHz and CMV-PGs according to subgroup II/IB/IA (Table 1.6).



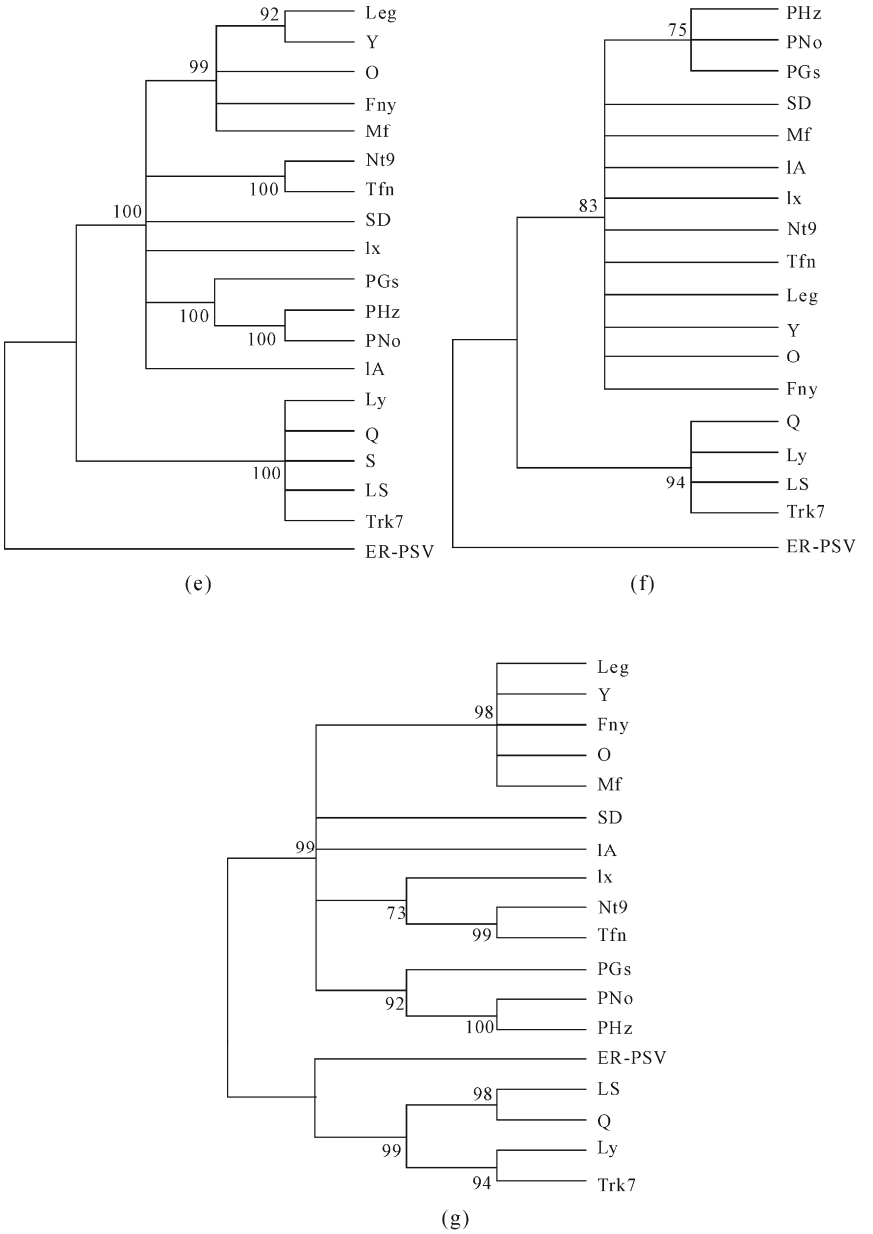


Fig. 1.8. Phylogenetic trees of nucleotide sequences of the five ORFs of isolates CMV-PGs and CMV-PHz and other CMV isolates as references

(a) 1a; (b) 2a; (c) 2b; (d) 3a; (e) CP; (f): 5′ UTR of RNA3; (g) 3′ UTR of RNA3. The dendrograms were conducted by NJ with 1,000 bootstrap replications. Bootstrap scores exceeding 70% are placed at major nodes. ER-PSV was used as an outgroup. The GenBank accession numbers of referenced strains are listed in Table 1.4

Table 1.5 The sequence divergence between different CMV isolates and subgroups. The ORFs or 5′ UTR or 3′ UTR of RNA3 are compared using Kimura's 2-parameter model, with minimum and maximum percentage numbers of different cases as shown

RNA	PHz/ PGs	PHz/ PNb	PNb/ PGs	PHz/ IA	PHz/ IB	PHz/ II	PGs/ IA	PGs/ IB	PGs/ II
1a	5.7			12.0— 13.2	13.2— 14.1	26.2— 27.0	11.2— 12.8	13.3— 14.1	26.8— 27.5
2a	3.4			11.1— 11.7	12.7— 14.5	34.5— 35.0	11.2— 12.0	12.5— 14.1	33.5— 33.9
2b	5.1			21.8— 26.0	20.2— 24.1	55.3— 58.0	21.4— 28.5	19.2— 23	55.5— 58.2
3a	3.6	0.4	3.7	10.3— 11.6	9.5— 11.2	21.7— 22	10.6— 11.8	10.1— 11.3	22.6— 22.9
CP	3.9	0.3	4.2	8.1— 9.6	8.5— 10.8	29.1— 30.7	9.0— 9.8	7.8— 10.1	28.6— 30.2
5′ UTR	1.4	0	1.4	4.2— 13.3	5.6— 7.1	49.8— 53.0	2.8— 11.7	4.2— 5.6	53.0— 56.5
3′ UTR	9.8	0.4	9.3	14.8— 16.4	15.4— 20.2	36.5— 38.8	10.6— 12.1	12.1— 17	34.3— 36.4

Table 1.6 Comparison of the pairwise comparison results between different CMV isolates and subgroups. The ORFs or 5′ UTR or 3′ UTR of RNA3 with remarkable difference are compared

PHz	1a	2a	2b	3a	CP	5′ UTR	3′ UTR	PGs	1a	2a	2b	3a	CP	5′UTR	3′UTR
II	a	a	a	a	a	a	a	II	a	a	a	a	A	a	a
IB	b	b	b	b	b	b	b	IB	b	b	b	b	B	b	b
IA	c	c	b	b	b	b	c	IA	c	c	b	b	B	c	c

The phylogenetic and sequence divergence analysis of 2a and 2b ORFs: As partially involved in virus replication, 2a phylogeny analysis shows that CMV-PHz and CMV-PGs are grouped in a separate branch while IA and IB form their own clade within subgroup I, radically. Subgroup II forms a single group (Fig. 1.8(b)). The *2a* gene of CMV-PHz and CMV-PGs has nucleotide divergence of 11.1% to 11.7% and 11.2% to 12.0% with subgroup IA, 12.7% to 14.5% and 12.5% to 14.1% with subgroup IB, 34.5% to 35.0% and 33.5% to 33.9% with subgroup II, respectively (Table 1.5). When CMV-PHz or CMV-PGs are pairwise compared with other subgroup strains, remarkable differences in the pairwise comparison between different subgroups are both a/b/c patterns, subgroup II, subgroup IB and subgroup IA, respectively. It is also shown in Table 1.6. As a newly discovered gene of CMV for determining symptom development, the 2b phylogeny analysis shows a radial pattern. They are clustered into a single clade within subgroup IB with high values (Fig. 1.8(c)). The 2b ORF shows the highest divergence with other subgroups. The nucleotide divergence for CMV-PHz and CMV-PGs sequences taken in pairs ranged from 21.8% to 26.0% and 21.4% to 28.5% with subgroup IA, 20.0% to 24.1% and 19.2% to 23.0% with subgroup IB, 55.3% to 58% and 55.5% to 58.2% with subgroup II, respectively. Remarkable difference models in both of them are a/b/b corresponding to subgroup II/IB/IA.

1.3.1 Phylogenetice and Sequence Divergence Analysis of 3a and CP ORFs

Phylogenetice analysis of 3a shows that more branches and compact trees within the groups, CMV-PHz, CMV-PGs and CMV-PNb (DQ512476) are clustered into a separate branch away from subgroup IA and IB (Fig. 1.8(d)). The 3a sequence divergence between CMV-PHz or CMV-PGs and other subgroups is found to range from 10.3% to 11.6% and 10.6% to 11.8% with subgroup IA, 9.5% to 11.2% and 10.1% to 11.3% with subgroup IB, 21.7% to 22% and 22.6% to 22.9% with subgroup II, respectively. Corresponding remarkable differences are both a/b/b patterns for subgroup II/IB/IA. As the most important determinator of the host range and symptom expression, the CP phylogeny analysis is quite similar to that of 3a, but absolutely different in the degree of the branch, and the three strains form a single clade within subgroup IB with high bootstrap scores (Fig. 1.8(e)).CP of CMV-PHz and CMV-PGs displays the degree of divergence from other subgroups: 8.1% to 9.6% and 9.0% to 9.8% with subgroup IA, 8.5% to 10.8% and 7.8% to 10.1% with subgroup IB, 29.1% to 30.7% and 28.6% to 30.2% with subgroup II, separately. The according remarkable difference patterns are both a/b/b models for subgroup II/IB/IA, as shown in Table 1.6.

1.3.2 Phylogenetice and Sequence Divergence Analysis of 5' UTR and 3' UTR, 2a and 2b ORFs of RNA3

In the tree of 5′ UTR it is hard to distinguish subgroup IA from IB; they are congregated in a whole clade, and the three *Pinellia* isolates are a branch within the group (Fig. 1.8(f)). The pairwise nucleotide-sequence divergence of 5′UTR of CMV-PHz and CMV-PGs RNA3 ranged from 4.2% to 13.3% and 2.8% to 11.7% with subgroup IA, 5.6% to 7.1% and 4.2% to 5.65% with subgroup IB, 49.8% to 54.0% and 53.0% to 56.5% with subgroup II, respectively (Table 1.5). The remarkable difference of CMV-PHz is different from CMV-PGs; a/b/b is for CMV-PHz; a/b/c is for CMV-PGs (Table 1.6). From the structure of their RNA3, it is found that the positions of RNA3 ORFs of *Pinellia* isolates are different from other documented subgroup I strains with about 25 base regions deleted. The "shorter" 5′ UTR structure is more similar to the position of subgroup II (Fig. 1.9). Thus, we consider the results for the alignment and regard the three *Pinellia* isolates as some intermediate between CMV subgroup I and subgroup II, since consensus sequences of IA, IB, II and RIB can be considered to be the closest ancestral 5′ UTR (Roossinck et al., 1999). The alignment can be divided into 7 motifs (A, B, C, D1, D2, E and F) as in (Roossinck et al., 1999). In their reports, boxes A and F are conserved in all strains, and are also found from CMV-PHz, CMV-PNb and CMV-PGs. Boxes B, C and E are found to vary in different subgroups, the motifs analysis supported the same conclusion, with CMV-PHz, CMV-PNb and CMV-PGs containing boxes B and C, but imperfect E. Only box D1 was found in CMV-PHz, CMV-PNb and CMV-PGs. However, box D is a fore-and-aft repeat in subgroups I, D1 and D2, respectively. Besides, CMV-PGs