Papaya ringspot potyvirus: isolate variability and the origin of PRSV type P (Australia)

Marion F. Bateson,1 Juliane Henderson,1 Worawan Chaleeprom,1 Adrian J. Gibbs2 and James L. Dale1*

1 Centre for Molecular Biotechnology, School of Life Science, Queensland University of Technology, GPO Box 2434, Brisbane, Queensland 4001, and 2 Research School of Biological Sciences, Australian National University, Canberra, ACT, Australia

We have sequenced the coat protein gene of nine isolates of papaya ringspot virus (PRSV) including six Australian and three Asian isolates and compared these with four previously reported sequences of PRSV. There was up to 12% sequence variation between isolates at the nucleotide level. However, there was no significant difference between the sequences obtained from Australian isolates irrespective of whether they were PRSV type P (cucurbit or papaya infecting) or PRSV type W (cucurbit infecting) and these isolates were more closely related to one another than to any other isolate. These results imply that PRSV-P, first recorded in Australia in 1991, arose locally from PRSV-W (first recorded in Australia in 1978) rather than being introduced. Further, there was no consistent sequence difference between PRSV-P and PRSV-W isolates that would obviously account for their host range difference.

Introduction

Papaya ringspot virus (PRSV), a species of the Potyviridae, is classified into two strains which can be distinguished only by host range. Papaya ringspot virus type W (PRSV-W), formerly watermelon mosaic virus 1, infects a number of species in the Cucurbitaceae, and is responsible for severe disease in a wide range of economically important cucurbit crops. Papaya ringspot virus type P (PRSV-P) infects papaya and has a limited experimental host range within the Cucurbitaceae. PRSV-P occurs in most tropical and subtropical countries where papaya is grown and has become a major limiting factor in papaya production in many countries, particularly in south-east Asia.

PRSV-W was first recorded in Australia in 1978 (Greber, 1978) and is now widespread throughout major cucurbit growing areas in Australia. PRSV-P, however, was not reported in Australia until 1991 (Thomas & Dodman, 1993). The origin of PRSV-P in Australia is not known; however, one possibility is that the virus was introduced to Australia from overseas in infected planting material or seeds. At present, PRSV-P is restricted to south-east Queensland.

Commercial papaya cultivars are not resistant to PRSV infection, so transgenic resistance must be considered as a way to develop PRSV resistant cucurbits and papaya varieties. Coat protein mediated resistance (Abel et al., 1986) has been successfully applied to a wide range of host–virus combinations (Nelson et al., 1988; Kaniewski et al., 1990; MacKenzie & Tremaine, 1990; Quemada et al., 1991; Brault et al., 1993) including several potyviruses (Da Camara Machado et al., 1992; Stark & Beachy, 1989; Namba et al., 1992; Ling et al., 1991) including PRSV in papaya (Fietch et al., 1992). However, this resistance can be very strain or isolate specific, probably depending on the relatedness of the transgene and the challenge virus (Sanders et al., 1992; Nelson et al., 1988; Quemada et al., 1991; Nakajima et al., 1993). Therefore, the variability between strains/isolates in the amino acid and/or nucleotide sequence of the coat protein of the challenge virus may have a significant effect on the level and stability of transgenic resistance.

Initial evidence suggested that there was little sequence variation within the coat protein gene among PRSV isolates. We previously reported the sequence of the coat protein gene of DB1, an Australian glasshouse isolate of PRSV-W (Bateson & Dale, 1992) and showed that there was considerable nucleotide (at least 96%) and amino acid (at least 98%) sequence similarity with U.S. isolates of PRSV-P and W (Quemada et al., 1990).

In this paper, we report that we have cloned and sequenced the coat protein gene of several isolates of PRSV from Australia and Asia. Comparison of the

The GenBank accession numbers for the sequence data reported in this paper are U14736 to U14744.
nucleotide and amino acid sequences of these isolates suggests that not only is PRSV far more variable than was at first recognized, but also that PRSV-P probably arose from PRSV-W in Australia.

Methods

Virus isolates. The names and origins of virus isolates are listed in Table 1. Australian isolates of PRSV-P and PRSV-W were obtained from field infected papaya and pumpkin or zucchini, respectively. Australian isolates GAT, BD, DAY, WP and BUN were collected from south-east Queensland by D. Persley while the NT (Northern Territory) isolate was supplied by Dr K. Gibb. Isolates of PRSV-P from Vietnam and Thailand were obtained as fresh and lyophilized leaf material, respectively, from field infected papaya. The Sri Lankan isolate of PRSV-P was obtained as a preparation purified from papaya from Dr M. Dassanayake. All isolates were stored at --80 °C before use.

RT-PCR from leaf tissue. Nucleic acids were extracted from 60 mg of fresh or frozen leaf tissue or 15 mg of lyophilized leaf tissue as described by Robertson et al. (1991). They were precipitated with ethanol, resuspended in 100 µl Tris-EDTA buffer and used directly for cDNA synthesis. Complementary primer MB12 was annealed to viral RNA by heating 10 µl of nucleic acid extract at 75 °C for 3 min and quenching on ice. First-strand cDNA was synthesized using AMV reverse transcriptase (Promega) and used immediately for PCR amplification or stored at --20 °C. A 905 bp fragment representing the coat protein gene of PRSV was amplified from 1 µl to 5 µl of first-strand cDNA mixture using 2.5 units of Taq polymerase (Cetus) and 35 pmol each of synthetic primers MB11 and MB12. The DNA-RNA hybrid was denatured at 94 °C for 2 min and the cDNA amplified for 35 cycles by extending at 72 °C for 1 min with a final extension cycle at 72 °C for 10 min. PCR products were analysed by electrophoresis in 1% agarose gels and stained with ethidium bromide. The PCR amplified DNA was cloned immediately without further purification.

Cloning of coat protein gene of PRSV isolates. PCR amplified DNA was cloned using a TA Cloning Kit (Invitrogen). A 1 µl aliquot of a standard 50 µl PCR reaction was ligated into the T-tailed plasmid and transformed into competent Escherichia coli strain INV-αF' according to the manufacturer’s instructions. Plasmid DNA from potential recombinants was purified by alkaline lysis and digested with restriction enzyme BamHI (Boehringer Mannheim). Clones containing a fragment of approximately 900 bp, corresponding to the coat protein gene, were selected. At least two clones of each isolate were selected for sequencing.

dsDNA sequencing. Double-stranded templates for sequencing were prepared as described by Hattori & Sakaki (1986). Denatured plasmids were sequenced using Sequenase version 2 (United States Biochemical) according to the manufacturer's instructions. Overlapping sequences were obtained using Universal primers (forward and reverse) and the specific primers MB11, MB12, MB13, MB14 and MB26.

Synthesis of oligonucleotides. Synthetic primers were synthesized in an Applied Biosystems PCR-Mate DNA synthesizer. Primers MB11 and MB12 were designed with six additional bases representing a BamHI restriction site at the 5' end. MB11 also included a start codon (ATG). Primers MB6, MB13, MB14 and MB26 were complementary to the PRSV-W(Aust-DB1) coat protein gene sequence. Primers had the following sequences: MB12, 5'GGATCCGCCCGACAAACAC-ACAAGTTGCGGATG3'; MB11, 5'GGATCCATGTCCAAAAATGA-AGCTGTGGATGCT3'; MB6, 5'GCAGCGCTGCGAAACACTAGCAAGTGCGATG3'; MB13, 5'GGAAATGGAATATTGTATGACCTTT3'; MB14, 5'ACCAAACATCAAGCCATTAGACT3'; MB26, 5'AATTGACATTCTAACAACCTGTCG3'.

Sequence analysis. Nucleotide sequences were aligned and translated, to obtain the encoded amino acid sequences, using the IBI Pustell sequence analysis program. The Wisconsin Genetics Computer Group (GCG) program package version 7.3 (Devereaux et al., 1984) was used to obtain sequence similarities. Nucleotide sequences were translated and the resulting amino acid sequences were aligned by the Clustal V program (Higgins & Sharp, 1989; Higgins et al., 1991); gaps were then added to the nucleotide sequences to align them with the amino acid alignments. Various distance matrices were then calculated from the aligned nucleotide sequences using the DISCALC program (G. F. Weiller, 1994 in preparation). These were compared by the DIPLOMO program (G. F. Weiller and A. J. Gibbs, in preparation) and some represented as 'trees' using the neighbour-joining method of Saitou & Nei (1987). Trees were also kindly computed by Dr John Hancock using the PAUP character state method (Swofford, 1991) and patristic distances of all trees compared by the DIPLOMO method.

Results

Synthetic primers MB11 and MB12, homologous to part of the PRSV-W(Aust-DB1) coat protein gene sequence, were used to amplify the coat protein gene of two Australian field isolates of PRSV-W and seven isolates of PRSV-P, including four Australian and three Asian isolates. All isolates of PRSV tested could be amplified with these primers, although SRI was amplified less efficiently than other isolates of PRSV. Up to four clones of each isolate were sequenced (data not shown); at most there were two nucleotide differences between the clones of each isolate. Single clones representing each isolate were selected for sequence comparison with other isolates of PRSV. The nucleotide sequences (Fig. 1) and encoded
Isolate variability in PRSV

<table>
<thead>
<tr>
<th>Isolate</th>
<th>3549</th>
</tr>
</thead>
</table>

Fig. 1. Nucleotide sequences of the coat protein genes of PRSV isolates aligned using the IBI Pustell sequence analysis program. Sequences are compared to DB1, an Australian glasshouse isolate of PRSV-W. Base differences are indicated, — indicates the same base is present; ▼ indicates approximate positions of N and C termini as suggested by Shukla et al. (1989); * represents deletions with respect to PRSV-W (Aust-DB1).
Fig. 2. Putative amino acid sequences of the coat protein genes of PRSV isolates compared to DB1, an Australian glasshouse isolate of PRSV-W. Amino acid differences are indicated below the sequence, ▼ indicates the same amino acid is present, ▼ indicates approximate positions of N and C termini as suggested by Shukla et al. (1989); * represents deletions with respect to PRSV-W(Aust-DB1).

Table 2. Percentage similarity between isolates of PRSV

<table>
<thead>
<tr>
<th></th>
<th>GAT</th>
<th>DBI</th>
<th>DAY</th>
<th>WP</th>
<th>BUN</th>
<th>BD</th>
<th>NT</th>
<th>USAP</th>
<th>USAW</th>
<th>TAIW</th>
<th>THAI</th>
<th>VIET</th>
<th>SRI</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAT</td>
<td>100</td>
<td>98.82</td>
<td>98.70</td>
<td>98.35</td>
<td>98.47</td>
<td>98.35</td>
<td>98.11</td>
<td>96.58</td>
<td>96.82</td>
<td>90.57</td>
<td>91.27</td>
<td>90.80</td>
<td>89.03</td>
</tr>
<tr>
<td>DBI</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DAY</td>
<td>100</td>
<td>98.70</td>
<td>98.35</td>
<td>98.47</td>
<td>98.35</td>
<td>98.11</td>
<td>96.58</td>
<td>96.82</td>
<td>90.45</td>
<td>90.92</td>
<td>90.68</td>
<td>89.68</td>
<td>88.80</td>
</tr>
<tr>
<td>WP</td>
<td>100</td>
<td>98.35</td>
<td>98.47</td>
<td>98.70</td>
<td>98.47</td>
<td>98.93</td>
<td>97.17</td>
<td>90.92</td>
<td>91.51</td>
<td>90.92</td>
<td>89.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BUN</td>
<td>100</td>
<td>98.47</td>
<td>98.35</td>
<td>98.23</td>
<td>98.11</td>
<td>96.82</td>
<td>97.17</td>
<td>91.04</td>
<td>91.27</td>
<td>91.16</td>
<td>88.80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BD</td>
<td>100</td>
<td>98.11</td>
<td>96.82</td>
<td>97.17</td>
<td>91.04</td>
<td>91.27</td>
<td>91.16</td>
<td>90.92</td>
<td>90.92</td>
<td>89.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NT</td>
<td>100</td>
<td>96.58</td>
<td>96.82</td>
<td>90.92</td>
<td>91.51</td>
<td>90.92</td>
<td>89.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>USAP</td>
<td>100</td>
<td>98.11</td>
<td>91.04</td>
<td>91.39</td>
<td>90.92</td>
<td>89.27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>USAW</td>
<td>100</td>
<td>90.57</td>
<td>90.92</td>
<td>90.80</td>
<td>89.50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAIW</td>
<td>100</td>
<td>93.75</td>
<td>93.51</td>
<td>89.39</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>THAI</td>
<td>100</td>
<td>93.04</td>
<td>89.39</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VIET</td>
<td>100</td>
<td>89.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SRI</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

It was found that, as with other potyviruses, the N-terminal regions of the virion proteins, and the portions of the genes encoding them, were more variable than the core regions, as is shown by a graph of the 'heterozygosity' of each position in the aligned sequences (Fig. 3). These differences seem to result from differences in selection against change; there was a mean of 5.4% nucleotide changes in the third codon positions of the region encoding the N terminus and 4.9% in that encoding the core, but mean differences of 3.2% and 1.9% in the first and second codon positions of the region encoding the N terminus, and only 0.5% and 0.4% respectively for the region encoding the core. Nonetheless, dendrograms calculated for the N-terminal and core regions using either nucleotide or amino acid differences, were closely similar, as too were dendrograms calculated for the same sets of data using the PAUP programs (Swofford, 1991). So, to illustrate the relatedness of the complete PRSV virion protein genes, the silent changes of the complete genes were used to calculate, by the neighbour-joining method, the dendrogram shown in Fig. 4.

The four Australian isolates of PRSV-P (BD, BUN, DAY and WP) were found to cluster with the two Australian PRSV-W isolates, and the Australian cluster was closest to the cluster of the W and P isolates of PRSV from the U.S.A. (USAW and USAP). There were no significant differences between the W and P isolates in the cluster of Australian isolates; their N-terminal-encoding and core-encoding regions had mean nucleotide
Isolate variability in PRSV

Fig. 3. Heterozygosity (Nei et al., 1975) of the codons of the virion protein genes of 13 isolates of PRSV. Upper and lower scatter diagrams show the average heterozygosity of each codon of the aligned nucleotides and the encoded amino acids respectively; gaps treated as the fifth nucleotide or twenty-first amino acid. Arrow indicates the junction of the N-terminal and core regions.

Fig. 4. Dendrogram showing the relationships of the virion protein genes of 13 isolates of PRSV calculated by the neighbour-joining method (Saitou & Nei, 1987) for the percentage non-coding (silent) nucleotide differences between them.

dissimilarities of 1.62% (±0.16) and 1.51% (±0.05), respectively. However, they differed from the U.S. isolates, in the same regions, by 4.44% (±0.18) and 2.86 (±0.06) respectively, and by larger amounts from the other Asian PRSV isolates.

To establish the root of the PRSV tree (Fig. 4), namely the possible phylogenetic pathway to other potyvirus virion proteins, we added two other potyvirus virion protein genes, potato virus Y (PVY) and pea seed-borne mosaic virus (PSbMV), as outgroups. The relationships of this larger group were calculated from the core region only, so that the resulting dendrogram (Fig. 5) could be directly compared with other published trees. The topology of this tree shows that there are no large differences in the rate of evolutionary change in different PRSV lineages.

Discussion

Initially, information from the coat protein gene sequences of PRSV-W and PRSV-P from the U.S.A. (Quemada et al., 1990; Wang & Yeh, 1992) and PRSV-W from Australia (Bateson & Dale, 1992) suggested that there was very little sequence variability within this gene between isolates of PRSV. However, in this study we have found considerable variability in coat protein gene sequences of PRSV from different parts of the world.
There was up to 12% nucleotide sequence variability, both within the Asian PRSV isolates as well as between Asian and Australian or U.S. PRSV isolates. Variability was also evident at the amino acid level, particularly between Asian isolates. Interestingly, this sequence variability did not correspond to strain differences; the coat protein gene sequences of PRSV-P(Aust) and PRSV-W(Aust) isolates were very similar as were those of PRSV-W(USA) and PRSV-P(USA). The variability did, however, correspond to geographical regions. At the present time, information on variability within a geographical region is only available for Australia. There was no significant difference between the six Australian isolates of PRSV. While the Australian isolates of PRSV-P and PRSV-W were most closely related to the U.S. isolates of PRSV-P and PRSV-W, they were more closely related to each other.

PRSV-W was first recorded in Australia in 1978 and PRSV-P in 1991. Our data suggest that PRSV-P(Aust) was derived, possibly by mutation, from PRSV-W(Aust) and was not introduced, as has been previously suggested. This is supported by the observation that PRSV-P is rarely, if ever, found naturally infecting cucurbits (S. Ferrara, personal communication) suggesting that, while PRSV-P can infect cucurbits experimentally, in nature it is better adapted to papaya. However, PRSV-P(Aust) has been recorded naturally infecting cucurbits (D. Persley, personal communication) although this was in field trials. Thus PRSV-P in Australia may well be a very recent mutational event and has yet to become fully adapted to papaya. It is therefore possible that PRSV-P has also arisen locally from PRSV-W within other regions rather than spreading between regions. Therefore, we are currently testing this hypothesis by comparing the coat protein gene sequences of PRSV-P and -W isolates from within particularly south-east Asian countries.

In PRSV, as has been shown for strains of other potyviruses, the coat protein core is highly conserved with most amino acid variability confined to the N terminus (Shukla & Ward, 1989). Our results confirm that between isolates of PRSV, the core is also conserved and the N terminus is most variable. Xiao et al. (1993) suggested that the N terminus may be involved in host specificity. However, we have found no obvious sequence changes, either nucleotide or amino acid, that would account for the difference between PRSV-P and PRSV-W. Interestingly, the C terminus of the coat protein was highly conserved between all PRSV isolates. Shukla et al. (1988) showed that the C terminus, like the N terminus, is exposed on the surface of the virus particle and therefore does not necessarily have a role in the structural stability of virions. The observed high level of conservation, however, suggests that the C terminus has an important role for which little amino acid sequence variability can be tolerated.

Sequence variability may have important implications for the use of coat protein gene constructs for transgenic resistance. Evidence already exists for several virus–host systems to suggest that such resistance can be highly sequence specific (Sanders et al., 1992; Nelson et al., 1988; Quemada et al., 1991; Nakajima et al., 1993). Recently, transgenic papaya, incorporating the coat protein gene of a mild strain (HA 5-1) of the U.S. isolate of PRSV-P (Tennant et al., 1993) has been reported to be protected specifically against infection by the severe U.S. isolate (HA) of PRSV-P but not against infection with other isolates of PRSV-P including an Australian and a Thai isolate; the reaction of transgenic plants ranged from delayed symptoms to complete susceptibility. This is surprising considering that Australian isolates are about 96% similar at the nucleotide and 98% similar at the amino acid levels to PRSV-P.

This work was, in part, supported by grants from the Horticultural Research and Development Corporation/Queensland Fruit and Vegetable Growers and a collaborative grant with the Department of Industry, Technology and Regional Development and Arthur Yates and Company. J.H. is supported by an Australian Postgraduate
Research Award and W.C. is supported by an Australian International Development Aid Bureau Scholarship. We thank D. Persley, Queensland Department of Primary Industry for supply of PRSV isolates and personal communication.

References


(Received 5 April 1994; Accepted 19 August 1994)