BRIEF REVIEW

The potyviruses of Australia

A. J. Gibbs · A. M. Mackenzie · K.-J. Wei ·
M. J. Gibbs

Abstract Many potyviruses have been found in Australia. We analyzed a selected region of the coat protein genes of 37 of them to determine their relationships, and found that they fall into two groups. Half were isolated from cultivated plants and crops, and are also found in other parts of the world. Sequence comparisons show that the Australian populations of these viruses are closely related to, but less variable than, those in other parts of the world, and they represent many different potyvirus lineages. The other half of the potyviruses have only been found in Australia, and most were isolated from native plants. The sequences of these potyviruses, which are probably endemic, are on average five times more variable than those of the crop potyviruses, but surprisingly, most of the endemic potyviruses belong to one potyvirus lineage, the bean common mosaic virus lineage. We conclude that the crop potyviruses entered Australia after agriculture was established by European migrants two centuries ago, whereas the endemic plant potyviruses probably entered Australia before the Europeans. Australia, like the U.K., seems recently to have had c. one incursion of a significant crop potyvirus every decade. Our analysis suggests it is likely that potyviruses are transmitted in seed more frequently than experimental evidence indicates, and shows that understanding the sources of emerging pathogens and the frequency with which they 'emerge' is essential for proper national biosecurity planning.

"The hypothesis is proposed that all plant viruses in Australia were introduced since European settlement of the Australia continent toward the end of the eighteenth century", N. H. White 1973.

A. J. Gibbs · A. M. Mackenzie · K.-J. Wei · M. J. Gibbs
Research School of Biological Sciences,
Australian National University,
Canberra, ACT 0200, Australia

Present Address:
A. J. Gibbs (✉)
7 Hutt St, Yarralumla, ACT 2600, Australia
e-mail: adrian_j_gibbs@hotmail.com

Present Address:
A. M. Mackenzie
Division of Plant Industry,
C.S.I.R.O., Black Mountain Laboratories,
Canberra, ACT 2601, Australia

Present Address:
M. J. Gibbs
79 Carruthers St, Curtin, ACT 2605, Australia

The question

In 1973 White [36] drew attention to the fact that very few viruses had been isolated from Australian native plants and, with a few possible exceptions, the viruses isolated from crops in Australia had been previously found in other parts of the world. In an attempt to stimulate research, he suggested that it was possible that all plant viruses now in Australia had been introduced from other parts of the world since European settlement in the eighteenth century. His reasons were that "no authentic virus peculiar to the indigenous species of the Australian flora has been proved, and with few exceptions, those viruses first recorded from Australia on introduced plant species, have been found in other parts of the world." He then discussed how the evidence to refute or support the idea might be obtained using

various "principles governing the ecology of plant viruses in the world at large". He noted, in particular, that Holmes [18] had shown that, among 36 species of *Nicotiana* from the Americas and Australia, only species from Chile and Peru were resistant or tolerant to tobacco mosaic virus (TMV) and had suggested that TMV may have been long enough in the Americas to select plants with those traits, whereas in other parts of the world, such as Australia, the virus was absent, and selection for resistance had not occurred.

Gibbs and Guy [11] suggested that, in addition to a more thorough search for viruses in native plants, Neville White's challenge might be answered from studies of variation of known viruses, because if viral populations evolved in a time-dependent manner, then analyses of their phylogeny might correlate with their provenance, and populations that had recently entered Australia would probably be less variable than those that had been evolving within Australia over a longer period.

Over the past third of a century, many more viruses have been isolated from plants in Australia, including several from wild plants. The gene sequences of a significant number of these have been determined [34], and, importantly, some have been found to be related to viruses overseas. So it is now possible to test Neville White's question and establish whether the relative geographic isolation of Australia until European colonization two centuries ago is reflected in its plant virus populations as it is in its flora [6] and fauna [5], including its aphids [3, 8]. A large proportion of the plant viruses isolated in Australia are from one family, the *Potyviridae*. Here, we report a phylogeographic analysis of them, seeking clues of their origins.

Potyviruses

The family *Potyviridae* is one of the two largest families of known plant viruses [9], and most of its species are from the genus *Potyvirus*. Potyviruses have been isolated from species of all the major angiosperm taxa in all parts of the world. Like other species of the *Potyviridae*, potyviruses have flexible filamentous virions. Potyviruses are transmitted by migrating aphids probing plants in search of their preferred host species; several aphid species may transmit each potyvirus. Some are also transmitted in seeds to the progeny of infected plants.

Early attempts to characterize and identify potyviruses relied on serological tests, but these gave unreliable results as antisera produced using different virion preparations of the same virus often gave significantly different results. Dharma Shukla, Colin Ward and their colleagues [29] showed that the major antigenic epitope of all potyvirus

virions was the N-terminal region of the virion protein (CP), which is exposed on the surface of the virion and often degrades while virions are being purified, so potyvirus virions are antigenically and immunogenically unstable. Furthermore, parts of the sequences from this genomic region are often repetitive, probably as a result of 'replicase slippage' during transcription [16]. It seems therefore to have evolved in a saltatory way [33]; however, other parts of potyvirus genomes seem to evolve in a coherent progressive time-dependent way, and taxonomies obtained from them mostly correlate broadly with groupings based on the most consistent serological tests.

Potyviruses were among the first viruses to be isolated from Australian crops. A survey of Australian viruses in 1988 [2] recorded around 150 potyviruses in crop species but only 25 in wild species. Other wild and weedy species of plants showing mosaic or mottling symptoms were subsequently found to contain potyvirus-like virions and react with group-specific potyvirus antisera [22]. Gene sequencing of samples of these and other plants has confirmed that potyviruses are widespread in Australia, certainly along the western [34] and eastern seaboards of Australia, but perhaps not in the center of the continent (A. Mackenzie and A. Gibbs, unpublished work).

Here, we review the phylogeography of all the potyviruses isolated from Australian plants for which gene sequences are publicly available, including several that have not been previously described. In these analyses, the phylogenies of various sets of potyviruses have been determined using gene sequences, and we have searched the trees for clusters of potyviruses with restricted geographic distributions to obtain clues about when they arrived in Australia.

Sequences and methods

Most recent reports of new potyvirus gene sequences include some form of phylogenetic analysis, but different sets of viruses, genes and parts of genes have been used in different analyses, and so comparisons between analyses cannot be made. Standardization is essential for making comparative phylogeographic analyses. Therefore, for ours, we used a single region of the potyvirus genome, and we used the same large outgroup of sequences for all analyses.

In late 2006 the Genbank database contained more than 5,000 sequences retrieved by the search term "potyviridae OR potyvirus". More than 300 of these sequences were of complete or near-complete genomes of members of 47 species and were at least 8,000 nucleotides in length. The majority of the others were of the 3′-terminal region, which encodes the coat protein (CP), and many also included part

of the NIb gene, as cDNA encoding this region is reliably obtained by RT-PCR using primers that target the motif -GNNSQ- [4, 12]. The sequences including this region were aligned using CLUSTALX [32] with default parameters. They varied greatly in relatedness along their length (Fig. 1), especially the region encoding the N-terminal part of the CP, which is often repetitive and requires a large number of gaps to align [33]. The N-terminal CP of papaya ringspot virus (PRSV) (AY027810), for example, has the amino acid sequence:

-DAGLNEKLKEKEKEKQKEKEKDEQKDKDNDGA SDGNDVSTSTKTGERDRDVNAG-, and its main repetitive region -KEKEKEKQKEKEKD- is encoded by the repetitive sequence:

-AAAGAAAAAGAAAAAGAAAAACAGAAAGAAA AAGAAAAAGAT-.

Other potyviruses have different, often quite unrelated, repetitive sequences in this region, indicating that this region has probably arisen in an independent saltatory way in different potyviruses by a process involving replicase slippage [33]. The remainder of the CP gene, that encoding its 'core' and C-terminus (Fig. 1, region D to E), shows no unusual sequences of this sort and, by contrast, seems to have evolved in a coherent way by point mutations and occasionally by homologous recombination. We therefore used the nucleotide sequences of this 'coherently evolving CP' (cCP) region for our analyses; nucleotide sequences contain both silent and non-silent differences and are
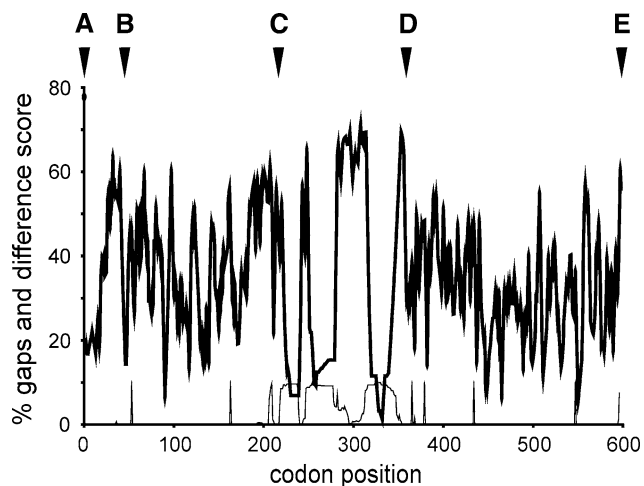
therefore usually more phylogenetically informative than the amino acid sequences they encode. The 5′ terminal region of the cCP region is variable and so is best identified using the amino acid sequence it encodes, which, for example in the cCP region of potato virus Y (PVY) (Genbank Reference Sequence NC_001616 [27]), is DVNAG.

A neighbor-joining (NJ) tree calculated from the cCP sequences of all potyvirids (Fig. 2) showed that they are of two types. Those of potyviruses (sensu stricto) form a radiation with a noticeably uniform branch length, namely a 'star-burst'. The other potyvirids are placed on longer hierarchical lineages; these are distinctly different in topology from the potyvirus cluster. These contrasting topologies may indicate that the potyviruses and the remaining potyvirids have evolved differently, and if so, then their trees cannot be compared directly. We therefore confined our analyses to the cCP sequences of potyviruses.

The 'star-burst' topology of the potyvirus cCP tree suggests that almost all the sequences have diverged a similar amount from the apparent center of the radiation so that all pairs of sequences linked through that center are similar distances from one another, and none are more distant. We used for our analyses a large representative set
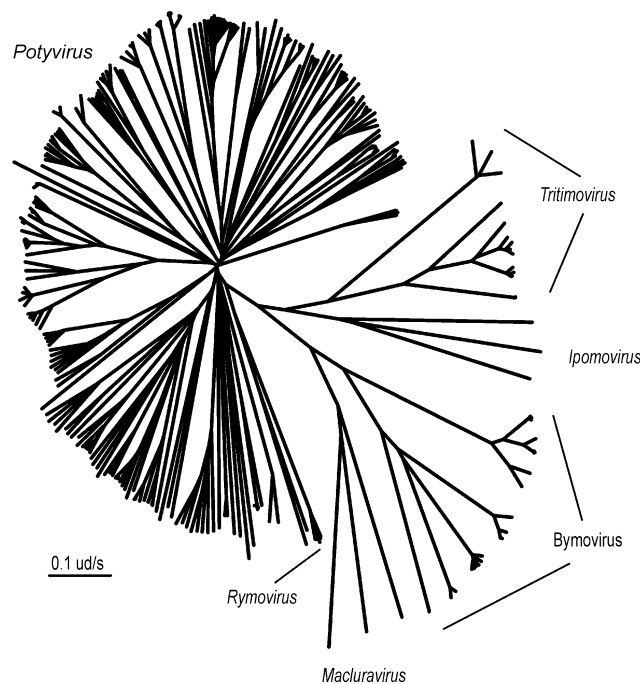

Fig. 1 Graph showing the variability of the GNNS 3′-terminal region of the aligned genomes of 365 potyviruses; this region is of 1,794 nucleotides and gaps (598 codons) after being aligned. Variability was determined by the DnDscan method [14] using a window of nine nucleotides and step of three nucleotides. The *upper line* is the mean pairwise divergences of the three codons in each window position. The *lower thin line* shows the positions and percentages (×0.1) of codon gaps. *Arrow A* indicates the position of the GNNSGQP motif, *B* the position of the GDD motif, *C* the 3′ codon of the NIb gene, *D* the position of the -DVNAG- motif, and *E* the 3′ terminus of the CP gene

Fig. 2 Neighbor-joining tree showing the relationships of the GNNS 3′-terminal region (defined in text) of the genomes of 304 potyvirids. The radial fan of sequences includes all those of species of the genus *Potyvirus*; the longer lineages include all those of the 'non-potyvirus' potyvirid genera: *Bymovirus*, *Macluravirus*, *Ipomovirus*, *Rymovirus* and *Tritimovirus*. The unit of divergence is the 'uncorrected pairwise sequence difference per site' (*ud/s*)

of 'outgroup potyvirus' cCP sequences. It consisted of the cCPs of members of the 47 different potyvirus species, each of which was represented in Genbank in late 2006 by at least one complete genomic sequence, together with that of ryegrass mosaic rymovirus (RGMV), which is the 'non-potyvirus' potyvirid most closely related to potyviruses. The relationships of the 'outgroup potyvirus' cCP sequences together with their accession codes are shown in Fig. 3. The pairwise patristic distances (i.e. branch lengths) presented as a histogram in Fig. 4 confirms the consistency of their divergence; the pairwise sequence divergences range from about 0.25 to over 0.40 uncorrected differences/site (ud/s), or 75–60% identity, but most formed a symmetrical peak centered at 0.373 ± 0.018 ud/s, which is the diameter of the 'star-burst'. This peak represents branches linking 24 individual species and five or so lineages of two to eleven species with each lineage having consistent inter-species divergences of less than 0.30 ud/s.

The lineages found by the NJ method were also found in trees made using the maximum-parsimony method (MP; PAUP version 4.0b10 with the tree branch reconnection method) [31], by the maximum-likelihood method using PhyML [15] and either the GTR + I + G evolutionary model [23] or the HKY + I + G model [17]. None of these methods or evolutionary models consistently resolved the relationships between the different lineages and species at the apparent center of the major initial radiation, nor were any of nodes at that center significantly supported by

boot-strap analyses. Boot-strap analyses did, however, support the membership and relationships within individual lineages, namely BCMV and ten other species in the BCMV lineage (NJ 97%, ML 57% and MP 88%), PVY and five other viruses (boot-strap support NJ 100%, ML 80% and MP 80%), sugarcane mosaic virus (SCMV) and two others (boot-strap support NJ 100%, ML 74% and MP 100%), also, although with smaller but still significant bootstrap support, turnip mosaic virus (TuMV) and two
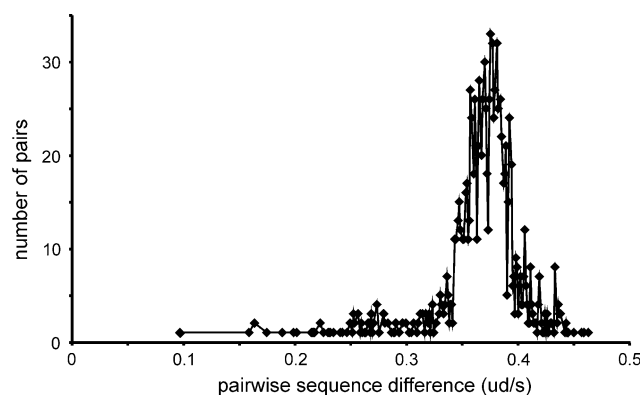

Fig. 4 Histogram of the pairwise sequence divergences of the cCPs in Fig. 3. The pairwise patristic distances (i.e. branch lengths) in the tree were obtained using PATRISTIC [10] and presented as a histogram (Fig. 4) using SBHistogram (http://www.sb-software.com/sbhisto/)

Fig. 3 The outgroup potyviruses. Neighbor-joining tree showing the relationships and accession codes of the cCP sequences of the 47 potyviruses and ryegrass mosaic rymovirus used as the outgroup potyviruses for all analyses. They were aligned via their encoded amino acid sequences using the Transalign program (kindly supplied by Georg Weiller) and CLUSTALX [19] with default parameters, and this gave sequences with 720 nucleotides and gaps (240 codons). The relationships of these cCP sequences were calculated by the neighbor-joining method
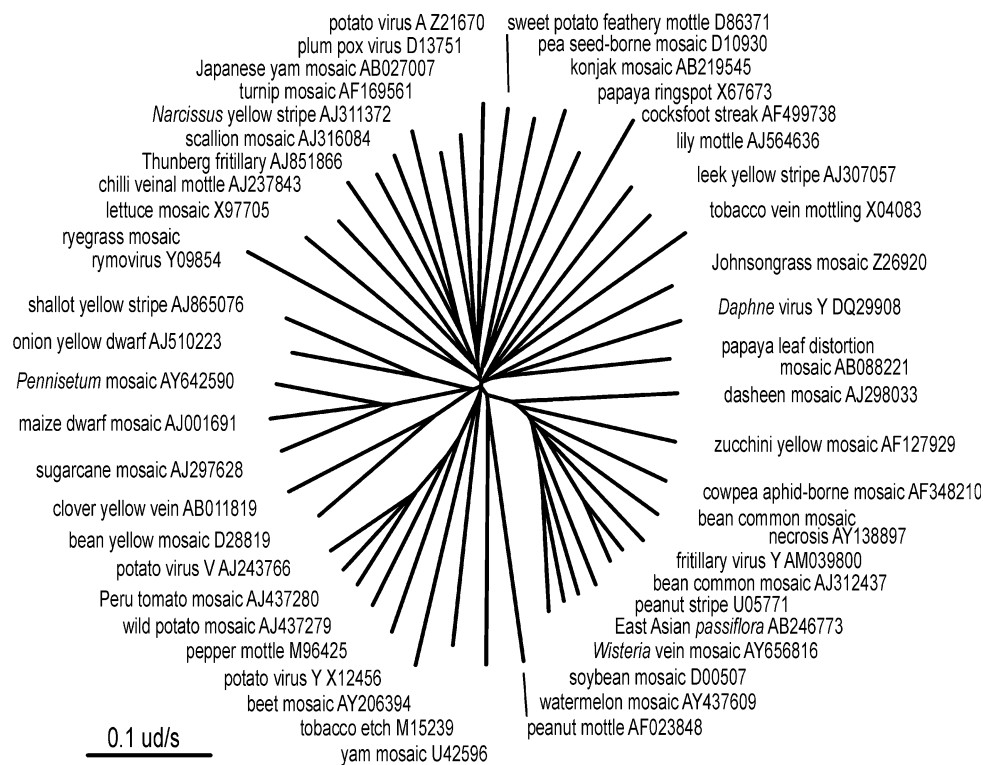
others, bean yellow mosaic and clover yellow vein viruses; and onion yellow dwarf and shallot yellow stripe viruses.

For our analyses, each known Australian cCP sequence was used as a query sequence and matched against the Genbank database using its BLAST facility (http://www.ncbi.nlm.nih.gov/BLAST/). The query sequence and all those that significantly matched it were then aligned with the 'outgroup potyvirus' cCP sequences using CLUSTALX [32] in its 'Profile Alignment' mode with default parameters; all cCP sequences with more than 5% of the nucleotides missing from either end were discarded. The trees were calculated by the NJ method and mostly viewed as radial phylogenetic trees to emphasize groupings, lineages and radiations. These comparisons showed how closely the Australian cCPs were related to those of the same and other species, and whether there was any 'nesting' of geographically restricted lineages within the phylogenies. Basal sequences and those with atypical long branches were checked for evidence of recombination using the RDP package version 3.22 [24] with default settings and also by the PHYLPRO program [35].

Potyviruses found in Australia

The occurrence in Australia of many of the viruses discussed in this paper has been recorded previously [1, 2, 25, 28]. Some have not previously been reported, and the extent to which they have been characterized varies greatly. PVY, for example, has been very fully characterized, but not in Australia, where only a few isolates of it have had their cCP genes partially determined. Others have not been characterized in a traditional sense, but their gene sequences have been recovered from samples of wild and weed plants showing chlorotic mosaics or mottling, using RT-PCR and potyvirid-specific primers [12]. Most of our attempts to transmit viruses from these plants by sap-inoculating a limited range of standard indicator plants grown in glasshouses were unsuccessful. These sequences were mostly deposited in Genbank in 2005, and we call them "Candidatus" potyviruses by analogy with, and using the style of naming of, bacteria that are known only from gene sequences and have not yet been cultured and maintained (see http://en.wikipedia.org/wiki/Candidatus [26, 30]).

Of the potyviruses isolated in Australia, 36, including the Candidatus species, have had the sequences of their cCP genes determined, and the partial cCP sequences of five others are known. These viruses, together with notes on their occurrence in Australia and overseas and their cCP sequence relationships, are listed in the Electronic Supplementary Material in alphabetical order. The phylogenetic analyses show that they fall into two groups.

Potyviruses found in Australia and also in other parts of the world

Most of these are well-studied crop pathogens that cause significant crop damage. Their phylogenies mostly place the Australian isolates among the terminal 'twigs' of trees of their world populations, and the world population of each virus is much more diverse than the Australian population. Typical of this pattern is that of SCMV, which has been reported from all the sugarcane-growing areas of Australia. It is widely distributed throughout the world, naturally infects many different grass species, and can be transmitted experimentally to even more. A dozen cCP sequences of Australian isolates have been determined (see Electronic Supplementary Material). They form a close-knit, and probably monophyletic, group of isolates (Fig. 5) with a mean divergence of 0.011 ud/s that is, in one part of the star cluster, formed by all known SCMVs. Thus, it is most likely that the Australian isolates were introduced to Australia on a single occasion, perhaps when sugarcane was imported in 1924 [2]. Within the same potytomous cluster are isolates from Brazil, China, India and the USA (smallest divergence 0.015 ud/s). The most diverse SCMV cCPs differ from one another by around 0.115 ud/s, and
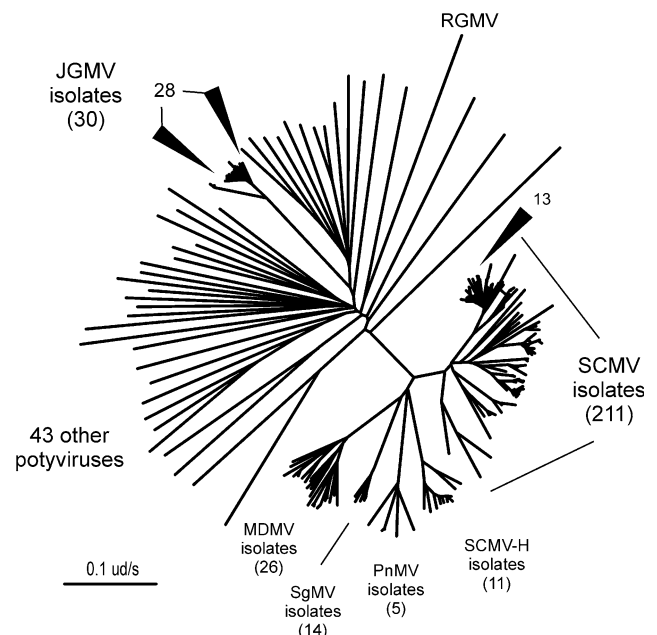

Fig. 5 Grass potyviruses. Neighbor-joining tree calculated from the aligned cCP sequences from various potyviruses from grasses together with those of the other 43 outgroup potyviruses and ryegrass mosaic rymovirus. *SCMV* sugarcane mosaic virus, *SCMV-H* sugarcane mosaic virus H isolates, *PnMV* pennisetum mosaic virus, *SgMV* sorghum mosaic virus, *MDMV* maize dwarf mosaic virus, *JGMV* Johnson grass mosaic virus. *Arrows* indicate the positions of cCPs from Australian isolates; the numbers of cCPs are given when there is more than one

they are clearly separated from the other viruses of the 'SCMV lineage', which are maize dwarf mosaic virus, Pennisetum mosaic virus and sorghum mosaic virus, and which have inter-species divergences of around 0.18–0.20 ud/s (Fig. 5). Johnson grass mosaic virus (JGMV), which was not originally distinguished from SCMV lineage viruses is distinct and differs from them by a mean divergence of 0.374 ud/s.

Another pattern of relationships is that shown by bean yellow mosaic virus (BYMV). This virus has been found worldwide, most often in legumes and various monocotyledonous bulb and orchid species, notably cultivated *Gladiolus* spp. BYMV is genetically variable, with divergences up to 0.16 ud/s. Figure 6 shows the relationships of the cCP sequences of 52 BYMV isolates and 19 isolates of the closely related clover yellow vein virus, including several from Australia (see Electronic Supplementary Material). There are several distinct lineages of BYMV cCPs, and Australian isolates are found in most of them. There is no correlation between the phylogenetic relationships of different isolates and their provenance or the host from which they were isolated. The simplest interpretation of this tree is that the world population of BYMV has several lineages, and isolates from at least four of them
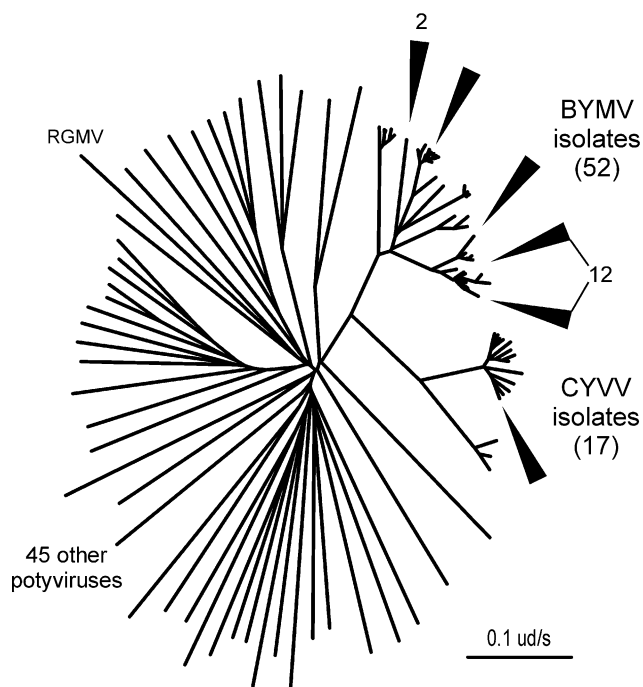

Fig. 6 Bean yellow mosaic and clover yellow vein viruses. Neighbor-joining tree calculated from the aligned cCP sequences from various isolates of bean yellow mosaic and clover yellow vein viruses together with those of the 45 outgroup potyviruses and ryegrass mosaic rymovirus. *Arrows* indicate the positions of cCPs from Australian isolates; the numbers of cCPs are given when there is more than one

have been imported into Australia, probably on separate occasions. International trade, especially of gladioli and other bulbs used as ornamentals, and also in forage legume seed, is probably responsible for the worldwide distribution of BYMV.

The other cosmopolitan viruses found in Australia are Apium virus Y (ApVY), bean common mosaic virus (BCMV), celery mosaic virus (CeMV), Ornithogalum mosaic virus (OrMV), PRSV, pea seed-borne mosaic virus (PSbMV), PVY, sweet potato feathery mottle viruses (SPFMV-C and SPFMV-RC), sweet potato virus Y (SPVY), TuMV, watermelon mosaic virus (WMV) and zucchini yellow mosaic virus (ZYMV); see "Electronic Supplementary Material".

The simplest interpretation of the phylogenetic trees of these viruses is that, in each instance, the Australian population is a small part of the world population rather than the reverse, indicating that these potyviruses probably migrated to Australia from overseas. Some of these phylogenies, those of BYMV, CYVV, PVY, OrMV and TuMV, place the Australian isolates in more than one lineage of the world population of the virus, indicating that they probably entered Australia on more than one occasion, unlike PRSV and SCMV, whose Australian populations form single clusters in their respective world trees and probably entered Australia on only one occasion. The trees of the remaining five cosmopolitan viruses, ApVY, CeMV, JGMV, PeMV and PleVY, give no clear indication about their origins. They are known from too few sequences to make sensible comparisons, and three, CeMV, JGMV and PeMV, have phylogenies which place the Australian and overseas sequences as sister clusters. Significantly, however, none of the trees of these cosmopolitan viruses show a diverse Australian population with an overseas population as mere twigs, so there is no indication that any of them originated in Australia; Australia has not been a source of potyviruses.

Potyviruses found only in Australia

Most of these have been isolated from wild or weed species, and only two from recently introduced crop species: passion fruit and carrots. Best studied of these is Hardenbergia mosaic virus (HarMV) [34]. Many are 'Candidatus' potyviruses, namely ceratobium mosaic virus (CerMV), Clitoria virus Y (CliVY), Dianella chlorotic mottle virus (DiCMV), Diuris virus Y (DiVY), Euphorbia ringspot virus (ERV), Eustrephus virus Y (EustVY), Glycine virus Y (GVY), Hibbertia virus Y (HibVY), Kennedya virus Y (KVY), Passiflora foetida virus Y (PfoVY), Pleione virus Y (PlVY), Pterostylis virus Y (PtVY), Rhopalanthe virus Y (RhVY), Sarcochilus virus Y (SarVY), Siratro 1 virus Y (Sir1VY) and Siratro 2 virus Y (Sir2VY).

The viruses infecting *Passiflora* spp. are noteworthy. Two infect passion fruit (see Electronic Supplementary Material), and one infects a weed *Passiflora*. All three of them are viruses of the BCMV lineage and most closely related to three *Candidatus* potyviruses found only in Australia, Hib-VY, S1VY and S2VY. Similar diseases of *Passiflora* spp. that are also caused by members of the BCMV lineage have been reported from other parts of the world; species of *Passiflora*, originally from the Americas, seem to be particularly susceptible to BCMV lineage potyviruses. Also noteworthy are four viruses from orchids (PlVY, PtVY, RhVY and SarVY), but only one of them (PtVY) is from the wild, and two of them from imported orchids (RhVY and SarVY), but again, most of them are members of the BCMV lineage (Fig. 7). Finally, there is carrot virus Y (CarYV), which is widespread and damaging in feral and crop carrots in all the southern States of Australia. This virus has only been reported in Australia, but its wide distribution but very small population divergence suggests that it is most likely to be a recent migrant to Australia, but its overseas parent population has not yet been identified.

Phylogenetic differences

The phylogenetic trees of the cosmopolitan and endemic viruses differ not only qualitatively but also quantitatively. First, the Australian populations of the cosmopolitan potyviruses are much less diverse (mean 0.015 ud/s, range 0.002–0.042 ud/s; Table 1) than those of the endemics (mean within-species divergence 0.053 ud/s; range 0.007–0.159 ud/s; Table 2). The mean divergence between the cCPs of Australian isolates of cosmopolitan viruses and the closest overseas isolates is also small (c. 0.019 ud/s), whereas the mean divergences of their overseas populations are much larger (0.091 ud/s; range 0.002–0.190 ud/s) (Table 1). By contrast, the cCP sequence closest to each endemic virus is, on average, much more distant (mean interspecies divergence 0.177 ud/s; range 0.110–0.260 ud/s). The third significant difference between the cosmopolitan and endemic potyviruses is that 14 of the 18 endemic potyviruses are members of the BCMV lineage, whereas the cosmopolitan viruses are a diverse selection of the genus *Potyvirus* representing 13 different lineages; ApVY and CeMV are members of one lineage, BCMV, WMV and ZYMV of another, and BYMV and CYVV of a third, and all the others are members of distinct species from other lineages.

The answer

We conclude from our analyses that potyviruses did not originate in Australia, but are immigrants. Half entered
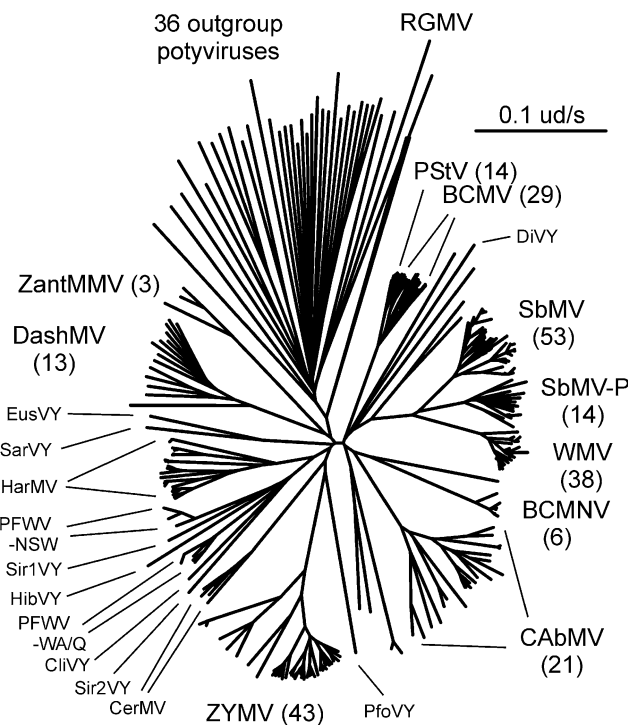

Fig. 7 The bean common mosaic virus lineage of potyviruses. Neighbor-joining tree calculated from the aligned cCP sequences from 316 isolates of BCMV lineage potyviruses together with those of the remaining 36 outgroup potyviruses and ryegrass mosaic rymovirus (*RGMV*). The endemic Australian potyviruses are: *DiVY* diuris virus Y, *PfoVY* passiflora foetida virus Y, *CerMV* ceratobium mosaic virus (five isolates), *Sir2VY* siratro 2 virus Y, *CliVY* clitoria virus Y, *PFWV-WA/Q* passionfruit woodiness viruses from Western Australia (four isolates) and one isolate from Queensland, HibVY hibbertia virus Y, *Sir1VY* siratro 1 virus Y, *PFWV-NSW* passionfruit woodiness viruses from New South Wales (three isolates), *HarMV* hardenbergia mosaic virus (34 isolates), *SarVY* sarcochilus virus Y, *EustVY* eustrephus virus Y. The clusters of major crop-infecting viruses are: *PStV* peanut stripe virus isolates of BCMV, *BCMV* bean common mosaic virus, *BCMNV* bean common mosaic necrosis virus, *CAbMV* cowpea aphid-borne mosaic virus, *DashMV* dasheen mosaic virus, *SbMV* soybean mosaic virus, *SbMV-P* soybean mosaic virus-Pinellia, *WMV* watermelon mosaic virus, *ZantMMV* zantedeschia mild mottle virus, *ZYMV* zucchini yellow mosaic virus. Only the BCMV, WMV and ZYMV clusters include Australian isolates, which have divergences from the nearest overseas isolates of only 0.03, 0.0015 and 0.002 ud/s, respectively. The numbers of cCPs in each cluster are in parentheses. The other unnamed species on long branches are, clockwise, Impatiens flower break virus between BCMV and Diuris virus Y; East Asian Passiflora virus and Wisteria vein mosaic virus between Diuris virus Y and SbMV; Florida passionflower virus between WMV and BCMNV; passionfruit woodiness virus-Thailand (two isolates), Tricyrtis virus Y and Fritillary virus Y between CAbMV and Passiflora foetida virus Y

Australia over the past two centuries, and their phylogenies show that they are part of their respective world populations. By contrast, the others entered Australia before Europeans arrived, they are much more divergent, and most are from a single lineage of potyviruses. Our

Table 1 Cosmopolitan potyviruses: viruses found both in Australia and overseas

Virus	Number of Australian sequences	Divergence[a] of Australian population	Divergence between of Australian and world population	Divergence of overseas population
Apium virus Y[b]	3	0.026	ND	ND
Bean common mosaic virus	2	–	0.030	0.089
Bean yellow mosaic virus	13	M[c]	0.003	0.159
Celery mosaic virus	4	0.002	0.020	0.002
Clover yellow vein virus	4	M	0.021	0.158
Johnson grass mosaic virus	25	0.042	0.075	0.009
Ornithogalum mosaic virus	1	–	0.005	0.190
Papaya ringspot virus	6	0.014	0.026	0.114
Pea seed-borne mosaic virus	1	–	0.003	0.083
Peanut mottle virus	1	–	0.025	0.020
Pleione virus Y	1	–	ND	ND
Sugarcane mosaic virus	12	0.011	0.015	0.138
Sweet potato feathery mottle virus				
Strain C	5	0.008	0.024	0.054
Strain RC	5	M	0.006	0.017
Sweet potato virus Y	5	0.004	0.011	0.135
Turnip mosaic virus	3	M	0.011	0.152
Watermelon mosaic virus	1	–	0.014	0.048

[a] Divergences were calculated as the mean pairwise nucleotide differences/site of sequences linked through the chosen node in NJ trees of the cCP sequences [10]

[b] Apium virus Y has been found in New Zealand, and the partial cVP sequence (EU127499) of one isolate is 0.03 ud/s divergent from Australian isolates. Partial cCP sequences have also been reported for Euphorbia ringspot virus (one sequence 0.02 ud/s from only overseas sequence); potato virus Y (three partial sequences); zucchini yellow mosaic virus (four sequences)

[c] M, phylogeny indicates that there have been multiple incursions

conclusions confirm and expand on those of [34] who studied potyviruses of Western Australia.

How is it that so many potyviruses have entered Australia recently, despite strict quarantine measures? The numbers of migratory aphids arriving from Southeast Asia are unlikely to have dramatically increased over the past two centuries, so the change is likely to be associated directly with human activity, such as trade. Some of the viruses, such as those infecting potatoes, sugarcane and sweet potatoes, were probably imported in live plants, tubers, etc., whereas others probably arrived in seeds; however, for many of these, there is no experimental record of seed transmission. There is, for example, no report of the seed transmission of PRSV, either in cucurbits or papaya; however, the phylogeographic evidence [13] indicates that PRSV probably originated in India and spread through Southeast Asia to the eastern Pacific islands, and a separate lineage migrated to the Americas, and recently from there to Australia, Hawaii and Taiwan. The most credible explanation of these transoceanic journeys is seed transmission in cucurbit seed.

Seed transmission tests usually involve a small number of seeds, far fewer than those needed to establish a crop, and so the few infected seeds required to establish infection in a crop might not be detected in tests of small samples of large seed populations. Wheat streak mosaic tritimovirus (WSMV) provides a cautionary tale. Much field and laboratory research has been done on WSMV, especially in North America, yet there had been only one report of its transmission in maize seed until recent thorough experiments showed that it is transmitted to a small but significant proportion of wheat seed [21]. This observation explained its recent arrival and subsequent spread in Australia [7].

Our assessment that at least 18 potyviruses have entered Australia in the past two centuries suggests that the long-term rate of entry and establishment of significant potyviruses in Australia is at least 0.9 incursions/decade. Interestingly, this is the same rate as that reported for the U.K., where pepino mosaic virus, PVY-NTN and ZYMV have entered and become established in the U.K. in the past 34 years [20]. Over the same period, 23 other viruses were recorded for the first time in the U.K., but these account for only 11% of all the newly recorded plant pathogens, two-thirds of which were fungi. Thus, it is likely that the total numbers of quarantine breaches is very large.

Effective biosecurity planning requires an understanding of both the dynamics and the sources of emerging

Table 2 Potyviruses found only in Australia

Virus	Number of Australian sequences	Divergence[a] of Australian population	Species of the BCMV lineage
Carrot virus Y[b]	3	0.007	N
Ceratobium mosaic virus	5	0.061	Y
Clitoria virus Y	1	–	Y
Dianella chlorotic mottle virus[c]	1	–	Y
Diuris virus Y	1	–	Y
Eustrephus virus Y	1	–	Y
Glycine virus Y[c]	1	–	N
Hardenbergia mosaic virus	30	0.159	Y
Hibbertia virus Y	1	–	Y
Kennedya virus Y	1	–	Y
Passiflora foetida virus Y	1	–	Y
Passion fruit woodiness—NSW	3	0.037	Y
Passion fruit woodiness—WA/Q[d]	5	0.041	Y
Pterostylis virus Y	2	0.011	N
Rhopalanthe virus Y	1	–	N
Sarcochilus virus Y	1	–	Y
Siratro 1 virus Y	1	–	Y
Siratro 2 virus Y	1	–	Y

[a] Divergences were calculated as the mean pairwise nucleotide differences/site of sequences linked through the chosen node in NJ trees of the cCP sequences [10]

[b] Carrot virus Y infects carrot crops and is probably a recent immigrant

[c] Partial cCP sequence

[d] Four sequences from WA and one closely related sequence isolated in Queensland

The single cCP sequences of Clitoria virus Y, Diuris virus Y, Eustrephus virus Y, Hibbertia virus Y, Kennedya virus Y, Passiflora foetida virusY, Rhopalanthe virus Y, Sarcochilus virus Y, Siratro 1 virus Y and Siratro 2 virus Y had identities of 84, 78, 83, 82, 74, 89, 80, 83, 86 and 84% to the nearest sequences in BLAST searches of the Genbank database; the partial single cCP sequences of Dianella chlorotic mottle virus and Glycine virus Y had 82 and 84% identities with the nearest sequences

pathogens. It is important to know whether new epidemics are of known pathogens that have extended their geographical range or host range or pathogens that have crossed the "ancient ecosystem-recent agrosystem interface" [34] and now infect crops. Most of the potyviruses found in Australian crops fall into the latter class, having recently extended their geographical range. It seems likely that, over the past two centuries, only two of them (BYMV and JGMV) have subsequently crossed the agrosystem-ecosystem interface (see Electronic Supplementary Material), whereas three potyviruses (PFWV-WA/Q, PFWV-NSW and PfoVY) have made the reverse journey.

All these emergences are genetically conservative, as none have involved large host range changes, except those to *Passiflora* species, but it seems that species of this genus have an inherent susceptibility to viruses of the BCMV lineage. Our analyses therefore indicate that biosecurity planning for potyviruses should focus on means to identify new species, rather than variants of known species, and support the findings of [37], who made an analysis of the progressive accumulation of information about potyvirus gene sequences in the international databases.

Coda

Our analyses support the hypothesis stated by Neville White in 1973. New evidence agrees with his suggestion that most of the crop potyviruses then known in Australia had been introduced from overseas since Europeans entered Australia and established agriculture. However, recent studies have also shown that there are many potyviruses in native plants and that most of these probably entered Australia before Europeans.

References

1. Brunt AA, Crabtree K, Dallwitz M, Gibbs A, Watson L (1996) Viruses of plants. C. A. B. International, UK
2. Buchen-Osmond C, Crabtree K, Gibbs A, McLean G (1988) Viruses of plants in Australia. Australian National University, Canberra
3. Carver M, Gross GF, Woodward TE (1991) Hemiptera (bugs, leafhoppers, cicadas, aphids, scale insects, etc). In: Naumann ID (ed) The insects of Australia, vol 1. Melbourne University Press, Melbourne
4. Chen J, Chen J, 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
5. Clemens WA, Richardson AJ, Baverstock PR (1989) Biogeography and phylogeny of the Metatheria. In: Walton DW, Richardson BJ (eds) Fauna of Australia vol IB Mammalia. Australian Government Publishing Service, Canberra, pp 527–548
6. Crisp MD (2001) Historical biogeography and patterns of diversity in plants, algae and fungi: introduction. J Biogeogr 28:153
7. Dwyer GI, Gibbs MJ, Gibbs AJ, Jones RAC (2007) Wheat streak mosaic virus in Australia: relationship to isolates from the Pacific Northwest of the USA and its dispersal via seed transmission. Plant Dis 91:164–170

8. Eastop V (1966) A taxonomic study of Australian Aphidoidea (Homoptera). Aust J Zool 14:399–592

9. Fauquet CM, Mayo MA, Maniloff J, Desselberger U, Ball LA (2005) Virus taxonomy: classification and nomenclature of viruses. 8th report of the international committee on the taxonomy of viruses. Elsevier-Academic Press, New York

10. Fourment M, Gibbs MJ (2006) PATRISTIC: a program for calculating patristic distances and graphically comparing the components of genetic change. BMC Evol Biol 6:1

11. Gibbs A, Guy P (1978) How long have there been plant viruses in Australia? Aust Plant Pathol 8:41–42

12. Gibbs AJ, Mackenzie AM, Gibbs MJ (2003) The 'potyvirid primers' will probably provide phylogenetically informative DNA fragments from all species of Potyviridae. J Virol Methods 112:41–44

13. Gibbs AJ, Ohshima K, Phillips MJ, Gibbs MJ (2008) The pre-history of potyviruses: their initial radiation was during the dawn of agriculture. PLoS ONE (in press)

14. Gibbs MJ, Wayper P, Fourment ML, Wood JT, Ohshima K, Armstrong JS, Gibbs AJ (2006) The variable codons of H3 influenza A virus haemagglutinin genes. Arch Virol 52:11–24

15. Guindon S, Gascuel O (2003) PhyML—a simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52:696–704

16. Hancock JM (1995) The contribution of slippage-like processes to genome evolution. J Mol Evol 41:1038–1047

17. Hasegawa M, Kishino H, Yano T-A (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J Mol Evol 22:160–174

18. Holmes FO (1951) Indications of a new world origin of tobacco mosaic virus. Phytopathology 41:341–349

19. Jeanmougin F, Thompson JD, Gibson TJ, Gouy M, Higgins DG (1998) Multiple sequence alignment with Clustal X. Trends Biochem Sci 23:403–405

20. Jones DR, Baker RHA (2007) Introductions of non-native plant pathogens into Great Britain, 1970–2004. Plant Pathol 56:891–910

21. Jones RAC, Coutts BA, Mackie AE, Dwyer GI (2005) Seed transmission of wheat streak mosaic virus shown unequivocally in wheat. Plant Dis 89:1048–1050

22. Kiratyi-Angul S, Gibbs A (1992) Group-specific antiserum detects potyviruses in several wild plant species in S.E. Australia. Aust Plant Pathol 21:118–119

23. Lanave C, Preparata G, Saccone C, Serio G (1984) A new method for calculating evolutionary substitution rates. J Mol Evol 20:86–93

24. Martin DP, Williamson C, Posada D (2005) RDP2: recombination detection and analysis from sequence alignments. Bioinformatics 21:260–262

25. McLean GD, Price LK (1984) Virus, viroid, mycoplasma and rickettsial diseases of plants in Western Australia. Technical Bulletin. Department of Agriculture, Western Australia, Perth, vol 68, p 22

26. Murray RGE, Stackebrandt E (1995) Taxonomic Note: implementation of the provisional status Candidatus for incompletely described procaryotes. Int J Syst Bacteriol 45:186–187

27. Robaglia C, Durand-Tardif M, Tronchet M, Boudazin G, Astier-Manifacier S, Casse-Delbart F (1989) Nucleotide sequence of potato virus Y (N Strain) genomic RNA. J Gen Virol 70:935–947

28. Sampson PJ, Walker J (1982) An annotated list of plant diseases in Tasmania. Department of Agriculture, Tasmania, p 121

29. Shukla DD, Ward CW, Brunt AA (1994) The potyviridae. CAB International, Wallingford

30. Stackebrandt E, Frederiksen W, Garrity GM, Grimont PAD, Kämpfer P, Maiden MCJ, Nesme X, Possello-Mora R, Swings J, Trüper HG, Vauterin L, Ward A, Whitman WB (2002) Report of the ad hoc committee for the re-evaluation of the species definition in bacteriology. Int J Syst Evol Microbiol 52:1043–1047

31. Swofford DL (2002) PAUP*: phylogenetic analysis using parsimony (*and other methods), version 4. Sinauer Associates, Sunderland

32. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 24:4876–4882

33. Ward CW, Weiller G, Shukla DD, Gibbs AJ (1995) Molecular systematics of the potyviridae, the largest plant virus family. In: Gibbs AJ, Calisher CH, Garcia-Arenal H (eds) Molecular basis of virus evolution. Cambridge University Press, Cambridge, pp 477–500

34. Webster CG, Coutts BA, Jones RAC, Jones MGK, Wylie SJ (2007) Virus impact at the ancient ecosystem-recent agroecosystem interface: studies on three legume-infecting potyviruses in the Southwest Australian floristic region. Plant Pathol 56(5):729–742

35. Weiller GF (1998) Phylogenetic profiles: a graphical method for detecting genetic recombinations in homologous sequences. Mol Biol Evol 15:326–335

36. White NH (1973) The geographic isolation of Australia in relation to plant viruses. Aust Plant Path Soc Newsl 2:32–33

37. Zheng L, Wayper PJ, Gibbs AJ, Fourment M, Rodoni BC, Gibbs MJ (2008) Accumulating variation at conserved sites in potyvirus genomes is driven by species discovery and affects degenerate primer design. PLoS ONE 3:e1586