Virus Research 141 (2009) 169–173

Contents lists available at ScienceDirect

Virus Research

journal homepage: www.elsevier.com/locate/virusres

Non-cultivated plants of the Tallgrass Prairie Preserve of northeastern Oklahoma frequently contain virus-like sequences in particulate fractions

Vijay Muthukumar[a], Ulrich Melcher[a,*], Marlee Pierce[a], Graham B. Wiley[b], Bruce A. Roe[b], Michael W. Palmer[c], Vaskar Thapa[c], Akhtar Ali[d], Tao Ding[a]

[a] Department of Biochemistry and Molecular Biology, Oklahoma State University, Stillwater, OK 74078, USA
[b] Department of Chemistry & Biochemistry, University of Oklahoma, Norman, OK 73019, USA
[c] Department of Botany, Oklahoma State University, Stillwater, OK 74078, USA
[d] Department of Biological Science, University of Tulsa, Tulsa, OK 73401-3189, USA

ARTICLE INFO

Article history:
Available online 17 January 2009

Keywords:
Plant virus
Metagenomics
Biodiversity
Retroelements
Tymovirus

ABSTRACT

The diversity of viruses associated with non-cultivated plants was assessed from plant samples collected in the Tallgrass Prairie Preserve of northeastern Oklahoma, USA. The samples were processed to determine the sequences of nucleic acids extracted from the virus-like particle fraction of plant homogenates. Sequences from 95 specimens of 52 plant species included those of probable origin from the genomes of plants (including retroelements), bacteria, fungi, other organisms, and viruses. Virus-like sequences were identified in sequences from 25% of the specimens, coming from 19% of the plant species. Evidence of a member of the genus *Tymovirus* was found in 16 specimens of 6 plant species, making it the most predominant virus associated with the sampled plants. There was evidence of the presence of more than one virus in each of six specimens.

1. Introduction

Viruses associated with plants are highly diverse with respect to the strategies they have evolved for transmission, replication and dispersal. Harrison (1981) categorized plant viruses into two groups based on their adaptation to existence primarily in cultivated plant species or primarily non-cultivated ("wild") species. "CULPAD" viruses are adapted to existence in cultivated species and do not require association with wild plants. In this view, "WILPAD" viruses are those whose primary existence is in wild plants. Harrison further suggested that some viral genera consist mostly of CULPAD viruses and other genera contain principally WILPAD viruses. Whether the spectrum of viruses associated with wild plants differs from that associated with cultivated plants cannot be assessed at present because the large majority of recognized plant viruses, including WILPAD viruses, have been discovered due to their causing disease in cultivated plants (Wren et al., 2006). The infection by viruses from introduced cultivated plants of native non-cultivated plants and the infection of introduced cultivated plants by viruses from native non-cultivated plants has been noted and discussed (Webster et al., 2007).

One way to test for a discordant distribution of viruses or viral genera in cultivated and non-cultivated plants is to survey the latter for evidence of the presence of viruses. Surveys for specific known viruses in non-cultivated plant populations have been reported (Cooper and Jones, 2006). For example, four viruses, members of the genera *Luteovirus*, *Caulimovirus*, *Potyvirus* and *Tymovirus* were found to be common in non-cultivated plants (Raybould et al., 1999). Previously unknown viruses have been discovered in obviously diseased non-cultivated plants (Ooi et al., 1997; Robertson, 2005). A survey for viruses (Wren et al., 2006) that examined specimens, not selected for symptom presence, was initiated at the Tallgrass Prairie Preserve (TPP) of northeastern Oklahoma (Hamilton, 2007). The Preserve consists of approximately 15 000 ha of unbroken tallgrass prairie, cross-timbers woodland and riparian areas. It hosts over 700 species of plants (Palmer, 2007). In the present survey, plant samples were screened for virus-like sequences in double-stranded RNA and in nucleic acids associated with particulate fractions of plant homogenates. The approach (virus-like particle-viral nucleic acid; VLP-VNA) has recently been described and its successful application to specimens of *Ambrosia psilostachya* was reported (Melcher et al., 2008). Application of this method to a further 95 specimens of 52 species of TPP plants is the subject of this report.

2. Experimental/Materials and methods

Specimen location, harvest, and homogenization, isolation of virus-like particle (VLP) fractions, extraction of nucleic acid from VLP to give VLP-VNA and its amplification, cloning and sequenc-

* Corresponding author. Tel.: +1 405 744 6210; fax: +1 405 744 7799.
E-mail address: u-melcher-4@alumni.uchicago.edu (U. Melcher).

0168-1702/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.virusres.2008.06.016

ing were performed as previously described (Melcher et al., 2008; Muthukumar, 2008). Briefly, leaf tissue was obtained from plants of the TPP (Hamilton, 2007) with locations and dates of harvest recorded and identified as to the plants' species (Palmer, 2007). Specimen collection was aimed at obtaining a specimen of each available species and multiple specimens of a select list of particularly abundant species. Collection was done without reference to symptoms and covered most areas of the TPP. Aliquots of tissue (100 mg) were homogenized by beating with glass beads in a citrate buffer. Following clarification of the homogenate by low speed centrifugation, VLP fractions were obtained by two ultracentrifugations over 20% sucrose pads. Resuspended VLP pellets were first treated with DNaseI to destroy DNA not encapsidated in particles and subsequently with proteinase K and SDS. The mixtures were extracted with phenol and the nucleic acids precipitated using ethanol. Amplification followed the procedure of Wang et al. (2002) and was followed by agarose gel electrophoresis. Products of successful amplifications were subjected to TA cloning and the sequences of inserts in approximately 100 plasmid clones per specimen were determined.

For the work reported here, insert sequences were used in BLASTn and BLASTx searches (Altschul et al., 1997) of the GenBank general databases (nt and nr) to categorize the sequences. Sequences and the results of BLASTn and BLASTx searches are available by ftp (ftp://ftp.genome.ou.edu/pub/PVBE/vijay_data) from the project's sequence website (http://www.genome.ou.edu/pvbe.html). Sequences identical to those of *Moloney murine leukemia virus* were excluded as these could have been contributed by the reverse transcriptase reagent used in amplification. Data handling and storage used MS Excel and MySQL. Statistical analysis was performed through SAS and included Tukey's Studentized Range Test and the *F*-test.

3. Results

3.1. Plant harvest and VNA amplification

Of the 687 plant specimens processed using the VLP-VNA method and subjected to PCR amplification, 59% produced amplification products visible in agarose gel electrophoresis (example in Fig. 1). The rest did not produce detectable products. Seventeen species were represented by multiple specimens. For only one of these species, *Medicago lupulina*, was amplification obtained for all specimens (Fig. 2). However, all species produced amplified products from at least one specimen. Only the VNA's isolated from plants that were PCR-positive were subjected to cloning and sequencing. Among the plant specimens analyzed for amplifiable VLP-VNA, only 2.3% were noted at collection as having any symptoms of disease. In each year of harvest the proportion of samples that were PCR-positive was the same among plants with and without symptoms. Thus, presence or absence of amplification was not an indicator of disease, manifested as symptoms.

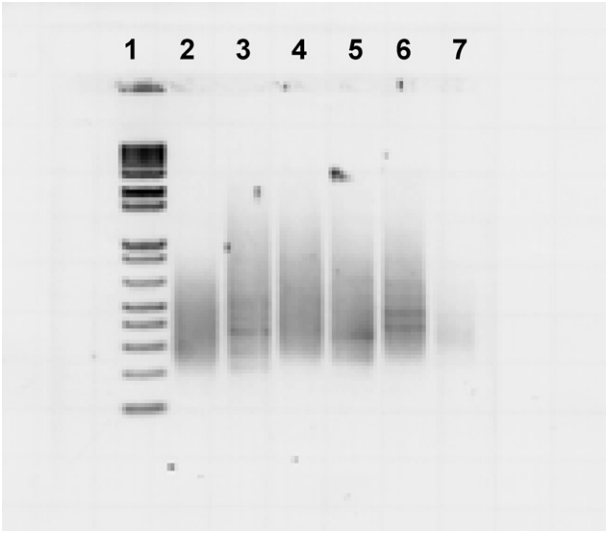

Fig. 1. Random PCR amplification of five plant specimens collected from the Tallgrass Prairie Preserve. Lane 1: 1 kb + ladder (InVitrogen), lane 2: *Ambrosia psilostachya* (05TGP00295), lane 3: *Asclepias viridis* (05TGP00248), lane 4: *Asplenium rhizophyllum* (05TGP00289), lane 5: *Dichanthelium oligosanthes* (05TGP00162), lane 6: *Panicum virgatum* (05TGP00296), lane 7: negative control (nuclease-free water).

3.2. Sequence classification of contigs—distributions

Of the 405 plant specimens yielding amplification, 95 have to date been processed to provide sequences from 52 species of plants. Four species were represented by multiple samples: two forbs *Ambrosia psilostachya* (5 specimens) and *Asclepias viridis* (16 specimens), and 2 grasses *Sorghastrum nutans* (8 specimens) and *Panicum virgatum* (18 specimens). The following analyses refer to these 95 samples.

Sequence reads were categorized as virus-like, retrovirus-like, bacteria-like, fungus-like, plant-like, animal-like, other or uncertain (Fig. 3). The "other" category consists of sequences whose best hits are to eukaryotic organisms other than plants, fungi and animals. The "uncertain" category consists of sequences that failed to recognize any database sequence under the conditions used. The largest proportion of reads (28.5%) was classified as uncertain because they could not be assigned to a source with any certainty. These sequences might be plant-derived mRNA sequences from genes that are not highly conserved in the genomes of sequenced higher plants. To test this possibility, the ratio of uncertain reads to plant-like reads was examined for three multiple specimen plant species with enough reads in these categories for statistical analysis. The ratios for *S. nutans* and *P. virgatum* were significantly different ($p < 0.05$) from the ratio for *A. psilostachya* by Tukeys' Studentized Range Test. Among the species for which only a single sample was isolated, *Vitis* sp. did not have any uncertain sequence reads, consistent with the completion of a *V. vinifera* genome project (Velasco et al., 2007), while very few reads were identified as from a plant source for the non-vascular plant *Chara globularis*.

Retrovirus-like sequences could originate from actual retrovirus-like particles or from retrotransposons known to be abundant in many plant genomes. The ratios of retrovirus-like sequence reads to the total of plant-like and uncertain reads were calculated. Among species with only a single representative (29), 16 lacked retrovirus-like sequences among the VLP-VNA sequences retrieved. In the other 13, the percentage of retrovirus sequence ranged from 2 to 18%. Comparison of these percentages for species with multiple specimens suggested that only one comparison, that between *A. psilostachya* and *P. virgatum*, produced a difference significant at $p < 0.05$. For each of the four species a wide range was found for individual samples. The variation was much larger than for the percent of uncertain and plant-like sequences that were plant-like. *F*-test evaluation showed that variances in the retrovirus-like ratios were significantly larger ($p < 0.001$) than those for the plant-like sequence ratios for *A. psilostachya*, *A. viridis*, and *S. nutans*.

The developed technique also yielded sequences highly similar to bacterial and fungal sequences. Variability, similar to that seen in the proportions of retrovirus-like sequences was observed also in the percentage of total reads due to bacteria-like sequences. For example, in *A. psilostachya*, four plants had from 0 to 2.1% bacteria-like reads, but one had 93.6% attributable to that category of source. For *P. virgatum*, values ranged from 0 to 55.5%. For those *A. viridis* specimens in which there were no abundant virus-like sequences

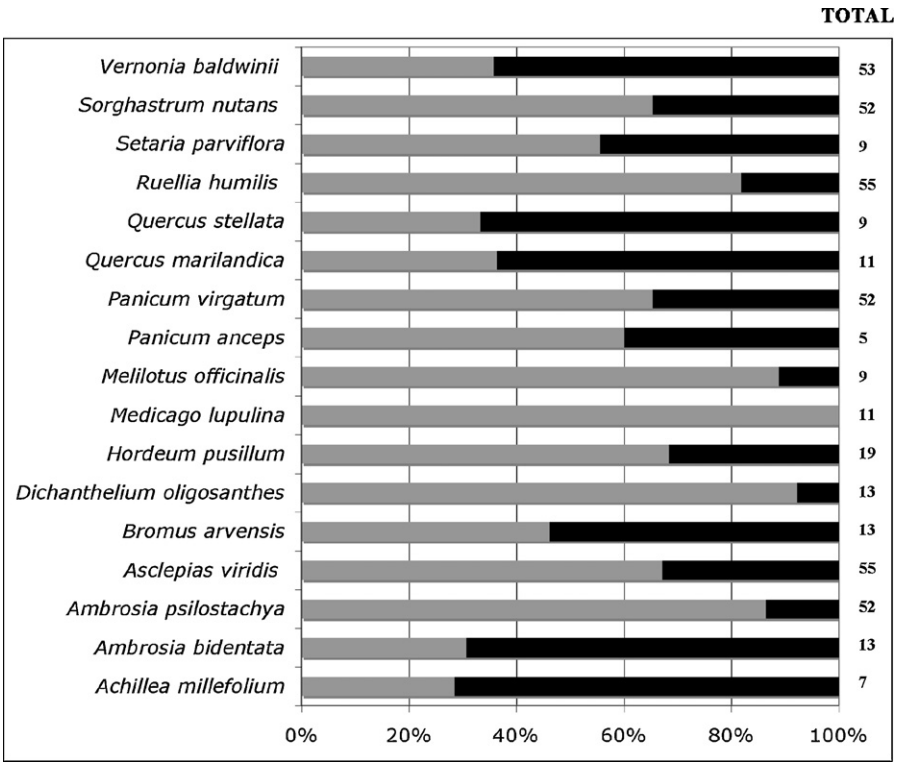

Fig. 2. PCR amplifiability of nucleic acids in virus-like particle fractions of plants from the Tallgrass Prairie Preserve. Shown are the percentages of specimens of selected plant species that exhibited (grey) or did not exhibit (black) evidence of PCR products as observed by agarose gel electrophoresis. The numbers of specimens examined are indicated to the right of each species bar.

to mask the other sequences present, one plant had 0.5% bacteria-like sequences, while the others had 22, 29, 97 and 100% for the reads from this source category. Thus the possible association of bacteria with these plants was not a uniform feature. Plants with high levels of bacteria-like sequences appeared to be located in a swath of sites on the TPP, but were not correlated with known geographical features. Most of the putative bacterial sequences were similar to ribosomal sequences. The principal genera appearing as top hits among bacteria-like sequences were *Azoarcus*, *Burkholderia*, *Dechloromonas*, *Deinococcus*, *Escherichia*, *Magnetospirillum*, *Ralstonia*, *Salmonella*, *Shigella*, *Streptococcus* and *Vibrio*.

Overall, fungus-like sequences accounted for 2.5% of all the reads encountered, substantially less than the 28% seen for bacteria-like sequences. Most plants did not yield any fungus-like sequences. Only four plants had more than 10% of the sequence reads as fungus-like sequences. The highest (32.7%) was seen in a single specimen of *P. virgatum* and was dominated by a relative of *Pycnoporus*. As with bacteria-like sequences, the predominant functional type of nucleic acid found was ribosomal. The other large contributions of fungus-like sequences were seen in *Ludwigia palustris* (11.5%) where a *Kluyveromyces* species most resembled the recovered sequence, in *Dioidia teres* (22.6%) and in one specimen of *S. nutans* (11.8%), with *Phaeosphoria*, a genus containing grass pathogens, being the predominant best hit species for both of the latter.

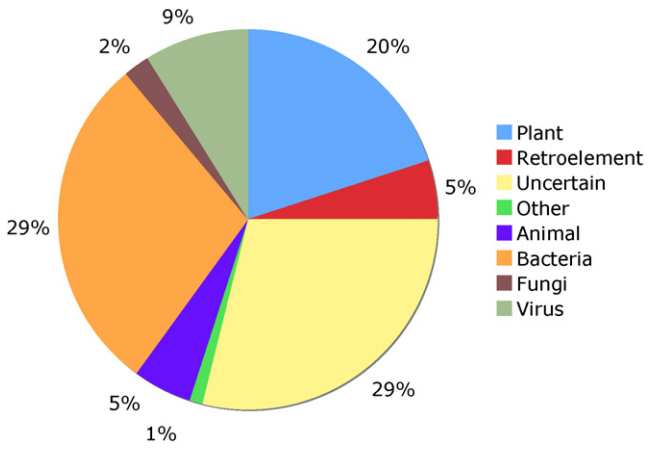

Fig. 3. Distribution among likely sources of nucleotide sequence reads obtained from virus-like particle fractions of 95 plant specimens from the Tallgrass Prairie Preserve.

3.3. Virus-like sequences

Virus-like sequences were identified by BLASTn and BLASTx searches of the general databases in sequences from 25% (24 of 95) specimens. They came from 19% (10 of 52) of the plant species included in the study. Within the four species that were sampled multiple times, the incidence of virus-like sequences varied: 17% for *P. virgatum*, 29% for *S. nutans* and 60% for *A. psilostachya* and *A. viridis*. It is notable that virus-like sequences were not totally absent from any of these species and neither did every specimen of a species yield virus-like sequences.

Table 1 lists the putative taxonomic assignments of virus-like sequences. Several virus-like sequences were found belonging to virus taxa not known to infect plants. With one exception, these identifications resulted from *E*-values close to the threshold for reporting (10^{-4}) and require further examination for confirmation. The exception is a chrysovirus-like sequence, identified by BLASTx with an *E*-value of 5×10^{-27}, from a fern, *Asplenium rhizophyllum*. Recognizing that chrysoviruses have so far been only isolated from fungi, it is notable that no sequences similar to those of fungi were identified from this plant specimen. Sequences related to known plant viral genera mapped to *Badnavirus* (two plant species),

Table 1
Categorization of virus-like sequences recovered from Tallgrass Prairie Preserve plant specimens.

Virus family	Virus genus	Plant species	No. of plant specimens	No. of reads	Read %
Ascoviridae	*Ascovirus*[a]	*Artemisia ludoviciana*	1	3	2.0
Caulimoviridae	*Badnavirus*	*Ambrosia psilostachya*	2	10	1.3
		Sorghastrum nutans	1	3	1.0
Chrysoviridae	*Chrysovirus*	*Asplenium rhizophyllum*	1	27	11.9
Comoviridae	*Comovirus*	*Asclepias viridis*	1	4	0.7
Flexiviridae	*Unassigned*	*Ambrosia psilostachya*	2	194	25.4
Paramyxoviridae	*Metapneumovirus*[a]	*Desmanthus illinoensis*	1	3	1.5
Phycodnaviridae	*Coccolithovirus*[a]	*Festuca subverticillata*	1	2	0.8
		Sorghastrum nutans	1	2	0.8
Ranaviridae	*Ranavirus*[a]	*Panicum virgatum*	1	2	0.7
Siphoviridae	*Lambda-like phage*[a]	*Ambrosia psilostachya*	1	2	0.4
Tymoviridae	*Tymovirus*	*Amorpha fruticosa*	1	2	0.7
		Asclepias viridis	10	3154	85.7
		Asplenium rhizophyllum	1	33	14.6
		Cephalanthus occidentalis	1	4	1.1
		Desmanthus illinoensis	1	7	3.5

[a] Identification based on *E*-values in BLASTx searches of the general database of 1×10^{-6} to 3×10^{-4}. All others had *E*-value less than 10^{-10}.

Comovirus, and *Tymovirus* (six plant species) and an unassigned genus in the *Flexiviridae* (Melcher et al., 2008). Previous analysis has shown that the *Flexiviridae*-like sequences originated from two distinct sequences, implying two separate viruses. Preliminary analysis of the multiple *Tymovirus*-related sequences suggests that they all arose from variants of the same virus.

The percentage of total sequence reads, per specimen, that appeared to be virus-derived was taken as a measure of the titre of the putative virus in the plant. The putative tymovirus was in high titre in most of the *A. viridis* specimens positive for this virus. The putative member of the *Flexiviridae* in *A. psilostachya* was present in moderate titre, while all other identifications were from putative viruses present in low titre. Many of the latter were present in specimens that had another virus in relatively high titre. Apparent tymovirus infections were associated with the *Metapneumovirus*-like sequence in *D. illinoensis*, the *Chrysovirus*-like sequences *in A. rhizophyllum* and the *Comovirus*-like sequences in *A. viridis*. The putative *Flexiviridae* member was accompanied by a *Badnavirus*-like sequence in *A. psilostachya* (Melcher et al., 2008). Overall, 6 of the 24 plants containing virus-like sequences had sequences similar to more than one virus genus. The data suggest that non-cultivated plants are frequently infected with multiple viruses.

4. Discussion and conclusion

4.1. Amplification of VLP-VNA

Not all plant specimens yielded VLP-VNA fractions with nucleic acids that could be amplified by the protocol employed. The absence of amplicons was unlikely to be due to the inherent properties of the plant species since, in all cases where multiple specimens of the same species were tested, at least some produced detectable amplification products. The absence of amplicons also cannot be taken as evidence of the absence of nucleic acid in the preparations. It is possible that nucleic acids are present in amplification-negative samples at concentrations too low to give products visible in the agarose gels under the limited number of amplification cycles used. Alternatively, the VNA preparations could have contained contaminants that inhibited the PCR reactions. Additional work using higher cycle numbers in amplification or further diluted templates will be needed to resolve this issue.

4.2. Categorization of sequences

For many plant specimens, a large percentage of the sequences retrieved were labelled “uncertain” in view of their failure to produce “hits” in BLAST searches of the general databases. It is likely that most of these were derived from the plant genome for several reasons. The proportion of uncertain sequences to recognized plant sequences was species-dependent and relatively constant among multiple specimens of the same species. BLASTn searches of the EST database with selected uncertain sequences resulted in a few significant hits. All were to plant EST sequences (data not shown). No uncertain sequences were identified from a member of one genus (*Vitis*) with an available complete genome sequence for a member, while the highest ratio of uncertain to plant-like sequences was from a taxon (*Chara*) which is under-represented among genome projects. Nonetheless, it is possible that a few uncertain sequences represent viruses whose sequences are completely unrelated to those of known viruses. A major fraction of virus sequences obtained from environmental samples showed no significant nucleotide and amino acid similarities to any sequences available in Genbank at the time (Edwards and Rohwer, 2005). The origin of these sequences could be of great interest as they could represent novel and highly distinct viruses.

Retrotransposon sequences were abundant in many VLP-VNA specimens. Some such sequences were more closely related to those of authentic viruses than to retroelements and were classified as virus sequences. The origin of the non-viral retrovirus-like sequences, abundant in some VLP-VNA preparations is uncertain. Retrotranposons are abundant in many genomes (Vitte and Panaud, 2005). Yet, it is doubtful that most of the retrovirus-like sequences detected were part of plant genomic DNA contained in the VLP-VNA fraction. In plant species for which multiple specimens were examined, the ratio of retrovirus-like sequences to the total of plant-like and uncertain sequences was much more variable than the ratio of uncertain to plant-like sequences. For example, retrovirus-like sequences were not detected in many specimens of *P. virgatum* whereas others of the same species had up to 27% of the recovered sequences in that category. These observations suggest that retroelements in these plants are variously active in response to some unknown condition or signal. The data do not allow distinction between two explanations for the high abundance of these sequences among those recovered from VLP-VNA. They simply may reflect a high level of transcription of retroelement genes such that recovery of these sequences represents their proportion of total mRNA. Alternatively, they may reflect production of endogenous retroviral particles.

In drawing the above conclusions some specimens were ignored because their content of microbial sequences was so large that the background of plant-derived sequences was severely depressed, prohibiting significant comparisons among plant-like, uncertain

and retrovirus-like categories. In addition to viruses, these microbe-like sequences putatively originated from bacteria and fungi associated with the plants. These sequences are referred to as having putative bacterial or fungal sources, since the microbes have not been isolated and characterized. Since multiple specimens of the same plant species varied widely in their content of putative microbial sequences, these sequences likely were not from obligate endophytes. Whether the putative microbes were facultative endophytes or surface-adsorbed microbes could not be determined with the procedures used. The sequences were principally ribosomal RNA sequences. Their appearance in the VLP-VNA fraction could be due to failure to remove completely small fungal cells and bacteria by centrifugation prior to the ultracentrifugal pelleting of VLP.

4.3. Viruses

As with bacteria and fungi, the identification of virus-like sequences is only indicative of the presence of viruses in the plants from which they were extracted and may not indicate a viral infection of the plant. Particularly distant relationships with viruses not known to infect plants merit careful scrutiny. Confirmation of the presence of a virus requires isolation, transmission, and reisolation. Such confirmation is in progress for selected putative viruses. It should also be recognized that the assignment of multiple sequences from a plant specimen to a single virus genus does not necessarily mean that there was a single virus species whose genome contains all the sequences reported. Such a conclusion can only be accepted after purification and characterization of the genome of the purified virus. Nevertheless, the simplest assumption is that a single virus of that genus is present, especially if identity of sequence is found in substantial overlap regions.

Three features of the virus-like sequence data stand out. First, among these sequences, the only ones that resembled closely the sequences of known viruses were those attributable to a reagent contamination (*Moloney murine leukemia virus*, not included in Table 1). The remainder, although assignable to recognized viral genera, had hitherto not been reported, which supports the view (Wren et al., 2006) that the number of virus species in the world is underestimated grossly by our current knowledge of viruses. Second, evidence of a variety of viruses was uncovered. Icosahedral (*Tymovirus*) and filamentous (*Flexiviridae*) particle morphologies were represented as were viruses with DNA (*Badnavirus*) as well as RNA (*Comovirus*, *Flexiviridae* inter alia) genomes. The method was not expected to detect viroids or viruses that do not form particles, such as members of the *Endornavirus* genus. Third, it is striking that what is apparently one VLP-VNA virus, a member of the genus *Tymovirus*, has such a dominant presence at the TPP. In several specimens of *A. viridis*, its sequences constitute all or close to all of the sequences retrieved. The next most abundant putative virus is that of an unclassified member of the *Flexiviridae* (Melcher et al., 2008) whose sequence constituted only one-quarter of those recovered. The remaining putative identifications were based on small numbers of reads from the specimens. Viruses present in plants at less than 1% of the titre of the tymovirus may escape detection by the VLP-VNA method without examination of substantially more clones. In the present study, evidence of a tymovirus was found in six diverse plant species, including dicotyledonous and monocotyledonous plants, legumes, and a fern. Sequence similarity (U. Melcher, unpublished) suggests strongly that variants of the same virus had infected multiple specimens and multiple species at the TPP.

The taxa putatively identified in this study were neither WILPAD nor CULPAD virus genera in the original Harrison (1981) listing. Indeed, he listed tymoviruses as having representatives of both kinds of viruses. Nevertheless, in favour of the hypothesis that the WILPAD–CULPAD distinction is a useful one, perhaps at the level of viral species, is the observation that no evidence of virus species known from diseases of crop plants was found in this study of 95 plants of 53 species at the TPP. In particular, no evidence of the agriculturally abundant potyviruses was found. Since the emergence of the *Potyvirus* genus coincided approximately with the development of agriculture (Gibbs et al., 2008), the viruses may be considered the quintessential CULPAD viruses, explaining their absence in this survey.

Acknowledgements

The authors express their appreciation to Dan McGlinn, Fumiko Shirakura, Matt Allen and Pete Earls for assistance in specimen collection, Fenqiu Zhang and Tsuyoshi Asako for specimen processing, Yan Song for database management, and Kay Scheets for review of the manuscript. This work was supported by National Science Foundation-EPSCoR award EPS-0447262, the Oklahoma State Regents for Higher Education and the Oklahoma Agricultural Experiment Station whose Director has approved the manuscript for publication.

References

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Cooper, I., Jones, R.A.C., 2006. Wild plants and viruses: under-investigated ecosystems. Adv. Virus Res. 67, 1–47.

Edwards, R.A., Rohwer, F., 2005. Viral metagenomics. Nat. Rev. 3, 504–510.

Hamilton, R.G., 2007. Restoring heterogeneity on the Tallgrass Prairie Preserve: applying the fire-grazing interaction model. In: Masters, R.E., Galley, K.E.M. (Eds.), Proceedings of the 23rd Tall Timbers Fire Ecology Conference: Fire in Grassland and Shrubland Ecosystems. Tall Timbers Research Station, Tallahassee, pp. 163–169.

Gibbs, A.J., Ohshima, K., Phillips, M.J., Gibbs, M.J., 2008. The prehistory of potyviruses: their initial radiation was during the dawn of agriculture. PLoS ONE 3, e2523.

Harrison, B.D., 1981. Plant virus ecology: ingredients, interactions and environmental influences. Ann. Appl. Biol. 99, 195–209.

Melcher, U., Muthukumar, V., Wiley, G.B., Min, B.E., Palmer, M.W., Verchot-Lubicz, J., Nelson, R.S., Roe, B.A., Ali, A., Thapa, V. Pierce, M.L., 2008. Evidence for novel viruses by analysis of nucleic acids in virus-like particle fractions from *Ambrosia psilostachya*. J. Virol. Methods.

Muthukumar, V., 2008. Metagenomics for the Identification of Plant Viruses in the Tallgrass Prairie Preserve. M.S. Thesis, Oklahoma State University.

Ooi, K., Ohshita, S., Ishii, I., Yahara, T., 1997. Molecular phylogeny of geminivirus infecting wild plants in Japan. J. Plant Res. 110, 247–257.

Palmer, M.W., 2007. Vascular Flora of the Tallgrass Prairie Preserve, Osage County, Oklahoma. Castanea 71, pp. 235–246.

Raybould, A.F., Maskell, L.C., Edwards, M.L., Cooper, J.I., Gray, A.J., 1999. The prevalence and spatial distribution of viruses in natural populations of *Brassica oleracea*. New Phytol. 141, 265–275.

Robertson, N.L., 2005. A newly described plant disease complex involving two distinct viruses in a native Alaskan lily, *Streptopus amplexifolius*. Can. J. Bot. 83, 1257–1267.

Velasco, R., Zharkikh, A., Troggio, M., Cartwright, D.A., Cestaro, A., Pruss, D., Pindo, M., Fitzgerald, L.M., Vezzulli, S., Reid, J., Malacarne, G., Iliev, D., Coppola, G., Wardell, B., Micheletti, D., Macalma, T., Facci, M., Mitchell, J.T., Perazzolli, M., Eldredge, G., Gatto, P., Oyzerski, R., Moretto, M., Gutin, N., Stefanini, M., Chen, Y., Segala, C., Davenport, C., Demattе, L., Mraz, A., Battilana, J., Stormo, K., Costa, F., Tao, Q., Si-Ammour, A., Harkins, T., Lackey, A., Perbost, C., Taillon, B., Stella, A., Solovyev, V., Fawcett, J.A., Sterck, L., Vandepoele, K., Grando, S.M., Toppo, S., Moser, C., Lanchbury, J., Bogden, R., Skolnick, M., Sgaramella, V., Bhatnagar, S.K., Fontana, P., Gutin, A., Van de Peer, Y., Salamini, F., Viola, R., 2007. A high quality draft consensus sequence of the genome of a heterozygous grapevine variety. PLoS ONE 2, e1326.

Vitte, C., Panaud, O., 2005. LTR retrotransposons and flowering plant genome size: emergence of the increase/decrease model. Cytogenet. Genome Res. 110, 91–107.

Wang, D., Coscoy, L., Zylberberg, M., Avila, P.C., Boushey, H.A., Ganem, D., DeRisi, J.L., 2002. Microarray-based detection and genotyping of viral pathogens. Proc. Natl. Acad. Sci. U.S.A. 99, 15687–15692.

Webster, C.G., Coutts, B.A., Jones, R.A.C., Jones, M.G.K., Wylie, S.J., 2007. Virus impact at the interface of an ancient ecosystem and a recent agroecosystem: studies on three legume-infecting potyviruses in the southwest Australian floristic region. Plant Pathol. 56, 729–742.

Wren, J.D., Roossinck, M.J., Nelson, R.S., Scheets, K., Palmer, M.W., Melcher, U., 2006. Plant virus biodiversity and ecology. PLoS Biol. 4, e80.