Viruses in Oklahoma *Euphorbia marginata*

Jennifer Hackett
Department of Biochemistry and Molecular Biology, 246 NRC, Oklahoma State University, Stillwater OK 74078, currently at Department of Molecular Biosciences, University of Kansas, Lawrence KS 66045

Vijay Muthukumar
Department of Biochemistry and Molecular Biology, 246 NRC, Oklahoma State University, Stillwater OK 74078, currently at Oklahoma Medical Research Foundation, Oklahoma City, OK 73104

Graham B. Wiley
Department of Chemistry and Biochemistry, 101 David L. Boren Blvd., University of Oklahoma, Norman OK 73019, currently at Oklahoma Medical Research Foundation, Oklahoma City, OK 73104

Michael W. Palmer
Department of Botany, 104 LSE, Oklahoma State University, Stillwater OK 74078

Bruce A. Roe
Department of Chemistry and Biochemistry, 101 David L. Boren Blvd., University of Oklahoma, Norman OK 73019

Ulrich Melcher
Department of Biochemistry and Molecular Biology, 246 NRC, Oklahoma State University, Stillwater OK 74078

A survey of non-cultivated plants of the Tallgrass Prairie Preserve of Oklahoma for viruses used young leaf tissue for analysis of virus sequence signature content. To assess whether tissue selection may affect the results of such a survey, Snow-on-the-mountain (*Euphorbia marginata* Pursh) plants growing along roadsides of Payne Co., Oklahoma, were sampled during the fall of 2007. Specimens of upper leaves, lower leaves, white foliar bracts and fruits were separately processed to obtain nucleic acid preparations from virus-like particle fractions of plant homogenates. Signatures of two viruses, a comovirus and a tymovirus, Asclepias asymptomatic virus (AsAV), both also found on the Tallgrass Prairie Preserve, were observed in the sequences of the nucleic acid preparations. Not all tissues of positive plants were consistently positive. In the sequences retrieved, the nucleotides at only five positions differed from those of AsAV obtained from *Asclepias viridis*. Sequences from bracts and from lower leaves of different plants were more similar to one another than to other sequences from the same plant.

INTRODUCTION

The vast majority of viruses of plants are known because they are pathogens of crop and horticultural plants (Wren and others 2006). Whether the known viruses reflect accurately the kinds of viruses present in association with non-cultivated plants has not been established. Thus, non-cultivated plants growing in The Nature Conservancy's Tallgrass Prairie Preserve (TPP) near Pawhuska OK were surveyed for nucleic

acid sequences indicative of virus presence (Melcher and others 2008, Muthukumar and others 2008). Evidence of several hundred species of viruses was discovered with almost all being previously unknown to humankind. One of the putative viruses, not previously described from any other location, tentatively named Asclepias asymptomatic virus (AsAV) and belonging to the *Tymovirus* genus, was found frequently and in abundance (Muthukumar and others 2008). It has been isolated and characterized in the laboratory (Min, Feldman, Wiley, Muthukumar, Roe, Roossinck, Ali, Melcher, Palmer and Nelson, in preparation). The virus has spherical particles containing a single strand of positive-sense RNA. It does not cause noticeable symptoms of disease when inoculated on *Asclepias viridis* plants, but, when tested on *Nicotiana benthamiana* produces leaf mosaic and distortion. Its closest known relative is *Kennedya yellow mosaic virus*. AsAV was found often in association with evidence of a second virus, TGP Comovirid 1, also a virus with spherical particles. Its genome consists of two single strands of positive-sense RNA. Infectivity of this putative virus has not been tested yet. Wherever possible in the TPP survey, young leaves, near growing tips of plants, were harvested selectively. The choice of tissue was determined by the ready accessibility of the tissue and the belief that viruses tend to accumulate best in younger leaves. The survey recognized that the sampling of young leaves might cause some viruses, such as those preferentially located in roots, to be overlooked. This communication reports an analysis of selected above-ground parts of multiple specimens of Snow-on-the-mountain (*Euphorbia marginata,* Pursh) plants growing along roadsides of Payne Co. Oklahoma in the fall of 2007.

EXPERIMENTAL PROCEDURES

Eight *E. marginata* plants were sampled on 19 September, 2007, mainly from along roadsides around Payne County, OK. Locations are given in Table 1. Samples of upper leaves, lower leaves, white foliar bracts, and fruits were harvested. The samples were processed by the virus-like particle virus-nucleic-acid (VLP-VNA) method, previously described (Melcher and others 2008). Briefly, a 100 mg sample of plant tissue was homogenized in a Bead Beater (BioSpec Products, Bartlesville, Oklahoma). Following low speed centrifugation, supernatants were centrifuged at high speeds over a sucrose pad to pellet viral particles. These were resuspended, digested with DNaseI and proteinase K before solvent extraction and alcohol precipitation. The nucleic acid was randomly amplified following the procedure of Wang *et al*. (Wang and others 2002), using four nucleotide long tags at the ends of the amplifying primers (Roossinck and others 2009), and monitoring amplification success by agarose gel electrophoresis. After ligation of adapters and additional amplification by emulsion PCR with 40 micron beads, the DNA was sequenced on the massively parallel 454/Roche GS-FLX pyrosequencer (Margulies and others 2005). After assembling the reads with the 454 de novo newbler assembler into a set of 1,748 consensus contigs containing 443,504 nucleotides, BLASTx searching (Altschul and others 1997) of the non-redundant

Table 1. Universal Transmercator (UTM) Coordinates of Sampled *E. marginata* Plants.

Plant	Northing (m)[a]	Easting (m)[a]
07JH012	665192	4000361
07JH013	664137	3998136
07JH014	664137	3998129
07JH015	664263	3998708
07JH016	664267	3998495
07JH017	664271	4003238
07JH020	664239	4003298
07JH021	664237	4002618

[a] UTM, Sector 14, using NAD27

protein sequence database resulted in identification of 64 sequence contigs resembling those of known viruses. Sequences of 63 contigs related to those of tymoviruses were aligned with one another and the consensus sequence of AsAV using CAP3 (Huang and Madan 1999) (alignment available as a supplemental file "EmAsAV.fa").

The alignment of the *E. marginata* AsAV sequences was used to identify and count phylogenetically informative positions (positions at which at least two of the four possible nucleotides were represented more than once). For each sample (representing one plant part from one plant), the string of informative nucleotides was extracted from the alignment (Figure 1) using "X" for positions not covered by the available sequence. Pairwise comparisons identified the number of identical residues / total residues compared (total informative residues minus residues where one or both samples had "X") for each possible pair.

RESULTS

Of the 32 samples, 25 produced amplified products that yielded 1,748 nucleotide sequence contigs. The sequences were used to query the nonredundant GenBank/EMBL/DDBJ nucleotide and protein sequence databases by BLASTn and BLASTx, respectively, revealing the presence of two kinds of viruses, one similar to members of the *Tymovirus* genus and the other to members of the *Comoviridae* family. The latter was represented by one sequence whose translation by BLASTx had 31% identity search to the polyprotein of *Broad bean wilt virus* of the *Fabavirus* genus. Individual fragments of the tymovirus genome exhibited 62-70% identity by BLASTx search to the virion pro-

```
S1 CCTCCCCCCCACTCTTCCAATCACCTACTTCAATCAGAAACTGAGAGACTCTCGCAGTT
S2 CCTCCCCCCCACTCTTCCAATCACCTACTTCAATCAGAAACTGAGAGACTCTCGCAGTT
S3 CCTCCCCCCCACTCTTCCAATCACCTCCTTCAATCAGAAACTGAGAGACTCTCGCAATT
S4 CTTCCCCCCCACTCTTCCAATCACCTTCTTCAATCAGAAACTGAGAGACTCTCGCAATT
S5 CTTCCCCCCCACTCTTCCAATCACCTTCTTCAATCAGAAACTGAGAGACTCTCGCAATT
S6 ----------------CAATCACCTCCTTCAATCAGAAACTGAGAGACTCTCGCAATT
S7 ------------------TCACCTACTACAATCAGAAACTGAGAGACTCTCGCAGTT
S8 -------------------------------------------------TTTCGCAATT
Q1  C                         A                               G
Q2  C                         A                               G
Q3  C                         C                               A
Q4  T                         T                               A
Q5  T                         T                               A
Q6  X                         C                               A
Q7  X                         A                               G
Q8  X                         X                               A

     Q2    Q3    Q4    Q5    Q6    Q7    Q8
Q1  3/3   1/3   0/3   0/3   0/2   2/2   0/1
Q2        1/3   0/3   0/3   0/2   2/2   0/1
Q3              1/3   1/3   2/2   0/2   1/1
Q4                    3/3   1/2   0/2   1/1
Q5                          1/2   0/2   1/1
Q6                                0/2   1/1
Q7                                      0/1
```

Figure 1. Illustration of the procedure used to measure sequence similarity between AsAV populations in different samples. Lines beginning with "S" are an excerpt of actual sequence data for eight samples, containing three phylogenetically informative positions. "-" represents missing information. Lines beginning with "Q" represent the extracted informative residues with "x" representing missing information. The matrix at the bottom shows the scores that result from comparing the informative residues.

tein of *Kennedya yellow mosaic virus*, 64-69% identity to its coat protein and 41-97% (mean 68%) identity to its replicase protein. Within the *Tymovirus* genus, sequences of the coat protein with less than 90% identity are regarded as belonging to different species, while in the *Fabavirus* genus of the *Comoviridae*, the cut-off value is 75% (Fauquet and others 2005). The *Comoviridae* sequence was found as 0.1% of the total sequence reads from analysis of the fruits of plant 07JH012B. Sequences putatively belonging to the same viral species, nicknamed TGP Comovirid 1, had been found also from several plant species of the TPP (Melcher et al., in preparation). Sequences of retro-transcribing elements were found also, but whether these resulted from pararetroviruses or genomic retroelements was not explored.

Of the 25 samples yielding sequence (Table 2), 18 had sequences that placed them as derived from AsAV, the most predominant virus of the TPP. From *E. marginata*, 4,953 nt of sequence were assembled by CAP3 into three segments of 1,263, 564, and 3,126 nt, in 5′ to 3′ direction. Since the genome reconstructed from *A. viridis* had 6,174 nt, the *E. marginata* sequencing data collectively produced 80% of the complete genome. Among well represented sites (four or more sequence contigs) where the *E. marginata* sequence differed from the consensus AsAV sequence (Table 3), only 5 of 33 differed in all contigs covering that position. This observation suggests strongly that the divergence from AsAV is due to natural sequence variation and is unlikely to be due in any major part to host species adaptation.

To consider adaptation to different tissues, we compared phylogenetically informative residues within samples from the same tissue and from different tissues in the same plant. Only 15 of 39 informative alleles (Table 4) were shared among sequences from the same plant but 63% (63 of 102) of the informative alleles were the same for comparison of the same tissues of different plants. This difference was noticeable for lower leaves and bracts, and perhaps also for fruits. The latter and upper leaves (for which no shared alleles were identified) had insufficient loci amenable for comparison (6 each).

DISCUSSION

The frequent finding of evidence of AsAV in Payne County *E. marginata* (Table 2) extends knowledge of the most prevalent virus of the TPP. Prior to its discovery in the TPP, the virus was not known elsewhere. The present study demonstrates that the virus is not limited to the TPP and suggests that it may have a still wider geographical range, yet

Table 2. Percentage Tymovirus Content of *E. marginata* VLP-VNA samples.

Plant	Fruits	Bracts	Lower leaves	Upper leaves	Total
07JH012	0.2[a]	15.3	0.5		4.6
07JH013	3.1	0.3	15.6	1.0	6.4
07JH014	0.5	0.8		2.2	0.7
07JH015	0.7	1.0	0.5	0.3	0.6
07JH016		0			0
07JH017	0	0.5	0		0.1
07JH020	0		7.6	0	1.9
07JH021	0.2	0	0	0.4	0.2
All plants	0.4	3.5	4.0	0.5	1.9

[a] Numbers are the percentage of reads that were sequences of AsAV. Reads per sample ranged from 382 to 2904. No entry was made for samples that failed to produce amplified products from the VLP-VNA fraction.

Table 3. Polymorphisms in AsAV Sequences from *E. marginata*.

No. of Sequences[a]	No. of positions	No. Polymorphic (%)	No. Informative (%)
1	489	n.p.[b]	n.p.
2	711	26 (3.7)	n.p.
3	1331	108 (8.1)	n.p.
4	873	127 (14.5)	10 (1.1)
5	667	46 (6.9)	10 (1.5)
6	724	83 (11.5)	22 (3.0)
7	99	14 (14)	8 (8)
8	60	12 (20)	2 (3)
9	15	5 (33)	2 (13)

[a]For positions evaluated, the number of sequence determinations containing the position.
[b] n.p. = not possible

Table 4. Shared Informative AsAV Sequence Positions among Plants and Tissues.

Sample	Shared	Total	Percent
Plant			
JH12	4	11	36
JH13	7	19	37
JH14	4	8	50
JH15	0	0	n.a.
JH21	0	0	n.a.
Total	15	39	38
Tissue			
Bracts	23	36	64
Lower Leaves	34	54	63
Upper Leaves	0	6	0
Fruits	6	6	100
Total	63	102	62

to be explored. Although, in the TPP, AsAV was found in a large number of specimens of a wide range of host species, in most of these it was found in relatively low amounts as judged by percentage of sequence reads attributable to the virus (unpublished results). *Asclepias viridis* was the clear exception with some samples yielding 100% viral sequence. In *E. marginata,* although the amounts of AsAV were variable, they ranged up to 15% of the sequences recovered, making this plant species more productive of the virus than most others. Thus, *A. viridis* and *E. marginata* could be the dominant hosts in nature. The genera *Asclepias* and *Euphorbia* are not closely related, and the plants are very different in their life histories: *A. viridis* is a spring-blooming geophytic perennial and *E. marginata* is a fall-blooming annual. However, the two species are similar in that they both occur in grasslands and have overlapping geographic ranges. Potentially more interesting, both species exude copious latex, and possess nonarticulated laticifers, a trait fairly rare in the plant kingdom (Hagel and others 2008). Perhaps these factors play an important role in the success and life history of AsAv.

As with AsAV from *A. viridis,* numerous positions of the *E. marginata* virus exhibited polymorphisms. However, only five positions were found to differ from the *A. viridis* version in all of the *E. marginata* specimens analyzed. These residues may be distinct because of a founder effect on the population (French and Stenger 2003) or because *A. viridis* and *E. marginata* may impose different selective pressures on these positions. Adaptation of viral sequences to particular plant host species or to other selective pressures such as temperature, vectors, resistance gene presence is known to occur (García-Arenal and Fraile 2008).

A large variability in detection of the virus in *E. marginata*, reflected in inconsistent recovery of sequences from all plant

tissues, was observed. This could be due to a patchy distribution of the virus in individual plants since, for each sample, only 0.1 g of tissue was used. On the other hand, it may reflect uneven spread of the virus through the plant from the initial site of inoculation. Thus surveys for viruses in natural plants that sample a small amount of tissue from one specimen of a species may miss detecting viruses with low incidences and low titers. The variability of detection precluded a conclusion as to which tissues served as the best hosts for the virus. Although no conclusions about plant tissue influence on virus quantity could be reached, the data did allow the suggestion of an influence of plant tissue on the quality of the virus. Certain residues at polymorphic positions appeared to correlate better with the type of plant tissue than with the individual plant (Table 4), providing moderate support for the concept that replication conditions in various parts of the plant confer different selective pressures on viral sequence. Other variables, not accounted for in this analysis, could also play a role in sequence selection. Nevertheless, the results are sufficiently intriguing to warrant further investigation.

ACKNOWLEDGEMENTS

The authors acknowledge a Wentz Scholarship to Jennifer Hackett, support from the Robert J. Sirny Professorship, funding from National Science Foundation EPSCoR award EPS-0447262 and the Oklahoma Agricultural Experiment Station whose Director has approved the manuscript for publication.

REFERENCES

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389-402.

Fauquet CM, Mayo MA, Maniloff J, Desselberger U, Ball LA, editors, 2005. Virus Taxonomy. London Elsevier.

French R, Stenger DC. 2003. Evolution of wheat streak mosaic virus: dynamics of population growth within plants may expalin limited variation. Annu Rev Phytopathol 41:199-214.

García-Arenal F, Fraile A. 2008. Questions and concepts in plant virus evolution: a historical perspective. In: Roossinck, MJ (Ed), Plant Virus Evolution, Springer, Berlin, p. 1-14.

Hagel JM, Yeung EC, Facchini PJ. 2008. Got milk? The secret life of laticifers. Trends in Plant Science 13:631-639.

Huang X, Madan A. 1999. CAP3: A DNA sequence assembly program. Genome Res 9:868-77.

Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV, Godwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer ML, Jarvie TP, Jirage KB, Kim JB, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, Lohman KL, Lu H, Makhijani VB, McDade KE, McKenna MP, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Ronan MT, Roth GT, Sarkis GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasz A, Vogt KA, Volkmer GA, Wang SH, Wang Y, Weiner MP, Yu P, Begley RF, Rothberg JM. 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature 437:376-80.

Melcher U, Muthukumar V, Wiley GB, Min BE, Palmer MW, Verchot-Lubicz J, Nelson RS, Roe BA, Ali A, Thapa V, Pierce ML. 2008. Evidence for novel viruses by analysis of nucleic acids in virus-like particle fractions from *Ambrosia psilostachya*. J Virol Methods 152:49-55.

Muthukumar V, Melcher U, Pierce ML, Wiley GB, Roe BA, Palmer MW, Thapa V, Ali A, Ding T. 2008. Non-cultivated plants of the Tallgrass Prairie Preserve of northeastern Oklahoma frequently contain virus-like sequences in particulate fractions. Virus Res 141:169-173.

Roossinck MJ, Saha P, Wiley GB, Quan J, White JD, Lai H, Chavarría F, Shen G, Roe BA. 2009. Ecogenomics: Using massively parallel pyrosequencing to understand virus ecology. Molecular Ecology in press.

Wang R, Beggs M, Robertson L, Cerniglia C. 2002. Design and evaluation of oligonucleotide-microarray method for the detection of human intestinal bacteria in fecal samples. FEMS Microbiol Lett 213:175.

Wren JD, Roossinck MJ, Nelson RS, Scheets K, Palmer MW, Melcher U. 2006. Plant Virus Biodiversity and Ecology. PLoS Biology 4:e80.

Received: September 14, 2009; Accepted October 19, 2009.