Short Communication Title: Sequence diversity of wheat mosaic virus isolates

Author: Lucy R. Stewart¹

1. Corn, Soybean and Wheat Quality Research Unit, USDA-ARS, 1680 Madison Ave.
Wooster, OH 44691. Ph. 330-263-3834. Lucy.Stewart@ars.usda.gov.

Summary:

Wheat mosaic virus (WMoV), transmitted by eriophyid wheat curl mites (*Aceria tosichella*) is the causal agent of High Plains disease in wheat and maize. WMoV and other members of the genus *Emaravirus* evaded thorough molecular characterization for many years due to the experimental challenges of mite transmission and manipulating multisegmented negative sense RNA genomes. Recently, the complete genome sequence of a Nebraska isolate of WMoV revealed eight segments, plus a variant sequence of the nucleocapsid protein-encoding segment. Here, near-complete and partial consensus sequences of five more WMoV isolates are reported and compared to the Nebraska isolate: an Ohio maize isolate (GG1), a Kansas barley isolate (KS7), and three Ohio wheat isolates (H1, K1, W1). Results show two distinct groups of WMoV isolates: Ohio wheat isolate RNA segments had 84% or lower nucleotide sequence identity to the NE isolate, whereas GG1 and KS7 had 98% or higher nucleotide sequence identity to the NE isolate. Knowledge of the sequence variability of WMoV isolates is a step toward understanding virus biology, and potentially explaining observed biological variation.

Keywords: *Wheat mosaic virus* (WMoV), High Plains disease, nucleoprotein, emaravirus

Wheat mosaic virus is grouped in the genus *Emaravirus*, a recently-named taxon of negative-sense RNA plant viruses with multiple genome segments (Mielke-Ehret and Muhlbach, 2012). The virus (WMoV) causes High Plains disease, first described in maize and wheat in Texas, Kansas, Colorado, Idaho, Nebraska, and Utah in the 1993-1994 growing seasons (Jensen et al., 1996). The reported disease distribution has expanded in recent years to include Oregon and Washington, where it is problematic in sweet corn (Gieck et al., 2007), and Ohio and Australia where disease has been detected (Coutts et al., 2014; Stewart et al., 2013; Stewart et al., 2014). WMoV is transmitted by the wheat leaf curl mite *Aceria tosichella* Keifer (Seifers et al., 1997). Other described emaravirus species include: *European mountain ash ringspot-associated virus*, the four segmented type member of the genus *Emaravirus*; *Fig mosaic virus*, with six reported RNA segments; *Rose rosette virus* now with seven reported segments; and *Raspberry leaf blotch virus* now with eight reported genome segments (Di Bello et al., 2015; Elbeaino et al., 2009; Laney et al., 2011; Lu et al., 2015; McGavin et al., 2012; von Bargen et al., 2013). WMoV, like other emaraviruses, has been recalcitrant to thorough characterization because of its mite transmission and multi-segmented negative sense RNA genome.

Initial efforts to characterize WMoV (variously named High Plains virus, maize red stripe virus, and wheat mosaic virus in the literature) included virion and nucleoprotein purification, development of antisera against the characteristic 32 kDa nucleoprotein for detection, transmission by vascular puncture inoculation (VPI) to recapitulate disease symptoms, and partial sequencing of the virus genome (Jensen et al., 1996; Louie et al., 2006; Seifers et al., 1997; Seifers et al., 2004). Genome segment number was estimated at four to six nucleic acid species based on gel separation of virion RNA, in which four bands

were usually visible, the center two broad and diffuse (Skare et al., 2006). Skare et al. reported obtaining 6.4kb of virus sequence, but these sequences were not made publicly available (Skare et al., 2006). Not until 2014 was the putative complete sequence of 8 genomic segments reported for a Nebraska WMoV isolate (Tatineni et al., 2014). This group used deep sequencing of purified virion RNA to identify highly enriched virus sequences that had no contemporary homologs in public sequence databases. In this study, near-complete and partial coding sequences of five more WMoV isolates are reported and compared to the Nebraska isolate and previously reported partial sequences of other isolates. In total, WMoV sequences from isolates collected from four states (Ohio, Kansas, Texas and Nebraska) and from three different host plants (barley, maize, and wheat) are compared (Table S1). New isolates sequenced are: Ohio wheat isolates H1, K1, and W1 from original samples and vascular puncture inoculation transmissions to ‘Spirit’ maize (Stewart et al., 2013); Ohio sweet corn (*Zea mays*) isolate GG1 original sample (Stewart et al., 2014); and Kansas ‘Westford’ barley (*Hordeum vulgare*) isolate KS7 (sample code U99-14) from Dr. Dallas Seifers (Kansas State University) maintained since 1999 by serial vascular puncture inoculation (VPI) in ‘Spirit’ maize as previously reported (Louie et al., 2006). Total RNA was extracted from fresh or -80°C frozen infected wheat or maize tissue using DirectZol™ RNA kits (Zymo Research, Irvine, CA). WMoV-positive samples were identified by ELISA using antisera produced against Kansas corn isolate HPV96KS5-PI (Louie et al., 2006) and KS7, and reverse transcription-polymerase chain reaction (RT-PCR) as previously described (Stewart et al., 2013).

Partial WMoV sequences were first identified in RNA-sequencing surveys of Ohio wheat and maize viruses (Stewart et al., 2013; Stewart et al., 2014). To complete WMoV

isolate genomes, RNA-Sequencing was performed on original survey samples (K1, GG1) or samples transmitted to maize (H1, W1). For KS7, RNA-Seq on RNA from virion preparations was also performed as described (Tatineni et al., 2014; Tatineni et al., 2015) but insufficient sample tissue for other isolates remained for this approach. RNA preparation, RNA-Sequencing, and sequence analyses were described previously for pooled survey samples (Stewart et al., 2013; Stewart et al., 2014). A total of 184 million 100 nt raw single-end reads were obtained from pooled wheat samples. 2.2 million raw 100 nt paired-end reads were obtained from pooled maize samples, and 29 million from a KS7-infected maize sample. Multiplexed samples from KS7 infected plant tissue or virion RNA (sample codes 51 and HPV respectively), H1 (sample code 53), and W1 (sample code 55), all transmitted to ‘Spirit’ maize, each produced 18.4 million 100 nt paired-end reads.

Illumina sequences were analyzed using CLC Genomics software (Cambridge, MA) using plant virus identification pipelines previously described (Stewart et al., 2013; Stewart et al., 2014) or similar thereto. Isolate sequences were completed by RT-PCR and Sanger sequencing at the Ohio State University Plant-Microbe Genomics Facility using primer-walking. Primers (Table S2) were designed from RNA-Seq contigs, generic emaravirus end primers (Mielke and Muehlbach, 2007), and from Nebraska isolate RNA5-8 sequences (Tatineni et al., 2014). Random amplification of cDNA ends (RACE) was attempted unsuccessfully from limited remaining sample RNA to obtain exact end sequences, so 5’ and 3’ ends were trimmed based on ends predicted from Nebraska isolate sequences and not experimentally validated. Consensus sequences for each segment were assembled in Sequencher (GeneCodes Corp. Ann Arbor, MI) using Sanger and RNA-Seq sequences.

Combining RNA-Seq and Sanger sequencing of RT-PCR products, all eight known WMoV segments were identified for Kansas KS7 and Ohio GG1 isolates. KS7 and GG1 segments shared over 98% nucleotide sequence identity with the NE isolate, and predicted proteins were 95-100% identical (Table 2). Thus, KS7 and GG1 were considered “NE-like isolates”. For Ohio wheat isolates H1, K1, and W1, partial or complete coding sequences were obtained for all eight segments, but no isolate sequence was complete (Table 1). We were unable to amplify RNA5 sequence from the Ohio wheat isolates H1 and W1 after transmission to maize (Stewart et al., 2013); however, near-complete RNA5 sequence was identified in RNA-Seq of pooled wheat samples, and was amplified from original sample from isolate K1 (Table 1). The Ohio wheat isolate RNAs 1-7 shared only 77-84% nucleotide sequence identity and 81-97% predicted amino acid sequence identity with corresponding sequences of the NE isolate (Table 2) but are very similar to each other (98% or higher identity for near-complete RNAs 3 and 6).

RNA1 is predicted to encode the 2272 amino acid viral RNA-dependent RNA polymerase (RdRP), with a Bunyavirus RdRP protein family domain found at aa 622-1406 in the isolate W1 RNA1 sequence (E-value = 1.1×10^{-40} , predicted by Simple Modular Architecture Research Tool, SMART) (Letunic et al., 2015; Schultz et al., 1998). GG1, NE, and KS7 isolates were >99% identical in predicted RNA1 amino acid sequences, whereas the predicted W1 RNA1 product was only 91.8% identical to the NE sequence (Table 2). Near-complete RNA1 sequences obtained from isolates GG1 and KS7 had 23 single nucleotide polymorphisms (present in one or both) compared to the 6981 nt NE RNA1 sequence. The 6991 nt W1 isolate RNA1 sequence had 10 nucleotide insertions (1, 2 and 7-mer) in the 5' untranslated region (utr) relative to the NE-like isolates.

RNA2 of emaraviruses is predicted to encode an envelope glycoprotein precursor. Consistent with this, SMART (Letunic et al., 2015; Schultz et al., 1998) analysis predicts two transmembrane domains (aa140-162; 202-219 in W1; and aa143-165 and 205-224 in GG1 predicted RNA2 proteins). The putative glycoprotein precursor protein signal peptide and precursor cleavage sites at NE I49/V50 and A244 and D225 were also predicted at homologous sites in the other isolates as previously predicted for rose rosette virus (RRV) RNA2 product (Laney et al., 2011). Wheat isolate RNA2 sequences contained several nucleotide insertion/deletions in the 5' and 3' utr regions compared to the NE-like RNA2 sequences.

RNA3 is predicted to encode the nucleoprotein (NP) that encapsidates viral RNA within double-membrane bound virions. As previously reported for the NE isolate (Tatineni et al., 2014), isolates GG1 and KS7 were each found to contain two distinct RNA3 variants encoding full-length NPs. The RNA3A and 3B predicted proteins shared 95-99% within-group sequence identity and 88-89% between-group identity. Only one RNA3 copy was identified for each of the Ohio wheat isolates H1, K1, and W1, and the pooled Ohio wheat samples (Table 1). The Ohio wheat isolate RNA3 sequences clustered separately from both the NE-like 3A and 3B sequences (Fig. 1). The predicted Ohio wheat isolate NP sequences are 81-86% identical to the NE-like 3A and 3B sequences. Comparisons with previously published WMoV NP amino acid sequences indicated that neither collection host nor location of collection predict WMoV NP variability. Texas maize isolate NP ABC58222 (Skare et al., 2006) clustered with Ohio wheat isolates H1, K1, and W1; but Texas isolate TX96 (Seifers et al., 2009) (unknown collection host) clustered with NE-like maize, barley and wheat-isolate 3B NP from Ohio, Kansas, and Nebraska (GG1, KS97, and NE). Only one

TX96 NP sequence was reported (Seifers et al., 2009; Seifers et al., 2004), although all the most similar isolates contained two. The Kansas barley isolate KS7 NP sequences cluster with the 3A/3B group whereas the Kansas wheat isolate referred to here as KS04 (U04-82, UniProt Accession no. P85309) (Seifers et al., 2009) NP clustered with the Ohio wheat isolates (Fig. 1, S1).

RNA-Seq contig assemblies of RNA3 and RT-PCR results showed multiple 5' ends for nucleoprotein sequences, all containing the predicted 5' emaravirus end sequence (5'-AGTAGTGWCTCC-3') (Mielke and Muehlbach, 2007; Mielke-Ehret and Muehlbach, 2012) with or without partial duplications within the 5' utr. We were not able to confirm whether variability in the 5' ends occurs *in planta* or is an artifact of contigging and RT-PCR sequencing since we were unable to confirm end sequences. For Ohio wheat isolate H1, two RNA3 sequences were obtained, one of which had a truncated 3' utr, but both were predicted to encode identical proteins, suggesting it may have been a subgenomic RNA, which have been reported for WMoV RNAs 3, 4, 7, and 8 (Tatineni et al., 2014).

WMoV and other emaravirus RNA4 sequences are predicted to encode movement protein (MP). The RNA4-encoded protein of raspberry leaf blotch virus (RLBV), the emaravirus most similar to WMoV in RNAs 1-4 sequences, localizes to plasmodesmata (McGavin et al., 2012). WMoV RNA4 sequences share some similarity with other emaraviruses and were the most highly conserved among WMoV isolates. Predicted Ohio wheat isolate RNA4 protein sequences were 96-97% identical to the NE-like sequences, and >99% identical to each other. Among the isolates KS7, NE, and GG1, 7 indels were present in the 5' utr of the RNA4 sequence. The 364 aa of the GG1 and KS7 isolates had only one aa polymorphism.

WMoV RNAs 5-8 are each predicted to encode a protein of unknown function. The putative wheat isolate RNA5 protein was only 83.7% identical to the NE isolate RNA5 protein, whereas KS7 and GG1 sequences were over 99% identical to NE RNA5 protein. RRV and fig mosaic virus (FMV) RNA5 products have been predicted to be homodimerizing RNA-binding proteins (Di Bello et al., 2015). The WMoV RNA5 predicted protein shares weak similarity with the WMoV RNA6 predicted protein, as previously reported (Tatineni et al., 2014).

RNA6 predicted products were 99% identical among NE-like isolates, 99% identical among Ohio wheat isolates, and 81% identical between the groups. RNA7 predicted products were 99% identical between NE-like isolates, 99% identical among Ohio wheat isolates, and 89-90% identical between the groups. The Ohio wheat isolate RNA7 sequences contained several utr indels relative to NE-like sequences, including a 49-nucleotide deletion at nt 207-256. SMART analysis predicted a signal peptide in the putative translation product of RNA7 at aa 1-23.

RNA8 sequences were unusual in that the predicted ORF (Tatineni et al., 2014) covered only 531/1339 nt (40%) of the segment RNA. No other ORF conserved across the isolates was found in the remaining sequence. While the predicted RNA8 ORF amino acid sequence shared 74.4% identity with the NE sequence (Table 2), the non-coding region of the wheat isolates was even more divergent. For example, the W1 RNA8 sequence shared only 30.7% nucleotide identity with the NE isolate RNA8 (Table 2). The W1 and H1 wheat isolate RNA8 nucleotide sequences were 89.9% identical to each other, with the H1 isolate sequence missing 28 5' nt and containing two deletions in the 5' utr (20 nt and a 55 nt) relative to the W1 isolate sequence and the W1 sequence missing 3 3' nt relative to H1.

Ohio wheat isolate sequences were 89.7% identical at the nucleotide level and 99.4% identical (100% similar) at the amino acid level.

The presence of multiple copies of the NP-coding segment is unusual, but segment duplication and divergence appears to be a pattern in the emaraviruses as more sequence data emerge (Di Bello et al., 2015; Lu et al., 2015). From a practical standpoint, since the nucleoprotein encoded by RNA3 is used for serological detection of WMoV, variations in antisera could be important in diagnostics, as was found by Seifers et al. (Seifers et al., 2009). However, the antiserum used here, prepared against both a 1996 Kansas maize isolate (not sequenced) and the KS7 Kansas barley isolate, detected all isolates described (Stewart et al., 2013; Stewart et al., 2014).

The first four WMoV RNA segments are predicted to encode proteins performing the expected major functions of plant virus proteins: replication, encapsidation, and movement. However, the functions of proteins encoded by RNAs 5-8 remain mysterious, since these have no strong sequence homologs that indicate function. Possible roles of proteins encoded on viral RNAs might include packaging all eight or nine genomic segments into one or more membrane-bound virions, formation of double membrane bodies, or association with and transmission by eriophyid mite vectors.

The relationship between sequence diversity and WMoV biology remains to be experimentally determined. Biological variability in wheat curl mite transmission of WMoV was reported for both vector and virus populations (Seifers et al., 2002), but virus transmission biology associations with specific viral sequence differences have yet to be determined. Differences in infectivity in maize have also been described for isolates with different partial NP sequences (Seifers et al., 2004). The sequenced isolates do not

encompass the full geographic range of known WMoV distribution, so additional sequences will help reveal whether the two groups described here encompass the full extent of WMoV diversity.

Acknowledgements:

Thanks to the Ohio Small Grains Marketing Program, USDA-ARS, and Ohio State University for funding the initial survey of Ohio wheat, and to Ohio wheat and corn growers for permitting surveys on their property. Thanks to the Molecular and Cellular Imaging Center (OSU) for RNA-sequencing and bioinformatics support, This project would not be possible without the expert technical assistance of Kristen Willie (USDA-ARS) for primer design and Sanger sequencing confirmation of RNA-Seq data. Special thanks also to Ellie Walsh (OSU), Jane Todd (USDA-ARS), and Dr. Margaret G. Redinbaugh (USDA-ARS) for critical reading of the manuscript.

Tables and Figure legends:

Table 1. Summary of *Wheat mosaic virus* (WMoV) sequences and major open reading frame (ORF) coordinates (GenBank accession nos. KT98860-KT988892, KT970499-KT970505, and KT995097-KT995113). Gray shading indicates complete protein coding sequence obtained.

Table 2: Nucleotide/amino acid sequence identity (and similarity using a BLOSUM60 matrix) compared to Nebraska WMoV isolate (GenBank accession no. KJ939623-KJ939631). Calculated from ClustalW alignments made using MacVector v.14.03 (Cary, NC).

Fig. 1. Best tree of *Wheat mosaic virus* nucleoprotein amino acid sequences. Tree generated in MacVector v.14.03 and rooted with *Raspberry leaf blotch virus* nucleoprotein sequence (GenBank accession no. FR823301). Sequences include isolates reported in this study GG1, W1, H1, K1, KS7, and pooled wheat (H1+K1+W1); and previously reported sequences of the Nebraska isolate (GenBank accession nos. KJ939625-6), Texas isolates TX96 and ABC58222, and Kansas isolate KS04 (U04082).

References:

- Coutts, B.A., Cox, B.A., Thomas, G.J. and Jones, R.A.C., 2014. First Report of Wheat mosaic virus Infecting Wheat in Western Australia. *Plant Dis* 98(2), 285-285.
- Di Bello, P.L., Ho, T. and Tzanetakis, I.E., 2015. The evolution of emaraviruses is becoming more complex: seven segments identified in the causal agent of Rose rosette disease. *Virus Res* 210, 241-244.
- Elbeaino, T., Digiario, M. and Martelli, G.P., 2009. Complete nucleotide sequence of four RNA segments of fig mosaic virus. *Arch Virol* 154(11), 1719-1727.
- Gieck, S.L., Hamm, P.B., Clough, G.H. and David, N.L., 2007. High plains virus: An emerging disease of sweet corn in the Columbia basin of Oregon and Washington. *Phytopathology* 97(7), S167-S168.
- Jensen, S.G., Lane, L.C. and Seifers, D.L., 1996. A new disease of maize and wheat in the high plains. *Plant Dis* 80(12), 1387-1390.
- Laney, A.G., Keller, K.E., Martin, R.R. and Tzanetakis, I.E., 2011. A discovery 70 years in the making: characterization of the Rose rosette virus. *J Gen Virol* 92, 1727-1732.
- Letunic, I., Doerks, T. and Bork, P., 2015. SMART: recent updates, new developments and status in 2015. *Nucleic Acids Res* 43(D1), D257-D260.
- Louie, R., Seifers, D.L. and Bradfute, O.E., 2006. Isolation, transmission and purification of the High Plains virus. *Journal of virological methods* 135(2), 214-222.
- Lu, Y., McGavin, W., Cock, P.J., Schnettler, E., Yan, F., Chen, J. and MacFarlane, S., 2015. Newly identified RNAs of Raspberry leaf blotch virus encoding a related group of proteins. *J Gen Virol*. doi: 10.1099/jgv.0.000277.
- McGavin, W.J., Mitchell, C., Cock, P.J.A., Wright, K.M. and MacFarlane, S.A., 2012. Raspberry leaf blotch virus, a putative new member of the genus Emaravirus, encodes a novel genomic RNA. *J Gen Virol* 93, 430-437.
- Mielke, N. and Muehlbach, H.P., 2007. A novel, multipartite, negative-strand RNA virus is associated with the ringspot disease of European mountain ash (*Sorbus aucuparia* L.). *J Gen Virol* 88(Pt 4), 1337-1346.
- Mielke-Ehret, N. and Muehlbach, H.P., 2012. Emaravirus: A Novel Genus of Multipartite, Negative Strand RNA Plant Viruses. *Viruses-Basel* 4(9), 1515-1536.
- Schultz, J., Milpetz, F., Bork, P. and Ponting, C.P., 1998. SMART, a simple modular architecture research tool: Identification of signaling domains. *P Natl Acad Sci USA* 95(11), 5857-5864.
- Seifers, D.L., Harvey, T.L., Louie, R., Gordon, D.T. and Martin, T.J., 2002. Differential transmission of isolates of the High Plains virus by different sources of wheat curl mites. *Plant Dis* 86(2), 138-142.
- Seifers, D.L., Harvey, T.L., Martin, T.J. and Jensen, S.G., 1997. Identification of the wheat curl mite as the vector of the high plains virus of corn and wheat. *Plant Dis* 81(10), 1161-1166.
- Seifers, D.L., Martin, T.J., Harvey, T.L., Haber, S., Krokhin, O., Spicer, V., Ying, S. and Standing, K.G., 2009. Identification of Variants of the High Plains virus Infecting Wheat in Kansas. *Plant Dis* 93(12), 1265-1274.
- Seifers, D.L., She, Y.M., Harvey, T.L., Martin, T.J., Haber, S., Ens, W., Standing, K.G., Louie, R. and Gordon, D.T., 2004. Biological and molecular variability among High Plains virus isolates. *Plant Dis* 88(8), 824-829.

- Skare, J.M., Wijkamp, I., Denham, I., Rezende, J.A.M., Kitajima, E.W., Park, J.W., Desvoyes, B., Rush, C.M., Michels, G., Scholthof, K.B.G. and Scholthof, H.B., 2006. A new eriophyid mite-borne membrane-enveloped virus-like complex isolated from plants. *Virology* 347(2), 343-353.
- Stewart, L.R., Paul, P.A., Qu, F., Redinbaugh, M.G., Miao, H., Todd, J. and Jones, M., 2013. Wheat mosaic virus (WMoV), the Causal Agent of High Plains Disease, is Present in Ohio Wheat Fields. *Plant Dis* 97(8), 1125-1125.
- Stewart, L.R., Teplier, R., Todd, J.C., Jones, M.W., Cassone, B.J., Wijeratne, S., Wijeratne, A. and Redinbaugh, M.G., 2014. Viruses in maize and johnsongrass in southern ohio. *Phytopathology* 104(12), 1360-1369.
- Tatineni, S., McMechan, A.J., Wosula, E.N., Wegulo, S.N., Graybosch, R.A., French, R. and Hein, G.L., 2014. An eriophyid mite-transmitted plant virus contains eight genomic RNA segments with unusual heterogeneity in the nucleocapsid protein. *J Virol* 88(20), 11834-11845.
- Tatineni, S., McMechan, A.J., Wosula, E.N., Wegulo, S.N., Graybosch, R.A., French, R. and Hein, G.L., 2015. Correction for Tatineni et al., An Eriophyid Mite-Transmitted Plant Virus Contains Eight Genomic RNA Segments with Unusual Heterogeneity in the Nucleocapsid Protein. *J Virol* 89(14), 7443.
- von Barga, S., Arndt, N., Robel, J., Jalkanen, R. and Buttner, C., 2013. Detection and genetic variability of European mountain ash ringspot-associated virus (EMARaV) in Sweden. *Forest Pathol* 43(5), 429-432.

Method: Neighbor Joining; Best Tree; tie breaking = Systematic
Distance: Uncorrected ("p")
Gaps distributed proportionally

Fig. 1. Best tree of *Wheat mosaic virus* nucleoprotein amino acid sequences. Tree generated in MacVector v.14.03 and rooted with *Raspberry leaf blotch virus* nucleoprotein sequence (GenBank accession no. FR823301). Included sequences include isolates reported in this study GG1, W1, H1, K1, KS7, and pooled wheat (H1+K1+W1); and previously reported sequences of the Nebraska isolate (GenBank accession nos. KJ939625-6), Texas isolates TX96 and ABC58222, and Kansas isolate KS04 (U04082).

Table 1. Summary of *Wheat mosaic virus* (WMoV) sequences and major open reading frame (ORF) coordinates (GenBank accession nos.: KT98860-KT988892, KT970499-KT970505, and KT995097-KT995113). Gray shading indicates complete protein coding sequence obtained.

Isolate	RNA1	RNA2	RNA3A	RNA3B	RNA4	RNA5	RNA6	RNA7	RNA8
Kansas-KS7	6980 nt ORF: 6912-94	2211 nt ORF: 2132-129	1440 nt ORF: 1217-357	1442 nt ORF: 1221-352	1681 nt ORF: 1569-475	1713 nt ORF: 1593-157	1751 nt ORF: 1633-155	1435 nt ORF: 1318-401	1338 nt ORF: 1245-715
Ohio-GG1	6972 nt ORF: 6904-86	2211 nt ORF: 2132-129	1070 nt ORF: 991-128	1444 nt ORF: 1222-353	1681 nt ORF: 1569-475	1715 nt ORF: 1595-159	1752 nt ORF: 1634-156	1435 nt ORF: 1318-401	1339 nt ORF: 1246-716
Ohio-H1	Partial 537, 565 nt	Partial 477 nt	1548 nt, ORF: 1328-462 1501 ¹ , ORF: 1288-422		1280 nt ORF: 1168-74	NS ²	Partial 1041 nt	Partial 749 nt	1250 nt ORF: 1165-635
Ohio-K1	Partial 562 nt	Partial 534 nt	1548 nt, ORF: 1328-462		NS	Partial 1230 nt	1747 nt ORF: 1627-152	1390 nt ORF: 1271-354	NS
Ohio-W1	6991 nt ORF: 6923-105	2170 nt ORF: 2091-97	1547 nt, ORF: 1327-461		1678 ³ nt ORF: 1565-471	NS	1728 nt ORF: 1619-144	1324 nt ORF: 1217-300	1350 nt ORF: 1268-738
Wheat Pool (H1, K1, W1)	Partial 3907, 1765, 1236 nt	2109 nt ORF: 2030-36	1167nt, ORF: 896-47 partial 204 nt		Partial 440, 324, 209 nt	1644 nt ORF: 1523-102	Partial 1277, 492, 277, 218 nt	Partial 582, 358 nt	Partial 377 nt

1. Truncation, identical amino acid translation predicted.
2. NS= No sequence obtained.

Table 2: Nucleotide/amino acid sequence identity (and similarity using a BLOSUM60 matrix) compared to a Nebraska WMoV isolate (GenBank accession nos. KT939623-KT939631, KT970499-KT970505, and XX-XX). Calculated from ClustalW alignments made in MacVector v.14.03 (Cary, NC).

	KS7	GG1	H1	K1	W1	Wheat (pooled)
RNA1	99.7/99.8(99.8)	99.5/99.6(99.8)	NC ¹	NC	83.4/91.8(96.2)	NC
RNA2	99.7/99.9(99.9)	99.4/99.1(99.4)	NC	NC	--/85.2(92.7)	--/85.6(93.1)
RNA3A²	98.7/99.7(100)	-- ³ /95.8(97.6)	NC	NC	NC	NC
RNA3B	99.3/99.7(100)	98.9/99.7(100)	--/86.5(92.4)	77.7/85.5(91.3)	77.6/85.5(92.0)	--/85.5(91.3)
RNA4	98.3/99.5(100)	98.0/99.2(99.7)	--/96.7(99.2)	NC	84.1/96.7(99.2)	NC
RNA5	99.3/99.4(99.6)	99.5/99.4(99.6)	NC	NC	NC	--/83.7(91.8)
RNA6	99.6/99.6(99.8)	99.7/99.4(99.6)	NC	82.0/81.3(89.0)	80.9/81.7(89.0)	NC
RNA7	98.1/98.7(100)	98.5/99.0(100)	NC	82.2/89.5(93.1)	--/90.2(93.8)	NC
RNA8	99.2/100(100)	99.3/100(100)	--/74.4(86.4)	NC	30.7/74.4(86.4)	NC

1. NC= No comparison
2. NE 3A and 3B are 87.3% identical at the nucleotide level and 88.6 (94.8)% identical (similar) in amino acid sequence.
3. -- = Incomplete nucleotide sequence; no accurate comparison can be made.