Alfalfa vein mottling virus, a novel potyvirid infecting *Medicago sativa* L.

[1] Lev G. Nemchinov*, [2] Olga A. Postnikova, [3] William M. Wintermantel, [4] John C. Palumbo and [1] Sam Grinstead

[1]USDA-ARS, Molecular Plant Pathology Laboratory, Beltsville, MD, USA; [2]USDA-ARS, Animal Biosciences and Biotechnology Laboratory, Beltsville MD; [3]USDA-ARS Crop Improvement and Protection Research, Salinas, CA; [4]University of Arizona Yuma Agricultural Center Yuma, AZ.

*Corresponding author: lev.nemchinov@usda.gov

Abstract

We have recently identified a novel virus detected in alfalfa seed material. The virus was tentatively named alfalfa-associated potyvirus 1, as its genomic fragments bore similarities with potyvirids. In this study, we obtained a complete genome sequence of the virus and classified it as a tentative species in the new genus, most closely related to the members of the genus *Ipomovirus* in the family *Potyviridae*. This assumption is based on the genome structure, phylogenetic relationships, transmission electron microscopy investigations and, in part, on serological cross-reactivity of the virus. We also demonstrated its mechanical transmission to the indicator plant *Nicotiana benthamiana* and to the natural host *Medicago sativa*, both of which developed characteristic symptoms therefore suggesting a pathogenic nature of the disease. Consistent with symptomatology, the virus was renamed to alfalfa vein mottling virus. A name Alvemovirus was proposed for the new genus in the family *Potyviridae*, of which alfalfa vein mottling virus is a tentative member.

Keywords: plant viruses; positive-strand RNA viruses; *Potyviridae*; alfalfa vein mottling virus; Alvemovirus; alfalfa (*Medicago sativa* L.).

Introduction

Research on alfalfa (*Medicago sativa* L.), a major forage crop worldwide, greatly benefited from the latest advances in high throughput sequencing (HTS) technologies that enabled the completion of its genome assembly [1] revealing new insights into the mechanisms of the crop's resistance to abiotic and biotic stresses [2,3,4,5], identification of genetic markers associated with important agronomic traits [6,7,8] and discovery of numerous microorganisms infecting alfalfa, primarily viruses [9-12]. Viruses were found to be an integral part of multi-pathogenic infections collectively forming the alfalfa pathobiome, a diverse community of pathogenic microbes within the biotic environment of the plant [13,14]. Recently, during the initial seed screenings of alfalfa germplasm accessions maintained by the USDA ARS National Plant Germplasm System (NPGS), we have detected a broad range of viruses in the mature seeds of different germplasm sources [15]. Among them were fragments of a poty-like virus with a genomic organization typical for *Potyviridae* that were ~ 26%-32% identical to a few members of the family. We provisionally named it alfalfa-associated potyvirus 1 [15]. In this research, aimed at surveying alfalfa varieties grown in the U.S. to identify, characterize, and prevent the spread of novel and emerging viruses in the country, we found AaPV1 in plant samples collected from commercial alfalfa fields in Arizona, USA. A coding-complete genome of the virus was assembled from overlapping sets of contigs and a phylogenetic relationship with other members of the family *Potyviridae* established. The virus was further categorized as a novel species, tentatively related to the genus *Ipomovirus* and potentially representing a new genus in the family *Potyviridae*. Consistent with symptomatology in the natural host *Medicago sativa* L., the virus was designated as **al**falfa **ve**in **mo**ttling virus (AVMV), a representative member of the new taxon, provisionally named Alvemovirus.

Materials and Methods

Plant material

47 Five alfalfa plants (stems and leaves) were sampled from each of the four different fields, 10-15 acres in
48 size, located in Yuma country, Arizona, USA. Geographic coordinates of the alfalfa fields and the
49 adjacent crops are shown in Table 1.

50 **Total RNA extraction, RNA sequencing, RT-PCR, and 5'RACE**
51
52 Five leaves were pooled from each of the 20 plants for RNA extraction. Total RNA was extracted using
53 Promega Maxwell® RSC Plant RNA Kit (Promega Corp., Fitchburg, WI) according to the manufacturer's
54 directions. cDNA libraries were prepared using Illumina TruSeq Stranded Total RNA with Ribo-Zero
55 Library Prep Kit (Illumina Inc., San Diego, CA USA) and RNA sequencing was performed by Psomagen
56 (Psomagen Inc., Rockville MD USA) on a NovaSeq6000 S4 platform (150PE, 1Gb, 20 million total reads).
57 Reverse transcription–polymerase chain reactions (RT-PCR) were performed using the SuperScript One-
58 Step RT-PCR System according to the manufacturer's directions (Thermo Fisher Scientific Inc., Waltham,
59 MA USA) or utilizing SuperScript™ III First-Strand Synthesis System Thermo Fisher Scientific, Waltham,
60 MA USA) following by PCR amplification with AmpliTaq Gold DNA Polymerase Thermo Fisher Scientific).
61 Total RNA was extracted using TRIzol Reagent (Thermo Fisher Scientific Inc., Waltham, MA USA).
62 Negative control reactions were performed with RNA obtained from uninfected *N. benthamiana* or with
63 sterile RNAse-free water processed for cDNA preparation the same way as RNA extracted from the
64 infected alfalfa plants. Three virus-specific sets of primers were designed based on the results of the HTS
65 analysis: 1) LN1077/LN1078 (forward, 5' CTCCCTTTGGTACTCGGTATTG, 3', pos. 7679-7701 and reverse,
66 5' CTCTAATGTGCGGACCTTTCT 3', pos. 8072-8093, size 414 bp); 2) LN1079/LN1080 (forward, 5'
67 TGACAGCGAGTTCTCATTCC 3', pos. 9165-9184, and reverse, 5' GTCAGTCAAACCAGCCTTTATTC 3', pos.
68 9800-9822, size 658 bp); 3) LN1095/LN1096 (forward, 5' CCAGGGTGTTGTTGATGATTTG 3', pos. 4176-
69 4197, and reverse, 5' TCTCAGTCACATTCCGCATAAA 3', pos. 4579-4600, size 425 bp). PCR products were
70 either sequenced directly or cloned into TOPO II vector (Thermo Fisher Scientific Inc., Waltham, MA

71 USA) and sequenced. 5'RACE was performed with SMARTer RACE 5'/3' Kit (Takara Bio USA, Inc. San Jose,
72 CA USA) according to the manufacturer's directions.

73 **Bioinformatic analysis**

74 A schematic representation of the bioinformatics pipeline used in this study is shown in **Fig.1**. Briefly,
75 sequence reads were first trimmed using Trimmomatic [16], followed by their assembly with SPAdes
76 [17]. The resulting contigs were screened using BLASTx searches [18] against a virus database containing
77 all plant virus protein sequences from the NCBI RefSeq database
78 (https://www.ncbi.nlm.nih.gov/refseq/). The potential plant viral hits were searched once again using
79 BLASTx against the full NCBI nr protein database. BBMap [19] was used to generate sequencing
80 coverage values for the identified contigs. For the prediction of cleavage sites for viral proteases,
81 polyproteins from 109 different members of the family *Potyviridae* were aligned by four different
82 Multiple Sequence Alignments (MSA) software: ClustalW, ClustalO, MUSCLE, and CLC Genomics
83 Workbench 23, very accurate setting (QIAGEN LLC, Germantown MD USA). Cleavage sites obtained by
84 each MSA method were compared and the most likely sites were selected by another comparison with
85 previously reported consensus sites [20].

86 **Phylogenetic analysis**

87 Phylogenetic analysis was performed with polyproteins of the selected members of the family
88 *Potyviridae*. Sequences were aligned with Clustal W and the unrooted tree was built with RAxML-NG
89 [21] using maximum likelihood algorithm, a maximum of 1000 bootstrap replicates, and bootstopping
90 (autoMRE, cutoff: 0.030000). Boostrapping converged after 150 replicates.

91 **Mechanical inoculation of *Nicotiana benthamiana* Domin. and alfalfa (*Medicago sativa* L.)**
92 **plants**

93 Symptomatic leaves from field alfalfa plants were homogenized in cold 20 mM potassium phosphate
94 buffer, pH 7.0, using a sterile pestle and mortar. The resulting suspension was briefly centrifuged to

95 precipitate plant debris and ten microliters of the supernatant were rubbed onto carborundum-dusted
96 leaves of three-weeks old *N. benthamiana* plants using sterile inoculating loops. Carborundum-dusted
97 leaves of the three weeks-old seedlings of two alfalfa cultivars, cv. CUF 101 and cv. Maverick, were rub-
98 inoculated with purified viral preparations using sterile inoculating loops.

99 **Virus purification**

100 The virus was partially purified from 20 grams of symptomatic *N. bentamiana* leaves using a protocol
101 developed for *Poinsettia mosaic virus* [22].

102 **Transmission electron microscopy**

103 For TEM observations, TEM grids were incubated on the drop of purified viral preparation for 2 minutes,
104 rinsed with one milliliter of sterile water and stained with 1% phosphotungstate (PTA) solution. The grids
105 were examined in a Hitachi H-7700 Electron Microscope at the Electron and Confocal Microscope Unit,
106 Beltsville Agricultural Research Center.

107 **Western blot assays**

108 Western blots (WB) were performed as previously described [23]. Membranes were probed with
109 polyclonal antiserum to sweet potato mild mottle virus (SPMMV; DSMZ, Braunschweig, Germany) in the
110 company's recommended dilution (1:1000).

111 ## Results

112 **Symptoms displayed on the infected alfalfa plants**

113 According to the results of HTS, out of 20 examined plants, eight plants (40%) originating from three
114 different fields, were infected with the virus. These plants displayed a variety of virus-like symptoms
115 such as mottling, chlorosis, vein clearing, and bright yellow blotches (**Fig. 2**). The symptoms were likely
116 triggered by multiple co-infecting pathogens, including the virus in question [14].

117

118 **Viral genome assembly and structure**

119 Assembled contigs generated from each of the eight plants exceeded 9-10 kb in length and represented
120 variants of the same virus (**Table S1**). The longest contig with 4940 x per-nt average coverage consisted
121 of 10,323 nucleotides (nt) and included a poly (A) tail, thus indicating that the 3'-terminal portion of the
122 viral genome is complete. Multiple 5'RACE reactions adjusted the HTS-derived length of the viral
123 genome to 10,037 nt, excluding the poly(A) tail (75 nt). According to the sequenced 5'RACE PCR
124 products, the 5' non-coding region of the virus contains 107 nt; an open reading frame (ORF) encoding a
125 large polyprotein comprises 9,732 nt (108-9,839 nt); and the 3' non-coding region of the virus appears
126 complete, enclosing 198 nt (9,840-10,037), without the poly (A) tail.

127 In PSI-BLAST search, the top three hits for the viral polyprotein (excluding AaPV1 submissions) were:
128 eggplant ipomovirus A (GenBank ID: BCG55390.1; 32.8% identity, query cover 78%; E-value=0.0); rose
129 yellow mosaic roymovirus (GenBank ID: BBI90117.2; 31.6% identity; query cover 87%; E-value=0.0); and
130 *Potyvirus* spp. polyprotein (GenBank ID: QHB15167.1; 31.3% identity, query cover 85%; E-value=0.0).
131 Among other hits were squash vein yellowing virus (genus *Ipomovirus*; GenBank ID: AEV45694.1; 34,4%
132 identity; query cover 65%, E-value=0.0); chili ringspot virus (genus Potyvirus; GenBank ID: APW85806.1;
133 28.7% identity; query cover 79%; E-value=0.0), and numerous species from different genera of the
134 family *Potyviridae*.

135 The InterPro tool for functional analysis of proteins [24], predicted nine protein domains in the viral
136 ORF: Potyviridae P1 protease domain (IPR002540); Helper component proteinase domain (PR001456);
137 Protein P3 of potyviral polyprotein (IPR039560); Superfamilies 1 and 2 helicase ATP-binding type-1
138 domain (IPR014001); Superfamilies 1 and 2 helicase C-terminal domain (IPR001650; Potyviridae
139 polyprotein domain (IPR013648); Potyvirus NIa protease (NIa-pro) domain (IPR001730); Viral RNA-
140 dependent RNA polymerase domain (IPR001205); and Potyvirus coat protein domain (IPR001592). The

141 tentative genome structure of the virus and predicted cleavage sites, obtained by multiple alignment of
142 109 different potyvirids, are shown in **Fig. 3**.

143 The virus does not appear to have a highly conserved G_{1-2} A_{6-7} motif preceding overlapping PIPO ORF
144 characteristic for *Potyviridae* and located within the P3 cistron [25], Instead, there is an A_7 motif, which
145 is embedded in the overlapping ORF, presumably (according to the ORF Finder;
146 https://www.ncbi.nlm.nih.gov/orffinder/) starting at the position 3122 nt (+3 ORF relative to the
147 polyprotein) and followed by 63 codons with a termination codon TGA at pos. 3311-3313.

148 **Phylogenetic analysis**

149 Phylogenetic analysis deduced from the alignment of the viral polyprotein and representative
150 polyproteins of different members of the family *Potyviridae*, clustered the virus into a distinct outgroup,
151 branched out of the genus *Ipomovirus* (**Fig. 4**). This concurs with the significant divergence of the viral
152 genomic and amino acid sequences from other members of the genus and likely indicates that the virus
153 is a member of a new, not yet established, genus in the family.

154 **Mechanical transmission to *N. benthamiana* plants**

155 Since potivirids are readily transmissible mechanically
156 (https://ictv.global/report/chapter/potyviridae/potyviridae), we attempted mechanical transmission of
157 the virus to *N. benthamiana*, a common indicator host for diagnosis of many plant viruses [26] (Hull,
158 2002). *N. benthamiana* plants mechanically inoculated by a sap produced from infected alfalfa leaves,
159 developed distinct symptoms two-three weeks post inoculation. The symptoms could be visually
160 characterized as chlorosis, mottling, and chlorotic lesions spots on non-inoculated leaves. The plants
161 were also stunted (**Fig. 5**). The infection of *N. bentamiana* plants was additionally confirmed by RT-PCR
162 assay with primers LN1095/96 developed based on the HTS data and specific for the virus (**Fig.6**). The

163 amplicon was of the correct size predicted by the analysis of the HTS-derived assembly and Sanger
164 sequencing verified it is a fragment of the virus.

165 **Transmission electron microscopy of the purified viral preparations**

166 Transmission electron microscopy (TEM) can provide reliable information on particle dimensions and
167 morphology, based on which a virus can be placed to a particular taxonomic group [27]. TEM
168 examinations showed that the virions were flexuous filaments with a modal length of 800-1050 nm and
169 the modal diameter of about 11–13 nm (**Fig.7**). This morphology is indicative of the members of the
170 family *Potyviridae* and the larger particle length suggests a relationship with the genus *Ipomovirus*.

171 **Mechanical transmission to alfalfa (*Medicago sativa* L.) plants**

172 Among several plants of two inoculated alfalfa cultivars, only a single plant of cv. CUF 101 developed
173 obvious symptoms ~ 20 days post inoculation. Symptoms can be visually described as mottling and vein
174 yellowing (**Fig. 8**). The symptomatic plant was tested, and the infection confirmed by RT-PCR assay with
175 two different sets of primers developed based on the HTS data and specific for the virus (**Fig.9**). PCR
176 products were sequenced and found to have originated from the virus in question.

177 **Serological examinations**

178 In Western blot (WB) assays, virus-infected tissues did not react with broad spectrum monoclonal
179 antibodies (Abs) to potyviruses (PTY 1), [28], (results not shown). However, they appeared to weakly
180 cross-react with polyclonal antisera to SPMMV (DSMZ, Braunschweig, Germany) when indirect WB assay
181 (membranes were probed with primary Abs to SPMMV following by secondary Ab-alkaline phosphatase
182 conjugate) was used. This reaction was accompanied by a high background and non-specific reactivity,
183 common for indirect serological detection protocols (**Fig. S1A**). When membranes were probed with
184 SPMMV primary Abs conjugated with AP (direct assay), the only reactive band derived from the purified
185 viral preparation, likely due to the large concentration of the viral coat protein rather than to cross-

186 reactivity with SPMM antiserum. No cross-reactivity was observed with virus-infected plant tissues (**Fig**
187 **S1B**). This experiment indicated a possible distant serological relationship between the novel virus and a
188 type species of the genus *Ipomovirus*. Although percent amino acid identity between the SPMMV CP and
189 presumptive CP of the alfalfa ipomovirus is quite low (31%), the latter may possess some epitopes
190 recognized by the SPMMV polyclonal Abs.

191 Discussion

192 We report here the identification and characterization of a new virus infecting alfalfa (*Medicago sativa*
193 L.) and belonging to the family *Potyviridae*. Genomic fragments of the same virus were previously found
194 in mature alfalfa seeds [15]. Since they were to some degree (~30%) identical to several potyvirids from
195 different genera of the family, we tentatively designated the virus AaPV1. In this study, based on the
196 complete genome sequence and structure, phylogenetic relationships, TEM observations,
197 symptomatology, and in part by serological reactivity, we reclassify and rename AaPV1 to alfalfa vein
198 mottling virus (AVMV), a novel species tentatively related to the genus *Ipomovirus*. However, considering
199 a distinct clustering of AVMV in the phylogenetic tree and low identity levels with other members of the
200 genus *Ipomovirus*, AVMV is likely to be a representative species of a different taxon, most closely related
201 to ipomoviruses, for which we propose a name Alvemovirus.

202 The characteristic symptoms caused by the virus on indicator plant *Nicotiana benthamiana* and on the
203 natural host alfalfa, suggest a pathogenic nature of the disease. Successful mechanical inoculation of *M.*
204 *sativa* with virus purified from the indicator plant followed by RT-PCR detection of the AVMV in
205 mechanically inoculated alfalfa, satisfies Koch's postulates and establishes the virus as cause of the
206 disease.

207 Ipomoviruses, to which AVMV is presumably related, infect a wide range of hosts, including economically
208 important crops such as tomato, cucumber, squash, sweet potato, watermelon, and others. To our
209 knowledge, they have not been diagnosed in alfalfa. Our findings indicate that *M. sativa* can be a natural

210 reservoir for the AVMV and thus, considering its possible relationship to ipomoviruses transmitted by
211 whiteflies, the virus could potentially be transmitted via the same (or different) vector to other crops
212 growing in the vicinity. Since all alfalfa samples used in this study were taken within 100 m of the melon
213 fields (**Table 1**), testing this crop would further clarify the virus origin and host range.

214 Conclusions

215 This work reports the discovery and characterization of a novel plant viral pathogen naturally infecting
216 alfalfa (*Medicago sativa* L.), a major forage crop worldwide. The virus identified is a member of the large
217 and economically important family *Potyviridae*, in which it tentatively belongs to the new genus
218 designated as Alvemovirus.

219

220 Funding information

221 This study was supported by the United States Department of Agriculture, Agricultural Research Service
222 (CRIS number 8042-21000-300-00D) and partially by the National Plant Disease Recovery System
223 (NPDRS) grants to LGN.

224 Author contributions

225 Conceptualization: LGN. Sample collection and evaluation: JCP and WMW; Methodology: OAP, LGN, SG.
226 Data analysis: SG and LGN. Experimentation: LGN and OAP. Writing – original draft preparation: LGN.
227 Writing – review and editing: all authors.

228 Data Availability Statement

229 The genomic sequence of the alfalfa vein mottling virus has been deposited in GenBank on 08/24/2023
230 under the accession number OR483877.

231 Conflicts of interest

232 The authors declare that there are no conflicts of interest.

233 ## References

234 1. Chen H, Zeng Y, Yang Y, Huang L, Tang B, et al. Allele-aware chromosome-level genome assembly
235 and efficient transgene-free genome editing for the autotetraploid cultivated alfalfa. *Nat*
236 *Commun* 2020; 11: 2494.
237 2. Postnikova OA, Shao J, Nemchinov LG. Analysis of the alfalfa root transcriptome in response to
238 salinity stress. *Plant Cell Physiol* 2013; 54:1041–1055.
239 3. Postnikova OA, Hult M, Shao J, Skantar A, Nemchinov LG. Transcriptome analysis of resistant and
240 susceptible alfalfa cultivars infected with root-knot nematode *Meloidogyne incognita*. *PLoS One*
241 2015;10(3): e0123157.
242 4. Nemchinov LG, Shao J, Lee MN, Postnikova OA, Samac DA. Resistant and susceptible responses
243 in alfalfa (*Medicago sativa*) to bacterial stem blight caused by *Pseudomonas syringae* pv.
244 *Syringae*. *PloS One*. 2017;12(12): e0189781.
245 5. Vieira P, Mowery J, Eisenback JD, Shao J, and Nemchinov LG. 2019. Cellular and transcriptional
246 responses of resistant and susceptible cultivars of alfalfa to the root lesion nematode,
247 *Pratylenchus penetrans. Front Plant Sci* 2019; 10:971.
248 6. Yu, L-X, Zheng P, Bhamidimarri S, Liu X-P, Main D. The impact of genotyping-by-sequencing
249 pipelines on snp discovery and identification of markers associated with verticillium wilt
250 resistance in autotetraploid alfalfa (*Medicago sativa* L.). Front Plant Sci 2017; 8:89.
251 7. Yu L-X, Zhang F, Culma CM, Lin S, Niu Y, et al. Construction of high-density linkage maps and
252 identification of quantitative trait loci associated with verticillium wilt resistance in
253 autotetraploid alfalfa (*Medicago sativa* L.). *Plant Dis* 2020; 104(5):1439-1444.
254 8. Lin S, Medina CA, Boge B, Hu J, Fransen S, et al. Identification of genetic loci associated with
255 forage quality in response to water deficit in autotetraploid alfalfa (*Medicago sativa* L.). *BMC*
256 *Plant Biol* 2020;20(1):303.

9. Bejerman N, Roumagnac P and Nemchinov LG. High-throughput sequencing for deciphering the virome of alfalfa (*Medicago sativa* L.). *Front Microbiol* 2020; 11:553109.
10. Nemchinov LG, Irish BM, Grinstead S, Shao J, Vieira P. Diversity of the virome associated with alfalfa (Medicago sativa L.) in the U.S. Pacific Northwest. *Sci Rep* 2020; 12(1):8726.
11. Jiang P, Shao J and Nemchinov, L. G. Identification of emerging viral genomes in transcriptomic datasets of alfalfa (Medicago sativa L.). *Virol J* 2019; 16, 153.
12. Nemchinov LG, Grinstead SC, Mollov DS. Alfalfa virus S, a new species in the family *Alphaflexiviridae*. *PLoS One*; 2017. 12 (5), e0178222
13. Vayssier-Taussat M, Albina E, Citti C, Cosson JF, Jacques MA, Lebrun MH, et al. Shifting the paradigm from pathogens to pathobiome: new concepts in the light of meta-omics. *Front Cell Infect Microbiol* 2014; 4:29.
14. Nemchinov LG, Irish BM, Uschapovsky IV, Grinstead, Jonathan Shao J, and Postnikova OA. Composition of the alfalfa pathobiome in commercial fields. *Front Microbiol* 2023; 14:1225781.
15. Nemchinov LG, Irish BM, Grinstead S, Postnikova OA. 2023. Characterization of the seed virome of alfalfa (*Medicago sativa* L). *Virol J* 2023; 20(1):96.
16. Bolger AM, Lohse M, and Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014; 30: 2114–2120.
17. Bankevich A, Nurk S, Antipov D, Gurevich A, Dvorkin M, et al. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012; 19: 455–477.
18. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol* 1990; 215:403-10.
19. Bushnell, B. BBMap: A fast, accurate, splice-aware aligner. 9th Annual Genomics of Energy & Environment Meeting, Walnut Creek, CA, USA. 2014. https://www.osti.gov/servlets/purl/1241166.

281 20. Goh CA and Hahn, Y. Analysis of proteolytic processing sites in potyvirus polyproteins revealed
282 differential amino acid preferences of Nia-Pro protease in each of seven cleavage sites. PloS One.
283 2021; 16(1): e0245853.
284 21. Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. RaxML-NG: a fast, scalable and user-
285 friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* 2019; 35: 4453–
286 4455.
287 22. Lesemann DE, Koenig R, Huth W, Brunt AA, Phillips S, Barton RJ. Poinsettia mosaic virus: A
288 tymovirus? *Phytopath Z* 1983; 107:250–262.
289 23. Nemchinov LG and Natilla A. Transient expression of the ectodomain of matrix protein 2 (M2e)
290 of avian influenza A virus in plants. *Protein Expr Purif* 2007; 56:153-9.
291 24. Paysan-Lafosse T, Blum M, Shuguransky S, Grego T, Pinto BL, et al. InterPro in 2022. *Nucleic Acids*
292 *Res* 2023; 51(D1): D418-D427.
293 25. Chung BYW, W Miller WA, Atkins JF, Firth AE. An overlapping essential gene in the Potyviridae.
294 *Proc Natl Acad Sci U S A* 2008; 105:5897-902.
295 26. Hull R. Matthews' Plant Virology. Fourth edition, Academic Press. 2002.
296 27. Bhat AI and Rao GP. 2020. Electron Microscopy and Utramicrotomy. In: Characterization of plant
297 viruses. Methods and Protocols. Humana Press, New York, NY U.S.A. p.173-185.
298 28. Jordan R, J Hammond J. Comparison and differentiation of potyvirus isolates and identification of
299 strain-, virus-, subgroup-specific and potyvirus group-common epitopes using monoclonal
300 antibodies. *J Gen Virol* 1991; 72; (Pt 1):25-36.

301

302

303

304 ## Figure legends

305 **Figure 1.** Bioinformatics pipeline

306 **Figure 2.** Diverse symptomatology of alfalfa plants in the field. **A**, asymptomatic plants. **B-D**, plants
307 displaying different virus-like symptoms.

308 **Figure 3.** Tentative genome structure of the new virus infecting alfalfa. Black arrow heads indicate
309 putative cleavage sites in the polyprotein. P1: protein 1 protease; P3: protein 3; PIPO: Pretty Interesting
310 Potyviridae ORF (Chung et al. 2008); 6K1: 6 kDa peptide 1; CI: cylindrical inclusion; 6K2: 6 kDa peptide 2;
311 VPg: virus protein genome-linked; Nia-Pro: nuclear inclusion a protease; NIb: nuclear inclusion b, RNA-
312 directed RNA polymerase; CP: coat protein.

313 **Figure 4.** Phylogenetic relationships of alfalfa vein mottling virus with other members of the family
314 *Potyviridae*. The unrooted tree was deduced from the Clustal W alignment and built with RaxML-NG [18]
315 (Kozlov et al. 2019) tool using maximum likelihood algorithm, 1000 bootstrap replicates and
316 bootstopping (autoMRE, cutoff: 0.030000).

317 **Figure 5**. Symptoms developed on *Nicotiana benthamiana* plants after mechanical inoculation with
318 AVMV-infected alfalfa. **A**, uninfected *N. benthamiana* plant. **B-D**, symptomatic plants.

319 **Figure 6.** RT-PCR assay with total RNA extracted from symptomatic *Nicotiana benthamiana* plants. Virus-
320 specific primers used in the assay were developed based on the HTS data. **M**, 1 Kb Plus DNA Ladder
321 (Thermo Fisher Scientific, Waltham, MA USA). **1**, RT-PCR amplicon obtained from the virus-infected plant
322 using primer pair LN1095/LN1096. **2**, no RT-PCR amplicons were obtained from the uninfected *N.*
323 *benthamiana* plants.

324 **Figure 7.** Transmission electron microscopy photograph of partially purified viral preparations.
325 Magnification: 30,000. Scale bar: 500 nm.

326 **Figure 8.** Symptoms developed on leaves of alfalfa (*Medicago sativa* L.) cv. CUF101 after mechanical
327 inoculation with partially purified viral preparations.

328 **Figure 9.** RT-PCR assay with symptomatic leaves of mechanically inoculated alfalfa plants. **M**, 1 Kb Plus
329 DNA Ladder (Thermo Fisher Scientific, Waltham, MA USA). **1**, PCR amplicon obtained from the
330 symptomatic leaves of alfalfa cv. CUF101 using primer pair LN1077/LN1078. **2**, no PCR amplicon was
331 obtained from control reaction using the same primers. **3**, PCR amplicon obtained from the symptomatic
332 alfalfa leaves using primer pair LN1079/LN1079. **4**, no PCR amplicon was obtained from control reaction
333 with the same primers.

334 **Table 1.** Geographic locations of alfalfa fields.

335 **Table S1**. Different viral contigs generated from each of the eight plant samples infected with the virus.

336 **Figure S1.** Western blot assays with polyclonal antibodies specific to sweet potato mild mottle virus
337 (SPMMV). **A: M**, ColorBurst protein marker (Sigma-Aldrich, St. Louis, MO USA); **1**, virus-infected tissues
338 of *Nicotiana benthamiana*; **2**, uninfected tissues of *N. benthamiana*; **3**, mechanically inoculated,
339 symptomatic tissues of alfalfa; **4**, uninfected alfalfa sample; 5, partially purified viral preparation. **B: M,**
340 ColorBurst protein marker (Sigma-Aldrich, St. Louis, MO USA); 1-4, virus-infected *N. benthamina* and
341 alfalfa leaf tissues; 5 and 6, healthy *N. benthamina* and alfalfa leaf tissues; 7, partially purified viral
342 preparation.

343

344

345

346

347

348

349

350

351

352

Table 1. Geographic locations of alfalfa fields

Alfalfa fields, №	GPS coordinates		Adjacent crops	№ of plants collected
24E	32° 40’ 56.58”N	114° 13’ 42.72”W	Cantaloupe, lettuce, cauliflower	5
32E	32° 42’ 38.62”N	114°05’ 28.21” W	Cantaloupe, cauliflower, spinach	5
43E	32° 47’ 11.44”N	113° 54’ 15.81”W	Cantaloupe, lettuce, broccoli, cauliflower	5
51E	32° 46’ 42.26”N	113° 45’ 20.89”W	Cantaloupe, lettuce	5

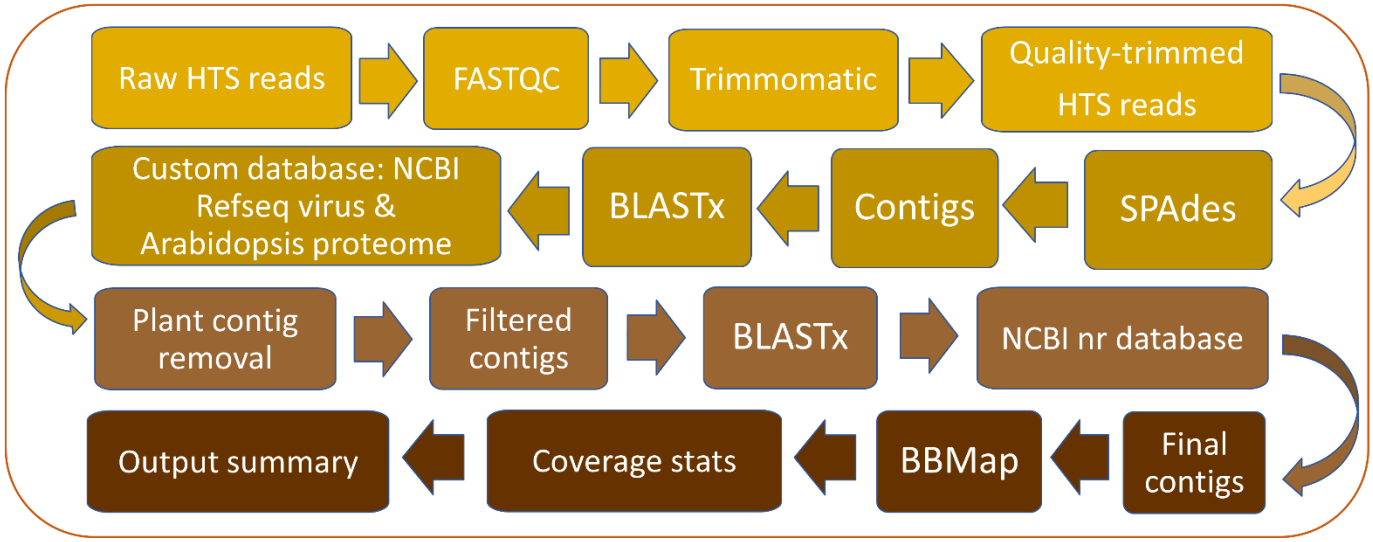

Raw HTS reads
FASTQC
Trimmomatic
Quality-trimmed HTS reads
SPAdes
Contigs
BLASTx
Custom database: NCBI Refseq virus & Arabidopsis proteome
Plant contig removal
Filtered contigs
BLASTx
NCBI nr database
Final contigs
BBMap
Coverage stats
Output summary

Figure 1

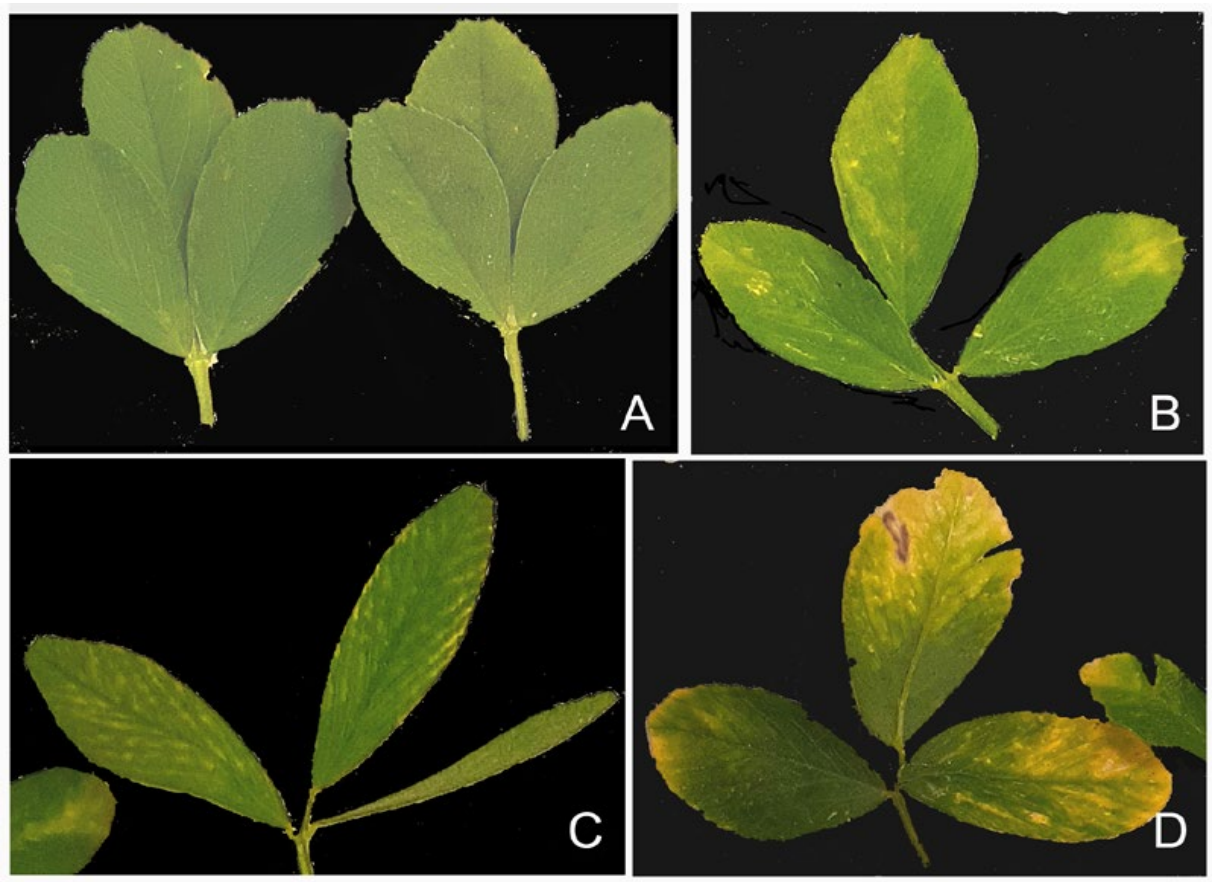


Figure 2

398

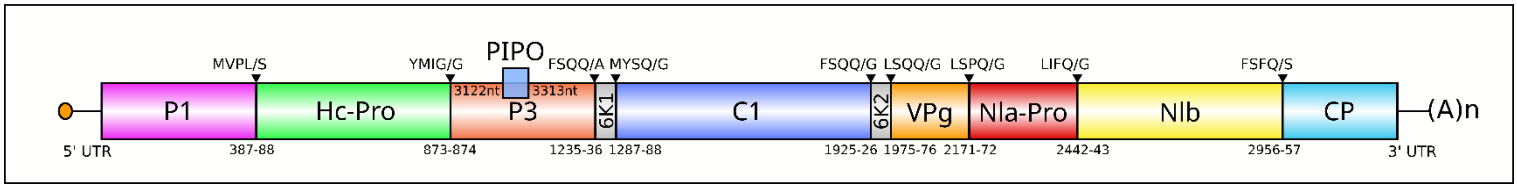
399

400

401

402

403



404

405 **Figure 3**

406

407

408

409

410

411

412

413

3.60
3.20
2.80
2.40
2.00
1.60
1.20
0.80
0.40
0.00
MH932227 Celery latent virus
MH330686 Areca palm necrotic spindle-spot virus
MH395371 Areca palm necrotic ringspot virus
KY491536 Bellflower veinal mottle virus
KP405232 Artichoke latent virus
KU053507 Broad-leafed dock virus A
MG755240 Yam chlorotic necrosis virus
AB710145 Chinese yam necrotic mosaic virus
KT724961 Yam chlorotic necrotic mosaic virus
MG978107 Alpinia oxyphylla mosaic virus
MF622947 Cardamom mosaic virus
MT110148 Amomum potyvirus A
LC055681 Rice necrosis mosaic virus
Y10973 Barley mild mosaic virus
AJ306718 Oat mosaic virus
AJ132268 Barley yellow mosaic virus Chinese
D86634 Wheat yellow mosaic virus
MH645040 Wheat spindle streak mosaic virus isolate WSSMV South
MN046368 Wheat spindle streak mosaic virus isolate WSSMV-OH2
JF280796 Rose yellow mosaic virus
MH379332 Passiflora edulis symptomless virus
Z48506 Brome streak mosaic rymovirus
EF608612 Wheat eqlid mosaic virus
AF057533 Wheat streak mosaic virus
AY377938 Oat necrotic mottle virus
KF260962 Tall oatgrass mosaic virus
KF984546 Yellow oat-grass mosaic virus
JX156425 Caladenia virus A
FJ263671 Triticum mosaic virus
GQ388116 Sugarcane streak mosaic virus
Alfalfa vein mottling virus
HE600072 Tomato mild mottle virus
HQ840786 Eggplant mild leaf mottle virus
Z73124 Sweetpotato mild mottle virus
EU259611 Squash vein yellowing virus
AY578085 Cucumber vein yellowing virus
KU935732 Coccinia mottle virus
GU563327 Cassava brown streak virus
FJ185044 Ugandan cassava brown streak virus strain Ugandan
HM181930 Ugandan cassava brown streak virus strain UG
AY994084 Blackberry virus Y
Y09854 Ryegrass mosaic virus
AY623626 Agropyron mosaic virus
AY623627 Hordeum mosaic virus
AJ001691 Maize dwarf mosaic virus
AJ297628 Sugarcane mosaic virus
AY206394 Beet mosaic virus
AJ507388 Soybean mosaic virus
AY112735 Bean common mosaic virus
X67673 Papaya ringspot virus
X12456 Potato virus Y
X16415 Plum pox virus
D83749 Bean yellow mosaic virus
AJ296311 Potato virus A
M11458 Tobacco etch virus
Celavirus
Arepavirus
Bevemovirus
Macluravirus
Bymovirus
Roymovirus
Tritimovirus
Poacevirus
Alvemovirus
Ipomovirus
Brambyvirus
Rymovirus
Potyvirus

Figure 4

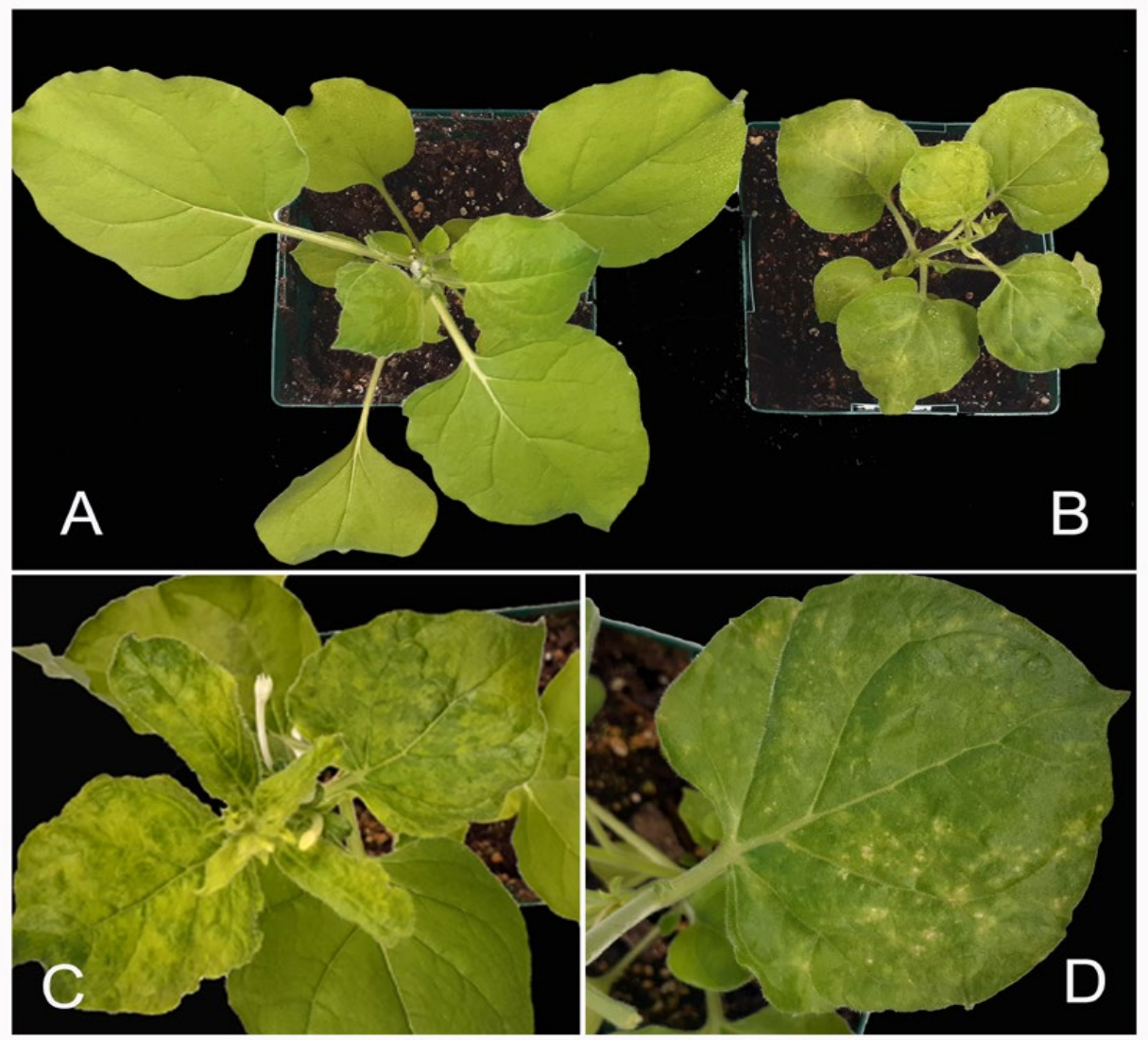


Figure 5

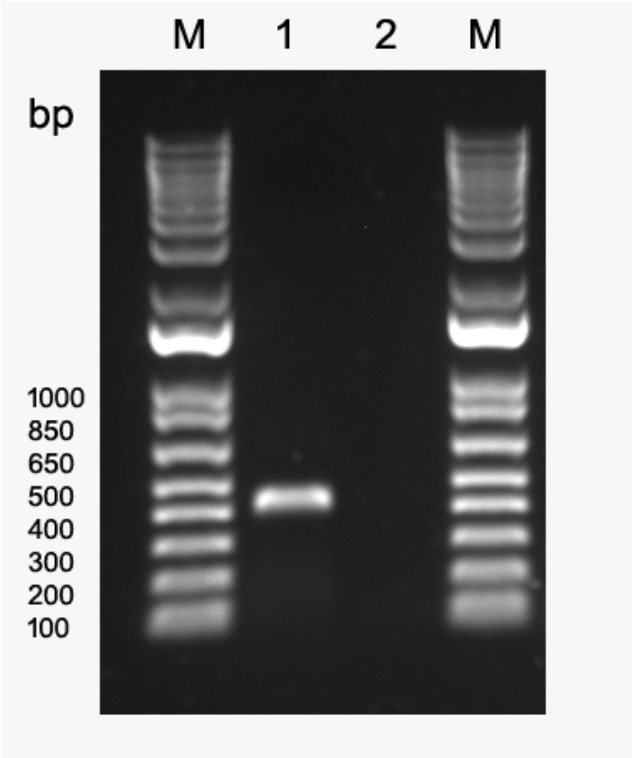
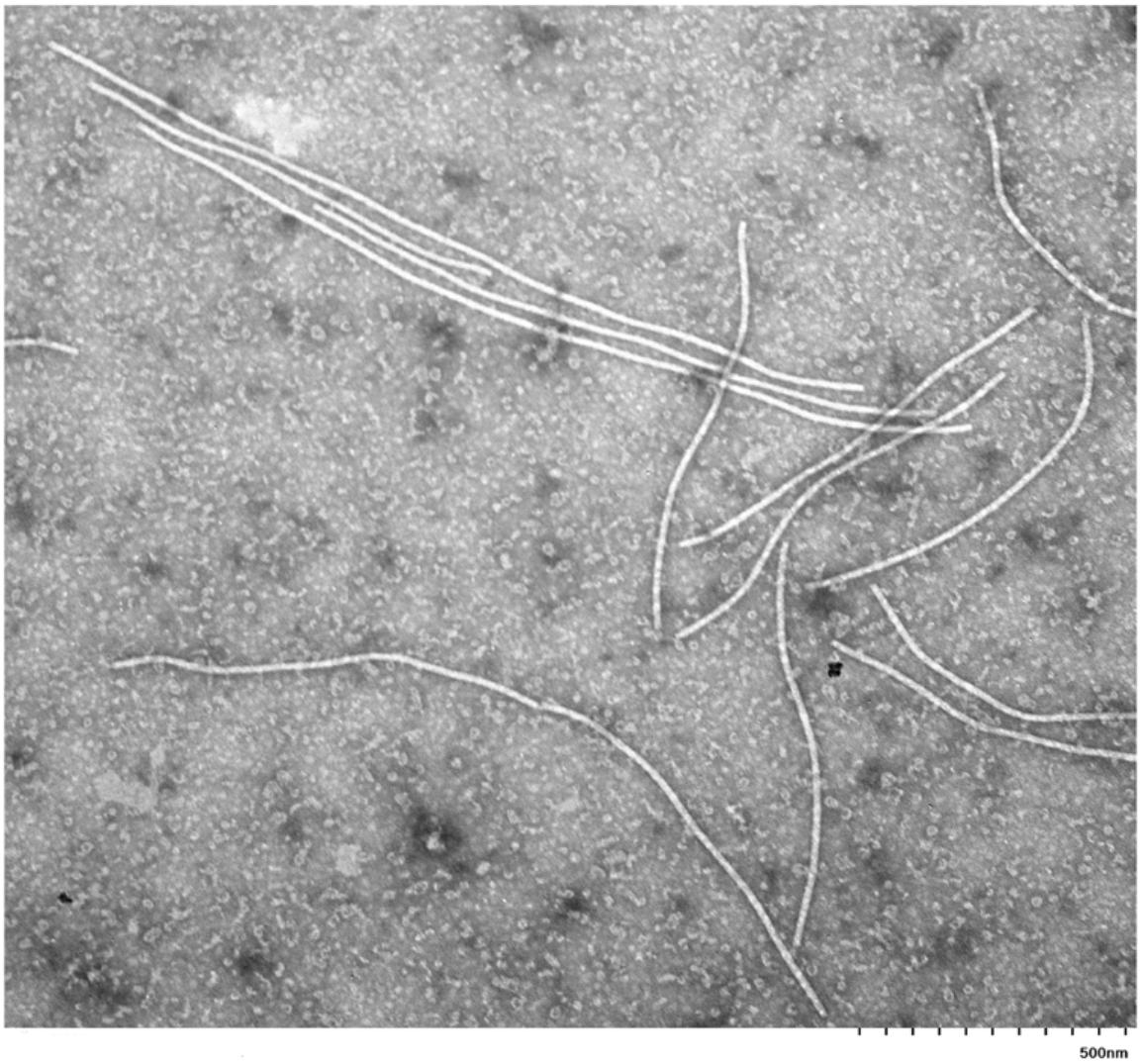


Figure 6

445



446
447

448 **Figure 7**

449

Figure 8

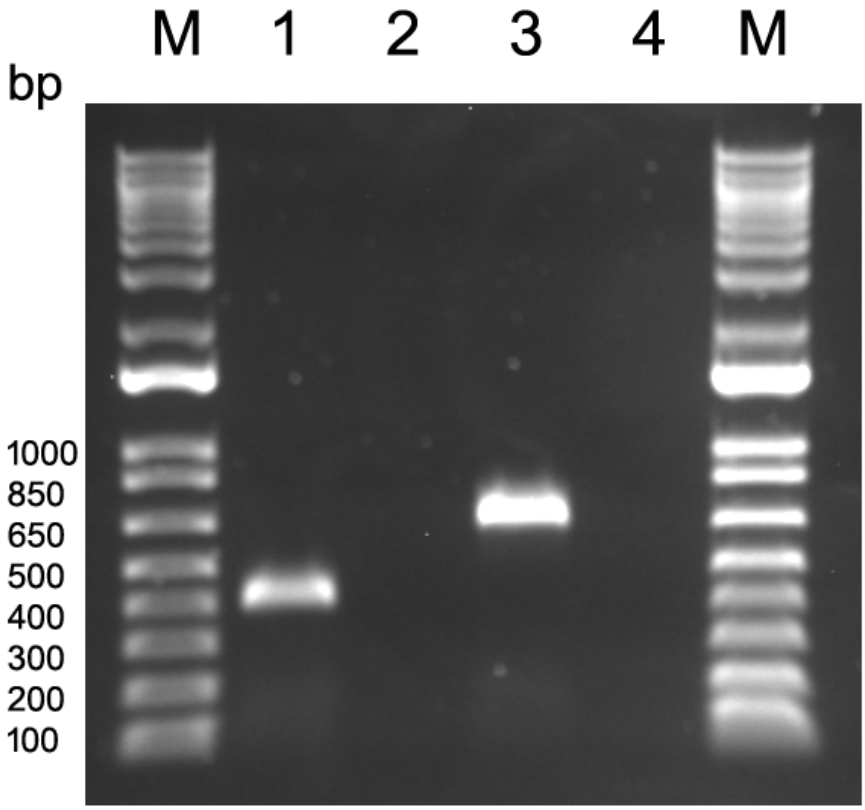


Figure 9