

1 Complete sequence and genomic annotation of Carrot torradovirus 2 1 (CaTV1)

3 Zuriñe Rozado-Aguirre^{1,2}, Ian Adams¹, Adrian Fox¹, Matthew Dickinson², and Neil Boonham^{1,3}

4 ¹Plant Protection Programme, Fera, Sand Hutton, York, YO41 1LZ, UK

5 ²University of Nottingham, Sutton Bonington, Leicestershire, LE12 5RD, UK

6 ³ IAFRI, Newcastle University, Newcastle upon Tyne, NE1 7RU

7

8 Corresponding author: zurine.rozado@fera.co.uk

9

10

11 Abstract

12 *Carrot torradovirus 1* (CaTV1) is a new member of the *Torradovirus* genus within the family
13 *Secoviridae*. CaTV1 genome sequences were obtained from a previous Next Generation Sequencing
14 (NGS) study and were compared to other members and tentative new members of the genus. The
15 virus is comprised of a bipartite genome and RACE was used to amplify and sequence each end of
16 RNA1 and RNA2. As a result RNA1 and RNA2 are estimated as containing 6944 and 4995
17 nucleotides respectively, with RNA1 encoding the proteins involved in virus replication, and RNA2
18 encoding the encapsidation and movement proteins. Sequence comparisons showed that CaTV1
19 clustered within the non-tomato infecting torradoviruses and is most similar to *Motherwort yellow*
20 *mottle virus* (MYMoV). The nucleotide identities of the Pro-Pol and coat protein regions were below
21 the criteria established by the ICTV for demarcating species, confirming that CaTV1 should be
22 classified as a new species within the *Torradovirus* genus.

23

24 Introduction

25 *Carrot torradovirus 1* (CaTV1) was an incidental finding discovered in a Next Generation Sequencing
26 (NGS) study seeking to elucidate the causes of internal necrosis in carrots (*Daucus carota*) in the UK

27 [1]. CaTV1 has been recently reported in carrot leaves collected in the Southwest of France,
28 indicating its presence also outside the UK [2]. The virus is similar to members of the *Torradovirus*
29 genus, first described in 2007 to place two new viruses affecting tomato crops, *Tomato torrado virus*
30 (*ToTV*) and *Tomato marchitez virus* (*ToMarV*) [3, 4]. Later more viruses affecting different crops have
31 been added to the genus, including, tomato chocolate virus (*ToChV*), Tomato chocolate spot virus
32 (*ToChSV*), Tomato necrotic dwarf virus (*ToNDV*), *Lettuce necrotic leaf curl virus* (*LNLCV*), *Motherwort*
33 *yellow mottle virus* (*MYMoV*), *Cassava torrado-like virus* (*CsTLV*) and *Squash chlorotic leaf spot virus*
34 (*SCLSV*) [5-11]. Torradoviruses are considered members of the *Secoviridae* family within the order
35 *Picornavirales* [12] and previous sequence comparisons within this genus established two different
36 clades for tomato-infecting (TI) and non-tomato infecting (NTI) members [13].

37 Although CaTV1 was a sequencing finding using NGS, the sequence of the virus was not fully
38 described. In this study we complete the characterization of the genome of CaTV1, including
39 completion of the 3' and 5' ends of both RNA fragments using RACE, annotation of the genome
40 identifying the location of translational features, and establishing the similarities between this and
41 other members of the *Torradovirus* genus.

42

43 **Completion of the CaTV1 genome**

44 The 3' and 5' ends of both RNA1 and RNA2 were amplified using the SMARTer RACE cDNA
45 amplification kit (Clontech) according to the manufacturer's protocols. The 3' and 5' PCR products
46 were analysed by direct sequencing. Results indicated that the 5' UTR and 3' UTR regions of RNA1
47 were 127 and 240 nt long respectively, and for RNA2, were 611 nt (5' UTR) and 327 (3' UTR) in
48 length. Completed sequences were deposited in GenBank with accession numbers KF533719.2 and
49 KF533720.2 for RNA1 and RNA2 respectively.

50

51 **CaTV1 genome organization**

52 CaTV1 - RNA1

53 As a member of the *Secoviridae* family, RNA1 is likely to code for a single ORF translated into a
54 polyprotein which is then likely to be processed by serine-like proteases into mature proteins [12].
55 Following analysis of RNA1, one predicted ORF (RNA1-ORF1) was identified (6944 nts), encoding a
56 putative polyprotein of 2192 amino acids (aa) with a molecular mass of 249 kDa. The translational
57 start (AUG) and stop (UAA) codons were found at nucleotide positions 127-129 and 6703-6705
58 respectively (Figure 1).

59 The complete nucleotide sequence was compared to other sequences available in GenBank using
60 BlastN, revealing 71 % identity to LNLCV (KC855266) and ToMarV (KT756874), 69 % to ToChSV
61 (GQ305131) and ToTV (KM091449) and 68 % to MYMoV (KM855266) and ToNDV (KC999058), all
62 members of the *Torradovirus* genus.

63 Identification of protein motifs were made based on previous characterization of ToTV and ToMarV [3,
64 4]. The polyprotein contains the conserved Hel-Pro-Pol replication block typical of picorna-like viruses
65 (nt positions 401-1516). Comparison of the aa sequence of the Pro-Pol region (1072-1516), limited
66 by the “CG” motif of the 3C-like proteinase and the “GDD” motif of the RNA-dependent RNA
67 polymerase (RdRp), suggested levels of similarity lower than those specified in the species
68 demarcation criteria (< 80 %) established in the ninth ICTV report [14]. This conserved domain is
69 typically used to determine differences among different picornavirales members. Pairwise
70 comparisons using the Hel-Pro-Pol protein sequences with other torradoviruses, showed that CaTV1
71 shares 58.5 %, 59.5 % and 41.2 % of the sequence with LNLCV, MYMoY and SCLSV respectively.
72 Lower levels were found when the same region was compared to tomato-infecting torradoviruses.
73 Typical helicase motifs (type III helicase), A (GKT), B (DD) and C (N) were found at positions 410, 456
74 and 507 respectively. The region, between aa 401-508, shares levels of identity up to 84 % with the
75 corresponding region of LNLCV and 82.9 % with MYMoV, while lower levels of identity were found in
76 the tomato-infecting torradoviruses (46.4 % ToTV and 50 % ToMarV). A histidine residue in the
77 putative protease substrate binding pocket is located at aa position 1075 and is required for
78 proteolytic processing in members of the *Secoviridae* family. Typical RdRp motifs (I-VII) were found
79 between aminoacid positions 1311-1596 [15]. Pairwise comparison revealed that the closest amino
80 acid identities in the RdRp domain were found with LNLCV (75.1 %) and MYMoV and ToChV (70.5
81 %).

82 To determine the relationship between CaTV1 and other viruses of the family, a neighbour-joining
83 phylogenetic tree with 500 bootstrap replications was constructed with the RdRp region using MEGA6
84 and ClustalX (figure 2a). The analysis revealed different clades for TI and NTI members confirming
85 previous results described [13]. It also indicates differences between NTI torradoviruses in the RNA1
86 with Squash chlorotic leaf spot virus (SCLSV), the latest proposed member of the genus, in an
87 independent clade.

88

89 CaTV1 - RNA2

90 RNA2 (4995 nt) has two predicted ORFs encoding two polyproteins. ORF1 encodes a putative
91 polyprotein of 202 aa with a predicted molecular weight of 22 kDa (figure 1). This ORF, partially
92 overlaps the large ORF2, and is a distinguishing feature of the *Torradovirus* genus. The closest
93 related member was MYMoV (83 % identity). No conserved motifs were found in the sequence and
94 several differences were seen between TI and NTI torradoviruses, confirming previous results
95 reported [13].

96 ORF2 encodes a large polyprotein of 1167 aa (130 kDa). The translational start (AUG) codon was
97 found at nt positions 1165-1167 while the stop (UAA) codon was at nt positions 4666-4668. When the
98 whole nucleotide sequence was compared to other members of the *Secoviridae* family, the most
99 closely related members were LNLCV (68 % identity) and MYMoV (66 % identity).

100 The region encoding the movement protein (MP) domain was found between aa 45-239 with the
101 typical LxxPxL motif in positions 211-216. Based on homology to related viruses and previous
102 identification of ToMarV and ToChSV cleavage sites [16], the cleavage site between the MP and the
103 first coat protein is likely to be located at position 487 (Q⁴⁸⁷/A⁴⁸⁸). The coat protein region shares
104 levels of identity of 42.2 % and 47.3 % with LNLCV and MYMoV respectively, which again are the
105 closest related members. This percentage also demarcates CaTV1 as a new species of the genus
106 according to the criteria established in the ICTV ninth report [14]. A glutamine residue (Q) at position -
107 1 is highly conserved in all the torradoviruses and could potentially act as a protease cleavage site
108 among the different proteins [16]. By comparison with ToMarV and ToChSV cleavage sites, a Q
109 residue has been identified at position 695 (Q⁶⁹⁵/S⁶⁹⁶) which could be a potential cleavage site

110 between the Vp35/Vp26 coat proteins. Additionally, the putative Vp26 and Vp23 cleavage site region
111 would be located at aa position 935 (Q⁹³⁵/I⁹³⁶). However, the actual number and size of the predicted
112 capsid proteins for CaTV1 have not been determined experimentally.

113 A phylogenetic tree was constructed with all the members of the *Torradovirus* genus using the whole
114 region with the three coat proteins (Figure 2b). In agreement with the results obtained with RNA1,
115 three different clades can be again differentiated: five viruses cluster in the tomato-infecting
116 torradovirus clade; CaTV1 is grouped with MYMoV and LNLCV in the non-tomato infecting (NTI)
117 torradovirus clade; and SCLSV clusters with Cassava torrado-like virus (CsTLV) sequences in a third
118 independent clade.

119 CaTV1 is transmitted by aphids [17] and a search of possible aphid transmission motifs was carried
120 out comparing the coat protein region of TI and NTI torradoviruses. DAG (Asp-Ala-Gly) is usually a
121 highly conserved motif in N-terminal of the coat protein of potyviruses and it has been demonstrated
122 that the change of any of the amino acid prevents transmission by aphids [18]. However this motif
123 could not be found within the CaTV1 sequence. TI torradoviruses have been described to be whitefly-
124 transmitted [19, 20], but no transmission motifs have been identified in the sequences for any of these
125 members so far.

126 The presence of a small (3–5 kDa) VPg linked to the 5' end of the RNAs has been confirmed for most
127 members of the order; comparative genomics strongly suggests that this property is universally
128 conserved among the genus and the *Secoviridae* family [13].

129 This study has described a complete characterization of the whole genome sequences of CaTV1 for
130 both RNAs by comparison with other members of the genus and confirms that CaTV1 is a new
131 species according to the criteria established by the ICTV. Phylogenetic studies using both RNAs have
132 also confirmed and given further evidence of the differences between non-tomato infecting
133 torradoviruses.

134

135 **Compliance with ethical standards**

136 This research was supported as project FV 382b through UK government funding under Defra Plant
137 Health Capability. All authors declare that they have no conflict of interest. This article does not
138 contain any studies with human participants or animals performed by any of the authors

139

140 **References**

141 1. Adams I.P., Skelton, A., Macarthur, R., Hodges, T., Hinds, H., Flint, L., Deb Nath, P., Boonham N.,
142 Fox, A., (2014). Carrot yellow leaf virus is associated with Carrot Internal Necrosis. PLOSone 9, Issue
143 11, e109125.

144 2. Rozado-Aguirre, Z., Marais, A., Svanella-Dumas, L, Faure, C., Latour, F., Villeneuve, F., Dickinson,
145 M., Fox, A., Boonham, N., Candresse, T. (2017). First Report of Carrot torradovirus 1 (CaTV1), a
146 member of the torradovirus genus, on carrots in France. Plant disease,
147 <http://dx.doi.org/10.1094/PDIS-01-17-0095-PDN>

148 3. Verbeek, M., Dullemans, A.M., van den Heuvel, J.F.J.M., Maris, P.C., van der Vlugt R., (2007).
149 Identification and characterization of tomato torrado virus, a new plant picorna-like virus from tomato.
150 Archives of Virology, 152, 881-890.

151 4. Verbeek, M., Dullemans, A.M., van den Heuvel, J.F.J.M., Maris, P.C., van der Vlugt R., (2008).
152 *Tomato marchitez virus*, a new plant picorna-like virus from tomato related to *tomato torrado virus*.
153 Archives of Virology, 153 (1), 127-134.

154 5. Verbeek, M., Dullemans, A.M., van den Heuvel, J.F.J.M., Maris, P.C., van der Vlugt R., (2010).
155 Tomato chocolàte virus: a new plant virus infecting tomato and a proposed member of the genus
156 *Torradovirus*. Archives of Virology, 155 (5), 751-755

157 6. Batuman, O., Kuo, Y.W., Palmieri, M., Rojas, M.R., Gilbertson, R.L., (2010). Tomato chocolàte spot
158 virus, a member of a new torradovirus species that causes a necrosis-associated disease of tomato in
159 Guatemala. Archives of Virology, 155, 857–869.

160 7. Pospieszny H, Borodynko N, Obrępalska-Stęplowska A, Hasiów B. (2007). The first report of
161 *Tomato torrado virus* in Poland. Plant Disease, Vol 91, 1364.

- 162 8. Verbeek, M., Dullemans, A., van Raaij, H.M., Verhoeven, J., van der Vlugt, R., (2013). Lettuce
163 necrotic leaf curl virus, a new plant virus infecting lettuce and a proposed member of the genus
164 *Torradovirus*. Archives of Virology, 159, (4), 801-805.
- 165 9. Seo, J.K., Kang, M., Kwak, H.R., Kim M.K., Kim, C.S., Lee, S.H., Kim, J.S., Choi H.S., (2014).
166 Complete genome sequence of motherwort yellow mottle virus, a novel putative member of the genus
167 *Torradovirus*. Archives of Virology, 160 (2), 587-90.
- 168 10. Carvajal-yepes, M., Olaya, C., Lozano, I., Cuervo, M., Castano, M., Cuellar, WJ., (2014).
169 Unraveling complex viral infections in cassava (*Manihot esculenta* Crantz) from Colombia. Virus
170 Research. 186, 76-86.
- 171 11. Lecoq H, Verdin E, Tepfer M, Wipf-Scheibel C, Millot P, Dafalla G, Desbiez C., (2016)
172 Characterization and occurrence of squash chlorotic leaf spot virus, a tentative new torradovirus
173 infecting cucurbits in Sudan. Archives of Virology, 161 (6), 1651-1655.
- 174 12. Sanfaçon, H., Wellink, J., Le Gall, O., Karasev, A., van der Vlugt, R., Wetzels, T., (2009).
175 Secoviridae: a proposed family of plant viruses within the order *Picornavirales* that combines the
176 families *Sequiviridae* and *Comoviridae*, the unassigned genera *Cheravirus* and *Sadwavirus*, and the
177 proposed genus *Torradovirus*. Archives of Virology, 154 (5), 899–907.
- 178 13. van der Vlugt, R., Verbeek, M., Dullemans, AM, Wintermantel, WM., Cuellar, WJ., Fox, A.,
179 Thompson, JR., (2015). Torradoviruses. Annual Review of Phytopathology, Vol 53, 485-512.
- 180 14. King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ. (2012). Ninth Report of the International
181 Committee on Taxonomy of Viruses. San Diego: Elsevier Acad.
- 182 15. Koonin, E.V., (1991). The phylogeny of RNA-dependent RNA polymerases of positive-strand RNA
183 viruses. J. Gen. Virol, 72 (Pt 9), p. 2197-2206.
- 184 16. Ferriol, I., Silva-Junior, D.M., Nigg, J.C., Zamora-Macorra, E.J., Falk, B.W., (2016). Identification
185 of the cleavage sites of the RNA2-encoded polyproteins for two members of the genus *Torradovirus*
186 by N-terminal sequencing of the virion capsid proteins. Virology, Vol 496, 109-115.

187 17. Rozado-Aguirre, Z., Adams, I., Collins, L., Fox, A., Dickinson, M., Boonham, N., (2016). Detection
188 and transmission of Carrot torrado virus, a novel putative member of the *Torradovirus* genus. Journal
189 of Virological Methods, Vol 235, 119-124.

190 18. López-Moya, JJ., Wangy, RY., Pirone, TP., (1999). Context of the coat protein DAG motif affects
191 potyvirus transmissibility by aphids. J Gen Virol. 80 (Pt 12), 3281-3288.

192 19. Verbeek, P.J., van Bekkum, A.M., Dullemans, A., van der Vlugt, R., (2013). Torradoviruses are
193 transmitted in a semi-persistent and stylet-borne manner by different whitefly vector. Virus Research
194 186, 55-60.

195 20. Amari, K., Gonzalez-Ibeas, D., Gómez, P., Sempere R.N., Sanchez-Pina, M.A., Aranda M.A.,
196 (2008). Tomato torrado virus is transmitted by *Bemisia tabaci* and Infects Pepper and Eggplant in
197 Addition to Tomato. Plant disease, Vol 92 (7), 1139.

198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224

225 Figure legends

226

227 **Figure 1:** Proposed genetic diagram for CaTV1 RNA1 and RNA2 with the positions of the ORFs
228 noted. Relative positions of regions containing helicase, protease and RNA-dependent RNA
229 polymerase motifs on RNA1, and movement protein and coat proteins on RNA2 are indicated.

230

231 **Figure 2:** (A) A phylogenetic tree of sequences of viruses within the *Secoviridae* family was
232 constructed using the RdRp amino acid region. (B) A second phylogenetic tree constructed using the
233 whole coat protein region of CaTV1 RNA2 of all the members of the *Torradovirus* genus. Alignments
234 were done using MEGA6 using the neighbour-joining algorithm. Sequences of all the viruses were
235 chosen from GenBank and accession numbers are shown. The numbers at the branch points are the
236 percentage bootstrap values following 500 bootstrap resampling and the scale indicates the number
237 of amino acid substitutions per site.

238

239

240

241

242

243

244

245

246

247

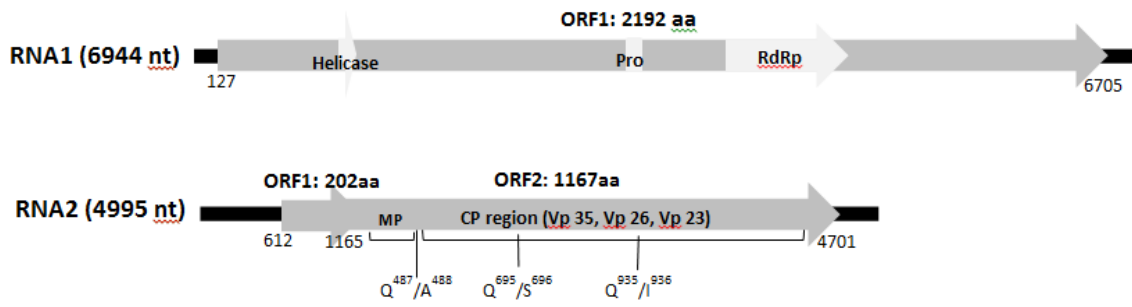
248

249

250

251 Figure 1

252



253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269 **Figure 2**

