Reverse Genetics with a Full-length Infectious cDNA Clone of Bovine Torovirus

Ujike Makoto*[P],1,2, Etoh Yuka[P]1,, Urushiyama Naoya1, Taguchi Fumihiro1, Enjuanes Luis 3 and Kamitani Wataru4.

1. Laboratory of Veterinary Infectious Diseases, Faculty of Veterinary Medicine, Nippon Veterinary and Life Science University, 1-7-1 Kyonan-cho, Musashino, Tokyo 180-8602, Japan

2. Research Center for Animal Life Science, Nippon Veterinary and Life Science University, 1-7-1 Kyonan-cho, Musashino, Tokyo 180-8602, Japan

3. Department of Molecular and Cell Biology, National Center of Biotechnology (CNB-CSIC), Campus Universidad Autónoma de Madrid, Darwin 3, 28049 Madrid, Spain

4. Department of Infectious Diseases and Host Defense, Gunma University Graduate School of Medicine, Gunma, Japan.

>

[P] equally contribute to this work

*Corresponding author:

Laboratory of Veterinary Infectious Diseases, Faculty of Veterinary Medicine, Nippon Veterinary and Life Science University, 1-7-1 Kyonan-cho, Musashino, Tokyo 180-8602, Japan

E-mail: ujike@nvlu.ac.jp tel: 81-422-31-4151 Ext. 3530

22   **Abstract**

23   Torovirus (ToV) has recently been classified in the new family Tobaniviridae, although it belonged to the

24   Coronavirus (CoV) family historically. Reverse genetics systems for many CoVs have been established, but

25   none exist for ToVs. Here, we describe a reverse genetics system using a full-length infectious cDNA clone of

26   bovine ToV (BToV) in a bacterial artificial chromosome (BAC). Recombinant BToV containing genetic markers

27   had the same phenotype as wild-type (wt) BToV. To generate two types of recombinant virus, the

28   Hemagglutinin-esterase (HE) gene was manipulated, since cell-adapted wtBToV generally loses the full-

29   length HE (HEf), resulting in soluble HE (HEs). First, recombinant viruses with HEf and HA-tagged HEf or

30   HEs genes were rescued; these showed no significant differences in cell growth, suggesting that HE is not

31   essential for viral growth in cells. Then, recombinant virus in which HE was replaced by the Enhanced Green

32   Fluorescent Protein (EGFP) gene expressed EGFP in infected cells, but showed significantly reduced viral

33   growth compared to wtBToV. Moreover, the recombinant virus readily deleted the EGFP gene after one

34   passage. Interestingly, one variant with mutations in non-structural proteins (NSPs) showed improved EGFP

35   expression and viral growth during serial passages, although it eventually deleted the EGFP gene, suggesting

36   that these mutations contributed to EGFP gene acceptance. These recombinant viruses provide new insights

37   regarding BToV and its reverse genetics will help advance understanding of this neglected pathogen.

38 **Importance**

39 ToVs are diarrhea-causing pathogens that have been detected in many species, including humans. BToV has

40 spread worldwide, leading to economic losses. We developed the first reverse genetics system for

41 Tobaniviridae using a BAC-based BToV. Using this system, we showed that recombinant BToVs with HEf

42 and HEs showed no significant differences in cell growth. In contrast, clinical BToVs generally lose the HE

43 gene after a few passages but some recombinant viruses retained the HE gene for up to 20 passages,

44 suggesting some benefits of HE retention. The EGFP gene of the recombinant viruses was unstable and was

45 rapidly deleted, likely via negative selection. Interestingly, one virus variant with mutations in NSPs was

46 more stable, resulting in improved EGFP-expression and viral growth, suggesting that the mutations

47 contributed to some acceptance of the exogenous EGFP gene without clear positive selection. The

48 recombinant BToVs and reverse genetics developed here are powerful tools for understanding fundamental

49 viral processes and their pathogenesis and for developing BToV vaccines.

Introduction

Torovirus (ToV) belongs to order *Nidovirales,* family *Tobaniviridae,* subfamily *Torovirinae* and genus *Torovirus* (1). ToVs are a causative agent of diarrheic and respiratory diseases, and are detected in many species including bovine, equine, swine, goat, and human species (2–7). Although ToV infections are generally asymptomatic or do not have severe symptoms, the bovine torovirus (BToV) (8–16) and porcine ToV (PToV) (17–22) are distributed globally, causing economic loss. At present, there are no drugs and vaccines available for treatment and prevention of this disease.

ToVs are enveloped, positive-sense single-stranded RNA viruses with a genome 25-30kb in length, comprising the conserved six open reading frames. The first two-thirds of the genome contains ORF1a and ORF1b, with an overlap by a frameshift, encoding the replicase/transcriptase proteins (22–25). The remaining one-third of the genome encodes four structural proteins: spike (S), membrane (M), hemagglutinin-esterase (HE), and nucleocapsid (N) (26–30). ToV genome seemed to have occasionally undergone inter-genotype recombination events (31–33).

ToVs were difficult to propagate in cultured cells, with the exception of equine ToV (6); however, in the last 13 years, a new cell line, human rectal tumor-18 (HRT18) cells, has been shown to be susceptible to BToV, and several BToVs in Japan were successfully isolated and propagated using HRT18 cells (33–37). BToV with the full length HE gene was initially isolated from diarrheal feces, whereas cell-adapted BToV, after several passages in HRT18 cells, generally lost a full-length HE protein as a result of HE gene mutation (34, 35). This finding suggests that the HE protein is not essential for replication in cell culture and may instead suppress it.

ToVs historically belonged to the family *Coronaviridae,* which was divided into the subfamilies coronavirus (CoV) and ToVs, since both viruses are structurally and morphologically similar despite having some differences (24, 38–40). CoVs are mainly associated with respiratory and enteric diseases and are regarded as important pathogens in humans and animals. In human CoVs, in addition to four human CoVs causing mild upper respiratory disease (41), three novel life-threating CoVs that cause acute lung injury have emerged in the 21st century, namely, severe acute respiratory syndrome coronavirus (SARS-CoV) in 2002-2003 (42, 43), Middle East respiratory syndrome coronavirus (MERS-CoV) in 2012 (44), and SARS-CoV-2 in 2019 (45, 46). In animals, some CoVs cause severe or lethal diseases in swine (porcine epidemic diarrhea virus [PEDV] and transmissible gastroenteritis virus [TGEV]), bovine (bovine CoV), avian (infectious

79 bronchitis virus [IBV]), mouse (mouse hepatitis virus [MHV]), and feline (Feline infectious peritonitis virus

80 [FIPV]) animals. As such, numerous studies have been conducted on CoVs. Notably, reverse genetics systems

81 for many human and animal CoVs, including the recently emerged SARS-CoV-2, have been established to

82 study these fundamental viral processes and pathogenesis or to aid in vaccine development (47–68). Full-

83 length cDNA-based reverse genetics systems of CoVs have been made, although there were obstacles such as

84 large genome sizes, thereby complicating genome engineering, and the instability of some CoV replicase

85 genes in bacteria. Currently, four systems are available: bacterial artificial chromosome (BAC) (48), in vitro

86 ligation of cDNA fragments (47), vaccinia vector with full-length CoV cDNA (59), and the recently developed

87 yeast artificial chromosome systems (60). Thanks to these systems, a greater understanding of CoVs has been

88 gained.

89 In contrast to CoVs, only few studies of ToVs have been conducted, and no reverse genetics systems have

90 been established since many ToVs are difficult to propagate in cells, and because less attention has been paid

91 to ToVs as ToV infections are generally asymptomatic or not severe. In this study, we describe a reverse

92 genetic system for BToV (Aichi strain) based on the cloning of a full-length genomic cDNA into BAC. Using

93 this system, recombinant viruses with HA-tagged HE gene, untagged full-length HE gene, or with EGFP

94 gene, were successfully rescued by manipulating the HE gene; these viruses were then characterized.

95

96

97 ## Results

98 ### Full-length genome sequence of wild-type (wt) BToV.

99 To determine the parent BToV (Aichi strain) for reverse genetics, the full-length genomic sequences of three

100 plaque-purified BToV were analyzed. We tried to determine the 5' terminal end of these viruses using the 5'

101 RACE method, but several different sequences were observed in the terminal three bases; 5'-TG<u>ACGT</u>-3', 5'-

102 GT<u>ACGT</u>-3', 5'-GG<u>ACGT</u>-3', 5'-G<u>ACGT</u>-3', 5'-<u>ACGT</u>-3', 5'-(-)<u>CGT</u>-3'. The sequence at the 5' terminal end of

103 three published BToVs (25, 33) and two published PToVs (Accession # JQ860350 and KM403390) all shared

104 the 5'-TG<u>ACGT</u>-3' sequence, whereas equine ToV showed 5'-<u>ACGT</u>-3' (74). Whether the differences observed

105 in the BToVs were due to the intrinsic features of BToV or due to suboptimal experimental condition of the

106 5' RACE method remained unclear. However, 5'-TG<u>ACGT</u>-3, previously reported in published BToVs and

107 PToVs, was found in only one out of seven in our analysis, whereas 5'-<u>ACGT</u>-3' reported in equine BToV is a

108 consensus sequence. Thus, in this study, the sequence at 5' terminal end of BToV was set as 5'-<u>ACGT</u>-3'. Two

109 BToV clones showed the same sequence; from these, one was selected as the parent virus (wtBToV).

110

111 ### Construction of a full-length BToV genome in BAC.

112 Full-length BToV (Aichi strain) genome cDNA was assembled into the pBeloBAC11 plasmid carrying

113 cytomegalovirus (CMV) immediate-early promoter, hepatitis delta virus ribozyme (Rz), and bovine growth

114 hormone (BGH) termination, and the cDNA was inserted downstream off the CMV promoter and flanked by

115 a 25-poly(A) sequence. The BToV genome cDNA was amplified into eight PCR fragments (BToV-A to H)

116 carrying homology arms (hms) at both terminals, and these were sequentially assembled (Fig. 1A) using the

117 Red/ET recombination method described in the Materials and Methods (Fig. 1B). After assembly of the eight

118 fragments, the BAC carrying the full-length BToV genome cDNA (pBAC-BToV^{mut1}) contained two single

119 mutations at nucleotide (nt) G3399Tand at nt T8469G, and the 1,350 bp *E. coli* chromosomal derived

120 sequence was inserted at nt 10,136, likely to counter the toxic regions in bacteria observed in some CoVs (47,

121 62, 63) (Fig. 2A). These two single mutations and *E. coli* sequence were reverted as described in the Materials

122 and Methods (Fig. 2A).

123

124 ### Rescue of recombinant BToV

125 To distinguish between the parent wtBToV and recombinant wtBToV (rBToV) generated from the BAC

126   plasmid, two gene makers were inserted into the HE gene at nt T26530C and T26554A (both silent) (pBAC-

127   BToV) (Fig. 2B). pBAC-BToV was first transfected into nonpermissive 293T, COS7, BHK, Vero, and Huh7

128   cells because of the very low transfection efficiency of permissive HRT18 cells. Of these cells, 293T showed

129   the best virus production with no clear cytopathic effect (CPE), and the viral titer of the supernatant 3 days

130   post-transfection (dpt) averaged $1.0 \times 10^{4.1}$ $TCID_{50}$/ ml (N=3), suggesting that the host range specificity of

131   BToV, which was difficult to propagate in cultured cells, was primarily determined by the entry step, which

132   was also observed in CoVs (47, 59, 63). The supernatant from transfected 293T was inoculated into permissive

133   HRT18 cells, after which the virus was plaque-purified three times, and then the purified virus was used to

134   characterize phonotypic properties. CPE and plaque morphology induced or formed by rBToV were identical

135   to parental wtBToV (Fig. 3B and 4A). Comparison of the full-length genome sequence between rBToV and

136   wtBToV showed complete sequence identities except for two gene markers (Fig. 3C). To compare growth

137   kinetics between rBToV and wtBToV, HRT18 cells were infected with a multiplicity of infection (MOI) of 0.001,

138   and supernatants were harvested at the indicated times. rBToV and wtBToV had indistinguishable growth

139   properties in HRT18 cells and had peaked titers of $>1.0 \times 10^{6.4}$ $TCID_{50}$/ml at 48 h post infection (pi) (Fig. 4D).

140   Thus, a full-length infectious cDNA clone of BToV in the BAC system was successfully constructed, and

141   features of the rescued rBToV and parental wtBToV had the same phenotypic properties in HRT18 cells.

142

143   **Characterization of recombinant BToV with full-length HE or HA-tagged HE gene**

144   Although BToV initially isolated from clinical samples had full-length HE (HEf) genes, cell adapted BToV

145   usually lost HEf due to a stop codon somewhere in HE gene. Cell-adapted wtBToV in our laboratory also has

146   a stop codon at nt 481 (the base number of HE gene), CAG (Q)→TAG (stop), resulting in soluble HE 160 aa

147   in length (HEs) (Fig. 3A). This suggested that HE protein is dispensable for virus replication in cells and may

148   have a negative effect on it (34–37). However, the precise functions of BToV HE protein in cellular viral

149   growth still remains elusive. To investigate HE gene stability and the effect of HE protein on viral growth in

150   HRT18 cells, recombinant BToVs with HEf, or with HA-tagged HEf and HEs genes to facilitate their detection

151   were generated (designated as rHEf, rHEf/HA, rHEs/HA, and rBToV was written as rHEs together here,

152   respectively) (Fig. 3A). All these recombinant BToVs were successfully rescued using the same method

153   described in rBToV. They displayed no significant differences in plaque morphology and CPEs (data not

154   shown) in HRT18 cells compared to parental wtBToV and rBToV (rHEs) (Fig. 4A). To detect HEf or HA-tagged

155   HEf and HEs proteins, HRT18 cells infected with recombinant BToVs were fixed at 24 hpi. The cells were

156   then subjected to an α-NA esterase assay or to indirect immunofluorescence (IF) using mouse anti-M and

157   anti-HE antiserum or using rabbit anti-HA and mouse anti-HA 12CA5 antibodies. Stained cells were

158   observed under confocal laser scanning microscopy (Fig. 4B). M proteins were detected similarly in all

159   cells infected with the recombinant BToVs, and HE proteins and α-NA esterase activity were observed

160   only in cells infected with rHEf and rHEf/HA. On the other hand, HEs/HA proteins were well stained

161   by both anti-HA antibodies (only rabbit anti-HA antibody was shown), whereas HEf/HA protein was

162   barely stained despite being well detected by mouse anti-HE antiserum (Fig. 4B). We next tried to detect

163   HEf/HA by immunoblotting. Infected cells were lysed, and detection was performed using rabbit anti-N

164   antiserum and two anti-HA antibodies. With similar detection of N proteins in all infected cells, HEs/HA

165   could be detected in both anti-HA antibodies similar to IF results, whereas HEf/HA could not be detected by

166   rabbit anti-HA but could only be detected by the mouse anti-HA 12CA5 antibody (Fig. 4C). To investigate

167   whether the different reactivity of the C-terminal HA-tag was due to the modification of the HA-tag in the

168   cytoplasm by either an infected cell metabolic change or viral proteins, HA-tagged proteins in COS7 cells

169   transiently expressing HEf/HA and HEs/HA were attempted to be detected in the same way. Both HA-tagged

170   proteins could be stained in IF (data not shown), but reactivity did not change in immunoblotting (Fig. 4C).

171   Although we could not rule out the possibility that differences were due to the modification or due to the

172   expression level of HA-tag of HEf/HA in infected cells from IF results, immunoblotting results at least

173   suggested that the C-terminal HA-tag of HEf/HA had the specific structure that affects reactivity of the anti-

174   HA antibodies. Finally, to compare growth kinetics, HRT18 cells were infected with these recombinant BToVs

175   with a MOI of 0.001. All these recombinant BToVs, wtBToV, and rBToV (rHEs) had no significant differences

176   in growth properties (Fig. 4D). These results indicated that recombinant BToVs could express full-length HE

177   and HA-tagged proteins in infected cells with growth ability comparable to those of the parental wtBToV and

178   to rBToV (rHEs)

179

180   **Full-length HE gene stability of rHEf and rHEf/HA**

181   To assess stability of the full-length HE gene of rHEf and rHEf/HA, HRT18 cells were infected with two

182   clones (No.1 and No.2) derived from different plaques of each virus, and viruses were serially passaged until

183   twenty passages (P20). Cells were infected with the viruses harvested at the indicated passage history and

184   were fixed or lysed at 24 hpi, and then subjected to either α-NA esterase assay or N protein-detection by

185   immunoblotting (Fig. 5A and 5B). N proteins were detected in all infected cells (Fig. 5B). On the other hand,

186   cells infected with rHEf No.2 and two rHEf/HA retained α-NA esterase activity until P19, whereas cells

187   infected with rHEf No.1 gradually decreased in activity during passages and completely lost activity at P19

188   (Fig. 5A). RT-PCR of the supernatants from infected cells (resulting in +1 passage history) using primers

189   flanking the HE gene showed that during serial passages, band sizes of rHEf No.1 without esterase activity

190   and rHEf No.2 with esterase activity were not different (Fig. 5C). On the other hand, although both rHEf/HA

191   exhibited esterase activity, a smaller band was observed in rHEf/HA No.2 from P10. To further study their

192   HE genes, all these viruses at P19 were plaque-purified once, and the HE gene of viruses from six well-

193   isolated plaques were sequenced. As summarized in Fig. 5D, all six rHEf No.1 sequences had one base

194   deletion at nt T19 in the HE gene, causing 17 aa short peptide with a stop codon by -1 frame shifting, and

195   three out of six had an additional deletion between nt 1,216 to nt 1,241. In contrast, all six rHEf No.2 retained

196   the full-length HE gene, and one out of six had an amino acid substitution at T321I in the HE protein. Out

197   of the six, five rHEf/HA No.1 retained the HA-tagged full-length HE gene, but one had an aa mutation within

198   the HA-tag (YPYDYPDYA→YLYDYPDYA), whereas three rHEf/HA No.2 retained the HA-tagged full-length

199   HE gene with a D247E substitution in the HE protein in one virus but the other three had a deletion between

200   nt 278 to nt 878, causing a short soluble HE protein 99 aa in length. These data partly supported the idea

201   that HE proteins are not essential or could have a negative effect on virus growth in HRT18 cells. However,

202   it is notable that each one clone could retain the full-length HE gene up to at least twenty passages. Thus,

203   cell adapted BToV may be able to retain full-length HE gene more stably in cells under certain conditions.

204

205   **Characterization of recombinant BToV carrying EGFP gene**

206   Since recombinant viruses carrying the reporter gene are helpful in understanding fundamental viral

207   processes and for screening therapeutic compounds, we attempted to generate recombinant BToV by

208   replacing the HE gene with the EGFP gene (rEGFP). Rescued rEGFP were plaque-purified three times and

209   expanded (clone No.1 and No.2) or expanded without plaque purification (No.3 to No.6) because of the

210   instability of the EGFP gene described below. rEGFP formed smaller and more heterogeneous plaques

211   compared to the parent wtBToV (Fig. 6A). To detect EGFP in infected cells, infected cells at 36 hpi were

212   observed under fluorescence microscope, or fixed, permeabilized, stained with mouse anti-M antiserum, and

213   observed under confocal laser scanning microscopy (Fig. 6B). EGFP-expressions were observed only in cells

214   infected with rEGFP but not with wtBToV and rBToV, and cells expressing EGFP always also expressed M

215   protein (Fig. 6B). These infected cells were next subjected to N protein and EGFP detection by

216   immunoblotting. Although the amount of N protein in the rEGFP-infected cells was significantly less than

217   that of cells infected with wtBToV and rBToV, EGFP was only detected in rEGFP-infected cells (Fig. 6C).

218   Comparison of growth kinetics among rEGFP, wtBToV, and rBToV showed that rEGFP significantly

219   decreased growth ability. In four independent experiments, no virus could be detected at 24 hpi, and peak

220   titer $1.0 \times 10^{3.8\text{-}4.6}$ TCID$_{50}$/ml was observed at 48 to 72 hpi in three experiments, which decreased by >2 log

221   compared to parent viruses (Fig. 6D). In one experiment, peak titer reached $1.0 \times 10^{6.6}$ TCID$_{50}$/ml at 72 hpi,

222   but the virus no longer expressed EGFP (Fig. 6D). These data indicated that rEGFP could express EGFP in

223   infected cells but significantly decreased its growth properties. Moreover, rEGFP seemed to readily lose

224   EGFP-expression during virus growth.

225   To study EGFP gene stability, HRT18 cells were infected with rEGFP, and viruses were serially passaged

226   until five passages (P5). Cells were infected with rEGFP harvested at the indicated passage history and were

227   observed under a fluorescence microscope (Fig. 7A). As a result, EGFP was not observed in most of the cells

228   infected with rEGFP (No.1 to No.4) after one passage. RT-PCR of P0 and P1 viruses using primers flanking

229   EGFP gene found that the all P0 rEGFP predominantly had the full-length EGFP gene, but smaller bands

230   were observed in all P1 rEGFP (No.1 to No.4) (Fig. 7B). Of these, the major bands of P1 rEGFP (No.2 to No.4)

231   were analyzed, and it was found that a large part of the EGFP gene was deleted (Fig. 7D), suggesting that

232   rEGFP readily lost EGFP-expression due to deletion of EGFP gene after one passage.

233

234   **Improvement of EGFP-expression and growth properties in rEGFP variant**

235   Interestingly, rEGFP No.5 and No.6 maintained the EGFP gene after one passage. No.6 gradually decreased

236   EGFP-expression until P5, while No.5 appeared to gradually increase (Fig. 7A). Thus, serial passages of

237   rEGFP No.5 was continued until P15, and the viruses were then harvested. Cells were infected with rEGFP

238   No.5 harvested at the indicated passage history and their supernatants (+1 passage history) were collected

239   and subjected to RT-PCR targeting EGFP gene (Fig. 7B), or the infected cells were lysed at 36 hpi and

240   subjected to N protein and EGFP detection by immunoblotting (Fig.7C). Although smaller bands were

241   observed, rEGFP No.5 retained full-length EGFP gene at least until P6 and mainly lost it by P9 (Fig. 7B),

242 but EGFP-expression was still observed in infected cells even at P9 to an extent (data not shown). The major

243 bands of P15 rEGFP No.5 (and P6 No.6: data not shown) were analyzed by sequencing, and a part of the

244 EGFP gene was deleted (Fig. 7D). Similar to RT-PCR results, immunoblotting of N protein and EGFP of

245 rEGFP showed that EGFP-expression level of No.5 was maximized at P5, and this gradually decreased until

246 P11 (Fig. 7C). It should be noted that the EGFP-expression level or growth property seemed to dramatically

247 increase compared to that of P0 rEGFP No.5.

248 To further investigate P5 rEGFP No.5, this was plaque-purified three times and three clones were expanded

249 (designated rEGFP c1-c3). These variants formed similar plaque morphologies to wtBToV and rBToV (Fig.

250 8A). Comparing growth kinetics among rEGFP c1-c3, wtBToV, and rBToV showed that their peak titer

251 reached $1.0 \times 10^{6.5-7.2}$ $TCID_{50}$/ ml at 72 hpi, which is comparable to that of the parental viruses, although the

252 peak was reached one day slower (Fig. 8B). Cells infected with rEGFP c1-c3 were lysed at 24 hpi, and these

253 lysates were then subjected to N protein and EGFP detection. The result indicated that EGFP-expression

254 levels in these rEGFP c1-c3 were dramatically improved compared to P0 original virus (Fig.8C). Analysis of

255 the full-length genome sequence of rEGFP c1 revealed five aa substitutions at G1001S, C1442F, T2129I in

256 NSP1, I3562T in NSP4, and I5327V in NSP10 (Helicase), respectively, strongly suggesting that these

257 accumulated mutations in NSPs were involved in this phenotype (Fig.8D). Finally, EGFP gene stability in

258 rEGFP c1-c3 were investigated. As shown in Fig. 8E, these variants tend to stably express EGFP until P4,

259 but smaller bands due to the deletion of the full-length HE gene were mainly observed, with EGFP-expression

260 decreasing until P6 (data not shown in rEGFP c3). Sequencing the major band of P7 rEGFP c1-c3 showed

261 deletion of the EGFP gene, similar to those observed in other rEGFP (Fig.7D). These data indicated that any

262 or all of the five mutations in NSPs would contribute to some acceptance of exogenous gene such as EGFP

263 gene, resulting in improved EGFP-expression and viral growth in cells, although the EGFP gene was

264 eventually lost

265

266   **Discussion**

267   In this study, we successfully constructed a full-length infectious BToV cDNA clone in a bacterial artificial

268   chromosome. To our knowledge, this is the first reverse genetics system for the new family *Tobaniviridae*.

269

270   **BAC construction**

271   Red/ET recombination method used to assemble eight BToV cDNA fragments into the BAC has been recently

272   utilized for full-length cDNA assembly of some CoVs, PEDV, and FIPV (58, 75) and for introducing mutations

273   (76–78). This method is more rapid and produces mutated clones more efficiently than traditional methods

274   (78); however, despite several tries, the unintentional insertion of an *E. coli*-derived sequence which may

275   counter toxic regions could not be removed. Red/ET seemed not to function in restoring the toxic region to

276   the host. Thus, a type IIS restriction enzyme, SapI, was inserted into both ends of the *E. coli*-derived sequence,

277   and this was removed by in vitro self-ligation. The resulting plasmid could be successfully propagated in *E.*

278   *coli* without pRed/ET enzyme. Similar observations were not seen in the full-length genome assembly of

279   PEDV and FIPV into low copy BAC plasmid (58, 75), but were seen in the subcloning of the TGEV cDNA

280   fragment into high copy plasmids (47). The position wherein the *E. coli* sequence is inserted is similar in both

281   (BToV at nt 10,136 and TGEV at nt 9,973), suggesting that these regions (BToV is in 3C-like protease) were

282   particularly toxic in bacteria in ToVs and CoVs. Though it remains unclear whether very small amounts of

283   3C-like protease in *E. coli* or, functional viral proteins cleaved by this protease causes toxicity, the *E. coli*-

284   derived sequence was inserted not when the BToV-C fragment containing the 3C-like protease gene was

285   assembled but when the full-length 1ab gene (BToV-F fragment) was assembled. For BToV, since insertion of

286   the *E. coli*-derived sequence was observed even in low copy BAC plasmid, this region may be more toxic than

287   that of CoV or Red/ET recombination method may be prone to this. Thus, to establish other BAC-based ToV

288   reverse genetics using this method, it may be better to counter the toxic region (3-C like protease gene)

289   beforehand with an *E. coli*-derived sequence flanked by type IIS restriction enzymes for subsequent in vitro

290   ligation.

291

Characterization of recombinant BToV with manipulating HE gene.

Since cell-adapted BToV generally lost the full-length HE gene, HE protein is not essential for viral growth in cells or could suppress growth (34–37, 69). To study this, we generated rHEf with full-length HE gene and rHEf/HA, rHEs/HA with HA-tagged HEf, or HEs gene. Comparing their growth kinetics to wtBToV and rBToV (rHEs) showed no significant differences; however, during serial passaging of two clones from each rHEf and rHEf/HA, each clone lost the full-length HE gene, supporting the above idea. In contrast, each the other clone could retain the full-length-HE gene at least until twenty passages. This was unexpected because another BToV (Niigata strain) isolated from a clinical sample in our laboratory easily lost the full-length HE gene somewhere due to a stop codon insertion until eight passages during cell-adaptation (Islam and Taguchi, unpublished data). Thus, the negative effect on HE protein in the cells of cell-adapted Aichi-BToV may be lower than that of non-cell-adapted Niigata-BToV. Furthermore, HE proteins of cell-adapted Aichi-BToV may have some advantage in terms of viral growth in the cells.

HE proteins are retained not only in ToV, but also in some of the beta CoVs (79). HE is a type I glycoprotein on the viral envelope with two reversible functional domains: one binds O-acetylated sialic acids (O-Ac-Sia) and the other destroys this binding. Both viruses also contain another type I glycoprotein S protein responsible for receptor-binding and fusion activity. Among some beta CoVs, the species *Betacoronavirus-1*, including human coronavirus OC43 (HCoV-OC43), bovine coronavirus, and porcine hemagglutinating encephalomyelitis virus, use O-Ac-Sia as the receptor via the S protein (80, 81). Thus, these viruses bind to O-Ac-Sia via both the S and HE proteins, while HE acts as a receptor-destroying enzyme that promotes release of progeny viruses from infected cells and helps avoid attachment to non-permissive cells or to receptor-like decoy. The importance of HE in these viruses has been reported, and HE of HCoV-OC43 plays an essential role in the efficient release of progeny viruses from infected cells (82) and balancing receptor-binding and receptor-destroying of HE contributes to host adaptation (83). In contrast, the species *Murine coronavirus*, a mouse hepatitis virus (MHV) in which the S protein uses CAECAM1a as receptor (84) and in which HE binds to O-Ac-Sia exclusively (85) does not require HE proteins for viral growth. In fact, many cell-adapted laboratory strains of MHV fail to produce HE (86), with MHV rapidly losing it during serial passages (87, 88), which is similar to BToV. On the other hand, experiments using recombinant MHV carrying the full-length HE gene showed about half of the HE deficiencies until six passages (88), while our experiment showed

320   that some recombinant BToVs could retain HE at least until twenty passages. This may be explained by the

321   hemagglutination activity of the S protein of cell-adapted Aichi-BToV (37). It is not known whether the

322   hemagglutination activity is mediated via O-Ac-Sia binding, but, if so, HE protein may contribute to the

323   regulation of S protein-mediated cell surface attachment to facilitate release of progeny viruses. Alternatively,

324   in twenty passages, accumulation of mutations in other regions may contribute to HE protein acceptance, as

325   observed in rEGFP. Although the precise function of BToV HE remains to be elucidated, further studies using

326   these recombinant BToVs will bring better understanding.

327

328   **Characterization of recombinant BToV with EGFP gene.**

329   Since recombinant viruses expressing reporter proteins such as GFP, Red Fluorescent Protein and luciferase

330   are useful for understanding fundamental viral processes and for high-throughput testing of therapeutic

331   compounds, some recombinant CoVs carrying these reporters were generated (50, 54, 55, 57, 58). These

332   recombinant CoVs have their ORF3, ORF4 or ORF5 accessory genes replaced with reporter genes, and these

333   could all stably express reporter proteins, with their growth abilities comparable to or slightly reduced

334   compared to the parent virus. In contrast, rFGFP in which the HE gene, which is also an accessory gene, was

335   substituted with EGFP gene, could express a fluorescent protein, but it showed significantly reduced growth

336   compared to parental wtBToV, as well as less EGFP-expression. Furthermore, rEGFP easily lost EGFP-

337   expression ability after only one passage due to the deletion of HE genes. Since EGFP itself is a neutral

338   protein, which probably does not give any advantage or disadvantage to viral growth in cells, whereas rEGFP

339   with deleted EGFP gene increased in terms of growth properties (Fig.6D), EGFP gene itself rather than

340   EGFP would lead to a significant decrease in viral growth. It is interesting to note that, despite eight

341   independent experiments, starting points of HE deletion in the five were concentrated at around specific

342   positions, namely nt 24, 27, 28, and 30, and likewise the end points of the four were concentrated at around

343   nt 668, 671, 673, and 690 (Fig.7D), suggesting that "specific negative positions" of the EGFP gene are easily

344   recognized and deleted. Deletion patterns of the EGFP gene showed that in many cases, a large part of or the

345   entire EGFP, even including a part of the transcription-regulating sequence of EGFP (HE) and, the

346   untranslated region between EGFP (HE) and the N genes, were deleted, whereas a relatively short deletion

347   leading to retention of more than 80% of the EGFP gene was also observed (Fig. 7D). All rEGFP with deleted

348   genes rapidly replaced intact rEGFP, likely under negative selection of EGFP genetic structure. These

349   suggested that either the entirety or a part of the EGFP gene structure may suppress viral genome replication

350   or transcription, resulting in decreased viral growth.

351   On the other hand, rEGFP variant obtained during serial passages showed remarkably improved EGFP-

352   expression and growth ability in the cells. This is noteworthy considering EGFP gene has a negative effect

353   on viral growth whereas a neutral EGFP protein exists without clear positive selection. rEGFP variant

354   contained accumulated five mutations in NSPs: three in NSP1, and one each in NSP4 and NSP10 (helicase),

355   respectively. Though functions of NSP1 and NSP4 of ToVs remain unknown, the primary function of helicase,

356   which is highly conserved among nidoviruses including ToVs and CoVs, is to unwind DNA or RNA duplexes

357   (89). The crystal structure of MERS-CoV helicase (corresponding to NSP13) showed that it comprises

358   multiple domains, an N-terminal Cys/His rich domain (CH) binding to three zinc atoms, a beta-barrel domain,

359   and a C-terminal helicase core with two Rec-A like domains (90). Protein sequence alignment between MERS

360   and BToV helicase by MEGA6 software (91) revealed that the mutation at I5327V (I68V in helicase) was

361   located next to Cys69 (corresponding to Cys72 of MERS-CoV helicase) in the CH, one of the cysteine residues

362   that coordinated with the third zinc atom. It was reported that alanine substitutions in highly conserved

363   Cys/His in CH of HCoV-229E and other nidovirus equine viral arteritis have big effects on helicase activity

364   (92, 93). Thus, one possibility is that mutation in the CH of rEGFP helicase may facilitate to unwind "specific

365   negative positions" of the EGFP gene, resulting in increased growth. However, even this variant showed

366   similar EGFP gene deletion patterns to other intact rFGFP until five passages. This finding that the virus

367   tolerates unfavorable foreign gene to some degree without clear positive selection may contribute in

368   understanding the mechanism by which the virus evolves by accepting other genes.

369

370   In this article, we descraibe a reverse genetics platform for BAC-based BToV, which is a useful tool for

371   understanding fundamental viral processes and pathogenesis and for BToV vaccine development. Moreover,

372   since the rEGFP variant carrying the middle stable EGFP gene can be plaque-purified and expanded for

373   practical use, this recombinant virus can be used for high-throughput testing of therapeutic compounds and

374   entry assays. Recently, in addition to a unique combination of discontinuous and continuous ToV RNA

375   synthesis(94) , differences between CoVs and ToVs such as "hidden" proteins (U1 and U2) of unknown

376   function suspected to be translated from non-AUG initiation codons (CUG-initiation) (95) and N protein

377   localization, have been highlighted. This reverse genetics system can provide future understanding of ToV

378   proteins and unique RNA synthesis mechanisms.

379

Materials and Methods

Cells and Viruses

Human rectal adenocarcinoma subcell line (HRT18-Aichi) (34, 36, 69), 293T, and COS7 cells were maintained at 37°C and 5% CO_2 in Dulbecco's Modified Minimum Eagle Medium (DMEM) (Sigma-Aldrich, MO, USA) supplemented with 10% fetal bovine serum (FBS) [DMEM(+)] and penicillin (50 units/mL)-streptomycin (50 µg/mL) (PS) (Sigma-Aldrich). Cell-adapted BToV (Aichi strain) was propagated in HRT18 cells in DMEM without FBS [DMEM(-)], as FBS inhibits BToV infection.

Antibodies

Anti-HE, anti-M polyclonal mouse (mouse anti-HE and anti-M) antiserum, and anti-N polyclonal rabbit (rabbit anti-N) antiserum were obtained by in vivo electroporation of mice with pCAGGS-BToV-HEf according to the method previously described (70), by immunizing mice with the M protein synthetic peptide (2-FETNYWPFPDQAPN-15), and by immunizing rabbit with the N protein synthetic peptide (141-EVSSGDQETPHKIA-154). Anti-HA-tag mouse monoclonal antibody (Clone: 12CA5) (mouse anti-HA 12CA5) was obtained from supernatants of the hybridoma cells. Anti-HA-tag polyclonal rabbit IgG (rabbit anti-HA) antibody (MBL Nagoya, Japan) and anti-GFP mouse monoclonal antibody (mFx75) (mouse anti-GFP) (Fujifilm Wako, Osaka, Japan) were purchased. Secondary antibodies, horseradish peroxidase (HRP)-conjugated goat anti-rabbit IgG (Thermo Fisher Scientific, MA, USA), HRP-conjugated goat anti-mouse IgG (Rockland Immunochemicals, PA, USA), Alexa Fluor™ 594-conjugated goat anti-mouse IgM and IgG (Thermo Fisher Scientific), FITC-conjugated goat anti-rabbit IgG (MP Biomedicals, CA, USA), and FITC-conjugated goat anti-mouse IgG+IgM+IgA (Rockland Immunochemicals) and IgG+IgM (Jackson Immuno Research, PA, USA) were purchased. Two Alexa Fluor™ 594-conjugated goat anti-mouse and two FITC-conjugated goat anti-mouse antibodies were mixed 1:1, respectively, and were used.

Plaque assay

HRT18 cells were seeded onto 6-well plates with 1.0×10^6 cells/well. The next day, the cells were inoculated with 400 µL of diluted virus in order to form 10–20 plaques per well and incubated at 37°C for 1.5 h. Unabsorbed virus were removed and cells were overlaid with MEM (Sigma-Aldrich) containing 0.65%

408   agarose (Nacalai tesque, Kyoto, Japan), 2nM L-glutamine (Sanko Junyaku Co., Tokyo, Japan), 10mM HEPES

409   (Dojindo, Kumamoto, Japan) and PS, and these were incubated at 37°C until the appropriate sized plaque

410   were observed (approx. three days). Cells were fixed with 10% formaldehyde (Fujifilm Wako) and stained

411   with crystal violet. For plaque purification of virus, without cell fixation, a well-isolated plaque was picked

412   and resuspended in 100 µL DMEM(-) and stored at -70°C. Then, a 50 µL aliquot from the 100 µL plaque

413   solution was expanded in HRT18 cells with DMEM(-) in a 10 cm dish, and then these viruses were harvested

414   after the appropriate CPE was observed and was stored at -70°C as a master stock. 200 µL of master virus

415   was expanded in the cells in an 18 cm dish and was used as a working wtBToV.

416

417   **RNA extraction, RT-PCR and sequencing**

418   Viral RNA was extracted from 100–250 µL of the supernatant using TRIzol® LS Reagent (Thermo Fisher

419   Scientific) according to the manufacturer's instructions. Complementary DNA was synthesized using a

420   SuperScript III® first-strand synthesis system (Thermo Fisher Scientific) and Oligo dT_{15} primer (TaKaRa,

421   Shiga, Japan) or Oligo dT-3 sites Adaptor Primer for 3' RACE (Takara). To analyze the complete BToV genome,

422   seven fragments of about 3500–4500 base pair (bp) overlapping 170–360 bp (which correspond to BToV-

423   derived sequence of fragments BToV-A to -G described below) and one fragment of about 320 bp of the 3'-

424   terminal region by 3' RACE were amplified using Phusion High-Fidelity PCR Kit (New England Biolabs, MA,

425   USA) under the following conditions [98°C, 50 s; 35 cycles of (98°C, 10 s; 60–65°C, 30 s; 72°C, 1 min/kb); 72°C

426   10 min], and these PCR fragments were purified with AMPure XP (Beckman Coulter, CA, USA). 5' ends of

427   the viral genome were analyzed by 5'-Full RACE Core Set (TaKaRa) and the PCR products of 5' RACE were

428   cloned using pGEM®-T Easy Vector Systems (Promega, WI, USA). Purified PCR fragments and pGEM

429   cloning 5'-terminal sequence of BToV were sequenced using primers designed for every ~300 bp of the BToV

430   genome. Sequences were assembled by MEGA6 (71) or Sequencher software (Hitachi High-Tech, Tokyo,

431   Japan). Even without 5' RACE, almost the entire full-length sequencing was achieved, except for 21 bp at the

432   5' terminal end derived from the primer, which was applied to the full-length sequencing of rBToV and rEGFP

433   c1.

434

435   **BAC construction of wtBToV**

436 pBeloBAC11 containing CMV immediate-early promoter, Rz, and BGH termination (48) was used. Full-

437 length cDNA of the BToV genome was assembled under the control of the CMV promoter and was flanked at

438 the 3'-terminal end by a poly(A) tail (Fig. 1A), using a Red/ET recombination system counterselection BAC

439 modification kit (Gene Bridges Heidelberg, Germany) according to the manufacturer's instructions (Fig. 1B).

440 Briefly, *E. coli* strain NEB10ß (Bio-Rad) carrying the BAC was electroporated with the pRed/ET expression

441 plasmid using an Eppendorf Eporator® (Eppendorf, Hamburg, Germany). The Red/ET recombinant enzyme

442 was induced by L-arabinose (Sigma-Aldrich) and a temperature shift from 30°C to 37°C, and then linear

443 rpsL/Neo counter-selection/selection cassette (prepared by PCR) flanked by hms was electroporated so that

444 the Red/ET recombination inserts the rpsL/Neo cassette into the target position through hms. Only *E. coli*

445 carrying the BAC with the rpsL/Neo cassette could be selected by kanamycin at 30°C. The Red/ET enzyme

446 was induced in the same way in *E. coli* carrying modified BAC and was electroporated with the linear cDNA

447 fragment of BToV flanked by hms. Red/ET recombination could replace the rpsL/Neo cassette with the cDNA

448 fragment of BToV. Only *E. coli* carrying the BAC with the target BToV cDNA fragment could be selected by

449 streptomycin at 37°C at which the pRed/ET plasmid was removed (Fig. 1B). To use this system, the full-

450 length genomic sequence of BToV was divided into eight fragments flanked by the hms and was assembled

451 into the BAC plasmid in a sequential order (Fig. 1A). To distinguish the base numbers between BToV and

452 BAC, BToV genome is 28,313 nt (without poly-A and was written as the normal number) while that of BAC

453 is 8,219 bp, written as the number in parentheses starting from the CMV promoter. The BToV-derived

454 sequences of each fragment correspond to nt 1 to 4,497 (BToV-A), nt 4,139 to 8,512 (BToV-B), nt 8,301 to

455 12,748 (BToV-C), nt 12,413 to 16,612 (BToV-D), nt 16,444 to 20,731(BToV-E), nt 20,512 to 24,877 (BToV-F), nt

456 24,675 to 28,196 (BToV-G), and nt 27,987 to 28,313 (BToV-H), respectively. The 5' terminal hm of the BToV-A

457 fragment and the other fragments were regions of a partial CMV promoter, (205) to (605), and regions

458 overlapping between the fragments, respectively. The 3' terminal hm of all fragments is the region including

459 Rz and BGH, which is (606) to (1,001) (Fig. 1A). Fragments BToV-A to -G were prepared using standard PCR

460 protocols with the overlap extension technique (72). Fragment BToV-H was prepared through PCR using a

461 plasmid template containing a chemically synthesized BToV sequence between nt 27,987 to 28,313, 25 nt of

462 adenine (pA), and the hm of the 3' terminal sequence described above (Eurofins, Luxembourg, Luxembourg).

463 All fragments were amplified using KOD Plus Neo (TOYOBO, Osaka, Japan) and purified with gel-extraction

464 and phenol-chloroform extraction before electroporation. After assembly of the eight fragments, BAC

465 carrying the cDNA clone of the full-length BToV (pBAC-BToV^{mut1}) contained three mutations, namely, two

466 single mutations at nt G3399T (E871D in NSP1 protein) and at nt T8469G (silent), and the *E. coli*

467 chromosomal derived sequence of 1,350 bp was inserted at nt 10,136 (Fig. 2A). pRed/ET recombination

468 method could revert two substitutions but failed to delete the *E. coli* chromosomal derived sequence. To delete

469 this sequence in vitro, type IIS restriction enzyme SapI (New England Biolabs), which recognizes asymmetric

470 DNA sequences and cleave outsides their 5'-GCTCTTCN ↓ NNN -3' recognition site, was used. First, two SapI

471 sites at nt (4,526) and (5,736) near or in the SopA gene of BAC plasmid were removed by introducing single

472 mutations, 5'-(4,526) GCTCTTC→CCTCTTC and 5'-(5,736) GCTCTTC→GGTCTTC respectively, using

473 NEBuilder HiFi DNA Assembly Master Mix (New England Biolabs) and two NgoMVI restriction sites located

474 outside the two SapI sites at nt (3,633) and (7,549). Next, SapI recognition sites were inserted on both sides

475 (one in the reverse direction) of the *E. coli* chromosomal derived sequence by pRed/ET recombination. Finally,

476 the resulting BAC plasmid (pBAC-BToV^{mut2}) was treated with SapI and was subjected to in vitro self-ligation

477 (pBAC-BToV^{pre}). pBAC-BToV^{pre} was electroporated into NEB10ß and was successfully propagated.

478

479 **BAC construction of mutant BToVs**

480 To introduce mutations into the HE gene of pBAC-BToV^{pre}, the rpsL/Neo cassette was inserted through the

481 5' terminal hm including a partial M gene, nt 26,037 to 26,249 and the 3' terminal hm including 3'-UTR,

482 poly(A), Rz, and BGH, 28,150 to (1,001), using pRed/ET recombination (Fig. 2B). The resulting BAC with the

483 rpsL/Neo cassette (pBAC-BToV$^{HE/N-rpsL/Neo}$) allows manipulation of not only the HE gene but also the N gene.

484 The rpsL/Neo cassette was replaced with the mutated HE gene or with the Enhanced Green Fluorescent

485 Protein (EGFP) gene (Fig. 2B).

486 Cell-adapted wtBToV in our laboratory lost full-length HE (HEf) due to a stop codon at nt 481 (the base

487 number of HE gene), CAG (Q)→TAG (stop), resulting in the soluble form HE 160 aa in length (HEs) (36, 37)

488 (Fig. 3A). Using pBAC-BToV$^{HE/N-rpsL/Neo}$, four pBAC plasmids were created; pBAC-BToV with two genetic

489 markers at nt T207C (T26530C) and nt T228A (T26554A), pBAC-BToV-HEf with the replacement of a stop

490 codon (TAG) to CAG to express full-length HE, pBAC-BToV-HEs/HA and pBAC-BToV-HEf/HA with the

491 addition of HA-tag at the C-terminal end of HEs and HEf proteins, respectively. pBAC-BToV-HEf/HA had the

492 HE gene mutation at nt 1,029, GAATTC→GAACTC (silent) to remove the EcoRI site. In addition, pBAC-

493 BToV-EGFP in which the HE gene was completely replaced with EGFP was created. All BAC plasmids were

494 purified with NucleoBond Xtra Maxi (Macherey-Nagel, Dueren, Germany) from 1,000mL cultured *E. coli*

495 cells.

496

497 **Rescue of recombinant BToVs**

498 293T and HRT18 cells were seeded into 24-well plates at 1.5×10^5 and 2.0×10^5 cells/well, respectively, on

499 the day before the experiment. 293T cells were transfected with 2.0 µg pBAC-BToV or -HE mutant and -

500 EGFP with 1.5 µL Lipofectamine 3000 and 1.0 µL P3000 Regent (Thermo Fisher Scientific) and incubated at

501 37°C. The next day, DMEM(+) was replaced with 500 µL DMEM(-) to remove FBS. After 3 dpt, supernatants

502 were harvested. The 50 µL supernatant was inoculated to susceptible HRT18 cells which had been washed

503 with phosphate buffered saline (PBS) twice after which 450 µL DMEM (-) was added. After 2 dpi,

504 supernatants were harvested and viruses were plaque-purified three times. A 20 µL aliquot from the 100 µL

505 plaque suspension was added to HRT18 cells with 480 µL DMEM(-) in 24-well plates, and the supernatant

506 was harvested after 2 dpi. This was stored at -70°C as a master virus. Next, 100 µL of the master viruses

507 were expanded on HRT18 cells with 10 mL DMEM(-) in a 10 cm dish, and these viruses were harvested after

508 the appropriate CPE was observed. These were stored as passage 0 (P0) viruses, which were used for

509 subsequent experiments. In the case of rEGFP, plaque purification of some recombinant viruses (No.3 to

510 No.6) were omitted because of EGFP gene instability. 293T cells prepared on a 3 cm dish were transfected

511 with 5 times reagent volumes and pBAC-BToV-EGFP. The 2.5 mL supernatants of transfected HRT18 cells

512 were inoculated into the cells in a 10 cm dish to expand the viruses, and these viruses were then harvested.

513 In this case, these were used as passage 0 (P0) viruses. The nomenclature of recombinant BToV rescued from

514 these BAC plasmids was rBToV(rHEs), rHEf, rHEs/HA, rHEf/HA and rEGFP, respectively (Fig. 3A).

515

516 **Immunofluorescence staining**

517 A 35 mm glass base dish (12 mm glass area) (Iwaki, Shizuoka, Japan) was seeded with 1.0×10^5 HRT18 cells,

518 the day before the experiment. Cells were washed twice with PBS and infected with 300 µL recombinant

519 BToV in DMEM(-) with a MOI of 0.05. After 24 hpi (and 36 hpi only for the rEGFP), infected cells were

520 washed with PBS, fixed with 4% paraformaldehyde for 20 min at room temperature (rt), and were

521   permeabilized with 0.2% Triton X-100 for 20 min at RT. Cells were incubated with mouse anti-HE, anti-M

522   antiserum, or rabbit and mouse anti-HA antibodies at 4°C overnight, and were then incubated with a mixed

523   FITC-conjugated goat anti-mouse or mixed AlexaFluor™ 594-conjugated goat anti-mouse for mouse

524   antiserum and FITC-conjugated goat anti-rabbit antibodies for rabbit antibody, for 2 h at rt. Cell nuclei were

525   stained with Hoechst 33258 solution (Dojindo) for 20 min at rt. PBS washing was done twice between each

526   step. Stained cells were observed under confocal laser scanning microscopy using an LSM 710 laser scanning

527   microscope (Carl Zeiss, Oberkochen, Germany).

528

529   α-naphthyl acetate esterase activity

530   HRT18 cells grown in 24-well plates were infected with recombinant BToVs with a MOI of 0.05. At 24 hpi,

531   cells were fixed with fixative solution, stained for a α-naphthyl acetate esterase with a α-NA esterase staining

532   kit (Muto Pure Chemicals, Tokyo, Japan) according to the manufacturer's instructions, and observed under

533   phase-contrast microscopy.

534

535   Immunoblotting

536   HRT18 cells grown in 24-well plates were infected with the recombinant BToVs with a MOI of 0.05. COS7

537   cells were transfected with 1.0 µg pCAGGS-HEf or pCAGGS-HEs with 1.5 µL Lipofectamine 3000 and 2.0 µl

538   P3000 Regent. At 24 hpi (and 36 hpi only for rEGFP), infected HRT18 cells and 40 hpt transfected COS7 cells

539   were lysed with 100 µL sample buffer (50 mM Tris, 2% SDS, 0.1% bromophenol blue, 10% glycerol, and 1%

540   2-mercaptoethanol) and were boiled for 5 min. N protein, EGFP, and HA-tagged proteins of samples were

541   analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and were transferred to

542   a PVDF membrane (Merck Millipore, Tokyo, Japan). The membrane was incubated with rabbit anti-N

543   antiserum or mouse anti-GFP antibody, and then with HRP-conjugated anti-rabbit and anti-mouse

544   antibodies, respectively. To detect HA-tagged proteins, mouse-anti-HA 12CA5 antibody for HA-tagged HEf

545   and rabbit anti-HA antibody for HA-tagged HEs were used because of differences in their reactivity. Protein

546   bands were visualized with ECL Prime Western Blotting Detection Reagent (GE Healthcare, IL, USA) on a

547   Light Capture II instrument (ATTO, Tokyo, Japan).

548

549 ## Growth kinetics of recombinant BToVs

550 HRT18 cells in 24-well plates were washed twice with PBS and were infected with recombinant BToVs with

551 a MOI of 0.001. After 1.5 h adsorption at 37°C, unadsorbed viruses were removed and cells were washed

552 twice with PBS, and 0.5 mL DMEM(-) was added. Viruses in the supernatant were collected at 2, 24, 48, and

553 72 hpi and the supernatant was clarified by centrifugation at 12,000 rpm for 5 min at 4°C. Virus titers in the

554 culture media were determined in a 96-well plate by $TCID_{50}$, as previously described (73).

555

556 ## Measurement of the stability of HE and EGFP gene

557 HRT18 cells in a 24-well plate were washed twice with PBS after which 0.5 mL DMEM(-) was added; these

558 were then infected with recombinant BToVs with a MOI of 0.001. After the appropriate CPE was observed (2

559 to 3 dpi), the supernatant was harvested and 0.5 uL was inoculated to the fresh cells, resulting in a serial

560 passages by 1,000-fold diluted virus. To estimate stability of the HEf and EGFP genes during serial passages,

561 HRT18 cells in a 24-well plate were infected with rHEf and rHEf/HA harvested at the indicated passage

562 history with a MOI of 0.05, and cells at 24 hpi were subjected to an α-NA esterase assay or to immunoblotting

563 of N proteins. EGFP-expression of the infected cells with rEGFP were monitored during serial passages under

564 a fluorescence microscope (Olympus CKX41 U-RFLT50; Olympus, Tokyo, Japan). Later, images of cells

565 infected with rEGFP harvested at the indicated passage history at the same conditions were taken, and

566 EGFP and N proteins were detected by immunoblotting. Supernatants of the infected cells (resulting in +1

567 passage history) were subjected to RNA extraction, RT-PCR, and sequencing described above. PCR fragments

568 of 1130 and 1630 bp containing HE and EGFP genes, respectively, were amplified using Takara ExTaq

569 (Takara) under the following conditions [94°C, 5min; 30 cycles of (94°C, 30 s; 50°C, 30 s; 72°C, 1 min/kb);

570 72°C 10min]. DNase treatment was also performed before RNA extraction to digest BAC plasmid from

571 transfection, depending on passage history.

572

573   **Figure legends**

574   **Fig. 1.** Assembly of a BToV full-length cDNA into a BAC and flowchart of Red/ET recombination. (A) The

575   genome sequence of BToV is divided into eight fragments (BToV-A to H), and each fragment contained

576   homology arms (hm) overlapping each fragment at the 3' and 5' terminal ends (vertical dashed line). To

577   distinguish the base number between BToV genome and BAC, the BToV genome of 28,313 nt without poly(A)

578   (25 nt) was written as the normal number, while BAC of 8,219 bp was written as the number in parentheses

579   starting with the CMV promoter. (B) (i) *E. coli* carrying the BAC was electroporated with the pRed/ET

580   plasmid, and the Red/ET recombinant enzyme was induced by L-arabinose. (ii) The linear rpsL/Neo counter-

581   selection/selection cassette flanked by hms was electroporated, and the Red/ET enzyme inserted the rpsL/Neo

582   cassette into the target position via hms. After selecting appropriate antibodies, (iii) the Red/ET enzyme in

583   *E. coli* carrying the modified BAC could replace the rpsL/Neo cassette with the cDNA fragment of BToV in

584   the same way. Only *E. coli* carrying BAC with the target cDNA fragment could be selected by antibodies and

585   37°C at which the pRed/ET plasmid was removed.

586

587   **Fig. 2.** Schematic diagram of pBAC-BToV construction strategy. (A) (i) BAC-BToV^{mut1} contained two

588   mutations at nt 3,399 (E→D) and 8,469 (silent), *E. coli* chromosomal insertion at 10,136 and two SapI sites

589   at nt (4,526) and (5,736) on BAC. Two mutations were reverted by the pRed/ET recombination method. To

590   remove the *E. coli*-derived sequence, two SapI sites on BAC were removed by single nt substitution of the

591   recognition site (which is shown below) using in vitro recombination, and (ii) the SapI recognition sites (one

592   is in reverse direction) were inserted on both sides of the E.coli derived sequence by pRed/ET recombination.

593   (iii) Finally, the resulting BAC plasmid (pBAC-BToV^{mut2}) was treated with SapI and subjected to in vitro self-

594   ligation (pBAC-BToV^{pre}). (B) To introduce genetic makers of pBAC-BToV^{pre}, rpsL/Neo cassette was inserted

595   between nt 26,249 and 28,150 containing partial M, HE, and N genes via hms (pBAC-BToV$^{HE/N-rpsL/Neo}$). The

596   rpsL/Neo cassette was replaced with HE carrying the genetic makers at nt 26,530 and 26,554, with the

597   surrounding sequence via the hms (pBAC-BToV).

598

599   **Fig. 3.** Schematic diagram and generation of recombinant BToVs (A) Cell-adapted wtBToV lost full-length

600   HE (HEf, 419aa) due to a stop codon at nt 481 (the base number of HE gene), resulting in the soluble form

601 HE (HEs, 160aa). rHEf/HA and rHEs/HA were added to the HA-tag at the C-terminal end of HEs and HEf

602 proteins, respectively. rEGFP carries a reporter EGFP gene in which HE gene was completely replaced. nt:

603 nucleotide, aa: amino acid. (B) CPE of HRT18 cells infected with wtBToV and rBToV. HRT18 cells infected

604 with a MOI of 0.05 were observed at 24 hpi under phase-contrast microscopy. (C) Sequence analysis of

605 wtBToV and rBToV at genetic maker sites (underlined) in HE gene.

606

607 **Fig.4.** Growth and characteristics of recombinant HE mutant BToVs. (A) Representative of plaque

608 morphology in the 6-well plate of recombinant BToVs at 3 dpi. (B) (Upper) Detection of the membrane (M),

609 HE, and HA-tagged proteins of recombinant BToVs by indirect immunofluorescence. HRT-18 cells were

610 infected with these viruses with a MOI of 0.05 and infected cells were fixed, permeabilized at 24 hpi, and

611 then stained with mouse anti-M and anti-HE antiserum or rabbit anti-HA antibodies and FITC-conjugated

612 secondary antibodies (green). The nuclei were stained with Hoechst (blue). Staining cells were observed

613 under confocal laser scanning microscopy. (Lower) α-NA esterase activity of the cells infected with

614 recombinant BToVs. Fixed infected cells described above were tested for α-NA esterase activity. (C)

615 Immunoblotting of HA-tagged and N proteins in HRT18 cells infected with recombinant BToVs and COS7

616 cells transfected with plasmid encoding HEs/HA and HEf/HA. Infected HRT18 cells at a MOI of 0.05 at 24

617 hpi and transfected COS7 cells at 40 hpt were lysed, and the lysates were subjected to immunoblotting.

618 Proteins were detected with rabbit anti-N antiserum and mouse-anti-HA 12CA5 antibody for HEf/HA and

619 rabbit anti-HA antibody for HEs/HA. (D) Growth kinetics of recombinant BToVs. HRT-18 cells were infected

620 with a MOI of 0.001. At the indicated time, supernatants of infected cells were harvested, and virus titers

621 were determined by $TCID_{50}$. The mean and standard deviation of three independent experiments are shown

622 (N = 3). Dashed line shows detection limit. N.D.: not detected.

623

624 **Fig.5.** Full-length HE gene stability of recombinant HE mutant BToVs. (A) α-NA esterase activity of infected

625 cells with recombinant BToVs carrying full-length HE gene during serial passages. HRT-18 cells in 24-well

626 plate infected with a MOI of 0.05 were fixed at 24 hpi, and then tested for α-NA esterase activity. Two clones

627 (No.1 and No.2) derived from different plaques of each virus were used. (B) Immunoblotting of N proteins in

628 HRT18 cells infected with recombinant BToVs described above at the indicated passage history. (C) RT-PCR

629   of the HE gene of recombinant BToVs. The supernatant from infected cells (resulting in +1 passage history)

630   were amplified by RT-PCR using primers flanking the HE gene (shown in figure D, 1670 bp). Amplicons were

631   analyzed by agarose gel electrophoresis. (D) Sequencing analysis results summary of recombinant BToVs HE

632   genes. Viruses at P19 were plaque-purified once and six well-isolated plaques were picked, and then these

633   viruses were inoculated into fresh HRT18 cells. The HE gene of each supernatant was amplified and

634   sequenced. 5/6 means that 5 out of 6 plaques have the HE proteins described in figure.

635

636   **Fig.6.** Growth and characteristics of recombinant BToV carrying EGFP gene. (A) Representative of plaque

637   morphology of rEGFP at 3 dpi. In these rEGFP, 293T cells in 3 cm dish were transfected with BAC-BToV-

638   EGFP, and the supernatants harvested after 2 dpt were inoculated into fresh HRT18 cells in a 10 cm dish

639   without plaque purification. This was stored as passage 0 (P0) virus, and the viruses obtained from each

640   independent experiment were numbered (No.3 to No.6) (B) Detection of EGFP-expression (green) in infected

641   cells with rEGFP under a fluorescence microscope. HRT18 cells were infected with rEGFP No.3 with a MOI

642   of 0.05, and infected cells were stained with Hoechst (Blue) at 36 hpi. Living infected cells were observed

643   under a fluorescence microscope (left). Infected cells were fixed at 36 hpi, permeabilized, stained with mouse

644   anti-M antiserum and AlexaFluor™ 594-conjugated secondary antibody (red), and observed by confocal laser

645   scanning microscopy (right). Cells infected with rBToV, wtBToV were observed or fixed at 24 hpi. (C)

646   Immunoblotting of EGFP and N proteins in infected cells with rEGFP. Infected cells described above were

647   lysed at 36 hpi. Cells infected with rBToV and wtBToV were lysed at 24 hpi. Proteins in the lysates were

648   detected with rabbit anti-N antiserum and mouse anti-GFP antibody. (D) Growth kinetics of rEGFP No.3.

649   Virus titers were determined as described in Fig. 4. Each result of four independent experiments of rEGFP

650   No.3 is shown. Dashed line shows detection limit. N.D.: not detected.

651

652   **Fig.7.** EGFP gene stability of rEGFP. (A) EGFP-expression of rEGFP during serial passages. HRT-18 cells

653   were infected with a MOI of 0.001, and then 1,000-fold diluted virus was inoculated into fresh cells. Viruses

654   were then harvested after 3 dpi. Two rEGFP (No.1 and No.2) were plaque-purified and cloned, while four

655   rEGFP (No.3 to No.6) described in Fig. 6 were not. Cells were infected with 1,000-fold diluted rEGFP

656   harvested at the indicated passage history and observed at 36 hpi under a fluorescence microscope. In P0

657 viruses, infected cells with a MOI of 0.05 (actual experiment was 0.001) are shown to easily observe EGFP-

658 expression. (B) RT-PCR of the EGFP gene of rEGFP. The supernatant viruses from infected cells (+1 passage

659 history) were amplified by RT-PCR using primers flanking the EGFP gene (Figure D, 1130bp). Amplicons

660 were analyzed by agarose gel electrophoresis. (C) Immunoblotting of EGFP and N proteins in cells infected

661 with rEGFP No.5 at indicated passage history. Cells were infected with 1,000-fold diluted rEGFP No.5 from

662 P1 and infected with P0 viruses (rEGFP No5, rBToV, wtBToV) at a MOI of 0.05. Infected cells with rBToV

663 and wtBToV and with rEGFP No.5 were lysed at 24 hpi and at 36 hpi, respectively. (D) Summary of EGFP

664 gene sequencing analysis of rEGFP. Major bands of EGFP gene of each supernatant were purified and

665 sequenced.

666

667 **Fig.8.** Improvement of the growth kinetics and EGFP-expression of rEGFP. (A) Representative of plaque

668 morphology of variant rEGFP. The P5 rEGFP No.5 was plaque-purified three times, and three clones were

669 expanded (designated c1 to c3). (B, C) Growth kinetics and immunoblotting of rEGFP (c1 to c3). Virus titers

670 (N=4) were determined and immunoblotting was performed as described in Fig. 6. Dash line shows detection

671 limit. N.D.: not detected. The cells were infected with a MOI of 0.05, and lysed at 24 hpi for variant rEGFP

672 rBToV, wtBToV and at 36 hpi for P0 rEGFP No.5. (D) Full-length genome sequence of rEGFP c1. Mutations

673 different from the pBAC-BToV-EGFP are shown. (E) Genetic stability of EGFP gene of rEGFP c1 and c2.

674 Analysis of EGFP-expression and EGFP gene of these viruses during serial passages are described in Fig.7

675

676

Acknowledgments

References

685

686   1.   ICTV (International Committee on Taxonomy of Virus). 2019. Virus Taxonomy: 2019 Release.

687   2.   Vanopdenbosch E, Wellemans G, Petroff K. 1991. Breda virus associated with respiratory disease in

688   calves. Vet Rec 129:203.

689   3.   Duckmanton L, Luan B, Devenish J, Tellier R, Petric M. 1997. Characterization of torovirus from human

690   fecal specimens. Virology 239:158–68.

691   4.   Kroneman A, Cornelissen LA, Horzinek MC, de Groot RJ, Egberink HF. 1998. Identification and

692   characterization of a porcine torovirus. J Virol 72:3507–11.

693   5.   Woode GN, Reed DE, Runnels PL, Herrig MA, Hill HT. 1982. Studies with an unclassified virus isolated

694   from diarrheic calves. Vet Microbiol 7:221–40.

695   6.   Weiss M, Steck F, Horzinek MC. 1983. Purification and Partial Characterization of a New Enveloped

696   RNA Virus (Berne Virus). J Gen Virol 64:1849–1858.

697   7.   Ito T, Okada N, Okawa M, Fukuyama S, Shimizu M. 2009. Detection and characterization of bovine

698   torovirus from the respiratory tract in Japanese cattle. Vet Microbiol 136:366–371.

699   8.   SH L, HY K, EW C, D K. 2019. Causative Agents and Epidemiology of Diarrhea in Korean Native Calves.

700   J Vet Sci 20.

701   9.   Nogueira JS, Asano KM, de Souza SP, Brandão PE, Richtzenhain LJ. 2013. First detection and molecular

702   diversity of Brazilian bovine torovirus (BToV) strains from young and adult cattle. Res Vet Sci 95:799–

703   801.

704   10. Lojkić I, Krešić N, Šimić I, Bedeković T. 2015. Detection and molecular characterisation of bovine corona

705   and toroviruses from Croatian cattle. BMC Vet Res 11:202.

706   11. Koopmans M, van Wuijckhuise-Sjouke L, Schukken YH, Cremers H, Horzinek MC. 1991. Association of

707   diarrhea in cattle with torovirus infections on farms. Am J Vet Res 52:1769–73.

708   12. Hoet AE, Saif LJ. 2004. Bovine torovirus (Breda virus) revisited. Anim Heal Res Rev 5:157–71.

709   13. Hoet AE, Nielsen PR, Hasoksuz M, Thomas C, Wittum TE, Saif LJ. 2003. Detection of Bovine Torovirus

710   and other Enteric Pathogens in Feces from Diarrhea Cases in Cattle. J Vet Diagnostic Investig 15:205–

711   212.

712   14. Gülaçtı İ, Işıdan H, Sözdutmaz İ. 2014. Detection of bovine torovirus in fecal specimens from calves with

713   diarrhea in Turkey. Arch Virol 159:1623–1627.

714   15.   Duckmanton L, Carman S, Nagy E, Petric M. 1998. Detection of bovine torovirus in fecal specimens of

715   calves with diarrhea from Ontario farms. J Clin Microbiol 36:1266–70.

716   16.   H L, B Z, H Y, C T. 2020. First Detection and Genomic Characteristics of Bovine Torovirus in Dairy

717   Calves in China. Arch Virol 165.

718   17.   Zhou L, Wei H, Zhou Y, Xu Z, Zhu L, Horne J. 2014. Molecular epidemiology of Porcine torovirus (PToV)

719   in Sichuan Province, China: 2011–2013. Virol J 11:106.

720   18.   Y F, Y K, F S, H A, R I, K S, Y K, T O, M O, T F, S T, Y O, J S, T M, T O, M N. 2020. Complete Genome

721   Sequencing and Genetic Analysis of a Japanese Porcine Torovirus Strain Detected in Swine Feces. Arch

722   Virol 165.

723   19.   Shin D-J, Park S-I, Jeong Y-J, Hosmillo M, Kim H-H, Kim H-J, Kwon H-J, Kang M-I, Park S-J, Cho K-O.

724   2010. Detection and molecular characterization of porcine toroviruses in Korea. Arch Virol 155:417–422.

725   20.   Pignatelli J, Grau-Roma L, Jiménez M, Segalés J, Rodríguez D. 2010. Longitudinal serological and

726   virological study on porcine torovirus (PToV) in piglets from Spanish farms. Vet Microbiol 146:260–268.

727   21.   Anbalagan S, Peterson J, Wassman B, Elston J, Schwartz K. 2014. Genome sequence of torovirus

728   identified from a pig with porcine epidemic diarrhea virus from the United States. Genome Announc 2.

729   22.   ZM H, YL Y, LD X, B W, P Q, YW H. 2019. Porcine Torovirus (PToV)-A Brief Review of Etiology,

730   Diagnostic Assays and Current Epidemiology. Front Vet Sci 6.

731   23.   SL S, EJ S, RJ   de G. 2006. Characterization of a Torovirus Main Proteinase. J Virol 80.

732   24.   Snijder EJ, Horzinek MC. 1993. Toroviruses: replication, evolution and comparison with other members

733   of the coronavirus-like superfamily. J Gen Virol 74:2305–2316.

734   25.   Draker R, Roper RL, Petric M, Tellier R. 2006. The complete sequence of the bovine torovirus genome.

735   Virus Res 115:56–68.

736   26.   Snijder EJ, Den Boon JA, Spaan WJM, Verjans GMGM, Horzinek MC. 1989. Identification and Primary

737   Structure of the Gene Encoding the Berne Virus Nucleocapsid Protein. J Gen Virol 70:3363–3370.

738   27.   Horzinek MC, Ederveen J, Weiss M. 1985. The Nucleocapsid of Berne Virus. J Gen Virol 66:1287–1296.

739   28.   Den Boon JA, Snijder EJ, Locker JK, Horzinek MC, Rottier PJ. 1991. Another triple-spanning envelope

740   protein among intracellularly budding RNA viruses: the torovirus E protein. Virology 182:655–63.

741 29. Snijder EJ, Den Boon JA, Spaan WJ, Weiss M, Horzinek MC. 1990. Primary structure and post-

742   translational processing of the Berne virus peplomer protein. Virology 178:355–63.

743 30. Cornelissen LA, Wierda CM, van der Meer FJ, Herrewegh AA, Horzinek MC, Egberink HF, de Groot RJ.

744   1997. Hemagglutinin-esterase, a novel structural protein of torovirus. J Virol 71:5277–86.

745 31. Smits SL, Lavazza A, Matiz K, Horzinek MC, Koopmans MP, de Groot RJ. 2003. Phylogenetic and

746   evolutionary relationships among torovirus field variants: evidence for multiple intertypic recombination

747   events. J Virol 77:9567–77.

748 32. Cong Y, Zarlenga DS, Richt JA, Wang X, Wang Y, Suo S, Wang J, Ren Y, Ren X. 2013. Evolution and

749   homologous recombination of the hemagglutinin–esterase gene sequences from porcine torovirus. Virus

750   Genes 47:66–74.

751 33. Ito M, Tsuchiaka S, Naoi Y, Otomaru K, Sato M, Masuda T, Haga K, Oka T, Yamasato H, Omatsu T,

752   Sugimura S, Aoki H, Furuya T, Katayama Y, Oba M, Shirai J, Katayama K, Mizutani T, Nagai M. 2016.

753   Whole genome analysis of Japanese bovine toroviruses reveals natural recombination between porcine

754   and bovine toroviruses. Infect Genet Evol 38:90–95.

755 34. Aita T, Kuwabara M, Murayama K, Sasagawa Y, Yabe S, Higuchi R, Tamura T, Miyazaki A, Tsunemitsu

756   H. 2012. Characterization of epidemic diarrhea outbreaks associated with bovine torovirus in adult cows.

757   Arch Virol 157:423–431.

758 35. Ito T, Okada N, Fukuyama S. 2007. Epidemiological analysis of bovine torovirus in Japan. Virus Res

759   126:32–37.

760 36. Kuwabara M, Wada K, Maeda Y, Miyazaki A, Tsunemitsu H. 2007. First Isolation of Cytopathogenic

761   Bovine Torovirus in Cell Culture from a Calf with Diarrhea. Clin Vaccine Immunol 14:998–1004.

762 37. Shimabukuro K, Ujike M, Ito T, Tsunemitsu H, Oshitani H, Taguchi F. 2013. Hemagglutination mediated

763   by the spike protein of cell-adapted bovine torovirus. Arch Virol 158:1561.

764 38. Weiss M, Horzinek MC. 1986. Morphogenesis of Berne Virus (Proposed Family Toroviridae). J Gen Virol

765   67:1305–1314.

766 39. Horzinek MC, Weiss M, Ederveen J. 1984. Berne Virus is Not "Coronavirus-like." J Gen Virol 65:645–649.

767 40. Fagerland JA, Pohlenz JFL, Woode GN. 1986. A Morphological Study of the Replication of Breda Virus

768   (Proposed Family Toroviridae) in Bovine Intestinal Cells. J Gen Virol 67:1293–1304.

769   41.   Corman VM, Muth D, Niemeyer D, Drosten C. 2018. Hosts and Sources of Endemic Human

770   Coronaviruses. Adv Virus Res 100:163–188.

771   42.   Drosten C, Gunther S, Preiser W, van der Werf S, Brodt HR, Becker S, Rabenau H, Panning M,

772   Kolesnikova L, Fouchier RA, Berger A, Burguiere AM, Cinatl J, Eickmann M, Escriou N, Grywna K,

773   Kramme S, Manuguerra JC, Muller S, Rickerts V, Sturmer M, Vieth S, Klenk HD, Osterhaus AD,

774   Schmitz H, Doerr HW. 2003. Identification of a novel coronavirus in patients with severe acute

775   respiratory syndrome. N Engl J Med2003/04/12. 348:1967–1976.

776   43.   Kuiken T, Fouchier RA, Schutten M, Rimmelzwaan GF, van Amerongen G, van Riel D, Laman JD, de

777   Jong T, van Doornum G, Lim W, Ling AE, Chan PK, Tam JS, Zambon MC, Gopal R, Drosten C, van der

778   Werf S, Escriou N, Manuguerra JC, Stohr K, Peiris JS, Osterhaus AD. 2003. Newly discovered

779   coronavirus as the primary cause of severe acute respiratory syndrome. Lancet2003/08/02. 362:263–270.

780   44.   Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus AD, Fouchier RA. 2012. Isolation of a novel

781   coronavirus from a man with pneumonia in Saudi Arabia. N Engl J Med2012/10/19. 367:1814–1820.

782   45.   Zhou P, Yang X-L, Wang X-G, Hu B, Zhang L, Zhang W, Si H-R, Zhu Y, Li B, Huang C-L, Chen H-D,

783   Chen J, Luo Y, Guo H, Jiang R-D, Liu M-Q, Chen Y, Shen X-R, Wang X, Zheng X-S, Zhao K, Chen Q-J,

784   Deng F, Liu L-L, Yan B, Zhan F-X, Wang Y-Y, Xiao G-F, Shi Z-L. 2020. A pneumonia outbreak associated

785   with a new coronavirus of probable bat origin. Nature 579:270–273.

786   46.   R L, X Z, J L, P N, B Y, H W, W W, H S, B H, N Z, Y B, X M, F Z, L W, T H, H Z, Z H, W Z, L Z, J C, Y M,

787   J W, Y L, J Y, Z X, J M, WJ L, D W, W X, EC H, GF G, G W, W C, W S, W T. 2020. Genomic

788   Characterisation and Epidemiology of 2019 Novel Coronavirus: Implications for Virus Origins and

789   Receptor Binding. Lancet (London, England) 395.

790   47.   Yount B, Curtis KM, Baric RS. 2000. Strategy for systematic assembly of large RNA and DNA genomes:

791   transmissible gastroenteritis virus model. J Virol 74:10600–11.

792   48.   F A, JM G, Z P, A I, E C, J P-D, L E. 2000. Engineering the Largest RNA Virus Genome as an Infectious

793   Bacterial Artificial Chromosome. Proc Natl Acad Sci U S A 97:5516–21.

794   49.   G T, R H-L, I S, V T, HJ T. 2008. Genome Organization and Reverse Genetic Analysis of a Type I Feline

795   Coronavirus. J Virol 82.

796   50.   Donaldson EF, Yount B, Sims AC, Burkett S, Pickles RJ, Baric RS. 2008. Systematic assembly of a full-

797   length infectious clone of human coronavirus NL63. J Virol 82:11948–57.

798   51.   Á B, A F, Z Z, Á H, L D, F A, L E, S B. 2012. Molecular Characterization of Feline Infectious Peritonitis

799   Virus Strain DF-2 and Studies of the Role of ORF3abc in Viral Cell Tropism. J Virol 86.

800   52.   Tekes G, Spies D, Bank-Wolf B, Thiel V, Thiel H-J. 2012. A reverse genetics approach to study feline

801   infectious peritonitis. J Virol 86:6994–8.

802   53.   F A, ML D, I S, S Z, JL N-T, S M-J, G A, L E. 2013. Engineering a Replication-Competent, Propagation-

803   Defective Middle East Respiratory Syndrome Coronavirus as a Vaccine Candidate. MBio 4:e00650-13.

804   54.   Scobey T, Yount BL, Sims AC, Donaldson EF, Agnihothram SS, Menachery VD, Graham RL, Swanstrom

805   J, Bove PF, Kim JD, Grego S, Randell SH, Baric RS. 2013. Reverse genetics with a full-length infectious

806   cDNA of the Middle East respiratory syndrome coronavirus. Proc Natl Acad Sci U S A 110:16157–62.

807   55.   Beall A, Yount B, Lin C-M, Hou Y, Wang Q, Saif L, Baric R. 2016. Characterization of a Pathogenic Full-

808   Length cDNA Clone and Transmission Model for Porcine Epidemic Diarrhea Virus Strain PC22A. MBio

809   7:e01451-15.

810   56.   Ehmann R, Kristen-Burmann C, Bank-Wolf B, König M, Herden C, Hain T, Thiel H-J, Ziebuhr J, Tekes

811   G. 2018. Reverse Genetics for Type I Feline Coronavirus Field Isolate To Study the Molecular

812   Pathogenesis of Feline Infectious Peritonitis. MBio 9.

813   57.   D M, B M, D N, S S, N O, MA M, C D. 2017. Transgene Expression in the Genome of Middle East

814   Respiratory Syndrome Coronavirus Based on a Novel Reverse Genetics System Utilizing Red-mediated

815   Recombination Cloning. J Gen Virol 98.

816   58.   Terada Y, Kuroda Y, Morikawa S, Matsuura Y, Maeda K, Kamitani W. 2019. Establishment of a Virulent

817   Full-Length cDNA Clone for Type I Feline Coronavirus Strain C3663. J Virol 93.

818   59.   Thiel V, Herold J, Schelle B, Siddell SG. 2001. Infectious RNA transcribed in vitro from a cDNA copy of

819   the human coronavirus genome cloned in vaccinia virus. J Gen Virol 82:1273–1281.

820   60.   T TNT, F L, N E, P V, H S, J P, J K, S S, M H, A K, M G, K S, L L, L H, M W, S P, D H, V C, S C-P, S S, D

821   M, D N, VM C, MA M, C D, R D, J J, V T. 2020. Rapid Reconstruction of SARS-CoV-2 Using a Synthetic

822   Genomics Platform. Nature 582:561–565.

823   61.   Hou YJ, Okuda K, Edwards CE, Martinez DR, Asakura T, Dinnon KH, Kato T, Lee RE, Yount BL,

824   Mascenik TM, Chen G, Olivier KN, Ghio A, Tse L V., Leist SR, Gralinski LE, Schäfer A, Dang H, Gilmore

825   R, Nakano S, Sun L, Fulcher ML, Livraghi-Butrico A, Nicely NI, Cameron M, Cameron C, Kelvin DJ, de

826   Silva A, Margolis DM, Markmann A, Bartelt L, Zumwalt R, Martinez FJ, Salvatore SP, Borczuk A, Tata

827   PR, Sontake V, Kimple A, Jaspers I, O'Neal WK, Randell SH, Boucher RC, Baric RS. 2020. SARS-CoV-2

828   Reverse Genetics Reveals a Variable Infection Gradient in the Respiratory Tract. Cell.

829   62.   Casais R, Thiel V, Siddell SG, Cavanagh D, Britton P. 2001. Reverse Genetics System for the Avian

830   Coronavirus Infectious Bronchitis Virus. J Virol 75:12359–12369.

831   63.   B Y, MR D, SR W, RS B. 2002. Systematic Assembly of a Full-Length Infectious cDNA of Mouse Hepatitis

832   Virus Strain A59. J Virol 76:11065–78.

833   64.   B Y, KM C, EA F, LE H, PB J, E P, MR D, TW G, RS B. 2003. Reverse Genetics With a Full-Length

834   Infectious cDNA of Severe Acute Respiratory Syndrome Coronavirus. Proc Natl Acad Sci U S A

835   100:12995–3000.

836   65.   Coley SE, Lavi E, Sawicki SG, Fu L, Schelle B, Karl N, Siddell SG, Thiel V. 2005. Recombinant mouse

837   hepatitis virus strain A59 from cloned, full-length cDNA replicates to high titers in vitro and is fully

838   pathogenic in vivo. J Virol 79:3097–106.

839   66.   F A, ML D, C G, D E, E A, J O, I S, S Z, S A, JL M, A N, C C, L E. 2006. Construction of a Severe Acute

840   Respiratory Syndrome Coronavirus Infectious cDNA Clone and a Replicon to Study Coronavirus RNA

841   Synthesis. J Virol 80:10900–6.

842   67.   St-Jean JR, Desforges M, Almazán F, Jacomy H, Enjuanes L, Talbot PJ. 2006. Recovery of a

843   neurovirulent human coronavirus OC43 from an infectious cDNA clone. J Virol 80:3670–4.

844   68.   MM B, RL G, EF D, B R, AC S, T S, RJ P, D C, RE J, RS B, MR D. 2008. Synthetic Recombinant Bat

845   SARS-like Coronavirus Is Infectious in Cultured Cells and in Mice. Proc Natl Acad Sci U S A 105.

846   69.   Ito T, Katayama S, Okada N, Masubuchi K, Fukuyama S i., Shimizu M. 2010. Genetic and Antigenic

847   Characterization of Newly Isolated Bovine Toroviruses from Japanese Cattle. J Clin Microbiol 48:1795–

848   1800.

849   70.   Kadowaki S, Chen Z, Asanuma H, Aizawa C, Kurata T, Tamura S. 2000. Protection against influenza

850   virus infection in mice immunized by administration of hemagglutinin-expressing DNAs with

851   electroporation. Vaccine 18:2779–88.

852   71.   K T, G S, D P, A F, S K. 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Mol Biol

853 Evol 30.

854 72. Horton MR, Pease LR. 1991. Recombination and mutagenesis of DNA-sequences using PCR. In Directed

855 mutagenesis: a practical approach. Edited by M. J. McPherson. 217-247.

856 73. REED LJ, MUENCH H. 1938. A SIMPLE METHOD OF ESTIMATING FIFTY PER CENT

857 ENDPOINTS12. Am J Epidemiol 27:493–497.

858 74. van Vliet ALW, Smits SL, Rottier PJM, Groot RJ de. 2002. Discontinuous and non-discontinuous

859 subgenomic RNA transcription in a nidovirus. EMBO J 21:6571–6580.

860 75. Suzuki T, Terada Y, Enjuanes L, Ohashi S, Kamitani W. 2018. S1 Subunit of Spike Protein from a

861 Current Highly Virulent Porcine Epidemic Diarrhea Virus Is an Important Determinant of Virulence in

862 Piglets. Viruses 10.

863 76. Terada Y, Kawachi K, Matsuura Y, Kamitani W. 2017. MERS coronavirus nsp1 participates in an

864 efficient propagation through a specific interaction with viral RNA. Virology 511:95–105.

865 77. Sakai Y, Kawachi K, Terada Y, Omori H, Matsuura Y, Kamitani W. 2017. Two-amino acids change in the

866 nsp4 of SARS coronavirus abolishes viral replication. Virology 510:165–174.

867 78. Muth D, Meyer B, Niemeyer D, Schroeder S, Osterrieder N, Müller MA, Drosten C. 2017. Transgene

868 expression in the genome of Middle East respiratory syndrome coronavirus based on a novel reverse

869 genetics system utilizing Red-mediated recombination cloning. J Gen Virol 98:2461–2469.

870 79. De Groot RJ. 2006. Structure, function and evolution of the hemagglutinin-esterase proteins of corona-

871 and toroviruses. Glycoconj J 23:59–72.

872 80. B S, HJ G, R B, HD K, G H. 1990. Hemagglutinating encephalomyelitis virus attaches to N-acetyl-9-O-

873 acetylneuraminic acid-containing receptors on erythrocytes: comparison with bovine coronavirus and

874 influenza C virus. Virus Res 16.

875 81. Vlasak R, Luytjes W, Spaan W, Palese P. 1988. Human and bovine coronaviruses recognize sialic acid-

876 containing receptors similar to those of influenza C viruses. Proc Natl Acad Sci 85:4526–4529.

877 82. Desforges M, Desjardins J, Zhang C, Talbot PJ. 2013. The Acetyl-Esterase Activity of the Hemagglutinin-

878 Esterase Protein of Human Coronavirus OC43 Strongly Enhances the Production of Infectious Virus. J

879 Virol 87:3097–3107.

880 83. Bakkers MJG, Lang Y, Feitsma LJ, Hulswit RJG, de Poot SAH, van Vliet ALW, Margine I, de Groot-

Mijnes JDF, van Kuppeveld FJM, Langereis MA, Huizinga EG, de Groot RJ. 2017. Betacoronavirus Adaptation to Humans Involved Progressive Loss of Hemagglutinin-Esterase Lectin Activity. Cell Host Microbe 21:356–366.

84. Williams RK, Jiang GS, Holmes K V. 1991. Receptor for mouse hepatitis virus is a member of the carcinoembryonic antigen family of glycoproteins. Proc Natl Acad Sci 88:5533–5536.

85. Langereis MA, Vliet ALW van, Boot W, Groot RJ de. 2010. Attachment of Mouse Hepatitis Virus to O-Acetylated Sialic Acid Is Mediated by Hemagglutinin-Esterase and Not by the Spike Protein. J Virol 84:8970–8974.

86. Luytjes W, Bredenbeek PJ, Noten AFH, Horzinek MC, Spaan WJM. 1988. Sequence of mouse hepatitis virus A59 mRNA 2: Indications for RNA recombination between coronaviruses and influenza C virus. Virology 166:415–422.

87. Yokomori K, Banner LR, Lai MMC. 1991. Heterogeneity of gene expression of the hemagglutinin-esterase (HE) protein of murine coronaviruses. Virology 183:647.

88. Lissenberg A, Vrolijk MM, Vliet ALW van, Langereis MA, Groot-Mijnes JDF de, Rottier PJM, Groot RJ de. 2005. Luxury at a Cost? Recombinant Mouse Hepatitis Viruses Expressing the Accessory Hemagglutinin Esterase Protein Display Reduced Fitness In Vitro. J Virol 79:15054–15063.

89. Lehmann KC, Snijder EJ, Posthuma CC, Gorbalenya AE. 2015. What we know but do not understand about nidovirus helicases. Virus Res 202:12–32.

90. Hao W, Wojdyla JA, Zhao R, Han R, Das R, Zlatev I, Manoharan M, Wang M, Cui S. 2017. Crystal structure of Middle East respiratory syndrome coronavirus helicase. PLoS Pathog 13:e1006474.

91. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol 30:2725–9.
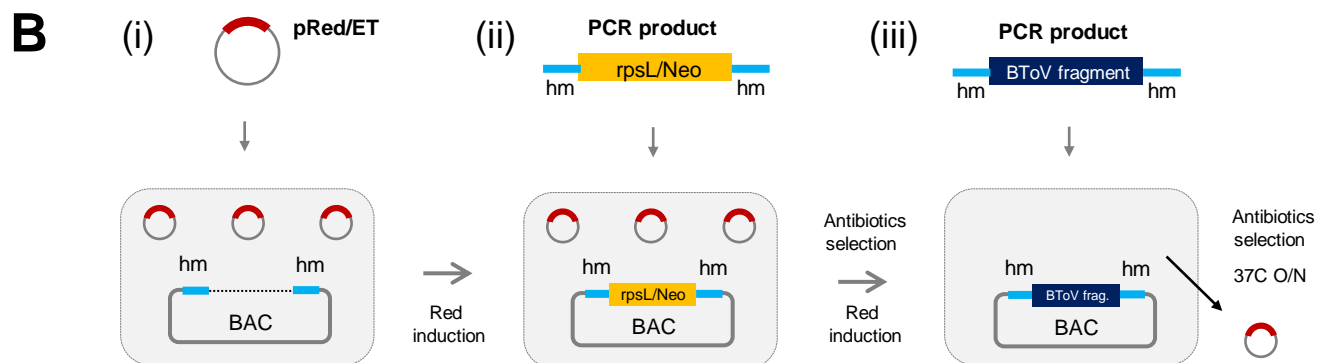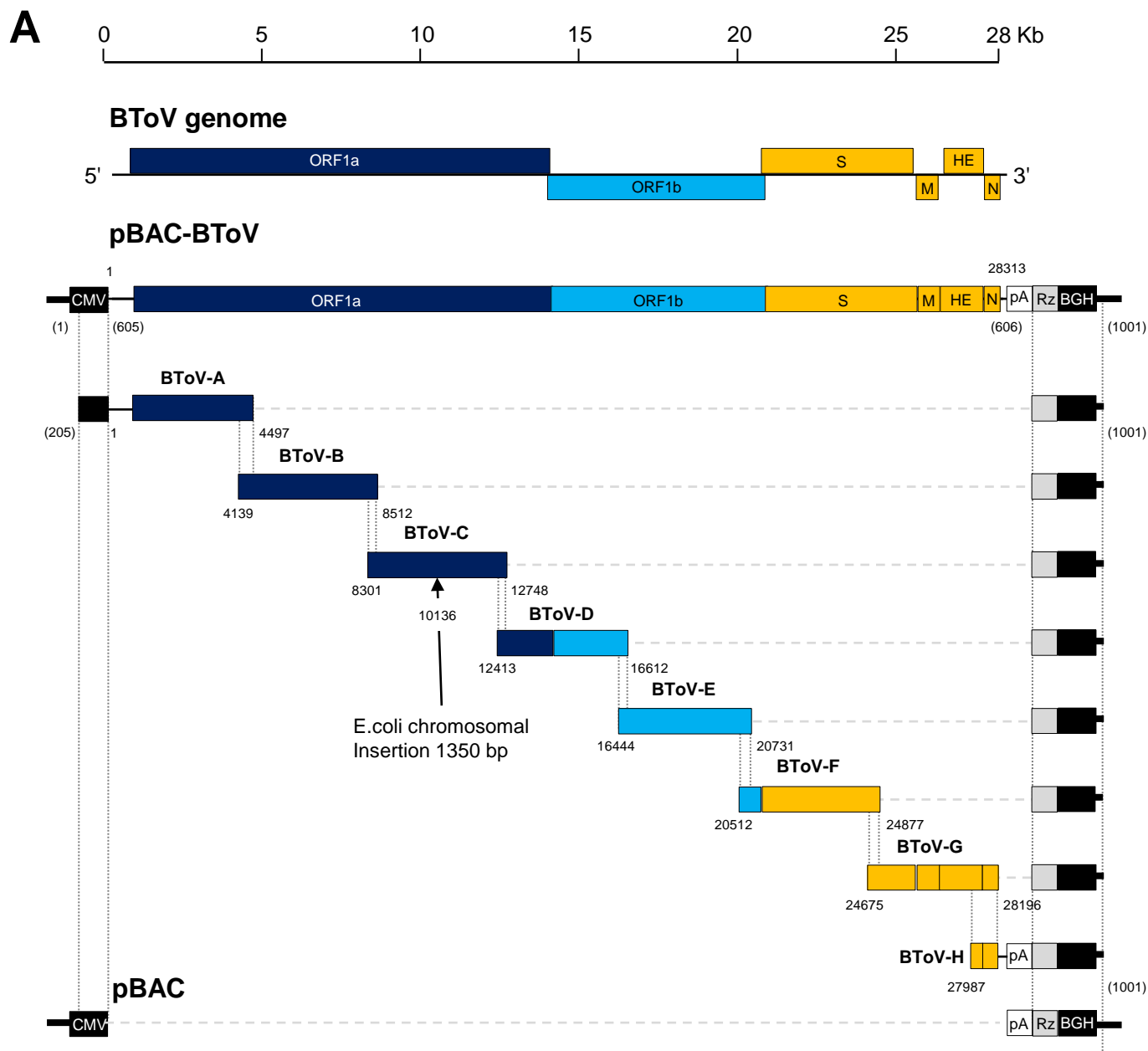
92. van Dinten LC, van Tol H, Gorbalenya AE, Snijder EJ. 2000. The predicted metal-binding region of the arterivirus helicase protein is involved in subgenomic mRNA synthesis, genome replication, and virion biogenesis. J Virol 74:5213–23.

93. Seybert A, Posthuma CC, van Dinten LC, Snijder EJ, Gorbalenya AE, Ziebuhr J. 2005. A complex zinc finger controls the enzymatic activities of nidovirus helicases. J Virol 79:696–704.

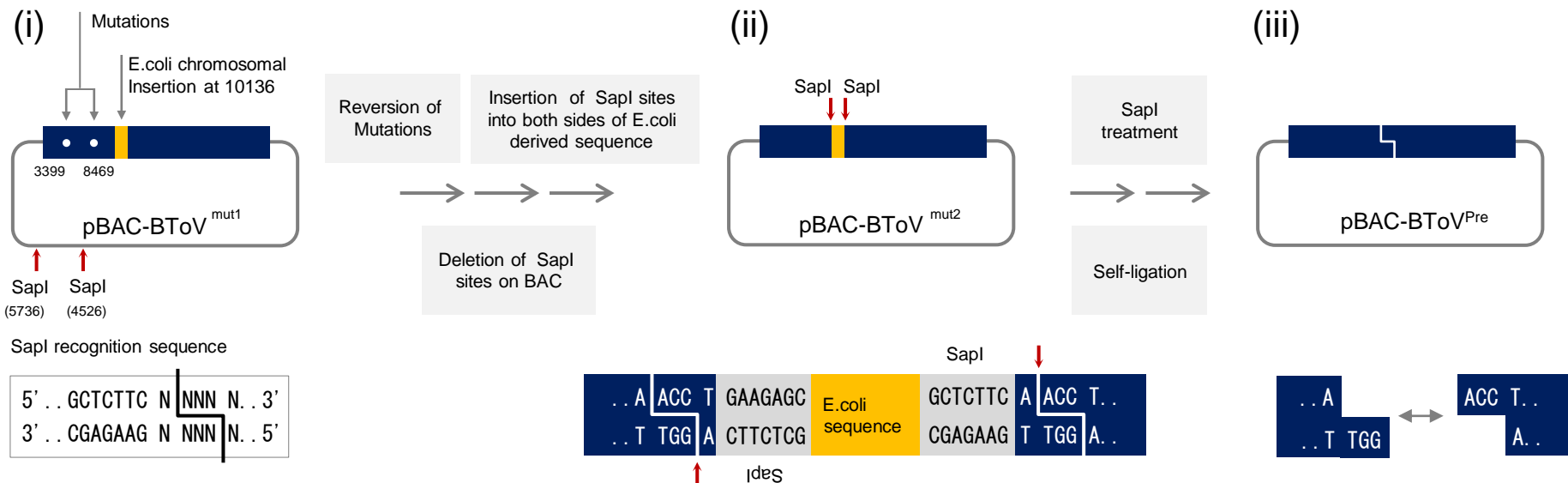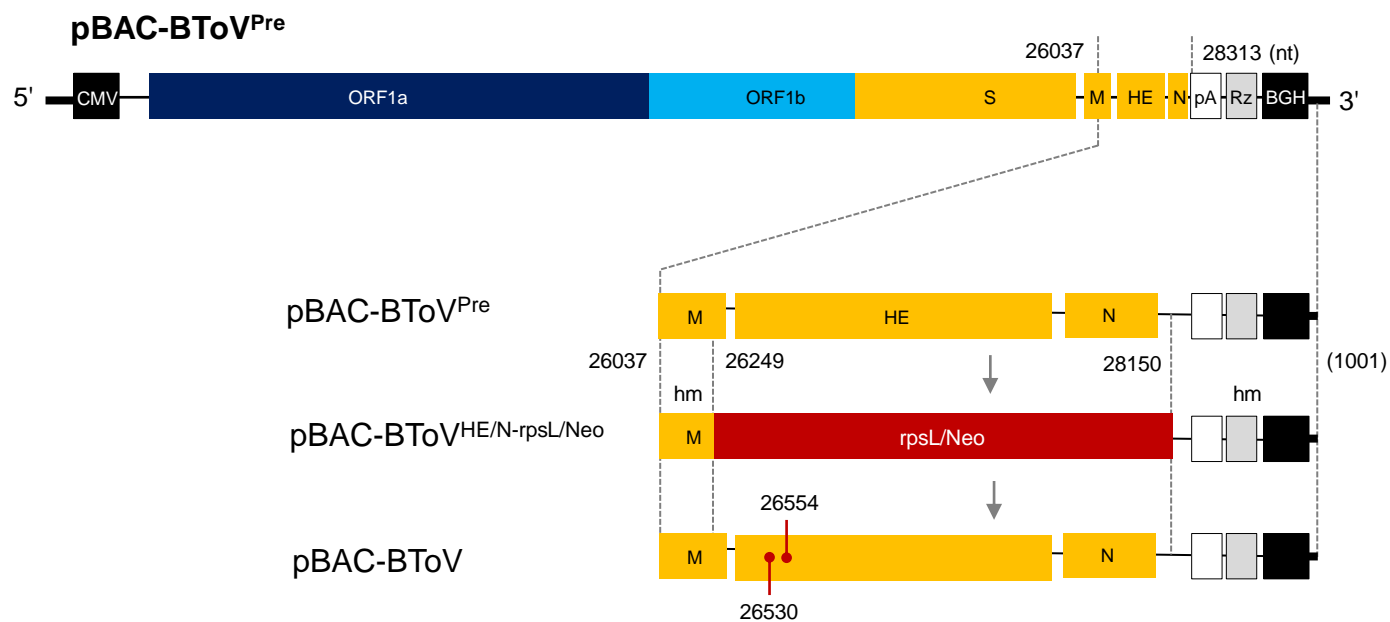94. van Vliet ALW, Smits SL, Rottier PJM, de Groot RJ. 2002. Discontinuous and non-discontinuous

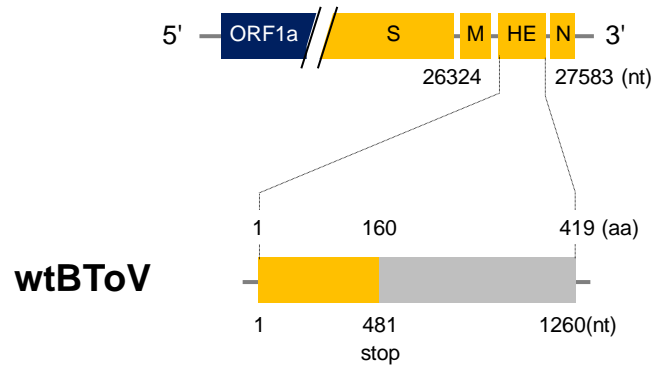909   subgenomic RNA transcription in a nidovirus. EMBO J 21:6571–80.

910   95.   Stewart H, Brown K, Dinan AM, Irigoyen N, Snijder EJ, Firth AE. 2018. Transcriptional and

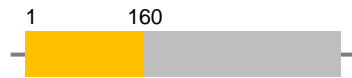911   Translational Landscape of Equine Torovirus. J Virol 92.

912

Fig.1

A

(i)

Mutations

E.coli chromosomal
Insertion at 10136

3399   8469

pBAC-BToV^mut1

SapI   SapI
(5736)   (4526)

SapI recognition sequence

```
5' . . GCTCTTC N NNN N . . 3'
3' . . CGAGAAG N NNN N . . 5'
```

Reversion of
Mutations

Insertion of SapI sites
into both sides of E.coli
derived sequence

Deletion of SapI
sites on BAC

(ii)

SapI   SapI

pBAC-BToV^mut2

SapI

```
. . A ACC T   GAAGAGC              GCTCTTC A ACC T . .
. . T TGG A   CTTCTCG  E.coli      CGAGAAG T TGG A . .
                       sequence
```

SapI

SapI
treatment

Self-ligation

(iii)

pBAC-BToV^Pre

```
. . A        ACC T . .
. . T TGG         A . .
```

B

pBAC-BToV^Pre

26037

28313 (nt)

5' CMV   ORF1a   ORF1b   S   M   HE   N pA Rz BGH   3'

pBAC-BToV^Pre

M   HE   N

26037   26249   28150   (1001)

hm

pBAC-BToV^HE/N-rpsL/Neo

M   rpsL/Neo   hm

pBAC-BToV

M   N

26554

26530

Fig.2

A

wtBToV

rBToV (rHEs)

rHEf

rHEs/HA

rHEf/HA

rEGFP

B
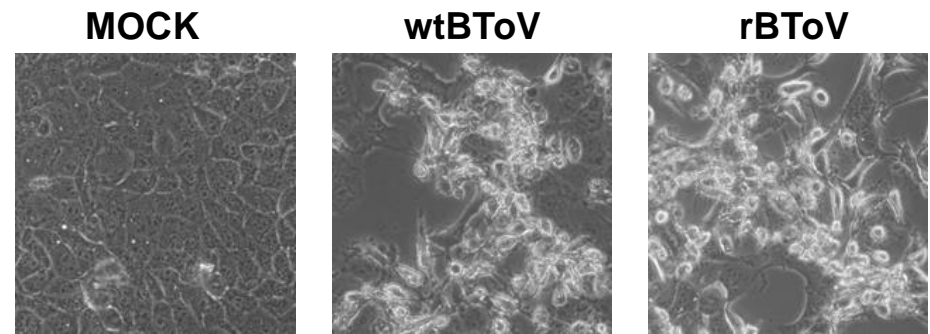
MOCK   wtBToV   rBToV

C

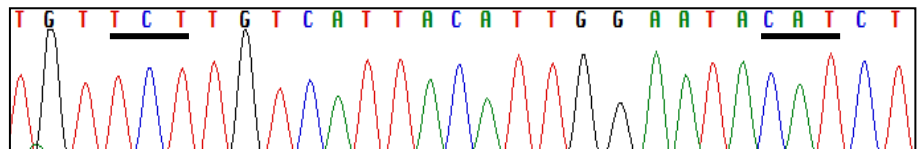   26526   26554
wtBToV: TGTT T TTGTCATTAC ATTGGAATAC T TCT
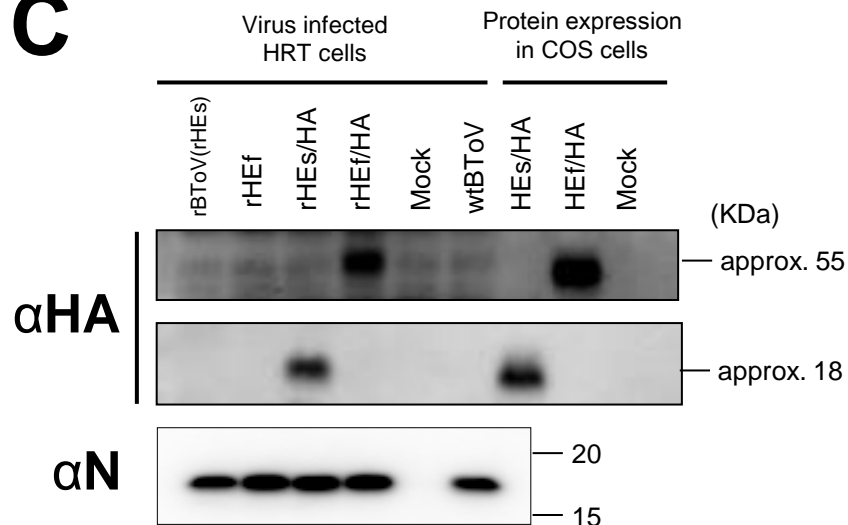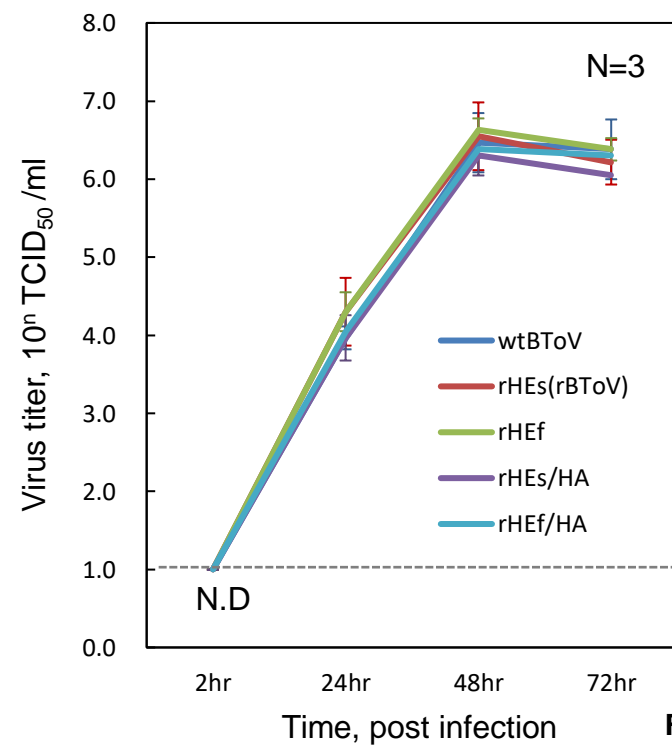
rBToV: TGTT C TTGTCATTAC ATTGGAATAC A TCT

wtBToV

TGTTTTTGTCATTACATTGGAATACTTCT

rBToV

TGTTCTTGTCATTACATTGGAATACATCT

Fig.3

A

wtBToV rBToV(rHEs) rHEf rHEs/HA rHEf/HA

B

wtBToV rBToV(rHEs) rHEf rHEs/HA rHEf/HA Mock

M

HE

HA

α-NA esterase activity

C

Virus infected HRT cells   Protein expression in COS cells

rBToV(rHEs) rHEf rHEs/HA rHEf/HA Mock wtBToV HEs/HA HEf/HA Mock

(KDa)

αHA — approx. 55

— approx. 18

αN — 20

— 15

D

N=3

Virus titer, 10^n TCID$_{50}$/ml

wtBToV
rHEs(rBToV)
rHEf
rHEs/HA
rHEf/HA

N.D

2hr   24hr   48hr   72hr

Time, post infection

Fig.4

A

Passage history

P0 P4 P9 P14 P19

rHEf No1 — wtBToV
rHEf No2 — rBToV(rHEs)
rHEf/HA No1 — Mock
rHEf/HA No2

B

αN

rBToV(rHEs) | rHEf No1 (P0 P4 P9 P14 P19) | rHEf No2 (P0 P4 P9 P14 P19) | Mock | wtBToV (KDa)

— 20
— 15

αN

rBToV(rHEs) | rHEf/HA No1 (P0 P4 P9 P14 P19) | rHEf/HA No2 (P0 P4 P9 P14 P19) | Mock | wtBToV

— 20
— 15

C

rHEf No1 | rHEf No2
M | P0 P1 P5 P10 P15 P20 | P0 P1 P5 P10 P15 P20

(Kbp)
1.5
1.0
0.5
0.1

rHEf/HA No1 | rHEf/HA No2
M | P0 P1 P5 P10 P15 P20 | P0 P1 P5 P10 P15 P20

(Kbp)
1.5
1.0
0.5
0.1

D

PCR (1670bp)

5' — ORF1a // S | M | HE | N — 3'
26324   27583 (nt)

1   17   419 (aa)

rHEf No1   6/6

rHEf No2   6/6

rHEf/HA No1   5/6   HA

rHEf/HA No2   3/6   99   HA

Fig.5

A

wtBToV rEGFP No 3 rEGFP No 6

B

Blight field	Fluorescent live cell		Confocal microscopy		
	Hoechst	EGFP	M	EGFP	Marge
wtBToV					
rBToV					
rEGFP					
Mock					

C

αGFP

αN

rBToV rEGFP No3 rEGFPNno6 Mock wtBTov (KDa)

30
20

20

15

D

Virus titer, 10^n TCID$_{50}$/ml

wtBToV
rBToV
rEGFP-1
rEGFP-2
rEGFP-3
rEGFP-4

N.D N.D

2hr 24hr 48hr 72hr

Time, post infection

Fig.6

A rEGFP

No1 No2 No3 No4 No5 No6

P0

P2

P5

B

P0 | P1 | rEGFP No5

M No1 No2 No3 No4 No5 No6 | No1 No2 No3 No4 No5 No6 | M P0 P1 P3 P6 P9 P12 P15

(Kbp)
1.5
1.0
0.5
0.1

(Kbp)
1.5
1.0
0.5

C

rBToV | rEGFP No5 | Mock wtBToV (KDa)
P0 P1 P2 P5 P8 P11

αGFP — 30
— 20

αN — 20
— 15

D

5′ ORF1a // S M EGFP N 3′
26324 27043 (nt)

PCR (1130bp)

1 720 (nt)

rEGFP — EGFP —

No2 P1 — 27 671 —

No3 P1 — -8 723 —

No4 P1 — 54 690 —

No5 P15 — 30 420 —

No6 P6 — 28 161 —

c1 P7 — 27 540 —

c2 P7 — 169 673 —

c3 P7 — 24 668 —

Fig.7

A

wtBToV rBToV P5 rEGFP No5 — c1 c2 c3

B

Virus titer, 10^n TCID$_{50}$/ml

N=4

N.D

Time, post infection

wtBToV
rBToV
c1
c2
c3

2hr 24hr 48hr 72hr

C

rEGFP No5
P5
rBToV P0 c1 c2 c3 Mock wtBToV (KDa)

αGFP — 30 — 20

αN — 20 — 15

Fig.8

D

P5 rEGFP No5 c1

5'— ORF1a — 1b — S — M — HE — N —3'

1ab

NSP1 2 3 4 5 6 7 8 9 10 11 12 13 14

G1001S T2129 I I 3562T I 5327V

C1442F

(aa)

E

P1 P2 P3 P4 P5 P6

c1

c2

rBToV c1 P0 P1 P2 P3 P4 P5 P6 Mock wtBToV   rBToV c2 P0 P1 P2 P3 P4 P5 P6 Mock wtBToV

αGFP

αN

c1   c2

(Kbp) M P0 P1 P2 P3 P4 P5 P6 P7   M P0 P1 P2 P3 P4 P5 P6 P7

1.5
1.0

0.5

0.1