

ORF18	ATGGCGTGGTACAGAAATTC AAGGTTTGTCTACAATGCTTTAAAAC TCAACTTGC GTTCC	60
ORF19	ATGGCGTGGTACAGAAATTC AAGGTTTGTCTACAATGCTTTAAAAC TCAACTTGC GTTCC	60
ORF20	ATGGCGTGGTACAGAAATTC AAGGTTTGTCTACAATGCTTTAAAAC TCAACTTGC GTTCC	60
ORF21	ATGGCGTGGTACAGAAATTC AAGGTTTGTCTACAATGCTTTAAAAC TCAACTTGC GTTCC	60
TK81-O	ATGGCGTGGTACAGAAATTC AAGGTTTGTCTACAATGCTTTAAAAC TCAACTTGC GTTCC	60
ORF18	AAAACATTTGGTACTATTC CAACTCCAAGAGTTCATT CGAATTCCTCATCTT TGTGTTTAC	120
ORF19	AAAACATTTGGTACTATTC CAACTCCAAGAGTTCATT CGAATTCCTCATCTT TGTGTTTAC	120
ORF20	AAAACATTTGGTACTATTC CAACTCCAAGAGTTCATT CGAATTCCTCATCTT TGTGTTTAC	120
ORF21	AAAACATTTGGTACTATTC CAACTCCAAGAGTTCATT CGAATTCCTCATCTT TGTGTTTAC	120
TK81-O	AAAACATTTGGTACTATTC CAACTCCAAGAGTTCATT CGAATTCCTCATCTT TGTGTTTAC	120
ORF18	AATCAATCTACTAATAAGT GTAGTGGGTTATTTGGGT CTGCAAAATCTGGGTAT TTTAAT	180
ORF19	AATCAATCTACTAATAAGT GTAGTGGGTTATTTGGGT CTGCAAAATCTGGGTAT TTTAAT	180
ORF20	AATCAATCTACTAATAAGT GTAGTGGGTTATTTGGGT CTGCAAAATCTGGGTAT TTTAAT	180
ORF21	AATCAATCTACTAATAAGT GTAGTGGGTTATTTGGGT CTGCAAAATCTGGGTAT TTTAAT	180
TK81-O	AATCAATCTACTAA---GT GTAGTGGGTTATTTGGGT CTGCAAAATCTGGGTAT TTTAAT	177
ORF18	GGGTTTAAACATCATCAAG AGATTAGCTCTTTCTCT GGGTTTTGCAAGGAGAA TTATCAT	240
ORF19	GGGTTTAAACATCATCAAG AGATTAGCTCTTTCTCT GGGTTTTGCAAGGAGAA TTATCAT	240
ORF20	GGGTTTAAACATCATCAAG AGATTAGCTCTTTCTCT GGGTTTTGCAAGGAGAA TTATCAT	240
ORF21	GGGTTTAAACATCATCAAG AGATTAGCTCTTTCTCT GGGTTTTGCAAGGAGAA TTATCAT	240
TK81-O	GGGTTTAAACATCATCAAG AGATTAGCTCTTTCTCT GGGTTTTGCAAGGAGAA CTATCAT	237
ORF18	GGTGATAAAACCGAAGTAAG TGTGAATCATGGCTGGAA AAATTCCTTGTTCCAATT GGA	300
ORF19	GGTGATAAAACCGAAGTAAG TGTGAATCATGGCTGGAA AAATTTACTTCTTC---TT GCA	297
ORF20	GGTGATAAAACCGAAGTAAG TGTGAATCATGGCTGGAA AAATTCCTTGTTCCAATT GGA	300
ORF21	GGTGATAAAACCGAAGTAAG TGTGAATCATGGCTGGAA AAATTCCTTGTTCCAATT GGA	300
TK81-O	GGTGATAAAACCGAAGTAAG TGTGAATTTCCGGTGGAA AAATTTACTTCTTGAATT GCA	297
ORF18	C-----TAATCTTGACTTT TGGTATACTTGGTTACCC TCATG <u>TGCACCCAGTAGTT</u>	351
ORF19	GTTGCAC---TAATCTTGA-----TTGCTTACCGTCATG <u>TGCACCCAGTAGTT</u>	342
ORF20	C-----TAATCTTGACTTT TGGTATACTTGGTTACCC TCATG <u>TGCACCCAGTAGTT</u>	351
ORF21	C-----TAATCTTGACTTT TGGTATACTTGGTTACCC TCATG <u>TGCACCCAGTAGTT</u>	351
TK81-O	C-----TAATAATCTCGCAT TCTGGTATGATTGCTTTCT TTTATTT <u>TGCACCCAGTAGTT</u>	351

ORF18	<u>GTGCCA</u> TATACAGGAAGGAAGCATTATGTGCTTATGTCAACAACCTCGTGAGAATGAAATT	411
ORF19	<u>GTGCCA</u> TATACAGGAAGGAAGCATTATGTGCTTATGTCAACAACCTCGTGAGAATGAAAAAT	402
ORF20	<u>GTGCCA</u> TATACAGGAAGGAAGCATTATGTGCTTATGTCAACAACCTCGTGAGAATGAAATT	411
ORF21	<u>GTGCCA</u> TATACAGGAAGGAAGCATTATGTGCTTATGTCAACAACCTCGTGAGAATGAAATT	411
TK81-O	<u>GTGCCA</u> TATACAGGAAGGAAGCATTATGTGATTTTGTCAACAACCTCATGAGAATGAAAAAT	411
	D-Fw	
ORF18	GGAGAAGTTGAGAAGCGGAAAAATACAACCTGCTACACACCCCTGATACTGATAGGGTTAGG	471
ORF19	GGAGAAGTTGAGAAGCGGAAAAATACAACCTGCTACACACCCCTGATACTGAGAGGGTTAGG	462
ORF20	GGAGAAGTTGAGAAGCGGAAAAATACAACCTGCTACACACCCCTGATACTGATAGGGTTAGG	471
ORF21	GGAGAAGTTGAGAAGCGGAAAAATACAACCTGCTACACACCCCTGATACTGATAGGGTTAGG	471
TK81-O	GGAGAATTGAGAAGCGGAAAAATACAACCTGCTACACACCCCTGATACTGAGAGGGTTAGG	471
ORF18	TCAATATTCCAACACATTCTTGAATCACTGGAAAGAGAGATTAATCACCATGAACTCGAA	531
ORF19	TCTATATTCCAACACATTATTGAATCACTGGAAAGAGAGATTAATCACCATGAACTCGAA	522
ORF20	TCAATATTCCAACACATTCTTGAATCACTGGAAAGAGAGATTAATCACCATGAACTCGAA	531
ORF21	TCAATATTCCAACACATTCTTGAATCACTGGAAAGAGAGATTAATCACCATGAACTCGAA	531
TK81-O	TCTATATTCCAACACATTCTTGAATCACTGGAAAGAGAGATTAATCACCATGAACTCGAA	531
ORF18	CTCGAACTCGAA-----AGAGATGAAACTTTCAAGGAGAAAACCATTGGAAGGAGGAG	585
ORF19	CTCGAA-----AGAGATGAAACTTTCAAGGAGAAAACCATTGGAAGGAGGAG	570
ORF20	CTCGAA-----AGAGATGAAACTTTCAAGGAGAAAACCATTGGAAGGAGGAG	579
ORF21	CTCGAACTCGAA-----AGAGATGAAACTTTCAAGGAGAAAACCATTGGAAGGAGGAG	585
TK81-O	CTCGAACTCGAACTCGAAAGAGATGAAACTTTCAAGGAGAAAACCATTGGAAGGAGGAG	591
ORF18	ACAGTTGATGATAAAGATAGTAGGAAGAAGCATAAGTGGGGCTAAGATAACTACTAACCAT	645
ORF19	ACAGTTGATGATAAAGATAGTAGGAAGAAGCATAAGTGGGGCTAAGATAACTACTAACCAT	630
ORF20	ACAGTTGATGATAAAGATAGTAGGAAGAAGCATAAGTGGGGCTAAGATAACTACTAACCAT	639
ORF21	ACAGTTGATGATAAAGATAGTAGGAAGAAGCATAAGTGGGGCTAAGATAACTACTAACCAT	645
TK81-O	ACAGATCATGATAAAGATAGTAGGAAGAAGCATAAGTGGGGCTAAGATAACTACTAACCAT	651
ORF18	TTGGAAGGGATGAATTGGGAAATTTTCGTTGTTGATAAACCGTTGGTTGAGTCCAGTTAT	705
ORF19	TTGGAAGGGTTGAATTGGGAAATTTTCGTTGTTGATAAACCGTTGGTTGAGTCCAGTTGT	690
ORF20	TTGGAAGGGATGAATTGGGAAATTTTCGTTGTTGATAAACCGTTGGTTGAGTCCAGTTAT	699
ORF21	TTGGAAGGGATGAATTGGGAAATTTTCGTTGTTGATAAACCGTTGGTTGAGTCCAGTTAT	705
TK81-O	---GAAGGGATGAATTGGGAAATTTTCGTTGTCGATAAACCGTTGGTTGAGTCCAGTTGT	708

ORF18 TTATTAGGTGGGAAGATTGTTGTTTACACCGGATTGCTCAACCATT-GCAACTCTGATG 763
 ORF19 TTATTTGATGGGAAGATTGTTGTTTACACCGGATTGCTCAACCATTT-CAACTCTGATG 748
 ORF20 TTATTAGGTGGGAAGATTGTTGTTTACACCGGATTGCTCAACCATT-GCAACTCTGATG 757
 ORF21 TTATTAGGTGGGAAGATTGTTGTTTACACCGGATTGCTCAACCATT-GCAACTCTGATG 763
 TK81-0 ATATTTGGTGGGAAGATTGTTGTTTACACTGGATTGCTCAACCATTTG-ATCTCTGATG 766

Intron 1

ORF18 CTGAATTGGCTACAATTATCGCGCATCAGGTTGGGCATGCTGTGGCTCGACATGAGGCAG 823
 ORF19 CTGAATTGGCTACAATTATCGCGCATCAGGTTGGGCATGCTGTGGCTCGACATGAGGCAG 808
 ORF20 CTGAATTGGCTACAATTATCGCGCATCAGGTTGGGCATGCTGTGGCTCGACATGAGGCAG 817
 ORF21 CTGAATTGGCTACAATTATCGCGCATCAGGTTGGGCATGCTGTGGCTCGACATGAGGCAG 823
 TK81-0 CTGAATTGGCTACAATTATCGCGCATCAGGTTGGGCATGCTGTGGCTCGACATGAGGCAG 826

Gre

ORF18 AGGATTGCAGACGATTTTCTGGTTGTTAATA---TCCCTCAACGTGATATTATTTAAAA 880
 ORF19 AGCATTGGACAGCATTGTTCTGGTGGTCAATGTTAGGGTTCACGTGACATTATTTGAAA 868
 ORF20 AGGATTGCAGACGATTTTCTGGTTGTTAATA---TCCCTCAACGTGATATTATTTAAAA 874
 ORF21 AGGATTGCAGACGATTTTCTGGTTGTTAATA---TCCCTCAACGTGATATTATTTAAAA 880
 TK81-0 AGCATTGGACAACATTGTTGTGGTCGATACTGTTAGTGATATACATGACAATATTTCAAT 886

ORF18 TTCTATTTACTGAGCCTGAATCTGCCAATGCAAGATCAAACACTACTCTTAAGGCATCCTC 940
 ORF19 TTCTATTTACTGCGCCTGAATTTGCCAATGCAAGATCAAACACTACTCTTAAGGCATCCTC 928
 ORF20 TTCTATTTACTGAGCCTGAATCTGCCAATGCAAGATCAAACACTACTCTTAAGGCATCCTC 934
 ORF21 TTCTATTTACTGAGCCTGAATCTGCCAATGCAAGATCAAACACTACTCTTAAGGCATCCTC 940
 TK81-0 ATCTATTTACTGCGCCTGAATTTGCCAATGCAATATCAAACACTACTCTCAAGGCATCCTC 946

Intron 2

ORF18 TCTTGCAAAAAGTTTGGGAAGATTATTCAGGCTAGAGCTCCACAATTACTGCCACGAACTA 1000
 ORF19 TCTTGCAAAAAGTTTGGGAAGATTATTCAGGCTAGATTTCATCAATTACTGCCACGAACTA 988
 ORF20 TCTTGCAAAAAGTTTGGGAAGATTATTCAGGCTAGAGCTCCACAATTACTGCCACGAACTA 994
 ORF21 TCTTGCAAAAAGTTTGGGAAGATTATTCAGGCTAGAGCTCCACAATTACTGCCACGAACTA 1000
 TK81-0 TCTTGCAAAAAGTTTGGGAAGATTATTCAGGCTAGATTTCATCAATTACTGCCACGAACTA 1006

D-Rv

ORF18 TCT---GCTTGTCCTTGTGGATTGTTTCCCTCGGTGTTTATTCTTTATTATGGTCGGA 1057
 ORF19 CCTTGCATTTGGGCTTTGTTGGATTGTCTTCCCTGGTGTATTCTTTATTTTGGTCGGA 1048
 ORF20 TCT---GCTTGTCCTTGTGGATTGTTTCCCTCGGTGTTTATTCTTTATTATGGTCGGA 1051
 ORF21 TCT---GCTTGTCCTTGTGGATTGTTTCCCTCGGTGTTTATTCTTTATTATGGTCGGA 1057
 TK81-0 CCTTGCACCTTGGGCTTTCTTGGATTGTCTTCCCTGGTGTATTCTTTATTTTGGTCGGA 1066

```

ORF18      AGGAAATAGAAGCAGATCACATTGGAGTGCTTCTGATGGCTTCTGCTGGATACGACCCGC 1117
ORF19      AGGAAATAGAAGCAGATCACATTGGAGTGCTTCTGATGGCTTCTGCTGGATACGACCCGC 1108
ORF20      AGGAAATAGAAGCAGATCACATTGGAGTGCTTCTGATGGCTTCTGCTGGATACGACCCGC 1111
ORF21      AGGAAATAGAAGCAGATCACATTGGAGTGCTTCTGATGGCTTCTGCTGGATACGACCCGC 1117
TK81-O     AGGAAATAGAAGCAGATCACATTGGAGTGCTTCTGATGGCTTCTGCTGGATACGACCCGC 1126

ORF18      GAGTTGCACCTCAAGTATATGACAAGCTTGCAAAGCCACTGGGCGACTGGAAGTGTTTAG 1177
ORF19      GAGTTGCACCTCAAGTATATGACAAGCTTGCAAAGCCACTGGGCGACTGGAAGTGTTTAG 1168
ORF20      GAGTTGCACCTCAAGTATATGACAAGCTTGCAAAGCCACTGGGCGACTGGAAGTGTTTAG 1171
ORF21      GAGTTGCACCTCAAGTATATGACAAGCTTGCAAAGCCACTGGGCGACTGGAAGTGTTTAG 1177
TK81-O     GAGTTGCACCTCAAGTATATGACAAGCTTGCAAAGCCACTGGGCGACTGGAAGTGTTTAG 1186

ORF18      CAACTCATCCATTTGCAAGAATGAGAGCAAAGTTGTTAGCTCGAGCTGATGTTATGAAGG 1237
ORF19      CAACTCATCCATTTGCAAGAATGAGAGCAAAGTTGTTAGCTCGAGCTGATGTTATGAAGG 1228
ORF20      CAACTCATCCATTTGCAAGAATGAGAGCAAAGTTGTTAGCTCGAGCTGATGTTATGAAGG 1231
ORF21      CAACTCATCCATTTGCAAGAATGAGAGCAAAGTTGTTAGCTCGAGCTGATGTTATGAAGG 1237
TK81-O     CAACTCATCCATTTGCAAGAATGAGAGCAAAGTTGTTAGCTCGAGCTGATGTTATGAAGG 1246

ORF18      AAGCAGATAAGATATACAATGAAGTTGTAGCAGGACGTGCAATTCAAGGTCTTCAGTAA 1296
ORF19      AAGCAGATAAGATATACAATGAAGTTGTAGCAGGACGTGCAATTCAAGGTCTTCAGTAA 1287
ORF20      AAGCAGATAAGATATACAATGAAGTTGTAGCAGGACGTGCAATTCAAGGTCTTCAGTAA 1290
ORF21      AAGCAGATAAGATATACAATGAAGTTGTAGCAGGACGTGCAATTCAAGGTCTTCAGTAA 1296
TK81-O     AAGCAGATAAGATATACAATGAAGTTGTAGCAGGACGTGCAATTCAAGGTCTTCAGTAA 1305

```

FIGURE S5.– Sequence alignment of *bvORF18* (ORF18), *bvORF19* (ORF19), *bvORF20* (ORF20), *bvORF21* (ORF21), and *bvORF20L* (TK81-O). Hyphens indicate gaps inserted for maximum matching. Residues of nucleotide sequences are numbered from the translational initiation codon. Positions of introns are shown with black triangles, but the intronic sequences are not shown. Exon/intron boundaries have been experimentally confirmed (H. Matsuhira, T. Mikami and T. Kubo, manuscript in preparation). Primer sequences are underlined. Nucleotide residues corresponding to Site 1 and Site 2 in Figure 4 are shown by red and blue letters, respectively. 5'-CTCGAA-3' repeated sequences are indicated by purple letters.