```
bvORF18/21 MAWYRNSRFVYNALKLNLRSKTFGTIPTPRVHSNSSSLFYNQSTNKCSGLFGSAKSGYFN 60
bvORF19    MAWYRNSRFVYNALKLNLRSKTFGTIPTPRVHSNSSSLFYNQSTNKCSGLFGSAKSGYFN 60
bvORF20    MAWYRNSRFVYNALKLNLRSKTFGTIPTPRVHSNSSSLFYNQSTNKCSGLFGSAKSGYFN 60
bvORF20L   MAWYRNSRFVYNALKLNLRSKTFGTIPTPRVHSNSSSLFYNQST-KCSGLFGSAKSGYFN 59
           ******************************************** ***************

bvORF18/21 GFKHHQEISSFSGFARRNYHGDKTEVSVESWLEKFLVPIGLILTFG-ILGYPHVHPVVVP 119
bvORF19    GFKHHQEISSFSGFARRNYHGDKTEVSAESLLEKLLL---LAVALI-LIAYRHVHPVVVP 116
bvORF20    GFKHHQEISSFSGFARRNYHGDKTEVSVESWLEKFLVPIGLILTFG-ILGYPHVHPVVVP 119
bvORF20L   GFKHHQEISSFSGFARRNYHGVKTEVSVEFRVEKLLLGIALIISHSGMIAFFYLHPVVVP 119
           ********************* *****.*  :**:*:   * ::   ::.: ::******

bvORF18/21 YTGRKHYVLMSTTRENEIGEVEKRKIQPATHPDTDRVRSIFQHILESLEREINHHELELE 179
bvORF19    YTGRKHYVLMSTTRENENGEVEKRKIQPATHPDTERVRSIFQHIIESLEREINHHELELE 176
bvORF20    YTGRKHYVLMSTTRENEIGEVEKRKIQPATHPDTDRVRSIFQHILESLEREINHHELELE 179
bvORF20L   YTGRKHYVILSTTHENENGEFEKRKIQPATHPDTERVRSIFQHILESLEREINHHELELE 179
           ********::***:*** **.*************:*********:***************

bvORF18/21 LE--RDETFKEKTIWKEETVDDKDSRKKHSGAKITTNHLEGMNWEIFVVDKPLVESSYLL 237
bvORF19    ----RDETFKEKTIWKEETVDDKDSRKKHSGAKITTNHLEGLNWEIFVVDKPLVESSCLF 232
bvORF20    ----RDETFKEKTIWKEETVDDKDSRKKHSGAKITTNHLEGMNWEIFVVDKPLVESSYLL 235
bvORF20L   LELERDETFKEKTIWKEETDHDKDSRKKHSGAKITTNH-EGMNWEIFVVDKPWVESSCIF 238
               *************** .***************** **:********** **** ::

bvORF18/21 GGKIVVYTGLLNHCNSDAELATIIAHQVGHAVARHEAEDSTAFFWL-LISLNVILFKILF 296
bvORF19    DGKIVVYTGLLNHFNSDAELATIIAHQVGHAVARHEAEHWTALFWWSMLGFYVTLFEILF 292
bvORF20    GGKIVVYTGLLNHCNSDAELATIIAHQVGHAVARHEAEDSTAFFWL-LISLNVILFKILF 294
bvORF20L   GGKIVVYTGLLNHCISDAELATIIAHQVGHAVARHEAEHWTTLLWSILLVIYMTIFQYLF 298
           .************  ************************. *:::*  :: : : :*: **

bvORF18/21 TEPESANARSKLLLRHPLLQKVWKIIQARAPQLLPR-TICLSLVGLFSSVFILYYGRKEI 355
bvORF19    TAPEFANARSKLLLRHPLLQKVWKIIQARFHQLLPRTTLRLGFVGLSSLVFILYFGRKEI 352
bvORF20    TEPESANARSKLLLRHPLLQKVWKIIQARAPQLLPR-TICLSLVGLFSSVFILYYGRKEI 353
bvORF20L   TAPEFANAISKLLSRHPLLQKVWKIIQARFHQLLPRTTLHLGFLGLSSLVFILYFGRKEI 358
           * ** *** **** ***************  ***** *: *.::** * *****:*****

bvORF18/21 EADHIGVLLMASAGYDPRVAPQVYDKLAKPLGDWNCLATHPFARMRAKLLARADVMKEAD 415
bvORF19    EADHIGVLLMASAGYDPRVAPQVYDKLAKPLGDWNCLATHPFARMRAKLLARADVMKEAD 412
bvORF20    EADHIGVLLMASAGYDPRVAPQVYDKLAKPLGDWNCLATHPFARMRAKLLARADVMKEAD 413
bvORF20L   EADHIGVLLMASAGYDPRVAPQVYDKLAKPLGDWNCLATHPFARMRAKLLARADVMKEAD 418
           ************************************************************

bvORF18/21 KIYNEVVAGRAIQGLQ 431
bvORF19    KIYNEVVAGRAIQGLQ 428
bvORF20    KIYNEVVAGRAIQGLQ 429
bvORF20L   KIYNEVVAGRAIQGLQ 434
           ****************
```

FIGURE S9.– Comparison of amino acid sequences of five bvORF20-related genes between NK–198 and TK–81mm–O. Note that bvORF18 and bvORF21 are identical (see Fig. S5). Amino acid residues are numbered from the first methionine residue. Asterisks (*) indicate positions that have a single, fully conserved residue; colons (:) indicate that one of the following 'strong' groups is fully conserved: STA, NEQK, NHQK, NDEQ, QHRK, MILV, MILF, HY, FYW. Points (.) indicate that one of the following 'weaker' groups is fully conserved: CSA, ATV, SAG, STNK, STPA, SGND, SNDEQK, NDEQHK, NEQHRK, FVLIM, HFY (after CLUSTAL W package). HQVGH motifs in the bvORF20-related ORFs are underlined.