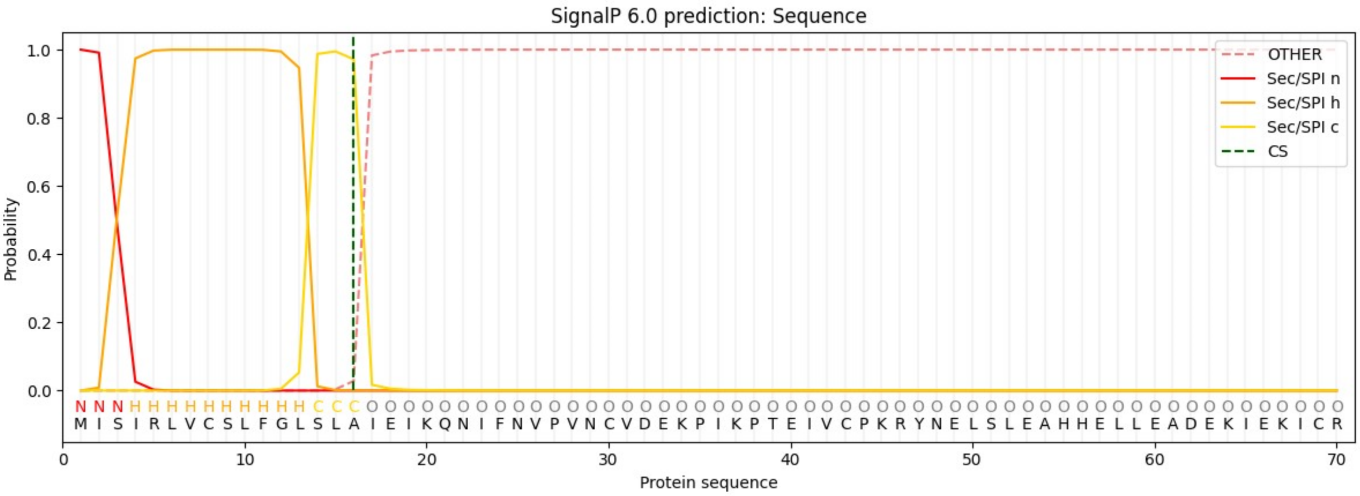
a

Cleavage site between pos. 16 and 17.
Probability 0.971839

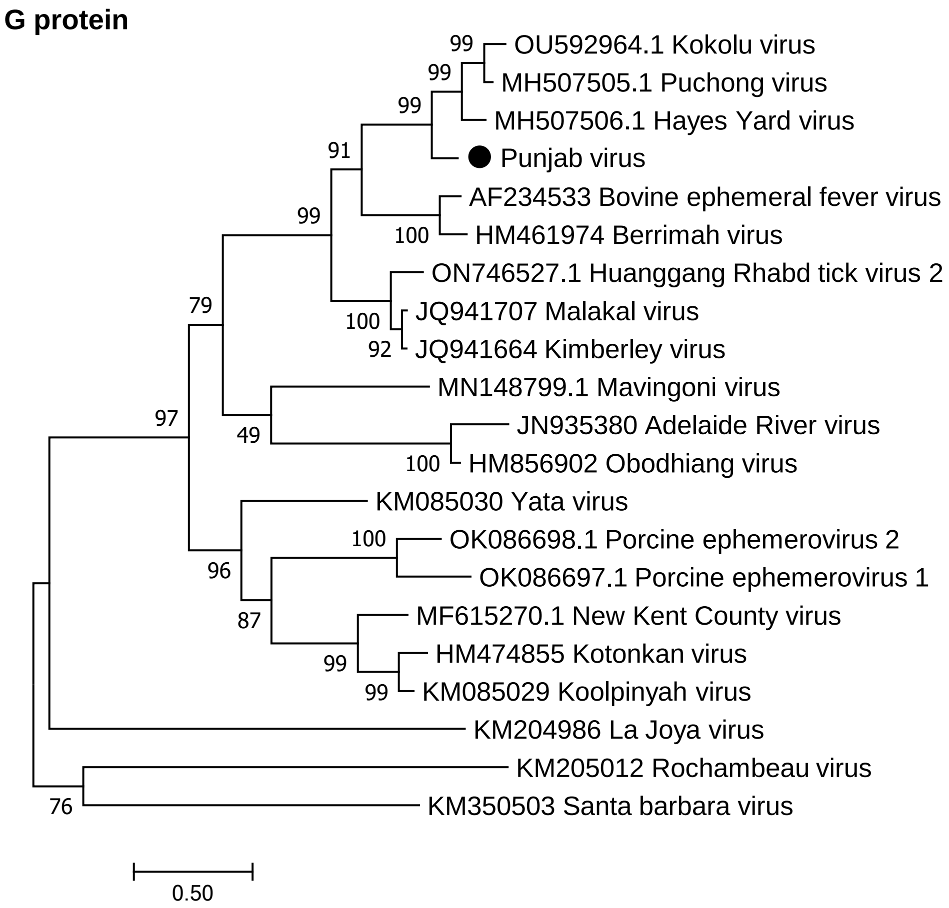


b

```
Query  1    MISIRLVCSLFGLS---LAIEIKQNIFNVPVNCVDEKPIKPTEIVCPKRYNELSLEAHHE  57
            MIS    C +FGL     + +I   +FNVPVNCV EKPIKPTEI CPKRYNELSLEAHH
Sbjct  1    MISKIQSCIIFGLCSRVFSADIGSKVFNVPVNCVGEKPIKPTEITCPKRYNELSLEAHHT  60

Query  58   LLEADEKIEKICRPALRDDDHVEGYICRQQYWETECEETWYFATVINYHIRETVPDHESC  117
            L+E DEK+E+ICRPALRDDDHVEGY+CR+Q+WETECEETWYFATVINYHIRE  P    C
Sbjct  61   LIEGDEKVEQICRPALRDDDHVEGYVCREQHWETECEETWYFATVINYHIREKPPSLSDC  120

Query  118  QNAVKQLKEGVLIPPYYPPAGCFWNTKIQESIKFMVLIRHKSLLNPMDNRIHDSQFVSSC  177
              AVKQ K+G+LIPPYYPPAGCFWNTKI+E+IKFMVLI+H+S+LNP+DN +HD+QF++ C
Sbjct  121  ITAVKQYKDGILIPPYYPPAGCFWNTKIKETIKFMVLIKHRSILNPIDNLVHDTQFINPC  180

Query  178  SLDEVKKRGCKLKDITGIWVPDLETGLESEHCSKKHWECIGIKSYKSELDKRTRIWEAPE  237
             L   +K GCKLKDITG+W+P+LE GL SEHCSKKHWECIGIKSYKSE+D++TRIWEAPE
Sbjct  181  DLKGAQKDGCKLKDITGLWIPELENGLSSEHCSKKHWECIGIKSYKSEVDEKTRIWEAPE  240

Query  238  IGIINVTKACKQDFCGYKGVVFDDGEWWGYTNDSEKELESSLVKPCKNRRPGIRVHNDHT  297
            IGIIN+TK+CKQ FCGY+GVVFDDGEWWGY N++E EL  + +  C  ++PGIRVHNDHT
Sbjct  241  IGIINITKSCKQSFCGYRGVVFDDGEWWGYANETEAELIDAHIPTCSGKKPGIRVHNDHT  300

Query  298  EYETYDIKAELENERCQNTISKILNSESINTIDMAYLSPTRPGRDFAYRFEQVNWTEEFC  357
            +YETYDI+AE+ENERCQNTISKILNSE INTIDM+YLSPTRPGRDFAYRF+QVNWTE FC
Sbjct  301  DYETYDIRAEMENERCQNTISKILNSEPINTIDMSYLSPTRPGRDFAYRFKQVNWTETFC  360

Query  358  LHWGTTGAVKNCKIPWTNNVRGGRVKKDYLGIGGYKRATCDFRPILDEDGNGYIDESELH  417
            L W  +G +K+C+  W  + RGG+V K+++GIGGY RA C+FRPI+D+DG+GYI ++EL
Sbjct  361  LRWAESGLIKDCRKHWKLSDRGGKVTKEHVGIGGYTRALCEFRPIVDQDGDGYISKTELG  420

Query  418  GHKISNNFSVMKLSFREGKDLSEAIEVGYNGMIKVKNENQNGSMYMVRTKSVYDGVEQRN  477
             H +S  FS MKL  R      ++++VG+NG++KV ++      YMVRT SVYDG+EQRN
Sbjct  421  RHNMSAEFSSMKLGTRSSNQPVDSLDVGFNGIVKVTDDRTREEKYMVRTTSVYDGLEQRN  480

Query  478  RFMKFEVMEFEDVLVRYRGEGELNTKELIVDLTPENKKNISRTDFVQTIAKGGREIVSGI  537
            R MKFEV+EFEDV+ +Y+GEG  + K   +DLTP+ +K+I+RTDFV+TIA GG+EIVSGI
Sbjct  481  RLMKFEVLEFEDVVSKYQGEGHYDVKGKSIDLTPDEQKSINRTDFVKTIANGGKEIVSGI  540

Query  538  VGWFTGAAKIVRWTIWGIGGIVTTYAIYKLHKMIRKKKSGENHNETEEEKVESG--NEWY  595
            VGWFTG AK+VRWTIW +G +VTTYAIYKLHK++R K   E      E++ E    + W+
Sbjct  541  VGWFTGTAKLVRWTIWAVGALVTTYAIYKLHKLVRGKPRSEEDRSQAEDRAEENKRSSWF  600

Query  596  HISPKNKRRPINLKKEG-LYEEIDLGNEYSPNEQKHMRKSPNHHGFFNV  643
            +   KN  R   LK +  LYEEIDL + YS  +++  +       FF+V
Sbjct  601  NSGKKNDLRKNPLKHDNPLYEEIDLEDTYSMRKKEGWKTDRTRDNFFDV  649
```

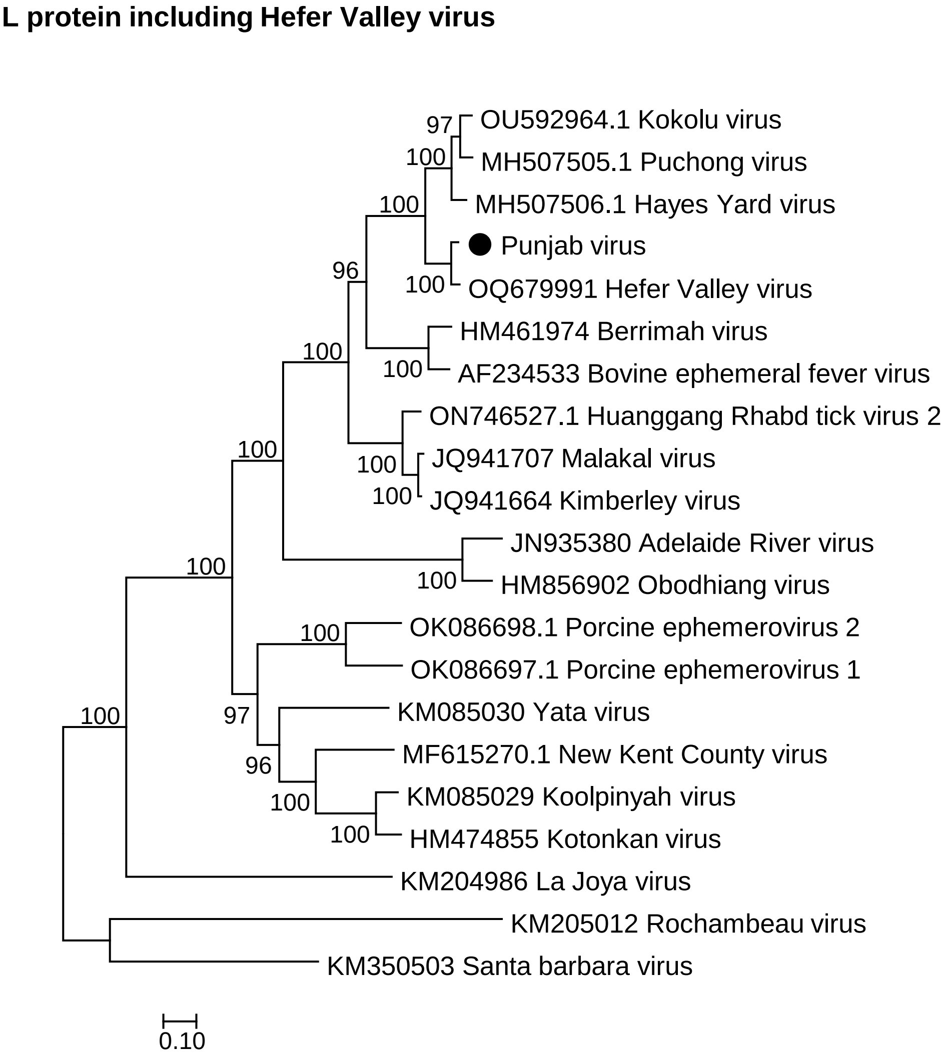
Query: Punjab virus G (643 aa)
Sbjct: Puchong virus G (649 aa)

Supplementary Figure 1. Characterization of Punjab virus G protein. (a) Signal peptide prediction by SignalP 6.0 server. n, h, c indicates N-terminal, center hydrophobic, and C-terminal region of the predicted signal peptide. CS, cleavage site. **(b)** BLASTp was performed against the NCBI nr database using the Punjab virus G protein sequence as a query. The resultant BLAST alignment of the best hit (Puchong virus G protein) is shown.



G protein

Supplementary Figure 2. Phylogenetic relationship of Punjab virus and ephemeroviruses. Phylogenetic trees were reconstructed by the maximum likelihood method using amino acid sequences of G protein of Punjab virus, ephemeroviruses, and outgroup rhabdoviruses. Bootstrap values equal or more than 70 are shown on each branch. The scale bar indicates the number of amino acid substitutions per site.



L protein including Hefer Valley virus

Supplementary Figure 3. Phylogenetic relationship of Punjab virus and ephemeroviruses including Hefer Valley virus. Phylogenetic trees were reconstructed by the maximum likelihood method using amino acid sequences of L protein of Punjab virus, ephemeroviruses, and outgroup rhabdoviruses. Bootstrap values equal or more than 70 are shown on each branch. The scale bar indicates the number of amino acid substitutions per site.