Figure S4. (A). Homology model of N-terminal domain (NTD) of spike protein from SARS-related coronavirus BtKY72 showing the locations of three CSIs specific for the *Sarbecovirus*-lineage 3. The homology models are based on available experimental structure of SARS coronavirus BJ012 spike protein (PDB: 5x58). The location of identified 1 aa deletion is highlighted. Detailed species distribution information for the 1 aa deletion in Figure 8C. (B). Homology model of RBD of SARS-related coronavirus BtKY72 spike protein (Acc no: APO40579) based on available experimental structure of SARS coronavirus BJ012 spike protein (PDB: 5x58). The receptor binding domain (RBD) domain is highlighted as green and extended C-terminal region as pale green. The locations of identified signature indels are highlighted red and labelled. Detailed species distribution information for the 1 aa insert and 2 aa insert in RBD in Figure 8A and 8B.

```
QHS34546_SARS-CoV-2/ human/ IND/        ICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFA
QJA16794_SARS-CoV-2/ human/ ID-U        ICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFA
QIG55945.1_Pangolin_CoV_MP789           ICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFA
AVP78031_Bat-SARS-like-CoVZC45          ICAQKFNGLTVLPPLLTDEMIAAYTAALISGTATAGWTFGAGAALQIPFA
AVP78042_Bat-SARS-like_CoVZXCC          ICAQKFNGLTVLPPLLTDEMIAAYTAALISGTATAGWTFGAGAALQIPFA
                                        *******************  **:**::** *:*.*************


YP_009724390_SARS-COV-2/ Wuhan-         MQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQ
QHS34546_SARS-CoV-2/ human/ IND/        MQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQ
QJA16794_SARS-CoV-2/ human/ ID-U        MQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQ
QIG55945.1_Pangolin_CoV_MP789           MQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQ
AVP78031_Bat-SARS-like-CoVZC45          MQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQESLTSTASALGKLQ
AVP78042_Bat-SARS-like_CoVZXCC          MQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQESLTSTASALGKLQ
                                        ***********************************:**:***********


YP_009724390_SARS-COV-2/ Wuhan-         DVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITG
QHS34546_SARS-CoV-2/ human/ IND/        DVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITG
QJA16794_SARS-CoV-2/ human/ ID-U        DVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITG
QIG55945.1_Pangolin_CoV_MP789           DVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITG
AVP78031_Bat-SARS-like-CoVZC45          DVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITG
AVP78042_Bat-SARS-like_CoVZXCC          DVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITG
                                        **************************************************


YP_009724390_SARS-COV-2/ Wuhan-         RLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHL
QHS34546_SARS-CoV-2/ human/ IND/        RLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHL
QJA16794_SARS-CoV-2/ human/ ID-U        RLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHL
QIG55945.1_Pangolin_CoV_MP789           RLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHL
AVP78031_Bat-SARS-like-CoVZC45          RLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHL
AVP78042_Bat-SARS-like_CoVZXCC          RLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHL
                                        **************************************************


YP_009724390_SARS-COV-2/ Wuhan-         MSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNG
QHS34546_SARS-CoV-2/ human/ IND/        MSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNG
QJA16794_SARS-CoV-2/ human/ ID-U        MSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNG
QIG55945.1_Pangolin_CoV_MP789           MSFPQSAPHGVVFLHVTYVPSQEKNFTTTPAICHEGKAHFPREGVFVSNG
AVP78031_Bat-SARS-like-CoVZC45          MSFPQSAPHGVVFLHVTYIPSQEKNFTTAPAICHEGKAHFPREGVFVSNG
AVP78042_Bat-SARS-like_CoVZXCC          MSFPQSAPHGVVFLHVTYIPSQEKNFTTAPAICHEGKAHFPREGVFVSNG
                                        ******************:*:*******:*****:*************


YP_009724390_SARS-COV-2/ Wuhan-         THWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFK
QHS34546_SARS-CoV-2/ human/ IND/        THWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFK
QJA16794_SARS-CoV-2/ human/ ID-U        THWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFK
QIG55945.1_Pangolin_CoV_MP789           THWFVTQRNFYEPQIITTDNTFVSGSCDVVIGIVNNTVYDPLQPELDSFK
AVP78031_Bat-SARS-like-CoVZC45          THWFVTQRNFYEPKIITTDNTFVSGNCDVVIGIINNTVYDPLQPELDSFK
AVP78042_Bat-SARS-like_CoVZXCC          THWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIINNTVYDPLQPELDSFK
                                        *************:*********** *******:****************


YP_009724390_SARS-COV-2/ Wuhan-         EELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLID
QHS34546_SARS-CoV-2/ human/ IND/        EELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLID
QJA16794_SARS-CoV-2/ human/ ID-U        EELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLID
QIG55945.1_Pangolin_CoV_MP789           EELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLID
AVP78031_Bat-SARS-like-CoVZC45          EELDKYFKNHTSPDIDLGDISGINASVVNIQKEIDRLNEVARNLNESLID
AVP78042_Bat-SARS-like_CoVZXCC          EELDKYFKNHTSPDIDLGDISGINASVVNIQKEIDRLNEVARNLNESLID
                                        *************:***************************:********


YP_009724390_SARS-COV-2/ Wuhan-         LQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCS
QHS34546_SARS-CoV-2/ human/ IND/        LQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCS
QJA16794_SARS-CoV-2/ human/ ID-U        LQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCS
QIG55945.1_Pangolin_CoV_MP789           LQELGKYEQYIKWPWYIWLGFIAGLIAIIMVTIMLCCMTSCCSCLKGCCS
AVP78031_Bat-SARS-like-CoVZC45          LQELGKYEQYIKWPWYVWLGFIAGLIAIVMVTILLCCMTSCCSCLKGCCS
AVP78042_Bat-SARS-like_CoVZXCC          LQELGKYEHYIKWPWYVWLGFIAGLIAIVMVTILLCCMTSCCSCLKGCCS
                                        ********:*******:***********:****:****************


YP_009724390_SARS-COV-2/ Wuhan-         CGSCCKFDEDDSEPVLKGVKLHYT
QHS34546_SARS-CoV-2/ human/ IND/        CGSCCKFDEDDSEPVLKGVKLHYT
QJA16794_SARS-CoV-2/ human/ ID-U        CGSCCKFDEDDSEPVLKGVKLHYT
QIG55945.1_Pangolin_CoV_MP789           CGSCCKFDEDDSEPVLKGVKLHYT
AVP78031_Bat-SARS-like-CoVZC45          CGSCCKFDEDDSEPVLKGVKLHYT
AVP78042_Bat-SARS-like_CoVZXCC          CGFCCKFDEDDSEPVLKGVKLHYT
                                        ** *********************
```

```
YP_009724390_SARS-COV-2/ Wuhan-      FVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNY
QHS34546_SARS-CoV-2/ human/ IND/     FVIRGDEVIQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNY
QJA16794_SARS-CoV-2/ human/ ID-U     FVRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNY
QIG55945.1_Pangolin_CoV_MP789        FVVRGDEVRQIAPGQTGRIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNY
AVP78031_Bat-SARS-like-CoVZC45       FLIRFSEVRQVAPGQTGVIADYNYKLPDDFTGCVIAWNTAKQDVG-----
AVP78042_Bat-SARS-like_CoVZXCC       FLIRFSEVRQVAPGQTGVIADYNYKLPDDFTGCVIAWNTAKQDTG-----
                                     *::* .** *:****** ********************: : *


YP_009724390_SARS-COV-2/ Wuhan-      NYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQP
QHS34546_SARS-CoV-2/ human/ IND/     NYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQP
QJA16794_SARS-CoV-2/ human/ ID-U     NYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQP
QIG55945.1_Pangolin_CoV_MP789        NYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFHP
AVP78031_Bat-SARS-like-CoVZC45       NYFYRSHRSTKLKPFERDLSSDE--------NGVR------TLSTYDFNP
AVP78042_Bat-SARS-like_CoVZXCC       HYFYRSHRSTKLKPFERDLSSDE--------NGVR------TLSTYDFNP
                                     :*:** .*.::*******:*:: ***. .*.:*.*:*


YP_009724390_SARS-COV-2/ Wuhan-      TNGVGYQPYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT
QHS34546_SARS-CoV-2/ human/ IND/     TNGVGYQPYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT
QJA16794_SARS-CoV-2/ human/ ID-U     TNGVGYQPYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT
QIG55945.1_Pangolin_CoV_MP789        TNGVGYQPYRVVVLSFELLKAPATVCGPKQSTNLVKNKCVNFNFNGLTGT
AVP78031_Bat-SARS-like-CoVZC45       NVPLEYQATRVVVLSFELLNAPATVCGPKLSTQLVKNQCVNFNFNGLKGT
AVP78042_Bat-SARS-like_CoVZXCC       NVPLEYQATRVVVLSFELLNAPATVCGPKLSTQLVKNQCVNFNFNGLKGT
                                     . : **. ********:********* **:****:********* .**


YP_009724390_SARS-COV-2/ Wuhan-      GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVIT
QHS34546_SARS-CoV-2/ human/ IND/     GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVIT
QJA16794_SARS-CoV-2/ human/ ID-U     GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVIT
QIG55945.1_Pangolin_CoV_MP789        GVLTESSKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVIT
AVP78031_Bat-SARS-like-CoVZC45       GVLTDSSKRFQSFQQFGKDASDFIDSVRDPQTLEILDITPCSFGGVSVIT
AVP78042_Bat-SARS-like_CoVZXCC       GVLTDSSKRFQSFQQFGKDASDFIDSVRDPQTLEILDITPCSFGGVSVIT
                                     ****:*.*:* .*****:* :* *:**********************


YP_009724390_SARS-COV-2/ Wuhan-      PGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGC
QHS34546_SARS-CoV-2/ human/ IND/     PGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGC
QJA16794_SARS-CoV-2/ human/ ID-U     PGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGC
QIG55945.1_Pangolin_CoV_MP789        PGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGC
AVP78031_Bat-SARS-like-CoVZC45       PGTNTSLEVAVLYQDVNCTDVPTTIHADQLTPAWRIYATGTNVFQTRAGC
AVP78042_Bat-SARS-like_CoVZXCC       PGTNTSSEVAVLYQDVNCTDVPTTIHADQLTPAWRIYAIGTSVFQTRAGC
                                     ****** :*******.****:** .:********:**:*: *:.****:***


YP_009724390_SARS-COV-2/ Wuhan-      LIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVASQSIIAYTMSL
QHS34546_SARS-CoV-2/ human/ IND/     LIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVASQSIIAYTMSL
QJA16794_SARS-CoV-2/ human/ ID-U     LIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVASQSIIAYTMSL
QIG55945.1_Pangolin_CoV_MP789        LIGAEHVNNTYECDIPIGAGICASYQTQTNS----RSVSSQAIIAYTMSL
AVP78031_Bat-SARS-like-CoVZC45       LIGAEHVNASYECDIPIGAGICASYHTASIL----RSTSQKAIVAYTMSL
AVP78042_Bat-SARS-like_CoVZXCC       LIGAEHVNASYECDIPIGAGICASYHTASIL----RSTGQKAIVAYTMSL
                                     ******** :****************:* : **...:*:*.******


YP_009724390_SARS-COV-2/ Wuhan-      GAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTEC
QHS34546_SARS-CoV-2/ human/ IND/     GAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTEC
QJA16794_SARS-CoV-2/ human/ ID-U     GAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTEC
QIG55945.1_Pangolin_CoV_MP789        GAENSVAYANNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSIEC
AVP78031_Bat-SARS-like-CoVZC45       GAENSIAYANNSIAIPTNFSISVTTEVMPVSMAKTSVDCTMYICGDSIEC
AVP78042_Bat-SARS-like_CoVZXCC       GAENSIAYANNSIAIPTNFSISVTTEVMPVSMAKTSVDCTMYICGDSIEC
                                     *****:**:**********:******::****:*********** **


YP_009724390_SARS-COV-2/ Wuhan-      SNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGG
QHS34546_SARS-CoV-2/ human/ IND/     SNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGG
QJA16794_SARS-CoV-2/ human/ ID-U     SNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGG
QIG55945.1_Pangolin_CoV_MP789        SNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGG
AVP78031_Bat-SARS-like-CoVZC45       SNLLLQYGSFCTQLNRALSGIAIEQDKNTQEVFAQVKQIYKTPPIKDFGG
AVP78042_Bat-SARS-like_CoVZXCC       SNLLLQYGSFCTQLNRALSGIAIEQDKNTQEVFAQVKQIYKTPPIKDFGG
                                     ******************:***:*******************


YP_009724390_SARS-COV-2/ Wuhan-      FNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDL
QHS34546_SARS-CoV-2/ human/ IND/     FNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDL
QJA16794_SARS-CoV-2/ human/ ID-U     FNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDL
QIG55945.1_Pangolin_CoV_MP789        FNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDL
AVP78031_Bat-SARS-like-CoVZC45       FNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGGISARDL
AVP78042_Bat-SARS-like_CoVZXCC       FNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDISARDL
                                     *******************************************.*:****


YP_009724390_SARS-COV-2/ Wuhan-      ICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFA
```

S3 Fig. **A multiple sequence alignment of the spike protein from pangolin-CoV_MP789 and representative CoV strains from SARS-CoV-2 and CoVZC clusters showing the chimeric nature of MP789 sequence**. In this alignment, the polymorphic positions where the sequence of pangolin-CoV_MP789 is identical to that of CoVZC cluster of viruses are highlighted in yellow, whereas the polymorphic positions where the sequence of pangolin-CoV_MP789 is similar to SARS-CoV-2 viruses are highlighted in blue.

```
YP_009724390_SARS-COV-2/ Wuhan-        -MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLH
QHS34546_SARS-CoV-2/ human/ IND/       -MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLH
QJA16794_SARS-CoV-2/ human/ ID-U       -MFVFLVLLPLVSSQCVNLTTRTQLPPSYTNSFTRGVYYPDKVFRSSVLH
QIG55945.1_Pangolin_CoV_MP789          MLFFFFLHFALVNSQCVNLTGRAAIQPSFTNSSQRGVYYPDTIFRSNTLV
AVP78031_Bat-SARS-like-CoVZC45         MLFFLFLQFALVNSQCVNLTGRTPLNPNYTNSSQRGVYYPDTIYRSDTLV
AVP78042_Bat-SARS-like_CoVZXCC         MLFFLFLQFALVNSQCD-LTGRTPLNPNYTNSSQRGVYYPDTIYRSDTLV
                                        :*.::: :.**.***   ** *: : * :***  *******.::**..*


YP_009724390_SARS-COV-2/ Wuhan-        STQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSN
QHS34546_SARS-CoV-2/ human/ IND/       STQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSN
QJA16794_SARS-CoV-2/ human/ ID-U       STQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSN
QIG55945.1_Pangolin_CoV_MP789          LSQGYFLPFYSNVSWYYALT-KTNSAEKRVDNPVLDFKDGIYFAATEKSN
AVP78031_Bat-SARS-like-CoVZC45         LSQGYFLPFYSNVSWYYSLT-TNNAATKRTDNPILDFKDGIYFAATEHSN
AVP78042_Bat-SARS-like_CoVZXCC         LSQGYFLPFYSNVSWYYSLT-TNNAATKRTDNPILDFKDGIYFAATEHSN
                                        :*. ****:***:::::   . . .  ** ***:* *:**:***:**:**


YP_009724390_SARS-COV-2/ Wuhan-        IIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNN
QHS34546_SARS-CoV-2/ human/ IND/       IIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVY-HKNN
QJA16794_SARS-CoV-2/ human/ ID-U       IIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNN
QIG55945.1_Pangolin_CoV_MP789          IVRGWIFGTTLDNTSQSLLIVNNATNVIIKVCNFQFCYDPYLSGYYHNN-
AVP78031_Bat-SARS-like-CoVZC45         IIRGWIFGTTLDNTSQSLLIVNNATNVIIKVCNFDFCYDPYLSGYYHNN-
AVP78042_Bat-SARS-like_CoVZXCC         IVRGWIFGTTLDNTSQSLLIVNNATNVIIKVCNFDFCYDPYLSGYYHNN-
                                        *:**********..:***********:****:*:** **:*. * *:*


YP_009724390_SARS-COV-2/ Wuhan-        KSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDG
QHS34546_SARS-CoV-2/ human/ IND/       KSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDG
QJA16794_SARS-CoV-2/ human/ ID-U       KSWMESEFRVYSSANNCTFEYVSQPFLMDXEGKQXVFKNLREFVFKNIDG
QIG55945.1_Pangolin_CoV_MP789          KTWSTREFAVYSSYANCTFEYVSKSFMLDIAGKSGLFDTLREFVFRNVDG
AVP78031_Bat-SARS-like-CoVZC45         KTWSIREFAVYSSYANCTFEYVSKSFMLNISGNGGLFNTLREFVFRNVDG
AVP78042_Bat-SARS-like_CoVZXCC         KTWSIREFAVYSFYANCTFEYVSKSFMLNISGNGGLFNTLREFVFRNVDG
                                        *.*   ** ***  ********:.*::: *:   *..*******:*:**


YP_009724390_SARS-COV-2/ Wuhan-        YFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYL
QHS34546_SARS-CoV-2/ human/ IND/       YFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYL
QJA16794_SARS-CoV-2/ human/ ID-U       YFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYL
QIG55945.1_Pangolin_CoV_MP789          YFKIYSKYTPVNVNSNLPIGFSALEPLVEIPAGINITKFRTLLTIHRGDP
AVP78031_Bat-SARS-like-CoVZC45         HFKIYSKFTPVNLNRGLPTGLSVLQPLVELPVSINITKFRTLLTIHRGDP
AVP78042_Bat-SARS-like_CoVZXCC         HFKIYSKFTPVNLNRGLPTGLSVLQPLVELPVSINITKFRTLLTIHRGDP
                                        :******.**:*:  .** *:*.:****:* .****:*:***:**.


YP_009724390_SARS-COV-2/ Wuhan-        TPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSET
QHS34546_SARS-CoV-2/ human/ IND/       TPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSET
QJA16794_SARS-CoV-2/ human/ ID-U       TPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSET
QIG55945.1_Pangolin_CoV_MP789          MPNN---GWTVFSAAYYVGYLAPRTFMLNYNENGTITDAVDCALDPLSEA
AVP78031_Bat-SARS-like-CoVZC45         MPNN---GWTAFSAAYFVGYLKPRTFMLKYNENGTITDAVDCALDPLSET
AVP78042_Bat-SARS-like_CoVZXCC         MSNN---GWTAFSAAYFVGYLKPRTFMLKYNENGTITDAVDCALDPLSET
                                        ..:   ***. :***:**** ****:*:********************:


YP_009724390_SARS-COV-2/ Wuhan-        KCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFAS
QHS34546_SARS-CoV-2/ human/ IND/       KCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFAS
QJA16794_SARS-CoV-2/ human/ ID-U       KCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFAS
QIG55945.1_Pangolin_CoV_MP789          KCTLKSLTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATTFAS
AVP78031_Bat-SARS-like-CoVZC45         KCTLKSLTVQKGIYQTSNFRVQPTQSVVRFPNITNVCPFHKVFNATRFPS
AVP78042_Bat-SARS-like_CoVZXCC         KCTLKSLSVQKGIYQTSNFRVQPTQSIVRFPNITNVCPFHKVFNATRFPS
                                        ******::*:***************:.*******:*** :***** *.*


YP_009724390_SARS-COV-2/ Wuhan-        VYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADS
QHS34546_SARS-CoV-2/ human/ IND/       VYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADS
QJA16794_SARS-CoV-2/ human/ ID-U       VYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADS
QIG55945.1_Pangolin_CoV_MP789          VYAWNRKRISNCVADYSVLYNSTSFSTFKCYGVSPTKLNDLCFTNVYADS
AVP78031_Bat-SARS-like-CoVZC45         VYAWERTKISDCIADYTVFYNSTSFSTFKCYGVSPSKLIDLCFTSVYADT
AVP78042_Bat-SARS-like_CoVZXCC         VYAWERTKISDCIADYTVFYNSTSFSTFKCYGVSPSKLIDLCFTSVYADT
                                        ****:*.:**:*:***:.:***.***********:** ***** .****: 
```

AAX16193.1 SARS-coronavirus-WH20
AAP51225.1 SARS coronavirus GD01
AAS00002.1 SARS coronavirus GZ02
AAR91584.1 SARS coronavirus NS-1
ACB69882.1 SARS coronavirus BJ182-4
ACZ71840.1 SARS coronavirus wtic-MB
AAP30028.1 SARS coronavirus BJ01
AAY60779.1 SARS coronavirus BJ162
87 AAR23251.1 SARS coronavirus Sino3-11
AAR87532.1 SARS coronavirus TW3
AFD56971.1 SARS coronavirus HKU-39849
AEA10622.1 SARS coronavirus MA15
AAR87543.1 SARS coronavirus TW4
BAE93399.1 SARS coronavirus Frankfurt 1
ACZ72179.1 SARScoronavirusExoN1
AAR87598.1 SARS coronavirus TW9
AAU04645.1 Civet SARS CoV 007/2004
48 AAP82978.1 SARS coronavirus ShanghaiQXC1
AIA62319.1 BtRs-BetaCoV/GX2013

Sarbecovirus
Lineage-1
SARS and SARS-related CoVs

ACU31044.1 SARS coronavirus Rs 672/2006
QDF43834.1 Coronavirus BtRs-BetaCoV/YN2018D
AID16712.1 Bat SARS-like coronavirus
ATO98167.1 BatSARS-like-coronavirus Rs4237
QDF43814.1 Coronavirus BtRl-BetaCoV/SC2018
73 AKZ19085.1 Bat SARS-like coronavirus YNLF 34C
QDF43829.1 Coronavirus BtRs-BetaCoV/YN2018C
AGZ48817.1 Bat SARS-like coronavirus Rs3367
61 QDF43824.1 Coronavirus BtRs-BetaCoV/YN2018B
AGZ48805.1 Bat SARS-like coronavirus RsSHC014
AIA62309.1 BtRs-BetaCoV/HuB2013
AVP78041.1 Bat SARS-like coronavirus-CoVZXC21
AVP78030.1 Bat SARS-like coronavirus-CoVZC45
32 ADE34811.1 Bat SARS coronavirus HKU3-12
60 79 ADE34754.1 Bat SARS coronavirus HKU3-7
87 ADE34778.1 Bat SARS coronavirus HKU3-9
54
ABG47059.1 Bat-CoV-273/2005
AIA62276.1 BtRf-BetaCoV/JL2012
94 ARI44808.1 Bat coronavirus Jiyuan-84
AIA62289.1 BtRf-BetaCoV/HeB2013
75 AIA62299.1 BtRf-BetaCoV/SX2013

100

APO40578 SARS-related coronavirus strain BtKY72
52 YP 003858583.1 Bat-coronavirus-BM48-31/BGR/2008

Sarbecovirus
Lineage-3

100 QIA48613.1 Pangolincoronavirus
QIA48622.1 Pangolin coronavirus PCoV GX-P1E
100 QHR63299.1 Batcoronavirus RaTG13
YP 009725307.1 SARS-Cov-Wuhan-Hu-1
100
98 QIU78833.1 SARS-Cov-2

SARS-Cov-2r cluster

AFE48811.1 Rabbit coronavirus HKU14
ABD75567.1 HumancoronavirusHKU1
AID16659.1 Longquan Rl rat coronavirus
61
AFO115140.1 Murine coronavirus
90
AFG25758.1 Rat coronavirus
100
90 YP 003029844.1 Rat coronavirus Parker

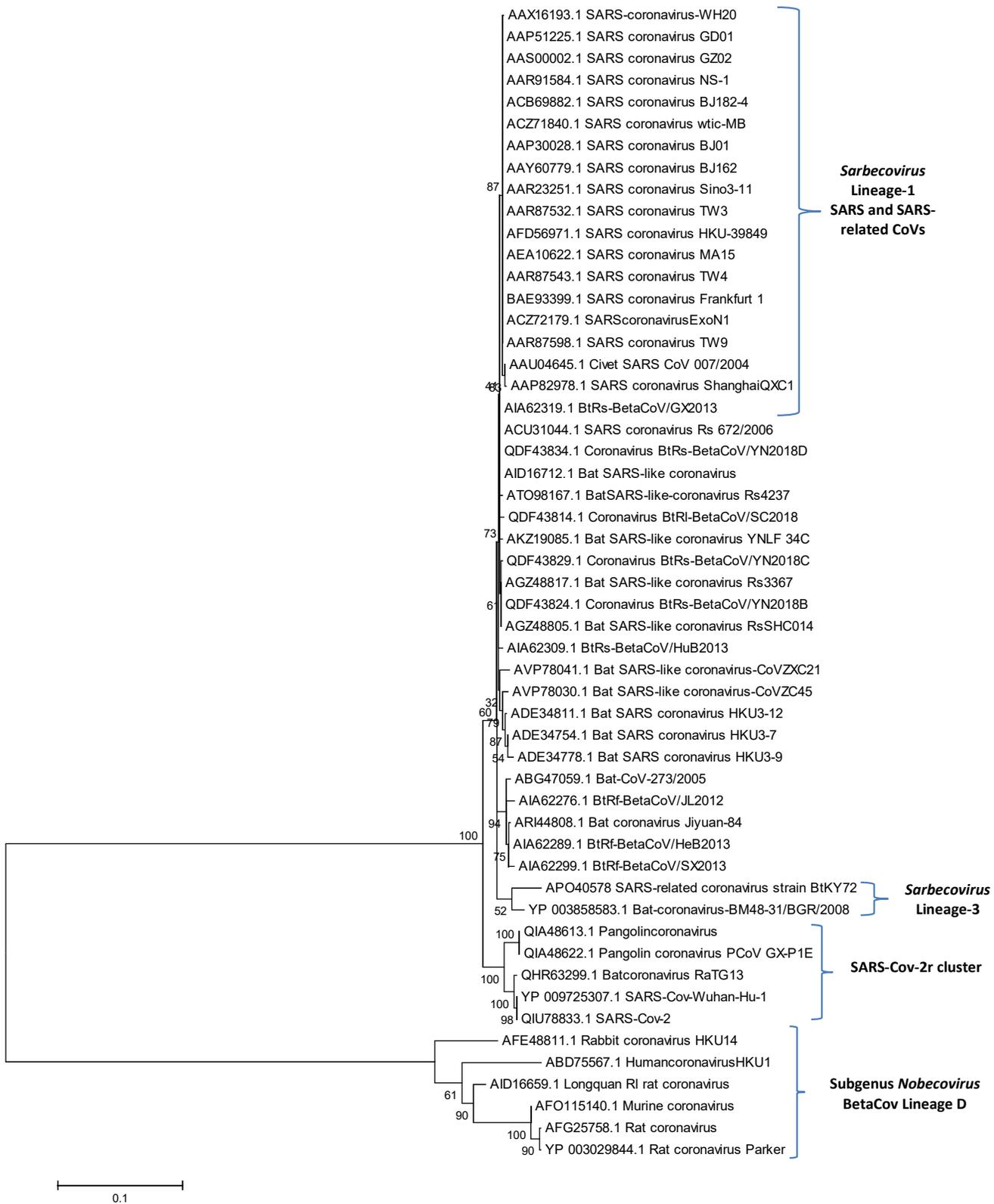**Subgenus *Nobecovirus***
BetaCov Lineage D

0.1

Figure S2. A maximum likelihood tree based on RNA dependent RNA polymerase protein sequences showing branching of different strains of Sarbecoviruses

Figure S1. A maximum likelihood tree based on spike protein sequences showing branching of different strains of Sarbecoviruses