|  |  |  |  |
| --- | --- | --- | --- |
| ClusterTable 1S-Nucleotide pairwise identity comparison among coronavirus lineages phylogenetically assigned to clusters and subclusters according to SplitTree4 software delineation. | Subgroup | Coronavirus species | Nucleotide sequences |
| Cluster I | % identity interval |  Most distant accessions |  Closest accessions |
| Subgroup I | SARS-CoV-1 | 89.99-100.00 | (MK062184-NC\_004718) ;(MK062182-NC\_004718) ; (MK062180- NC\_004718) ;(MK062184-AY310120) ; (MK062182-AY310120) ;(MK062180-AY310120) ;(AY291315-MK062184) ;(AY291315-MK062182) ;(AY291315- MK062180)  | (AY291315-AY310120) ;(JX163928-AP006561) ;(JX163924-AP006561) ;(AY291451-AP006561) ;(MK062182-MK062184) ;(MK062180-MK062184) ;(MK062181-MK062183) ;(MK062179-MK062183) ;KF514412-MK062183) ;(MK062180-MK062182) ;(MK062179-MK062181) ;(KF514412-MK062181) ;(KF514412-MK062179) ;(JX163924-JX163928) ;(AY291451-JX163928) ;(AY291451-JX163924) |
| Subgroup II | SARS-CoV-2 (TN) | 99.86-100.00 | (MT955171-MT955174) | (MT955168-MT955172) ;(MT955168-MT559037) ;(MT955168-MT499219) ;(MT955168-MT499220) ;(MT955168-MT365033) ;(MT955172-MT559037) ;(MT955172-MT499219) ;(MT955172-MT499220) ;(MT955172-MT365033) ;(MT559037-MT499219) ;(MT559037-MT499220) ;(MT559037-MT35033) |
|  |  |  |  |  |  |
| ClusterTable 1S continued | Subgroup | Coronavirus species | Nucleotide sequences |
| Cluster I | % identity interval |  Most distant accessions |  Closest accessions |
| Subgroup II | SARS-CoV-2 (TN) | 99.86-100.00 |  | (MT499219-MT499220) ;(MT499219-MT365033) ;(MT499220-MT365033) |
| Subgroup III | Animal SARS-CoV. +SARS-CoV-1+ SARS-CoV-2 | 26.76-99.92 | (MT040333-DQ022305)  | (MT559038-MT955170) |
| Cluster II | None | MERS-CoV. Homo sapiens and Camel | 99.40-99.97 | (MH013216-MG011360) ;(MH013216-MG011361)  | (MG011362-MG011360) ;(MG011362-MG011361) ;(KT806047-KT806053)  |
| Cluster III | Subgroup I | HCoV-229E | 40.35-100.00 | (NC\_002645-KY621348) ;(NC\_002645-KY674914) ;(AF304460-KY621348) ;(AF304460-KY674914)  | (KY674914-KY621348) ;(NC\_002645-AF304460)  |
| Subgroup II | HCoV-NL63 | 35.22-99.97 | (KF530105-MK334047) | (KF530109-KF530114) ;(KF530107-KF530114) ;(KF530113-KF530114) |
| Cluster IV | Subgroup I | HCoV-OC43 | 31.12-100.00 | (KF530098-KF530085) ;(KF530098-KF530090) ;(KF50098-KF530092) ;(KF50098-KF530097)  | KF530090-KF530097) ;(KF530095-KF530096)  |
| Subgroup II | HCoV-HKU1 | 54.47-100.00 | (KY674921-KF686340) | (KF686342-KF686346) ;(KY674942-KY674943) ;(KY674941-KY674943) ;(KY674941-KY674942) |

Table 2S Detection of putative recombination signatures by means of RECCO algorithm and determination of their positions and frequencies along the sequences of spike glycoprotein-coding gene of coronavirus lineages distributed in clusters and subclusters (MaxSavings ≥ 5.0, dataset *p*-val ≤ 1.0). Nucleotide numbering corresponds to the aligned sequences. Abbreviations: NRS: number of recombination sites.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Cluster | Subgroup | Coronavirusspecies | Accession | Position of the breakpoint in Spike glycoprotein-coding gene sequence | NRS | Length of breakpoint | Genomic position of the longest breakpoint (size in nt) |
|  |  |  |  | Start | End | Savings |  | 1 residue | 2 residues | 3 residues | >3 residues |
| Cluster I | Subgroup III | SARS-CoV-1 + SARS-CoV-2 + SARS-CoV in Animals | DQ0022305 | 274 | 276 | 9.0 | 23 | 5 | 5 | 9 | 4 | 2419-2430 (12) |
| 352 | 352 | 9.0 |
| 382 | 384 | 9.0 |
| 626 | 630 | 5.0 |
| 633 | 633 | 5.0 |
| 655 | 656 | 5.0 |
| 684 | 685 | 5.0 |
| 712 | 720 | 8.0 |
| 737 | 738 | 8.0 |
| 753 | 754 | 7.0 |
| 769 | 770 | 8.0 |
| 823 | 825 | 7.0 |
| 827 | 829 | 7.0 |
| 834 | 836 | 7.0 |
| 997 | 999 | 20.5 |
| 1034 | 1034 | 20.5 |
| 1081 | 1083 | 20.5 |
| 1108 | 1110 | 9.0 |
| 1404 | 1404 | 5.7 |
| 1461 | 1463 | 5.7 |
| 2419 | 2430 | 20.5 |
| 2783 | 2783 | 10.0 |
| 2935 | 2940 | 10.0 |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| ClusterTable 2S continued | Subgroup | Coronavirus species | Accession | Position of the breakpoint in Spike glycoprotein-coding gene sequence | NRS | Length of breakpoint | Genomic position of the longest breakpoint (size in nt) |
|  |  |  |  | Start | End | Savings |  | 1 residue | 2 residues | 3 residues | >3 residues |
| Cluster I | Subgroup III | SARS-CoV-1 + SARS-CoV-2 + SARS-CoV in Animals | KF367457 | 85 | 85 | 5.6 | 9 | 2 | 1 | 1 | 5 | 1373-1389 (17) |
| 268 | 268 | 10.4 |
| 510 | 522 | 5.5 |
| 587 | 588 | 5.5 |
| 1368 | 1371 | 6.5 |
| 1373 | 1389 | 6.5 |
| 1399 | 1401 | 6.5 |
| 1426 | 1437 | 6.5 |
| 1513 | 1520 | 6.5 |
| MT040333 | 168 | 177 | 5.0 | 14 | 2 | 2 | 2 | 8 | 2419-2475 (57) |
| 774 | 775 | 5.15 |
| 790 | 795 | 5.15 |
| 1259 | 1266 | 5.15 |
| 1318 | 1320 | 5.15 |
| 1331 | 1333 | 5.15 |
| 1353 | 1353 | 5.15 |
| 1992 | 1992 | 5.00 |
| 2085 | 2086 | 5.00 |
| 2419 | 2475 | 5.66 |
| 2575 | 2583 | 5.66 |
| 3473 | 3477 | 5.66 |
| 3481 | 3486 | 5.66 |
| 3658 | 3663 | 5.66 |

Table 2S continued

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Cluster | Subgroup | Coronavirus species | Accession | Position of the breakpoint in Spike glycoprotein-coding gene sequence | NRS | Length of breakpoint | Genomic position of the longest breakpoint (size in nt) |
|  |  |  |  | Start | End | Savings |  | 1 residue | 2 residues | 3 residues | >3 residues |
| Cluster III | Subgroup I | HCoV/229E | KY996417 | 1059 | 1059 | 8.40 | 1 | 1 | 0 | 0 | 0 | 1059-1059 (1) |
| KY674919 | 1384 | 1407 | 7.00 | 1 | 0 | 0 | 0 | 1 | 1384-1407 (24) |
| Subgroup II | HCoV/NL63 | MK334046 | 3617 | 3744 | 11.00 | 1 | 0 | 0 | 0 | 1 | 3617-3744 (128) |
| MK334044 | 2408 | 3408 | 5.00 | 1 | 0 | 0 | 0 | 1 | 2408-3408 (1001) |
| MK334043 | 2245 | 2298 | 5.00 | 2 | 0 | 0 | 0 | 2 | 3284-3408 (125) |
| 3284 | 3408 | 5.00 |
| KF530112 | 2245 | 2277 | 7.00 | 1 | 0 | 0 | 0 | 1 | 2245-2277 (33) |
| KF530105 | 731 | 924 | 8.00 | 1 | 0 | 0 | 0 | 1 | 731-924 (194) |
| Cluster IV | Subgroup I | HCoV/OC43 | KX334031 | 863 | 1218 | 5.00 | 2 | 0 | 0 | 0 | 2 | 863-1218 (356) |
| 1264 | 1544 | 5.00 |
| Subgroup II | HCoV/HKU1 | MK167038 | 1252 | 1253 | 30.5 | 2 | 0 | 1 | 0 | 1 | 1491-1500 (10) |
| 1491 | 1500 | 30.5 |

Table 3S-Detection of positively selected signatures at *p* < 0.1 across HCoV/NL63 (Cluster III, Subgroup II) and HCoV/HKU1 (Cluster IV, Subgroup II) coronavirus lineages using PARRIS algorithm where inferred rate distribution (synonymous rate and ω ratio) for the null (M1) and the alternative (M2) models, were determined. St. Dev.= Standard Deviation ; - = None.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Cluster | Subgroup | Coronavirus species |  | Synonymous rate | *d*N/*d*S ratio (ω) |
| Cluster III | Subgroup II | HCoV/NL63 | Inferred rate distributions | Rate class | 1 | 2 | 3 |  | Summary | Rate class | 1 | 2 | 3 |  | Summary |
| Null model(M1) | *d*S | 0.000 | 1.050 | 1.050 | Mean | 1.000 | ω | 0.350 | 1.000 | - | Mean | 0.353 |
| Probability | 0.049 | 0.758 | 0.192 | St.Dev. | 0.228 | Probability | 1.000 | 0.000 | - | St.Dev. | 0.000 |
| AlternativeModel (M2) | *d*S | 0.000 | 1.050 | 1.540 | Mean | 1.000 | ω | 0.350 | 1.000 | 7.610 | Mean | 0.379 |
| Probability | 0.049 | 0.951 | 0.000 | St.Dev. | 0.227 | Probability | 0.996 | 0.000 | 0.004 | St.Dev. | 0.485 |
| Cluster IV | Subgroup II | HCoV/HKU1 | Null model(M1) | *d*S | 0.04 | 0.16 | 1.620 | Mean | 1.000 | ω | 0.210 | 1.000 | - | Mean | 0.211 |
| Probability | 0.269 | 0.131 | 0.600 | St.Dev. | 0.756 | Probability | 1.000 | 0.000 | - | St.Dev. | 0.000 |
| AlternativeModel (M2) | *d*S | 0.020 | 0.360 | 1.380 | Mean | 1.000 | ω | 0.020 | 1.000 | 1.100 | Mean | 0.817 |
| Probability | 0.185 | 0.125 | 0.690 | St.Dev. | 0.571 | Probability | 0.186 | 0.814 | 0.000 | St.Dev. | 0.383 |