

Figure S1 Nucleotide sequence alignment of the 5' UTRs of the KNU16 isolates with reference PDCoV strains. A 7-nt insertion (dashed box) and the conserved ACACCA transcription regulatory sequence (TRS) involved in genome transcription and replication (solid box) are indicated

```

HKU15-155 ACATCAATCTATGTCTC---ACCCACCCAACCTCCACTCAAATATCTCTTTGGTTCCAGAGAGTCGTAGTGTATAGCCAGAGAGCCAGTCAGAGGGCGCT 100
HKU15-44  ....T.....GT TAA..... 100
HB        ....T.....GT TAA..... 100
AH2004    .....T---..... 100
NH        .....C.....GT TAA.....T.....A..... 100
IL2768    .....GT TAA..... 100
IN2847    .....GT TAA..... 100
KY4813    .....GT TAA.....A..... 100
KNU14-04  .....T.....GT TAA..... 100
KNU16-07  .....GT TAA..... 100
KNU16-11  .....GT TAA.....A.....C..... 100
**** *  ***** * ***** 100

```

```

HKU15-155 ATCATGCAAAC TAGGGCTGGCTACTCTAGCACAGAATCACATCCCGATAATCAACAGTGCTAGAAGGTTGATTATACCATTTAATATGCCAGGCCACGC 200
HKU15-44  ..... 200
HB        ..... 200
AH2004    ..... 200
NH        ..... 200
IL2768    ..... 200
IN2847    ..... 200
KY4813    ..... 200
KNU14-04  ..... 200
KNU16-07  ..... 200
KNU16-11  ..... 200

```

```

HKU15-155 GGAGTACGATCGAGGGTACAGCATAATCTCAACTTTTGTGGGCCACAATTTAATCCTAATTGGAGAAGGCCAAAGGACTGTACTACTTTTGTAGGTGT 300
HKU15-44  ..... 300
HB        ..... 300
AH2004    .....C.....GT TAA..... 300
NH        ..... 300
IL2768    .....G..... 300
IN2847    .....G..... 300
KY4813    .....G..... 300
KNU14-04  .....G..... 300
KNU16-07  .....G..... 300
KNU16-11  .T.....G..... 300
* * ***** 300

```

```

HKU15-155 AGCAGTCGCCAGTGGGAAAGCCCAACTAGGTTACAATTGTGGTGGGGACAAATTAGGGGAAATTAATTTGGCTTATAGGGGGATGGAGC 392
HKU15-44  ..... 392
HB        ..... 392
AH2004    ..... 392
NH        ..... 392
IL2768    ..... 392
IN2847    ..... 392
KY4813    ..... 392
KNU14-04  ..... 392
KNU16-07  ..... 392
KNU16-11  ..... 392

```

Figure S2 Nucleotide sequence alignment of the 3' UTRs of the KNU16 isolates with reference PDCoV strains. A 3-nt (TAA) insertion (solid box) at position 25,048 to 25,051 corresponding to the 3' UTR compared to the Hong Kong strain HKU15-155 is shown

| | | |
|-----------|---|-----|
| HKU15-155 | RQRLYVTAVEGQTHGTTSVHATDTSVVITDVCTDYIYGVSGTGIIKPSDLLLNHNGIAFTSPTGELYAFKNIITGKTLQVLPCETPSQLIVINNTVVGA | 500 |
| HKU15-44 | | 500 |
| HB | | 500 |
| AH2004 | | 500 |
| NH | | 500 |
| IL2768 | | 500 |
| IN2847 | | 500 |
| KY4813 | | 500 |
| KNU14-04 | | 500 |
| KNU16-07 |V..... | 500 |
| KNU16-08 |P.....K..... | 500 |
| KNU16-11 |P.....E.....K..... | 500 |

***** ** ***** **

| | | |
|-----------|--|-----|
| HKU15-155 | ITSSNSTENNRFTTTIVTPTFFYSTNATTFNCTKPVLSYGPISVCSGAIAGTSTLQNRPSIVSLYDGEVEIPSAFSLSVQTEYLQVQSEQVIVDCPQY | 600 |
| HKU15-44 | | 600 |
| HB |A..... | 600 |
| AH2004 |A..... | 600 |
| NH |A..... | 600 |
| IL2768 |V..... | 600 |
| IN2847 |T.....V..... | 600 |
| KY4813 |V..... | 600 |
| KNU14-04 |V..... | 600 |
| KNU16-07 |V..... | 600 |
| KNU16-08 |VEI..... | 600 |
| KNU16-11 |VEI..... | 600 |

***** ** ***** **

| | | |
|-----------|---|-----|
| HKU15-155 | VCNGNSRCLQLLAQYTSACSNIEAALHSSAQLDSREIINMFQTSTQSLQLANITNFKGDYNFSSILTRLGGRSAIEDLLFNKVVTSGLGTVDQDYKACS | 700 |
| HKU15-44 | | 700 |
| HB | | 700 |
| AH2004 |V.....S..... | 700 |
| NH |A..... | 700 |
| IL2768 |I.....S..... | 700 |
| IN2847 |R.....I.....S..... | 700 |
| KY4813 |V.....I.....S..... | 700 |
| KNU14-04 |I.....S..... | 700 |
| KNU16-07 |K.....I.....S..... | 700 |
| KNU16-08 |RI.....I.....S..... | 700 |
| KNU16-11 |RI.....I.....S..... | 700 |

***** ** ***** **

| | | |
|-----------|--|-----|
| HKU15-155 | RDMAIADLVCSQYNGIMVLPGVVDAEKMMYTGSLTGAMVFGGLTAAAAIPFATAVQARLNYVALQTNVLQENQKILAESFNQAVGNISLALSSVNDAI | 800 |
| HKU15-44 | | 800 |
| HB | | 800 |
| AH2004 | | 800 |
| NH | | 800 |
| IL2768 | | 800 |
| IN2847 | | 800 |
| KY4813 | | 800 |
| KNU14-04 | | 800 |
| KNU16-07 | | 800 |
| KNU16-08 | | 800 |
| KNU16-11 | | 800 |

***** ** ***** **

```

HKU15-155 QQTSEALNTVAIAIKKIQTVVNQGEALSHLTAQLSNFQAISTSIQDIYNRLEEVEANQQVDRLITGRLAALNAYVTQLLNQMSQVRSRLLAQQKINE 900
HKU15-44 .....I..... 900
HB .....I..... 900
AH2004 .....I..... 900
NH .....I..... 900
IL2768 .....I..... 900
IN2847 .....I..... 900
KY4813 .....I..... 900
KNU14-04 .....I..... 900
KNU16-07 .....I..... 900
KNU16-08 .....I..... 900
KNU16-11 .....I..... 900
*****

HKU15-155 CVKSQSSRYGFCGNGTHIFSLTQTAPNGIFFMHAVLVPNKFTRVNASAGICVDNTRGYSLQPQLILYQFNNSWRVTPRNMYPEPRLPRQADFIQLTDCSVT 1000
HKU15-44 .....P..... 1000
HB .....P..... 1000
AH2004 .....D..... 1000
NH .....I..... 1000
IL2768 .....I..... 1000
IN2847 .....I..... 1000
KY4813 .....I..... 1000
KNU14-04 .....I..... 1000
KNU16-07 .....I..... 1000
KNU16-08 .....I..... 1000
KNU16-11 .....I..... 1000
*****

HKU15-155 FYNTTAANLPNIIPDVIDVNTQTVSDIIDNLPATATPPQWDVGIYNNITILNLTVEINDLQERSKNLSQIADRLQNYIDNLTNTLVLDLEWLNRVETYLKWPWY 1100
HKU15-44 .....I..... 1100
HB .....I..... 1100
AH2004 .....I..... 1100
NH .....I..... 1100
IL2768 .....I..... 1100
IN2847 .....I..... 1100
KY4813 .....I.....E..... 1100
KNU14-04 .....I..... 1100
KNU16-07 .....I..... 1100
KNU16-08 .....I.....D..... 1100
KNU16-11 .....I.....D..... 1100
*****

HKU15-155 IWLAIALALIAFVTILITIFLCTGCCGGCGFCGGCGFLFSKKRSTDDQPTPSFKFKEW 1159
HKU15-44 .....Y..... 1160
HB .....Y..... 1159
AH2004 .....Y..... 1160
NH .....Y..... 1159
IL2768 .....Y..... 1160
IN2847 .....Y..... 1160
KY4813 .....Y..... 1160
KNU14-04 .....Y..... 1160
KNU16-07 .....Y..... 1160
KNU16-08 .....S.....-Y.....N..... 1159
KNU16-11 .....S.....-Y.....N..... 1159
*****

```

Figure S3 Amino acid alignment of the complete S glycoprotein genes among PDCoV strains. Putative signal sequences (dotted line box) are indicated and potential *N*-glycosylation sites predicted by the software tool NetNGlyc 1.0 (<http://www.cbs.dtu.dk/services/NetNGlyc/>) are shown in boldface type. A 1-aa (asparagine, N) insertion (solid box) in the S genes of the KNU16 series and other reference strains is indicated. The dashes (-) indicate deleted sequences. A 1-aa (lysine, K) deletion (shaded box) in the S genes of KNU16-08 and KNU16-11 is indicated, and the amino acid changes identified in these two KNU16 strains are shaded

Table S1 Results of retrospective detection of PDCoV in 10 provinces from 2014 to 2016

| Region | 2014 | | 2015 | | 2016 | |
|-----------|-----------------------|------------------|----------------|-----|----------|-----|
| | <i>n</i> ^a | (+) ^b | <i>n</i> | (+) | <i>n</i> | (+) |
| Gyeonggi | 32 | 1 | 18 | 7 | 36 | 7 |
| Gangwon | 1 | 1 | 0 ^c | 0 | 0 | 0 |
| Daejeon | 45 | 7 | 38 | 6 | 25 | 4 |
| Chungbuk | 15 | 7 | 19 | 4 | 12 | 5 |
| Chungnam | 69 | 5 | 69 | 19 | 69 | 15 |
| Gyeongbuk | 17 | 2 | 28 | 6 | 40 | 2 |
| Gyeongnam | 18 | 4 | 15 | 6 | 29 | 2 |
| Jeonbuk | 20 | 6 | 16 | 4 | 9 | 0 |
| Jeonnam | 7 | 1 | 5 | 1 | 4 | 0 |
| Jeju | 5 | 0 | 7 | 2 | 15 | 6 |
| Total | 229 | 34 | 215 | 55 | 239 | 41 |

^a Number of diarrheal fecal/intestinal samples.

^b Number of positive samples.

^c No samples collected.

Table S2 List of primers used in this study

| Primer name | Nucleotide sequence (5'–3') | Purpose | Location (nt) |
|-------------|--------------------------------|-------------|---------------|
| 1F-Fwd | ACATGGGGACTAAAGATAAAAATTATAGCA | PCR | 1–30 |
| 1F-Rev | TTGGTTAAGCTCATATGCAGCGAGT | RT and PCR | 2390–2414 |
| 2F-Fwd | GAGCTCTGGAGGCTGTTGATGTCG | PCR | 2200–2223 |
| 2F-Rev | CGTACGACTCACAAACCCATTTATGG | RT and PCR | 4070–4094 |
| 3F-Fwd | GTGTGGCTCCTATCTCAGGACC | PCR | 3541–3562 |
| 3F-Rev | CTTATACTCGACGTGCGGAAGC | RT and PCR | 7142–7163 |
| 4F-Fwd | CCAGCAGGTGTCATTCATAGAGACG | PCR | 6843–6867 |
| 4F-Rev | TTCGAAGACTTTAGCTGCACACTGCTTAGT | RT and PCR | 11615–11644 |
| 5F-Fwd | GGTATGGTTATTGGCCACATCTCTG | PCR | 10956–10980 |
| 5F-Rev | GCATCAAGGATACCGTCCTGTAAC | RT and PCR | 14068–14091 |
| 6F-Fwd | GCTTGCTAGCACGAAAACACACTTG | PCR | 13299–13323 |
| 6F-Rev | GGTCCAATTAATGCTTCTGCCG | RT and PCR | 17183–17205 |
| 7F-Fwd | GGTATCCCTATCGTGCGCAAAC | PCR | 16991–17012 |
| 7F-Rev | GTTAACGTTCTTAGTGGGGTAGCC | RT and PCR | 19501–19524 |
| 8F-Fwd | GGGCGCCTTCTTGTGAGGCCAC | PCR | 19298–19319 |
| 8F-Rev | GTCCAGGGATGGTAACGGCCC | RT and PCR | 22816–22836 |
| 9F-Fwd | GGCTTGCCATTGCCCTGGCTC | PCR | 22628–22648 |
| 9F-Rev | GCTCCATCCCCCTATAAGCC | RT and PCR | 25401–25422 |
| 5'RACE-R1 | GGCAGTATCACGAGCTAACTCAGCAAG | 5'-RACE-PCR | 969–995 |
| 5'RACE-R2 | CTCCTATGGGTATCACACTTGCGCAAC | 5'-RACE-PCR | 469–495 |
| 3'RACE-F1 | GCCCAGCTCAAGGTTTCAGAGTTG | 3'-RACE-PCR | 24399–24422 |
| 3'RACE-F2 | CTGCTAAACCCAAACAGCAGAAGAAACC | 3'-RACE-PCR | 24921–24948 |
| 543F | GCCAAGAACAAGTCCAAGCGCGAC | Sequencing | 544–566 |
| 1094F | GGTTGTTGGCGTAACTTTTACTGCC | Sequencing | 1094–1118 |
| 1649F | CCTTAAATGCTTTGAAACCCCTGAG | Sequencing | 1649–1673 |
| 2753F | GGAACGGAAGACCAGCAGCGACG | Sequencing | 2753–2775 |
| 4089F | CGTACGAACGGGTTGATGCATCTAAAACAA | Sequencing | 4089–4118 |
| 4639F | CTCAACTTGTCGAGGACGTGGTTGC | Sequencing | 4639–4663 |
| 5173F | CTGGCTGGCCTTGTAAATTACAATGC | Sequencing | 5173–5197 |
| 5743F | CGCTCTCACGAGCTACCAGGCTTAG | Sequencing | 5743–5767 |
| 6293F | CAATGCAGAAGACCAGTCTACCATT | Sequencing | 6293–6317 |
| 6941F | GTCACTCAAGACAGCAAATCTCTTTAAC | Sequencing | 6941–6968 |
| 7394F | GGGCACTGTATCTGGTATTCACATC | Sequencing | 7394–7418 |
| 8155F | GGCTTAAGAATGTTGTCTATTGCCAC | Sequencing | 8155–8181 |
| 8743F | GTATCGAGGATTTCAACTCATGGGC | Sequencing | 8743–8767 |
| 9288F | GGTTACACTGTTGCTCTTTATATGCTG | Sequencing | 9288–9314 |
| 9843F | GCTGTTGTTGTTGCTAACCTTCTTG | Sequencing | 9843–9867 |
| 10404F | GATAGGGTAAAAGCTCTGTTTGAGTG | Sequencing | 10404–10429 |
| 11639F | TTCGAAATCGAGGAGGAATGCTACAACGCT | Sequencing | 11639–11668 |
| 12190F | CCCCTTGAGTATGATGGATTCAG | Sequencing | 12190–12214 |
| 12741F | CTATCACAGACTACAGCTACTACAG | Sequencing | 12741–12765 |
| 13873F | CTTACCTTACCCAGATGTCTCTCG | Sequencing | 13873–13896 |
| 14421F | GCATACCCATTGTAGCTAATGGTTC | Sequencing | 14421–14445 |
| 14993F | GATAGCTATGCAGAGGGTTACAACAG | Sequencing | 14992–15017 |
| 15566F | GCGGCGAAACCGAACTCAAGACAG | Sequencing | 15566–15589 |
| 16083F | GTGAGGAGCCACTTGCTAAATTAG | Sequencing | 16083–16106 |
| 17542F | CAACTGTTATTATAAGAAGTCAATG | Sequencing | 17542–17567 |
| 18145F | GCAACATCAATTGCAGGAAGTGGGTG | Sequencing | 18145–18170 |
| 18646F | CAACTACATAGTTAAAAAGGATAGG | Sequencing | 18646–18670 |

| | | | |
|--------|--------------------------------|------------|-------------|
| 19858F | CTTTATTGCTTTACTAATAACACCTAC | Sequencing | 19858–19884 |
| 20413F | GACCTGTCAGCAGTAAACAATGGCATG | Sequencing | 20413–20439 |
| 20965F | GGTGCAATTGTGGGAACATCCAC | Sequencing | 20965–20987 |
| 21259F | CAGTCCTTGCAGTTAGCTAATATTACCAAC | Sequencing | 21259–21288 |
| 21933F | GGCTGCACTTAATGCATATGTTACTC | Sequencing | 21933–21958 |
| 23032F | GAAACACTCACTATGTCTGACGCAG | Sequencing | 23032–23056 |
| 23583F | CAACGATCCAGAATTTGCTGTTCTG | Sequencing | 23583–23607 |
| 24131F | GAAAACCATGGCTATTGGCTGCGTT | Sequencing | 24131–24155 |

Table S3 The number of nucleotide and amino acid changes between the global strains and KNU16-07 or KNU16-11.

| Encoded proteins | No. of nucleotide (amino acid) changes | | | | | | |
|------------------|--|-------------------------|-------------------|---------------------|-------------------------|-------------------------|-------------------------|
| | HKU15-155 /China/2012 | HKU15-44 /China/2012 | NH /China/2016 | IN2847 /USA/2014 | KNU14-04 /Korea/2014 | KNU16-07 /Korea/2016 | KNU16-11 /Korea/2016 |
| 5' UTR | 13/17 | 9/16 | 3/10 | 1/8 | 0/7 | 7 | 7 |
| ORF1ab | 128(27)/205(42) | 203(34)/278(48) | 163(27)/236(42) | 24(7)/103(22) | 30(6)/97(21) | 105(27) | 105(27) |
| S | 56(1)/80(27) | 49(15)/73(22) | 64(18)/90(27) | 14(3)/34(16) | 10(8)/34(17) | 44(25) | 44(25) |
| E | 1(0)/2(1) | 1(0)/2(1) | 2(0)/3(1) | 0(0)/1(1) | 0(0)/1(1) | 1(1) | 1(1) |
| M | 6(1)/5(1) | 8(0)/7(0) | 6(0)/5(0) | 1(0)/0(0) | 2(0)/1(0) | 1(0) | 1(0) |
| NS6 | 2(0)/3(1) | 1(1)/2(2) | 4(0)/5(1) | 1(1)/2(2) | 0(0)/1(1) | 1(1) | 1(1) |
| N | 10(1)/17(2) | 10(0)/17(1) | 11(0)/18(1) | 3(0)/10(1) | 5(0)/12(1) | 9(1) | 9(1) |
| NS7 | 5(4)/10(8) | 5(5)/10(9) | 6(6)/11(10) | 2(2)/7(6) | 3(3)/8(7) | 5(4) | 5(4) |
| 3' UTR | 5/8 | 3/6 | 3/6 | 0/3 | 1/4 | 3 | 3 |