Background: Coronaviruses (CoVs) can be classified into alphacoronavirus (group 1), betacoronavirus (group 2), and gammacoronavirus (group 3) based on diversity of the protein sequences. Their 3C-like protease (3CLpro), which catalyzes the proteolytic processing of the polyproteins for viral replication, is a potential target for anti-coronaviral infection.

Methodology/Principal Findings: Here, we profiled the substrate specificities of 3CLpro from human CoV NL63 (group 1), human CoV OC43 (group 2a), severe acute respiratory syndrome coronavirus (SARS-CoV) (group 2b) and infectious bronchitis virus (IBV) (group 3), by measuring their activity against a substrate library of 19X8 of variants with single substitutions at P5 to P3' positions. The results were correlated with structural properties like side chain volume, hydrophobicity, and secondary structure propensities of substituting residues. All 3CLpro prefer Gin at P1 position, Leu at P2 position, basic residues at P3 position, small hydrophobic residues at P4 position, and small residues at P1' and P2' positions. Despite 3CLpro from different groups of CoVs share many similarities in substrate specificities, differences in substrate specificities were observed at P4 positions, with IBV 3CLpro prefers P4-Pro and SARS-CoV 3CLpro prefers P4-Val. By combining the most favorable residues at P3 to P5 positions, we identified super-active substrate sequences ‘VARLQ SGF’ that can be cleaved efficiently by all 3CLpro with relative activity of 1.7 to 3.2, and ‘VPRLQ SGF’ that can be cleaved specifically by IBV 3CLpro with relative activity of 4.3.

Conclusions/Significance: The comprehensive substrate specificities of 3CLpro from each of the group 1, 2a, 2b, and 3 CoVs have been profiled in this study, which may provide insights into a rational design of broad-spectrum peptidomimetic inhibitors targeting the proteases.
three to four of their own autocleavage sequences have been measured by Hegyi et al. [27]. Comprehensive study on substrate specificities of group 1, 2a and 3 3CL\textsuperscript{pro} is lacking. Here, we profiled the substrate specificities of selected 3CL\textsuperscript{pro} from group 1, 2a, 2b and 3 CoVs. Activities of 3CL\textsuperscript{pro} from HCoV-NL63 (group 1), HCoV-OC43 (group 2a), SARS-CoV (group 2b) and IBV (group 3) against a substrate library of 19 x 8 variants were measured by fluorescence resonance energy transfer (FRET) assay [26]. Similarities and differences in substrate specificities among different 3CL\textsuperscript{pro} are discussed.

Results

Profiling substrate specificities of 3CL\textsuperscript{pro} from group 1, 2a, 2b, and 3 CoVs

We have previously created a 19 x 8 substrate library by performing saturation mutagenesis at P5 to P3' positions on the wild type (WT) sequence (SAVLQ\textsuperscript{Q}SGF), which corresponds to the autocleavage sequence at the N-terminus of SARS-CoV 3CL\textsuperscript{pro} [26]. The values of k\textsubscript{obs}/[3CL\textsuperscript{pro}] of the proteases against this WT sequence were 443 ± 11, 124 ± 13, 180 ± 5 and 174 ± 19 mM\textsuperscript{-1} min\textsuperscript{-1} for HCoV-NL63 (group 1), HCoV-OC43 (group 2a), SARS-CoV (group 2b), and IBV (group 3), respectively. That all proteases can cleave the WT sequence efficiently justifies that we can use our substrate library to profile the substrate specificities of 3CL\textsuperscript{pro} from other groups of CoVs. Based on the FRET assay we developed, we measured the activities of 3CL\textsuperscript{pro} from HCoV-NL63, HCoV-OC43, SARS-CoV and IBV against the 19 x 8 substrate variants (Figure 1, Table S1) [26].

Differences in substrate specificities among 3CL\textsuperscript{pro}

We then tested if the relative activities of 3CL\textsuperscript{pro} from any CoV strains were significantly different from the other by analysis of variance. Substitutions that resulted in significantly higher relative activities (p < 0.001) were indicated as filled symbol in Figure 1. IBV 3CL\textsuperscript{pro} (Figure 1, triangles) was the most efficient in cleaving A4P and A4F with relative activities of 1.09 ± 0.24 and 0.58 ± 0.14, respectively, while SARS 3CL\textsuperscript{pro} (Figure 1, diamonds) preferred A4V with relative activity of 1.39 ± 0.19. HCoV-OC43 3CL\textsuperscript{pro} (Figure 1, squares) appeared to be the most versatile in accepting substitutions at P1 and P2 positions, and could cleave Q1H, Q1M, L2M and L2C, significantly better than 3CL\textsuperscript{pro} from other strains. No significant differences were observed for other substitutions, suggesting that 3CL\textsuperscript{pro} from different CoVs shares many similarities in substrate preferences.

Substrate preferences that are common to all 3CL\textsuperscript{pro}

The most preferred P1 residue is Gln (Figure 1), which forms hydrogen-bonds with the side-chain of an invariant His residue and the backbone carbonyl group of an invariant Phe residue (His-163 and Phe-140 in SARS-CoV 3CL\textsuperscript{pro}) in the P1 binding pocket. Interestingly, our results showed that 3CL\textsuperscript{pro} from all groups of CoVs can cleave His at P1 position reasonably well. The relative activities for 3CL\textsuperscript{pro} from HCoV-NL63, HCoV-OC43, SARS-CoV, and IBV were 0.26 ± 0.08, 0.47 ± 0.08, 0.19 ± 0.03 and 0.23 ± 0.12, respectively (Table S1). Consistent with this observation, His is found natively at P1 positions in the polyproteins from group 1 and 2a CoVs (Table S3). Taken together, the ability to cleave His at P1 position is a conserved property for all 3CL\textsuperscript{pro}. Moreover, we showed that all 3CL\textsuperscript{pro} can cleave Q1M, albeit at
an even lower rate, and all other substitutions resulted in undetected activity.

The protease activities correlate positively with the hydrophobicity of substituting residues at P2 position (Figure 2). In fact, among the P2 variants, only L2M, L2C, L2F, L2I and L2V were cleavable, suggesting that P2 position favors hydrophobic residues. However, substitution with β-branched residues, Val or Ile, led to 10-folds decreases in the activity (Figure 1, Table S1). Considering that Leu, Val and Ile share similar hydrophobicity and side chain volume, the large differences in activities suggest that β-branched residues are not preferred in all 3CL pro, probably due to steric clashes with the P2 binding pocket. Taken together, P2 position prefers hydrophobic residues without β-branch, and the most preferred residue is Leu.

At P3 position, the protease activities on Arg/Lys-substituting variants were 5 to 14 fold higher than that on Asp/Glu-substituting variants (Figure 1, Table S1). This observation suggests that P3 position prefers positively charged residues over negatively charged one. In the active site of 3CL pro, there is no substrate-binding pocket for P3 residue. Molecular modeling showed that there is an invariant Glu residue (Glu-166 in SARS-CoV 3CL pro) in the active site of 3CL pro that may form favorable charge-charge interactions with a positively charged residue at the P3 position, which may explain why Arg/Lys are favored over Asp/Glu at this position [31]. Moreover, no cleavage was observed for substrate containing Pro-substitution at P3 position.

The protease activities correlate negatively with side chain volume, and positively with the hydrophobicity of substituting residues at P4 position (Figure 2). The correlations with hydrophobicity were more evident (with correlation coefficients >0.89) when only small residues (Ala, Asn, Asp, Cys, Gly, Ser, and Thr) with side chain volumes <70 Å³ (Figure 3) were included in the analysis. This result suggests that as long as the side chain can fit into the P4 binding pocket, the protease activity is directly proportional to the hydrophobicity of the substituting residues. On the other hand, charged residues like Lys, Arg, His, Asp and Glu were not cleavable, presumably due to the unfavorable burial of charges in the hydrophobic P4 pocket.

In general, the activities of 3CL pro correlate positively with the hydrophobicity and β-sheet propensity of substituting residues at P5 position (Figure 2). The correlations are significant (p<0.05) for group 2a, 2b, and 3 CoVs, but are weaker for group 1 CoV. Like the P3 position, there is no substrate-binding pocket for P5 residue. In the crystal structure of SARS-CoV 3CL pro in complex with a peptide substrate, the P5 residue adopts an extended β-strand conformation to avoid clashing of P5-P6 residues with the protease [31]. Residues with high β-sheet propensity may stabilize the extended conformation at P5 and improve enzyme-substrate interaction. As shown in Figure 1, a number of substitutions at P5 position resulted in a substrate better than the WT sequence (i.e. with relative activity >1). Consistent with the suggestion that P5 position favors residues with high hydrophobicity and β-sheet

Figure 2. Correlation between 3CL pro activities and structural properties of substituting residues. The relative protease activities of 3CL pro from HCoV-NL63 (shaded, group 1), HCoV-OC43 (white, group 2a), SARS-CoV (black, group 2b) and IBV (grey, group 3), were correlated with structural properties of substituting residue properties, including side chain volume [28], hydrophobicity [29] and α-helix and β-sheet propensities [30]. Correlation coefficients of +/-0.56 and +/-0.44 correspond to p-values of 0.01 and 0.05 respectively. doi:10.1371/journal.pone.0027228.g002
propensity, Val-substitution consistently yielded substrates with higher than WT activities for all 3CLpro. On the other hand, negatively charged residues (Asp/Glu) were not favored at P5 position, with significantly lower activities (0.16 to 0.50).

At P1’ position, the protease activities correlate negatively with the side chain volume of substituting residues (Figure 2). In fact, the relative activities for substrates with the smallest residues (Gly, Ala, Ser, and Cys) at P1’ position were in the range of 0.64 to 1.40, which were consistently higher than those for other larger residues (Figure 1). At P2’ position, all variants, except G2’P, could be cleaved with relative activities of 0.17 to 1.04 (Figure 1). The protease activities also correlate negatively with the side chain volume (Figure 2), but the difference in the protease activities was relatively small (Figure 1). At P3’ position, no obvious substrate preference was observed.

The effect of combining multiple favorable substitutions

Our profiling analysis showed that all CoV 3CLpro prefer P5-Val and P3-Arg (Figure 1). To test if we can combine two favorable substitutions to create a more active substrate, we have created a doubly-substituted substrate variant ‘VARLQ ↓ SGF’. The protease activities of HCoV-NL63, HCoV-OC43, SARS-CoV and IBV against the doubly-substituted sequence were 1.70±0.07, 1.87±0.17, 1.70±0.12 and 3.24±0.37, respectively (Table 1). The results suggest that the increase in activity is additive, and the sequence ‘VARLQ ↓ SGF’ can represent a good broad-spectrum substrate for all 3CLpro.

On the other hand, our profiling analysis suggests that 3CLpro from SARS-CoV and IBV have different substrate preferences at P4 position – SARS-CoV prefers P4-Val (relative activity = 1.09±0.24) while IBV prefers P4-Pro (relative activity = 1.39±0.10 (Table 1). To see if we can exploit this distinct substrate preference at P4 position to create a substrate more specific for IBV 3CLpro, we have created the triply-substituted variant ‘VPRLQ ↓ SGF’. The protease activity of IBV 3CLpro against this sequence was boosted to 4.33±0.98, while that of the other strains were significantly reduced, demonstrating that this substrate sequence can represent a specific substrate-sequence for IBV 3CLpro (Table 1). Similarly, the protease activity of SARS-CoV 3CLpro against the triply-substituted sequence ‘VVRLQ ↓ SGF’ was boosted to 2.50±0.51, while that of the other strains were reduced (Table 1). Taken together, these results suggest that one can combine the substrate preference profiled in this study to create a better substrate sequences.

Discussion

This study provides the first comprehensive profiling of substrate specificities of 3CLpro from group 1, 2a, and 3 CoVs. We showed that the substrate specificities of these 3CLpro share many similarities to those of 3CLpro from SARS-CoV (group 2b) reported previously by us [26]. Table 2 summarizes the substrate specificities that are common among all 3CLpro.

**Table 1.** 3CLpro activities against doubly- and triply-substituted substrate variants. WT substrate was substituted at P3 to P5 positions to generate doubly- and triply-substituted variants. The relative activities of 3CLpro on these substrate variants are reported.

| Variant sequence | HCoV-NL63 | HCoV-OC43 | SARS-CoV | IBV |
|------------------|-----------|-----------|----------|-----|
| VAVLQ ↓ SGF      | 1.23±0.40 | 1.55±0.30 | 1.80±0.31 | 1.58±0.27 |
| SARLQ ↓ SGF      | 1.14±0.24 | 1.36±0.17 | 0.97±0.12 | 1.72±0.22 |
| VARTQ ↓ SGF      | 1.70±0.07 | 1.87±0.17 | 1.70±0.17 | 3.24±0.37 |
| SPVLQ ↓ SGF      | 0.06±0.01 | 0.29±0.07 | 0.61±0.10 | 1.09±0.24 |
| VPRRLQ ↓ SGF     | 0.13±0.04 | 0.91±0.12 | 0.99±0.12 | 4.33±0.98 |
| SVALQ ↓ SGF      | 0.26±0.10 | 0.59±0.07 | 1.39±0.19 | 0.59±0.09 |
| VVRRLQ ↓ SGF     | 1.23±0.06 | 0.60±0.05 | 1.97±0.19 | 0.86±0.05 |
| VVPRLQ ↓ SGF     | 1.63±0.07 | 0.55±0.04 | 2.50±0.51 | 2.19±0.13 |

**Table 2.** Summary of substrate specificities that are common among all 3CLpro.

| Position | Substrate preferences |
|----------|-----------------------|
| P5       | No strong preference  |
| P4       | Small hydrophobic residues |
| P3       | Positively charged residues |
| P2       | High hydrophobicity and absence of β-branch |
| P1'      | Small residues |
| P2'      | Small residues |
| P3'      | No strong preference |

doi:10.1371/journal.pone.0027228.t001
doi:10.1371/journal.pone.0027228.t002
specificities that are common to all 3CL\textsuperscript{pro}. Although the substrate specificities for 3CL\textsuperscript{pro} from different groups of CoVs share a number of similarities, unique substrate preferences were identified in this study. In particular, we showed that only IBV 3CL\textsuperscript{pro}, but not other proteases, prefers P4-Pro (Figure 3).

To understand the structural basis of this unique substrate preference, we compared the structures of IBV 3CL\textsuperscript{pro} with other coronaviral 3CL\textsuperscript{pro}. We noticed that strand-11 of IBV 3CL\textsuperscript{pro} is positioned further away from the P4 and P5 substrate-binding site compared to other 3CL\textsuperscript{pro} (Figure 4) [31,32,33]. This results in a wider substrate-binding pocket in IBV 3CL\textsuperscript{pro}. We further docked the substrate variant A4P into the substrate-binding pocket of IBV 3CL\textsuperscript{pro}. Due to the cyclic structure of Pro residue, the backbone O dihedral angle of the P4 residue is restrained to ca. −60°, which causes the substrate peptide to bend towards the strand-11 of 3CL\textsuperscript{pro}. Such conformation of substrate is much better accommodated by IBV 3CL\textsuperscript{pro}, which has a wider substrate-binding pocket near the P4 and P5 positions. This observation justifies why only IBV 3CL\textsuperscript{pro} cleaves P4-Pro efficiently.

Similarities in substrate specificity suggest that it is feasible to create a broad-spectrum inhibitor that targets all 3CL\textsuperscript{pro}. A broad-spectrum inhibitor is desirable for a first line defense against coronaviral infection because CoVs are capable of generating novel strains with high virulence through high frequency of mutations and recombination [34,35,36,37]. Based on the autocleavage sequence of SARS-CoV 3CL\textsuperscript{pro} (i.e. AVLQ\textsuperscript{↓}), Rao and co-workers designed broad-spectrum peptidomimetic inhibitors that can inhibit 3CL\textsuperscript{pro} from different groups of CoVs [20]. Their results are consistent with our observation that the autocleavage sequence of SARS-CoV 3CL\textsuperscript{pro} can be well cleaved by all 3CL\textsuperscript{pro}. The substrate preferences profiled in this study will provide a rational basis to improve the broad-spectrum 3CL\textsuperscript{pro} inhibitors. For example, by combining favorable substitutions at P3 to P5 positions, we identified a substrate sequence ‘VARLQ↓SGF’ that can be cleaved with high relative activities by 3CL\textsuperscript{pro} from all groups of CoVs (Table 1). This substrate sequence may serve as a good starting point of the design of broad-spectrum peptidomimetic inhibitors for 3CL\textsuperscript{pro}.

Although it is generally accepted that substrate specificity provides insights into the design of peptidomimetic protease inhibitors, there are exceptions to the dogma that good peptidomimetic inhibitors should be derived from good substrate sequences. For example, Hilgenfeld and co-workers showed that the P2 position of peptide aldehyde inhibitors can accommodate aspartate or serine, which are poor substrates for SARS-CoV 3CL\textsuperscript{pro} [38].

In the FRET assay developed by us, all 3CL\textsuperscript{pro} can efficiently cleave the WT sequence of ‘SAVLQ↓SGF’ with activity of 120–440 mM\textsuperscript{–1} min\textsuperscript{−1}, and the activity can be further improved by 1.7 to 3.2 fold using the substrate sequence of ‘VARLQ↓SGF’. Because the substrate sequences can be cleaved by all 3CL\textsuperscript{pro} with high efficiency, one could use the FRET assay to screen for broad-spectrum inhibitors targeting 3CL\textsuperscript{pro} from all groups of CoVs.

Materials and Methods

Cloning, Expression and Purification of 3CL\textsuperscript{pro} and the Substrate Library

Cloning, expression and purification of SARS-CoV 3CL\textsuperscript{pro} and HCoV-229E 3CL\textsuperscript{pro} were described previously [26]. Codon-optimized DNA sequences encoding HCoV-NL63 (GenBank AY567487) and HCoV-OC43 (GenBank AAXS5666), and IBV (GenBank M95169) 3CL\textsuperscript{pro} were purchased from Mr. Gene (http://mrgene.com). The coding sequences of 3CL\textsuperscript{pro} from HCoV-NL63, HCoV-OC43 and IBV were sub-cloned and expressed in E. coli strain BL21 (DE3) pLYS S as fusion proteins with N-terminal tags of poly-histidine-small ubiquitin-related modifier (His\textsubscript{6}-SUMO) or poly-histidine-maltose binding protein (His\textsubscript{6}-MBP). Protein expression was induced by addition of 0.1 mM of isopropyl β-D-thiogalactopyranoside. After overnight incubation at 25°C, cells were harvested by centrifugation and resuspended in buffer A (20 mM Tris, pH 7.8, 150 mM NaCl and 1 mM tris(2-carboxyethyl)phosphine) with 30 mM imidazole and disrupted by sonication. Soluble fraction was subject to immobilized metal ion affinity chromatography for purification as described for SARS-CoV 3CL\textsuperscript{pro} [26]. The His\textsubscript{6}-SUMO or His\textsubscript{6}-MBP tags were removed by protease digestion using sentrin-specific protease 1 or factor Xa, respectively, followed by immobilized metal ion affinity chromatography. Native 3CL\textsuperscript{pro} were finally purified by G75 size exclusion column and stored in buffer A. Elution profiles of size exclusion chromatography indicated that all 3CL\textsuperscript{pro} purified were dimeric.

The construction, expression and purification of the substrate library were described previously [26]. In brief, the WT substrate sequence ‘TSAVLQ↓SGFRKM’ was inserted between the cyan fluorescent protein and the yellow fluorescent protein to create the substrate protein. Saturation mutagenesis was performed at each of the P5 to P3’ positions to generate a substrate library of 19 x 8 variants.

FRET assay for 3CL\textsuperscript{pro} activity measurement

The protease activity of 3CL\textsuperscript{pro} was measured by the FRET assay we developed previously [26]. Purified 3CL\textsuperscript{pro} at 0.2 to 2 μM were mixed with 35 μM of the substrate protein in buffer A. Cleavage of the substrate protein leads to a decrease in fluorescence at 530 nm when the reaction mixture was excited at 430 nm. The fluorescence intensity, monitored by EnVision 2101 Multilabel Plate Reader, was fitted to single exponential
decay to obtain the observed rate constant \( k_{\text{obs}} \). The protease activity against variant substrates was normalized against the WT activity to yield the relative activity. The assay was repeated in triplicate.

**Correlation analysis**

Structural properties of substituting residues, including side chain volume [28], hydrophobicity [29], and \( \alpha \)-helix and \( \beta \)-sheet propensities [30], were correlated with relative activity to determine correlation coefficients \((r)\) and \(p\)-values.

**Supporting Information**

Figure S1 Molecular modeling showing P3-Arg may interact with Glu-166 of 3CL\(_{pro}\). The model was based on the crystal structure of 3CL\(_{pro}\) (grey) in complex with a peptide substrate ‘TSAVLQ \( \downarrow \) SGFRK’ (yellow). P3-Val was replaced by P3-Arg using the program PyMOL. As shown, the invariant Glu-166 is in close proximity to P3-Arg, and may form favorable charge-charge interaction to P3-Arg. (TIF)

**References**

1. van der Hoek L, Pyk R, Jobling MK, Vermeulen-Oost W, Berkhourt RJ, et al. (2004) Identification of a new human coronavirus. Nat Med 10: 368–373.
2. Hamre D, Procknow J (1966) A new virus isolated from the human respiratory tract. Proc Soc Exp Biol Med 121: 190–193.
3. Woo PC, Lau SK, Chu CM, Chan KH, Tsoi HW, et al. (2005) Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. J Virol 79: 884–893.
4. Kuiken T, Fouquier RA, Schutten R, Rimmelzwaan GF, van Amerongen G, et al. (2005) Newly discovered coronavirus as the primary cause of severe acute respiratory syndrome. Lancet 362: 263–270.
5. Peiris JS, Lai ST, Poon LL, Guan Y, Yam LY, et al. (2003) Coronavirus as a possible cause of severe acute respiratory syndrome. Lancet 361: 1319–1325.
6. Cheever FS, Daniels JB, Pappenheimer AM, Bailey OT (1949) A murine virus [HJM] causing disseminated encephalomyelitis with extensive destruction of myelin. J Exp Med 90: 181–210.
7. Binn LN, Lazar EC, Keenan KP, Huxcoll DL, Marshwick RH, et al. (1974) Recovery and characterization of a coronavirus from military dogs with diarrhea. Proc Annu Meet U S Anil Health Assoc, pp 359–366.
8. Poon LL, Chu DK, Chan KH, Wong OK, Ellis TM, et al. (2005) Identification of a novel coronavirus in human bronchial biopsy. J Virol 79: 2001–2009.
9. Beaumette FR, Hudson CB (1937) Cultivation of the virus of infectious diarrhea. Proc Annu Meet U S Anim Health Assoc. pp 359–366.
10. Doyle LP, Hutchings LM (1946) A transmissible gastroenteritis virus in pigs. J Exp Med 84: 313–319.
11. Beaudette FR, Hudson CB (1937) Cultivation of the virus of infectious gastroenteritis coronavirus, mouse hepatitis corona-virus. MHV, PHEV stand for porcine epidemic diarrhoea coronavirus, transmissible gastroenteritis coronavirus, mouse hepatitis coronavirus and porcine hemagglutinating encephalomyelitis coronavirus respectively. (DOC)

**Author Contributions**

Conceived and designed the experiments: C-PC DC-CW K-BW. Performed the experiments: C-PC. Analyzed the data: C-PC K-BW. Contributed reagents/materials/analysis tools: H-FC DC-CW. Wrote the paper: C-PC KBW.

**Table S1**

| Reference | Description |
|-----------|-------------|
| 1. van der Hoek L, Pyk R, Jobling MK, Vermeulen-Oost W, Berkhourt RJ, et al. (2004) Identification of a new human coronavirus. Nat Med 10: 368–373. | N/A |
| 2. Hamre D, Procknow J (1966) A new virus isolated from the human respiratory tract. Proc Soc Exp Biol Med 121: 190–193. | N/A |
| 3. Woo PC, Lau SK, Chu CM, Chan KH, Tsoi HW, et al. (2005) Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. J Virol 79: 884–893. | N/A |
| 4. Kuiken T, Fouquier RA, Schutten R, Rimmelzwaan GF, van Amerongen G, et al. (2005) Newly discovered coronavirus as the primary cause of severe acute respiratory syndrome. Lancet 362: 263–270. | N/A |
| 5. Peiris JS, Lai ST, Poon LL, Guan Y, Yam LY, et al. (2003) Coronavirus as a possible cause of severe acute respiratory syndrome. Lancet 361: 1319–1325. | N/A |

**Table S2**

| Reference | Description |
|-----------|-------------|
| 1. van der Hoek L, Pyk R, Jobling MK, Vermeulen-Oost W, Berkhourt RJ, et al. (2004) Identification of a new human coronavirus. Nat Med 10: 368–373. | N/A |
| 2. Hamre D, Procknow J (1966) A new virus isolated from the human respiratory tract. Proc Soc Exp Biol Med 121: 190–193. | N/A |
| 3. Woo PC, Lau SK, Chu CM, Chan KH, Tsoi HW, et al. (2005) Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. J Virol 79: 884–893. | N/A |
| 4. Kuiken T, Fouquier RA, Schutten R, Rimmelzwaan GF, van Amerongen G, et al. (2005) Newly discovered coronavirus as the primary cause of severe acute respiratory syndrome. Lancet 362: 263–270. | N/A |
| 5. Peiris JS, Lai ST, Poon LL, Guan Y, Yam LY, et al. (2003) Coronavirus as a possible cause of severe acute respiratory syndrome. Lancet 361: 1319–1325. | N/A |

**Table S3**

| Reference | Description |
|-----------|-------------|
| 1. van der Hoek L, Pyk R, Jobling MK, Vermeulen-Oost W, Berkhourt RJ, et al. (2004) Identification of a new human coronavirus. Nat Med 10: 368–373. | N/A |
| 2. Hamre D, Procknow J (1966) A new virus isolated from the human respiratory tract. Proc Soc Exp Biol Med 121: 190–193. | N/A |
| 3. Woo PC, Lau SK, Chu CM, Chan KH, Tsoi HW, et al. (2005) Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. J Virol 79: 884–893. | N/A |
| 4. Kuiken T, Fouquier RA, Schutten R, Rimmelzwaan GF, van Amerongen G, et al. (2005) Newly discovered coronavirus as the primary cause of severe acute respiratory syndrome. Lancet 362: 263–270. | N/A |
| 5. Peiris JS, Lai ST, Poon LL, Guan Y, Yam LY, et al. (2003) Coronavirus as a possible cause of severe acute respiratory syndrome. Lancet 361: 1319–1325. | N/A |

**Author Contributions**

Conceived and designed the experiments: C-PC DC-CW K-BW. Performed the experiments: C-PC. Analyzed the data: C-PC K-BW. Contributed reagents/materials/analysis tools: H-FC DC-CW. Wrote the paper: C-PC KBW.