Global prevalence of SARS-CoV-2 3CL protease mutations associated with nirmatrelvir or ensitrelvir resistance

Ip, Jonathan Daniel; Chu, Allen Wing-Ho; Chan, Wan-Mui; Leung, Rhoda Cheuk-Ying; Abdullah, Syed Muhammad Umer; Sun, Yanni; To, Kelvin Kai-Wang

Published in:
eBioMedicine

Published: 01/05/2023

Document Version:
Final Published version, also known as Publisher’s PDF, Publisher’s Final version or Version of Record

License:
CC BY-NC-ND

Publication record in CityU Scholars:
Go to record

Published version (DOI):
10.1016/j.ebiom.2023.104559

Publication details:

Citing this paper
Please note that where the full-text provided on CityU Scholars is the Post-print version (also known as Accepted Author Manuscript, Peer-reviewed or Author Final version), it may differ from the Final Published version. When citing, ensure that you check and use the publisher's definitive version for pagination and other details.

General rights
Copyright for the publications made accessible via the CityU Scholars portal is retained by the author(s) and/or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights. Users may not further distribute the material or use it for any profit-making activity or commercial gain.

Publisher permission
Permission for previously published items are in accordance with publisher's copyright policies sourced from the SHERPA RoMEO database. Links to full text versions (either Published or Post-print) are only available if corresponding publishers allow open access.

Take down policy
Contact lbscholars@cityu.edu.hk if you believe that this document breaches copyright and provide us with details. We will remove access to the work immediately and investigate your claim.
Global prevalence of SARS-CoV-2 3CL protease mutations associated with nirmatrelvir or ensitrelvir resistance

Jonathan Daniel Ip,a Allen Wing-Ho Chu,a Wan-Mui Chan,a Rhoda Cheuk-Ying Leung,a Syed Muhammad Umer Abdullah,b Yanni Sun,b and Kelvin Kai-Wang To,c,d,e,∗

aState Key Laboratory for Emerging Infectious Diseases, Carol Yu Centre for Infection, Department of Microbiology, School of Clinical Medicine, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Pokfulam, Hong Kong Special Administrative Region, China
bDepartment of Electrical Engineering, City University of Hong Kong, Tat Chee Avenue, Kowloon, Special Administrative Region, China
cDepartment of Microbiology, Queen Mary Hospital, Pokfulam, Hong Kong Special Administrative Region, China
dDepartment of Clinical Microbiology and Infection Control, The University of Hong Kong-Shenzhen Hospital, Shenzhen, China
eCentre for Virology, Vaccinology and Therapeutics, Hong Kong Science and Technology Park, Hong Kong Special Administrative Region, China

Summary
Background Nirmatrelvir-ritonavir (Paxlovid) and ensitrelvir are 3-chymotrypsin-like cysteine protease (3CLpro) inhibitors which have been approved for the treatment of COVID-19 in 2021 and 2022, respectively. Previous studies have identified 3CLpro mutations that are associated with reduced susceptibility to these antivirals. The aim of the current study was to estimate the global prevalence of 3CLpro inhibitor-resistant SARS-CoV-2 strains.

Methods We compiled a list of 3CLpro mutations which have been associated with nirmatrelvir or ensitrelvir resistance based on either viral replication or 3CLpro activity assays, and determined their prevalence among 13.4 million sequences deposited in GISAID as of December 14, 2022, about 1 year after the approval of nirmatrelvir-ritonavir. We analyzed the prevalence for different time periods, SARS-CoV-2 lineages and geographical locations.

Findings Overall, 0.5% (67,095/13,446,588) of the sequences contained at least one mutation that was shown to affect the inhibitory activity of nirmatrelvir or ensitrelvir on viral replication or 3CLpro activity. We did not observe any increasing trend of resistance after the widespread clinical use of nirmatrelvir-ritonavir. G15S (2070 per million) and T21I (1386 per million) were the most prevalent mutations, and these mutations were dominant in some SARS-CoV-2 lineages. E166V and S144E, previously shown to affect the inhibitory activity of nirmatrelvir on viral replication or protease activity by > 100-folds, were found in <1 per million sequences.

Interpretation Our data suggest that 3CLpro inhibitor resistance is currently rare. However, continuous global genotypic and phenotypic surveillance would be crucial in the early detection of resistant mutants.

Funding Richard and Carol Yu, May Tam Mak Mei Yin, The Shaw Foundation Hong Kong, Michael Tong, Marina Lee, Government Consultancy Service, the Emergency Key Program of Guangzhou Laboratory (See acknowledgements for full list).

Copyright © 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Keywords: COVID-19; SARS-CoV-2; Antiviral resistance; Nirmatrelvir; Ensitrelvir; Protease inhibitor

Introduction Coronavirus Disease 2019 (COVID-19) pandemic contributed to an estimated 14.83 million excess deaths globally between 2020 and 2021.1 The case-fatality rate is higher among older adults, individuals with comorbidities and those who have not received COVID-19 vaccines.2 Pharmacological therapies, including antivirals and immunomodulatory agents, have shortened the duration of hospitalization or reduced death rates.1,3

Among the virus-targeting antivirals, remdesivir and monoclonal antibodies were the mainstay of treatment in 2020 and 2021. However, these antivirals are given via the intravenous route and are therefore only suitable for treating hospitalized patients. Furthermore, most
monoclonal antibodies are now ineffective due to the emergence of immune escape mutants. In late 2021, two oral antivirals, nirmatrelvir-ritonavir (Paxlovid) and molnupiravir, were approved for clinical use, which have made outpatient treatment possible. Since then, several other oral antivirals have been found to be clinically effective and have been approved in different countries, including entrelvir furamic acid (S-217622; approved in Japan in November 2022), azvudine (approved in China in August 2022), and deuremidevir hydrobromide (VV116; approved in China in January 2023).6

Both nirmatrelvir-ritonavir and ensitrelvir furamic acid inhibit the 3-chymotrypsin-like cysteine protease (3CL\textsuperscript{pro}), also known as main protease (M\textsuperscript{pro}) and non-structural protein 5 (nsp5), cleaves the polyproteins pp1a and pp1ab, a critical step in SARS-CoV-2 replication.7 In a phase 3 clinical trial, nirmatrelvir-ritonavir was demonstrated to reduce the risk of severe disease.8 Ensitrelvir furamic acid was associated with lower viral load in a hamster model and in a phase 2/3 clinical trial.9,10

Antiviral resistance can result in treatment failure. Although naturally-occurring nirmatrelvir-resistant SARS-CoV-2 strains have not been reported, nirmatrelvir-resistant mutants have been found after serial passage in cell lines.11-13 Several studies have demonstrated the effect of individual mutations on the inhibitory activity of 3CL\textsuperscript{pro}.13,14-20 Previous studies assessing sequences deposited into public databases up to June 2022 showed that mutations that affect nirmatrelvir antiviral activity or mutations located at the 3CL\textsuperscript{pro} inhibitor binding site were rarely found.14,15,21 In the current study, we determined the prevalence of SARS-CoV-2 mutants with reduced susceptibility to nirmatrelvir or ensitrelvir based on 3CL\textsuperscript{pro} genetic mutations.

Methods

Compilation of a list of 3CL\textsuperscript{pro} inhibitor mutation

First, based on studies or FDA reports published on or before January 11, 2023, we compiled a list of mutations which showed reduced susceptibility to nirmatrelvir or ensitrelvir in terms of viral replication or 3CL\textsuperscript{pro} activity (Tables 1 and 2).13,15-20 We excluded mutations that were only shown to affect binding or occurred in resistant mutants but had not been experimentally verified to affect viral replication or 3CL\textsuperscript{pro} activity.

SARS-CoV-2 sequence analysis

We downloaded the metadata of SARS-CoV-2 sequences from GISAID on 14th December 2022.22 A SARS-CoV-2 sequence was included if it had been marked as complete in the metadata file, and was directly obtained from a human clinical specimen without in vitro passage. A sequence was excluded if collection date information was not complete. All processing was performed on Python v3.9.1223 (pandas v1.4.2,24 numpy v1.22.4,25 openpyxl v3.0.10,26 json v2.0.9) running on Anaconda Software Distribution v4.11.0. The script started by loading the metadata into RAM in chunks of 10,000 sequences. The sequences in each chunk were selected for further analysis based on the inclusion and exclusion criteria. The date strings were then converted into pandas datetime objects, and each strain was then assigned with the respective quarter that it was collected in.
To calculate the prevalence of each mutation, the strains containing the mutation identifier in their “AA Substitutions” column were extracted. Then, the number of strains belonging to each continent was counted using the value_counts() function. If we were to search for the prevalence of a set of mutations, we would use a pipe symbol “|” to separate all the possible mutations, as pandas.series.str.contains default supports regular expression (regex) searching. The number of mutations was then normalized to strain per million.

To extract the metadata of strains with 3CLpro E166V or both E166V and L50F, each strain was searched for the target mutation(s) within their “AA Substitutions” column. If a target was found, the GISAID accession ID was then normalized to strain per million.

Ethics
Ethical approval is not required as anonymised data retrieved from GISAID were used in this study.

Role of the funding source
The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results
A total of 14,157,266 SARS-CoV-2 sequences were obtained from GISAID on 14th December 2022, of which 13,446,588 satisfied our selection criteria. 67,095 (4990/million) sequences contained at least one amino acid mutation that has been shown to affect the inhibitory activity of nirmatrelvir or ensitrelvir on viral replication or 3CLpro activity, including 36,072 (2683/million) sequences containing mutations that have been shown to affect the inhibitory activity on viral replication or 3CLpro activity, including 36,072 (2310/million) sequences with mutations that have been shown to affect the inhibitory activity of nirmatrelvir or ensitrelvir on viral replication or 3CLpro activity. Including the Variant of Interest (VOI) Lambda [C.37] (Supplementary Fig. S2a), currently, there is no published data on the effect of G15S mutation on the inhibitory effect of 3CLpro inhibitor on viral replication.11 The second most frequently found mutation was T21I (1386/million [18,637/13,446,588]), which was shown to affect the inhibitory effect of nirmatrelvir on viral replication.15 T21I was present in almost all strains in the B.1.1.318 lineage, a VOI prevalent in West Africa in early 2021.17,20 and the Gamma variant sublineage

<table>
<thead>
<tr>
<th>Mutation</th>
<th>Reduction in susceptibility of SARS-CoV-2 replication (fold change)</th>
<th>Reduction in inhibition of 3CLpro activity (fold change)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>G15S</td>
<td>N/A</td>
<td>4.4</td>
<td>11</td>
</tr>
<tr>
<td>T21I</td>
<td>1.1–4.6</td>
<td>N/A</td>
<td>11,13,14</td>
</tr>
<tr>
<td>L50F</td>
<td>1.4–4.2</td>
<td>N/A</td>
<td>11,13,14</td>
</tr>
<tr>
<td>Y54A</td>
<td>N/A</td>
<td>24.0</td>
<td>14</td>
</tr>
<tr>
<td>Y54C</td>
<td>2.6–8.3</td>
<td>N/A</td>
<td>15</td>
</tr>
<tr>
<td>T135I</td>
<td>N/A</td>
<td>3.2</td>
<td>11</td>
</tr>
<tr>
<td>F340A</td>
<td>N/A</td>
<td>39.0</td>
<td>11</td>
</tr>
<tr>
<td>F340L</td>
<td>N/A</td>
<td>5.4</td>
<td>11</td>
</tr>
<tr>
<td>S144A</td>
<td>2.2–12.2</td>
<td>20.5–92</td>
<td>12,13,19</td>
</tr>
<tr>
<td>S144E</td>
<td>N/A</td>
<td>470.0</td>
<td>14</td>
</tr>
<tr>
<td>S144F/G/M/Y</td>
<td>N/A</td>
<td>19.2–38.0</td>
<td>14</td>
</tr>
<tr>
<td>S144T</td>
<td>N/A</td>
<td>160.0</td>
<td>14</td>
</tr>
<tr>
<td>H164N</td>
<td>N/A</td>
<td>6.4</td>
<td>14</td>
</tr>
<tr>
<td>M165T</td>
<td>N/A</td>
<td>29.9</td>
<td>14</td>
</tr>
<tr>
<td>E166A</td>
<td>3.3</td>
<td>10.0–33.0</td>
<td>14,17</td>
</tr>
<tr>
<td>E166G</td>
<td>N/A</td>
<td>16.0–16.4</td>
<td>14,18</td>
</tr>
<tr>
<td>E166V</td>
<td>25.0–288</td>
<td>N/A</td>
<td>14–15</td>
</tr>
<tr>
<td>L150F</td>
<td>3.6–16.5</td>
<td>N/A</td>
<td>14,17</td>
</tr>
<tr>
<td>Del P168</td>
<td>5.1</td>
<td>N/A</td>
<td>13</td>
</tr>
<tr>
<td>H172Q/F</td>
<td>N/A</td>
<td>&gt;42.0</td>
<td>18</td>
</tr>
<tr>
<td>A173Y</td>
<td>24</td>
<td>&gt;332.7</td>
<td>14,18,20</td>
</tr>
<tr>
<td>A173T</td>
<td>4.1</td>
<td>N/A</td>
<td>19</td>
</tr>
<tr>
<td>A173V</td>
<td>0.9–11.6</td>
<td>26.0</td>
<td>13,14,19</td>
</tr>
<tr>
<td>V286G</td>
<td>N/A</td>
<td>13.0</td>
<td>14</td>
</tr>
<tr>
<td>Q189K</td>
<td>N/A</td>
<td>65.0</td>
<td>14</td>
</tr>
<tr>
<td>Q192L</td>
<td>4.3</td>
<td>28.0</td>
<td>14,18</td>
</tr>
<tr>
<td>Q192P</td>
<td>7.6</td>
<td>33.0</td>
<td>14,18</td>
</tr>
<tr>
<td>Q192TS/A/H/V/W/C/F</td>
<td>N/A</td>
<td>&gt;22.2</td>
<td>18</td>
</tr>
<tr>
<td>D248E</td>
<td>N/A</td>
<td>3.7</td>
<td>14</td>
</tr>
<tr>
<td>P252L</td>
<td>5.9</td>
<td>N/A</td>
<td>13,14</td>
</tr>
<tr>
<td>T304I</td>
<td>2.1–5.5</td>
<td>N/A</td>
<td>13,13,14</td>
</tr>
</tbody>
</table>

Table 1: 3CLpro mutations that were associated with reduced nirmatrelvir inhibitory activity of SARS-CoV-2 replication or 3CLpro activity. Overall, the most frequent mutation was G15S (2070/million [27,828/13,446,588]) (Fig. 2 and Supplementary Fig. S1), which conferred reduced susceptibility to nirmatrelvir 3CLpro inhibitory activity.24 G15S was present in almost all strains in the lineages B.1.1.1, B.1.1.369, B.1.1.372, B.1.1.375, and lineage C, including the Variant of Interest (VOI) Lambda [C.37] (Supplementary Fig. S2a). There are also differences in the geographical distribution of the prevalence of mutations. South America has the highest proportion of sequences carrying mutations associated with nirmatrelvir or ensitrelvir resistance (29,032 per million), followed by Africa (24,437 per million) and North America (4809 per million) (Supplementary Table S1).
P.1.7.1, which was prevalent in South America (Supplementary Fig. S2b). In a study assessing the emergence of nirmatrelvir-resistant strains during serial passage, T21I was shown to be one of the five precursor mutations that were present in all resistant strains.11

Among mutations that have been shown to affect the inhibitory activity of nirmatrelvir or ensitrelvir on viral replication, E166V conferred the greatest reduction in susceptibility (25-288-fold reduction for nirmatrelvir; 23-fold reduction for ensitrelvir) (Tables 1 and 2). E166V was found in 13 sequences (0.97/million) (Supplementary Table S2 and S3). Four were collected from a 65-year-old man from Japan at 1-week intervals between 30 May and 17 June 2022, and belonged to BA.1.1 lineage. All other strains with E166V mutations belonged to different lineages and were collected from different places. Amino acid 166 is a key residue that interacts with nirmatrelvir, including polar contact with the pyrrolidone group and hydrogen bonds between the tert-butyl moiety of nirmatrelvir.29,30 Previous studies showed that E166V confers a loss of viral fitness, but can be compensated by L50F mutation.11,12 Among the 13 sequences with E166V mutations, one (7.7%) sequence also contained L50F mutation (0.07/million). Among mutations that have been shown to affect the inhibitory activity of nirmatrelvir or ensitrelvir on 3CLpro activity, S144E confers the greatest reduction in the inhibition of nirmatrelvir on 3CLpro activity (470-fold reduction in 3CLpro activity) (Table 1). However, it occurs in less than 1 per million (Supplementary Figure S1).

Table 2: 3CLpro mutations that were associated with reduced ensitrelvir inhibitory activity of SARS-CoV-2 replication or 3CLpro activity.

<table>
<thead>
<tr>
<th>Mutation</th>
<th>Reduction in susceptibility of SARS-CoV-2 replication (fold change)</th>
<th>Reduction in inhibition of 3CLpro activity (fold change)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>T21I</td>
<td>1.7</td>
<td>N/A</td>
<td>11</td>
</tr>
<tr>
<td>T45I</td>
<td>4.1</td>
<td>N/A</td>
<td>19</td>
</tr>
<tr>
<td>D48Y</td>
<td>5.0</td>
<td>N/A</td>
<td>19</td>
</tr>
<tr>
<td>M49I/L/T/V</td>
<td>2.6-25.4</td>
<td>N/A</td>
<td>19</td>
</tr>
<tr>
<td>L50F</td>
<td>2.8</td>
<td>N/A</td>
<td>11</td>
</tr>
<tr>
<td>S144A</td>
<td>13-16.9</td>
<td>N/A</td>
<td>11,19</td>
</tr>
<tr>
<td>E166V</td>
<td>23</td>
<td>N/A</td>
<td>11</td>
</tr>
<tr>
<td>Del P168</td>
<td>6.8</td>
<td>N/A</td>
<td>19</td>
</tr>
<tr>
<td>A173V</td>
<td>1.7</td>
<td>N/A</td>
<td>11</td>
</tr>
<tr>
<td>P252L</td>
<td>1.9</td>
<td>N/A</td>
<td>11</td>
</tr>
<tr>
<td>T304I</td>
<td>1.6</td>
<td>N/A</td>
<td>11</td>
</tr>
</tbody>
</table>

This table only shows the effect of single mutations.

There are several limitations in this study. First, there could be other mutations that confer 3CLpro inhibitor resistance which have not been described previously. In the study, we compiled a list of mutations that influence viral replication or 3CLpro activity, which are associated with reduced susceptibility to nirmatrelvir approval of Paxlovid. Only 0.5% of analyzed sequences had mutations associated with 3CLpro inhibitor resistance. There are several unique features in our study. First, we have examined mutations or deletions at 23 amino acid positions of the 3CLpro protein which have been experimentally verified to affect either the viral replication or 3CLpro activity. Second, we have assessed the prevalence of mutations up to January 2023, which is more than 1 year after the approval of nirmatrelvir-ritonavir. Third, we have performed subgroup analysis on the prevalence of mutations in different sublineages.

While it is reassuring that 3CLpro inhibitors will likely remain clinically effective, there are several reasons why we should continue to monitor for 3CLpro inhibitor resistance. First, the most frequent mutations, G15S and T21I, were the dominant mutations within some lineages, suggesting that these mutations do not affect viral fitness and can be widely disseminated. These mutations first appeared before the clinical use of 3CLpro inhibitors, suggesting that they can arise spontaneously and are not selected by the usage of protease inhibitors. However, with higher selective pressure after widespread use, these mutations may spread and become fixed in a population. Second, the low prevalence of 3CLpro resistant mutations may be related to the low consumption of nirmatrelvir-ritonavir after clinical approval.11 Possible reasons for low consumption include the fear of viral rebound and side effects. With increasing use of 3CLpro inhibitors, resistance may emerge in the future. Third, antiviral resistance can occur suddenly. For example, oseltamivir was first introduced in late 1990s as an antiviral for influenza, but a major surge in oseltamivir resistance only occurred about 10 years later, when oseltamivir resistance for seasonal influenza A (H1N1) in 2008 suddenly jumped from <10% before the 2006–2007 season to almost 100% in the 2008–2009 season.32 Fourth, there is a growing concern that widespread use of molnupiravir, another widely used oral antiviral for the treatment of SARS-CoV-2 infection that acts by introducing viral genome mutations, may accelerate the mutation rate of circulating viruses.33,34

We found a difference in the geographical distribution of SARS-CoV-2 3CLpro mutations, with South America and Africa having the highest prevalence. The high prevalence in these two continents is because some mutations were present in widely circulating lineage in those continents. Specifically, G15S was present in Lambda variant which circulates widely in South America in 2021, while T21I was present in the B.1.1.318 which circulated widely in West Africa in early 2021.

There are several unique features in our study. First, we have examined mutations or deletions at 23 amino acid positions of the 3CLpro protein which have been experimentally verified to affect either the viral replication or 3CLpro activity. Second, we have assessed the prevalence of mutations up to January 2023, which is more than 1 year after the approval of nirmatrelvir-ritonavir. Third, we have performed subgroup analysis on the prevalence of mutations in different sublineages.

While it is reassuring that 3CLpro inhibitors will likely remain clinically effective, there are several reasons why we should continue to monitor for 3CLpro inhibitor resistance. First, the most frequent mutations, G15S and T21I, were the dominant mutations within some lineages, suggesting that these mutations do not affect viral fitness and can be widely disseminated. These mutations first appeared before the clinical use of 3CLpro inhibitors, suggesting that they can arise spontaneously and are not selected by the usage of protease inhibitors. However, with higher selective pressure after widespread use, these mutations may spread and become fixed in a population. Second, the low prevalence of 3CLpro resistant mutations may be related to the low consumption of nirmatrelvir-ritonavir after clinical approval.11 Possible reasons for low consumption include the fear of viral rebound and side effects. With increasing use of 3CLpro inhibitors, resistance may emerge in the future. Third, antiviral resistance can occur suddenly. For example, oseltamivir was first introduced in late 1990s as an antiviral for influenza, but a major surge in oseltamivir resistance only occurred about 10 years later, when oseltamivir resistance for seasonal influenza A (H1N1) in 2008 suddenly jumped from <10% before the 2006–2007 season to almost 100% in the 2008–2009 season.32 Fourth, there is a growing concern that widespread use of molnupiravir, another widely used oral antiviral for the treatment of SARS-CoV-2 infection that acts by introducing viral genome mutations, may accelerate the mutation rate of circulating viruses.33,34

We found a difference in the geographical distribution of SARS-CoV-2 3CLpro mutations, with South America and Africa having the highest prevalence. The high prevalence in these two continents is because some mutations were present in widely circulating lineage in those continents. Specifically, G15S was present in Lambda variant which circulates widely in South America in 2021, while T21I was present in the B.1.1.318 which circulated widely in West Africa in early 2021. 

Discussion

3CLpro inhibitors represent a major breakthrough in the treatment of COVID-19, especially for outpatient treatment since these antivirals can be taken orally. Our analysis indicates that 3CLpro inhibitor resistant mutations remain rare and mostly sporadic up to one year after the
or ensitrelvir from previous studies and FDA reports published on or before January 11, 2023 (Tables 1 and 2). We only included the mutations that were previously experimentally investigated with fold change in viral replication and inhibition of 3CLpro activity. Some mutations, such as substitutions at G143, were not experimentally tested, therefore fold change results were not available. The effects of multiple mutations were also not considered in our study. Second, as the genomic surveillance is poorer in some countries, emergence of resistance in these places could be missed if the sequences and metadata are not published online. Third, as GISAID is an open public database, it is possible that the sequences retrieved can be duplicated or from the same person. Fourth, the 3CLpro sequences obtained from clinical specimens available on GISAID do not contain information regarding the viability of the viral strains. Our analysis would not be able to estimate the proportion of strains with mutations that are from infectious virus particles.

In conclusion, our study shows that the prevalence of protease inhibitor resistance remain low one year after the approval of nirmatrelvir-ritonavir. Early detection of the emergence of 3CLpro inhibitor resistance would require a global effort in genomic and phenotypic characterization with prompt dissemination of information.

Fig. 1: Global temporal trend of 3CLpro mutations that have been shown to affect the inhibitory activity of nirmatrelvir or ensitrelvir on viral replication or 3CLpro activity. Only the 5 most frequently found mutations are shown. Solid line indicates 3CLpro mutations which have been demonstrated to affect inhibitory activity of nirmatrelvir and/or ensitrelvir on viral replication. Dotted line indicates 3CLpro mutations which have been demonstrated to affect inhibitory activity of nirmatrelvir or ensitrelvir on protease activity only, but without data on viral replication.

Fig. 2: Frequency of 3CLpro mutations (sequences per million) among sequences deposited onto GISAID up to 14th December 2022. Only the 10 most frequently found mutations affecting the inhibitory activity of nirmatrelvir or ensitrelvir are shown.
Contributors
JDI and KKWT contributed to the conception and design of the study. JDI, AWHC, WMG, SMUA, RCYL contributed to data acquisition. JDI, SMUA, YS, RCYL and KKWT contributed to data analysis and interpretation. All authors critically revised the manuscript for important intellectual content. KKWT obtained the funding. KKWT supervised the research. JDI and KKWT have accessed and verified the data. All authors read and approved the final version of the manuscript.

Data sharing statement
The Python source code used in this study have been deposited at https://github.com/Jonathan-D-ly/Global-prevalence-of-SARS-CoV-2-3CL-protease-mutations-associated-with-nirmatrelvir-or-ensitrelvir.

Declaration of interests
KKWT report collaboration with Sinovac, Sinopharm and Wantai BioPharm. Other authors declare no competing interests.

Acknowledgments
We gratefully acknowledge the authors originating and submitting laboratories of the SARS-CoV-2 genetic sequences and metadata made available through GISAID from which this research is based. This work was partly supported by the Consultancy Service for Enhancing Laboratory Surveillance of Emerging Infectious Diseases and Research Capability on Antimicrobial Resistance for Department of Health of the Hong Kong, and the Emergency Key Program of Guangzhou Laboratory (EKPG22-01), and the Hong Kong Innovation and Technology Fund MPR/071/20X. We acknowledge funding received from private donors including the Shaw Foundation Hong Kong, Richard Yu and Carol Yu, May Tam Mak Mei Yin, Michael Seak-Kam Tong, Respiratory Viral Research Foundation Limited, Hui Ming, Hui Hoy and Chow Sin Lan Charity Fund Limited, Chan Yin Chuen Memorial Charitable Foundation, and Marina Man-Wai Lee.

Appendix A. Supplementary data
Supplementary data related to this article can be found at https://doi.org/10.1016/j.cebion.2023.104559.

References