Transmissible SARS-CoV-2 variants with resistance to clinical protease inhibitors

Seyed Arad Moghadasi[1], Emmanuel Heilmann[2], Sofia N. Moraes[1], Fiona L. Kearns[3], Dorothee von Laer[2], Rommie E. Amaro[3] & Reuben S. Harris[1,4,5,*]

[1] Department of Biochemistry, Molecular Biology, and Biophysics, University of Minnesota – Twin Cities; Minneapolis, Minnesota, USA, 55455

[2] Institute of Virology, Medical University of Innsbruck; Innsbruck, Austria

[3]Department of Chemistry and Biochemistry, University of California San Diego, La Jolla, California, USA, 92093

[4] Department of Biochemistry and Structural Biology, University of Texas Health San Antonio; San Antonio, Texas, USA, 78229

[5] Howard Hughes Medical Institute, University of Texas Health San Antonio; San Antonio, Texas, USA, 78229

*Corresponding author. Email: rsh@uthscsa.edu

Keywords: antiviral drugs; antiviral drug resistance; main viral protease (M^{pro}/3CLpro); natural genetic variation; Paxlovid (nirmatrelvir); protease inhibitor resistance; SARS-CoV-2

One Sentence Summary: Resistance to protease inhibitor drugs, nirmatrelvir (Paxlovid) and ensitrelvir, exists in SARS-CoV-2 variants in the human population.

Abstract: First-generation vaccines and drugs have helped reduce disease severity and blunt the spread of SARS-CoV-2. However, ongoing virus transmission and evolution and increasing selective pressures have the potential to yield viral variants capable of resisting these interventions. Here, we investigate the susceptibility of natural variants of the main protease (M^{pro}/3CLpro) of SARS-CoV-2 to protease inhibitors. Multiple single amino acid changes in M^{pro} confer resistance to nirmatrelvir (the active component of Paxlovid). An additional inhibitor in clinical development, ensitrelvir, shows a different resistance mutation profile. Importantly, phylogenetic analyses indicate that nirmatrelvir-resisting variants have pre-existed the introduction of this drug into the human population and are capable of spreading. A similarly strong argument can be made for ensitrelvir. These results caution against broad administration of protease inhibitors as stand-alone therapies and encourage the development of additional

protease inhibitors and other antiviral drugs with different mechanisms of action and resistance profiles.

INTRODUCTION

The main protease (M^{pro}/3CLpro) of coronaviruses (CoV) emerged as a strong candidate for inhibitor development in response to the first severe acute respiratory syndrome (SARS) pandemic in 2002 and was swiftly revisited in response to the more recent COVID-19 pandemic (*1-4*). M^{pro} activity is essential for virus replication and, combined with precedents set by the successful development of human immunodeficiency virus-type 1 (HIV-1) and hepatitis C virus (HCV) protease inhibitors, drugs targeting this enzyme are likely to help treat SARS-CoV-2 infections (*5, 6*). Many groups have embarked on campaigns to target M^{pro} with multiple chemical series being advanced into potent inhibitors at unprecedented speeds, mainly owing to the wealth of biochemical and structural information that has accumulated on coronaviruses proteases over the past two decades (*7-11*). For instance, prior efforts were leveraged to accelerate the development of PF-00835231 into PF-07321332 (nirmatrelvir), the active ingredient in Paxlovid and the first FDA approved M^{pro} inhibitor currently prescribed in outpatient settings (*12*). Another M^{pro} inhibitor in clinical-stage development is S-217622 (ensitrelvir), a non-covalent, non-peptidic inhibitor developed through computational and medicinal chemistry (*13*). Ensitrelvir and other molecules in various stages of development may soon provide alternatives to Paxlovid or present opportunities for combinatorial therapy.

While Paxlovid is already proving useful in blunting SARS-CoV-2 disease pathogenesis, the long-term consequences of its wide-spread prescription are unknown. Resistance is a major concern given the relatively rapid rates at which SARS-CoV-2 is changing (Alpha, Beta, Delta, Omicron, *etc*) and the fact that the potency of nirmatrelvir and ensitrelvir varies widely against other coronavirus species (*14*). For instance, the main proteases of the human α-coronaviruses NL63 and 229E are less susceptible to these drugs suggesting that natural mechanisms of resistance are indeed possible and may already exist in nature (*12, 15*). In addition, during the clinical development of Paxlovid, the murine coronavirus MHV was used to study nirmatrelvir resistance (https://www.fda.gov/media/155050/download). One of the selected amino acid changes, S144A, causes a >90-fold reduction in the binding efficacy of nirmatrelvir to recombinant M^{pro} *in vitro*. To our knowledge, this mutant has yet to be assessed in cellular assay or animal models.

We recently reported a cell-based gain-of-signal assay for SARS-CoV-2 M^{pro} function in which wildtype (WT) protease activity suppresses luminescent signal and genetic or chemical inhibition yields signal increases proportional to inhibitor efficacy or mutant severity (*15*). This system enabled us to show that a single amino acid change (P168S) in an active site-adjacent loop region of M^{pro} improves susceptibility to the HCV protease inhibitor boceprevir yet has no effect on the inhibitory capacity of the tool compound GC376 (*15*) (repeated below). P168S is a naturally occurring variant that comprises 76% of changes at this amino acid position in SARS-CoV-2 clinical isolates based on sequences deposited in the GISAID database [1-July-2022 (*16*)]. This striking example of differential drug responsiveness inspired us to ask whether this natural variation at position 168, other natural changes at position 168, and other naturally occurring variants in the vicinity of the active site cavity may confer resistance to nirmatrelvir and/or ensitrelvir. Our results combine to demonstrate that multiple drug resistance mutations

already exist in transmissible isolates of SARS-CoV-2 in the global population. However, the longer-term view remains optimistic given our observations of differential drug susceptibilities, which suggests that future compounds may be developed to avoid cross-resistance and help combat the COVID-19 pandemic and future coronavirus outbreaks.

RESULTS

A natural M^{pro} variation ΔP168 confers resistance to nirmatrelvir and ensitrelvir

As introduced above, P168S is an existing SARS-CoV-2 M^{pro} variant. The next most frequent change at this position in the GISAID database is a single residue deletion, ΔP168 (**Fig. 1A**). High resolution structures demonstrate that P168 is located close to the binding sites of boceprevir and nirmatrelvir (4.0Å and 3.3Å, respectively) and approximately twice as far from that of ensitrelvir [8.8Å; (*13, 17*)] (**Fig. 1B**). Based on our prior work with P168S (*15*), we predicted that changes at this amino acid position in M^{pro} will more strongly affect the efficacy of boceprevir and nirmatrelvir and have little effect on ensitrelvir. In confirmation of our prior studies, P168S causes a 5.5-fold increase in susceptibility to boceprevir (**Fig. 1C**, left). We were therefore surprised to find that the ΔP168 mutant protease has no detectable effect with boceprevir, as its dose response curve is indistinguishable from that of WT M^{pro}. In contrast, the ΔP168 variant causes 5.1- and 6.8-fold increased resistance to nirmatrelvir and ensitrelvir, and the P168S variant shows WT-like responsiveness to these two drugs (**Fig. 1C**, middle and right). As a positive control, the S144A mutant described in the Introduction also shows a strong nirmatrelvir resistance phenotype with a 14-fold increase in IC_{50} (**Fig. S1**).

An additional metric of our cell-based assay for M^{pro} function is luminescent signal in the absence of drug (*15*). The WT construct emits very low luminescence and any decrease in M^{pro} activity results in increased signal with a maximum of approximately a 40-fold increase as defined by a catalytic residue mutation (C145A, **Fig. 1D**). Consistent with a high prevalence in global isolates, P168S shows no change in background luminescence relative to WT, whereas the selected mutant S144A causes a 3.5-fold increase (**Fig. 1D**), consistent with a reported diminution of biochemical activity (*18*). In comparison, ΔP168 elicits a more modest 2-fold increase in background luminescence relative to WT (**Fig. 1D**). These results suggest that the ΔP168 enzyme remains active and likely capable of supporting virus replication. In support of both resistance and activity results, the increased resistance of the M^{pro} ΔP168 variant to both nirmatrelvir and ensitrelvir was confirmed using an orthologous VSV-based M^{pro} *cis*-cleavage assay in which inhibition of catalytic activity enables VSV replication (*19*) (**Fig. 1E**).

Drug resistance profiles of additional naturally occurring single amino acid M^{pro} variants

Encouraged by analyzing the resistance phenotypes of amino acid changes at a single position in global sequences, we extended our analyses to include an additional nine naturally variable residues that localize to two regions in proximity to the M^{pro} active site (**Fig. 2A-B**). Given the strong phenotypes of P168 variants, we hypothesized that mutations at additional residues within this beta hairpin might also cause resistance. We therefore first focused on residues 165, 169, 171, and 173, which show variability across coronavirus species and, importantly, also within circulating SARS-CoV-2 variants (**Fig. 2C**). M165, T169, and V171 were each substituted with isoleucine because this is a recurrent change at these positions, and

A173 was substituted to valine as the most observed change at this position in SARS-CoV-2 and also the residue found in other CoV strains (HCoV-229E and NL63). The second region encompasses M^{pro} residues 45-49, which form a helix that spans the hydrophobic S2 subsite through the sidechain of M49. This region is hypervariable across different coronaviruses species both in amino acid identity as well as overall length. M49I is a highly frequent change in circulating viruses (>1000 occurrences) and a substitution we have shown previously to cause little difference in GC376 or boceprevir susceptibility, but its impact on nirmatrelvir and ensitrelvir efficacy has yet to be analyzed. Furthermore, we were intrigued by the hydrogen bonding pattern formed by T45 and D48 and curious whether disrupting this network might impact inhibitor susceptibility (**Fig. 2B**). Therefore, in addition to M49I, T45I and D48Y were selected for analysis due to the likelihood of disrupting the aforementioned hydrogen bond network, and S46F and E47K were chosen as non-conservative changes also likely to be disruptive (**Fig. 2C**).

Of nine amino acid substitution mutants, S46F, E47K, M165I, T169I, and V171I have little affect (<2-fold) on M^{pro} susceptibility to either nirmatrelvir or ensitrelvir and, moreover, did not cause substantial increases in background luminescence consistent with near-WT catalytic activity (**Fig. 2D-E**). Amongst all mutants tested, A173V immediately stands out as a separation-of-function variant by causing a 11.6-fold increase in resistance to nirmatrelvir and no change in susceptibility to ensitrelvir (**Fig. 2E**). Furthermore, this resistance mutation is unlikely to impact protease activity because there is only a modest, 2-fold increase in background luminescence in the absence of inhibitor (**Fig. 2D**). A173T is also observed frequently at this position and causes an intermediate phenotype with 4.5-fold resistance to nirmatrelvir and unchanged susceptibility to ensitrelvir (**Fig. S2**). The resistance phenotype shown by A173V is surprising considering this residue does not make direct contact with the inhibitor and, therefore, a mechanism based on dynamic local reorganization of binding site residues impacting nirmatrelvir binding is possible (**Fig. 2B** and see below). These contrasting results highlight the potential for chemically distinct M^{pro} inhibitors to have different drug resistance mutation profiles. In further support of this interpretation, additional contrast is seen with resistance mutations surrounding the S2 subsite. T45I and D48Y cause 4.5- and 6-fold increases in resistance to ensitrelvir relative to the WT, and both have more modest effects on nirmatrelvir susceptibility (approximately 2-fold; **Fig. 2E**). Most strikingly, M49I causes no shift in the nirmatrelvir dose response while causing a 12.4-fold increase in resistance to ensitrelvir (**Fig. 2E**). Together, these data demonstrate the ability of naturally occurring variants to exhibit differential responses to current front-line and potential next-generation M^{pro} drugs.

A double mutant shows synergistic resistance to nirmatrelvir

The two changes with the largest effect on nirmatrelvir resistance are ΔP168 and A173V. This prompted us to test whether the combination would be tolerated and, if so, whether they might be additive or multiplicative in terms of drug resistance. Remarkably, the ΔP168/A173V double mutant shows a 51-fold increase in resistance to nirmatrelvir (**Fig. 3A**). In contrast, this double mutant elicits only a 2.8-fold increase in resistance to ensitrelvir, which is less than that of the ΔP168 mutant alone (compare response curves in **Fig. 3A** and **Fig. 1C**). As shown above, both of these mutations have modest effects on background luminescence levels and the double mutant elicits a roughly additive effect causing a less than 3-fold increase in overall luminescence, which is indicative of intact protease functionality (**Fig. 3B**). The double mutant

was also tested in the orthologous VSV based M^{pro} assay yielding similarly striking separation-of-function results (**Fig. S3**). Taken together, these results demonstrate that strong resistance to nirmatrelvir can be achieved by combining two naturally occurring amino acid changes. Moreover, our observations with nirmatrelvir resistance (synergy) and ensitrelvir resistance (epistasis) provide further support for the likelihood that these compounds have largely non-overlapping binding mechanisms and therefore distinct resistance profiles.

To provide a structure-based rationale for the observed resistance of the ΔP168/A173V mutant to nirmatrelvir, these two amino acid changes were modelled using RosettaCM (*20*) and compared to existing structures of the WT protein (*13*) (**Fig. 3C**). Deletion of P168 is predicted to cause a tightening of the apical loop of the beta-hairpin, resulting in the main chain carbonyl of L167 to fold downward and form hydrogen bonds with the main chain amino group of V171 (**Fig. 3C**). This rearrangement of the peptide backbone is also predicted to cause the sidechain of L167 to swing out toward and likely clash with the trifluoroacetamide group of nirmatrelvir. A173 does not directly interact with nirmatrelvir, and instead lies behind M165 which makes hydrophobic contacts with the dimethyl-3-azabicyclohexane moiety at the P2 position of nirmatrelvir. Modeling of A173V shows the side chain of valine points in toward the hydrophobic core formed by the sidechains of M165, L167, and V171. The bulkier side chain of valine compared to alanine pushes M165 out towards the active site which could potentially clash with the P2 position of nirmatrelvir. The intermediate resistance phenotype of A173T is in line with this explanation, given that a threonine side chain is less bulky than valine and likely to be more accommodating of the side chain of M165. Thus, both ΔP168 and A173V appear to cause resistance to nirmatrelvir through sidechain rearrangements that indirectly impact active site residues and cause steric clashes with nirmatrelvir.

In contrast, the selective resistance of M49I to ensitrelvir is likely a more direct effect caused by the loss of hydrophobic contacts with the inhibitor (**Fig. S4**). Unlike nirmatrelvir, ensitrelvir engages the S1' subsite through its 6-chloro-2-methyl-2H-indazole moiety, which projects towards T25 and makes extensive hydrophobic contacts with the side chain of M49. Modeling of M49I shows a large gap that is formed due to the branched sidechain of isoleucine being unable to extend towards T25 and thereby decreasing contact with the P1' ligand of ensitrelvir (**Fig. S4**). Other small amino acids such as valine or threonine at position 49 would be predicted to have a similar or larger resistance phenotypes to M49I while changes to leucine or phenylalanine could still make favorable contacts with ensitrelvir and thereby not cause substantial resistance. These results suggest that reliance on hydrophobic contacts in the S1' subsite may be a vulnerability when considering escape mechanisms for next generation M^{pro} inhibitors. However, the non-overlapping drug susceptibility profiles of A173V and M49I are encouraging for the prospect of combinatorial therapy with chemically distinct M^{pro} inhibitors.

Global variant distributions and evidence for transmission

The global frequency and distribution of an individual mutant provides an indication as to whether a particular amino acid change might be tolerated in nature. However, the current sequence volume of SARS-CoV-2 genomes in the GISAID database is so large that low/no stringency filtering indicates mutations at every position of M^{pro}, even at the absolutely conserved catalytic dyad, despite the fact that amino acid changes in this region prevent proteolytic activity and, consequently, virus replication (**Fig. S5**). This strongly suggests that the

database contains a certain level of sequences that are not viable and therefore not transmitting through the human population. To address this problem, we used Ultrafast Sample placement on Existing tRee (UShER) to determine phylogenetic relationships between genomes harboring drug resistance mutations in M^{pro} (*21*).

Viral genomes containing ΔP168 have arisen multiple times independently, with the majority of these sequences falling within the Delta lineage (**Fig. 4A**). A distinct cluster of 49 genomes deposited between September and December of 2021 is derived from a single founder event followed by multiple regional transmissions in Germany, in addition to evidence suggesting its spread to England, USA, Austria, and Romania (**Fig. 4A**, **Fig. S6A**). The ΔP168 variant has yet to be documented in currently circulating Omicron strains; however, given the evidence for multiple independent introductions of this variant as well as clear transmission capacity, there is strong potential for reemergence. It is further notable that these ΔP168 variant sequences were deposited prior to Paxlovid release, and widespread prescription of this drug could provide sufficient selective pressure to increase the allele frequency of ΔP168.

A173V has occurred multiple times during the pandemic on different viral backbones including Alpha, Delta, and Omicron (**Fig. 4B**). A173V distribution and more frequent introduction relative to ΔP168 is expected, given that single base substitutions are more likely to occur than in-frame deletions. Moreover, A173V appears in multiple Omicron lineages, suggesting that current strains with this change may already be spreading in localized clusters. Of more immediate concern is a July 2022 report of two genetically identical cases in the USA of A173V in the emerging BA.5 Omicron sub-lineage (*22*) (**Fig. 4B**, **Fig. S6B**). Moreover, A173V cases are present in multiple continents including North America, Europe, and Asia, providing further evidence that this variant is able to readily arise and transmit—with additional spread likely ongoing.

M49I, which confers 12.4-fold resistance to ensitrelvir (**Fig. 2E**), has been reported in multiple viral backbones including Omicron and in multiple geographic locations (**Fig. 4C**). Importantly, of all naturally occurring variants studied here, M49I is the second most frequent change present in circulating viruses with >1000 entries in the GISAID database (5-Aug-2022). Owing to its high prevalence, we only inspected M49I harboring genomes within the multiple Omicron sublineages. These phylogenetic analyses demonstrate multiple independent occurrences and clear evidence for transmission chains, including related cases currently circulating in the BA.5 sublineage (**Fig. 4C**, **Fig. S6C**). Although ensitrelvir has yet to be approved for clinical use, the high level of resistance conferred by M49I and the already high frequency of this allele in the circulating viral population are concerning and combine to suggest that treatments with this drug may need to be coupled with variant identification using viral genome sequencing to ensure that prescriptions are not provided to patients whose viruses already have this resistance mutation.

Although the double mutant described here has yet to be reported in circulating genomes, the emergence of combinatorial mutants is inevitable—especially under the selective pressures imposed by drugs. In support of this likelihood, two recently deposited sequences in Germany suggest the existence of a new variant with three mutations in M^{pro}, namely S46F, A94T, and A173T (**Fig. S7**). As shown here, A173T causes a 4.5-fold increase in resistance to nirmatrelvir and S46F appears to have no significant effect. However, we have yet to analyze all possible double and triple mutant combinations, and it is possible that these three amino acid substitutions could confer a selective advantage *in vivo*. It will be interesting to determine whether the patients

infected with this triple mutant were treated with Paxlovid.

DISCUSSION

Major efforts continue to be made to develop antiviral drugs to complement vaccination-based strategies for treating patients infected by SARS-CoV-2 with the ultimate hopes of ending the COVID-19 pandemic and fortifying against future outbreaks. M^{pro} inhibitors are at the forefront of coronavirus antiviral drug development with Paxlovid (nirmatrelvir) already authorized for clinical use and several other compounds including ensitrelvir in various stages of development (*23*). However, drug resistance mutations have the potential to rapidly undermine these and future therapies. Here, we show that several naturally occurring M^{pro} variants already elicit resistance to nirmatrelvir and ensitrelvir (results summarized in **Table 1**). The highest levels of resistance for single amino acid substitution mutants are A173V to nirmatrelvir (11.6-fold) and M49I to ensitrelvir (12.4-fold). Phylogenetic analyses show that these (and other) variants have already occurred multiple independent times in different parts of the globe with regional clusters and genetic linkage providing compelling evidence for transmission.

Importantly, nirmatrelvir is a substrate-mimicking covalent drug and ensitrelvir is a non-peptide/non-covalent inhibitor and, consistent with chemically distinct mechanisms of action, our studies indicate that these drugs will have largely distinct resistance profiles. For instance, several naturally occurring amino acid substitutions in the helix adjacent to the active site (T45I, D48Y, and M49I) confer increased resistance to ensitrelvir, whereas A173V confers selective resistance to nirmatrelvir. Only one naturally occurring change, ΔP168, confers resistance to both drugs. Although structural modeling provides plausible explanations for several of these phenotypes, additional studies will be needed to establish precise mechanisms of action. For instance, ΔP168 confers similar levels of resistance to both drugs, yet structures show that ensitrelvir binds much further away.

During the preparation of this manuscript, multiple preprints reported M^{pro} mutants with resistance to nirmatrelvir (*24-28*). One of these studies selected for a virus harboring A173V during serial passage experiments in the presence of boceprevir, which also showed cross-resistance to nirmatrelvir (*25*). Coupled with our results, we suggest that A173 may be a resistance hotspot. Four of these studies focus on E166 as a hotspot, with E166V, E166A, and E166Q eliciting a range of resistance phenotypes. However, we did not pursue mutations at E166 because this residue is conserved across all available coronavirus M^{pro} sequences and is therefore likely essential for viral fitness. Accordingly, these reports contrast on how E166 variants may impact viral fitness. Nevertheless, during our analysis of M^{pro} mutants in circulating virus we did notice an E166Q allele, even after filtering-out low coverage sequences. Curiously, we found that H163W and M165Y mutations occur at the exact same frequency as E166Q in the same deposited sequences (**Fig. S8A**). We suggest that this is due to a single guanosine deletion in a poly-T stretch, which causes a frameshift after F160 and leads simultaneously to "detection" of H163W, M165Y, and E166Q in addition to a downstream stop codon (**Fig. S8B**). These findings combine to suggest that mutations at E166 may be technical artifacts of sequencing and are not circulating variants. Furthermore, E166V and E166A, which can only be identified without filtering, are found in GISAID <10 times–none of which are genetically related–providing further evidence that changes at this position are not transmissible (**Fig. S9A-B**). However, we cannot eliminate the possibility that some subset of E166 changes may be selected *in vivo*,

especially if coupled with suppressor mutations that somehow restore fitness. We therefore suggest manual inspection of full genome sequences with variants of concern, especially variants that appear at highly conserved amino acid positions.

It is presently unclear what magnitude of resistance will lead to treatment failure. Precedents with HCV NS3/4A show that single amino acid changes can elicit selective resistance of multiple orders of magnitude towards different inhibitors with minimal impact on viral fitness (*29*). However, resistance to HIV protease inhibitors typically requires two or more mutations, with single amino acid changes typically showing modest changes in inhibitor susceptibility (*30, 31*). The naturally occurring SARS-CoV-2 M^{pro} variants described here may be sufficient for resistance *in vivo* and/or serve as evolutionary stepping-stones for intermediary resistance and provide a permissive environment enabling selection of secondary mutations that confer full drug resistance.

Although circulating viruses have yet to manifest combinations of the protease-resistant mutations described here, the already high single-variant allele frequencies and the fact that most arise from a single nucleotide change combine to suggest that double mutants are inevitable. In support, combining the two changes that confer resistance to nirmatrelvir, ΔP168 and A173V, confers a near-synergistic level of resistance (51-fold versus 5- and 11-fold, respectively, for each single mutant). This result suggests that strategies should be taken to minimize the widespread development of resistance including the careful design of M^{pro} inhibitory drugs with different mechanisms of action and, by analogy to the success of combinatorial therapy for HCV and HIV-1, developing therapies that combine different drugs to simultaneously target multiple viral processes.

MATERIALS AND METHODS

Cell culture and M^{pro} reporter assays

The pcDNA5/TO-Src-M^{pro}-Tat-fLuc reporter construct has been described (*15*). M^{pro} variants were generated by site-directed mutagenesis (primers available upon request), and all mutations were confirmed by Sanger sequencing. 293T cells were maintained at 37°C and 5% CO_2 in DMEM (Gibco catalog number 11875093) supplemented with 10% fetal bovine serum (ThermoFisher catalog number 11965084) and penicillin-streptomycin (Gibco catalog number 15140122). For each M^{pro} variant, $3x10^6$ 293T cells were plated in a 10cm dish and transfected 24h later with 2µg of the corresponding Src-M^{pro}-Tat-fLuc plasmid using TransIT-LT1(Mirus catalog number MIR 2304) transfection reagent. 4h post transfection, cells were washed once with phosphate buffered saline (PBS), trypsinized, resuspended in fresh media, counted and subsequently diluted to a final concentration of $4x10^5$ cells/ml. Dilution series of inhibitors were prepared in fresh media at twice the final desired concentration of the reaction and 50µL was pipetted into a 96 well cell culture plate. 50µL of the cell suspension was added directly to the 96 well plate with inhibitor containing media to yield a final cell concentration of 20,000 cells per well. 44h after plating into 96 well plates, media was removed and 50µL of Bright-Glo reagent was added to each well and incubated at room temperature in the dark for 5m before transferring to white flat 96-well plate for measuring luminescence on a Biotek Synergy H1 plate reader. Percent inhibition at each concentration of inhibitor was derived with the formula below using the relative luminescence (RL) of an inhibitor treated sample to the untreated control.

$$\%\ inhibition = \%100 - (100/(RL)$$

Results were plotted using GraphPad Prism 9 and fit using a four-parameter non-linear regression to calculate IC_{50}. Resistance of mutants was calculated by the fold change in IC_{50} of the mutant relative to WT M^{pro}.

The *cis*-cleaving VSV based M^{pro} assay was performed as described (*19*). Briefly, 293T cells were transfected with the phosphoprotein (P)-M^{pro} fusion construct for each variant of interest and after an overnight incubation, resuspended and plated into 96 well plates. Transfected cells were then treated with the intended inhibitor and infected with VSV-ΔP-RFP at an MOI of 0.1. 48h post-infection, fluorescence was measured using a Fluoro/ImmunoSpot counter (CTL Europe GmbH, Bonn, Germany). Data were plotted as relative fluorescence using GraphPad Prism 9 and fit using a four-parameter non-linear regression to calculate IC_{50}.

Structural modeling

Structural models of candidate resistance mutants were generated using RosettaCM on the Robetta web server. The ΔP168/A173V model was generated using the nirmatrelvir bound structure of M^{pro} (PDB: 7SI9) as the comparative structure while the ensitrelvir bound structure (PDB: 7VU6) was used for the M49I mutant. Structural models of the mutant proteins were overlayed on the inhibitor bound structures using PyMOL for the respective resistance mutations to visualize inhibitor interactions with the structural models. All structural figures of M^{pro} were generated using PyMOL (Version 2.5.3, Schrödinger, LLC).

SARS-CoV-2 variant analyses

Relative distributions of amino acid changes at the amino acid positions of interest in M^{pro} were counted using the GISAID EpiCoV web server and filtered based on viral genome sequences that are considered "Complete" and "High Coverage". To generate phylogenetic trees of viral genomes containing the M^{pro} variants of interest, full length viral genomes were first retrieved from the GISAID database and filtered to exclude sequences with low coverage and analyzed using UShER. The generated phylogenetic trees were visualized using Auspice.us from NextStrain (*32*). Metadata for the viral genomes were retrieved from GISAID and overlayed on the phylogenetic tree using the Auspice.us web application.

List of Supplementary Materials

Fig S1 to S9

References and Notes

1. K. Anand, J. Ziebuhr, P. Wadhwani, J. R. Mesters, R. Hilgenfeld, Coronavirus main proteinase (3CLpro) structure: basis for design of anti-SARS drugs. *Science* **300**, 1763-1767 (2003).
2. L. Zhang *et al.*, Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved alpha-ketoamide inhibitors. *Science* **368**, 409-412 (2020).
3. B. Boras *et al.*, Preclinical characterization of an intravenous coronavirus 3CL protease inhibitor for the potential treatment of COVID19. *Nat Commun* **12**, 6055 (2021).
4. R. F. Service, A call to arms. *Science* **371**, 1092-1095 (2021).
5. C. Flexner, HIV-protease inhibitors. *N Engl J Med* **338**, 1281-1292 (1998).

6. J. Anderson, C. Schiffer, S. K. Lee, R. Swanstrom, Viral protease inhibitors. *Handbook of experimental pharmacology*, 85-110 (2009).
7. A. Luttens *et al.*, Ultralarge Virtual Screening Identifies SARS-CoV-2 Main Protease Inhibitors with Broad-Spectrum Activity against Coronaviruses. *J Am Chem Soc* **144**, 2905-2920 (2022).
8. C. H. Zhang *et al.*, Potent Noncovalent Inhibitors of the Main Protease of SARS-CoV-2 from Molecular Sculpting of the Drug Perampanel Guided by Free Energy Perturbation Calculations. *ACS Cent Sci* **7**, 467-475 (2021).
9. J. Qiao *et al.*, SARS-CoV-2 M(pro) inhibitors with antiviral activity in a transgenic mouse model. *Science* **371**, 1374-1378 (2021).
10. A. D. Rathnayake *et al.*, 3C-like protease inhibitors block coronavirus replication in vitro and improve survival in MERS-CoV-infected mice. *Sci Transl Med* **12**, (2020).
11. Covid Moonshot Consortium, Open Science Discovery of Oral Non-Covalent SARS-CoV-2 Main Protease Inhibitor Therapeutics. *BioRxiv* https://doi.org/10.1101/2020.10.29.339317 (2022).
12. D. R. Owen *et al.*, An oral SARS-CoV-2 M(pro) inhibitor clinical candidate for the treatment of COVID-19. *Science* **374**, 1586-1593 (2021).
13. Y. Unoh *et al.*, Discovery of S-217622, a Noncovalent Oral SARS-CoV-2 3CL Protease Inhibitor Clinical Candidate for Treating COVID-19. *Journal of medicinal chemistry* **65**, 6499-6512 (2022).
14. A. Telenti, E. B. Hodcroft, D. L. Robertson, The Evolution and Biology of SARS-CoV-2 Variants. *Cold Spring Harbor perspectives in medicine* **12**, (2022).
15. S. A. Moghadasi *et al.*, Gain-of-Signal Assays for Probing Inhibition of SARS-CoV-2 M(pro)/3CL(pro) in Living Cells. *mBio* **13**, e0078422 (2022).
16. X. Shu, N. C. Shaner, C. A. Yarbrough, R. Y. Tsien, S. J. Remington, Novel chromophores and buried charges control color in mFruits. *Biochemistry* **45**, 9639-9647 (2006).
17. D. W. Kneller *et al.*, Covalent narlaprevir- and boceprevir-derived hybrid inhibitors of SARS-CoV-2 main protease. *Nat Commun* **13**, 2268 (2022).
18. J. Barrila, U. Bacha, E. Freire, Long-range cooperative interactions modulate dimerization in SARS 3CLpro. *Biochemistry* **45**, 14908-14916 (2006).
19. E. Heilmann *et al.*, A VSV-based assay quantifies coronavirus Mpro/3CLpro/Nsp5 main protease activity and chemical inhibition. *Commun Biol* **5**, 391 (2022).
20. Y. Song *et al.*, High-resolution comparative modeling with RosettaCM. *Structure* **21**, 1735-1742 (2013).
21. Y. Turakhia *et al.*, Ultrafast Sample placement on Existing tRees (UShER) enables real-time phylogenetics for the SARS-CoV-2 pandemic. *Nat Genet* **53**, 809-816 (2021).
22. Y. Cao *et al.*, BA.2.12.1, BA.4 and BA.5 escape antibodies elicited by Omicron infection. *Nature*, (2022).
23. Z. Lv *et al.*, Targeting SARS-CoV-2 proteases for COVID-19 antiviral development. *Front Chem* **9**, 819165 (2021).
24. Y. Hu *et al.*, Naturally occurring mutations of SARS-CoV-2 main protease confer drug resistance to nirmatrelvir. *BioRxiv* https://doi.org/10.1101/2022.06.28.497978 (2022).
25. Y. Zhou *et al.*, Nirmatrelvir Resistant SARS-CoV-2 Variants with High Fitness in Vitro. *BioRxiv* https://doi.org/10.1101/2022.06.06.494921 (2022).
26. S. Iketani *et al.*, The Functional Landscape of SARS-CoV-2 3CL Protease. *BioRxiv* https://doi.org/10.1101/2022.06.23.497404 (2022).

27. D. Jochmans *et al.*, The substitutions L50F, E166A and L167F in SARS-CoV-2 3CLpro are selected by a protease inhibitor in vitro and confer resistance to nirmatrelvir. *BioRxiv* https://doi.org/10.1101/2022.06.07.495116 (2022).
28. E. Heilmann, F. Costacurta, A. Volland, D. von Laer, SARS-CoV-2 3CLpro mutations confer resistance to Paxlovid (nirmatrelvir/ritonavir) in a VSV-based, non-gain-of-function system. *BioRxiv* https://doi.org/10.1101/2022.07.02.495455 (2022).
29. E. Lontok *et al.*, Hepatitis C virus drug resistance-associated substitutions: State of the art summary. *Hepatology* **62**, 1623-1632 (2015).
30. N. M. King, M. Prabu-Jeyabalan, E. A. Nalivaika, C. A. Schiffer, Combating susceptibility to drug resistance: lessons from HIV-1 protease. *Chem Biol* **11**, 1333-1338 (2004).
31. A. Molla *et al.*, Ordered accumulation of mutations in HIV protease confers resistance to ritonavir. *Nat Med* **2**, 760-766 (1996).
32. J. Hadfield *et al.*, Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics* **34**, 4121-4123 (2018).

Acknowledgments: We thank members of the Harris laboratory for support and constructive feedback. RSH is an Investigator of the Howard Hughes Medical Institute and the Ewing Halsell President's Council Distinguished Chair.

Funding:

National Institute of Allergy and Infectious Disease grant U19-AI171954 (RSH, REA)
Austrian Science Fund (FWF) grant in the special call "SARS urgent funding" (EH, DvL)

Author contributions:

Conceptualization: SAM, RSH
Methodology: SAM, EH, SNM, FLK
Investigation: SAM, EH, SNM, FLK
Visualization: SAM
Funding acquisition: SAM, EH, DvL, RSH, REA
Project administration: RSH
Supervision: RSH, DvL, REA
Writing – original draft: SAM, RSH
Writing – review & editing: All authors

Competing interests: DvL is founder of ViraTherapeutics GmbH and serves as a scientific advisor to Boehringer Ingelheim and Pharma KG. The other authors have no competing interests to declare.

Data and materials availability: All data are available in the main text or the supplementary materials.

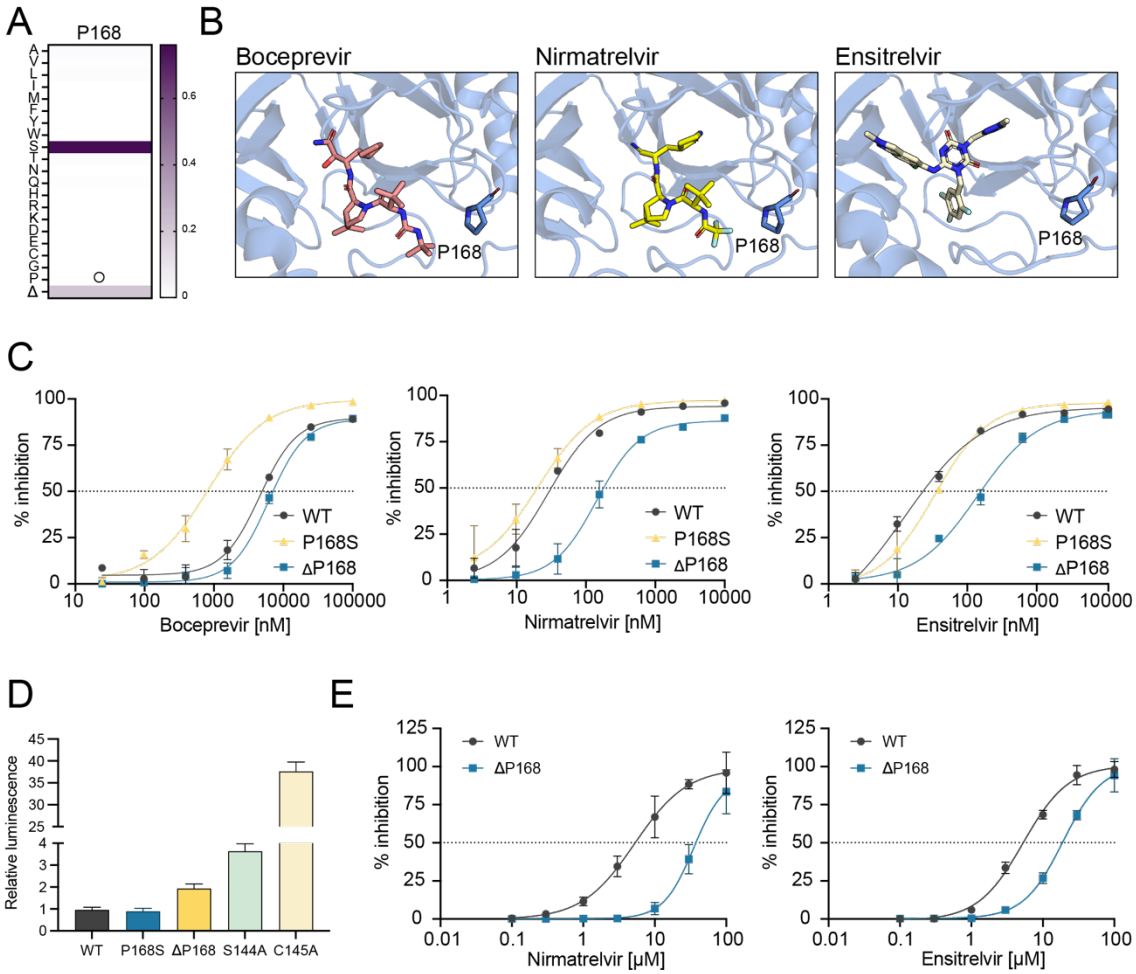

Fig. 1. **ΔP168 confers resistance to nirmatrelvir and ensitrelvir.**

(**A**) Relative frequency of amino acid changes at P168, excluding proline, in viral genomes deposited in the GISAID database as of 07-01-22.

(**B**) Co-crystal structures of SARS-CoV-2 M^{pro} in complex with boceprevir (PDB: 6WNP), nirmatrelvir (PDB: 7SI9) and ensitrelvir (PDB: 7VU6).

(**C**) Dose-response curves of WT, P168S and ΔP168 M^{pro} variants using the live cell Src-M^{pro}-Tat-fLuc assay with 4-fold serial dilution of inhibitor beginning at 10 μM (nirmatrelvir and ensitrelvir) or 100 μM for boceprevir (data are mean +/- SD of biologically independent triplicate experiments).

(**D**) Relative luminescence of cells expressing Src-M^{pro}-Tat-fLuc variants in the absence of inhibitor.

(**E**) Dose-response curves of WT and ΔP168 against nirmatrelvir and ensitrelvir in an orthologous VSV-based M^{pro} *cis*-cleavage assay (data are mean +/- SD of biologically independent triplicate experiments).

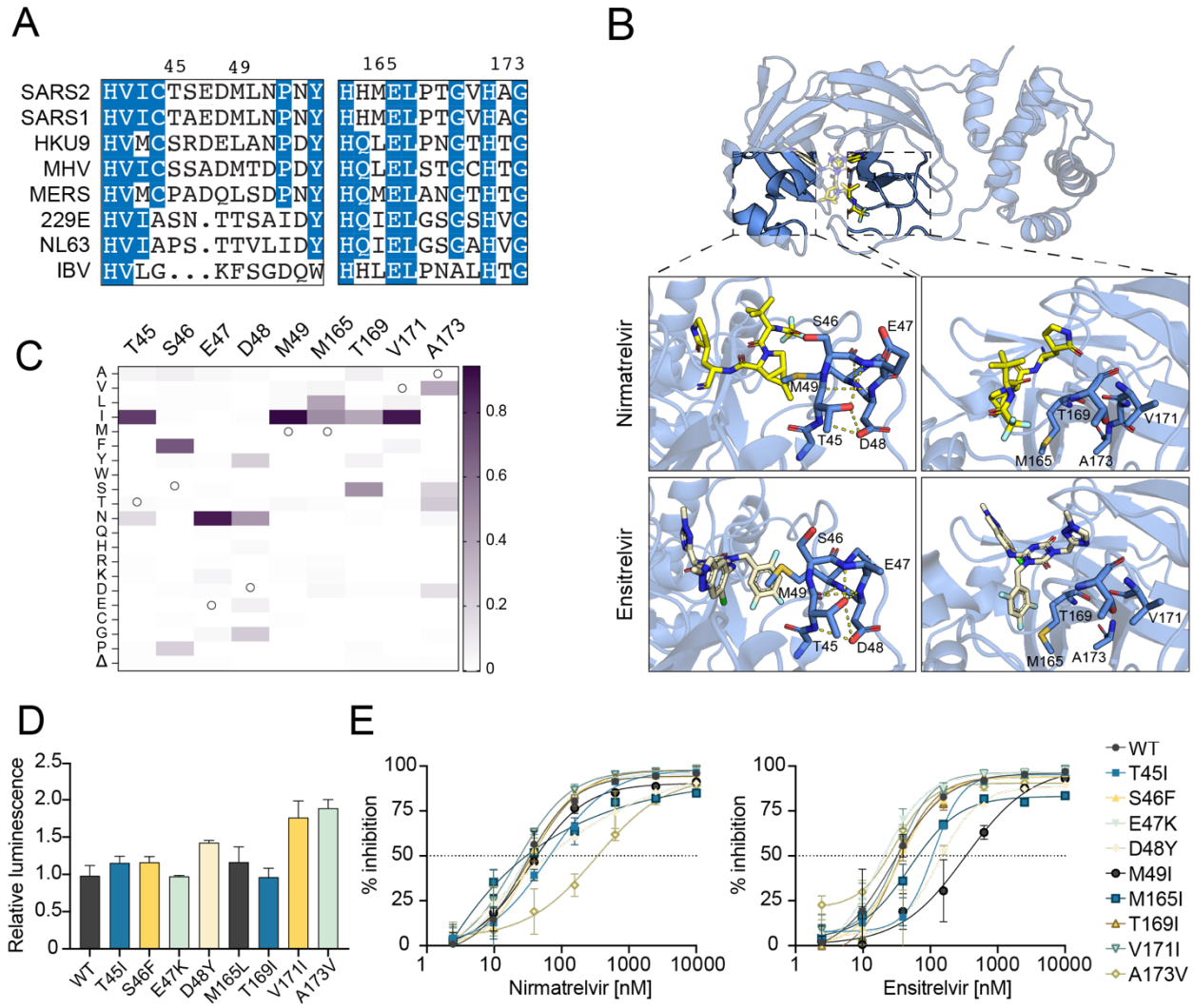

Fig. 2. Variable active site residues elicit differential resistance to nirmatrelvir and ensitrelvir.

(**A**) Alignment of coronavirus M^{pro} amino acid sequences spanning residues 41-54 and 163-174 (based on SARS-CoV-2 M^{pro} residue position).

(**B**) Structure of SARS-CoV-2 M^{pro} and inhibitors highlighting variable residues that were mutated (PDB: 7SI9 and 7VU6 for nirmatrelvir and ensitrelvir respectively).

(**C**) Relative frequency of amino acid changes at tested variable residues, excluding the respective WT residue, in viral genomes deposited in the GISAID database as of 1-July-2022.

(**D**) Relative luminescence of cells expressing Src-M^{pro}-Tat-fLuc variants in the absence of inhibitor.

(**E**) Dose-response curves of variants using the live cell Src-M^{pro}-Tat-fLuc assay with 4-fold serial dilution of inhibitor beginning at 10 µM (data are mean +/- SD of biologically independent triplicate experiments).

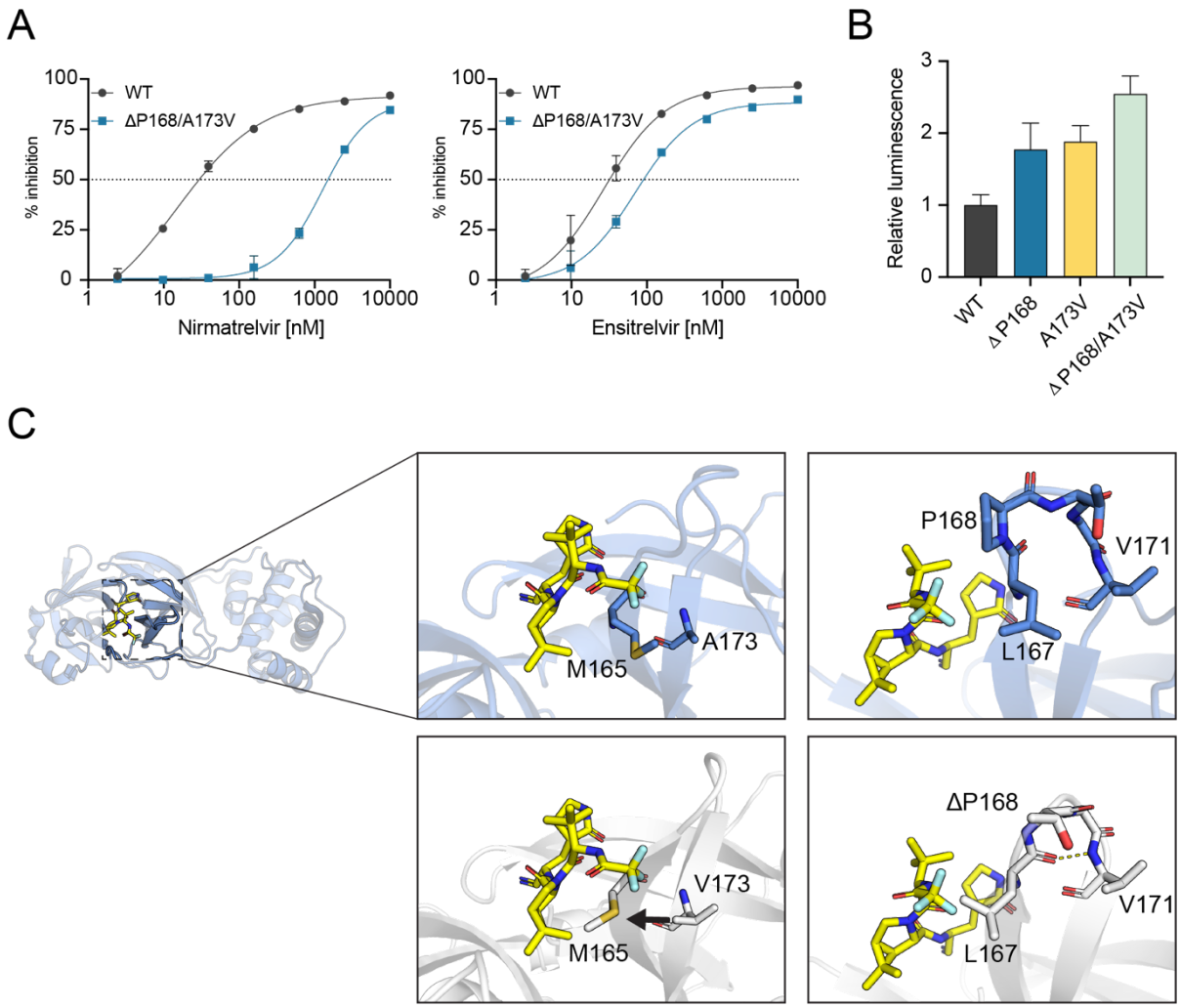

Fig. 3. ΔP168/A173V double mutant elicits synergistic selective resistance to nirmatrelvir.

(**A**) Dose-response of double ΔP168/A173V mutant vs WT using using the live cell Src-M^{pro}-Tat-fLuc assay with 4-fold serial dilution of inhibitor beginning at 10 µM.

(**B**) Relative luminescence of cells expressing respective Src-M^{pro}-Tat-fLuc variants in the absence of inhibitor.

(**C**) Structural model of ΔP168/A173V mutant SARS-CoV-2 M^{pro} generated using RosettaCM and overlayed on nirmatrelvir bound structure (PDB: 7SI9, WT M^{pro} in blue and double mutant model in white).

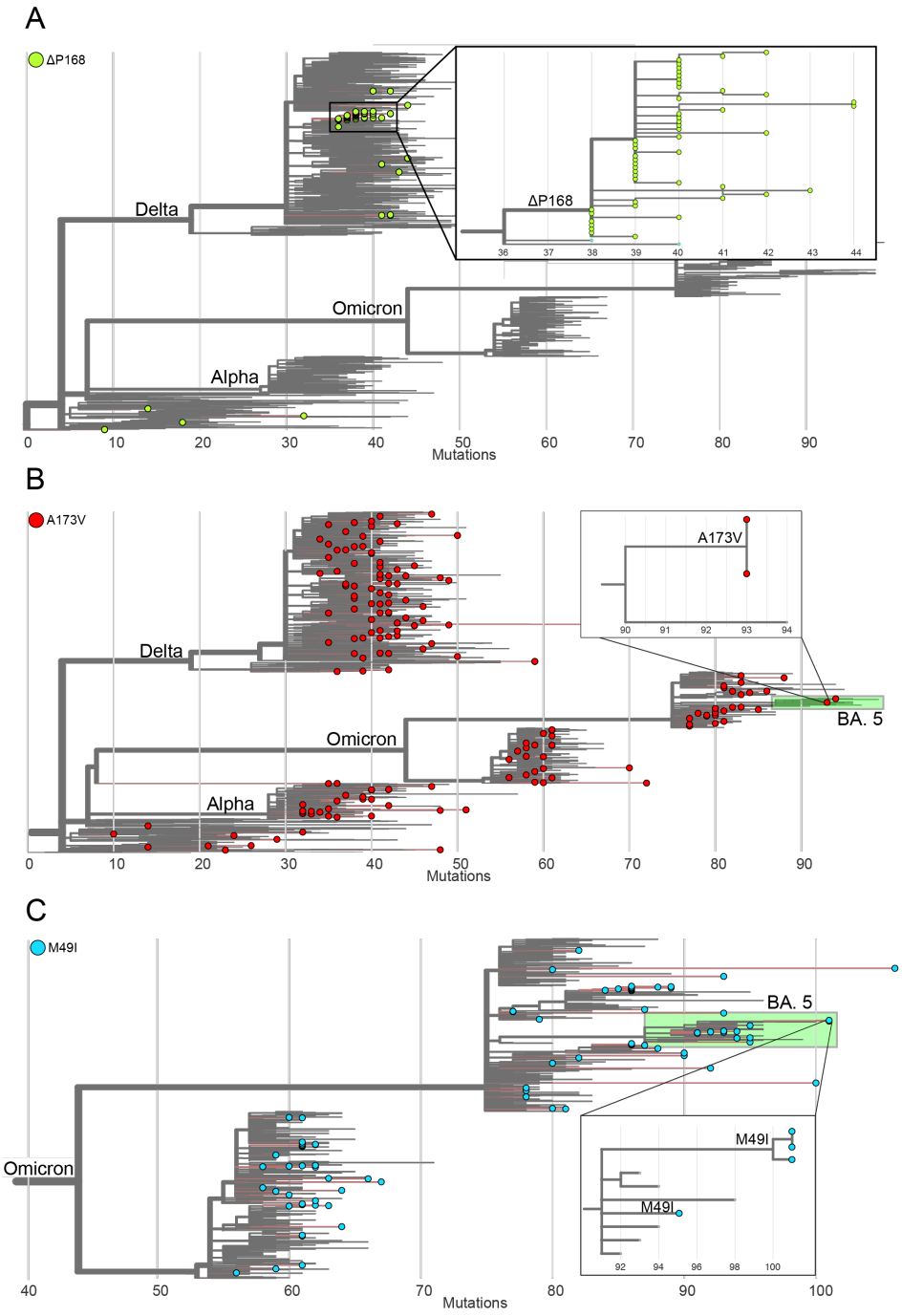

Fig. 4. Phylogenetic relationship of viruses harboring drug resistant M^{pro} variants.

(**A-C**) Phylogenetic trees containing viral genomes with ΔP168, A173V, and M49I, resistance mutations from the GISAID sequence database (5-Aug-2022). Full length viral genomes containing mutations of interest were filtered to exclude low coverage sequences, phylogenetic trees were generated using UShER and visualized with Auspice.us.

Table 1. Resistance of SARS-CoV-2 M^{pro} variants to nirmatrelvir and ensitrelvir. Fold resistance for each mutant tested were calculated based on relative IC_{50} versus WT in assays ran in parallel to limit potential variability in transfection efficiency.

M^{pro} variant	Nirmatrelvir		Ensitrelvir	
	IC_{50} nM (95% CI)	Fold resistance relative to WT	IC_{50} nM (95% CI)	Fold resistance relative to WT
WT	31.6 (26.1 to 37.7)	-	23.0 (20.4 to 25.9)	-
T45I	64.2 (46.2 to 85.7)	~2	111 (92.4 to 132)	4.1
S46F	34.5 (28.3 to 41.5)	<2	34.5 (28.3 to 41.5)	<2
E47K	27.3 (10.0 to 44.3)	<2	17.4 (12.8 to 21.8)	<2
D48Y	62.7 (44.0 to 88.9)	~2	135 (99.3 to 184)	5
M49I	33.5 (20.6 to 46.7)	<2	335 (197 to 664)	12.4
S144A	314.1 (220 to 469)	14.3	N.D.	N.D.
M165I	29.9 (22.4 to 39.5)	<2	49.3 (33.3 to 67.2)	~2
P168S	19.9 (15.5 to 25.0)	<2	36.4 (30.1 to 43.6)	<2
ΔP168	180 (154 to 219)	5.1	157 (132 to 188)	6.8
T169I	34.0 (23.7 to 45.0)	<2	30.9 (24.1 to 38.7)	<2
V171I	21.1 (13.6 to 28.3)	<2	33.4 (26.4 to 41.2)	<2
A173V	328 (186 to 796)	11.6	29.6 (24.2 to 35.4)	<2
A173T	117 (104 to 132)	4.5	34.0 (27.5 to 37.1)	<2
ΔP168 + A173V	1507 (1380 to 1660)	51	90.5 (80.9 to 102)	2.8