



HAL
open science

Study of the entry and replication of two human coronaviruses: SARS-CoV-2 and HKU1

Nell Saunders

► **To cite this version:**

Nell Saunders. Study of the entry and replication of two human coronaviruses: SARS-CoV-2 and HKU1. Virology. Université Paris Cité, 2024. English. NNT : 2024UNIP5071 . tel-04952408

HAL Id: tel-04952408

<https://theses.hal.science/tel-04952408v1>

Submitted on 17 Feb 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



UNIVERSITÉ PARIS CITÉ

École doctorale Bio Sorbonne Paris Cité (ED BioSPC 562)

Institut Pasteur

Département de Virologie

Unité Virus et Immunité

Study of the entry and replication of two human coronaviruses : SARS-CoV-2 and HKU1

Par **NELL SAUNDERS**

Thèse de doctorat d' **INFECTIOLOGIE**

Dirigée par **OLIVIER SCHWARTZ**

Présentée et soutenue publiquement le 13/09/2024

Devant un jury composé de :

CAROLINE GOUJON, DR, Université de Montpellier

NICOLAS MEUNIER, DR, Université Paris Saclay

ALI AMARA, DR, Université Paris Cité

MELANIE OTT, PROFESSOR , University California San Francisco

JULIAN BUCHRIESER, CR, Université Paris Cité

TIMOTHY WAI, DR, Université Paris Cité

OLIVIER SCHWARTZ, DR, Université Paris Cité

Rapportrice

Rapporteur

Examineur

Examinatrice

Membre invité

Membre invité

Directeur de thèse

Abstract (English)

Title:

Study of the entry and replication of two human coronaviruses: SARS-CoV-2 and HKU1

Keywords:

Coronavirus, receptor, antibody, fusion, microscopy, HKU1, SARS-CoV-2, TMPRSS2

Abstract:

To this day, seven coronaviruses have spilled over from the animal reservoir into humans. The COVID19 pandemic brought back the attention of the scientific community to these viruses. This thesis focuses on the entry and replication of two human coronaviruses, SARS-CoV-2 and HKU1. It is structured around three main axes:

- Two variants of concern of SARS-CoV-2 that emerged in 2021, AY.4.2 and the first Omicron variant (BA.1), are characterized. We found that AY.4.2 does not present major differences in fusogenicity and affinity for ACE2 compared to its parental strain the Delta variant (B.1.617.2). However, AY.4.2 is 1.3 to 3-fold less neutralized by sera from vaccinated individuals. We also observe a partial loss of neutralization by the therapeutic antibody Imdevimab. We then show that the BA.1 variant escapes neutralization by antibodies induced by vaccination and by almost all therapeutic antibodies.
- We show TMPRSS2 acts as a receptor for seasonal coronavirus HKU1, a coronavirus causing common colds. First, we show that TMPRSS2 triggers HKU1 spike-mediated cell-cell fusion and pseudovirus infection in a wide range of human cell lines. Catalytically inactive TMPRSS2 mutants do not cleave HKU1 spike but allow pseudovirus infection. There is a high affinity interaction *in vitro* between TMPRSS2 and HKU1 receptor binding domain (RBD). We use VHH nanobodies targeted against TMPRSS2 to block this interaction. We then determine the structure of HKU1's RBD in complex with TMPRSS2. Finally, we study the use of animal TMPRSS2 by HKU1 to gain further knowledge on putative hosts of HKU1.
- We study cellular remodeling following SARS-CoV-2 infection, using label free real-time images of cells infected by SARS-CoV-2 acquired with the holotomography microscope commercialized by Nanolive. We analyze these images in partnership with them, as they developed algorithms for organelle segmentation. This enables us to quantify the remodeling of different organelles during SARS-CoV-2 infection.

Our findings illustrate the various strategies adopted by human coronavirus to propagate and evade recognition by the host immune system, from entry to replication.

Résumé (Français)

Titre :

Etude de l'entrée et de la réplication de deux coronavirus humains : SARS-CoV-2 et HKU1.

Mots-clefs :

Coronavirus, récepteur, anticorps, fusion, microscopie, HKU1, SARS-CoV-2, TMPRSS2

Résumé :

À ce jour, sept coronavirus ont franchi les barrière inter-espèce pour infecter l'homme. La pandémie COVID19, a ramené l'attention de la communauté scientifique sur ces virus. Cette thèse se concentre sur l'entrée et la réplication de deux coronavirus, SARS-CoV-2 et HKU1. Elle est structurée autour de trois axes principaux :

- Deux variants préoccupants du SARS-CoV-2 apparus en 2021 (AY.4.2 et BA.1) sont caractérisés. Nous montrons notamment qu'AY.4.2 ne présente pas de différences majeures en termes de fusogénicité et d'affinité pour ACE2 par rapport à sa souche parentale B.1.617.2. Cependant, AY.4.2 est 1,3 à 3 fois moins neutralisé par les sérums des individus vaccinés que B.1.617.2. Nous observons également une perte partielle de la neutralisation par Imdevimab, un anticorps thérapeutique. Nous montrons ensuite que le variant BA.1 n'est plus neutralisé par les anticorps induits par la vaccination, ni par les anticorps thérapeutiques.
- Nous montrons que TMPRSS2 agit comme récepteur pour le coronavirus saisonnier HKU1. Tout d'abord, nous démontrons que l'expression de TMPRSS2 permet la fusion cellule-cellule médiée par la protéine Spike de HKU1 et l'infection par le pseudovirus dans un large éventail de lignées cellulaires humaines. Les mutants catalytiquement inactifs de TMPRSS2 ne clivent pas le spicule de HKU1 mais permettent l'infection par des pseudovirus HKU1. Il y a une interaction à haute affinité in vitro entre TMPRSS2 et le domaine de liaison au récepteur (RBD) de HKU1. Nous utilisons des nanocorps ciblant TMPRSS2 pour bloquer cette interaction. Nous déterminons ensuite la structure du RBD de HKU1 en complexe avec TMPRSS2. Enfin, nous étudions l'utilisation des TMPRSS2 de différentes espèces par HKU1.
- Nous étudions le remodelage cellulaire induit par l'infection par SARS-CoV-2. Nous utilisons le microscope holotomographique commercialisé par Nanolive, qui nous permet de faire des images en « label-free ». Un collaborateur de Nanolive a développé des algorithmes pour segmenter les organelles. Cela nous permet de quantifier l'évolution de différentes organelles pendant l'infection par le SARS-CoV-2.

Nos résultats illustrent les différentes stratégies adoptées par les coronavirus humains pour entrer dans les cellules et se répliquer, tout en échappant à la reconnaissance par le système immunitaire.

Résumé substantiel (Français)

À ce jour, sept coronavirus ont franchi les barrières inter-espèce pour infecter l'homme. Cette thèse se concentre sur l'entrée et la réplication de deux de ces coronavirus, SARS-CoV-2 et HKU1. SARS-CoV-2 est le virus responsable de la pandémie COVID19. HKU1 est un coronavirus saisonnier peu étudié responsable de rhumes, dont l'origine et la date d'émergence dans l'espèce humaine sont inconnues. J'utilise les outils développés dans le laboratoire pour l'analyse de SARS-CoV-2 afin d'étudier l'entrée et la réplication d'HKU1. Une introduction générale sur les coronavirus ouvre cette thèse. Mes travaux sont ensuite structurés autour de trois axes principaux.

1) Étude des variants préoccupants de SARS-CoV-2

Lors de la pandémie COVID19, des virus mutés – des variants – ont commencé à apparaître, et remplacer progressivement le virus initial. La transmission du virus se déroule en trois étapes : le relargage du virus par l'individu infecté, la survie et propagation du virus dans l'environnement et l'infection de nouveaux hôtes. Ainsi, des mutations augmentant l'entrée et la réplication du virus dans les cellules augmenteront la charge virale. Des mutations aggravant les symptômes permettent une production accrue d'aérosol. Des mutations affectant la stabilité du virus dans les aérosols et sur les surfaces permettent une survie prolongée dans l'environnement. Enfin des mutations qui confèrent un échappement à la réponse immunitaire innée et adaptative, ou une résistance aux traitements utilisés permettent au virus de persister plus longtemps chez les individus infectés, ou de réinfecter des individus préalablement infectés.

Delta (B.1.617.2) était le variant dominant de juin à décembre 2021 ; il possède 9 mutations dans son spicule par rapport à la souche ancestrale. Certaines permettent l'échappement de la reconnaissance par des anticorps neutralisants, d'autres augmentent la stabilité du spicule, sa liaison au récepteur cellulaire ACE2 et sa capacité à induire la fusion membranaire. Delta a donné naissance à des sous-lignées dont AY.4.2, qui possède trois mutations supplémentaires dans le domaine N-terminal de son spicule. En novembre 2021, AY.4.2 représentait 15% des séquences au Royaume-Uni et se répandait mondialement. Nous avons donc décidé d'étudier l'effet de ces trois mutations.

Nous montrons notamment qu'AY.4.2 ne présente pas de différences majeures en termes de fusogénicité et d'affinité pour ACE2 par rapport à sa souche parentale Delta. Cependant, AY.4.2 est 1,3 à 3 fois moins neutralisé par les sérums des individus vaccinés que Delta. Nous observons également une perte partielle de la neutralisation par Imdevimab, un anticorps thérapeutique. Avec ces travaux, nous montrons que des mutations présentes dans la partie N-

terminale du spicule affectent également la reconnaissance du domaine de liaison au récepteur (RBD), qui est le domaine ciblé par Imdevimab.

En novembre 2021, le variant Omicron (BA.1) est apparu. BA.1 est devenu prédominant mondialement en moins d'un mois, supplantant à une vitesse inégalée les autres variants. BA.1 possède 30 mutations dans son spicule par rapport à la souche ancestrale, alors que les variants précédant en comportait une dizaine.

Nous montrons que ce variant échappe à la neutralisation par les anticorps générés suite à 2 doses de vaccins, le schéma vaccinal majoritaire en Décembre 2021. Les sérums de patients ayant reçu un rappel 6 mois post-première dose ou ayant été vacciné après une infection ont une activité neutralisante contre BA.1 un mois post-rappel, bien qu'elle soit 6 à 20 fois plus faible que contre Delta, soulignant l'importance de ce rappel. BA.1 n'est plus lié ni neutralisé par un panel d'anticorps thérapeutiques. Sotrovimab présente une perte de neutralisation d'un facteur 3, Cilgavimab et Adintrevimab d'un facteur 20. Les 6 autres anticorps thérapeutiques ne sont plus neutralisants aux doses testées, et n'ont donc plus d'utilité clinique contre ce variant.

D'autres travaux du laboratoire et d'autres équipes montrent que durant les deux premières années de la pandémie des variants plus transmissibles ont été sélectionnés. En novembre 2021, il y avait environ 400 millions de cas cumulés déclarés mais les experts estiment que plus de 3 milliards d'infections avaient eu lieu. En outre, la couverture vaccinale était d'environ 60%. L'échappement aux réponses immunitaires induites par la vaccination et les infections préalables ont contribué à l'avantage du variant Omicron BA.1 par rapport à Delta et ces sous-variants. Ces derniers présentaient certes des mutations d'échappement, mais étaient tout de même neutralisés par des individus vaccinés ou préalablement exposés. Depuis, certaines des mutations observées dans les variants pré-Omicron, dont celles que nous avons caractérisé dans nos travaux sur AY.4.2 sont réapparues indépendamment dans des nouveaux variants d'Omicron, confirmant l'avantage qu'elles conféraient. Les études parues sur les variants qui n'ont cessé d'émerger depuis seront discutées.

2) Identification de TMPRSS2 comme le récepteur d'HKU1

Les coronavirus saisonniers sont responsables de 15 à 30% des rhumes (pré-COVID). HKU1 est l'un de ces quatre coronavirus saisonniers endémiques. Il a été découvert en 2005 mais sa date initiale d'émergence dans l'espèce humaine est inconnue. Le virus le plus proche retrouvé dans le réservoir animal est le virus de l'hépatite murine (MHV), mais on ne sait pas si d'autres hôtes ont joué un rôle lors du passage à l'espèce humaine. La séropositivité à HKU1 chez les adultes est estimée à 95% en France.

Pour rentrer dans les cellules hôtes, le spicule d'HKU1 s'attache d'abord via les acides sialiques 9-O-acétylés. Cela permet une ouverture du spicule, et l'exposition du RBD qui était préalablement enfoui. Ensuite, le RBD se lie à un récepteur protéique non-identifié. Enfin, le spicule est clivé par des protéases, soit à la surface (TMPRSS2/4/11D, MMP2/9, KLK13 ...) soit dans les endosomes (cathepsines). Cela permet une projection du peptide de fusion et une fusion des membranes cellulaires et virales.

Nous montrons qu'en plus de cliver HKU1, TMPRSS2 agit comme récepteur pour ce virus. L'expression de TMPRSS2 permet la fusion cellule-cellule médiée par le spicule d'HKU1 et l'infection par des pseudovirus HKU1 dans un large éventail de lignées cellulaires humaines. Les mutants catalytiquement inactifs de TMPRSS2 ne clivent pas le spicule de HKU1 mais permettent l'infection par des pseudovirus HKU1. Le spicule peut en effet être clivé par d'autres protéases à la surface ou dans les endosomes.

Nous isolons du virus infectieux HKU1 à partir de prélèvements nasopharyngés de patients infectés, en utilisant des lignées bronchiques primaires différenciées à l'interface air-liquide. Nous étudions l'effet de différentes drogues et conditions de culture dans ce système. Nous testons également différentes lignées immortalisées afin d'isoler le virus, sans succès.

Nos collaborateurs, Dr Ignacio Fernandez et Prof. Félix Rey, montrent une interaction à haute affinité *in vitro* entre TMPRSS2 et le domaine de liaison au récepteur (RBD) de HKU1. L'équipe du Dr Pierre Lafaye a développé des nanocorps ciblant TMPRSS2 par immunisation d'un alpaca. Nous caractérisons ces nanocorps et les utilisons pour bloquer l'interaction du RBD de HKU1 avec TMPRSS2. Nous démontrons l'efficacité d'un nanocorps pour bloquer l'infection par un isolat infectieux d'HKU1 sur des cellules bronchiques primaires différenciées à l'interface air-liquide.

Dr Ignacio Fernandez, Dr Stéphane Duquerroy et Prof. Félix Rey déterminent la structure du RBD de HKU1 en complexe avec TMPRSS2. Nous caractérisons l'importance des résidus situés à l'interface du complexe, à la fois sur la fusion cellulaire, l'entrée de pseudovirus et l'interaction du spicule et de TMPRSS2. Nous étudions l'utilisation des TMPRSS2 de différentes espèces de mammifères par HKU1. Déterminer les résidus à l'interface entre TMPRSS2 et HKU1 nous permet de prédire et valider la capacité des TMPRSS2 de différentes espèces à interagir avec HKU1.

Nos travaux sur TMPRSS2 et HKU1 permettent de mieux comprendre l'étape d'entrée du cycle infectieux d'HKU1. Néanmoins il y a potentiellement d'autres étapes que nous ne comprenons pas, étant donné notre incapacité à ce jour à cultiver du virus infectieux sur les lignées humaines surexprimant le récepteur. Nous aborderons différentes expériences non publiées sur la culture de ce virus.

Ce rôle de récepteur pour TMPRSS2 n'a jamais été montré auparavant. TMPRSS2 est impliqué dans le clivage de nombreux virus respiratoires (coronavirus, influenza, para influenza, paramyxovirus, métapneumovirus, virus Sendai...). Nous discutons de sa possible implication en tant que récepteur pour d'autres virus.

3) Utilisation de la microscopie « label-free » pour étudier le remodelage cellulaire induit par SARS-CoV-2

Afin de regarder les modifications subcellulaires, la microscopie à fluorescence est la technique la plus communément utilisée. Elle permet d'imager différentes organelles et protéines simultanément, grâce à des agents de contraste, des protéines fluorescentes ou des anticorps. Les signaux générés ont un bon contraste et sont simples à analyser. Les progrès techniques permettent également d'obtenir des images à des résolutions en dessous de la longueur d'onde. L'imagerie par fluorescence présente néanmoins différents désavantages : les échantillons ne sont pas dans leur environnement natif en raison des agents de contrastes ou fluorophores qui peuvent interférer avec des processus biologiques et il peut y avoir de la phototoxicité ou du photoblanchiment si l'on image sur des longues durées.

Les techniques de microscopie classiques « label-free », telles que la microscopie à fond clair, à contraste de phase ou à contraste interférentiel différentiel sont intrinsèquement limitées par le manque de contraste des échantillons biologiques. Certaines approches basées sur l'intelligence artificielle permettent maintenant d'obtenir une segmentation subcellulaire, ce qui n'était pas possible auparavant.

D'autres méthodes label-free permettant un meilleur contraste ont été développées dans les dernières années, dont l'holotomographie et la microscopie Raman. L'holotomographie est basée sur la mesure de l'indice de réfraction des échantillons et permet de distinguer les composants cellulaires dont l'indice de réfraction est suffisamment différent de celui du cytosol : les gouttelettes lipidiques, les mitochondries, les membranes nucléaires et cellulaires et les nucléoles. Leur segmentation n'est pas aussi directe que les résultats obtenus par microscopie à fluorescence. L'entreprise qui commercialise le microscope que nous utilisons, Nanolive, développe des algorithmes pour la segmentation des différentes organelles, avec nos données. Cela nous permet de quantifier l'évolution de la masse sèche, de la taille et de la localisation de différentes organelles au cours du temps. Nous couplons cette approche avec de la microscopie à fluorescence et de la microscopie électronique afin de valider les prédictions du modèle.

Cette nouvelle méthode est utilisée afin d'étudier les conséquences de l'infection par SARS-CoV-2. Nous montrons que l'infection par le SARS-CoV-2 et la formation de syncytia induit

un remodelage cellulaire important. Notre étude est descriptive et permet principalement de valider l'utilisation de ce microscope et de ces algorithmes ; elle n'a pas une visée d'explication moléculaire des phénomènes observés. Néanmoins un de phénomènes observés est la forte accumulation de gouttelettes lipidique dans les cellules infectées. Ces gouttelettes ont déjà été observées par d'autres équipes.

Nos résultats illustrent les différentes stratégies adoptées par les coronavirus humains pour entrer dans les cellules et se répliquer, tout en échappant à la reconnaissance par le système immunitaire de leur hôte. Bien que ces travaux soient principalement axés sur l'entrée du virus et la protéine de spicule du coronavirus, le rôle d'autres protéines virales est également abordé.

Acknowledgments

Je souhaite remercier les **rapporteurs et membres du jury** de prendre le temps de revoir mes travaux de thèse. J'espère que vous apprendrez quelque chose en la lisant. J'ai pris un plaisir que je n'aurai pas soupçonné à écrire cette thèse, tout en essayant d'être succincte sur toutes les parties sauf ces remerciements, ils font 6 pages, n'hésitez pas à les sauter.

Olivier. Merci de m'avoir accepté dans ton laboratoire pour ce stage de M2, en plein COVID, et de m'avoir fait confiance pour me présenter à des bourses de thèses alors que ça ne faisait qu'un mois que j'étais au labo. Merci de m'avoir poussée à effectuer des raisonnements scientifiques toujours plus poussés et rigoureux. Ces trois belles années et demie ont été marquées par des discussions passionnantes. Tu as toujours eu du temps pour discuter de résultats quand j'en avais besoin. On n'était pas forcément d'accord sur tout, mais la science naît aussi du désaccord, car ça m'a incité à faire des manips pour te prouver que mon hypothèse était juste (ou fausse haha). J'admire ta capacité à savoir quand insister, et quand laisser du mou, dans des moments où j'avais besoin de plus d'espace. Je te remercie également pour tout le temps que tu as passé sur des demandes de financement pour qu'on puisse faire de la belle science sans se poser la question des moyens. J'ai appris énormément sur le monde de la publication et de la diplomatie où tu excelles, et j'espère un jour être aussi balèze que toi là-dessus. J'ai apprécié tes petites touches d'humour, nos discussions les midis ou tu venais manger avec nous, même si parfois tu voulais trop parler science ;p. En somme, même si de temps en temps je me plaignais auprès de mes chers collègues (je suis quand même française), je n'aurai pas pu rêver d'un meilleur directeur de thèse, alors MERCI du fond du cœur.

I would like to thank you, **Tim** for your adaptability throughout my PhD. Although it didn't turn out quite as mitochondrial as you and I expected, you have given me precious training, professional and personal advice throughout my PhD. You are a great mentor, and I really appreciated your role in my PhD. I wish I had a second PhD to carry out some projects to completion, but Yining is doing a great job on ORF9b, and I am glad I could be of some help.

Maintenant j'aimerais remercier tous les membres de l'UVI, désolée c'est long et peu structuré. Ces trois années sans vous c'est même pas imaginable, nos repas, nos pots, nos rires, nos cafés et nos manips. **Julian**, quel plaisir de travailler avec toi pendant ces trois années. Tu m'as appris beaucoup d'un point de vue expérimental, mais aussi sur les raisonnements scientifiques, les figures, les présentations, et ce qui est important à considérer pour la suite de ma carrière. J'ai également apprécié ta gentillesse, tes tours de magie, discuter de tout et de rien, tes conseils

en jeu, et ton soutien quand ça n'allait pas. **Delphine** puis **Isa**, on a bien travaillé et rigolé dans ce bureau, et merci d'avoir supporté mon bordel (et **Isa** ma musique). **Delphine**, merci pour ton soutien ces trois années, j'ai apprécié nos conversations sur à peu près tout et n'importe quoi, la science, les femmes, les problèmes, l'avenir... **Flo**, tu es aussi bavarde que moi, alors évidemment qu'on allait bien s'entendre. Merci pour ta présence (au labo et aussi à Montigny ;)), nos conversations, tes gestes attentionnés ces trois dernières années, l'aide sur les manip et les prods. **Françoise**, merci pour ta gentillesse, ton écoute, ton aide et le travail immense que tu fais pour nous tous, et désolée je ne mets toujours pas ma blouse quand je suis dans le rush (souvent) ^^ **Nico** merci pour les cadeaux spontanés, les gâteaux, les conseils scientifiques et pro, et désolée encore pour les cristaux de NaCl. **Tim**, merci pour tout le temps que tu as pris pour aider une petite thésarde perdue dans sa vie à faire des choix, pour ta dispo à peu près tout le temps pour me conseiller et papoter. **Will** (qui ça qui ça) thanks for sharing my daily PhD life, withstanding my doubtful humor and my chitchat, it was great sharing this PhD with you. **Mariam**, la force tranquille du labo, j'ai apprécié t'avoir à mes côtés pendant notre stage de M2, puis la thèse, parler d'encadrement, de vie et d'avenir. **Martin**, the big Austrian dude, you brought a lot of fun (and waffles) to the lab. **Mathieu**, c'est une grande fierté pour moi d'avoir réussi à te faire marcher plusieurs fois, en trois ans, tu n'as pas réussi à me faire aimer le foot, mais tu étais néanmoins de bonne compagnie, quand tu ne consommais pas tout nos réactifs sans en recommander :p. **Jeanne**, en un mois au labo, c'est comme si tu avais toujours été là. Qu'est-ce que j'ai apprécié nos discussions sur la vie, la thèse, la fantasy, ton écoute quand je radotais parce que j'étais énervée, et aussi merci pour tout ce que tu m'as appris sur la pop-culture Corrèenne, on sait jamais, ça peut servir ! **Andréa**, merci d'avoir partagé cette fin de thèse avec moi, tu as illuminé ces derniers mois au labo. **Raphaël**, voici tes acknowledgements tant attendus ;p thank you to Raphaël Jeger-Madiot for constantly disturbing us in our office to talk about nothing or get gum, but bringing cookies in exchange. **Elodie** and **Mike**, thanks for initially training me in the lab, and sharing those first few months in the lab with me. **Ludivine**, merci pour toute l'aide que tu m'as apporté et pour cette première année de thèse passée à tes côtés. **Jocelyne**, merci pour tout ce que tu fais au quotidien pour nous rendre la vie plus simple. **Isma** et **Fabienne**, merci pour votre aide administrative ces trois années. **Tim**, **Mathieu**, **Lisa** et **Andréa**, merci pour vos questions qui nous poussent tous à faire de la plus belle science. **Delphine**, **Tim**, **Nico** et **Julian**, merci pour les relectures diverses. **Flo** et **Nico**, merci de votre investissement dans Vir'o Vert, on en a besoin ! Je n'ai pas réussi à prendre le temps ces dernières années, et ça j'en suis désolée. **Tim** et **Julian**, grâce à vous j'en ai appris beaucoup sur la vie de jeunes parents, et ce qui est sûr c'est que je vais attendre quelques années encore ! **Alice**, **Hunter** et **Eva**, merci d'avoir été mes stagiaires et d'avoir supporté mon encadrement un peu bordélique, j'espère que vous avez apprécié ces mois au labo. Enfin merci à Stacy, Barbara, Jeremy, Remi, Jérôme, Samuel,

Hippolyte, Donatella, Lou-Léna, Augustin, Chloé, Yan, Ivanka, Amélie, Mathilde et Mathilde qui ont partagé un bout de chemin au labo avec moi.

Je voudrais également remercier les membres du Mitochondrial Biology Group. **Claire**, merci pour tes conseils et ta gentillesse durant ces trois années, c'était un réel plaisir de te voir quand je traversais la rue. **Martin** thank you for this year we shared. You did a tremendous amount of work, I enjoyed talking and working with you, and I have very good memories of us in the P3 the whole day waiting for the Oroboros to stabilize. **Yining** and **Juan-Diego**, I enjoyed working with you guys. **Elodie**, **Erminia**, **Marcio** and **Marie**, c'était un plaisir de vous rencontrer et de passer du temps avec vous.

Now is time to thank all the people who collaborated with us, and without whom this work wouldn't have gone far. **Ignacio**, thank you for all your hard work these last two years, it was a pleasure working with you; you are efficient and fun. You stuck with my one thousand requests, and I am so glad we got to work together, hope it will happen again in the future. **Felix**, thank you for all your input and work, I have learned a lot and you have made me see the beauty of structures. **Mathieu**, merci pour ton travail colossal et simplement magnifique sur les images du Nanolive. J'ai beaucoup apprécié nos réunions et nos discussions au long de ces trois années, et confronter notre vision biologique à tes descriptions mathématiques de ce qui se passe dans les cellules. C'était un plaisir d'enfin te voir en vrai cette année. **Cyril**, merci pour les anticorps, les ELISA, les protéines, mais surtout pour nos cafés et tes conseils. **Hugo**, merci pour tes inputs lors de nos réunions sur HKU1, et les anticorps. **Vincent**, merci pour tes images sublimes et tout ton travail sur les épithélix, malgré mes échantillons parfois mal annotés. **Nadège**, merci pour ces belles images de microscopie électronique, cette corrélation Nanolive-EM c'était pas évident, et désolée d'avoir parfois oublié nos rdv. **Catherine**, merci pour toutes les prods de pseudovirus. **Etienne**, **Matthieu** et **Artem** merci pour les séquençages, que ce soit de l'eau, de l'Herpes (oops) ou HKU1. Merci à tous les **collaborateurs hospitaliers** qui nous ont fourni des échantillons nasopharyngés pour nos isolations d'HKU1 (on ne retient que les derniers qui ont marché, mais y'en a eu pas mal qui ont échoué avant aussi), de SARS-CoV-2 et des sérums pour nos études. Merci à toute la **plateforme d'imagerie** pour les formations et l'aide au cours de ces trois années, et plus particulièrement **Nathalie**, **Nassim**, **Julien** et **Christelle**.

Merci à toutes les belles personnes que j'ai rencontré à l'étage ou à Pasteur, et avec qui j'ai partagé des déjeuners, des cafés, des conversations d'ascenseur ou de SAS, des discussions scientifiques. Selen (nos vies de thésardes), Philippe (surtout nos discussions de gauchistes sur l'écologie ou le féminisme), Kyle (my american fellow in Pasteur), Felix and Angelianna (thanks

for your work as PhD reps, and also just for being fun), Nolwenn, Ségo, Maxime, Vincent, Jeanne (ma go to pour les questions ED), Patricia Phiphi, et Pierre. Merci aux **Enragés**, Juliette, Lexi, Sol, Stephane, Val, Louis, Manu, Michael, Samuel, Markel, Kamouk, c'était fun ces deux années à vos côtés.

J'aimerais aussi remercier ceux qui ont planté des graines, d'une manière ou d'une autre lors de mon parcours. Mes profs de lycée et de prépa qui m'ont donné des bonnes bases scientifiques et de la rigueur, **M. Bouvry**, **M. Kaczmarek**, **M. Schlösser**, **M. Vanderbrook**, **M. Frajman**. Tous les encadrants des IChO, **Hakim**, **Aurélien**, **Lucas**, **Elise**, et **Clément**, pour tout ce que vous m'avez apporté, et surtout **Clément** pour ton amitié. Mes profs de l'X, qui m'ont fait redécouvrir la bio en bac+3 alors qu'après le lycée je m'étais dit plus jamais. **Nicolas Barraud** et **Francesco Villa**, qui m'ont accueilli dans leurs équipes pour des stages. **Marion**, cette première expérience à tes côtés m'a fait tomber amoureuse de la recherche ; merci de m'avoir formé et d'être devenue une amie.

Merci à tous mes amis d'avant, qui m'ont suivi de proche ou de loin dans cette thèse et aussi dans tout l'avant qui m'a permis d'arriver jusqu'en thèse. Vous avez supporté mes lapins, mes retards, mes crises existentielles, mes silences radios, mes absences, désolée ; j'espère vous avoir apporté en échange de la joie et un brin de folie. **Louise**, nos vendredis soirs pizza depuis toujours, **Edwige**, tu as été une des premières avec qui j'ai parlé de tout et de rien, **Claracha**, pour ton soutien inconditionnel et nos rires qui nous ramènent en enfance, **Tess**, je suis tellement heureuse de t'avoir dans ma vie et j'aime le fait que ce soit si simple de te parler, **Marion**, tu es une des personnes les plus attentionnée que je connaisse et j'adore les moments où l'on se retrouve, **Lisa**, pour ces soirées **Delphine**, on s'est toujours posé un peu les mêmes questions, heureusement que t'étais là pour en discuter, **MH**, tu m'aides toujours dans mes questionnements sur la recherche, l'environnement les femmes et la société, par tes démarches et tes lectures, en plus d'être de bonne compagnie ;) la team des expats du Volley, **Chloé**, on en a parcouru du chemin depuis ce premier projet de bio, merci pour ton amitié à l'X, et même si on se voit moins depuis, j'espère qu'on continuera à se voir, **Caro** depuis la sup, tu as toujours été là pour moi, **Phiphi**, nos kawas, **Paul** et ton humour un peu lourd, **Gautier**, nos discussions sur la politique, **Dodo**, ah ces chorés, **Ben**, toutes ces après-midi jeu, **Camille**, ton soutien quand ça n'allait pas et tous ces beaux moments passés ensemble, **Nico**, et toutes les fois où tu m'as encouragé même si je jouai faux, **Léo**, en retenant le quart de la moitié du commencement de tout ce que tu m'as raconté je suis sacrément moins bête, **Quentin**, ces tournois de volley, ces discussions dans ton casert et ces parties endiablées de jeu interminables, **PE** t'auras le droit à ton propre paragraphe un peu plus tard. La willy, **Nico**, l'autre électron très fidèle, boum boum sur la grosse caisse, nos rocks endiablés, ton comique (parfois un peu

trop ?) de répétition, Alix, Omar, Jo, Hugo et Caro, ces vacances et soirées à vos côtés. Le **PBB, Louis, Louis, Tristan, Théo, Nico, Rémi**, ma première expérience en fanfare, on était tout de même assez médiocres mais on rigolait bien. Les copains de Lausanne, nos sorties en montagnes et nos pyj party ont quand même bien égayé cette année COVID. **Chloé** encore et toujours ;), **Auriane**, la spécialiste pour me proposer des plans randoms et me sortir de ma zone de confort, **Eloi**, entre BforCure et Pasteur on en a passé des heures à discuter de tout et de rien, **Thib**, et tes bons plans éternels.

Ces trois années de thèse ont aussi été remplies d'activités et de nouvelles (re)rencontres. Que de belles aventures avec les ~~Toucan-Boucan~~ Canards Scéniques (coïn coïn). Merci à tous les membres du bureau qui ont donné tellement de leur temps pour fonder cette fanfare et faire qu'elle tourne, et qu'on soit devenu aussi RAS. Nos weekends et nos répés m'ont permis de pouvoir devenir amie avec certains d'entre vous et j'espère que ça va durer. **Nico** un plaisir de partager tous ces moments avec toi, **Théo**, merci pour tous les fous rires, **Raph**, désolée pour les fausses notes et les notes pas en places, merci pour ta patience, **Loupio**, merci pour les encouragements, les calins et les discussions, **Grégoire**, un pote très utile ;p, **Rémi**, nos we à Gre, **Victor** et **Côme**, pour nos papotages en weekend, et tous les autres.

Les SSA girls et SSA boys, merci pour TOUT. Un esprit sain dans un corps (pas toujours) sain (j'en profite au passage pour remercier mes kinés, Claire, Damien et Florian). En débutant il y a trois ans, je ne savais pas que je deviendrai aussi mordue. Merci tout d'abord à **Alex**, tout le **CA** et surtout toute **cette team de bénévoles** du feu de dieu pour tout ce que vous faites pour ce club, qu'est-ce que j'ai aimé m'engager à vos côtés. Merci aussi **Alex** de m'avoir convaincue de devenir adjointe, puis coach, j'ai beaucoup appris aussi en étant entraîneur, et ça m'a permis de garder le contact quand j'étais blessée. Merci à toutes mes adjointes quand j'ai été coach, votre bonne humeur, vos feedbacks, vos lancés, **Ali, Deborah, Chouchouette, Elli** et **Tatiana**. Merci à tous mes coaches, **Igor, Manu, Johan, Maxime, Tristan** et **Thor** qui m'ont fait quand même sacrément progresser. Merci à **la team place d'It** pour avoir animé tous ces retours, Nico, Robin, Julien, Nari et Tatiana. Merci à tous mes partenaires d'un tournoi ou plus, d'entraînements et ami.e.s au cours de ces années ; **Lisa**, nos soirées, nos vacances, l'universitaire, quand je te voyais galérer avec ta thèse, ça me rappelait à quel point on est quand même chanceux en bio, **Chouchouette**, la seule qui joue mieux quand elle sort la veille, **Laurène**, et l'accueil à Montpel.

Je crois qu'il est enfin temps de remercier ma famille. Mom, dad, thanks for your unconditional love all these years. **Maman**, merci d'avoir toujours accepté la personne que j'étais sans questionnement, d'avoir passé tant de temps à t'occuper de moi et d'avoir répondu aux cent-

milles questions que je posais à longueur de journée. Merci sur ces dernières années de m'avoir proposé régulièrement des places de spectacles afin que la culture et la beauté continuent de s'immiscer dans mon esprit si rationnel. **Dad**, one thing you taught me well is to say thank you, I hope these 6 pages are to your liking. Thanks for understanding how my minds work non-stop on a problem until I find the solution and for pushing me to be always better, and bragging all the time about me it does make me feel nice. That should be an easy one, d1tU4e&IIUvmd2. **Emma**, thanks for taking care of me and always making me feel loved even though I'm shitty at long distance relationships & asking questions. Thanks for being a great mediator in our family, merci de venir me titiller dans ma zone de confort, de m'aider à changer d'avis sur tout, et à exprimer mes sentiments. **Sarah**, je suis tellement heureuse de la relation qu'on a pu développer au cours de ces années à Paris, merci pour tous ces moments qu'on a partagé, ces sauts de confiance quand je te dis, ça te dirait d'aller en Slovénie avec moi, ton écoute, parfois aussi les moments où tu me remets en question, parce que ça me fait devenir une meilleure personne. Merci **Franswaz**, d'avoir toujours cru en mes rêves, mêmes s'ils étaient un peu saugrenus pour toi, et pour tes petites attentions. Merci **Léo**, pour les rires, les belotes et les cafés, **Patrick** pour les jeux de mots plus ou moins réussis et la musique, **Pierre** pour les discussions scientifiques, **Lucas** pour l'amour immense que tu as à offrir, **Camille** pour les petits messages attentionnés que tu m'envoies. Merci à mes grands-parents, **Claudie**, **Marceau** et **Marie**, pour l'amour que vous m'avez apporté, chacun à votre façon.

PE, plus de 6 ans que tu me supportes, tu dois avoir récolté assez de données pour écrire une thèse sur la Nell dans tous ces états. Merci d'avoir été à mes côtés pendant cette thèse, j'ai hâte de passer les prochaines années avec toi. Merci mille milliards de fois pour toute la joie que tu apportes dans ma vie, nos rires, nos repas, nos soirées flemmes, nos sorties, nos aprèms beach... Merci de supporter mon bordel, de faire la vaisselle et de sortir les poubelles. Merci aussi pour le soutien et les câlins quand c'est la fin du monde (à peu de chose près, un jour tous les 24 jours). Tu le sais, mais ça coûte rien de le redire, je t'aime.

Ceux qui lisent toujours, vous l'avez compris, je n'en serai jamais arrivé là sans votre soutien, votre amitié ou votre amour, alors merci encore.

List of publications

- Fernández I*, **Saunders N***, Duquerroy S, Bolland WH, Arbabian A, Baquero Salazar E, Blanc C, Lafaye P, Haouz A, Buchrieser J, Schwartz O, Rey FA. *Structural basis of TMPRSS2 zymogen activation and recognition by the HKU1 seasonal coronavirus*. Cell. In press, 2024. *equal contributions
- **Saunders N**, Monel B, Cayet N, Archetti L, Moreno H, Jeanne A, Marguier A, Wai T, Schwartz O, Frechin M. *Dynamic label-free analysis of SARS-CoV-2 infection reveals virus-induced subcellular remodeling*. Nature Communications. 2024 Jun (4996)
- **Saunders N**, Schwartz O. *TMPRSS2 est le récepteur cellulaire du coronavirus saisonnier HKU1*. Medecine/Science. 2024 Apr;40(4):335-337. French.
- **Saunders N**, Fernandez I, Planchais C, Michel V, Rajah MM, Baquero Salazar E, Postal J, Porrot F, Guivel-Benhassine F, Blanc C, Chauveau-Le Fric G, Martin A, Grzelak L, Oktavia RM, Meola A, Ahouzi O, Hoover-Watson H, Prot M, Delaune D, Cornelissen M, Deijs M, Meriaux V, Mouquet H, Simon-Lorière E, van der Hoek L, Lafaye P, Rey F, Buchrieser J, Schwartz O. *TMPRSS2 is a functional receptor for human coronavirus HKU1*. Nature. 2023 Dec;624(7990):207-214.
- Gellenoncourt S, **Saunders N**, Robinot R, Auguste L, Rajah MM, Kervevan J, Jeger-Madiot R, Staropoli I, Planchais C, Mouquet H, Buchrieser J, Schwartz O, Chakrabarti LA. *The Spike-Stabilizing D614G Mutation Interacts with S1/S2 Cleavage Site Mutations To Promote the Infectious Potential of SARS-CoV-2 Variants*. Journal of Virology. 2022 Oct 12;96(19):e0130122.
- **Saunders N***, Planas D*, Bolland WH, Rodriguez C, Fourati S, Buchrieser J, Planchais C, Prot M, Staropoli I, Guivel-Benhassine F, Porrot F, Veyer D, Péré H, Robillard N, Saliba M, Baidaliuk A, Seve A, Hocqueloux L, Prazuck T, Rey FA, Mouquet H, Simon-Lorière E, Bruel T, Pawlotsky JM, Schwartz O. *Fusogenicity and neutralization sensitivity of the SARS-CoV-2 Delta sublineage AY.4.2*. EBioMedicine. 2022 Mar;77:103934. *equal contributions
- Planas D*, **Saunders N***, Maes P*, Guivel-Benhassine F, Planchais C, Buchrieser J, Bolland WH, Porrot F, Staropoli I, Lemoine F, Péré H, Veyer D, Puech J, Rodary J, Baele G, Dellicour S, Raymenants J, Gorissen S, Geenen C, Vanmechelen B, Wawina-Bokalanga T, Martí-Carreras J, Cuypers L, Sève A, Hocqueloux L, Prazuck T, Rey FA, Simon-Lorière E, Bruel T, Mouquet H, André E, Schwartz O. *Considerable escape of SARS-CoV-2 Omicron to antibody neutralization*. Nature. 2022 Feb;602(7898):671-675. *equal contributions
- Rajah MM, Hubert M*, Bishop E*, **Saunders N***, Robinot R, Grzelak L, Planas D, Dufloo J, Gellenoncourt S, Bongers A, Zivaljic M, Planchais C, Guivel-Benhassine F, Porrot F, Mouquet H, Chakrabarti LA, Buchrieser J, Schwartz O. *SARS-CoV-2 Alpha, Beta, and Delta variants display enhanced Spike-mediated syncytia formation*. The EMBO Journal. 2021 Dec 15;40(24):e108944. *equal contributions

Conferences

- **Poster:** Journées départementales de virologie, May 2024
- **Oral:** Keystone Symposia, HIV and (Re)Emerging Viruses: Aligning Lessons Across Pandemics, Hannover, April 2024
- **Oral:** Forum de l'École Doctorale BioSPC, November 2023
- **Oral:** Stapa-ADIC Retreat, October 2023
- **Poster:** 20th Smögen Symposium on Virology, August 2023
- **Oral:** Young Researchers in Life Science, July 2023
- **Oral:** Journées départementales de virologie, May 2022

Table of content

Abstract (English)	1
Résumé (Français)	2
Résumé substantiel (Français)	3
Acknowledgments	8
List of publications	14
Conferences	15
Table of content	16
List of Figures	19
List of Tables	20
List of abbreviations	21
Chapter 1 – Introduction	23
1.1. Coronaviruses emergence and epidemiology	24
1.1.1. Coronavirus discovery	24
1.1.2. Seasonal coronaviruses	26
1.1.2.1. Classification and Origin	26
1.1.2.2. Disease	27
1.1.2.3. Epidemiology	27
1.1.3. Epidemic coronaviruses	27
1.1.3.1. SARS-CoV	27
1.1.3.2. MERS-CoV	28
1.1.3.3. SARS-CoV-2	28
Initial emergence	28
Virus spread	29
Variants of concern	29
Disease	30
Treatments	30
Vaccines	31
Burden of COVID19	31
Long COVID	32
1.1.4. Animal coronaviruses	32
1.2. Coronavirus description	34
1.2.1. Structural proteins	35
1.2.1.1. Spike protein	36
1.2.1.2. Membrane protein	36
1.2.1.3. Envelope protein	37
1.2.1.4. Nucleocapsid protein	37
1.2.1.5. Hemagglutinin-esterase	37
1.2.2. Non-structural proteins	37
1.2.3. Accessory proteins	38
1.3. Coronavirus entry and membrane fusion	41
1.3.1. Saccharide binding	41
1.3.2. Receptor binding	42
1.3.3. Spike cleavage	42

1.3.4.	Syncytia formation.....	44
1.4.	Coronavirus replication cycle	45
1.4.1.	Initial genome translation	45
1.4.2.	Genome replication and transcription.....	45
1.4.3.	Virion formation	46
1.5.	Immune response to coronaviruses	48
1.5.1.	Innate immune response	48
1.5.1.1.	Intrinsic cellular response.....	48
1.5.1.2.	Cellular mediated innate immunity.....	49
1.5.2.	Adaptive immunity	50
1.5.2.1.	T-cell immunity.....	50
1.5.2.2.	Humoral immunity	50
1.5.2.3.	Pre-existing immunity	51
	Research Objectives	52
	Chapter 2 - Characterization of SARS-CoV-2 VOCs.....	53
2.1.	Introduction	55
2.2.	Fusogenicity and neutralization sensitivity of the SARS-CoV-2 Delta sublineage AY.4.2.....	57
2.3.	Considerable escape of SARS-CoV-2 Omicron to antibody neutralization.	75
2.4.	Discussion.....	87
2.4.1.	What are the impacts of Spike mutations?	87
2.4.1.1.	Increased infectivity	87
2.4.1.2.	Alteration in the entry routes and impact	87
2.4.1.3.	Humoral immunity escape.....	88
2.4.1.4.	Mutations present in AY.4.2	88
2.4.2.	Links between infection, immunity and viral evolution	89
2.4.2.1.	Natural, vaccine and hybrid immunity.....	89
2.4.2.2.	Immune imprinting	89
2.4.3.	What about mutations in the rest of the genome?	90
2.4.3.1.	Innate immunity modulations	90
ORF9b	90	
ORF6	91	
N	91	
NSP6.....	92	
2.4.3.2.	Humoral immune escape	92
2.4.3.3.	Cellular immune response escape	92
2.4.3.4.	Changes in viral replication	92
2.4.3.5.	Increased vRNA packaging	93
	Chapter 3 – The secrets of HKU1 entry	94
3.1.	Introduction	96
3.1.1.	HKU1.....	96
3.1.2.	TMPRSS2.....	97
3.2.	TMPRSS2 is a functional receptor for human coronavirus HKU1.....	99
3.3.	Structural basis of TMPRSS2 zymogen activation and recognition by the HKU1 seasonal coronavirus.....	129
3.4.	Discussion.....	159
3.4.1.	Anti-TMPRSS2 nanobodies	159
3.4.2.	Other studies performed on HKU1-TMPRSS2 structure.....	160

3.4.3.	Struggles with HKU1 authentic virus	161
3.4.3.1.	Isolation on HAE	162
3.4.3.2.	The role of sialic acids.....	163
3.4.3.3.	Are there differences between the two TMPRSS2 isoforms?.....	164
3.4.3.4.	Isolation on immortalized cell lines.....	164
3.4.3.5.	Alternative methods to cultivate the virus	165
3.4.4.	Where does HKU1 come from?.....	166
3.4.4.1.	Could other coronaviruses use TMPRSS2 as a receptor?.....	166
3.4.4.2.	What have we learned of the potential origin of HKU1?	167
Chapter 4 – Subcellular remodeling during SARS-CoV-2 infection.....		169
4.1.	Introduction	171
4.1.1.	Classical imaging techniques.....	171
4.1.2.	Technical improvements in imaging	171
4.1.2.1.	Phase Imaging.....	171
4.1.2.2.	Raman microscopy	172
4.1.2.3.	Other label-free imaging techniques.....	172
4.1.3.	Analysis pipelines.....	173
4.1.3.1.	For cellular characterization.....	173
4.1.3.2.	For organelle segmentation.....	173
4.2.	Dynamic label-free analysis of SARS-CoV-2 infection reveals virus-induced subcellular remodeling	175
4.3.	Discussion.....	200
4.3.1.	Strengths and limits of our approach.....	200
4.3.2.	The biology behind the observed phenomenon	200
Conclusion		202
References		205
Appendix.....		224
Appendix 1		225
Appendix 2		226
Appendix 3		227

List of Figures

Figure 1: Electron microscopy (EM) images of coronaviruses..	24
Figure 2: Phylogenetic analysis of coronaviruses	25
Figure 3: The origins and intermediate hosts of coronaviruses.....	26
Figure 4: Evolution of variants frequency worldwide over the course of the pandemic	30
Figure 5: Human coronavirus description.....	34
Figure 6: Spike protein domains	36
Figure 7: Coronavirus entry in target cells.....	41
Figure 8: Description of the Hemifusion process	44
Figure 9: Coronavirus genome replication	46
Figure 10: Coronavirus replication cycle.....	47
Figure 11: Innate immune response antagonism by SARS-CoV-2.....	48
Figure 12: Main residues present in the epitopes targeted by anti-NTD, and class 1, 2, 3 and 4 anti-RBD antibodies.....	88
Figure 13: VHH affinity for TTSPs and animal TMPRSS2	159
Figure 14: Superimposition of HKU1-TMPRSS2 structures.....	161
Figure 15: SEM images of mock and infected HAE.....	162
Figure 16: IF staining of mock or HKU1 infected HAE	163
Figure 17: Plasmid-based assembly of SARS-CoV-2	165
Figure 18: Multiple alignment of the sequences of different Spike of embecoviruses	166
Figure 19: AlphaFold predictions of Longquan Rat coronavirus and OC43 RBD folding, overlapped with HKU1 structure.....	167
Figure 20: Sialic acid usage by embecoviruses	167
Figure 21: Specie specific TMPRSS2 use by HKU1A and HKU1B.....	168
Figure 22: Comparison of classical label-free imaging approaches.....	171
Figure 23: Comparison of state of the art HTM segmentation of mitochondria our approach	200

List of Tables

Table 1: Number of doses administered by the major vaccine manufacturer	31
Table 2: Homology between human coronavirus structural proteins (%).	35
Table 3: Main roles of the non-structural proteins of coronaviruses	38
Table 4: Main roles of the accessory proteins of human coronaviruses	40
Table 5: Main characteristics of human coronavirus spike proteins, and cellular factors involved in their entry	43
Table 6: TTSPs distribution and their role in viral glycoprotein cleavage	97

List of abbreviations

3CLpro	3-chromotrysin-like proteinase
aa	Amino-acid
ACE2	Angiotensin-converting enzyme 2
ADAM	A disitegrin and metalloprotease
ADCC	Antibody-dependent cellular cytotoxicity
ADCP	Antibody-dependent cellular phagocytosis
AI	Artificial intelligence
APN	Amino-peptidase N
BCoV	Bovine coronavirus
CoV	Coronavirus
COVID	Coronavirus disease
COVID19	Coronavirus disease 2019
CTD	C-terminal domain
DC	Dendritic cells
DMV	Double membrane vesicles
DPP4	Dipeptidyl peptidase 4
dsRNA	Double-stranded RNA
E	Envelope protein
EM	Electron microscopy
ER	Endoplasmic reticulum
ERGIC	ER-Golgi intermediate compartment
FP	Fusion peptide
GISAID	Global Initiative on Sharing All Influenza Data
gRNA	Genomic RNA
HAE	Human airway epithelium
hCoV	Human coronavirus
HE	Hemagglutinin-esterase
HR	Heptad-repeat
HTM	Holotomography
I	Internal protein
IBV	Avian infectious bronchitis virus
IFITM	Interferon induced transmembrane protein
IFN	Interferon
ISG	Interferon stimulated gene
KLK	Kallikrein
LD	Lipid droplets
LDLDR	Low-density lipoprotein receptor class A domain
M	Membrane protein
MERS-CoV	Middle-east respiratory coronavirus
MHC	Major-histocompatibility complex
MHV	Mouse hepatitis virus
MMP	Matrix metalloprotease
N	Nucleocapsid protein
NET	Neutrophil extracellular traps

NSP	Non-structural protein
NTD	N-terminal domain
ORF	Open-reading frame
PAMP	Pathogen-associated molecular pattern
PANGO	Phylogenetic Assignment of Named Global Outbreak
PDCoV	Porcine delta coronavirus
PEDV	Porcine epidemic diarrhea virus
PLpro	Parpain-like protease
pp	Polyprotein
QPI	Quantitative Phase Imaging
RDB	Receptor binding domain
RdRP	RNA-dependent RNA polymerase
ROS	Reactive oxygen species
RTC	Replication transcription complex
S	Spike protein
SARS-CoV	Severe acute respiratory syndrome coronavirus
sgRNA	Subgenomic RNA
ssRNA	Single-stranded RNA
SRCR	Scavenger receptor cysteine-rich domain
TGEV	Transmissible gastroenteritis virus
TLR	Toll-like receptors
TMPRSS	Type II transmembrane serine protease
TRS	Transcription regulatory sequence
TTSP	Type II transmembrane serine protease
UTR	Untranslated regions
VOC	Variant of concern
VOI	Variant of interest
WHO	World Health Organization

Chapter 1 – Introduction



1.1. Coronaviruses emergence and epidemiology

1.1.1. Coronavirus discovery

The first coronavirus, Avian Infectious Bronchitis Virus (IBV), was identified in 1931 in chickens suffering respiratory disease. It was subsequently cultivated in 1937. 10 years later, the mouse hepatitis virus (MHV) was discovered. In 1965, several independent groups identified and cultivated infectious agents isolated from nasal samples of a adults with a cold¹⁻⁴. Electron microscopy allowed them to show the strains they isolated had similar morphology to IBV, MHV and transmissible gastroenteritis virus of swine⁵. This new family of virus was named coronavirus after the crown like structure they displayed by electron microscopy (**Figure 1**).

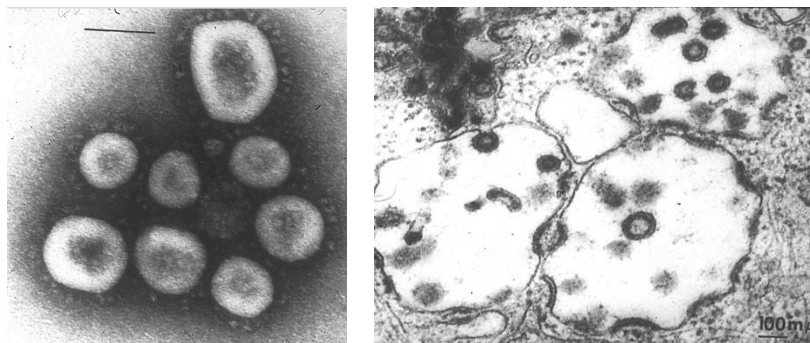


Figure 1: Electron microscopy (EM) images of coronaviruses. Right: Coronavirus OC16². Left: Strain 229E in WI-38 cells⁴.

This family is now subdivided in four genera: alphacoronavirus (α -CoV), betacoronavirus (β -CoV), deltacoronavirus (δ -CoV) and gammacoronavirus (γ -CoV) (**Figure 2**). The common ancestor reservoir to all coronavirus is unknown. The first two families have evolved from a bat coronavirus, while the later evolved in birds⁶. α -CoV and β -CoV common ancestor jumped in two different bat species, giving rise to the two genera⁷, before jumping into different mammalian species. The scenario is similar for the δ -CoV and γ -CoV in birds, although spillovers into mammals are scarcer.

Molecular clock analysis predicts the last common ancestor between bat and bird coronaviruses existed 10000 years ago but these analyses do not consider selection and recombination⁶. Using assumptions on species split and isolations, it is estimated that these viruses split around 150 million years ago^{8,9}.



Figure 2: Phylogenetic analysis of coronaviruses. The phylogenetic analysis was based on S, N and ORF1ab sequences. Colors indicates genera and, for β -CoV, subgenera. From Iruegas *et al.* 2021¹⁰.

Over 30 different strains causing common colds were isolated in the 1960s. However, only two were further cultivated and identified: hCoV-229E and hCoV-OC43. In 2003, severe acute respiratory syndrome coronavirus (SARS-CoV) caused the first known human epidemic^{11,12}. This brought back the attention of researchers to coronaviruses; two other human coronaviruses causing common colds were subsequently identified: hCoV-NL63¹³ and hCoV-HKU1¹⁴. NL63 was retrospectively found in samples from 1983¹⁵. In 2012, the middle-east respiratory coronavirus (MERS-CoV) caused the second coronavirus epidemic¹⁶. Finally, 2019 was marked by the emergence of SARS-CoV-2, the causative agent of the coronavirus disease 2019 (COVID19) pandemic^{17,18}. All the human coronaviruses are respiratory viruses, that are transmitted through secretions or aerosolized droplets of infected individuals¹⁹. The human coronaviruses are divided into two groups related to their pathogenicity: seasonal and epidemic. Myriads of animal coronaviruses have been identified; a small non-exhaustive description of them is presented subsequently.

1.1.2. Seasonal coronaviruses

1.1.2.1. Classification and Origin

hCoV-229E and hCoV-NL63 are α -CoV²⁰, likely of a bat origin^{21,22}. Evidence suggests that hCoV-229E intermediate host are dromedary camels (**Figure 3**: The origins and known intermediate hosts of coronaviruses. From Shore 2020^{26,27}).²³ hCoV-NL63’s intermediate host is unknown. hCoV-OC43 and hCoV-HKU1 are β -CoV, of the embecovirus subgenus (**Figure 1**), likely of rodent origin²⁴. For OC43, bovine are the suspected intermediate host²⁵, while HKU1’s is unknown (**Figure 3**).

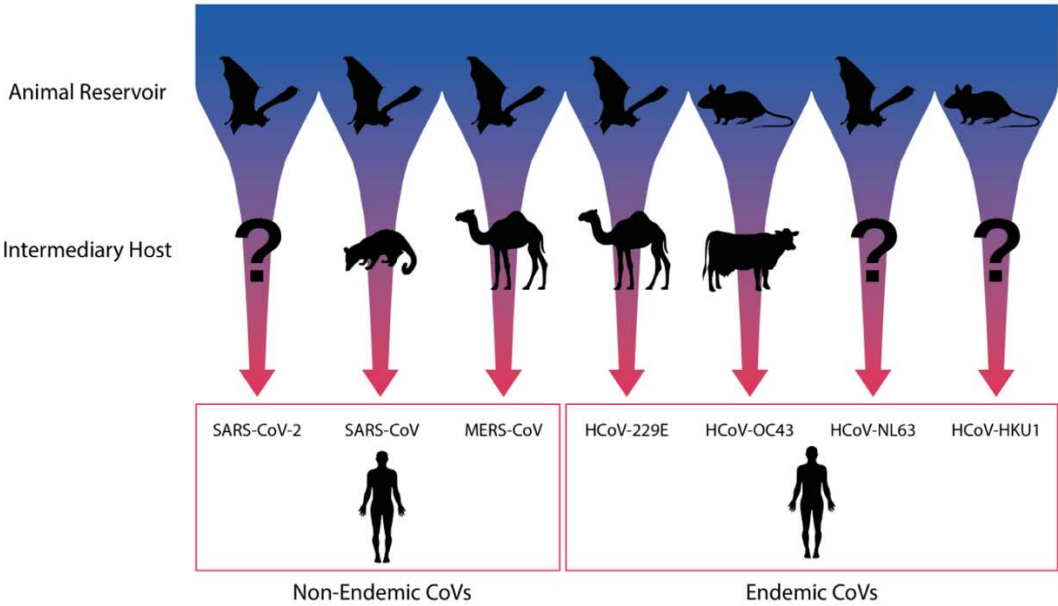


Figure 3: The origins and known intermediate hosts of coronaviruses. From Shore 2020^{26,27}.

1.1.2.2. Disease

Infection by seasonal coronaviruses generally results in mild cold-like symptoms¹⁹: fever, runny or stuffy nose, sore throat, headache, cough, fever and body aches. Rarely, infection leads to ear infection or asthma exacerbation in patients suffering asthma. However, in neonates, the elderly, patients with cardiopulmonary disease or the immunocompromised, pneumonia, chronic bronchitis and death can occur. Cases of encephalitis have also been described for OC43²⁸.

1.1.2.3. Epidemiology

These coronaviruses display seasonality. Infections generally peak in winter^{29–33}. Depending on the year and the place sampled, the relative distribution of the four coronaviruses varies, although OC43 is on average the most common.

In France, the seroprevalence of those four coronaviruses is estimated to be between 75 and 95%³⁴. It is estimated 20%–40% of previously uninfected children become infected every year³⁴. Antibodies half-life is estimated between 1 and 5 years, although cross-reactivity between different coronavirus make it hard to precisely estimate these parameters³⁴.

1.1.3. Epidemic coronaviruses

1.1.3.1. SARS-CoV

SARS-CoV is the virus responsible for the first known coronavirus outbreak. It is a β -CoV of the Sarbecovirus subgenus (**Figure 2**). It emerged in the Guangdong province (China) in February 2003^{11,12,35}. Horseshoe bats are the natural reservoir of SARS-CoV like viruses³⁶. Sampling of the animals sold on live food market and their traders showed that civets were the most likely intermediate host^{37,38}.

SARS-CoV spread throughout the world in February 2003. Over 8000 cases were reported³⁹ before the outbreak was contained in July 2003 thanks to sanitary measures and international collaboration. The last cases were reported in May 2004⁴⁰. The epidemic cause 774 reported deaths. The estimated case fatality ratio is 10% but it might be overestimated due to the lack of diagnosis.

SARS-CoV incubation period is estimated around 6 days⁴¹. Infection can be asymptomatic or induce flu-like symptoms, including muscle ache, fever, diarrhea, chills, cough, headaches and shortness of breath^{12,42}. Infected individuals can develop pneumonia, severe acute respiratory distress as well as heart or liver failure. The disease is less severe in children than adults, and although hospitalization can be required, no fatality in children have been reported⁴³. On the other hand, the fatality rate is much higher in the elderly; it is estimated at 40%^{41,42}. Treatments used during the pandemic include the use of a mechanical respirator, antiviral agents (ribavirin,

lopinavir in combination with ritonavir), immunosuppressive agents (corticosteroids), Chinese traditional medicine and in most severe cases the administration of convalescent plasma or immunoglobulin⁴⁴. Large and controlled studies lack on the efficacy of these treatments and retrospective meta-analyses of the studies performed were inconclusive on their efficacy⁴⁵.

1.1.3.2. MERS-CoV

MERS-CoV is the etiological agent of Middle East Respiratory Syndrome (MERS). It is a β -CoV of the Merbecovirus subgenus (**Figure 2**). It was first isolated from a patient in Saudi Arabia in April 2012¹⁶. It then spread throughout Saudi Arabia and neighboring countries. Cases were imported in over 27 countries, although the only significant outbreak outside the Middle East took place in Korea, where nearly 200 cases were identified⁴⁶. More than 2600 cases were reported, with 85% of them in Saudi Arabia⁴⁷. Sporadic cases still occur nowadays in Saudi Arabia. Over 900 deaths were reported, making MERS the deadliest coronavirus with a case fatality ratio of 35%⁴⁷.

As for SARS-CoV, infection ranges from asymptomatic to severe pneumonia with acute respiratory distress syndrome and multi-organ failure resulting in death, infections being milder in children. Guidance on treatments was mainly provided by the previous SARS-CoV outbreak; convalescent plasma, interferons, lopinavir and mono- or polyclonal antibodies are recommended when the benefits are likely to exceed the risks.

1.1.3.3. SARS-CoV-2

Initial emergence

SARS-CoV-2 is the etiological agent of the COVID19 pandemic. It is a β -CoV of the Sarbecovirus subgenus (**Figure 2**). The first SARS-CoV-2 official case occurred in a patient in Wuhan (China) in early December 2019. Two thirds of the initial cases clustered around a live animal and Seafood market in Huanan. To this day, the origin of SARS-CoV-2 remains debated⁴⁸. The first possibility is, as for SARS-CoV, that a spillover from bats to human occurred, maybe through an intermediate host sold in a live animal food market⁴⁹. Three bat coronavirus with very high similarity to SARS-CoV-2 (96,8%), notably in their receptor binding domain have been isolated in Laos⁵⁰. Other closely related viruses were found in bats and pangolins. Yet, there is a notable difference between all these virus and SARS-CoV-2; they lack the furin cleavage site, which is associated with an increased pathogenicity in humans. Neither SARS-CoV-2 nor a very closely related virus were found in any of the animals sampled on the Huanan food market. This, as well as phylogenetic alignment of the initial cases point towards the hypothesis that this market could have been the location of the first superspreading event, rather than the initial place of emergence⁵¹. The second possibility is that of a laboratory leak,

either of a coronavirus preserved in its natural form, or passaged in the laboratory⁵², although no clear evidence supports this theory.

Virus spread

The virus first spread throughout Wuhan until the first containment measures were taken on the 23rd of January 2020. However, the virus had already disseminated to the rest of China and the world through airline travels in the weeks preceding the lockdown. The World Health Organization (WHO) declared a Public Health Emergency of International Concern on 30 January 2020. The international measures taken were insufficient to prevent the worldwide spread of SARS-CoV-2. Lockdowns and sanitary measures all over the world strongly reduced the impact of the pandemic during the first year. A meta-analysis of 8 studies showed a reduction in incidence of COVID-19 associated with handwashing, mask wearing and physical distancing⁵³. The efficacy of lockdown to stop the increase of hospitalizations, deaths and diminish the burden on the intensive care units has been clearly demonstrated⁵⁴⁻⁵⁷.

Variants of concern

Spontaneous mutations that occur in the viral genome are selected for if they are advantageous for the virus. Thanks to international sequencing effort and a shared database (GISAID), it has been possible to closely monitor the appearance and the frequency of mutations in the SARS-CoV-2 genome. Over the course of the pandemic, numerous SARS-CoV-2 variants have emerged all over the globe. The most common nomenclature is their classification by lineages (hierarchical nomenclature with abbreviations, PANGO). They can also be classified by clades (year and order of appearance, Nextstrain). Emerging variants can be classified by the WHO as variants under monitoring, variants of interest (VOIs), or variants of concern (VOCs). Since May 2021, the WHO attributes Greek letters to the lineages of variants of interest and concern. This enables a simple, international nomenclature for both scientists and non-scientists. Furthermore, it prevents any stigma, as variants were initially named after their country of origin. This thesis uses the WHO and the PANGO nomenclatures.

The SARS-CoV-2 strain containing the D614G substitution in the spike protein was the earliest variant (B.1) to emerge; it supplanted the ancestral strain in under a month. In December 2020, the alpha (B.1.1.7)⁵⁸, beta (B.1.351)⁵⁹ and gamma (P.1)⁶⁰ variants supplanted the B.1 strain respectively in the UK, South-Africa and Brazil⁶¹, thanks to increased transmissibility and antibody escape^{62,63}. They were classified by the WHO as VOCs. The delta strain (B.1.617.2) which had first been detected in India in October 2020, started supplanting the other VOCs worldwide in May 2021⁶⁴, due to further increased transmissibility and antibody escape^{65,66} (**Figure 4**). It was the predominant variant until December 2021. In the end of November 2021, a new variant appeared in Botswana, Omicron (BA.1) (**Figure 4**). It rapidly spread, supplanting delta worldwide in a month, reinfecting individuals that had been infected

with Delta a month before. Previous infection did not confer protection, nor did vaccination without a booster⁶⁷. Since then, we have seen a never-ending succession of Omicron variants, BA.2, BA.5, BQ.1, XBB, XBB.1.5, XBB.1.6, XBB.1.9, EG.5.1, BA.2.86, JN.1, to cite just the main ones.

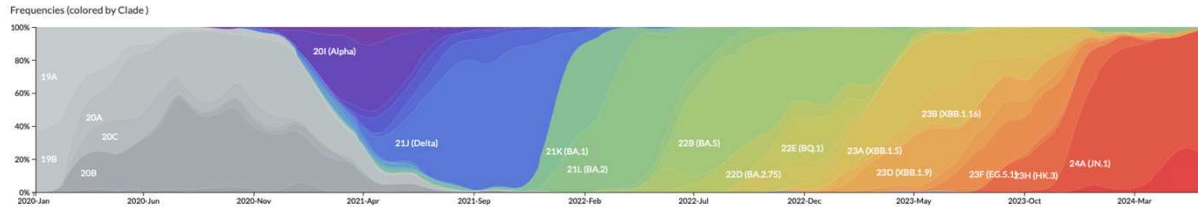


Figure 4: Evolution of variants frequency worldwide over the course of the pandemic. Available GISAID sequences were used to calculate frequencies. Nexstrain nomenclature and PANGO lineages for omicron variants are used. From NextStrain⁶¹.

The origin of these VOCs remains debated, but the main hypothesis is that the virus evolved in chronically infected immunocompromised individuals, allowing antibody escape mutations to occur^{68,69}.

Disease

As for SARS-CoV and MERS-CoV, SARS-CoV-2 infection can be asymptomatic or induce flu-like symptoms, muscle ache, fever, diarrhea, chills, cough, headaches and shortness of breath^{70,71}. A partial or complete loss of smell can occur in infected individuals. Infected individuals can develop pneumonia, severe acute respiratory distress as well as heart or liver failure^{70,71}. The different variants have distinct tropism, incubation periods and symptoms. Previous infection and vaccination generally protect against severe forms of the disease^{72,73}.

Treatments

Severe cases of the disease were initially treated as SARS-CoV and MERS-CoV. However, the unprecedented scale of the pandemic allowed large scale studies to take place, and the development of new treatments⁷⁴. Monoclonal therapeutic antibodies have been approved either as a preexposure prophylaxis in immunocompromised individuals (tixagevimab co-packaged with cilgavimab), or to prevent severe forms in high-risk individuals post-exposure (bamlanivimab, etesevimab, bebtelovimab, casirivimab, imdevimab, sotrovimab)⁷⁵. These antibodies initially neutralized the virus with high potency. However most Omicron variants are no longer neutralized by these antibodies, which are thus no longer used⁷⁵. An antiviral drug, Paxlovid (nirmatrelvir with ritonavir) is now the standard of care to prevent severe forms. Nirmatrelvir inhibits a viral protease that is essential for viral replication while ritonavir increases nirmatrelvir's plasma concentrations by inhibiting its degradation⁷⁶. Two other antivirals that interact with the viral RNA-dependent RNA polymerase have been approved: remdesivir induces premature transcription arrest while molnupiravir induces viral

mutagenesis⁷⁵. Their use is recommended when Paxlovid cannot be used⁷⁵. Several immunomodulators have shown to reduce the risk of severe forms (abatacept, baricitinib, tofacitinib, dexamethasone, infliximab, tocilizumab and sarilumab)⁷⁵.

Vaccines

To this day, over 50 vaccines have received an emergency or a full approval in at least one country against SARS-CoV-2. They are based on different technologies, the most produced worldwide as of January 2022 are presented in **Table 1**. There are seven types of vaccines: mRNA (Pfizer/BioNTech, Moderna), protein subunit (Novavax), attenuated virus, inactivated virus (Sinovac, Sinopharm, Bharat Biotech), non-replicating virus (Oxford/AstraZeneca, Janssen, Gamaleya Research), virus-like particle and DNA vaccines. It is estimated that vaccines prevented 14-20 million deaths worldwide in 2021⁷⁷. They show various levels of protection, but meta-analyses show on average their efficacy is around 70% for preventing symptomatic infections and around 90% for preventing severe cases and death⁷⁸. This efficacy decreases after 6 months to a year⁷⁹, and variant adapted boosters are recommended for the most fragile⁸⁰. Pfizer, Moderna and Novavax now produce regularly updated boosters.

Brand	Number of doses (million)	Countries
Pfizer/BioNTech	2647	EU, USA
AstraZeneca	2563	India, EU, China, South Korea, UK, USA, Japan, Thailand, Australia, Argentina
Sinovac	2466	China
Sinopharm	2274	China
Moderna	752	USA, Switzerland
Janssen Pharmaceutical	321	USA, EU
Gamaleya Research	277	Russia, Kazakhstan, Belarus, Serbia, Argentina
Bharat Biotech	250	India

Table 1: Number of doses administered by the major vaccine manufacturer as of January 2022 (over 50 million doses)⁸¹.

Burden of COVID19

In 2020, a total of 80 million cases and 1.7 million deaths were reported. These numbers are a large underestimation of the true number of cases and lethality. According to the WHO, a better estimation of the excess mortality due to COVID would be 3 million individuals. The case fatality ratio worldwide is estimated at 0.45%, although it is hard to estimate properly due to the differences in access to tests, hospitals and treatment, to asymptomatic forms of the disease and to the absence of reporting of both deaths and cases in some areas⁸². Vaccination

started in the end of 2020, conferring protection from SARS-CoV-2, altering the course of the pandemic and reducing severity of the disease.

In June 2022, projections of total deaths were estimated at 17 million, while reported deaths were around 7 million⁸³. In June 2022, 60% of the world population was vaccinated at least once, 12 billion doses had been injected⁸³. At this time, about 10000 weekly deaths were reported⁸⁴. Since then, the fatality of COVID19 has drastically reduced, thanks to vaccination, and because most cases are now reinfections. The number of weekly deaths reported is around 3000⁸⁴.

To this day, there has been 800 million cases and 8 million deaths reported. However, in November 2021, studies estimated the true number of cases was close to 4 billion⁸⁵. 3 years have passed, people do not report positive results of self-antigenic test or do not test at all, the case fatality ratio has gone down thanks to vaccination and previous protection. One can perform studies on water waste to evaluate the prevalence; however, these models already depend on an accurate prediction of the number of cases at certain timepoints, shedding can be affected by vaccination or variant specificities, and water RNA concentration by parameters such as precipitations, industrial effluents ... Although they yield important information about trends, they do not provide accurate reports of cases. It is thus very hard to evaluate the true number of cases, which is likely closer to 15 billion than to 1; a survey in my entourage reveals most people have caught COVID between 2 and 4 times. Estimations of the total number of cases are no longer published.

Long COVID

Individuals can develop long COVID following infection. Symptoms, their gravity and their onset are very diverse; they include general, respiratory, neurological and digestive affections⁸⁶. Worst cases show a myalgic encephalomyelitis/chronic fatigue syndrome, which can seriously affect the lifestyle of those affected⁸⁶. Symptoms can persist for months or years after the acute phase of the disease. The causes of this pathology and the risk-factors are poorly understood, and no treatment are known to this date. Estimates of the risk of developing some form of long COVID range between 2 and 12%⁸⁷: it is estimated that 65 million people suffer long COVID worldwide⁸⁶.

1.1.4. Animal coronaviruses

Coronaviruses can infect a wide range of hosts, including feline, primates, canine, bovine, birds, rodents, porcine and dromedary camels. Animal coronaviruses can cause a wide range of diseases; respiratory, enteric, hepatic and neurological⁸⁸. Coronaviruses are a major burden in the farming industry. Bovine (BCoV) and porcine coronavirus (PEDV, TGEV, PDCoV) cause

severe diarrhea and/or respiratory disease resulting in significant mortality in piglets and bovines⁸⁸. IBV lead to dramatic drops in poultry egg production⁸⁸ and mortality ranges from 0 to 80% depending on the strain. Coronaviruses have a moderate to high mutation rates and can undergo recombination; this leads to fast genotype expansion and diversification. This can in turn lead to cross-species transmission; identifying animal hosts of coronavirus is essential to prevent and understand future spillovers in the human population⁸⁹. Major anthropogenic modifications have taken place in the last century: urbanization, extension of cropland in wildlife habitats, bushmeat hunting, wildlife trade, deforestation and intensive farming. Those changes can be associated with increased pathogen transmission and zoonosis⁹⁰⁻⁹³; however these effects are complex and remain an important field of study⁹⁴.

1.2. Coronavirus description

Coronaviruses are single stranded positive RNA viruses. Coronaviruses have a genome of 27-32kb⁹⁵. The virions' diameters range from 80 to 120 nm⁹⁵. The genomic RNA (gRNA) is encapsidated by the Nucleocapsid protein (N). The virions are enveloped in a lipid bilayer, in which the Spike protein (S), the Membrane protein (M) and the Envelope protein (E) are embedded. Embecoviruses OC43 and HKU1 have a fifth structural protein, Hemagglutinin-esterase (HE) (**Figure 5**).

The viral genome is capped on its 5'-end by a m⁷GpppA₁. On its 3'-end it possesses a 30-60 nucleotide long poly-A tail⁹⁵. These modifications protect the viral RNA from cellular exoribonucleases, enhancing its stability.

Coronaviruses have two large open-reading frames (ORF): ORF1a and ORF1b. They account for 70% of the viral genome and are further processed into 16 non-structural proteins (NSP)⁹⁵. The rest of the genome is comprised of the structural proteins and a varying number of smaller ORFs which encode accessory proteins (**Figure 5**). The genome also contains two untranslated regions (UTR) in 3' and 5'⁹⁵.

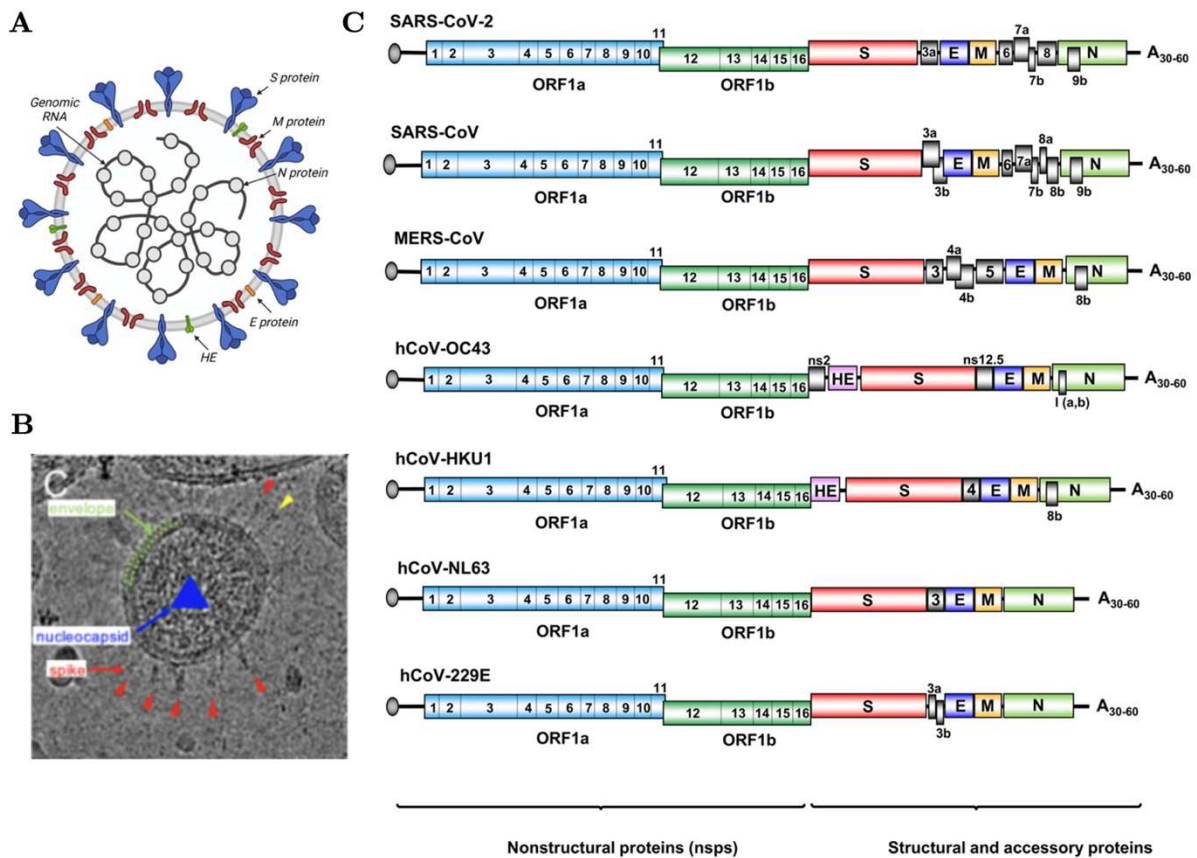


Figure 5: Human coronavirus description. A. Coronavirus virion structure. The S protein protrudes from the surface. HE, when present forms a second ring. E and M are embedded in the membrane. N packages the genome. From Mingaleeva *et al.* 2022⁹⁶. B. Cryo-EM image of SARS-CoV-2 virion. From Liu *et al.* 2020⁹⁷. C. Organization of human coronaviruses genomes. From Brant *et al.* 2021⁹⁸.

1.2.1. Structural proteins

Amino acid (aa) conservations of the structural proteins between human coronaviruses are reported in **Table 2**. S is involved in entry, E, M and N in virion formation, and HE in viral egress. For MHV⁹⁹, TGEV¹⁰⁰, BCoV¹⁰⁰ and IBV¹⁰¹, co-expressing E and M is sufficient to drive virus-like particles (VLP) formation. For SARS-CoV and SARS-CoV-2 studies have shown that M+N^{102,103} or E+M^{104,105} or E+M+N¹⁰⁶ are sufficient and necessary to drive assembly. Evidence shows that E's absence strongly diminishes viral titers and infectivity¹⁰⁷.

The following sections describe their main characteristics.

Membrane	NL63	229E	HKU1	OC43	SARS-CoV	SARS-CoV-2	MERS-CoV
	NL63	100	62	34	32	29	27
229E	62	100	33	32	27	29	30
HKU1	34	33	100	80	36	36	45
OC43	32	32	80	100	40	39	43
SARS-CoV	29	27	36	40	100	90	42
SARS-CoV-2	27	29	36	39	90	100	40
MERS-CoV	31	30	45	43	42	40	100

Spike	NL63	229E	HKU1	OC43	SARS-CoV	SARS-CoV-2	MERS-CoV
	NL63	100	64	28	29	25	26
229E	64	100	28	29	28	28	28
HKU1	28	28	100	65	31	30	33
OC43	29	29	65	100	32	31	34
SARS-CoV	25	28	31	32	100	77	32
SARS-CoV-2	26	28	30	31	77	100	32
MERS-CoV	26	28	33	34	32	32	100

Envelope	NL63	229E	HKU1	OC43	SARS-CoV	SARS-CoV-2	MERS-CoV
	NL63	100	47	16	18	19	19
229E	47	100	22	22	27	28	24
HKU1	16	22	100	51	29	29	23
OC43	18	22	51	100	19	20	27
SARS-CoV	19	27	29	19	100	96	37
SARS-CoV-2	19	28	29	20	96	100	37
MERS-CoV	20	24	23	27	37	37	100

Nucleocapsid	NL63	229E	HKU1	OC43	SARS-CoV	SARS-CoV-2	MERS-CoV
	NL63	100	45	25	26	28	28
229E	45	100	29	26	25	26	27
HKU1	25	29	100	64	34	34	34
OC43	26	26	64	100	35	35	36
SARS-CoV	28	25	34	35	100	90	48
SARS-CoV-2	28	26	34	35	90	100	49
MERS-CoV	29	27	34	36	48	49	100

Hemagglutinin-Esterase	HKU1	OC43
	HKU1	100
OC43	58	100

Table 2: Homology between human coronavirus structural proteins (%). Alignments were performed with ProteinBlast by retrieving sequences from Uniprot.

1.2.1.1. Spike protein

S is expressed at the surface of the enveloped virion. It is the protein responsible for the virus' crown-like, 'corona' structure. The spike is a glycoprotein of 1255-1356 aa in hCoVs. It is 128–160 kDa before glycosylation and 150–200 kDa after N-linked glycosylations¹⁰⁸. S assemble in trimers; each virus particle expresses 50-100 trimers of spike¹⁰⁹, which form a 20 nm crown on the virion. It is the main protein enabling coronavirus entry into cells; it binds to a cellular receptor, before inducing membrane fusion. It is composed of a small intravirion domain, a transmembrane domain and an ectodomain composed of two subunits, S1 and S2^{110,111} (**Figure 6**). S1 is composed of an N-terminal (NTD) and C-terminal domain (CTD), which is the receptor binding domain (RBD) (**Figure 6**). The S2 subunit is composed of a cytosolic C-terminal domain, a transmembrane anchor, two α helices (Heptad-repeat (HR) 1 and 2) and an amphiphilic fusion peptide (FP) (**Figure 6**). One or two cleavage sites are present depending on the coronavirus. MERS-CoV, SARS-CoV-2, HKU1 and OC43 S contain a polybasic RXXR|S motif in at the S1/S2 junction, that is cleaved by furin¹¹². The two subunits remain bound by non-covalent interactions until the spike is further cleaved at the S2/S2' site, which is present in all hCoV¹¹³. Removing the S1/S2 cleavage site from human coronaviruses that possess it decreases their fitness *in vitro* and *in vivo*¹¹⁴.

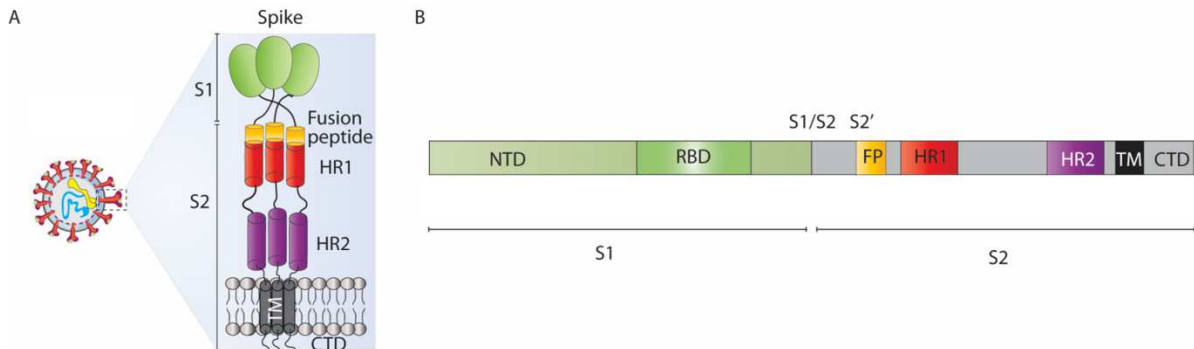


Figure 6: Spike protein domains. The S1 domain is composed of the NTD and the RBD. The S2 domain is composed of a fusion peptide and the HR, as well as a transmembrane domain. The Spike assembles in trimers. From Ali *et al.* 2023¹¹⁵.

1.2.1.2. Membrane protein

M is the most abundant structural protein in the virion; there is about 20-fold more M than S on a virion. M has three transmembrane domains. M assembles as homodimers, to drive virus assembly, morphogenesis and membrane budding, through interaction with the other structural proteins^{98,108,116,117}.

M inhibits the production of type I and III interferon (IFN), by interacting with NF- κ B, TBK1, MAVS and RIG-I^{118–121}.

1.2.1.3. Envelope protein

E is a small protein (8-12kDa) found in low amounts in virions. E proteins are highly divergent between coronaviruses (Table 2) in terms of aa composition but are structurally conserved. E is involved in virus assembly, morphogenesis and release of mature virions^{98,108,122}. E's deletion results in the formation of abnormal non-spherical virions. One study shows that E is not responsible for membrane curvature directly but through its interaction with M¹²³, while another shows E itself generates bending in a lipid bilayer¹²⁴.

E assembles in pentamers that form ion-channels (viroporins) in the virion membrane. These viroporins contribute to the pathogenesis by disrupting the homeostasis of the cell and activation the inflammasome¹²⁵. E also interacts with PALS1¹²⁶, a host-protein that is essential to maintain tight junctions and thus an intact lung epithelium.

1.2.1.4. Nucleocapsid protein

N is found underneath the viral envelope. N binds to the gRNA through positively charged regions in its NTD and then self-oligomerizes thanks to its CTD. This results in the formation of helicoidal structure^{98,108,127}. N can then interact with M which results in viral assembly.

N also promotes viral transcription; it is recruited to the replication transcription complexes and participates in RNA chaperoning^{128,129}. Furthermore, it inhibits the stress granule dependent antiviral response¹³⁰ and type I and III IFN production by acting both on RIG-I and IRF3¹³¹. N can also interfere with translation by interacting with EF1 α ¹³².

1.2.1.5. Hemagglutinin-esterase

HE is only present in embecoviruses. It is the result of a horizontal gene transfer, likely from influenza C or D^{133,134}. HEs are generally composed of a lectin domain which binds to sialic acid and of an esterase domain which cleaves sialic acids, allowing virion release. However, in OC43 and HKU1, HE seems to lack lectin binding ability. This is not the case in animal embecoviruses. OC43 phylogenetic analysis show that after its spillover into humans, OC43 HE binding to sialic acids was selected against and ultimately lost¹³⁵. The HE of OC43 and HKU1 seemed to have kept a functional esterase domain¹³⁶. HE strongly enhances OC43 infectivity¹³⁷.

1.2.2. Non-structural proteins

All human coronaviruses possess 16 NSPs. A non-exhaustive list of their identified roles, notably for SARS-CoV-2, is presented in **Table 3**. Many of them act at several levels, and all their roles are not elucidated yet. Some of their functions will be explored in the following sections.

Protein	Role
NSP1	Suppresses host translation by binding to the ribosome ¹³⁸ , preventing interferon and interferon stimulated genes (ISG) expression.
NSP2	Inhibits IFN β expression ¹³⁹ .
NSP3 (PLpro)	Contains one or two protease domain ¹⁰⁸ . Cleaves NSP1, 2 and 3 from polyprotein (pp) 1a and 1ab. Involved in replication organelle morphogenesis ¹⁴⁰ . Induces proteasomal degradation of p53, inhibiting p53-dependant IFN production ^{141,142} . Antagonizes the innate immune response ¹⁴³⁻¹⁴⁵ .
NSP4	Involved in replication organelle morphogenesis ¹⁴⁰ .
NSP5 (3CLpro)	Processes pp1a and pp1ab to release NSP4-16. Suppresses innate immunity ^{144,146} .
NSP6	Involved in replication organelle morphogenesis ¹⁴⁰ . Interferes with interferon response ¹⁴⁷ .
NSP7	RNA polymerase co-factor ¹⁴⁸ . Interferes with IFN α signalling ^{149,150} .
NSP8	RNA polymerase co-factor ^{148,151} . Likely polyadenylates the genome. Interferes with protein trafficking to the cellular membrane ¹³⁸ .
NSP9	Binds to viral RNA, involved in RNA modifications and capping ¹⁵² . Interferes with protein trafficking to the cell membrane ¹³⁸ and nuclear transport, and thus with the immune response ¹⁵³ .
NSP10	Co-factor of the RNA capping machinery ¹⁵⁴ .
NSP11	Unknown role.
NSP12	RNA dependent RNA polymerase ^{155,156} .
NSP13	RNA helicase and RNA 5'-phosphatase, involved in RNA capping. Interferes with interferon signaling ^{147,157} .
NSP14	Proofreading 3'-5' exoribonuclease and N ⁷ -methyltransferase involved in RNA capping ¹⁵⁸ . Blocks host translation. Interferes with interferon signaling ¹⁵⁷ .
NSP15	Endoribonuclease that removes 5'-polyuridine from the negative strand RNA, removing potential PAMPs ¹⁵⁹ . Blocks the innate immune response ¹⁵⁷ .
NSP16	Ribose 2'-O-methyltransferase, involved in RNA capping ¹⁶⁰ . Blocks the host mRNA splicing activity ¹³⁸ .

Table 3: Main roles of the non-structural proteins of coronaviruses. Adapted from Steiner *et al*¹⁶¹.

1.2.3. Accessory proteins

Human coronaviruses possess between 1 and 10 accessory proteins. They are translated from separate open reading frames in the 3' end of the genome. They have various roles, among which interfering with the innate immune response and adaptive immune response, interfering with the host gene expression and promoting viral egress¹⁶¹.

One common feature between 5 of the 7 human coronaviruses is the presence of an internal protein (I) which overlaps with N and is translated from the same subgenomic RNA (sgRNA) with a frameshift. Studies have shown this I protein interferes with the interferon response¹⁶²⁻

Interestingly, some of these genes were acquired independently but evolved toward the same function. Another common feature between several human coronaviruses is the presence of an accessory gene coding for a viroporin in between S and E^{167–169}. The ones from SARS-CoV, 229E and NL63 share some sequence similarity, while the one from OC43 is totally unrelated to that of the other human coronavirus. OC43 viruses lacking this viroporin display growth defect, which are complemented by adding the viroporin from SARS-CoV, NL63 or 229E, suggesting they play similar roles, even though they were acquired independently¹⁶⁹. Several of the human coronaviruses encode a phosphodiesterase, that antagonizes RNase-L. RNase-L induces transcriptional arrest and cell death upon double-stranded RNA (dsRNA) detection. This phosphodiesterase is essential for viral replication in MHV. For embecoviruses this phosphodiesterase could come from a vertebrate protein AKAP7¹⁷⁰, while its origin for MERS-CoV is unknown.

The other known role of these proteins for human coronaviruses are described in **Table 4**.

Protein	Coronavirus	Role
Ns2	OC43	Phosphodiesterase, prevents cell transcriptional arrest ¹⁷¹ .
ORF3	MERS-CoV	Induces apoptosis ¹⁷² .
ORF3	NL63	Transmembrane protein. Unknown role but 60% similarity with 229E ORF3 (viroporin).
ORF3 / ORF3a / ORF3b	229E	Single ORF in clinical isolates, 2 ORFs in lab adapted strains (also called ORF4a/4b) ¹⁷³ . Viroporin with ion-channel activity ¹⁶⁸ .
ORF3a	SARS-CoV, SARS-CoV-2	Viroporin. Reduces protein (including major-histocompatibility complex (MHC)) trafficking. Promotes lysosomal exocytosis, induces inflammatory response and cell death ^{167,174–176} . Prevents STAT1 phosphorylation ¹⁴⁷ .
ORF3b	SARS-CoV	Antagonizes IFN response ^{177,178} .
ORF3c	SARS-CoV-2	Interferes with IFN- β production ¹⁷⁹
ORF4	HKU1	Unknown role.
ORF4a	MERS-CoV	Prevents stress granule formation ^{180–182} . Inhibits interferon induction.
ORF4b	MERS-CoV	Phosphodiesterase, prevents cell transcriptional arrest ¹⁸³ . Inhibits the type I IFN and NF- κ B signaling pathways ^{184,185} .
NS12.9	OC43	Viroporin involved in virion morphogenesis ¹⁶⁹ .
ORF5	MERS-CoV	Interferes with NF- κ B activation and the host inflammation response ¹⁸⁶ .
ORF6	SARS-CoV, SARS-CoV-2	Blocks nuclear import and export which suppresses innate immune response ^{187–190} .
ORF7a	SARS-CoV, SARS-CoV-2	Blocks interferon response by blocking STAT2 phosphohorylation ¹⁴⁷ . Interacts with the MHC-I heavy chain, impairing its surface expression ¹⁹¹ .
ORF7b	SARS-CoV, SARS-CoV-2	Blocks interferon response by blocking STAT1 and 2 phosphohorylation ¹⁴⁷ .

ORF8	SARS-CoV-2	Suppresses IFN β signaling. Disrupts host cell epigenetic regulation ¹⁹² . Leads to degradation of MHC-I ¹⁹³ .
ORF8a/ ORF8b	SARS-CoV	The two ORFs result from a deletion in the original ORF8 which occurred early in the pandemic. Unclear role ¹⁹⁴ .
ORF8b/ I protein	MERS-CoV	Suppresses type I IFN response ^{163,165} .
ORF8b / I protein	HKU1	Unknown role.
ORF9b / I protein	SARS-CoV, SARS-CoV-2	Suppresses type I and III IFN response ^{162,164,166} . Affects mitochondrial protein import ¹⁹⁵ .
I protein	OC43	Single ORF in clinical isolates, two (a/b) in culture adapted strain ¹⁹⁶ . Unknown role.
ORF10	SARS-CoV-2	Suppresses type I interferon response ¹⁹⁷ .

Table 4: Main roles of the accessory proteins of human coronaviruses.

1.3. Coronavirus entry and membrane fusion

To enter cells, coronaviruses S first attaches to mono or polysaccharides present on the surface of the cell (**Figure 7**). It then binds to a receptor, before being primed at the surface or in endosomes by different proteases (**Figure 7**). This triggers membrane fusion and release of the viral genome in the cell cytoplasm (**Figure 7**). The steps are described in more details in the following sections, and the cellular partners involved for each coronavirus are summarized in **Table 5**.

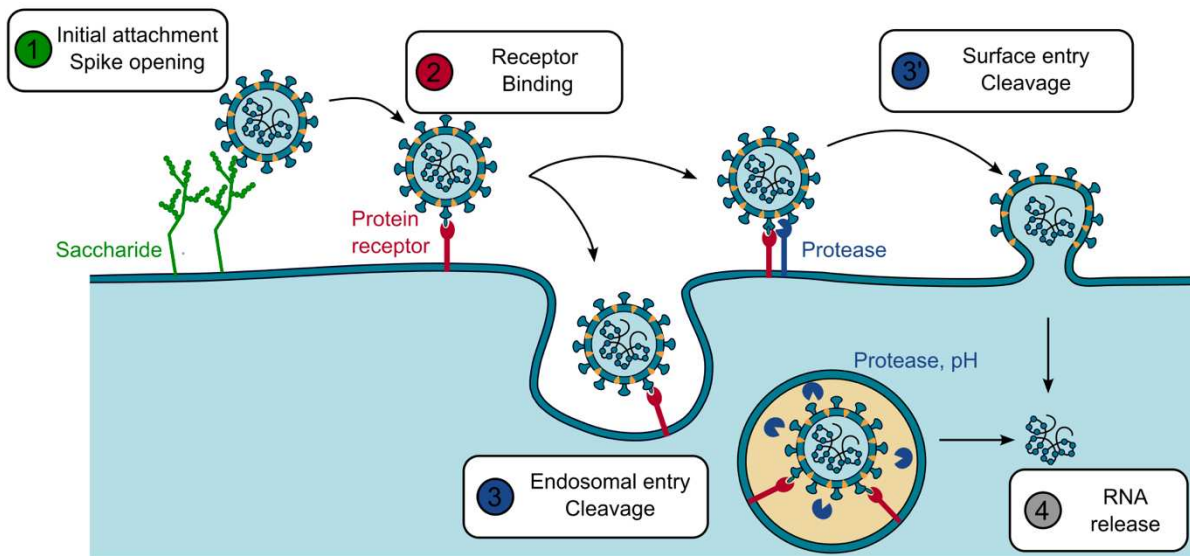


Figure 7: Coronavirus entry in target cells. S first binds to saccharides, before binding to its receptor. It then undergoes cleavage either at the membrane or in endosomes. The RNA is then released in the cytoplasm.

1.3.1. Saccharide binding

S can either be in an opened or closed conformation. In the closed conformation, the RBD is not accessible and is protected from potential neutralizing antibodies. It can be further shielded by glycans. In the opened conformation, the RBD is exposed and can bind to cells. For SARS-CoV, SARS-CoV-2 and MERS-CoV, these states appear to co-exist on virions; S trimers can transition from one to the other or exist in intermediate states where one or two RBDs are exposed^{198–200}. For HKU1, NL63, OC43 and 229E only the closed conformations of S have been observed so far on virions¹³⁶. Whether this difference between open/closed state between seasonal coronavirus and epidemic coronaviruses is due to a longer adaptation period in humans is currently unknown. The link with pathogenicity has not been studied either.

The NTD of S1 is accessible in the closed conformation of S and allows a primary attachment to the target cells. NL63, SARS-CoV and SARS-CoV-2 interact with Heparan-Sulfate through their NTD^{201,202}. MERS-CoV, OC43 and HKU1 interact with sialosides. MERS-CoV has a

preference for α 2,3-linked sialosides, the interaction is blocked by 9-O-acetylation or 5-N-glycolylation of the sialic acid²⁰³. On the other hand, OC43 and HKU1 bind to 9-O-acetylated-sialic acids, HKU1 has a preference for α 2,8-linked sialosides, while OC43 binds both α 2,3 and α 2,8-linked sialosides²⁰⁴. No attachment factors have been identified for 229E.

For HKU1A, the binding to 9-O-acetylated-sialic acids has recently been shown to allow opening of S and exposure of the receptor binding domain²⁰⁵. This opening is essential for viral entry, and mutation of a residue involved in this change of conformation completely prevents entry¹³⁶. However, opened forms of purified HKU1B spike have been observed in absence of sialosides very recently²⁰⁶. Whether this difference is an artefact resulting of mutations introduced in the Spike by different groups to synthesize it, or a true difference between HKU1A and HKU1B is unclear. Similar mechanism of opening upon NTD binding to cellular factors might exist for other coronaviruses, but it has not been shown yet. So far, we do not know how the OC43, NL63 and 229E spike open to expose their RBD.

1.3.2. Receptor binding

Once the RBD is exposed, it can bind to a protein receptor expressed on the surface of cells. SARS-CoV, SARS-CoV-2 and NL63 bind to Angiotensin-converting enzyme 2 (ACE2)²⁰⁷⁻²⁰⁹. MERS-CoV utilizes dipeptidyl peptidase 4 (DPP4)²¹⁰. 229E interacts with amino-peptidase N (APN)²¹¹. Two chapters of this thesis show that HKU1's receptor is type II transmembrane serine protease 2 (TMPRSS2). The receptor for OC43 is currently unknown.

Other proteins such as L-SIGN, DC-SIGN, TMEM106b and Neuropilin 1 increase the viral entry of SARS-CoV-2, potentially through enhancing of the initial attachment of S, or functioning as secondary receptors for S²¹²⁻²¹⁴. Their role for other coronaviruses has not been studied yet.

1.3.3. Spike cleavage

Upon binding to its cellular receptors, S undergoes conformational changes that expose the secondary S2/S2' cleavage site. Then, entry can occur at the membrane or through endosomes:

- S can be cleaved at the S2/S2' site directly at the cell surface by surface serine proteases. TMPRSS1, 2, 4, 11D, 11E, 11F, 13, 14^{215,216}, matrix metalloprotease 2, 9, 14, 16, 17 (MMP)^{217,218}, A disitegrin and metalloprotease 8, 10, 17 (ADAM)^{218,219} and Kallikrein 13 (KLK13)²²⁰ cleave at least one coronavirus. TMPRSS2 cleaves all human coronaviruses^{221,222} and is highly expressed in the respiratory tract.

- The virus can be endocytosed, by clathrin^{223,224} or caveolin²²⁵ dependent pathways. In endosomes, S can be cleaved by cathepsins^{226–228}, or undergo conformational changes due to the acidification of endosomes^{229,230,225}.

Following cleavage or conformational changes, the S1 subdomain of the spike is shed. The S2 domain is revealed. HR1 becomes helical, and the FP is projected in the cellular membrane (plasma or endosome membrane). The FP disrupts the lipid bilayer and anchors the virus to the membrane²³¹. Upon insertion of the FP in the cellular membrane, highly conserved hydrophobic grooves are exposed on HR1. HR1 and HR2 interact through these grooves; this brings the viral and cellular membranes together²³¹. This overcomes the repulsion between the two lipid bilayer and allows a process called hemifusion to occur (**Figure 8**)²³¹. Both membrane merge, and a pore is formed between the cellular and viral membranes. The viral content is released in the cytosol.

Virus	Described state of S	S1/S2 cleavage	Initial Attachment factor	Receptor	S2/S2' Cleavage
SARS-CoV	(o), (c)	(n)	Heparan-Sulfate	ACE2 ²⁰⁷	TMPRSS2/11A/11D/11E/13 ²³² , catL ²²⁷
SARS-CoV-2	(o), (c)	(y)	Heparan-Sulfate ²⁰²	ACE2 ²³³	TMPRSS2/4/11D/11E/11F/13/14 ^{216,218,232} , MMP2/9/14/16/17 ^{217,218} , ADAM8/10/17 ^{218,219} , catL ²²⁸
NL63	(c)	(n)	Heparan-Sulfate ²⁰¹	ACE2 ²⁰⁸	TMPRSS2 ²³²
MERS-CoV	(o), (c)	(y)	α 2,3-linked sialic acids ²⁰³	DPP4 ²¹⁰	TMPRSS2/11D/11E/13 ²³² , catL ²³⁴
229E	(c)	(n)	Unknown	APN ²¹¹	TMPRSS2 ²³² , catL ²³⁵
OC43	(c)	(y)	9-O-acetylated- α 2,3/2,8-linked sialic acids ²³⁶	Unknown	TMPRSS2 ²³²
HKU1	(c)	(y)	9-O-acetylated- α 2,8-linked sialic acids ²³⁶	Unknown	TMPRSS2 ²³² , KLK13 ²²⁰

Table 5: Main characteristics of human coronavirus spike proteins, and cellular factors involved in their entry. Opened (o), closed (n), Yes (y), No (n)

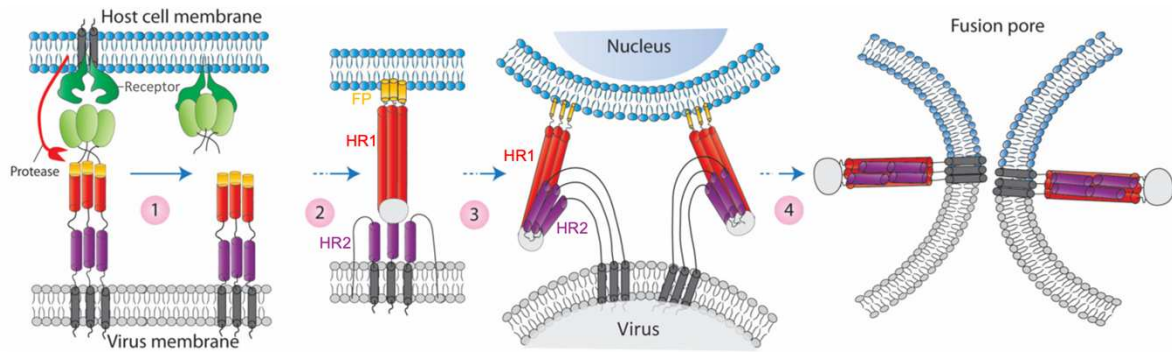


Figure 8: Description of the Hemifusion process. After cleavage of the S1 domain, the FP is projected in the adjacent membrane. HR1 and HR2 then interact and bring the two membranes together, allowing the formation of a pore. From Ali *et al.* 2023¹¹⁵

1.3.4. Syncytia formation

Upon infection by coronaviruses, the coronavirus spike can be expressed at the plasma membrane. As for the spike expressed on viruses, this spike expressed at the membrane can bind to cellular receptors expressed on neighboring cells and undergo cleavage. This induces fusion of the infected cell and its neighbor, and appearance of a syncytia, a multinucleated cell. These syncytia have been observed in post-mortem biopsies of SARS-CoV-2 infected patients²³⁷. Their contribution to the COVID19 pathology is poorly understood. They permit cell to cell transmission of the virus and thus immune evasion²³⁸. A study shows syncytia can trap lymphocytes, thus dampening the immune response²³⁹. Syncytia death likely contributes to cytokine release and inflammation^{240,241}. Their effect on viral replication is controversial; premature cell death could allow the release of viruses, but they might be immature²⁴².

1.4. Coronavirus replication cycle

1.4.1. Initial genome translation

After entry in the cell, cellular proteases remove N from the viral genome. The viral genome modifications allow it to be recognized as an mRNA by the cell. It is thus translated by the cellular machinery. ORF1a is translated into pp1a. ORF1a and ORF1b partially overlap, but there is a -1 frameshift between them. During translation, a ribosomal -1 frameshift can occur, thanks to a slippery sequence composed of several identical nucleotides on which the ribosome can shift out of frame^{95,243}. Furthermore, a pseudoknot is present in the RNA directly after this slippery sequence; the ribosome pauses, making a frameshift more likely²⁴³. The probability of frameshift for coronavirus ranges between 45 and 70%^{244,245}, resulting in about two-fold more abundance of NSP1-10 than of NSP11-16. pp1a and pp1ab are then auto-processed into the different NSPs. SARS-CoV, MERS-CoV and SARS-CoV-2 have one papain-like protease (PLpro) within NSP3¹⁰⁸. The four seasonal coronaviruses have two PLpro within NSP3¹⁰⁸. PLpro(s) cleave NSP1-3 from the pp^{108,161}. The 3-chromotrysin-like proteinase (3CLpro) NSP5 then processes the rest of the pp (NSP4-11/16)^{108,161}.

1.4.2. Genome replication and transcription.

NSP3, 4 and 6 then remodel extensively the endoplasmic reticulum (ER) to form convoluted membranes and double membrane vesicles (DMVs), which remain linked with the ER through their outer membrane¹⁴⁰. NSP3 and 4 form pores in these DMVs that are wide enough to accommodate the export of mRNA²⁴⁶⁻²⁴⁸. The viral RNA-dependent RNA polymerase (RdRP), formed of NSP7, 8, 9, 10, 12, 13, 15 and 16, assembles and is recruited to those DMVs^{95,148,155}.

The viral RNA is incorporated in those newly formed DMVs where it serves as a template for the RdRP transcription. The DMVs protect the dsRNA that is formed subsequently from recognition by host cell factors¹⁴⁰. First, full-length and subgenomic negative strand RNA are produced by the RdRP which used the viral RNA as a template. The viral genome has translation regulatory sequences (TRS) upstream of every ORF except ORF1ab. These TRS can halt the RdRP^{95,249}. The 3' end of the nascent RNA can then anneal with the TRS-loop (TRS-L) 70 nucleotide upstream of the 5' end of the positive strand RNA (**Figure 9**)⁹⁵. The transcription starts over at this TRS-L. This results in the production of subgenomic negative strand RNAs (sgRNA) with a leader sequence in their 3'-end. Those negative sense, full-length and sgRNAs are the templates used to produce positive sense genomic RNAs (gRNA) and subgenomic RNAs (**Figure 9**)⁹⁵.

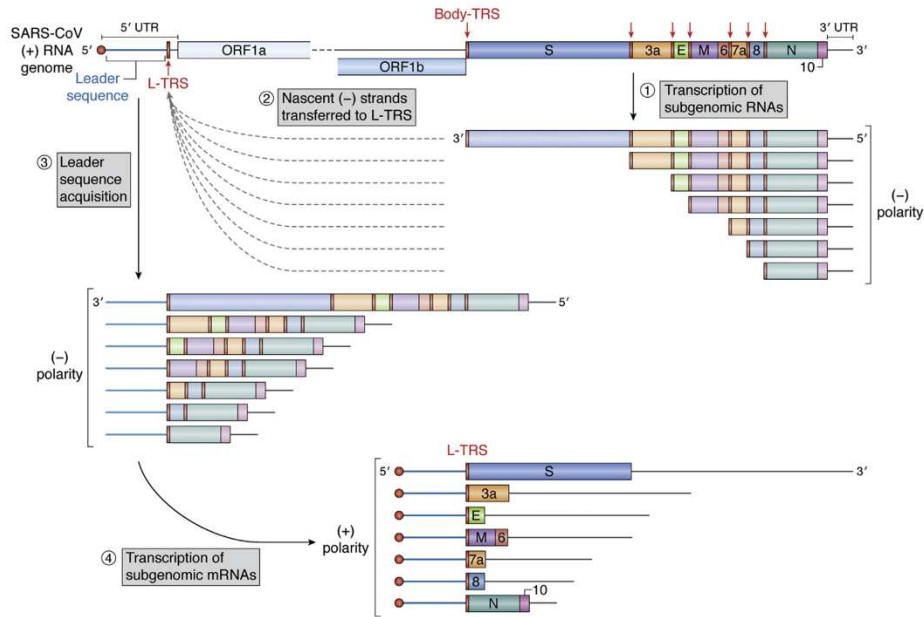


Figure 9: Coronavirus genome replication. A negative sense RNA is first synthesized from the incoming viral RNA. The presence of TRS along the genome results in the production of multiple sgRNAs, which are then transcribed into subgenomic mRNAs. From Grellet *et al.* 2022²⁴⁹

NSP1 blocks the host translation by binding to the 18S subunit of the ribosome next to the mRNA entry channel¹³⁸. However, the presence of a leader sequence in the 5'-end of the mRNA is sufficient to abrogate this inhibition¹³⁸. The sgRNA all possess this leader sequence and are thus translated in structural and accessory proteins in the ER. These newly synthesized proteins further support the replication of the virus and silence various pathways of the immune response.

1.4.3. Virion formation

The newly translated structural proteins are targeted to the ER-Golgi intermediate compartment (ERGIC). E and M interact with S, to induce S retention in the ERGIC and drive its integration in virions^{250,251}. M then recruits the N-RNA complex. The virion can then egress and exit the cells by exocytosis²⁵² or through the lysosome pathway²⁵³.

The full replication cycle is summarized in **Figure 10**.

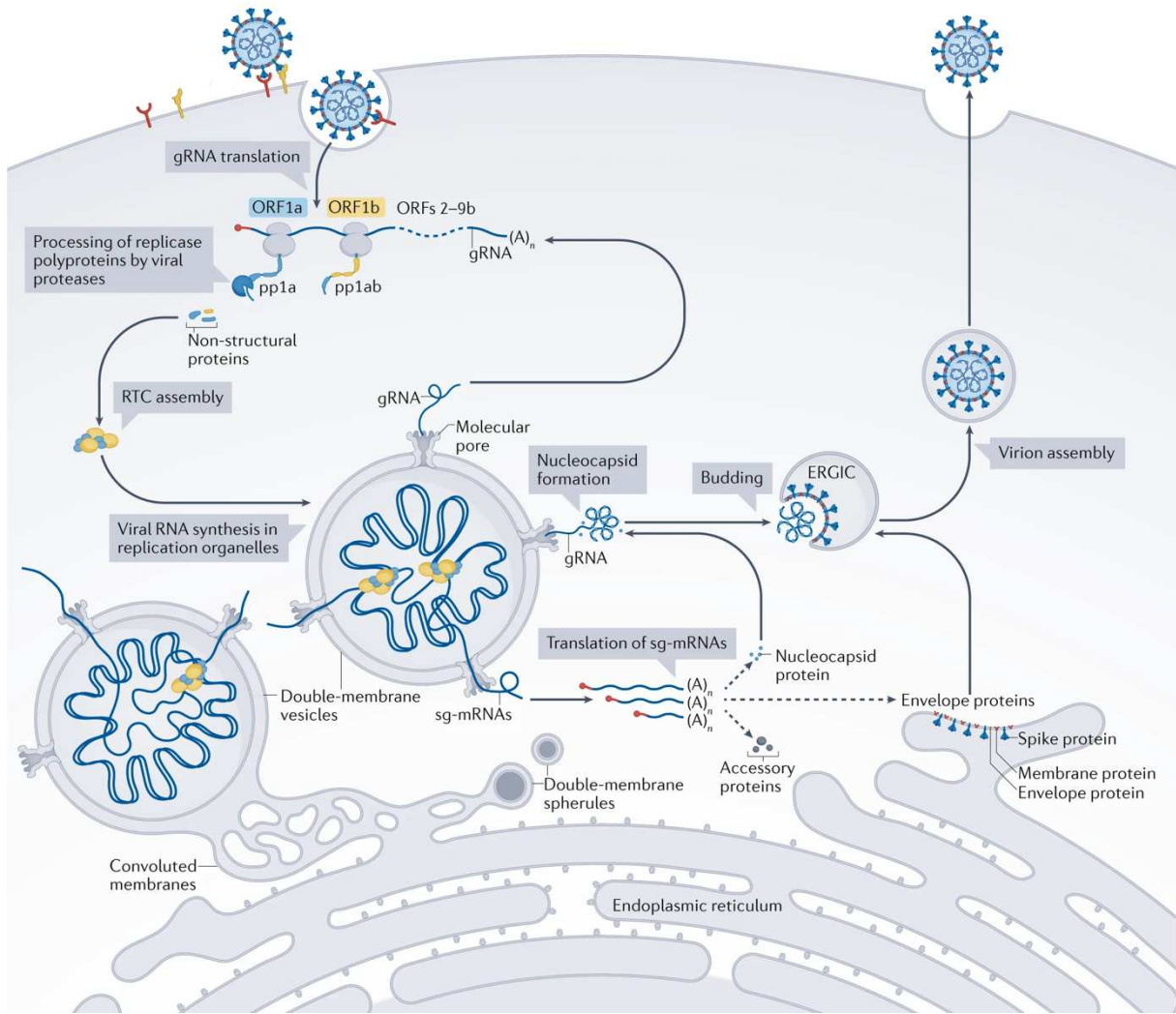


Figure 10: Coronavirus replication cycle. The incoming viral RNA is translated by the cell. This results in the production of NSPs which form the replication transcription complex (RTC) and induce the formation of DMVs. The viral RNA is replicated in those DMVs, gRNA and sgRNA are formed. Accessory and structural proteins are translated from these sgRNA and enable the formation of viral particles from the ERGIC. These viruses are released in the extracellular space. From Malone *et al.* 2022²⁵⁴.

1.5. Immune response to coronaviruses

1.5.1. Innate immune response

1.5.1.1. Intrinsic cellular response

The infection produces pathogen-associated molecular pattern (PAMP), among which ss and dsRNA. These PAMPs can be sensed by toll-like receptors (TLR) 3 and 7²⁵⁵, MDA5, LGP2 and RIG-I^{256–258}. Several viral proteins can also be recognized as PAMP: for instance, TLR2 recognizes E. The infection can also trigger damage to the mitochondria and release of mtDNA. This DNA is recognized by cGAS, which activates the cGAS-STING pathway²⁵⁹.

This induces the translocation of the NF- κ B, IRF3 and/or IRF7 transcription factors to the nucleus, which induce the production of cytokines as well as type I and III IFN. These interferons are secreted and can bind to IFN receptors in an autocrine or paracrine manner; this activates the JAK/STAT pathway and the production of ISGs. Several coronavirus proteins antagonize this immune response¹⁶¹. Their identified role in SARS-CoV-2 infection is shown in **Figure 11**.

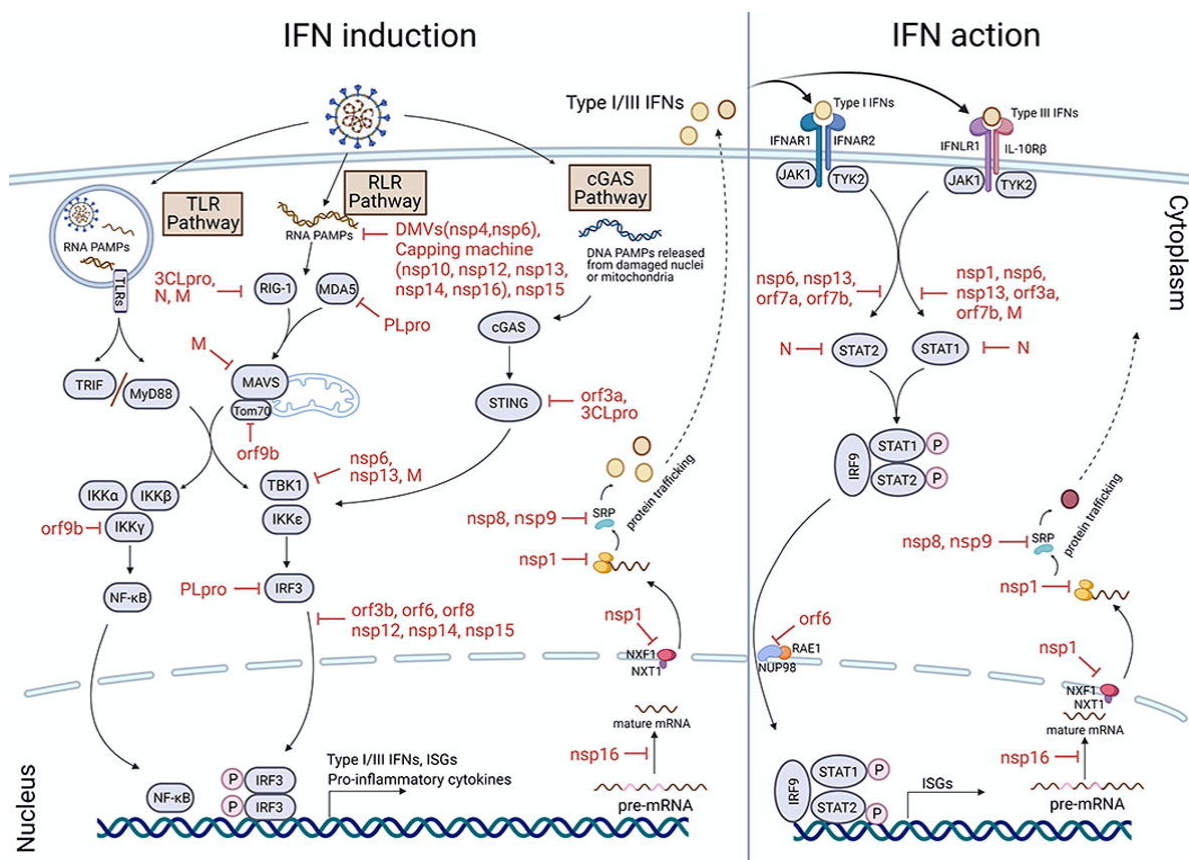


Figure 11: Innate immune response antagonism by SARS-CoV-2. Several structural, non-structural and accessory proteins from SARS-CoV-2 block the IFN induction and the IFN sensing pathways. From Min *et al.* 2022²⁶⁰.

The ISGs can impact viral entry, replication, or egress. For instance, Ly6E blocks entry²⁶¹, IFIT 3 and 5 block replication²⁶², BST2 prevents egress²⁶³, and the role of many more ISGs has been shown²⁶². The role of interferon induced transmembrane proteins (IFITMs) is debated, as both restricting and enhancing effects take place depending on the system studied and the virus^{240,264-268}.

1.5.1.2. Cellular mediated innate immunity

The innate immune response to coronavirus infection is also mediated by immune cells. Infected cells produce cytokines that lead to the recruitment of macrophages, monocytes, neutrophils, and dendritic cells (DC) to the lungs²⁶⁹.

Macrophages and monocytes participate in the inflammation and the tissue repair. However, their effects are not necessarily beneficial for the infected host. Indeed, patients that present severe disease have dysregulated myeloid cells which produce high levels of inflammatory cytokines, notably of IL-6 and TNF²⁷⁰⁻²⁷²; this cytokine storm is partly responsible for the lung damage and the multiple organ failure observed in critically ill patients. For SARS-CoV-2, macrophages can be infected; they then undergo pyroptosis which further participates in the inflammation²⁷³.

SARS-CoV-2 N can activate the lectin pathway of the complement, and S the alternative route of the complement. The immune complex produced at later stages of the disease can also activate the complement. The complement can impact the course of the disease negatively by being overactivated and dysregulated in severe forms caused by epidemic coronaviruses^{274,275}.

Neutrophils are also activated during the infection by coronaviruses. They form neutrophil extracellular traps (NETs) and produce reactive oxygen species (ROS). While the NETs' role is to trap the pathogen, their excessive generation contributes to immunothrombosis in severely ill patients²⁷⁶. ROS also contributes to tissue damage in critically ill patients²⁷⁷.

DCs are recruited to the lung and participate in the restriction of coronaviruses spread by secreting IFN-I and III²⁷⁸. The ability of DCs to secrete IFN α and λ is affected in severely ill patients and inversely correlates with the severity of the disease^{279,280}. DCs also participate in the adaptive immune response by presenting coronavirus peptides in complex with MHC molecules. SARS-CoV-2 and 229E can infect DCs, which triggers their death and likely contributes to the immune evasion^{281,282}.

1.5.2. Adaptive immunity

1.5.2.1. T-cell immunity

Specific CD4 and CD8 T cells against coronaviruses can be detected in the blood of patients a couple of days after the onset on symptoms²⁸³⁻²⁸⁷. They target all viral proteins, although S, M and N specific T cells are more prevalent. CD4 specific T cells play a role in the activation of B cells and the recruitment of innate immune cells repair by producing cytokines (IFN γ , IL-2 and TNF- α) and chemokines (CXCL-9, 10 and 11). They are also essential for the establishment of B memory lymphocytes. CD8 T cells secrete cytotoxic molecules such as granzyme B and perforin to lyse infected cells.

In the severe cases of SARS-CoV, MERS-CoV and SARS-CoV-2, there is a reduction in the number CD4 and CD8 T cells in the blood^{288,289}. There could be several mechanisms at play, including the infiltration of T-cells in inflamed tissue, T-cell apoptosis and pyroptosis²⁹⁰. This lymphopenia can contribute to viral persistence and tissue damage in severely ill patients^{290,291}.

A long-lasting T cell memory response is established during the infection^{283,292,287,285}, or vaccination^{293,294}. Several studies show the role of those memory T-cells to protect against subsequent infection^{295,296}, but their relative contribution compared to the antibody response is unknown.

1.5.2.2. Humoral immunity

Specific antibodies against coronaviruses are found in the blood of patients one to two weeks after the onset of symptoms^{297,298}. For SARS-COV-2, 90% of these antibodies target the S protein²⁹⁹. IgGs, IgAs and IgMs are present in the blood and nasal mucosa of recovered patients but IgGs are most long-lived; IgAs and IgMs are no longer detected in the blood 2 months post-onset of symptoms³⁰⁰, while IgGs can persist for over 8 months³⁰¹. Patients affected with severe disease have higher antibodies titers than patients that have mild disease^{302,303}. However, patients that die from the disease have lower antibody levels; this could be explained by defect in the production of neutralizing antibodies and a less good control of the infection^{304,305}. Vaccination by mRNA vaccines generates or boosts IgGs both in the blood and in the nasal mucosa, but has been shown to have little effect of mucosal IgAs^{306,307}.

These antibodies have different roles. They can directly neutralize the virus or participate in antibody-dependent cellular cytotoxicity (ADCC) and antibody-dependent cellular phagocytosis (ADCP)^{308,309}. The levels of neutralizing antibodies correlate with the protection against infection^{310,311}.

1.5.2.3. Pre-existing immunity

During a lifetime, one is very likely to be exposed multiple times to coronaviruses. As mentioned previously, sero-epidemiological studies indicate that 90% of the population has been exposed to several seasonal coronaviruses infections^{34,312}. These infections can produce cross-reactive antibodies and T-cells. Those antibodies and T-cells could also mediate immunity against a coronavirus that has not been encountered before, whether seasonal or epidemic. It has for instance been shown that 35% of individuals that were naïve to SARS-CoV-2 had specific SARS-CoV-2 T-cells, likely due to the infection by endemic coronaviruses²⁸⁹. The role of these cross-reactive antibody and memory T-cells in seasonal coronavirus pathology is poorly studied; it is nearly impossible to track the exposure and colds of individuals over a lifetime.

For SARS-COV-2, contradictory results have been published: several studies report cross-reactive responses do not protect against infection or severe disease³¹³ while others show they do³¹⁴⁻³¹⁷. One article demonstrates that cross-reactive immunity actually impairs the production of neutralizing antibodies against SARS-CoV-2³¹⁸. Meta-analysis show that pre-existing T cell immunity might be partially protective³¹⁹.

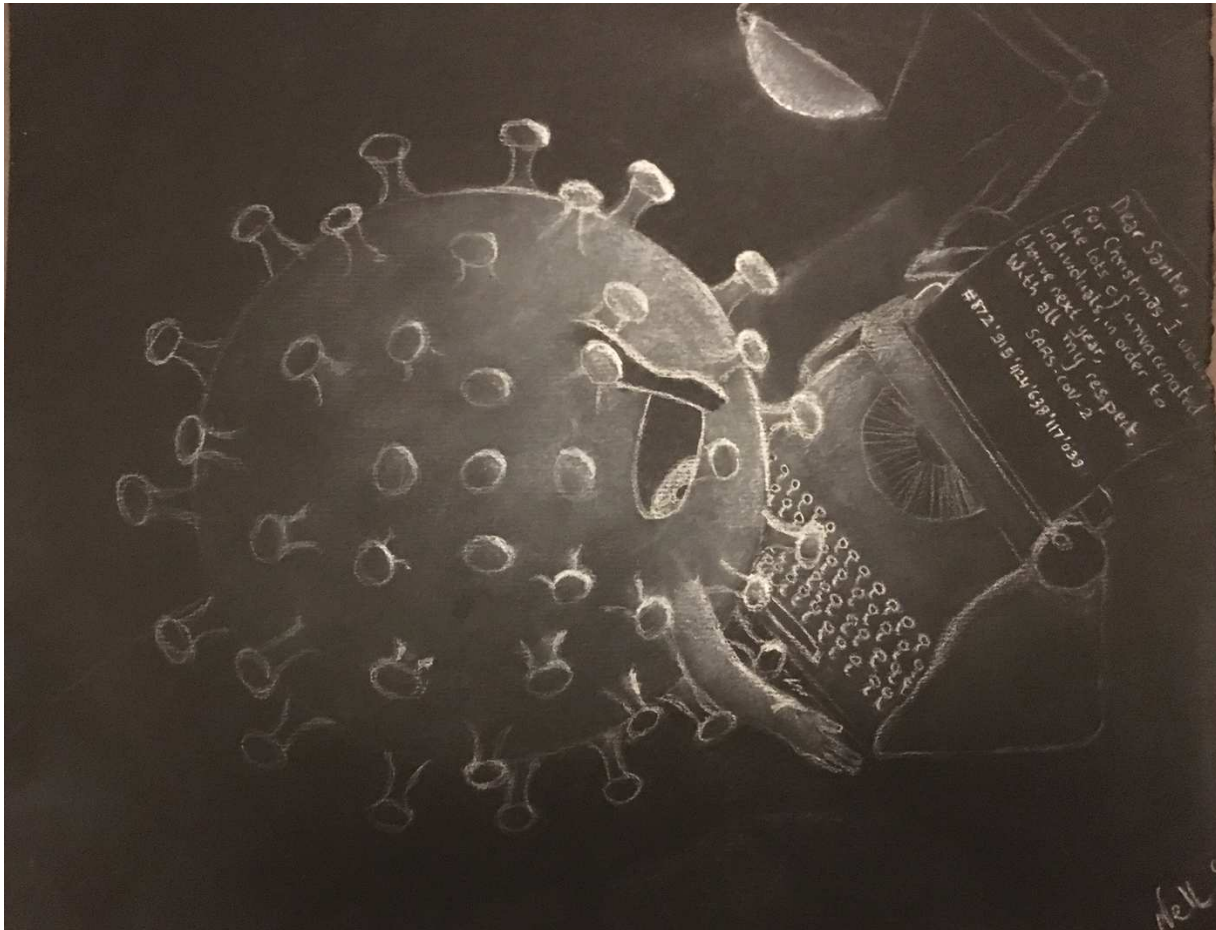
Research Objectives

My thesis has three main axes:

- During the peak of the SARS-CoV-2 pandemic, the characterization of the properties of new variants of SARS-CoV-2 and of their sensitivity to antibodies elicited by vaccination was of key importance for sanitary measures and clinical treatment of severe forms. Mapping the role of different mutations and their impact on different properties of the virus also participated in preparedness to the emergence of new variants. I thus studied for several variants of SARS-CoV-2 (Alpha, Beta, Delta: **Appendix 1, AY.4.2: Chapter 2, Part 1**, Omicron BA.1: **Chapter 2, Part 2**) their affinity for ACE2, their fusogenicity and their ability to escape recognition by antibodies present in the sera of vaccinated individuals or by monoclonal antibodies (therapeutic and non-therapeutic). I also performed fusion assays for a study on the importance of the D614G mutation for the appearance of mutations in the S1/S2 cleavage site (**Appendix 2**).
- I took on an observation made by a former PhD student, Michael Maaran Rajah. He showed that TMPRSS2 was sufficient to induce the cell-cell fusion mediated by HKU1. The initial goal of my project was to show the role of TMPRSS2 on the cleavage of the spike and on a putative unknown receptor. However, as experiences moved forward, we realized that TMPRSS2 might be acting as a receptor for HKU1. Showing TMPRSS2 acts as a receptor for HKU1 occupied the main part of my PhD (**Chapter 3, Part 1**). Collaborators at Institut Pasteur obtained the structure of HKU1 in complex with TMPRSS2. I then worked on the role of residues present at the interface of the complex, as well as on the conservation of TMPRSS2 among mammals (**Chapter 3, Part 2**).
- I monitored organelle remodeling in SARS-CoV-2 infected cells. Nanolive has commercialized a microscope based on holotomography that allows label-free live imaging in cells. Mitochondria, nuclei, nucleoli and lipid droplets are visible without any dyes. I used this microscope to acquire time-lapse images of SARS-CoV-2 infected cells. In collaboration with us, Nanolive developed a pipeline to analyze our images. We thus have a descriptive study of cellular remodeling in infected cells. (**Chapter 4**).

A general introduction on human coronaviruses opens this dissertation. Each chapter is preceded by a specific introduction and concluded by a specific discussion.

Chapter 2 - Characterization of SARS-CoV-2 VOCs



The articles presented in this chapter are the following:

- **Saunders N***, Planas D*, Bolland WH, Rodriguez C, Fourati S, Buchrieser J, Planchais C, Prot M, Staropoli I, Guivel-Benhassine F, Porrot F, Veyer D, Péré H, Robillard N, Saliba M, Baidaliuk A, Seve A, Hocqueloux L, Prazuck T, Rey FA, Mouquet H, Simon-Lorière E, Bruel T, Pawlotsky JM, Schwartz O. *Fusogenicity and neutralization sensitivity of the SARS-CoV-2 Delta sublineage AY.4.2*. EBioMedicine. 2022 Mar;77:103934. *equal contributions.

This project was the result of a group effort, supervised by Prof. Olivier Schwartz. I present this work with full acknowledgment of the work performed by my co-author Dr Delphine Planas and all other authors. I performed the antibody binding to spike transfected and infected cells, the ACE2 binding and the study of the fusogenicity of the spikes. I wrote the manuscript with my supervisor Prof. Olivier Schwartz and did the work following peer-review.

- Planas D*, **Saunders N***, Maes P*, Guivel-Benhassine F, Planchais C, Buchrieser J, Bolland WH, Porrot F, Staropoli I, Lemoine F, Péré H, Veyer D, Puech J, Rodary J, Baele G, Dellicour S, Raymenants J, Gorissen S, Geenen C, Vanmechelen B, Wawina-Bokalanga T, Martí-Carreras J, Cuypers L, Sève A, Hocqueloux L, Prazuck T, Rey FA, Simon-Lorière E, Bruel T, Mouquet H, André E, Schwartz O. *Considerable escape of SARS-CoV-2 Omicron to antibody neutralization*. Nature. 2022 Feb;602(7898):671-675. *equal contributions.

This project was the result of a group effort supervised by Prof. Olivier Schwartz. I present this work with full acknowledgment of the work performed by my co-author Dr Delphine Planas and Dr Piet Maes and all other authors. I performed the antibody binding to infected cells.

These articles are reproduced with the authorization of the journals.

2.1. Introduction

The first variant, B.1, appeared in February 2020. Compared to the ancestral strain, referred to as Wuhan strain, it contained a single substitution in the Spike protein (D614G). This substitution has been extensively studied; it stabilizes the S trimer^{320,321}, increasing its fusogenic properties³²², which results in enhanced infectivity^{323–325} and transmissibility^{326,327}.

The Alpha, Beta and Gamma variants then slowly supplanted this B.1 strain in December 2020, before being themselves replaced by the Delta variant in May 2021 (**Figure 4**). The Delta variant caused a massive outbreak in India, as well as waves all over the world. We and others extensively studied these variants. The laboratory developed pipelines to assess their neutralization by serum of previously infected or vaccinated individuals^{63,65}, as well as methods to characterize the effect of single mutations in the spike. These variants were less neutralized by sera from previously infected or vaccinated individuals than the D614G strain, but vaccination remained protective. I participated in a study to assess the effects of the single point mutations present in these variants during my master thesis (**Appendix 1**)⁶². We notably identified single point mutations that enable increased antibody escape (K417N, Δ Y144, Δ 242-244), increased affinity for ACE2 (N501Y) or increased fusogenicity (D614G, P681R, D1118H). Furthermore, I participated in another study where we show that the P681R/H mutations increase the cleavage of the spike, but that they are dependent on the D614G mutation which stabilizes the S1/S2 interaction, suggesting this first mutation was a prerequisite to the appearance of these variants with increased fusogenicity (**Appendix 2**)³²⁸.

The Delta variant remained dominant worldwide for 7 months. Its spike contained 8 mutations compared to the B.1 ancestral strain that explain its increased transmissibility and decreased susceptibility to neutralization^{65,329}. T19R, G142D and Δ 156-157+R158G are predicted to significantly alter the NTD epitopes, and are thus expected to affect the recognition of S by neutralizing antibodies^{65,330}. L452R and T478K affect antibody escape and ACE2 binding³³¹. P681R increases spike fusogenicity and cleavage^{328,332}. D950N slightly increases spike fusogenicity³³².

The Delta further gave rise to different sublineages, among which AY.1 and AY.4.2. The first contained the K417N mutation that we characterized both in the B.1 and Delta background; it increases antibody escape⁶². The latter contained 2 additional mutations in the NTD of the spike (Y145H and A222V). Y145 is part of an epitope targeted by neutralizing antibodies. A222V was also present in a variant that emerged in Spain in the summer of 2020. It was uncharacterized in November 2021. An additional mutation, T95I was present in 93% of AY.4.2 sequences. This mutation arose independently in numerous variants, including Kappa and Iota, and its frequency slowly increased in the parental Delta strain as well (40% in November 2021). The frequency of sequences from this sublineage started rising in the UK. In November 2021,

AY.4.2 represented 15% of the sequence in the UK and had started spreading worldwide. We thus decided to study this sublineage.

In December 2021, a new variant, Omicron (BA.1), with over 30 mutations in the spike emerged. This variant supplanted all the other variants at an unprecedented speed and became predominant worldwide in under a month. We studied the capacity of this variant to escape neutralization by the antibodies generated in response to vaccination or infection by previous variants, as well as by therapeutic antibodies.

2.2. Fusogenicity and neutralization sensitivity of the SARS-CoV-2 Delta sublineage AY.4.2.

Nell Saunders,^{a,b,*} Delphine Planas,^{a,c,*} William H. Bolland,^{a,b} Christophe Rodriguez,^{d,e} Slim Fourati,^{d,e} Julian Buchrieser,^a Cyril Planchais,^f Matthieu Prot,^g Isabelle Staropoli,^a Florence Guivel-Benhassine,^a Françoise Porrot,^a David Veyer,^{h,i} H       P    ,^{h,i} Nicolas Robillard,^h Madelina Saliba,^h Artem Baidaliuk,^g Aymeric Seve,^j Laurent Hocqueloux,^j Thierry Prazuck,^j Felix A. Rey,^k Hugo Mouquet,^f Etienne Simon-Lori    ,^g Timoth     Bruel,^{a,c,**} Jean-Michel Pawlotsky,^{d,e,**} and Olivier Schwartz^{a,c,**†}

^aVirus & Immunity Unit, Department of Virology, Institut Pasteur; CNRS UMR 3569, Paris, France

^bUniversit     de Paris, Sorbonne Paris Cit    , Paris, France

^cVaccine Research Institute, Creteil, France

^dDepartment of Virology, H       Henri Mondor (AP-HP), Universit     Paris-Est, Cr      , France

^eInstitut Mondor de Recherche Biom      , INSERM U955, Cr      , France

^fLaboratory of Humoral Immunology, Department of Immunology, Institut Pasteur, INSERM U1222, Paris, France

^gG5 Evolutionary genomics of RNA viruses, Department of Virology, Institut Pasteur, Paris, France

^hH       Europ     Georges Pompidou, Laboratoire de Virologie, Service de Microbiologie, Paris, France

ⁱINSERM, Functional Genomics of Solid Tumors (FunGeST), Centre de Recherche des Cordeliers, Universit     de Paris and Sorbonne Universit    , Paris, France

^jCHR d'Orl    ns, service de maladies infectieuses, Orl    ns, France

^kStructural Virology Unit Institut Pasteur, Universit     de Paris, CNRS UMR3569, 75015 Paris, France

Summary

Background SARS-CoV-2 lineages are continuously evolving. As of December 2021, the AY.4.2 Delta sub-lineage represented 20 % of sequenced strains in the UK and had been detected in dozens of countries. It has since then been supplanted by Omicron. The AY.4.2 spike displays three additional mutations (T95I, Y145H and A222V) in the N-terminal domain when compared to the original Delta variant (B.1.617.2) and remains poorly characterized.

Methods We compared the Delta and the AY.4.2 spikes, by assessing their binding to antibodies and ACE2 and their fusogenicity. We studied the sensitivity of an authentic AY.4.2 viral isolate to neutralizing antibodies.

Findings The AY.4.2 spike exhibited similar binding to all the antibodies and sera tested, and similar fusogenicity and binding to ACE2 than the ancestral Delta spike. The AY.4.2 virus was slightly less sensitive than Delta to neutralization by a panel of monoclonal antibodies; noticeably, the anti-RBD Imdevimab showed incomplete neutralization. Sensitivity of AY.4.2 to sera from vaccinated individuals was reduced by 1.3 to 3-fold, when compared to Delta.

Interpretation Our results suggest that mutations in the NTD remotely impair the efficacy of anti-RBD antibodies. The spread of AY.4.2 was not due to major changes in spike fusogenicity or ACE2 binding, but more likely to a partially reduced neutralization sensitivity.

Funding The work was funded by Institut Pasteur, Fondation pour la Recherche M      cale, Urgence COVID-19 Fundraising Campaign of Institut Pasteur, ANRS, the Vaccine Research Institute, Labex IBEID, ANR/FRM Flash Covid PROTEO-SARS-CoV-2 and IDISCOVER.

Copyright      2022 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: SARS-CoV-2; Delta; AY.4.2; Vaccines; Fusogenicity; Neutralization

†Corresponding author.

E-mail address: olivier.schwartz@pasteur.fr (O. Schwartz).

* Co-first authors

** co-last authors

Introduction

The pandemic circulation of SARS-CoV-2 is associated with emergence of variants with increased inter-individual transmission or immune evasion properties. The Delta Variant of Concern (VOC), originally identified in India in 2020, has supplanted pre-existing strains worldwide in less than 6 months.^{1,2} The spike protein

Research in context

Evidence before this study

SARS-CoV-2 lineages are continuously evolving. Monitoring the emergence of lineages using the GISAID database or cov-lineages.com website, and assessing the reports of public health agencies allowed us to note the rise of the Delta sublineage AY.4.2 in UK and other countries. Very little was known regarding the biology and antibody neutralization sensitivity of this strain.

Added value of this study

This study uses well established tools to characterize three mutations that arouse in the spike protein of SARS-CoV-2. These three mutations did not impact the fusogenic properties of the Delta spike. The AY.4.2 virus was slightly less sensitive than the Delta ancestral strain to neutralization by therapeutic antibodies and sera from vaccinated individuals. Noticeably, the efficacy of the therapeutic anti-RBD antibody Imdevimab was reduced.

Implications of all the available evidence

Mutations and deletions of amino-acids 95, 141-145 and/or 222 of the spike appeared independently in several variants of concern and of interest (Omicron, Kappa, Alpha, Iota, B.1.177 or AY.4.2). The appearance of these mutations highlights the importance of convergent evolution in different variants. Understanding the role of these mutations in the biology of SARS-CoV-2 and their potential effect on antibody escape is essential to assess the spread and the sensitivity of SARS-CoV-2 variants to vaccines and therapeutic antibodies.

of Delta contains 9 mutations, when compared to the B.1 ancestral strain (D614G), including five changes in the NTD (T19R, G142D, Δ156, Δ157, R158G), two in the receptor binding domain (RBD) (L452R, T478K), one mutation close to the furin cleavage site (P681R) and one in the S2 region (D950N).³ This set of mutations reduces sensitivity to antibody neutralization, enhances the fusogenicity of the spike and improves viral fitness.^{3,4,5,6,7,8} The increased transmissibility of VOCs may also be due to mutations in other viral proteins, such as R203N in the nucleocapsid (N).⁹

The Delta lineage is heterogeneous and has continued to evolve. It can be divided into sublineages or clades¹⁰ (preprint)^{11,12} (preprint). Different classifications exist. Next strain has classified the Delta variant into 3 main clades (21A, 21I and 21J). The Pangolin nomenclature is more resolute and has designed almost 180 sublineages within these clades, all named AY as aliases to the B.1.617.2 lineages.¹³ Mutations fixed in one sublineage (e.g. spike: T19R, G142D or D950N) are also present at low frequencies in other sublineages.

This may reflect founder effects or similar selective pressures on these variants. One sublineage, termed AY.4.2 (or VUI-21OCT-01) has drawn attention due to its slow but continuous rise in UK between July and December 2021^{14,15} (preprint). AY.4.2 sequences from 45 countries have been uploaded to the GISAID database. As of Dec 18, 2021, around 62,000 genomes have been reported in the UK on GISAID, representing around 15% of reported Delta cases in this country between December 1 and 18, 2021. Its occurrence has since then drastically diminished, as the Delta lineages have been replaced by Omicron strains worldwide.^{16,17,18}

The AY.4.2 sub-lineage is notably defined by the presence of Y145H and A222V mutations that lie within the N-terminal Domain (NTD) of the spike. Their impact on spike function is poorly characterized. Through modelling, the Y145H substitution has been predicted to decrease spike stability, but this has not been experimentally demonstrated.¹⁹ The mutation is located close to residue 144, which is deleted in the Alpha variant. A 141-144 deletion has also been reported in several chronically SARS-CoV-2 infected immunocompromised individuals.²⁰ Furthermore, a 143-145 deletion is also observed in the Omicron variant.²¹ Deletions of aa 144 and adjacent residues may drive antibody escape.^{22,23} The A222V mutation was noted in the B.1.177 (or 20A.EU1) lineage that emerged in Spain and spread throughout Europe in summer 2020.²⁴ This lineage did not have obvious transmission advantage and its spread was mostly explained by epidemiological factors such as travelling.²⁴ When introduced into the D614G spike, the A222V substitution slightly but not significantly impacted neutralization of pseudoviruses by human convalescent sera.²⁵ The effect of combined Y145H and A222V mutations on the Delta spike background remains unknown. Of note, most AY.4.2 sequences (93%) now include the T95I mutation in the NTD of the spike, a substitution that was rarely observed in the original Delta B.1.617.2 lineage, but which gradually appeared and is now present in 40% of Delta sequences on GISAID. The T95I substitution was previously detected in the close B.1.617.1 lineage (also termed Kappa)²⁶. It was also present in the B.1.526 lineage (also termed Iota) that accounted for up to 30% of sequenced cases in New York City in early 2021.²⁷ It is also present in the Omicron variant.²¹ This substitution was found in two vaccinated individuals with breakthrough infections and selected in immunocompromised individuals with chronic COVID-19 treated with convalescent plasma and monoclonal antibodies.^{28,29} The T95 residue is located outside the NTD antigenic supersite and its contribution to immune evasion is poorly characterized.²⁶

Here, we studied the AY.4.2 spike by assessing its fusogenic activity, affinity to ACE2 and recognition by antibodies. We also isolated an infectious AY.4.2 strain and examined its sensitivity to a panel of monoclonal

antibodies and sera from individuals having received two or three vaccine doses.

Methods

Orléans Cohort of convalescent and vaccinated individuals. Since August 27, 2020, a prospective, monocentric, longitudinal, interventional cohort clinical study enrolling 59 non-infected healthy controls is ongoing, aiming to describe the persistence of specific and neutralizing antibodies over a 24-months period. This study was approved by the ILE DE FRANCE IV ethical committee. At enrolment, written informed consent was collected and participants completed a questionnaire which covered sociodemographic characteristics, and data related to anti-SARS-CoV-2 vaccination, if received (brand product, date of first and second vaccination). Following anti-SARS-CoV-2 vaccination blood sampling was performed monthly (ClinicalTrials.gov Identifier: NCT04750720). For the present study, we selected 27 vaccinated participants (11 with Pfizer and 16 with AstraZeneca), at 5 or 7 months post second dose for the AstraZeneca and Pfizer vaccines respectively and at one month post third dose for the Pfizer vaccine. Study participants did not receive any compensation.

Plasmids

A codon-optimized version of the reference Wuhan SARS-CoV-2 Spike (GenBank: QHD43416.1) was ordered as a synthetic gene (GeneArt, Thermo Fisher Scientific) and was cloned into a phCMV backbone (GenBank: AJ318514), by replacing the VSV-G gene. The mutations for the Alpha and Delta spikes were added in silico to the codon-optimized Wuhan strain and ordered as synthetic genes (GeneArt, Thermo Fisher Scientific) and cloned into the same backbone. The D614G spike plasmid was generated by introducing the mutation into the Wuhan reference strain via Q5 site-directed mutagenesis (NEB). The T95I, Y145H and A222V were successively introduced into the Delta spike by the same process. Plasmids were sequenced prior to use. The primers used for sequencing and the site-directed mutagenesis are listed in the tables S3A and S3B.

Cell lines

HEK 293T cells (CCLV Cat# CCLV-RIE 1018, RRID: CVCL_0063), U2OS cells (KCLB Cat# 30096, RRID: CVCL_0042), Vero E6 (ECACC Cat# 85020206, RRID: CVCL_0574) cells and derivatives were cultured in DMEM with 10% Fetal Calf Serum and 1% Penicillin/Streptomycin. Cell lines transduced with GFP1-10/11 and ACE2 expression vectors were previously described^{30,31} and grown with 1 µg/ml puromycin, 10 µg/ml blasticidin, respectively (InvivoGen). GFP-

split cells were validated by their ability to generate GFP expressing syncytia after infection by SARS-CoV-2 or transfection by SARS-CoV-2 spike. ACE2 expression in U2OS was validated by flow cytometry. Cells were either purchased from ATCC or gifts from members of the Institut Pasteur.

Cell-cell fusion assay

For cell–cell fusion assays, 3.5×10^5 293T cell lines stably expressing GFP1-10 were transfected in suspension with 50 ng of phCMV-SARS-CoV2-spike and 450 ng of pQCXIP-Empty for 30 min at 37°C. Cells were washed twice. For imaging, they were seeded at a confluency of 3×10^4 cells per well in a 96 well plate. Vero GFP-11 cells were added at a confluency of 1.5×10^4 cells per well. The GFP area and the number of nuclei were quantified 18h post-transfection using Harmony High-Content Imaging and Analysis Software, as previously described.^{30,31} For surface staining, they were seeded at a confluency of 6×10^4 cells per well and stained as described below using mAb 129.

S-Fuse neutralization assay

U2OS-ACE2 GFP1-10 or GFP 11 cells, also termed S-Fuse cells, become GFP+ when they are productively infected by SARS-CoV-2.^{30,31} Cells were tested negative for mycoplasma. Cells were mixed (ratio 1:1) and plated at 8×10^3 per well in a µClear 96-well plate (Greiner Bio-One). The indicated SARS-CoV-2 strains were incubated with mAb, sera or nasal swabs at the indicated concentrations or dilutions for 15 minutes at room temperature and added to S-Fuse cells. The nasal swabs and sera were heat-inactivated 30 min at 56°C before use. 18 hours later, cells were fixed with 2% PFA, washed and stained with Hoechst (dilution 1:10,000, Invitrogen). Images were acquired with an Opera Phenix high content confocal microscope (PerkinElmer). The GFP area and the number of nuclei were quantified using the Harmony software (PerkinElmer). The percentage of neutralization was calculated using the number of syncytia as value with the following formula: $100 \times (1 - (\text{value with serum} - \text{value in "non-infected"}) / (\text{value in "no serum"} - \text{value in "non-infected"}))$. Neutralizing activity of each serum was expressed as the half maximal effective dilution (ED50). ED50 values (in µg/ml for mAbs and in dilution values for sera) were calculated with a reconstructed curve using the percentage of the neutralization at the different concentrations. We previously reported a correlation between neutralization titres obtained with the S-Fuse assay and a pseudovirus neutralization assay.³²

Clinical history of the patient infected with AY.4.2

A nasopharyngeal swab collected from a 10-year-old boy tested positive for SARS CoV-2 on October 20th 2021,

was sent to Hôpital Henri Mondor sequencing platform in the context of a nationwide survey. Briefly, private and public diagnostic laboratories in France participate to the national SARS-CoV-2 genomic surveillance by providing a random subsampling of positive SARS CoV-2 samples to national sequencing platforms weekly.³³

Virus sequencing

The full-length SARS-CoV-2 genome of the virus isolated from the patient was sequenced using next-generation sequencing. Viral RNA was extracted from the nasopharyngeal swab in viral transport medium. Sequencing was performed with the Illumina COVID-Seq Test (Illumina, San Diego, California), using 98-target multiplex amplifications along the full SARS-CoV-2 genome. The libraries were sequenced with NextSeq 500/550 High Output Kit v2.5 (75 Cycles) on a NextSeq 500 device (Illumina). The sequences were demultiplexed and assembled as full-length genomes using the DRAGEN COVIDSeq Test Pipeline on a local DRAGEN server (Illumina). The sample was identified as AY4.2 according to the Pangolin nomenclature, before being submitted to the GISAID database.³⁴

Virus strains

The variant strains were isolated from nasopharyngeal swabs on Vero cells and amplified by one or two passages on Vero cells. The delta strain was isolated from a nasopharyngeal swab of a hospitalized patient returning from India. The swab was provided and sequenced by the laboratory of Virology of Hôpital Européen Georges Pompidou (Assistance Publique – Hôpitaux de Paris). Both patients provided informed consent for the use of the biological materials. Titration of viral stocks was performed on Vero E6, with a limiting dilution technique allowing a calculation of TCID₅₀, or on S-Fuse cells. Viruses were sequenced directly on nasal swabs, and after one or two passages on Vero cells. Sequences were deposited on GISAID immediately after their generation, with the following ID: B.1.617.2: ID: EPI_ISL_2029113; AY4.2: EPI_ISL_5748228.

3D representation of mutations on the surface of spike of Delta and AY.4.2

The panels in Figure S3.B were prepared using The PyMOL Molecular Graphics System, v.2.1 (Schrödinger). The atomic model used (Protein Data Bank: 6XR8) has previously been described.³⁵

Flow Cytometry

Vero cells were infected with the indicated viral strains at a multiplicity of infection (MOI) of 0.1. At 48h post-infection, cells were detached using PBS-EDTA and

transferred into U-bottom 96-well plates (20,000 cell/well). HEK293T cells were transfected in suspension using lipofectamine 2000 as per manufacturer's instruction (ThermoFischer), using 25% of phCMV-SARS-CoV2-spike and 75% of pQCXIP-Empty. 24h post-transfection, cells were detached using PBS-EDTA and transferred into U-bottom 96-well plates (50,000 cell/well). Cells were then incubated for 30 min at 4°C with the indicated mAbs (1 µg/mL) or Serum (1:300 dilution or as indicated for dose response) in MACS buffer (PBS, 5g/L BSA, 2mM EDTA). Cells were washed with PBS, and stained using anti-IgG AF647 (1:600 dilution in MACS, 30 min at 4°C) (ThermoFisher). Cells were then fixed for 30 min using PFA 4%. Data were acquired on an Attune Nxt instrument (Life Technologies).

For ACE2 binding, 293T cells transfected with S proteins for 24 hours were stained with soluble biotinylated ACE2 diluted in MACS buffer at indicated concentrations (from 20 to 0.08 µg/ml) for 30 min at 4°C. The cells were then washed twice with PBS and then incubated with Alexa Fluor 647-conjugated streptavidin (Thermo Fisher Scientific, 1:400) for 30 min at 4°C. Cells were then fixed for 30 min using PFA 4%. Data were acquired on an Attune Nxt instrument (Life Technologies). Analysis was performed with FlowJo 10.7.1 (Becton Dickinson).

Antibodies

The four therapeutic antibodies were kindly provided by CHR Orleans. Human anti-SARS-CoV2 mAbs were cloned from S-specific blood memory B cells of Covid19 convalescents (Planchais et al, manuscript in preparation). Recombinant human IgG1 mAbs were produced by co-transfection of Freestyle 293-F suspension cells (Thermo Fisher Scientific) as previously described³⁶ and purified by affinity chromatography using protein G sepharose 4 fast flow beads (GE Healthcare). Antibodies were validated by flow cytometry, by measuring their ability to bind spike expressing cells and not control cells.

Statistical analysis

Flow cytometry data were analyzed with FlowJo v10 software (Becton Dickinson). Calculations were performed using Excel 365 (Microsoft). Figures were drawn on Prism 9 (GraphPad Software). Statistical analysis was conducted using GraphPad Prism 9. Statistical significance between different groups was calculated using the tests indicated in each figure legend.

No statistical methods were used to predetermine cohort size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

Ethics

Our research complies with all relevant ethical regulation. The Orléans study was approved by a national external committee (CPP Ile de France IV, IRB No. 00003835). At enrolment a written informed consent was collected for all participants.

Role of funders

The funders of this study had no role in study design, data collection, analysis and interpretation, or writing of the article.

Results

Antibody recognition of the AY.4.2 variant spike

To characterize the function of the AY.4.2 spike, we introduced the T95I, Y145H and A222V signature mutations in an expression plasmid coding for the Delta spike protein.³⁷ We first examined the ability of the Delta and AY.4.2 spikes to bind to a panel of 14 anti-SARS-CoV-2 monoclonal antibodies targeting either the RBD or the NTD. We tested 4 clinically approved antibodies, Bamlanivimab (LY-CoV555), Etesevimab (LY-CoVo16), Casirivimab (REGN10933) and Imdevimab (REGN10987) targeting the RBD^{38,39} as well as 4 other anti-RBD (RBD-48, RBD-85, RBD-98 and RBD-109) and 6 anti-NTD (NTD-18, NTD-20, NTD-32, NTD-45, NTD-69 and NTD-71) antibodies derived from convalescent individuals (Planchais et al, in preparation). Neutralizing anti-SARS-CoV-2 mAbs targeting the RBD can be classified into 4 main categories depending on their binding epitope.^{40,41} RBD-48 and RBD-85 belong to the first category ('Class 1') and act by blocking binding of the 'up' conformation of RBD to ACE2.⁴¹ The precise epitopes of RBD-98 and RBD-109 are not yet defined but overlap with those of RBD-48 and RBD-85. Casirivimab and Imdevimab are mixed in the REGN-COV2 cocktail from Regeneron (RonapreveTM) and target different domains of the RBD. Casirivimab is a Class 1 antibody whereas Imdevimab binds to a lateral domain and belongs to the Class 3.³⁹ The anti-NTD antibodies bind uncharacterized epitopes within this domain, as assessed by Elisa (not shown).

We previously assessed the ability of most of these antibodies to recognize the spikes of Alpha, Beta and Delta variants.^{3,37} To study their activity against AY.4.2, we first transfected the plasmids expressing the Delta and AY.4.2 spike proteins into 293T cells and analyzed antibody binding by flow cytometry (Figure 1a). In line with our previous results, the Delta spike was recognized by 9 of the 16 antibodies.^{3,37} The AY.4.2 spike displayed the same binding profile as Delta's (Figure 1a).

Since the three mutations lie in the NTD, we extended our analysis to nine additional monoclonal antibodies targeting this domain. These antibodies were

also cloned from SARS-CoV-2 infected individuals and bind to uncharacterized epitopes (Planchais et al, in preparation). As a control we used mAbro, a pan-coronavirus antibody that targets an unknown but conserved epitope within the S2 region²¹ (Planchais, manuscript in preparation). They do not display any neutralizing activity against the ancestral Wuhan SARS-CoV-2 (not shown). Six out of the nine antibodies bound to the Delta and AY.4.2 spikes expressed at the cell surface, with various intensities (Figure 1b). There was no major difference in their binding to Delta and AY.4.2 spikes, except for NTD-53 which bound slightly more to AY.4.2 spike protein than to Delta's and, conversely, NTD-105 which bound slightly more to Delta spike protein than to AY.4.2's (Figure 1b).

We next examined the binding of antibodies present in the sera of vaccinated individuals to Delta and AY.4.2 spikes. We selected individuals that received either two doses of Pfizer vaccine, sampled 7 months post second dose (n=10), or three doses, sampled at least one month after the third dose (n = 10) (Table S1A). We also studied individuals immunized with two doses of AstraZeneca vaccine, sampled at 5 months post second dose (n=16) (Table S1B). Sera were tested at a 1:300 dilution, which allows a quantitative assessment of the antibody levels by flow cytometry.^{42,43} Overall antibody levels were similar after two doses of Pfizer or AstraZeneca vaccines, and increased by 8-fold after the boost of Pfizer vaccine (Figure 1c). There was no major difference in the binding to the Delta and the AY.4.2 spikes (Figure 1c). We then performed a titration of the antibody levels in a subset of 8 sera by serial dilutions and obtained similar binding titres for the two spikes (Figure S1a), confirming the results obtained at the 1:300 dilution.

Altogether, these results indicate that the T95I, Y145H and A222V mutations are not associated with significant changes in recognition of the spike by a panel of 24 monoclonal antibodies and by sera from vaccine recipients.

Fusogenicity and ACE2 binding of the AY.4.2 variant spike

We previously established a quantitative GFP-Split based cell-cell fusion assay to compare the fusogenic potential of mutant or variant spike proteins.^{30,37} In this assay, 293-T cells expressing part of the GFP protein (GFP1-10) are transfected with the spike plasmid. The transfected donor cells are then co-cultured with acceptor Vero cells expressing the other part of GFP (GFP11).³⁷ Upon cell-cell fusion, the syncytia become GFP positive and the fluorescent signal is scored with an automated confocal microscope.^{30,37} Of note, 293T cells were chosen as donors because they lack ACE2 and do not fuse with each other upon spike expression. Vero cells were selected as targets because they endogenously express ACE2 and are naturally sensitive to

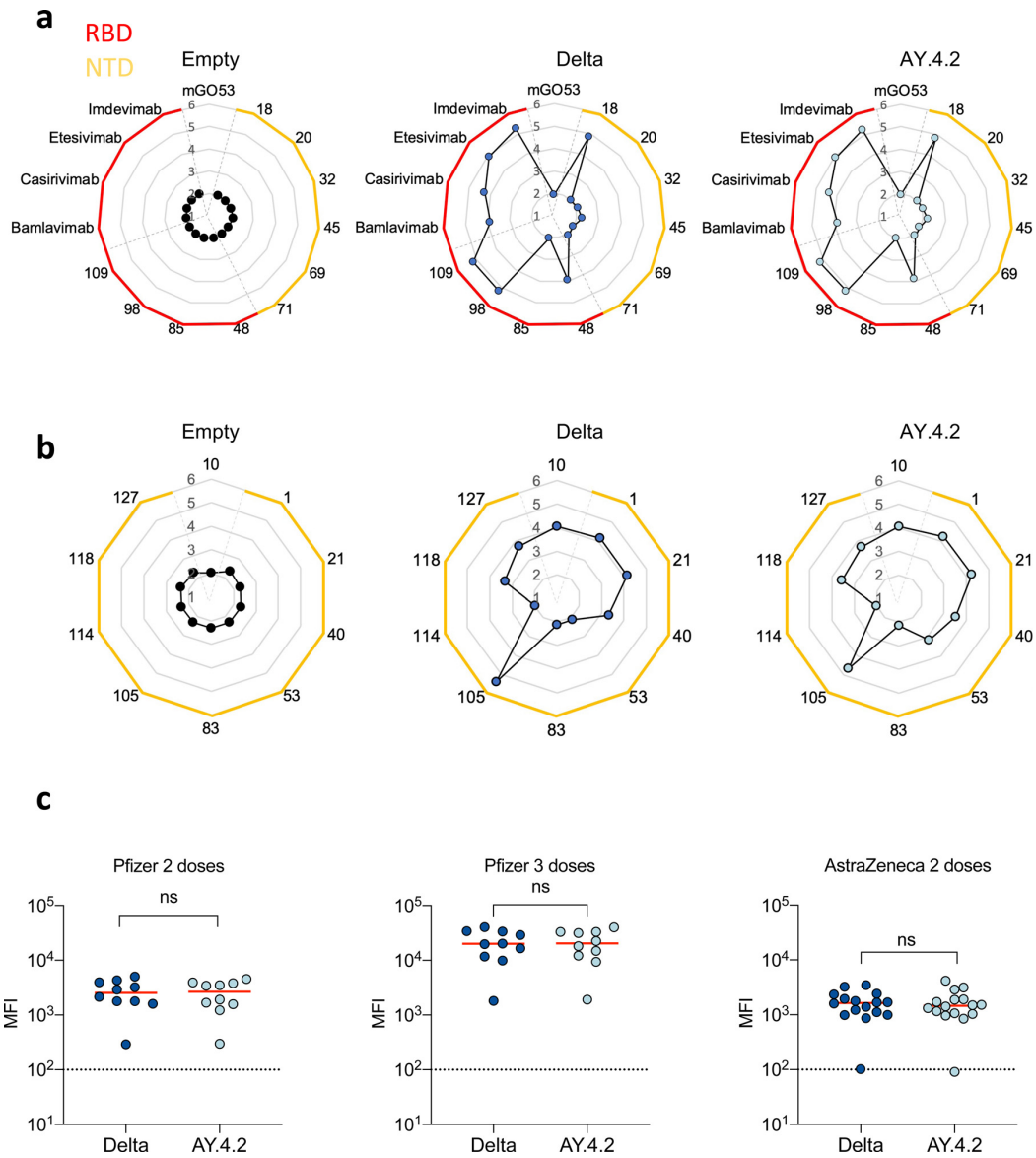


Figure 1. Antibody binding to cells expressing the Delta or AY.4.2 spike protein. (a-c) HEK293T cells were transiently transfected with plasmids expressing the Delta or AY.4.2 spike protein. (a-b) Binding of a panel of monoclonal antibodies targeting either the spike NTD or RBD. After 24h, cells were stained with the indicated antibody (1µg/mL). Radar charts represent for each antibody the logarithm of the median fluorescent intensity of the staining. (c) Binding of a panel of sera from vaccinated individuals. Sera from Pfizer vaccinated recipients were sampled at month 7 (M7) post- 2nd dose (n=10) and at month 8 (M8), 1 month post-third dose (n=10). Sera from AstraZeneca vaccinated individuals were sampled at M5 post full vaccination (n=16). After 24h, cells were stained with Sera (1:300 dilution). Statistical analysis: Wilcoxon signed rank test to compare the two viral strains. The comparisons were performed among the same participants. ns: non-significant [Wilcoxon signed rank test].

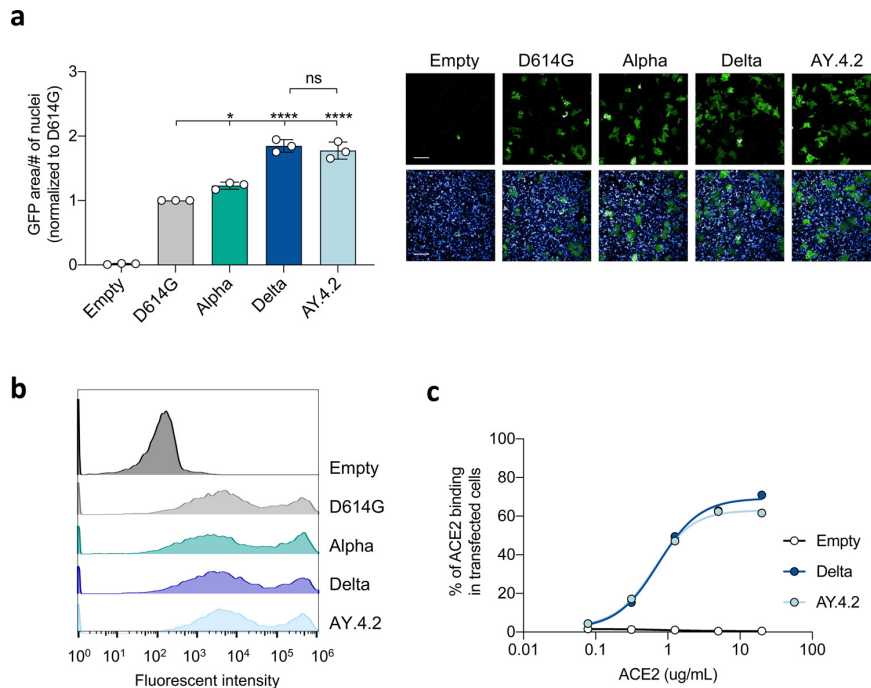


Figure 2. Comparison of Delta and AY.4.2 spike's fusogenicity and ACE2 affinity

(a-b) Donor 293T GFP1-10 cells were transfected with the indicated spike encoding plasmid. (a) Donor cells were added to Vero GFP11 acceptor cells to assess fusion, using an Opera Phenix microscope (Perkin Elmer). Left Panel: Fusion was quantified by using the total GFP area/number of nuclei before normalizing to D614G for each experiment. Data are mean \pm SD of three independent experiments. Statistical analysis: One-way ANOVA, each strain is compared to D614G or delta. ns: non-significant, *** $p < 0.001$ [One way ANOVA]. Right Panel: Representative images of one out of the three experiment. Green: GFP-Split, Blue: Hoechst. Scale bars: 200 μ m. (b) Donor cells were surface stained with a monoclonal anti-S antibody (mAb129) to quantify spike expression. The data was then acquired by flow cytometry. Data are mean \pm SD of three independent experiments. Statistical analysis: One-way ANOVA, each strain is compared to D614G or delta. ns: non-significant. (c) 293T cells were transfected with the indicated spike encoding proteins. After 24 h, they were stained with biotinylated ACE2 and fluorescent streptavidin before analysis by flow cytometry.

SARS-CoV-2. We thus analyzed the fusogenic activity of the AY.4.2 spike and compared it to the D614G, Alpha and Delta spikes. As previously reported, the D614G and Alpha spike variants were less fusogenic than the Delta spike (Figure 2a, b, Figure S2). The combination of T95I, Y145H and A222V substitutions did not modify the fusogenic activity of the Delta spike (Figure 2a).

We next explored AY.4.2 spike binding to the ACE2 receptor. To this aim, we transiently expressed the Delta and AY.4.2 spike proteins in 293T cells. Cells were then stained with a serial dilution of soluble biotinylated ACE2 and revealed with fluorescent streptavidin before analysis by flow cytometry (Figure 2c). We previously reported using this assay that the spike protein of Alpha had the highest affinity to ACE2, followed by Delta and then by D614G.^{3,37} Titration binding curves were generated with the Delta and AY.4.2 spikes, showing no difference between the spikes' affinity for ACE2 (Figure 2c).

Therefore, the fusogenicity and ACE2 binding of the AY.4.2 spike are similar to the ones of the parental Delta variant.

Isolation and characterization of an infectious AY.4.2 strain

We isolated the AY.4.2 variant from the nasopharyngeal swab of a symptomatic individual from the Paris region. The isolate was amplified by two passages on Vero E6 cells. Sequences of the swab and the outgrown viruses were identical and identified the AY.4.2 variant (GISAID accession ID: EPI_ISL_5748228, also termed hCoV-19/France/GES-HMN-21102260073/2021) (Figure S3a). In particular, the spike protein contains the 3 expected mutations in the NTD (T95I, Y145H and A222V) when compared to the Delta strain used here as a reference. It contains several mutations outside of the spike, all of them are characteristic of the AY.4.2.3

sublineage, except for nspr14 G143R (Figure S3a). The mutations present in both Delta and AY.4.2, or only AY.4.2 were mapped on the surface of the spike (Figure S3b). The 95 and 222 residues are buried in the NTD, while the Y145H is exposed on the surface of the NTD, in an epitope which is known to be targeted by neutralizing antibodies.³⁷ Viral stocks were titrated using S-Fuse reporter cells and Vero cells.^{30,31} S-Fuse cells allow rapid titration and measurement of neutralizing antibodies. They generate a GFP signal as soon as 6 hours post infection and the number of GFP+ cells correlates with the viral inoculum.^{30,31} Viral titres were similar in the two target cells and reached 10⁵ to 10⁶ infectious units/ml for the two strains. Syncytia were observed in infected Vero and S-Fuse cells (not shown). As expected, the syncytia were positive for spike staining (not shown).

We asked whether the spike present at the surface of infected cells displays the same characteristics as upon expression by transfection. We examined by flow cytometry the binding of neutralizing and non-neutralizing monoclonal antibodies to Vero cells infected with the Delta and AY.4.2 isolates. We observed the same profile of binding (Figure 3a,b) for the two strains, and no noticeable difference with transfected 293T cells.

Altogether, these results indicate that the profile of antibody binding is similar in spike-expressing transfected 293-T cells and Vero infected cells. AY.4.2 and Delta infected cells display the same affinity to the panel of monoclonal antibodies we tested.

Neutralization of AY.4.2 by monoclonal antibodies

We next compared the sensitivity of Delta and AY.4.2 strains to the previously described panel of neutralizing mAbs using the S-Fuse assay (Figure 4a). 8 out of 14 antibodies neutralized both strains. With most of the neutralizing antibodies, we observed a slightly increased IC₅₀s against AY.4.2 (median 2.2-fold increase when compared to Delta, Figure 4a and Table S2). Bamlanivimab was inactive against AY.4.2, in agreement with previous results with Delta.^{35,44} Imdevimab displayed an incomplete neutralization. The maximum neutralization plateaued at 60% against AY.4.2, even at high antibody concentrations (1 µg/mL), whereas it reached almost 100% against Delta (Figure 4a). This resulted in a statistically significant decrease in the maximal neutralization and an increase in the IC₅₀ (p<0.0001 and p<0.01 respectively [extra sum of tests F-test]). We obtained similar results with two different batches of Imdevimab (not shown). Therefore, AY.4.2 displays a slightly more elevated resistance to neutralization by the monoclonal antibodies tested than Delta. This resistance is more marked for Imdevimab.

Sensitivity of AY.4.2 to sera from vaccine recipients

We next asked whether vaccine-elicited antibodies neutralized AY.4.2. We used the same set of sera that were characterized by flow cytometry in Figure 1 and compared their neutralizing activity against Delta and AY.4.2.

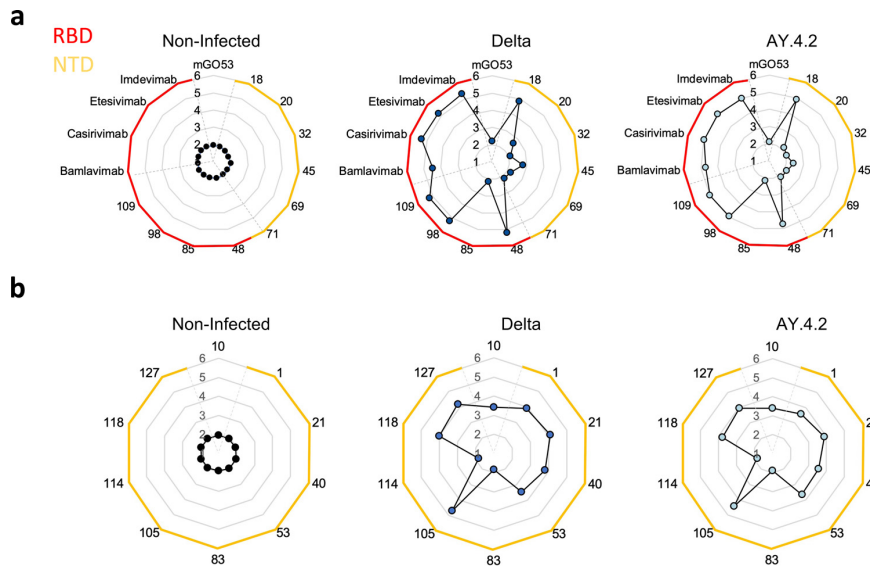


Figure 3. Antibody binding of the two viral isolates (a-b) Vero cells were transiently infected with Delta or AY.4.2 and harvested 48 hours post-infection for surface staining. (a-b) Binding of a panel of monoclonal antibodies targeting either the spike NTD or RBD, four of which are commercial therapeutic antibodies (1 µg/mL). Radar charts represent for each antibody the logarithm of the median fluorescent intensity of the staining.

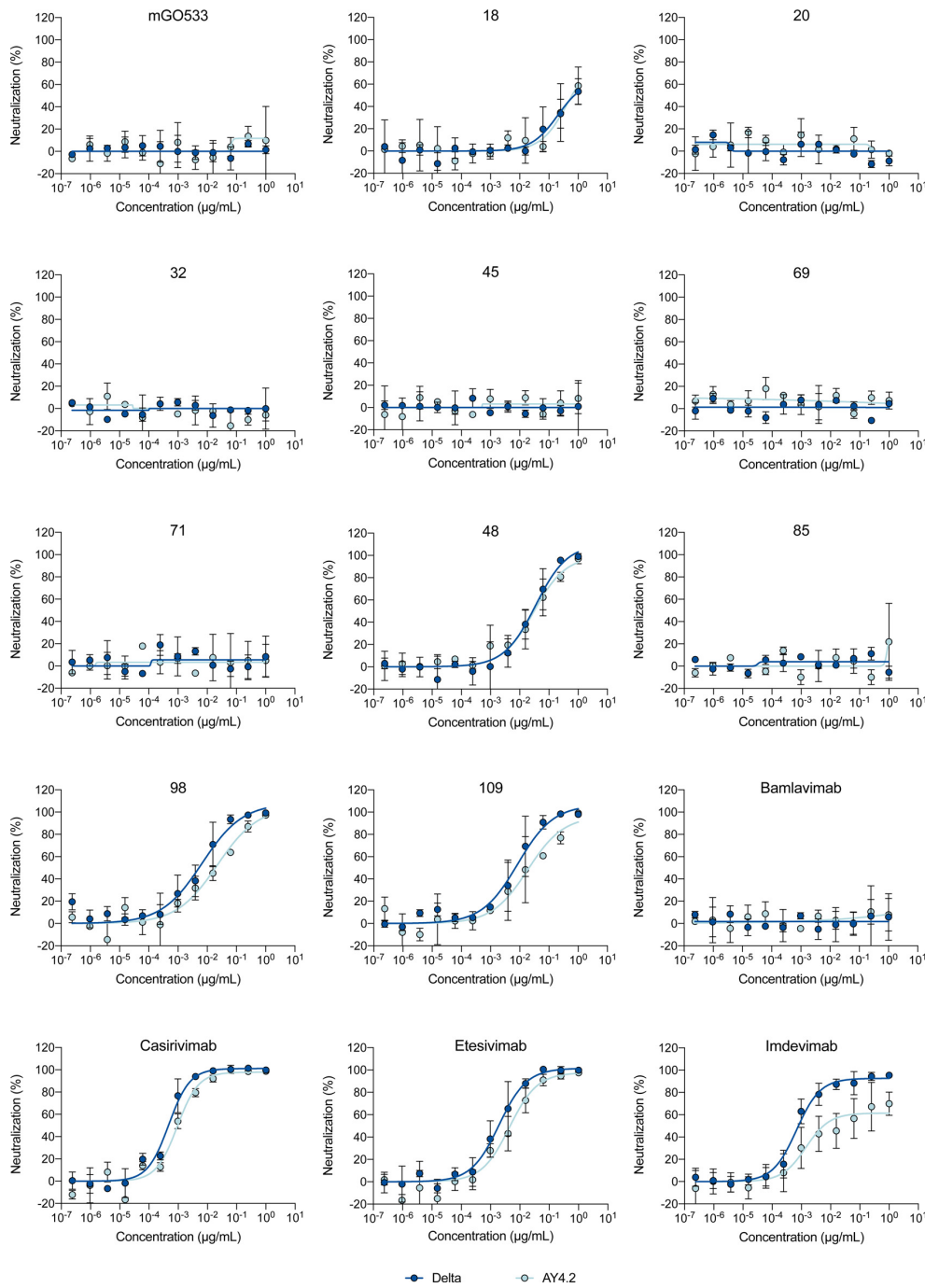


Figure 4. Neutralizing activity of monoclonal antibodies against the two viral isolates
 Dose response analysis of the neutralizing activity of a panel of monoclonal against the Delta and AY.4.2 viral isolates. Data are mean \pm SD of two independent experiments.

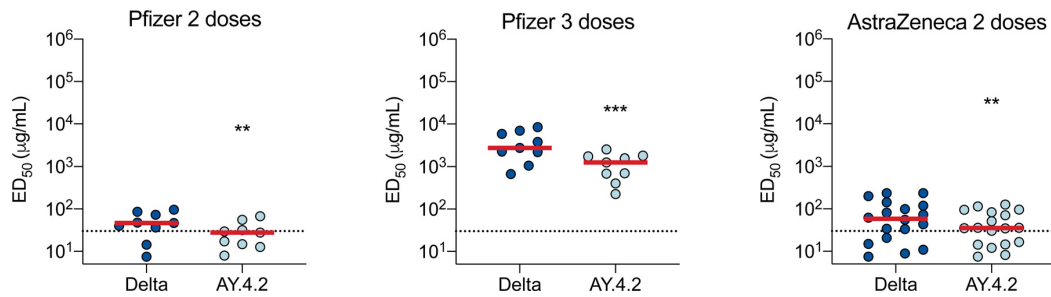


Figure 5. Neutralizing activity of vaccinated individuals' sera against the two viral isolates

ED50 of neutralization of the Delta and AY.4.2 viral isolates by sera from vaccine recipients. Sera from Pfizer vaccinated recipients were sampled at month 7 (M7) post-2nd dose (n=9) and at month 8 (M8), 1 month post-third dose (n=9). Sera from AstraZeneca vaccinated individuals were sampled at M5 post full vaccination (n=16). Data are mean from two independent experiments. Statistical analysis: Wilcoxon signed rank test to compare the two viral strains. The comparisons were performed among the same participants. **p<0.01, ***p<0.001 [Wilcoxon signed rank test].

With the Pfizer vaccine, seven months after the second dose, the levels of neutralizing antibodies were relatively low against Delta (median ED₅₀ of neutralization of 47), reflecting the waning of the humoral response at this time point³ (Figure 5a). These titres were slightly lower against AY.4.2 (ED₅₀ of 28). The median fold change was 1.6 (CI 96% [1.1:2.7]). One month after the booster dose (administered at M7 post vaccination), titres strongly increased (25–50 fold), reaching 2716 and 1260 for Delta and AY.4.2 strains, respectively (Figure 5b). The median fold change was 3.0 (CI 96% [2.4:3.4]).

A similar pattern was observed with the AstraZeneca vaccine. Five months after the second dose, the neutralizing titres against Delta and AY.4.2 were low (ED₅₀ of 58 and 35, respectively) (Figure 5c). The median fold change was 1.3 (CI 96% [1.1:2.1]).

Therefore, by using a set of sera with either low or high neutralizing antibody titres, we consistently observed a slight (1.3 to 3.0 median fold reduction) but significant decrease (p<0.01, [Wilcoxon signed rank test]) of their activity against AY.4.2, when compared to the parental Delta variant.

Discussion

Diversification of the Delta variant was regularly reported. The AY.4.2 sublineage was first identified in July 2021 and accounted for 15% and 20% sequenced Delta cases in UK, during the first and third weeks of November, respectively.¹⁴ This corresponds to an AY.4.2 logistic growth rate of 15% per week in this country.¹⁴ AY.4.2 has also been detected in dozens of countries. AY.4.2 was slowly but continuously rising and may thus have displayed a slight selective advantage compared to the parental Delta strain. An increase of the growth rate may depend on the context and should not be necessary interpreted as a change in biological transmissibility.¹⁴ Preliminary lines of evidence

indicated that hazard ratios for hospitalization or death were similar for Delta and AY.4.2, indicating that outcomes of AY.4.2 cases are not more severe than those of Delta cases. Since December 2021, Delta and its sublineages, including AY.4.2, have been replaced by Omicron. It remains however of interest to understand the parameters that may have favoured the spread of AY.4.2, compared to the original Delta strain.

The AY.4.2 strain remains poorly characterized. It carries 3 main substitutions, T95I, Y145H and A222V, compared to the parental Delta lineage. Here, we show that the AY.4.2 spike is functionally very close to that of Delta. By using a panel of 24 monoclonal antibodies targeting either the RBD or the NTD, we did not detect major differences in antibody recognition, when the spikes are expressed by transient transfection in 293T cells. Polyclonal sera from individuals having received either Pfizer or AstraZeneca vaccines similarly recognized the two spikes. Their fusogenic activity, when measured in a syncytia formation assay,³⁷ and the binding affinity to ACE2 were also similar for AY.4.2 and Delta. This is consistent with results showing little or no increase in the entry of vesicular stomatitis virus (VSV) pseudotyped with the spike protein of Delta and AY.4.2 in 293T, Caco2, Calu3 and Huh7 cells.⁴⁵

We isolated an authentic AY.4.2 strain from an infected patient and examined its sensitivity to antibody neutralization. We analyzed the profile of binding of a panel of monoclonal antibodies to infected cells. We did not observe major differences between the two strains.

We then studied the neutralization of the two viral isolates by a panel of monoclonal antibodies. Imdevimab, a therapeutic antibody used in combination with Casirivimab in the commercially approved REGN-COV2 cocktail from Regeneron and Roche (RonapreveTM), incompletely neutralized AY.4.2. Even at high concentration, the neutralizing activity plateaued at 60%. For instance, incomplete neutralization and

deviation from sigmoidal neutralization curves have been previously observed with some HIV broadly neutralizing antibodies (bNAbs).⁴⁶ This process has been attributed to heterogeneity in glycosylation of the HIV gp120/gp41 Env complex.⁴⁶ Our results suggest that the conformation of AY.4.2 and Delta differ slightly, which results in a decrease of Imdevimab neutralization. Our observation differs with results obtained in VeroE6 cells using VSV pseudotyped with the spike protein of Delta and AY.4.2., where no difference was observed in Imdevimab neutralization. This could be due for instance to differences in glycosylation or other post-translational modifications, which differ from a cell type to another.⁴⁵ It is also likely that neutralization assays performed with VSV-based pseudotypes may provide slightly different results than those using authentic viruses.

As AY.4.2 does not harbour mutations within the epitope of Imdevimab, our results also indicate that mutations in the NTD of the spike may remotely impact the accessibility of anti-RBD antibodies. The 3D structure of the spike shows that some regions of the NTD are in close proximity to the RBD.^{44,47} Imdevimab binds to a lateral region of the RBD and (Class 3 antibody) and may thus be more affected by changes in the NTD than other anti-RBD antibodies binding to the apex of the spike. Furthermore, the other neutralizing anti-RBD antibodies that we tested displayed a slight decrease in sensitivity to AY.4.2, when compared to Delta (1.8 to 3.3 fold increase of the IC₅₀), except for antibody 48. Of the 6 monoclonal antibodies targeting the NTD we tested, only NTD-r8³ neutralized Delta. NTD-r8 similarly neutralized both strains at high concentration. It will be worth determining whether other antibodies targeting the NTD super antigenic site^{39,48} are less active against AY.4.2.

We further show that sera from individuals having received two or three doses of Pfizer vaccine, or two doses of AstraZeneca, remained active against AY.4.2 despite a 1.3 to 3.0 fold reduction in neutralization titres. These results are consistent with a report showing a slight but non-significant decrease in AY.4.2 neutralization titres (1.5 fold) by sera from BNT162b2 vaccinated individuals.⁴⁹ This decrease may be attributed to the slight reduction of the efficacy of some Imdevimab-like antibodies in the serum, or targeting other RBD and NTD regions in the spike.

Preliminary epidemiology results of vaccine effectiveness in UK, for both symptomatic and non-symptomatic breakthrough infections, indicated no significant differences between AY.4.2 and non-AY.4.2 cases.⁵⁰ Our results indicate that the slight decrease in neutralizing titres reported here did not significantly impact vaccine effectiveness against AY.4.2, at least within 6-7 months post-vaccination.

Caveats and limitations

Our study lacks analysis of the AY.4.2 variant in more relevant cellular models. Future work in primary

human bronchial epithelium⁵¹ or viral competition experiments will help determining whether AY.4.2 is more fit than the parental delta lineage in cell culture systems. Furthermore, our study is limited by the relatively low number of tested sera. However, we obtained similar results with sera from patients having received two different vaccines and the decrease of sensitivity to neutralization observed between Delta and AY.4.2 was statistically significant.

Declaration of interests

C.P., H.M., O.S, T.B., F.R. have a pending patent application for an anti-RBD mAb not used in the present study (PCT/FR2021/070522).

Contributors

Experimental strategy design, experiments: NS, DP, WB, JB, CP, MP, IS, FGB, FP, AB, ESL, TB, OS

Verification of data: NS, DP, WB, JB, TB, OS

Vital materials: CR, SF, CP, DV, HP, NR, MS, JP, MP, AS, LH, TP, HM, JP

Manuscript writing: NS, OS

Manuscript editing: NS, DP, CR, SF, TB, JP, OS

All authors read and approved the final version of the manuscript.

Acknowledgments

We thank patients who participated to this study, members of the Virus and Immunity Unit for discussions and help, Nathalie Aulner and the UtechS Photonic Bio-Imaging (UPBI) core facility (Institut Pasteur), a member of the France BioImaging network, for image acquisition and analysis. We thank Julien Puech, Julien Rodary et Dhiaeddine Edriss from the Hopital Européen Georges Pompidou for their help with sequencing. We thank Fabienne Peira, Vanessa Legros and Laura Courtelmont for their help with the cohorts.

Work in OS lab is funded by Institut Pasteur, Fondation pour la Recherche Médicale (FRM), Urgence COVID-19 Fundraising Campaign of Institut Pasteur, ANRS, the Vaccine Research Institute (ANR-10-LABX-77), Labex IBEID (ANR-10-LABX-62-IBEID), ANR/FRM Flash Covid PROTEO-SARS-CoV-2 and IDIS-COVR. Work in UPBI is funded by grant ANR-10-INSB-04-01 and Région Ile-de-France program DIM-Health. NS is supported by the French Ministry of Higher Education, Research and Innovation. DP is supported by the Vaccine Research Institute. HM lab is funded by the Institut Pasteur, the Milieu Intérieur Program (ANR-10-LABX-69-01), the INSERM, REACTing, EU (RECOVER) and Fondation de France (#00106077) grants. ESL lab is funded by Institut Pasteur and the French Government's Investissement d'Avenir programme, Laboratoire d'Excellence "Integrative Biology

of Emerging Infectious Diseases” (grant n°ANR-10-LABX-62-IBED).

Data Sharing Statement

All data supporting the findings of this study are available within the paper and are available from the corresponding author upon request.

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.ebiom.2022.103934.

References

- Tao K, Tzou PL, Nouhin J, Gupta RK, de Oliveira T, Kosakovsky Pond SL, et al. The biological and clinical significance of emerging SARS-CoV-2 variants. *Nature Reviews Genetics*. 2021;22(12):757–773.
- Davies NG, Abbott S, Barnard RC, Jarvis CI, Kucharski AJ, Munday JD, et al. Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. *Science*. 2021;372(6538):eabg3055.
- Planas D, Veyer D, Baidaliuk A, Staropoli I, Guivel-Benhassine F, Rajah MM, et al. Reduced sensitivity of SARS-CoV-2 variant Delta to antibody neutralization. *Nature*. 2021.
- Mlcochova P, Kemp SA, Dhar MS, Papa G, Meng B, Ferreira IATM, et al. SARS-CoV-2 B.1.617.2 Delta variant replication and immune evasion. *Nature*. 2021;599(7883):114–119.
- Liu J, Liu Y, Xia H, et al. BNT162b2-elicited neutralization of B.1.617 and other SARS-CoV-2 variants. *Nature*. 2021.
- Liu C, Ginn HM, Dejnirattisai W, et al. Reduced neutralization of SARS-CoV-2 B.1.617 by vaccine and convalescent serum. *Cell*. 2021;184(16):4220–36.e13.
- Wall EC, Wu M, Harvey R, et al. Neutralising antibody activity against SARS-CoV-2 VOCs B.1.617.2 and B.1.351 by BNT162b2 vaccination. *The Lancet*. 2021;397(10292):2331–2333.
- Zhang J, Xiao T, Cai Y, et al. Membrane fusion and immune evasion by the spike protein of SARS-CoV-2 Delta variant. *Science*. 2021;0(0):eabl9463.
- Syed AM, Taha TY, Tabata T, et al. Rapid assessment of SARS-CoV-2 evolved variants using virus-like particles. *Science*. 2021;0(0):eabl6184.
- England PH. SARS-CoV-2 variants of concern and variants under investigation in England. *Technical briefing*. 2021;28. <https://www.gov.uk/government/publications/investigation-of-sars-cov-2-variants-technical-briefings>.
- Rono EK. Covid-19 genomic analysis reveals clusters of emerging sublineages within the delta variant. *Biorxiv*. 2021. 2021.10.08.46334.
- Stern A, Fleishon S, Kustin T, et al. The unique evolutionary dynamics of the SARS-CoV-2 Delta variant. *medRxiv*. 2021. 2021.08.05.21261642.
- WHO. Weekly epidemiological update on COVID-19 - 26 October 2021, Edition 63. <https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19-26-october-2021>. 2021.
- England PH. SARS-CoV-2 variants of concern and variants under investigation in England. *Technical briefing*. 2021;29. <https://www.gov.uk/government/publications/investigation-of-sars-cov-2-variants-technical-briefings>.
- Lassauniere R, Polacek C, Fonager J, et al. Neutralisation of SARS-CoV-2 Delta sub-lineage AY.4.2 and B.1.617.2+E484K by BNT162b2 mRNA vaccine-elicited sera. *medRxiv*. 2021. 2021.11.08.21266075.
- WHO. [https://www.who.int/news/item/26-11-2021-classification-of-omicron-\(b.1.1.529\)-sars-cov-2-variant-of-concern](https://www.who.int/news/item/26-11-2021-classification-of-omicron-(b.1.1.529)-sars-cov-2-variant-of-concern). 2021.
- Karim SSA, Karim QA. Omicron SARS-CoV-2 variant: a new chapter in the COVID-19 pandemic. *The Lancet*. 2021;398(10317):2126–2128.
- Viana R, Moyo S, Amoako DG, et al. Rapid epidemic expansion of the SARS-CoV-2 Omicron variant in southern Africa. *Nature*. 2022.
- Aljindan RY, Al-Subaie AM, Al-Ohali AI, Kumar D T, Doss C GP, Kamaraj B. Investigation of nonsynonymous mutations in the spike protein of SARS-CoV-2 and its interaction with the ACE2 receptor by molecular docking and MM/GBSA approach. *Computers in Biology and Medicine*. 2021;135:104654.
- Borges V, Isidro J, Cunha M, et al. Long-Term Evolution of SARS-CoV-2 in an Immunocompromised Patient with Non-Hodgkin Lymphoma. *mSphere*. 2021;6(4):e00244-21.
- Planas D, Saunders N, Maes P, et al. Considerable escape of SARS-CoV-2 Omicron to antibody neutralization. *Nature*. 2021.
- McCarthy KR, Rennick LJ, Nambulli S, et al. Recurrent deletions in the SARS-CoV-2 spike glycoprotein drive antibody escape. *Science*. 2021;371(6534):1139–1142.
- McCallum M, De Marco A, Lempp FA, et al. N-terminal domain antigenic mapping reveals a site of vulnerability for SARS-CoV-2. *Cell*. 2021;184(9):2332–47.e16.
- Hodcroft EB, Zuber M, Nadeau S, et al. Spread of a SARS-CoV-2 variant through Europe in the summer of 2020. *Nature*. 2021;595(7869):707–712.
- Wu J, Zhang L, Zhang Y, et al. The Antigenicity of Epidemic SARS-CoV-2 Variants in the United Kingdom. *Frontiers in Immunology*. 2021;12(2205).
- McCallum M, Walls AC, Sprouse KR, et al. Molecular basis of immune evasion by the Delta and Kappa SARS-CoV-2 variants. *Science*. 2021;0(0):eabl8506.
- West AP, Wertheim JO, Wang JC, et al. Detection and characterization of the SARS-CoV-2 lineage B.1.526 in New York. *Nature Communications*. 2021;12(1):4886.
- Bailly B, Péré H, Veyer D, et al. Persistent Coronavirus Disease 2019 (COVID-19) in an Immunocompromised Host Treated by Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2)-Specific Monoclonal Antibodies. *Clinical Infectious Diseases*. 2021.
- Hacisuleyman E, Hale C, Saito Y, et al. Vaccine Breakthrough Infections with SARS-CoV-2 Variants. *New England Journal of Medicine*. 2021.
- Buchrieser J, Duffoo J, Hubert M, et al. Syncytia formation by SARS-CoV-2 infected cells. *The EMBO Journal*. 2020 e106267.
- Planas D, Bruel T, Grzelak L, et al. Sensitivity of infectious SARS-CoV-2 B.1.1.7 and B.1.351 variants to neutralizing antibodies. *Nature Medicine*. 2021.
- Sterlin D, Mathian A, Miyara M, et al. IgA dominates the early neutralizing antibody response to SARS-CoV-2. *Science Translational Medicine*. 2020:eabd2223.
- France SP. <https://www.santepubliquefrance.fr/etudes-et-enquetes/enquetes-flash-evaluation-de-la-circulation-des-variants-du-sars-cov-2-en-france>. 2021.
- Shu Y, McCauley J. GISAID: Global initiative on sharing all influenza data – from vision to reality. *Eurosurveillance*. 2017;22(13):30494.
- Cai Y, Zhang J, Xiao T, et al. Distinct conformational states of SARS-CoV-2 spike protein. *Science*. 2020;369(6511):1586–1592.
- Lorin V, Mouquet H. Efficient generation of human IgA monoclonal antibodies. *J Immunol Methods*. 2015;422:102–110.
- Rajah MM, Hubert M, Bishop E, et al. SARS-CoV-2 Alpha, Beta, and Delta variants display enhanced Spike-mediated syncytia formation. *The EMBO Journal*. 2021 e108944. n/a(n/a).
- Taylor PC, Adams AC, Hufford MM, de la Torre I, Winthrop K, Gottlieb RL. Neutralizing monoclonal antibodies for treatment of COVID-19. *Nature Reviews Immunology*. 2021.
- Starr TN, Greaney AJ, Addetia A, et al. Prospective mapping of viral mutations that escape antibodies used to treat COVID-19. *Science*. 2021;371(6531):850–854.
- Liu L, Wang P, Nair MS, et al. Potent neutralizing antibodies against multiple epitopes on SARS-CoV-2 spike. *Nature*. 2020;584(7821):450–456.
- Barnes CO, Jette CA, Abernathy ME, et al. SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies. *Nature*. 2020;588(7839):682–687.
- Grzelak L, Temmam S, Planchais C, et al. A comparison of four serological assays for detecting anti-SARS-CoV-2 antibodies in human serum samples from different populations. *Science Translational Medicine*. 2020;12(559):eabc3103.
- Duffoo J, Grzelak L, Staropoli I, et al. Asymptomatic and symptomatic SARS-CoV-2 infections elicit polyfunctional antibodies. *Cell Rep Med*. 2021;2(5):100275.
- Lucas C, Vogels CBF, Yildirim I, et al. Impact of circulating SARS-CoV-2 variants on mRNA vaccine-induced immunity. *Nature*. 2021.

- 45 Arora P, Kempf A, Nehlmeier I, et al. No evidence for increased cell entry or antibody evasion by Delta sublineage AY.4.2. *Cellular & Molecular Immunology*. 2022.
- 46 McCoy LE, Falkowska E, Doores KJ, et al. Incomplete Neutralization and Deviation from Sigmoidal Neutralization Curves for HIV Broadly Neutralizing Monoclonal Antibodies. *PLOS Pathogens*. 2015;11(8):e1005110.
- 47 Wrapp D, Wang N, Corbett KS, et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science*. 2020;367(6483):1260–1263.
- 48 Starr TN, Greaney AJ, Dingens AS, Bloom JD. Complete map of SARS-CoV-2 RBD mutations that escape the monoclonal antibody LY-CoV555 and its cocktail with LY-CoV016. *Cell Reports Medicine*. 2021;2(4).
- 49 Lassaunière R, Polacek C, Fonager J, et al. Neutralisation of SARS-CoV-2 Delta sub-lineage AY.4.2 and B.1.617.2+E484K by BNT162b2 mRNA vaccine-elicited sera. *medRxiv*. 2021. 2021.11.08.21266075.
- 50 England PH. SARS-CoV-2 variants of concern and variants under investigation in England. *Technical briefing*. 2021;27. <https://www.gov.uk/government/publications/investigation-of-sars-cov-2-variants-technical-briefings>.
- 51 Robinot R, Hubert M, de Melo GD, et al. SARS-CoV-2 infection induces the dedifferentiation of multiciliated cells and impairs mucociliary clearance. *Nature Communications*. 2021;12(1):4354.

Fusogenicity and neutralization sensitivity of the SARS-CoV-2 Delta sublineage AY.4.2

Supplementary Information

Table S1. Clinical Data regarding the Pfizer and Astra Zeneca vaccinated individuals.....	2
Table S2. EC50 of neutralizing monoclonal antibodies.....	3
Table S3. Primers used for Site Directed Mutagenesis and sequencing.	4
Figure S1. Antibody binding to cells expressing the Delta or AY.4.2 spike protein. (A) HEK293T cells were transiently transfected with plasmids expressing the different spike proteins. Dose response binding of a panel of sera from vaccinated individuals in the Astra Zeneca cohort.....	5
Figure S2. Comparison of Delta and AY.4.2 fusogenicity and ACE2 affinity.(A-B) Donor 293T GFP1-10 cells were transfected with the indicated spike encoding plasmid. (A) Donor cells were surface stained with a monoclonal anti-S antibody (129) to quantify spike expression. The data was then acquired by flow cytometry. Left panel: Percentage of positive cells. Right panel: Median fluorescent intensity in the positive cells. Data are mean \pm SD of three independent experiments. Statistical analysis: One-way ANOVA, each strain is compared to D614G or delta ns: non-significant.....	6
Figure S3. Comparison of the Delta and the AY.4.2 strains.(A) Schematic overview of the Delta and AY4.2 sublineages consensus sequences with a focus on the spike built with the Sierra tool 58. Amino acid modifications in comparison to the ancestral Wuhan-Hu-1 sequence (NC_045512) are indicated.....	7

a

ID	Sex	Age	Vaccine	1st dose	2nd dose	3rd dose	M7 sampling post-2nd dose		M1 sampling post-3rd dose	
							Date	Days	Date	Days
VAC PF #1	Female	64	Pfizer	7/Jan/21	28/Jan/21	1/Jul/21	/	/	31/Aug/21	61
VAC PF #2	Male	69	Pfizer	7/Jan/21	28/Jan/21	31/Aug/21	31/Aug/21	215	28/Sep/21	28
VAC PF #3	Male	60	Pfizer	8/Jan/21	29/Jan/21	31/Aug/21	31/Aug/21	214	12/Oct/21	42
VAC PF #4	Male	65	Pfizer	6/Jan/21	28/Jan/21	6/Sep/21	26/Aug/21	210	8/Oct/21	32
VAC PF #5	Male	53	Pfizer	6/Jan/21	26/Jan/21	31/Aug/21	30/Aug/21	216	6/Oct/21	36
VAC PF #6	Female	36	Pfizer	6/Jan/21	28/Jan/21	16/Nov/21	31/Aug/21	215	/	/
VAC PF #7	Male	74	Pfizer	6/Jan/21	29/Jan/21	13/Sep/21	7/Sep/21	221	11/Oct/21	28
VAC PF #8	Male	59	Pfizer	6/Jan/21	27/Jan/21	31/Aug/21	31/Aug/21	216	6/Oct/21	36
VAC PF #9	Male	52	Pfizer	7/Jan/21	28/Jan/21	3/Nov/21	31/Aug/21	215	/	/
VAC PF #10	Male	62	Pfizer	8/Jan/21	29/Jan/21	31/Aug/21	31/Aug/21	214	5/Oct/21	35
VAC PF #11	Female	63	Pfizer	31/Jan/21	31/Jan/21	6/Sep/21	3/Sep/21	215	11/Oct/21	35

b

ID	Sex	Age	Vaccine	1st dose	2nd dose	M6 sampling post-2nd dose	
						Date	Days
VAC AZ #1	Female	57	AstraZeneca	6/Mar/21	5/May/21	4/Oct/21	152
VAC AZ #2	Female	64	AstraZeneca	11/Feb/21	5/May/21	1/Oct/21	149
VAC AZ #3	Male	58	AstraZeneca	11/Feb/21	6/May/21	1/Oct/21	148
VAC AZ #4	Female	55	AstraZeneca	11/Feb/21	6/May/21	30/Sep/21	147
VAC AZ #5	Female	61	AstraZeneca	9/Feb/21	15/Apr/21	11/Oct/21	179
VAC AZ #6	Female	63	AstraZeneca	15/Feb/21	5/May/21	5/Oct/21	153
VAC AZ #7	Female	59	AstraZeneca	9/Feb/21	4/May/21	5/Oct/21	154
VAC AZ #8	Female	59	AstraZeneca	12/Feb/21	19/May/21	8/Oct/21	142
VAC AZ #9	Male	61	AstraZeneca	9/Feb/21	4/May/21	4/Oct/21	153
VAC AZ #10	Female	61	AstraZeneca	10/Feb/21	5/May/21	1/Oct/21	149
VAC AZ #11	Female	55	AstraZeneca	9/Feb/21	4/May/21	1/Oct/21	150
VAC AZ #12	Male	73	AstraZeneca	26/Mar/21	28/May/21	8/Oct/21	133
VAC AZ #13	Male	57	AstraZeneca	5/Feb/21	3/May/21	1/Oct/21	151
VAC AZ #14	Female	60	AstraZeneca	14/Feb/21	4/May/21	30/Sep/21	149
VAC AZ #15	Female	60	AstraZeneca	7/Apr/21	4/May/21	7/Oct/21	156
VAC AZ #16	Female	63	AstraZeneca	12/Feb/21	7/May/21	7/Oct/21	153
VAC AZ #17	Male	56	AstraZeneca	18/Feb/21	11/May/21	8/Oct/21	150
VAC AZ #18	Male	64	AstraZeneca	7/Feb/21	19/May/21	6/Sep/21	110

Table S1. Clinical Data regarding the (a) Pfizer and (b) Astra Zeneca vaccinated individuals.

Antibody	Delta EC50 ($\mu\text{g/mL}$)	AY.4.2 EC50 ($\mu\text{g/mL}$)	Fold change
48	34.8	30.5	0.9
98	6.6	22.0	3.3
102	1.3	2.4	1.9
109	8.4	18.7	2.2
Casirivimab	0.5	0.9	2.0
Etesivimab	1.9	4.4	2.3
Imdevimab	0.7	1.9	2.9

Table S2. EC50 of neutralizing monoclonal antibodies.

a

T95I_F	TTTGCCAGCATCGAGAAGTCC
T95I_R	GTACACGCCATCGTTGAAG
Y145H_F	GGACGTCTACCACCACAAGAACAAC
Y145H_R	AGGAAGGGGTCGTTGCAG
A222V_F	GGGATTCAGTGTGCTGGAACCCCTGGTG
A222V_R	TGTGGCAGATCGCGCACG
D614G_F	CTGTACCAGGGCGTGAATTGCACAGAGGTG
D614G_R	AACGGCCACCTGGTTGCT

b

phCMV For	CTCTTTCCTACAGTCCTGG
phCMV Int1 For	AGCGAGTTCGCGGTGTACAG
phCMV Int1_Delta For	GCCACCAACGTGGTCATCAA
phCMV Int2 For	CGCAAGCGCATTAGCAACTG
phCMV Int3 For	GCGTGCTGACCGAGAGTAAC
phCMV Int4 For	GCAACCTGCTGCTGCAGTAC
phCMV Int5 For	GTGGTCAACCAGAACGCTCAG
phCMV Int6 For	AGAACCACACAAGCCCCGAC
phCMV Rev	TAGCCAGAAGTCAGATGCTC

Table S3. Primers used for (a) Site Directed mutagenesis (b) Sequencing

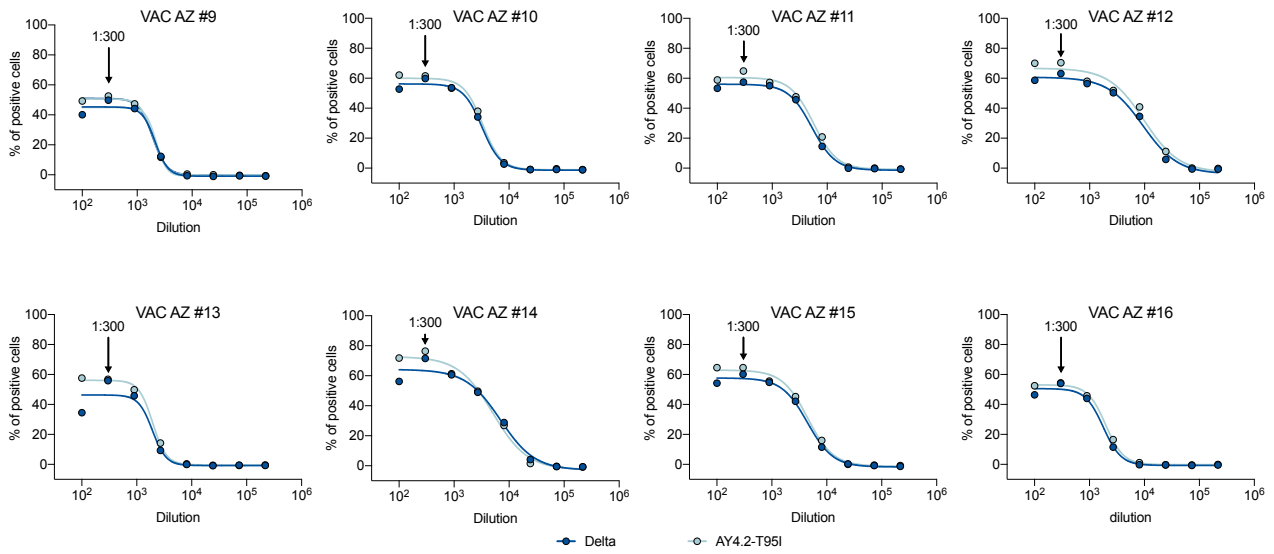


Fig. S1. Dose response binding of sera from vaccinated individuals.

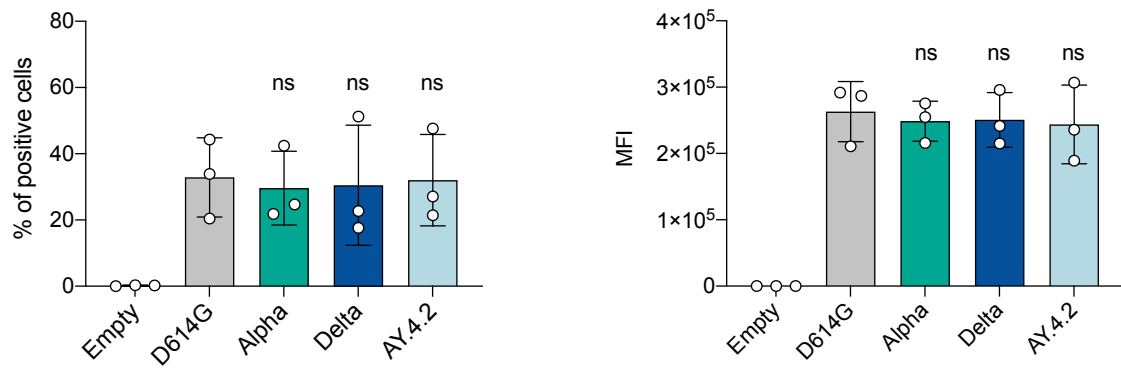
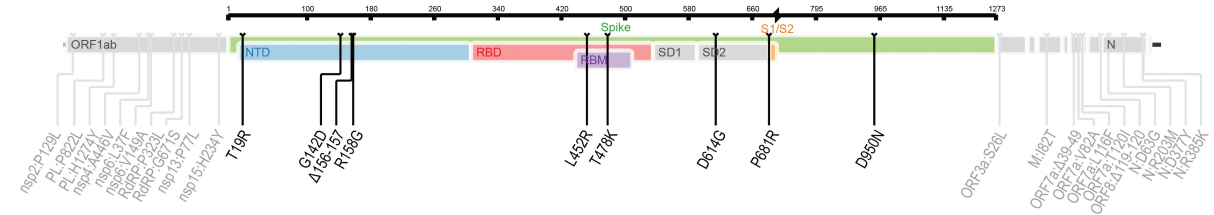


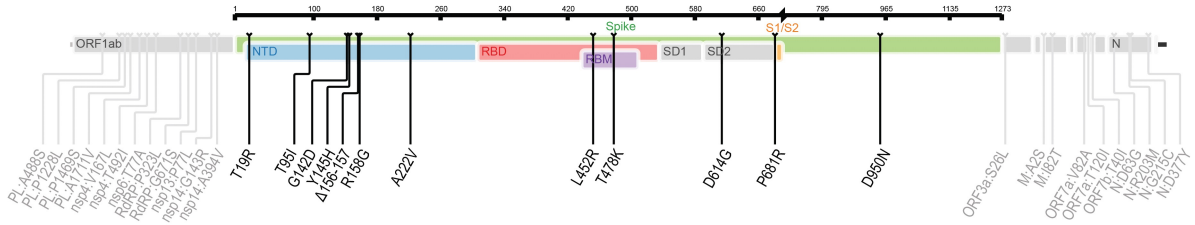
Fig. S2. Comparison of Delta and AY.4.2 spike surface expression.

a

B.1.617.2



AY.4.2



b

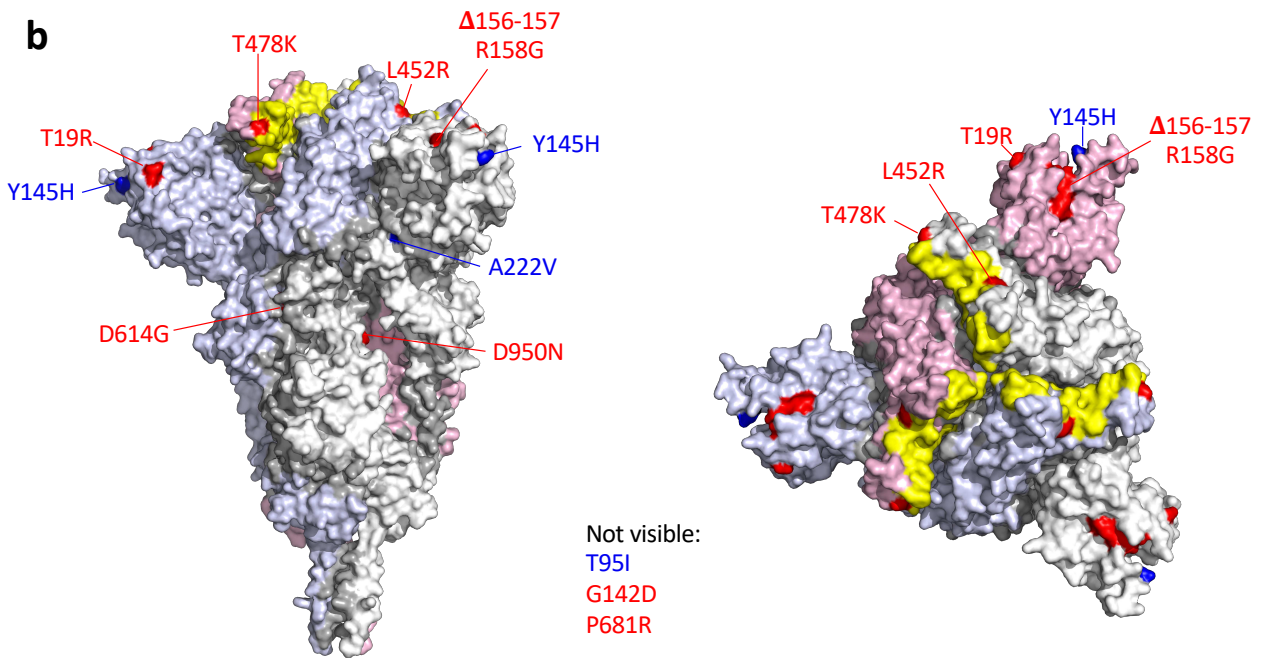


Fig. S3. Comparison of the Delta and the AY.4.2 strains.

2.3. Considerable escape of SARS-CoV-2 Omicron to antibody neutralization.

<https://doi.org/10.1038/s41586-021-04389-z>

Received: 10 December 2021

Accepted: 23 December 2021

Published online: 23 December 2021

 Check for updates

Delphine Planas^{1,2,15}, Nell Saunders^{1,3,15}, Piet Maes^{4,15}, Florence Guivel-Benhassine¹, Cyril Planchais⁵, Julian Buchrieser¹, William-Henry Bolland^{1,3}, Françoise Porrot¹, Isabelle Staropoli¹, Frederic Lemoine⁶, Hélène Péré^{7,8}, David Veyer^{7,8}, Julien Puech⁷, Julien Rodary⁷, Guy Baele⁴, Simon Dellicour^{4,9}, Joren Raymenants¹⁰, Sarah Gorissen¹⁰, Caspar Geenen¹⁰, Bert Vanmechelen⁴, Tony Wawina-Bokalanga⁴, Joan Martí-Carreras⁴, Lize Cuypers¹¹, Aymeric Sève¹², Laurent Hocqueloux¹², Thierry Prazuck¹², Félix A. Rey¹³, Etienne Simon-Loriere¹⁴, Timothée Bruel^{1,2,16}✉, Hugo Mouquet^{4,5,16}✉, Emmanuel André^{10,11,16}✉ & Olivier Schwartz^{1,2,16}✉

The SARS-CoV-2 Omicron variant was first identified in November 2021 in Botswana and South Africa^{1–3}. It has since spread to many countries and is expected to rapidly become dominant worldwide. The lineage is characterized by the presence of around 32 mutations in spike—located mostly in the N-terminal domain and the receptor-binding domain—that may enhance viral fitness and enable antibody evasion. Here we isolated an infectious Omicron virus in Belgium from a traveller returning from Egypt. We examined its sensitivity to nine monoclonal antibodies that have been clinically approved or are in development⁴, and to antibodies present in 115 serum samples from COVID-19 vaccine recipients or individuals who have recovered from COVID-19. Omicron was completely or partially resistant to neutralization by all monoclonal antibodies tested. Sera from recipients of the Pfizer or AstraZeneca vaccine, sampled five months after complete vaccination, barely inhibited Omicron. Sera from COVID-19-convalescent patients collected 6 or 12 months after symptoms displayed low or no neutralizing activity against Omicron. Administration of a booster Pfizer dose as well as vaccination of previously infected individuals generated an anti-Omicron neutralizing response, with titres 6-fold to 23-fold lower against Omicron compared with those against Delta. Thus, Omicron escapes most therapeutic monoclonal antibodies and, to a large extent, vaccine-elicited antibodies. However, Omicron is neutralized by antibodies generated by a booster vaccine dose.

In less than three weeks after its discovery, the Omicron variant was detected in dozens of countries. The WHO classified this lineage (previously known as Pango lineage B.1.1.529) as a variant of concern (VOC) on 26 November 2021 (ref. ¹). Preliminary estimates of its doubling time range between 1.2 days and 3.6 days in populations with a high rate of SARS-CoV-2 immunity^{2,5}. Omicron is expected to supplant the currently dominant Delta lineage during the next weeks or months. Little is known about its sensitivity to the humoral immune response. Recent reports indicated that Omicron has a reduced sensitivity to certain monoclonal and polyclonal antibodies^{6–10}, and CD8⁺ T cell epitopes that were previously characterized in other variants seem to be conserved in Omicron¹¹.

Isolation and characterization of an Omicron variant

We isolated an Omicron variant from a nasopharyngeal swab of an unvaccinated individual who developed moderate symptoms 11 days after returning to Belgium from Egypt. The virus was amplified by one passage in Vero E6 cells. The sequences of the swab and the outgrown virus were identical, and were identified as the Omicron variant (Pango lineage BA.1, GISAID: EPI_ISL_6794907 (swab) and EPI_ISL_7413964 (outgrown)) (Fig. 1a). The spike protein contained 32 changes compared with the D614G strain (belonging to the basal B.1 lineage), which we used here as a reference, including 7 changes in the N-terminal domain (NTD), with substitutions, deletions and a 3-amino-acid insertion (A67V, Δ69–70,

¹Virus and Immunity Unit, Institut Pasteur, Université de Paris, CNRS UMR3569, Paris, France. ²Vaccine Research Institute, Créteil, France. ³École Doctorale BioSPC 562, Université de Paris, Paris, France. ⁴Department of Microbiology, Laboratory of Clinical and Epidemiological Virology, Immunology and Transplantation, Rega Institute, KU Leuven, Leuven, Belgium. ⁵Humoral Immunology Laboratory, Institut Pasteur, Université de Paris, INSERM U1222, Paris, France. ⁶Hub de Bioinformatique et Biostatistique, Institut Pasteur, Université de Paris, CNRS USR 3756, Paris, France. ⁷Laboratoire de Virologie, Service de Microbiologie, Hôpital Européen Georges Pompidou, Paris, France. ⁸Functional Genomics of Solid Tumors (FunGeST), Centre de Recherche des Cordeliers, INSERM, Université de Paris, Sorbonne Université, Paris, France. ⁹Spatial Epidemiology Lab (SpELL), Université Libre de Bruxelles, Brussels, Belgium. ¹⁰Laboratory of Clinical Microbiology, Department of Microbiology, Immunology and Transplantation, KU Leuven, Leuven, Belgium. ¹¹Department of Laboratory Medicine, National Reference Centre for Respiratory Pathogens, University Hospitals Leuven, Leuven, Belgium. ¹²Service de Maladies Infectieuses, CHR d'Orléans, Orléans, France. ¹³Structural Virology Unit, Institut Pasteur, Université de Paris, CNRS UMR3569, Paris, France. ¹⁴G5 Evolutionary Genomics of RNA Viruses, Institut Pasteur, Université de Paris, Paris, France. ¹⁵These authors contributed equally: Delphine Planas, Nell Saunders, Piet Maes. ¹⁶These authors jointly supervised this work: Timothée Bruel, Hugo Mouquet, Emmanuel André, Olivier Schwartz. ✉e-mail: timothee.brue@pasteur.fr; hugo.mouquet@pasteur.fr; emmanuel.andre@uzleuven.be; olivier.schwartz@pasteur.fr

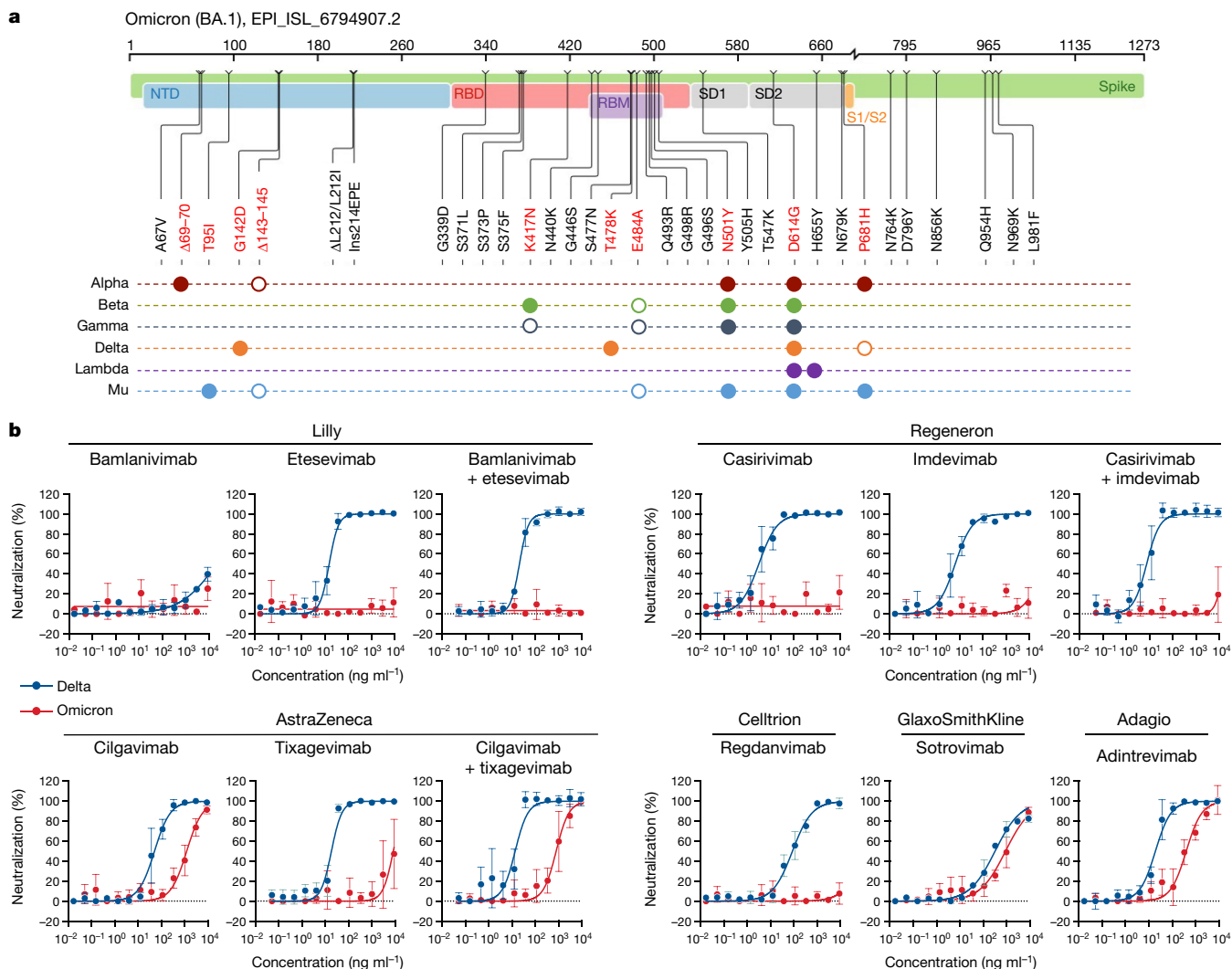


Fig. 1 | Neutralization of SARS-CoV-2 variants Delta and Omicron by clinical and preclinical monoclonal antibodies. a, The mutational landscape of the Omicron spike protein. The amino acid modifications are indicated in comparison to the ancestral Wuhan-Hu-1 sequence (GenBank: NC_045512). Consensus sequences of the spike protein were built using the Sierra tool³⁸. The Omicron sequence corresponds to the viral strain that was isolated in Belgium and used in the study (GISAID: EPI_ISL_6794907). Mutations are compared to some pre-existing VOCs and variants of interest. The filled circles indicate changes identical to Omicron. The open circles indicate different substitutions at the same position. **b**, Neutralization curves of monoclonal antibodies.

T95I, G142D, Δ141–143, Δ211L212I and Ins214EPE), 15 mutations in the receptor-binding domain (RBD) (G339D, S371L, S373P, S375F, K417N, N440K, G446S, S477N, T478K, E484A, Q493R, G496S, Q498R and N501Y, and Y505H), the T574K mutation, 3 mutations close to the furin cleavage site (H655Y, N679K and P681H) and 6 mutations in the S2 region (N764K, D796Y, N856K, Q954H, N969 and L981F) (Fig. 1a). This extensive constellation of changes is unique, but includes at least 11 modifications that have been observed in other lineages and VOCs or at sites that are mutated in other variants (Fig. 1a). Viral stocks were titrated using S-Fuse reporter cells and Vero cells. S-Fuse cells become GFP⁺ after infection, enabling rapid assessment of infectivity and the measurement of neutralizing antibody levels^{12–14}. Syncytia were observed in Omicron-infected S-Fuse cells (Extended Data Fig. 1). Syncytia were smaller after infection with Omicron, relative to Delta (Extended Data Fig. 1). Future experiments will help to determine whether the fusogenic potential of Omicron is different from that of other variants¹⁵.

Dose–response analysis of the neutralization by clinical or preclinical monoclonal antibodies (bamlanivimab, etesevimab, casirivimab, imdevimab, adintrevimab, cilgavimab, tixagevimab, regdanvimab and sotrovimab) and the indicated combinations (bamlanivimab + etesevimab, casirivimab + imdevimab (corresponding to monoclonal antibodies present in Ronapreve), cilgavimab + tixagevimab (corresponding to monoclonal antibodies present in Evusheld)) on Delta (blue dots) and Omicron (red dots) variants. Data are mean ± s.d. of three independent experiments. The IC₅₀ values for each antibody are presented in Extended Data Table 1.

Phylogenetic analysis of the Omicron lineage

We inferred a global phylogeny by subsampling SARS-CoV-2 sequences that are available on the GISAID EpiCoV database. To better contextualize the isolated virus genome, we performed a focused phylogenetic analysis using all of the Omicron samples deposited on GISAID on 6 December 2021 as background (Extended Data Fig. 2). The tree topology indicates that the Omicron lineage does not directly derive from any of the previously described VOCs. The very long branch of the Omicron lineage in the time-calibrated tree (Extended Data Fig. 2) might reflect a cryptic and potentially complex evolutionary history. At the time of writing, no Omicron genomic sequences from Egypt were available on GISAID, nor do we know of any sequences of travellers that used the same planes. The isolated strain genome showed no close connection to other Belgian Omicron infections. Follow-up analyses with additional

genomic data will improve the phylogenetic resolution to determine whether the patient was infected before or after returning to Belgium.

The mutational landscape of Omicron

We highlighted the 29 amino acid substitutions, the 3 amino-acid deletions and a 3-residue insertion that are present in the Omicron spike, with respect to the Wuhan strain, in a 3D model of the protein (Extended Data Fig. 3a). The 15 mutations in the RBD cluster around the trimer interface. The RBD is the target of the most potently neutralizing monoclonal antibodies against SARS-CoV-2, which have been divided into four classes depending of the location of their epitope^{4,16,17} (Extended Data Fig. 3b). Monoclonal antibodies in classes 1 and 2 compete for binding to human ACE2 (hACE2), whereas those from classes 3 and 4 bind away from the hACE2 interaction surface (Extended Data Fig. 3b). The epitopes of the class 2 and 3 monoclonal antibodies are exposed irrespective of the conformation of the RBD on the spike ('up' or 'down' configuration¹⁸), whereas those of classes 1 and 4 require an RBD in the up conformation. Whereas the previous VOCs displayed mutations only in the region targeted by class 1 and 2 monoclonal antibodies, Omicron mutations are located within the epitopes of all four classes of monoclonal antibodies. The mutations, insertion and deletions in the NTD might also affect the recognition of this domain by antibodies.

Neutralization of Omicron by monoclonal antibodies

We next assessed the sensitivity of Omicron to a panel of human monoclonal antibodies using the S-Fuse assay. We tested nine antibodies that are in clinical use or in development^{19–25}. These monoclonal antibodies belong to the four main classes of anti-RBD antibodies^{4,16,17}. Bamlanivimab and etesevimab (class 2 and class 1, respectively) are mixed in the Lilly cocktail. Casirivimab and imdevimab (class 1 and class 3, respectively) form the REGN-COV2 cocktail from Regeneron and Roche (Ronapreve). Cilgavimab and tixagevimab (class 2 and class 1, respectively) from AstraZeneca are also used in combination (Evusheld). Regdanvimab (Regkirona; Celltrion) is a class 1 antibody. Sotrovimab (Xevudy) by GlaxoSmithKline and Vir Biotechnology is a class 3 antibody that displays activity against diverse coronaviruses. It targets an RBD epitope outside the receptor-binding motif that includes N343-linked glycans. Adintrevimab (ADG20), developed by Adagio, binds to an epitope that is located between the class 1 and class 4 sites.

We measured the activity of the nine antibodies described above against Omicron and included the Delta variant for comparison purposes (Fig. 1b). As previously reported, bamlanivimab did not neutralize Delta^{14,26,27}. The other antibodies neutralized Delta with a 50% inhibitory concentration (IC_{50}) varying from 3.1 to 325 $ng\ ml^{-1}$ (Fig. 1b and Extended Data Fig. 4). Five antibodies (bamlanivimab, etesevimab, casirivimab, imdevimab and regdanvimab) lost antiviral activity against Omicron. The four other antibodies displayed a 2.8-fold to 453-fold increase in IC_{50} (ranging from 403 to 8,305 $ng\ ml^{-1}$) against Omicron. Sotrovimab was the only antibody that displayed a rather similar activity against both strains, with an IC_{50} of 325 and 917 $ng\ ml^{-1}$ against Delta and Omicron, respectively. We also tested the antibodies in combination to mimic the therapeutic cocktails. Bamlanivimab + etesevimab (Lilly) or casirivimab + imdevimab (Ronapreve) were inactive against Omicron. Cilgavimab + tixagevimab (Evusheld) neutralized Omicron with an IC_{50} of 773 $ng\ ml^{-1}$, corresponding to a 58-fold increase relative to Delta (Fig. 1b and Extended Data Fig. 4).

Next, using flow cytometry, we examined the binding of each monoclonal antibody to Vero cells infected with the Delta and Omicron variants (Extended Data Fig. 4). The five clinical antibodies that lost antiviral activity (bamlanivimab, etesevimab, casirivimab, imdevimab and regdanvimab) displayed a strong reduction (8-fold to 47-fold and 11-fold to 242-fold at 1 and 0.1 $\mu g\ ml^{-1}$, respectively) in their binding to Omicron-infected cells compared with Delta-infected cells, as

measured by the median fluorescence intensity of the signal (Extended Data Fig. 4). Cilgavimab, sotrovimab, tixagevimab and adintrevimab, which remained partly active, were less impaired in their binding to Omicron-infected cells (2-fold to 9-fold and 1.6-fold to 11-fold decrease at 1 and 0.1 $\mu g\ ml^{-1}$, respectively) (Extended Data Fig. 4).

Thus, Omicron completely or partially escapes neutralization by the tested antibodies. Our results are consistent with findings reported in recent preprints^{7,8,10}. The neutralization escape is correlated with a decrease in the binding of the antibodies to Omicron spike.

Sensitivity of Omicron to sera from vaccinees

We next examined whether vaccine-elicited antibodies neutralized Omicron. To achieve this, we randomly selected 54 individuals from a cohort established in the French city of Orléans, comprising vaccinated individuals who were not previously infected with SARS-CoV-2. The characteristics of the vaccinated individuals are shown in Extended Data Table 1. Sixteen individuals received the Pfizer two-dose vaccine regimen and 18 received the AstraZeneca two-dose vaccine regimen. Twenty individuals who were vaccinated with the Pfizer vaccine received a booster dose. We measured the potency of the sera of these individuals against the Delta and Omicron strains. We used the D614G ancestral strain (belonging to the basal B.1 lineage) as a control (Fig. 2a). We calculated the 50% effective dilution (ED_{50}) for each combination of serum and virus. Sera were first sampled five months after the full two-dose vaccination. For the Pfizer vaccine, the levels of neutralizing antibodies were relatively low against D614G and Delta (median ED_{50} of neutralization of 329 and 91), reflecting the waning of the humoral response¹⁴ (Fig. 2a). We did not detect any neutralization against the Omicron variant with these sera, except one, which displayed a low antiviral activity (Fig. 2a). The percentage of sera with detectable neutralizing activity is shown in Extended Data Fig. 5.

A similar pattern was observed for the AstraZeneca vaccine. Five months after vaccination, the levels of antibodies neutralizing Delta were low (ED_{50} of 187 and 68 against D614G and Delta, respectively). No antiviral activity was detected against Omicron in 90% of the sera (Fig. 2a and Extended Data Fig. 5).

We next examined the effect of a Pfizer booster dose, administered seven months after Pfizer vaccination. The sera were collected one month after the third dose. The booster dose enhanced neutralization titres against D614G and Delta by 39-fold and 49-fold (ED_{50} of 12,739 and 4,489, respectively, compared with the sampling time 5 months after the full two-dose vaccination). The sera from 1 month after the booster dose were also associated with a strong increase in the neutralization activity against Omicron (ED_{50} of 722) (Fig. 2b). At this time point, 100% of the tested sera displayed neutralizing activity (Extended Data Fig. 5).

Taken together, these results indicate that Omicron is poorly or not neutralized by sera of vaccinated individuals sampled five months after vaccination. The booster dose triggered detectable cross-neutralization activity against Omicron. However, even after the booster dose, the variant displayed a decrease in ED_{50} of 18-fold and 6-fold compared with D614G and Delta, respectively.

Sensitivity of Omicron to convalescent sera

We subsequently examined the neutralization ability of sera from convalescent individuals. We randomly selected 61 longitudinal samples from 40 donors in a cohort of infected individuals from Orléans. Individuals were diagnosed with SARS-CoV-2 infection by quantitative PCR with reverse transcription (RT-qPCR) (Extended Data Table 1b). We previously studied the potency of these sera against D614G, Alpha, Beta and Delta isolates^{13,14}. We analysed individuals sampled at a median of 6 and 12 months (M6 and M12) after the onset of symptoms. For the D614G and Delta variants, the neutralization titres were stable or slightly decreased over time (569 and 580 for D614G, and 235 and 143

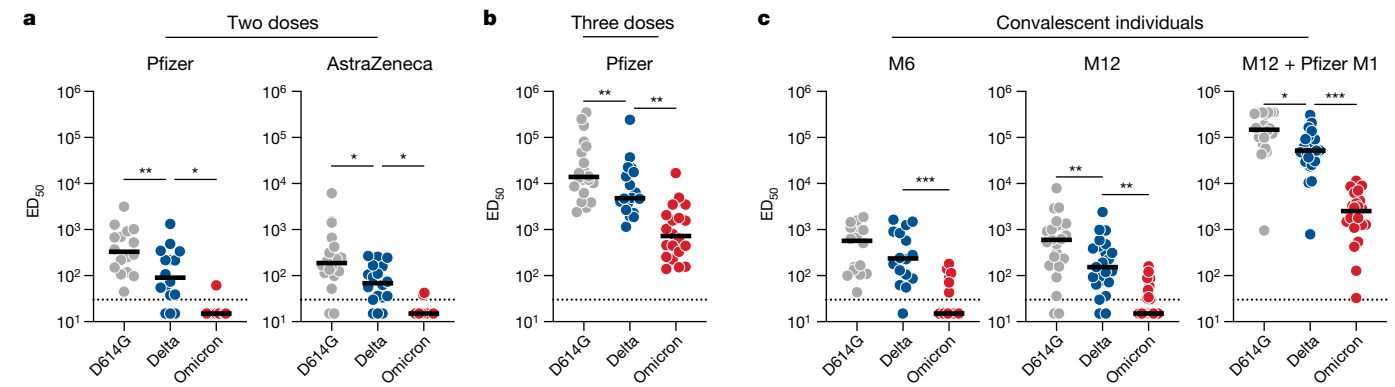


Fig. 2 | Sensitivity of the SARS-CoV-2 variants D614G, Delta and Omicron to sera from vaccinated, convalescent or infected-then-vaccinated individuals. Neutralization titres of the sera against the three indicated viral variants are expressed as ED_{50} . **a**, Neutralizing activity of sera from individuals who were vaccinated with the AstraZeneca ($n = 18$; left) and Pfizer ($n = 16$; right) vaccines, sampled at 5 months after the second dose. **b**, Neutralizing activity of sera from Pfizer-vaccinated recipients sampled one month (M1) after the third injection. $n = 20$. The dotted line indicates the limit of detection ($ED_{50} = 30$). **c**, The neutralizing activity of sera from convalescent individuals ($n = 16$),

sampled at 6 months after the onset of symptoms (left). Middle, the neutralizing activity of sera from convalescent individuals ($n = 23$), sampled at 12 months after the onset of symptoms. Right, the neutralizing activity of sera from individuals who were infected then vaccinated ($n = 22$), sampled one month after the first injection (right). In each panel, data are the mean values from two to three independent experiments. Two-sided Friedman tests with Dunn's multiple-comparison correction was performed to compare D614G and Omicron with the Delta variant; * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

for Delta, at M6 and M12, respectively)¹³ (Fig. 2c). The convalescent sera barely neutralized Omicron or did not inhibit Omicron at all at these time points. Only 36% and 39% of the samples displayed neutralizing activity against Omicron at M6 and M12, respectively, whereas the majority (91–94%) were active against Delta (Extended Data Fig. 5).

Twenty-two individuals were vaccinated at M12 with a Pfizer dose. Sera sampled 1 month after vaccination showed a considerable increase in neutralizing antibody titres against the D614G and Delta variants, reaching a median ED_{50} of 78,162 and 33,536, respectively (Fig. 2d). These sera also neutralized Omicron, with a median ED_{50} of 1,466 (Fig. 2d). Thus, as shown for other variants^{13,28,29}, a single dose of vaccine boosted cross-neutralizing antibody responses to Omicron in previously infected individuals. However, the neutralization titres were reduced by 53-fold and 23-fold compared with D614G and Delta, respectively.

Discussion

Omicron has opened a new chapter in the COVID-19 pandemic^{2,30}. The principal concerns about this variant include its high transmissibility, as underlined by its rapid spread in different countries, and the presence of more than 55 mutations spanning the whole viral genome. Omicron contains 32 mutations in its spike protein in the NTD, RBD and in vicinity of the furin cleavage site. Some mutations were already present in other VOCs and variants of interest, and have been extensively characterized^{30–32}. Owing to their position, they are expected to affect the binding of natural or therapeutic antibodies, to increase affinity to ACE2 and to enhance the fusogenic activity of the spike. Future work will help to determine how this association of mutations affects viral fitness in culture systems and their contribution to the high transmissibility of the variant.

Here we studied the cross-reactivity of clinical or preclinical monoclonal antibodies, as well as of 115 sera from vaccine recipients and long-term convalescent individuals against an infectious Omicron isolate. We report that, among nine monoclonal antibodies in clinical use or in development, six (bamlanivimab, etesevimab, casirivimab, imdevimab, tixagevimab and regdanvimab) were inactive against Omicron. Two other antibodies (cilgavimab and adintrevimab) displayed about a 20-fold increase in IC_{50} . Sotrovimab was less affected by Omicron's mutations, with an increase in IC_{50} by only threefold. We also show that Omicron was barely neutralized by sera from vaccinated

individuals who were sampled 5 months after the administration of two doses of the Pfizer or AstraZeneca vaccine. Sera from convalescent individuals at 6 or 12 months after infection barely neutralized or did not detectably neutralize Omicron.

The decrease in antibody efficacy helps to explain the high number of breakthrough infections and reinfection cases, and the spread of Omicron in both non-immune and immune individuals³³. There is currently no evidence of increased disease severity associated with Omicron compared with Delta, either among naive or immunized individuals. It is probable that, even if pre-existing SARS-CoV-2 antibodies may poorly prevent Omicron infection, anamnestic responses and cellular immunity will be operative to prevent severe forms of the disease³⁴.

We further report that a booster dose of Pfizer vaccine, as well as vaccination of previously infected individuals, strongly increased overall levels of anti-SARS-CoV-2 neutralizing antibodies, well above a threshold allowing inhibition of Omicron. Affinity maturation of antibodies is known to improve the efficacy of the humoral anti-SARS-CoV-2 response over time^{35,36}. This process helps to explain the efficacy of booster doses. However, sera with high antibody levels displayed a 6-fold to 23-fold reduction in neutralization efficacy against Omicron compared with the currently predominant Delta strain.

Potential limitations of our work include the low number of sera analysed from vaccine recipients and convalescent individuals, and the lack of characterization of cellular immunity, which is known to be more cross-reactive than the humoral response. Our results may therefore partly underestimate the residual protection offered by vaccines and previous infections against Omicron infection, in particular with regard to the severity of disease. We analysed only sera sampled one month after the booster dose, or after vaccination of infected individuals. Future work with more individuals and longer survey periods will help to characterize the duration of the humoral response against Omicron. We focused on immune responses elicited by Pfizer and AstraZeneca vaccination. It will be worth determining the potency of other vaccines against this variant.

We focused our analyses on one single viral isolate, corresponding to the archetype Omicron variant sequence (Pango BA.1 lineage). Two related lineages with additional mutations (BA.2 and BA.3) have recently emerged and are less widely spread. It will be worth comparing the behaviour of viral isolates from these more recent lineages with the main BA.1 Omicron strain.

Our results have important public health consequences regarding the use of therapeutic monoclonal antibodies and vaccines. Clinical indications of monoclonal antibodies include pre-exposure prophylaxis in individuals who are unable to mount an immune response, as well as the prevention of COVID-19 in infected individuals who are at high risk for progressing towards severe disease. Antibody-based treatment strategies need to be rapidly adapted to Omicron. Experiments in preclinical models or clinical trials are warranted to assess whether the drops in IC_{50} are translated into impaired clinical efficacy of the monoclonal antibodies that retain efficacy against Omicron. Most low-income countries display a low vaccination rate, a situation that probably facilitates SARS-CoV-2 spread and continuous evolution. A booster dose improves the quality and the level of the humoral immune response, and is associated with a strong protection against severe forms of the disease³⁷. An accelerated deployment of vaccines and boosters throughout the world is necessary to counteract viral spread. Our results also suggest that there is a need to update and complete the current pharmacopoeia, in particular with regard to vaccines and monoclonal antibodies.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-021-04389-z>.

- Classification of Omicron (B.1.1.529): SARS-CoV-2 Variant of Concern, [https://www.who.int/news/item/26-11-2021-classification-of-omicron-\(b.1.1.529\)-sars-cov-2-variant-of-concern](https://www.who.int/news/item/26-11-2021-classification-of-omicron-(b.1.1.529)-sars-cov-2-variant-of-concern) (WHO, 2021).
- Karim, S. S. A. & Karim, Q. A. Omicron SARS-CoV-2 variant: a new chapter in the COVID-19 pandemic. *Lancet* **398**, 2126–2128 (2021).
- Viana, R. et al. Rapid epidemic expansion of the SARS-CoV-2 Omicron variant in southern Africa. *Nature*, <https://doi.org/10.1038/s41586-022-04411-y> (2022).
- Taylor, P. C. et al. Neutralizing monoclonal antibodies for treatment of COVID-19. *Nat. Rev. Immunol.* **21**, 382–393 (2021).
- Grabowski, F., Kochańczyk, M. & Lipniacki, T. Omicron strain spreads with the doubling time of 3.2–3.6 days in South Africa province of Gauteng that achieved herd immunity to Delta variant. Preprint at <https://doi.org/10.1101/2021.12.08.21267494> (2021).
- Cele, S. et al. Omicron extensively but incompletely escapes Pfizer BNT162b2 neutralization. *Nature*, <https://doi.org/10.1038/s41586-021-04387-1> (2021).
- Cao, Y. et al. Omicron escapes the majority of existing SARS-CoV-2 neutralizing antibodies. *Nature*, <https://doi.org/10.1038/s41586-021-04385-3> (2021).
- Cameron, E. et al. Broadly neutralizing antibodies overcome SARS-CoV-2 Omicron antigenic shift. *Nature*, <https://doi.org/10.1038/s41586-021-04386-2> (2021).
- Dejnirattisai, W. et al. Reduced neutralisation of SARS-CoV-2 Omicron-B.1.1.529 variant by post-immunisation serum. *Lancet* **399**, 234–236 (2021).
- Aggarwal, A. et al. SARS-CoV-2 Omicron: evasion of potent humoral responses and resistance to clinical immunotherapeutics relative to viral variants of concern. Preprint at <https://doi.org/10.1101/2021.12.14.21267772> (2021).
- Redd, A. D. et al. Minimal cross-over between mutations associated with Omicron variant of SARS-CoV-2 and CD8⁺ T cell epitopes identified in COVID-19 convalescent individuals. Preprint at <https://doi.org/10.1101/2021.12.06.471446> (2021).
- Buchrieser, J. et al. Syncytia formation by SARS-CoV-2 infected cells. *EMBO J.* **39**, e106267 (2020).
- Planas, D. et al. Sensitivity of infectious SARS-CoV-2 B.1.1.7 and B.1.351 variants to neutralizing antibodies. *Nat. Med.* **27**, 917–924 (2021).
- Planas, D. et al. Reduced sensitivity of SARS-CoV-2 variant Delta to antibody neutralization. *Nature* **596**, 276–280 (2021).
- Rajah, M. M. et al. SARS-CoV-2 Alpha, Beta, and Delta variants display enhanced spike-mediated syncytia formation. *EMBO J.* **40**, e108944 (2021).
- Liu, L. et al. Potent neutralizing antibodies against multiple epitopes on SARS-CoV-2 spike. *Nature* **584**, 450–456 (2020).
- Barnes, C. O. et al. SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies. *Nature* **588**, 682–687 (2021).
- Wrapp, D. et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **367**, 1260–1263 (2020).
- Rappazzo, C. G. et al. Broad and potent activity against SARS-like viruses by an engineered human monoclonal antibody. *Science* **371**, 823–829 (2021).
- Zost, S. J. et al. Potently neutralizing and protective human antibodies against SARS-CoV-2. *Nature* **584**, 443–449 (2020).
- Kim, C. et al. A therapeutic neutralizing antibody targeting receptor binding domain of SARS-CoV-2 spike protein. *Nat. Commun.* **12**, 288 (2021).
- Shi, R. et al. A human neutralizing antibody targets the receptor-binding site of SARS-CoV-2. *Nature* **584**, 120–124 (2020).
- Jones, B. E. et al. The neutralizing antibody, LY-CoV555, protects against SARS-CoV-2 infection in nonhuman primates. *Sci. Transl. Med.* **13**, eabf1906 (2021).
- Hansen, J. et al. Studies in humanized mice and convalescent humans yield a SARS-CoV-2 antibody cocktail. *Science* **369**, 1010–1014 (2020).
- Pinto, D. et al. Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody. *Nature* **583**, 290–295 (2020).
- Liu, J. et al. BNT162b2-elicited neutralization of B.1.617 and other SARS-CoV-2 variants. *Nature* **596**, 273–275 (2021).
- Lucas, C. et al. Impact of circulating SARS-CoV-2 variants on mRNA vaccine-induced immunity. *Nature* **600**, 523–529 (2021).
- Gallais, F. et al. Evolution of antibody responses up to 13 months after SARS-CoV-2 infection and risk of reinfection. *EBioMedicine* **71**, 103561 (2021).
- Krammer, F. et al. Antibody responses in seropositive persons after a single dose of SARS-CoV-2 mRNA vaccine. *New Engl. J. Med.* **384**, 1372–1374 (2021).
- McCallum, M. et al. SARS-CoV-2 immune evasion by the B.1.427/B.1.429 variant of concern. *Science* **373**, 648–654 (2021).
- Plante, J. A. et al. The variant gambit: COVID-19's next move. *Cell Host Microbe* **29**, 508–515 (2021).
- Starr, T. N. et al. Prospective mapping of viral mutations that escape antibodies used to treat COVID-19. *Science* **371**, 850–854 (2021).
- Pulliam, J. R. C. et al. Increased risk of SARS-CoV-2 reinfection associated with emergence of the Omicron variant in South Africa. Preprint at <https://doi.org/10.1101/2021.11.11.21266068> (2021).
- Gagne, M. et al. Protection from SARS-CoV-2 Delta one year after mRNA-1273 vaccination in rhesus macaques is coincident with anamnestic antibody response in the lung. *Cell* **185**, 113–130 (2021).
- Gaebler, C. et al. Evolution of antibody immunity to SARS-CoV-2. *Nature* **591**, 639–644 (2021).
- Wang, Z. et al. Naturally enhanced neutralizing breadth against SARS-CoV-2 one year after infection. *Nature* **595**, 426–431 (2021).
- Barda, N. et al. Effectiveness of a third dose of the BNT162b2 mRNA COVID-19 vaccine for preventing severe outcomes in Israel: an observational study. *Lancet* **398**, 2093–2100 (2021).
- Tzou, P. L. et al. Coronavirus Antiviral Research Database (CoV-RDB): an online database designed to facilitate comparisons between candidate anti-coronavirus compounds. *Viruses* **12**, 1006 (2020).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2021

Methods

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment. Our research complies with all relevant ethical regulation.

Orléans Cohort of convalescent and vaccinated individuals

Since 27 August 2020, a prospective, monocentric, longitudinal, interventional cohort clinical study enrolling 170 individuals infected with SARS-CoV-2 with different disease severities, and 59 non-infected healthy control individuals is ongoing, aiming to describe the persistence of specific and neutralizing antibodies over a 24-month period. This study was approved by the ILE DE FRANCE IV ethical committee. At enrolment, written informed consent was collected and the participants completed a questionnaire that covered sociodemographic characteristics, virological findings (SARS-CoV-2 RT-PCR results, including date of testing), clinical data (date of symptoms onset, type of symptoms and hospitalization), and data related to anti-SARS-CoV-2 vaccination if ever (brand product, and date of first and second doses). The serological status of the participants was assessed every three months. Those who underwent anti-SARS-CoV-2 vaccination had regular blood sampling after first dose of vaccine (ClinicalTrials.gov: NCT04750720). The primary outcome was the presence of antibodies against the SARS-CoV-2 spike protein, as measured using the S-Flow assay. The secondary outcome was the presence of neutralizing antibodies as measured using the S-Fuse assay. For this study, we selected 61 convalescent and 54 vaccinated participants. Some individuals were sampled multiple times. We analysed a total of 115 sera. Study participants did not receive any compensation. The characteristics of each individual from the two cohorts are presented in Supplementary Table 2. The cohorts were constituted before the occurrence of the Omicron variant.

Phylogenetic analysis

To contextualize the isolated Omicron genome, all SARS-CoV-2 sequences available on the GISAID EpiCov database as of 6 December 2021 were retrieved. A subset of complete and high-coverage sequences, as indicated in GISAID, assigned to lineages B.1.529 or BA.1 and BA.2, were randomly subsampled. This subset was included in a global SARS-CoV-2 phylogeny reconstructed with augur and visualized with auspice as implemented in the Nextstrain pipeline (<https://github.com/nextstrain/ncov>, version from 6 May 2021)³⁹. Within Nextstrain, a random subsampling approach capping a maximum number of sequences per global region was used. The acknowledgment of contributing and originating laboratories for all sequences used in the analysis is provided in Supplementary Table 1.

3D representation of mutations on the surface of spike of B.1.617.2 and other variants

The panels in Extended Data Fig. 3 were prepared using The PyMOL Molecular Graphics System, v.2.1 (Schrödinger). The atomic model used (Protein Data Bank: 6XR8) has previously been described⁴⁰.

S-Fuse neutralization assay

U2OS-ACE2 GFP1-10 or GFP11 cells, also termed S-Fuse cells, become GFP⁺ when they are productively infected by SARS-CoV-2 (refs. ^{12,13}). Cells tested negative for mycoplasma. Cells were mixed (ratio 1:1) and plated at 8×10^3 per well in a μ Clear 96-well plate (Greiner Bio-One). The indicated SARS-CoV-2 strains were incubated with serially diluted monoclonal antibodies or sera for 15 min at room temperature and added to S-Fuse cells. The sera were heat-inactivated for 30 min at 56 °C before use. Then, 18 h later, cells were fixed with 2% PFA, washed and stained with Hoechst (dilution of 1:1,000, Invitrogen). Images were acquired using an Opera Phenix high-content confocal microscope (PerkinElmer). The GFP area and the number of nuclei were quantified

using the Harmony software (PerkinElmer). The percentage of neutralization was calculated using the number of syncytia as value with the following formula: $100 \times (1 - (\text{value with serum} - \text{value in 'non-infected'}) / (\text{value in 'no serum'} - \text{value in 'non-infected'}))$. Neutralizing activity of each serum was expressed as the half maximal effective dilution (ED₅₀). ED₅₀ values (in $\mu\text{g ml}^{-1}$ for monoclonal antibodies and in dilution values for sera) were calculated with a reconstructed curve using the percentage of the neutralization at the different concentrations.

Characteristics of the patient infected with Omicron

The 32-year-old woman was unvaccinated and developed moderate symptoms on 22 November 2021, 11 days after returning to Belgium from Egypt via Turkey (stop-over to switch flights, without having left the airport). She did not display any risk factor for severe COVID-19 and rapidly recovered. She transmitted the virus to her husband but not to their children. She provided informed written consent to use the swab for future studies. The nasopharyngeal swab tested positive for SARS-CoV-2 on this date. The leftover material of the sample was used in this study after performing routine diagnostics, within the context of the mandate that was provided to UZ/KU Leuven as National Reference Center (NRC) of respiratory pathogens, as described in detail in the Belgian Royal Decree of 09/02/2011.

Virus strains

The reference D614G strain (hCoV-19/France/GE1973/2020) was supplied by the National Reference Centre for Respiratory Viruses hosted by Institut Pasteur and headed by S. van der Werf. This viral strain was supplied through the European Virus Archive goes Global (Evag) platform, a project that has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement no 653316. The variant strains were isolated from nasal swabs using Vero E6 cells and amplified by one or two passages. Delta was isolated from a nasopharyngeal swab of a hospitalized patient returning from India¹⁴. The swab was provided and sequenced by the laboratory of Virology of Hôpital Européen Georges Pompidou (Assistance Publique, Hôpitaux de Paris). The Omicron-positive sample was cultured on Vero E6 cells as previously described⁴¹. Viral growth was confirmed by RT-qPCR three days after infection. At day 6 after infection, a cytopathic effect was detected and a full-length sequencing of the virus was performed. The Omicron strain was supplied and sequenced by the NRC UZ/KU Leuven. Both patients provided informed consent for the use of the biological materials. Titration of viral stocks was performed on Vero E6 cells, with a limiting dilution technique enabling the calculation of the median tissue culture infectious dose, or on S-Fuse cells. Viruses were sequenced directly on nasal swabs, and after one or two passages on Vero cells. The sequences were deposited on GISAID immediately after their generation (D614G: EPI_ISL_414631; Delta ID: EPI_ISL_2029113; Omicron ID: EPI_ISL_6794907).

Flow cytometry

Vero cells were infected with the indicated viral strains at a multiplicity of infection of 0.01. Two days after, cells were detached using PBS-0.1% EDTA and transferred into U-bottom 96-well plates (50,000 cells per well). Cells were then incubated for 30 min at 4 °C with the indicated monoclonal antibodies (1 or 0.1 $\mu\text{g ml}^{-1}$) in MACS buffer (PBS, 5 g l⁻¹ BSA, 2 mM EDTA). Cells were washed with PBS and stained using anti-IgG AF647 (1:600 dilution in MACS buffer) (Thermo Fisher Scientific). Stainings were also performed on control non-infected cells. Cells were then fixed in 4% PFA for 30 min at room temperature. Data were acquired on the Attune Nxt instrument using Attune Nxt Software v.3.2.1 (Life Technologies) and analysed using FlowJo v.10.7.1 (Becton Dickinson).

Antibodies

Four clinically available antibodies (bamnanimab, casirivimab, etesevimab and imdevimab) were provided by CHR Orleans.

The other human SARS-CoV-2 anti-RBD neutralizing antibodies (ADG20 or adintrevimab, AZD1061 (COV2-2130) or cilgavimab, AZD8895 (COV2-2196) or tixagevimab, CT-P59 or regdanvimab, LY-CoV016 (CB6) or etesevimab, LY-CoV555 or bamlanivimab, REGN10933 or casirivimab, REGN10987 or imdevimab, and VIR-7831 (S309) or sotrovimab^{19–25} were produced as follows. DNA fragments encoding their IgH and IgL variable domains were synthesized (Life Technologies, Thermo Fisher Scientific). Purified digested DNA fragments were cloned into human Igy1- and Igκ-/Igλ-expressing vectors⁴² and recombinant IgG1 antibodies were produced by transient co-transfection of Freestyle 293-F suspension cells (Thermo Fisher Scientific) using the PEI-precipitation method as previously described⁴³. IgG1 antibodies were purified by batch/gravity-flow affinity chromatography using protein G Sepharose 4 Fast Flow Beads (Cytiva) according to the manufacturer's instructions, dialysed against PBS using Slide-A-Lyzer dialysis cassettes (Thermo Fisher Scientific), quantified using the NanoDrop 2000 instrument (Thermo Fisher Scientific), and checked for purity and quality on a silver-stained SDS-PAGE gel (3–8% Tris-Acetate Novex, Thermo Fisher Scientific). The pan-coronavirus anti-S2 non-neutralizing antibody Ab-10 was previously described^{13,14}.

Statistical analysis

Flow cytometry data were analysed using FlowJo v.10 (TriStar). Calculations were performed using Excel 365 (Microsoft). Figures were generated using Prism 9 (GraphPad Software). Statistical analysis was conducted using GraphPad Prism 9. Statistical significance between different groups was calculated using the tests indicated in each figure legend.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

All data supporting the findings of this study are available within the Article or from the corresponding authors on request. Viral sequences are available on request and were deposited at GISAID (<https://www.gisaid.org/>) under the following numbers: D614G (EPI_ISL_414631), Delta ID (EPI_ISL_2029113) and Omicron ID (EPI_ISL_6794907). Source data are provided with this paper.

39. Hadfield, J. et al. Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics* **34**, 4121–4123 (2018).
40. Cai, Y. et al. Distinct conformational states of SARS-CoV-2 spike protein. *Science* **369**, 1586–1592 (2020).

41. Van Cleemput, J. et al. Organ-specific genome diversity of replication-competent SARS-CoV-2. *Nat. Commun.* **12**, 6612 (2021).
42. Tiller, T. et al. Efficient generation of monoclonal antibodies from single human B cells by single cell RT-PCR and expression vector cloning. *J. Immunol. Methods* **329**, 112–124 (2008).
43. Lorin, V. & Mouquet, H. Efficient generation of human IgA monoclonal antibodies. *J. Immunol. Methods* **422**, 102–110 (2015).
44. Walls, A. C. et al. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell* **181**, 281–292 (2020).

Acknowledgements We thank the staff at the European Health Emergency Preparedness and Response Authority (HERA) for their support; S. Cole for his help in initiating the collaboration between Institut Pasteur and KU Leuven through the HERA network; A. Baidaliuk and F. Gambaro for their help with the sequencing data analysis; the patients who participated to this study; the members of the Virus and Immunity Unit and other teams for discussions and help; N. Aulner and the staff at the UtechS Photonic Bioluminescence (UPBI) core facility (Institut Pasteur)—a member of the France Bioluminescence network—for image acquisition and analysis; the members of the KU Leuven University authorities, and J. Arnout, B. Lambrecht, C. Van Geet and L. Sels for their support; L. Belec, N. Robillard and M. Saliba for their help with sequencing; and F. Peira, V. Legros and L. Courtellemont for their help with the cohorts. The Opera system was co-funded by Institut Pasteur and the Région Ile de France (DIM1Health). Work in the O.S. laboratory is funded by Institut Pasteur, Urgence COVID-19 Fundraising Campaign of Institut Pasteur, Fondation pour la Recherche Médicale (FRM), ANRS, the Vaccine Research Institute (ANR-10-LABX-77), Labex IBEID (ANR-10-LABX-62-IBEID), ANR/FRM Flash Covid PROTEO-SARS-CoV-2 and IDISCOVER. Work in the UPBI is funded by grant ANR-10-INSB-04-01 and Région Ile-de-France program DIM1-Health. D.P. is supported by the Vaccine Research Institute. The H.M. laboratory is funded by the Institut Pasteur, the Milieu Intérieur Program (ANR-10-LABX-69-01), the INSERM, REACTing, EU (RECOVER) and Fondation de France (00106077) grants. The E.S.-L. laboratory is funded by Institut Pasteur, the INCEPTION program (Investissements d'Avenir grant ANR-16-CONV-0005) and the French Government's Investissement d'Avenir programme, Laboratoire d'Excellence 'Integrative Biology of Emerging Infectious Diseases' (grant no. ANR-10-LABX-62-IBEID). G.B. acknowledges support from the Internal Funds KU Leuven under grant agreement C14/18/094, and the Research Foundation—Flanders (Fonds voor Wetenschappelijk Onderzoek—Vlaanderen, GOE1420N, G098321N). P.M. acknowledges support from a COVID-19 research grant of 'Fonds Wetenschappelijk Onderzoek'/Research Foundation—Flanders (grant no. G0H4420N). S.D. is supported by the Fonds National de la Recherche Scientifique (FNRS, Belgium) and also acknowledges support from the Research Foundation—Flanders (Fonds voor Wetenschappelijk Onderzoek—Vlaanderen, G098321N) and from the European Union Horizon 2020 project MOOD (grant no. 874850). The funders of this study had no role in study design, data collection, analysis and interpretation, or writing of the article.

Author contributions Experimental strategy design, experiments: D.P., N.S., F.G.-B., C.P., J.B., W.-H.B., F.P., I.S., F.A.R., E.S.-L., T.B., H.M. and O.S. Vital materials: P.M., C.P., F.L., H.P., D.V., J.P., J. Rodary, G.B., S.D., J. Raymenants, S.G., C.G., B.V., T.W.-B., J.M.-C., L.C., A.S., L.H., T.P., H.M. and E.A. Phylogenetic analysis: G.B., E.S.-L., F.L. and S.D. Manuscript writing: D.P., F.A.R., E.S.-L., T.B., H.M., E.A. and O.S. Manuscript editing: D.P., N.S., P.M., G.B., L.C., F.A.R., E.S.-L., T.B., H.M., E.A. and O.S.

Competing interests C.P., H.M., O.S., T.B. and F.A.R. have a pending patent application for an anti-RBD monoclonal antibody that was not used in this study (PCT/FR2021/070522).

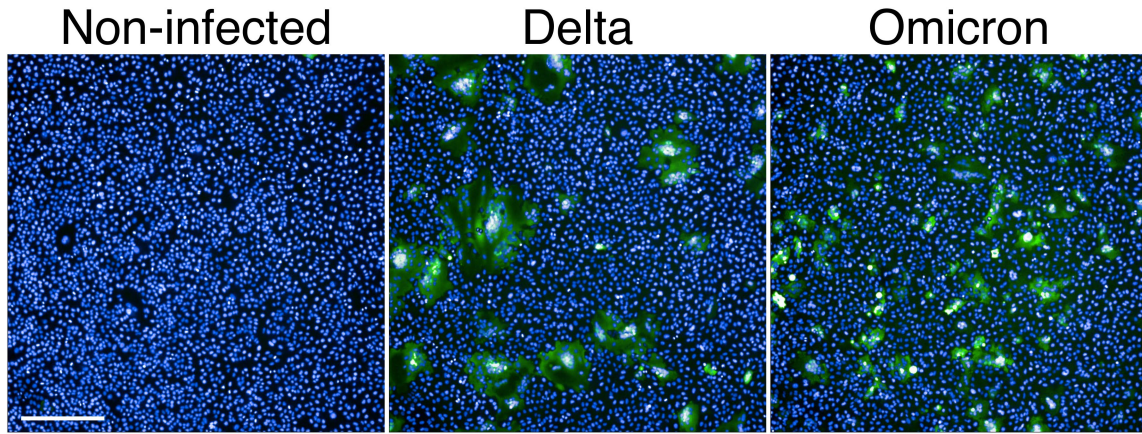
Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41586-021-04389-z>.

Correspondence and requests for materials should be addressed to Timothée Bruel, Hugo Mouquet, Emmanuel André or Olivier Schwartz.

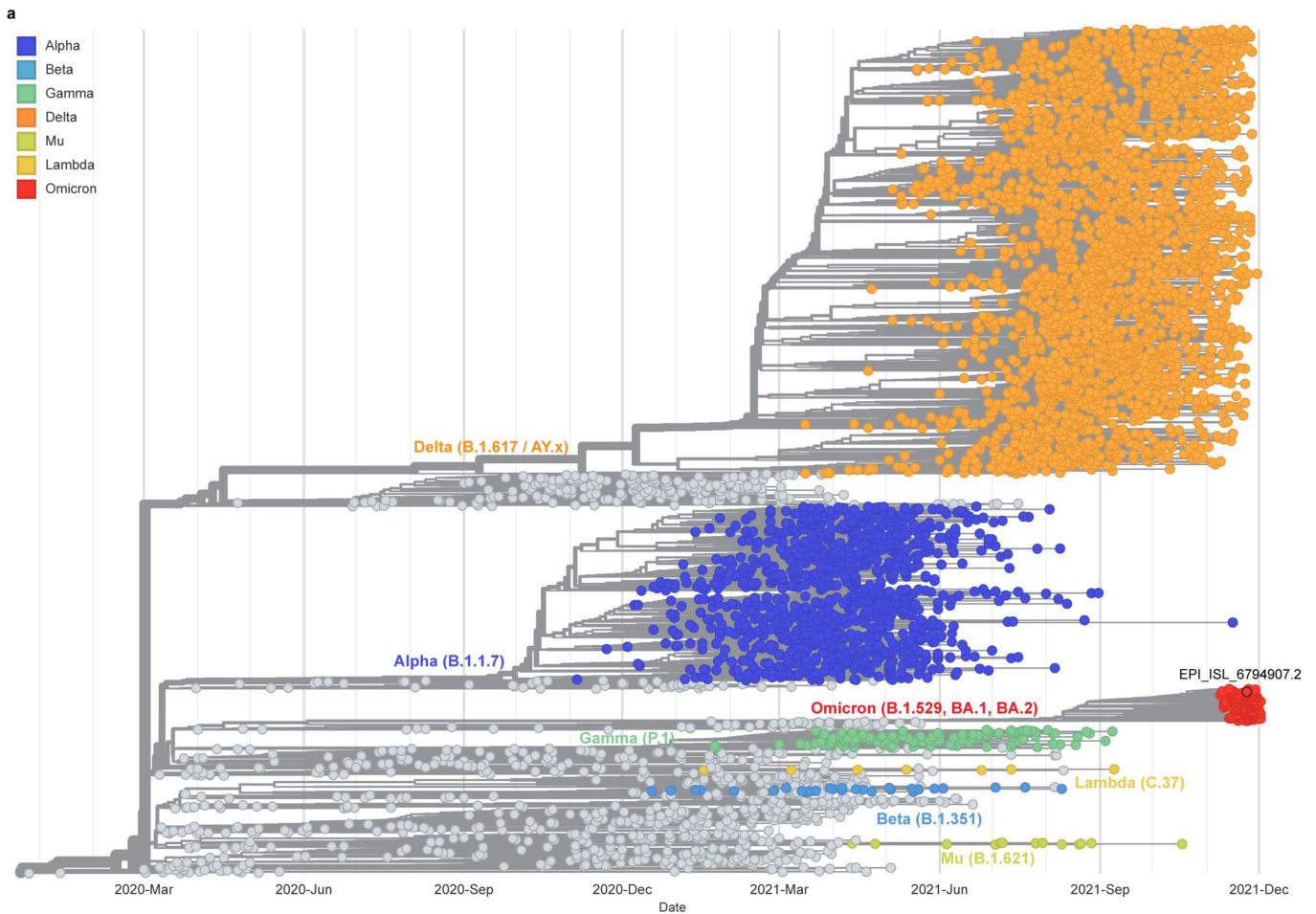
Peer review information Nature thanks the anonymous reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at <http://www.nature.com/reprints>.



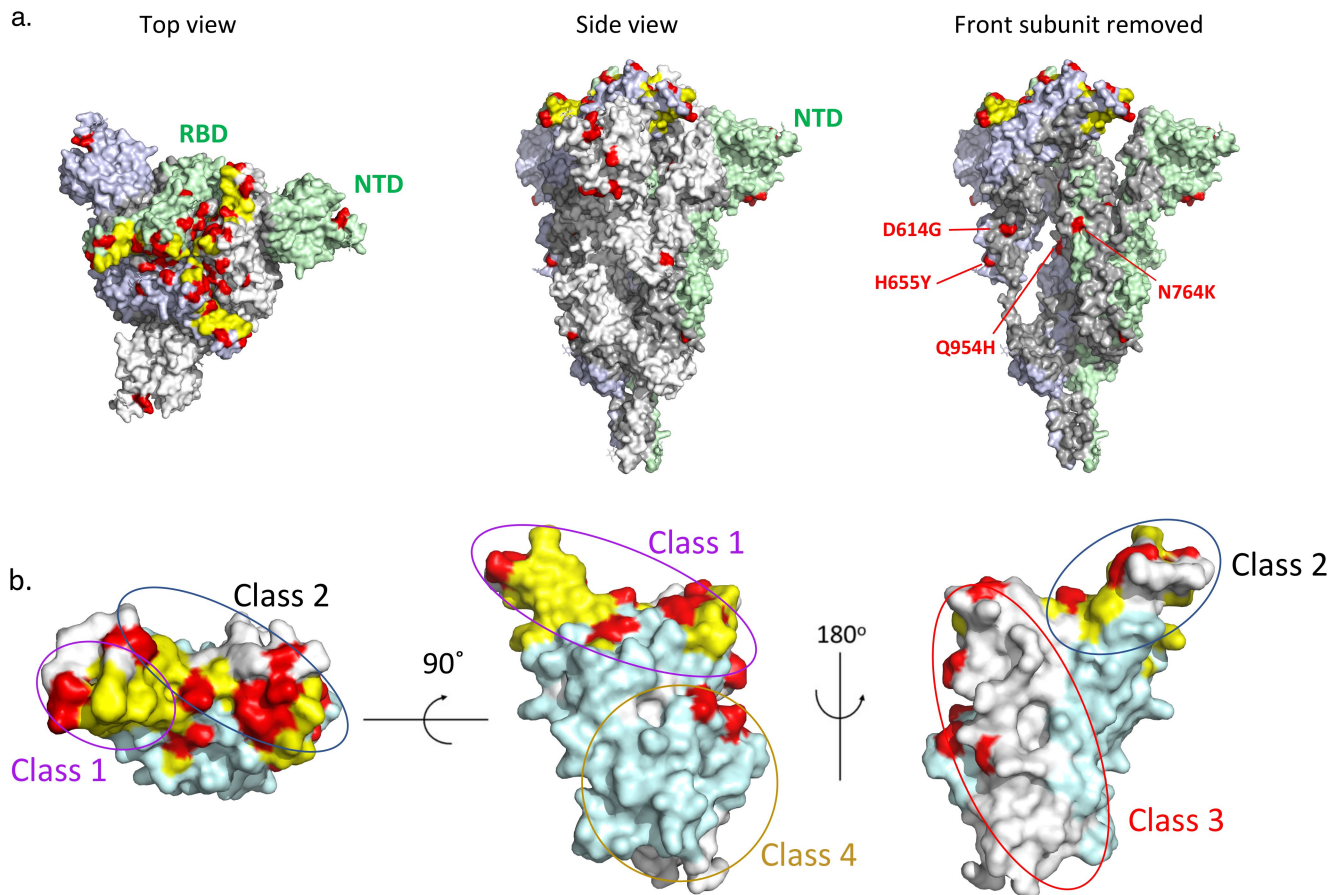
Extended Data Fig. 1 | SARS-CoV-2 variants induce syncytia in S-Fuse cells. S-Fuse cells were exposed to the indicated SARS-CoV-2 strains, at a multiplicity of infection (MOI) of 10^{-3} . The cells become GFP+ when they fuse together.

After 20 h, infected cells were stained with Hoechst to visualize nuclei. Syncytia (green) and nuclei (blue) are shown. Representative images from three independent experiments are shown. Scale bar, 500 μ m.



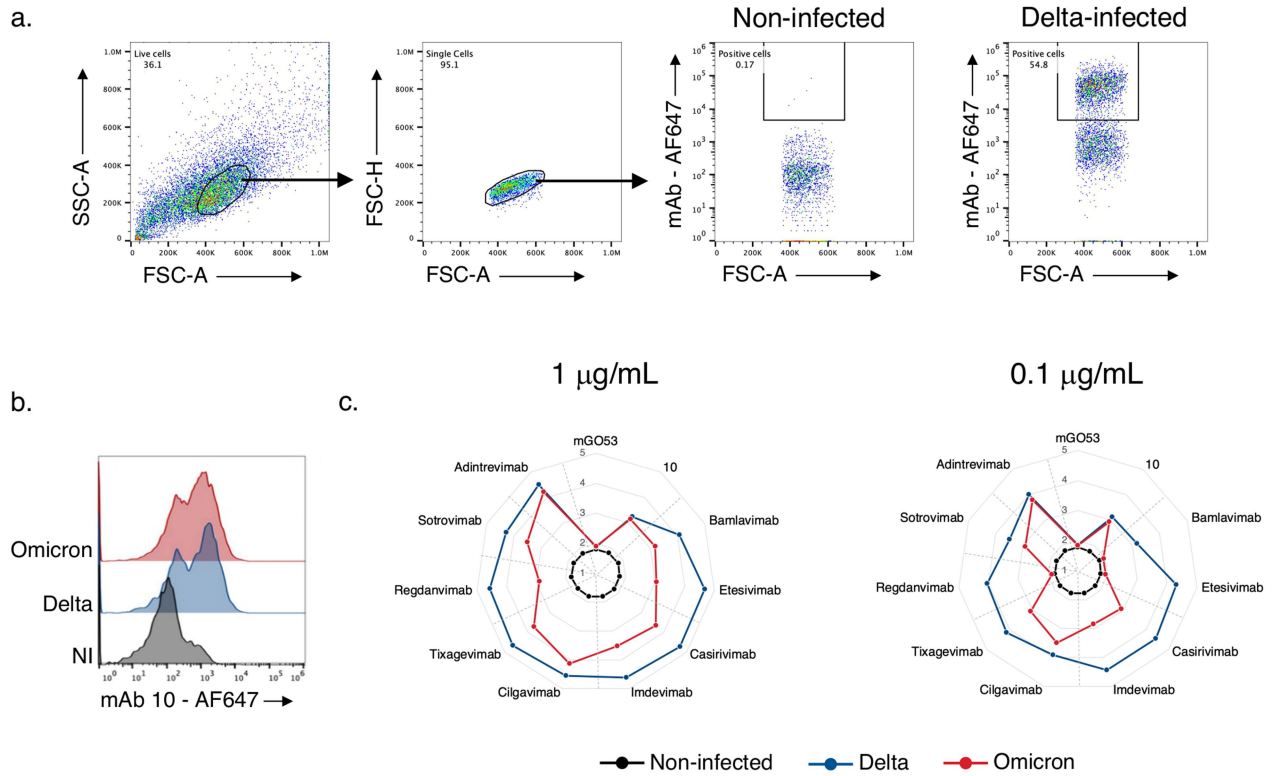
Extended Data Fig. 2 | Global phylogeny of SARS-CoV-2 highlighting the Omicron lineage. Time calibrated global SARS-CoV-2 phylogeny available from the Nextstrain platform (<https://nextstrain.org/ncov/gisaid/global>)³⁹.

The position of the isolated Omicron variant is highlighted, and the variants of concern (VOCs) (Alpha, Beta, Gamma, Delta and Omicron) and variants of interest (VOIs) (Lambda, Mu) are coloured as indicated.



Extended Data Fig. 3 | Mapping of the mutations present in Omicron to the spike's surface. **a.** The spike shown in top (left panel) and in side view (middle and right panels). The spike trimer is shown in surface representation with the three protomers coloured in light grey, light blue and light green. N-terminal and the receptor-binding (NTD and RBD) domains are labelled for the protomer in green only. The represented spike (PDB: 6XR8) is in the closed conformation, i.e., with all three RBDs in the "Down" conformation⁴⁴. The RBD surface of interaction with hACE2 (which is partially occluded in a closed spike) is coloured in yellow. The amino acid differences in the spike of the Omicron variant with respect to the initial Wuhan sequence are marked in red. In the

right panel, the front subunit was removed to show changes in S2 and in the C-terminal segment of S1 (labelled) that map to the trimer interface, which could impact the stability of the spike trimer. **b.** The RBD view down the hACE2 binding surface (left panel) and in two other orthogonal orientations (middle and right panel), as indicated. The hACE2 binding surface is coloured in yellow and the residues altered in Omicron are in red. The RBD surfaces that are buried and exposed in a closed spike are coloured in light cyan and white, respectively. The ovals outline the location of the epitopes of neutralizing antibodies of the various classes that have been described¹⁷.

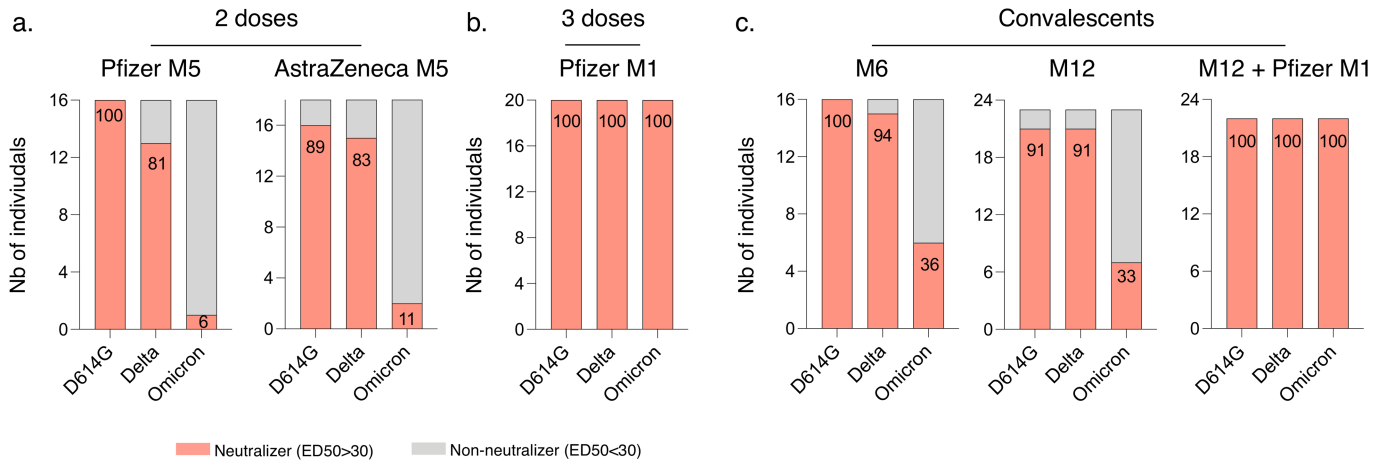


d.

		Neutralization (IC50 ng/ml)		Binding (Fold-decrease Omicron/Delta)			
		Delta	Omicron	1 µg/mL		0.1 µg/mL	
				Delta	Omicron	Delta	Omicron
Lilly	Bamlavimab	>9000	>9000	1	8	1	17
	Etesivimab	15.4	>9000	1	42	1	242
	Bamla/Ete	20.5	>9000	1	n.d	1	n.d
Regeneron	Casirivimab	3.1	>9000	1	12	1	32
	Imdevimab	6.5	>9000	1	12	1	39
	Casirivimab/Imdevimab	7.6	>9000	1	n.d	1	n.d
AstraZeneca	Cilgavimab	50.3	1213	1	3	1	3
	Tixagevimab	18.3	8305	1	9	1	11
	Cilgavimab/Tixagevimab	13.4	773	1	n.d	1	n.d
Celltrion	Regdanvimab	86.6	>9000	1	47	1	145
GlaxoSmithKline	Sotrovimab	325	917	1	6	1	4
Adagio	Adintrevimab	19.6	403	1	2	1	2

Extended Data Fig. 4 | Binding of anti-SARS-CoV-2 monoclonal antibodies to Vero cells infected with Delta and Omicron variants. Vero cells were infected with the indicated variants at an MOI of 0.01. After 48 h, cells were stained with 1 or 0.1 µg ml⁻¹ of the indicated anti-SARS-CoV-2 monoclonal antibodies (Bamlavimab, Etesivimab, Casirivimab, Imdevimab, Adintrevimab, Cilgavimab, Tixagevimab, Regdanvimab, Sotrovimab) and analysed by flow-cytometry. **a.** Gating strategy and example of gates on negative (non-infected) or positive (Delta-infected) samples. **b.** The anti-S2 pan-coronavirus mAb 10 was used to measure the percentage of infected cells. Histograms show binding of mAb 10 to Vero cells infected with the indicated

variants. **c.** Radar charts represent for each antibody the logarithm of the median fluorescent intensity (MFI) of the staining. Data are representative of two or three independent experiments. **d.** Inhibitory Concentrations 50% (IC50) of mAbs against Delta and Omicron variants. The IC50 of the indicated mAbs and some of their combinations were calculated from the neutralization curves displayed in Fig. 1b. Results are in ng ml⁻¹. Colour code: Grey: inactive mAbs. Green: mAbs displaying a neutralizing activity. The binding activity was measured by flow cytometry on Vero cells infected with the indicated variants. Results are presented as the fold-decrease of binding to Omicron-infected cells relative to Delta-infected cells.



Extended Data Fig. 5 | Fraction of neutralizers in the cohorts of vaccinated or convalescent individuals. Individuals with an ED50 of neutralization above 30 were categorized as neutralizers and are indicated in pink. Non-neutralizers are in grey. The numbers indicate the percentage of neutralizers. **a.** Fraction of neutralizers in sera from Pfizer (n = 16) (left panel) and AstraZeneca (n = 18) (right panel) vaccinated recipients sampled 5 months after the second dose (results related to Fig. 2a). **b.** Fraction of neutralizers in sera from Pfizer

vaccinated recipients sampled one month after the 3rd injection (n = 20; (results related to Fig. 2b). **c.** Fraction of neutralizers in sera from convalescent individuals, sampled at 6 months post onset of symptoms (M6) (n = 16) (right panel), at 12 months (M12) (n = 23) (middle panel) and one month after the 1st injection (n = 22) (left panel; results related to Fig. 2c). In each panel, data are mean from 2 to 3 independent experiments.

Extended Data Table 1 | Characteristics of the two cohorts of vaccinated and convalescent individuals

a. Vaccinees

	AstraZeneca M5 (2 doses)	Pfizer M5 (2 doses)	Pfizer M1 (3 doses)
	n=18	n=16	n=20
Sex			
Female	12	6	9
Male	6	10	11
Age (median; range)	60 (55;78)	53 (33;74)	53 (33;74)
Immune deficiency	0	0	0
Previous COVID-19	0	0	0
Anti-N	0	0	0
1st shot	Feb 5 – April 7, 2021	Jan 6- Feb 4,2021	Jan 6- April12,2021
2nd shot	May3-19, 2021	Jan 28 - March 3, 2021	Jan 28 – May 8, 2021
3rd shot	-	-	July 1 –Nov 25, 2021
Sampling days post-2 nd dose; M5 (median; range)	150 (110-178)	161 (138-176)	-
Sampling days post-3 rd dose; M1 (median; range)	-	-	33 (8-61)

b. Convalescents

	6M POS n=16	12M POS n=23	12M POS then vaccinated n=22
Sex			
Female	10	11	11
Male	6	12	11
Age (Median; range)	56 (32;77)	52 (23;82)	52 (23 ;82)
Severity			
Severe	8	13	9
Mild-Moderate	7	8	4
Asymptomatic	1	2	1
PCR	15	21	20
Anti-S (S-Flow)	16	23	22
Sampling days POS (median; range)	184 (129;195)	368 (344-454)	-
Days between the vaccination and onset of symptoms (median; range)	-	-	394 (353-444)
Sampling days post-vaccination (median; range)	-	-	32 (21-48)

2.4. Discussion

2.4.1. What are the impacts of Spike mutations?

Mutations in the spike can be selected because they increase the affinity for ACE2, enable escape from neutralizing antibodies or cellular immunity, increase spike opening or stability or increase infectivity. The literature on the effects of different mutations in the spike is extensive; this part will put our work (**Chapter 2**, **Appendix 1** and **Appendix 2**) in perspective.

2.4.1.1. Increased infectivity

As discussed in the introduction of this chapter, D614G is a mutation that enhances the stability^{320,321} and the fusogenic properties³²² of the spike. P681H/R which is present in Delta and Omicron variants increases spike fusogenicity and cleavage^{328,332}. Δ 69-70 which is present in Alpha and some Omicron sublineages increases spike incorporation into virions and thus infectivity³³³.

Mutations that increase the affinity for ACE2, such as N439K, Y453F, S486P and N501Y, are also frequent (**Appendix 1**)^{62,334}. Some of the antibody escape mutations reduce the affinity for ACE2 and are thus only selected in presence of mutations that increase ACE2 affinity. All variants have increased affinity for ACE2 compared to the D614G ancestral strain, but affinity for ACE2 is not necessarily optimal in all variants³³⁵. Interestingly, BA.1 and BA.2 had the R493Q mutation, which diminishes the affinity for ACE2; position 493 was then reverted to its ancestral amino-acid in subsequent sublineages which had more escape mutations in the RBD that reduced ACE2 affinity^{336,337}, and thus had to be “compensated for”.

2.4.1.2. Alteration in the entry routes and impact

As the spike evolved, so have its pathways to enter cells.

The E484D substitution increases the affinity of the spike for an alternative receptor TMEM106b. This increases the capacity of SARS-CoV-2 variants that carry this mutation to infect ACE2 negative cells²¹⁴.

Omicron lineages enter cells independently of serine transmembrane proteases and rely more on metalloproteases and endosomal entry. One study shows BA.1 is more sensitive to inhibition by IFITMs and Guanylate-binding proteins, which are restriction factors, and that this sensitivity is due to the spike³³⁸. Others report Omicron lineages are more resistant than the ancestral strain to inhibition by interferons and that this resistance to interferons is mediated

by the spike³³⁹⁻³⁴¹. The authors suggest this enables evasion from constitutive and interferon-induced antiviral factors, among which IFITMs³⁴¹.

2.4.1.3. Humoral immunity escape

Many of the mutations that arise can be related to epitopes targeted by neutralizing antibodies (**Figure 12**), as we and others have shown (**Appendix 1**)^{334,342}. Many of the mutations predicted to affect recognition by different class of antibodies appeared during the pandemic³⁴², often repeatedly and independently. Mutations that increase infectivity were likely necessary for the virus to acquire mutations that enable antibody escape even if they are deleterious for the infectivity, as these variants would not otherwise be selected.

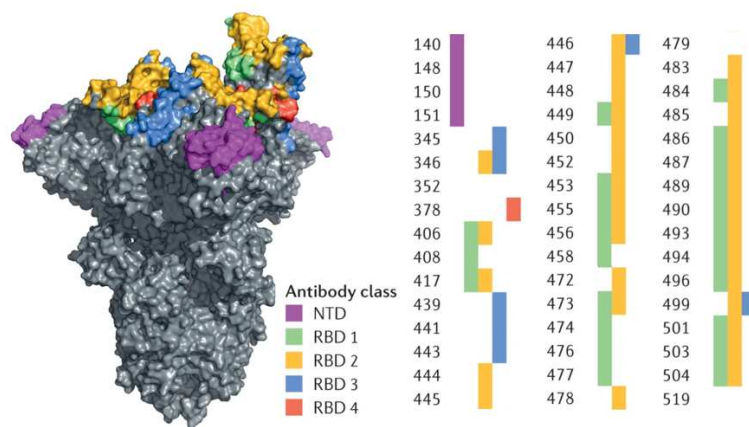


Figure 12: Main residues present in the epitopes targeted by anti-NTD, and class 1, 2, 3 and 4 anti-RBD antibodies. Mutations in those positions have been shown to reduce neutralization by monoclonal and/or polyclonal antibodies. From Harvey *et al.* 2021³³⁴.

2.4.1.4. Mutations present in AY.4.2

Our results on the slight reduction in neutralization of AY.4.2 were further confirmed by another study that was published around the same time³⁴³. The T95I mutation was also present in the BA.1 lineage and has reappeared in some of the most recent variants such as JN.1 and XBC. One preprint suggests that this mutation, in combination with G142D which is present in the lineages where T95I is present, affects the conformation of the NTD site targeted by neutralizing antibodies³⁴⁴. The Y145H mutation is part of the NTD epitope targeted by neutralizing antibodies, and mutations in this region arose in many variants. A study following our work showed that the A222V mutation slightly increases affinity for ACE2, by conferring more flexibility to the opened spike³⁴⁵.

One of the results that we have not further investigated or explained is the early plateau observed in Imdevimab neutralization of AY.4.2 compared to Delta. Imdevimab is an anti-RBD class 3 antibody. It sterically interferes with ACE2 binding. None of the mutations in AY.4.2

are predicted to affect this region. It is possible that the increase in RBD flexibility conferred by the A222V mutation enables partial escape from Imdevimab.

2.4.2. Links between infection, immunity and viral evolution

2.4.2.1. Natural, vaccine and hybrid immunity

Individuals can acquire natural or vaccine immunity to a pathogen, depending on whether they survive infection or are vaccinated against this pathogen.

Natural infection by SARS-CoV-2 confers protection thanks to antibodies and memory CD4+ T cells, CD8+ T cells and B cells. Furthermore, natural infection confers an antibody response in the nasal mucosa³⁴⁶. This immunity lasts at least 6-8 months³⁴⁷, but this was not tested on longer timescales; fast variant turnover enable reinfection and immunity is restimulated by subsequent infection or vaccination. Available SARS-CoV-2 vaccines induce both an antibody and a T-cell response^{348–350}.

For SARS-CoV-2, one can also acquire hybrid immunity, either by being infected before vaccination, or being infected after vaccination. Vaccination post-infection boosts the memory B cells, T cell immunity as well as the antibody response³⁵¹. It also increases the breadth of the antibody response, conferring better protection against the first variants, notably Beta, than vaccination alone^{352,353}.

In accordance with these findings, we have shown that serum from vaccinated convalescent individuals had about 10-fold higher neutralizing titers against D614G and Delta, and 2-fold higher against BA.1, compared to participants vaccinated thrice. Titers from unvaccinated convalescent or twice-vaccinated individuals were much lower. Our results are biased by the fact that the serum we tested were sampled at least 5 months after infection/vaccination for the convalescent or twice-vaccinated while they were collected 1 month post last dose for the thrice-vaccinated or vaccinated convalescent participants, but other studies report similar effects.

2.4.2.2. Immune imprinting

The emergence of Omicron variants that have over 30 mutations in their spike, against which previous vaccination and infection are poorly protective as demonstrated in our results, has brought the question of immune imprinting to the front scene. Individuals that have been exposed to or vaccinated against the Wuhan strain develop an immune response directed to the antigens present in this ancestral strain. Upon exposure to an Omicron strain, pre-existing B cells are boosted, and cross-reactive antibodies are produced, but no *de novo* response

occurs³⁵⁴⁻³⁵⁶. Antibodies that recognize conserved epitopes between the infecting/vaccinating strain and the ancestral strain are produced^{355,357}.

Some studies showed that individuals previously infected, triple vaccinated and subsequently infected with BA.1 generated less Omicron specific antibodies than triple vaccinated naïve individuals^{358,359}, suggesting the immune imprinting by the ancestral Wuhan strain is deleterious to mount specific response against Omicron. However, there are also several studies that show the cross-reactive antibodies generated in case of repeated exposure to and/or vaccination by Omicron strains expands the cross-neutralizing antibody response, which confers protection against emerging variants and other sarbecoviruses³⁶⁰⁻³⁶². The bivalent vaccines are now broadly used.

Interestingly, many of the newest subvariants of Omicron have distinct evolutionary courses but convergent mutations on the RBD. While antibody escape converging mutations had been observed and described previously, in the NTD (140-156, 240-260), or in the RBD (417, 444-446, 452, 484)³³⁴ it is their speed of appearance and number in the Omicron subvariants that is unprecedented³³⁵. There is evidence that because of immune imprinting, BA.2 and BA.5 breakthrough infections reduce the diversity of sites targeted by neutralizing antibodies, as antibodies target sites that are shared with the ancestral strain³³⁵. This in turns focuses the immune response and favors emerging variants that escape this recognition³³⁵. Convergent RBD mutations are occurring in Omicron subvariants: mutations at position 346, 356, 444, 445, 446, 450, 452, 460, 486, 490, 493 and 494 appeared independently in at least 5 subvariants. They are mutations that were predicted to confer antibody escape by deep-mutational scanning^{342,363,364}. Their appearance in the newest Omicron subvariants allows escape from recognition in individuals that had BA.2 and BA.5 breakthrough infections³³⁵. Immune imprinting thus likely impacts SARS-CoV-2 evolution.

2.4.3. What about mutations in the rest of the genome?

2.4.3.1. Innate immunity modulations

During the infection by the earlier VOCs, proinflammatory genes are less expressed than during infection by the ancestral strain³⁶⁵. On the other hand, BA.1 and BA.2 induce more potent interferon response than the ancestral strain³⁶⁶⁻³⁶⁸. This could be compensated by the fact that the Omicron are less sensitive to IFN as discussed previously^{340,341}. BA.4, BA.5 and subsequent Omicron variants reverted to less immunogenic phenotypes³⁶⁸. Several genes are thought to participate in this innate immune evasion and are differentially expressed or mutated in different variants.

ORF9b

ORF9b suppresses innate immunity by interacting with TOM70 and preventing MAVS activation^{164,166}. In Alpha, Beta, Gamma, Delta and Omicron strains, ORF9b is overexpressed. In Alpha and Delta, there is a deletion 3 amino acids before the start of N. For Omicron there is a substitution in the same position. This deletion or mutation is predicted to affect N's Kozak sequence and increase leaky ribosomal scanning^{365,367}. Omicron's ORF9b is upregulated compared to the ancestral strain, but less than Alpha's, as expected from the sequences. It is not understood so far how ORF9b is overexpressed in the Beta and Gamma variants³⁶⁷. ORF9b overexpression dampens the innate immune response.

Other mutations arose in ORF9b, notably in the Omicron variants: I5T, P10S, del27-29, N55S and T60A. Their roles have not been characterized yet; it is possible they were selected because they have a positive effect for N and no deleterious effect for ORF9b.

ORF6

ORF6 is a potent inhibitor of the interferon response; it blocks STAT1 and IRF3's translocation into the nucleus^{188,189}.

ORF6 is another accessory protein that is strongly overexpressed in some VOCs compared to the ancestral strain. In Alpha, Beta, Gamma and BA.1 there is a high expression of ORF6 sgRNA³⁶⁷, which is translated in a higher expression of ORF6 for Alpha, Beta and Gamma. However, in BA.1 and BA.2, there is a reduction of ORF6 protein levels compared to the ancestral strain³⁶⁷. BA.4 and BA.5 express levels of ORF6 protein that are comparable to those of the ancestral strain³⁶⁷. Most Omicron sublineages, but not BA.5, have a D61L substitution in ORF6. This mutation impairs binding with Nup98-Rae1, immune evasion and host mRNA import, diminishing the pathogenesis *in vivo*^{367,369}. It is unclear how ORF6 sgRNA and protein expression are regulated in the different variants as there are no mutations other than D61L, nor in the gene or its upstream sequence³⁷⁰.

N

N inhibits the stress granule dependent antiviral response¹³⁰ and type I and III IFN production by acting both on RIG-I and IRF3¹³¹. The expression of N follows similar trends as for ORF6: increased sgRNA for all VOCs except Delta, increased expression for Alpha, Beta and Gamma, reduced expression for BA.1 and BA2, similar for Delta, BA.4 and BA.5.

It is not understood so far which mutations regulate N sgRNA expression and translation. One of the first mutation to appear convergently in different variants was in the 203-205 region of N. These deletions increase the phosphorylation of N in its Serine/Aspartate rich domain (175-247); hyperphosphorylated N is in the cytoplasm (thus capable of antagonizing the innate immune response), while hypophosphorylated N is incorporated in virions³⁷¹. An optimal

equilibrium between the two forms might increase replication. This mutation in N also induces the production of a novel sgRNA and protein, N*; its role is unknown^{365,367}.

NSP6

Deletions 105-107 or 106-108 in NSP6 occurred independently in the Alpha, Beta, Gamma and Omicron variants. These deletions enhance the ability of mutant nsp6 to zipper the ER, inducing more and more regular DMVs³⁷², which likely contributes to the reduction of the innate immune response. This mutation also enhances the ability of nsp6 to suppress STAT1 and STAT2 phosphorylation and increases virulence in mice³⁷³.

2.4.3.2. Humoral immune escape

Although S is the main target of the humoral immune response as it has been described previously, N can also be targeted in case of natural infection. It is thus possible that some of its mutations have been selected to evade recognition by antibodies, but it has not been shown so far.

2.4.3.3. Cellular immune response escape

T cell epitopes are more conserved than those of neutralizing antibodies^{294,374}. Among 20 immunoprevalent epitopes defined in a study³⁷⁵, only 3 have been mutated in some of the latest Omicron subvariants: BA.2.75 has a mutation in an epitope in ORF1a/b, BQ.1 in E, and BA.2.86 in M. Another study reports Q213K in ORF3a, P13L/S/T, T362I and P365S in N confer escape from T-cell recognition³⁷⁶. Apart from P13L which is present in all Omicron VOCs, the other mutations have only been observed in sporadic variants. Another study reports that aa 417 and 981 are in targeted T cell epitope³⁷⁷; the K417N mutation additionally confers escape from the humoral response, and mutation 981 was only seen in BA.1. The breadth of the T-cell response in vaccinated or previously infected individuals, and the polymorphism in HLA genes make it unlikely that a variant with a mutation within a T-cell epitope that does not confer any other advantage has a worldwide spread³⁷⁶. T cell immunity escape by mutating epitopes does not appear to be a major driver of SARS-CoV-2 evolution.

2.4.3.4. Changes in viral replication

Very early after the start of the pandemic, a mutation in NSP12, a component of the RdRp was selected. The P323L substitution increases the interaction with NSP8, favorizing RdRp assembly, viral replication and thus transmissibility.

While a study previously mentioned shows the 105-107 deletion in nsp6 increases ER zippering and enhances the ability of NSP6 to suppress STAT1 and STAT2 phosphorylation, another study shows that RNA replication is lower with BA.1 NSP6 than with Delta NSP6³⁷⁸.

Furthermore, authors show BA.1 NSP6 is less efficient than Delta NSP6 to use lipid droplets to fuel replication organelles³⁷⁸. BA.1 has two mutations in NSP6, Δ 105-107 and I189V, their contribution to this phenotype is not elucidated. BA.2 and subsequent variants only have Δ 106-108, it is thus not known whether they have impaired lipid droplet consumption. However, they have the T223I mutation in ORF3a, which impairs lipid droplet accumulation and viral replication. This shows two different strategies adopted by the virus to lower replication by limiting lipid influx into DMVs, and thus pathogenesis and inflammation.

ORF8 expression reduces viral particle formation, through two distinct mechanisms mediated by its N-terminal and C-terminal regions³⁷⁹. It induces Golgi fragmentation and diminishes Spike expression³⁷⁹. While ORF8 reduces MHC-I expression on the cell surface¹⁹³ and Spike expression on the surface of infected cells which could limit their fusion and antibody recognition³⁸⁰, it is detrimental for viral particle assembly. The latest variants (XBB strains, EG5.1 and HK.3) contain a nonsense mutation that results in a truncated ORF8; restoring the full ORF8 diminishes viral particle production³⁷⁹. ORF8 deletion thus increases viral replication.

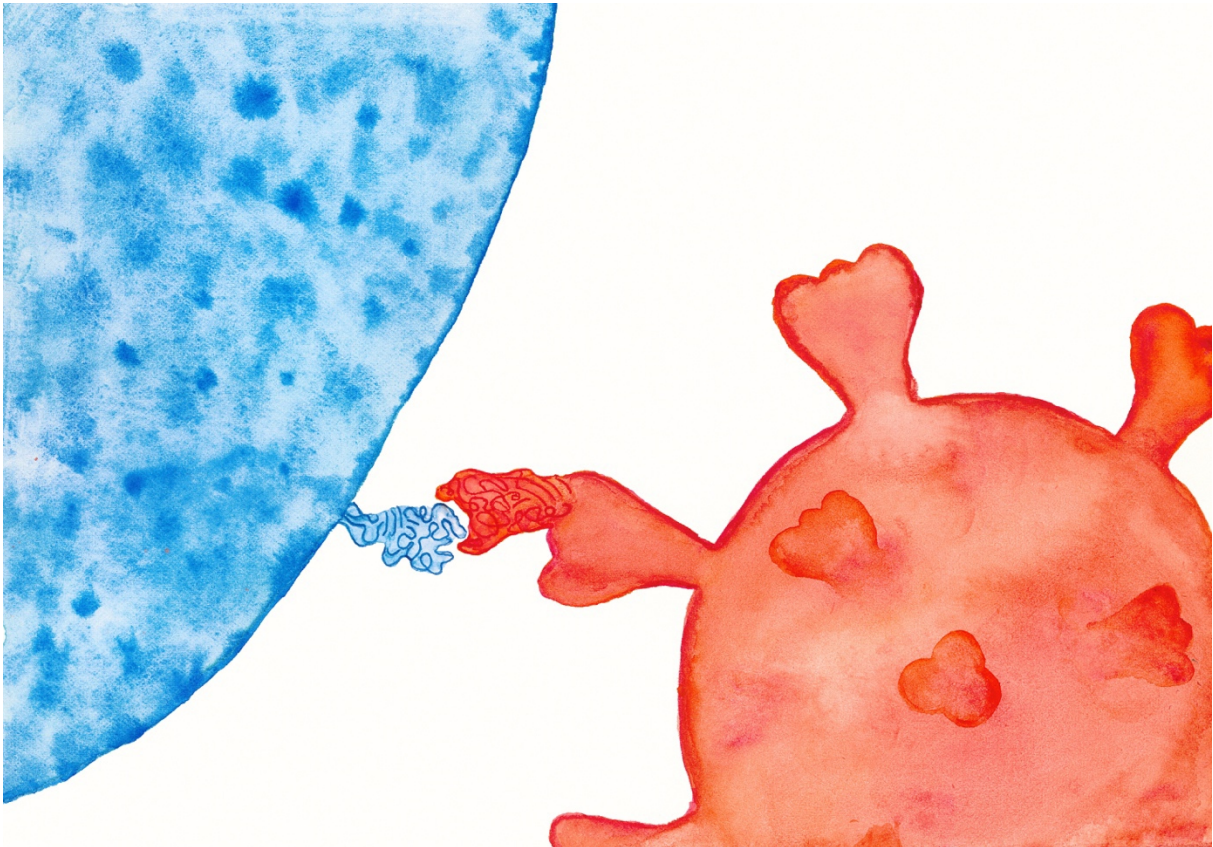
While some variants that are selected have increased replication, others have lower replication and pathogenicity. Indeed, too much pathogenicity could be detrimental for the spread of the virus.

2.4.3.5. Increased vRNA packaging

SARS-CoV-2 virus-like particles can be obtained by co-transfecting cells with E, S, M, N plasmids as well as a plasmid containing Luciferase and the SARS-CoV-2 genomic region containing the packaging signal³⁸¹. By introducing the R203K/M mutations in N, the authors show an increase in mRNA packaging³⁸¹. As mentioned previously, this mutation increases phosphorylation of N³⁷¹. VLPs produced with the mutant N seem to incorporate more RNA, but less N³⁸¹. Furthermore, by using reverse genetics on the ancestral strain, they show these mutations increase infectivity by 10-fold³⁸¹, which could be due both to increased vRNA packaging and an increase in the innate immune response antagonism as mentioned previously.

To conclude, SARS-CoV-2 evolution is dictated by many parameters and most notably innate and adaptive immunity escape. While some's mutation impact, notably in the Spike are well characterized, by us and others, the role of many of them remains a mystery. Deep mutational predictive screens have proven to be important to predict the impact of potential mutations in S and N on adaptive immunity escape or escape of recognition by antigenic tests. For the rest of the genome, who knows what evolution will bring...

Chapter 3 – The secrets of HKU1 entry



The articles presented in this chapter are the following:

- **Saunders N**, Fernandez I, Planchais C, Michel V, Rajah MM, Baquero Salazar E, Postal J, Porrot F, Guivel-Benhassine F, Blanc C, Chauveau-Le Fric G, Martin A, Grzelak L, Oktavia RM, Meola A, Ahouzi O, Hoover-Watson H, Prot M, Delaune D, Cornelissen M, Deijs M, Meriaux V, Mouquet H, Simon-Lorière E, van der Hoek L, Lafaye P, Rey F, Buchrieser J, Schwartz O. *TMPRSS2 is a functional receptor for human coronavirus HKU1*. Nature. 2023 Dec;624(7990):207-214.

This project was the result of a group effort supervised by Prof. Olivier Schwartz and Dr Julian Buchrieser. I present this work with full acknowledgment of the work performed by all authors. I was involved in cloning, fusion experiments, enzymatic activity measurement, expression stainings, Western Blots, pseudovirus infections, Spike binding assessment by flow cytometry, clinical isolate viral isolation, VHH testing in cell lines and on HAE cells. I trained an M1 student, Hunter Hoover Watson, during this project. I wrote the manuscript with Dr Julian Buchrieser and Prof. Olivier Schwartz. I performed the work necessary after the peer-review of the manuscript with the help of colleagues.

I wrote a short review of this paper in the French journal Médecines/Sciences. It is presented in **Appendix 3**.

- Fernandez I*, **Saunders N***, Duquerroy S, Bolland W, Arbabian A, Baquero E, Blanc C, Lafaye P, Haouz A, Buchrieser J, Schwartz O, Rey F. *Structural basis of TMPRSS2 zymogen activation and recognition by the HKU1 seasonal coronavirus*. Cell. In press, 2024. *equal contributions.

This project was led by Dr Ignacio Fernandez, under the supervision of Prof. Félix Rey and Prof. Olivier Schwartz, following our first collaboration on the identification of TMPRSS2 as HKU1's receptor. I performed the cloning of the mutant proteins and experimental validation of the roles of the residues situated at the RBD-TMPRSS2 interface. I performed part of the work concerning the animal TMPRSS2.

The results mentioned in the discussion on TMPRSS2 isoforms and GD3 synthase effects were obtained by Eva Thuillier, a master student I supervised. Home-made HAE were grown by Andréa Cottignies-Calamarte.

3.1. Introduction

3.1.1. HKU1

HKU1 is an embecovirus identified in 2005¹⁴, which does not have any close relatives. It is thought to originate from rodents, but intermediate hosts are unknown. It's main proteins (3CLpro, Pol, Hel, HE, S, E, M and N) have between 50 and 90% homology with those of other embecoviruses¹⁴. Compared to other embecoviruses, HKU1 has a deletion of approximately 800 bases in between the ORF1ab and HE¹⁴. This region encodes a phosphodiesterase that prevents the translation arrest and apoptosis induced by RNase L upon sensing of dsRNA³⁸². Furthermore, it has an addition of 2-20 tandem repeats of 30 aa, upstream of PL1pro, that encode a 10 aa fragment NDDEDVVTGD³⁸³. These repeated sequences are not present in any other coronavirus. Their role is uncharacterized.

There are three genotypes of HKU1: HKU1 A, B and C. HKU1 C clade stems from a recombination between HKU1A and HKU1B³⁸³. There are two possibilities to explain the presence of HKU1 A and B clades: either two close but separate events of spillover into the human species occurred, or HKU1 comes from one event of spillover but there was a long undetected circulation in humans (>100 years)³⁸⁴.

HKU1 is poorly characterized. Although it was first identified in 2005, to this date there are less than 90 articles containing "HKU1" in the title, two-thirds of which are case-reports, or epidemiological studies. Most of the knowledge on HKU1 is inferred from studies published on SARS-CoV-2, MHV or OC43, two other embecoviruses.

As for the other human seasonal coronavirus, their low pathogenicity has also contributed to the lack of interest of the scientific community in these viruses, until the COVID19 pandemic. This lack of data is also due to the difficulty to isolate this virus. To this date, it has only been isolated in primary human airway epithelial cells differentiated at the air-liquid interface. No permissive cell line has been identified, while it is the case for the 6 other human coronaviruses. In a study performed using lung-derived organoids, the authors were able to isolate high titers of the other three seasonal coronaviruses, but do not study HKU1³⁸⁵.

As described in the **Introduction**, HKU1 first binds to cells via 9-O-acetylated, α 2,8-linked sialic acids. This triggers an opening of the spike and exposure of the RBD. A groove in this RBD has been previously described to be the potential receptor binding motif (RBM); it is conserved between the two genotypes, it is targeted by neutralizing antibodies, and mutating it abolishes the competition in between soluble RBD and real virus. However, the receptor was unknown. After receptor binding, the spike is cleaved either at the surface or in endosomes,

which triggers membrane fusion and infection²²¹. Clinical isolates of HKU1 rely on TMPRSS2 for entry, rather than go through endosomes²²¹.

3.1.2. TMPRSS2

The type II transmembrane serine protease family (TTSP) has 20 known members in mouse, and 18 in humans²³², over half of which are expressed in the respiratory system (**Table 6**). Many of them play a pivotal role in viral entry, through priming of viral surface proteins (**Table 6**).

Protein	Expression in the respiratory system	Plays a role in the viral entry
Hepsin/TMPRSS1	lung	Avian metapneumovirus
TMRSS2	air, bronch, lung, trach	All human coronaviruses, IAV, IBV, paramyxoviruses, metapneumovirus, parainfluenza virus, Sendai virus
TMRSS3	lung	Not shown
TMRSS4	bronch, lung	IAV, SARS-CoV-2
Spinesin/TMPRSS5	none	Not shown
Matriptase-2/ TMPRSS6	lung	Not shown
Matriptase-3/ TMPRSS7	lung, trach	Not shown
Polyserase-1/TMPRSS9	none	Not shown
Corin/TMPRSS10	none	Not shown
HATL1/TMPRSS11A	trach	SARS-CoV
HATL5/TMPRSS11B	trach	Not shown
HAT/TMPRSS11D/AsP	air, bronch, trach	SARS-CoV, MERS-CoV, SARS-CoV-2, IAV, IBV
DESC1/TMPRSS11E	air, trach	IAV, SARS-CoV, MERS-CoV
HATL4/TMPRSS11F	trach	SARS-CoV-2
TMRSS12	none	Avian metapneumovirus
MSPL/MSPS/TMPRSS13	lung	IAV, SARS-CoV, MERS-CoV, SARS-CoV-2
Matriptase/ST14/PRSS14	lung	IAV
Enterokinase/TMPRSS15	none	Not shown

Table 6: TTSPs distribution and their role in viral glycoprotein cleavage. Upper airway (air), bronchioles (bronch), lung, trachea (trach), none. Influenza A/B virus (IAV/IBV). Adapted from Böttcher-Friebertshäuser, 2018²³².

They all have a N-terminal transmembrane domain, a stem region and a serine protease domain which has a catalytic triad composed of a serine, an aspartic acid, and a histidine³⁸⁶. The stem

region can include different types of protein domain. Several TTSPs have different isoforms because of alternative splicing; some isoforms lack part of the cytosolic domain. TTSPs cleave their substrate after an arginine or a lysine³⁸⁶. All those TTSPs are synthesized in an inactive form; they undergo cleavage in a conserved domain after an arginine or a lysine, between the stem region and the serine protease domain³⁸⁶. This allows a change of conformation into an active form, where both subunits remain attached by a disulfide bond³⁸⁶. In some cases, the protease domain can also be released. Several TTSPs autoactivate themselves *in vitro*³⁸⁶.

TMPRSS2 is expressed in the respiratory and digestive tracts, endocrine tissues, the kidneys, liver and pancreas as well as in male tissues and to a lesser extent in female tissues³⁸⁷. Its expression is regulated by androgens, but its role is unknown. TMPRSS2 is overexpressed in metastatic prostate cancer, and plays a role in metastasis³⁸⁸⁻³⁹⁰. In about 50% of prostate cancer TMPRSS2 promoter is fused with ERG fusion; this gene fusion plays a significant role in oncogenesis and disease progression^{391,392},

In its stem region, TMPRSS2 has a low-density lipoprotein receptor class A domain (LDLRA) which binds to calcium and a scavenger receptor cysteine-rich domain (SRCR) which binds to other surface or extracellular proteins^{222,232,386}. TMPRSS2 undergoes autocleavage after the arginine in position 255³⁹³. Mutating the cleavage site (R255Q) results in a poorly active protease, while mutating the active site (S441A) results in a full loss of the catalytic activity of TMPRSS2³⁹³.

TMPRSS2 has two main isoforms. Isoform 1 possesses an additional 37aa cytoplasmic tail compared to isoform 2. Their respective distribution is poorly studied, although on report shows they are both expressed in lung tissues³⁹⁴. Both isoforms cleave viruses' glycoproteins, allowing membrane fusion. Notably TMPRSS2 plays a role in the entry of human and animal coronaviruses, mammalian and avian influenza A, influenza B, paramyxoviruses, metapneumoviruses, parainfluenza and Sendai virus²³².

While the role of TMPRSS2 as an activating factor for many viruses has been extensively described, here we show it can also act as a receptor for the seasonal coronavirus HKU1. We then characterize the HKU1/TMPRSS2 complex as well as discuss the use of TMPRSS2 from different hosts.

3.2. TMPRSS2 is a functional receptor for human coronavirus HKU1

<https://doi.org/10.1038/s41586-023-06761-7>

Received: 7 April 2023

Accepted: 18 October 2023

Published online: 25 October 2023

 Check for updates

Nell Saunders¹, Ignacio Fernandez², Cyril Planchais³, Vincent Michel⁴, Maaran Michael Rajah¹, Eduard Baquero Salazar⁵, Jeanne Postal¹, Francoise Porrot¹, Florence Guivel-Benhassine¹, Catherine Blanc⁶, Gaëlle Chauveau-Le Friec⁷, Augustin Martin¹, Ludivine Grzelak¹, Rischa Maya Oktavia², Annalisa Meola², Olivia Ahouzi², Hunter Hoover-Watson¹, Matthieu Prot⁸, Deborah Delaune^{8,9}, Marion Cornelissen^{10,11}, Martin Deijs^{11,12}, Véronique Meriaux⁷, Hugo Mouquet³, Etienne Simon-Lorière^{8,13}, Lia van der Hoek^{11,12}, Pierre Lafaye⁷, Felix Rey², Julian Buchrieser^{1,15} & Olivier Schwartz^{1,14,15}✉

Four endemic seasonal human coronaviruses causing common colds circulate worldwide: HKU1, 229E, NL63 and OC43 (ref. 1). After binding to cellular receptors, coronavirus spike proteins are primed for fusion by transmembrane serine protease 2 (TMPRSS2) or endosomal cathepsins^{2–9}. NL63 uses angiotensin-converting enzyme 2 as a receptor¹⁰, whereas 229E uses human aminopeptidase-N¹¹. HKU1 and OC43 spikes bind cells through 9-O-acetylated sialic acid, but their protein receptors remain unknown¹². Here we show that TMPRSS2 is a functional receptor for HKU1. TMPRSS2 triggers HKU1 spike-mediated cell–cell fusion and pseudovirus infection. Catalytically inactive TMPRSS2 mutants do not cleave HKU1 spike but allow pseudovirus infection. Furthermore, TMPRSS2 binds with high affinity to the HKU1 receptor binding domain (Kd 334 and 137 nM for HKU1A and HKU1B genotypes) but not to SARS-CoV-2. Conserved amino acids in the HKU1 receptor binding domain are essential for binding to TMPRSS2 and pseudovirus infection. Newly designed anti-TMPRSS2 nanobodies potently inhibit HKU1 spike attachment to TMPRSS2, fusion and pseudovirus infection. The nanobodies also reduce infection of primary human bronchial cells by an authentic HKU1 virus. Our findings illustrate the various evolution strategies of coronaviruses, which use TMPRSS2 to either directly bind to target cells or prime their spike for membrane fusion and entry.

HKU1 was first identified from an elderly patient with severe pneumonia in Hong Kong in 2005 (ref. 13). HKU1 was later shown to cause the common cold and benign respiratory symptoms, but complications include severe lower respiratory infections, particularly in young children, the elderly and immunocompromised individuals¹⁴. It is estimated that 70% of children seroconvert before the age of 6 years¹⁵. The global seroprevalence of HKU1 is similar to that of other seasonal human coronaviruses (hCoV) and is between 75% and 95%^{15–17}. Three main viral genotypes have been identified: HKU1A, HKU1B and HKU1C. HKU1A and B spikes have 85% identity (Fig. 1a). HKU1C is a HKU1A/B recombinant, and its spike shares 99% identity with HKU1B. Conserved regions in the HKU1A and B spikes include a putative receptor binding domain (RBD) with a receptor binding motif (RBM)^{18,19} and S1/S2 and S2/S2' cleavage sites²⁰ (Fig. 1a).

Both HKU1 and OC43 spike bind to 9-O-acetylated α ,2,8 linked disialosides on target cells^{12,21,22}. Their protein receptors have not been identified. After receptor binding, the coronavirus spike is cleaved at the S2' site by membrane-bound proteases, such as transmembrane serine protease 2 (TMPRSS2), or by endosomal cathepsins, resulting in a conformational change of S2 that projects the fusion peptide into the target membrane to drive fusion²³. TMPRSS2 belongs to the type II transmembrane serine protease (TTSP) family, which comprises 19 cell-surface enzymes^{24,25}. TTSPs are involved in many processes, including epithelial homeostasis, extracellular matrix degradation, hormone and growth factor activation and initiation of proteolytic cascades through cleavage of membrane cellular proteins²⁴. TMPRSS2 primes coronavirus spikes and other viral-envelope glycoproteins, such as influenza haemagglutinin, enabling fusion²⁶. TTSPs are synthesized

¹Virus & Immunity Unit, Institut Pasteur, Université de Paris Cité, CNRS UMR 3569, Paris, France. ²Structural Virology Unit, Institut Pasteur, Université de Paris Cité, CNRS UMR 3569, Paris, France. ³Humoral Immunology Unit, Institut Pasteur, Université de Paris Cité, INSERM U1222, Paris, France. ⁴Pathogenesis of Vascular Infections Unit, Institut Pasteur, INSERM, Paris, France. ⁵Nanoimaging core, Institut Pasteur, Université de Paris Cité, INSERM U1222, Paris, France. ⁶Pasteur-TheraVectys Joint Lab, Institut Pasteur, Université de Paris Cité, Paris, France. ⁷Antibody Engineering Platform, C2RT, Institut Pasteur, Université de Paris Cité, CNRS UMR 3528, Paris, France. ⁸G5 Evolutionary Genomics of RNA Viruses, Institut Pasteur, Paris, France. ⁹Institut de Recherche Biomédicale des Armées, Brétigny-sur-Orge, France. ¹⁰Department of Medical Microbiology and Infection Prevention, Amsterdam UMC, Molecular Diagnostic Unit, University of Amsterdam, Amsterdam, The Netherlands. ¹¹Amsterdam Institute for Infection and Immunity, Amsterdam, The Netherlands. ¹²Department of Medical Microbiology and Infection Prevention, Amsterdam UMC, Laboratory of Experimental Virology, University of Amsterdam, Amsterdam, The Netherlands. ¹³National Reference Center for viruses of respiratory infections, Institut Pasteur, Paris, France. ¹⁴Vaccine Research Institute, Creteil, France. ¹⁵These authors contributed equally: Julian Buchrieser, Olivier Schwartz. ✉e-mail: julian.buchrieser@pasteur.fr; olivier.schwartz@pasteur.fr

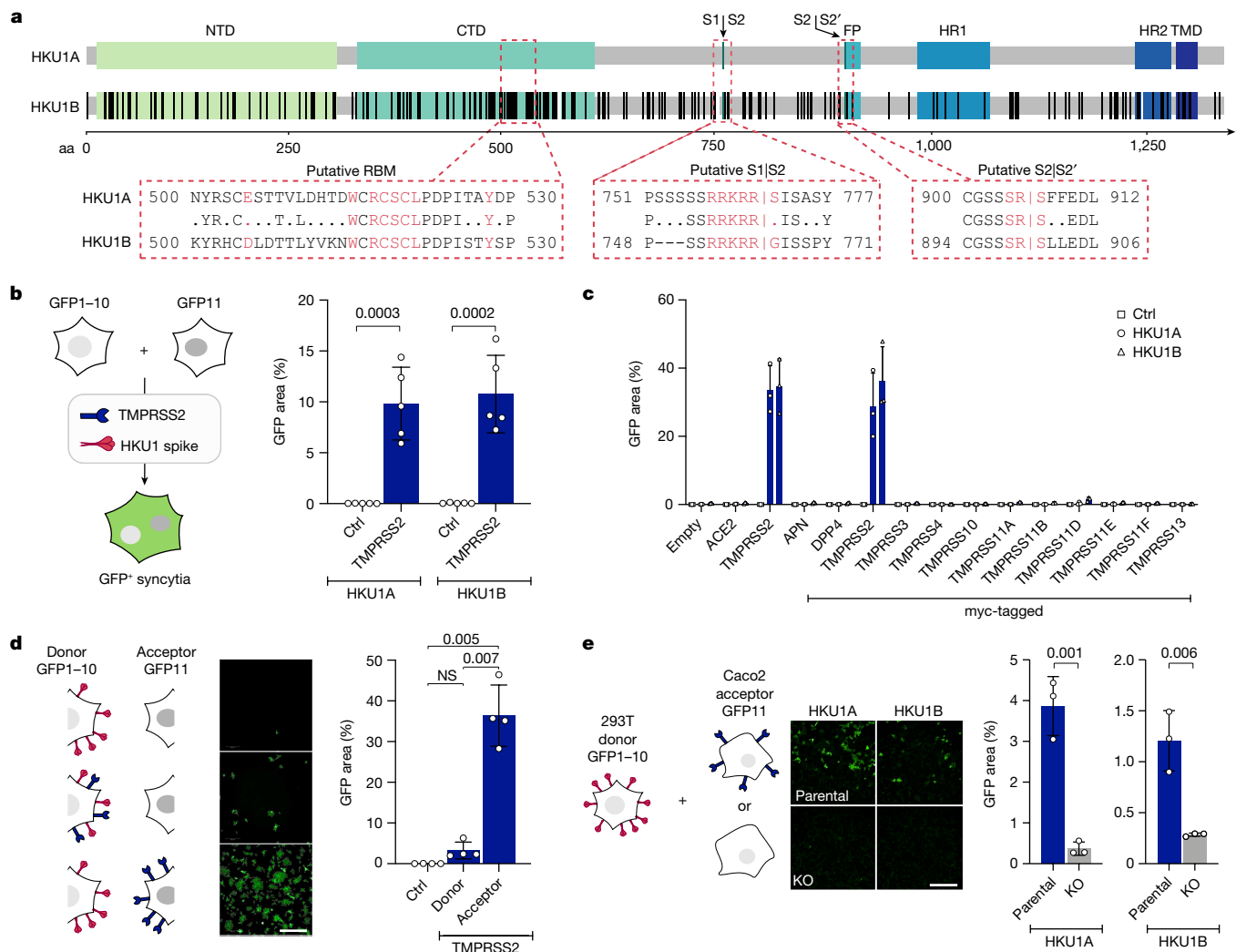


Fig. 1 | TMPRSS2 triggers HKU1 spike fusion. **a**, Alignment of HKU1A and B spikes. Black indicates a mismatch and white indicates deletion. Red boxed text indicates amino acids of the putative RBM and S1/S2 and S2/S2' cleavage sites. aa, amino acid position. **b**, TMPRSS2 mediates HKU1 cell–cell fusion. 293T cells expressing either GFP1-10 or GFP11 (293T-GFP-split cells) were transfected with HKU1 spike and TMPRSS2 or pQCXIP-empty control (Ctrl) plasmids, and fusion was quantified by measuring the GFP area after 20 h. Data are mean \pm s.d. of five independent experiments. **c**, Effect of a panel of proteases on cell–cell fusion. 293T-GFP-split cells were transfected with HKU1 spike and the indicated protease plasmids, fusion was quantified by measuring the GFP area after 48 h. Data are mean \pm s.d. of three independent experiments. **d**, TMPRSS2 has to be on the acceptor cell. TMPRSS2 was transfected either in donor cells and HKU1A spike or in acceptor cells. Fusion was quantified after 20 h. Left, experimental

design. Middle, representative images of GFP⁺ cells. Right, fusion quantification. Scale bar, 400 μ m. Data are mean \pm s.d. of four independent experiments. **e**, Role of endogenous TMPRSS2. 293T donor cells expressing HKU1A spike were mixed with Caco2 acceptor cells knocked out or not for the *tmprss2* gene. Left, experimental design. Middle, representative images of GFP⁺ cells. Right, fusion was quantified by measuring the GFP area after 20 h of coculture. Scale bar, 400 μ m. Data are mean \pm s.d. of three independent experiments. Statistical analysis: **b, d, e**, two-sided unpaired *t*-test compared to control/parental condition; **c**, one-way analysis of variance (ANOVA) on non-normalized data with Tukey's multiple comparisons. NTD, N-terminal domain; CTD, C-terminal domain; FP, fusion peptide; HR1/2, heptad repeat 1/2; NS, not significant; TMD, transmembrane domain.

as single-chain proenzymes that require proteolytic activation²⁴. TMPRSS2 undergoes autocleavage into two subunits that remain attached by a disulfide bond²⁵. In the respiratory tract, TMPRSS2 is expressed in nasal, bronchial and small-airways tissues and more particularly in ciliated cells that are the main target of HKU1 (refs. 27,28). Here, using cell–cell fusion assays, infections with pseudovirus and live virus, as well as in vitro binding tests, we identify TMPRSS2 as a high-affinity receptor for HKU1.

Generation of anti-TMPRSS2 nanobodies

To our knowledge, there are at present no available anti-TMPRSS2 antibodies that allow cell-surface staining, as most of the existing ones are directed towards cytosolic fragments of the protein. We thus

immunized an alpaca with the soluble ectodomain of a catalytically inactive TMPRSS2 (S441A mutant) to produce variable heavy domain of heavy chain (VHH) single-chain nanobodies. We isolated five VHHs (A01, F05, A07, C11 and D01) that bind with an affinity in the nanomolar range to soluble S441A TMPRSS2 (Extended Data Fig. 1a and Extended Data Table 1). We selected one VHH with the highest affinity (A01) to generate a dimeric antibody with a human Fc domain that efficiently stained TMPRSS2⁺ cells (Extended Data Fig. 1b).

TMPRSS2 triggers spike-dependent fusion

To investigate the effect of TMPRSS2 on HKU1 spike-mediated fusion, we generated plasmids encoding for the HKU1A (isolate N1, National Center for Biotechnology Information (NCBI) txid443239) and B

(isolate N5, NCBI txid443241) spikes (Fig. 1a). In 293T cells that do not express TMPRSS2, transient transfection led to surface expression of the two proteins, as assessed by flow cytometry using mAb10, a pan S2 coronavirus antibody²⁹ (Extended Data Fig. 2a). To study cell–cell fusion, we used a green fluorescent protein (GFP)-split-based model in which fused cells become GFP⁺ (Fig. 1b)³⁰. HKU1A and B spikes did not induce fusion alone but were highly fusogenic when co-expressed with TMPRSS2 in 293T cells (Fig. 1b and Extended Data Fig. 2b). By contrast, in experiments run in parallel, the SARS-CoV-2 spike induced fusion when co-expressed with angiotensin-converting enzyme 2 (ACE2) and not with TMPRSS2 (Extended Data Fig. 2c). HKU1 also triggered syncytia when co-expressed with TMPRSS2 in U2OS cells (Extended Data Fig. 2d).

As a control of specificity for HKU1 fusion, we tested a panel of 12 other surface proteases, including the coronavirus receptors aminopeptidase-N, dipeptidyl peptidase-4 and ACE2, as well as TMPRSS4 and TMPRSS11D, which have been reported to cleave the SARS-CoV-2 spike^{31–33}. The proteases were correctly expressed, as assessed by flow cytometry (Extended Data Fig. 2e–i), but only TMPRSS2 triggered cell–cell fusion (Fig. 1c).

We next generated fluorescent TMPRSS2-mNeonGreen and HKU1A Spike-mScarlet-I to follow fusion by video microscopy. When mixed 24 h after transfection, cells expressing spike and cells expressing TMPRSS2 formed syncytia in less than 1 h, indicating that this process is rapid (Extended Data Fig. 2j and Supplementary Videos 1 and 2). By expressing TMPRSS2 on either 293T GFP1-10 donor cells (spike-transfected) or 293T GFP11 acceptor cells, we observed that TMPRSS2 had to be present on acceptor cells, opposite of the cells expressing spike, to induce high levels of fusion (Fig. 1d).

We then investigated whether endogenous levels of TMPRSS2 were sufficient to induce HKU1 spike-dependent fusion. We mixed HKU1A spike-expressing 293T donor cells with Caco2 acceptor cells, which endogenously express low levels of TMPRSS2 (ref. 34) (Extended Data Fig. 3). Spike-expressing cells fused with parental Caco2 cells but not with Caco2 TMPRSS2 knockout (KO) cells, a Caco2 derivative in which the *tmprss2* gene was knocked down by CRISPR–Cas9 (Fig. 1e and Extended Data Fig. 3a–c). Silencing of TMPRSS2 using siRNA in Caco2 cells also significantly reduced fusion with spike-expressing 293T cells (Extended Data Fig. 3d).

Altogether, these results indicate that TMPRSS2 expression enables HKU1A and B spike-mediated cell–cell fusion.

Inactive TMPRSS2 allows HKU1 infection

To further analyse the role of TMPRSS2 in HKU1 spike fusion, we generated two well-characterized TMPRSS2 mutants, R255Q and S441A²⁵ (Fig. 2a). R255 is in the autocleavage site, and S441 is in the catalytic site; both mutations prevent TMPRSS2 autocleavage and activity. As expected, R255Q and S441A were correctly expressed in 293T cells but lacked catalytic activity, measured with a substrate generating a fluorescent signal on cleavage³⁵ (Extended Data Fig. 4a). We then studied the cleavage profile of TMPRSS2 and HKU1 spike by western blot (Fig. 2b). We focused our analysis on HKU1A spike because none of the polyclonal or monoclonal antibodies tested recognized HKU1B spike by western blot. When expressed alone, wild-type (WT) TMPRSS2 was cleaved but the mutants were not. When the HKU1A spike was present, a similar profile of TMPRSS2 processing was observed. Without TMPRSS2, the spike was partially cleaved into S1 and S2 subunits (Fig. 2b), most likely at the polybasic furin cleavage S1/S2 site (Fig. 1a). WT TMPRSS2, but not the catalytically inactive R255Q and S441A mutants, generated further cleavage bands in the spike, including a 100 kDa band below the S2 band, which likely corresponds to the S2' fragment.

We then examined how the spike was cleaved by TMPRSS2 present on adjacent cells. To this aim, we mixed spike-expressing cells with

TMPRSS2-expressing cells. We assessed the processing of the spike by western blot 3 and 6 h after coculture (Extended Data Fig. 4b). The spike cleavage by WT TMPRSS2 was slightly visible 3 h after coculture, and this process was more marked at 6 h. Of note, syncytia were detected at 3 and 6 h post-coculture in the cells transfected with WT TMPRSS2.

WT and mutant TMPRSS2 had slightly different levels of expression on transfection in 293T cells when assessed by flow cytometry (Extended Data Fig. 4c,d). We thus adjusted the amounts of plasmids to reach similar levels. The catalytically inactive TMPRSS2 mutants induced less cell–cell fusion despite correct expression with either HKU1A or B spike (Fig. 2c), confirming the importance of spike cleavage for cell–cell fusion.

We next generated single-cycle HKU1 pseudoviruses—lentiviral particles pseudotyped with HKU1 spikes—to further investigate the role of TMPRSS2 in HKU1 entry. The pseudovirus strategy has been successfully used to study SARS-CoV-2 entry and to identify ACE2 as a receptor for this virus³⁶. HKU1A and B pseudoviruses did not infect parental 293T cells efficiently, whereas transient expression of TMPRSS2 enabled high viral entry (Fig. 2d).

In striking contrast with their inability to trigger cell–cell fusion, the catalytically inactive R255Q and S441A TMPRSS2 mutants readily allowed infection of 293T cells with HKU1A or HKU1B pseudovirus (Fig. 2d). The TMPRSS2 mutants behave as expected in response to SARS-CoV-2 spike (D614G ancestral strain), which requires the enzymatic activity of the protease to enhance infectivity³⁶. In 293T-ACE2 cells, WT TMPRSS2 increased SARS-CoV-2 pseudotype entry by eightfold, but this was not the case for the R255Q and S441A mutants (Extended Data Fig. 4e). Together, our data show that catalytic activity of TMPRSS2 is not required for HKU1-mediated pseudovirus entry.

Coronavirus spike can be processed either by proteases at the surface of target cells, allowing for membrane fusion, or by cathepsins in endosomes, enabling entry after internalization⁸. As TMPRSS2 catalytic activity was not required for HKU1 entry, we examined the cytoplasmic access route of pseudovirus in 293T cells expressing WT or S441A TMPRSS2. To this aim, we performed infections in the presence of SB412515, a cathepsin L inhibitor or E64d, a pan-cysteine protease inhibitor. The drugs were added 2 h before infection and maintained for 48 h. SB412515 and E64d reduced the entry of HKU1A pseudovirus in S441A TMPRSS2 cells and not in WT TMPRSS2 cells (Extended Data Fig. 4f). This strongly indicates that viral entry occurs through endosomes with the catalytically inactive TMPRSS2 and at the surface with the WT protease. With HKU1B, SB412515 and E64d had little or no effect on entry (Extended Data Fig. 4g), indicating that HKU1B might be cleaved by other proteases or requires less cleavage to induce fusion, although it still uses TMPRSS2 as a receptor. Infection by both HKU1A and HKU1B pseudoviruses was inhibited by hydroxychloroquine in S441A TMPRSS2 cells, confirming an endocytic route with the mutant protease (Extended Data Fig. 4f,g).

We next asked whether TMPRSS2 could confer sensitivity to HKU1 pseudovirus in other cell lines. We generated U2OS and A549 stably expressing WT or S441A TMPRSS2. We reintroduced the two proteins in Caco2 TMPRSS2 KO cells. Flow cytometry demonstrated that TMPRSS2 was expressed in this panel of cells (Extended Data Fig. 4h,i). The presence of WT or S441A TMPRSS2 enabled infection to various extents of U2OS, A549 and Caco2 cells by HKU1A and HKU1B pseudoviruses (Fig. 2e). KO of TMPRSS2 significantly reduced entry into Caco2 cells. Furthermore, in parental Caco2 cells, preincubation with Camostat, a serine protease inhibitor known to inhibit TMPRSS2, did not affect HKU1 pseudovirus infection, whereas, as expected, it diminished SARS-CoV-2 entry by twofold (Extended Data Fig. 4j), confirming the role of endogenous TMPRSS2 as a receptor.

The Vero E6-TMPRSS2 cell line³⁷ remained insensitive to HKU1 pseudovirus infection (Extended Data Fig. 4k). Vero E6 cells are African green

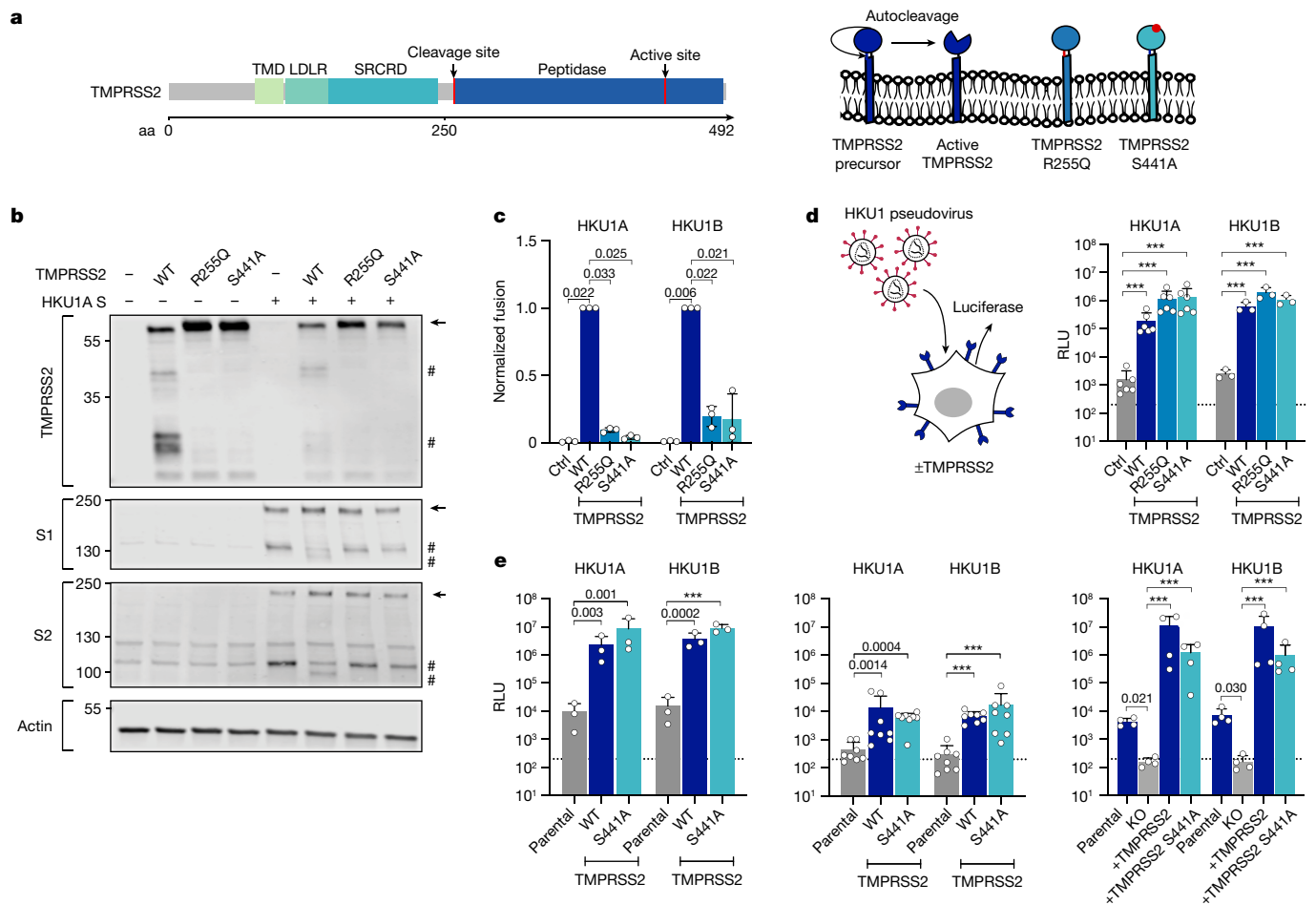


Fig. 2 | Effect of WT and mutant TMPRSS2 on HKU1 cell–cell fusion and pseudovirus infection. **a**, TMPRSS2 protein. TMPRSS2 is composed of a TMD, a class A low-density lipoprotein receptor (LDLR) domain, a scavenger receptor cysteine-2 rich domain (SRCRD) and a serine peptidase. TMPRSS2 precursor autocleaves at R255-L256, resulting in an active protease. **b**, TMPRSS2 and spike cleavage. Western blot of 293T transfected with HKU1A spike and indicated TMPRSS2 mutants. One membrane was probed for S1, TMPRSS2 and actin, another for S2 and actin. For gel source data, see Supplementary Information 1. Representative blots of three independent experiments. Molecular weights, kDa. Arrows and # denote the uncleaved and cleaved proteins, respectively. **c**, Catalytically inactive TMPRSS2 mutants do not trigger HKU1 cell–cell fusion. Fusion of 293T-GFP split transfected with HKU1 spike and mutant TMPRSS2 was quantified after 20 h and normalized to WT TMPRSS2. Ctrl, pQCXIP-Empty. Left, HKU1A. Right, HKU1B. Data are mean \pm s.d. of three independent

experiments. **d**, Catalytically inactive TMPRSS2 mutants mediate HKU1 pseudovirus infection. 293T cells transfected with WT or mutant TMPRSS2 were infected by Luc-encoding HKU1 pseudovirus. Luminescence was read 48 h postinfection. RLU, relative light unit. Dotted line indicates background in non-infected cells. Ctrl, pQCXIP-Empty. Data are mean \pm s.d. of six (HKU1A) or three (HKU1B) independent experiments. **e**, HKU1 pseudovirus infection in cell lines stably expressing TMPRSS2. Left, U2OS cells. Middle, A549 cells. Right, Caco2 cells. The TMPRSS2 KO Caco2 cells were stably transduced with indicated TMPRSS2. Data are mean \pm s.d. of three (U2OS), eight (A549) or four (Caco2) independent experiments. Statistical analysis: **c**, one-way ANOVA on non-normalized data with Dunnett’s multicomparison test compared to WT TMPRSS2-expressing cells; **d,e**, one-way ANOVA on log-transformed data with Dunnett’s multicomparison test compared to the untransfected (**d**), parental (**e**, U2OS and A549) or KO (**e**, Caco2) cells; *** $P < 0.0001$.

monkey kidney cells, so it is possible that other parameters (glycosylation, sialylation, adhesion molecules or other cellular proteins and so on) may regulate their sensitivity to HKU1.

We then analysed the role of sialic acids during TMPRSS2-mediated HKU1 entry. We treated U2OS TMPRSS2⁺ cells with Neuraminidase, an enzyme that removes (α 2,3), (α 2,6) and (α 2,8) linked sialic acids. Neuraminidase decreased binding of two sialic acid ligands, Lectin SNA and Siglec-E, without affecting surface levels of TMPRSS2 (Extended Data Fig. 5). Neuraminidase reduced sensitivity of target cells to HKU1 pseudovirus infection in a dose-dependent manner, indicating that sialic acids are necessary to trigger efficient HKU1 entry in TMPRSS2⁺ cells (Extended Data Fig. 5).

Taken together, our results show that TMPRSS2 is required for HKU1 pseudovirus infection but that its catalytic activity is dispensable, possibly because it can be rescued by other proteases such as cathepsins.

HKU1 spike binds to TMPRSS2

We next generated recombinant ectodomains of HKU1A and B spikes to investigate their binding to TMPRSS2-expressing cells (Fig. 3a,b and Extended Data Fig. 6a). HKU1 spike bound weakly to WT TMPRSS2 and more strongly to TMPRSS2 S441A and R255Q, as assessed by flow cytometry. Addition of Camostat increased binding to WT TMPRSS2 (Fig. 3b and Extended Data Fig. 6a), indicating that the proteolytic activity of TMPRSS2 somehow interfered with our readout or decreased binding by degrading or shedding the bound spike. Alternatively, WT TMPRSS2’s turnover might be faster than the mutants’. As expected, with a soluble SARS-CoV-2 spike, we detected binding to cells expressing ACE2 but not WT or mutant TMPRSS2 (Fig. 3b and Extended Data Fig. 6a,b), highlighting the different behaviours of HKU1 and SARS-CoV-2.

We then showed binding of the recombinant TMPRSS2 ectodomain to the HKU1A or B spikes by enzyme-linked immunosorbent assay (ELISA)

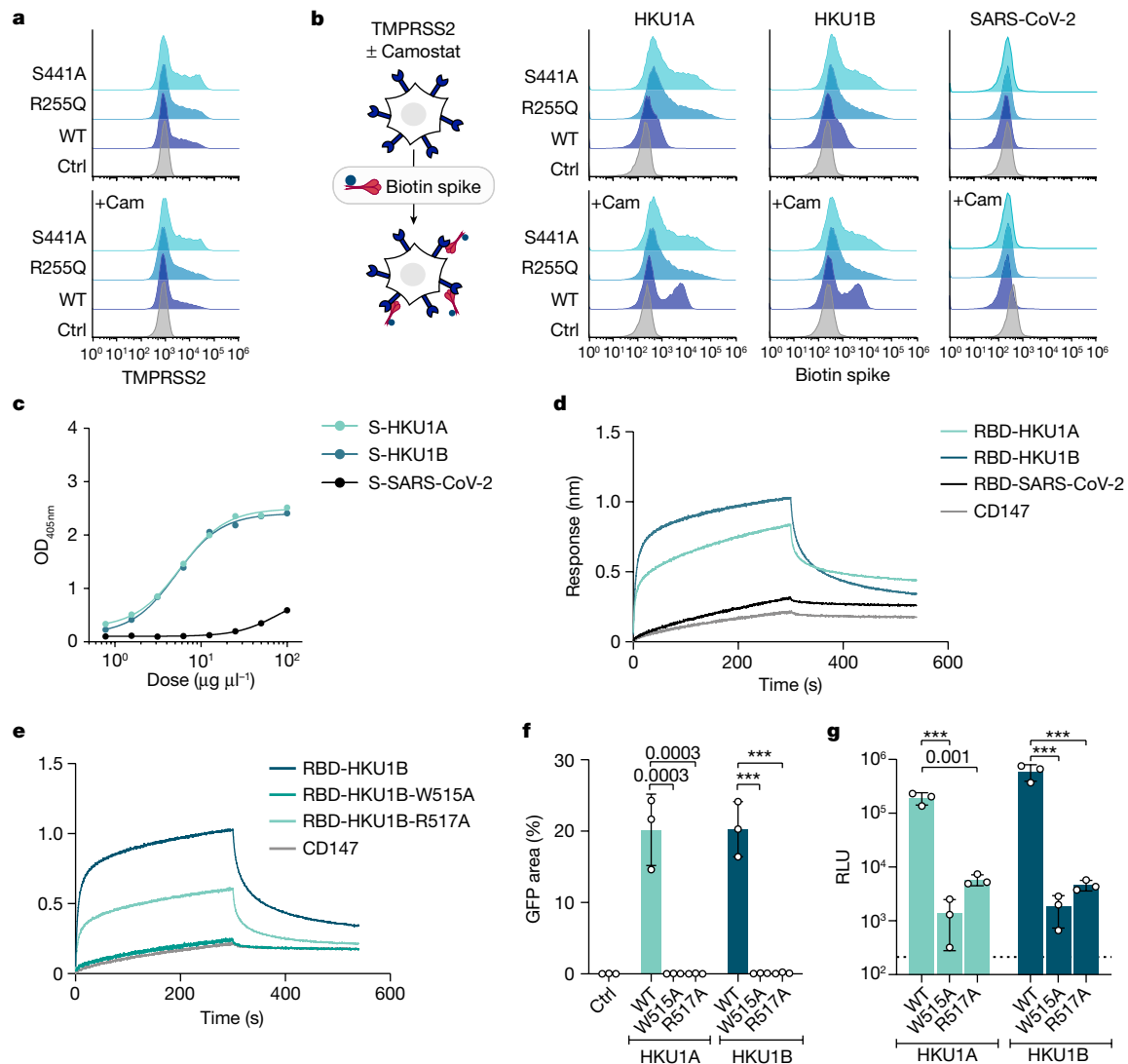


Fig. 3 | Analysis of HKU1 spike binding to TMPRSS2. **a, b**, Binding of soluble HKU1 spikes to 293T cells expressing TMPRSS2. Cells were transfected with TMPRSS2 mutants and incubated or not with 10 μ M of Camostat. TMPRSS2 levels (assessed with a commercial antibody) (**a**). Binding of soluble biotinylated trimeric spikes measured by flow cytometry (**b**). One representative experiment of three is shown. Ctrl, pQCXIP-Empty control plasmid. **c**, Binding of HKU and SARS-CoV-2 spikes on immobilized WT TMPRSS2 measured by ELISA. Mean of two independent experiments. **d**, Binding of S441A TMPRSS2 to RBD-coated receptors quantified by BLI. One representative experiment of four is shown. **e–g**, Properties of W515A and R517A mutant HKU1 spikes. Binding of TMPRSS2 to WT or mutant RBD-coated receptors quantified by BLI (**e**). One representative

experiment of four is shown. 293T GFP-split cells were transfected with TMPRSS2 and the indicated HKU1 mutant spikes, and fusion was quantified by measuring the GFP area after 24 h (**f**). 293T cells expressing TMPRSS2 were infected by Luc-encoding mutant HKU1 pseudovirus (**g**). Luminescence was read 48 h postinfection. Dotted line indicates background. Data are mean \pm s.d. of 3 independent experiments. Statistical analysis: **f**, one-way ANOVA on non-normalized data with Dunnett's multicomparison test compared to WT TMPRSS2-expressing cells; **g**, one-way ANOVA on log-transformed data with Dunnett's multicomparison test compared to WT spike pseudotypes * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

but not to the SARS-CoV-2 spike (Fig. 3c). This indicates a direct interaction between TMPRSS2 and HKU1 spike. Conversely, soluble ACE2 bound SARS-CoV-2 but not HKU1 spike by ELISA (Extended Data Fig. 6c). Because of the low yield of soluble WT TMPRSS2, we generated a soluble S441A mutant that could be obtained in amounts that allowed biophysical experiments. We expressed RBDs of the different viruses (residues 323–609 for HKU1A, 323–607 for HKU1B, 331–528 for SARS-CoV-2) to measure their affinity for S441A TMPRSS2 by biolayer interferometry (BLI). The RBDs from HKU1A and B interacted strongly with TMPRSS2, whereas the interaction of the SARS-CoV-2 RBD with TMPRSS2 was comparable to that of an irrelevant control protein (CD147) (Fig. 3d and Extended Data Fig. 6d,e). Using a range of S441A TMPRSS2 concentrations, affinity constants (K_d) of 334 nM and 137 nM were determined for HKU1A and B RBDs, respectively (Table 1, Extended Data Fig. 6d,e

and Supplementary Information 5 and 6). The SARS-CoV-2 RBD bound to the ectodomain of ACE2 with a K_d of 92 nM (Extended Data Fig. 6f and Table 1).

HKU1 spike RBM

The residues W515 and R517 in the HKU1 RBD¹⁹ have been reported to be critical for binding to an unknown cellular receptor¹⁸. We produced the recombinant HKU1B RBD with the W515A and R517A mutations. As assessed by BLI, the W515A mutation abrogated interaction with TMPRSS2, reaching response levels comparable to those obtained with a control protein, whereas the R517A mutation reduced binding by 2.8-fold ($K_d = 376$ nM) (Fig. 3e, Extended Data Fig. 6g,h, Table 1 and Supplementary Information 5 and 6). We also generated plasmids coding

Table 1 | Affinity (K_d) of the indicated RBD for TMPRSS2 or ACE2

Sensor	Ligand	K_d (nM)
RBD-HKU1A	TMPRSS2	334.3±66.2
RBD-HKU1B	TMPRSS2	136.7±20.8
RBD-HKU1B-W515A	TMPRSS2	ND
RBD-HKU1B-R517A	TMPRSS2	376.2±96.4
RBD-SARS-CoV-2	TMPRSS2	ND
RBD-HKU1B	ACE2	ND
RBD-SARS-CoV-2	ACE2	92

ND (non-detectable) denotes proteins for which interaction with the loaded sensor was similar to or less than interaction with an empty sensor.

for the HKU1A and HKU1B spikes with the W515A or R517A mutations. These mutants were correctly expressed at the cell surface, as assessed by flow cytometry (Extended Data Fig. 6i), but they lost their cell–cell fusion properties in the presence of TMPRSS2 (Fig. 3f). Their ability to trigger pseudovirus entry was decreased by two to three, resulting in infection levels close to background (Fig. 3g). Therefore, the conserved W515 and R517 residues in the HKU1A and B RBDs are critical for binding to TMPRSS2, viral fusion and entry.

Anti-TMPRSS2 VHHs block spike binding

We then examined whether the anti-TMPRSS2 VHH inhibited the HKU1 receptor function or the enzymatic activity of the protease. The five VHHs (A01, F05, A07, C11 and D01) efficiently bound soluble S441A TMPRSS2 as assessed by BLI (Extended Data Fig. 1a and Extended Data Table 1) and flow cytometry (Extended Data Fig. 7a). Three VHHs (A07, C11 and D01) inhibited HKU1B RBD interaction with S441A TMPRSS2 measured by BLI (Fig. 4a). The same three VHHs reduced TMPRSS2 catalytic activity (Extended Data Fig. 7b). Those nanobodies also inhibited HKU1A and B spike TMPRSS2-mediated cell–cell fusion (Fig. 4b and Extended Data Fig. 7c). We then examined their effect on viral entry independently of their inhibition of the catalytic activity of TMPRSS2. They reduced pseudovirus infection of 293T cells expressing S441A TMPRSS2 in a dose-dependent manner (Fig. 4c and Extended Data Fig. 7d). The two other nanobodies (A01 and F05), despite showing efficient binding to TMPRSS2, did not interfere with its enzymatic activity and proviral roles (Fig. 4b,c and Extended Data Fig. 7c,d), indicating that they bind to regions different from those of the three active VHHs and are not involved in HKU1 spike binding.

Anti-TMPRSS2 VHH blocks HKU1 infection

HKU1 does not grow in any cell line tested so far, but viral amplification in human ciliated airway epithelial cell cultures has been reported^{28,38,39}. We isolated an HKU1 virus from a nasal swab of an individual suffering from a respiratory tract infection. To this end, we used primary human bronchial epithelial (HBE) cells differentiated at the air–liquid interface for more than four weeks. We first examined whether these cells were positive for TMPRSS2. Immunofluorescence of HBE cells with anti-TMPRSS2 VHH A01-Fc revealed a preferential staining of ciliated cells, with a positive signal accumulating at the cilia (Fig. 5a). We then amplified the virus from the clinical sample by one passage on HBE cells. We observed an increase of HKU1 viral RNA in apical culture supernatants, with concentrations peaking at 5×10^6 viral RNA copies per μ l at 2–3 days postinfection (Extended Data Fig. 8a). Metagenomic sequencing of the viral supernatant identified a HKU1 genotype B (Extended Data Fig. 8b), and no other virus was detected. For reasons that remain to be elucidated, our attempts to further grow this HKU1B isolate in various cell lines expressing WT or S441 TMPRSS2 were unsuccessful. However, the virus grew on a second passage on HBE cells, indicating that it was infectious (Extended Data Fig. 8a).

We then asked whether the anti-TMPRSS2 nanobodies inhibited infection of HBE cells. We pre-incubated target cells with the A07 VHH or a control nanobody and measured the spike intensity by immunofluorescence 48 h postinfection (Fig. 5b). The A07 VHH strongly reduced the appearance of infected cells, indicating that the spike binding and/or cleavage activities of TMPRSS2 are necessary for a productive HKU1 infection.

Discussion

Here we demonstrate that TMPRSS2 is a receptor for HKU1. TMPRSS2 triggers HKU1-mediated cell–cell fusion and viral entry and binds with high affinity to both HKU1A and HKU1B RBDs. The enzymatic activity of TMPRSS2 is required for HKU1-dependent cell–cell fusion but not for entry of HKU1 pseudovirus. Two cathepsin inhibitors decrease HKU1A pseudovirus infection mediated by the catalytically inactive TMPRSS2 but not by the WT protease. This strongly indicates that after viral binding to TMPRSS2, viral particles expressing HKU1A can either fuse at the plasma membrane if the protease is active or be internalized and processed in the endosomal compartment when the protease is inactive. For HKU1B, other mechanisms could be at play.

TMPRSS2 binds directly to HKU1 but not to SARS-CoV-2 spike. Auto-cleavage of TMPRSS2 is not required for binding. The affinity of HKU1A and HKU1B ($K_d = 334$ and 131 nM, respectively) is slightly below what we

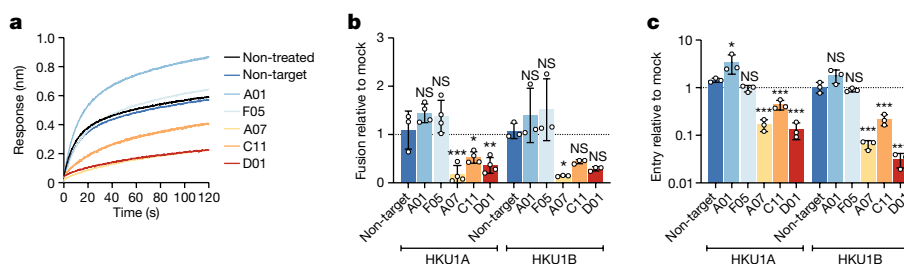


Fig. 4 | Anti-TMPRSS2 VHH nanobodies inhibit HKU1 binding to TMPRSS2, cell–cell fusion and pseudovirus infection. a, Effect of VHHs on binding of HKU1B RBD to TMPRSS2 measured by BLI. One representative experiment of two is shown. **b**, Effect of VHHs on HKU1-mediated cell–cell fusion. 293T GFP-split cells were transfected with TMPRSS2 and HKU1 spike in the presence of 1μ M of VHH, and fusion was quantified 20 h later. Data were normalized to the non-VHH treated condition (dotted line). Data are mean \pm s.d. of four independent experiments. **c**, Effect of VHHs on HKU1 pseudovirus infection. 293T cells transfected with S441A TMPRSS2 were treated with 1μ M VHH 2 h

before infection. Luminescence was read 48 h postinfection. Data were normalized to the non-treated condition for each virus (dotted line). Data are mean \pm s.d. of three independent experiments. Statistical analysis: **b**, one-way ANOVA data with Dunnett’s multiple comparisons compared to non-target VHH; **c**, one-way ANOVA on log-transformed data with Dunnett’s multiple comparisons compared to non-target VHH; * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. Exact P values of (b), HKU1A, A07 0.0005, C11 0.025, D01 0.004; HKU1B, A07 0.029; (c) HKU1A, A01 0.016, A07 < 0.0001, C11 0.0009, D01 < 0.0001; HKU1B, A07/C11/D01 < 0.0001.

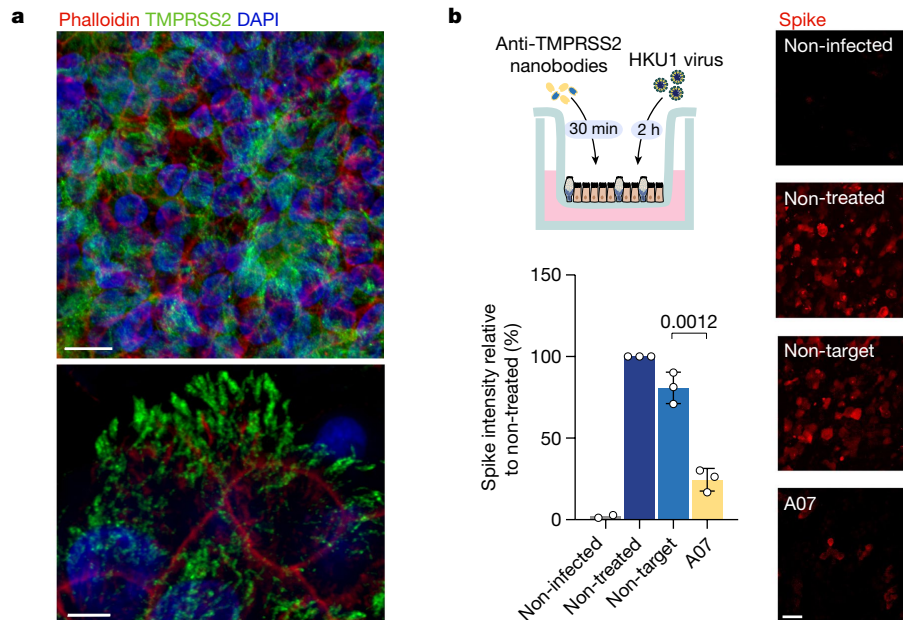


Fig. 5 | Live HKU1 virus infection of HBE cells. **a**, TMPRSS2 staining of HBE cells. Red, phalloidin; blue, 4',6'-diamidino-2-phenylindole (DAPI); green, TMPRSS2 stained with VHH A01-Fc. Scale bars, 10 μ m (top), 5 μ m (bottom). Images are representative of three independent experiments. **b**, Effect of the anti-TMPRSS2 VHH (A07) on HKU1 infection. The experimental design is represented. Infected cells were visualized with an anti-spike antibody and

scored. Representative images of spike staining 48 h postinfection are shown. Spike pixel intensity in five random fields per experiment was measured and normalized to the intensity in the infected but non-treated condition. Data are mean \pm s.d. of three independent experiments for infected conditions and mean of two for uninfected condition. Scale bar, 20 μ m. Statistical analysis: two-sided unpaired *t*-test compared to non-target VHH.

measured for the SARS-CoV-2 RBD interaction with ACE2 ($K_d = 92$ nM). A conserved groove in a region of HKU1A and B spikes composed of amino acid 505, 515, 517–521 and 528 has been proposed to be involved in binding to an unknown receptor^{18,19}. We show that the 515 and 517 residues of the spike are essential for efficient binding to TMPRSS2, viral fusion and entry, thus adding evidence that this region may be part of the RBM. Further structural studies are warranted to precisely map the amino acids and regions allowing the interactions between the spike and its receptor. Another protease, Kalikrein 13, has been reported to cleave HKU1 spike⁴⁰. Future work will help address how Kalikrein 13 may modulate HKU1 interaction with TMPRSS2.

We further designed anti-TMPRSS2 VHHs that inhibit TMPRSS2 binding to HKU1 spikes and prevent cell–cell fusion and viral entry. The anti-TMPRSS2 nanobodies inhibit the enzymatic activity of TMPRSS2 and hinder HKU1 pseudovirus entry using S441A TMPRSS2. This strongly indicates that the spike binds close to the catalytic site of the enzyme. These nanobodies confirmed the role of TMPRSS2 as a HKU1 receptor and provide useful tools to interfere with TMPRSS2 function. TMPRSS2 also plays a role in the development of certain cancers⁴¹. It will be worth studying how these nanobodies may target tumour cells or inhibit infection of viruses relying on this enzyme, as an alternative to small-molecule inhibitors⁴².

Finally, we isolated an HKU1B live virus and report that TMPRSS2 is required for productive infection of primary HBE cells, which are natural targets of the virus. It will be of interest to further determine whether the HKU1 binding property, the enzymatic activity or, most likely, both functions of TMPRSS2 are necessary to trigger HKU1 infection in these primary cells. We have so far been unable to amplify the virus in cell lines. Future work will be necessary to understand which step of the viral life cycle is impaired in TMPRSS2⁺ cell lines or if another factor is required. It will also be worth examining whether soluble TMPRSS2, known to be shed in the extracellular space, may interfere with HKU1 entry.

Human coronavirus receptors that have been identified so far and allow productive viral entry are cell-surface proteases. TMPRSS2 and

other TTSPs are known to prime human coronaviruses for fusion by cleaving their spikes, generally after viral binding to target cells. Of note, Omicron strains rely less on TMPRSS2 than previous SARS-CoV-2 lineages, reflecting a constant adaptation of coronaviruses to their hosts⁴³. Our study provides the first evidence, to our knowledge, that TMPRSS2 also acts as a direct physical receptor for HKU1 entry, in addition to its previously known role in cleaving the spike. ACE2 binding is an ancestral and evolvable trait of sarbecoviruses^{44,45}. TMPRSS2 binding may represent another parameter of coronavirus evolution. Whether coronaviruses have lost affinity or sensitivity to TMPRSS2 during evolution or whether HKU1 has gained affinity for human or animal TMPRSS2 remains to be determined. Our results highlight the critical role of TMPRSS2 and other proteases as determinants of coronavirus tropism and pathogenesis²³.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-023-06761-7>.

1. Park, S., Lee, Y., Michelow, I. C. & Choe, Y. J. Global seasonality of human coronaviruses: a systematic review. *Open Forum Infect. Dis.* **7**, ofaa443 (2020).
2. Bertram, S. et al. TMPRSS2 activates the human coronavirus 229E for cathepsin-independent host cell entry and is expressed in viral target cells in the respiratory epithelium. *J. Virol.* **87**, 6150–6160 (2013).
3. Glowacka, I. et al. Evidence that TMPRSS2 activates the severe acute respiratory syndrome coronavirus spike protein for membrane fusion and reduces viral control by the humoral immune response. *J. Virol.* **85**, 4122–4134 (2011).
4. Shirato, K., Kawase, M. & Matsuyama, S. Wild-type human coronaviruses prefer cell-surface TMPRSS2 to endosomal cathepsins for cell entry. *Virology* **517**, 9–15 (2018).
5. Shirato, K., Kawase, M. & Matsuyama, S. Middle East respiratory syndrome coronavirus infection mediated by the transmembrane serine protease TMPRSS2. *J. Virol.* **87**, 12552–12561 (2013).
6. Milewska, A. et al. Human coronavirus NL63 utilizes heparan sulfate proteoglycans for attachment to target cells. *J. Virol.* **88**, 13221–13230 (2014).

7. Millet, J. K. & Whittaker, G. R. Host cell entry of Middle East respiratory syndrome coronavirus after two-step, furin-mediated activation of the spike protein. *Proc. Natl Acad. Sci. USA* **111**, 15214–15219 (2014).
8. Shang, J. et al. Cell entry mechanisms of SARS-CoV-2. *Proc. Natl Acad. Sci. USA* **117**, 11727–11734 (2020).
9. Shulla, A. et al. A transmembrane serine protease is linked to the severe acute respiratory syndrome coronavirus receptor and activates virus entry. *J. Virol.* **85**, 873–882 (2011).
10. Hofmann, H. et al. Human coronavirus NL63 employs the severe acute respiratory syndrome coronavirus receptor for cellular entry. *Proc. Natl Acad. Sci. USA* **102**, 7988–7993 (2005).
11. Yeager, C. L. et al. Human aminopeptidase N is a receptor for human coronavirus 229E. *Nature* **357**, 420–422 (1992).
12. Hulswit, R. J. G. et al. Human coronaviruses OC43 and HKU1 bind to 9-O-acetylated sialic acids via a conserved receptor-binding site in spike protein domain A. *Proc. Natl Acad. Sci. USA* **116**, 2681–2690 (2019).
13. Woo, P. C. Y. et al. Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. *J. Virol.* **79**, 884–895 (2005).
14. Kahn, J. S. & McIntosh, K. History and recent advances in coronavirus discovery. *Pediatr. Infect. Dis. J.* **24**, S223–S227 (2005).
15. Sayama, Y. et al. Seroprevalence of four endemic human coronaviruses and, reactivity and neutralization capability against SARS-CoV-2 among children in the Philippines. *Sci. Rep.* **13**, 2310 (2023).
16. Killerby, M. E. et al. Human coronavirus circulation in the United States 2014–2017. *J. Clin. Virol.* **101**, 52–56 (2018).
17. Woudenberg, T. et al. Humoral immunity to SARS-CoV-2 and seasonal coronaviruses in children and adults in north-eastern France. *EBioMedicine* **70**, 103495 (2021).
18. Ou, X. et al. Crystal structure of the receptor binding domain of the spike glycoprotein of human betacoronavirus HKU1. *Nat. Commun.* **8**, 15216 (2017).
19. Kirchdoerfer, R. N. et al. Pre-fusion structure of a human coronavirus spike protein. *Nature* **531**, 118–121 (2016).
20. Bestle, D. et al. TMPRSS2 and furin are both essential for proteolytic activation of SARS-CoV-2 in human airway cells. *Life Sci. Alliance* **3**, e202000786 (2020).
21. Huang, X. et al. Human coronavirus HKU1 spike protein uses o-acetylated sialic acid as an attachment receptor determinant and employs hemagglutinin-esterase protein as a receptor-destroying enzyme. *J. Virol.* **89**, 7202–7213 (2015).
22. Li, Z. et al. Synthetic O-acetylated sialosides facilitate functional receptor identification for human respiratory viruses. *Nat. Chem.* **13**, 496–503 (2021).
23. Millet, J. K., Jaimes, J. A. & Whittaker, G. R. Molecular diversity of coronavirus host cell entry receptors. *FEMS Microbiol. Rev.* **45**, fuao057 (2020).
24. Bugge, T. H., Antalis, T. M. & Wu, Q. Type II transmembrane serine proteases. *J. Biol. Chem.* **284**, 23177–23181 (2009).
25. Fraser, B. J. et al. Structure and activity of human TMPRSS2 protease implicated in SARS-CoV-2 activation. *Nat. Chem. Biol.* **18**, 963–971 (2022).
26. Böttcher-Friebertshäuser, E. in *Activation of Viruses by Host Proteases* (eds Böttcher-Friebertshäuser, E., Garten, W. & Klenk, H.) Ch. 8 (Springer, 2018); https://doi.org/10.1007/978-3-319-75474-1_8.
27. Lukassen, S. et al. SARS-CoV-2 receptor ACE2 and TMPRSS2 are primarily expressed in bronchial transient secretory cells. *EMBO J.* **39**, e105114 (2020).
28. Pirc, K. et al. Culturing the unculturable: human coronavirus HKU1 infects, replicates, and produces progeny virions in human ciliated airway epithelial cell cultures. *J. Virol.* **84**, 11255–11263 (2010).
29. Planchais, C. et al. Potent human broadly SARS-CoV-2-neutralizing IgA and IgG antibodies effective against Omicron BA.1 and BA.2. *J. Exp. Med.* **219**, e20220638 (2022).
30. Buchrieser, J. et al. Syncytia formation by SARS-CoV-2-infected cells. *EMBO J.* **39**, e106267 (2020).
31. Zang, R. et al. TMPRSS2 and TMPRSS4 promote SARS-CoV-2 infection of human small intestinal enterocytes. *Sci. Immunol.* **5**, eabc3582 (2020).
32. Kishimoto, M. et al. TMPRSS11D and TMPRSS13 activate the SARS-CoV-2 spike protein. *Viruses* **13**, 384 (2021).
33. Hoffmann, M. et al. Camostat mesylate inhibits SARS-CoV-2 activation by TMPRSS2-related proteases and its metabolite GBPA exerts antiviral activity. *EBioMedicine* **65**, 103255 (2021).
34. Koch, J. et al. TMPRSS2 expression dictates the entry route used by SARS-CoV-2 to infect host cells. *EMBO J.* **40**, e107821 (2021).
35. Afar, D. E. H. et al. Catalytic cleavage of the androgen-regulated TMPRSS2 protease results in its secretion by prostate and prostate cancer epithelia. *Cancer Res.* **31**, 1686–1692 (2001).
36. Hoffmann, M. et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* **181**, 271–280.e8 (2020).
37. Planas, D. et al. Resistance of Omicron subvariants BA.2.75.2, BA.4.6, and BQ.1.1 to neutralizing antibodies. *Nat. Commun.* **14**, 824 (2023).
38. Dominguez, S. R. et al. Isolation, propagation, genome analysis and epidemiology of HKU1 betacoronaviruses. *J. Gen. Virol.* **95**, 836–848 (2014).
39. Dijkman, R. et al. Isolation and characterization of current human coronavirus strains in primary human epithelial cell cultures reveal differences in target cell tropism. *J. Virol.* **87**, 6081–6090 (2013).
40. Milewska, A. et al. Kallikrein 13 serves as a priming protease during infection by the human coronavirus HKU1. *Sci. Signal.* **13**, eaba9902 (2020).
41. Koistinen, H. et al. The roles of proteases in prostate cancer. *IUBMB Life* <https://doi.org/10.1002/iub.2700> (2023).
42. Shrimp, J. H. et al. An enzymatic TMPRSS2 assay for assessment of clinical candidates and discovery of inhibitors as potential treatment of COVID-19. *ACS Pharmacol. Transl. Sci.* **3**, 997–1007 (2020).
43. Suzuki, R. et al. Attenuated fusogenicity and pathogenicity of SARS-CoV-2 Omicron variant. *Nature* **603**, 700–705 (2022).
44. Li, W. et al. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* **426**, 450–454 (2003).
45. Starr, T. N. et al. ACE2 binding is an ancestral and evolvable trait of sarbecoviruses. *Nature* **603**, 913–918 (2022).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature Limited 2023

Methods

Plasmids

Codon-optimized HKU1A (RefSeq: YP_173238.1) and B/C isolate N5P8 referred to as HKU1B (UniProtKB/Swiss-Prot: QOZME7.1) full spikes were ordered as synthetic genes (GeneArt, Thermo Fisher Scientific) and cloned into a phCMV backbone (GeneBank, AJ318514) by replacing the VSV-G gene. pQCXIP-Empty control plasmid was previously described³⁰. pQCXIP-BSR-GFP11 and pQCXIP-GFP1-10 were a kind gift from Yutaka Hata⁴⁶ (Addgene plasmid no. 68716 and no. 68715). pCSDest-TMPRSS2 was a kind gift from Roger Reeves⁴⁷ (Addgene plasmid no. 53887). Mutations in the HKU1 spike and TMPRSS2 were introduced using the NEB Q5 Site-Directed Mutagenesis Kit. Plasmids were sequenced before use. phCMV-HKU1-S-mNeonGreen and pCDEST-TMPRSS2-mScarlet-I were generated by Gibson assembly. pCAGGS-based expression vectors N-terminal MYC-epitope tagged TMPRSS2, TMPRSS3, TMPRSS4, TMPRSS10, TMPRSS11A, TMPRSS11B, TMPRSS11D, TMPRSS11E, TMPRSS11F and TMPRSS13 were a kind gift from Stefan Pöhlmann³³. pLV-EF1a-IRES-Hygro was a gift from Tobias Meyer (Addgene plasmid no. 85134)⁴⁸. pLV-EF1a-DEST-IRES-Hygro was generated by Gibson assembly. pLV-TMPRSS2-Hygro was generated by cloning TMPRSS2 from pCSDest-TMPRSS2 into pLV-EF1a-DEST-IRES-Hygro using gateway cloning. pLV-TMPRSS2-S441A-Hygro and pLV-TMPRSS2-R255Q-Hygro were generated by Q5 site-directed mutagenesis (New England Biolabs).

Cells

HEK293T (293T), U2OS, Vero E6, A549 and Caco2/TC7 (Caco2) were from the American Type Culture Collection and cultured in Dulbecco's Modified Eagle's Medium (DMEM) with 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin. GFP-split cells were previously described³⁰ and cultured with 1 $\mu\text{g ml}^{-1}$ of puromycin (InvivoGen). Cells stably expressing TMPRSS2 were cultured with 100 $\mu\text{g ml}^{-1}$ hygromycin. Cells were routinely screened for mycoplasma. Cells were authenticated by genotyping (Eurofins).

Reagents

SB412515 was purchased from Cayman Chemical, E64d and Camostat mesylate from Sigma-Aldrich, hydroxychloroquine from Merck and Neuraminidase (or Sialidase from *Arthrobacter ureafaciens*) from Roche.

Sequence alignments

Alignments of HKU1A and B spikes in Fig. 1a were performed using Protein Blast with default settings (NCBI). Alignment figures were generated using the seqvisr package in R (Github: <https://doi.org/10.5281/zenodo.6583981>).

GFP-split fusion assay

Cell-cell fusion assays were performed as previously described^{30,49}. Briefly, 293T cells stably expressing GFP1-10 and GFP11 were cocultured at a 1:1 ratio (6×10^4 cells per well) and transfected in suspension with Lipofectamine 2000 (Thermo Fisher Scientific) in a 96-well plate (uClear, no. 655090) (10 ng of spike plasmid, indicated amounts of TMPRSS2 plasmids adjusted to 100 ng DNA with pQCXIP-Empty). For acceptor-donor experiments, 293T GFP-split cells (GFP1-10 and GFP11) were transfected separately in suspension with 500 ng of DNA (10% spike, indicated amounts of TMPRSS2 plasmid adjusted to 500 ng with pQCXIP-Empty) for 30 min at 37 °C. Cells were washed twice, and acceptor and donor cells were mixed and seeded at 6×10^4 cells per well. For Caco2 knockdown experiments, Caco2 GFP11 cells were transfected with control siRNA-directed against Luciferase 5'-CGUACGCGGAUACUUCGA-3' or siGENOME Human TMPRSS2 SMARTpool (no. 7113, Horizon Discovery) at 50 nM using Lipofectamine RNAiMAX (Thermo Fisher Scientific) in a six-well dish for 48 h. 293T GFP1-10 cells were transfected with 10% of spike plasmid in a six-well

dish for 24 h. Caco2 knockdown or Caco2 KO cells and 293T spike cells were mixed at a 1:1.5 ratio in a 96-well plate. The remaining Caco2 cells were used for RNA extraction for quantitative polymerase chain reaction (qPCR). For all experiments, at 20 h posttransfection, images covering 90% of the well surface were acquired per well on an Opera Phenix High-Content Screening System (PerkinElmer). The GFP area was quantified on Harmony High-Content Imaging and Analysis Software.

RNA extraction, reverse transcription and qPCR

At 48 h post-siRNA transfection, 5×10^5 Caco2 cells were lysed using RNA lysis buffer (QIAGEN) supplemented with 10 μl of β -mercaptoethanol. RNA extraction was performed using the RNeasy Plus Mini Kit (QIAGEN) according to the manufacturer's protocol. Reverse transcription was performed using SuperScript II (Thermo Fisher Scientific) according to the manufacturer's protocol. qPCR was performed using iTaq universal SYBR green supermix (Bio-Rad) on a QuantStudio 6 Real-Time PCR machine (Thermo Fisher Scientific). The following primers were used: β -Tubulin forward, 5'-CTTCGGCCAGATCTTCAGAC-3', reverse 5'-AGAGAGTGGGTCAGCTGGAA-3'; TMPRSS2 forward, 5'-GGG GATACAAGCTGGGGTTC-3', reverse 5'-GATTAGCCGTCTGCCCTCAT-3'.

HKU1-spike-TMPRSS2 video microscopy

One million 293T cells were transfected with either 1 μg of pCDEST-TMPRSS2-Scarlet-I or phCMV-HKU1A-S-NeonGreen plasmid in a six-well plate using lipofectamine 2000 according to the manufacturer's protocol. At 24 h posttransfection, cells were harvested with phosphate buffered saline (PBS) + 0.1% ethylenediaminetetraacetic acid (EDTA) and resuspended at 6×10^5 cells per ml in DMEM 10% FBS. The cells were mixed in suspension at a 50:50 ratio, and 100 μl were plated per well in a 96-well plate (uClear, no. 655090). 30 min after plating, cells were imaged every 150 s for 2.5 h at 37 °C, 5% CO₂ on an Opera Phenix High-Content Screening System (PerkinElmer).

Pseudovirus generation and infection

Pseudoviruses were produced by transfection of 293T cells as previously described⁵⁰. Briefly, cells were cotransfected with plasmids encoding for lentiviral proteins, a luciferase reporter and the HKU1 spike plasmid. Pseudotyped virions were harvested two and three days after transfection. Production efficacy was assessed by measuring infectivity or HIV Gag p24 concentration. For transfected cells (293T), infection was performed 24 h posttransfection, in suspension using 10 ng of p24 and 2×10^4 cells in 100 μl . For the stable cell lines, infection was performed on plated cells, with or without spinoculation (2 h at 2,000g), using 10–50 ng of p24 per well and 2×10^4 cells in 100 μl in a 96-well plate. The next day, 100 μl of media was added. Then, 48 h postinfection, 100 μl of media was carefully removed and 100 μl of Bright-Glo lysis buffer (ProMega) was added. After 10 min, luminescence was acquired using the EnSpire in a white plate (PerkinElmer).

Isolation and characterization of a live HKU1 virus

Origin of the sample. Respiratory material of sample MAL21 0303 was obtained in the winter of 2022–2023 at Amsterdam UMC. Material was collected in Universal Transport Medium (Copan) and stored at –80 °C. Details on the disease of the patient are unknown as materials were donated for research anonymously. Presence of HCoV-HKU1 was determined by means of real-time qPCR as described⁵¹ with primers HKU1 forward, 5'-TCCTACTAYTCAAGAAGCTATCC-3', HKU1 reverse 5'-AATGAACGATTATTGGGTCCAC-3' and HKU1-probe CY5BHQ2 TYCGCCTGGTACGATTTTGCTCA.

HKU1 isolation and culture on human nasal epithelium. MucilAir, reconstructed human nasal epithelium cultures that had been differentiated for four weeks, were purchased from Epithelix and cultured in 700 μl MucilAir media on the basal side of the air/liquid interface cultures and monitored for healthy cilia movements. Cultures were

Article

kept at 34 °C under a 5% CO₂ atmosphere. One hour before infection, mucus was removed from the apical side of the culture by washing with Hanks' Balanced Salt Solution (HBSS) (Gibco). For viral isolation, cells were infected with nasal swabs diluted at a 1:4 ratio in Universal Transport Medium (Copan) for 2 h at 34 °C. Viral input was removed and cells were washed three times with 200 µl of HBSS (Gibco). The apical side was harvested twice every 24 h for 7 days by adding 200 µl HBSS for 10 min at 34 °C, collecting the liquid on the apical side and diluting it (1:1) in Universal Transport Medium (Copan). The harvest was centrifuged at 1,000g for 5 min to remove debris, 50 µl was used for qPCR, and the remaining samples were stored at -80 °C. Passage 1 and passage 2 viral stocks were sequenced with identical results.

Sequencing of the HKU1 strain. RNA was extracted from the cell supernatant using the QIAamp Viral RNA Kit (Qiagen) following the manufacturer's instructions. Extracted RNA was treated with Turbo DNase (Ambion) followed by ribosomal RNA⁵². RNA was reverse-transcribed into double-stranded cDNA using random hexamers and libraries prepared using the Nextera XT kit (Illumina), before sequencing on an Illumina NextSeq500 (2 × 75 cycles). Raw sequence data (human reads removed) were deposited in the European Nucleotide Archive portal (<https://www.ebi.ac.uk/ena/>) under bioproject accession number PRJEB64017. To determine the sequence of the acidic tandem repeat region of the genome, the region was amplified using external primers (HKU1_ATR_L1 5'-ATGAAGCAATGGCCTCTCGT-3' and HKU1_ATR_R1 5'-CACAGAACGCAACCAACAGT-3') before Sanger sequencing.

Genome assembly. Adaptors and low-quality sequences of raw reads were removed using Trimmomatic v.0.39 (ref. 53). We assembled the trimmed reads using megahit v.1.2.9 (ref. 54) with default parameters. The contigs were queried against the NCBI non-redundant protein database using DIAMOND v.2.0.4 (ref. 55) to look for potential contaminants in addition to the detected HKU1 genome. The Sanger data was used with the assembled contigs to generate the final HKU1 scaffold on which the trimmed reads were mapped to generate the final consensus (Extended Data Fig. 8). The mapping data were visually checked to confirm the accuracy of the obtained genome using Geneious Prime 2023. The sequence of the isolated virus was deposited in GenBank, ID HCoV-HKU1/NDL/IPP01/2022, accession number OR260091.

Phylogenetic and recombination analysis. All available complete HKU1 genome sequence data and metadata were retrieved from the Bacterial and Viral Bioinformatics Resource Center⁵⁶ (<https://www.bv-brc.org/>) in June 2023. Sequences were aligned using MAFFT v.7.467 (ref. 57), and the alignment was checked for accuracy using BioEdit v.7.2.5. We used a combination of six methods implemented in RDP4 (ref. 58) (RDP, GENECONV, MaxChi, Bootscan, SisScan and 3SEQ) to detect potential recombination events in the newly reported genome. The ModelFinder application⁵⁹, as implemented in IQ-TREE v.2.0.6 (ref. 60), was used to select the best-fitting nucleotide substitution model, and maximum-likelihood phylogenies were inferred using complete genomes or the spike coding sequences. Branch support was calculated using ultrafast bootstrap approximation with 1,000 replicates⁶¹. The phylogenies were visualized using the auspice module from Nextstrain⁶². Interactive phylogenies are available at <https://github.com/Simon-LorierLab/HKU1>.

Analysis of HKU1 replication by real-time qPCR

Viral RNA was extracted from culture supernatants using the Quick-DNA/RNA Viral 96 Kit (Zymo) according to manufacturer's instructions. qPCR was run using the Luna Universal Probe One-Step kit on a QuantStudio 6 (Thermo Fisher Scientific) according to manufacturer's instructions using the following primers: HKUqPCR5, 5'-CTGGTACGATTTTGCTCAA-3' and HKUqPCR3,

5'-ATTATTGGGTCCACGTGATTG-3' (ref. 28) and TaqMan QSY probe 5'-FAM-TTGAAGGCTCAGGAAGGCTGCTCTCTAA-QSY7-3'. A DNASTring was used to generate a standard curve (5'-GGATCTACTATTCAAGAAGCTATCCCTACTAGGTTTTTCGCTGGTACGATTTTGCCTCAAGGCTATTATGTTGAAGGCTCAGGAAGGCTGCTCTCTAATAGCCGGCCAGGTTTCAGTTCTCAATCAGCTGGACCAATAATCGTTTCATTAAAGTAGAAGTAATTCTAATTTAGACATTCTGATTCTATAGTAAACCTG-3').

Human nasal epithelium VHH inhibition experiments

One hour before infection, mucus was removed from the apical side of the culture by washing with HBSS (Gibco). For viral-infectivity assays, cells were incubated for 30 min with 50 µl of universal transport medium (Yocon) containing monomeric non-target VHH93 (also called N-G9-3 (ref. 63)) or anti-TMPRSS2 VHH A07 at 5 µM. Then 5 × 10⁶ viral copies were added in 50 µl of universal transport medium for 2 h at 34 °C (final VHH concentration 2.5 µM). Viral input was removed and cells were washed three times with 200 µl of HBSS (Gibco). After 48 h, cells were fixed on the apical and basal sides with 4% PFA for 30 min. For imaging, fixed cells were stained intracellularly with mAb10 at 1 µg ml⁻¹, a pan-anti-coronavirus spike antibody²⁹, anti-TMPRSS2 VHH A01-Fc at 6 µg ml⁻¹ and phalloidin Atto-647 (Sigma-Aldrich) at 1:200, as described previously⁶⁴, and imaged using the LSM-700 confocal microscope (Zeiss).

Image quantification

To quantify the effect of the nanobodies on HKU1 infection of epithelia, the fluorescence intensity of spike-positive pixels of five fields (160 µm² each) spread over the surface of the sample was measured using ImageJ-Fiji.

CRISPR-Cas9 KO

crRNAs TMPRSS2 (CGGATGCACCTCGTAGACAG) and predesigned unspecific crRNA used as control were ordered from Integrated DNA Technologies (IDT). crRNA and tracrRNA were resuspended in IDT Duplex Buffer according to the manufacturer's instructions. On the day of the nucleofection, duplexes were formed by mixing equimolar concentrations of crRNA and tracrRNA, followed by 5 min annealing at 95 °C. Then, 100 pmol of RNA duplexes were mixed (1:2) with 50 pmol TrueCut Cas9 Protein v.2 (Thermo Fisher Scientific) for 10 min in real time to generate ribonucleoprotein complexes. Caco2 cells were resuspended in SE Cell Line buffer solution (Lonza), mixed with ribonucleoprotein and Alt-R Cas9 Electroporation Enhancer (90 pmol, IDT) and nucleofected in a 4D-Nucleofector System (Lonza) using the SE Cell line 4D-Nucleofector X Kit S (programme DG-113). After nucleofection, cells were seeded in DMEM 10% FBS. 48 h posttransfection, cells were subcloned. Clones were screened for TMPRSS2 KO by staining using VHH A01-Fc and flow cytometry (immunostaining section). KO was confirmed by PCR and sequencing the target region using the following primers: forward 5'-AAGACGGAGGAGAAGGGTCA-3' and reverse 5'-AGTTGTAGACACCTAGGGAGAA-3'.

Protein production

Construct design. Spike, RBD, TMPRSS2 and ACE2 constructs were obtained from Genscript as codon-optimized synthetic genes. Ectodomains from HKU1A (residues 14–1281) and B (residues 14–1276) were cloned into pcDNA3.1(+), downstream of a murine Ig kappa signal peptide (METDTLLLWVLLWVPGSTG) and upstream of a thrombin cleavage site followed by a His tag. Spikes were stabilized by mutating the furin cleavage site (⁷⁵⁶RRKRR⁷⁶⁰⁻⁷⁵⁶GGSGS⁷⁶⁰ in HKU1A; ⁷⁵²RRKRR⁷⁵⁶⁻⁷⁵²GGSGS⁷⁵⁶ in HKU1B) and two residues in the S2 subunit (¹⁰⁷¹AL¹⁰⁷²⁻¹⁰⁷¹PP¹⁰⁷² in HKU1A; ¹⁰⁶⁷NL¹⁰⁶⁸⁻¹⁰⁶⁷PP¹⁰⁶⁸ in HKU1B) and adding a Foldon trimerization motif at the C terminus. The ectodomain from the Wuhan SARS-CoV-2 spike (residues 1–1208) was cloned into pcDNA3.1(+), and stabilized with six proline mutations (F817P, A892P, A899P, A942P, K986P, V987P), as reported⁶⁵. The furin site was also

replaced as above (⁶⁸²RRAR⁶⁸⁵⁻⁶⁸²GSAS⁶⁸⁵) and a C-terminal Foldon motif was introduced, as well as Hisx8, Strep and Avi tags.

The RBDs (residues 323–609 for HKU1A, 323–607 for HKU1B, 331–528 for SARS-CoV-2 Wuhan) were cloned into pCAGGS (HKU1) or pcDNA3.1(+) (SARS-CoV-2) following a murine immunoglobulin kappa signal peptide and upstream of a thrombin cleavage site and in tandem Hisx8, Strep and Avi tags. The WT TMPRSS2 ectodomain (residues 107–492) followed by C-terminal tags (8xHis-tag and AviTag) was synthesized and cloned into pcDNA3.1/Zeo(+) expression vector (Thermo Fisher Scientific). The TMPRSS2 ectodomain with the S441A mutation was cloned into a modified pMT/BIp plasmid (Invitrogen, hereafter called pT350), which translates the protein in frame with an enterokinase cleavage site and a double Strep tag at the C-terminal end. The ACE2 peptidase domain (residues 19–615) was cloned in pcDNA3.1(+) with a murine immunoglobulin kappa signal peptide and a C-terminal thrombin cleavage site followed by a Strep tag. The coding sequences of the selected VHHs in the vector pHEN6 were subcloned into a bacterial expression vector pET23 encoding a C-terminal His tag using NcoI and NotI restriction sites.

Protein expression and purification. RBD, spike and ACE2-encoding plasmids were transiently transfected into Expi293F cells (Thermo Fisher Scientific) using FectoPro DNA transfection reagent (Poly-Plus) or the polyethylenimine precipitation method, as previously described⁶⁶. After five days at 37 °C, cells were harvested by centrifugation and the supernatants were concentrated. The spike proteins used for flow cytometry were purified from culture supernatants by high-performance chromatography using Ni Sepharose Excel Resin according to the manufacturer's instructions (GE Healthcare) and dialysed against PBS using Slide-A-Lyzer dialysis cassettes (Thermo Fisher Scientific). The spike proteins used for ELISA were further purified by size-exclusion chromatography (SEC) on a Superose6 10/300 column (Cytiva). Eluted fractions were analysed with sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE), and those containing bands of the expected molecular weight were pooled, concentrated and further purified by SEC on a Superdex 200 10/300 column (Cytiva). AviTagged SARS-CoV-2 tri-S used for flow cytometry was biotinylated using an Enzymatic Protein Biotinylation Kit (Sigma-Aldrich). HKU1A and HKU1B spike proteins used for ELISA were biotinylated using an EZ-Link Sulfo-NHS-Biotinylation Kit (Thermo Fisher Scientific). The HKU1B RBD used for the BLI experiment with nanobodies was biotinylated with the BirA Biotin-Protein Ligase Standard Reaction Lit (Avidity).

The pT350 plasmid encoding S441A TMPRSS2 was used to perform a stable transfection on *Drosophila* S2 cells with the pCoPuro plasmid for puromycin selection. The cell line was selected and maintained in serum-free insect cell medium (HyClone, GE Healthcare) containing 7 µg ml⁻¹ puromycin and 1% penicillin/streptomycin. For protein production, the cells were grown in spinner flasks until the density reached 10⁷ cells per ml, at which point the protein expression was induced with 4 µM CdCl₂. After six days, the culture was centrifuged and the supernatant was concentrated and used for affinity purification using a Strep-Tactin column (IBA). The eluate was concentrated and applied onto a Superdex 200 16/60 column (Cytiva) equilibrated with 10 mM Tris-HCl (pH 8.0), 100 mM NaCl.

E. coli BL21pLysS cells were transformed with the plasmids encoding the different VHHs, which were expressed in the cytoplasm after overnight induction with 0.5 mM isopropyl β-D-1-thiogalactopyranoside at 16 °C. The cultures were centrifuged, the bacterial pellets were resuspended in 40 ml of lysis buffer (20 mM Tris-HCl, 200 mM NaCl, 20 mM imidazole, pH 8.0) containing cOmplete Protease Inhibitor Cocktail (Roche), and they were frozen at –80 °C until used. On the purification day, the resuspended pellets were thawed, sonicated (15 min, 9 s on-pulse, 5 s off-pulse), centrifuged and loaded onto a HisTrap column. Bound proteins were eluted with a linear gradient of buffer B (20 mM

Tris-HCl, 200 mM NaCl, 500 mM imidazole, pH 8.0) and analysed by SDS–PAGE. Fractions with higher purity were pooled, concentrated and further purified by SEC on a Superdex 75 16/60 column (Cytiva) pre-equilibrated in 10 mM Tris-HCl, 100 mM NaCl, pH 8.0.

The purity of the final protein samples was analysed by SDS–PAGE followed by Coomassie Blue staining or silver staining. For gels, see Supplementary Information 2.

Flow cytometry

For spike binding, 293T cells were transiently transfected with TMPRSS2 and maintained in the presence or absence of Camostat (10 µM) for 24 h. The cells were incubated with soluble biotinylated spike diluted in magnetic-activated cell sorting (MACS) buffer (PBS, 5 g/L BSA, 2 mM EDTA) at 2 µg ml⁻¹ for 30 min at 4 °C. The cells were washed twice with PBS and then incubated with Alexa Fluor 647-conjugated streptavidin (Thermo Fisher Scientific, S21374, 1:400) for 30 min at 4 °C. The cells were washed once with PBS and fixed with 4% paraformaldehyde. The results were acquired using an Attune Nxt Flow Cytometer (Life Technologies, software v.3.2.1). Transfection efficiency for TMPRSS2 was assessed using a commercial anti-TMPRSS2 antibody (for the experiments performed at the initiation of the study), the anti-TMPRSS2 monomeric VHH or the dimeric VHH A01-Fc (see below).

Staining with the commercial anti-TMPRSS2 antibody was performed on fixed cells by staining intracellularly with rabbit anti-TMPRSS2 (Atlas HPA035787), for 30 min in real time in PBS/BSA/Azide/0.05% Saponin followed by a Alexa Fluor 647 Goat anti-Rabbit Antibody (Thermo Fisher Scientific, A-21245, 1:500).

For the spike, transfection efficiency was measured at the surface of live cells using mAb10 diluted in MACS buffer for 30 min at 4 °C and a human secondary IgG. mAb10 is an antibody generated from a SARS-CoV-2 infected patient that cross-reacts with HKU1 (ref. 29).

Surface expression of TMPRSS2 was assessed on live cells by staining with anti-TMPRSS2 VHH A01-Fc (described in the study) at 1 µg ml⁻¹ for 30 min at 4 °C in MACS buffer, followed by staining with Alexa Fluor 647-conjugated Goat anti-Human Antibody (Thermo Fisher Scientific, A-21445, 1:500). The control VHH-Fc ctrl (R3VQFc) recognizes an unrelated protein (phosphorylated Tau protein)⁶³.

The monomeric VHHs were used at 0.5 µg ml⁻¹ in MACS buffer for 30 min at 4 °C, and staining was revealed using a coupled Anti-His Antibody (R&D Systems, IC0501R, 1:1,000).

Transfection efficiency for myc-tagged constructs was assessed on fixed cells by staining intracellularly with mouse anti-c-myc antibody, clone 9E10 (Thermo Fisher Scientific, M4439) for 30 min in real time in PBS/BSA/Azide/0.05% Saponin followed by an Alexa Fluor 647-conjugated Goat anti-Mouse Antibody (Thermo Fisher Scientific, A-21242, 1:500). Transfection efficiency of aminopeptidase-N (CD13) and dipeptidyl peptidase-4 (CD26) was assessed on live cells by surface staining in MACS buffer with CD13-PE (130-120-312, Miltenyi Biotec, 1:50) and CD26-PE (130-126-41, Miltenyi Biotec, 1:50) for 30 min at 4 °C. The cells were washed twice with PBS and fixed with 4% paraformaldehyde. The results were acquired using an Attune Nxt Flow Cytometer (Life Technologies). Gating strategies are described in Supplementary Information 3.

Sialic acid staining

Cells were harvested using PBS/0.1% EDTA for 5 min at 37 °C, washed in PBS and stained 1 h at 4 °C in PBS/1% SVF containing either 10 µg ml⁻¹ Sambucus Nigra Lectin-FITC (Lectin NSA) (Thermo Fisher Scientific, L32479) or 2.5 µg ml⁻¹ recombinant mouse Siglec-E-Fc (BioLegend, 551504). Cells were washed twice in PBS. Lectin-stained cells were fixed for 10 min in 4% paraformaldehyde. Siglec-E-Fc stained cells were further incubated with Alexa Fluor 647-conjugated Goat anti-Human Antibody (Thermo Fisher Scientific, A-21445, 1:500) for 30 min at 4 °C before being fixed for 10 min in 4% paraformaldehyde.

ELISA assay

ELISAs were performed as previously described⁶⁷. Briefly, high-binding 96-well ELISA plates (Costar, Corning) were coated overnight with 250 ng per well of purified TMPRSS2 or ACE2. After washing with 0.05% Tween 20-PBS (washing buffer), the plates were blocked for 2 h with 2% BSA, 1 mM EDTA, 0.05% Tween 20-PBS (blocking buffer), washed and then incubated with serially diluted soluble biotinylated spike proteins. Recombinant spike proteins were tested at 100 $\mu\text{g ml}^{-1}$ and at seven consecutive 1:2 dilutions in PBS. After washings, the plates were revealed by incubation for 1 h with HRP-conjugated streptavidin (BD Biosciences) in blocking buffer and by adding 100 μl of HRP chromogenic substrate (ABTS solution, Euromedex) after washing steps. Optical densities were measured at 405 nm ($\text{OD}_{405\text{nm}}$), and background values given by incubation of PBS alone in coated wells were subtracted. Experiments were performed using a HydroSpeed microplate washer and a Sunrise microplate absorbance reader (Tecan).

BLI assay

Affinity of recombinant RBDs towards the purified ectodomains of S441A TMPRSS2 or ACE2 was assessed in real time using a Bio-Layer Interferometry Octet RED384 device (Pall ForteBio). Nickel-NTA capture sensors (Sartorius) were loaded for 10 min at 1,000 rpm shaking speed with the Wuhan RBD at 100 nM or the HKU1A/B RBDs at 200 nM in PBS. The sensors were then blocked with PBS containing BSA at 0.2 mg ml^{-1} (assay buffer) and were incubated at 1,000 rpm with twofold serially diluted concentrations (800 nM to 3.12 nM) of S441A TMPRSS2 or ACE2 ectodomains in assay buffer. Association and dissociation were monitored for 300 s and 240 s, respectively. Measurements for a reference sensor were recorded using a sensor loaded with an unrelated protein (CD147) that was dipped at each analyte concentration. A sample reference measurement was recorded from a sensor loaded with either RBD and dipped in the assay buffer. Specific signals were calculated by double referencing, subtracting non-specific signals obtained for the sensor and sample references from the signals recorded for the RBD-loaded sensors dipped in S441A TMPRSS2 solutions. The steady-state signal was plotted against the analyte concentration, and the curve was fitted assuming a 1:1 binding model.

The affinity of the nanobodies for the S441A TMPRSS2 ectodomain was determined using a similar procedure. Nickel-NTA capture sensors were loaded with each nanobody at 100 nM in PBS and then blocked with assay buffer and incubated at 1,000 rpm with twofold serially diluted concentrations (400 nM to 3.12 nM) of S441A TMPRSS2. Association and dissociation were monitored for 240 s and 180 s, respectively. A sample reference measurement was recorded from a sensor loaded with each nanobody and dipped in the assay buffer. Association and dissociation profiles were fitted assuming a 1:1 binding model.

Experiments to identify anti-TMPRSS2 nanobodies that block binding to the HKU1 RBD were performed by immobilizing the biotinylated HKU1B RBD on streptavidin capture sensors. They were blocked in assay buffer and dipped into solutions containing a pre-incubated mixture of TMPRSS2 S441A (200 nM) and a nanobody (400 nM). The signal corresponding to the association was recorded.

TMPPRSS2 enzymatic activity

Enzymatic activity was measured using Boc-QAR-AMC (R&D Systems, ES014), a substrate of TMPRSS2 that fluoresces when cleaved. For WT and mutant TMPRSS2, cells were transfected in a black-bottom 96-well plate as described above. After 24 h, the medium was replaced with 100 μl FBS-free, phenol-red-free medium containing 100 μM of fluorogenic substrate. Indicated concentrations of inhibitors were added. When the assay was performed with soluble TMPRSS2, 60 nM of soluble protein was added to the well and mixed with VHH nanobodies for 15 min, and then 100 μM of fluorogenic substrate was added.

Western blot

Cells were lysed in TXNE buffer (1% Triton X-100, 50 mM Tris-HCl (pH 7.4), 150 mM NaCl, 5 mM EDTA, 1X Roche cOmplete protease inhibitors) for 30 min on ice. Equal amounts (10 μg) of cell lysates were analysed by western blot. The following antibodies were diluted in western blot buffer (PBS, 1% BSA, 0.05% Tween, 0.01% Na Azide): Rabbit anti-Human TMPRSS2 (Atlas Antibodies, HPA035787, 1:1,000), Mouse anti-Beta Actin (Abcam, 60008-1-Ig, 1:2,000), Rabbit anti-HKU1 S1 Polyclonal Antibody (Thermo Fisher Scientific, PA5-120768, 1:2,000) and Rabbit anti-HKU1 S2 Polyclonal Antibody (Thermo Fisher Scientific, PA5-120769, 1:1,000). Species-specific secondary DyLight-coupled antibodies were used (1:10,000), and proteins were revealed using a LI-COR imager. Images were processed using Image Studio Lite software v.5.2.5.

Nanobody isolation and production

Alpaca immunization. Animal procedures were performed according to French legislation and in compliance with the European Communities Council Directives (2010/63/UE, French Law 2013-118, 6 February 2013). The Animal Experimentation Ethics Committee of the Pasteur Institute (CETEA 89) approved this study (2020-27412). One young adult male alpaca (*Lama pacos*) was immunized at days 0, 17 and 24 with 150 μg of S441A TMPRSS2. The immunogen was mixed with Freund complete adjuvant for the first immunization and with Freund incomplete adjuvant for the following immunizations. The immune response was monitored by titration of serum samples by ELISA on coated TMPRSS2. The bound alpaca antibodies were detected with polyclonal rabbit anti-alpaca IgG⁶⁸.

Library construction and phage display. The blood of the immunized animal (about 300 ml) was collected, and peripheral blood lymphocytes were isolated by centrifugation on a Ficoll (Cytiva) discontinuous gradient and stored at -80°C until further use. Total RNA and cDNA were obtained as previously described⁶⁸. The VHH repertoires were amplified from the cDNA by two successive PCR reactions, and the VHH fragments were cloned into the SfiI/NotI restriction sites of the pHEN6 phagemid vector⁶³. The selection of specific phage-VHVs was performed by phage display. Many phage-VHVs (10^{13}) were used to perform three rounds of panning. Briefly, Phage-VHVs were incubated for 1 h with TMPRSS2 that was previously coated on an immunotube (Nunc). To remove non-specific binders, an extensive washing procedure was performed, and specific phage-VHVs were eluted in 100 mM triethylamine. *E. coli* TG1 at exponential growth phase were then infected with eluted phage-VHVs. Phage-VHVs were produced from individual colonies, and binding of the phages to TMPRSS2 on plate was revealed with an anti-M13 monoclonal antibody conjugated to peroxidase (Abcam). The VHH nucleotide sequences were determined using M13-40 primer (Eurofins).

Production of VHVs. The coding sequences of the selected VHVs in the vector pHEN6 were subcloned into a bacterial expression vector pET23 encoding a C-terminal His tag using NcoI and NotI restriction sites. Sequences of the VHH are available in Supplementary Information 4. Transformed *E. coli* BL21pLysS cells expressed VHVs in the cytoplasm after overnight induction with 0.5 mM isopropyl β -D-1-thiogalactopyranoside at 16°C . Purified VHVs were isolated on Co⁺⁺ affinity columns from cytoplasmic extracts treated with 10 U ml^{-1} Benzonase Nuclease (Merck) and Complete protease inhibitor (Roche) according to the manufacturer's instructions, followed by SEC with a Superdex 75 column (Cytiva).

Production of dimeric VHH-Fc. The VHVs' engineered genes were cloned into a pFUSE-derived vector (InvivoGen); this vector harbours a human IgG1-Fc domain. Consequently, the VHH was expressed as a Fc-fusion bivalent antibody. The vector was used to transform Expi293F mammalian cells (Thermo Fisher Scientific), and protein expression

was carried out according to the manufacturer's recommendations. Protein was purified from the expression medium by affinity chromatography on a 1 ml protein G column (Cytiva). After sample application, the column was washed with 20 column volumes of PBS, and the protein was subsequently eluted with 10 column volumes of PBS supplemented with 0.1 M Glycine (pH 2.3). Affinity-eluted VHH-Fcs were finally polished on a HiLoad 16/600 Superdex 200 pg prepacked column (Cytiva) using PBS buffer⁶⁹.

Statistical analysis. Flow cytometry data were analysed with FlowJo v.10.8 software (BD Life Science). Calculations were all performed with Microsoft Excel 365. GraphPad Prism 9 for Mac was used to generate figures and for statistical analysis (GraphPad Software). Statistical significance between different conditions was calculated using the tests indicated in the corresponding figure legends.

Reporting summary

Further information on the research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Raw sequence data (human reads removed) for the HKU1 virus were deposited in the European Nucleotide Archive portal (<https://www.ebi.ac.uk/ena/>) under bioproject accession number PRJEB64017. The raw data of the main figures are available in the Supplementary Information. The raw data of the Extended Data are available from the corresponding authors on reasonable request. Source data are provided with this paper.

46. Kodaka, M. et al. A new cell-based assay to evaluate myogenesis in mouse myoblast C2C12 cells. *Exp. Cell. Res.* **336**, 171–181 (2015).
47. Survey of human chromosome 21 gene expression effects on early development in *Danio rerio*. *G3* **8**, 2215–2223 (2018).
48. Hayer, A. et al. Engulfed cadherin fingers are polarized junctional structures between collectively migrating endothelial cells. *Nat. Cell Biol.* **18**, 1311–1323 (2016).
49. Rajah, M. M. et al. SARS-CoV-2 Alpha, Beta, and Delta variants display enhanced spike-mediated syncytia formation. *EMBO J.* **40**, e108944 (2021).
50. Iglesias, M. C. et al. Lentiviral vectors encoding HIV-1 polyepitopes induce broad CTL responses in vivo. *Mol. Ther.* **15**, 1203–1210 (2007).
51. Loens, K. et al. Performance of different mono- and multiplex nucleic acid amplification tests on a multipathogen external quality assessment panel. *J. Clin. Microbiol.* **50**, 977–987 (2012).
52. Matranga, C. B. et al. Enhanced methods for unbiased deep sequencing of Lassa and Ebola RNA viruses from clinical and biological samples. *Genome Biol.* **15**, 519 (2014).
53. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
54. Li, D. et al. MEGAHIT v1.0: a fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods* **102**, 3–11 (2016).
55. Buchfink, B., Xie, C. & Huson, D. H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **12**, 59–60 (2015).
56. Olson, R. D. et al. Introducing the Bacterial and Viral Bioinformatics Resource Center (BV-BRC): a resource combining PATRIC IRD and ViPR. *Nucleic Acids Res.* **51**, D678–D689 (2023).
57. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
58. Martin, D. P., Murrell, B., Khoosal, A. & Muhire, B. in *Bioinformatics: Volume I: Data, Sequence Analysis, and Evolution* (ed. Keith, J. M.) Ch. 17 (Springer, 2017).
59. Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A. & Jermiin, L. S. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589 (2017).

60. Nguyen, L.-T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015).
61. Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q. & Vinh, L. S. UFBoot2: improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* **35**, 518–522 (2017).
62. Hadfield, J. et al. Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics* **34**, 4121–4123 (2018).
63. Gransagne, M. et al. Development of a highly specific and sensitive VHH-based sandwich immunoassay for the detection of the SARS-CoV-2 nucleoprotein. *J. Biol. Chem.* **298**, 101290 (2022).
64. Robinot, R. et al. SARS-CoV-2 infection induces the dedifferentiation of multiciliated cells and impairs mucociliary clearance. *Nat. Commun.* **12**, 4354 (2021).
65. Hsieh, C.-L. et al. Structure-based design of prefusion-stabilized SARS-CoV-2 spikes. *Science* **369**, 1501–1505 (2020).
66. Lorin, V. & Mouquet, H. Efficient generation of human IgA monoclonal antibodies. *J. Immunol. Methods* **422**, 102–110 (2015).
67. Mouquet, H. et al. Memory B cell antibodies to HIV-1 gp140 cloned from individuals infected with clade A and B viruses. *PLoS ONE* **6**, e24078 (2011).
68. Lafaye, P., Achour, I., England, P., Duyckaerts, C. & Rougeon, F. Single-domain antibodies recognize selectively small oligomeric forms of amyloid beta, prevent Abeta-induced neurotoxicity and inhibit fibril formation. *Mol. Immunol.* **46**, 695–704 (2009).
69. Li, Q. et al. Generation of nanobodies acting as silent and positive allosteric modulators of the $\alpha 7$ nicotinic acetylcholine receptor. *Cell. Mol. Life Sci.* **80**, 164 (2023).
70. Grubaugh, N. D. et al. An amplicon-based sequencing framework for accurately measuring intrahost virus diversity using PrimalSeq and iVar. *Genome Biol.* **20**, 8 (2019).

Acknowledgements We thank A. Amara, T. Bruel and N. Casartelli for critical reading of the manuscript, members of the Virus and Immunity Unit for discussions and help, N. Aulner, N. Mahtal and the UtechS Photonic Bioluminescence core facility (Institut Pasteur), P. Charneau for help in lentiviral pseudotype preparation and S. Pöhlmann for the kind gift of plasmids. N.S. is funded by the Ministère de l'Enseignement supérieur et de la Recherche. O.S.'s laboratory is funded by Institut Pasteur, Fondation pour la Recherche Médicale, the National Agency for AIDS Research—Emerging Infectious Diseases, the Vaccine Research Institute (ANR-10-LABX-77), Humanities in the European Research Area programme (DURABLE consortium), Labex Integrative Biology of Emerging Infectious Diseases (IBEID) (ANR-10-LABX-62-IBEID), Agence Nationale de la Recherche (ANR)/Fondation pour la Recherche Médicale Flash Covid PROTEO-SARS-CoV-2 and IDISCOVER. E.S.-L.'s lab is funded by the INCEPTION programme (Investissements d'Avenir grant no. ANR-16-CONV-0005), the National Institutes of Health Pasteur International Center for Research on Emerging Infectious Diseases programme (award no. U01AI151758), the Health Emergency Preparedness and Response European programme DURABLE (101102733) and the Labex IBEID (ANR-10-LABX-62-IBEID). H.M.'s laboratory is funded by the Institut Pasteur, the Milieu Intérieur Programme (ANR-10-LABX-69-01) and l'Institut national de la santé et de la recherche médicale, REACTing and European Union (Rapid European COVID-19 Emergency Response (RECOVER)) grants. M.M.R. was supported by the Pasteur-Paris University International Doctoral Programme and the Institut Pasteur Department of Virology 'Bourse de Soudure' fellowship. F.R.'s laboratory is funded by ANR grant nos. ANR-13-ISV8-0002-01 and ANR-10-LABX-62-10 IBEID, Wellcome Trust collaborative grant no. UNS22082 and the Institut Pasteur and the Centre national de la recherche scientifique. Work with UtechS Photonic Bioluminescence is funded by grant no. ANR-10-INSEB-04-01 and Région Ile-de-France programme DIM1-Health. The funders of this study had no role in study design, data collection, analysis and interpretation or writing of this Article.

Author contributions The experimental strategy was designed by N.S., I.F., C.P., M.M.R., H.M., F.R., J.B. and O.S. Data acquisition and analysis were performed by N.S., I.F., C.P., V.M., M.M.R., E.B.S., F.P., F.G.-B., J.P., C.B., A. Martin, L.G., R.M.O., A. Meola, O.A., H.H.-W., M.P., D.D., E.S.-L. and J.B. Vital materials and expertise were provided by G.C.-L.F., M.C., M.D., V.M., L.V.D.H. and P.L. The manuscript was written by N.S., J.B. and O.S. and edited by N.S., I.F., C.P., E.B.S., H.M., F.R., P.L., J.B. and O.S. All authors read and approved the final version of the manuscript.

Competing interests N.S., J.B., O.S., I.F., E.B.S., F.R. and P.L. have a provisional patent on anti-TMPRSS2 nanobodies. The other authors declare no competing interests.

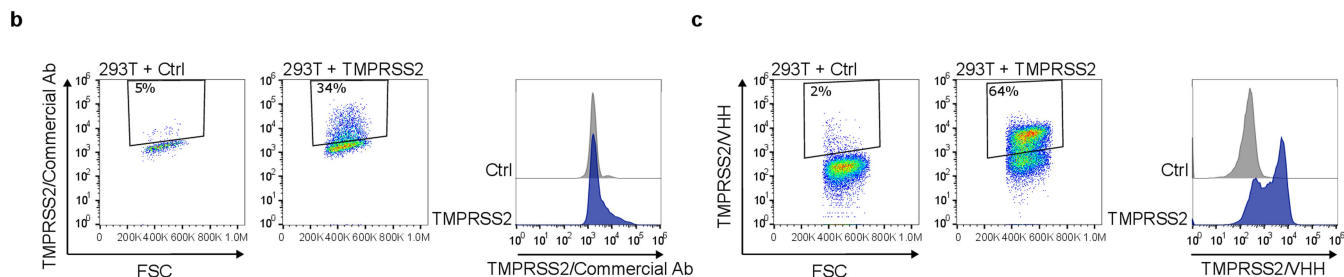
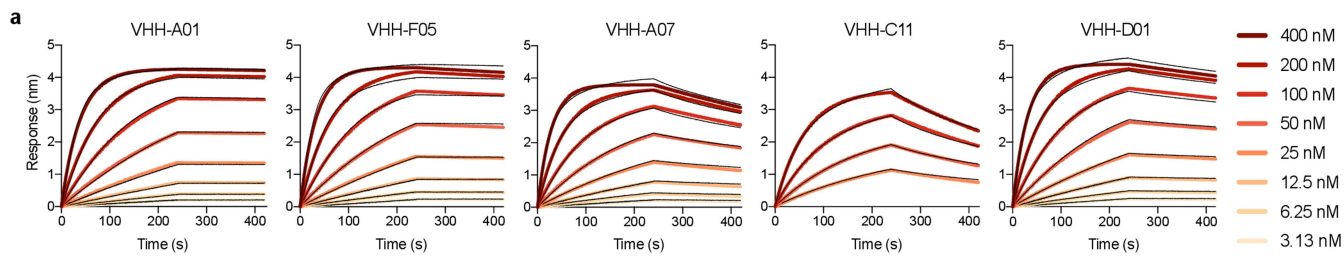
Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41586-023-06761-7>.

Correspondence and requests for materials should be addressed to Julian Buchrieser or Olivier Schwartz.

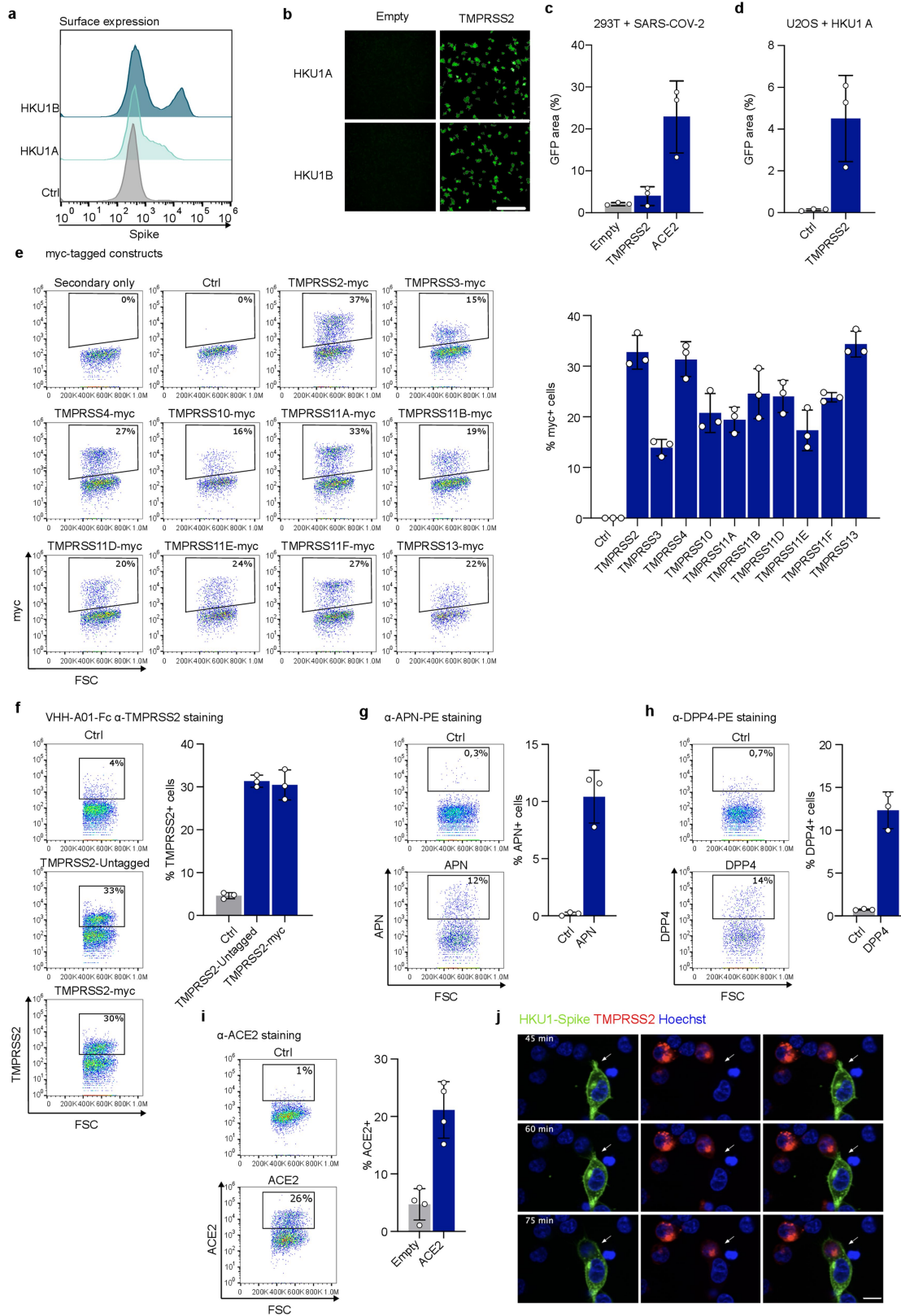
Peer review information Nature thanks Zhengli Shi, Hyeryun Choe and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at <http://www.nature.com/reprints>.



Extended Data Fig. 1 | a. Binding of the 5 anti-TMPRSS2 VHH on recombinant TMPRSS2 measured by BLI. b, c. Comparison of one commercial TMPRSS2 antibody and of the dimeric anti-TMPRSS2 VHH A01-Fc. 293 T cells were

transfected with WT TMPRSS2. 24 h later, cells were analysed by flow-cytometry. Staining with the commercial antibody was performed on fixed cells (b) while surface staining with the VHH was performed on live cells (c).



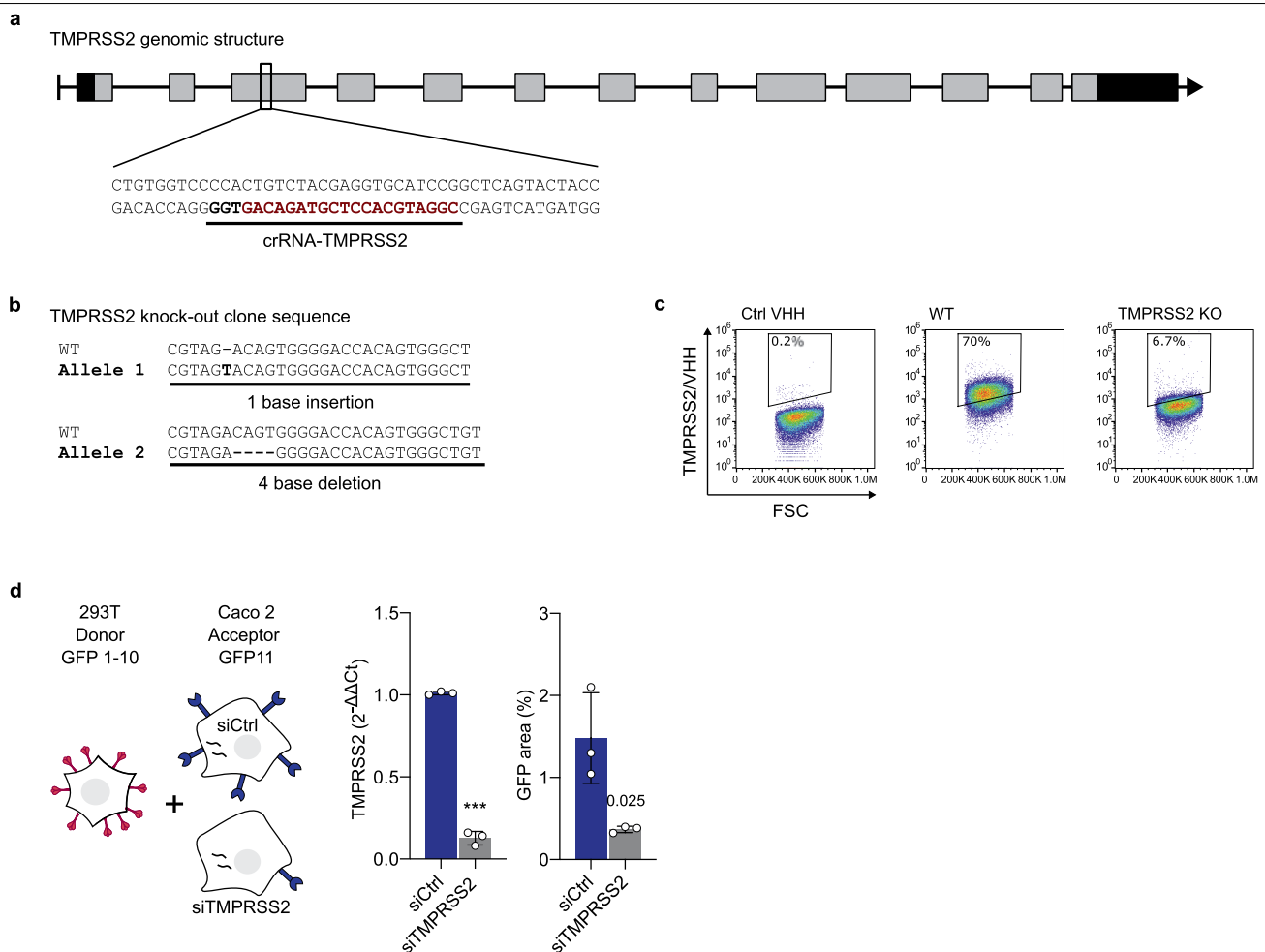
Extended Data Fig. 2 | See next page for caption.

Article

Extended Data Fig. 2 | Effect of TMPRSS2 and other proteases on HKU1 cell-cell fusion. **a, b. Surface expression and fusogenic activity of HKU1 spikes.**

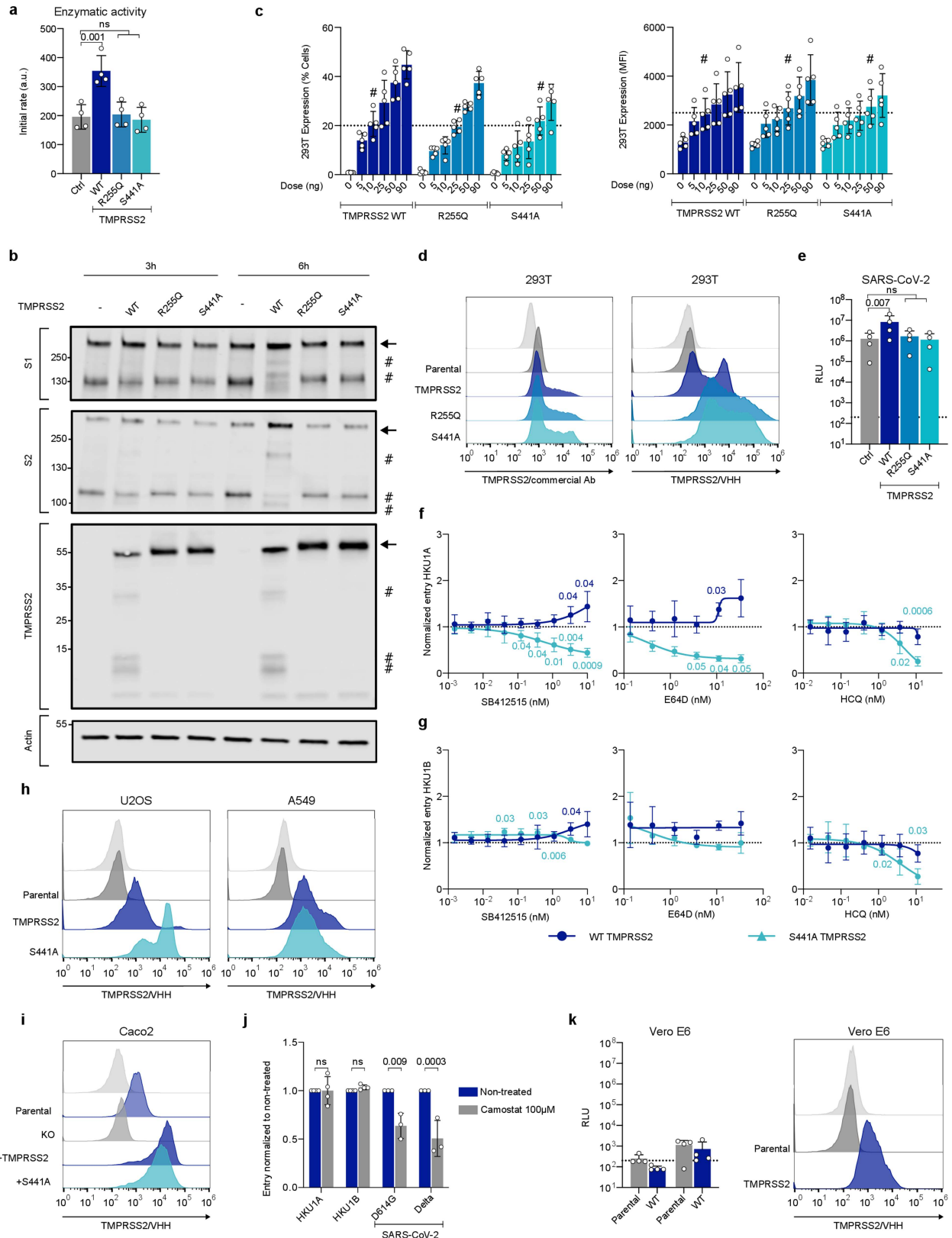
a. 293 T cells were transfected with plasmids encoding for HKU1A or HKU1B spikes and stained 24 h later with mAb10, a pan-anti-coronavirus spike antibody. Data are representative of 3 independent experiments. **b.** Cell-cell fusion mediated by HKU1A or HKU1B spikes. 293 T GFP-Split cells were transfected with HKU1 spikes and TMPRSS2, fusion was visualized by the appearance of GFP+ cells. Data are representative of 6 independent experiments. Scale bar: 400 μ m **c. Cell-cell fusion mediated by the SARS-CoV-2 spike.** 293 T GFP-Split cells were transfected with SARS-CoV-2 spike, in the presence of ACE2 or TMPRSS2, fusion was visualized by the appearance of GFP+ cells. Data are mean \pm SD of 3 independent experiments **d. U2OS cell-cell fusion mediated by HKU1A spike.** U2OS GFP-Split cells were transfected with HKU1A spike and TMPRSS2, fusion was visualized by the appearance of GFP+ cells 24 h later. Data are mean \pm SD of 3 independent experiments **e. Expression levels of myc-tagged proteases.** 293 T cells were transfected with a control plasmid or the

indicated myc-tagged TMPRSS constructs and stained 24 h later with myc antibody 9E10. Left: Representative dot plots. Right: Percentage of positive cells. Data are mean \pm SD of 3 independent experiments. **f. Surface expression of tagged and untagged TMPRSS2.** 293 T cells were transfected with WT TMPRSS2 (TMPRSS2-Untagged) or a myc-tagged TMPRSS2 and surface stained for TMPRSS2 using VHH-A01-Fc. Left: Representative dot plots. Right: Percentage of positive cells. Data are mean \pm SD of 3 independent experiments. **g, h, i. Surface expression of APN, DPP4 and ACE2.** 293 T cells were transfected with APN, DPP4 or ACE2 plasmids, and surface stained with the respective antibodies 24 h later. Left: Representative dot plots. Right: Percentage of positive cells. Data are mean \pm SD of 3 (TMPRSS2, APN, DPP4) or 4 (ACE2) independent experiments. **j. Images of time lapse microscopy of HKU1-mediated cell-cell fusion** (extracted from Supp. video 1). 293 T cells were transfected either with TMPRSS2-scarlet-l or HKU1A S-NeonGreen. After 24 h, cells were mixed and imaged every 2.5 min for 2 h at 37 $^{\circ}$ C.



Extended Data Fig. 3 | Knock out and silencing of TMPRSS2 in Caco2 cells. **a. Linear diagram of the organization of the TMPRSS2 gene.** The CRISPR-Cas9 targeting site is underlined and the proto-spacer recognition motif (PAM) is in bold. Rectangles represent exons, black for 5' and 3' untranslated regions, grey for coding regions. **b. Sequence analysis of the knock-out Caco2 clone.** Both alleles compared to the original wild-type sequence are shown. The knockout clone harbors an out-of-frame deletion and an insertion, resulting in a frameshift on both alleles. **c. Validation of TMPRSS2 knockout by TMPRSS2 surface staining.** Representative dot plots of VHH anti-TMPRSS2 (VHH-A01-Fc) surface staining of a WT and KO Caco2 obtained following CRISPR gene

targeting. **d. Role of endogenous TMPRSS2 on cell-cell fusion assessed by siRNA.** 293 T donor cells expressing HKU1A spike were mixed with Caco2 acceptor cells, silenced or not for TMPRSS2. Left: experimental design. Middle: relative expression of TMPRSS2 in Caco2 cells, assessed by RT-qPCR. Data were normalized to β -Tubulin levels. Relative mRNA expression normalized to siCtrl condition ($2^{-\Delta\Delta Ct}$) was plotted. Right: fusion was quantified by measuring the GFP area after 20 h of coculture. Data are mean \pm SD of 3 independent experiments. Statistical analysis: two-sided unpaired t-test compared to siCtrl cells. ***p < 0.0001.

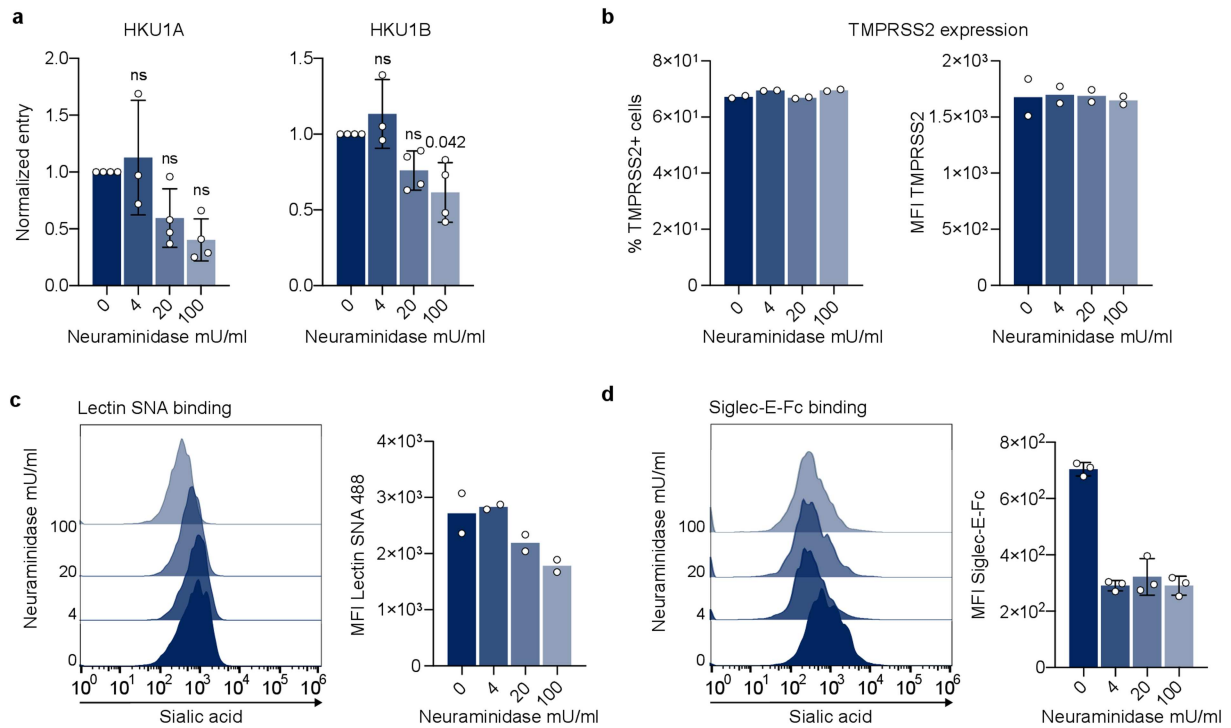


Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Effect of mutant TMPRSS2 on cell-cell fusion and pseudovirus infection a. Enzymatic activity of WT and mutant TMPRSS2.

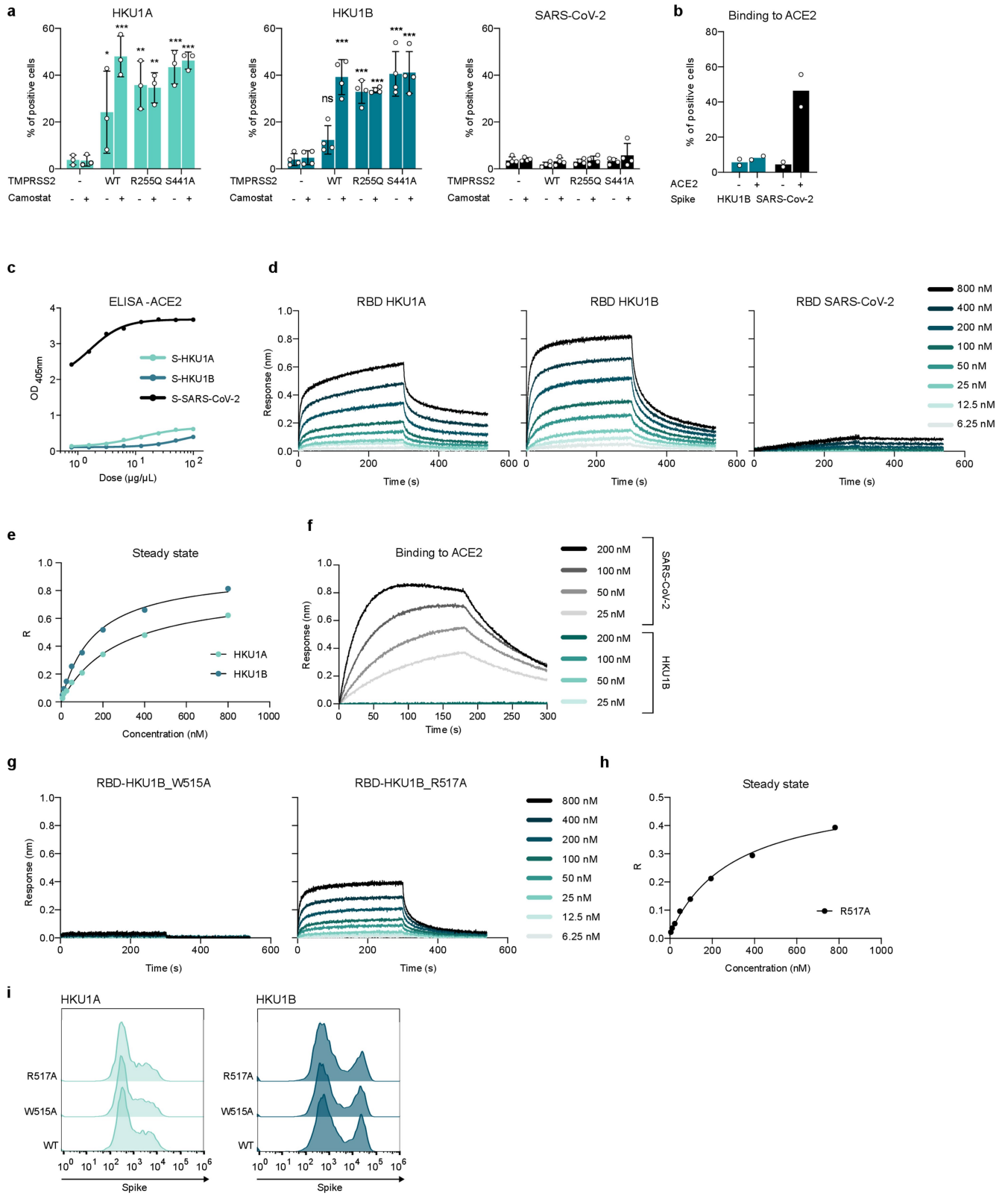
293 T cells were transfected with WT and indicated TMPRSS2 mutants. 24 h post transfection, the catalytic activity was assessed using BoC-QAR-AMC fluorogenic substrate. Data are mean \pm SD of 4 independent experiments. Statistical analysis: a: one-way ANOVA with Dunnett's multiple comparisons compared to Ctrl cells (transfected with pQCXIP-Empty). ** $p < 0.01$. **b. Effect of TMPRSS2 on HKU1 spike expressed on adjacent cells.** 293 T cells were transfected either with HKU1A spike or with the different TMPRSS2. 20 h post-transfection, cells were mixed at a 1:1 ratio, and let to settle. 3 or 6 h post-mixing, cells were harvested and lysed for WB. One membrane was probed for S1, TMPRSS2 and actin, another membrane was probed for S2 and actin. For gel source data, see SI 1c, d. Representative blots of 2. Molecular weights: kDa. The arrows and # denote the uncleaved and cleaved protein products, respectively. **c. Surface levels of WT and mutant TMPRSS2.** 293 T cells were transfected with the indicated doses of TMPRSS2 plasmid. 18 h post-transfection they were stained intracellularly for TMPRSS2 using the commercial antibody. Left: % of TMP positive cells. Right: Median fluorescent intensity (MFI). #: indicates the chosen plasmid ratios to achieve similar TMPRSS2 levels with WT and indicated mutants. Data are mean \pm SD of 3 independent experiments. **d. Comparison of the anti-TMPRSS2 commercial antibody and VHH staining in 293 T cells.** 293 T cells were transfected with WT, R255Q and S441A TMPRSS2 plasmids. Cells were stained 24 h later with the indicated antibodies and analysed by flow cytometry. One experiment representative of 3 is shown. Light grey curves correspond to staining with control antibodies or VHH. Parental: untransfected cells. **e. Effect of WT and mutant TMPRSS2 on SARS-CoV-2 pseudovirus infection.** 293 T cells expressing ACE2 were transfected with WT and indicated TMPRSS2 mutants. 24 h later, cells were infected by Luc-encoding SARS-CoV-2 pseudovirus. Luminescence was measured after 48 h. Data are mean \pm SD of 4 independent experiments. Statistical analysis: RM one-way ANOVA with Geisser-Greenhouse correction on log-transformed data, with Dunnett's

multiple comparisons compared to Ctrl cells (transfected with pQCXIP-Empty). ** $p < 0.01$. **f, g. Effect of SB412515, E64d or hydroxychloroquine (HCQ) on f. HKU1A or g. HKU1B pseudovirus infection.** 293 T cells expressing WT or S441A TMPRSS2 were incubated for 2 h with the indicated drugs, before infection with HKU1A or HKU1B pseudoviruses. Luminescence was read 48 h post infection. Left: SB412515. Middle: E64d. Right: HCQ. Data are mean \pm SD of 3 (E64d, HKU1B), 4 (E64d HKU1A, SB142515 HKU1B), 5 (HCQ HKU1A) or 6 (HCQ HKU1B, SB412515 HKU1A) independent experiments. Statistical analysis: RM two-way ANOVA with Geisser-Greenhouse correction on non-normalized log transformed data, with Dunnett's multiple comparisons compared to the non-treated conditions. **h, i. Surface levels of TMPRSS2 in different cell lines stably expressing WT or S441 TMPRSS2.** Cells were stained for TMPRSS2 using VHH-A01-Fc and analyzed by flow cytometry. Representative histograms are shown. **h.** Left: U2OS. Right: A549. Light grey: cells stained with a non-target control VHH-Fc. Dark grey: Unmodified parental cell lines. Dark and light blue: Cells transduced with either TMPRSS2 or TMPRSS2 S441A mutant. **i.** Caco2. Light grey: cells stained with a non-target control VHH-Fc. Blue: Unmodified parental cell line. Dark Grey: TMPRSS2 KO Caco2. Dark and light blue: TMPRSS2 KO caco2 stably transduced with TMPRSS2 WT or S441A mutant expression vectors. **j. Effect of Camostat on endogenous TMPRSS2 in Caco2 cells.** Caco2 cells were incubated in the presence of 100 μ M Camostat for 2 h, before infection with HKU1A, HKU1B, or SARS-CoV-2 (D614G or Delta) pseudovirus. Data are normalized to the infection in the absence of the drug. Data are mean \pm SD of 3 (SARS-CoV-2) or 4 (HKU1) independent experiments. Statistical analysis: RM two-way ANOVA on non-normalized log transformed data, with Sidak's multiple comparisons compared to the non-treated conditions. **k. Susceptibility of Vero E6 and Vero E6-TMPRSS2 cells to HKU1 pseudovirus infection.** Left: pseudovirus infection. Right: TMPRSS2 surface levels. Dark grey: Parental cells. Dark blue: Cells transduced with TMPRSS2. Data are mean \pm SD of 4 independent experiments.



Extended Data Fig. 5 | Effect of neuraminidase on HKU1 pseudovirus infection. U2OS-TMPRSS2 cells were treated with indicated concentration of neuraminidase from *Arthrobacter ureafaciens* for 24 h. **a. HKU1A and HKU1B pseudovirus infection in neuraminidase treated cells.** Cells were infected with HKU1A or B pseudoviruses. Luminescence was read 48 h post infection. Data were normalized to the non-treated condition. Data are mean \pm SD of 3 (4 mU/mL) or 4 (0, 20, 100 mU/mL) independent experiments. Statistical analysis: Mixed-effect analysis with Geisser-Greenhouse corrections (handles missing values) on non-normalized log transformed data, with Dunnett's multi-comparison test compared to non-treated cells. **b. Surface levels of**

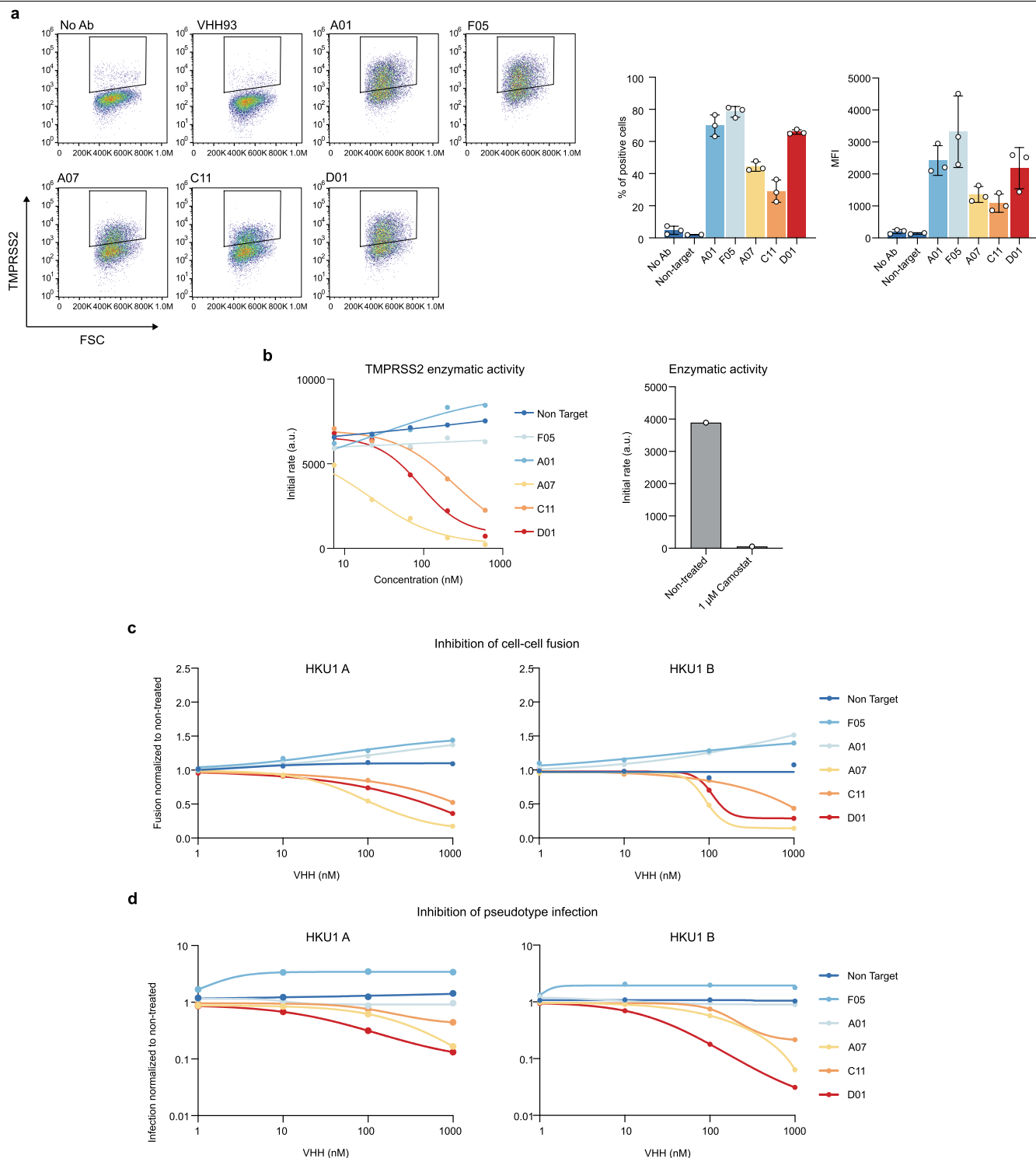
TMPRSS2 in neuraminidase treated cells. Left: % of TMP positive cells. Right: Median fluorescent intensity (MFI). Data are mean of 2 independent experiments. **c. Sambucus Nigra Lectin (SNA) binding on neuraminidase treated cells.** Treated cells were stained with fluorescent Lectin SNA that preferentially binds α -2,6- over α -2,3 linked sialic acids. Left: Representative histograms. Right: MFI. Data are mean of 2 independent experiments. **d. Siglec-E binding on neuraminidase treated cells.** Treated cells were stained with recombinant Siglec-E-Fc protein that preferentially binds α 2,8- over α 2,3- and α 2,6-linked sialic acids. Left: Representative histograms. Right: MFI. Data are mean \pm SD of 3 independent experiments.



Extended Data Fig. 6 | See next page for caption.

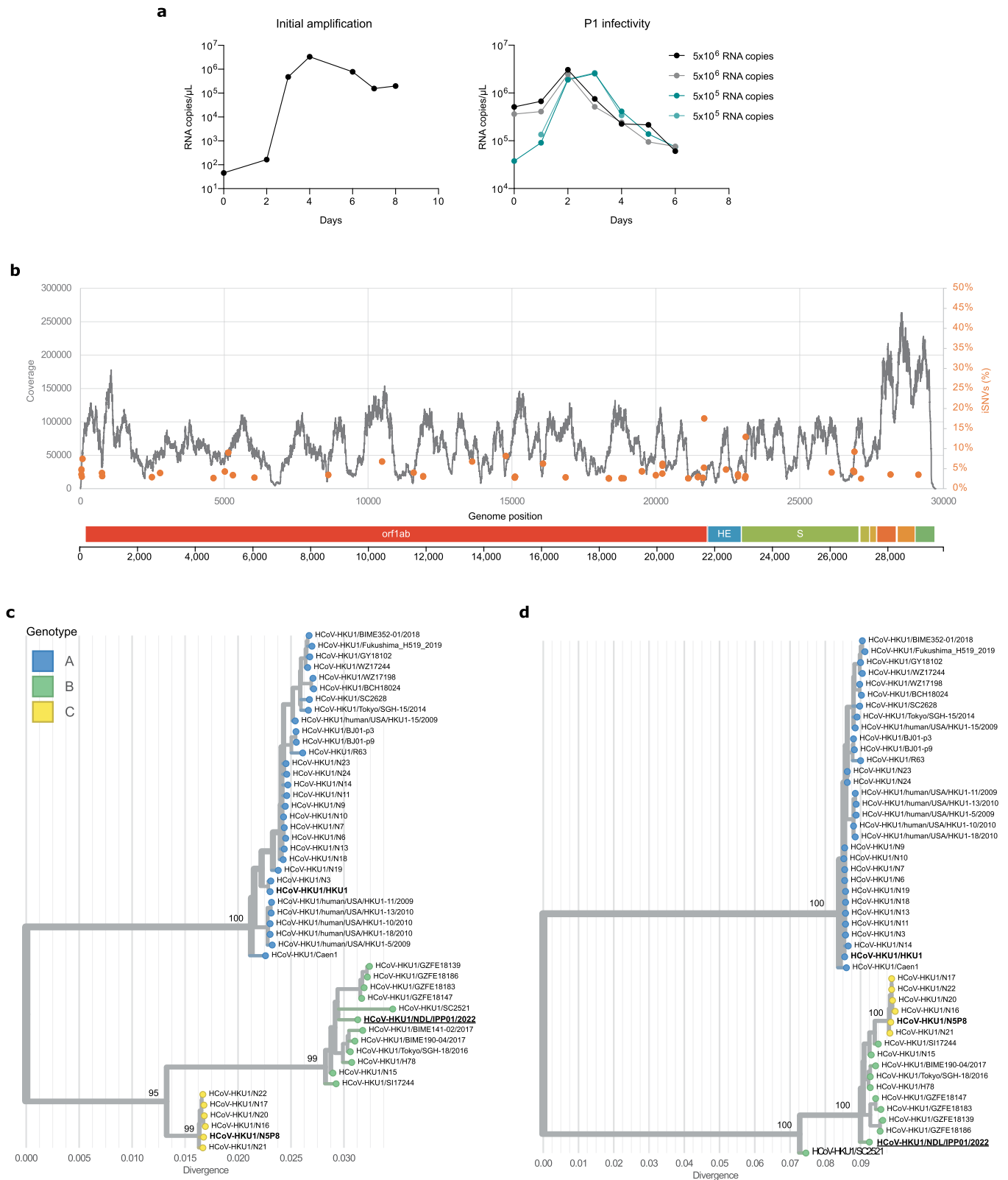
Extended Data Fig. 6 | Binding of HKU and SARS-CoV-2 spikes to TMPRSS2 or ACE2. a. Binding of the indicated recombinant spikes to 293 T cells expressing TMPRSS2. Cells were transfected with WT or mutant TMPRSS2 and incubated or not overnight with 10 μ M of Camostat. The spikes were then incubated for 0.5 h and their binding was revealed with streptavidin-647 and measured by flow cytometry. The % of cells binding to TMPRSS2 was quantified. Data are mean \pm SD of 3 (HKU1A) or 4 (HKU1B, SARS-CoV-2) independent experiments. Two Way ANOVA with Dunnett's multiple comparisons compared to control cells with or without Camostat. Exact p-values: HKU1A: TMPRSS2 WT: *0.029, TMPRSS2 WT + :*** < 0.0001, TMPRSS2 R255Q: **0.0010, TMPRSS2 R255Q + :**0.0013, TMPRSS2 S441A: ***0.0001, TMPRSS2 S441A + : ***<0.0001. HKU1B: ***<0.0001. **b. Binding of the indicated recombinant spikes to 293 T cells expressing ACE2.** Cells were transfected with ACE2. The spikes were then incubated for 0.5 h and their binding was revealed with streptavidin-647 and measured by flow cytometry. The % of cells binding to ACE2 was quantified. Data are mean of 2 independent experiments. **c. Binding**

of the indicated soluble spikes on immobilized ACE2 measured by ELISA. d. Binding of S441A TMPRSS2 to HKU1A, HKU1B or SARS-CoV-2 RBD measured by BLI. The response was measured at the indicated concentrations of spikes. Left: HKU1A. Middle: HKU1B. Right: SARS-CoV-2. One representative experiment of 4 is shown. **e. Determination of the affinity of HKU1A and B RBD for TMPRSS2 using the steady state method.** Circles: Experimental values. Black: Fitting of the experimental data **f. Binding of ACE2 to SARS-CoV-2 or HKU1B RBD quantified by BLI,** at different concentrations of spikes. **g. Binding of S441A TMPRSS2 to HKU1B mutants.** Response was measured by BLI at different concentrations of spikes. Left: HKU1B RBD mutant W515A. B: HKU1B RBD mutant R517A. **h. Determination of the affinity of HKU1B-R517A RBD for TMPRSS2 using the steady state method.** Circles: Experimental values. Black: Fitting of the experimental data. **i. Cell surface levels of WT and mutant HKU1 spikes.** 293 T were transfected with the indicated WT or mutant HKU1A or B spikes, expression was measured by flow cytometry after 24 h, using the anti-spike mAb10.



Extended Data Fig. 7 | Characterization of anti-TMPRSS2 VHHs. a. Binding of VHHs on TMPRSS2-expressing cells. 293 T cells were transfected with TMPRSS2 S441A. Binding of the indicated VHHs (0.5 μM) was assessed by flow cytometry. Left: representative dot plots. Right: Quantification of the percentage of positive cells and MFI (Median Fluorescent Intensity of positive cells). Data are mean ± SD of three independent experiments **b. Effect of VHHs or Camostat on TMPRSS2 enzymatic activity.** Recombinant soluble TMPRSS2 was incubated with the indicated VHHs at different concentrations or with Camostat (1 μM). The initial rate of enzymatic activity, measured with a fluorescent substrate is plotted. One representative experiment of 3 is shown. **c. Effect of VHHs on HKU1A or HKU1B cell-cell fusion.** 293 T GFP-Split cells were co-transfected with TMPRSS2 and HKU1 spike in the presence of indicated

amounts of VHH. Fusion was quantified by measuring the GFP area after 20 h. Data were normalized to the non-treated condition for each spike. **d. Effect of VHHs on HKU1A or HKU1B pseudovirus infection.** 293 T cells transfected with S441A TMPRSS2 were incubated with the indicated amounts of VHH for 2 h and infected by Luc-encoding pseudovirus. Luminescence was read 48 h post infection. Data were normalized to the non-treated condition for each virus. Data are from one experiment representative of 3. **Statistical analysis:** c: One Way ANOVA with Dunnett's multiple comparisons compared to cells stained with a non-target antibody (VHH93). d: One Way ANOVA on log-transformed data with Dunnett's multiple comparisons compared to cells stained with a non-target antibody (VHH93).

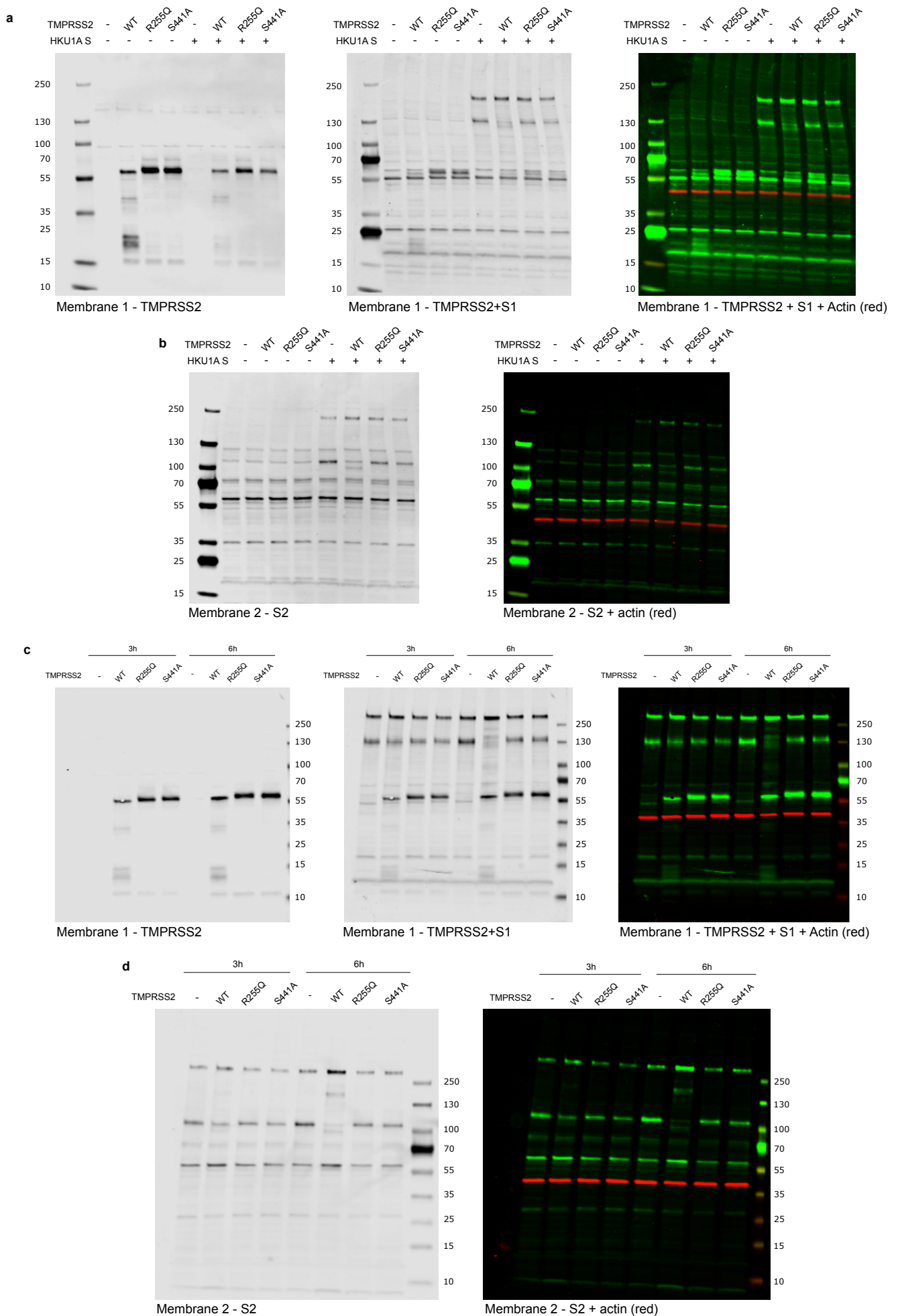


Extended Data Fig. 8 | Isolation and characterization of a live HKU1B virus.
a. Viral RNA copies in the supernatant of HBE cells were quantified by RT-qPCR. The values at day 0 are the content of the input. Left: initial amplification (first passage). Right: Second amplification of the first passage virus (harvested at day 4) used at two dilutions (1/10 and 1/100) in duplicates.
b. Reads coverage and frequency of minor variants frequencies of the isolate sequencing data (first passage virus). Intra-sample single-nucleotide variants frequencies were estimated using iVAR⁷⁰. The genome organization of HKU1 is shown below the plot.
c, d. Phylogenetic analysis of HKU1. Maximum

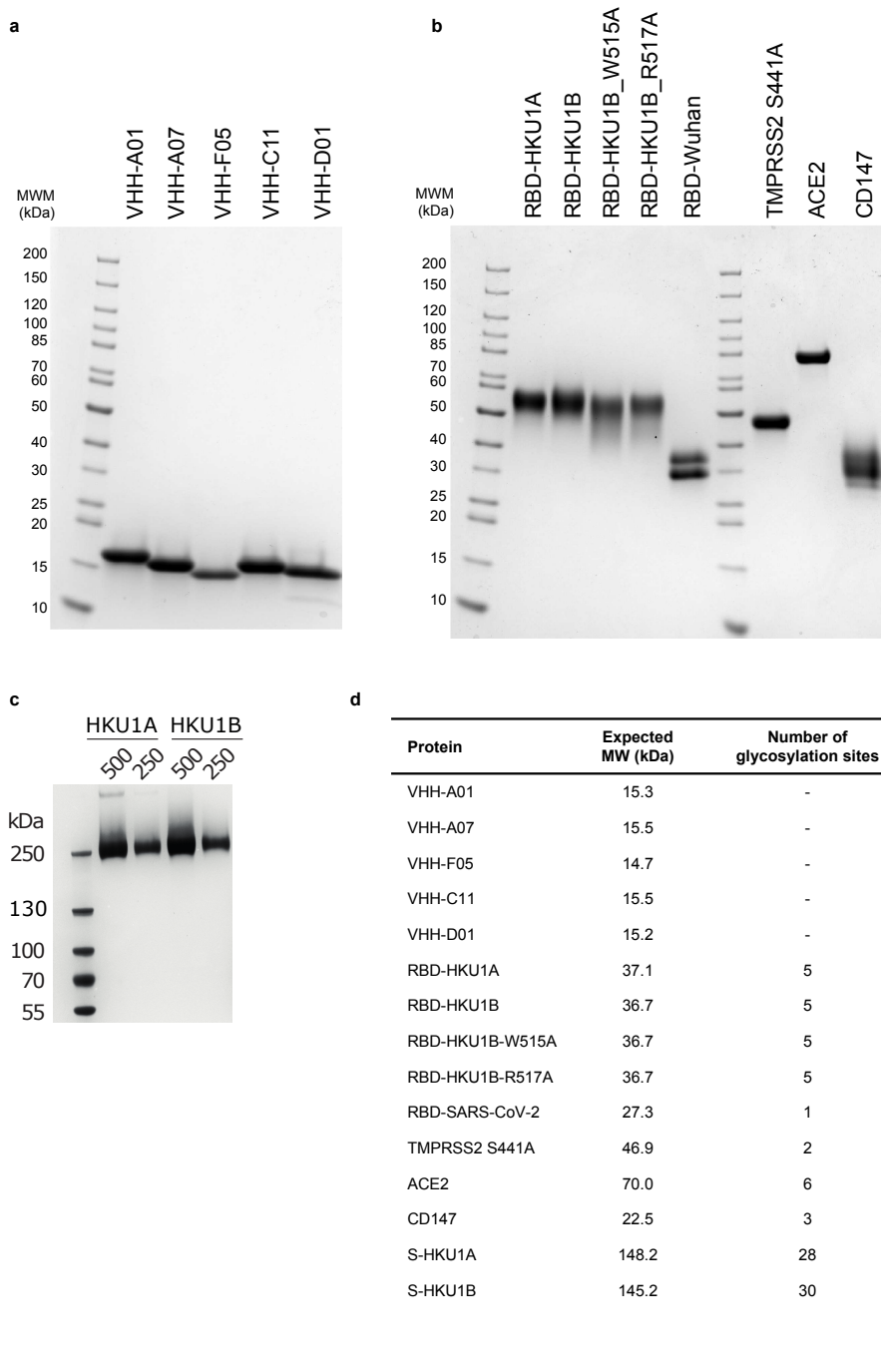
likelihood phylogenies of human coronavirus HKU1 (n = 48) estimated using IQ-TREE v2 with 1000 replicates from **c.** the complete genome and **d.** spike coding sequences. The tree is midpoint rooted and ultrafast bootstraps values are shown on the main branches. A similar topology was obtained when rooting the tree using another embecovirus (OC43). The tip name corresponding to the newly isolated virus sequence is highlighted in bold and underlined. The tips names corresponding to the recombinant spikes used in this study are shown in bold.

Extended Data Table 1 | Affinity of the nanobodies for TMPRSS2 S441A and summary of their effect

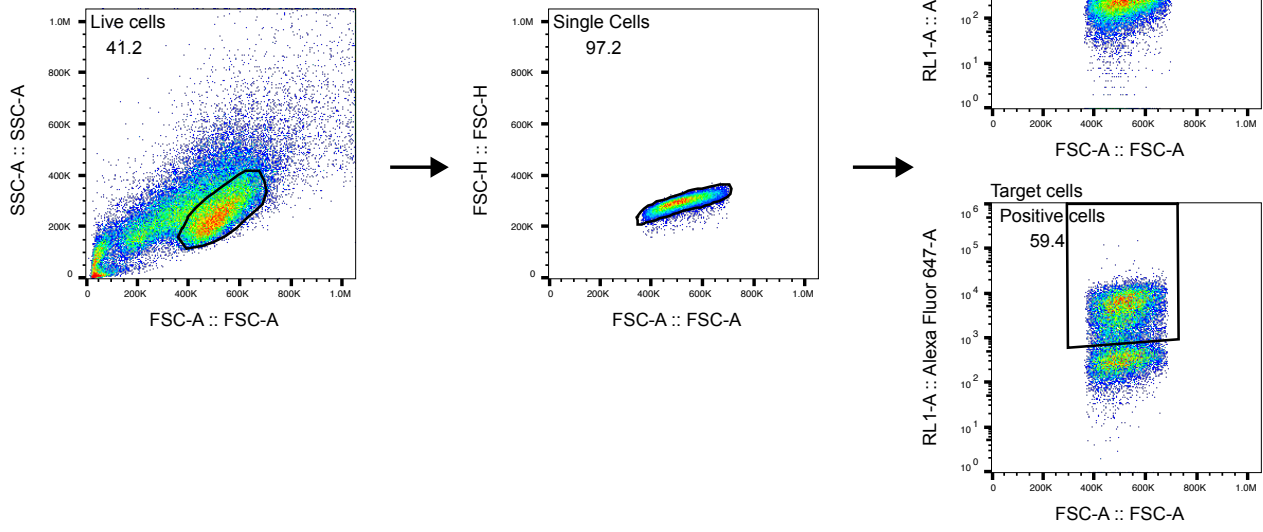
Nanobody	k_{on} (1/Ms)	k_{off} (1/s)	K_d (nM)	Enzymatic activity inhibition	Cell-cell fusion inhibition	Pseudotype infection inhibition
VHH-A01	$6.4 \cdot 10^4$	$5.1 \cdot 10^{-5}$	0.78	-	-	-
VHH-F05	$7.6 \cdot 10^4$	$1.9 \cdot 10^{-4}$	2.6	-	-	-
VHH-A07	$8.6 \cdot 10^4$	$1.1 \cdot 10^{-3}$	13.3	++	++	++
VHH-C11	$7.0 \cdot 10^4$	$2.3 \cdot 10^{-4}$	32.5	+	+	+
VHH-D01	$7.9 \cdot 10^4$	$4.8 \cdot 10^{-4}$	6.0	+	+	++



Supplementary Information 1: Full western blot membranes (a,b, Fig. 2, c,d. Extended Data Fig. 4)



Supplementary Information 2: a. b. c. Quality control gels of the purified proteins. d. Expected molecular weights and glycosylation sites



Supplementary Information 3: Gating strategies

VHH Sequences

Monomeric VHH sequences :

>A01
QVQLVESGGGLVQPGGSLRLSCVVSGFSLDYYAIGWFRQAPGKEREVSCIGSSGDKTNYADSVKGRFTISRDN
NTVYLQMNSLKPEDTAVYYCAAESALYSDCTEEQNPLYDYWGQGTQVTVSS

>A07
QVQLVESGGGLVQPGGSLRLSCTS SGPLEHYDIIWFRQAPGREREGVSSITSSGGHTNYADSVKGRFTISRDN
VVYLQMNSLKPEDTAVYYCAGRVGRRNWIVPLDGYDNAYWGQGTQVTVSS

>C11
QVQLVESGGGLVQPGGSLRLSCAA SGFTLDYYDIYWFRQAPGKEREVSSITSSGGRTNYADSVKGRFTISRDN
NTVYLQMNSLKPEDTAVYYCAA AVGRRNWIAPLNGYENALWGKGLTVTS

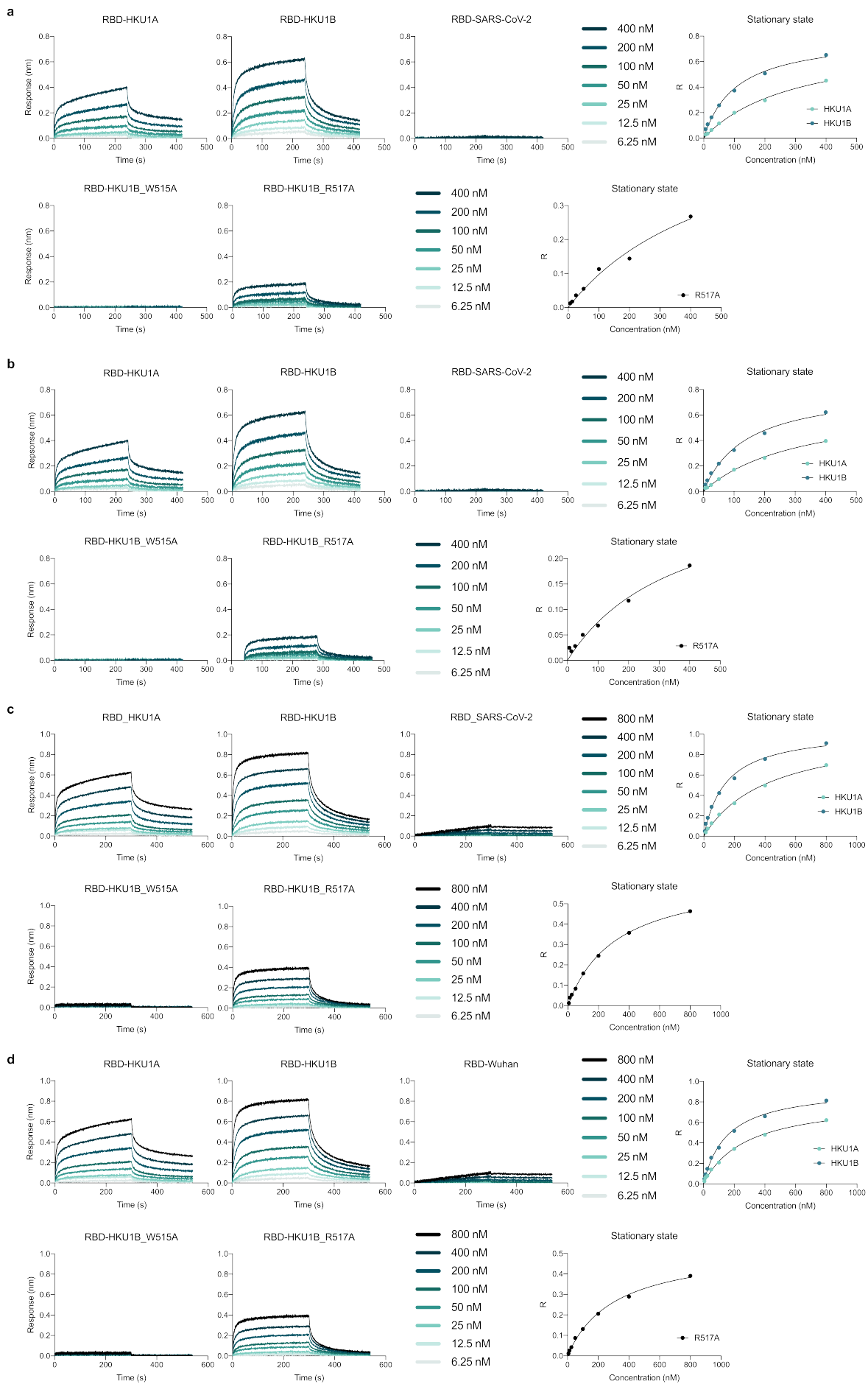
>D01
EVQLVESGGGLVQPGGPLRLSCAS SGSTLEHYDIGWFRQVPGGLREGVSSITASSGRTNYADSVKGRFTISRDN
AVYLQMNSLKPEDTAVYYCAGKIGRRNWVAPLDGFENAYWGQGTQVTVSS

>F05
EVQLVESGGGLVQPGGSLRLSCAA SGFTLDYYAIGWFRQAPGKEREVSCISSGDSIKYVDSVKGRFTISRDN
TVYLQMNSLKPEDTAVYYCAA DALGSGCLTGNVDYWGQGTQVTVSS
CDR1-CDR2-CDR3

VHH-Fc sequences :

>TMPR-A1-Fc-*humain* (VHH-A01-Fc)
QVQLVESGGGLVQPGGSLRLSCVVSGFSLDYYAIGWFRQAPGKEREVSCIGSSGDKTNYADSVKGRFTISRDN
NTVYLQMNSLKPEDTAVYYCAAESALYSDCTEEQNPLYDYWGQGTQVTVSSSEPKTPKQPAAARSDKTHCPPCP
APELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQ
DWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYK
TTPPVLDSDGSFFLYSKLTVDKSRWQQGNVVFSCVMHEALHNHYTQKSLSLSPGK

Supplementary Information 4: VHH Sequences



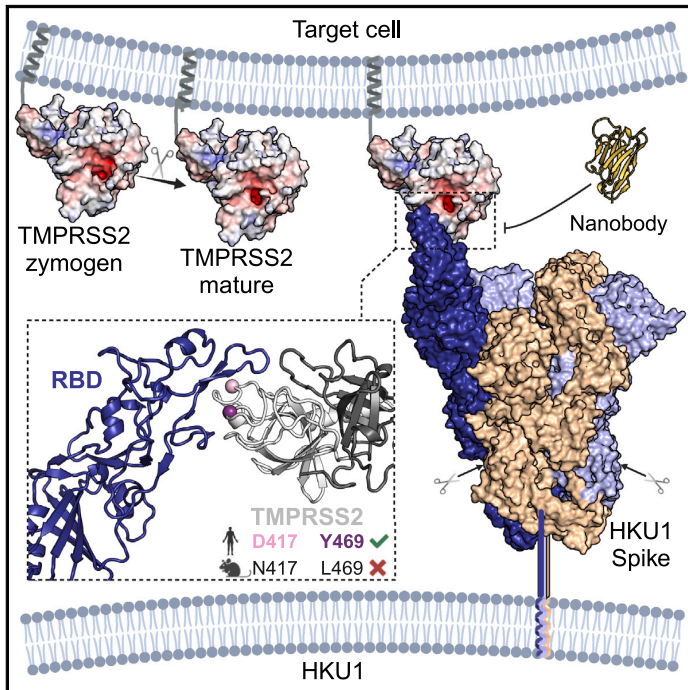
Supplementary Information 5: a. b. c. d. Complete BLI data used for fitting and determination of the affinity of Tmprss2 for HKU1 RBD

Replicate	RBD	Kd	95% CI	R ²	Range tested
1	HKU1A	304.5	226.6 to 433.3	0.9967	6.25-400 nM
2	HKU1A	232.1	245.1 to 440	0.9975	6.25-400 nM
3	HKU1A	430.1	328.6 to 574.4	0.9965	6.25-800 nM
4	HKU1A	279.6	221.1 to 356.9	0.9964	6.25-800 nM
1	HKU1B	106.9	75.9 to 152.6	0.9922	6.25-400 nM
2	HKU1B	140.6	94.9 to 213.5	0.9913	6.25-400 nM
3	HKU1B	144.2	107.6 to 194.2	0.9928	6.25-800 nM
4	HKU1B	155.1	119.8 to 201.5	0.9945	6.25-800 nM
1	HKU1B_R517A	515.1	211.9 to 3327	0.9812	6.25-400 nM
2	HKU1B_R517A	363.8	150.8 to 1523	0.9735	6.25-400 nM
3	HKU1B_R517A	316.9	267.6 to 377.3	0.9983	6.25-800 nM
4	HKU1B_R517A	308.0	246.7 to 388.0	0.9971	6.25-800 nM

Supplementary Information 6: Kd, 95% confidence interval, R² of the fit, and range of the concentration of TMRSS2 tested for all BLI replicates

3.3. Structural basis of TMPRSS2 zymogen activation and recognition by the HKU1 seasonal coronavirus

Graphical abstract



Authors

Ignacio Fernández, Nell Saunders, Stéphane Duquerroy, ..., Julian Buchrieser, Olivier Schwartz, Félix A. Rey

Correspondence

olivier.schwartz@pasteur.fr (O.S.), felix.rey@pasteur.fr (F.A.R.)

In brief

TMPRSS2 is a cell-surface protease synthesized as a zymogen that undergoes autocleavage to reach its active conformation, which involves conformational changes on key exposed activation loops. The HKU1 seasonal coronavirus uses TMPRSS2 as entry receptor, with the spike protein recognizing the activation loops and inserting a motif into the protease's catalytic groove.

Highlights

- HKU1-HCoV RBD binds at the periphery of TMPRSS2 catalytic groove
- HKU1-HCoV Spike opening is required for TMPRSS2 recognition
- TMPRSS2 residues D417 and Y469 are main determinants of host tropism
- TMPRSS2 autocleavage increases affinity toward HKU1-HCoV RBD

Fernández et al., 2024, *Cell* 187, 1–15
August 8, 2024 © 2024 Elsevier Inc. All rights are reserved, including those for text and data mining, AI training, and similar technologies.
<https://doi.org/10.1016/j.cell.2024.06.007>



Article

Structural basis of TMPRSS2 zymogen activation and recognition by the HKU1 seasonal coronavirus

Ignacio Fernández,^{1,9} Nell Saunders,^{2,9} Stéphane Duquerroy,^{1,3} William H. Bolland,² Atousa Arbabian,¹ Eduard Baquero,⁴ Catherine Blanc,⁵ Pierre Lafaye,⁶ Ahmed Haouz,⁷ Julian Buchrieser,² Olivier Schwartz,^{2,8,*} and Félix A. Rey^{1,10,*}

¹Institut Pasteur, Université de Paris Cité, CNRS UMR 3569, Structural Virology Unit, 75015 Paris, France

²Institut Pasteur, Université de Paris Cité, CNRS UMR 3569, Virus & Immunity Unit, 75015 Paris, France

³Université Paris-Saclay, Faculté des Sciences, Orsay, France

⁴Institut Pasteur, Université de Paris Cité, INSERM U1222, Nanoimaging core, 75015 Paris, France

⁵Institut Pasteur, Université de Paris Cité, Pasteur-TheraVectys Joint Lab, Paris, France

⁶Institut Pasteur, Université Paris Cité, CNRS UMR 3528, Antibody Engineering Facility-C2RT, 75015 Paris, France

⁷Institut Pasteur, Université Paris Cité, CNRS UMR 3528, Crystallography Facility-C2RT, 75015 Paris, France

⁸Vaccine Research Institute, Créteil, France

⁹These authors contributed equally

¹⁰Lead contact

*Correspondence: olivier.schwartz@pasteur.fr (O.S.), felix.rey@pasteur.fr (F.A.R.)

<https://doi.org/10.1016/j.cell.2024.06.007>

SUMMARY

The human seasonal coronavirus HKU1-CoV, which causes common colds worldwide, relies on the sequential binding to surface glycans and transmembrane serine protease 2 (TMPRSS2) for entry into target cells. TMPRSS2 is synthesized as a zymogen that undergoes autolytic activation to process its substrates. Several respiratory viruses, in particular coronaviruses, use TMPRSS2 for proteolytic priming of their surface spike protein to drive membrane fusion upon receptor binding. We describe the crystal structure of the HKU1-CoV receptor binding domain in complex with TMPRSS2, showing that it recognizes residues lining the catalytic groove. Combined mutagenesis of interface residues and comparison across species highlight positions 417 and 469 as determinants of HKU1-CoV host tropism. The structure of a receptor-blocking nanobody in complex with zymogen or activated TMPRSS2 further provides the structural basis of TMPRSS2 activating conformational change, which alters loops recognized by HKU1-CoV and dramatically increases binding affinity.

INTRODUCTION

The cell surface transmembrane serine protease 2 (TMPRSS2) proteolytically primes the spike protein of multiple coronaviruses for entry into cells of the respiratory tract and was recently identified as the entry receptor for the human HKU1 coronavirus (HKU1-CoV).¹ TMPRSS2 belongs to the “type 2 transmembrane serine proteases” (TTSP) involved in proteolytic remodeling of the extracellular matrix. Dysregulation of TMPRSS2 and other TTSPs has been observed in malignancies and is associated to tumor proliferation and invasiveness.^{2–5} Moreover, TMPRSS2 is an androgen-regulated protease that activates prostate cancer metastatic cascades.⁶

HKU1-CoV is a seasonal beta-coronavirus initially isolated in 2005 at the University of Hong Kong,⁷ hence its name. It causes common colds worldwide and can develop complications in young children, the elderly, and immunocompromised individuals.⁸ Entry of HKU1-CoV into cells relies on its trimeric

spike surface glycoprotein, which catalyzes the fusion of the viral envelope with the target cell membrane. The spike is synthesized as a precursor cleaved into two subunits, S1 and S2. S1 harbors an N-terminal domain (NTD), a receptor binding domain (RBD), and subdomains SD1 and SD2. The RBD adopts conformations known as “up” and “down”, which determine different forms on the spike (closed, 1-RBD-up, 2-RBD-up, 3-RBD-up, or “open”). The S2 subunit is the membrane fusion effector and requires further cleavage at a second site, the S2' site, to be functional.⁹ As an important entry factor, one of the roles of TMPRSS2 is to cleave HKU1-CoV S2 at the S2' site,¹ priming it for driving membrane fusion.

The HKU1-CoV spike NTD binds α 2,8-linked 9-O-acetylated disialosides (9-O-Ac-Sia(α 2,8)Sia)^{10–12} causing a conformational change that exposes the RBD and opens the spike.¹³ The transition into open states has been associated to the activation of the protein to trigger membrane fusion.^{14,15} The surface sialo glycan is thus a primary receptor of HKU1-CoV.

The activity of a hemagglutinin esterase (HE) anchored in the viral envelope^{10,16} allows dissemination of the virus to relevant tissues by acting as a receptor-destroying enzyme.¹⁷ We have shown that TMPRSS2 recognizes the HKU1-CoV RBD in a so far uncharacterized mode.¹ Even though sialic acid is important to trigger efficient HKU1-CoV entry into TMPRSS2 expressing cells,¹ no experimental evidence has demonstrated the interplay between the two receptors. We also developed nanobodies against TMPRSS2, some of which bind without affecting recognition of the RBD (i.e., nanobody A01), while others block the interaction, reducing infection of HKU1-CoV susceptible cells, and inhibit TMPRSS2 proteolytic activity (i.e., nanobody A07).¹

TMPRSS2 is a type II single-pass transmembrane protein with an N-terminal cytosolic tail and an ectodomain containing a low-density lipoprotein receptor type-A (LDLR-A) domain followed by a class A scavenger receptor cysteine-rich (SRCR) domain and a C-terminal trypsin-like serine peptidase (SP) domain.¹⁸ It is synthesized as a proenzyme (zymogen) that undergoes auto cleavage to reach its mature, active conformation.^{19,20} The active site contains the catalytic triad H296-D345-S441, which cleaves after arginine 255 of the zymogen for autolytic activation, at the peptide sequence (RQSR₂₅₅↓IVGG), where the arrow denotes the scissile bond. The substrate sequence, designated P4-P3-P2-P1↓P1'-P2'-P3'-P4', is recognized through specific TMPRSS2 residues that form the corresponding S4-S3-S2-S1-S1'-S2'-S3'-S4' sites along the catalytic groove.²¹ TMPRSS2 substrates include proteins of the tumor microenvironment,⁶ such as the hepatocyte growth factor (cleaved sequence KQLR₄₉₄↓VVNG), tissue plasminogen activator (PQFR₃₁₀↓IKGG), or human glandular kallikrein 2 (IQSR₂₄↓IVGG), as well as the S2' site of the coronavirus spike,^{22–26} at the sequence motifs SSSR₉₀₀↓SLLS (HKU1-CoV isolate N5) or PSKR₈₁₅↓SFIE (SARS-CoV-2 Wuhan). The amino acid cleavage preferences at positions P1–P4 of the substrate have been determined by positional scanning of synthetic peptide libraries.⁶

Here, we determined the X-ray structure of TMPRSS2 in its zymogen form as a ternary complex with the HKU1-CoV genotype B RBD (from here on, HKU1B RBD) and the non-competing nanobody A01. TMPRSS2 binds exclusively to RBDs in the “up” conformation, making spike opening mandatory for binding. Sialo glycan binding to the spike at the NTD is a required first step that allows fusion with TMPRSS2-expressing cells. We further identified key TMPRSS2 residues determining host tropism through extensive functional validation of the RBD-TMPRSS2 interface. We also determined a high-resolution X-ray structure of TMPRSS2 in complex with nanobody A07, showing that it inserts its long CDR3 into the protease’s active site, explaining its role in blocking HKU1-CoV entry and TMPRSS2 activity. The crystals of the TMPRSS2/A07 complex contained two forms, the zymogen and activated protease, illustrating the activation mechanism of TMPRSS2 and providing key information for the development of specific, zymogen-targeting drugs. The conformational change caused by the activating cleavage of TMPRSS2 involves loops that are recognized by the HKU1-CoV RBD, and protease activation is required for high-affinity binding.

RESULTS

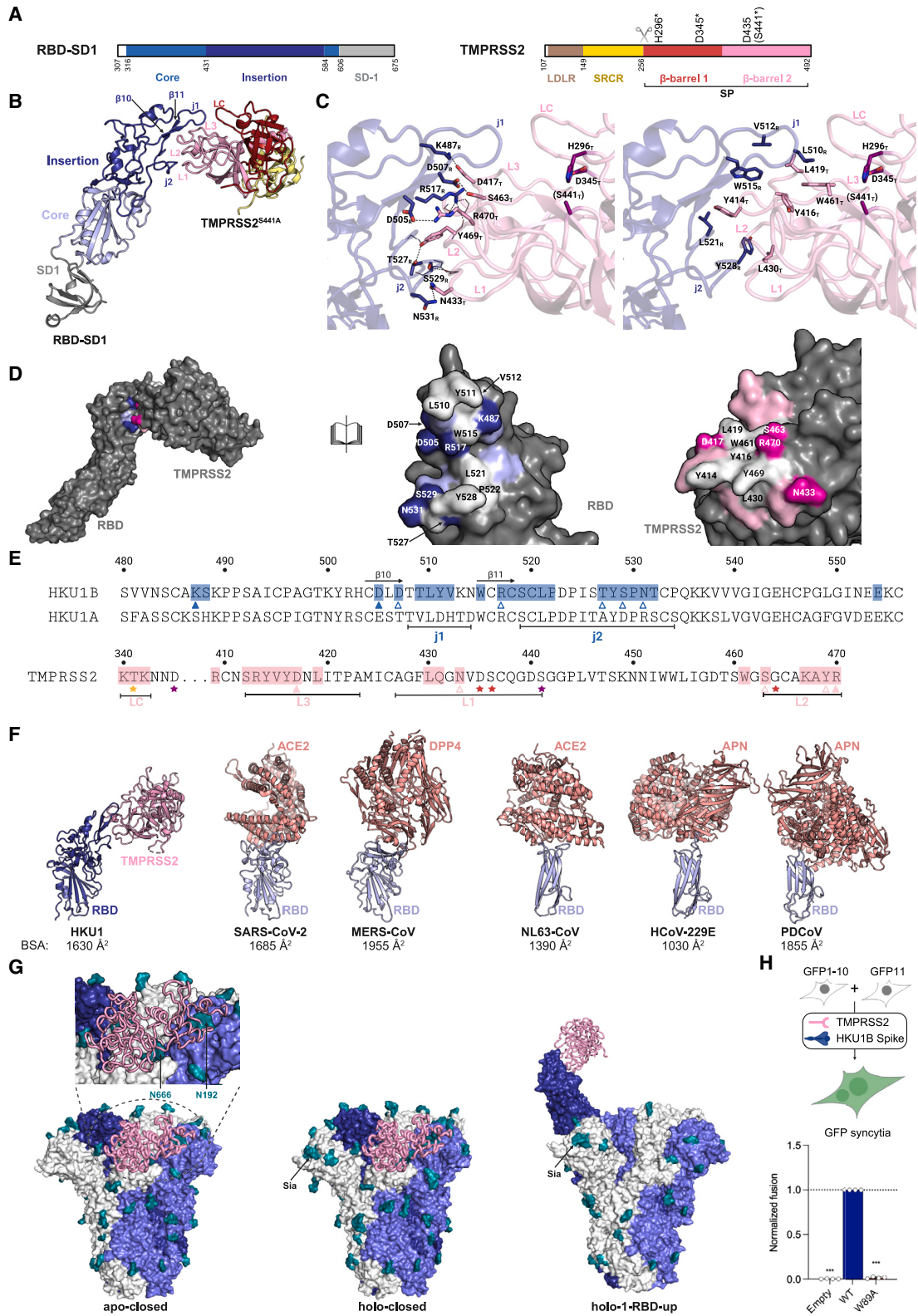
The HKU1B RBD partially blocks the TMPRSS2 catalytic groove

We determined the X-ray structure of the ternary TMPRSS2^{S441A}/RBD/A01 complex (Figure S1) to 3.55 Å resolution, displayed in Figures 1A–1E, with the crystallographic statistics listed in Table 1. We used the inactive TMPRSS2^{S441A} mutant to increase protein yield and avoid protein degradation during purification and crystallization. As described previously, the SP domain of TMPRSS2 has a characteristic chymotrypsin-like fold,²¹ featuring two β-barrels with the substrate binding groove in between.²⁸ This groove is surrounded by eight loops: loops 1, 2, and 3 in the C-terminal β-barrel controlling specificity for the P1 position of the substrate, and loops A, B, C, and E, (N-terminal β-barrel) as well as loop D (C-terminal β-barrel), which affect specificity at more distal positions.^{29,30} The structure of the complex (Figure 1B) shows that the elongated HKU1B RBD has a structure resembling a pincer plier at its distal end (in the “insertion” domain, Figure 1A), with two jaws (j1 and j2) that grab the TMPRSS2 SP domain at a surface involving loops 1–3 (L1–L3) as well as loop C (LC). The RBD pincer plier’s base is made by a distal β-hairpin of the RBD (residues 503–519, β10–β11), with the loop connecting the two β-strands forming one of the jaws (j1, aa 509–512) while the other is formed by the segment immediately downstream the β-hairpin (j2, residues 519–533) (Figure 1B). The base of the plier contacts loops 2 and 3 of the SP domain, while the j1 jaw inserts into the catalytic groove and contacts loops 3 and C and j2 interacts with loop 1 (Figure 1B). The RBD thus interacts with both β-barrels of the TMPRSS2 SP domain at either side of the groove between them, thereby obstructing access of the substrate to the catalytic site. We performed an *in vitro* TMPRSS2 activity assay and showed that incubation with increasing RBD concentrations inhibits the protease activity (Figure S1C).

Nanobody A01 binds at the interface between the SRCR and SP domains, and contacts loops of the SP C-terminal β-barrel at the side opposite to the catalytic groove (Figure S1B), with a large, buried surface area (BSA) of 2,500 Å² (1,300 Å² on the nanobody and 1,200 Å² on TMPRSS2) (Table 2).

Highly complementary interacting surfaces

The HKU1B RBD/TMPRSS2 interface buries a surface area of about 1,600 Å² (~800 Å² on each side, Table 2) and has a shape complementarity (Sc) value of 0.72, which is substantially higher than the typical 0.64–0.68 range observed for antibody–antigen complexes.³¹ The polar network at the interface involves salt bridges and several hydrogen bonds (listed in Table S1 and displayed in Figure 1C, left). The polar interactions surround a central hydrophobic patch in the surfaces of both RBD and TMPRSS2 (Figure 1D) that gives rise to a non-polar cluster at the center of the interface (Figure 1C, right). We refer from here on to RBD or TMPRSS2 residues using suffix R or T, respectively. Of note, residues W515_R and R517_R, which, as we showed previously,¹ are crucial for the interaction with TMPRSS2, are located at the interface. The catalytic triad (H296_T, D345_T, S441_T) is not directly contacted by the RBD, which instead reaches the substrate binding site (residues K342_T from the S1 site



(legend on next page)

and T341_T at S2, in loop C) or nearby residues (i.e., N433_T and S463_T; Figure 1E). Comparison with the mode of binding of the RBD of other coronaviruses to their receptor (Figure 1F) showed that the HKU1B RBD shares a similar buried surface area at the interface, although in the other coronaviruses the RBDs lack the pincer plier structure observed in HKU1.

Only an RBD in the “up” conformation can bind TMPRSS2

To understand the interactions of the HKU1B RBD with TMPRSS2 in the context of the trimeric spike, we superimposed the X-ray structure of the complex on the RBD of the HKU1-CoV spike.^{13,32} This exercise showed that TMPRSS2 cannot bind the RBD in the closed conformation because of clashes with the adjacent S1 protomer in the trimer (Figure 1G). In contrast, TMPRSS2 binding is unencumbered with the RBD in the “up” conformation, which is achieved only upon binding a disialoside glycan (9-*O*-Ac-Sia(α2,8)-Sia) in an allosteric pocket present in the NTD.¹³ Our RBD/TMPRSS2 structure therefore predicts that mutations that abolish binding of sialo glycans would result in a spike protein unable to bind TMPRSS2. We tested this prediction by using a split-GFP assay for syncytia formation in TMPRSS2-expressing cells.¹ We introduced the W89A mutation in the NTD, which prevents sialo glycan binding¹¹ and generates spike proteins that display the same closed conformation as the wild-type spike in the apo form, as determined by cryo-EM.¹³ In line with these observa-

tions, the HKU1B W89A spike protein did not induce syncytia, despite normal surface expression (Figures 1H and S2). These data confirmed that TMPRSS2 acts as entry receptor only after the allosteric conformational change induced on the spike by sialo glycan binding to the NTD.

Functional analysis of the RBD-TMPRSS2 interface Mutagenesis of the HKU1-CoV spike protein

To establish the importance of the residues at the interface of the HKU1B RBD/TMPRSS2 complex, we used a battery of assays (Figure 2). We introduced mutations in plasmids coding for the HKU1B spike, generating 16 mutants that were properly expressed in 293T cells as assessed by flow cytometry (Figure S2A). We co-transfected the spike mutants with TMPRSS2 to evaluate their fusogenic properties in the syncytia formation assay. Most of the tested spike variants showed a significantly reduced ability to induce cell-cell fusion, especially when the mutations involved residues that form salt bridges (K487_R and D505_R) or are part of the hydrophobic cluster formed at the interface (L510_R, V512_R, L521_R, Y528_R) (Figure 2A). The effects on cell-cell fusion were further confirmed by introducing selected mutations on soluble RBDs and testing the binding to TMPRSS2^{S441A} by biolayer interferometry (BLI) (Figure 2B). Overall, the mutants showed diminished binding to the receptor measured at the stationary state, with L510_R, Y528_R, and K487_R having the most significant effects.

Figure 1. The HKU1B-RBD recognizes TMPRSS2 with a unique binding mode

(A) Schematic representation of the HKU1B SD1-RBD and TMPRSS2 constructs used for crystallization experiments. The different (sub)domains are indicated with colors, and the residues at the boundaries are below the scheme. TMPRSS2 serine-peptidase (SP) domain is formed by two β-barrels, shown in different colors. Important TMPRSS2 residues, such as H296-D345-S441 (catalytic triad, indicated with a star) and D435 (part of the S1 site) are on top of the scheme. (S441) indicates that this active site residue was mutated to alanine in the crystallized structure. The scissor indicates the TMPRSS2 autocleavage site.

(B) Crystal structure of the HKU1B-RBD in complex with TMPRSS2^{S441A}. Crystals were obtained for the ternary complex with VHH-A01, but for better clarity, the nanobody is not displayed (the structure of the ternary complex is displayed in Figure S1B). Both proteins are colored according to the (sub)domain code presented in (A). Important elements on the RBD, such as the loops that form jaws 1 (j1) and 2 (j2), are indicated. TMPRSS2 loops at the interface are labeled L1 (loop 1, residues 427–441), L2 (loop 2, residues 462–471), L3 (loop 3, residues 412–423), and LC (loop C, residues 334–346).

(C) Description of relevant contacts at the interface between the RBD and TMPRSS2^{S441A}. The left panel shows polar residues that form salt bridges or hydrogen bonds (dashed lines), while the right panel shows hydrophobic and aromatic residues. The subscript of each residue indicates if it is part of the RBD (R) or TMPRSS2 (T). The residues that form the catalytic triad (H296, D345, S441A) are colored in purple.

(D) Open-book representation of the RBD-TMPRSS2^{S441A} complex. The contact surface was colored indicating residues that form polar interactions (dark blue for RBD residues, light magenta for TMPRSS2), hydrophobic and aromatic residues (white), and residues that are at the interface and establish van der Waals contacts (light blue for the RBD and pink for TMPRSS2^{S441A}).

(E) Scheme highlighting residues buried at the RBD-TMPRSS2^{S441A} interface (blue/pink shade). Amino acids that form salt bridges are indicated with filled triangles, while those that form hydrogen bonds are indicated with the empty symbol. For comparison, the sequence from the HKU1A RBD (isolate N1) is aligned below the one from HKU1B. Residues that are important for TMPRSS2 activity are indicated with stars (purple: catalytic triad; red: S1 site; orange: S2 site). Loops and β-strands are indicated, as in the previous panels.

(F) Comparison between the structures of different coronavirus RBDs (light blue) in complex with their receptors (dark pink). Two examples of betacoronaviruses (SARS-CoV-2 in complex with ACE2, PDB: 6M0J; MERS-CoV in complex with DPP4, PDB: 4L72), two of alphacoronaviruses (NL63-CoV with ACE2, PDB: 3KBH; HCoV-229E with APN, PDB: 6U7G), and one of deltacoronaviruses (PDCoV and human APN, PDB: 7VPQ) were selected. The buried surface area (BSA) of each complex is indicated below.

(G) The structure of the RBD-TMPRSS2^{S441A} complex was aligned to the RBD from the closed HKU1A Spike (apo-closed, PDB: 8OHN), the closed Spike with a disialoglycan (Sia) in the NTD (holo-closed, PDB: 8OPM), and the Spike in an open form with a disialoglycan (Sia) (holo-1-RBD-up, PDB: 8OPN). An inset on the first panel zooms into the region indicated with an oval to better visualize the clashes between TMPRSS2 bound to a protomer in the “down” conformation and amino acids and glycans from the other chains. Each protomer of the Spike is indicated with different colors (dark blue, light blue, and white), and glycans are colored in green.

(H) 293T cells expressing either GFP1-10 or GFP11 (GFP split system) were transfected with TMPRSS2 and HKU1B (wild type or harboring the W89A mutation). Cell-cell fusion was quantified by measuring the GFP area after 20 h. Data are mean ± SD of three independent experiments. The dotted line indicates a normalized fusion of 1.0 relative to wild-type Spike. Statistical analysis: (H) One-way ANOVA with Dunnett’s multiple comparison test compared to the WT TMPRSS2 on non-normalized log-transformed data. ****p* < 0.001. See also Figure S1 and Table S1.

Table 1. Data collection and refinement statistics

	RBD+TMPRSS2+VHH-A01	TMPRSS2+VHH-A07	TMPRSS2+VHH-A07 ¹
PDB code	8S0M	8S0L	8S0N
Data collection	–	–	–
Space Group	P 6 ₁	C 1 2 1	C 1 2 1
a, b, c (Å)	201.9, 201.9, 210.3	182.7, 53.8, 65.5	161.9, 54.5, 165.8
a, b, g (deg)	90.0, 90.0, 120.0	90.0, 100.7, 90.0	90.0, 108.4, 90.0
Resolution (Å)	25–3.55 (3.64–3.55)	25–1.80 (1.85–1.80)	25–2.30 (2.39–2.30)
Rmerge	0.265 (12.876)	0.106 (2.643)	0.240 (2.236)
Mean(I)/sd(I)	14.6 (0.4)	9.2 (0.6)	6.6 (0.9)
Number of reflections	58457 (4358)	58045 (4194)	50880 (2545)
Completeness (%)	99.6 (100.0)	99.8 (98.3)	82.7 (39.4)
Comp. (ellipsoidal)	–	–	95.0 (79.2)
Multiplicity	43.3 (41.4)	6.4 (6.3)	7.0 (6.9)
CC(1/2)	1.00 (0.223)	0.998 (0.293)	0.990 (0.316)
Refinement			
Resolution (Å)	25–3.55	25–1.80	25–2.30
Number of reflections	58058	57982	50824
Rwork/Rfree	19.29/22.13	17.69/21.54	20.87/24.71
NCS	2	–	2
No. of atoms			
Protein	12567	4003	7032
Ligand/Carb/Water	95	389	207
B-factor			
Clashscore	6.67	3.70	3.92
RMS deviations			
Bond lengths (Å)	0.004	0.010	0.002
Bond angles (°)	0.788	1.013	0.493
Ramachandran			
Favored (%)	96.05	96.15	96.52
Outliers (%)	0.13	0.20	0.11
Rotamer outliers	1.00	0.46	0.26
Diffraction limit #1: 2.796:(0.8883, 0.0000, 0.4594): 0.984 a* + 0.177 c*.			
Diffraction limit #2: 2.101:(0.0000, 1.0000, 0.0000): b*.			
Diffraction limit #3: 2.241:(-0.4594, 0.0000, 0.8883): -0.413 a* + 0.911 c*.			
¹ Dataset treated with staraniso (http://staraniso.globalphasing.org/cgi-bin/staraniso.cgi). ²⁷			

Design and analysis of TMPRSS2 mutants

We assessed the effects of TMPRSS2 mutations within the identified HKU1B interface on syncytia formation (Figure 2C), binding of soluble HKU1B spike (Figure 2D), and HKU1B pseudovirus entry (Figure 2E). We generated 19 mutants that had similar or slightly higher cell surface expression levels than the wild-type protein, as determined by flow cytometry (Figure S2B). To assess their proteolytic activity, we used another seasonal coronavirus (229E) as a reporter system. When co-expressed with its receptor, aminopeptidase N (APN),³³ the 229E spike is unable to trigger cell-cell fusion in 293T (Figure S2B). Addition of wild-type TMPRSS2, but not catalytically inactive TMPRSS2^{S441A}, triggers cell-cell fusion in 293T cells co-transfected with APN and 229E spike (Figure S2B). All the TMPRSS2 mutants studied induced fusion in cells co-expressing APN and 229E spike, indicating

that they cleave the 229E spike and are catalytically active (Figure S2B). Co-transfecting the wild-type HKU1B spike with TMPRSS2 variants showed that essentially all mutations at the receptor binding interface had a significantly reduced capacity to induce HKU1B-mediated cell-cell fusion (Figure 2C). As expected, the control mutant R316A_T, located in loop E and not involved in receptor binding, induced normal cell-cell fusion.

We also analyzed by flow cytometry the binding of soluble HKU1B spikes to 293T cells expressing the TMPRSS2 variants. Most of the mutants impaired in cell-cell fusion showed decreased binding (Figure 2D). The most affected mutants were those that abrogated salt bridges (D417A_T and R470A_T) or hydrogen bonds (Y469A_T) or disrupted the hydrophobic cluster at the interface (L410R_T, Y414A_T, L430R_T, W461A_T). The control R316A_T showed normal binding levels. We also used

Table 2. Buried surface area and number of polar interactions in the crystal structures

TMPRSS2 Partner	Partner Δ ASA (\AA^2)	TMPRSS2 Δ ASA (\AA^2)	Number of H-Bonds	Number of salt bridges
HKU1B RBD ^a	843	820	10	2
–	850	840	8	1
VHH-A01 ^a	1242	1143	11	1
–	1287	1195	10	1
VHH-A07 crystal 1 Cleaved	1439	1219	13	5
VHH-A07 crystal 1 Zymogen	1303	1145	9	3
VHH-A07 crystal 2 ^a	1044	1193	11	3
–	1047	1182	11	4

^a2 complexes in the crystal asymmetric unit.

TMPRSS2-expressing cells to determine their sensitivity to lentiviral pseudotypes bearing the HKU1B spike (Figure 2E). We previously reported that mutations abrogating TMPRSS2 protease activity (R225Q and S441A) mediate viral infection at similar levels as the wild-type protein because the pseudoviruses can follow an endocytic route of entry.¹ While the mutation R316A_T did not affect pseudovirus infection, the TMPRSS2 variants that showed reduced binding to the soluble spike were also inefficient for viral entry. The Y416A_T mutant did not trigger pseudovirus entry (Figure 2E) or induce cell-cell fusion (Figure 2C).

Overall, as summarized in the heatmaps (Figure 2F), our characterization of more than 30 point-mutants indicates that changes in residues at the interface of the HKU1B-RBD/TMPRSS2 complex impair binding, fusogenicity, and infection, validating the crystal structure.

TMPRSS2 residues at position 417 and 469 are key to define host tropism

We identified six positions in the TMPRSS2 SP domain—Y414_T, Y416_T, D417_T, L419_T, L430_T, W461_T, L469_T, and R470_T—as the most relevant for the interaction with the HKU1B RBD. We then examined the pattern of conservation of these residues across different TMPRSS2 orthologs by aligning 201 mammalian TMPRSS2 sequences between amino acids 340 and 473 (numbering corresponding to human TMPRSS2), which span the interface with the HKU1B RBD. The resulting sequence logo (Figure 3A) indicated that among the residues listed above, Y416_T, L419_T, L430_T, and W461_T are strictly conserved, while Y414_T is conserved in more than 80% of the analyzed species. In contrast, the variability is high at positions 417 and 469, where D and Y, the respective residues in humans, are not frequently found. Many species have a non-charged residue in the place of human D417, while polar (N) or non-polar (L, F) residues are found in the place of human Y469. These observations suggest that residues at positions 417 and 469 may be determinants of TMPRSS2 function as HKU1 receptor in different mammalian species. We tested this hypothesis by assessing the capacity of TMPRSS2 from selected mammals (macaque, mouse, hamster, and ferret), to induce HKU1-CoV cell-cell fusion and pseudovirus infection. The respective TMPRSS2s display [D; N], [N; L], [N; L], and [N; N] instead of [D; Y] as in humans at positions [417; 469] (Figure 3D). Despite similar expression levels (Figure S2C), only human TMPRSS2 induced

syncytia in 293T cells co-expressing the HKU1B spike (Figure 3B, left). The macaque TMPRSS2 (which differs only at position 469), allowed reduced pseudovirus entry (Figure 3B, right). Introducing the N469Y mutation into macaque TMPRSS2 restored syncytia formation and increased pseudovirus infection by ~10-fold (Figure 3C). Changing Q467 on the macaque protein to the mostly conserved lysine did not show further significant effects on cell-cell fusion or pseudovirus entry (Figure 3C). With mouse TMPRSS2, the double mutant engineered to change the [N; L] to the human [D; Y] motif restored syncytia formation and pseudovirus entry, while the two single mutants were poorly or non-functional (Figure 3C).

Nanobody A07 inserts its CDR3 into the TMPRSS2 substrate-binding groove

We next crystallized the complex between TMPRSS2^{S441A} and the inhibitory nanobody A07 (Figures 4A and S3) and determined its structure to 1.8 Å resolution (Table 1). A07 covers the active site cleft and buries a large surface area (about 2,600 Å², ~1,400 Å² on the VHH, and ~1,200 Å² on TMPRSS2), interacting with residues in exposed loops of the SP domain (loops 1, 2, 3, A, B, C, D), some also involved in the interaction with HKU1B RBD (Figure 4A). Superimposing the TMPRSS2^{S441A}/RBD structure on the TMPRSS2^{S441A}/A07 complex showed clashes between the nanobody and the RBD (Figure 4B), explaining the blocking activity and further validating the interaction site that we report here for the RBD.

Nanobody A07 contacts TMPRSS2 almost exclusively through its complementarity-determining regions (CDRs). The most notorious feature is the insertion of its long (21 residues) CDR3 in the substrate binding cleft in between the two lobes of the SP domain. Superposing the complex with the peptide-bound structures of the homologous hepsin and TMPRSS13 proteases shows that the side chain of R103_{A07} occupies the TMPRSS2 P1 position, making contacts with residues D435_T, S436_T, and G464_T (Figure 4C), which form the S1 site.²¹

The TMPRSS2^{S441A}/A07 crystals displayed electron density for the TMPRSS2 LDLR-A domain, which was not resolved in the previous structure of TMPRSS2.²¹ The LDLR-A domain includes a calcium-binding site formed by the side chains of D134_T, H138_T, D144_T, and E145_T and the main chain carbonyl groups of N131_T and V136_T (Figure S4A).

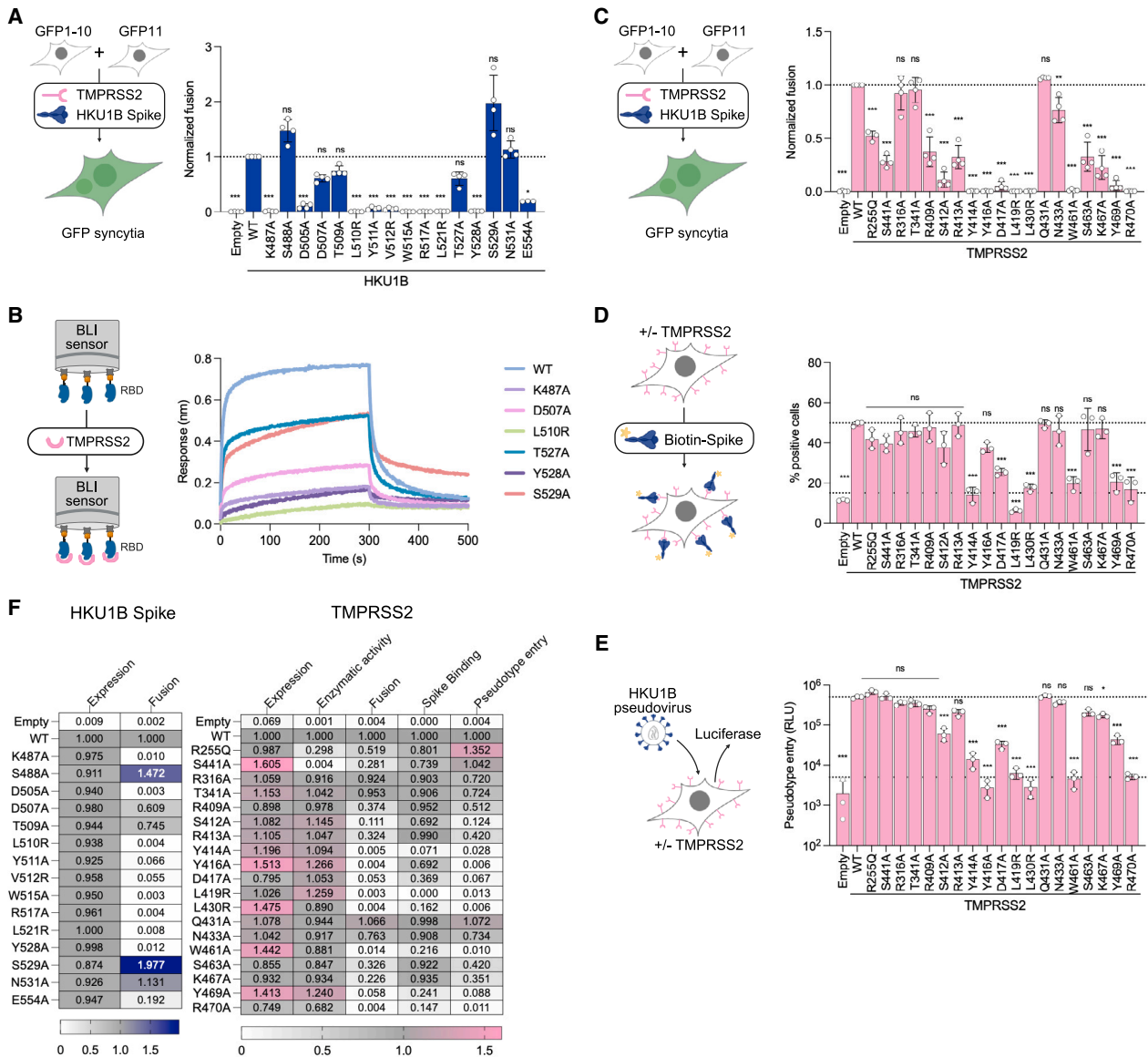


Figure 2. Functional experiments validate the RBD-TMPRSS2 crystal structure

(A) 293T cells expressing either GFP1-10 or GFP11 (GFP split system) were transfected with HKU1B Spike and TMPRSS2, and fusion was quantified by measuring the GFP area after 20 h. Data are mean \pm SD of three independent experiments. The dotted line indicates a normalized fusion of 1.0 relative to wild-type TMPRSS2.

(B) Biolayer interferometry (BLI) experiment performed with soluble HKU1B RBDs containing mutations in the binding interface. The RBDs were immobilized to Ni-NTA sensors and their interaction with soluble TMPRSS2^{S441A} (800 nM) was followed in real time. One representative experiment of three is presented.

(C) Cell-cell fusion experiment carried out transfecting the wild-type HKU1B Spike with TMPRSS2 WT or different mutants. Fusion was quantified by measuring the GFP area after 20 h. Data are mean \pm SD of four independent experiments. The dotted line indicates a normalized fusion of 1.0 relative to wild-type TMPRSS2.

(D) 293T cells were transfected with plasmids coding for TMPRSS2 or mutated variants, they were incubated with biotinylated soluble HKU1B Spike and binding to cells was determined by flow cytometry using labeled streptactin. Data are mean \pm SD of three independent experiments. The upper dotted line indicates the percentage of transfected cells with wild-type TMPRSS2 that were considered positive, while the bottom dotted line shows the background levels in non-transfected cells.

(E) 293T cells transfected with TMPRSS2 WT or mutant variants were infected by luciferase-encoding HKU1B pseudoviruses. Luminiscence was read 48 h postinfection. RLU, relative light unit. The upper dotted line indicates the mean RLU obtained when cells were transfected with wild-type TMPRSS2, while the lower dotted line shows the background levels in non-transfected cells. Data are mean \pm SD of three independent experiments.

(F) Heatmaps summarizing the functional data from (A) to (E), along with the results from controls of expression and enzymatic activity (supplementary information). Statistical analysis: (A, C, and E) One-way ANOVA with Dunnett's multiple comparison test compared to the WT on non-normalized log-transformed data. (D) One-way ANOVA with Dunnett's multiple comparison test compared to the WT. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

See also [Figures S2](#) and [S7](#).

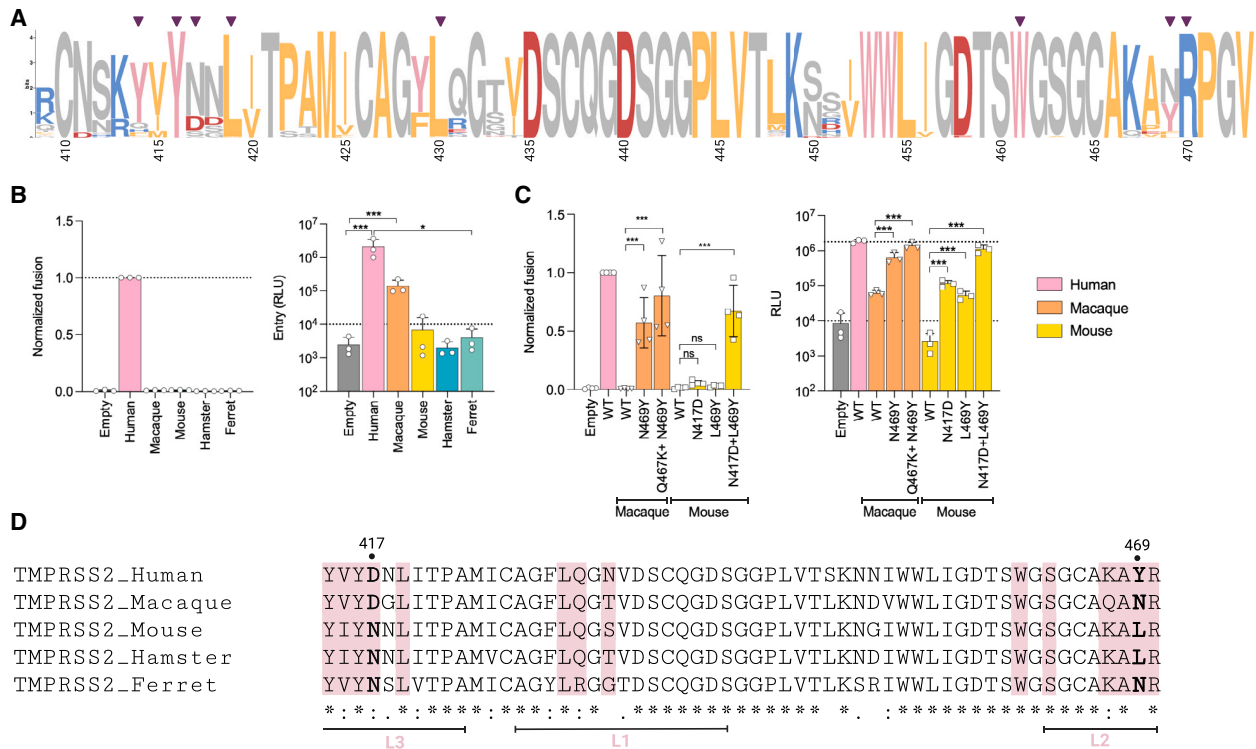


Figure 3. TMPRSS2 residues at positions 417 and 469 influence binding to different homologues

(A) Frequency of different amino acids occupying positions 409 to 473 (numbers at the bottom) of TMPRSS2 orthologs. The logo was obtained using 201 sequences from mammals. Purple inverted triangles indicate the positions identified as relevant for human TMPRSS2 in functional experiments (Figure 2). (B) Cell-cell fusion (left) and pseudovirus entry (right) experiments performed with cells transfected with TMPRSS2 from different species. Data are mean \pm SD of three independent assays. The dotted line indicates a normalized fusion of 1.0 relative to human TMPRSS2 or the mean RLU obtained with human TMPRSS2. (C) Cell-cell fusion (left) and pseudovirus entry (right) experiments performed with cells transfected with wild-type or mutant TMPRSS2 from selected mammals. Data are mean \pm SD of three independent experiments. The dotted line indicates a normalized fusion of 1.0 relative to human TMPRSS2 or the mean RLU obtained with human TMPRSS2.

(D) Sequence alignment of the TMPRSS2 homologues used for the experiments in (B), highlighting the residues that are part of the interface formed with the HKU1-RBD (pink shade) and indicating TMPRSS2 loops L1 (loop 1), L2 (loop 2), and L3 (loop 3). Statistical analysis: (B and C) One-way ANOVA with Šídák’s multiple comparison test * $p < 0.05$, *** $p < 0.001$.

See also Figure S2.

The SP domain of cleaved TMPRSS2 (TMPRSS2-CI) crystallized previously in complex with the Nafamostat inhibitor²¹ superposes very well with its counterpart in the TMPRSS2^{S441A}/A07 crystals, with root-mean-square (RMS) deviation of 0.315 Å for 1,502 atoms in 236 residues. In this structure, the residues immediately downstream from the autocleavage site (I256_T and V257_T) are found in an internal pocket where the free amino group of I256_T forms a salt bridge with the side chain of D440_T (Figure 4D). This interaction can only be established after cleavage of the protease and is a characteristic feature of serine proteases in the active conformation.³⁴ This observation strongly suggests that TMPRSS2^{S441A} in the crystals of the complex with A07 underwent cleavage by a contaminating protease. The crystallographic data further indicated that the maturation cleavage occurred in 53% of the molecules forming the crystal. Another conformation was detected, for which the high-resolution diffraction allowed the refinement of an atomic model (Table 1). In this second model, which refined to an occupancy of 47%, the loop bearing the cleavage site was disordered,

with the first residue with clear electron density after the cleavage site being G259_T, as expected for an uncleaved form of TMPRSS2^{S441A}. We confirmed that this is indeed the case by crystallizing the TMPRSS2^{S441A}/A07 complex under different conditions, which yielded crystals that diffracted to 2.4 Å resolution. The residues I256_T and V257_T were not visible in the structure determined from these new crystals, in which TMPRSS2^{S441A} showed no evidence of proteolysis and was present with an occupancy of 100%. The resulting model aligned very well with the second conformation described above, supporting the hypothesis that it corresponds to the TMPRSS2 zymogen. Furthermore, this structure has the characteristic “zymogen triad” initially observed for chymotrypsinogen,³⁵ corresponding to D440_T-H279_T-S272_T in TMPRSS2 (Figure S4B). Upon activation, D440_T is released from this triad to make the salt bridge/hydrogen bond with the newly formed N terminus at I256_T. This rearrangement leads to formation of the oxyanion hole required for cleavage of the scissile peptide bond of the substrate. Proteases that do not have a zymogen triad, such

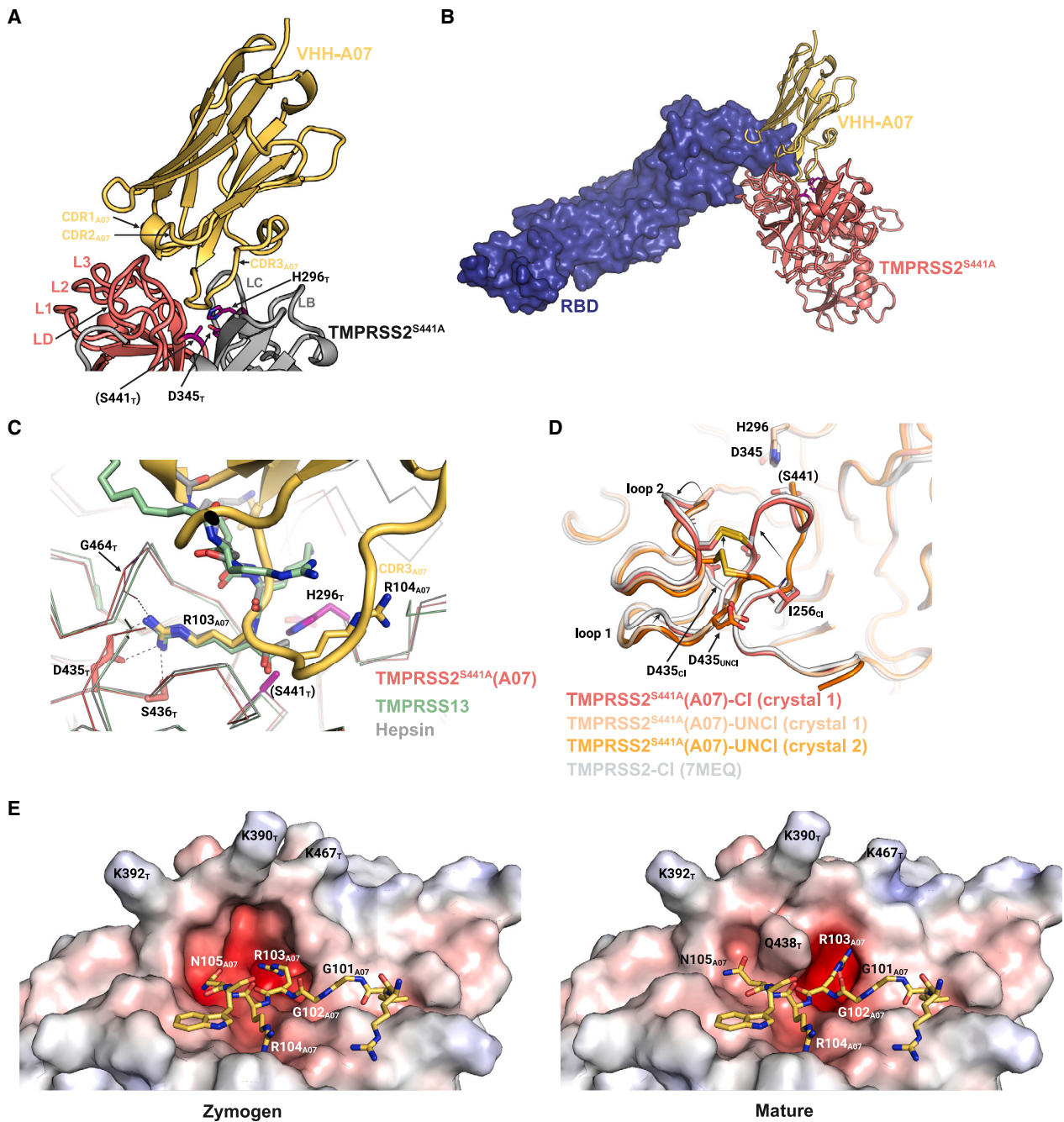


Figure 4. VHH-A07 blocks TMPRSS2 activity and binding to the HKU1-RBD by inserting its CDR3 in the active site

(A) Crystal structure of the nanobody VHH-A07 (yellow) complexed to TMPRSS2^{S441A}. The two β-barrels that form the SP domain are colored in gray (β-barrel 1) and light red (β-barrel 2). Important structural elements are indicated in the nanobody (CDR1_{A07}, CDR2_{A07}, CDR3_{A07}), as well as in TMPRSS2^{S441A} (loops LB, LC, LD, L1, L2, L3). Residues from the catalytic triad are shown in purple. Subscripts in the labels identify the protein. The active site S441 was mutated to alanine in the crystallized construct, so it is annotated between parentheses.

(B) Superposition of the TMPRSS2^{S441A} + RBD and TMPRSS2^{S441A} + VHH-A07 complexes. For simplicity, TMPRSS2 from the complex with the RBD is not shown.

(C) Superposition of TMPRSS2^{S441A} in complex with VHH-A07 (protease in light red, nanobody in yellow) with the structures of TMPRSS13 (green, PDB: 6KD5) and Hepsin (gray, PDB: 1Z8G) bound to substrate-analog inhibitors (decanoyl-RVKR-chloromethylketone shown in green sticks and Ac-KQLR-chloromethylketone in gray sticks, respectively). For clarity, the proteases are shown in ribbon representation and the nanobody in cartoon. Residues from the TMPRSS2 catalytic triad are indicated in purple, with S441 labeled with parenthesis to indicate that it is mutated to alanine in the crystallized construct. Residues that form the S1 site (D435, S436, G464) are indicated, as well as R103 from nanobody A07.

(legend continued on next page)

as the tissue-type plasminogen activator, have a high level of catalytic activity in the zymogen form.³⁶

Comparison of the structures of cleaved and zymogen TMPRSS2^{S441A} revealed that the main changes are localized to loops 1 (L430_T-D440_T) and 2 (G462_T-K467_T). In the zymogen, residues 430–440 occupy an external position, and S1 site residues (D435_T, S436_T) are away from the active site (Figure 4D).

As mentioned above, in TMPRSS2^{S441A}-CI, the free N terminus of I256_T flips to the interior of the molecule, displacing the 430–440 segment toward the core of the domain (Figure 4D). The disulfide bond between C437_T and C465_T propagates this movement to loop 2 (G462_T-K467_T), which also adopts a different conformation (Figure 4D). Consequently, the S1 site is formed without alteration of the catalytic triad. Of note, the conformation of loops 1 and 2 in TMPRSS2^{S441A}-CI is the same as that observed in the active forms of hepsin³⁷ and TMPRSS13³⁸ (Figure S4C).

TMPRSS2 maturation affects binding affinity toward the HKU1-CoV RBD

As described above, maturation of TMPRSS2 into an active form implies changes in the conformation of loops 1 and 2. Since they are part of the HKU1B RBD binding site, we hypothesized that TMPRSS2 cleavage may impact the interaction with the RBD. We therefore conducted BLI experiments with immobilized RBD, soluble TMPRSS2^{S441A}, and cleaved soluble TMPRSS2^{S441A} (TMPRSS2^{S441A}-CI). The latter was obtained by digestion of the zymogen with wild-type TMPRSS2. Sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) confirmed the complete digestion of TMPRSS2^{S441A}, generating fragments of ~25–30 kDa associated by a disulfide bond (Figure S5A). The BLI curves showed that TMPRSS2^{S441A}-CI displays a dramatic increase in the association signal and reduced the dissociation rate (Figure 5A). The curves corresponding to TMPRSS2^{S441A}-CI fitted well to a 1:1 binding model. We determined the kinetic association and dissociation rates and measured a dissociation constant (K_d) of 30 nM (Figure 5C). This value is ~5-fold higher than the one obtained for TMPRSS2^{S441A}-zymogen.¹ We also used TMPRSS2^{S441A}-CI to determine the affinity of some of the RBD mutants (Figure 5B). This confirmed that K487_R, L510_R, and Y528_R were the most impaired (Figure 5C), while S529_R showed increased affinity, mostly due to a lower dissociation rate. Additionally, we tested the effect of the RBD mutants on the *in vitro* TMPRSS2 activity and determined that their ability to inhibit the protease correlated with the affinity toward TMPRSS2 (Figure S5B).

We then compared the structures of TMPRSS2^{S441A}-CI and TMPRSS2^{S441A}-zymogen (obtained with A07) with that of

TMPRSS2^{S441A} in complex with the RBD (Figure 5D). In the latter structure, residues I256_T and V257_T were not visible, indicating that TMPRSS2^{S441A} is in the uncleaved form. Overall, the protease conformation matched TMPRSS2^{S441A}-zymogen, except for local differences in the loops at residues around G432_T and S463_T. In the complex with the RBD, these loops adopt a conformation closer to the one found in TMPRSS2^{S441A}-CI, where G432_T moves toward the active site, further away from the RBD interface, and S463_T flips outwards to form a hydrogen bond with the RBD (Figure 5D), explaining why the affinity of TMPRSS2^{S441A}-CI for the RBD is higher than that of the zymogen.

DISCUSSION

We report the structure of HKU1B RBD bound to its receptor TMPRSS2, revealing a unique binding mode among coronaviruses. The HKU1-CoV RBD uses a pincer plier motif at its distal end to recognize the TMPRSS2 substrate specificity loops. The interaction blocks access of substrates to the catalytic groove, inhibiting the proteolytic activity. Our results are in line with a recent report on the structure of the HKU1A RBD complexed with cleaved TMPRSS2 and the inhibition of TMPRSS2 proteolytic activity by HKU1A RBD.³⁹

The HKU1-CoV spike protein binds cell-surface disialoside glycans (9-O-Ac-Sia(α2,8)Sia) through a binding pocket in the NTD, causing the allosteric opening of the spike.¹³ This observation led to the conclusion that the sialo glycan is a primary receptor, while a proteinaceous receptor would be secondary. This mechanism ensures that the spike will not undergo premature activation before reaching a cell, limiting exposure of the receptor binding motif to neutralizing antibodies. Moreover, when interacting with an up-RBD, the active site of TMPRSS2 is far from the spike core, indicating that a single TMPRSS2 molecule cannot simultaneously cleave the spike and act as receptor. By binding to up-RBD conformations, TMPRSS2 traps partially open intermediates of the HKU1 spike, displacing the conformational equilibrium toward the open form, where S2 is unshielded and primed for fusion, acting by the same mechanism as the angiotensin-converting enzyme 2 (ACE2) with the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) spike.⁴⁰ Considering that ~45% of HKU1-CoV spike with 9-O-Ac-Sia(α2,8)Sia were found closed or with only one RBD “up,”¹³ a secondary protein receptor with high affinity toward the RBD would guarantee the unidirectionality of the conformational change toward the open spike.

We also determined the crystal structure of TMPRSS2 bound to nanobody A07, which competes with the RBD for binding.¹

(D) Superposition of different TMPRSS2^{S441A} structures obtained in complex with VHH-A07. A first crystal form (crystal 1) allowed building two atomic models, where one has the characteristic features of a cleaved (CI, light red) serine protease, while the other represents the zymogen (UNCI, wheat) form. The second crystal form (crystal 2) allowed building the model shown in orange that corresponds to uncleaved TMPRSS2^{S441A}. A previously reported structure of active TMPRSS2 is shown in white for comparison (PDB: 7MEQ). The I256 residue, which follows the autocleavage site, is indicated. Important TMPRSS2 loops are designated as well as the positions of the catalytic triad (H296, D345, S441A) and D435 from the S1 site. The disulfide C437–C465 is shown in yellow sticks. Arrows with tapered lines indicate the repositioning on relevant segments upon cleavage.

(E) Comparison between the TMPRSS2^{S441A} active site in the uncleaved (left) and cleaved (right) forms found in crystal 1. The surface of the protease was colored according to the electrostatic potential, and the residues R99_{A07}-W106_{A07} from A07 CDR3 are shown in sticks.

See also Figures S3 and S4.

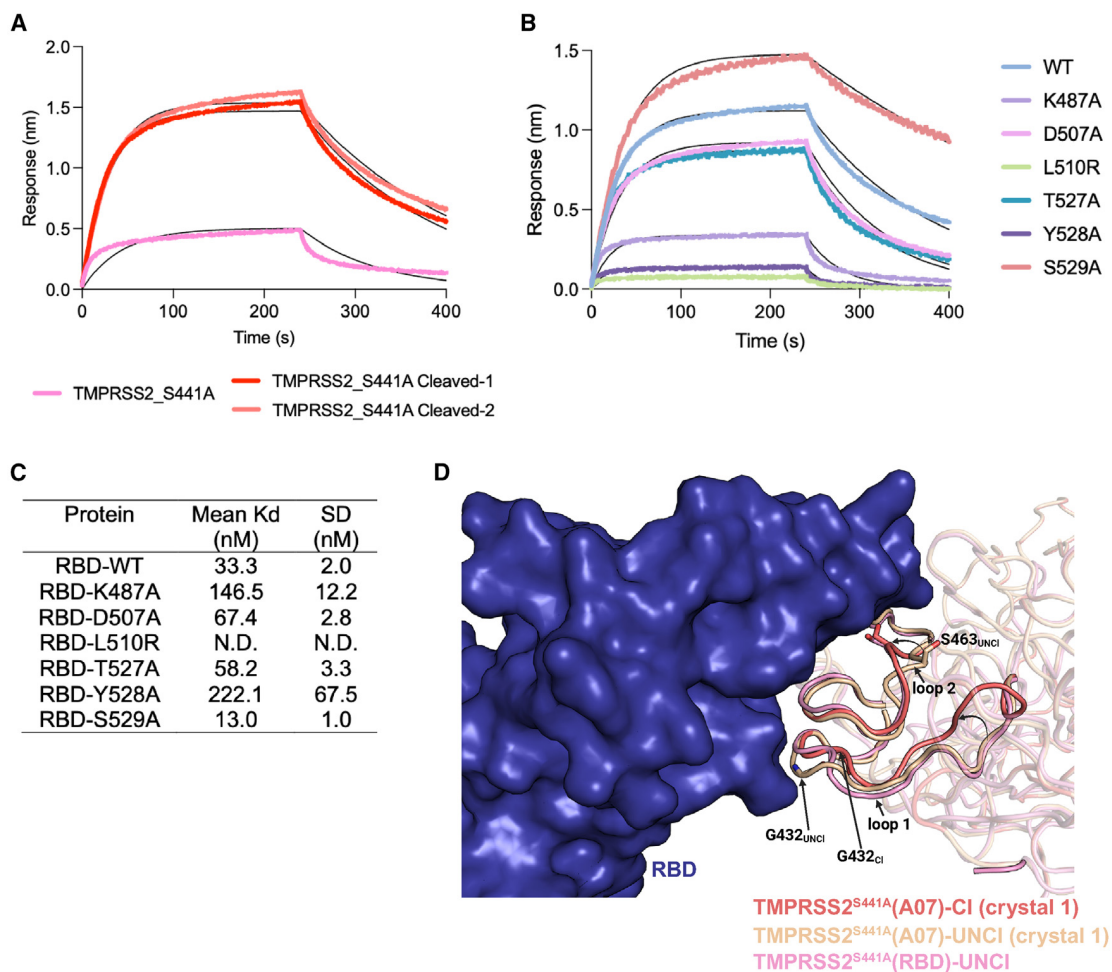


Figure 5. TMPRSS2 cleavage affects binding to the HKU1 RBD

(A) BLI experiment performed by immobilizing the HKU1B RBD and measuring the response upon incubation with 150 nM TMPRSS2^{S441A} or with 150 nM TMPRSS2^{S441A} previously cleaved by incubation with the wild-type protease. Two independent cleavage reactions were carried out. The colored curves correspond to the experimental data, and the thin black lines are the curves fitted to a 1:1 binding model. One representative experiment of three is presented. (B) BLI experiments performed by immobilizing different RBDs and measuring the response after incubation with 120 nM cleaved TMPRSS2^{S441A}. The colored curves correspond to the experimental data and the thin black lines are the curves fitted to a 1:1 binding model. One representative experiment of three is presented.

(C) Kd values of different RBDs toward cleaved TMPRSS2^{S441A}. Three experiments were performed for each RBD using a TMPRSS2 concentration range from 7.5 to 240 nM.

(D) Superposition of the TMPRSS2^{S441A} structure determined in complex with the HKU1-RBD (TMPRSS2 in pink, RBD in blue surface) or in complex with VHH-07 (crystal 1: atomic model of the cleaved form shown in light red, and the zymogen form in wheat color). Loops that form the binding site and change upon cleavage (loops 1 and 2) are labeled. The positions where conformational changes were observed are indicated with the residues that are implicated (backbone atoms of G432 and side chain of S463). Arrows with tapered lines indicate the movement of relevant segments upon TMPRSS2 autocleavage.

See also [Figures S5](#) and [S6](#).

A07 inserts its long CDR3 in the active site cleft, with R103_{A07} occupying the P1 position ([Figure 4E](#)). It has been suggested that the predicted S1' site (V280_T, H296_T, C297_T) accepts small hydrophobic P1' residues.²¹ Our structure shows R104_{A07} in the S1' site, and its side chain is well accommodated, interacting with main chain atoms from H296_T and C297_T. Given that the peptide bond R103_{A07}-R104_{A07} lies in the right position to undergo the attack of the active site S441, it is possible that the nanobody is cleaved by TMPRSS2. Nevertheless, we expect

the extensive network of polar interactions and the large BSA to sustain binding even if the antibody is cleaved.

The structures of cleaved and uncleaved forms of TMPRSS2 highlight the internal reorganization of the protease upon autocleavage. In the zymogen form, loop 1 adopts a conformation where S1 residues are away from the active site and are not in an optimal position to allow substrate binding, as illustrated in [Figure 4E](#) (left). The proteolytic cleavage releases the N terminus of I256_T, which flips toward the interior of the protein, with

formation of a salt bridge between the charged amino group of I256_T and D440_T. In turn, loop 1 is displaced and adopts a new conformation, transmitting this movement to loop 2 via a disulfide bond connecting them. Consequently, the S1 site residues (D435_T and S436_T from loop 1 and G464_T from loop 2) are brought into place to allow substrate binding. These changes agree with the general activation mechanism proposed for other serine proteases.⁴¹ Furthermore, the TMPRSS2 loops of the cleaved form adopt the same position as in other enzymes of the same family (hepsin and TMPRSS13), suggesting a stable conformation in the active form. The structure of the A07/TMPRSS2 zymogen further shows how the substrate mimicking CDR3 is accommodated in the groove in the absence of the S1 site (Figure 4E, right), which forms only after activation. Together with the inherent plasticity of the zymogen, these data can potentially inform the development of TMPRSS2-specific small molecules targeting this site.

We also established that TMPRSS2 activation affects the interaction with the HKU1-CoV RBD. In zymogen TMPRSS2^{S441A}, residues 431–433 (loop 1) adopt a conformation that may not favor binding to the RBD, which has the j2 jaw of the pincer plier facing it. Moreover, S463_T in loop 2 is too far to interact with D507_R. Upon cleavage, the 431–433 segment moves toward the core of TMPRSS2, facilitating the approach of the RBD, and S463 moves outwards, forming a hydrogen bond with D507_R. These changes in loops 1 and 2 appear to be important for binding, since the structure of zymogen TMPRSS2^{S441A} with the RBD shows that they adopt a position similar to that observed in activated TMPRSS2^{S441A}-Cl. This “mixed” conformation, with local changes restricted to segments 431–433 and 463–464, might represent a high-energy state, which would decrease the affinity. In addition, the zymogen TMPRSS2^{S441A} loops could adopt the ensemble of observed conformations, having a higher entropic content than TMPRSS2^{S441A}-Cl and resulting in lower affinity toward the HKU1-CoV RBD. The BLI curves show that zymogen TMPRSS2^{S441A} does not fit to a 1:1 binding model, while TMPRSS2^{S441A}-Cl does, in line with a heterogeneous population of conformations that becomes homogeneous upon cleavage.

In sum, the structure of the HKU1B RBD in complex with TMPRSS2 provides insights into the mode of action of the receptor and on the mechanism of spike activation. We also identify the most important residues of the interface, highlighting how they contribute to host species tropism, paving the way to characterization of potential animal reservoirs. Finally, we describe the maturation of TMPRSS2 from a zymogen to an active protease and how it impacts binding to HKU1-CoV. The observed plasticity of the TMPRSS2 zymogen stands out as a vulnerable feature of the protease that can be specifically targeted to prevent activation. Our results therefore open the way to the development of specific drugs against dysregulated TMPRSS2 in tumors without affecting other serine proteases to avoid toxicity.

Limitations of the study

We did not study the interaction of HKU1A with TMPRSS2. We previously reported that both HKU1A and HKU1B bind to and use TMPRSS2 as a functional receptor.¹ It is thus likely that the two HKU1 spikes similarly interact with TMPRSS2. The

recent description by cryo-EM of a TMPRSS2/HKU1A RBD complex³⁹ revealed a similar mode of interaction for both HKU1A and HKU1B. Also, we did not determine the structure of the trimeric HKU1-CoV spike in association with the receptor. We circumvented this problem by superposing the structure of the RBD in complex with TMPRSS2 to the known structure of the HKU1 spike.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
 - Lead contact
 - Materials availability
 - Data and code availability
- EXPERIMENTAL MODEL
 - Cells
- METHOD DETAILS
 - Construct design
 - Protein expression and purification
 - Crystallization and structural determination
 - Biolayer interferometry (BLI)
 - Sequence alignment
 - GFP-split fusion assay
 - Pseudovirus generation and infection
 - Flow cytometry
 - *In vitro* TMPRSS2 enzymatic activity
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.cell.2024.06.007>.

ACKNOWLEDGMENTS

We thank the staff from the Crystallography platform at Institut Pasteur and the synchrotron source SOLEIL (Saint-Aubin, France) for granting access to the facility. We thank the staff of the beamlines Proxima 1, especially Pierre Legrand, and Proxima 2A for their advice and assistance during X-ray data collections.

We thank Pablo Guardado-Calvo for access to the Octet-R8 instrument and Patrick England from the Molecular Biophysics platform at Institut Pasteur for useful discussions on biolayer interferometry data analysis.

This work was supported by the “URGENCE COVID-19” fundraising campaign of Institut Pasteur “Allospike” project. F.A.R. is also funded by ANR grants ANR-18-CE92-0006 PHLEBO, ANR-10-LABX-62-10 IBEID, Wellcome Trust collaborative grant no. UNS22082, and the Institut Pasteur and the Center national de la recherche scientifique. O.S. is funded by Institut Pasteur, Fondation pour la Recherche Médicale, Humanities in the European Research Area programme (DURABLE consortium and LEAPS consortium), the National Agency for AIDS Research–Emerging Infectious Diseases, the Vaccine Research Institute (ANR-10-LABX-77), Agence Nationale de la Recherche (ANR)/Fondation pour la Recherche Médicale Flash Covid PROTEO-SARS-CoV-2 and IDISCOVER, and IBEID Labex. Work with UtechS Photonic Bio-Imaging is funded by grant no. ANR-10-INSB-04-01 and Région Ile-de-France program DIM1-Health.

AUTHOR CONTRIBUTIONS

The experimental strategy was designed by I.F., N.S., S.D., J.B., O.S., and F.A.R. Construct design, protein expression, and protein purification were performed by I.F., A.A., and E.B. Crystallization and X-ray diffraction data

collection were carried out by I.F. and A.H., while data processing was performed by I.F. and S.D. The nanobodies were produced by P.L. Cell-cell fusion, spike binding, and pseudovirus entry experiments, as well as the mutagenesis of the plasmids used for these assays, were performed by N.S., W.H.B., C.B. and J.B. The biolayer interferometry experiments were performed by I.F. The original manuscript draft was written by I.F. and F.A.R., with input from N.S., S.D., J.B., and O.S. The funding was acquired by O.S. and F.A.R.

DECLARATION OF INTERESTS

I.F., N.S., E.B., P.L., J.B., O.S., and F.A.R. have a provisional patent on anti-TMPRSS2 nanobodies.

F.A.R. is a founder of *Meletios Therapeutics* and a member of its scientific advisory board.

Received: February 14, 2024

Revised: March 16, 2024

Accepted: June 5, 2024

Published: July 3, 2024

REFERENCES

- Saunders, N., Fernandez, I., Planchais, C., Michel, V., Rajah, M.M., Baquero Salazar, E., Postal, J., Porrot, F., Guivel-Benhassine, F., Blanc, C., et al. (2023). TMPRSS2 is a functional receptor for human coronavirus HKU1. *Nature* 624, 207–214. <https://doi.org/10.1038/s41586-023-06761-7>.
- Ko, C.J., Huang, C.C., Lin, H.Y., Juan, C.P., Lan, S.W., Shyu, H.Y., Wu, S.R., Hsiao, P.W., Huang, H.P., Shun, C.T., and Lee, M.S. (2015). Androgen-Induced TMPRSS2 Activates Matriptase and Promotes Extracellular Matrix Degradation, Prostate Cancer Cell Invasion, Tumor Growth, and Metastasis. *Cancer Res.* 75, 2949–2960. <https://doi.org/10.1158/0008-5472.CAN-14-3297>.
- Mukai, S., Yorita, K., Kawagoe, Y., Katayama, Y., Nakahara, K., Kamibeppu, T., Sugie, S., Tukino, H., Kamoto, T., and Kataoka, H. (2015). Matriptase and MET are prominently expressed at the site of bone metastasis in renal cell carcinoma: immunohistochemical analysis. *Hum. Cell* 28, 44–50. <https://doi.org/10.1007/s13577-014-0101-3>.
- Tomlins, S.A., Rhodes, D.R., Perner, S., Dhanasekaran, S.M., Mehra, R., Sun, X.W., Varambally, S., Cao, X., Tchinda, J., Kuefer, R., et al. (2005). Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science* 310, 644–648. <https://doi.org/10.1126/science.1117679>.
- Wang, Z., Wang, Y., Zhang, J., Hu, Q., Zhi, F., Zhang, S., Mao, D., Zhang, Y., and Liang, H. (2017). Significance of the TMPRSS2:ERG gene fusion in prostate cancer. *Mol. Med. Rep.* 16, 5450–5458. <https://doi.org/10.3892/mmr.2017.7281>.
- Lucas, J.M., Heinlein, C., Kim, T., Hernandez, S.A., Malik, M.S., True, L.D., Morrissey, C., Corey, E., Montgomery, B., Mostaghel, E., et al. (2014). The androgen-regulated protease TMPRSS2 activates a proteolytic cascade involving components of the tumor microenvironment and promotes prostate cancer metastasis. *Cancer Discov.* 4, 1310–1325. <https://doi.org/10.1158/2159-8290.CD-13-1010>.
- Woo, P.C.Y., Lau, S.K.P., Chu, C.M., Chan, K.H., Tsoi, H.W., Huang, Y., Wong, B.H.L., Poon, R.W.S., Cai, J.J., Luk, W.K., et al. (2005). Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. *J. Virol.* 79, 884–895. <https://doi.org/10.1128/JVI.79.2.884-895.2005>.
- Kahn, J.S., and McIntosh, K. (2005). History and recent advances in coronavirus discovery. *Pediatr. Infect. Dis. J.* 24, S223–S226. discussion S226. <https://doi.org/10.1097/01.inf.0000188166.17324.60>.
- Tortorici, M.A., and Veesler, D. (2019). Structural insights into coronavirus entry. *Adv. Virus Res.* 105, 93–116. <https://doi.org/10.1016/bs.aivir.2019.08.002>.
- Huang, X., Dong, W., Milewska, A., Golda, A., Qi, Y., Zhu, Q.K., Marasco, W.A., Baric, R.S., Sims, A.C., Pyrc, K., et al. (2015). Human Coronavirus HKU1 Spike Protein Uses O-Acetylated Sialic Acid as an Attachment Receptor Determinant and Employs Hemagglutinin-Esterase Protein as a Receptor-Destroying Enzyme. *J. Virol.* 89, 7202–7213. <https://doi.org/10.1128/JVI.00854-15>.
- Hulswit, R.J.G., Lang, Y., Bakkers, M.J.G., Li, W., Li, Z., Schouten, A., Ophorst, B., van Kuppeveld, F.J.M., Boons, G.J., Bosch, B.J., et al. (2019). Human coronaviruses OC43 and HKU1 bind to 9-O-acetylated sialic acids via a conserved receptor-binding site in spike protein domain A. *Proc. Natl. Acad. Sci. USA* 116, 2681–2690. <https://doi.org/10.1073/pnas.1809667116>.
- Li, Z., Lang, Y., Liu, L., Bunyatov, M.I., Sarmiento, A.I., de Groot, R.J., and Boons, G.J. (2021). Synthetic O-acetylated sialosides facilitate functional receptor identification for human respiratory viruses. *Nat. Chem.* 13, 496–503. <https://doi.org/10.1038/s41557-021-00655-9>.
- Pronker, M.F., Creutzmacher, R., Drulyte, I., Hulswit, R.J.G., Li, Z., van Kuppeveld, F.J.M., Snijder, J., Lang, Y., Bosch, B.J., Boons, G.J., et al. (2023). Sialoglycan binding triggers spike opening in a human coronavirus. *Nature* 624, 201–206. <https://doi.org/10.1038/s41586-023-06599-z>.
- Benton, D.J., Wrobel, A.G., Xu, P., Roustan, C., Martin, S.R., Rosenthal, P.B., Skehel, J.J., and Gamblin, S.J. (2020). Receptor binding and priming of the spike protein of SARS-CoV-2 for membrane fusion. *Nature* 588, 327–330. <https://doi.org/10.1038/s41586-020-2772-0>.
- Walls, A.C., Xiong, X., Park, Y.J., Tortorici, M.A., Snijder, J., Quispe, J., Cameroni, E., Gopal, R., Dai, M., Lanzavecchia, A., et al. (2020). Unexpected Receptor Functional Mimicry Elucidates Activation of Coronavirus Fusion. *Cell* 183, 1732. <https://doi.org/10.1016/j.cell.2020.11.031>.
- Hurdiss, D.L., Drulyte, I., Lang, Y., Shamorkina, T.M., Pronker, M.F., van Kuppeveld, F.J.M., Snijder, J., and de Groot, R.J. (2020). Cryo-EM structure of coronavirus-HKU1 haemagglutinin esterase reveals architectural changes arising from prolonged circulation in humans. *Nat. Commun.* 11, 4646. <https://doi.org/10.1038/s41467-020-18440-6>.
- de Groot, R.J. (2006). Structure, function and evolution of the hemagglutinin-esterase proteins of corona- and toroviruses. *Glycoconj. J.* 23, 59–72. <https://doi.org/10.1007/s10719-006-5438-8>.
- Chen, Y.W., Lee, M.S., Lucht, A., Chou, F.P., Huang, W., Havighurst, T.C., Kim, K., Wang, J.K., Antalis, T.M., Johnson, M.D., and Lin, C.Y. (2010). TMPRSS2, a serine protease expressed in the prostate on the apical surface of luminal epithelial cells and released into semen in prostatesomes, is misregulated in prostate cancer cells. *Am. J. Pathol.* 176, 2986–2996. <https://doi.org/10.2353/ajpath.2010.090665>.
- Afar, D.E., Vivanco, I., Hubert, R.S., Kuo, J., Chen, E., Saffran, D.C., Raitano, A.B., and Jakobovits, A. (2001). Catalytic cleavage of the androgen-regulated TMPRSS2 protease results in its secretion by prostate and prostate cancer epithelia. *Cancer Res.* 61, 1686–1692.
- Tseng, C.C., Jia, B., Barndt, R., Gu, Y., Chen, C.Y., Tseng, I.C., Su, S.F., Wang, J.K., Johnson, M.D., and Lin, C.Y. (2017). Matriptase shedding is closely coupled with matriptase zymogen activation and requires de novo proteolytic cleavage likely involving its own activity. *PLoS One* 12, e0183507. <https://doi.org/10.1371/journal.pone.0183507>.
- Fraser, B.J., Beldar, S., Seitova, A., Hutchinson, A., Mannar, D., Li, Y., Kwon, D., Tan, R., Wilson, R.P., Leopold, K., et al. (2022). Structure and activity of human TMPRSS2 protease implicated in SARS-CoV-2 activation. *Nat. Chem. Biol.* 18, 963–971. <https://doi.org/10.1038/s41589-022-01059-7>.
- Bertram, S., Dijkman, R., Habjan, M., Heurich, A., Gierer, S., Glowacka, I., Welsch, K., Winkler, M., Schneider, H., Hofmann-Winkler, H., et al. (2013). TMPRSS2 activates the human coronavirus 229E for cathepsin-independent host cell entry and is expressed in viral target cells in the

- respiratory epithelium. *J. Virol.* 87, 6150–6160. <https://doi.org/10.1128/JVI.03372-12>.
23. Glowacka, I., Bertram, S., Müller, M.A., Allen, P., Soilleux, E., Pfefferle, S., Steffen, I., Tsegaye, T.S., He, Y., Gnirss, K., et al. (2011). Evidence that TMPRSS2 activates the severe acute respiratory syndrome coronavirus spike protein for membrane fusion and reduces viral control by the humoral immune response. *J. Virol.* 85, 4122–4134. <https://doi.org/10.1128/JVI.02232-10>.
 24. Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., Schiergens, T.S., Herrler, G., Wu, N.H., Nitsche, A., et al. (2020). SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* 181, 271–280.e8. <https://doi.org/10.1016/j.cell.2020.02.052>.
 25. Limburg, H., Harbig, A., Bestle, D., Stein, D.A., Moulton, H.M., Jaeger, J., Janga, H., Hards, K., Koepke, J., Schulte, L., et al. (2019). TMPRSS2 Is the Major Activating Protease of Influenza A Virus in Primary Human Airway Cells and Influenza B Virus in Human Type II Pneumocytes. *J. Virol.* 93, e00649-19. <https://doi.org/10.1128/JVI.00649-19>.
 26. Shirato, K., Kawase, M., and Matsuyama, S. (2013). Middle East respiratory syndrome coronavirus infection mediated by the transmembrane serine protease TMPRSS2. *J. Virol.* 87, 12552–12561. <https://doi.org/10.1128/JVI.01890-13>.
 27. Tickle, I.J., Flensburg, C., Keller, P., Paciorek, W., Sharff, A., Vornheim, C., and Brice, G. (2016). STARANISO (<http://staraniso.globalphasing.org/cgi-bin/staraniso.cgi>) (Cambridge, United Kingdom: Global Phasing Ltd.).
 28. Perona, J.J., and Craik, C.S. (1995). Structural basis of substrate specificity in the serine proteases. *Protein Sci.* 4, 337–360. <https://doi.org/10.1002/pro.5560040301>.
 29. Goettig, P., Brandstetter, H., and Magdolen, V. (2019). Surface loops of trypsin-like serine proteases as determinants of function. *Biochimie* 166, 52–76. <https://doi.org/10.1016/j.biochi.2019.09.004>.
 30. Perona, J.J., and Craik, C.S. (1997). Evolutionary divergence of substrate specificity within the chymotrypsin-like serine protease fold. *J. Biol. Chem.* 272, 29987–29990. <https://doi.org/10.1074/jbc.272.48.29987>.
 31. Lawrence, M.C., and Colman, P.M. (1993). Shape complementarity at protein/protein interfaces. *J. Mol. Biol.* 234, 946–950. <https://doi.org/10.1006/jmbi.1993.1648>.
 32. Kirchdoerfer, R.N., Cottrell, C.A., Wang, N., Pallesen, J., Yassine, H.M., Turner, H.L., Corbett, K.S., Graham, B.S., McLellan, J.S., and Ward, A.B. (2016). Pre-fusion structure of a human coronavirus spike protein. *Nature* 537, 118–121. <https://doi.org/10.1038/nature17200>.
 33. Yeager, C.L., Ashmun, R.A., Williams, R.K., Cardellicchio, C.B., Shapiro, L.H., Look, A.T., and Holmes, K.V. (1992). Human aminopeptidase N is a receptor for human coronavirus 229E. *Nature* 357, 420–422. <https://doi.org/10.1038/357420a0>.
 34. Pasternak, A., Ringe, D., and Hedstrom, L. (1999). Comparison of anionic and cationic trypsinogens: the anionic activation domain is more flexible in solution and differs in its mode of BPTI binding in the crystal structure. *Protein Sci.* 8, 253–258. <https://doi.org/10.1110/ps.8.1.253>.
 35. Freer, S.T., Kraut, J., Robertus, J.D., Wright, H.T., and Xuong, N.H. (1970). Chymotrypsinogen: 2.5-angstrom crystal structure, comparison with alpha-chymotrypsin, and implications for zymogen activation. *Biochemistry* 9, 1997–2009. <https://doi.org/10.1021/bi00811a022>.
 36. Madison, E.L., Kobe, A., Gething, M.J., Sambrook, J.F., and Goldsmith, E.J. (1993). Converting tissue plasminogen activator to a zymogen: a regulatory triad of Asp-His-Ser. *Science* 262, 419–421. <https://doi.org/10.1126/science.8211162>.
 37. Herter, S., Piper, D.E., Aaron, W., Gabriele, T., Cutler, G., Cao, P., Bhatt, A.S., Choe, Y., Craik, C.S., Walker, N., et al. (2005). Hepatocyte growth factor is a preferred in vitro substrate for human hepsin, a membrane-anchored serine protease implicated in prostate and ovarian cancers. *Biochem. J.* 390, 125–136. <https://doi.org/10.1042/BJ20041955>.
 38. Ohno, A., Maita, N., Tabata, T., Nagano, H., Arita, K., Ariyoshi, M., Uchida, T., Nakao, R., Ulla, A., Sugiura, K., et al. (2021). Crystal structure of inhibitor-bound human MSPL that can activate high pathogenic avian influenza. *Life Sci. Alliance* 4, e202000849. <https://doi.org/10.26508/lsa.202000849>.
 39. McCallum, M., Park, Y.J., Stewart, C., Sprouse, K.R., Brown, J., Tortorici, M.A., Gibson, C., Wong, E., Ieven, M., Telenti, A., and Veesler, D. (2024). Human coronavirus HKU1 recognition of the TMPRSS2 host receptor. Preprint at bioRxiv, 2024.01.09.574565. <https://doi.org/10.1101/2024.01.09.574565>.
 40. Ke, Z., Oton, J., Qu, K., Cortese, M., Zila, V., McKeane, L., Nakane, T., Zivanov, J., Neufeldt, C.J., Cerikan, B., et al. (2020). Structures and distributions of SARS-CoV-2 spike proteins on intact virions. *Nature* 588, 498–502. <https://doi.org/10.1038/s41586-020-2665-2>.
 41. Kerr, M.A., Walsh, K.A., and Neurath, H. (1976). A proposal for the mechanism of chymotrypsinogen activation. *Biochemistry* 15, 5566–5570. <https://doi.org/10.1021/bi00670a022>.
 42. Buchrieser, J., Duffoo, J., Hubert, M., Monel, B., Planas, D., Rajah, M.M., Planchais, C., Porrot, F., Guivel-Benhassine, F., Van der Werf, S., et al. (2020). Syncytia formation by SARS-CoV-2-infected cells. *EMBO J.* 39, e106267. <https://doi.org/10.15252/embj.2020106267>.
 43. Kodaka, M., Yang, Z., Nakagawa, K., Maruyama, J., Xu, X., Sarkar, A., Ichimura, A., Nasu, Y., Ozawa, T., Iwasa, H., et al. (2015). A new cell-based assay to evaluate myogenesis in mouse myoblast C2C12 cells. *Exp. Cell Res.* 336, 171–181. <https://doi.org/10.1016/j.yexcr.2015.06.015>.
 44. Edie, S., Zaghloul, N.A., Leitch, C.C., Klinedinst, D.K., Lebron, J., Thole, J.F., McCallion, A.S., Katsanis, N., and Reeves, R.H. (2018). Survey of Human Chromosome 21 Gene Expression Effects on Early Development in *Danio rerio*. *G3 (Bethesda)* 8, 2215–2223. <https://doi.org/10.1534/g3.118.200144>.
 45. Hoffmann, M., Hofmann-Winkler, H., Smith, J.C., Krüger, N., Arora, P., Sørensen, L.K., Søgaard, O.S., Hasselstrøm, J.B., Winkler, M., Hempel, T., et al. (2021). Camostat mesylate inhibits SARS-CoV-2 activation by TMPRSS2-related proteases and its metabolite GBPA exerts antiviral activity. *EBioMedicine* 65, 103255. <https://doi.org/10.1016/j.ebiom.2021.103255>.
 46. Iwaki, T., Figuera, M., Ploplis, V.A., and Castellino, F.J. (2003). Rapid selection of *Drosophila* S2 cells with the puromycin resistance gene. *Bio-techniques* 35, 482–486. <https://doi.org/10.2144/03353bm08>.
 47. Chavas, L.M.G., Gourhant, P., Guimaraes, B.G., Isabet, T., Legrand, P., Lener, R., Montaville, P., Sirigu, S., and Thompson, A. (2021). PROX-IMA-1 beamline for macromolecular crystallography measurements at Synchrotron SOLEIL. *J. Synchrotron Radiat.* 28, 970–976. <https://doi.org/10.1107/S1600577521002605>.
 48. Kabsch, W. (2010). Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. D Biol. Crystallogr.* 66, 133–144. <https://doi.org/10.1107/S0907444909047374>.
 49. Kabsch, W. (2010). Xds. *Acta Crystallogr. D Biol. Crystallogr.* 66, 125–132. <https://doi.org/10.1107/S0907444909047337>.
 50. Evans, P.R., and Murshudov, G.N. (2013). How good are my data and what is the resolution? *Acta Crystallogr. D Biol. Crystallogr.* 69, 1204–1214. <https://doi.org/10.1107/S0907444913000061>.
 51. Liebschner, D., Afonine, P.V., Baker, M.L., Bunkóczi, G., Chen, V.B., Croll, T.I., Hintze, B., Hung, L.W., Jain, S., McCoy, A.J., et al. (2019). Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D Struct. Biol.* 75, 861–877. <https://doi.org/10.1107/S2059798319011471>.
 52. Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* 66, 486–501. <https://doi.org/10.1107/S0907444910007493>.

53. Williams, C.J., Headd, J.J., Moriarty, N.W., Prisant, M.G., Videau, L.L., Deis, L.N., Verma, V., Keedy, D.A., Hintze, B.J., Chen, V.B., et al. (2018). MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci.* 27, 293–315. <https://doi.org/10.1002/pro.3330>.
54. Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* 372, 774–797. <https://doi.org/10.1016/j.jmb.2007.05.022>.
55. DeLano, W.L. (2002). *The PyMOL Molecular Graphics System (DeLano Scientific)*.
56. Madeira, F., Pearce, M., Tivey, A.R.N., Basutkar, P., Lee, J., Edbali, O., Madhusoodanan, N., Kolesnikov, A., and Lopez, R. (2022). Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Res.* 50, W276–W279. <https://doi.org/10.1093/nar/gkac240>.
57. Crooks, G.E., Hon, G., Chandonia, J.M., and Brenner, S.E. (2004). WebLogo: a sequence logo generator. *Genome Res.* 14, 1188–1190. <https://doi.org/10.1101/gr.849004>.
58. Planchais, C., Fernandez, I., Bruel, T., de Melo, G.D., Prot, M., Beretta, M., Guardado-Calvo, P., Duffoo, J., Molinos-Albert, L.M., Backovic, M., et al. (2022). Potent human broadly SARS-CoV-2-neutralizing IgA and IgG antibodies effective against Omicron BA.1 and BA.2. *J. Exp. Med.* 219, e20220638. <https://doi.org/10.1084/jem.20220638>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
mAb10	Planchais C. et al. ⁵⁸	N/A
anti-TMPRSS2 A01-Fc	Saunders N. et al. ¹	N/A
Alexa Fluor 647-conjugated Goat anti-Human antibody	Thermo Fisher Scientific	Cat#A-21445; RRID: AB_2535862
anti-cMyc 9E10	Thermo Fisher Scientific	Cat#M4439; RRID: AB_439694
Alexa Fluor 647-conjugated Goat anti-Human antibody	Thermo Fisher Scientific	Cat#A-21242; RRID: AB_2535811
Bacterial and virus strains		
Escherichia coli BL21(DE3)pLysS competent cells	Promega	Cat# L1195
Chemicals, peptides, and recombinant proteins		
Puromycin	InvivoGen	Cat#ant-pr-1
Penicillin-streptomycin	Gibco	Cat#15140122
Serum-free insect cell medium	HyClone	Cat#SH30913.02
Enterokinase light chain	New England Biolabs	Cat#P8070L
Endoglycosidase D	New England Biolabs	Cat#P0742L
Endoglycosidase H	New England Biolabs	Cat#P0702L
Isopropyl β-d-1-thiogalactopyranoside (IPTG)	EUROMEDEX	Cat# EU0008-B
Complete protease inhibitor cocktail	Roche	Cat#11873580001
Effectene transfection reagent	Qiagen	Cat#301427
FectoPro® DNA transfection reagent	Polyplus	Cat#101000007
Lipofectamine 2000	Thermo Fisher Scientific	Cat#11668019
Camostat	Sigma-Aldrich	Cat#SML0057
Alexa Fluor 647-conjugated streptavidin	Thermo Fisher Scientific	Cat#S21374
t-Butyloxycarbonyl-Glutamine-Alanine-Arginine-7-Amino-4-methylcoumarin (Boc-QAR-AMC)	Biotechne	Ca#ES014
Recombinant protein: HKU1B RBD	This study	N/A
Recombinant protein: HKU1B RBD K487A	This study	N/A
Recombinant protein: HKU1B RBD D507A	This study	N/A
Recombinant protein: HKU1B RBD L510R	This study	N/A
Recombinant protein: HKU1B RBD W515A	Saunders N. et al. ¹	N/A
Recombinant protein: HKU1B RBD T527A	This study	N/A
Recombinant protein: HKU1B RBD Y528A	This study	N/A
Recombinant protein: HKU1B RBD S529A	This study	N/A
Recombinant protein: SARS-CoV-2 RBD	Saunders N. et al. ¹	N/A
Recombinant protein: HKU1B RBD-SD1	This study	N/A
Recombinant protein: HKU1B Spike ectodomain	This study	N/A
Recombinant protein: Human TMPRSS2 ectodomain	This study	N/A
Recombinant protein: Human TMPRSS2 ^{S441A} ectodomain	This study	N/A
Recombinant protein: Nanobody A01	This study	N/A
Recombinant protein: Nanobody A07	This study	N/A
HIV Gag p24 concentration kit	Saunders N. et al. ¹	N/A
Critical commercial assays		
Bright-Glo™ lysis buffer	Promega	Cat#E2620
Q5 Site-Directed mutagenesis kit	New England Biolabs	Cat# E0554S
Lenti-X p24 Rapid titer kit	Takara	Cat#632200

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
EZ-Link Sulfo-NHS-Biotinylation kit	Thermo Fisher Scientific	Cat#21217
Deposited data		
Crystal structure of the HKU1B RBD in complex with TMPRSS2 ^{S441A} and nanobody A01	This study	PDB: 8S0M
Crystal structure of TMPRSS2 ^{S441A} in the zymogen and mature forms complexed to nanobody A07 (Crystal form 1)	This study	PDB: 8S0L
Crystal structure of TMPRSS2 ^{S441A} in the zymogen form complexed to nanobody A07 (Crystal form 2)	This study	PDB: 8S0N
Experimental models: Cell lines		
Drosophila melanogaster S2 cell line	Thermo Fisher Scientific	Cat#R690-07
Expi293™ cells	Thermo Fisher Scientific	Cat#A14527
293T cells	ATCC	CRL-3216
293T GFP-split cells (GFP1-10 and GFP11)	Buchrieser J. et al. ⁴²	N/A
Recombinant DNA		
HKU1B RBD (residues 330–614) with Hisx8, Strep and Avi-tags cloned in pCAGGs	Saunders N. et al. ¹ and this study	N/A
HKU1B RBD (residues 330–614) K487A with Hisx8, Strep and Avi-tags cloned in pCAGGs	This study	N/A
HKU1B RBD (residues 330–614) D507A with Hisx8, Strep and Avi-tags cloned in pCAGGs	This study	N/A
HKU1B RBD (residues 330–614) L510R with Hisx8, Strep and Avi-tags cloned in pCAGGs	This study	N/A
HKU1B RBD (residues 330–614) W515A with Hisx8, Strep and Avi-tags cloned in pCAGGs	Saunders N. et al. ¹ and this study	N/A
HKU1B RBD (residues 330–614) T527A with Hisx8, Strep and Avi-tags cloned in pCAGGs	This study	N/A
HKU1B RBD (residues 330–614) Y528A with Hisx8, Strep and Avi-tags cloned in pCAGGs	This study	N/A
HKU1B RBD (residues 330–614) S529A with Hisx8, Strep and Avi-tags cloned in pCAGGs	This study	N/A
SARS-CoV-2 RBD (residues 331–528) with Hisx8, Strep and Avi-tags cloned in pcDNA3.1(+)	Saunders N. et al. ¹	N/A
HKU1B RBD-SD1 (residues 307–675) with an enterokinase site and double-Strep tag cloned in a modified pMT/BiP plasmid	This study	N/A
HKU1B stabilized Spike ectodomain (residues residues 14–1276) with a His-tag, cloned in pcDNA3.1(+)	Saunders N. et al. ¹	N/A
Human TMPRSS2 ^{S441A} ectodomain (residues 107–492) with an enterokinase site and double-Strep tag cloned in a modified pMT/BiP plasmid	This study	N/A
Human TMPRSS2 ectodomain (residues 107–492) with an enterokinase site and double-Strep tag cloned in a modified pMT/BiP plasmid	This study	N/A
pCoPURO plasmid	Iwaki et al. ⁴⁶	Addgene #17533
Nanobody A01 with a His-tag cloned in pET23	Saunders N. et al. ¹	N/A
Nanobody A07 with a His-tag cloned in pET23	Saunders N. et al. ¹	N/A
Human TMPRSS2 with an N-terminal cMYC-epitope, cloned in pCCAGS	Hoffmann, M. et al. ⁴⁵	N/A
Human TMPRSS2 with an N-terminal cMYC-tag, cloned into a pHCMV backbone	This study	N/A
Mouse TMPRSS2 with an N-terminal cMYC-tag, cloned into a pHCMV backbone	This study	N/A

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Ferret TMPRRS2 with an N-terminal cMYC-tag, cloned into a phCMV backbone	This study	N/A
Hamster TMPRRS2 with an N-terminal cMYC-tag, cloned into a phCMV backbone	This study	N/A
Macaque TMPRRS2 with an N-terminal cMYC-tag, cloned into a phCMV backbone	This study	N/A
pQCXIP-Empty	Buchrieser J. et al. ⁴²	N/A
pCSDest-TMPRRS2	Edie S. et al. ⁴⁴	Addgene #53887

Software and algorithms

XDS	Kabsch, W. ^{48,49}	https://xds.mr.mpg.de
Aimless	Evans, P.R. and Murshudov, G.N. ⁵⁰	https://www.ccp4.ac.uk/html/aimless.html
Phenix suite	Liebschner, D. et al. ⁵¹	https://www.phenix-online.org/
AlphaFold2	Deepmind	https://www.deepmind.com
Coot	Emsley, P. et al. ⁵²	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/
Molprobit	Williams, C.J. ⁵³	http://molprobit.biochem.duke.edu/
PDBePISA	Krissinel, E., and Henrick, K. ⁵⁴	https://www.ebi.ac.uk/pdbe/pisa/
PyMOL	Schrodinger, LLC	https://pymol.org/2
Clustal Omega web server	Madeira, F. et al. ⁵⁶	https://www.ebi.ac.uk/jdispatcher/msa/clustalo
WebLogo web server	Crooks, G.E. et al. ⁵⁷	https://weblogo.berkeley.edu
Prism Version 9	GraphPad	https://www.graphpad.com

Other

Strep-Tactin Superflow high-capacity column	IBA Life Sciences	Cat#2-1238-001
HiLoad™ 16/600 Superdex™ 200 pg	Cytiva	Cat#28989335
HiLoad™ 16/600 Superdex™ 75 pg	Cytiva	Cat#28989333
HisTrap™ HP	Cytiva	Cat#17524801
HisTrap™ Excel	Cytiva	Cat#17371206
Superdex 200 increase 10/300 GL	Cytiva	Cat#28990944
Octet® NTA capture biosensors	Sartorius	Cat#18-5101

RESOURCE AVAILABILITY**Lead contact**

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Félix Rey (felix.rey@pasteur.fr).

Materials availability

Reagents generated in this study are available on request from the [lead contact](#) with a completed Materials Transfer Agreement.

Data and code availability

Coordinates and structure factors of the crystal structures have been deposited in the Protein DataBank (PDB) and are publicly available as of the data of publication. Accession numbers are listed in the [key resources table](#). This paper does not report original code. Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

EXPERIMENTAL MODEL**Cells**

Functional experiments were performed using human female embryonic kidney 293 T cells (HEK293T or 293T) from the American Type Culture Collection (ATCC) that were cultured in Dulbecco's modified Eagle medium (DMEM) with 10% fetal bovine serum

(FBS) and 1% penicillin/streptomycin (100X stock: 10,000 units penicillin and 10 mg/mL streptomycin) at 37°C. GFP-split cells were previously described⁴² and cultured with 1 µg/mL of puromycin (InvivoGen) at 37°C. Cells were routinely screened for mycoplasma. Cells were authenticated by genotyping (Eurofins).

Recombinant proteins were expressed by stably transfected *Drosophila* Schneider line 2 (S2) cell lines or by transiently transfected Expi293F cells (Thermo-Fischer), as indicated in Method details. S2 cell lines were cultured in serum-free insect cell medium (HyClone, GE Healthcare) containing 7 µg/mL puromycin and 1% penicillin/streptomycin at 28°C. Expi293F cells were maintained in Expi293 expression medium (Gibco) without antibiotics at 37°C. These cell lines were not authenticated. The supplier does not inform their sex and this information is not relevant for our study.

METHOD DETAILS

Construct design

For production of recombinant proteins

Codon-optimized synthetic genes coding for the HKU1B RBD (residues 330–614 of the spike protein, isolate N5P8, NCBI accession Q0ZME7), the RBD and subdomain-1 (RBD-SD1, residues 307–675), the spike ectodomain (residues 14–1276) and the human TMPRSS2 ectodomain (residues 107–492, NCBI accession O15393) were obtained from Genscript. Cloning and mutagenesis of these genes were also performed by Genscript. The RBD was cloned into pCAGGS, following a murine immunoglobulin kappa signal peptide, and upstream of a thrombin cleavage site and *in-tandem* Hisx8, Strep and Avi-tags. The Spike ectodomain was stabilized in the pre-fusion form by introducing mutations in the furin site (⁷⁵²RRKRR⁷⁵⁶ to ⁷⁵²GGSGS⁷⁵⁶) and in the S2 subunit (¹⁰⁶⁷NL¹⁰⁶⁸ to ¹⁰⁶⁷PP¹⁰⁶⁸) and adding a Foldon trimerization motif at the C-terminus. This construct was cloned into pcDNA3.1(+), with an Ig kappa signal peptide, a thrombin cleavage site at the C-terminus followed by a His-tag. RBD-SD1, TMPRSS2 wild-type (WT) and TMPRSS2^{S441A} were cloned into a modified pMT/BiP plasmid (Invitrogen; hereafter termed pT350), which translates the protein in frame with an enterokinase cleavage site and a double strep-tag at the C-terminal end. Selected VHHs have been previously obtained¹ and were cloned into the bacterial expression vector pET23 with a C-terminal His-tag and N-terminal Met-Ala residues introduced during subcloning.

For cell transfection and functional assays

Codon-optimized synthetic genes coding for the full-length spike of HKU1 B/C isolate N5P8 (referred to as HKU1B, UniProtKB/Swiss-Prot: Q0ZME7.1) and those coding for TMPRSS2 from mouse (Mouse C57BL/6 - UniProt: Q3UKE3), ferret (Ferret - UniProt: A0A8U0SMZ2), hamster (Syrian Hamster Isoform X1 - UniProt: A0A1U8BWQ2) and macaque (Macaque - UniProt: F6SVR2) with an N-terminal cMYC-tag were ordered to GeneArt (Thermo Fisher Scientific) and cloned into a phCMV backbone (GeneBank: AJ318514) by replacing the VSV-G gene. pQCXIP-Empty control plasmid was previously described.⁴² pQCXIP-BSR-GFP11 and pQCXIP-GFP1-10 were a kind gift from Yutaka Hata⁴³ (Addgene plasmid #68716 and #68715). pCSDest-TMPRSS2 was a kind gift from Roger Reeves⁴⁴ (Addgene plasmid # 53887). Human pCCAGS N-terminal cMYC-epitope tagged TMPRSS2 was a kind gift from Stefan Pöhlmann.⁴⁵ Mutations in the HKU1 spike and TMPRSS2 were introduced using the NEB Q5 Side-Directed mutagenesis kit. Plasmids were sequenced before usage.

Protein expression and purification

Protein expression and purification for X-Ray crystallography

Plasmids encoding RBD-SD1 or TMPRSS2^{S441A} were co-transfected with the pCoPuro plasmid⁴⁶ for puromycin selection in *Drosophila* Schneider line 2 cells (S2) using the Effectene transfection reagent (Qiagen). The cell lines underwent selection in serum-free insect cell medium (HyClone, GE Healthcare) containing 7 µg/mL puromycin and 1% penicillin/streptomycin. For protein production, the cells were grown in spinner flasks until the density reached 10⁷ cells/mL, at which point the protein expression was induced with 4 µM CdCl₂. After 6 days, the cultures were centrifuged, and the supernatants were concentrated and used for affinity purification in a Strep-Tactin column (IBA). The strep tags were removed by incubating the proteins with 64 units of Enterokinase light chain (BioLabs) in 10 mM Tris, 100 mM NaCl, 2 mM CaCl₂, pH 8.0, at room temperature, overnight. The proteolysis reactions were buffer-exchanged into 10 mM Tris, 100 mM NaCl, pH 8.0, and subjected to a second affinity purification, recovering the flow-through fraction containing the untagged proteins. The proteins were concentrated and the enzymatic deglycosylation with endoglycosidase D (EndoD) and endoglycosidase H (EndoH) was set up at room temperature following overnight incubation with 1000 units of each enzyme in 50 mM Na-acetate, 200 mM NaCl, pH 5.5. The proteins were further purified on a size exclusion chromatography (SEC) Superdex 200 16/600 (Cytiva) column in 10 mM Tris, 100 mM NaCl, pH 8.0, and concentrated in VivaSpin concentrators.

Escherichia coli BL21 pLysS cells were transformed with the plasmids encoding the VHHs, which were expressed in the cytoplasm after overnight induction with 0.5 mM isopropyl β-*D*-1-thiogalactopyranoside (IPTG) at 16°C. The cultures were centrifuged, the bacterial pellets were resuspended in 40 mL of lysis buffer (20 mM Tris-HCl, 200 mM NaCl, 20 mM imidazole, pH 8.0) containing complete protease inhibitor cocktail (Roche) and they were frozen at –80°C until used. On the purification day, the resuspended pellets were thawed, sonicated (15 min, 9s on-pulse, 5s off-pulse), centrifuged and loaded onto a HisTrap column. Bound proteins were eluted with a linear gradient of buffer B (20 mM Tris-HCl, 200 mM NaCl, 500 mM imidazole, pH 8.0) and analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). Fractions with higher purity were pooled, concentrated and further purified by SEC on a Superdex 75 16/600 column (Cytiva) pre-equilibrated in 10 mM Tris-HCl, 100 mM NaCl, pH 8.0.

The purity of the final protein samples was analyzed by SDS-PAGE followed by Coomassie Blue staining.

Purification of complexes used for crystallization screenings

The RBD-SD1 construct was incubated with TMPRSS2^{S441A} and A01 at final concentrations of 47.4 μ M, 71.1 μ M and 107 μ M, respectively. After over-night incubation at 4°C, the reaction was loaded onto a Superdex 200 10/300 column (Cytiva) equilibrated in 10 mM Tris-HCl, 100 mM NaCl (pH 8.0) to isolate the complex by SEC. Eluted fractions were analyzed by SDS-PAGE and those corresponding to the ternary complex were pooled, concentrated to 8.5 mg/mL and used in crystallization trials.

TMPRSS2^{S441A} was incubated with A07 at final concentrations of 68 μ M and 102 μ M, respectively, over-night at 4°C. Then, the mix was loaded onto a Superdex 200 10/300 column (Cytiva) equilibrated in 10 mM Tris-HCl, 100 mM NaCl (pH 8.0) and eluted fractions were analyzed by SDS-PAGE. Fractions of the binary complex were pooled, concentrated to 6.1 mg/mL and used in crystallization trials.

Protein expression and purification for biophysical assays

RBD and Spike ectodomain-encoding plasmids were transiently transfected into Expi293F cells (Thermo-Fischer) using FectoPro DNA transfection reagent (PolyPlus). After 5 days at 37°C, cells were harvested by centrifugation and proteins from the supernatants were purified using a HisTrap-Excel column (Cytiva). Eluted fractions were pooled, concentrated and injected onto a Superdex 200 10/300 column (Cytiva) equilibrated in 10 mM Tris-HCl, 100 mM NaCl (pH 8.0) to perform size-exclusion chromatography. Fractions from the main peak were concentrated and frozen. The HKU1B spike protein used for flow cytometry was biotinylated using an EZ-Link Sulfo-NHS-Biotinylation Kit (Thermo Fisher Scientific).

TMPRSS2 (WT) was expressed from a stable S2 cell line as indicated before, and it was purified by affinity chromatography (the tag and glycans were not removed).

The purity of the final protein samples was analyzed by SDS-PAGE followed by Coomassie Blue staining.

Cleavage of TMPRSS2^{S441A} for biophysical assays

To prepare the first batch of cleaved TMPRSS2^{S441A}, 600 μ g were incubated at room temperature with 1 μ g of TMPRSS2 WT in 600 μ L of buffer (10 mM Tris-HCl, 100 mM NaCl, pH 8.0) for 21 h. The final concentration of TMPRSS2^{S441A} in the reaction was 21 μ M. Then, the mix was stored at 4°C for 8 h, aliquoted, flash-frozen in liquid nitrogen, and stored at -80°C until used.

The second batch of cleaved TMPRSS2^{S441A} was prepared by incubating at room temperature 230 μ g with 0.3 μ g of TMPRSS2 WT in 200 μ L of buffer (10 mM Tris-HCl, 100 mM NaCl, pH 8.0) for 7 h. The final concentration of TMPRSS2^{S441A} in the reaction was 25 μ M. Then, the mix was aliquoted, flash-frozen in liquid nitrogen, and stored at -80°C until used.

Crystallization and structural determination

The RBD-SD1/TMPRSS2^{S441A}/A01 complex crystallized in 0.35 M NaH₂PO₄, 0.65 M K₂HPO₄ at 4°C using the sitting-drop vapor diffusion method. The TMPRSS2^{S441A}/A07 complex crystallized in 20 %w/v polyethylene glycol (PEG) 3350, 0.05 M 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) (pH 7.0), 1 %w/v Tryptone, 0.001 %w/v NaN₃ (crystal form 1) and in 10 %w/v PEG 3000, 0.1 M imidazole (pH 8.0), 0.2 M lithium sulfate (crystal form 2) at 18°C using the sitting-drop vapor diffusion method. Crystals were flash-frozen by immersion into a cryo-protectant containing the crystallization solution supplemented with 33% (v/v) glycerol, followed by rapid transfer into liquid nitrogen.

The X-ray diffraction data of both complexes were collected at the SOLEIL synchrotron source (Saint Aubin, France). Collections were carried out at 100 K at the Proxima-1 beamline.⁴⁷

Data were processed, scaled and reduced with XDS^{48,49} and AIMLESS.⁵⁰ The structures were determined by molecular replacement using Phaser from the PHENIX suite⁵¹ with search ensembles obtained from AlphaFold2 (HKU1B-RBD-SD1 and A07) or from previously deposited structures (7MEQ for TMPRSS2^{S441A}, 7KN5 for A01). The final models were built by combining real space model building in Coot⁵² with reciprocal space refinement with phenix.refine.

The RBD/TMPRSS2^{S441A}/A01 complex crystallized in hexagonal P6₁ space group and diffracted to 3.55 Å resolution with two ternary complexes in the asymmetric unit. Refinement was carried out using constraints provided by non-crystallographic symmetry (NCS) and model targets (5KWB for the RBD, our high-resolution structure of TMPRSS2^{S441A} in complex with A07 for the TMPRSS2^{S441A} protein, and 7KN5 and 7VOA for the nanobody A01) allowing us to build a model with good geometry (Table 1).

The TMPRSS2^{S441A}/A07 complex crystallized in two forms of the C2 orthorhombic space group. One form diffracted to 1.8 Å with one complex in the asymmetric unit, and the other diffracted to 2.4 Å with 2 complexes in the asymmetric unit (Table 1). From the high-resolution crystal we built a model of cleaved TMPRSS2^{S441A}, while the second crystal form allowed us to build the uncleaved form. Nevertheless, many strong positive peaks remained in the difference electron density map of the high-resolution data, particularly near the active site. These densities were easily explained by superposing the uncleaved TMPRSS2^{S441A} structure, clearly showing the presence of an alternative conformation on two loops in this crystal (residues 431–440 and 462–467, linked by the disulfide bond C437-C465), which refined with final occupancies of 0.53 and 0.47 (cleaved and uncleaved forms, respectively). In addition, the cleavage released a new N-terminus (residues 256–258) that became ordered and inserted deeply into the core of the protease to interact with the side chain of the buried D440, as it is observed in this family of proteases upon activation.

The final models were validated with Molprobit.⁵³ The analyses of the macromolecular surfaces were carried out in PDBePISA.⁵⁴ Figures were created using Pymol⁵⁵ and BioRender.com.

Biolayer interferometry (BLI)

Affinity of recombinant RBDs toward the purified ectodomain of TMPRSS2^{S441A} was assessed in real-time using a bio-layer interferometry Octet-R8 device (Sartorius). Nickel-nitriloacetic acid (Ni-NTA) capture sensors (Sartorius) were loaded for 10 min at 1,000 rpm

shaking speed with the different RBDs at 400 nM in phosphate-buffered saline (PBS). The sensors were then blocked with PBS containing bovine serum albumin (BSA) at 1.0 mg/mL (assay buffer) and were incubated at 1,000 rpm with 2-fold serially diluted concentrations (800 nM–25 nM) of TMPRSS2^{S441A} in assay buffer. Association and dissociation were monitored for 300 s and 240 s, respectively. Measurements for a reference were recorded using a sensor loaded with an unrelated protein (CD147) that was dipped at each analyte concentration. A sample reference measurement was recorded from a sensor loaded with each RBD and dipped in the assay buffer. Specific signals were calculated by double referencing, subtracting nonspecific signals obtained for the sensor and sample references from the signals recorded for the RBD-loaded sensors dipped in TMPRSS2^{S441A} solutions. Three independent experiments were performed but only the curves obtained at 800 nM in the first experiment were chosen for preparing Figure 2.

Affinity of the recombinant RBDs toward cleaved TMPRSS2^{S441A} (first batch) was determined following a similar protocol, although the range of ligand concentrations assayed went from 240 nM to 7.5 nM. Association and dissociation were monitored for 240 s and 180 s, respectively. Specific signals were calculated by subtracting the nonspecific signal of the sample reference from the signals recorded for the RBD-loaded sensors. Association and dissociation profiles were fitted assuming a 1:1 binding model. Three independent experiments were performed (Figure S6) and the dissociation constant (K_d) values from each of them were averaged and used to calculate the standard deviation.

Sequence alignment

Multiple sequence alignments were performed using Clustal Omega.⁵⁶ The sequence logo was created with WebLogo⁵⁷ (<https://weblogo.berkeley.edu/>).

GFP-split fusion assay

Cell–cell fusion assays were performed as previously described.⁴² Briefly, 293T cells stably expressing GFP1-10 and GFP11 were co-cultured at a 1:1 ratio (6×10^4 cells/well) and transfected in suspension with Lipofectamine 2000 (Thermo) in a 96-well plate (uClear, #655090) (20 ng of spike plasmid, 20 ng of TMPRSS2 plasmids adjusted to 100 ng DNA with pQCXIP-Empty). At 20 h post-transfection, images covering 90% of the well surface, were acquired per well on an Opera Phenix High-Content Screening System (PerkinElmer). The GFP area was quantified on Harmony High-Content Imaging and Analysis Software.

Pseudovirus generation and infection

Pseudoviruses were produced by transfection of 293T cells as previously described.¹ Briefly, cells were cotransfected with plasmids encoding for lentiviral proteins, a luciferase reporter, and the HKU1 spike plasmid. Pseudotyped virions were harvested 2 and 3 days after transfection. Production efficacy was assessed by measuring infectivity or HIV Gag p24 concentration using the commercial ELISA Lenti-X p24 Rapid titer kit (Takara). For infection assays, 293T cells (6×10^4) were transfected in suspension with Lipofectamine 2000 (Thermo) in a 96-well white plate (20 ng of TMPRSS2 plasmids adjusted to 100 ng DNA with pQCXIP-Empty). 24 h post-transfection, cells were passed in 2 wells, and infected with the indicated amount of virus (5–10 ng of p24) in 100 μ L. The next day, 100 μ L of media was added. 48 h post-infection, 125 μ L of media was carefully removed, and 75 μ L of Bright-Glo lysis buffer (Promega) was added. After 10 min, luminescence was acquired using the EnSpire (PerkinElmer).

Flow cytometry

For spike binding, 293T cells were transiently transfected with TMPRSS2 and incubated with Camostat (10 μ M) for 2 h. The cells were incubated with soluble biotinylated spike diluted in magnetic-activated cell sorting (MACS) buffer (PBS, 5 g/L BSA, 2 mM ethylenediaminetetraacetic acid (EDTA) at 2 μ g/mL) for 30 min at 4°C. The cells were then washed twice with PBS and then incubated with Alexa Fluor 647-conjugated streptavidin (Thermo Fisher Scientific, S21374, 1:400 in MACS buffer) for 30 min at 4°C.

For the spike, transfection efficiency was measured at the surface of live cells using mAb10 diluted in MACS buffer for 30 min at 4°C, and Alexa Fluor 647 anti-human IgG (Thermo Scientific, 1:500 in MACS buffer). mAb10 is an antibody generated from a SARS-CoV-2 infected patient which cross-reacts with HKU1.⁵⁸

Surface expression of TMPRSS2 was assessed on live cells by staining with anti-TMPRSS2 A01-Fc¹ at 1 μ g/mL, for 30 min at 4°C in MACS buffer, followed by staining with Alexa Fluor 647-conjugated Goat anti-Human antibody (Thermo Fisher Scientific, A-21445, 1:500 in MACS buffer).

All cells were washed twice with PBS and fixed with 4% paraformaldehyde. The results were acquired using an Attune NXT Flow Cytometer (Life Technologies, software v3.2.1). Gating strategies are described in Figure S7.

TMPRSS2-myc expression was assessed on fixed cells by staining intracellularly with anti-cMyc 9E10 (Thermo - M4439, 1:400), for 30 min at RT in PBS/BSA 1%/NaN₃ 0.05% with saponin 0.05% followed by Alexa Fluor 647 Goat anti-mouse antibody (Thermo Fisher Scientific, A-21242, 1:500 in PBS/BSA 1%/NaN₃ with saponin 0.05%).

In vitro TMPRSS2 enzymatic activity

For enzymatic assays black 96 well plates were used. Soluble TMPRSS2 WT (5 nM final concentration) was mixed with the indicated RBDs at different concentrations in a buffer containing 50 mM Tris-HCl pH 8.0, 150 mM NaCl and 0.01% Tween 20. The plate was incubated 15 min at 37°C. The substrate Boc-QAR-AMC (t-Butyloxycarbonyl-Glutamine-Alanine-Arginine-7-Amino-4-methylcoumarin) was added (100 μ M final concentration, 100 μ L final volume). Fluorescence was read every 3 min for 3 h (excitation

wavelength: 380 nm; emission wavelength: 460 nm) at 37°C, using the EnSpire (PerkinElmer). The initial kinetic slope was taken over the 12 first minutes.

QUANTIFICATION AND STATISTICAL ANALYSIS

All statistical analysis were performed using GraphPad Prism. Details of tests and number of replicates can be found in the figure legends. We defined statistical significance as $p < 0.05$.

Supplemental figures

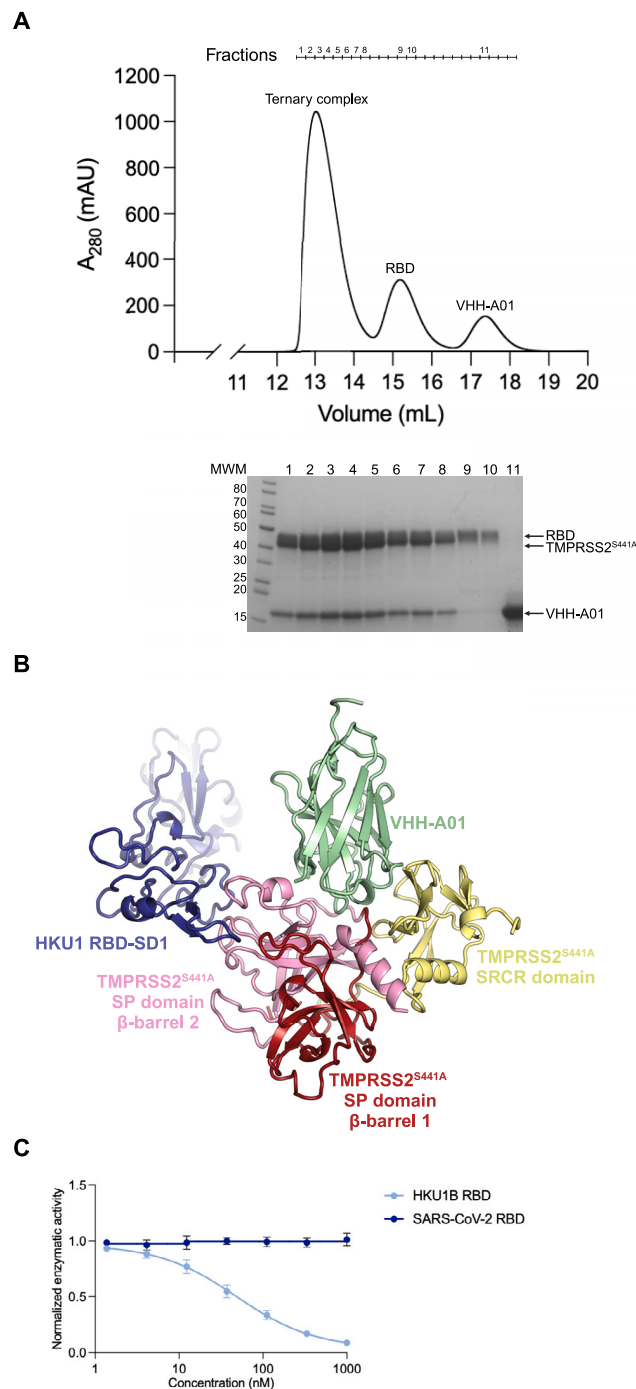


Figure S1. Purification of the ternary complex TMPRSS2^{S441A}/RBD-SD1/VHH-A01, related to Figure 1

(A) Size-exclusion chromatography of the binding reaction prepared with HKU1B RBD-SD1, TMPRSS2^{S441A} and VHH-A01. The eluate was collected in different fractions (numbers indicated above the chromatogram) and 10 μ L aliquots of some of them were analyzed under reducing SDS-PAGE (bottom panel). MWM: molecular weight marker.

(B) Crystal structure of the ternary complex.

(C) *In vitro* TMPRSS2 enzymatic activity. Cleavage of a fluorescent substrate (Boc-QAR-AMC) by soluble TMPRSS2 was evaluated upon incubation with different concentrations of the HKU1B RBD or SARS-CoV-2 RBD (control). The initial kinetic slope of the reaction was measured and normalized to the non-treated condition. Data are mean \pm SD of three independent assays. The plot represents a subset of a larger panel of RBDs presented in Figure S5.

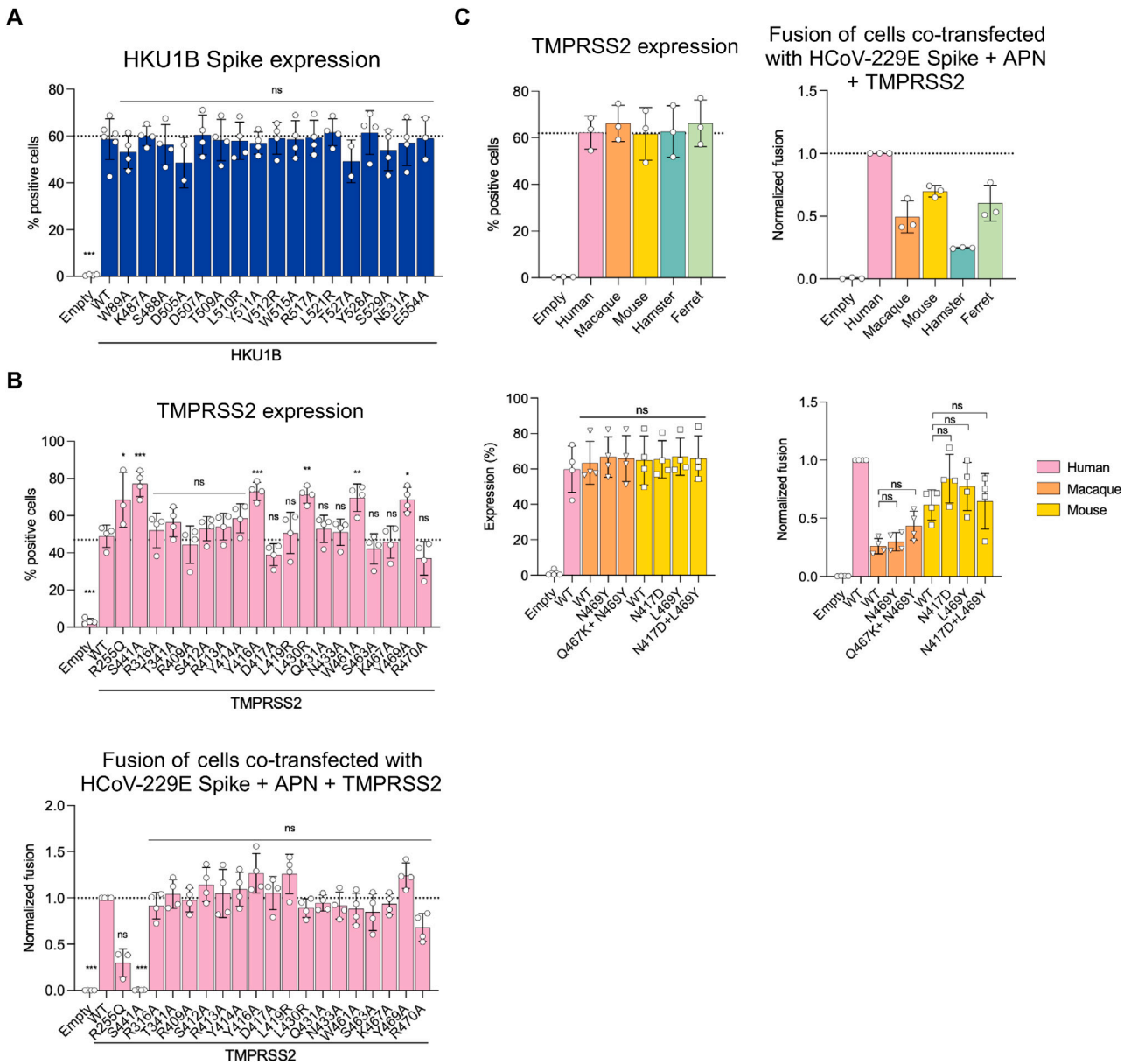


Figure S2. Expression and activity of Spike mutants, TMPRSS2 mutants and TMPRSS2 orthologs, related to Figures 2 and 3
 (A) Expression of HKU1B Spike constructs determined by FACS using an anti-S2 antibody. Data are mean \pm SD of two to four independent assays.
 (B) Surface expression of TMPRSS2 mutants assessed on live cells by FACS with anti-TMPRSS2 VHH-A01-Fc (top). Cell-cell fusion assay performed by co-transfecting cells with HCoV-229E Spike, APN (aminopeptidase N, the receptor) and a TMPRSS2 variant (bottom panel). Data are mean \pm SD of three to four independent experiments.
 (C) Expression of TMPRSS2 orthologs and their point mutants evaluated by intracellular staining of their c-myc epitope (left panel). Cell-cell fusion assay performed by co-transfecting cells with HCoV-229E Spike, APN and one TMPRSS2 ortholog or point mutant (right panel). Data are mean \pm SD of three (top panels) or four (bottom panels) independent experiments.
 Statistical analysis: (A, B, C) One-way ANOVA with Dunnett's multiple comparison test compared to the WT. For fusion assay One-Way ANOVA was performed on the non-normalized log transformed data. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

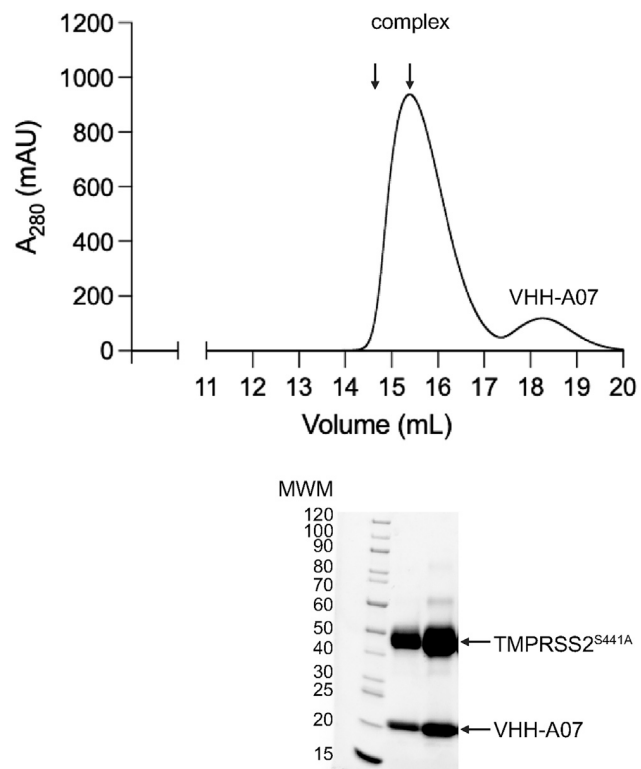


Figure S3. Purification of the TMPRSS2^{S441A}+VHH-A07 complex, related to Figure 4

TMPRSS2^{S441A} was incubated with an excess of VHH-A07 and the binding reaction was injected onto a size-exclusion chromatography column. The elution profile is shown, with two arrows indicating fractions that were analyzed by reducing SDS-PAGE (bottom). MWM: molecular weight marker.

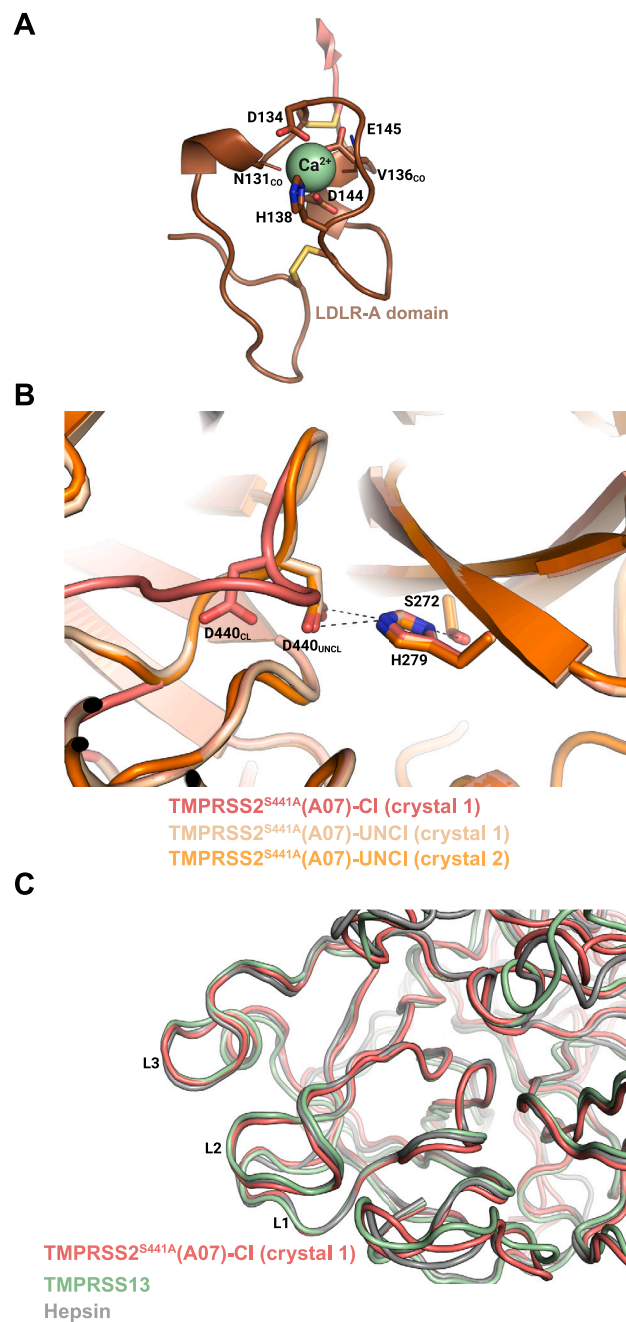


Figure S4. TMRSS2^{S441A} structure, related to Figure 4

(A) Structure of the LDLR-A domain obtained from crystal form 1 of the TMRSS2^{S441A}-VHH-A07 complex. A central calcium ion is shown, as well as the two backbone carbonyls (subscript CO) and side chains that coordinate it. Disulfides C133-C148 and C120-C139 are shown with yellow sticks.

(B) Superposition of the structures obtained for cleaved TMRSS2^{S441A} in the crystal 1 (TMRSS2^{S441A}-CI, red), TMRSS2^{S441A} zymogen in crystal 1 (TMRSS2^{S441A}-UNCI, wheat) and TMRSS2^{S441A} zymogen from crystal 2 (TMRSS2^{S441A}-UNCI, orange) showing the zymogen triad (S272, H279, D440). Upon autocleavage, the conformational change in loop 1 shifts the position of D440 (indicated by the subscript 'CI').

(C) Superposition of cleaved TMRSS2^{S441A} (from crystal 1) with active proteases from its subfamily (TMRSS13, green; hepsin, gray). The loops 1, 2 and 3 (L1, L2 and L3, respectively) are labeled.

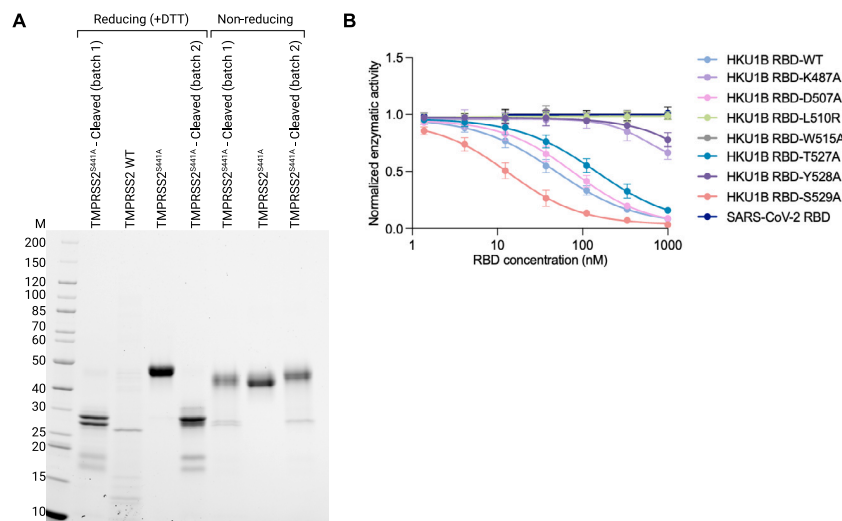


Figure S5. TMPRSS2^{S441A} proteolysis by TMPRSS2 wild-type and effect of mutations on the RBD ability to inhibit TMPRSS2 activity, related to Figure 5

(A) SDS-PAGE under reducing and non-reducing conditions to show the cleavage of TMPRSS2^{S441A} upon incubation with the wild-type protease. Two reactions were performed under different conditions (producing 'batch 1' and 'batch 2') and 2 μ L of each reaction (corresponding to \sim 2.5 μ g of substrate) were analyzed by SDS-PAGE. For comparison, wild-type (WT) TMPRSS2 and untreated TMPRSS2^{S441A} were also loaded in the gel. M: molecular weight marker.

(B) *In vitro* TMPRSS2 enzymatic activity evaluated upon incubation with different concentrations of the HKU1B RBDs harboring mutations at the binding interface. SARS-CoV-2 RBD was used as a control. The initial kinetic slope of cleavage of a fluorescent substrate (Boc-QAR-AMC) was measured and normalized to the non-treated condition. Data are mean \pm SD of three independent assays.

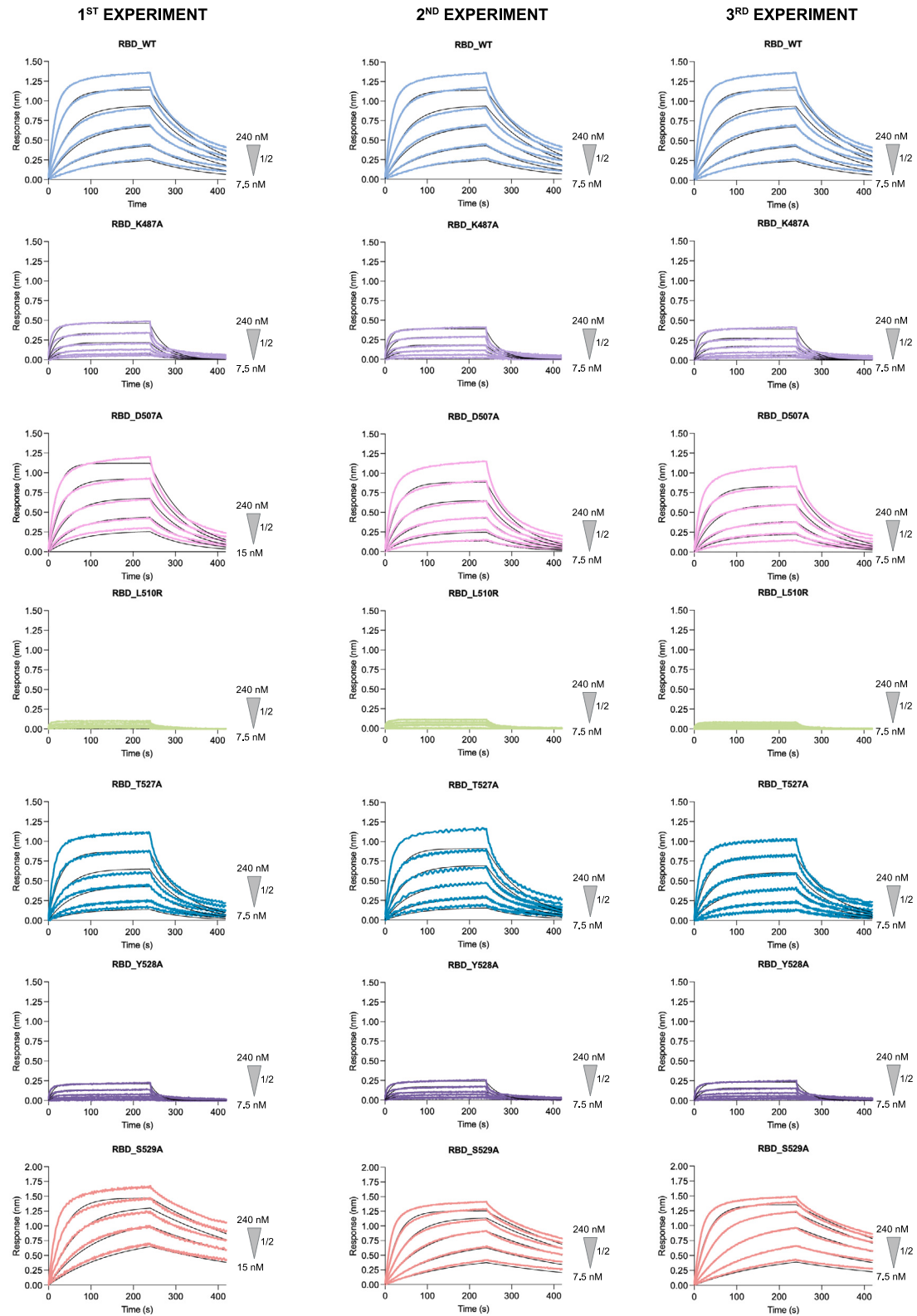


Figure S6. Biolayer interferometry (BLI) experiments to determine the K_d of RBD mutants against cleaved TMPRSS2^{S441A}, related to Figure 5 Three independent assays (each represented in a column) were performed. The initial and final concentrations tested are indicated to the right of each plot, as well as the serial dilutions between them.

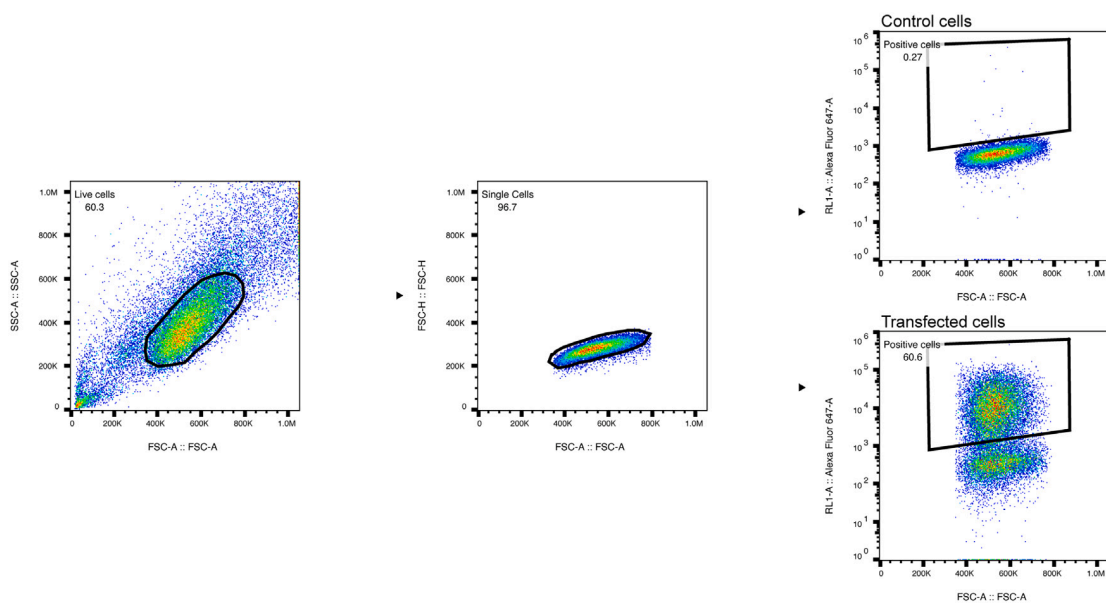


Figure S7. Flow cytometry Gating strategies, related to STAR Methods and to Figure 2
Gating strategy for the detection of surface expression of the transfected proteins by flow cytometry.

3.4. Discussion

3.4.1. Anti-TMPRSS2 nanobodies

A patent was deposited on the anti-TMPRSS2 nanobodies A01 and A07.

As shown in our studies, A01 binds at the interface between the SCRC and SP domains. It does not interfere with HKU1 binding. A07 inserts one of its moieties in the catalytic groove of TMPRSS2 and interacts with some residues of the SP domain that are at the interface of the TMPRSS2/HKU1 complex. This clash explains the inhibition of HKU1 pseudovirus and virus infection mediated by A07. Furthermore, it also explains why A07 inhibits TMPRSS2 enzymatic activity. A07 could thus also be useful to block TMPRSS2 enzymatic activity in the context of other viral infections, such as Influenza or other coronaviruses.

A01 and A07 can also be used as tools to stain TMPRSS2, especially when dimerized with an Fc chain. We tested a panel of commercial anti-TMPRSS2 antibodies; none allowed staining of TMPRSS2 extracellularly. A01-Fc is highly specific to human TMPRSS2 and does not recognize the other human TTSPs tested or TMPRSS2 from other species (**Figure 13**). A07 does not recognize the other human TTSPs tested but recognizes TMPRSS2 from primates and to a lesser extent from mouse and cow (**Figure 13**). They are therefore useful tool to study TMPRSS2 expression and localization.

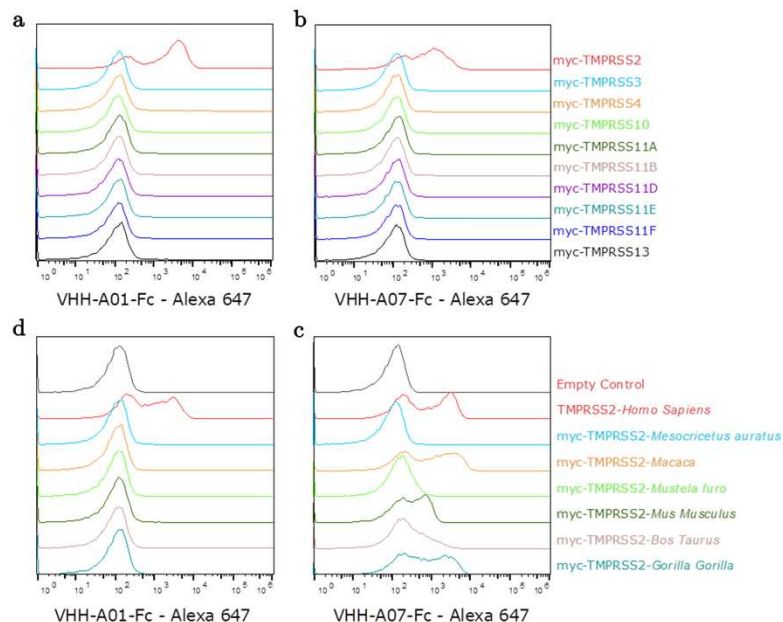


Figure 13: VHH affinity for TTSPs and animal TMPRSS2. Staining of 293T cells transfected with (a,b) human TTSPs, (c,d) animal TMPRSS2, by (a,c) A01-Fc, (b,d) A07-Fc, analyzed by flow cytometry, fluorescence intensity is plotted.

Interestingly, our collaborators only obtained diffracting crystals of the HKU1B RBD in complex with TMPRSS2 in the presence of the A01 nanobody. We had seen that A01 increased the entry of HKU1 pseudovirus, likely by increasing the affinity of HKU1 for TMPRSS2-A01. This increased affinity could favor the crystallization. Crystallization might also be favored in presence of A01 for geometric reasons.

The potential use of TMPRSS2 nanobodies to classify prostate cancers depending on their TMPRSS2 expression, or to target therapy to metastatic cells expressing high levels of TMPRSS2 remains to be investigated.

3.4.2. Other studies performed on HKU1-TMPRSS2 structure

Other studies came out in parallel to ours on HKU1/TMPRSS2 structure.

McCallum *et al.*³⁹⁵ resolved the structure of HKU1A RBD with TMPRSS2 by cryo-EM. As our collaborators, they struggled to isolate sufficient quantities of WT TMPRSS2 ectodomain. They thus optimized TMPRSS2 isolation by introducing several mutations in TMPRSS2 and managed to get higher yields of mature, active TMPRSS2. They performed studies to validate the importance of the residues at the interface for the RBD and TMPRSS2. They also validate that HKU1 binding interferes with TMPRSS2 catalytic activity. Finally, they study the ability of a range of other species' TMPRSS2 to enable HKU1 entry, which we will discuss subsequently.

Xia *et al.*²⁰⁶ resolved the structure of HKU1B full spike with uncleaved TMPRSS2 by cryo-EM. They used the R255Q TMPRSS2 mutant. Interestingly they do not use 9-O-acetylated sialic acids to trigger opening of S. They mention that this might be due to difference between HKU1A, HKU1B and HKU1C. However, in another study on HKU1C, which shares 99% homology with HKU1B, opened forms of the spike were not observed. These differences between studies could be due to different mutations introduced in the Spike to stabilize and purify it. They do not validate the importance of residues situated at the interface.

Overall, the three structures obtained superimpose nicely at the HKU1-TMPRSS2 interface (**Figure 14**). One of the notable differences between HKU1A and HKU1B is the Salt bridge between Asp417_{TMPRSS2} is formed with His488_{HKU1A} or Lys487_{HKU1B}. We found another salt-bridge between Arg470_{TMPRSS2} and Asp505_{HKU1A/B}, like McCallum *et al.*, while Xia *et al.* find this salt bridge is formed with Asp507_{HKU1B}. All studies agree an H-bond is formed between Tyr469_{TMPRSS2} and Leu521_{HKU1}. The interactions found between TMPRSS2 and residues 527-531_{HKU1} slightly differ between the studies.

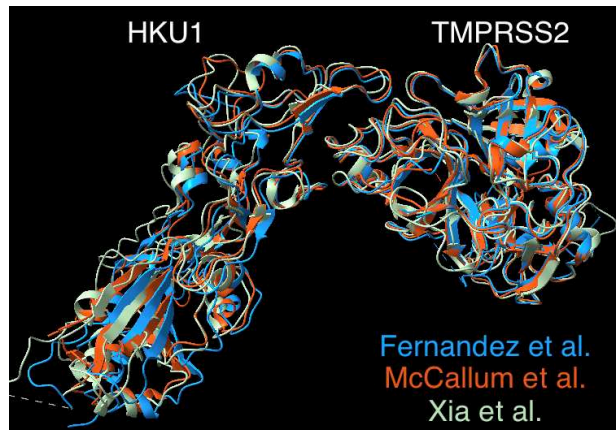


Figure 14: Superimposition of HKU1-TMPRSS2 structures obtained in part 3.3, by McCallum *et al.*³⁹⁵, and Xia *et al.*²⁰⁶.

In the structure obtained by our collaborators the loop 2 in the HKU1-TMPRSS2(Uncleaved)-A01 crystal superimposes with the conformation that is observed in cleaved TMPRSS2 and not in uncleaved TMPRSS2. It is flexible and can move between the two conformations. Upon cleavage this loop is displaced, which enables it to interact with HKU1. This difference in loop 2 conformation explains why HKU1 has a three-fold higher affinity for the cleaved form of TMPRSS2 than for the uncleaved form (**Part 3.2, Figure 5**).

Interestingly, even if the affinity of HKU1 is higher for cleaved TMPRSS2 than for uncleaved TMPRSS2, we saw much lower binding by flow cytometry of soluble HKU1 spike to cells expressing WT TMPRSS2 than to cells expressing uncleaved TMPRSS2 or incubated in the presence of camostat mesylate (**Part 3.2, Figure 3**). This suggests that mature TMPRSS2 either cleaves the streptavidin used to reveal the spike, is less stable than inactive TMPRSS2, is shed or internalized upon binding. Surface stainings reveal higher amount of TMPRSS2 in the presence of camostat or when cells are transfected with inactive TMPRSS2; whether this is due to higher stability of uncleaved TMPRSS2, to a higher affinity of the nanobodies for uncleaved TMPRSS2 or to cleavage of the secondary antibodies by TMPRSS2 is unclear. These results question the use of camostat mesylate to prevent HKU1 infection: camostat mesylate could stabilize TMPRSS2 at the cell surface, and even if it reduces the affinity of HKU1 for TMPRSS2 the overall effect may not be beneficial. While studies indicate that real virus is dependent on the enzymatic activity of TMPRSS2²²¹, it is conceivable low doses of camostat mesylate could stabilize TMPRSS2, while leaving enough unbound TMPRSS2 to cleave the virus.

3.4.3. Struggles with HKU1 authentic virus

To our knowledge, no one has managed to isolate authentic virus on immortalized cell-lines, even though a myriad of cells typically used for respiratory virus isolation were tested, including

RD (human rhabdomyosarcoma), HRT-18 (colorectal adenocarcinoma), HEP-2 (human epithelial carcinoma), MRC-5 (human lung fibroblast), A549 (human lung epithelial adenocarcinoma), Caco2 (human colorectal adenocarcinoma), Huh-7 (human hepatoma), B95a (marmoset B-lymphoblastoid), mixed neuron-glia culture, LLC-MK2 (rhesus monkey kidney), FRhK-4 (rhesus monkey kidney), BSC-1 (African green monkey kidney), Vero E6 (African green monkey kidney), MDCK (Madin-Darby canine kidney), I13.35 (murine macrophage) and L929 (murine fibroblast)¹⁴.

3.4.3.1. Isolation on HAE

We have so far managed, as others^{396,397}, to cultivate both HKU1A and HKU1B on commercial bronchial and nasal HAE (EpithelixTM). These cells are ciliated and express high levels of TMPRSS2 on their cilia³⁹⁸. However, when we compare the levels of vRNA harvested in the same system, they are 100-fold lower than those of SARS-CoV-2. We also managed to amplify the virus on “home-made” HAE, where primary bronchial cells are differentiated at the air-liquid interface for 4-6 weeks^{396,397}, although titers were lower.

We collaborated with Vincent Michel who performed Surface Electron Microscopy on our infected HAE. Interestingly, viral particles are not clearly visible, especially when one compares to similar images obtained with SARS-CoV-2 (**Figure 15**). There might be some virus budding on the cilia, but in low amounts. This is consistent with the fact less virus is harvested.

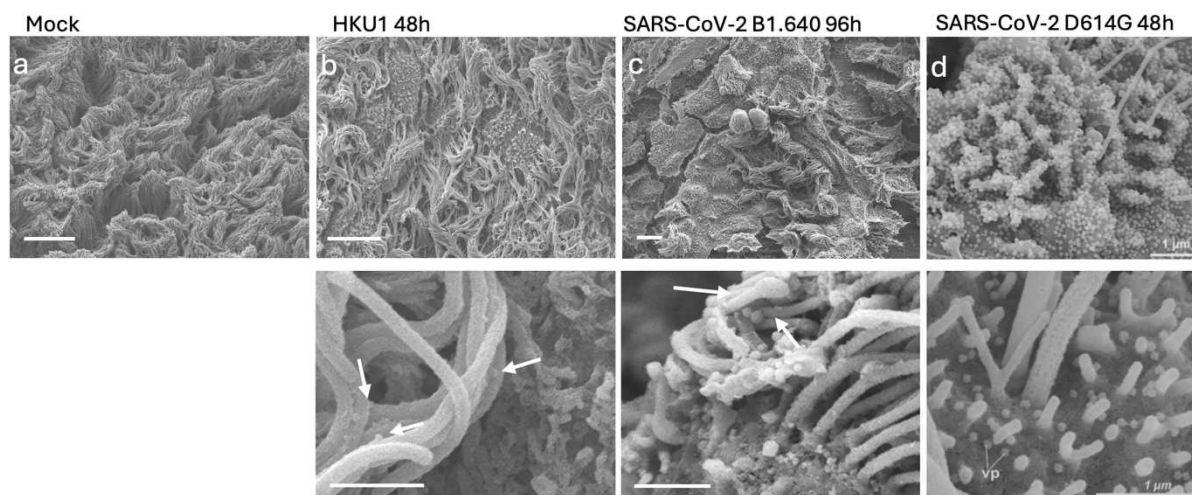


Figure 15: SEM images of mock and infected HAE. (a, b, c) Images obtained by Vincent Michel on Mock (a), HKU1 (b), SARS-CoV-2 B1.640 (c) infected epithelix. Scale bar top 10 μ m, bottom 1 μ m. (d) Published SEM images of HAE infected by SARS-CoV-2 D614G³⁹⁹. Arrows denote asperities which could be virus budding from the cilia.

There is a loss of cilia during infection that is visible by SEM and IF (**Figure 15, Figure 16**). We observe a complete loss of beating that correlates with the peak in viral titers. It is unclear

whether ciliated cells lose their cilia or die. It would be interesting to perform trans-epithelial electrical resistance measurements to measure loss of integrity in the epithelium.

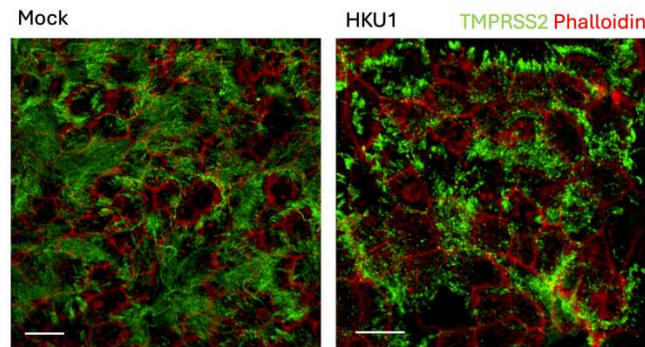


Figure 16: IF staining of mock³⁹⁸ or HKU1 infected (unpublished) HAE (EpithelixTM) 48 hours post-infection. Images obtained by Vincent Michel. Green: TMPRSS2, Red: Phalloidin. Scale bars 10 μ m.

Furthermore, it would be of interest to keep the HAE in culture for longer periods of time after infection to see if the barrier is regenerated or if the infection is controlled. One study performed similar experiments, but on a single viral isolate. They observed HKU1 infection was controlled, and titers decreased over time, which was not the case for OC43 and NL63⁴⁰⁰.

HKU1 infection of primary cells induces a type III IFN response⁴⁰¹. Preliminary experiments we performed indicate that JAK/STAT inhibitors Ruxolitinib and Tofacitinib mildly enhance viral replication in HAE, while IFN β diminishes replication. These results will have to be confirmed. Furthermore, HKU1 is the only known embecovirus that lacks a phosphodiesterase (ns2 in OC43); it will be worth studying the effect of RNase L inhibitors, as RNase L is one of the targets of this phosphodiesterase for OC43. Finally, SARS-CoV-2 replicates more efficiently at 33°C, the temperature of the upper respiratory tract, and the interferon response is delayed at 33°C versus 37°C⁴⁰². We isolated the virus at 33°C, the effect of the temperature on HKU1 replication should be studied.

Finally studies performed on SARS-CoV-2 show that between the TCID50 or PFU determined per mL and the amount of viral genome copies per mL, there is a 10'000 to 100'000 fold^{403,404}. This would mean that our harvests have a titer between 10³ and 10⁵ PFU or TCID50/mL. Although these titers are low and might be optimized as described previously, they should be sufficient to propagate the virus in immortalized cell lines, if they were susceptible to the virus. One could concentrate the virus by ultracentrifugation but given the low volumes (200-400 μ L) harvested at each time point, we would have to infect a large number of HAE.

3.4.3.2. The role of sialic acids

We and others clearly demonstrated the role of sialic acids in HKU1 entry^{236,204,205}. Pretreating the cells with Neuraminidase reduces sialic acid expression and HKU1 pseudovirus infection

(**Part 3.2, Extended Data 5**). Mutating the HKU1 S W89 residue that interacts with sialic acids completely abrogates cell-cell fusion and pseudovirus entry (**Part 3.3, Figure 1**). Most of the cell lines we tested were susceptible to HKU1 induced cell-cell fusion (293T, U2OS, Caco2 are described in **part 3.2, Figure 1**, Huh7 were subsequently tested) and HKU1 pseudovirus entry (293T, U2OS, Caco2, A549 are described in **part 3.2, Figure 2**, Huh7, Hela and MDCK were subsequently tested). One of the notable exceptions was Vero E6 cells, which never fused or enabled pseudovirus entry despite expressing high amount of TMPRSS2. It would be of interest to see if Vero cells lack specific sialic acids that are necessary for HKU1 attachment.

We have performed trials to enhance sialic acid expression in 293T cells to see if that affected pseudovirus entry. We used GD3 synthase (ST8SiaI), that is responsible for synthesizing an α 2,8-linked sialic acid (GD3) that can be 9-O-acetylated and are highly expressed in the human respiratory tract²⁰⁴. Its overexpression has previously been shown to enable HKU1 S binding and HKU1 pseudovirus entry. In our hands, it did not significantly affect entry, suggesting that the sialic acids expressed in our 293T cells are sufficient to enable HKU1 entry.

3.4.3.3. Are there differences between the two TMPRSS2 isoforms?

Two TMPRSS2 isoforms can be expressed in cells because of alternative splicing. Both isoforms are expressed in the lungs³⁹⁴. However, we did not find published single-cell RNAseq data that distinguish the expression of the two isoforms. We have seen no difference in the surface expression of both isoforms. Furthermore, they induce similar levels of HKU1 cell-cell fusion or pseudovirus entry. It is possible that the two isoforms are trafficked, recycled, or shed differently in the extracellular media. This may not affect pseudovirus entry but could affect authentic virus, as there are some differences between pseudovirus and virus which we do not understand yet. For instance, we have seen that for pseudoviruses the enzymatic activity of TMPRSS2 is dispensable in the tested cell lines, while authentic virus relies on TMPRSS2 activity for entry²²¹.

3.4.3.4. Isolation on immortalized cell lines

We have tried to propagate without success HKU1 authentic virus in a range of cell lines. We did not manage to obtain 293T cells that stably expressed TMPRSS2, as they rapidly lost it after transduction. It will thus be of interest to generate 293T that express the different isoforms upon induction, as transient transfection is not appropriate for viral isolations that can last 7-10 days. We used LnCap cells that endogenously express high levels of TMPRSS2, but no replication was observed. We tried using Caco2 cells that also endogenously express TMPRSS2, albeit less than LnCap cells, without success. We tried differentiating Caco2 cells but they were not susceptible to HKU1 infection. We also tried differentiating A549 cells at the air-liquid interface as it has been reported A549 express high levels of TMPRSS2 upon differentiation,

but our differentiations were not successful, this will thus be repeated. Finally, we stably transduced MDCK cells with the two isoforms of TMPRSS2, but no infection was observed.

3.4.3.5. Alternative methods to cultivate the virus

Unfortunately, we do not obtain enough viral RNA from our HAE to directly transfect cells, as it has been described for SARS-CoV-2⁴⁰⁵.

Several approaches to obtain SARS-CoV-2 infectious clones from subgenomic DNA fragments have been described. We launched a collaboration with Prof. Ott and Dr Taha, that recently developed a method to obtain SARS-CoV-2 infectious virions or single round infectious particles³⁷⁸ (**Figure 17**). They use 10 plasmids that encode the whole SARS-CoV-2 genome. To obtain single round virions, the Spike is replaced by a construct containing luciferase and GFP. These plasmids are then be assembled by Golden-Gate assembly. The assembled DNA can then be directly transfected in the cells or *in vitro* transcribed and electroporated. Single round infectious virus is obtained by adding a plasmid encoding the desired spike protein. Fully infectious or single-round infectious viruses are then obtained. They are now working on obtaining infectious clones of HKU1. This method, if it works would enable us to introduce mutations of interest into HKU1 to study their effects. To avoid gain-of-functions studies, these mutations would be introduced in single-round virus.

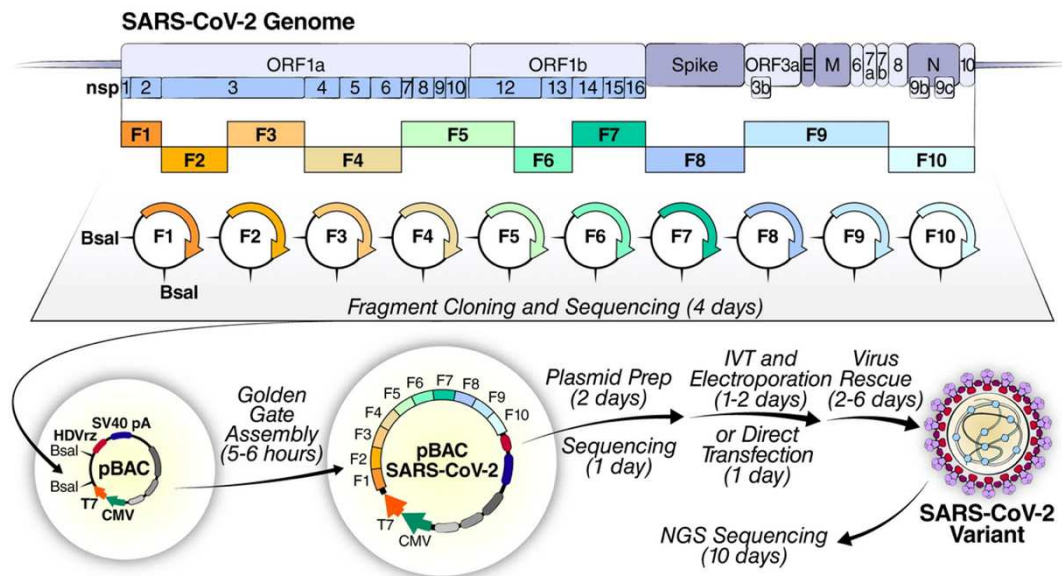


Figure 17: Plasmid-based assembly of SARS-CoV-2. Plasmids coding for different part of the genome are assembled by Golden Gate assembly. The resulting pDNA is either directly transfected or *in vitro* transcribed, before electroporation of the resulting RNA in appropriate cells. Virus is then harvested and sequenced. From Taha *et al.*³⁷⁸

3.4.4. Where does HKU1 come from?

For HKU1, no close relatives were found in other mammals. This is likely explained by lack of sampling. Phylogenetic analysis does not reveal the origin of HKU1. Could receptor usage give us some clues into HKU1's origin?

3.4.4.1. Could other coronaviruses use TMPRSS2 as a receptor?

When aligning embecoviruses RBD to see if there is sequence similarity, one can see the residues that are essential for the HKU1-TMPRSS2 interaction are not conserved in embecoviruses other than HKU1 (**Figure 18**). As there are divergences among the TMPRSS2 from different animal species, it is still possible that another animal coronavirus could bind to its respective TMPRSS2.

We thus used AlphaFold to see if we could predict any interaction of Longquan rat coronavirus, one of the coronavirus with the closest RBD to HKU1, with rat TMPRSS2 (**Figure 19**). AlphaFold predicts a structure that is similar to the one of the HKU1 RBD, with a pincer plier. However, no interaction with Rat TMPRSS2 is predicted, suggesting it does not use TMPRSS2 as a receptor.

For OC43, whose receptor is unknown, TMPRSS2 was not sufficient to induce fusion in 293T cells, no binding of OC43 spike was observed on cells expressing TMPRSS2, all the interface residues are mutated, and AlphaFold predicts a structure closer to a fork, than a plier (**Figure 19**). All evidence points to the fact that TMPRSS2 is not OC43's receptor.



Figure 18: Multiple alignment of the sequences of different Spike of embecoviruses (COBALT). Residues are colored by their frequency of conservation. The blue arrows denote the aa at the interface between HKU1 RBD and TMPRSS2 that are essential for the interaction.

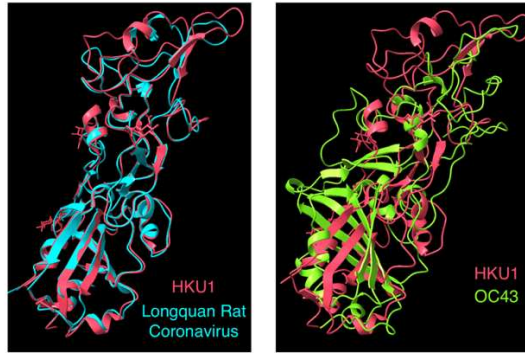


Figure 19: AlphaFold predictions of Longquan Rat coronavirus and OC43 RBD folding, overlapped with HKU1 structure (8vgt.pdb) performed with ChimeraX⁴⁰⁶.

3.4.4.2. What have we learned of the potential origin of HKU1?

We have seen that there are two essential conditions for HKU1 to infect cells: the sialic acids and TMPRSS2.

HKU1 and OC43 are much more specific in their affinity for sialic acids than other animal coronaviruses (**Figure 20**); this is quite notable for OC43 that is much more specific than BCoV, its ancestor⁴⁰⁷. It is thus possible that this high specificity is also a result of prolonged circulation in humans, as the loss of the lectin domain the HE of OC43 and HKU1¹³⁵. Looking at sialic acids present in different species is thus unlikely to provide clues on HKU1 origin.

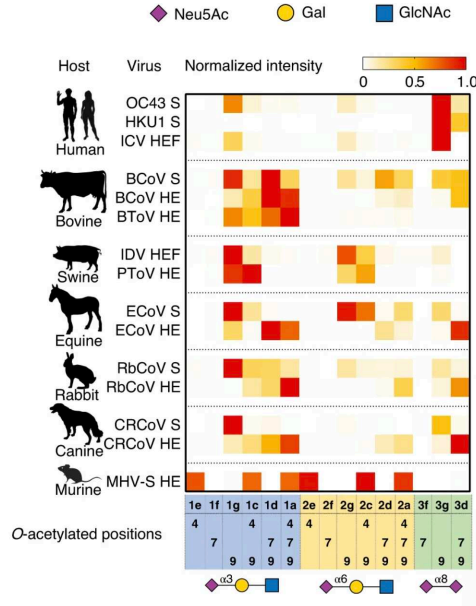


Figure 20: Sialic acid usage by different embecoviruses. Binding of soluble S and HE proteins to the different sialic acids is measured and normalized to the highest affinity obtained for each protein. From Li *et al.* 2021²⁰⁴.

We investigated the use of TMPRSS2 from different species by HKU1. We tested mouse, hamster, ferret and macaque TMPRSS2. Macaque TMPRSS2 enable some entry, although less than the human TMPRSS2. Furthermore, we observed some difference between HKU1A and

HKU1B; in our hands HKU1A was able to partly use mouse TMPRSS2, while it was not the case for HKU1B (**Figure 21**). McCallum *et al.*³⁹⁵ published data on showing HKU1A could use Hamster TMPRSS2 but not mouse TMPRSS2 which is not in agreement with our data (**Figure 21**). They test a larger panel of animal TMPRSS2 than us.

We then showed that the residues 417 and 469 were not conserved among the different species, and that reverting them to the human residues was sufficient to restore entry (**Part 3.3, Figure 3**). When looking at other species, about half of primates and all Bovidae have the same residues as the human in position 417 and 469. It will thus be of interest to test if those species TMPRSS2 can be used by HKU1.

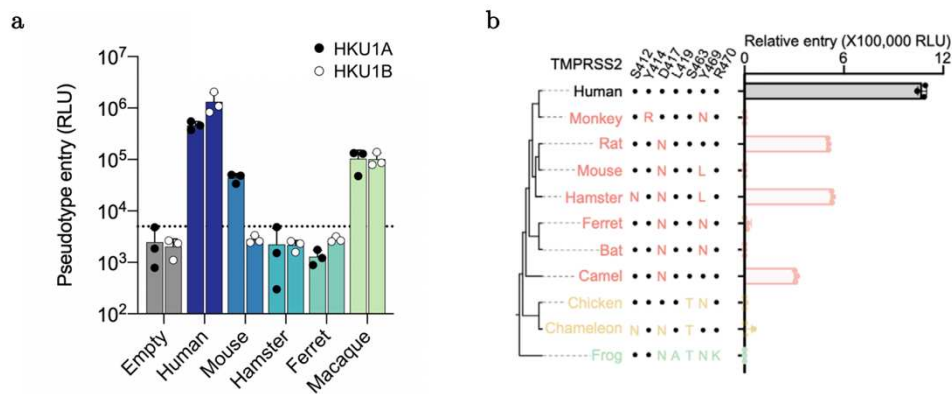
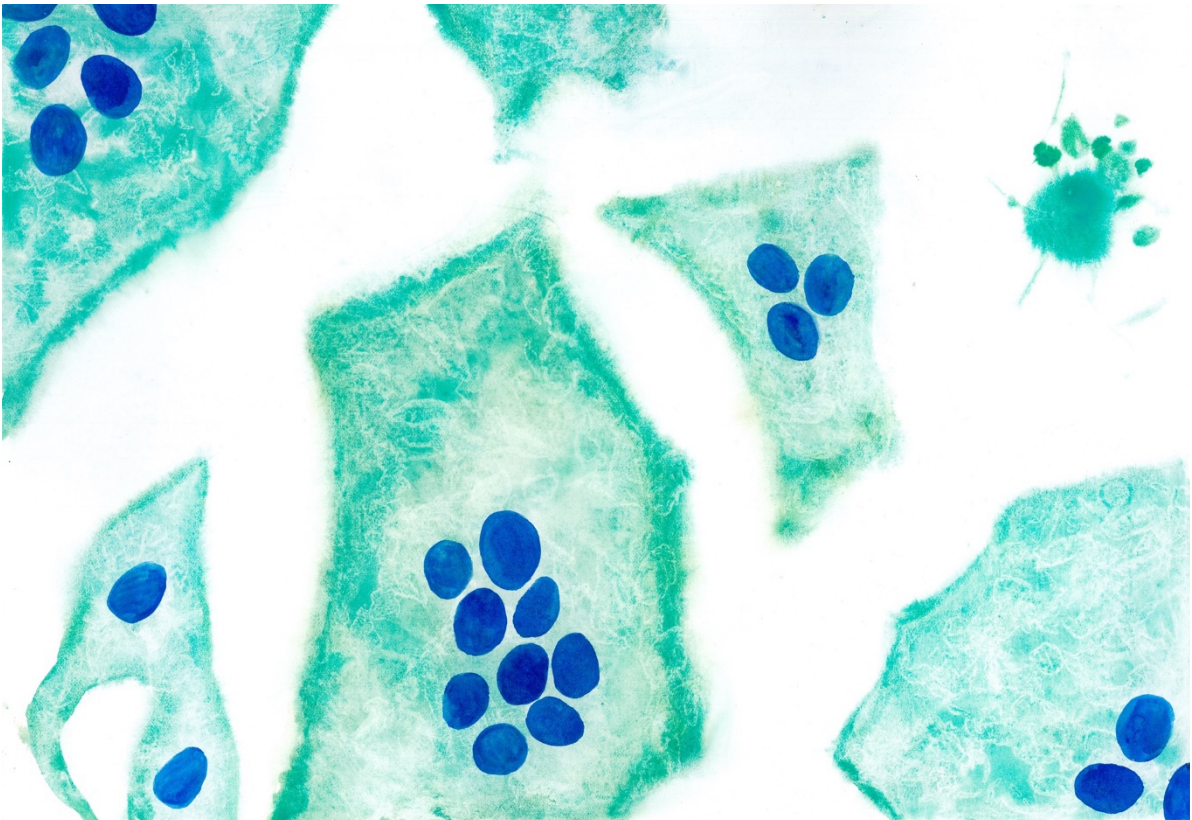


Figure 21: Species specific TMPRSS2 use by HKU1A and HKU1B. a. Data we obtained by transfecting the indicated TMPRSS2 in 293T cells and infecting with HKU1A or HKU1B pseudoviruses. **b.** Data obtained by McCallum *et al.*³⁹⁵ by transfecting the indicated TMPRSS2 in 293T cells and infecting with HKU1A pseudoviruses.

To conclude, although our work answers one of the main questions of the scientific community on HKU1, much work remains to be done on this poorly studied virus. Identifying a susceptible cell line or developing a replicon system to amplify and share HKU1 stocks will likely accelerate research on this virus. Increased sampling of animals to find closely related virus might yield insights on the origin of this virus.

Chapter 4 – Subcellular remodeling during SARS-CoV-2 infection



The article presented in this Chapter is the following:

- Saunders N, Monel B, Cayet N, Archetti L, Moreno H, Jeanne A, Marguier A, Wai T, Schwartz O, Fréchin M. *Dynamic label-free analysis of SARS-CoV-2 infection reveals virus-induced subcellular remodeling*. Nature communications, 2024 Jun (4996).

This project was done in collaboration with Nanolive, a Swiss-based company that commercialized a Holotomographic microscope. It was supervised by Dr Timothy Wai, Prof. Olivier Schwartz and Dr Mathieu Fréchin. I present this work with full acknowledgment of the work performed by all authors. I performed all the experimental work present in this manuscript except the cryo-EM experiments. All the computational analysis was carried out by Dr Mathieu Fréchin and his team.

4.1. Introduction

4.1.1. Classical imaging techniques

Classical microscopy relies on the light transmitted through a sample. As most cellular components are transparent, the contrast in Brightfield imaging is poor. To improve the contrast, one can use phase contrast or differential interference contrast microscopy which rely on the difference of velocity of light depending on the refractive index of the sample (**Figure 22**). One can also use darkfield which relies on the reflected light instead of the transmitted one (**Figure 22**). In cases where biological structures are birefringent, such as spindles, collagen, actin or microtubules, one can use Polarized light microscopy. For cells the most widely used techniques are phase contrast and DIC. However, they are not quantitative, and their analysis is limited due to the low contrast and low signal to noise ratio.

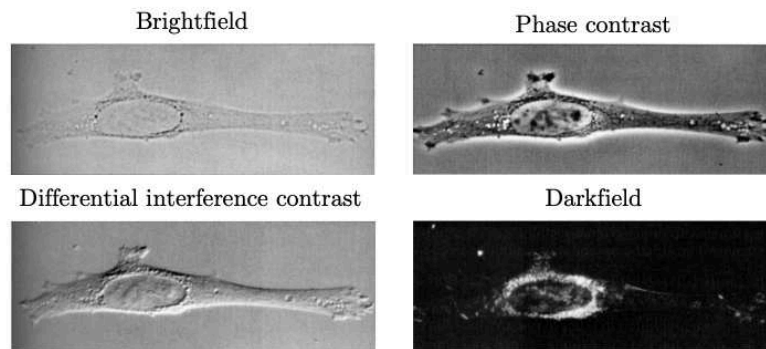


Figure 22: Comparison of classical label-free imaging approaches.

Most optical microscopy techniques rely on fluorescence. Adding fluorescent tags to protein or specific dyes enables quantitative, multicolor imaging of the processes of interest, which has led to the widespread use of fluorescence microscopy. Furthermore, recent advances in super-resolution microscopy, and thick sample imaging increased the potential of fluorescence microscopy. However, fluorophores can interfere with biological processes, and photobleaching or phototoxicity after high intensity, repeated and prolonged excitation of the sample can occur. This limits the use of fluorescence microscopy with live samples. New microscopy techniques and the use of artificial intelligence (AI) in analysis brought label-free microscopy to the front scene, notably for live imaging *in vitro* or *in vivo*.

4.1.2. Technical improvements in imaging

4.1.2.1. Phase Imaging

Phase contrast and DIC can be improved thanks to different techniques.

Quantitative Phase imaging (QPI) and quantitative DIC are improvements of the classical phase-contrast and DIC microscopes⁴⁰⁸. Thanks to the comparison of the beam going through the sample with a reference beam and computation, the pixel intensity observed is proportional to the phase-shift induced by the sample. Although this concept is simple and started being investigated a hundred years ago, its wide use was made possible by digital camera and computers.

Holotomography (HTM) is a 3D QPI. In 2D QPI, the phase-shift is linked to both the thickness and the refractive index of the sample, and it is not straight-forward to decouple them. In 3D QPI, instead of sending a beam of light perpendicular to the sample, a laser goes through the sample from different angles and at several inclinations. The phase-shift in this imaging laser is compared to the reference beam. This allows the computational reconstruction of a 3D tomogram, a Z-stack of QPI images. Two companies commercialized holotomography microscopes: Nanolive and Tomocube. This technique is not widely used yet but allows the study of many biological systems without any labels.

Gradient light interference microscope enables 3D quantitative phase imaging of samples that are hundreds of microns thick, opening new perspectives for biomedical imaging.

4.1.2.2. Raman microscopy

Raman spectroscopy relies on the Raman scattering; when a laser goes through a sample and encounters molecules, a very small proportion of the light (0.0000001%) is scattered at a higher wavelength than the incoming light. This wavelength depends on the nature of the bond (C-H, O-H, C=C...). When coupling this approach with microscopy, one obtains a spectrum from which the chemical composition of the analyzed region can be inferred. This can be useful to monitor biochemical changes in cells, drug uptake, or visualize newly synthesized protein or lipids when adding stably labeled isotopes to the cells^{409,410}. Raman spectroscopy was initially limited by the fact that samples have to be illuminated for prolonged periods of time because Raman scattering is so low. Significant technological progress has been made in the last years to reduce imaging time, cost and make analysis more user-friendly.

4.1.2.3. Other label-free imaging techniques

Rotating coherent scattering microscopy enables the detection of the scattered light by small objects at high frequency, to image phenomenon such as millisecond reorganization of macrophage actin structures, degranulation and pore formation, nanotube dynamics between cells, nanoparticle binding, and bacterial lectin dynamics⁴¹¹.

Coherent brightfield microscopy and interferometric scattering detect scattered light enabling nanoparticle/virus tracking at nanometric resolution^{412,413}.

Photoacoustic microscopy detects pigments such as hemoglobin and can be used in clinical settings⁴¹⁴.

Autofluorescence microscopy can be used to gain information on metabolism thanks to NAD(P)H fluorescence, identify the presence of cellular markers, such as flavins, or analyze extracellular matrix remodeling thanks to the collagen and elastin⁴¹⁵.

Second Harmonic Generation imaging enables the visualization of fibers such as collagen, actomyosin and microtubules⁴¹⁶.

4.1.3. Analysis pipelines

4.1.3.1. For cellular characterization

Statistical methods and machine learning can be used in QPI or holotomography to classify micro-organisms^{417,418} or cell-types^{419,420}, characterize cancer pathology^{421–423}, identify cell-death pathway⁴²⁴, grade spermatozoa quality⁴²⁵, identify red-blood cells diseases^{426,427}. An approach coupling QPI and polarized light imaging enables identification of brain regions in tissue⁴²⁸.

More than 20 studies develop approaches to enable a virtual transformation of QPI or brightfield images into bright-field histologically stained samples^{429–431}, and will likely diminish the requirements of histopathological staining for biomedical analysis.

4.1.3.2. For organelle segmentation

Several pipelines based on AI have emerged in the last few years to provide organelle segmentation in classical transmitted-light microscopy. These techniques rely on training deep-learning models with a set of pairs of images: a classical brightfield or phase-contrast image and a fluorescent image of the same sample where the organelle of interest is labeled.

One approach enabled prediction from brightfield images of nuclei, nuclear membrane, microtubules, actin filaments, mitochondria, ER, DNA and cell membrane with a Pearson correlation coefficient ranging between 60 and 90%. Another study enabled tracking of nuclei and cytoplasm dry mass with quantitative phase-imaging⁴³². Several studies focus on segmentation of mitochondria from phase contrast or bright-field images, and have Pearson correlation coefficients or similarity metrics above 80%^{433,434}. Using QPI also enables segmentation of the lysosomes⁴³⁵. Finally such techniques can also be used to predict the presence of protein aggregates in cells⁴³⁶.

Nanolive had previously published a method based purely on refractive index values, without machine learning, to segment lipid droplets (LD) and mitochondria from HTM images from the refractive index, with Pearson correlation coefficients around 60%⁴³⁷.

A study used a deep learning approach to actin, nucleus, nucleoli, cell membranes, mitochondria and LD for 3D tomograms⁴³⁸. However, they only use one metric to quantify their accuracy, SSIM, which does not analyze pixel but overall similarity, and their resolution, especially for mitochondria is quite low.


Here, we present a novel approach to segment organelles in HTM over time. We validate our predictions with fluorescence. We show the potential use of this method to quantify organelle remodeling over time. This method can be coupled with other imaging modalities such as live wide-field fluorescence imaging and confocal microscopy or electron microscopy after fixation.



4.2. Dynamic label-free analysis of SARS-CoV-2 infection reveals virus-induced subcellular remodeling

Received: 9 December 2023

Accepted: 30 May 2024

Published online: 11 June 2024

 Check for updates

Nell Saunders¹, Blandine Monel¹, Nadège Cayet², Lorenzo Archetti^{3,6}, Hugo Moreno^{3,6}, Alexandre Jeanne^{3,6}, Agathe Marguier³, Julian Buchrieser¹, Timothy Wai⁴, Olivier Schwartz^{1,5} & Mathieu Fréchin³  

Assessing the impact of SARS-CoV-2 on organelle dynamics allows a better understanding of the mechanisms of viral replication. We combine label-free holotomographic microscopy with Artificial Intelligence to visualize and quantify the subcellular changes triggered by SARS-CoV-2 infection. We study the dynamics of shape, position and dry mass of nucleoli, nuclei, lipid droplets and mitochondria within hundreds of single cells from early infection to syncytia formation and death. SARS-CoV-2 infection enlarges nucleoli, perturbs lipid droplets, changes mitochondrial shape and dry mass, and separates lipid droplets from mitochondria. We then used Bayesian network modeling on organelle dry mass states to define organelle cross-regulation networks and report modifications of organelle cross-regulation that are triggered by infection and syncytia formation. Our work highlights the subcellular remodeling induced by SARS-CoV-2 infection and provides an Artificial Intelligence-enhanced, label-free methodology to study in real-time the dynamics of cell populations and their content.

The COVID-19 pandemic is caused by the severe acute respiratory syndrome-coronavirus-2 (SARS-CoV-2)¹, inducing a broad spectrum of syndromes from a light cold to life-threatening pneumonia². The search for SARS-CoV-2 treatments is continuing³ and reductionist approaches vastly dominate experimental efforts. A stop-motion view of the SARS-CoV-2 infection cycle has emerged⁴, where its impact on the host cell is understood through key host/virus molecular entanglements⁵. Previous studies tackled the impact of the virus on a global cellular scale, employing fluorescence and electron microscopy^{6,7}, and as such were lacking the dimension of time. Filming the impact of SARS-CoV-2 on an entire cellular system from early infection to death would greatly improve our understanding of infection sequences and dynamics, yet the efforts to obtain such knowledge

are precluded by the limitations of live microscopy. The various types of fluorescence microscopy induce non-neglectable phototoxicity and molecular perturbations, due to the use of chemical or genetic labeling⁸⁻¹³. This limits the capacity to observe multiple targets over hours-long periods, which is the time scale necessary to capture the cellular changes induced by SARS-CoV-2. Classical label-free imaging techniques such as phase contrast or differential interference contrast (DIC), while less invasive, provide images plagued by optical aberrations, poor contrast, and limited spatial resolution. AI-augmented label-free microscopic methods have emerged¹⁴, emulating fluorescence staining for key cellular structures in the absence of fluorescence^{15,16} or bolstering the usage of lower-content, label-free images to detect specific cellular states¹⁷⁻¹⁹. Holotomographic

¹Virus & Immunity Unit, Institut Pasteur, Université Paris Cité, CNRS, UMR 3569 Paris, France. ²Institut Pasteur, Université Paris Cité, Ultrastructural Bioimaging Unit, 75015 Paris, France. ³Deep Quantitative Biology Department, Nanolive SA, Tolochenaz, Switzerland. ⁴Mitochondrial Biology Group, Institut Pasteur, Université Paris Cité, CNRS, UMR 3691 Paris, France. ⁵Vaccine Research Institute, Creteil, France. ⁶These authors contributed equally: Lorenzo Archetti, Hugo Moreno, Alexandre Jeanne. ✉ e-mail: Olivier.schwartz@pasteur.fr; Mathieu.frechin@nanolive.ch

microscopy (HTM) provides high-content refractive index (RI) images able to capture complex biological processes and multiple cellular structures at unprecedented spatial resolution and ultralow-power illumination²⁰. When combined with computer vision, HTM can support image-based quantitative investigations of cell dynamics over hours at relevant temporal resolutions²⁰.

In this study, we developed a high-content imaging pipeline combining live HTM, machine learning, and Bayesian network modeling to provide a quantitative and dynamic vision of the impact of SARS-CoV-2 on the organelle system of hundreds of infected cells in culture.

Results

Label-free microscopy shows virus-induced cellular alterations

Through key host–viral protein interactions, SARS-CoV-2 reshapes the subcellular organization and the organelles of its target cells^{6,7}. SARS-CoV-2 reroutes lipid metabolism^{21–23}, fragments the Golgi apparatus⁷, promotes the formation of double-membrane vesicles^{7,24,25} and alters mitochondrial function²⁶, with the goal of boosting virus production while delaying antiviral responses^{27,28}. Our aim was to capture the kinetics and extent of such alterations in living cells by recording and quantifying cellular and organellar dynamics in real time using HTM²⁰. We selected U2OS-ACE2 cells as targets because of their high sensitivity to SARS-CoV-2 and their flat shape that facilitates imaging²⁹. Cells were first infected with the Wuhan strain and imaged with HTM (Fig. 1A and Supplementary Fig. 1a, b). Time-lapse experiments were carried out for up to 2 days or until the death of infected cells (Fig. 1B). Non-infected cells were recorded as a control. The most obvious event visible as soon as 10 h post-infection (pi) was the formation of syncytia, a known phenomenon where infected cells expressing the viral spike (S) protein at their surface fuse with neighboring cells^{29,30}.

Formation of syncytia was used as a marker of productively infected cells. In such cells, we noticed a quick clustering of nuclei, visible as soon as two or more cells started to fuse. The zone of nuclei clustering apparently hosted groups of growing lipid droplets (LD), accumulating over time, while mitochondria were moving away from this region and redistributed across the cytoplasm. Within the nuclei, nucleoli appeared denser and rounder upon infection (Fig. 1C). We next determined which of these cellular events were due to the infection itself or the result of syncytia formation. To differentiate between these possibilities, we recorded cells that fused together in the absence of SARS-CoV-2, after transient expression of Syncytin-1, a fusogenic protein involved in the formation of placental syncytiotrophoblasts³¹. Infection-independent syncytia did not show the same features. LD remained small and rare, and nucleoli were not altered (Fig. 1B and C).

However, mitochondrial movements in both SARS-CoV-2- and Syncytin-1-induced syncytia seemed globally similar. As demonstrated before²⁰, we did not detect the Golgi network, the cytoskeleton, or DMVs with HTM since these structures show little RI contrast with their surroundings.

To go beyond the qualitative nature of these observations and to quantify the cellular alterations triggered by SARS-CoV-2, we designed an HTM image quantification pipeline in which cells and organelles were detected in time-lapse recordings by tailored machine learning (ML) approaches (Fig. 1D).

Machine learning detects cellular organelles in high-resolution label-free images

We adapted our ML strategy to the different characteristics of the biological objects of interest. Mitochondria that are small, pixel-scale objects were detected using a two-class pixel categorization³², where a trained extra-tree classifier attributes a class to each pixel based on its position in a derived feature space. Such an approach allowed precise and accurate detection of these sparse objects within the highly

textured HTM images (Fig. 2A). Its large hyperparameter space was not explored through a *human-in-the-loop* process but using the optuna optimization framework³³. Larger objects such as nuclei, nucleoli, and whole cells were optimally segmented with an adapted U-NET³⁴ fully convolutional network (Fig. 2B, Supplementary Fig. 2). For whole cell segmentation, a sharpening of the outlines was performed by object propagation within the RI signal³⁵. LD was segmented using Nanolive's dedicated assay, which automatically detects LD based on their high refractive index, unique signal distribution, and roundness.

We then validated the automatic segmentations of organelles within RI images by labeling the cells with organelle-specific fluorescent markers. Nuclei, nucleoli, LD, and mitochondria were stained respectively with Hoechst, Green Nucleolar staining (ab139475), lipid spot, and Mitotracker DeepRed (Fig. 2C–G, Supplementary Fig. 3a–d). We used the standard F1 and intersection over union (IoU) scores for the strict binary evaluation of masks versus references and the structural similarity index measure (SSIM)³⁶ for a quantification of similarity perception. For nuclei and nucleoli that are large and simple objects, matches between masks and fluorescent signals were high (Figs. 2C, G and 3B, Supplementary Fig. 3a). The LD masks did not perfectly match with the lipid spot fluorescent signal. This was expected since HTM resolving power is better than epifluorescence²⁰ (Fig. 2E, Supplementary Fig. 3c). This illustrates the challenge of objectively quantifying the quality of few-pixel object masks, especially in a live context where biological structures move through the succession of acquisition regimes. For these reasons, the scores of LD predictions were very good yet lower than those of nuclei and nucleoli (Fig. 2G).

Similarly, our RI-based mitochondrial predictions were sharper and better resolved than the fluorescent signal generated by Mitotracker DeepRed (Fig. 2F, Supplementary Fig. 3d). The scores obtained from comparing our RI-based ML predictions against fluorescence-derived references were good (Fig. 2G) yet lower than those of the other organelles because of unavoidable mismatches between predictions and ground truth. In addition to organelle motion, the typical crowding of mitochondria in the perinuclear region generates unresolved³⁷ Mitotracker epifluorescence signal. Such a signal is not optimal for comparison purposes. We thus used an expert-generated segmentation of mitochondria within an RI image to assess further the quality of our ML-generated mitochondrial mask (Fig. 2G).

To validate the biological relevance of our mitochondrial detection workflow, we silenced *OPA1*, a dynamin-like GTPase protein required for mitochondrial fusion³⁸, whose ablation causes mitochondrial fragmentation and inherited optic neuropathy³⁹. Inspection of the label-free HTM images revealed an obvious fragmentation of the mitochondrial network and alteration of mitochondria shapes, and our ML-based mitochondrial detection system reported a reduction of mitochondria size distribution (Supplementary Figs. 4 and 5). This confirmed our capacity to automatically detect and quantify a broad range of mitochondrial morphology under basal and pathological conditions in a label-free manner.

SARS-CoV-2 infected cells display specific organelle dynamics

We then examined in real-time the subcellular changes induced by infection with two different SARS-CoV-2 strains. We recorded movies of U2OS cells infected with either the SARS-CoV-2 Wuhan ancestral virus or the Omicron BA.1 variant. As controls, we used uninfected cells and cells that underwent intercellular fusion upon Syncytin-1 expression. We used our algorithms to identify over time single or fused cells with a precise segmentation even in time of transition from the single cell to syncytium state (Supplementary Fig. 6 and Supplementary Movies 1–19) and to segment nuclei, nucleoli, LD, mitochondria, and the cytosol. The cytosol is defined by the removal of all organelle segmentations from the single cells or syncytium segmentations. (Supplementary Fig. 6 and Supplementary Movies 1–19). This analysis was performed every 12 min, from 2 to 48 h pi.

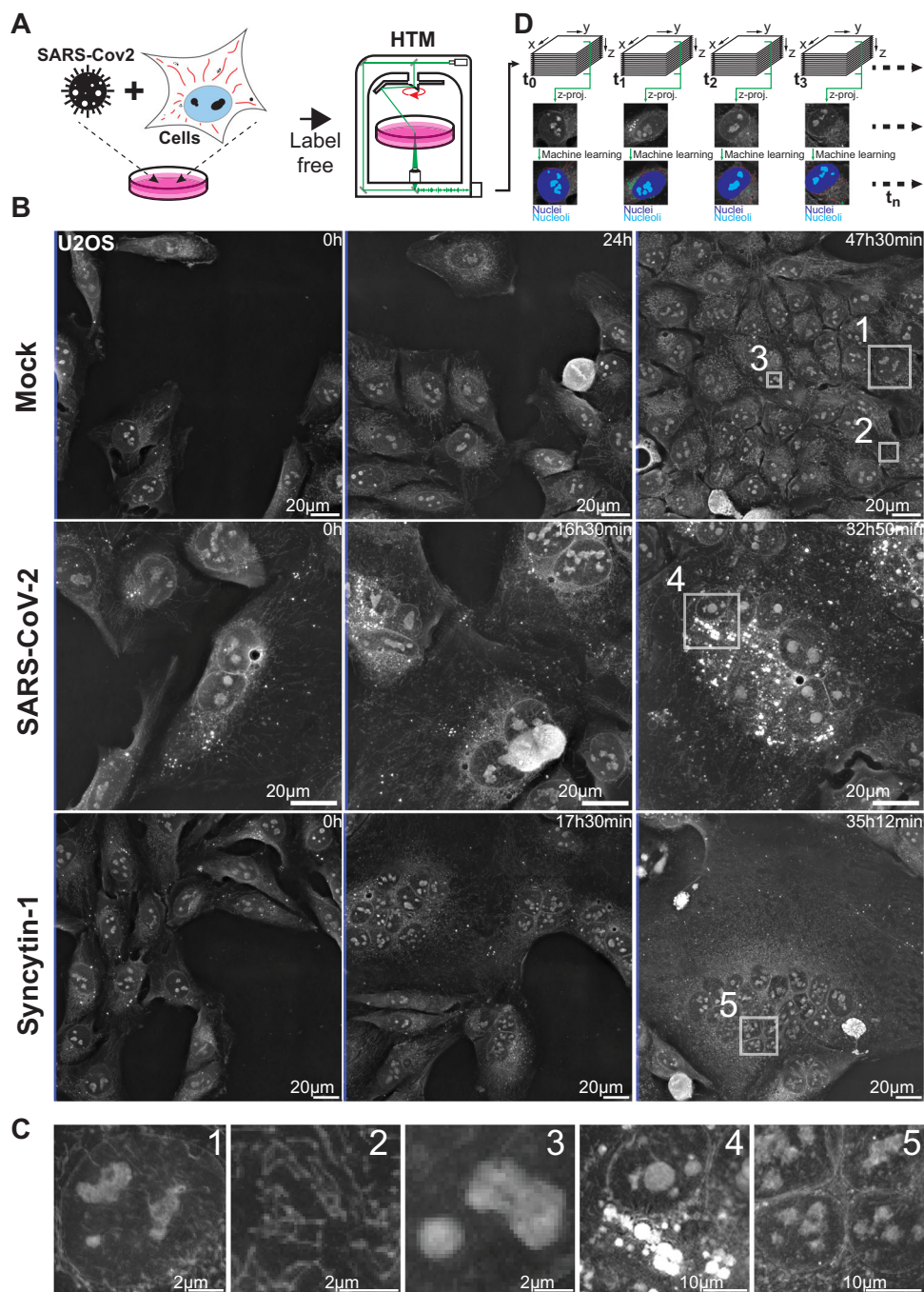


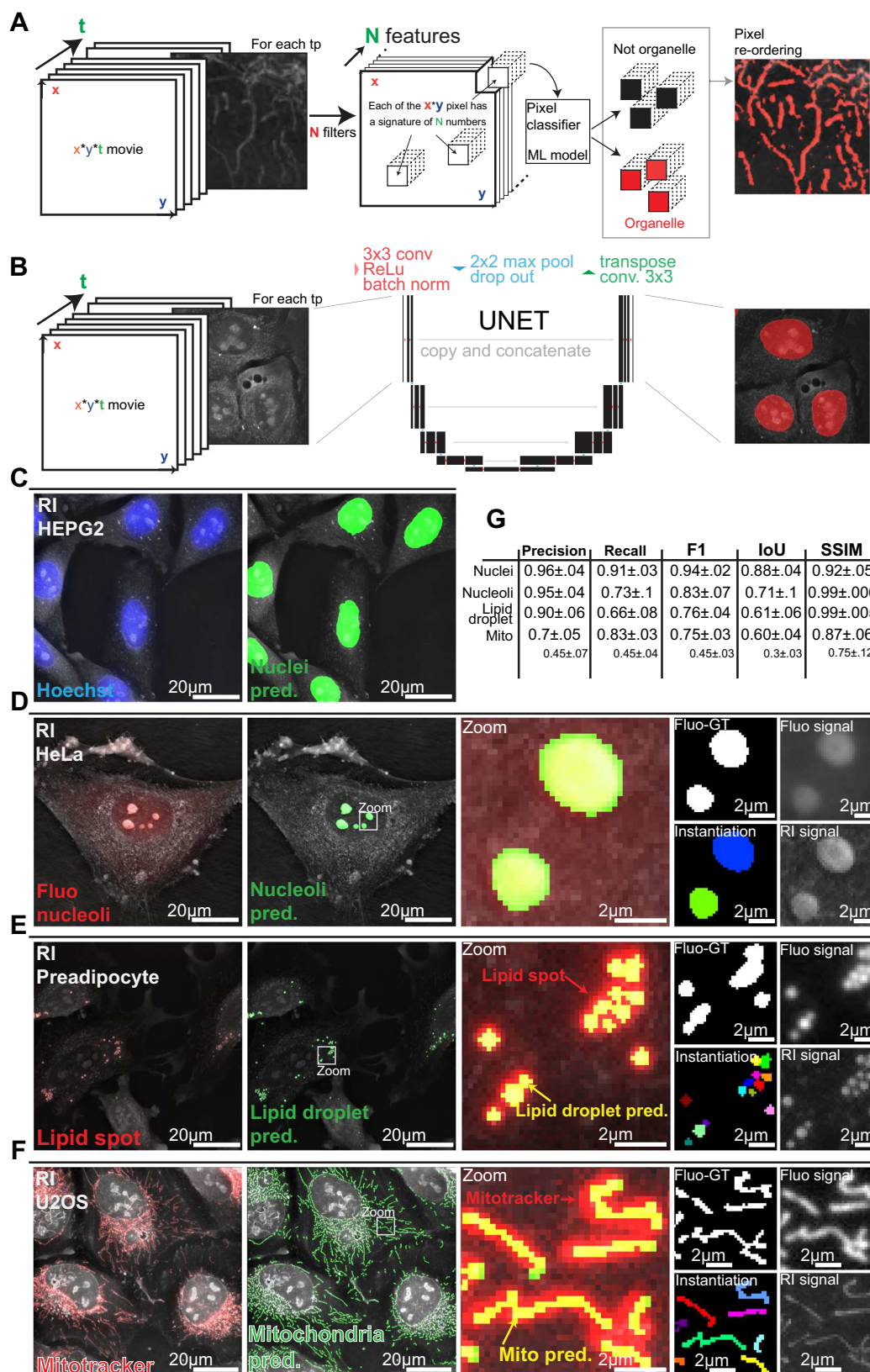
Fig. 1 | SARS-Cov-2 infected cells show subcellular dynamic changes.

A Refractive index (RI) map of U2OS cells infected with SARS-Cov-2, was acquired using holotomographic microscopy (HTM). **B** Representative images of non-infected cells (Mock), SARS-CoV-2 infected cells, or cells forming syncytia upon expression of the fusogenic protein Syncytin-1 (# of independent time-lapse

acquisitions ≥ 4). **C** Magnifications of cellular details available for further ML-aided image analysis, such as nuclei (1), mitochondria (2), nucleoli (3), lipid droplets (4), or syncytia nuclei cluster (5). **D** Outline of our time-lapse imaging analysis pipeline ensuring data processing, computer vision, and quantitative assessment of cells.

Our object segmentation was robust, allowing analysis over time within each movie, and across movies (Fig. 3 and Supplementary Movies 1–19). Of note, the formation of syncytia did not alter the quality of our predictions. Nuclei and nucleoli were properly detected despite the appearance of compact clusters of nuclei in syncytia (Fig. 3A and B). LD remained well-defined even when their size increased or when they moved from the cytosol to the perinuclear region (Fig. 3C and D and Supplementary Fig. 7). Mitochondria were accurately segmented, despite a large variety in length and distribution in infected cells (Fig. 3C and D).

We represented the data as time series of violin plots based on single cells or organelles points, to visualize the evolution of the various parameters over time and to allow explicit statistical assessment of infection-induced changes (Fig. 4). The progression of nucleoli/nucleus size ratio was similar in Wuhan-infected and non-infected cells, as well as in syncytia triggered by Syncytin-1 (Fig. 4A). In agreement with a recent report⁴⁰, the nucleoli of Omicron-infected cells became larger, especially at a late time point of infection. This might reflect the recruitment of the nucleolar machinery to facilitate viral translation^{41,42}. A massive



increase in LD number and size was observed in SARS-CoV-2 Wuhan- or Omicron-infected cells but not in control or Syncytin-1-expressing cells (Fig. 4B). These observations are in line with SARS-CoV-2-induced remodeling of lipid metabolism²¹ to support provirus signaling²³ and provide material for the formation of double-membrane vesicles (DMV) (Fig. 4B).

The obvious effect of SARS-CoV-2 on LD indicated a broad metabolic impact and thus led to the question of mitochondrial alterations. In control cells, we detected a modest yet significant reduction of mitochondrial size over the 1.5 days of culture, which may reflect the changing metabolic state of proliferating cells that progressively consumed culture medium (Fig. 4C). In both Wuhan and

Fig. 2 | Machine learning detects key organelles within HTM images. RI-refractive index, Mito pred-mitochondria predictions. **A** Mitochondria and lipid droplets detection using pixel classification: a feature space of size x^*y^*N is calculated by applying N convolution filters on each image time point (tp) of size x^*y . An extra tree classifier decides for each pixel if it belongs or not to an organelle signal based on its position in the feature space. **B** Nuclei, nucleoli, and cells detection using the convolutional network UNET. **C–F** Comparison of **C** nuclei, **D** nucleoli, **E** lipid droplet, and **F** mitochondria detection within refractive index images with

their respective fluorescent label signal made partially transparent such that underlying cellular structures are visible. Structures are thicker when visualized with epifluorescence microscopy compared to holotomography. **G** Precision, recall, and F1 score, as well as intersection over union (IoU) score and structural similarity index measure (SSIM) for each organelle. Nuclei and nucleoli prediction scores are calculated against fluo-derived ground truths (Fluo-GT). Lipid droplets and mitochondria are evaluated against expert-corrected or raw fluo-derived ground truth.

Omicron infected cells, there was a marked mitochondria size decrease around 15 h (900 min) pi (Fig. 4C), that is likely an infection-induced imbalance of mitochondrial dynamics⁴³. At later points, the length of the mitochondria increased again in infected cells. Contradictory findings of impaired mitochondrial fission and fusion in response to SARS-CoV-2 infection have been reported^{44–46}. Our results suggest that these events may be temporally distinct.

We then quantified the dry mass, defined as the bulk content in biomolecules (mainly proteins, lipids, and nucleic acid) that are not the water of hundreds of thousands of mitochondria. In fact, HTM returns the refractive index of the observed biological structures, which is linearly linked to the content in biomolecules of the observed structure. In control cells, we observed a stable dry mass per unit of mitochondria size overtime (Fig. 4D). This dynamic was changed in SARS-CoV-2-infected or Syncytin-1-expressing syncytia, which were characterized by an overall slight reduction of the dry mass of single mitochondria. This is in line with previous reports^{46,47} suggesting that SARS-CoV-2 down-regulates the translation of mitochondrial genes. The observation that Syncytin-1-expressing cells also displayed a decreased mitochondrial dry mass suggests that SARS-CoV-2 could use syncytia formation to promote mitochondrial alterations more efficiently to promote mitochondrial alterations more efficiently through molecular and structural mechanisms that still need to be discovered.

We next took advantage of our capacity to localize organelles in space and relative to each other to measure the distance between LD and mitochondria, a marker of the rate of fatty acid oxidation and thus of energy production^{33,34}. In uninfected or Syncytin-1-expressing cells, the proportion of LD in proximity (<400 nm) of mitochondria increased over time (Fig. 4E). In contrast, this ratio stayed stable or even decreased in Wuhan and Omicron-infected cells. Therefore, infection separates LDs from mitochondria, reflecting a probable impact of SARS-CoV-2 on cell metabolism.

Large lipid droplets form in infected cells only

We next examined the links that may exist between LD, mitochondria, and viral production zones, which we identified by immuno-staining with anti-NSP3 or anti-double stranded (ds) RNA antibodies. The viral NSP3 protein plays many roles in the virus life cycle, including viral polyprotein processing, formation of the viral replication compartment, viral RNA trafficking, and innate immunity antagonism⁴⁸. AntidsRNA antibody specifically recognizes double-stranded RNA (dsRNA) of >40 bp in length generated during the replication of positive sense genome viruses and thus selectively recognizes viral RNAs and not cellular RNAs⁴⁹, validated further by the absence of dsRNA signal in cells that do not display an NSP3 signal (Fig. 5A–C and Supplementary Fig. 7). NSP-3 and dsRNA, always both present in single cells showing fluorescent signal (Fig. 5A–D), rarely colocalized at the subcellular level as demonstrated by colocalisation experiments and by the absence of correlation of both signal shown by a Pearson coefficient of 0.34 (Fig. 5B and C). Correlative HTM and confocal microscopy show that at 24 h pi, cells that accumulated large perinuclear LD were also positive for dsRNA and NSP3 (Fig. 5A and Supplementary Fig. 3), confirming that the accumulation of large LD is a signature of SARS-CoV-2 infection. dsRNA accumulated next to LD, while NSP3 was excluded from them (Fig. 5B–D). We performed correlative HTM and electron

microscopy (CHEM) of non-infected or infected cells and focused on areas displaying large LD. The mitochondria surrounded large LD in control but not in SARS-CoV-2 infected cells (Fig. 5E, F and Supplementary Fig. 4). Altogether, the HTM quantitative analysis, combined with qualitative correlative microscopy indicates that LD alteration together with mitochondria relocation is a hallmark of effective virus production.

SARS-CoV-2 alters organelles cross-regulations

We next investigated the dependencies between LD, mitochondria, nuclei, nucleoli, and cytosol dry masses using Bayesian networks (BN) established on ten thousand different datasets of the same size as the original dataset, created through the random sampling of the original dataset while allowing replacement. BN is an established tool for modeling biological datasets⁵⁰ in fields such as signaling⁵¹, genomics^{52,53}, or immunology⁵⁴, and to model protein-nuclear bodies interactions⁵⁵, but not, to the best of our knowledge, to model the hierarchy and regulation existing between cell organelles. Established BN methods⁵⁶ allows us to search for the direct dependencies (while filtering out spurious correlations) and probabilities between factors of interest. These methods provide an intuitive visual representation in the form of a directed acyclic graph. We thus investigated, over all the bootstrapped datasets, the probability of discretized organelles' dry mass state, given the co-existing dry mass of the other organelles. We then reported under the form of inter-organelle networks the occurrence of specific directional dependencies. Considering that the dry mass variation of a subcellular compartment reflects regulated variations of its protein, lipid, and/or nucleic acid contents, we propose that the relationship between organelle dry mass captures an integrated level of organelle-dependent regulation. Henceforth, we will employ the term *organelle cross-regulation* (OCR) rather than referring to *dry mass influence diagrams*, the latter being less intuitive. As our approach relies on organelle detection and not on gene or protein level measurements, it is blind to organelle-independent global regulations that manifest during infection or cellular dysfunction, such as modulations of protein expression and other gene regulations that cannot be captured by assessing dry mass variations.

We also attributed a “regular” or “syncytium” identifying tag, to each cell of our control, SARS-COV-2-infected and Syncytin-1 expressing conditions. This allowed us to observe the specific impact of syncytia formation on OCR compared to infection. The differences between similarly established networks for the three conditions provide unique insights into how SARS-CoV-2 infection or syncytia formation impacts OCR (Fig. 6A–C). The nucleus has a strong influence on nucleoli, in all conditions, irrespective of infection or syncytia formation. This indicates that our networks can capture relevant functional relationships between organelles. Moreover, the nucleus dry mass sits in the center of the unperturbed OCR, which is expected since, in freely dividing cells, the cell cycle and thus the nucleus DNA content, synchronizes the rest of the OCRs (Fig. 6A).

We observed that SARS-CoV-2 infection rewired the OCR network (Fig. 6B), which can be explained by the fact that it can alter organelles involved in numerous processes, such as lipid metabolism^{21–23} and the cell cycle^{57–60}. When compared to the unperturbed condition (Fig. 6A), one can observe that the infection-induced OCR (Fig. 6B) lost the link between the nucleus and LDs, and mitochondria and LDs, confirming

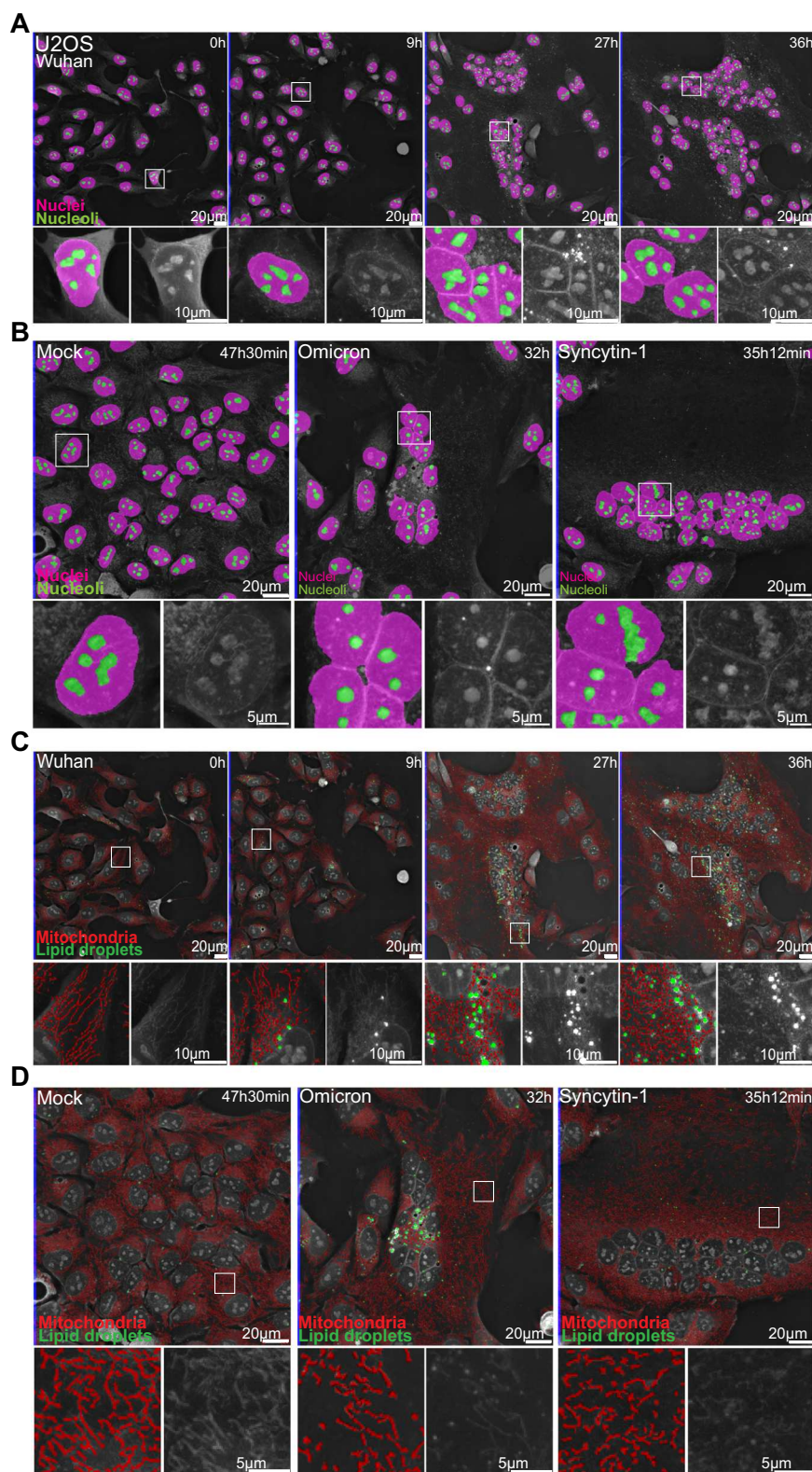


Fig. 3 | Automated detections of cellular organelles capture dynamics of SARS-CoV-2-induced changes. **A** 36 hours-long holotomographic microscopy (HTM) time-lapse acquisition of U2OS cells infected by the Wuhan SARS-CoV-2 strain. Pink: nuclei. Green: nucleoli. **B** Late time point images of time-lapse imaging experiments of non-infected cells (Mock), Omicron-infected cells, and Syncytin-1-expressing cells. Pink: nuclei. Green: nucleoli. **C** 36 hours-long holotomographic microscopy

(HTM) time-lapse acquisition of U2OS cells infected by the Wuhan SARS-CoV-2 strain. Red: mitochondria. Green: lipid droplet. **D** Late time point images of time-lapse imaging experiments of non-infected (Mock), Omicron-infected, and Syncytin-1-expressing cells. Red: mitochondria, Green: LD (# of independent time-lapse acquisitions ≥ 4).

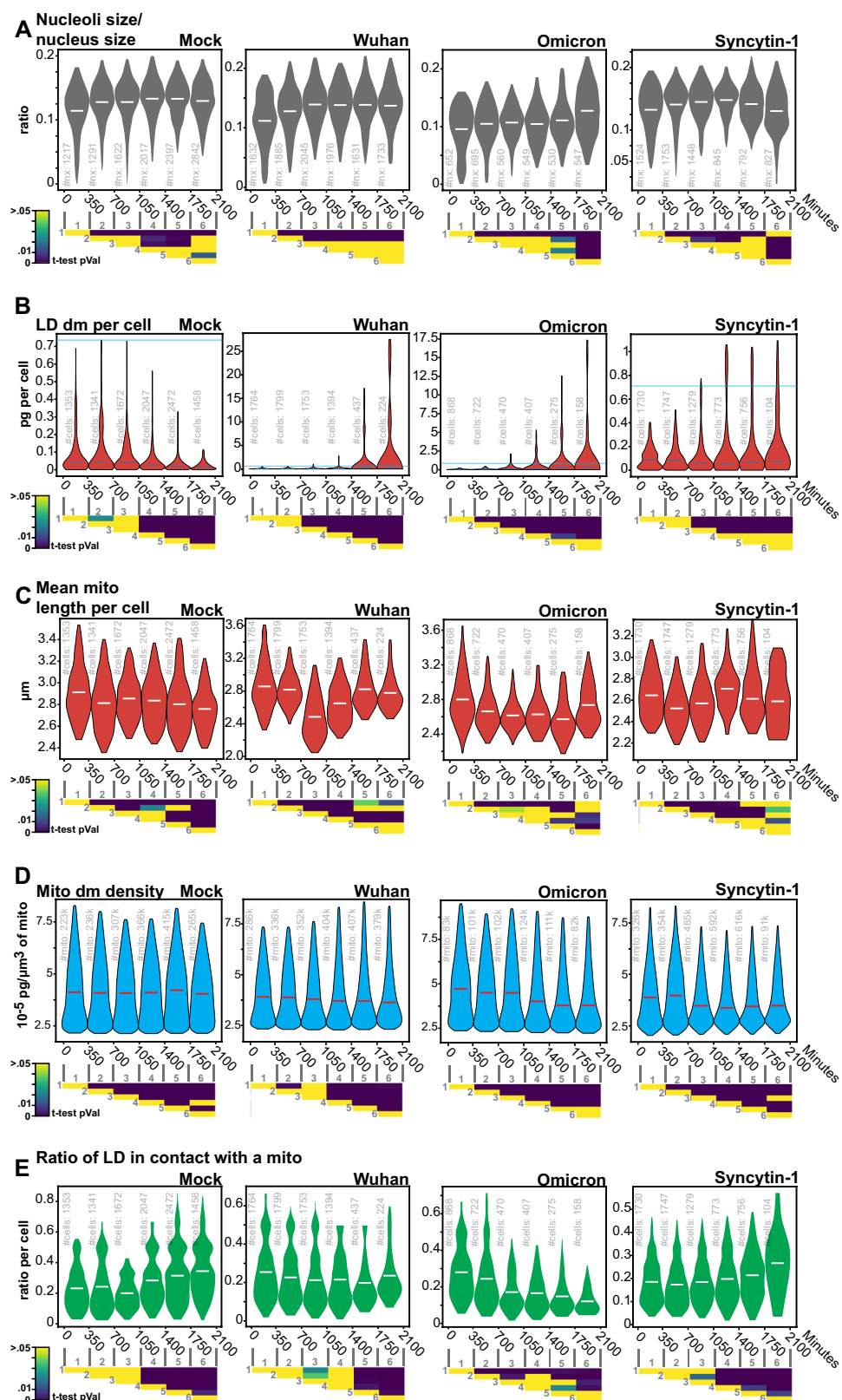


Fig. 4 | Label-free quantifications of SARS-CoV-2-induced alterations of organelle dynamics. **A** Nucleoli/nuclei size-ratio per cell over time (#of quantified nuclei in light-gray), **B** Lipid droplet dry mass per cell (#of quantified single cells in light-gray), **C** Mitochondria dry mass density (#of quantified mitochondria in light-gray), **D** Mean mitochondria length per cell (#of quantified cells in light-gray), **E** the ratio per cell of lipid droplets <400 nm away from the nearest mitochondria (#of quantified cells in light-gray). Violin plot representation of non-infected, Wuhan-,

Omicron-infected, and Syncytin-1-expressing cells. Each violin plot represents the distribution of the segmented single cells or organelles contained within the indicated period. Bin-to-bin two-sided *t*-tests *p*-values are indicated below each experiment. Single cells and organelles studied in each of the mock, Wuhan, Omicron, and Syncytin-1 coming from at least three different movies (see Supplementary Movies 1–19).

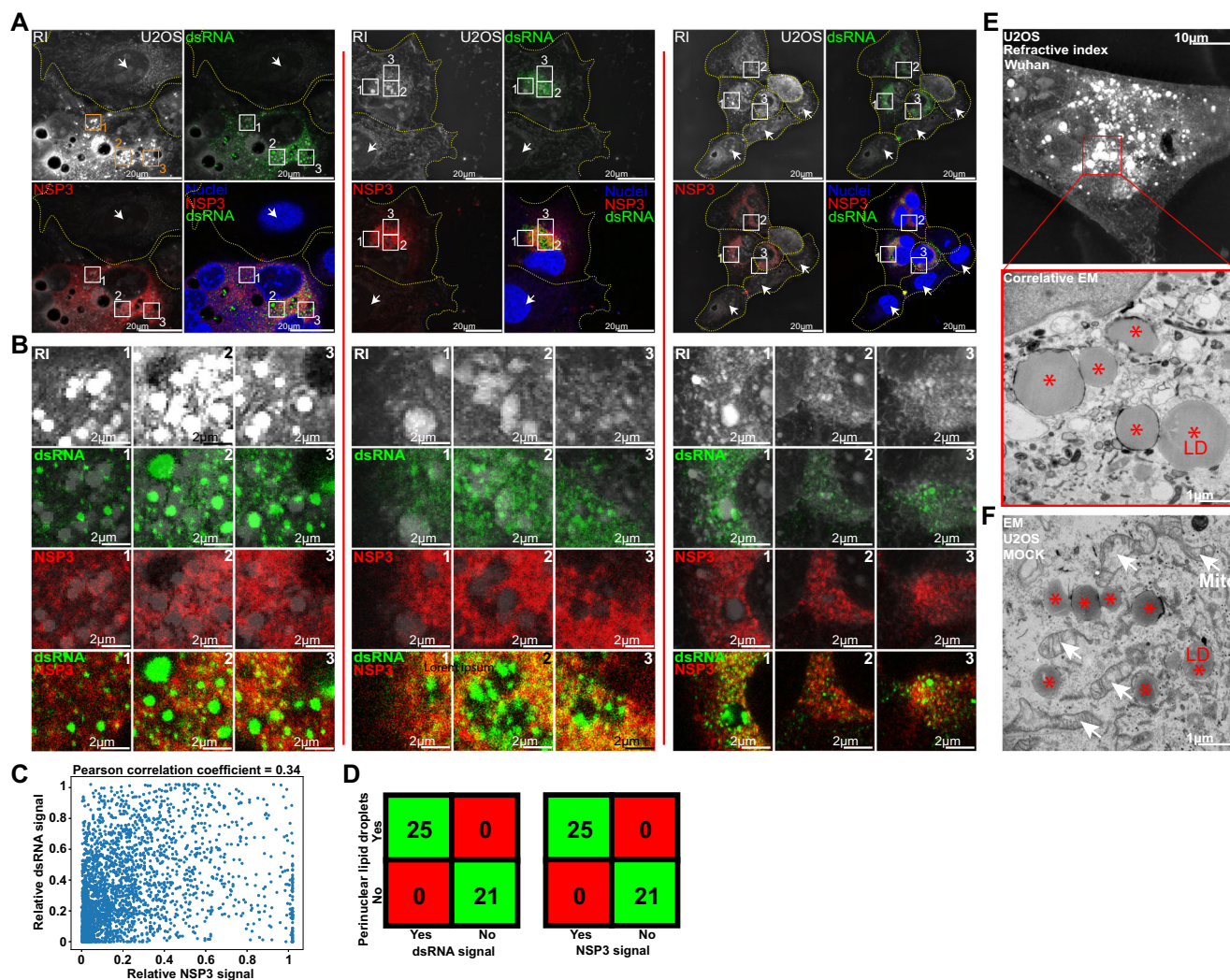


Fig. 5 | Perinuclear lipid droplet accumulation is a marker of infection. A and B Comparison of the refractive index (RI) signal acquired using holotomographic microscopy (HTM) with the double-stranded (ds) RNA and NSP3 immunofluorescent signals acquired with confocal microscopy. dsRNA spots are detected over and around lipid accumulations while the homogenous NSP3 signal is excluded from them (# of independent acquisitions = 8). **C** Scatter plot of NSP3 against dsRNA fluorescent signal and Pearson correlation coefficient of 0.34 confirm the absence of signal colocalization. **D** Confusion matrix. Cells with perinuclear lipid

droplet accumulation have SARS-CoV-2-induced dsRNA and NSP3 signals. Cells with no LD (arrow) show no dsRNA and NSP3 signals (see Supplementary Fig. 8). **E** Correlative holotomography/electron microscopy (EM) in infected cells. No mitochondria surround infection-induced lipid droplets (# of independent acquisitions = 5) (see also Supplementary Fig. 4). **F** Electron microscopy in non-infected cells, where LD is close to mitochondria (red stars: LD white arrows: mitochondria, # of independent acquisitions = 3).

our observations (Figs. 4E, 5D and E, Supplementary Figs. 8, 9) that SARS-CoV-2 lead to the spatial separation of LDs and mitochondria and suggesting a profound alteration of LDs metabolism. The infection-induced OCR network (Fig. 6B) was mostly altered around LDs, with minor to no changes elsewhere. Thus, our approach captures organelle connections, occurring during SARS-CoV-2 infection: LDs are specifically perturbed; the syncytium state is linked to LDs, and the OCR structure points towards LDs through the cytosol node (Fig. 6B).

The OCR of Syncytin-1-expressing cells (Fig. 6C) helps in understanding the relative roles of cell fusion versus viral infection.

In Syncytin-expressing cells, the syncytial state had no direct impact on one specific organelle and was only loosely related to the rest of the network through the cytosol (Fig. 6C). The formation of a syncytium is a complex, broad phenomenon that implies more than the fusion of plasma membranes on which we rely to identify it in our dataset. The molecular mechanisms at play, including cytoskeleton reorganization, nucleus clustering, and rapid mixing of different trafficking, signaling, and metabolic systems, are expected to have broad

consequences that will not be captured as directional dependencies but rather be seen in the way organelles are wired together. The core OCR linking cytosol, mitochondria, and the nuclei-nucleoli duo remains, confirmed the robustness of our approach. In these syncytia, the link between LDs and the rest of the OCR was no longer visible. Thus, one can speculate that one role of the syncytium is to change how organelles interact together. The virus may thus trigger the syncytium to establish a favorable OCR system to facilitate its spread.

Our bootstrapped BN analysis (Fig. 6B) suggested a direct link between the formation of the infection-induced syncytium and LDs. To further validate our systems approach, we performed a high-frequency HTM time-lapse experiment with one image being recorded every 2 min (Fig. 7A). The early moments of syncytium formation were generally followed by LD growth. While infectious and non-infectious syncytia looked globally similar, the non-infectious syncytia did not promote LD growth on its own (Figs. 3D and 4B).

We finally used confocal microscopy of tubulin immunostaining as a complementary approach to further investigate potential

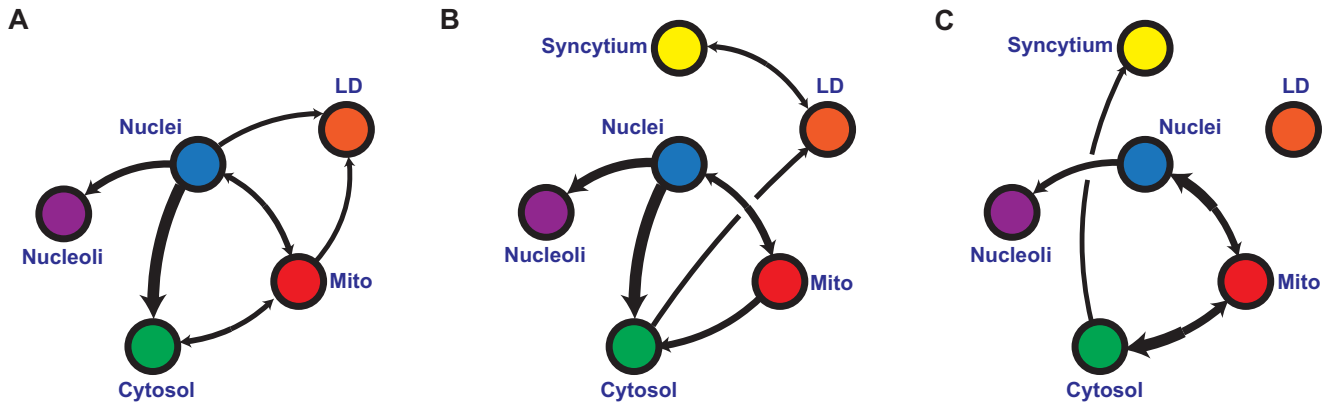


Fig. 6 | Bootstrapped Bayesian networks indicate that SARS-CoV-2 infection changes organelle cross-regulation. A–C Averaged Bayesian networks representing the dependencies (equivalence classes) between mitochondria, nuclei, nucleoli, lipid droplets, and cytosol dry mass states and the syncytial state of **A** non-infected cells, **B** infected cells, and **C** Syncytin-1-expressing cells over 10,000

randomly resampled datasets. All edges at each iteration were validated by two-sided chi-squared, *g*-squared, log-likelihood, and Cressie-Read pairwise tests, all displaying a *p*-value < 10⁻⁶ over discretized and non-discretized data (see Supplementary Tables 1–4 for two-sided chi-squared test details) (# of independent time-lapse acquisitions for each condition ≥ 4).

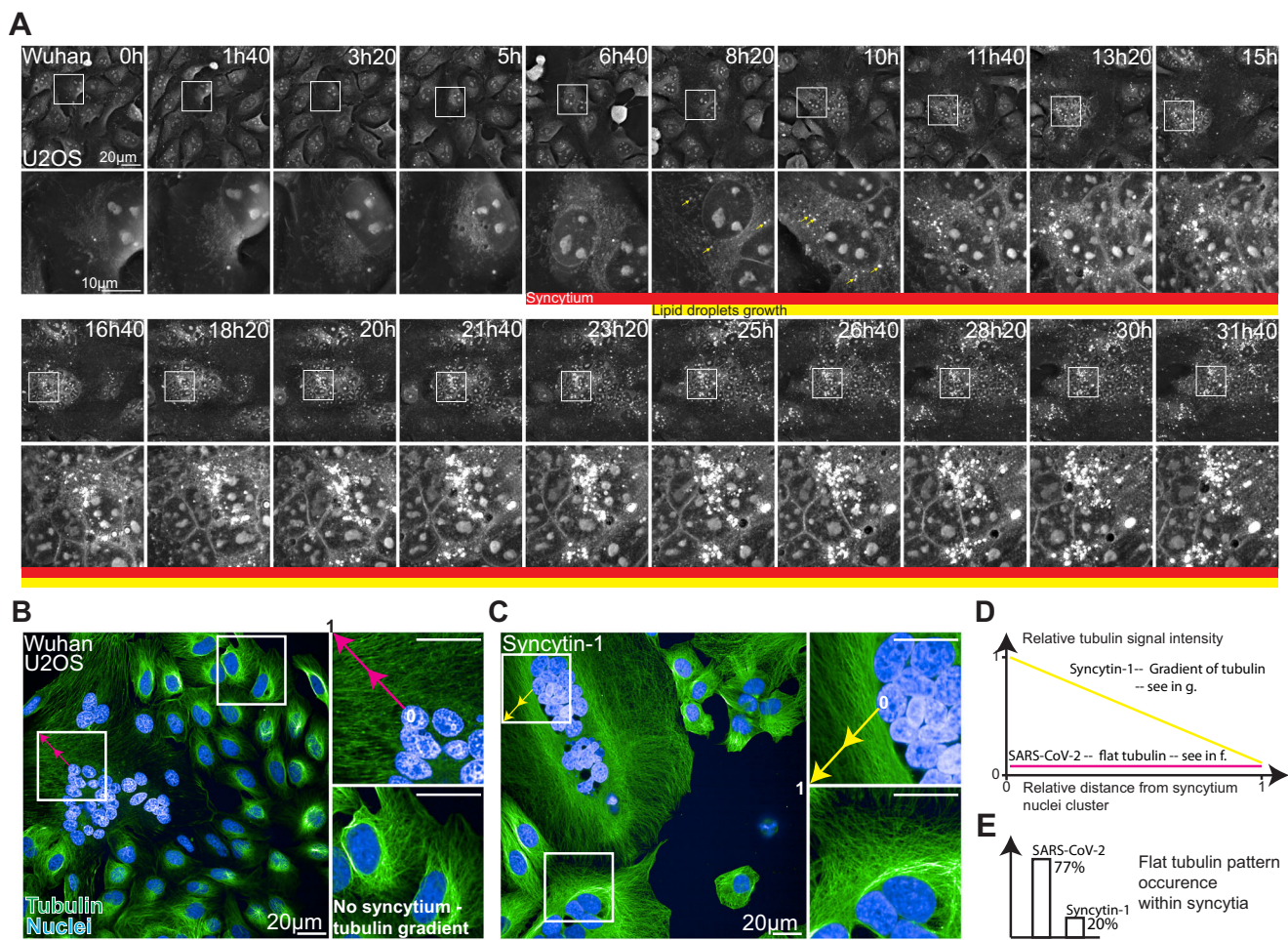


Fig. 7 | SARS-CoV-2-induced syncytia displays a flat radial tubulin signal and accelerates lipid droplet growth. **A** High-frequency time-lapse imaging of U2OS cells infected by SARS-CoV-2 Wuhan strain (see also Supplementary Movie 20). **B, C** Representative images of confocal immunofluorescence microscopy in SARS-

CoV-2 Wuhan- or Syncytin-1-induced syncytia. Green: Tubulin, Blue: DAPI. Pink and yellow arrows are the sections along which signals plotted in **D** are sampled. **D** Representative tubulin signal profiles from a nuclei cluster towards cell boundaries. **E** Occurrence of a flat radial actin pattern.

differences between the two types of syncytia. Compared to Syncytin-1 expressing cells, SARS-CoV-2-induced syncytia displayed a flat microtubule signal profile (Fig. 7B and C) from the clustered nuclei towards the syncytium boundaries (Fig. 7D). This flat microtubule signal

distribution was detected in 77% of the infected syncytia and only in 20% of Syncytin-induced fused cells (Fig. 7E). In the 80% remaining cells, a gradient was seen, with more tubulin surrounding the nuclei than in the cell boundaries. Thus, in addition to the events captured by

HTM, the cytoskeleton is differentially modified in SARS-CoV-2-infected or Syncytin1-induced syncytia. These differences could modulate LD accumulation and mitochondrial redistribution, opening future investigative directions.

Discussion

We demonstrate here that AI-enhanced label-free microscopy holds great promise for the investigation of pathological processes such as virus infection and offers possibilities to quantitative cell biology in general⁶¹. The emergence of high-content fluorescence microscopy⁶² and massively multiplexed labeling methods⁶³ revolutionized our understanding of cells' systems⁶⁴, whether they are unperturbed, adapting to changing micro-environments⁶⁵, or reacting to drug treatments⁶⁶. Yet, these methods perturb living cells and are not always suited for monitoring fine biological dynamics over long periods of time.

Among currently available label-free techniques, holotomographic microscopy (HTM) produces images whose quality enables effective computer vision solutions^{20,67}. To achieve this aim, it is important to overcome the laborious task of ground truth generation^{68,69} and to choose the adequate object detection technique, where deep learning is not always the most effective solution. The ground truth that we produced allowed for a U-NET segmentation of simple, large objects like cells, nuclei, or nucleoli, but was poorly suited for the detection of mitochondria that are sparser and thinner objects. A more explicit control of the receptive field in the case of the pixel classification approach, which depends on the blurring steps applied to images before feature extraction, was more adapted for objects spanning only a few pixels only. Our work is thus particularly relevant to the field of quantitative mitochondrial biology^{37,70,71}, given the susceptibility of that organelle to phototoxic stress and label-induced perturbations^{8,72,73}.

Our object detection strategy transforms the limitation of HTM, which is the complex nature of the images it provides, into an advantage: having access to biologically relevant, direct, simultaneous measurements of organelles within single cells. This opens a vast landscape of possibilities for system investigations. Such a dataset is particularly adapted for causal investigations due to the reasonable number of dependencies to evaluate (Supplementary Fig. 10). We could thus define in this study organelle cross-regulation (OCR) networks. There are pending questions: will the same OCR network topologies be observed in other cell types or in response to other viruses? Future work will help determine whether we have uncovered conserved regulation networks for synchronizing organelles operating in different cellular conditions.

There are limitations to our study. Firstly, Bayesian networks are directed acyclic graphs and, as such, cannot be circular or contain feedback loops, which are natural components of many biological systems. Thus, the use of causal inference provides a partial view of OCR networks and of their potential rewiring upon perturbations. Targeted metabolic or genetic perturbations to challenge the hypothesis gathered through network inference will help further characterize OCR dependencies. Secondly, key subcellular structures such as the Golgi or the endoplasmic reticulum are not captured by the current sensitivity of HTM and are therefore absent from our analysis. HTM could be associated with other label-free⁷⁴, fluorescence, or electron microscopy in correlative approaches. Such association is essential to advance both label-free and label-based imaging worlds. This is illustrated here by the assessment of mitochondrial localization by CHEM, the colocalization of LD, dsRNA, and NSP3 by correlative HTM/confocal microscopy, or the characterization of tubulin network modifications during SARS-CoV-2 infection and syncytia formation. Thirdly, we used only one cell type, U2OS cells expressing the ACE2 receptor, to establish the first set of discoveries in an uncharted territory. It will be important to extend such analysis to other cell types,

for instance, those that do not or poorly form syncytia upon infection. The use of viral strains carrying GFP-tagged viral proteins or other markers will facilitate the identification of infected cells⁷⁵. This will allow us to further explore the whole infection process and better characterize the isolated role of syncytia during infection such as the impact of almost instantaneous small molecules redistribution and the mixing of multiple signaling or metabolic states. Fourthly, to ensure capturing all biological structures despite dynamic variations in z distribution, we work with a typical 2D projection procedure that is applied similarly to all 3D sample reconstructions (Supplementary Fig. 11a, b). Thus, our quantifications approach the true dry mass content rather than being absolute (Supplementary Fig. 11c). It is important to mention that a full 3D-based quantification is not an absolute reference either, even with the best possible microscopic device, absolute quantitative knowledge is out of reach. This is why controlled experiments like those made in our study, are and will remain essential. Importantly, a 3D-based quantitative investigation at a similar experimental scale represents a massive technical and computational challenge that is a matter of future studies. Finally, our work quantitatively describes the multiple cellular events associated with SARS-CoV-2 infection in real-time. An analysis of cells expressing individual viral proteins, such as, for instance, ORF9b⁷⁶, that interacts with the mitochondrial protein TOMM70⁷⁷, or NSP6, that mediates contact between DMVs and LDs⁷⁸, will provide clues on the impact of SARS-CoV-2 on the dynamics and shape of mitochondria.

In summary, we developed an approach combining HTM, AI-assisted analysis, and causality inferences using Bayesian network modeling, to assess the impact of SARS-CoV-2 on the dynamics of cellular organelles and OCR. We report that the virus directly alters LDs, mitochondria, nuclei, and nucleoli. We describe how those organelles can influence each other. We also propose that the infectious syncytium is likely favoring a pro-virus cellular environment. This approach opens exciting possibilities to analyze any pathogen, drug effects, and physiological or pathological events affecting cell life, including nutrient variations, metabolic adaptation, and malignant transformation. It holds promise to lead to insights into the dynamics of a vast range of biological processes.

Methods

Holotomographic microscopy time-lapse acquisitions

All label-free images were acquired using a 3D-Cell Explorer-fluo (Nanolive SA, Tolochenaz, Switzerland) microscope. This microscope is equipped with a 520 nm laser for tomographic phase microscopy; the irradiance of the laser is 0.2 nW/μm², and the acquisition time per frame is 45 ms. It is equipped with a Blaser ace aca1920-155um camera, and an air objective lens (NA = 0.8, magnification ×60). The microscope is equipped with a fluorescent module (pE-300ultra, CoolLED) for standard epifluorescence images of the sample in three different channels: Cy5 (excitation peak 635 nm), TritC (excitation peak 554 nm) and FitC (excitation peak 474 nm). The microscope is equipped for long-term live cell imaging: temperature, humidity, and gas composition. The incubator chamber (Okolab) keeps the sample at 37 °C, is closed by a heating glass lid to prevent condensation, and is connected to a gas mixer (2GF-Mixer, Okolab) to maintain 5% of CO₂. The humidity module ensures a 90% relative humidity within the chamber. The 520 nm laser is divided into two beams to create an interferometer setup⁷⁹. One of the beams, referred to as the object beam, interacts with the sample before being gathered by the 60× objective lens. Simultaneously, the other beam, left unperturbed, functions as the reference beam. When meeting on a dichroic mirror, the two beams create an interference pattern a.k.a. a hologram, further captured by a CMOS camera (Supplementary Fig. 1a). This conventional holographic technique is augmented with the rotational scanning of the sample⁸⁰ (Supplementary Fig. 1b) using a rotating mirror that reduces the problem of coherent diffraction noise⁸¹ to imperceptible levels

(Fig. 1B and C). The series of holograms are then assembled by complex field deconvolution^{82,83}. Our HTM device incorporates dynamically adjustable mirrors that enable its optics to self-regulate throughout the acquisition and accommodate variations in the sample, such as medium evaporation or variations in meniscus curvature. Combined with an automated stage and a laser based autofocus, our HTM device has stable performances over long periods of time. Finally, the HTM device possesses an epifluorescence module allowing for simultaneous fluorescence and RI imaging (Supplementary Fig. 1c).

For imaging, 50,000 cells were plated in a 35 mm No.1.5 ibidi polymer coverslip bottom dish. After 24 h, the media was replaced with fresh media containing the indicated dose of the virus. Dish were then placed in the incubator chamber of the microscope and imaging was started 3 h post-infection.

Live fluorescent controls

For mitochondria labeling, TMRE was added to the media for 30 min, before washing and cell imaging. For siRNA experiments, Dharmacon smartpool siRNA directed against human-OPA1 siRNA (Dharmacon SmartPool M-005273-00-0005) or a control siRNA (Dharmacon SmartPool D-001210-02-05) was used. For nucleoli labeling, cells were transfected using lipo3000, as described by the manufacturer, with GFP-GRI, a subunit of the Simian Foamy virus that contains a nucleolar localization signal fused with GFP⁸⁴. 24 h post-transfection cells were imaged using the Nanolive. For lipid droplet labeling, Bodipy 493/503 was added in serum-free media for 15 min, before washing and imaging.

Immunofluorescent labeling

Infected cells were fixed with 4% PFA for 30 min. For correlation with Nanolive microscopy, the dish was scratched to provide a visual landmark. They were then imaged on the Nanolive. They were permeabilized with 1% Triton for 10 min and blocked overnight with PBS/1% BSA/0.05% sodium azide. Mouse anti-dsRNA J2 (1:100, RNT-SCI-10010200; Jena Bioscience), sheep anti-SARS-CoV-2 nsp3 (1:200, DU67768, MRC PPU) were used overnight in blocking buffer. Donkey anti-sheep 488 (#A-11015, Thermo Fischer Scientific) and Donkey anti-mouse 647 (#A-31571, Thermo Fischer Scientific) were used at a 1:500 dilution in blocking buffer for 1 h. Imaging was performed using a Leica SP8 confocal microscope. For cytoskeleton imaging, infection was performed in a 96-well plate (Perkin Elmer). Infected cells were fixed with 4% PFA for 30 minutes, permeabilized with 1% Triton for 10 min, and blocked overnight with PBS/1% BSA/0.05% sodium azide. Mouse anti-tubulin (1:100, ProteinTech 66031-1-Ig) was used overnight in the blocking buffer. Goat anti-mouse 488 (#A-11015, Thermo Fischer Scientific, #A-11001) was used at a 1:500 dilution in a blocking buffer for 1 h. Imaging was performed using the Operetta Phenix (Perkin Elmer).

Electron microscopy

50'000 cells were plated in a MatTek 35 mm Dish (P35G-1.5-14-C-GRD). The next day, the media was replaced with fresh media containing the indicated dose of the virus. After 24 h, they were fixed with 4% PFA for 30 min and imaged using holotomographic microscopy. Cells were then fixed in 2.5% glutaraldehyde (Sigma) in 1X PHEMS buffer overnight at 4 °C. Samples were washed 3 times in 1X PHEM buffer and post-fixed in 2% OsO₄ (Electron Microscopy Sciences) + 1.5% potassium ferrocyanide in water 1 h in the dark. After three washes in water, they were incubated for 20 min in 1% uranyl acetate in ethanol 25%. Samples were gradually dehydrated in an ethanol series (50%, 75%, 95%, 3 × 100%) and then embedded in EMBED-812 epoxy resin (Electron Microscopy Sciences), followed by polymerization for 48 h at 60 °C. Thin sections (70 nm) of the region of interest were cut with a Leica Ultramicrotome Ultracut UCT stained with uranyl acetate and lead citrate. Images were acquired using a Tecnai T12 120 kV (Thermo Fisher) with a bottom-mounted EAGLE camera.

Cells

HeLa, 3T3-derived preadipocytes, HEPG2, and U2OS cells were purchased from ATCC. U2OS cells were stably transduced with pLenti6-human-ACE2 as previously described²⁹ or with pCW57.1_Syn1. The doxycycline-inducible PCW57.1_Syn1 plasmid was obtained by cloning the previously described pHCMV_Syn1⁸⁵ into pCW57.1 (Addgene) using the NheI + AgeI restriction for both plasmids. U2OS cells stably expressing PCW57.1_Syn1 were generated by lentiviral transduction and puromycin selection. Synectin-1 expression was triggered by treatment with doxycycline. All cells were cultured in DMEM with 10% fetal bovine serum (FBS), 1% penicillin/streptomycin (PS). For imaging and infections, 25 nM HEPES was added to the media. 10 µg/ml blasticidin was added to U2OS cell cultures. Cells were routinely screened for mycoplasma.

Virus

The reference Wuhan strain and the Omicron BA.1 variant have been described previously^{86,87}. The viruses were isolated from nasal swabs of two patients using Vero E6 cells and amplified by one or two passages. Ethical regulations were followed. Both patients provided informed consent for the use of the biological materials. Titration of viral stocks was performed on Vero E6 cells, with a limiting dilution technique enabling the calculation of the median tissue culture infectious dose, or on S-Fuse cells⁸⁸.

Holotomographic image processing

All the time points composing the time-lapse imaging experiments were reconstructed and stored as 3D volumes by Eve, the software that controls the 3D cell explorer microscope and early data management. Custom routines were used to convert 3D volumes into 2D maximal projections along the z-axis from -2 to +6 µm around the point of focus, which ensures always capture of entire cells and their organelles while those biological objects may move along the z-axis.

Deep Learning object detection

A custom U-Net architecture (Supplementary Fig. 2, model available on https://github.com/SlowProspero/SARS_AI/) was used to segment nuclei, nucleoli, and cells.

The U-Net consists of an encoder-decoder structure meeting at a bottleneck layer. The encoder component has four blocks, each comprising three 3 × 3 convolutional layers followed by rectified linear unit (ReLU) activation and batch normalization. Between these blocks, 2 × 2 max-pooling layers are used to down-sample the feature maps, and dropout layers of 0.5 are used to prevent overfitting. The input layer is the full 448 × 448 image. Because of the 3 × 3 convolutional layer, and 2 × 2 max pooling layer, the following blocks are then 224 × 224 pixels and 32 feature maps, 112 × 112 and 64 feature maps, 56 × 56 and 128 feature maps, 28 × 28 and 256 feature maps.

The decoder component is symmetric to the encoder, it consists of four blocks each like the encoder blocks. However, instead of max-pooling layers, transposed convolutional layers (also known as deconvolutional layers) are utilized to up-sample the feature maps. The first decoder block is in common with the last encoder block. Skip connections are established by concatenating the feature maps from the corresponding encoder part to its symmetrical decoder part. Finally, the size of the bottleneck layer is 28 × 28 and 512 feature maps.

Predictions in images larger than 448 × 448 involve making patches of prediction of size 448 × 448. Such window is then moved across the image stepwise 50 by 50 pixels, horizontally and vertically, and the overlapping predictions are averaged.

Training processes were run in a tensor-flow 2.10.1/Python 3.8.10 environment for 500 epochs for the nuclei and nucleoli models on an Nvidia RTX3060 GPU with 12 GB of VRAM and for 50 epochs for the cell models on an Nvidia RTX A4000 16GB of VRAM. The ground truths for the cells, nuclei, and nucleoli, were made semi-manually with the

help of a custom labeling tool and the guidance of fluorescently stained cells. No ground truth was generated for the cytosol compartment as this mask type is generated from subtracting organelles masks from the cell mask, as described below in the dedicated section.

Nucleus

Our nucleoli model was trained with a dataset of 655 holotomographic microscopy images of mammalian cells images randomly split into a training set (589 images) and a testing set (66 images). These images come from 58 acquisitions and include multiple cell lines. Each image contains from 1 to 23 nuclei; all have a size of 480×480 pixels achieved by cropping or zero-padding. Probability maps provided by the model are binarized with a 50% probability threshold. All objects smaller than 10 pixels are rejected. The training took 4 h and 19 min.

Nucleoli

Our nucleoli model was trained with a dataset of 495 holotomographic microscopy images of mammalian cells randomly split into a training set (445 images) and a testing set (50 images). Each image contains at least a dozen nucleoli and has a size of 480×480 pixels achieved by cropping or zero-padding. Probability maps provided by the model are binarized with a 50% probability threshold. All objects smaller than 10 pixels are rejected. Potential unspecific signal outside the nucleus is removed thanks to the nucleus masks predicted by our nucleus model. The training took 4 h and 13 min.

Cells segmentation

The first part of our cell segmentation process aims at obtaining rough cell segmentation without precise cell border detection. This U-Net model was trained with a dataset of 1445 holotomographic microscopy images of mammalian cells randomly split into a training set (1295 images) and a testing set (150 images). Each image contains from 1 to 10 cells and has a size of 480×480 pixels achieved by cropping or zero-padding. Probability maps provided by the model are binarized with a 50% probability threshold. All objects smaller than 10 pixels are rejected. The training took 51 min. The second part of our cell segmentation process takes the U-Net produced cell blobs binary masks as a seed for a precise cell edges detection using a propagation algorithm approach.

Ensemble pixel classification for mitochondria detection

Mitochondria were segmented using a tailored pixel classification approach inspired by previous seminal work³². Our code can be found at https://github.com/SlowProspero/SARS_AI/. Because the simple refractive index value is not a reliable organelle signature, we increased the pixel feature space dimensionality by applying on refractive index images a set of convolution filters using the VIGRA Computer Vision Library (v1.11.1) for a total of 21 dimensions. Thus, each pixel is described by its respective refractive index value and by its value in the 20 convolved images created by applying the following 10 filters on the refractive index image at two different sigma values (1.4 and 2.0): gaussian smoothing, difference of gaussian, gaussian gradient, gaussian gradient magnitude, hessian of gaussian, hessian of gaussian eigenvalue, Laplacian of gaussian, structure tensor eigenvalue, tensor determinant, and tensor trace. The probability for each augmented pixel of our images to be part of a mitochondria or lipid droplet was then evaluated by an extra-tree classifier (scikit-learn v1.2.2) diverging from the default hyperparameter setup only by the number of estimators that is equal to 200. Each resulting probability image is then transformed into a binary mask using an adaptive background removal approach (OpenCV, cv2.threshold). Our mitochondria extra-tree classifier was trained thanks to the labeling of 149 images from two different cell lines (CHO and Preadipocytes) coming from 14 different acquisitions.

Detection of lipid droplets

Lipid droplets were automatically segmented using Nanolive lipid droplet assay software that identifies spherical structures of even signal distribution over specific local refractive index maxima. The produced lipid droplet segmentations were further used for custom metrics calculations.

Definition of the cytosol compartment

The cytosol compartment is defined as the mask created by the subtraction of all organelles masks (Nucleus, nucleoli, lipid droplets, mitochondria) from the cell segmentation mask.

Metrics calculation

The instantiated masks of cells, mitochondria, lipid droplets, nuclei, and nucleoli and cytosol were used to extract the spatial (shape and size descriptors, centroid position) and RI-derived (dry mass, textures) features of each of the mentioned biological objects in each frame of each time-lapse experiment (Supplementary Movies 1–19). This was performed using a scientific Python environment and the library scikit-image. Each object's dry mass content was calculated from its refractive index value using a linear calibration model³². The data were exported as a.csv file into a Python environment and plotted with the Python library matplotlib for figure making or used through the bnlearn python library to establish our organelle cross-regulation networks.

Bayesian network structure learning

We performed our Bayesian network (BN) structure learning using the bnlearn⁵⁶ python package, to determine which BN captures the best the directional dependencies that exist between our dataset variables, lipid droplets, mitochondria, nucleus, nucleoli, and cytosol dry mass densities, as well as the nature of the single cell, normal or syncytium. This was performed 10,000 times in a typical bootstrapping approach where each of the 10,000 datasets has the same size as the original dataset and is generated by random sampling with replacement. This approach allows the capturing of the most robust network structure through BN averaging and diminishes the impact of potential experimental and time-related technical biases.

The best BN for each condition over each dataset was defined using greedy score-based structure learning. For this study, we used the K2 and hill climb search-and-score algorithms, which incrementally test BN alternatives to improve the default scoring function BIC, which itself determines the probability of the BN structure given its training data.

We performed multiple independence tests to define the edge strengths of our BN using each of the three models and their related discretized and non-discretized data. For each pair of our learned BN, a chi-squared, g -squared, loglikelihood, and Cressie–Read statistical test is performed to determine its p -value and associated score. The mean chi squared scores over all bootstrapped iterations are then used as a quantification of the strength of the tested edges. Only edges satisfying the pairwise statistical test and observed at a frequency superior to 0.15 are displayed (see Supplementary Tables 1–4).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Source data are provided with this paper. All data supporting the findings of this study are available within the paper and its supplementary information. Further information can be requested from the lead contact at mathieu.frechin@nanolive.ch. Source data are provided with this paper.

Code availability

The Bayesian network code is available at https://github.com/SlowProspero/SARS_AI/.

References

- Rothan, H. A. & Byrareddy, S. N. The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak. *J. Autoimmun.* **109**, 102433 (2020).
- Lamers, M. M. & Haagmans, B. L. SARS-CoV-2 pathogenesis. *Nat. Rev. Microbiol.* **20**, 270–284 (2022).
- Perlman, S. & Peiris, M. Coronavirus research: knowledge gaps and research priorities. *Nat. Rev. Microbiol.* **21**, 125–126 (2023).
- V'kovski, P., Kratzel, A., Steiner, S., Stalder, H. & Thiel, V. Coronavirus biology and replication: implications for SARS-CoV-2. *Nat. Rev. Microbiol.* **19**, 155–170 (2021).
- Yang, S. L. et al. Comprehensive mapping of SARS-CoV-2 interactions in vivo reveals functional virus–host interactions. *Nat. Commun.* **12**, 5113 (2021).
- Mendonça, L. et al. Correlative multi-scale cryo-imaging unveils SARS-CoV-2 assembly and egress. *Nat. Commun.* **12**, 4629 (2021).
- Cortese, M. et al. Integrative Imaging Reveals SARS-CoV-2-Induced Reshaping of Subcellular Morphologies. *Cell Host Microbe* **28**, 853–866.e5 (2020).
- Phototoxicity revisited. *Nat. Methods* **15**, 751 (2018). <https://doi.org/10.1038/s41592-018-0170-4>.
- Ettinger, A. & Wittmann, T. Fluorescence live cell imaging. *Methods Cell Biol.* **123**, 77–94 (2014).
- Douthwright, S. & Sluder, G. Live cell imaging: assessing the phototoxicity of 488 and 546 nm light and methods to alleviate it. *J. Cell. Physiol.* **232**, 2461–2468 (2017).
- Trewin, A. J. et al. Light-induced oxidant production by fluorescent proteins. *Free Radic. Biol. Med.* **128**, 157–164 (2018).
- Kilian, N. et al. Assessing photodamage in live-cell STED microscopy. *Nat. Methods* **15**, 755–756 (2018).
- Reiche, M. A. et al. When light meets biology—how the specimen affects quantitative microscopy. *J. Cell Sci.* **135**, jcs259656 (2022).
- Park, J. et al. Artificial intelligence-enabled quantitative phase imaging methods for life sciences. *Nat. Methods* **20**, 1645–1660 (2023).
- Ounkomol, C., Seshamani, S., Maleckar, M. M., Collman, F. & Johnson, G. R. Label-free prediction of three-dimensional fluorescence images from transmitted-light microscopy. *Nat. Methods* **15**, 917–920 (2018).
- Guo, S.-M. et al. Revealing architectural order with quantitative label-free imaging and deep learning. *eLife* **9**, e55502 (2020).
- Goswami, N. et al. Label-free SARS-CoV-2 detection and classification using phase imaging with computational specificity. *Light Sci. Appl.* **10**, 176 (2021).
- Hu, C. et al. Live-dead assay on unlabeled cells using phase imaging with computational specificity. *Nat. Commun.* **13**, 713 (2022).
- Pirone, D. et al. Stain-free identification of cell nuclei using tomographic phase microscopy in flow cytometry. *Nat. Photonics* **16**, 851–859 (2022).
- Sandoz, P. A., Tremblay, C., Gisou van der Goot, F. & Frechin, M. Image-based analysis of living mammalian cells using label-free 3D refractive index maps reveals new organelle dynamics and dry mass flux. *PLoS Biol.* **17**, 1–22 (2019).
- Farley, S. E. et al. A global lipid map reveals host dependency factors conserved across SARS-CoV-2 variants. *Nat. Commun.* **13**, 3487 (2022).
- Tabata, K. et al. Convergent use of phosphatidic acid for hepatitis C virus and SARS-CoV-2 replication organelle formation. *Nat. Commun.* **12**, 7276 (2021).
- Dias, S. S. G. et al. Lipid droplets fuel SARS-CoV-2 replication and production of inflammatory mediators. *PLoS Pathog.* **16**, e1009127 (2020).
- Taha, T. Y. et al. Rapid assembly of SARS-CoV-2 genomes reveals attenuation of the Omicron BA.1 variant through NSP6. *Nat. Commun.* **14**, 2308 (2023).
- Pahmeier, F. et al. Identification of host dependency factors involved in SARS-CoV-2 replication organelle formation through proteomics and ultrastructural analysis. *J. Virol.* **97**, e00878–23 (2023).
- Flynn, R. A. et al. Discovery and functional interrogation of SARS-CoV-2 RNA–host protein interactions. *Cell* **184**, 2394–2411.e16 (2021).
- Seth, R. B., Sun, L., Ea, C.-K. & Chen, Z. J. Identification and characterization of MAVS, a mitochondrial antiviral signaling protein that activates NF- κ B and IRF3. *Cell* **122**, 669–682 (2005).
- Knoops, K. et al. SARS-coronavirus replication is supported by a reticulovesicular network of modified endoplasmic reticulum. *PLoS Biol.* **6**, e226 (2008).
- Buchrieser, J. et al. Syncytia formation by SARS-CoV-2-infected cells. *EMBO J.* **39**, e106267 (2020).
- Rajah, M. M. et al. SARS-CoV-2 Alpha, Beta, and Delta variants display enhanced spike-mediated syncytia formation. *EMBO J.* **40**, e108944 (2021).
- Mi, S. et al. Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. *Nature* **403**, 785–789 (2000).
- Sommer, C., Straehle, C., Kothe, U. & Hamprecht, F. A. Ilastik: interactive learning and segmentation toolkit. In *2011 IEEE International Symposium on Biomedical Imaging: from Nano to Macro* 230–233 (IEEE, 2011).
- Akiba, T., Sano, S., Yanase, T., Ohta, T. & Koyama, M. Optuna: a next-generation hyperparameter optimization framework. In *Proc. 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* 2623–2631 (ACM, 2019).
- Ronneberger, O., Fischer, P. & Brox, T. U-Net: convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science* (eds Navab, N., Hornegger, J., Wells, W. & Frangi, A.) 234–241 (Springer, 2015).
- Carpenter, A. E. et al. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.* **7**, R100–R100 (2006).
- Wang, Z., Bovik, A. C., Sheikh, H. R. & Simoncelli, E. P. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* **13**, 600–612 (2004).
- Lefebvre, A. E. Y. T., Ma, D., Kessenbrock, K., Lawson, D. A. & Digmán, M. A. Automated segmentation and tracking of mitochondria in live-cell time-lapse images. *Nat. Methods* **18**, 1091–1102 (2021).
- Olichon, A. et al. Loss of OPA1 perturbs the mitochondrial inner membrane structure and integrity, leading to cytochrome c release and apoptosis. *J. Biol. Chem.* **278**, 7743–7746 (2003).
- Cretin, E. et al. High-throughput screening identifies suppressors of mitochondrial fragmentation in OPA1 fibroblasts. *EMBO Mol. Med.* **13**, e13579 (2021).
- Merino, V. F. et al. Nucleolin mediates SARS-CoV-2 replication and viral-induced apoptosis of host cells. *Antivir. Res.* **211**, 105550 (2023).
- Neumüller, R. A. et al. Conserved regulators of nucleolar size revealed by global phenotypic analyses. *Sci. Signal.* **6**, ra70 (2013).
- Ma, T.-H. et al. Genetic control of nucleolar size: an evolutionary perspective. *Nucleus* **7**, 112–120 (2016).
- Wang, P. et al. A cross-talk between epithelium and endothelium mediates human alveolar–capillary injury during SARS-CoV-2 infection. *Cell Death Dis.* **11**, 1042 (2020).

44. Bhowal, C., Ghosh, S., Ghatak, D. & De, R. Pathophysiological involvement of host mitochondria in SARS-CoV-2 infection that causes COVID-19: a comprehensive evidential insight. *Mol. Cell. Biochem.* <https://doi.org/10.1007/s11010-022-04593-z> (2022).
45. Shi, C.-S. et al. SARS-coronavirus open reading frame-9b suppresses innate immunity by targeting mitochondria and the MAVS/TRAF3/TRAF6 signalosome. *J. Immunol.* **193**, 3080–3089 (2014).
46. Singh, K. et al. Network analysis and transcriptome profiling identify autophagic and mitochondrial dysfunctions in SARS-CoV-2 infection. *Front. Genet.* **12**, 599261 (2021).
47. Miller, B. et al. Host mitochondrial transcriptome response to SARS-CoV-2 in multiple cell models and clinical samples. *Sci. Rep.* **11**, 3 (2021).
48. Lei, J., Kusov, Y. & Hilgenfeld, R. Nsp3 of coronaviruses: structures and functions of a large multi-domain protein. *Antivir. Res.* **149**, 58–74 (2018).
49. Weber, F., Wagner, V., Rasmussen, S. B., Hartmann, R. & Paludan, S. R. Double-stranded RNA is produced by positive-strand RNA viruses and DNA viruses but not in detectable amounts by negative-strand RNA viruses. *J. Virol.* **80**, 5059–5064 (2006).
50. Needham, C. J., Bradford, J. R., Bulpitt, A. J. & Westhead, D. R. Inference in Bayesian networks. *Nat. Biotechnol.* **24**, 51–53 (2006).
51. Sachs, K., Gifford, D., Jaakkola, T., Sorger, P. & Lauffenburger, D. A. Bayesian network approach to cell signaling pathway modeling. *Sci. Signal.* **2002**, pe38 <https://doi.org/10.1126/stke.2002.148.pe38> (2002).
52. Wang, L., Audenaert, P. & Michoel, T. High-dimensional Bayesian network inference from systems genetics data using genetic node ordering. *Front. Genet.* **10**, 1196 (2019).
53. Angelopoulos, N., Chatziplis, A., Nangalia, J., Maura, F. & Campbell, P. J. Bayesian networks elucidate complex genomic landscapes in cancer. *Commun. Biol.* **5**, 306 (2022).
54. Rodin, A. S. et al. Dissecting response to cancer immunotherapy by applying Bayesian network analysis to flow cytometry data. *Int. J. Mol. Sci.* **22**, 2316 (2021).
55. Bodén, M., Dellaire, G., Burrage, K. & Bailey, T. L. A Bayesian network model of proteins' association with promyelocytic leukemia (PML) nuclear bodies. *J. Comput. Biol. J. Comput. Mol. Cell Biol.* **17**, 617–630 (2010).
56. Nagarajan, R., Scutari, M. & Lèbre, S. *Bayesian Networks in R: with Applications in Systems Biology* (Springer, New York, 2013).
57. Sui, L. et al. Host cell cycle checkpoint as antiviral target for SARS-CoV-2 revealed by integrative transcriptome and proteome analyses. *Signal Transduct. Target. Ther.* **8**, 21 (2023).
58. Surjit, M., Liu, B., Chow, V. T. K. & Lal, S. K. The nucleocapsid protein of severe acute respiratory syndrome-coronavirus inhibits the activity of cyclin-cyclin-dependent kinase complex and blocks S phase progression in mammalian cells. *J. Biol. Chem.* **281**, 10669–10681 (2006).
59. Yuan, X. et al. SARS coronavirus 7a protein blocks cell cycle progression at G0/G1 phase via the cyclin D3/pRb pathway. *Virology* **346**, 74–85 (2006).
60. Yuan, X., Shan, Y., Zhao, Z., Chen, J. & Cong, Y. G0/G1 arrest and apoptosis induced by SARS-CoV 3b protein in transfected cells. *Virol. J.* **2**, 66 (2005).
61. Bagheri, N., Carpenter, A. E., Lundberg, E., Plant, A. L. & Horwitz, R. The new era of quantitative cell imaging—challenges and opportunities. *Mol. Cell* **82**, 241–247 (2022).
62. Boutros, M., Heigwer, F. & Laufer, C. Microscopy-based high-content screening. *Cell* **163**, 1314–1325 (2015).
63. Kramer, B., Castillo, J., Pelkmans, L. & Gut, G. Iterative indirect immunofluorescence imaging (4i) on adherent cells and tissue sections. *BIO-Protoc.* **13**, e4712 (2023).
64. Valm, A. M. et al. Applying systems-level spectral imaging and analysis to reveal the organelle interactome. *Nature* **546**, 162–167 (2017).
65. Frechin, M. et al. Cell-intrinsic adaptation of lipid composition to local crowding drives social behaviour. *Nature* **523**, 88–91 (2015).
66. Spitzer, H., Berry, S., Donoghoe, M., Pelkmans, L. & Theis, F. J. Learning consistent subcellular landmarks to quantify changes in multiplexed protein maps. *Nat. Methods* **20**, 1058–1069 (2023).
67. Ivanov, I. E. et al. Mantis: high-throughput 4D imaging and analysis of the molecular and physical architecture of cells. *Cell Biol.* <https://doi.org/10.1101/2023.12.19.572435> (2023).
68. Hoffman, D. P., Slavitt, I. & Fitzpatrick, C. A. The promise and peril of deep learning in microscopy. *Nat. Methods* **18**, 130–132 (2021).
69. Karabağ, C., Ortega-Ruiz, M. A. & Reyes-Aldasoro, C. C. Impact of training data, ground truth and shape variability in the deep learning-based semantic segmentation of HeLa cells observed with electron microscopy. *J. Imaging* **9**, 59 (2023).
70. Waters, C. S., Angenent, S. B., Altschuler, S. J. & Wu, L. F. A PINK1 input threshold arises from positive feedback in the PINK1/Parkin mitophagy decision circuit. *Cell Rep.* **42**, 113260 (2023).
71. Kleele, T. et al. Distinct fission signatures predict mitochondrial degradation or biogenesis. *Nature* **593**, 435–439 (2021).
72. Evanko, D. Is phototoxicity compromising experimental results? <https://blogs.nature.com/methagora/2013/11/is-phototoxicity-compromising-experimental-results.html> (2013).
73. Tosheva, K. L., Yuan, Y., Matos Pereira, P., Culley, S. & Henriques, R. Between life and death: strategies to reduce phototoxicity in super-resolution microscopy. *J. Phys. Appl. Phys.* **53**, 163001 (2020).
74. Küppers, M., Albrecht, D., Kashkanova, A. D., Lühr, J. & Sandoghdar, V. Confocal interferometric scattering microscopy reveals 3D nanoscopic structure and dynamics in live cells. *Nat. Commun.* **14**, 1962 (2023).
75. Miserey-Lenkei, S., Trajkovic, K., D'Ambrosio, J. M. & Patel, A. J. A comprehensive library of fluorescent constructs of SARS-CoV-2 proteins and their initial characterisation in different cell types. *Biol. Cell.* **113**, 311–328 (2021).
76. Thorne, L. G. et al. Evolution of enhanced innate immune evasion by SARS-CoV-2. *Nature* **602**, 487–495 (2022).
77. Gordon, D. E. et al. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* **583**, 459–468 (2020).
78. Ricciardi, S. et al. The role of NSP6 in the biogenesis of the SARS-CoV-2 replication organelle. *Nature* **606**, 761–768 (2022).
79. Zetie, K. P., Adams, S. F. & Tocknell, R. M. How does a Mach-Zehnder interferometer work? *Phys. Educ.* **35**, 46–48 (2000).
80. Devaney, A. J. A filtered backpropagation algorithm for diffraction tomography. *Ultrason. Imaging* **4**, 336–350 (1982).
81. Yu, H., Jia, S., Liu, Y. & Dong, J. Phase coherent noise reduction in digital holographic microscopy based on adaptive total variation. *Opt. Lasers Eng.* **134**, 106204 (2020).
82. Wolf, E. Three-dimensional structure determination of semi-transparent objects from holographic data. *Opt. Commun.* **1**, 153–156 (1969).
83. Cotte, Y., Toy, M. F., Pavillon, N. & Depeursinge, C. Microscopy image resolution improvement by deconvolution of complex fields. *Opt. Express* **18**, 19462–19462 (2010).
84. Paris, J. et al. The invariant arginine within the chromatin-binding motif regulates both nucleolar localization and chromatin binding of Foamy virus Gag. *Retrovirology* **15**, 48 (2018).
85. Blond, J.-L. et al. An envelope glycoprotein of the human endogenous retrovirus HERV-W is expressed in the human placenta and fuses cells expressing the type D mammalian retrovirus receptor. *J. Virol.* **74**, 3321–3329 (2000).

86. Grzelak, L. et al. A comparison of four serological assays for detecting anti-SARS-CoV-2 antibodies in human serum samples from different populations. *Sci. Transl. Med.* **12**, eabc3103 (2020).
87. Planas, D. et al. Considerable escape of SARS-CoV-2 Omicron to antibody neutralization. *Nature* **602**, 671–675 (2022).
88. Planas, D. et al. Sensitivity of infectious SARS-CoV-2 B.1.1.7 and B.1.351 variants to neutralizing antibodies. *Nat. Med.* **27**, 917–924 (2021).

Acknowledgements

We thank Timothée Bruel, Nicoletta Casartelli, and Sasha Legrosdidier for critical reading of the manuscript, members of the Virus and Immunity Unit for discussions and help, Perrine Bomme and the Ultra-structural Bioimaging Unit (UBI) and UtechS Photonic Bioimaging (UPBI) core facilities (Institut Pasteur). We thank Kyle Van Der Langemheen for establishing Optuna. N.S. is funded by the ministère de l'Enseignement supérieur et de la Recherche. O.S. laboratory is funded by Institut Pasteur, Fondation pour la Recherche Médicale (FRM), ANRS-MIE, the Vaccine Research Institute (ANR-10-LABX-77), HERA European program (DURABLE consortium), Labex IBEID (ANR-10-LABX-62-IBEID), ANR/FRM Flash Covid PROTEO-SARS-CoV-2 and IDISCOVER. We are grateful for support for equipment from the French Government Program Investissements d'Avenir France Bioimaging (FBI, No. ANR-10-INSB-04-01) and the French gouvernement (Agence Nationale de la Recherche) Investissement d'Avenir program, Laboratoire d'Excellence "Integrative Biology of Emerging Infectious Diseases" (ANR-10-LABX-62-IBEID). We thank Thibault Corteoux and Lisa Polaro for connecting Nanolive's Deep Quantitative Biology team with the Institut Pasteur. We thank all Nanolive's employees for their relentless appreciation, encouragement, and support.

Author contributions

Conceptualization, N.S., B.M., T.W., O.S., and M.F.; Methodology, N.S., B.M., N.C., L.A., H.M., A.J., J.B., A.M., O.S., and M.F.; Data curation, L.A., H.M., A.J., A.M., and M.F.; Software, L.A., H.M., A.J., A.M., and M.F.; Investigation, N.S., N.C., B.M., and M.F.; Visualization, N.S. and M.F.; Writing—original draft, M.F.; Writing—review & editing, N.S., T.W., O.S., and M.F.; Funding acquisition, O.S.; Resources, O.S. and M.F.; Supervision, T.W., O.S., and M.F.

Competing interests

L.A., H.M., A.J., A.M., and M.F. are employees of Nanolive SA. The remaining authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-49260-7>.

Correspondence and requests for materials should be addressed to Olivier Schwartz or Mathieu Fréchin.

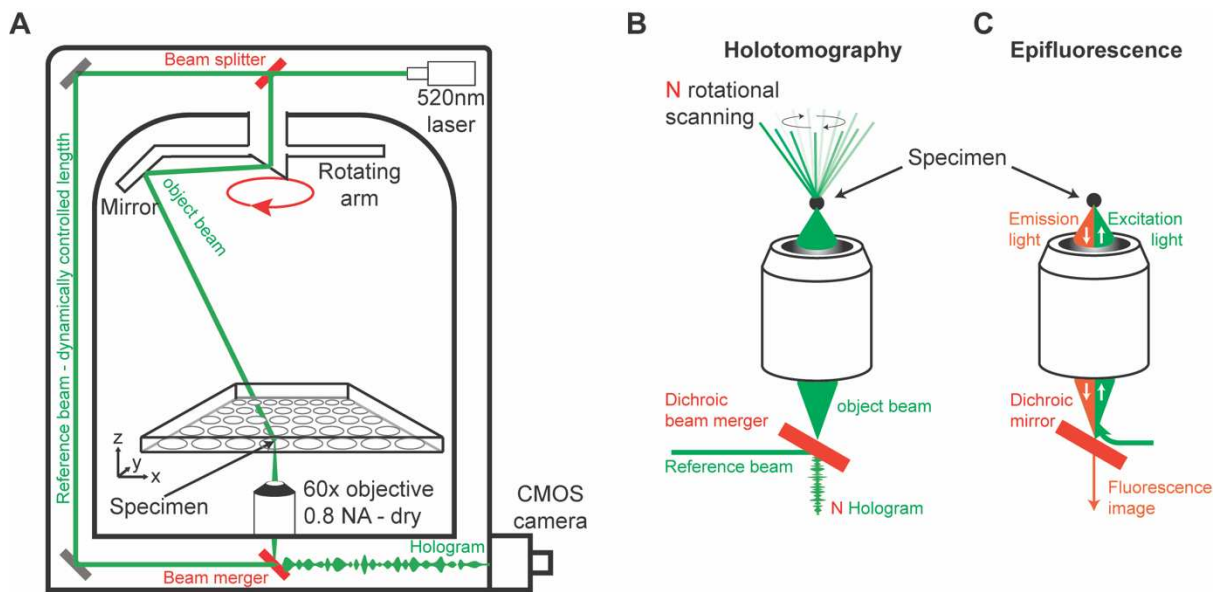
Peer review information *Nature Communications* thanks Rudranil De, Shalin Mehta, who co-reviewed with Soorya Pradeep, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

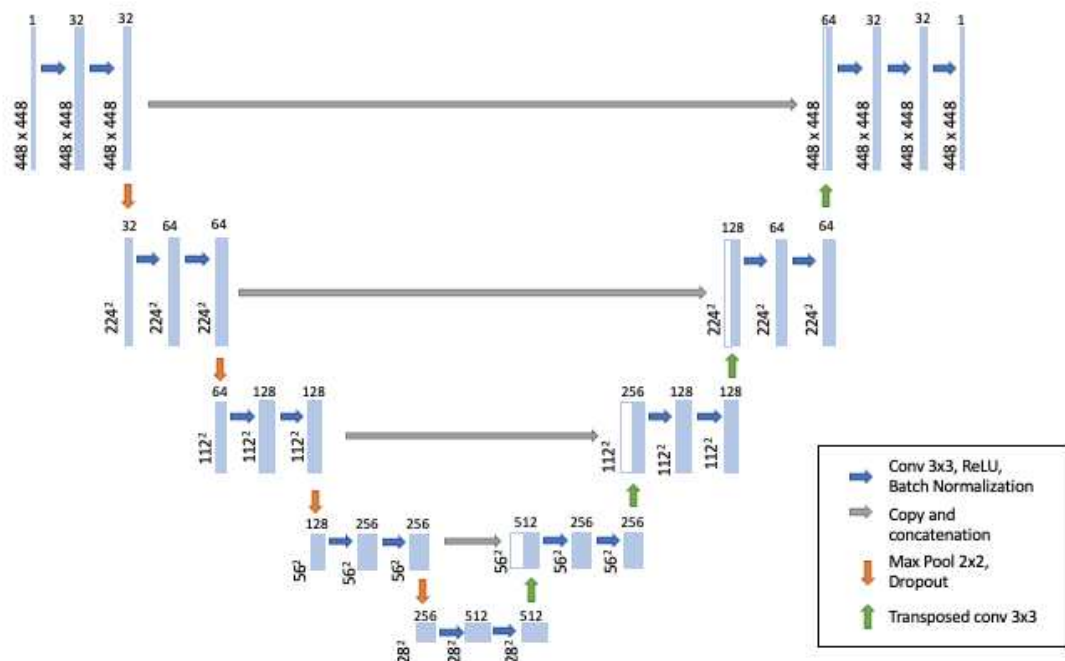
Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024

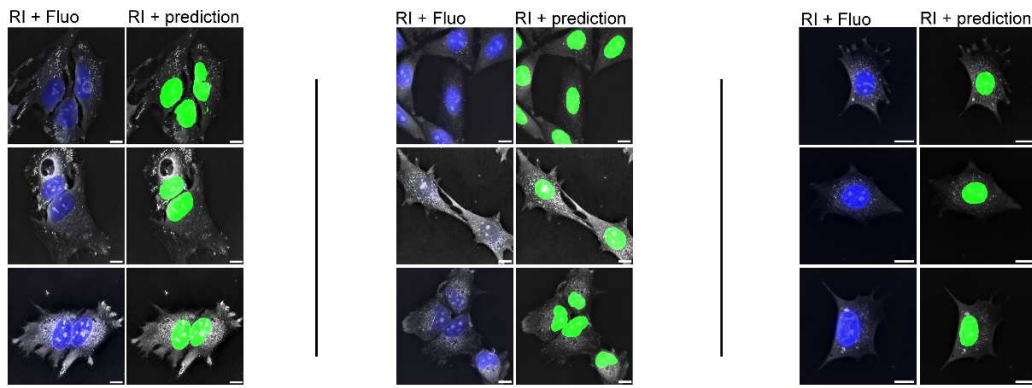


Supplementary figure 1 | Holotomographic microscope device schematics. (A) Overview of the microscope inner setting summarizing the holographic optical setup, the rotational scanning system, the adaptive optical path of the reference beam and the automated stage. (B) The rotation all over the specimen allows for the light to interact with the sample from different angles. The diffused light is collected by a 60x objective and brought to interference with the reference beam on a 50/50 dichroic beam merger, which creates as many holograms as there is illumination of the sample during the rotational scanning. (C). The device is equipped with an epifluorescence system.

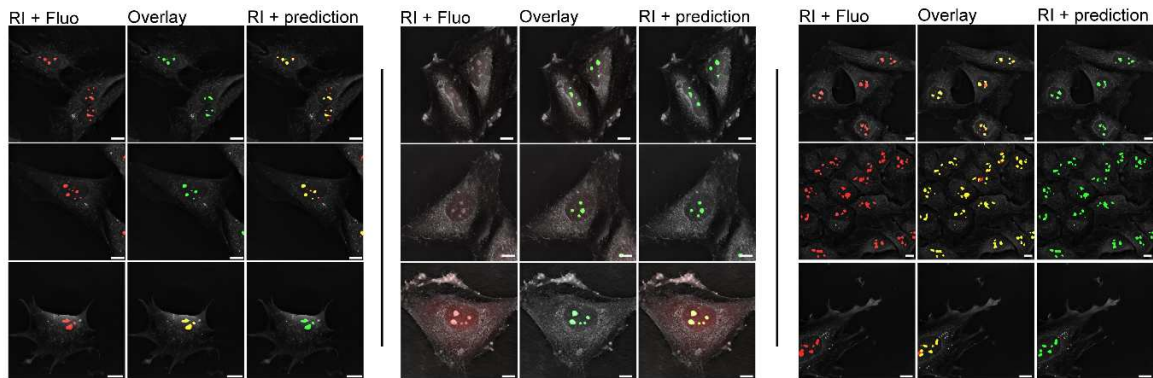


Supplementary figure 2 | Adapted UNET architecture used for cell, nuclei and nucleoli detection. Our deep learning segmentation model uses the well-established UNET architecture³⁴. We modified it further by adding batch normalization as well as dropout steps that were key to our organelle detection performances. Batch normalization³⁶ reduces internal covariate shift via the fixing of the means and variances of layer inputs. It improves gradient flow and removes its dependence to parameters scale or initial values. The expected impact of batch normalization is quicker learning and less divergence. The dropout³⁷ step is also a regularization technique where each neuron of the network has a probability to be switched off (or dropped) at each training cycle. It aims at avoiding that few connections overtake the network's behavior, a process known as co-adaptation leading to overfitting.

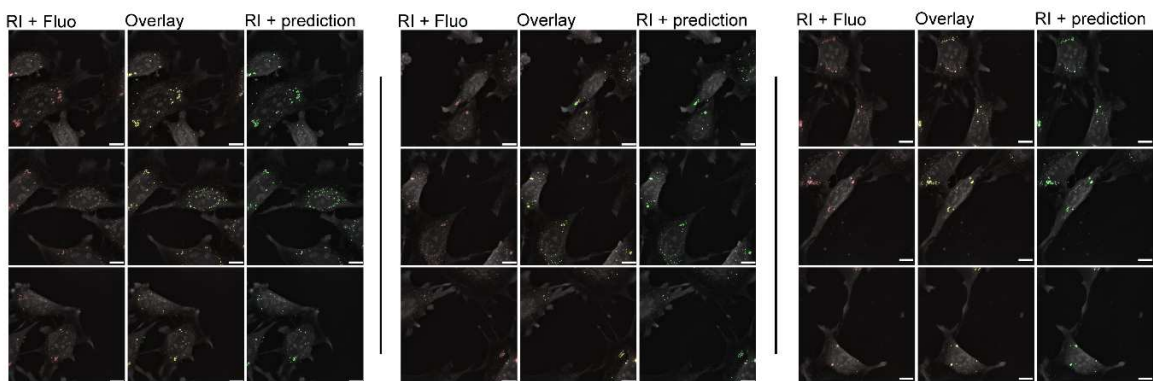
A Nuclei fluorescence signal Vs prediction mask



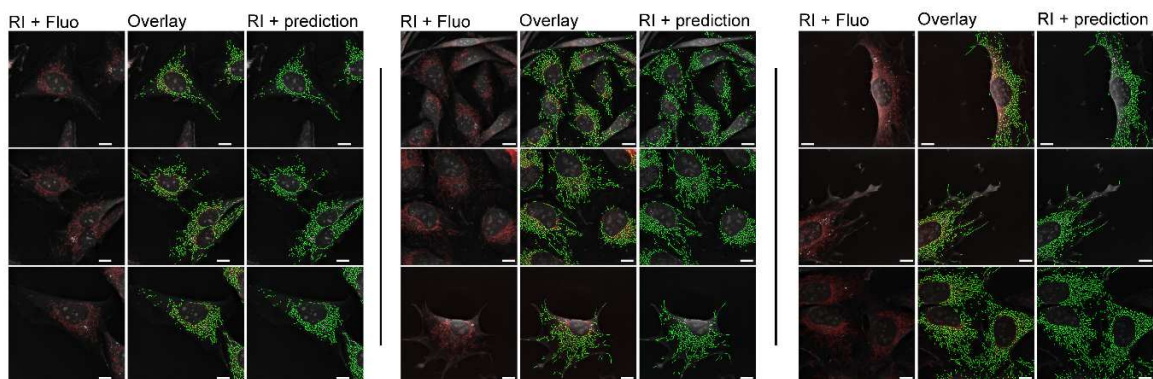
B Nucleoli fluorescence signal Vs prediction mask



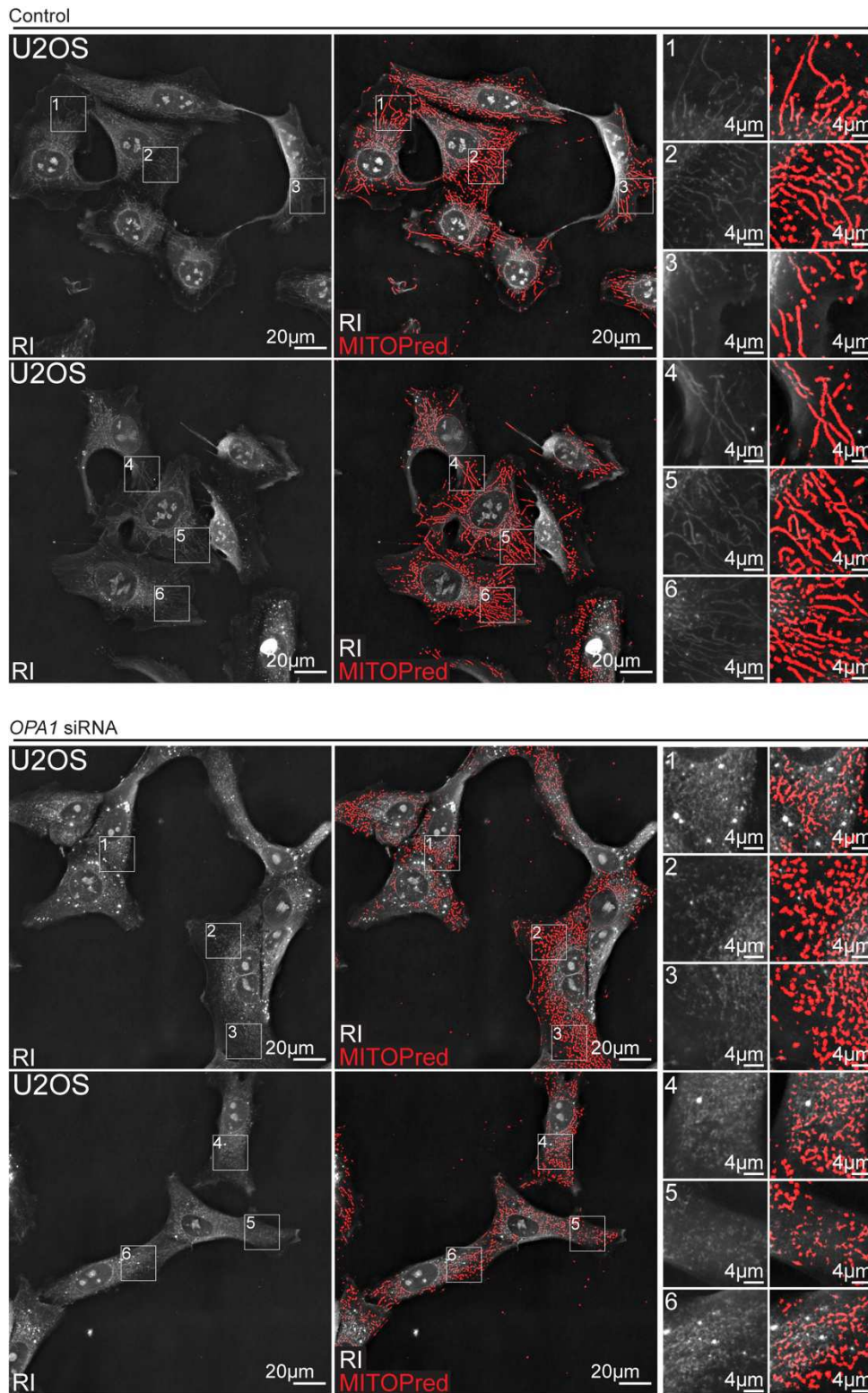
C Lipid droplet fluorescence signal Vs prediction mask



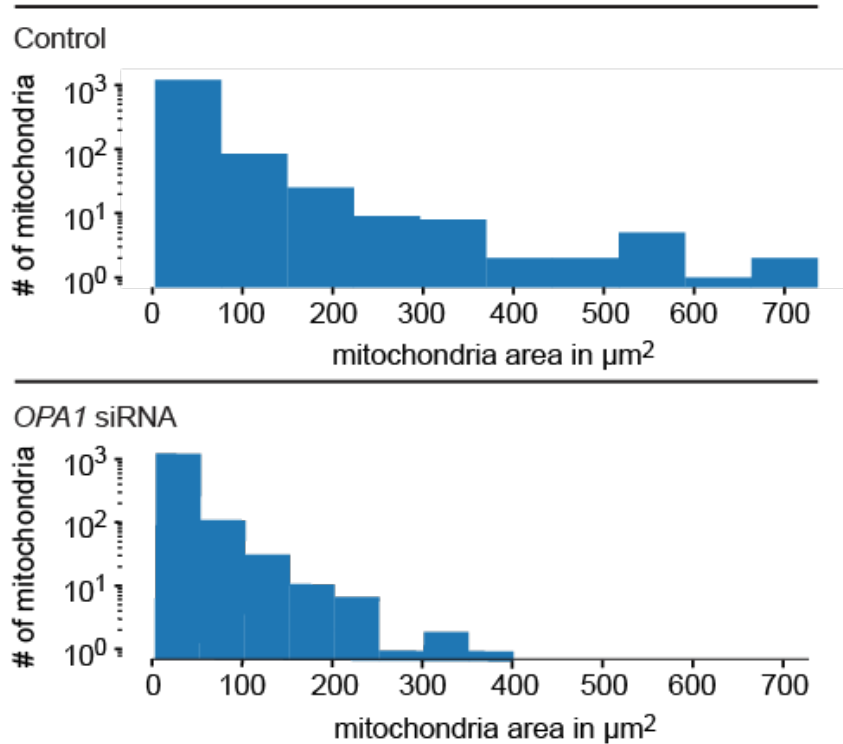
D Mitochondria fluorescence signal Vs prediction mask



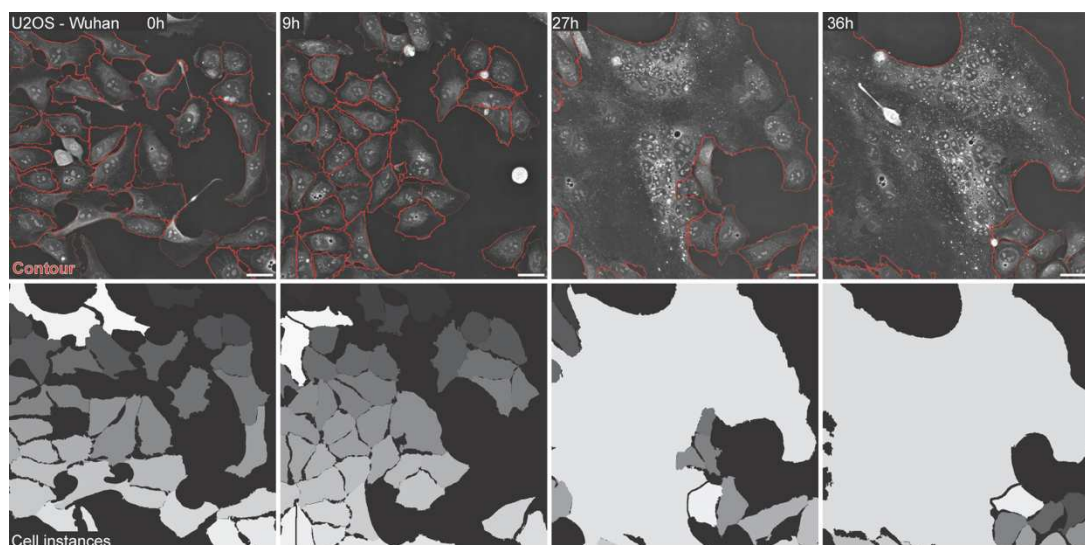
Supplementary figure 3 | Comparison of organelle fluorescent signal and machine learning-based organelle prediction. RI-refractive index, Fluo-Fluorescence signal (A) Various forms of Nuclei, (B) nucleoli, (C) lipid droplets and (D) mitochondria, are detected efficiently in cells of various forms and confluency. For metrics quantifying the quality of our predictions see figure 2G. Scale bar, 10 μm .



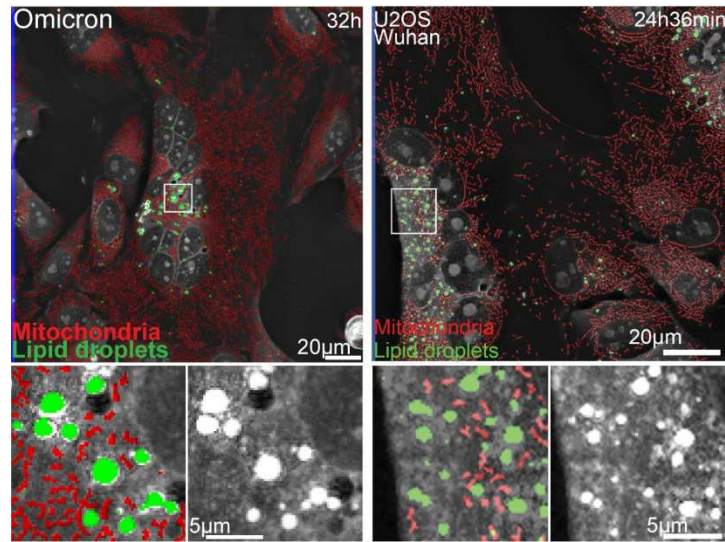
Supplementary figure 4 | Machine learning detects unperturbed or fragmented mitochondria after silencing of *OPA1*. Control U2OS cells display long, tubular, connected mitochondria, while U2OS cells experiencing *OPA1* silencing have punctuated, over-fragmented mitochondria. Zooms show that mitochondria are detected across a wide range of various morphologies. MITOPred-mitochondria predictions.



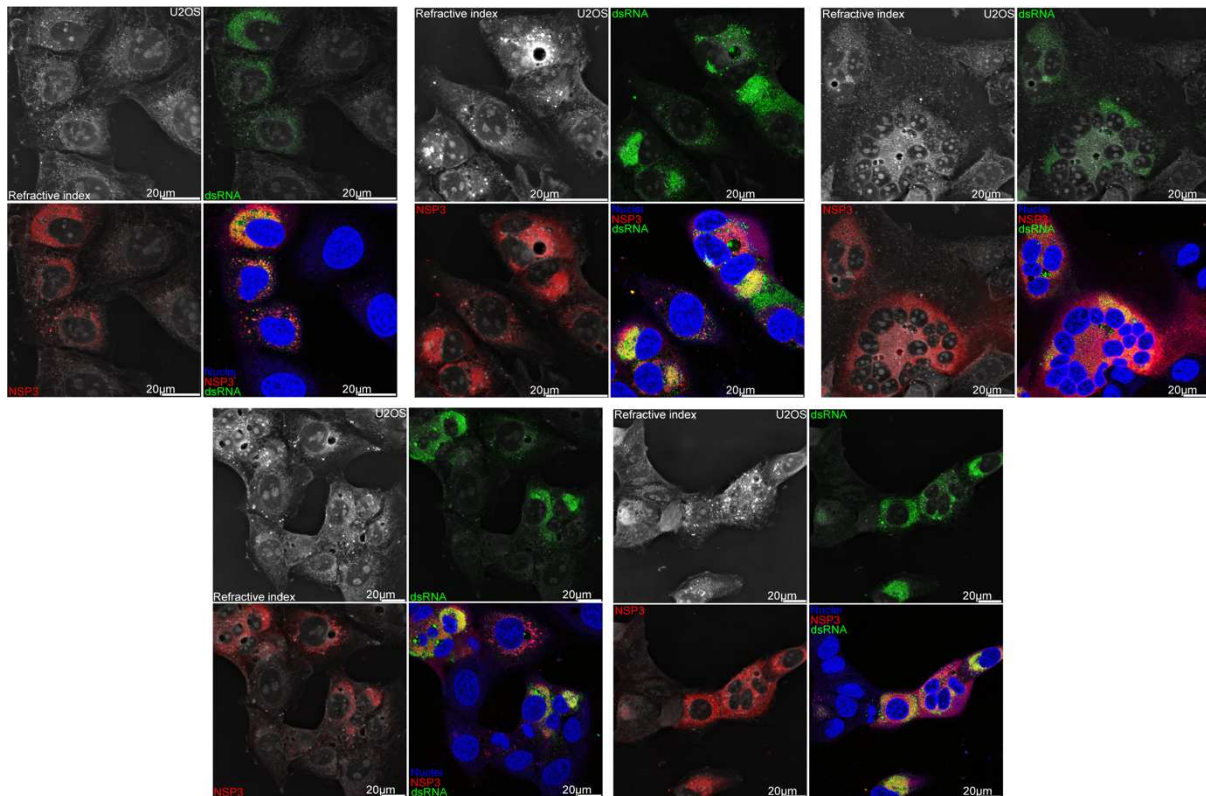
Supplementary figure 5 | The effect of *OPA1* silencing is quantified through automated mitochondria segmentation.



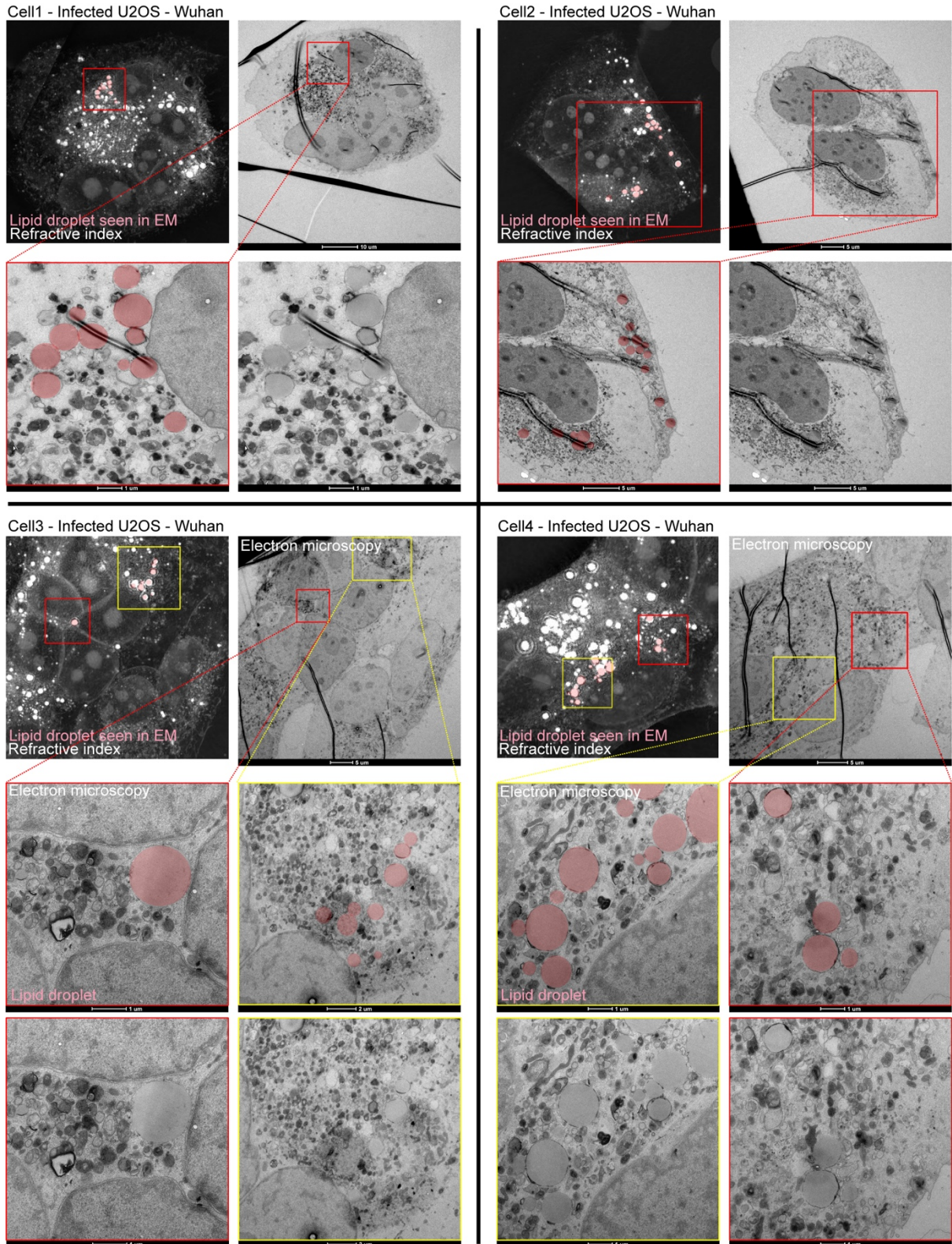
Supplementary figure 6 | Deep learning-based label-free cell segmentation detects single cells and syncytium with precision. At 0 and 9 hours, we can see that cells are all efficiently detected. At 27 and 36 hours, one can observe that the large syncytium is well segmented as well as the co-existing single cells that did not fuse with the syncytium, validating our proper cell vs syncytium segmentation robustness. All our cell vs syncytium proper segmentations can be verified in our Supplementary Movies 1-19.



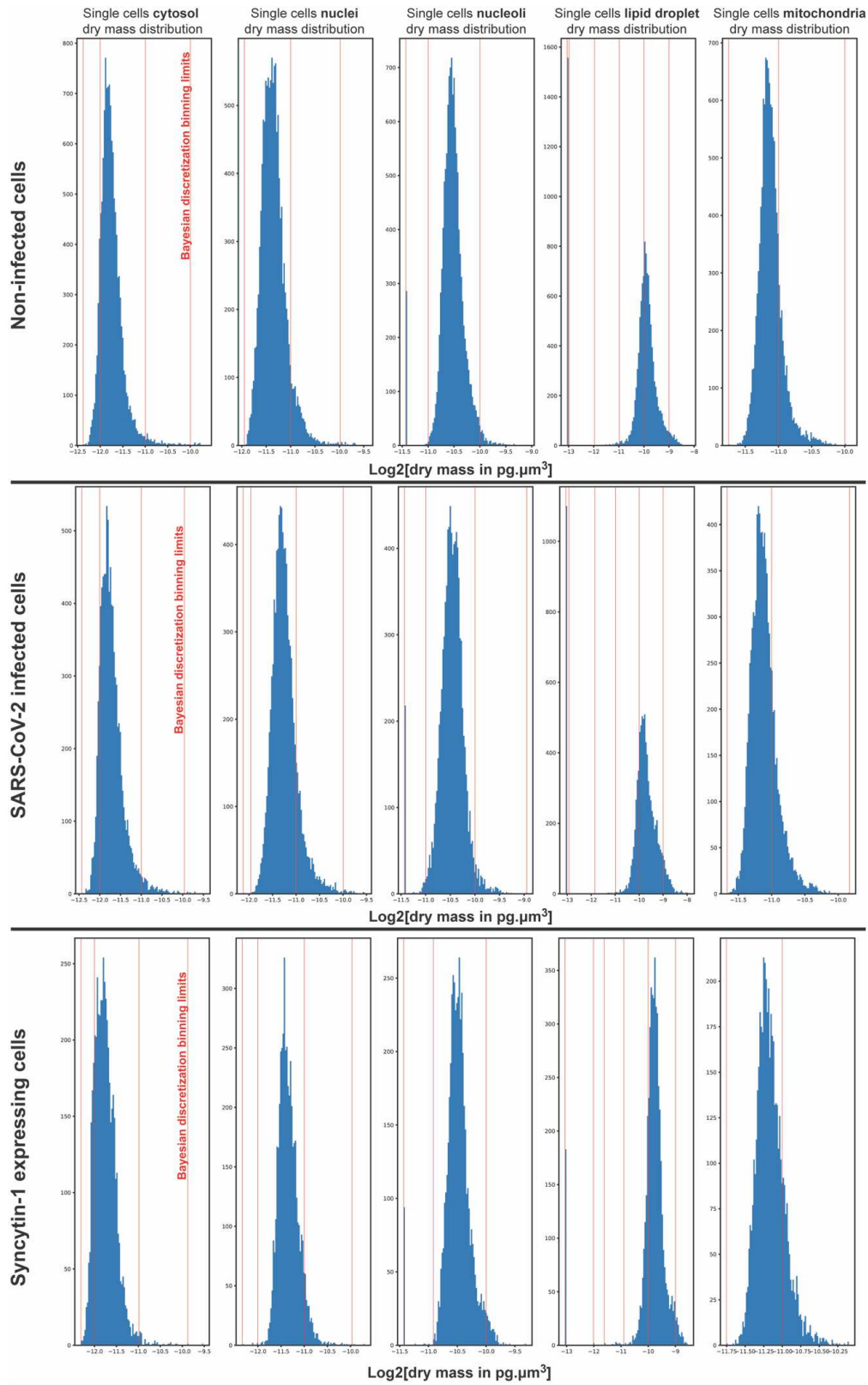
Supplementary figure 7 | Detection of SARS-CoV-2-induced lipid droplets in the perinuclear region of infectious syncytium.



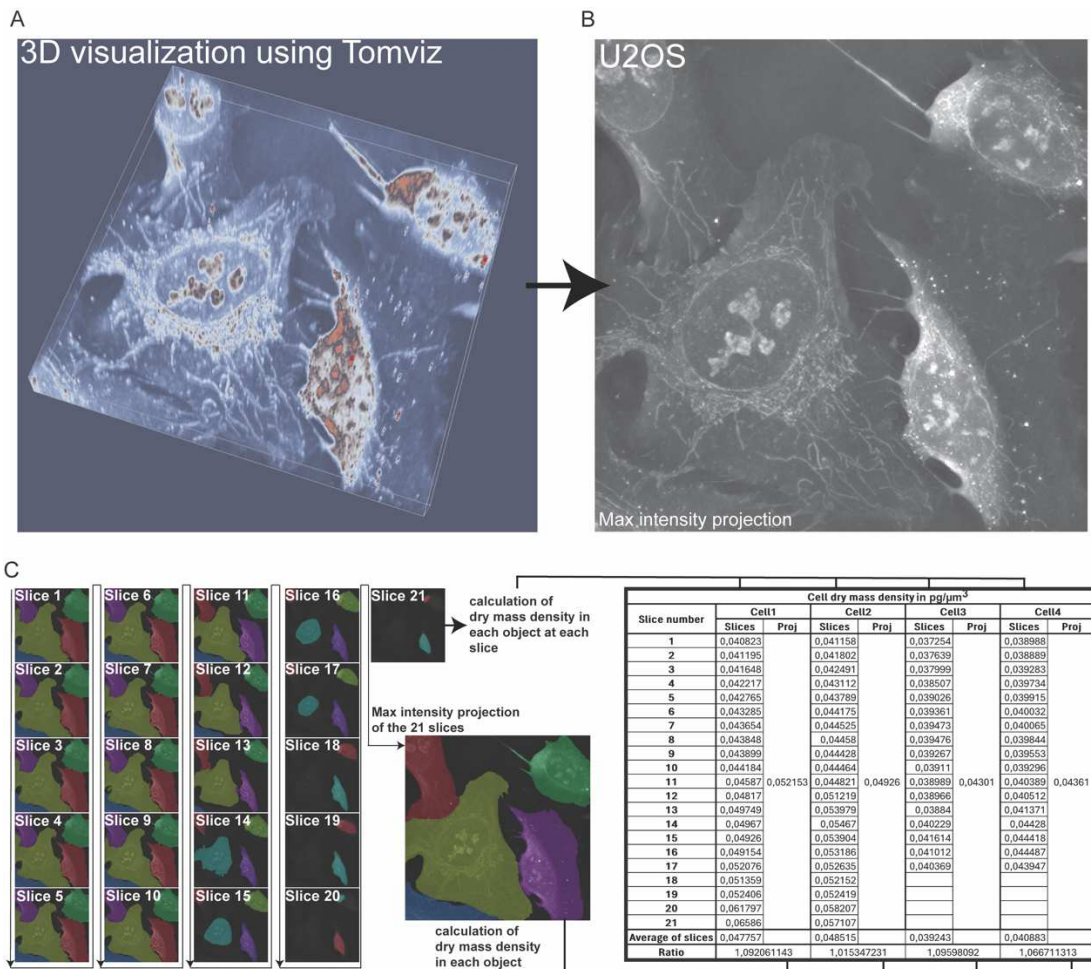
Supplementary figure 8 | Perinuclear lipid droplet accumulation is a marker of infection. Each of the five panels show that cells with large lipid droplets all show NSP3 and dsRNA immuno-fluorescent signal, while those with no lipid droplets do not display NSP3 and dsRNA immuno-fluorescent signal. See also figure 5.



Supplementary figure 9 | Lipid droplets of infected cells are not surrounded by mitochondria. The four U2OS cells (cell 1, cell 2, cell 3, cell 4) and syncytium infected with the SARS-CoV-2 Wuhan strain and developing large lipid droplets do not show any mitochondria close to lipid droplets when observed by electron microscopy. See also figure 5E and 5F.



Supplementary figure 10 | Distribution of single cells according to their respective dry mass located within the cytosol, nuclei, nucleoli, lipid droplets and mitochondria. Each line of five histograms represents single cells distributions (in blue) of a specific experimental condition (non-infected, infected, and syncytin-1-expressing cells) as well as their respective Bayesian binning represented as overlaying red lines.



Supplementary figure 11 | Maximum intensity projection of holotomographic 3D volumes recapitulate volume-derived dry mass densities. (A) Holotomographic microscopy (HTM) returns 3D refractive index maps (B) are transformed into maximal intensity projection images. (C) The dry mass densities calculated from manual 3D cell segmentations, or automated 2D cell segmentations provide similar results, confirming that 2D maximal intensity projections recapitulate well the whole cell content.

Edge	Covid		Syncytin 1	
	score	pVal	score	pVal
Cy>LD	260(15)	0	4(15)	0.9
Sy>LD	200(5)	0	2(5)	1
Cy>Sy	1(3)	0.9	33(3)	0

Supplementary table 1 | Pairwise statistical tests strengthen key organelle cross regulation (OCR) rewiring differences between infected cells and cells expressing Syncytin 1. Two-sided Chi-squared score and p-values reinforce the observation that cytosol and syncytium edges towards LD exists when cells are infected by SARS-CoV-2 but not when cells only express Syncytin 1.

Edge	freq.	score	DoF	pVal
Nx>Ni	0.84	2503	≤6	<10 ⁻⁶
Nx>Cy	0.80	6700	≤6	<10 ⁻⁶
Nx>LD	0.50	377	≤12	<10 ⁻⁶
Nx>Mi	0.64	1326	≤6	<10 ⁻⁶
Cy>Mi	0.32	481	≤9	<10 ⁻⁶
Mi>Nx	0.36	747	≤6	<10 ⁻⁶
Mi>LD	0.34	209	≤18	<10 ⁻⁶
Mi>Cy	0.68	1041	≤9	<10 ⁻⁶

Supplementary table 2 | Key metrics of unperturbed cell organelle cross regulation (OCR) network. Each edge of the OCR is described by its frequency over bootstrapping iterations, its mean two-sided chi-squared score, its maximum observed degree of freedom and p-Value.

Edge	freq.	score	DoF	pVal
Nx>Ni	0.81	1843	≤12	<10 ⁻⁶
Nx>Cy	0.76	2786	≤9	<10 ⁻⁶
Nx>Mi	0.60	1194	≤6	<10 ⁻⁶
Cy>LD	0.67	260	≤15	<10 ⁻⁶
Mi>Nx	0.40	764	≤6	<10 ⁻⁶
Mi>Cy	0.68	1318	≤6	<10 ⁻⁶
LD>Sy	0.75	644	≤5	<10 ⁻⁶
Sy>LD	0.25	200	≤5	<10 ⁻⁶

Supplementary table 3 | Key metrics of SARS-CoV-2-infected cell organelle cross regulation (OCR) network. Each edge of the OCR is described by its frequency over bootstrapping iterations, its mean two-sided chi-squared score, its maximum observed degree of freedom and p-Value.

Edge	freq.	score	DoF	pVal
Nx>Ni	0.91	456	≤9	<10 ⁻⁶
Nx>Mi	0.33	571	≤6	<10 ⁻⁶
Mi>Nx	0.67	1308	≤6	<10 ⁻⁶
Mi>Cy	0.67	1647	≤6	<10 ⁻⁶
Cy>Mi	0.33	812	≤6	<10 ⁻⁶
Cy>Sy	0.70	33	≤3	<10 ⁻⁶

Supplementary table 4 | Key metrics of Syncytin-1-expressing cell organelle cross regulation (OCR) network. Each edge of the OCR is described by its frequency over bootstrapping iterations, its mean two-sided chi-squared score, its maximum observed degree of freedom and p-Value.

4.3. Discussion

4.3.1. Strengths and limits of our approach

One of the main strengths of our study is the accuracy of the predictions. When comparing with the state-of-the-art method in HTM, one can see that our methods yield much higher resolution⁴³⁹. Furthermore, our approach can be used to monitor organelles over long periods of time which has not been done before using HTM.

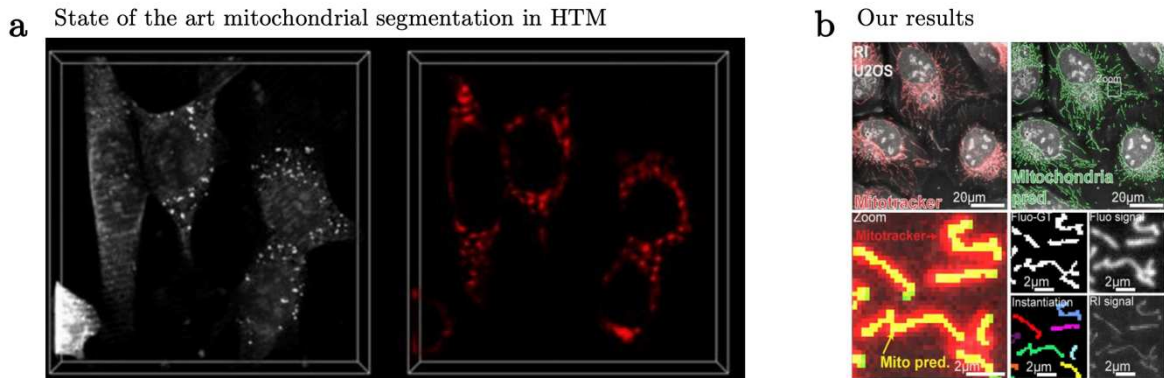


Figure 23: Comparison of (a) state of the art HTM segmentation of mitochondria⁴³⁸ and (b) our approach.

During our analysis, we do not exploit the 3D stack obtained by the microscope. The microscope generates 90 slices per image, with an axial resolution of 200nm. The whole cell can thus theoretically be analyzed, after inferring the region in between two images. However, this greatly complexifies the analysis, and the calculation times. For this first work, we decided it was not necessary, as we preferred focusing on a higher time resolution.

We did not segment lysosomes, although they possess a refractive index that is different from the cytosol's. Furthermore, not all organelles are visible, and we can thus miss critical information of the ER and the Golgi apparatus. This is limiting to perform studies on organelle cross-regulation.

4.3.2. The biology behind the observed phenomenon

Our study is a proof of concept of the use of AI for organelle segmentation over long periods of time in HTM images and is purely descriptive. This approach enables a global, non-biased observation of infected cells. Experiments using single viral proteins or modified viruses are needed to unravel the biological mechanisms of the observed phenomenon. Additional studies with other cell lines should also be performed before inferring any phenomenon we observe is characteristic of SARS-CoV-2 infection.

We cannot compare variants effects, as the phenotype observed likely depends on the infection levels, and those are affected by syncytia formation and viral replication which both depend on the variant as shown in **Chapter 2**.

Some experiments were performed with a strain lacking the furin cleavage site in the spike to assess the effect of the infection independently of syncytia formation. However, questions then arose on how to label infected cells for analysis, as we used syncytia as a marker of infection. In the end, those data were not exploited. Viruses modified to express GFP could be used to this end, but we did not have a virus lacking the cleavage site and expressing GFP.

Data published by other groups corroborate our observations on LDs accumulation^{440,441}. One study showed that LDs accumulation in infected cells was due to ORF3a expression⁴⁴². Interestingly, a mutation in ORF3a occurred in the beta variant, that completely abolished LD accumulation, while another one occurred in BA.2 that diminishes LD accumulation⁴⁴². This could be confirmed with HTM by transfecting cells with ORF3a (WT or mutant). Furthermore, nsp6 plays a role in tethering LDs to the replication organelle and allowing flow from lipid droplets to DMVs³⁷², which confirms the proximity we see between LDs and viral factories by confocal microscopy.

For mitochondria, no clear phenotype was observed. Studies report “an increase of both intracrystal space and matrix density”⁴⁴³, “enlarged cristae and increased absorption contrast, suggesting matrix condensation”⁴⁴⁴, shorter mitochondria⁴⁴⁵, mitochondria fragmentation⁴⁴⁶, mitochondria fusion⁴⁴⁷; mitochondrial dynamics seem very dependent on the model studied and likely depend of the stage of the infection.

We show an innovative approach to image cells in their native environment and quantify organelle remodeling over long periods of time. This approach can enable monitoring of organelles localization, shape, dry-mass and interaction. It can thus yield interesting insights on infection, drug’s mechanisms, the effect of mutations... AI guided label-free image analysis has made significant progress in the last few years; it is just a question of time before user-friendly robust analysis software are made widely accessible.

Conclusion



This thesis explores different aspects of coronavirus biology, among which entry, replication, cell-cell fusion and neutralization by the adaptive immune system and monoclonal antibodies.

We explore the effects of mutation in SARS-CoV-2 S on entry, fusogenicity, neutralization by the serum of previously vaccinated or infected individuals or by monoclonal antibodies. Our pipeline to characterize variant spike was also used for many subsequent variants. The effect of mutations in the rest of the genome is also discussed. As new variants constantly emerge, it is important to study their resistance to existing immune response, therapeutic antibodies and anti-viral drugs, as well as their pathogenicity and transmissibility. For instance, mutations enabling paxlovid resistance have been described in immunocompromised individuals, however it has been shown they decrease the fitness of the virus, and they are thus not expected to spread. Molecular mechanisms of the infection are now better understood, and several treatments are available, yet SARS-CoV-2 still kills elderly and immunocompromised patients and causes long-COVID in previously healthy individuals. Further efforts are necessary to better understand the virus and the disease and develop new drugs or therapeutic approaches.

We then focus on HKU1, and on the role of TMPRSS2 as a receptor for this virus. Thanks to the identification of the receptor of HKU1, we are one step closer to developing a cell-line system to propagate the virus; this will greatly accelerate the research on HKU1 as one could then study the susceptibility of different animal models to the virus, the neutralization of the virus, and the effect of drugs and cellular protein on replication. Furthermore, developing replicon system for HKU1 as well as other seasonal coronaviruses will enable investigating the role of different proteins. Nanobodies that bind TMPRSS2 are a useful tool to study TMPRSS2 in a laboratory setting. Their effect on coronavirus but also influenza or paramyxovirus entry should be investigated as some of them block TMPRSS2 activity. Increased animal sampling will hopefully bring additional insight on HKU1 origin. Much work remains to be done to fully comprehend this understudied virus.

Finally, we present a new method to monitor organelle remodeling using HTM. Such an approach can be used to observe cellular phenomenon in an unbiased manner. For instance, we observe lipid droplet accumulation upon infection. This technique can be adapted to study a wide range of virus and cellular processes with high throughput. While this approach currently has limits, such as the use of 2D datasets while 3D images are acquired, and the invisibility of several subcellular compartments such as the ER, increased computing power, development in AI and technical improvements might fill these gaps in the next years. Our work paves the way for the development of new HTM analysis algorithms.

While each human coronavirus has its specific set of accessory proteins and unique receptor usage, they have many shared features on a pathological – common cold, severity in the elderly, short span immune response ... – and a molecular point of view – opened and closed spike conformation, use of proteases as receptors, cleavage by cellular proteases, sensitivity of ISGs and induction of an immune response, similar NSPs ... Much of the knowledge gained on SARS-CoV-2 also applies to other coronaviruses, and we have seen drastic improvement in our comprehension of seasonal coronaviruses in the last years. Moreover, the techniques developed to study SARS-CoV-2 can be adapted to study other coronaviruses, whether emerging or seasonal, as we did for HKU1. Technical progress in HTM imaging and analysis could also be valuable to study emerging viruses. The knowledge can go both ways; elucidating seasonal coronavirus adaptation to humans over decades might help us understand changes that occur in SARS-CoV-2, as it evolves in the human reservoir.

A global understanding of all coronaviruses, and the development of broadly active therapeutic strategies and vaccines against coronaviruses is paramount in case a new coronavirus emerges.

References



1. Kendall, E. J., Bynoe, M. L. & Tyrrell, D. A. Virus isolations from common colds occurring in a residential school. *Br Med J* **2**, 82–6 (1962).
2. Tyrrell, D. A. & Bynoe, M. L. Cultivation of a novel type of common-cold virus in organ cultures. *Br Med J.* (1965).
3. Tyrrell, D. A. & Bynoe, M. L. Cultivation of viruses from a high proportion of patients with colds. *Lancet* **1**, 76–7 (1966).
4. McIntosh, K., Dees, J. H., Becker, W. B., Kapikian, A. Z. & Chanock, R. M. Recovery in tracheal organ cultures of novel viruses from patients with respiratory disease. *Proc. Natl. Acad. Sci. U.S.A.* **57**, 933–40 (1967).
5. Kahn, J. S. & McIntosh, K. History and recent advances in coronavirus discovery. *Pediatr Infect Dis J* **24**, S223–7, discussion S226 (2005).
6. Woo, P. C. *et al.* Discovery of seven novel Mammalian and avian coronaviruses in the genus deltacoronavirus supports bat coronaviruses as the gene source of alphacoronavirus and betacoronavirus and avian coronaviruses as the gene source of gammacoronavirus and deltacoronavirus. *J Virol* **86**, 3995–4008 (2012).
7. Tabibzadeh, A. *et al.* Evolutionary study of COVID-19, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) as an emerging coronavirus: Phylogenetic analysis and literature review. *Veterinary Medicine and Science* **7**, 559–571 (2021).
8. Hayman, D. T. S. & Knox, M. A. Estimating the age of the subfamily Orthocoronavirinae using host divergence times as calibration ages at two internal nodes. *Virology* **563**, 20–27 (2021).
9. Wertheim, J. O., Chu, D. K., Peiris, J. S., Kosakovsky Pond, S. L. & Poon, L. L. A case for the ancient origin of coronaviruses. *J Virol* **87**, 7039–45 (2013).
10. Iruegas, R. *et al.* The evolutionary making of SARS-CoV-2. 2021.01.29.428808 Preprint at <https://doi.org/10.1101/2021.01.29.428808> (2021).
11. Drosten, C. *et al.* Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *N Engl J Med* **348**, 1967–76 (2003).
12. Peiris, J. S. *et al.* Coronavirus as a possible cause of severe acute respiratory syndrome. *Lancet* **361**, 1319–25 (2003).
13. van der Hoek, L. *et al.* Identification of a new human coronavirus. *Nat Med* **10**, 368–73 (2004).
14. Woo, P. C. Y. *et al.* Characterization and Complete Genome Sequence of a Novel Coronavirus, Coronavirus HKU1, from Patients with Pneumonia. *Journal of Virology* **79**, 884–895 (2005).
15. NL63 isolate KF530110. <https://www.ncbi.nlm.nih.gov/nuccore/KF530110>.
16. Zaki, A. M., van Boheemen, S., Bestebroer, T. M., Osterhaus, A. D. & Fouchier, R. A. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N Engl J Med* **367**, 1814–20 (2012).
17. Zhu, N. *et al.* A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med* **382**, 727–733 (2020).
18. Zhou, P. *et al.* A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270–273 (2020).
19. van der Hoek, L. Human coronaviruses: what do they cause? *Antivir Ther* **12**, 651–8 (2007).
20. International Committee on Taxonomy of Viruses (ICTV). <https://ictv.global/taxonomy/>.
21. Huynh, J. *et al.* Evidence supporting a zoonotic origin of human coronavirus strain NL63. *J Virol* **86**, 12816–25 (2012).
22. Pfefferle, S. *et al.* Distant relatives of severe acute respiratory syndrome coronavirus and close relatives of human coronavirus 229E in bats, Ghana. *Emerg Infect Dis* **15**, 1377–84 (2009).
23. Corman, V. M. *et al.* Link of a ubiquitous human coronavirus to dromedary camels. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 9864–9 (2016).
24. Lau, S. K. *et al.* Discovery of a novel coronavirus, China Rattus coronavirus HKU24, from Norway rats supports the murine origin of Betacoronavirus 1 and has implications for the ancestor of Betacoronavirus lineage A. *J Virol* **89**, 3076–92 (2015).
25. Vijgen, L. *et al.* Complete genomic sequence of human coronavirus OC43: molecular clock analysis suggests a relatively recent zoonotic coronavirus transmission event. *J Virol* **79**, 1595–604 (2005).
26. Corman, V. M., Muth, D., Niemeyer, D. & Drosten, C. Hosts and Sources of Endemic Human Coronaviruses. *Adv Virus Res* **100**, 163–188 (2018).
27. Shore, J. The Endemic Coronaviruses and What They Might Tell us About COVID-19. *The*

- Native Antigen Company* <https://thenativeantigencompany.com/the-endemic-coronaviruses-and-what-they-might-tell-us-about-covid-19/> (2020).
28. Morfopoulou, S. *et al.* Human Coronavirus OC43 Associated with Fatal Encephalitis. *N Engl J Med* **375**, 497–498 (2016).
 29. Park, S., Lee, Y., Michelow, I. C. & Choe, Y. J. Global Seasonality of Human Coronaviruses: A Systematic Review. *Open Forum Infectious Diseases* **7**, (2020).
 30. Rucinski, S. L., Binnicker, M. J., Thomas, A. S. & Patel, R. Seasonality of Coronavirus 229E, HKU1, NL63, and OC43 From 2014 to 2020. *Mayo Clin Proc* **95**, 1701–1703 (2020).
 31. Nichols, G. L. *et al.* Coronavirus seasonality, respiratory infections and weather. *BMC Infectious Diseases* **21**, 1101 (2021).
 32. Shah, M. M. *et al.* Seasonality of Common Human Coronaviruses, United States, 2014–20211. *Emerg Infect Dis* **28**, 1970–1976 (2022).
 33. Moriyama, M., Hugentobler, W. J. & Iwasaki, A. Seasonality of Respiratory Viral Infections. *Annu Rev Virol* **7**, 83–101 (2020).
 34. De Thoisy, A. *et al.* Seroepidemiology of the Seasonal Human Coronaviruses NL63, 229E, OC43 and HKU1 in France. *Open Forum Infect Dis* **10**, ofad340 (2023).
 35. Zhong, N. S. *et al.* Epidemiology and cause of severe acute respiratory syndrome (SARS) in Guangdong, People’s Republic of China, in February, 2003. *Lancet* **362**, 1353–8 (2003).
 36. Li, W. *et al.* Bats are natural reservoirs of SARS-like coronaviruses. *Science* **310**, 676–9 (2005).
 37. CDC. Prevalence of IgG Antibody to SARS-Associated Coronavirus in Animal Traders - Guangdong Province, China, 2003.
 38. Guan, Y. *et al.* Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China. *Science* **302**, 276–8 (2003).
 39. WHO. Summary of probable SARS cases with onset of illness from 1 November 2002 to 31 July 2003. <https://www.who.int/publications/m/item/summary-of-probable-sars-cases-with-onset-of-illness-from-1-november-2002-to-31-july-2003>.
 40. CDC. SARS Update—May 19, 2004. <https://www.cdc.gov/sars/media/2004-05-19.html>.
 41. Donnelly, C. A. *et al.* Epidemiological determinants of spread of causal agent of severe acute respiratory syndrome in Hong Kong. *Lancet* **361**, 1761–1766 (2003).
 42. Stadler, K. *et al.* SARS--beginning to understand a new virus. *Nat Rev Microbiol* **1**, 209–218 (2003).
 43. Li, A. & Ng, P. Severe acute respiratory syndrome (SARS) in neonates and children. *Arch Dis Child Fetal Neonatal Ed* **90**, F461–F465 (2005).
 44. Lai, S. T. Treatment of severe acute respiratory syndrome. *Eur J Clin Microbiol Infect Dis* **24**, 583–591 (2005).
 45. Stockman, L. J., Bellamy, R. & Garner, P. SARS: Systematic Review of Treatment Effects. *PLoS Med* **3**, e343 (2006).
 46. Bleibtreu, A., Bertine, M., Bertin, C., Houhou-Fidouh, N. & Visseaux, B. Focus on Middle East respiratory syndrome coronavirus (MERS-CoV). *Med Mal Infect* **50**, 243–251 (2020).
 47. WHO. MERS-CoV. [https://www.who.int/news-room/fact-sheets/detail/middle-east-respiratory-syndrome-coronavirus-\(mers-cov\)](https://www.who.int/news-room/fact-sheets/detail/middle-east-respiratory-syndrome-coronavirus-(mers-cov)).
 48. Holmes, E. C. *et al.* The origins of SARS-CoV-2: A critical review. *Cell* (2021) doi:10.1016/j.cell.2021.08.017.
 49. Lytras, S., Xia, W., Hughes, J., Jiang, X. & Robertson, D. L. The animal origin of SARS-CoV-2. *Science* **373**, 968–970 (2021).
 50. Temmam, S. *et al.* Bat coronaviruses related to SARS-CoV-2 and infectious for human cells. *Nature* **604**, 330–336 (2022).
 51. Li, J., Lai, S., Gao, G. F. & Shi, W. The emergence, genomic diversity and global spread of SARS-CoV-2. *Nature* **600**, 408–418 (2021).
 52. Scientific Advisory Group for the Origins of Novel Pathogens (SAGO) Report.
 53. Talic, S. *et al.* Effectiveness of public health measures in reducing the incidence of covid-19, SARS-CoV-2 transmission, and covid-19 mortality: systematic review and meta-analysis. *BMJ* **375**, e068302 (2021).
 54. Cauchemez, S., Kiem, C. T., Paireau, J., Rolland, P. & Fontanet, A. Lockdown impact on COVID-19 epidemics in regions across metropolitan France. *The Lancet* **396**, 1068–1069 (2020).

55. Flaxman, S. *et al.* Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature* **584**, 257–261 (2020).
56. Salje, H. *et al.* Estimating the burden of SARS-CoV-2 in France. *Science* **369**, 208–211 (2020).
57. Dehning, J. *et al.* Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions. *Science* **369**, eabb9789 (2020).
58. Kirby, T. New variant of SARS-CoV-2 in UK causes surge of COVID-19. *Lancet Respir Med* **9**, e20–e21 (2021).
59. Tegally, H. *et al.* Detection of a SARS-CoV-2 variant of concern in South Africa. *Nature* **592**, 438–443 (2021).
60. Faria, N. R. *et al.* Genomics and epidemiology of a novel SARS-CoV-2 lineage in Manaus, Brazil. *medRxiv* 2021.02.26.21252554 (2021) doi:10.1101/2021.02.26.21252554.
61. Hadfield, J. *et al.* Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics* **34**, 4121–4123 (2018).
62. Rajah, M. M. *et al.* SARS-CoV-2 Alpha, Beta, and Delta variants display enhanced Spike-mediated syncytia formation. *The EMBO Journal* **n/a**, e108944 (2021).
63. Planas, D. *et al.* Sensitivity of infectious SARS-CoV-2 B.1.1.7 and B.1.351 variants to neutralizing antibodies. *Nature Medicine* (2021) doi:10.1038/s41591-021-01318-5.
64. Nextstrain. <https://nextstrain.org/ncov/gisaid>.
65. Planas, D. *et al.* Reduced sensitivity of SARS-CoV-2 variant Delta to antibody neutralization. *Nature* **596**, 276–280 (2021).
66. Yadav, P. D. *et al.* Neutralization of variant under investigation B.1.617 with sera of BBV152 vaccinees. *Clinical Infectious Diseases* (2021) doi:10.1093/cid/ciab411.
67. Planas, D. *et al.* Considerable escape of SARS-CoV-2 Omicron to antibody neutralization. *Nature* **602**, 671–675 (2022).
68. Borges, V. *et al.* Long-Term Evolution of SARS-CoV-2 in an Immunocompromised Patient with Non-Hodgkin Lymphoma. *mSphere* **6**, e00244-21 (2021).
69. Choi, B. *et al.* Persistence and Evolution of SARS-CoV-2 in an Immunocompromised Host. *New England Journal of Medicine* **383**, 2291–2293 (2020).
70. Cevik, M., Kuppalli, K., Kindrachuk, J. & Peiris, M. Virology, transmission, and pathogenesis of SARS-CoV-2. *BMJ* **371**, m3862 (2020).
71. Wu, Z. & McGoogan, J. M. Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72 314 Cases From the Chinese Center for Disease Control and Prevention. *JAMA* **323**, 1239–1242 (2020).
72. de Gier, B. *et al.* Effect of COVID-19 vaccination on mortality by COVID-19 and on mortality by other causes, the Netherlands, January 2021–January 2022. *Vaccine* (2023) doi:10.1016/j.vaccine.2023.06.005.
73. Huang, Y.-Z. & Kuan, C.-C. Vaccination to reduce severe COVID-19 and mortality in COVID-19 patients: a systematic review and meta-analysis. *Eur Rev Med Pharmacol Sci* **26**, 1770–1776 (2022).
74. Drożdżal, S. *et al.* An update on drugs with therapeutic potential for SARS-CoV-2 (COVID-19) treatment. *Drug Resist Updat* **59**, 100794 (2021).
75. NIH. COVID-19 treatment guidelines. <https://www.covid19treatmentguidelines.nih.gov/therapies/anti-sars-cov-2-antibody-products/> (2022).
76. Marzi, M., Vakil, M. K., Bahmanyar, M. & Zarenezhad, E. Paxlovid: Mechanism of Action, Synthesis, and In Silico Study. *Biomed Res Int* **2022**, 7341493 (2022).
77. Watson, O. J. *et al.* Global impact of the first year of COVID-19 vaccination: a mathematical modelling study. *Lancet Infect Dis* **22**, 1293–1302 (2022).
78. Yang, Z.-R. *et al.* Efficacy of SARS-CoV-2 vaccines and the dose–response relationship with three major antibodies: a systematic review and meta-analysis of randomised controlled trials. *The Lancet Microbe* **4**, e236–e246 (2023).
79. Wu, N. *et al.* Long-term effectiveness of COVID-19 vaccines against infections, hospitalisations, and mortality in adults: findings from a rapid living systematic evidence synthesis and meta-analysis up to December, 2022. *The Lancet Respiratory Medicine* **11**, 439–452 (2023).
80. Hogan, A. B. *et al.* Estimating long-term vaccine effectiveness against SARS-CoV-2 variants: a model-based approach. *Nat Commun* **14**, 4325 (2023).
81. Global Commission for Post-Pandemic Policy. COVID-19 vaccine production, to January 31st

2022. <https://globalcommissionforpostpandemicpolicy.org/covid-19-vaccine-production-to-january-31st-2022/>.
82. Variation in the COVID-19 infection–fatality ratio by age, time, and geography during the pre-vaccine era: a systematic analysis. *The Lancet* **399**, 1469–1488 (2022).
83. Institute for Health Metrics and Evaluation. COVID-19 Projections. <https://covid19.healthdata.org/global?view=cumulative-deaths&tab=trend>.
84. WHO. WHO COVID Dashboard. <https://data.who.int/dashboards/covid19/deaths?n=c>.
85. Barber, R. M. *et al.* Estimating global, regional, and national daily and cumulative infections with SARS-CoV-2 through Nov 14, 2021: a statistical analysis. *The Lancet* **399**, 2351–2380 (2022).
86. Davis, H. E., McCorkell, L., Vogel, J. M. & Topol, E. J. Long COVID: major findings, mechanisms and recommendations. *Nature Reviews Microbiology* (2023) doi:10.1038/s41579-022-00846-2.
87. Hastie, C. E. *et al.* True prevalence of long-COVID in a nationwide, population cohort study. *Nat Commun* **14**, 7892 (2023).
88. Alluwaimi, A. M., Alshubait, I. H., Al-Ali, A. M. & Abohelaika, S. The Coronaviruses of Animals and Birds: Their Zoonosis, Vaccines, and Models for SARS-CoV and SARS-CoV2. *Frontiers in Veterinary Science* **7**, (2020).
89. Ye, Z.-W. *et al.* Zoonotic origins of human coronaviruses. *Int J Biol Sci* **16**, 1686–1697 (2020).
90. Rulli, M. C., D’Odorico, P., Galli, N. & Hayman, D. T. S. Land-use change and the livestock revolution increase the risk of zoonotic coronavirus transmission from rhinolophid bats. *Nat Food* **2**, 409–416 (2021).
91. Daszak, P. *et al.* Collaborative research approaches to the role of wildlife in zoonotic disease emergence. *Curr Top Microbiol Immunol* **315**, 463–475 (2007).
92. Espinosa, R., Tago, D. & Treich, N. Infectious Diseases and Meat Production. *Environ Resour Econ (Dordr)* **76**, 1019–1044 (2020).
93. Young, H., Griffin, R. H., Wood, C. L. & Nunn, C. L. Does habitat disturbance increase infectious disease risk for primates? *Ecol Lett* **16**, 656–663 (2013).
94. Gottdenker, N. L., Streicker, D. G., Faust, C. L. & Carroll, C. R. Anthropogenic land use change and infectious diseases: a review of the evidence. *Ecohealth* **11**, 619–632 (2014).
95. Masters, P. S. The Molecular Biology of Coronaviruses. *Advances in Virus Research* **66**, 193 (2006).
96. Mingaleeva, R. N. *et al.* Biology of the SARS-CoV-2 Coronavirus. *Biochemistry (Mosc)* **87**, 1662–1678 (2022).
97. Liu, C. *et al.* The Architecture of Inactivated SARS-CoV-2 with Postfusion Spikes Revealed by Cryo-EM and Cryo-ET. *Structure* **28**, 1218-1224.e4 (2020).
98. Brant, A. C., Tian, W., Majerciak, V., Yang, W. & Zheng, Z.-M. SARS-CoV-2: from its discovery to genome structure, transcription, and replication. *Cell & Bioscience* **11**, 136 (2021).
99. Vennema, H. *et al.* Nucleocapsid-independent assembly of coronavirus-like particles by co-expression of viral envelope protein genes. *EMBO J* **15**, 2020–2028 (1996).
100. Baudoux, P., Carrat, C., Besnardeau, L., Charley, B. & Laude, H. Coronavirus pseudoparticles formed with recombinant M and E proteins induce alpha interferon synthesis by leukocytes. *J Virol* **72**, 8636–8643 (1998).
101. Corse, E. & Machamer, C. E. The cytoplasmic tails of infectious bronchitis virus E and M proteins mediate their interaction. *Virology* **312**, 25–34 (2003).
102. Huang, Y., Yang, Z., Kong, W. & Nabel, G. J. Generation of Synthetic Severe Acute Respiratory Syndrome Coronavirus Pseudoparticles: Implications for Assembly and Vaccine Production. *J Virol* **78**, 12557–12565 (2004).
103. Hatakeyama, S. *et al.* Dissection and identification of regions required to form pseudoparticles by the interaction between the nucleocapsid (N) and membrane (M) proteins of SARS coronavirus. *Virology* **380**, 99–108 (2008).
104. Hsieh, P.-K. *et al.* Assembly of severe acute respiratory syndrome coronavirus RNA packaging signal into virus-like particles is nucleocapsid dependent. *J Virol* **79**, 13848–13855 (2005).
105. Ho, Y., Lin, P.-H., Liu, C. Y. Y., Lee, S.-P. & Chao, Y.-C. Assembly of human severe acute respiratory syndrome coronavirus-like particles. *Biochem Biophys Res Commun* **318**, 833–838 (2004).
106. Siu, Y. L. *et al.* The M, E, and N Structural Proteins of the Severe Acute Respiratory Syndrome

- Coronavirus Are Required for Efficient Assembly, Trafficking, and Release of Virus-Like Particles. *Journal of Virology* **82**, 11318–11330 (2008).
107. Ye, Y. & Hogue, B. G. Role of the coronavirus E viroporin protein transmembrane domain in virus assembly. *J Virol* **81**, 3597–3607 (2007).
 108. Liu, D. X., Liang, J. Q. & Fung, T. S. Human Coronavirus-229E, -OC43, -NL63, and -HKU1 (Coronaviridae). in *Encyclopedia of Virology* 428–440 (2021). doi:10.1016/b978-0-12-809633-8.21501-x.
 109. Neuman, B. W. *et al.* Supramolecular Architecture of Severe Acute Respiratory Syndrome Coronavirus Revealed by Electron Cryomicroscopy. *J Virol* **80**, 7918–7928 (2006).
 110. Li, F., Li, W., Farzan, M. & Harrison, S. C. Structure of SARS Coronavirus Spike Receptor-Binding Domain Complexed with Receptor. *Science* **309**, 1864–1868 (2005).
 111. Walls, A. C. *et al.* Cryo-electron microscopy structure of a coronavirus spike glycoprotein trimer. *Nature* **531**, 114–117 (2016).
 112. Bestle, D. *et al.* TMPRSS2 and furin are both essential for proteolytic activation of SARS-CoV-2 in human airway cells. *Life Science Alliance* **3**, e202000786 (2020).
 113. Walls, A. C. *et al.* Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **181**, 281–292 e6 (2020).
 114. Johnson, B. A. *et al.* Loss of furin cleavage site attenuates SARS-CoV-2 pathogenesis. *Nature* **591**, 293–299 (2021).
 115. Ali, H., Naseem, A. & Siddiqui, Z. I. SARS-CoV-2 Syncytium under the Radar: Molecular Insights of the Spike-Induced Syncytia and Potential Strategies to Limit SARS-CoV-2 Replication. *Journal of Clinical Medicine* **12**, 6079 (2023).
 116. Tseng, Y.-T., Chang, C.-H., Wang, S.-M., Huang, K.-J. & Wang, C.-T. Identifying SARS-CoV membrane protein amino acid residues linked to virus-like particle assembly. *PLoS One* **8**, e64013 (2013).
 117. Jahirul Islam, Md., Nawal Islam, N., Siddik Alom, Md., Kabir, M. & Halim, M. A. A review on structural, non-structural, and accessory proteins of SARS-CoV-2: Highlighting drug target sites. *Immunobiology* **228**, 152302 (2023).
 118. Fang, X. *et al.* The membrane protein of SARS-CoV suppresses NF- κ B activation. *Journal of Medical Virology* **79**, 1431–1439 (2007).
 119. Fu, Y.-Z. *et al.* SARS-CoV-2 membrane glycoprotein M antagonizes the MAVS-mediated innate antiviral response. *Cell Mol Immunol* **18**, 613–620 (2021).
 120. Sui, L. *et al.* SARS-CoV-2 Membrane Protein Inhibits Type I Interferon Production Through Ubiquitin-Mediated Degradation of TBK1. *Front Immunol* 662989–662989 (2021).
 121. Zheng, Y. *et al.* Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) membrane (M) protein inhibits type I and III interferon production by targeting RIG-I/MDA-5 signaling. *Sig Transduct Target Ther* **5**, 1–13 (2020).
 122. Zhou, S. *et al.* SARS-CoV-2 E protein: Pathogenesis and potential therapeutic development. *Biomed Pharmacother* **159**, 114242 (2023).
 123. Yuan, Z. *et al.* The E3 Ubiquitin Ligase RNF5 Facilitates SARS-CoV-2 Membrane Protein-Mediated Virion Release. *mBio* **13**, e0316821 (2021).
 124. Mehregan, A. *et al.* Probing effects of the SARS-CoV-2 E protein on membrane curvature and intracellular calcium. *Biochim Biophys Acta Biomembr* **1864**, 183994 (2022).
 125. Nieto-Torres, J. L. *et al.* Severe acute respiratory syndrome coronavirus E protein transports calcium ions and activates the NLRP3 inflammasome. *Virology* **485**, 330–339 (2015).
 126. Teoh, K.-T. *et al.* The SARS coronavirus E protein interacts with PALS1 and alters tight junction formation and epithelial morphogenesis. *Mol Biol Cell* **21**, 3838–3852 (2010).
 127. Cubuk, J. *et al.* The SARS-CoV-2 nucleocapsid protein is dynamic, disordered, and phase separates with RNA. *Nat Commun* **12**, 1936 (2021).
 128. Zúñiga, S. *et al.* Coronavirus nucleocapsid protein facilitates template switching and is required for efficient transcription. *J Virol* **84**, 2169–2175 (2010).
 129. Cong, Y. *et al.* Nucleocapsid Protein Recruitment to Replication-Transcription Complexes Plays a Crucial Role in Coronaviral Life Cycle. *J Virol* **94**, e01925-19 (2020).
 130. Zheng, Z.-Q., Wang, S.-Y., Xu, Z.-S., Fu, Y.-Z. & Wang, Y.-Y. SARS-CoV-2 nucleocapsid protein impairs stress granule formation to promote viral replication. *Cell Discov* **7**, 1–11 (2021).

131. Oh, S. J. & Shin, O. S. SARS-CoV-2 Nucleocapsid Protein Targets RIG-I-Like Receptor Pathways to Inhibit the Induction of Interferon Response. *Cells* **10**, 530 (2021).
132. Zhou, B. *et al.* The nucleocapsid protein of severe acute respiratory syndrome coronavirus inhibits cell cytokinesis and proliferation by interacting with translation elongation factor 1alpha. *J Virol* **82**, 6962–6971 (2008).
133. Zeng, Q., Langereis, M. A., van Vliet, A. L. W., Huizinga, E. G. & de Groot, R. J. Structure of coronavirus hemagglutinin-esterase offers insight into corona and influenza virus evolution. *Proc Natl Acad Sci U S A* **105**, 9065–9069 (2008).
134. Luytjes, W., Bredenbeek, P. J., Noten, A. F., Horzinek, M. C. & Spaan, W. J. Sequence of mouse hepatitis virus A59 mRNA 2: indications for RNA recombination between coronaviruses and influenza C virus. *Virology* **166**, 415–422 (1988).
135. Bakkers, M. J. G. *et al.* Betacoronavirus Adaptation to Humans Involved Progressive Loss of Hemagglutinin-Esterase Lectin Activity. *Cell Host Microbe* **21**, 356–366 (2017).
136. Hurdiss, D. L. *et al.* Cryo-EM structure of coronavirus-HKU1 haemagglutinin esterase reveals architectural changes arising from prolonged circulation in humans. *Nat Commun* **11**, 4646 (2020).
137. Lang, Y. *et al.* Coronavirus hemagglutinin-esterase and spike proteins coevolve for functional balance and optimal virion avidity. *Proceedings of the National Academy of Sciences* **117**, 25759–25770 (2020).
138. Banerjee, A. K. *et al.* SARS-CoV-2 Disrupts Splicing, Translation, and Protein Trafficking to Suppress Host Defenses. *Cell* **183**, 1325–1339.e21 (2020).
139. Xu, Z. *et al.* SARS-CoV-2 impairs interferon production via NSP2-induced repression of mRNA translation. *Proc Natl Acad Sci U S A* **119**, e2204539119 (2022).
140. Angelini, M. M., Akhlaghpour, M., Neuman, B. W. & Buchmeier, M. J. Severe Acute Respiratory Syndrome Coronavirus Nonstructural Proteins 3, 4, and 6 Induce Double-Membrane Vesicles. *mBio* **4**, 10.1128/mbio.00524-13 (2013).
141. Ma-Lauer, Y. *et al.* p53 down-regulates SARS coronavirus replication and is targeted by the SARS-unique domain and PLpro via E3 ubiquitin ligase RCHY1. *Proc Natl Acad Sci U S A* **113**, E5192–5201 (2016).
142. Yuan, L. *et al.* p53 degradation by a coronavirus papain-like protease suppresses type I interferon signaling. *J Biol Chem* **290**, 3172–3182 (2015).
143. Shin, D. *et al.* Papain-like protease regulates SARS-CoV-2 viral spread and innate immunity. *Nature* **587**, 657–662 (2020).
144. Moustaqil, M. *et al.* SARS-CoV-2 proteases PLpro and 3CLpro cleave IRF3 and critical modulators of inflammatory pathways (NLRP12 and TAB1): implications for disease presentation across species. *Emerg Microbes Infect* **10**, 178–195 (2021).
145. Liu, G. *et al.* ISG15-dependent activation of the sensor MDA5 is antagonized by the SARS-CoV-2 papain-like protease to evade host innate immunity. *Nat Microbiol* **6**, 467–478 (2021).
146. Liu, Y. *et al.* SARS-CoV-2 Nsp5 Demonstrates Two Distinct Mechanisms Targeting RIG-I and MAVS To Evade the Innate Immune Response. *mBio* **12**, e0233521 (2021).
147. Xia, H. *et al.* Evasion of Type I Interferon by SARS-CoV-2. *Cell Rep* **33**, 108234 (2020).
148. Gao, Y. *et al.* Structure of the RNA-dependent RNA polymerase from COVID-19 virus. *Science* **368**, 779–782 (2020).
149. Zhang, J. *et al.* Porcine Epidemic Diarrhea Virus nsp7 Inhibits MDA5 Dephosphorylation to Antagonize Type I Interferon Production. *Microbiology Spectrum* **11**, e05017-22 (2023).
150. Deng, J. *et al.* SARS-CoV-2 NSP7 inhibits type I and III IFN production by targeting the RIG-I/MDA5, TRIF, and STING signaling pathways. *J Med Virol* **95**, e28561 (2023).
151. Imbert, I. *et al.* A second, non-canonical RNA-dependent RNA polymerase in SARS coronavirus. *EMBO J* **25**, 4933–4942 (2006).
152. Park, G. J. *et al.* The mechanism of RNA capping by SARS-CoV-2. *Nature* **609**, 793–800 (2022).
153. Makiyama, K. *et al.* NSP9 of SARS-CoV-2 attenuates nuclear transport by hampering nucleoporin 62 dynamics and functions in host cells. *Biochem Biophys Res Commun* **586**, 137–142 (2022).
154. Bouvet, M. *et al.* In vitro reconstitution of SARS-coronavirus mRNA cap methylation. *PLoS Pathog* **6**, e1000863 (2010).

155. Xu, X. *et al.* Molecular model of SARS coronavirus polymerase: implications for biochemical functions and drug design. *Nucleic Acids Res* **31**, 7117–7130 (2003).
156. de Velthuis, A. J. W., Arnold, J. J., Cameron, C. E., van den Worm, S. H. E. & Snijder, E. J. The RNA polymerase activity of SARS-coronavirus nsp12 is primer dependent. *Nucleic Acids Res* **38**, 203–214 (2010).
157. Yuen, C.-K. *et al.* SARS-CoV-2 nsp13, nsp14, nsp15 and orf6 function as potent interferon antagonists. *Emerg Microbes Infect* **9**, 1418–1428 (2020).
158. Chen, Y. *et al.* Functional screen reveals SARS coronavirus nonstructural protein nsp14 as a novel cap N7 methyltransferase. *Proc Natl Acad Sci U S A* **106**, 3484–3489 (2009).
159. Frazier, M. N. *et al.* Flipped over U: structural basis for dsRNA cleavage by the SARS-CoV-2 endoribonuclease. *Nucleic Acids Res* **50**, 8290–8301 (2022).
160. Decroly, E. *et al.* Coronavirus Nonstructural Protein 16 Is a Cap-0 Binding Enzyme Possessing (Nucleoside-2'O)-Methyltransferase Activity. *Journal of Virology* **82**, 8071–8084 (2008).
161. Steiner, S. *et al.* SARS-CoV-2 biology and host interactions. *Nat Rev Microbiol* **22**, 206–225 (2024).
162. Shi, C. S. *et al.* SARS-coronavirus open reading frame-9b suppresses innate immunity by targeting mitochondria and the MAVS/TRAF3/TRAF6 signalosome. *J Immunol* **193**, 3080–9 (2014).
163. Lee, J. Y., Bae, S. & Myoung, J. Middle East respiratory syndrome coronavirus-encoded ORF8b strongly antagonizes IFN- β promoter activation: its implication for vaccine design. *J Microbiol.* **57**, 803–811 (2019).
164. Jiang, H. W. *et al.* SARS-CoV-2 Orf9b suppresses type I interferon responses by targeting TOM70. *Cell Mol Immunol* **17**, 998–1000 (2020).
165. Wong, L.-Y. R. *et al.* Middle East Respiratory Syndrome Coronavirus ORF8b Accessory Protein Suppresses Type I IFN Expression by Impeding HSP70-Dependent Activation of IRF3 Kinase IKK ϵ . *J Immunol* **205**, 1564–1579 (2020).
166. Han, L. *et al.* SARS-CoV-2 ORF9b antagonizes type I and III interferons by targeting multiple components of the RIG-I/MDA-5-MAVS, TLR3-TRIF, and cGAS-STING signaling pathways. *J Med Virol* **93**, 5376–5389 (2021).
167. Xu, H. *et al.* SARS-CoV-2 viroporin encoded by ORF3a triggers the NLRP3 inflammatory pathway. *Virology* **568**, 13–22 (2022).
168. Zhang, R. *et al.* The ORF4a protein of human coronavirus 229E functions as a viroporin that regulates viral production. *Biochim Biophys Acta Biomembr* **1838**, 1088–1095 (2014).
169. Zhang, R. *et al.* The ns12.9 Accessory Protein of Human Coronavirus OC43 Is a Viroporin Involved in Virion Morphogenesis and Pathogenesis. *J Virol* **89**, 11383–11395 (2015).
170. Gusho, E. *et al.* Murine AKAP7 has a 2',5'-phosphodiesterase domain that can complement an inactive murine coronavirus ns2 gene. *mBio* **5**, e01312-01314 (2014).
171. Goldstein, S. A. *et al.* Lineage A Betacoronavirus NS2 Proteins and the Homologous Torovirus Berne pp1a Carboxy-Terminal Domain Are Phosphodiesterases That Antagonize Activation of RNase L. *J Virol* **91**, e02201-16 (2017).
172. Zhou, Y. *et al.* Host E3 ligase HUWE1 attenuates the proapoptotic activity of the MERS-CoV accessory protein ORF3 by promoting its ubiquitin-dependent degradation. *J Biol Chem* **298**, 101584 (2022).
173. Dijkman, R. *et al.* Human coronavirus 229E encodes a single ORF4 protein between the spike and the envelope genes. *Virol J* **3**, 106 (2006).
174. Freundt, E. C. *et al.* The Open Reading Frame, 3a Protein of Severe Acute Respiratory Syndrome-Associated Coronavirus Promotes Membrane Rearrangement and Cell Death. *Journal of Virology* **84**, 1097–1109 (2010).
175. Chen, D. *et al.* ORF3a of SARS-CoV-2 promotes lysosomal exocytosis-mediated viral egress. *Dev Cell* **56**, 3250-3263.e5 (2021).
176. Ren, Y. *et al.* The ORF3a protein of SARS-CoV-2 induces apoptosis in cells. *Cellular & Molecular Immunology* (2020) doi:10.1038/s41423-020-0485-9.
177. Zhou, P., Li, H., Wang, H., Wang, L.-F. & Shi, Z. Bat severe acute respiratory syndrome-like coronavirus ORF3b homologues display different interferon antagonist activities. *J Gen Virol* **93**, 275–281 (2012).
178. Konno, Y. *et al.* SARS-CoV-2 ORF3b Is a Potent Interferon Antagonist Whose Activity Is

- Increased by a Naturally Occurring Elongation Variant. *Cell Rep* **32**, 108185 (2020).
179. Stewart, H. *et al.* The SARS-CoV-2 protein ORF3c is a mitochondrial modulator of innate immunity. *iScience* **26**, 108080 (2023).
180. Y, Y. *et al.* The structural and accessory proteins M, ORF 4a, ORF 4b, and ORF 5 of Middle East respiratory syndrome coronavirus (MERS-CoV) are potent interferon antagonists. *Protein & cell* **4**, (2013).
181. Niemeyer, D. *et al.* Middle East respiratory syndrome coronavirus accessory protein 4a is a type I interferon antagonist. *J Virol* **87**, 12489–12495 (2013).
182. Rabouw, H. H. *et al.* Middle East Respiratory Coronavirus Accessory Protein 4a Inhibits PKR-Mediated Antiviral Stress Responses. *PLoS Pathog* **12**, e1005982 (2016).
183. Thornbrough, J. M. *et al.* Middle East Respiratory Syndrome Coronavirus NS4b Protein Inhibits Host RNase L Activation. *mBio* **7**, e00258 (2016).
184. Matthews, K. L., Coleman, C. M., van der Meer, Y., Snijder, E. J. & Frieman, M. B. The ORF4b-encoded accessory proteins of Middle East respiratory syndrome coronavirus and two related bat coronaviruses localize to the nucleus and inhibit innate immune signalling. *J Gen Virol* **95**, 874–882 (2014).
185. Yang, Y. *et al.* Middle East respiratory syndrome coronavirus ORF4b protein inhibits type I interferon production through both cytoplasmic and nuclear targets. *Sci Rep* **5**, 17554 (2015).
186. Menachery, V. D. *et al.* MERS-CoV Accessory ORFs Play Key Role for Infection and Pathogenesis. *mBio* **8**, 10.1128/mbio.00665-17 (2017).
187. Frieman, M. *et al.* Severe acute respiratory syndrome coronavirus ORF6 antagonizes STAT1 function by sequestering nuclear import factors on the rough endoplasmic reticulum/Golgi membrane. *J Virol* **81**, 9812–9824 (2007).
188. Miorin, L. *et al.* SARS-CoV-2 Orf6 hijacks Nup98 to block STAT nuclear import and antagonize interferon signaling. *Proc Natl Acad Sci U S A* **117**, 28344–28354 (2020).
189. Addetia, A. *et al.* SARS-CoV-2 ORF6 Disrupts Bidirectional Nucleocytoplasmic Transport through Interactions with Rae1 and Nup98. *mBio* **12**, e00065-21 (2021).
190. Shemesh, M. *et al.* SARS-CoV-2 suppresses IFN β production mediated by NSP1, 5, 6, 15, ORF6 and ORF7b but does not suppress the effects of added interferon. *PLoS Pathog* **17**, e1009800 (2021).
191. Arshad, N. *et al.* SARS-CoV-2 accessory proteins ORF7a and ORF3a use distinct mechanisms to down-regulate MHC-I surface expression. *Proc Natl Acad Sci U S A* **120**, e2208525120 (2023).
192. Kee, J. *et al.* SARS-CoV-2 disrupts host epigenetic regulation via histone mimicry. *Nature* **610**, 381–388 (2022).
193. Zhang, Y. *et al.* The ORF8 protein of SARS-CoV-2 mediates immune evasion through down-regulating MHC-I. *Proc Natl Acad Sci U S A* **118**, e2024202118 (2021).
194. Liu, D. X., Fung, T. S., Chong, K. K.-L., Shukla, A. & Hilgenfeld, R. Accessory proteins of SARS-CoV and other coronaviruses. *Antiviral Res* **109**, 97–109 (2014).
195. Lenhard, S. *et al.* The Orf9b protein of SARS-CoV-2 modulates mitochondrial protein biogenesis. *J Cell Biol* **222**, e202303002 (2023).
196. Vijgen, L. *et al.* Circulation of genetically distinct contemporary human coronavirus OC43 strains. *Virology* **337**, 85–92 (2005).
197. Li, X. *et al.* SARS-CoV-2 ORF10 suppresses the antiviral innate immune response by degrading MAVS through mitophagy. *Cell Mol Immunol* **19**, 67–78 (2022).
198. Yuan, Y. *et al.* Cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins reveal the dynamic receptor binding domains. *Nat Commun* **8**, 15092 (2017).
199. Gui, M. *et al.* Cryo-electron microscopy structures of the SARS-CoV spike glycoprotein reveal a prerequisite conformational state for receptor binding. *Cell Res* **27**, 119–129 (2017).
200. Yan, R. *et al.* Structural basis for the different states of the spike protein of SARS-CoV-2 in complex with ACE2. *Cell Res* **31**, 717–719 (2021).
201. Milewska, A. *et al.* Human coronavirus NL63 utilizes heparan sulfate proteoglycans for attachment to target cells. *J Virol* **88**, 13221–30 (2014).
202. Clausen, T. M. *et al.* SARS-CoV-2 Infection Depends on Cellular Heparan Sulfate and ACE2. *Cell* **183**, 1043-1057 e15 (2020).
203. Park, Y.-J. *et al.* Structures of MERS-CoV spike glycoprotein in complex with sialoside

- attachment receptors. *Nat Struct Mol Biol* **26**, 1151–1157 (2019).
204. Li, Z. *et al.* Synthetic O-acetylated sialosides facilitate functional receptor identification for human respiratory viruses. *Nat. Chem.* **13**, 496–503 (2021).
205. Pronker, M. F. *et al.* Sialoglycan binding triggers spike opening in a human coronavirus. *Nature* **624**, 201–206 (2023).
206. Xia, L., Zhang, Y. & Zhou, Q. Structural basis for the recognition of HCoV-HKU1 by human TMPRSS2. *Cell Res* 1–4 (2024) doi:10.1038/s41422-024-00958-9.
207. Li, W. *et al.* Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* **426**, 450–454 (2003).
208. Hofmann, H. *et al.* Human coronavirus NL63 employs the severe acute respiratory syndrome coronavirus receptor for cellular entry. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 7988–93 (2005).
209. Lan, J. *et al.* Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **581**, 215–220 (2020).
210. Raj, V. S. *et al.* Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC. *Nature* **495**, 251–254 (2013).
211. Curtis L. Yeager, R. A. A., Richard K. Williams, C. B. C. & Linda H. Shapiro, A. T. L., Kathryn V. Holmes. Human aminopeptidase N is a receptor for human coronavirus 229E. *Nature* (1992).
212. Amraie, R. *et al.* CD209L/L-SIGN and CD209/DC-SIGN act as receptors for SARS-CoV-2 and are differentially expressed in lung and kidney epithelial and endothelial cells. *bioRxiv* 2020.06.22.165803 (2020) doi:10.1101/2020.06.22.165803.
213. Cantuti-Castelvetri, L. *et al.* Neuropilin-1 facilitates SARS-CoV-2 cell entry and infectivity. *Science* **370**, 856–860 (2020).
214. Baggen, J. *et al.* TMEM106B is a receptor mediating ACE2-independent SARS-CoV-2 cell entry. *Cell* **186**, 3427–3442.e22 (2023).
215. Zang, R. *et al.* TMPRSS2 and TMPRSS4 promote SARS-CoV-2 infection of human small intestinal enterocytes. *Sci Immunol* **5**, (2020).
216. Kishimoto, M. *et al.* TMPRSS11D and TMPRSS13 Activate the SARS-CoV-2 Spike Protein. *Viruses* **13**, (2021).
217. Benlarbi, M. *et al.* Identification and differential usage of a host metalloproteinase entry pathway by SARS-CoV-2 Delta and Omicron. *iScience* **25**, 105316 (2022).
218. Chan, J. F.-W. *et al.* Altered host protease determinants for SARS-CoV-2 Omicron. *Sci Adv* **9**, eadd3867 (2023).
219. Jocher, G. *et al.* ADAM10 and ADAM17 promote SARS-CoV-2 cell entry and spike protein-mediated lung cell fusion. *EMBO Rep* **23**, e54305 (2022).
220. Milewska, A. *et al.* Kallikrein 13 serves as a priming protease during infection by the human coronavirus HKU1. *Science Signaling* **13**, eaba9902 (2020).
221. Shirato, K., Kawase, M. & Matsuyama, S. Wild-type human coronaviruses prefer cell-surface TMPRSS2 to endosomal cathepsins for cell entry. *Virology* **517**, 9–15 (2018).
222. Fraser, B. J. *et al.* Structure and activity of human TMPRSS2 protease implicated in SARS-CoV-2 activation. *Nat Chem Biol* **18**, 963–971 (2022).
223. Milewska, A. *et al.* Entry of Human Coronavirus NL63 into the Cell. *Journal of virology* (2018).
224. Bayati, A., Kumar, R., Francis, V. & McPherson, P. S. SARS-CoV-2 infects cells after viral entry via clathrin-mediated endocytosis. *J Biol Chem* **296**, 100306 (2021).
225. Owczarek, K. *et al.* Early events during human coronavirus OC43 entry to the cell. *Sci Rep* **8**, 7124 (2018).
226. Simmons, G. *et al.* Inhibitors of cathepsin L prevent severe acute respiratory syndrome coronavirus entry. *Proc Natl Acad Sci U S A* **102**, 11876–11881 (2005).
227. Huang, I.-C. *et al.* SARS coronavirus, but not human coronavirus NL63, utilizes cathepsin L to infect ACE2-expressing cells. *J Biol Chem* **281**, 3198–3203 (2006).
228. Zhao, M.-M. *et al.* Cathepsin L plays a key role in SARS-CoV-2 infection in humans and humanized mice and is a promising target for new drug development. *Signal Transduction and Targeted Therapy* **6**, 134 (2021).
229. Blau, D. M. & Holmes, K. V. Human coronavirus HCoV-229E enters susceptible cells via the endocytic pathway. *Adv Exp Med Biol* **494**, 193–198 (2001).
230. Eifart, P. *et al.* Role of endocytosis and low pH in murine hepatitis virus strain A59 cell entry.

J Virol **81**, 10758–10768 (2007).

231. Rey, F. Structure-function relations of the SARS-CoV-2 spike protein and impact of mutations in the variants of concern. *C R Biol* **344**, 77–110 (2021).
232. Böttcher-Friebertshäuser, E. Membrane-Anchored Serine Proteases: Host Cell Factors in Proteolytic Activation of Viral Glycoproteins. *Activation of Viruses by Host Proteases* 153–203 (2018) doi:10.1007/978-3-319-75474-1_8.
233. Tai, W. *et al.* Characterization of the receptor-binding domain (RBD) of 2019 novel coronavirus: implication for development of RBD protein as a viral attachment inhibitor and vaccine. *Cellular & Molecular Immunology* **17**, 613–620 (2020).
234. Park, J.-E. *et al.* Proteolytic processing of Middle East respiratory syndrome coronavirus spikes expands virus tropism. *Proceedings of the National Academy of Sciences* **113**, 12262–12267 (2016).
235. Bertram, S. *et al.* TMPRSS2 activates the human coronavirus 229E for cathepsin-independent host cell entry and is expressed in viral target cells in the respiratory epithelium. *J Virol* **87**, 6150–60 (2013).
236. Hulswit, R. J. G. *et al.* Human coronaviruses OC43 and HKU1 bind to 9-O-acetylated sialic acids via a conserved receptor-binding site in spike protein domain A. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 2681–2690 (2019).
237. Bussani, R. *et al.* Persistence of viral RNA, pneumocyte syncytia and thrombosis are hallmarks of advanced COVID-19 pathology. *EBioMedicine* **61**, 103104 (2020).
238. Cifuentes-Muñoz, N., Dutch, R. E. & Cattaneo, R. Direct cell-to-cell transmission of respiratory viruses: The fast lanes. *PLoS Pathog* **14**, e1007015 (2018).
239. Zhang, Z. *et al.* SARS-CoV-2 spike protein dictates syncytium-mediated lymphocyte elimination. *Cell Death & Differentiation* (2021) doi:10.1038/s41418-021-00782-3.
240. Buchrieser, J. *et al.* Syncytia formation by SARS-CoV-2-infected cells. *EMBO J* **39**, e106267 (2020).
241. Ma, H. *et al.* Pyroptosis of syncytia formed by fusion of SARS-CoV-2 spike and ACE2-expressing cells. *Cell Discov* **7**, 73 (2021).
242. Rajah, M. M., Bernier, A., Buchrieser, J. & Schwartz, O. The Mechanism and Consequences of SARS-CoV-2 Spike-Mediated Fusion and Syncytia Formation. *J Mol Biol* **434**, 167280 (2022).
243. Cavanagh, D. & Britton, P. Coronaviruses: General Features. *Encyclopedia of Virology* 549–554 (2008) doi:10.1016/B978-012374410-4.00370-8.
244. Irigoyen, N. *et al.* High-Resolution Analysis of Coronavirus Gene Expression by RNA Sequencing and Ribosome Profiling. *PLoS Pathog* **12**, e1005473 (2016).
245. Finkel, Y. *et al.* The coding capacity of SARS-CoV-2. *Nature* **589**, 125–130 (2021).
246. Oudshoorn, D. *et al.* Expression and Cleavage of Middle East Respiratory Syndrome Coronavirus nsp3-4 Polyprotein Induce the Formation of Double-Membrane Vesicles That Mimic Those Associated with Coronaviral RNA Replication. *mBio* **8**, e01658-17 (2017).
247. Wolff, G. *et al.* A molecular pore spans the double membrane of the coronavirus replication organelle. *Science* **369**, 1395–1398 (2020).
248. Zimmermann, L. *et al.* SARS-CoV-2 nsp3 and nsp4 are minimal constituents of a pore spanning replication organelle. *Nat Commun* **14**, 7894 (2023).
249. Grellet, E., L'Hôte, I., Goulet, A. & Imbert, I. Replication of the coronavirus genome: A paradox among positive-strand RNA viruses. *J Biol Chem* **298**, 101923 (2022).
250. Voß, D. *et al.* Studies on membrane topology, N-glycosylation and functionality of SARS-CoV membrane protein. *Virol J* **6**, 79 (2009).
251. Cavanagh, D. Coronaviridae: a review of coronaviruses and toroviruses. *Coronaviruses with Special Emphasis on First Insights Concerning SARS* 1–54 (2005) doi:10.1007/3-7643-7339-3_1.
252. de Haan, C. A. M. & Rottier, P. J. M. Molecular Interactions in the Assembly of Coronaviruses. *Adv Virus Res* **64**, 165–230 (2005).
253. Ghosh, S. *et al.* β -Coronaviruses Use Lysosomes for Egress Instead of the Biosynthetic Secretory Pathway. *Cell* **183**, 1520-1535.e14 (2020).
254. Malone, B., Urakova, N., Snijder, E. J. & Campbell, E. A. Structures and functions of coronavirus replication–transcription complexes and their relevance for SARS-CoV-2 drug design. *Nat Rev Mol Cell Biol* **23**, 21–39 (2022).
255. Bortolotti, D. *et al.* TLR3 and TLR7 RNA Sensor Activation during SARS-CoV-2 Infection.

Microorganisms **9**, 1820 (2021).

256. Rebendenne, A. *et al.* SARS-CoV-2 triggers an MDA-5-dependent interferon response which is unable to control replication in lung epithelial cells. *J Virol* **95**, e02415-20, JVI.02415-20 (2021).
257. Yin, X. *et al.* MDA5 Governs the Innate Immune Response to SARS-CoV-2 in Lung Epithelial Cells. *Cell Rep* **34**, 108628 (2021).
258. Thorne, L. G. *et al.* SARS-CoV-2 sensing by RIG-I and MDA5 links epithelial infection to macrophage inflammation. *EMBO J* **40**, e107826 (2021).
259. Domizio, J. D. *et al.* The cGAS-STING pathway drives type I IFN immunopathology in COVID-19. *Nature* **603**, 145–151 (2022).
260. Min, Y.-Q. *et al.* Immune evasion of SARS-CoV-2 from interferon antiviral system. *Computational and Structural Biotechnology Journal* **19**, 4217–4225 (2021).
261. Pfaender, S. *et al.* LY6E impairs coronavirus fusion and confers immune control of viral disease. *Nat Microbiol* **5**, 1330–1339 (2020).
262. Martin-Sancho, L. *et al.* Functional landscape of SARS-CoV-2 cellular restriction. *Molecular Cell* **81**, 2656-2668.e8 (2021).
263. Wang, S.-M., Huang, K.-J. & Wang, C.-T. BST2/CD317 counteracts human coronavirus 229E productive infection by tethering virions at the cell surface. *Virology* **449**, 287–296 (2014).
264. Winstone, H. *et al.* The polybasic cleavage site in the SARS-CoV-2 spike modulates viral sensitivity to Type I interferon and IFITM2. *Journal of Virology* JVI.02422-20 (2021) doi:10.1128/jvi.02422-20.
265. Zani, A. *et al.* Interferon-induced transmembrane protein 3 (IFITM3) limits lethality of SARS-CoV-2 in mice. *bioRxiv* (2021) doi:10.1101/2021.12.22.473914.
266. Zhao, X. *et al.* Interferon induction of IFITM proteins promotes infection by human coronavirus OC43. *Proceedings of the National Academy of Sciences* **111**, 6756–6761 (2014).
267. Shi, G. *et al.* Opposing activities of IFITM proteins in SARS-CoV-2 infection. *EMBO J* **40**, e106501 (2021).
268. Prelli Bozzo, C. *et al.* IFITM proteins promote SARS-CoV-2 infection and are targets for virus inhibition in vitro. *Nat Commun* **12**, 4584 (2021).
269. Ong, E. Z. *et al.* A Dynamic Immune Response Shapes COVID-19 Progression. *Cell Host Microbe* **27**, 879-882 e2 (2020).
270. Schulte-Schrepping, J. *et al.* Severe COVID-19 Is Marked by a Dysregulated Myeloid Cell Compartment. *Cell* **182**, 1419-1440.e23 (2020).
271. Silvin, A. *et al.* Elevated Calprotectin and Abnormal Myeloid Cell Subsets Discriminate Severe from Mild COVID-19. *Cell* **182**, 1401-1418.e18 (2020).
272. Wang, X. *et al.* The role of IL-6 in coronavirus, especially in COVID-19. *Front. Pharmacol.* **13**, (2022).
273. Sefik, E. *et al.* Inflammasome activation in infected macrophages drives COVID-19 pathology. *Nature* **606**, 585–593 (2022).
274. Zelek, W. M. & Harrison, R. A. Complement and COVID-19: Three years on, what we know, what we don't know, and what we ought to know. *Immunobiology* **228**, 152393 (2023).
275. Chouaki Benmansour, N., Carvelli, J. & Vivier, É. Implication de la cascade du complément dans les formes sévères de COVID-19. *Médecine/Sciences* **37**, 333–341 (2021).
276. Janiuk, K., Jabłońska, E. & Garley, M. Significance of NETs Formation in COVID-19. *Cells* **10**, 151 (2021).
277. Veenith, T. *et al.* High generation of reactive oxygen species from neutrophils in patients with severe COVID-19. *Sci Rep* **12**, 10484 (2022).
278. Cervantes-Barragan, L. *et al.* Control of coronavirus infection through plasmacytoid dendritic-cell-derived type I interferon. *Blood* **109**, 1131–1137 (2006).
279. Zhou, R. *et al.* Acute SARS-CoV-2 Infection Impairs Dendritic Cell and T Cell Responses. *Immunity* (2020) doi:10.1016/j.immuni.2020.07.026.
280. Venet, M. *et al.* Severe COVID-19 patients have impaired plasmacytoid dendritic cell-mediated control of SARS-CoV-2. *Nat Commun* **14**, 694 (2023).
281. Mesel-Lemoine, M. *et al.* A human coronavirus responsible for the common cold massively kills dendritic cells but not monocytes. *J Virol* **86**, 7577–7587 (2012).
282. Pérez-Gómez, A. *et al.* Dendritic cell deficiencies persist seven months after SARS-CoV-2

- infection. *Cell Mol Immunol* **18**, 2128–2139 (2021).
283. Welch, S. K., Saif, L. J. & Ram, S. Cell-mediated immune responses of suckling pigs inoculated with attenuated or virulent transmissible gastroenteritis virus. *Am J Vet Res* **49**, 1228–1234 (1988).
284. Wang, Y.-D. *et al.* T-cell epitopes in severe acute respiratory syndrome (SARS) coronavirus spike protein elicit a specific T-cell immune response in patients who recover from SARS. *J Virol* **78**, 5612–5618 (2004).
285. Alhabbab, R. Y. *et al.* MERS-CoV Infection Elicits Long-lasting Specific Antibody, T and B Cell Immune Responses in Recovered Individuals. *Clin Infect Dis* ciac456 (2022) doi:10.1093/cid/ciac456.
286. Grifoni, A. *et al.* Targets of T Cell Responses to SARS-CoV-2 Coronavirus in Humans with COVID-19 Disease and Unexposed Individuals. *Cell* **181**, 1489-1501.e15 (2020).
287. Sekine, T. *et al.* Robust T Cell Immunity in Convalescent Individuals with Asymptomatic or Mild COVID-19. *Cell* **183**, 158-168.e14 (2020).
288. Li, T. *et al.* Significant changes of peripheral T lymphocyte subsets in patients with severe acute respiratory syndrome. *J Infect Dis* **189**, 648–651 (2004).
289. Braun, J. *et al.* SARS-CoV-2-reactive T cells in healthy donors and patients with COVID-19. *Nature* **587**, 270–274 (2020).
290. Jafarzadeh, A., Jafarzadeh, S., Nozari, P., Mokhtari, P. & Nemati, M. Lymphopenia an important immunological abnormality in patients with COVID-19: Possible mechanisms. *Scandinavian Journal of Immunology* **93**, e12967 (2021).
291. Channappanavar, R., Zhao, J. & Perlman, S. T cell-mediated immune response to respiratory coronaviruses. *Immunol Res* **59**, 118–128 (2014).
292. Yang, L.-T. *et al.* Long-lived effector/central memory T-cell responses to severe acute respiratory syndrome coronavirus (SARS-CoV) S antigen in recovered SARS patients. *Clin Immunol* **120**, 171–178 (2006).
293. Jin, H. *et al.* Induction of Th1 type response by DNA vaccinations with N, M, and E genes against SARS-CoV in mice. *Biochem Biophys Res Commun* **328**, 979–986 (2005).
294. Keeton, R. *et al.* T cell responses to SARS-CoV-2 spike cross-recognize Omicron. *Nature* **603**, 488–492 (2022).
295. Pei, J., Briles, W. E. & Collisson, E. W. Memory T cells protect chicks from acute infectious bronchitis virus infection. *Virology* **306**, 376–384 (2003).
296. McMahan, K. *et al.* Correlates of protection against SARS-CoV-2 in rhesus macaques. *Nature* **590**, 630–634 (2021).
297. Premkumar, L. *et al.* The receptor binding domain of the viral spike protein is an immunodominant and highly specific target of antibodies in SARS-CoV-2 patients. *Sci Immunol* **5**, eabc8413 (2020).
298. Röltgen, K. & Boyd, S. D. Antibody and B cell responses to SARS-CoV-2 infection and vaccination. *Cell Host & Microbe* **29**, 1063–1075 (2021).
299. Piccoli, L. *et al.* Mapping Neutralizing and Immunodominant Sites on the SARS-CoV-2 Spike Receptor-Binding Domain by Structure-Guided High-Resolution Serology. *Cell* **183**, 1024-1042.e21 (2020).
300. Iyer, A. S. *et al.* Persistence and decay of human antibody responses to the receptor binding domain of SARS-CoV-2 spike protein in COVID-19 patients. *Science Immunology* **5**, eabe0367 (2020).
301. Anand, S. P. *et al.* Longitudinal analysis of humoral immunity against SARS-CoV-2 Spike in convalescent individuals up to 8 months post-symptom onset. *Cell Rep Med* **2**, 100290 (2021).
302. Röltgen, K. *et al.* Defining the features and duration of antibody responses to SARS-CoV-2 infection associated with disease severity and outcome. *Sci Immunol* **5**, eabe0240 (2020).
303. Weisberg, S. P. *et al.* Distinct antibody responses to SARS-CoV-2 in children and adults across the COVID-19 clinical spectrum. *Nature Immunology* **22**, 25–31 (2021).
304. Lucas, C. *et al.* Longitudinal analyses reveal immunological misfiring in severe COVID-19. *Nature* **584**, 463–469 (2020).
305. Lucas, C. *et al.* Impact of circulating SARS-CoV-2 variants on mRNA vaccine-induced immunity. *Nature* (2021) doi:10.1038/s41586-021-04085-y.
306. Bladh, O. *et al.* Mucosal immune responses following a fourth SARS-CoV-2 vaccine dose. *The Lancet Microbe* **4**, e488 (2023).

307. Puhach, O. *et al.* SARS-CoV-2 convalescence and hybrid immunity elicits mucosal immune responses. *eBioMedicine* **98**, (2023).
308. Schäfer, A. *et al.* Antibody potency, effector function, and combinations in protection and therapy for SARS-CoV-2 infection in vivo. *J Exp Med* **218**, e20201993 (2020).
309. Mackin, S. R. *et al.* Fc- γ R-dependent antibody effector functions are required for vaccine-mediated protection against antigen-shifted variants of SARS-CoV-2. *Nat Microbiol* **8**, 569–580 (2023).
310. Khoury, D. S. *et al.* Neutralizing antibody levels are highly predictive of immune protection from symptomatic SARS-CoV-2 infection. *Nature Medicine* (2021) doi:10.1038/s41591-021-01377-8.
311. Lingas, G. *et al.* Neutralizing Antibody Levels as a Correlate of Protection Against SARS-CoV-2 Infection: A Modeling Analysis. *Clin Pharmacol Ther* **115**, 86–94 (2024).
312. Gorse, G. J., Patel, G. B., Vitale, J. N. & O'Connor, T. Z. Prevalence of Antibodies to Four Human Coronaviruses Is Lower in Nasal Secretions than in Serum. *Clinical and Vaccine Immunology* **17**, 1875–1880 (2010).
313. Anderson, E. M. *et al.* Seasonal human coronavirus antibodies are boosted upon SARS-CoV-2 infection but not associated with protection. *Cell* **184**, 1858–1864.e10 (2021).
314. Sagar, M. *et al.* Recent endemic coronavirus infection is associated with less-severe COVID-19. *J Clin Invest* **131**, e143380, 143380 (2021).
315. Aran, D., Beachler, D. C., Lanes, S. & Overhage, J. M. Prior presumed coronavirus infection reduces COVID-19 risk: A cohort study. *J Infect* **81**, 923–930 (2020).
316. Mateus, J. *et al.* Low-dose mRNA-1273 COVID-19 vaccine generates durable memory enhanced by cross-reactive T cells. *Science* **0**, eabj9853.
317. Loyal, L. *et al.* Cross-reactive CD4+ T cells enhance SARS-CoV-2 immune responses upon infection and vaccination. *Science* **374**, eabh1823 (2021).
318. Lin, C.-Y. *et al.* Pre-existing humoral immunity to human common cold coronaviruses negatively impacts the protective SARS-CoV-2 antibody response. *Cell Host Microbe* **30**, 83–96.e4 (2022).
319. Murray, S. M. *et al.* The impact of pre-existing cross-reactive immunity on SARS-CoV-2 infection and vaccine responses. *Nat Rev Immunol* **23**, 304–316 (2023).
320. Yurkovetskiy, L. *et al.* Structural and Functional Analysis of the D614G SARS-CoV-2 Spike Protein Variant. *Cell* **183**, 739–751.e8 (2020).
321. Zhang, J. *et al.* Structural impact on SARS-CoV-2 spike protein by D614G substitution. *Science* **372**, 525–530 (2021).
322. Cheng, Y.-W. *et al.* D614G Substitution of SARS-CoV-2 Spike Protein Increases Syncytium Formation and Virus Titer via Enhanced Furin-Mediated Spike Cleavage. *mBio* **12**, e0058721 (2021).
323. Hou, Y. J. *et al.* SARS-CoV-2 D614G variant exhibits efficient replication ex vivo and transmission in vivo. *Science* **370**, 1464–1468 (2020).
324. Plante, J. A. *et al.* Spike mutation D614G alters SARS-CoV-2 fitness. *Nature* (2020) doi:10.1038/s41586-020-2895-3.
325. Korber, B. *et al.* Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. *Cell* **182**, 812–827.e19 (2020).
326. Zhao, S. *et al.* Modelling the association between COVID-19 transmissibility and D614G substitution in SARS-CoV-2 spike protein: using the surveillance data in California as an example. *Theoretical Biology and Medical Modelling* **18**, 10 (2021).
327. Zhou, B. *et al.* SARS-CoV-2 spike D614G change enhances replication and transmission. *Nature* **592**, 122–127 (2021).
328. Gellenoncourt, S. *et al.* The Spike-Stabilizing D614G Mutation Interacts with S1/S2 Cleavage Site Mutations To Promote the Infectious Potential of SARS-CoV-2 Variants. *J Virol* **96**, e0130122 (2022).
329. Dhawan, M. *et al.* Delta variant (B.1.617.2) of SARS-CoV-2: Mutations, impact, challenges and possible solutions. *Hum Vaccin Immunother* **18**, 2068883.
330. Zhang, J. *et al.* Membrane fusion and immune evasion by the spike protein of SARS-CoV-2 Delta variant. *Science* **0**, eabl9463 (2021).
331. Motozono, C. *et al.* SARS-CoV-2 spike L452R variant evades cellular immunity and increases infectivity. *Cell Host Microbe* **29**, 1124–1136.e11 (2021).
332. Furusawa, Y. *et al.* In SARS-CoV-2 delta variants, Spike-P681R and D950N promote

- membrane fusion, Spike-P681R enhances spike cleavage, but neither substitution affects pathogenicity in hamsters. *eBioMedicine* **91**, 104561 (2023).
333. Meng, B. *et al.* Recurrent emergence of SARS-CoV-2 spike deletion H69/V70 and its role in the Alpha variant B.1.1.7. *Cell Rep* **35**, 109292 (2021).
334. Harvey, W. T. *et al.* SARS-CoV-2 variants, spike mutations and immune escape. *Nature Reviews Microbiology* **19**, 409–424 (2021).
335. Cao, Y. *et al.* Imprinted SARS-CoV-2 humoral immunity induces convergent Omicron RBD evolution. *Nature* <https://doi.org/10.1038/s41586-022-05644-7> (2022) doi:10.1038/s41586-022-05644-7.
336. Cao, Y. *et al.* Characterization of the enhanced infectivity and antibody evasion of Omicron BA.2.75. *Cell Host & Microbe* **30**, 1527–1539.e5 (2022).
337. Wang, Q. *et al.* Antigenic characterization of the SARS-CoV-2 Omicron subvariant BA.2.75. *Cell Host & Microbe* doi:10.1016/j.chom.2022.09.002.
338. Mesner, D. *et al.* SARS-CoV-2 evolution influences GBP and IFITM sensitivity. *Proc Natl Acad Sci U S A* **120**, e2212577120.
339. Guo, K. *et al.* Interferon resistance of emerging SARS-CoV-2 variants. *Proc Natl Acad Sci U S A* **119**, e2203760119 (2022).
340. Nchioua, R. *et al.* Reduced replication but increased interferon resistance of SARS-CoV-2 Omicron BA.1. *Life Sci Alliance* **6**, e202201745 (2023).
341. Shi, G. *et al.* Omicron Spike confers enhanced infectivity and interferon resistance to SARS-CoV-2 in human nasal tissue. *Nat Commun* **15**, 889 (2024).
342. Greaney, A. J. *et al.* Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition. *Cell Host Microbe* **29**, 44–57 e9 (2021).
343. Lassauniere, R. *et al.* Neutralisation of SARS-CoV-2 Delta sub-lineage AY.4.2 and B.1.617.2+E484K by BNT162b2 mRNA vaccine-elicited sera. *medRxiv* 2021.11.08.21266075 (2021) doi:10.1101/2021.11.08.21266075.
344. Shen, L. *et al.* Spike Protein NTD mutation G142D in SARS-CoV-2 Delta VOC lineages is associated with frequent back mutations, increased viral loads, and immune evasion. 2021.09.12.21263475 Preprint at <https://doi.org/10.1101/2021.09.12.21263475> (2021).
345. Ginex, T. *et al.* The structural role of SARS-CoV-2 genetic background in the emergence and success of spike mutations: The case of the spike A222V mutation. *PLoS Pathog* **18**, e1010631 (2022).
346. Smith, N. *et al.* Distinct systemic and mucosal immune responses during acute SARS-CoV-2 infection. *Nature Immunology* **22**, 1428–1439 (2021).
347. Dan, J. M. *et al.* Immunological memory to SARS-CoV-2 assessed for up to 8 months after infection. *Science* **371**, eabf4063 (2021).
348. Skelly, D. T. *et al.* Two doses of SARS-CoV-2 vaccination induce robust immune responses to emerging SARS-CoV-2 variants of concern. *Nat Commun* **12**, 5061 (2021).
349. Sahin, U. *et al.* BNT162b2 vaccine induces neutralizing antibodies and poly-specific T cells in humans. *Nature* **595**, 572–577 (2021).
350. Alter, G. *et al.* Immunogenicity of Ad26.COV2.S vaccine against SARS-CoV-2 variants in humans. *Nature* **596**, 268–272 (2021).
351. Crotty, S. Hybrid immunity. *Science* **372**, 1392–1393 (2021).
352. Reynolds, C. J. *et al.* Prior SARS-CoV-2 infection rescues B and T cell responses to variants after first vaccine dose. *Science* **372**, 1418–1423 (2021).
353. Stamatatos, L. *et al.* mRNA vaccination boosts cross-variant neutralizing antibodies elicited by SARS-CoV-2 infection. *Science* eabg9175 (2021) doi:10.1126/science.abg9175.
354. Park, Y.-J. *et al.* Imprinted antibody responses against SARS-CoV-2 Omicron sublineages. *Science* **378**, 619–627 (2022).
355. Johnston, T. S. *et al.* Immunological imprinting shapes the specificity of human antibody responses against SARS-CoV-2 variants. *medRxiv* 2024.01.08.24301002 (2024) doi:10.1101/2024.01.08.24301002.
356. Pušnik, J. *et al.* Vaccination impairs de novo immune response to omicron breakthrough infection, a precondition for the original antigenic sin. *Nat Commun* **15**, 3102 (2024).
357. Röltgen, K. *et al.* Immune imprinting, breadth of variant recognition, and germinal center response in human SARS-CoV-2 infection and vaccination. *Cell* **185**, 1025–1040.e14 (2022).

358. Reynolds, C. J. *et al.* Immune boosting by B.1.1.529 (Omicron) depends on previous SARS-CoV-2 exposure. *Science* **0**, eabq1841.
359. Hornsby, H. *et al.* Omicron infection following vaccination enhances a broad spectrum of immune responses dependent on infection history. *Nat Commun* **14**, 5065 (2023).
360. Yisimayi, A. *et al.* Repeated Omicron exposures override ancestral SARS-CoV-2 immune imprinting. *Nature* **625**, 148–156 (2024).
361. Tortorici, M. A. *et al.* Persistent immune imprinting occurs after vaccination with the COVID-19 XBB.1.5 mRNA booster in humans. *Immunity* **57**, 904-911.e4 (2024).
362. Liang, C.-Y. *et al.* Imprinting of serum neutralizing antibodies by Wuhan-1 mRNA vaccines. *Nature* 1–3 (2024) doi:10.1038/s41586-024-07539-1.
363. Starr, T. N. *et al.* Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding. *Cell* **182**, 1295-1310.e20 (2020).
364. Greaney, A. J. *et al.* Mapping mutations to the SARS-CoV-2 RBD that escape binding by different classes of antibodies. *Nat Commun* **12**, 4196 (2021).
365. Thorne, L. G. *et al.* Evolution of enhanced innate immune evasion by SARS-CoV-2. *Nature* **602**, 487–495 (2022).
366. Alfi, O. *et al.* SARS-CoV-2 Omicron Induces Enhanced Mucosal Interferon Response Compared to other Variants of Concern, Associated with Restricted Replication in Human Lung Tissues. *Viruses* **14**, 1583 (2022).
367. Bouhaddou, M. *et al.* SARS-CoV-2 variants evolve convergent strategies to remodel the host response. *Cell* **186**, 4597-4614.e26 (2023).
368. Reuschl, A.-K. *et al.* Evolution of enhanced innate immune suppression by SARS-CoV-2 Omicron subvariants. *Nat Microbiol* **9**, 451–463 (2024).
369. Kehrer, T. *et al.* Impact of SARS-CoV-2 ORF6 and its variant polymorphisms on host responses and viral pathogenesis. *Cell Host & Microbe* **31**, 1668-1684.e12 (2023).
370. SARS-CoV-2 mutations. http://sarscov2-mutation-portal.urv.cat/SARS-CoV-2_mutation-portal/mut.php.
371. Surjit, M. *et al.* The Severe Acute Respiratory Syndrome Coronavirus Nucleocapsid Protein Is Phosphorylated and Localizes in the Cytoplasm by 14-3-3-Mediated Translocation. *J Virol* **79**, 11476–11486 (2005).
372. Ricciardi, S. *et al.* The role of NSP6 in the biogenesis of the SARS-CoV-2 replication organelle. *Nature* **606**, 761–768 (2022).
373. Bills, C. J. *et al.* Mutations in SARS-CoV-2 variant nsp6 enhance type-I interferon antagonism. *Emerg Microbes Infect* **12**, 2209208.
374. Tarke, A. *et al.* SARS-CoV-2 vaccination induces immunological T cell memory able to cross-recognize variants from Alpha to Omicron. *Cell* (2022) doi:10.1016/j.cell.2022.01.015.
375. Quadeer, A. A., Ahmed, S. F. & McKay, M. R. Landscape of epitopes targeted by T cells in 852 individuals recovered from COVID-19: Meta-analysis, immunoprevalence, and web platform. *Cell Rep Med* **2**, 100312 (2021).
376. de Silva, T. I. *et al.* The impact of viral mutations on recognition by SARS-CoV-2 specific T cells. *iScience* **24**, 103353 (2021).
377. Agerer, B. *et al.* SARS-CoV-2 mutations in MHC-I-restricted epitopes evade CD8+ T cell responses. *Sci Immunol* **6**, eabg6461 (2021).
378. Taha, T. Y. *et al.* Rapid assembly of SARS-CoV-2 genomes reveals attenuation of the Omicron BA.1 variant through NSP6. *Nat Commun* **14**, 2308 (2023).
379. Khalid, M. M. *et al.* Regulation of virion production by the ORF8 signal peptide across SARS-CoV-2 variants. 2024.03.05.583578 Preprint at <https://doi.org/10.1101/2024.03.05.583578> (2024).
380. Kim, I.-J. *et al.* SARS-CoV-2 protein ORF8 limits expression levels of Spike antigen and facilitates immune evasion of infected host cells. *J Biol Chem* **299**, 104955 (2023).
381. Syed, A. M. *et al.* Rapid assessment of SARS-CoV-2 evolved variants using virus-like particles. *Science* **0**, eabl6184 (2021).
382. Forni, D., Cagliani, R., Clerici, M. & Sironi, M. Molecular Evolution of Human Coronavirus Genomes. *Trends Microbiol* **25**, 35–48 (2017).
383. Woo, P. C. Y. *et al.* Comparative Analysis of 22 Coronavirus HKU1 Genomes Reveals a Novel Genotype and Evidence of Natural Recombination in Coronavirus HKU1. *J Virol* **80**, 7136–7145

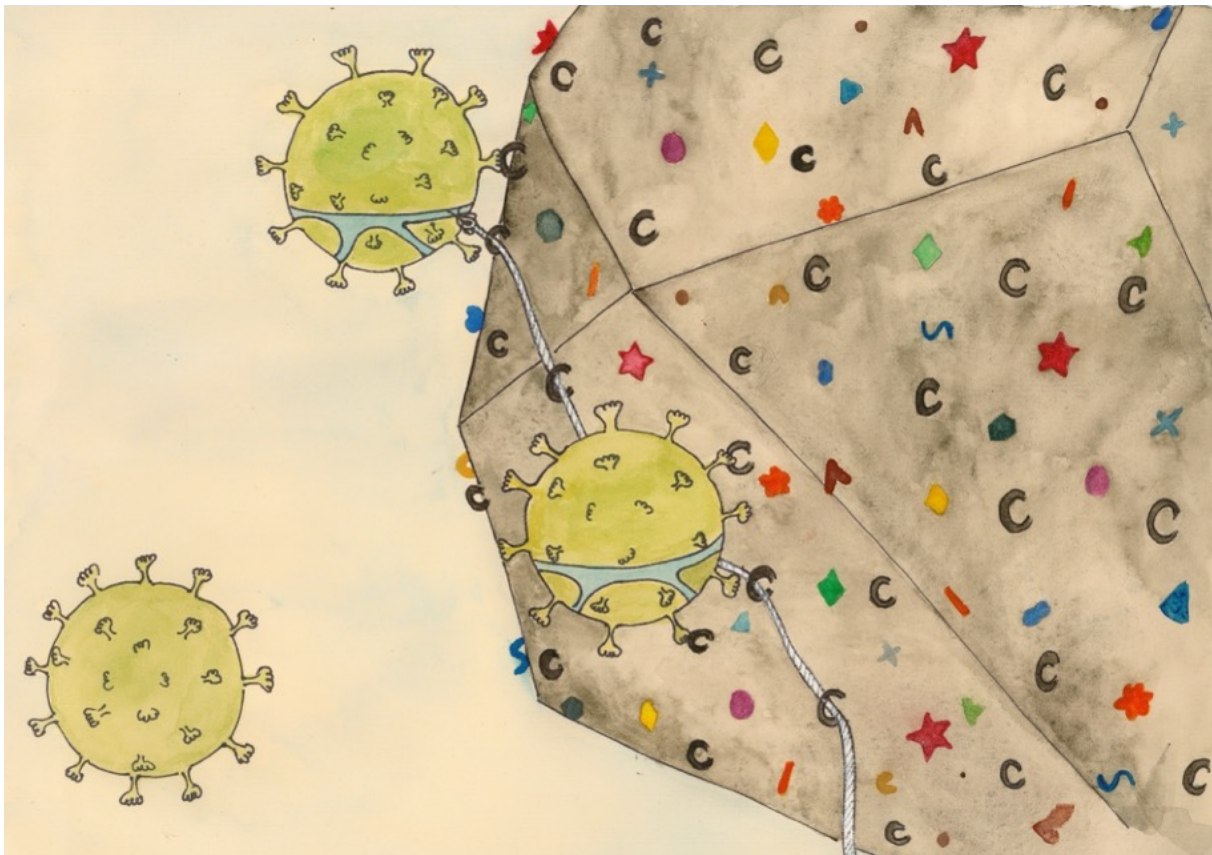
(2006).

384. Otieno, J. R., Cherry, J. L., Spiro, D. J., Nelson, M. I. & Trovao, N. S. Origins and Evolution of Seasonal Human Coronaviruses. *Viruses* **14**, (2022).
385. Li, P. *et al.* Recapitulating infection, thermal sensitivity and antiviral treatment of seasonal coronaviruses in human airway organoids. *eBioMedicine* **81**, 104132 (2022).
386. Hooper, J. D., Clements, J. A., Quigley, J. P. & Antalis, T. M. Type II Transmembrane Serine Proteases: Insights into an emerging class of cell surface proteolytic enzymes. *Journal of Biological Chemistry* **276**, 857–860 (2001).
387. The human Protein Atlas. <https://www.proteinatlas.org/>.
388. Wilson, S. *et al.* The membrane-anchored serine protease, TMPRSS2, activates PAR-2 in prostate cancer cells. *Biochem J* **388**, 967–972 (2005).
389. Lucas, J. M. *et al.* The Androgen-Regulated Protease TMPRSS2 Activates a Proteolytic Cascade Involving Components of the Tumor Microenvironment and Promotes Prostate Cancer Metastasis. *Cancer Discov* **4**, 1310–1325 (2014).
390. Ko, C.-J. *et al.* Androgen-Induced TMPRSS2 Activates Matrilysin and Promotes Extracellular Matrix Degradation, Prostate Cancer Cell Invasion, Tumor Growth, and Metastasis. *Cancer Res* **75**, 2949–2960 (2015).
391. Delliaux, C. *et al.* *TMPRSS2:ERG* gene fusion expression regulates bone markers and enhances the osteoblastic phenotype of prostate cancer bone metastases. *Cancer Letters* **438**, 32–43 (2018).
392. Brandi, F. *et al.* High concordance of TMPRSS-ERG fusion between primary prostate cancer and its lymph node metastases. *Oncology Letters* **16**, 6238–6244 (2018).
393. Daniel E. H. Afar *et al.* Catalytic Cleavage of the Androgen-regulated TMPRSS2 Protease Results in Its Secretion by Prostate and Prostate Cancer Epithelia. *cancer research* **31**, 1686–1692 (2001).
394. Zmora, P., Moldenhauer, A.-S., Hofmann-Winkler, H. & Pöhlmann, S. TMPRSS2 Isoform 1 Activates Respiratory Viruses and Is Expressed in Viral Target Cells. *PLoS One* **10**, e0138380 (2015).
395. McCallum, M. *et al.* Human coronavirus HKU1 recognition of the TMPRSS2 host receptor. Preprint at <https://doi.org/10.1101/2024.01.09.574565> (2024).
396. Pyrc, K. *et al.* Culturing the Unculturable: Human Coronavirus HKU1 Infects, Replicates, and Produces Progeny Virions in Human Ciliated Airway Epithelial Cell Cultures. *Journal of Virology* **84**, 11255–11263 (2010).
397. Dijkman, R. *et al.* Isolation and Characterization of Current Human Coronavirus Strains in Primary Human Epithelial Cell Cultures Reveal Differences in Target Cell Tropism. *Journal of Virology* **87**, 6081–6090 (2013).
398. Saunders, N. *et al.* TMPRSS2 is a functional receptor for human coronavirus HKU1. *Nature* **624**, 207–214 (2023).
399. Robinot, R. *et al.* SARS-CoV-2 infection induces the dedifferentiation of multiciliated cells and impairs mucociliary clearance. *Nature Communications* **12**, 4354 (2021).
400. Komabayashi, K. *et al.* Isolation of Human Coronaviruses OC43, HKU1, NL63, and 229E in Yamagata, Japan, Using Primary Human Airway Epithelium Cells Cultured by Employing an Air-Liquid Interface Culture. *Jpn J Infect Dis* **74**, 285–292 (2021).
401. Dominguez, S. R., Travanty, E. A., Qian, Z. & Mason, R. J. Human Coronavirus HKU1 Infection of Primary Human Type II Alveolar Epithelial Cells: Cytopathic Effects and Innate Immune Response. *PLoS ONE* **8**, e70129 (2013).
402. V'kovski, P. *et al.* Disparate temperature-dependent virus-host dynamics for SARS-CoV-2 and SARS-CoV in the human respiratory epithelium. *PLoS Biol* **19**, e3001158 (2021).
403. Brandolini, M. *et al.* Correlating qRT-PCR, dPCR and Viral Titration for the Identification and Quantification of SARS-CoV-2: A New Approach for Infection Management. *Viruses* **13**, 1022 (2021).
404. Zapata-Cardona, M. I. *et al.* Comparison among plaque assay, tissue culture infectious dose (TCID₅₀) and real-time RT-PCR for SARS-CoV-2 variants quantification. *Iran J Microbiol* **14**, 291–299 (2022).
405. Väisänen, E. *et al.* Infectious viruses from transfected SARS-CoV-2 genomic RNA. *Front Bioeng Biotechnol* **11**, 1129111 (2023).
406. Goddard, T. D. *et al.* UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci* **27**, 14–25 (2018).

407. Vijgen, L. *et al.* Complete Genomic Sequence of Human Coronavirus OC43: Molecular Clock Analysis Suggests a Relatively Recent Zoonotic Coronavirus Transmission Event. *J Virol* **79**, 1595–1604 (2005).
408. Nguyen, T. L. *et al.* Quantitative Phase Imaging: Recent Advances and Expanding Potential in Biomedicine. *ACS Nano* **16**, 11516–11544 (2022).
409. Palonpon, A. F., Sodeoka, M. & Fujita, K. Molecular imaging of live cells by Raman microscopy. *Current Opinion in Chemical Biology* **17**, 708–715 (2013).
410. Smith, R., L. Wright, K. & Ashton, L. Raman spectroscopy: an evolving technique for live cell studies. *Analyst* **141**, 3590–3600 (2016).
411. Jünger, F. *et al.* 100 Hz ROCS microscopy correlated with fluorescence reveals cellular dynamics on different spatiotemporal scales. *Nat Commun* **13**, 1758 (2022).
412. Ortega-Arroyo, J. & Kukura, P. Interferometric scattering microscopy (iSCAT): new frontiers in ultrafast and ultrasensitive optical microscopy. *Phys Chem Chem Phys* **14**, 15625–15636 (2012).
413. Huang, Y.-F. *et al.* Coherent Brightfield Microscopy Provides the Spatiotemporal Resolution To Study Early Stage Viral Infection in Live Cells. *ACS Nano* **11**, 2575–2585 (2017).
414. Ghosh, B. & Agarwal, K. Viewing life without labels under optical microscopes. *Commun Biol* **6**, 1–12 (2023).
415. Croce, A. C. & Bottiroli, G. Autofluorescence Spectroscopy and Imaging: A Tool for Biomedical Research and Diagnosis. *European Journal of Histochemistry : EJH* **58**, (2014).
416. Campagnola, P. j. & Dong, C.-Y. Second harmonic generation microscopy: principles and applications to disease diagnosis. *Laser & Photonics Reviews* **5**, 13–26 (2011).
417. Moon, I. & Javidi, B. Volumetric three-dimensional recognition of biological microorganisms using multivariate statistical method and digital holography. *J Biomed Opt* **11**, 064004 (2006).
418. Memmolo, P. *et al.* Identification and classification of biological micro-organisms by holographic learning. in *Optical Methods for Inspection, Characterization, and Imaging of Biomaterials IV* vol. 11060 51–56 (SPIE, 2019).
419. Lam, V. K. *et al.* Machine Learning with Optical Phase Signatures for Phenotypic Profiling of Cell Lines. *Cytometry A* **95**, 757–768 (2019).
420. Lam, V. *et al.* Quantitative scoring of epithelial and mesenchymal qualities of cancer cells using machine learning and quantitative phase imaging. *J Biomed Opt* **25**, 1–17 (2020).
421. Uttam, S. *et al.* Early Prediction of Cancer Progression by Depth-Resolved Nanoscale Mapping of Nuclear Architecture from Unstained Tissue Specimens. *Cancer Research* **75**, 4718–4727 (2015).
422. Majeed, H., Okoro, C., Kajdacsy-Balla, A., Toussaint, K. C. & Popescu, G. Quantifying collagen fiber orientation in breast cancer using quantitative phase imaging. *JBO* **22**, 046004 (2017).
423. Takabayashi, M., Majeed, H., Kajdacsy-Balla, A. & Popescu, G. Disorder strength measured by quantitative phase imaging as intrinsic cancer marker in fixed tissue biopsies. *PLOS ONE* **13**, e0194320 (2018).
424. Barker, K. L., Boucher, K. M. & Judson-Torres, R. L. Label-Free Classification of Apoptosis, Ferroptosis and Necroptosis Using Digital Holographic Cytometry. *Applied Sciences* **10**, 4439 (2020).
425. Butola, A. *et al.* High spatially sensitive quantitative phase imaging assisted with deep neural network for classification of human spermatozoa under stressed condition. *Sci Rep* **10**, 13118 (2020).
426. Javidi, B. *et al.* Sickle cell disease diagnosis based on spatio-temporal cell dynamics analysis using 3D printed shearing digital holographic microscopy. *Opt Express* **26**, 13614–13627 (2018).
427. Go, T., Kim, J. H., Byeon, H. & Lee, S. J. Machine learning-based in-line holographic sensing of unstained malaria-infected red blood cells. *J Biophotonics* **11**, e201800101 (2018).
428. Guo, S.-M. *et al.* Revealing architectural order with quantitative label-free imaging and deep learning. *eLife* **9**, e55502 (2020).
429. Rivenson, Y. *et al.* PhaseStain: the digital staining of label-free quantitative phase microscopy images using deep learning. *Light Sci Appl* **8**, 23 (2019).
430. Rivenson, Y. *et al.* Virtual histological staining of unlabelled tissue-autofluorescence images via deep learning. *Nat Biomed Eng* **3**, 466–477 (2019).
431. Bai, B. *et al.* Deep learning-enabled virtual histological staining of biological samples. *Light Sci Appl* **12**, 57 (2023).
432. Kandel, M. E. *et al.* Phase imaging with computational specificity (PICS) for measuring dry mass changes in sub-cellular compartments. *Nat Commun* **11**, 6256 (2020).









433. Guo, S., Ma, Y., Pan, Y., Smith, Z. J. & Chu, K. Organelle-specific phase contrast microscopy enables gentle monitoring and analysis of mitochondrial network dynamics. *Biomed Opt Express* **12**, 4363–4379 (2021).
434. Somani, A. *et al.* Virtual labeling of mitochondria in living cells using correlative imaging and physics-guided deep learning. *Biomed Opt Express* **13**, 5495–5516 (2022).
435. Fang, J., Zhang, H., Pan, Y., Smith, Z. J. & Chu, K. Label-Free Analysis of Organelle Interactions Using Organelle-Specific Phase Contrast Microscopy (OS-PCM). *ACS Photonics* **10**, 1093–1103 (2023).
436. Ibrahim, K. A. AI-driven detection and analysis of label-free protein aggregates. *Nat Rev Mol Cell Biol* **25**, 337–337 (2024).
437. Sandoz, P. A., Tremblay, C., van der Goot, F. G. & Frechin, M. Image-based analysis of living mammalian cells using label-free 3D refractive index maps reveals new organelle dynamics and dry mass flux. *PLoS Biol* **17**, e3000553 (2019).
438. Choi, J. *et al.* Label-free three-dimensional analyses of live cells with deep-learning-based segmentation exploiting refractive index distributions. (2021) doi:10.1101/2021.05.23.445351.
439. Ou, X. *et al.* Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. *Nature Communications* **11**, 1620 (2020).
440. Dias, S. S. G. *et al.* Lipid droplets fuel SARS-CoV-2 replication and production of inflammatory mediators. *PLoS Pathogens* **16**, e1009127 (2020).
441. Nardacci, R. *et al.* Evidences for lipid involvement in SARS-CoV-2 cytopathogenesis. *Cell Death Dis* **12**, 1–12 (2021).
442. Wang, W. *et al.* Genetic variety of ORF3a shapes SARS-CoV-2 fitness through modulation of lipid droplet. *Journal of Medical Virology* **95**, e28630 (2023).
443. Cortese, M. *et al.* Integrative Imaging Reveals SARS-CoV-2-Induced Reshaping of Subcellular Morphologies. *Cell Host Microbe* **28**, 853–866 e5 (2020).
444. Castro, V., Pérez-Berna, A. J., Calvo, G., Pereiro, E. & Gastaminza, P. Three-Dimensional Remodeling of SARS-CoV2-Infected Cells Revealed by Cryogenic Soft X-ray Tomography. *ACS Nano* **17**, 22708–22721 (2023).
445. Mendonça, L. *et al.* Correlative multi-scale cryo-imaging unveils SARS-CoV-2 assembly and egress. *Nat Commun* **12**, 4629 (2021).
446. Ramachandran, K. *et al.* SARS-CoV-2 infection enhances mitochondrial PTP complex activity to perturb cardiac energetics. *iScience* **25**, 103722 (2022).
447. Shin, H. J. *et al.* SARS-CoV-2 aberrantly elevates mitochondrial bioenergetics to induce robust virus propagation. *Sig Transduct Target Ther* **9**, 1–14 (2024).

Appendix



Appendix 1

SARS-CoV-2 Alpha, Beta, and Delta variants display enhanced Spike-mediated syncytia formation

Maaran Michael Rajah^{1,2} , Mathieu Hubert^{1,3,†}, Elodie Bishop^{1,4,†}, Nell Saunders^{1,†} , Remy Robinot¹, Ludivine Grzelak^{1,2}, Delphine Planas^{1,3}, Jérémy Dufloo^{1,2} , Stacy Gellenoncourt^{1,2} , Alice Bongers^{1,2}, Marija Zivaljic^{5,6}, Cyril Planchais⁷, Florence Guivel-Benhassine¹, Françoise Porrot¹, Hugo Mouquet⁷ , Lisa A Chakrabarti¹ , Julian Buchrieser^{1,*;‡}  & Olivier Schwartz^{1,2,**;‡} 

Abstract

Severe COVID-19 is characterized by lung abnormalities, including the presence of syncytial pneumocytes. Syncytia form when SARS-CoV-2 spike protein expressed on the surface of infected cells interacts with the ACE2 receptor on neighboring cells. The syncytia forming potential of spike variant proteins remain poorly characterized. Here, we first assessed Alpha (B.1.1.7) and Beta (B.1.351) spread and fusion in cell cultures, compared with the ancestral D614G strain. Alpha and Beta replicated similarly to D614G strain in Vero, Caco-2, Calu-3, and primary airway cells. However, Alpha and Beta formed larger and more numerous syncytia. Variant spike proteins displayed higher ACE2 affinity compared with D614G. Alpha, Beta, and D614G fusion was similarly inhibited by interferon-induced transmembrane proteins (IFITMs). Individual mutations present in Alpha and Beta spikes modified fusogenicity, binding to ACE2 or recognition by monoclonal antibodies. We further show that Delta spike also triggers faster fusion relative to D614G. Thus, SARS-CoV-2 emerging variants display enhanced syncytia formation.

Keywords coronavirus; fusion; SARS-CoV-2; spike; syncytia

Subject Categories Immunology; Microbiology, Virology & Host Pathogen Interaction

DOI 10.15252/embj.2021108944 | Received 10 June 2021 | Revised 27 September 2021 | Accepted 28 September 2021 | Published online 25 October 2021

The EMBO Journal (2021) 40: e108944

See also: **J Koch et al** (December 2021)

Introduction

SARS-CoV-2 was initially discovered during an outbreak in Wuhan, China, before it became pandemic (Huang *et al*, 2020a). Since its emergence, the ancestral Wuhan strain has been supplanted by variants harboring a variety of mutations. Several of these mutations occur in the highly antigenic Spike (S) protein which endowed many of the variants with the ability to evade part of the neutralizing antibody response (Weisblum *et al*, 2020; Planas *et al*, 2021a; Liu *et al*, 2021b; Rees-Spear *et al*, 2021; Starr *et al*, 2021). Individual amino acid changes in the S protein also affect viral fitness. One of the earliest identified variants contained the D614G mutation in S protein, which increased infectivity without significantly altering antibody neutralization (Yurkovetskiy *et al*, 2020). Several other variants have since emerged and have become globally dominant, including Alpha (B.1.1.7) first identified in the United Kingdom, Beta (B.1.351) identified in South Africa, Gamma (P.1 & P.2) identified in Brazil, and Delta (B.1.617.2) identified in India (preprint: Tegally *et al*, 2020; Buss *et al*, 2021; Frampton *et al*, 2021; Planas *et al*, 2021b; Sabino *et al*, 2021; preprint: Yadav *et al*, 2021). Some variants are more transmissible but their impact on disease severity is debated (Korber *et al*, 2020; Davies *et al*, 2021; Meng *et al*, 2021).

Clinically, SARS-CoV-2 infections range from asymptomatic or febrile respiratory disorders to severe lung injury characterized by vascular thrombosis and alveolar damage (Bussani *et al*, 2020). The deterioration of respiratory tissue is likely a result of both virus-induced cytopathicity and indirect immune-mediated damage (Buchrieser *et al*, 2020; Zhang *et al*, 2020; Zhou *et al*, 2020; Zhu *et al*, 2020). A peculiar dysmorphic cellular feature is the presence of large infected multinucleated syncytia, predominately comprised of pneumocytes (Bussani *et al*, 2020; Braga *et al*, 2021; Sanders

1 Virus & Immunity Unit, Department of Virology, Institut Pasteur, CNRS UMR 3569, Paris, France

2 Université de Paris, Sorbonne Paris Cité, Paris, France

3 Vaccine Research Institute, Creteil, France

4 Sorbonne Université, Paris, France

5 Integrative Neurobiology of Cholinergic Systems, Department of Neuroscience, Institut Pasteur, CNRS UMR 3571, Paris, France

6 Sorbonne Université, ED3C "Brain, Cognition, Behavior", Paris, France

7 Laboratory of Humoral Immunology, Department of Immunology, Institut Pasteur, INSERM U1222, Paris, France

*Corresponding author. Tel: +33 1 45 68 85 76; E-mail: julian.buchrieser@pasteur.fr


**Corresponding author. Tel: +33 1 45 68 83 53; E-mail: olivier.schwartz@pasteur.fr

†These authors contributed equally to this work as second authors

‡These authors contributed equally to this work as last authors

Appendix 2

The Spike-Stabilizing D614G Mutation Interacts with S1/S2 Cleavage Site Mutations To Promote the Infectious Potential of SARS-CoV-2 Variants

Stacy Gellenoncourt,^a Nell Saunders,^a Rémy Robinot,^a Lucas Auguste,^{a,b} Maaran Michael Rajah,^a Jérôme Kervevan,^a Raphaël Jeger-Madiot,^a Isabelle Staropoli,^a Cyril Planchais,^c Hugo Mouquet,^c Julian Buchrieser,^a Olivier Schwartz,^{a,d}  Lisa A. Chakrabarti^a

^aVirus & Immunity Unit, Institut Pasteur, Université Paris Cité, CNRS UMR 3569, Paris, France

^bUniversité de Nîmes, Nîmes, France

^cLaboratory of Humoral Immunology, Institut Pasteur, Université Paris Cité, Paris, France

^dVaccine Research Institute, Créteil, France

ABSTRACT Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) remained genetically stable during the first 3 months of the pandemic, before acquiring a D614G spike mutation that rapidly spread worldwide and then generating successive waves of viral variants with increasingly high transmissibility. We set out to evaluate possible epistatic interactions between the early-occurring D614G mutation and the more recently emerged cleavage site mutations present in spike of the Alpha, Delta, and Omicron variants of concern. The P681H/R mutations at the S1/S2 cleavage site increased spike processing and fusogenicity but limited its incorporation into pseudoviruses. In addition, the higher cleavage rate led to higher shedding of the spike S1 subunit, resulting in a lower infectivity of the P681H/R-carrying pseudoviruses compared to those expressing the Wuhan wild-type spike. The D614G mutation increased spike expression at the cell surface and limited S1 shedding from pseudovirions. As a consequence, the D614G mutation preferentially increased the infectivity of P681H/R-carrying pseudoviruses. This enhancement was more marked in cells where the endosomal route predominated, suggesting that more stable spikes could better withstand the endosomal environment. Taken together, these findings suggest that the D614G mutation stabilized S1/S2 association and enabled the selection of mutations that increased S1/S2 cleavage, leading to the emergence of SARS-CoV-2 variants expressing highly fusogenic spikes.

IMPORTANCE The first SARS-CoV-2 variant that spread worldwide in early 2020 carried a D614G mutation in the viral spike, making this protein more stable in its cleaved form at the surface of virions. The Alpha and Delta variants, which spread in late 2020 and early 2021, respectively, proved increasingly transmissible and pathogenic compared to the original strain. Interestingly, Alpha and Delta both carried the mutations P681H/R in a cleavage site that made the spike more cleaved and more efficient at mediating viral fusion. We show here that variants with increased spike cleavage due to P681H/R were even more dependent on the stabilizing effect of the D614G mutation, which limited the shedding of cleaved S1 subunits from viral particles. These findings suggest that the worldwide spread of the D614G mutation was a prerequisite for the emergence of more pathogenic SARS-CoV-2 variants with highly fusogenic spikes.

KEYWORDS SARS-CoV-2, epistasis, proteolytic cleavage, spike, variants

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) pandemic has remained a major public health issue during the past 2 years, due to the repeated emergence of viral variants endowed with increased transmissibility and/or immune

Editor Tom Gallagher, Loyola University Chicago

Copyright © 2022 American Society for Microbiology. All Rights Reserved.

Address correspondence to Lisa A. Chakrabarti, chakra@pasteur.fr.

The authors declare no conflict of interest.

Received 19 August 2022

Accepted 30 August 2022

Published 19 September 2022

Appendix 3

NOUVELLE

TMPRSS2 est le récepteur cellulaire du coronavirus saisonnier HKU1

Nell Saunders, Olivier Schwartz

Institut Pasteur, Université Paris Cité, CNRS UMR3569, Unité virus & immunité, Paris, France.
nell.saunders@pasteur.fr

> À ce jour, sept coronavirus capables d'infecter les êtres humains ont été identifiés. Trois d'entre eux, SARS-CoV (identifié en 2003), MERS-CoV (identifié en 2012) et SARS-CoV-2 (identifié en 2019) sont hautement pathogènes, et ont provoqué trois récentes épidémies ou pandémies. Quant aux quatre coronavirus saisonniers endémiques, 229E, NL63, OC43 et HKU1, ils causent des maladies bénignes dans la majorité des cas (ils sont impliqués dans 15 à 30 % des rhumes), mais peuvent occasionnellement provoquer des infections sévères chez les nourrissons ou les personnes âgées [1].

Le coronavirus HKU1 a été isolé pour la première fois à Hong-Kong en 2005 chez une personne souffrant d'une pneumopathie [2]. Il fait partie du lignage A des β -coronavirus. Le coronavirus le plus proche de HKU1 du point de vue de l'évolution est le virus de l'hépatite murine (*murine hepatitis virus*, MHV) [2]. Les séquences d'acides aminés de la protéine du spicule¹ de ces deux virus ont 65 % d'identité. La date d'émergence de HKU1, son origine animale et les potentiels hôtes intermédiaires sont inconnus. Le taux de séropositivité contre HKU1 est estimé entre 75 % et 95 % dans la population française, ce qui est comparable aux autres coronavirus saisonniers [3]. Il existe trois génotypes de HKU1 (A, B et C). HKU1C est issu d'une recombinaison entre HKU1A et HKU1B, et son spicule est quasiment identique à celui

de HKU1B (99 % d'identité de séquence peptidique) [4], tandis que les spicules de HKU1A et HKU1B ont 85 % d'identité de séquence peptidique.

Le spicule (S) des coronavirus est nécessaire à leur entrée dans les cellules hôtes. La protéine S de HKU1 s'attache aux acides sialiques 9-O-acétylés de la membrane plasmique des cellules hôtes par son domaine N-terminal, ce qui permet une ouverture de la protéine et l'exposition du domaine de liaison à son récepteur (*receptor-binding domain*, RBD) [5, 6]. La structure de la protéine S dans ses conformations fermée et ouverte a été élucidée récemment [6]. L'état fermé protège notamment le RBD de potentiels anticorps neutralisants. Après la liaison à son récepteur, non identifié jusqu'à récemment, la protéine S peut être clivée par des protéases à sérine², soit à la surface de la cellule par la protéase transmembranaire TMPRSS2 (*type II transmembrane serine protease*) ou la kallikréine KLK13, soit dans les endosomes par des cathepsines [7]. D'après les données disponibles pour le virus SARS-CoV-2 et d'autres coronavirus, ce clivage déclenche la projection du peptide de fusion et la fusion de la membrane virale avec la membrane cellulaire (plasmique ou endosomique), ce qui permet l'entrée du virus (*Figure 1*). La cellule infectée par un coronavirus produit différentes protéines virales, dont la protéine S, qu'elle expose à sa surface, et qui entraîne la fusion des cellules

infectées avec les cellules avoisinantes, formant des cellules multinucléées (ou syncytium).

Nous avons récemment montré que TMPRSS2 est un récepteur membranaire fonctionnel pour HKU1A et HKU1B [8].

TMPRSS2 permet l'entrée de pseudovirus exprimant la protéine S de HKU1

L'expression de TMPRSS2 est suffisante pour induire la fusion de cellules exprimant le spicule de HKU1, et pour l'entrée de pseudovirus (lentivirus) exprimant ce spicule. Nous avons montré le rôle de TMPRSS2 dans différents types de cellules, dont celles des lignées HEK293T et U2OS³. Sur un ensemble de quinze protéases impliquées dans l'entrée des coronavirus ou de la même famille que TMPRSS2, seule TMPRSS2 permet la fusion membranaire impliquant la protéine S de HKU1. Les cellules de la lignée Caco2⁴ expriment le gène *TMPRSS2* de façon endogène et fusionnent spontanément lors de l'expression de la protéine S, mais la délétion de *TMPRSS2* empêche la fusion. De plus, ces cellules sont sensibles à l'infection par des pseudovirus de HKU1, mais la délétion de *TMPRSS2* empêche l'entrée de ces pseudovirus [8].

¹ Le spicule désigne une glycoprotéine pointant à la surface de certains virus et qui leur sert de « clé d'entrée » dans les cellules qu'ils infectent.

² Cette catégorie de protéases est ainsi dénommée car leur site actif contient un résidu sérine qui joue un rôle essentiel dans la catalyse.

³ La lignée cellulaire HEK293T a été établie à partir de rein embryonnaire humain, et la lignée U2OS, à partir d'un ostéosarcome humain.

⁴ La lignée cellulaire Caco-2 est dérivée d'un adénocarcinome du côlon humain.



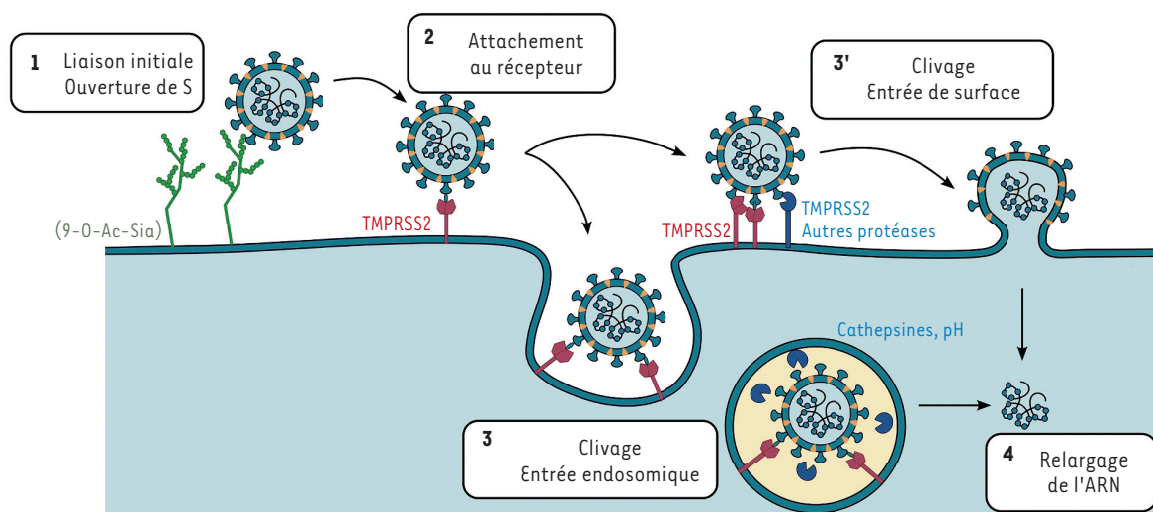


Figure 1. Schéma récapitulatif des deux voies d'entrée du coronavirus HKU1 dans la cellule hôte. Les étapes successives (numérotées de 1 à 4) de l'entrée des virions par la voie de surface ou par la voie des endosomes sont représentées.

L'activité catalytique de TMPRSS2 n'est pas requise pour l'entrée de pseudovirus exprimant la protéine S de HKU1

TMPRSS2 est une protéase à sérine qui devient fonctionnelle par autoclivage [9]. Afin d'exclure la possibilité que le rôle de TMPRSS2 soit dû à son activité enzymatique de clivage de la protéine S et qu'une autre protéine cellulaire serve de récepteur à la protéine S, nous avons produit deux mutants enzymatiquement inactifs de TMPRSS2 : TMPRSS2_{Arg255Gln}, mutée dans le site de clivage, et TMPRSS2_{Ser441Ala}, mutée dans le site catalytique.

Le mutant catalytiquement inactif de TMPRSS2 ne clive pas la protéine S. Son effet sur la fusion cellulaire est réduit par rapport à TMPRSS2 non mutée. Néanmoins, tout comme celle-ci, il permet l'entrée des pseudovirus HKU1 [8]. En effet, le clivage de la protéine S à la surface est nécessaire pour la fusion cellulaire, mais le virus peut pénétrer dans la cellule via les endosomes, dans lesquels d'autres protéases ou l'envi-

ronnement acide pourront induire la projection du peptide de fusion.

TMPRSS2 et la protéine S de HKU1 interagissent directement

Nous avons également montré l'existence d'une interaction directe entre la protéine S de HKU1 et TMPRSS2 par la technique ELISA. L'affinité de cette interaction a été calculée par interférométrie biocouche⁵, en utilisant le domaine RBD de la protéine S de HKU1 afin de s'affranchir de la nécessité de disposer de la protéine entière dans sa conformation ouverte. Nous avons mesuré des affinités (K_d) de 334 nM et de 137 nM respectivement pour HKU1A et HKU1B [8]. Ces valeurs sont plus élevées que celle obtenue pour SARS-CoV-2 et son récepteur ACE2 (enzyme de conversion de l'angiotensine 2) ($K_d = 37$ nM), mais elles restent de l'ordre de grandeur des interactions dites fortes.

⁵ Technique de biodétection optique permettant d'analyser les interactions entre molécules en temps réel, sans nécessiter leur marquage fluorescent.

Une étude structurale avait précédemment conclu à l'importance d'une région particulière du domaine RBD de la protéine S de HKU1, conservée entre HKU1A et HKU1B, dans la liaison à un récepteur alors inconnu, notamment en montrant que cette région était celle ciblée par des anticorps neutralisant HKU1 [10]. Nous avons montré que cette région était en effet impliquée dans la liaison à TMPRSS2 en mutant deux acides aminés, Trp515 et Arg517 : les deux protéines S mutées ne permettaient plus la fusion cellulaire ou l'entrée des pseudovirus, et avaient une affinité réduite ou non détectable pour TMPRSS2 [8].

L'interaction est bloquée par les nanocorps anti-TMPRSS2

Nous avons aussi isolé des nanocorps⁶ anti-TMPRSS2 après immunisation d'un alpaga. Trois d'entre eux bloquaient l'interaction de TMPRSS2 avec

⁶ Un nanocorps, ou anticorps à domaine unique, est un fragment d'anticorps composé d'un seul domaine variable d'anticorps monomère. Comme un anticorps entier, il est capable de se lier sélectivement à un antigène spécifique. Ces nanocorps sont produits naturellement par les Camélidés.



HKU1, ainsi que la fusion cellulaire et l'entrée des pseudovirus dans les cellules hôtes. Enfin, nous avons isolé du virus HKU1B infectieux à partir de prélèvements nasopharyngés d'une personne infectée, sur une culture primaire de cellules bronchiques différenciées en interface air-liquide. Ces cellules expriment TMPRSS2 sur leurs cils. L'infection de ces cellules par HKU1 était inhibée efficacement par l'un des nanocorps dirigés contre cette protéine [8].

Perspectives

L'ensemble des résultats présentés permet de conclure que TMPRSS2 est le récepteur du coronavirus HKU1 dans l'espèce humaine. Décrypter les voies d'entrée cellulaire des différents coronavirus permet de lutter plus efficacement contre l'infection par des thérapies ciblées. Cela permet également d'isoler et de caractériser plus rapidement de nouveaux coronavirus ou variants des coronavirus connus qui pourraient infecter l'homme. Les résultats de deux études concernant la structure du complexe formé par HKU1 et TMPRSS2

sont en cours de publication [11, 12]. Ils éclairent le mécanisme de l'entrée de HKU1, l'action d'anticorps neutralisants, et l'utilisation, par HKU1, de TMPRSS2 dans d'autres espèces animales qui pourraient être sensibles à l'infection par ce coronavirus. Par ailleurs, il conviendra d'examiner si TMPRSS2 joue également le rôle de récepteur pour d'autres virus, car cette protéase clive non seulement la protéine S des coronavirus, mais aussi les enveloppes d'autres virus respiratoires, tels que les virus influenza et para-influenza, les paramyxovirus, les métapneumovirus et le virus Sendai [13]. ♦

TMPRSS2 is the receptor of seasonal coronavirus HKU1

LIENS D'INTÉRÊT

Les auteurs déclarent n'avoir aucun lien d'intérêt concernant les données publiées dans cet article.

RÉFÉRENCES

1. Liu DX, Liang JQ, Fung TS. Human coronavirus-229E, -OC43, -NL63, and -HKU1 (coronaviridae). *Encyclopedia of virology* 2021 ; 428-40.
2. Woo PCY, Lau SKP, Chu C, et al. Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. *J Virol* 2005 ; 79 : 884-95.

3. De Thoisy A, Woudenberg T, Pelleau S, et al. Seroepidemiology of the seasonal human coronaviruses NL63, 229E, OC43 and HKU1 in France. *Open Forum Infect Dis* 2023 ; 10 : ofad340.
4. Woo PCY, Lau SKP, Yip CCY, et al. Comparative analysis of 22 coronavirus HKU1 genomes reveals a novel genotype and evidence of natural recombination in coronavirus HKU1. *J Virol* 2006 ; 80 : 7136-45.
5. Hulswit RJG, Lang Y, Bakkers MJG, et al. Human coronaviruses OC43 and HKU1 bind to 9-O-acetylated sialic acids via a conserved receptor-binding site in spike protein domain A. *Proc Natl Acad Sci USA* 2019 ; 116 : 2681-90.
6. Pronker MF, Creutzmacher R, Drulyte I, et al. Sialoglycan binding triggers spike opening in a human coronavirus. *Nature* 2023 ; 624 : 201-6.
7. Shirato K, Kawase M, Matsuyama S. Wild-type human coronaviruses prefer cell-surface TMPRSS2 to endosomal cathepsins for cell entry. *Virology* 2018 ; 517 : 9-15.
8. Saunders N, Fernandez I, Planchais C, et al. TMPRSS2 is a functional receptor for human coronavirus HKU1. *Nature* 2023 ; 624 : 207-14.
9. Afar DEH, Vivanco I, Hubert RS, et al. Catalytic cleavage of the androgen-regulated TMPRSS2 protease results in its secretion by prostate and prostate cancer epithelia. *Cancer Res* 2001 ; 31 : 1686-92.
10. Ou X, Guan H, Qin B, et al. Crystal structure of the receptor binding domain of the spike glycoprotein of human betacoronavirus HKU1. *Nat Commun* 2017 ; 8 : 15216.
11. McCallum M, Park YJ, Stewart C, et al. Human coronavirus HKU1 recognition of the TMPRSS2 host receptor. *BioRxiv* 2024,01.09.574565.
12. Fernández I, Saunders N, Duquerry S, et al. Structural basis of TMPRSS2 zymogen activation and recognition by the HKU1 seasonal coronavirus. *BioRxiv* 2024,02.21.581378.
13. Böttcher-Friebertshäuser E, Garten W, Klenk HD. *Activation of viruses by host proteases*. Springer Cham, 2018 : 348 p.



Avec m/s, vivez en direct les progrès et débats de la biologie et de la médecine

CHAQUE MOIS / AVEC LES ARTICLES DE RÉFÉRENCE DE M/S
CHAQUE JOUR / SUR WWW.MEDECINESCIENCES.ORG

Abonnez-vous sur
www.medecinesciences.org