SARS-CoV-2-related bat virus behavior in human-relevant models sheds light on the origin of COVID-19

Sarah Temmam, Xavier Montagutelli, Cécile Herate, Flora Donati, Béatrice Regnault, Mikael Attia, Eduard Baquero Salazar, Delphine Chretien, Laurine Conquet, Grégory Jouvion, et al.

SARS-CoV-2-RELATED BAT VIRUS BEHAVIOR IN HUMAN-RELEVANT MODELS SHEDS LIGHT ON THE ORIGIN OF COVID-19

Sarah Temmam*[1,2], Xavier Montagutelli*[3], Cecile Herate*[4], Flora Donati*[5,6], Beatrice Regnault[1,2], Mikael Attia[5], Eduard Baquero Salazar[7], Delphine Chretien[1,2], Laurine Conquet[3], Grégory Jouvion[8,9], Juliana Pipoli Da Fonseca[10], Thomas Cokelaer[10], Faustine Amara[5], Francis Relouzat[4], Thibaut Naninck[4], Julien Lemaitre[4], Nathalie Derreudre-Bosquet[4], Quentin Pascal[4], Massimiliano Bonomi[11], Thomas Bigot[1,12], Sandie Munier[5], Felix A. Rey[7], Roger Le Grand[4], Sylvie van der Werf[5,6], Marc Eloit[1,2,13,+]

*These authors contributed equally to this work.

[+] Corresponding author: Institut Pasteur, Paris, France. marc.eloit@pasteur.fr

1 Institut Pasteur, Université Paris Cité, Pathogen Discovery Laboratory, Paris, France.

2 Institut Pasteur, Université Paris Cité, The OIE Collaborating Center for the detection and identification in humans of emerging animal pathogens, Paris, France.

3 Institut Pasteur, Université Paris Cité, Mouse Genetics Laboratory, Paris, France.

4 Center for Immunology of Viral, Auto-immune, Hematological and Bacterial Diseases (IMVA-HB/IDMIT), Université Paris-Saclay, Inserm, CEA, Fontenay-aux-Roses, France.

5 Institut Pasteur, Université Paris Cité, CNRS UMR 3569, Molecular Genetics of RNA Viruses Unit, Paris, France.

6 Institut Pasteur, Université Paris Cité, National Reference Center for Respiratory Viruses, Paris, France.

7 Institut Pasteur, Université Paris Cité, CNRS UMR 3569, Structural Virology Unit, Paris, France.

8 Ecole Nationale Vétérinaire d'Alfort, Unité d'Histologie et d'Anatomie Pathologique, Maisons-Alfort, France.

9 Université Paris Est Créteil, EnvA, ANSES, Unité DYNAMYC, Créteil, France.

10 Biomics Platform, C2RT, Institut Pasteur, Université Paris Cité, Paris, France.

11 Institut Pasteur, Université Paris Cité, CNRS UMR 3528, Structural Bioinformatics Unit, Paris, France.

12 Institut Pasteur, Université Paris Cité, Bioinformatic and Biostatistic Hub - Computational Biology Department, Paris, France.

13 Ecole Nationale Vétérinaire d'Alfort, University of Paris-Est, Maisons-Alfort, France.

Abstract:

Bat sarbecovirus BANAL-236 is highly related to SARS-CoV-2 and infects human cells, albeit lacking the furin cleavage site in its spike protein. BANAL-236 replicates efficiently and pauci-symptomatically in humanized mice and in macaques, where its tropism is enteric, strongly differing from that of SARS-CoV-2. BANAL-236 infection leads to protection against superinfection by a virulent strain. We find no evidence of antibodies recognizing bat sarbecoviruses in populations in close contact with bats in which the virus was identified, indicating that such spillover infections, if they occur, are rare. Six passages in humanized mice or in human intestinal cells, mimicking putative early spillover events, select adaptive mutations without appearance of a furin cleavage site and no change in virulence. Therefore, acquisition of a furin site in the spike protein is likely a pre-spillover event that did not occur upon replication of a SARS-CoV-2-like bat virus in humans or other animals. Other hypotheses regarding the origin of the SARS-CoV-2 should therefore be evaluated, including the presence in bats of sarbecoviruses carrying a spike with a furin cleavage site.

Key Words:

Bat coronavirus, animal model, pathogenesis, adaptive mutations, serology.

Introduction

The origin of SARS-CoV-2, as well as its mode of introduction into the human population, is currently unknown. Various SARS-CoV-2-related viruses have been described in *Rhinolophus shameli* (isolated in Cambodia in 2010 (Delaune *et al*, 2021)), *R. pusillus* and *R. malayanus* (China, 2020 and 2019 respectively(Zhou *et al*, 2021)), in *R. acuminatus* (Thailand, 2020(Wacharapluesadee *et al*, 2021)), *R. cornutus* (Japan, 2013 (Murakami *et al*, 2020)) and *R. affinis,* RaTG13 (China, 2013) (Zhou *et al*, 2020; Rahalkar & Bahulikar, 2020). Very recently we reported the discovery in fecal swabs of bats from Laos coronaviruses highly similar to SARS-CoV-2: BANAL-52 in *R. malayanus,* BANAL-103 in *R. pusillus* and BANAL-236 in *R. marshalli*(Temmam *et al*, 2022). These were shown to bind human angiotensin-converting enzyme 2 (hACE2) more efficiently than early strains of SARS-CoV-2, to enter primate cells by an ACE2-dependent mechanism and, for the unique virus that could be isolated (BANAL-236), to replicate in human Caco-2 and Calu-3 cells expressing physiological levels of hACE2(Pommerenke *et al*, 2021). They lack a furin cleavage site that has been shown to reduce the pathogenicity and transmissibility of SARS-CoV-2 in animal models(Johnson *et al*, 2021), suggesting that these bat viruses could be less pathogenic in humans. These results do not rule out these viruses from having initiated the pandemic, as SARS-CoV-2 showed very high adaptability after the pandemic start. The emergence of variants of concern (VOC) paralleled their increase in transmissibility and/or pathogenicity in humans without any decrease of the host range, as evaluated in wild-type mice(Hoffmann *et al*, 2021; Cameroni *et al*, 2021; Halfmann *et al*, 2022). Recently, two publications suggested that the epicenter of the pandemic was the Huanan Seafood Wholesale Market(Worobey *et al*, 2022) and that the source of SARS-CoV-2 was very likely the live mammals sold in the market(Pekar *et al*, 2022).

Our objective was to investigate the possibility that infection of human populations, or more

generally permissive animals, with strains like BANAL-236 with potential low pathogenicity and/or transmissibility occurred prior to the detection of the first COVID-19 clinical cases (end of 2019), preceding adaptation to humans of more transmissible and pathogenic strains. We asked whether such infections could have been detected clinically, could have led to a cycle of human-to-human transmission, and whether adaptive mutations could have occurred in a first phase of silent circulation before the emergence and detection of symptomatic cases.

Results

Bat BANAL-236 replicates in transgenic mice expressing hACE2

In this study, we aimed to analyze whether mice that express hACE2 (K18-hACE2) were susceptible to BANAL-236. To this end, K18-hACE2 mice (on a C57BL/6J background, jax.org/strain/034860) were inoculated with BANAL-236 or SARS-CoV-2 strains (strain BetaCoV/France/IDF0372/2020, GISAID accession number EPI_ISL_406596, referred below as "Wuhan-372"), while BALB/cJRj mice were used as controls (Fig. 1A). These latter were chosen as non-transgenic controls because they tend to be more susceptible to SARS-CoV-2 infection(Leist *et al*, 2020) with a mouse-adapted virus than C57BL/6J mice, therefore increasing the chance to detect symptoms due to BANAL-236 infection.

Mice inoculated with BANAL-236 showed no weight loss at 10^4 or 10^3 PFU (Fig. 1B), while an important but transient weight loss was observed in mice infected with the Wuhan-372 strain at 10^4 PFU. Compared with Wuhan-372, BANAL-236-infected K18-hACE2 mice showed lower average genome copy number and virus titer in the lungs (-1.66 and -1.65 $\log_{10}$, respectively, Fig. 1C). Infectious BANAL-236 virus, however, could not be detected in the liver and spleen of K18-hACE2 mice, and the genome copy numbers were 2.53 $\log_{10}$ (spleen) and 5.53 $\log_{10}$ (liver) lower than in the lungs. Significant differences between the two viruses were not observed (-0.53 $\log_{10}$ in the spleen and -0.44 $\log_{10}$ in the liver between BANAL-236 and Wuhan-372). BALB/c

mice could be infected by BANAL-236, but at lower levels compared with K18-hACE2 mice and without any clinical sign of infection (Fig. 1B, 1C). All mice infected with BANAL-236 developed an antibody response, which neutralized the homologous strain more efficiently than the Wuhan-372 strain (Fig. 1D, day D30; Fig. 2; Appendix Figure S1).

K18-hACE2 mice infected by BANAL-236 at 10^4 PFU were fully protected against lethality and weight loss after challenge with 10^4 PFU of a virulent SARS-CoV-2 strain (hereafter called "Wuhan-D614G", strain hCoV-19/France/GES-1973/2020, GISAID accession number EPI_ISL_414631), while naive K18-hACE2 mice infected by Wuhan-D614G presented a 100% mortality rate at day 8 post-infection (Fig. 1E, logrank test, p=0.00095). At 10^3 PFU, 1 of 6 mice primo-infected by BANAL-236 and challenged by Wuhan-D614G was euthanized due to severe symptoms while 5/6 experienced moderate and transient body weight loss (Appendix Figure S2). This cross-protection (comparison with Wuhan-D614G: logrank test, p=0.0019) was less efficient than that conferred by the Wuhan-372 strain (logrank test, p=0.00095), for which, at the 10^3 PFU dose, all mice survived after challenge (Fig. 1E). The neutralizing antibody responses against BANAL-236 and Wuhan-372 were boosted after the challenge with the Wuhan-D614G strain (Fig. 1D, D45).

Histopathological analysis of the lung at day 3 post-inoculation of K18-hACE2 mice infected with 10^4 PFU revealed less severe lesions in BANAL-236- than in Wuhan-372-infected mice, with interstitial inflammation often centered on blood vessels or bronchioles and endothelial cell injury (Fig. 1F, Appendix Figure S3).

Therefore, BANAL-236 replicates in hACE2 transgenic mice, but its pneumotropism and its pathogenicity are lower than those of the SARS-CoV-2 Wuhan strain.

Pneumotropism of BANAL-236 is not increased by serial passages in hACE2 transgenic mice

We conducted six passages of the initial BANAL-236 strain by inoculating three K18-hACE2 mice intranasally with pooled lung homogenates from the previous passage (Fig. 1G). The pooled lung homogenates from the last passage (P6, 10^4 PFU) were then inoculated into K18-hACE2 mice to compare the pathogenicity of P6 with that of P1. We found no significant differences in body weight variation between the two strains (Fig. 1G). The number of genome copies per g of lung increased upon passages up to 9.7 log10 gc/g at P4 and remained constant during the next passages, never reaching the levels observed for the Wuhan-372 strain at the first passage (11.3 log10 gc/g). We did not detect any viral RNA in the liver of mice inoculated with P6, whereas the viral load was 4.1 log10 gc/g for P1. Genome copy numbers in the spleen decreased by 1.4 $\log_{10}$ between mice infected by P1 and P6. One mouse of three infected by BANAL-236 P6 presented a detectable but non quantifiable viral load in the feces or in the rectum (positive RT-qPCR at a RNA dilution of 1/10 [Ct=39.13] but negative when RNA was not diluted). These results show that the six serial passages in humanized mice did not modify the pneumotropism of BANAL-236 nor steer the virus towards a more virulent phenotype.

BANAL-236 behaves as an enteric virus in macaques

Two cynomolgus macaques (#MF1 and # MF2) were exposed to 6.10^5 PFU of BANAL-236 by the intranasal and intratracheal routes simultaneously and followed for 43 days (Fig. 3A). Body temperature and clinical score remained within normal values but #MF2 showed progressive weight loss (Fig. 3B). None of the animals had detectable genomic viral loads in broncho-alveolar lavage (BAL) and in nasopharyngeal and tracheal samples (swabs), as assessed by RT-qPCR for viral genomic and sub-genomic RNAs, except a low transient blip for the two macaques at D2-3 post-infection in nasopharyngeal fluids (2.9-3.2 $\log_{10}$ copies/mL) and in BALs (4.1-3.4 $\log_{10}$ copies/mL) (Fig. 3C). Viral sub-genomic RNAs (sgRNAs), which are indirect evidence for viral replication, were not detectable in nasopharyngeal swabs nor BALs. In

contrast, the two macaques shed the virus in the feces (rectal swabs) at D11 (3.52 log_{10} copies/mL) for animal #MF1 and from D7 to D14 for animal #MF2 that peaked at D11 at high level (7.84 log_{10} copies/mL). Both animals had undetectable viral loads at D21. Viral sub-genomic RNA peaked at D11 for animal #MF2 (Fig. 3C). Of note, this individual presented progressive weight loss from D3 to D28, suggesting a possible link with viral excretion of BANAL-236 virus in the feces. Results were compared to historical data obtained for SARS-CoV-2 (Wuhan-372 strain) when administered by the same standardized protocol to 33 or 11 macaques at 1.10^6 or 1.10^5 PFU, respectively. The replication profile of SARS-CoV-2 and BANAL-236 was strikingly different, with lung tropism (BAL) and respiratory shedding (tracheal and nasopharyngeal swab samples) being dominant for SARS-CoV-2, and enteric tropism for BANAL-236 (Fig. 3C).

The two macaques infected by BANAL-236 showed very mild pulmonary lesions (Appendix Figure S4) characterized by non-extended ground-glass opacities detected by chest computed tomography (CT score of 1), while macaques infected by SARS-CoV-2 showed CT scores of 5 or more (Fig. 3B)(Maisonnasse *et al*, 2020). BANAL-236-infected animals experienced a transient lymphopenia at D2 or D3 while other hematological parameters remained within normal ranges (Appendix Figure S5). Sera of the immunized macaques exhibited BANAL-236 and cross-reactive SARS-CoV-2 neutralizing responses (Fig. 3B).

Therefore BANAL-236 largely behaves like an enteric virus and displays low pathogenicity in macaques.

Mutations selected upon passages in human cellular or animal models

We attempted to model the human transmission cycles of BANAL-236 by passaging it 6 times in Caco-2 intestine cells and in transgenic mice expressing hACE2, beginning with the strain initially isolated and amplified by 2 passages in Vero E6 cells (Fig. 4A). Several studies reported

significant increases in lung viral load of transgenic mice after a small number of passages(Huang *et al*, 2021; Montagutelli *et al*, 2021; Gu *et al*, 2020). Therefore, we did not passage the virus further, as the objective was not to adapt the virus to human cells but to mimic a worst-case scenario considering the low transmissibility of the virus as informed by previous animal studies for SARS-CoV-2 deleted for the furin cleavage site(Johnson *et al*, 2021), and as developed below for BANAL-236. We did not find evidence for any variation within or around the furin cleavage site despite high vertical sequencing depth (Appendix Table S1). We confirmed with standard PCR that at the same locus no insertion was detectable (Appendix Figure S6).

However, we identified a set of mutations that were not detectable in the original BANAL-236 fecal swab despite a high sequencing depth (Appendix Table S2). Notably, a missense mutation V391I (V395 in SARS-CoV-2 Wuhan strains), already detected after the first passages in Vero E6 cells, became dominant after the first passage in intestinal Caco-2 cells. The same mutation was positively selected in the course of macaque infection (from day 7 to day 14) while being counter-selected in K18-hACE2 mice (Fig. 4B, Appendix Figure S7). This mutation was associated with other mutations appearing during passages in Caco-2 cells (NSP10 N4350K [N4358 in SARS-CoV-2], endoRNAse NSP15 N6448T [N6456 in SARS-CoV-2], ORF3a silent mutation T248 [T248 in SARS-CoV-2]) or during macaque infection (spike NTD S52R [T51 in SARS-CoV-2], NSP14 3'-5' exonuclease H5943Q [H5951 in SARS-CoV-2]), suggesting a possible complementarity of these mutations affecting different proteins or genes.

On the other hand, passaging the virus in transgenic mice selected a cluster of mutations located in the spike (P627L [P631 in SARS-CoV-2], silent T108 [T108 in SARS-CoV-2]); in the RdRp (silent N4998 [N4992 in SARS-CoV-2]; in ORF6 (A12T [A12 in SARS-CoV-2]), and in ORF3a (L101R [L101 in SARS-CoV-2]) that were specific to the mouse model (Fig. 4B). These

mutations were identified after the first passage and became predominant thereafter. Remarkably, these mutations were different from those observed in Caco-2 cells (Fig. 4C), suggesting that they probably do not impact infection pathways shared by transgenic mice and Caco-2 cells. Silent mutations were also identified but the codon used in SARS-CoV-2 for the corresponding residue was most generally the original codon of BANAL-236 sequenced from the bat feces and not the mutated one (Appendix Table S3).

To gain insight into the phenotype of the missense mutations of the spike RBD-V391I [V395 in SARS-CoV-2 Wuhan strains] and in particular its effect on binding to hACE2, we performed all-atom explicit solvent Molecular Dynamics (MD) simulations of the wild type (WT) BANAL-236 RBD/hACE2 complex and the mutant (Appendix Table S4). The MD simulations were initiated from the X-ray crystal structure of the WT complex (PDB code 7PKI)(Temmam *et al*, 2022) and an homology model of the mutant and were extended for a total aggregated simulation time of 6 µs. The MD simulations did not reveal significant differences upon mutation in the stability of the RBD or in the interaction strength with hACE2 in the microsecond timescale (Appendix Figure S8). The FoldX scoring function predicted similar binding affinities for all the RBD-hACE2 complexes studied (Appendix Figure S9). Moreover, persistent hydrogen bonds (Appendix Figure S10) and salt bridges (Appendix Figure S11) observed at the WT RBD-hACE2 interface were conserved in the mutated complex.

Regarding the S1-P627L mutation, this residue belongs to a loop in S1 that shares high amino acid identity with SARS-CoV-2. The loop is at the interface with the NTD and is not resolved in most cryo-EM structures of the spike, except for the Omicron variant(Yin *et al*, 2022), which indicates a high flexibility in this region of the spike. Therefore, we do not expect that the P627L mutation will drastically affect the overall structure and function of the spike besides inducing minor local conformational changes.

Finally, the S52R mutation (T51 in SARS-CoV-2) located in the NTD, which was positively selected during the course of macaque infection, is located at the interface of the trimer but is not fully surface exposed. This residue does not seem to interact with any other residue in the vicinity but may possibly affect the spike overall dynamic.

Sarbecoviruses serosurvey in humans exposed to bats

In order to determine whether humans in close contact with bats could have been infected by bat sarbecoviruses, we compared reactivity of sera from a population in close contact with bats in this sampling area (guano collectors, bat hunters or sellers) and with that of sera from the general population in Laos collected at the same period (2020) or earlier, before the emergence of COVID-19 in China (2019)(Virachith *et al*, 2021; Xaydalasouk *et al*, 2021). Pre-pandemic sera collected in France and matched for sex and age with the 2019 Lao pre-pandemic sera served as a negative control for subjects exposed to these viruses, and sera from Lao people infected with SARS-CoV-2 were used as a positive control. Sera were tested for their ability to neutralize pseudotypes of BANAL-236, -52 and -103 viruses or to bind their S proteins by two independent techniques (LIPS and LuLISA). None of the sera from the three Lao non-COVID populations showed binding or neutralizing activity against any of the three viruses above the background signals identified in the control French population (Fig. 5A-C), except for the 2019 Lao pre-pandemic sera that showed slight reactivity in LIPS-52 compared to the French pre-pandemic group. Because this finding was not supported by any other technique (PNT or LuLISA) and because the signal in bat-exposed population was lower to the one observed in the 2019 Lao pre-pandemic population, we considered this finding as non-significant. Control sera from patients infected with SARS-CoV-2 showed cross-neutralization and cross-binding to bat viruses BANAL-236, -52 and -103 as expected, given the proximity of their sequences (Fig. 5A-C). Cross-neutralization assays showed that the outlier (present in the bat-exposed Laotian

population) that neutralized BANAL-52 pseudotype was likely infected by SARS-CoV-2, reflecting the cross-reactivity of SARS-CoV-2 neutralizing antibodies for BANAL viruses (Appendix Figure S12, Appendix Table S6). Therefore, no human infection by bat sarbecoviruses close to SARS-CoV-2 (such as BANAL viruses) was detected, with an upper limit of infection rate detection of 4.1% (bat-exposed population, 2020, n=74, p=0.05) and 1.5% (general population 2019 and 2020, n=200, p=0.05).

Discussion

The origin of the COVID-19 pandemic could have proceeded from direct human infection by bat viruses like BANAL-236 or BANAL-52, which is supported by their very close proximity with SARS-CoV-2 and the failure to identify intermediate hosts contrary to their rapid detection for SARS-CoV(Guan *et al*, 2003) and MERS-CoV(Müller *et al*, 2015; Reusken *et al*, 2013). According to this hypothesis, people may have been infected in the South of China or neighboring countries by bats while visiting caves such as for trapping purposes, or by their products like meat or guano, and travelled to Wuhan. Another possibility could be the infection of terrestrial mammals with an ACE2 receptor presenting also a very high affinity for the RBD of such bat strains(Temmam *et al*, 2022) that could have acted as an intermediate host between the bat reservoir and humans. In this context, it is noteworthy that the affinity of human ACE2 for SARS-CoV-2 RBD is better than that predicted for the 54 other ACE2 orthologues evaluated(Kaushik *et al*, 2022) with the exception of the house cat (*Felis catus*), not commercialized on this market(Xiao *et al*, 2021). Another circumstance could be leakage from a laboratory that propagated bat sarbecoviruses in Wuhan.

To aid in evaluating the likelihood of these scenarios, we have taken advantage of our successful isolation of the BANAL-236 virus to anticipate the clinical consequence of such infection and the subsequent evolution of the virus. Our results add three important elements. First, data in

transgenic mice and macaques suggest that human infections, if any, could have been sub-clinical, or led to mild symptoms that could have been easily missed. Second, we show that BANAL-236 was mainly enterotropic and shed at high titer in the feces of macaques infected by the nasal and tracheal routes. This is reminiscent of SARS-CoV, which was enterotropic in addition to being pneumotropic(Leung *et al*, 2003; Peiris *et al*, 2003). Notably, viral load in mouse lungs was around 2 logs lower than upon SARS-CoV-2 infection, suggesting that transmission through the respiratory route was less efficient for BANAL-236 than for SARS-CoV-2. Therefore, if a similar strain had circulated initially in humans, or in other terrestrial mammals with similar ACE2 affinity as the human one, it was likely an enteric virus that remained undetected until acquiring a more efficient spreading ability, becoming also more pathogenic because of its adaptation to the respiratory tract. This spreading ability became further optimized in the variants that emerged later (delta, omicron). Spreading of enteric viruses is generally much less efficient than for respiratory viruses and does not lead to pandemics. In addition, people infected by such strains might be protected against more virulent strains (Fig. 1E). Third, we found no evidence of human infection by SARS-CoV-2-like bat viruses in a population in contact with infected bats, in contrast with previous observations that used less specific nucleocapsid-derived antigen(Li *et al*, 2019), which supports the hypothesis that BANAL viruses are highly transmissible neither from bats to humans nor between humans. It was recently estimated that approximately 66000 people were infected annually in Southeast Asia with SARS-like sarbecoviruses and that the frequency of infections could even be higher in people who have high contact with infected bats or intermediate hosts(Sánchez *et al*, 2022). These results were based on data from the literature on human-bat contacts, a few serological surveys among humans describing bat contact, and human SARS antibody duration. Our results, which were obtained with very specific tests are not in favor of high infection rates in such

highly exposed populations.

We have modeled the likely evolution of these bat viruses in humans by performing 6 passages in transgenic mice through the respiratory route and in Caco-2 cells, a human intestinal epithelial cell line. We evidenced mutations along the genome, but none was found among hACE2 contacting residues of the RBD, nor were predicted by our MD simulations to increase affinity for hACE2, which suggests that the very high affinity of BANAL-236 RBD for hACE2 did not benefit from additional selection pressure. This confirms that the repertoire of adaptive mutations at the RBD interface appears to be limited(Rochman *et al*, 2022). Within the spike, V391I, which has no impact on RBD binding to hACE2(Starr *et al*, 2020), was positively selected in intestinal Caco-2 cells and in monkey feces but not in the lungs of transgenic mice. This mutation could therefore reflect either an adaptation to the host species or the intestinal tropism of BANAL-236 in primates. Other mutations selected in transgenic mice (NSP12 RdRP N4984N, S1 P627L, NS6 A12T) and in monkey feces (NSP14 ExoN H5943Q and S1-NTD S52R) warrant further investigation regarding their phenotype and complementarity as it is noteworthy that they were selected synchronously during passages in mice or in the course of infection in macaques. Nevertheless, all residues mutated during passages, like most of the silent mutations, are identical in the Wuhan strain and in BANAL-236 sequenced from bat fecal swabs. Moreover, these mutations were rare in the variants described in the GISAID database (Appendix Table S5), suggesting that they could correspond to an early adaptive process but were not stabilized in the consensus of early and late SARS-CoV-2 strains. Therefore, the increase in fitness that they impart in human cells could only occur when the spike lacks a furin cleavage site.

The acquisition of a furin cleavage site could therefore have been the key event leading to the SARS-CoV-2 epidemic and pathogenic Wuhan strain. Furin cleavage sites have a well-known role in avian influenza subtypes H5 and H7 pathogenesis. Indeed, acquisition of a furin cleavage

site composed of arginine (R) or lysine (K) in avian influenza virus hemagglutinin after replication in chickens is a documented event leading to the emergence of highly pathogenic avian influenza viruses (HPAIVs) from low pathogenic avian influenza viruses (LPAIVs) after spillover from wild aquatic birds. Such a mechanism can be reproduced by passages *in ovo*(Horimoto & Kawaoka, 1995; Laleye & Abolnik, 2020; Seekings *et al*, 2020). The furin site is thought to be created by duplication of lysine and arginine residues by polymerase slippage, as these amino acids are encoded by purine-rich codons(Perdue *et al*, 1997). In our experiments, 6 passages in mice by the respiratory route or in human intestinal cells did not lead to the selection of a furin site. It can be argued that more passages could have allowed the selection of mutants with a polybasic site, as for the influenza virus which requires up to 11 *in vivo* passages from LPAIV strains to see HPAIVs strains emerge(Laleye & Abolnik, 2020). This high number of passages is consistent with the efficient circulation of LPAIVs in high-density poultry flocks. Regarding bat sarbecoviruses, in the context of the demonstration of, at worst, a low circulation in humans exposed to bats, our results make the spontaneous appearance of a furin site during silent circulation in humans unlikely.

The hypothesis of intense circulation of a bat coronavirus close to BANAL viruses in a different animal species remains open, but there is no reason to believe that such circulation would have been more efficient than in humans given the high affinity of the spike of BANAL viruses for human ACE2(Temmam *et al*, 2022). It has been proposed that a virus was circulating in bats, and likely after gaining the ability to bind ACE2, jumped into an intermediate host, from which two different lineages (named A and B) were separately but nearly concomitantly introduced into humans shortly thereafter(Pekar *et al*, 2022). Our results suggest that such a low-spreading strain is unlikely to diffuse in susceptible mammals (modeled by humanized transgenic mice and macaques) and to acquire a furin site, associated with pneumotropism at least on two occasions

during this short timeframe. This appears difficult to reconcile with, on the one hand, the quasi identity of sequences between BANAL-236 and SARS-CoV-2 even in the RBD, and on the other hand, the time necessary for this low-spreading virus to acquire by successive mutations a furin cleavage site while we did not identify such event even after six passages under selection pressure in transgenic mice or in human cells. Unfortunately, we could not use an evolutionary speed calibrated by known trees to estimate the time necessary for the appearance of a furin site as it corresponds to an indel.

Independently from the questions regarding the distant localization of infected bats from the epicenter of the epidemic(Harrison & Sachs, 2022), our work shows that the low diffusion, if any, of these enteric bat viruses between terrestrial mammals harboring high affinity receptors (like humans) renders unlikely the probability of co-infection of cells of these hosts by two coronaviruses, which would be required for the acquisition of the furin site by recombination, especially during such a short timeframe. Therefore, the most probable hypothesis is that the insertion of these PRRA amino-acids was a pre-spillover event that occurred in the bat reservoir and that generated a virus able to spread rapidly by the respiratory route in human populations (or through a not yet identified intermediate host). A natural origin of a SARS-CoV-2 supposes the existence of bat viruses harboring a furin site(Sander *et al*, 2022) that could have recombined with BANAL-like viruses. Interestingly, infection with BANAL-236 conferred cross-protection in mice against a SARS-CoV-2 challenge, which suggests that herd immunity against putative more virulent bat sarbecovirus strains resulting from such recombination could exist in local human populations frequently infected by bat viruses, albeit not detected in our study. To investigate the existence of bat coronaviruses bearing a furin cleavage site, we are pursuing investigations in caves located within the vast karstic biotope shared by China and neighboring countries.

Materials and Methods

Ethics

All work done in mice was approved by the Institut Pasteur Ethics Committee (project dap 210050) and authorized by the French Ministry of Research under #31816 in compliance with the European and French regulations on the protection of live vertebrates and the Standards for Human Care and Use of Laboratory Animals, of the Office for Laboratory Animal Welfare (OLAW, assurance number #F16-00110 A5476-01).

Cynomolgus macaques (*Macaca fascicularis*) originating from Mauritian AAALAC certified breeding centers were used in this study. All animals were housed within IDMIT animal facilities at CEA, Fontenay-aux-Roses under BSL-3 containment when necessary (Animal facility authorization #D92-032-02, Préfecture des Hauts de Seine, France) and in compliance with European Directive 2010/63/EU, the French regulations and the Standards for Human Care and Use of Laboratory Animals, of the Office for Laboratory Animal Welfare (OLAW, assurance number #A5826-01, US). Animals tested negative for *Campylobacter*, *Yersinia*, *Shigella* and *Salmonella* before being used in the study.

The protocols were approved by the institutional ethical committee "Comité d'Ethique en Expérimentation Animale du Commissariat à l'Energie Atomique et aux Energies Alternatives" (CEtEA #44) under statement number A20-037. The study was authorized by the "Research, Innovation and Education Ministry" under registration number APAFIS#24434-2020030216532863 v3. All information on the ethics committee is available at https://cache.media.enseignementsup-recherche.gouv.fr/file/utilisation_des_animaux_fins_scientifiques/22/1/comiteethiqueea17_juin2013_257221.pdf.

Biosafety

All experiments with animals (mice and macaques) and cell culture have been conducted in BSL-3 laboratories using procedures approved by each institution's health and security department. The project was monitored by an Institut Pasteur's internal committee in charge of the dual use research of concern monitoring: experiments in mice and cell culture were approved by the committee under the number #DURC-2021-01.

Human samples

Pre-COVID-19 human serum samples (n=100) were collected in the context of a hospital-based serostudy in Saravan province, southern Laos, 2017 (Lao National Ethics Committee for Health Research reference 018/NECHR/2017 (Xaydalasouk *et al*, 2021). Serum samples from late 2020 (n=100) were collected from the general Lao population, healthcare workers and bat contacts from six provinces (Ref 052/NECHR/2020 (Virachith *et al*, 2021)). The bat contacts group (n=74) consisted of guano collectors and bat hunters or sellers from four villages located in areas alongside isolated limestone karsts where BANAL-20-236 and other SARS-like coronaviruses were detected from bats. SARS-CoV-2 positive serum samples (n=15) were obtained from patients during routine COVID-19 surveillance.

Cells

Vero E6 and Caco-2 cells (ATCC) were grown in Dulbecco's modified Eagle's medium (DMEM) with 5% fetal bovine serum (Thermo Fisher Scientific) and 1% penicillin-streptomycin (Thermo Fisher Scientific). All cell lines were tested negative for mycoplasma.

Mice infection

B6.Cg-Tg(K18-ACE2)2Prlmn/J transgenic mice were imported from The Jackson Laboratory (SN #34860) and bred at the Institut Pasteur. BALB/cJRj were purchased from Janvier Labs (Le

Genest St Isle, France). Mice were maintained under specific-pathogen-free conditions with a 14-h light and 10-h dark cycle and ad libitum food and water in the Institut Pasteur animal facility.

Infection studies were performed in animal biosafety level 3 (BSL-3) facilities at the Institut Pasteur, on 11 to 19-week-old female mice. Forty to 60 microliters of either viral suspension or lung homogenate were inoculated intranasally to ketamine-xylazine-anesthetized mice. Group size and inoculation dose are indicated in the figure legends. Clinical signs of disease and weight loss were monitored daily. Mice reaching predefined humane endpoints, the end of the observation period or the time point for sample collection were euthanized by ketamine/xylazine overdose. Blood was collected by puncture of the retroorbital sinus. For viral quantification, tissue samples were placed in lysing matrix D tubes (MP Biomedical) and stored at -80°C.

For serial passaging, three mice (P1) were inoculated with 10^3 PFU of BANAL-236 and euthanized at 3 days post-infection. The right lung lobe was placed on a 70µ cell strainer (Falcon), minced with fine scissors and ground with a syringe plunger using 400 µL of PBS. The three lung homogenates were mixed in equal volumes and 40 µL of the mix was inoculated to each of three P2 mice. Viral load and viral titers were measured on the mix. The procedure was repeated until P6. Lung homogenates from P6 mice were frozen until viral titer was obtained. P7 mice were inoculated with 10^4 PFU of P6 mix.

Histology

Mouse lungs were fixed for one week in 10% neutral-buffered formalin for biosafety reasons and transferred into 70% ethanol. Samples were embedded in paraffin. Four µm-thick sections were cut and stained with hematoxylin and eosin (HE).

Macaques infection

According to preliminary experiments, to the literature and also to the recommendation of the Bill and Melinda Gates Foundation, we routinely use 10^5 or 10^6 PFU of SARS-CoV-2 Wuhan/animal to evaluate the efficacy of vaccine and treatment. Both doses are efficient to obtain mild symptoms after SARS-CoV-2 Wuhan exposure. For this study, two female cynomolgus macaques aged 4-5 years were exposed to a total dose of 6.10^5 PFU of BANAL-236 virus isolated and passaged twice in VeroE6 cells via the combination of intranasal and intra-tracheal routes (0.25 mL in each nostril and 4.5 mL in the trachea, i.e. a total of 5 mL; day 0), using atropine (0.04 mg/kg) for pre-medication and ketamine (5 mg/kg) with medetomidine (0.05 mg/kg) for anesthesia, with sample collection as previously described(Maisonnasse *et al*, 2021). Blood cell counts, haemoglobin and haematocrit were determined from EDTA blood using a DXH800 analyzer (Beckman Coulter). We have compared the results with those obtained in 44 macaques inoculated in independent challenges with the Wuhan strain using the same exposure (including titrations made in the same laboratory), sampling protocol and read out.

BANAL-236 genomic and sub-genomic RNA RT-qPCR for viral load assessment in NHPs

Upper respiratory (nasopharyngeal and tracheal) and rectal specimens were collected with swabs (Viral Transport Medium, CDC, DSR-052-01). Tracheal swabs were performed by insertion of the swab above the tip of the epiglottis into the upper trachea at approximately 1.5 cm of the epiglottis. All specimens were stored between 2°C and 8°C until analysis by RT-qPCR with a plasmid standard concentration range containing an RdRp gene fragment including the RdRp-IP4 RT-PCR target sequence. The limit of detection was estimated to be 2.67 $\log_{10}$ copies of SARS-CoV-2 gRNA per mL and the limit of quantification was estimated to be 3.67 $\log_{10}$ copies per mL. The protocol describing the procedure for the detection of SARS-CoV-2 is available on the WHO website (https://www.who.int/docs/default-source/coronaviruse/real-time-rt-pcr-assays-for-the-detection-of-sars-cov-2-institut-pasteur-paris.pdf?sfvrsn=3662fcb6_2). SARS-CoV-2 sub-genomic

mRNA (sgRNA) levels were assessed by RT-qPCR targeting the E gene using primers and probes previously described(Corman *et al*, 2020): leader-specific primer sgLeadSARSCoV2-F CGATCTCTTGTAGATCTGTTCTC, E-Sarbeco-R primer ATATTGCAGCAGTACGCACACA and E-Sarbeco probe HEX579 ACACTAGCCATCCTTACTGCGCTTCG-BHQ1. The limit of detection was estimated to be 2.87 $\log_{10}$ copies of SARS-CoV-2 sgRNA per mL and the limit of quantification was estimated to be 3.87 $\log_{10}$ copies per mL.

Chest CT Protocol

All imaging acquisitions were performed on the Digital Photon Counting (DPC) PET-CT system (Vereos-Ingenuity, Philips) implemented in BSL3 laboratory. For imaging sessions, animals were first anesthetized with Ketamine (10mg/kg) + Metedomidine (0.05mg/kg) and then maintained under isofluorane 2% in a supine position on a patient warming blanket (Bear Hugger, 3M) on the machine bed with cardiac rate, oxygen saturation and temperature monitoring.

CT was performed under breath-hold with a detector collimation of 64 × 0.6 mm, the tube voltage was 120 kV and intensity of about 150mAs. Automatic dose optimization tools (Dose Right, Z-DOM, 3D-DOM by Philips Healthcare) regulated the intensity. CT images were reconstructed with a slice thickness of 1.25 mm and an interval of 0.25 mm.

Images were analyzed using INTELLISPACE PORTAL 8 (Philips healthcare) and 3DSlicer (Open source tool). Pulmonary lesions were defined as Ground Glass Opacity, Crazy603 paving pattern or consolidation as previously described. Lesion features detected by CT imaging were assessed by two analyzers independently and final CT score results were obtained by consensus (Naninck *et al*, 2022).

Generation of lentiviral pseudoviruses and neutralization assay

The lentiviruses expressing BANAL-236 and Wuhan-372 synthetic spike genes were described previously(Temmam *et al*, 2022). Those expressing the spikes of BANAL-52 and BANAL-103 were constructed similarly. Neutralization assays were conducted as described(Temmam *et al*, 2022).

Luciferase Immunoprecipitation Assays

Recombinant antigens comprising the RBD or the ectodomain of the full spike of BANAL-236, -52, and -103 viruses were designed in fusion with the nanoluciferase as follows: the foldon domain (YIPEAPRDGQAYVRKDGEWVLLSTFL) was added to the C-terminus of each ectodomain to allow the S protein to trimerize(Stevens *et al*, 2004), resembling the native spike state of the virion. The nanoluciferase was added to the carboxy-terminal end of each construct spaced by a 3-residues GSG linker. An exogenous signal peptide was added to RBD constructs to ensure efficient protein secretion. Synthetic genes were ordered from GenScript with codon usage optimized for protein expression in mammalian cells and cloned in the pcDNA3.1(+) vector. Plasmids were transfected in Expi293F Gn-TI cells using the ExpiFectamine293 Transfection kit (Fisher Scientific) according to the manufacturer's recommendations, recombinant proteins were harvested at day 4 in the supernatant without any purification step and quantified by adding the substrate of the nanoluciferase to serial dilutions of the supernatant. Luminescence was measured onto a Centro XS^3 LB 960 luminometer (Berthold Technologies, France). LIPS assay was conducted as previously described(Grzelak *et al*, 2020) except that 10 µL of sera diluted to ¼ in fetal calf serum (FCS, Fisher scientific) were engaged per reaction. The signal-to-noise ratio of each serum was calculated by dividing the RLU by the mean of 10 negative controls (with 10 µL of FCS) + five standard deviations.

LuLISA (Luciferase-Linked Immunoscorbent Assay)

The purified full-length N and S protein stabilized in prefusion conformation of BANAL-236, BANAL-52 and BANAL-103 as described previously for SARS-CoV-2 (Woudenberg *et al*, 2021) was diluted in PBS 1X at 50 ng/µL. White plates (Thermo Scientific 437796 /470372) were coated with 50 µL/well during 4 hours at room temperature. After 3 washes in PBS 1X 0,1%Tween, 50 µL of heat-inactivated serums diluted 1/200 in PBS 1X, 0,1%Tween 1% milk were added/well, and incubated for 1 hour at room temperature. After 3 washes in PBS 1X 0,1%Tween, anti-IgG nanobody fused to the nanoluciferase was added diluted in PBS 1X, 0,1%Tween, 1% milk (dilution depending upon the preparation) and incubated for 20 minutes at room temperature. After 3 washes in PBS 1X 0,1%Tween, bioluminescent detection of antibody levels was performed by using 50 µL of 12 µM of the nanoluciferase substrate Furimazine, and measuring RLU for 0,5sec/well in a luminometer (Bertold centro XS).

Mutation analysis

The quantification of viral loads in genome copy per g of lung of each infected transgenic mouse expressing hACE-2 was carried out to constitute seven pools of replicates, each of 3-4 samples at equivalent concentration while individual macaque rectal samples were sequenced. A custom panel (251 probes) based on the MZ9370003.2 BANAL-236 genome was designed by TWIST Bioscience for NGS-based Target Capture. The double stranded cDNA (dscDNA) of each pool were constructed using the Protoscript II First strand cDNA synthesis kit followed by NEBNext Ultra II Non-Directional RNA Second Strand Synthesis Module. Purified dscDNA was then used as input for the Illumina DNA Prep with Enrichment library preparation protocol. Briefly, each purified dscDNA sample was tagmented and fragments were barcoded. The barcoded libraries were pooled and hybridized with the TWIST custom panel for BANAL-236 following the

specifications for third-party panels of the Illumina DNA Prep with Enrichment protocol. The captured libraries were sequenced on an Illumina MiSeq Micro format in a paired-end 300-cycle run. For the infected VeroE6 and Caco-2 cells, the libraries were performed using the SMARTer Stranded Total RNA-Seq Kit v3 - Pico Input Mammalian kit (TaKaRa Bio, Ozyme) was used according to the manufacturer instructions and sequenced on an Illumina NextSeq 500 in a paired-end 300-cycle run, except for the Caco-2_C4 sample (single-read sequencing).

Reads were trimmed with AlienTrimmer, then cropped using a custom python script in order to clean the sequence ends. Cleaned reads were mapped on the BANAL-20-236 genome (GenBank: MZ9370003.2) using CLC Genomics Workbench 20.0.4 with read Length fraction=1 and Similarity fraction=0.985. After adding InDel qualities, variants were called using LoFreq v.2.1.4 and variant annotation was performed with SnpEff v4_5covid19_core. Variants were checked for depth and strand bias. SNVs detected in at least three serial passages were first selected. Mutations with low frequency and weak variation were discarded.

RT-qPCR on mice tissues

Organs were disrupted in 1 mL of DMEM with the TissueLyser LT homogenizer (Qiagen) and 2 mL-reinforced tubes containing beads (Bertin, France). Total RNAs were extracted from 100 µL of clarified supernatant with the Maxwell® RSC simplyRNA Tissue kit (Promega) with DNAse treatment, according to the manufacturer's instructions. Samples were tested with a one-step RT-qPCR Taqman system using a set of primers and probe targeting the E gene, as previously described(Corman *et al*, 2020) and who.int/docs/default-source/coronaviruse/real-time-rt-pcr-assays-for-the-detection-of-sars-cov-2-institut-pasteur-paris.pdf?sfvrsn=3662fcb6_2.

Titration

Approximately 10^6 Vero E6 cells were seeded to each well of a 6-well plate and cultured overnight at 37°C, 5% CO_2. The next day, the cells were infected with 400 µL of 10-fold serial dilutions of viral supernatants in DMEM without fetal bovine serum (FBS), supplemented with 1 µg/mL TPCK-Trypsin and 1% penicillin-streptomycin. The viruses were incubated with the cells at 37 °C with 5% CO_2 for 1 hour. An overlay of MEM-avicel supplemented with 1 µg/mL TPCK-Trypsin was added to the cells. Three days later, the plates were stained with 30 % crystal violet, 20% ethanol and 10% formaldehyde. Plaques were counted and viral titers in PFU/mL were calculated.

Molecular Dynamics simulations of BANAL-236 RBD/human ACE2 complexes

We performed 6 all-atom explicit-solvent molecular dynamics (MD) simulations of the human ACE2 (hACE2) in complex with the wild type BANAL-236 RBD (WT-hACE2) and V391I mutant (V391I-hACE2). All simulations were performed with GROMACS 2021.4(Abraham *et al*, 2015) for a total aggregated time of 6 µs. The GROMACS topology and input files as well as the analysis scripts used are available on PLUMED-NEST (plumed-nest.org)(Bonomi *et al*, 2019) under accession ID plumID:22.027. This work was granted access to the HPC resources of IDRIS under the allocation 2022-102408 made by GENCI.

Details of the MD simulations: setup, equilibration, and production.

The X-ray structure of the wild type BANAL-236 RBD/hACE2 complex (PDB code 7PKI) was used as input to the CHARMM-GUI server(Jo *et al*, 2008). The zinc and chloride atoms present in 7PKI as well as the NAG and water residues were retained. The system was solvated in a triclinic box of initial x-y-z dimensions equal to 12.8 nm * 8.9 nm * 8.7 nm. Potassium and chloride ions were added to ensure charge neutrality at a salt concentration of 0.15 M. The total number of atoms were 96937. CHARMM-GUI was also used to construct a homology model of

the V391I mutant starting from the X-ray structure of the wild type. Additional details of the systems are reported in Appendix Table S4. The CHARMM36m force field(Huang *et al*, 2017) was used for the protein and ions and the TIP3P model(Jorgensen *et al*, 1983) was used for the water molecules. The CHARMM-GUI models were first energy-minimized using the steepest descent algorithm. After minimization, the wild type and mutant complexes were equilibrated using a 10-ns long simulation in the NPT ensemble followed by a 10-ns long simulation in the NVT ensemble. The temperature T was set at 300 K and the pressure P at 1 atm using the Bussi-Donadio-Parrinello thermostat(Bussi *et al*, 2007) and the Berendsen barostat(Berendsen *et al*, 1984) respectively. During equilibration, harmonic restraints on the positions of the protein backbone and sidechains heavy atoms were applied. From the NVT equilibration runs, 3 conformations of the wild type and V391I mutant were extracted and used as starting point for production simulations. The production simulations were performed in the NVT ensemble for 1 μs. A time step of 2 fs was used together with LINCS constraints on h-bonds(Hess *et al*, 1997). The van der Waals interactions were gradually switched off at 1.0 nm and cut off at 1.2 nm; the particle-mesh Ewald method was used to calculate electrostatic interactions with cutoff at 1.2 nm(Essmann *et al*, 1995).

Details of the analysis.

To evaluate the stability of the starting model during the production runs, we calculated the backbone Root Mean Square Deviation (RMSD) with respect to the energy-minimized structure for each frame of the trajectories. The RMSD was calculated separately for: RBD, hACE2, the residues at the RBD-hACE2 interface, which were defined as the residues in one subunit closer than 0.8 nm to the residues in the other subunit in the X-ray structure 7PKI, and the entire complex (Appendix Figure 8). RMSD calculations were performed using the *driver* utility of PLUMED v. 2.7(Bonomi *et al*, 2019). To estimate the binding energy between RBD and

hACE2, we used the *AnalyseComplex* tool in FoldX v. 4(Schymkowitz *et al*, 2005) (Appendix Figure S9). To identify relevant interactions at the RBD-hACE2 interface, we quantified the frequency of formation of inter-subunits hydrogen bonds (Appendix Figure S10) and salt bridges (Appendix Figure S11) during the course of the MD simulations. To monitor the formation of hydrogen bonds, we used the Hydrogen Bond Analysis module of the MDAnalysis library v. 2.1.0(Michaud-Agrawal *et al*, 2011). A donor–acceptor distance and angular cutoff of 0.3 nm and 150°, respectively, were used to define the formation of a hydrogen bond. For salt bridges, for each frame of the trajectories we used MDAnalysis to compute the distances between the sidechain charged groups of aspartic acids (OD1/OD2), glutamic acids (OE1/OE2), lysines (NZ), and arginines (NH1/NH2). A inter-subunits salt bridge was defined as formed if the distance between groups with opposite charge was lower than 0.32 nm.

Statistical analyses

Statistical analysis was performed with Prism (GraphPad Software, LLC) and R software. Survival curves were compared with a logrank test. Body weight loss curves were compared day-by-day with a Student's t-test.

Data availability

The NGS data generated in the study (Figure 3) were deposited in the NCBI/SRA archive, under the BioProject PRJNA796968 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA796968/). BioSample accession numbers: SAMN32126682-SAMN32126699.

Acknowledgments

We thank the PF3PR platform of Institut Pasteur for the production of recombinant proteins used in LuLISA assay; and G. Haustant and L. Lemée of the Biomics platform of Institut Pasteur for their help in mutations sequencing and analyses. We thank C. Manet, B. Delache, S. Langlois, Q.

Sconosciuti, V. Magneron, M. Potier, J. M. Robert, E. Burban, N. Dhooge T. Prot, and C. Dodan for the NHP experiments; L. Bossevot, M. Galpin-Lebreau, L. Pintore, , L. Moenne-Loccoz and J. Morin for RT-qPCR assays and preparation of reagents; R. Marlin and V. Contreras for scientific discussion and help; J. Dinh and E. Guyon for NHP sample processing; C. Chapon and N. Kahlaoui from the imagery facility; F. Ducancel, A. Pouget and Y. Gorin for their help with the logistics and safety management; I. Mangeot for her help with resources management and B. Targat for his contribution to data management. We thank A. Nougairede for sharing the plasmid used for the sgRNA assays standardization. We thank the Institut Pasteur BSL3 animal facility (R. Chennouf) and Histology Platform (D. Hardy and M. Tichit).

Next-generation sequencing was performed with the help of Biomics Platform, C2RT, Institut Pasteur, Paris, France, supported by France Génomique (ANR-10-INBS-09-09), IBISA, and the Illumina COVID-19 Projects' offer. The work was funded by an Institut Pasteur "Covid Taskforce" and in part by the H2020 project 101003589 (RECOVER) and Labex IBEID (ANR-10-LABX62-IBEID) grants. The Infectious Disease Models and Innovative Therapies (IDMIT) research infrastructure is supported by the "Programme Investissements d'Avenir", managed by the ANR under reference ANR-11-INBS-0008. The Fondation Bettencourt Schueller and the Region Ile-de-France contributed to the implementation of IDMIT's facilities and imaging technologies. The NHP model of SARS-CoV-2 infection have been developed thanks to the support from REACTing, the Fondation pour la Recherche Medicale (FRM; AM-CoV-Path).

Illustrations for mice and primate infections were created with BioRender.com.

Disclosure and Competing Interests Statement

Authors declare that they have no competing interests.

References

Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B & Lindahl E (2015) GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* 1–2: 19–25

Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A & Haak JR (1984) Molecular dynamics with coupling to an external bath. *J Chem Phys* 81: 3684–3690

Bonomi M, Bussi G, Camilloni C, Tribello GA, Banáš P, Barducci A, Bernetti M, Bolhuis PG, Bottaro S, Branduardi D, *et al* (2019) Promoting transparency and reproducibility in enhanced molecular simulations. *Nat Methods* 16: 670–673

Bussi G, Donadio D & Parrinello M (2007) Canonical sampling through velocity rescaling. *J Chem Phys* 126: 014101

Cameroni E, Bowen JE, Rosen LE, Saliba C, Zepeda SK, Culap K, Pinto D, VanBlargan LA, De Marco A, di Iulio J, *et al* (2021) Broadly neutralizing antibodies overcome SARS-CoV-2 Omicron antigenic shift. *Nature*: 1–9

Corman VM, Landt O, Kaiser M, Molenkamp R, Meijer A, Chu DK, Bleicker T, Brünink S, Schneider J, Schmidt ML, *et al* (2020) Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Euro Surveill* 25

Delaune D, Hul V, Karlsson EA, Hassanin A, Ou TP, Baidaliuk A, Gámbaro F, Prot M, Tu VT, Chea S, *et al* (2021) A novel SARS-CoV-2 related coronavirus in bats from Cambodia. *Nat Commun* 12: 6563

Essmann U, Perera L, Berkowitz ML, Darden T, Lee H & Pedersen LG (1995) A smooth particle mesh Ewald method. *J Chem Phys* 103: 8577–8593

Grzelak L, Temmam S, Planchais C, Demeret C, Tondeur L, Huon C, Guivel-Benhassine F, Staropoli I, Chazal M, Dufloo J, *et al* (2020) A comparison of four serological assays for detecting anti-SARS-CoV-2 antibodies in human serum samples from different populations. *Science Translational Medicine* 12

Gu H, Chen Q, Yang G, He L, Fan H, Deng Y-Q, Wang Y, Teng Y, Zhao Z, Cui Y, *et al* (2020) Adaptation of SARS-CoV-2 in BALB/c mice for testing vaccine efficacy. *Science* 369: 1603–1607

Guan Y, Zheng BJ, He YQ, Liu XL, Zhuang ZX, Cheung CL, Luo SW, Li PH, Zhang LJ, Guan YJ, *et al* (2003) Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China. *Science* 302: 276–278

Halfmann PJ, Iida S, Iwatsuki-Horimoto K, Maemura T, Kiso M, Scheaffer SM, Darling TL, Joshi A, Loeber S, Singh G, *et al* (2022) SARS-CoV-2 Omicron virus causes attenuated disease in mice and hamsters. *Nature*: 1–1

Harrison NL & Sachs JD (2022) A call for an independent inquiry into the origin of the SARS-CoV-2 virus. *Proceedings of the National Academy of Sciences* 119: e2202769119

Hess B, Bekker H, Berendsen HJC & Fraaije JGEM (1997) LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry* 18: 1463–1472

Hoffmann M, Krüger N, Schulz S, Cossmann A, Rocha C, Kempf A, Nehlmeier I, Graichen L, Moldenhauer A-S, Winkler MS, *et al* (2021) The Omicron variant is highly resistant against antibody-mediated neutralization – implications for control of the COVID-19 pandemic. *Cell*

Horimoto T & Kawaoka Y (1995) Molecular changes in virulent mutants arising from avirulent avian influenza viruses during replication in 14-day-old embryonated eggs. *Virology* 206: 755–759

Huang J, Rauscher S, Nawrocki G, Ran T, Feig M, de Groot BL, Grubmüller H & MacKerell AD (2017) CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat Methods* 14: 71–73

Huang K, Zhang Y, Hui X, Zhao Y, Gong W, Wang T, Zhang S, Yang Y, Deng F, Zhang Q, *et al* (2021) Q493K and Q498H substitutions in Spike promote adaptation of SARS-CoV-2 in mice. *EBioMedicine* 67: 103381

Jo S, Kim T, Iyer VG & Im W (2008) CHARMM-GUI: a web-based graphical user interface for CHARMM. *J Comput Chem* 29: 1859–1865

Johnson BA, et al, Bailey AL, Kalveram B, Lokugamage KG, Muruato A, Zou J, Zhang X, Juelich T, Smith JK, *et al* (2021) Loss of furin cleavage site attenuates SARS-CoV-2 pathogenesis. *Nature* 591: 293–299

Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW & Klein ML (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79: 926–935

Kaushik R, Kumar N, Zhang KYJ, Srivastava P, Bhatia S & Malik YS (2022) A novel structure-based approach for identification of vertebrate susceptibility to SARS-CoV-2: Implications for future surveillance programmes. *Environ Res* 212: 113303

Laleye AT & Abolnik C (2020) Emergence of highly pathogenic H5N2 and H7N1 influenza A viruses from low pathogenic precursors by serial passage in ovo. *PLoS One* 15: e0240290

Leist SR, Dinnon KH, Schäfer A, Tse LV, Okuda K, Hou YJ, West A, Edwards CE, Sanders W, Fritch EJ, *et al* (2020) A Mouse-Adapted SARS-CoV-2 Induces Acute Lung Injury and Mortality in Standard Laboratory Mice. *Cell* 183: 1070-1085.e12

Leung WK, To K-F, Chan PKS, Chan HLY, Wu AKL, Lee N, Yuen KY & Sung JJY (2003) Enteric involvement of severe acute respiratory syndrome-associated coronavirus infection. *Gastroenterology* 125: 1011–1017

Li H, Mendelsohn E, Zong C, Zhang W, Hagan E, Wang N, Li S, Yan H, Huang H, Zhu G, *et al* (2019) Human-animal interactions and bat coronavirus spillover potential among rural residents in Southern China. *Biosaf Health* 1: 84–90

Maisonnasse P, Aldon Y, Marc A, Marlin R, Dereuddre-Bosquet N, Kuzmina NA, Freyn AW, Snitselaar JL, Gonçalves A, Caniels TG, *et al* (2021) COVA1-18 neutralizing antibody protects against SARS-CoV-2 in three preclinical models. *Nat Commun* 12: 6097

Maisonnasse P, Guedj J, Contreras V, Behillil S, Solas C, Marlin R, Naninck T, Pizzorno A, Lemaitre J, Gonçalves A, *et al* (2020) Hydroxychloroquine use against SARS-CoV-2 infection in non-human primates. *Nature* 585: 584–587

Michaud-Agrawal N, Denning EJ, Woolf TB & Beckstein O (2011) MDAnalysis: a toolkit for the analysis of molecular dynamics simulations. *J Comput Chem* 32: 2319–2327

Montagutelli X, Prot M, Jouvion G, Levillayer L, Conquet L, Reyes-Gomez E, Donati F, Albert M, Werf S van der, Jaubert J, *et al* (2021) A mouse-adapted SARS-CoV-2 strain replicating in standard laboratory mice. 2021.07.10.451880 doi:10.1101/2021.07.10.451880 [PREPRINT]

Müller MA, Meyer B, Corman VM, Al-Masri M, Turkestani A, Ritz D, Sieberg A, Aldabbagh S, Bosch B-J, Lattwein E, *et al* (2015) Presence of Middle East respiratory syndrome coronavirus antibodies in Saudi Arabia: a nationwide, cross-sectional, serological study. *Lancet Infect Dis* 15: 559–564

Murakami S, Kitamura T, Suzuki J, Sato R, Aoi T, Fujii M, Matsugo H, Kamiki H, Ishida H, Takenaka-Uema A, *et al* (2020) Detection and Characterization of Bat Sarbecovirus Phylogenetically Related to SARS-CoV-2, Japan. *Emerg Infect Dis* 26: 3025–3029

Naninck T, Kahlaoui N, Lemaitre J, Maisonnasse P, De Mori A, Pascal Q, Contreras V, Marlin R, Relouzat F, Delache B, *et al* (2022) Computed tomography and [18F]-FDG PET imaging provide additional readouts for COVID-19 pathogenesis and therapies evaluation in non-human primates. *iScience* 25: 104101

Peiris JSM, Chu CM, Cheng VCC, Chan KS, Hung IFN, Poon LLM, Law KI, Tang BSF, Hon TYW, Chan CS, *et al* (2003) Clinical progression and viral load in a community outbreak of coronavirus-associated SARS pneumonia: a prospective study. *Lancet* 361: 1767–1772

Pekar JE, Magee A, Parker E, Moshiri N, Izhikevich K, Havens JL, Gangavarapu K, Malpica Serrano LM, Crits-Christoph A, Matteson NL, *et al* (2022) SARS-CoV-2 emergence very likely resulted from at least two zoonotic events Zenodo

Perdue ML, García M, Senne D & Fraire M (1997) Virulence-associated sequence duplication at the hemagglutinin cleavage site of avian influenza viruses. *Virus Res* 49: 173–186

Pommerenke C, Rand U, Uphoff CC, Nagel S, Zaborski M, Hauer V, Kaufmann M, Meyer C, Denkmann SA, Riese P, *et al* (2021) Identification of cell lines CL-14, CL-40 and CAL-51 as suitable models for SARS-CoV-2 infection studies. *PLoS One* 16: e0255622

Rahalkar MC & Bahulikar RA (2020) Lethal Pneumonia Cases in Mojiang Miners (2012) and the Mineshaft Could Provide Important Clues to the Origin of SARS-CoV-2. *Front Public Health* 8

Reusken CB, Haagmans BL, Müller MA, Gutierrez C, Godeke G-J, Meyer B, Muth D, Raj VS, Vries LS-D, Corman VM, *et al* (2013) Middle East respiratory syndrome coronavirus neutralising serum antibodies in dromedary camels: a comparative serological study. *The Lancet Infectious Diseases* 13: 859–866

Rochman ND, Faure G, Wolf YI, Freddolino PL, Zhang F & Koonin EV (2022) Epistasis at the SARS-CoV-2 Receptor-Binding Domain Interface and the Propitiously Boring Implications for Vaccine Escape. *mBio* 13: e00135-22

Sánchez CA, Li H, Phelps KL, Zambrana-Torrelio C, Wang L-F, Zhou P, Shi Z-L, Olival KJ & Daszak P (2022) A strategy to assess spillover risk of bat SARS-related coronaviruses in Southeast Asia. *Nat Commun* 13: 4380

Sander A-L, Moreira-Soto A, Yordanov S, Toplak I, Balboni A, Ameneiros RS, Corman V, Drosten C & Drexler JF (2022) Genomic determinants of Furin cleavage in diverse European SARS-related bat coronaviruses. *Commun Biol* 5: 1–8

Schymkowitz J, Borg J, Stricher F, Nys R, Rousseau F & Serrano L (2005) The FoldX web server: an online force field. *Nucleic Acids Res* 33: W382-388

Seekings AH, Howard WA, Nuñéz A, Slomka MJ, Banyard AC, Hicks D, Ellis RJ, Nuñéz-García J, Hartgroves LC, Barclay WS, *et al* (2020) The Emergence of H7N7 Highly Pathogenic Avian Influenza Virus from Low Pathogenicity Avian Influenza Virus Using an in ovo Embryo Culture Model. *Viruses* 12: E920

Starr TN, Greaney AJ, Hilton SK, Ellis D, Crawford KHD, Dingens AS, Navarro MJ, Bowen JE, Tortorici MA, Walls AC, *et al* (2020) Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding. *Cell* 182: 1295-1310.e20

Stevens J, Corper AL, Basler CF, Taubenberger JK, Palese P & Wilson IA (2004) Structure of the uncleaved human H1 hemagglutinin from the extinct 1918 influenza virus. *Science* 303: 1866–1870

Temmam S, Vongphayloth K, Salazar EB, Munier S, Bonomi M, Regnault B, Douangboubpha B, Karami Y, Chrétien D, Sanamxay D, *et al* (2022) Bat coronaviruses related to SARS-CoV-2 and infectious for human cells. *Nature*: 1–10

Virachith S, Pommelet V, Calvez E, Khounvisith V, Sayasone S, Kounnavong S, Mayxay M, Xangsayarath P, Temmam S, Eloit M, *et al* (2021) Low seroprevalence of COVID-19 in Lao PDR, late 2020. *The Lancet Regional Health - Western Pacific* 13: 100197

Wacharapluesadee S, Tan CW, Maneeorn P, Duengkae P, Zhu F, Joyjinda Y, Kaewpom T, Chia WN, Ampoot W, Lim BL, *et al* (2021) Evidence for SARS-CoV-2 related coronaviruses circulating in bats and pangolins in Southeast Asia. *Nat Commun* 12: 972

Worobey M, Levy JI, Serrano LMM, Crits-Christoph A, Pekar JE, Goldstein SA, Rasmussen AL, Kraemer MUG, Newman C, Koopmans MPG, *et al* (2022) The Huanan market was the epicenter of SARS-CoV-2 emergence Zenodo

Woudenberg T, Pelleau S, Anna F, Attia M, Donnadieu F, Gravet A, Lohmann C, Seraphin H, Guiheneuf R, Delamare C, *et al* (2021) Humoral immunity to SARS-CoV-2 and seasonal coronaviruses in children and adults in north-eastern France. *EBioMedicine* 70: 103495

Xaydalasouk K, Sayasinh K, Hübschen JM, Khounvisith V, Keomany S, Muller CP & Black AP (2021) Age-stratified seroprevalence of vaccine-preventable infectious disease in Saravan, Southern Lao People's Democratic Republic. *Int J Infect Dis* 107: 25–30

Xiao X, Newman C, Buesching CD, Macdonald DW & Zhou Z-M (2021) Animal sales from Wuhan wet markets immediately prior to the COVID-19 pandemic. *Sci Rep* 11: 11898

Yin W, Xu Y, Xu P, Cao X, Wu C, Gu C, He X, Wang X, Huang S, Yuan Q, *et al* (2022) Structures of the Omicron spike trimer with ACE2 and an anti-Omicron antibody. *Science* 375: 1048–1053

Zhou H, Ji J, Chen X, Bi Y, Li J, Wang Q, Hu T, Song H, Zhao R, Chen Y, *et al* (2021) Identification of novel bat coronaviruses sheds light on the evolutionary origins of SARS-CoV-2 and related viruses. *Cell*

Zhou P, Yang X-L, Wang X-G, Hu B, Zhang L, Zhang W, Si H-R, Zhu Y, Li B, Huang C-L, *et al* (2020) A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 579: 270–273

Figure legends

Fig. 1. BANAL-236 infection in mice expressing hACE2.

A. Schematic representation of infection of K18-hACE2 transgenic mice expressing human ACE2 (hACE2, black) or wild-type BALB/c mice (white) with BANAL-236 (red) or SARS-CoV-2 Wuhan-372 (blue) viruses. Sample size shown as 'x'. Lung, liver, spleen and brain were collected at day 3 (D3) post-infection. Blood was collected at D30 post-infection before challenge with SARS-CoV-2 Wuhan-D614G (yellow) virus. A second blood sample was collected at D14 post-challenge. Number of mice per group is shown as "x".

B. Body weight variation of K18-hACE2 or BALB/c mice infected by BANAL-236 (red) or Wuhan-372 (blue) viruses at 10^4 PFU (solid lines) or 10^3 PFU (dotted lines) expressed as the mean (+/- SEM) body weight variation. Significant differences observed between K18-hACE2 mice infected at 10^4 PFU by BANAL-236 or Wuhan-372 at day 4 is noted by an asterisk (Student test, $P<0.05$). "#" artifactual body weight loss due to an incident occurred in the animal core facility on that day.

C. Quantification of BANAL-236 and Wuhan-372 in lung, spleen and liver of mice infected at 10^4 PFU. Viral loads are expressed as the mean (+/- SD) genome copies of genomic RNA per g organ (solid bars) or as PFUs per g organ (dotted bars). The limit of detection of the RT-qPCR and viral titration are represented as a plain or dashed lines, respectively.

D. Neutralizing antibody titer (expressed in ED_{50}) of mice primo-infected at 10^4 PFU by BANAL-236 (red) or SARS-CoV-2 Wuhan-372 (blue) viruses. The antibody response against the Wuhan-D614G challenge strain of mice primo-infected by BANAL-236 is presented as a dashed line. PNT means pseudo-neutralization. Error bars represent the standard error of the mean (SEM), $p = 0.05$.

E. Survival curves of mice challenged with SARS-CoV-2 Wuhan-D614G at 10^4 PFU, after a primo-infection by BANAL-236 or Wuhan-372 viruses at 10^3 or 10^4 PFU. Survival of K18-hACE2 naive mice infected by SARS-CoV-2 Wuhan-D614G at 10^4 PFU is presented as a reference (yellow).

F. Histopathological analysis, 3 days post-inoculation, of the lung of K18-hACE2 mice infected with 10^4 PFU of BANAL-236 (a-d) or SARS-CoV-2 Wuhan-372 (e,f) viruses. Interstitial pneumonia characterized by interstitial inflammation, often centered on blood vessels (perivasculitis, black arrows) or bronchi/bronchioles (black arrowheads), and endothelial cell injury and inflammation (endothelitis). Lesions were globally more severe after Wuhan-372 infection (n=4; all of moderate severity) than after BANAL-236 infection (n=4; 3/4 of minimal severity (a-c) and 1/4 of moderate severity (d)). Scale bar: 100 µ. Low magnification images are provided in Figure S3.

G. Clinical and biological results of the 6 serial passages of BANAL-236 in K18-hACE2 mice as a pool of lung homogenates of the previous passage. The last passage (P6) was then inoculated into K18-hACE2 mice at 10^4 PFU. *Middle panel*: Viral load in the lung is expressed as the mean copies (+/- SEM) of genomic RNA per g of lung (solid line) or as PFUs per g of lung (dashed line). *Right panel*: Body weight variation of K18-hACE2 mice infected by BANAL-236 P1 (red, same results as those of Fig. 1B) or P6 (green) viruses at 10^4 PFU. Significant differences observed between K18-hACE2 mice infected by P1 or P6 are noted by an asterisk (**: $p<0.01$).

Fig. 2. Compared neutralizing antibody titer of mice infected by BANAL-236 or SARS-CoV-2 at two doses.

Neutralizing antibody titer (expressed in ED_{50}) of mice primo-infected at 10^3 (left) or 10^4 (right) PFU by BANAL-236 (red) or SARS-CoV-2 Wuhan-372 (blue) viruses. The antibody response against the Wuhan-D614G challenge strain of mice primo-infected by BANAL-236 is presented as a dashed line. PNT means pseudo-neutralization.

Fig. 3. BANAL-236 infection in cynomolgus macaques.

A. Schematic representation of infection of cynomolgus macaques by BANAL-236 virus. CT scan, clinical scoring and biological samples were realized 10 and 3 days before infection, respectively, to serve as baseline (BL) reference. The body temperature and weight were monitored at each anesthesia. Clinical scoring was realized at BL and then from D2 to D11 post-infection and CT scan at BL and at D2 and D11 post-infection. Broncho-alveolar lavage (BAL) was realized at D3 and D11 post-infection and biological sampling of nasopharyngeal, tracheal and rectal swabs at D2 to 28 post-infection to monitor the virus load and antibody profile. Blood

samples were collected 10 days before infection to 43 days post-infection and were used for hematology, cytokine quantification and antibody profiling.

B. Effects of infection by BANAL-236 virus in non-human primates. Animal #MF1 is presented in green and animal #MF2 in purple. Clinical score, body weight variation (in %), CT scores (compared to Wuhan-372) and neutralizing antibodies are measured from 10 days before infection to 43 days post-infection.

C. Viral load expressed in $\log_{10}$ copies of genomic RNA (solid line) or sub-genomic RNA (dashed line) per mL of tracheal, nasopharyngeal, BAL and rectal fluids. Lower limit of quantification is presented in red for both genomic (solid line) and sub-genomic (dashed line) RNA, and limits of detection are presented in green for both genomic (solid line) and sub-genomic (dashed line) RNA. Results were compared to historical data of SARS-CoV-2 Wuhan-372 virus inoculated at 10^6 or 10^5 PFU (presented as the mean virus load +/- 95% CI expressed in $\log_{10}$ copies of genomic RNA/mL).

Fig. 4. Mutations selected upon passages in K18-hACE2 mice and Caco-2 cells.

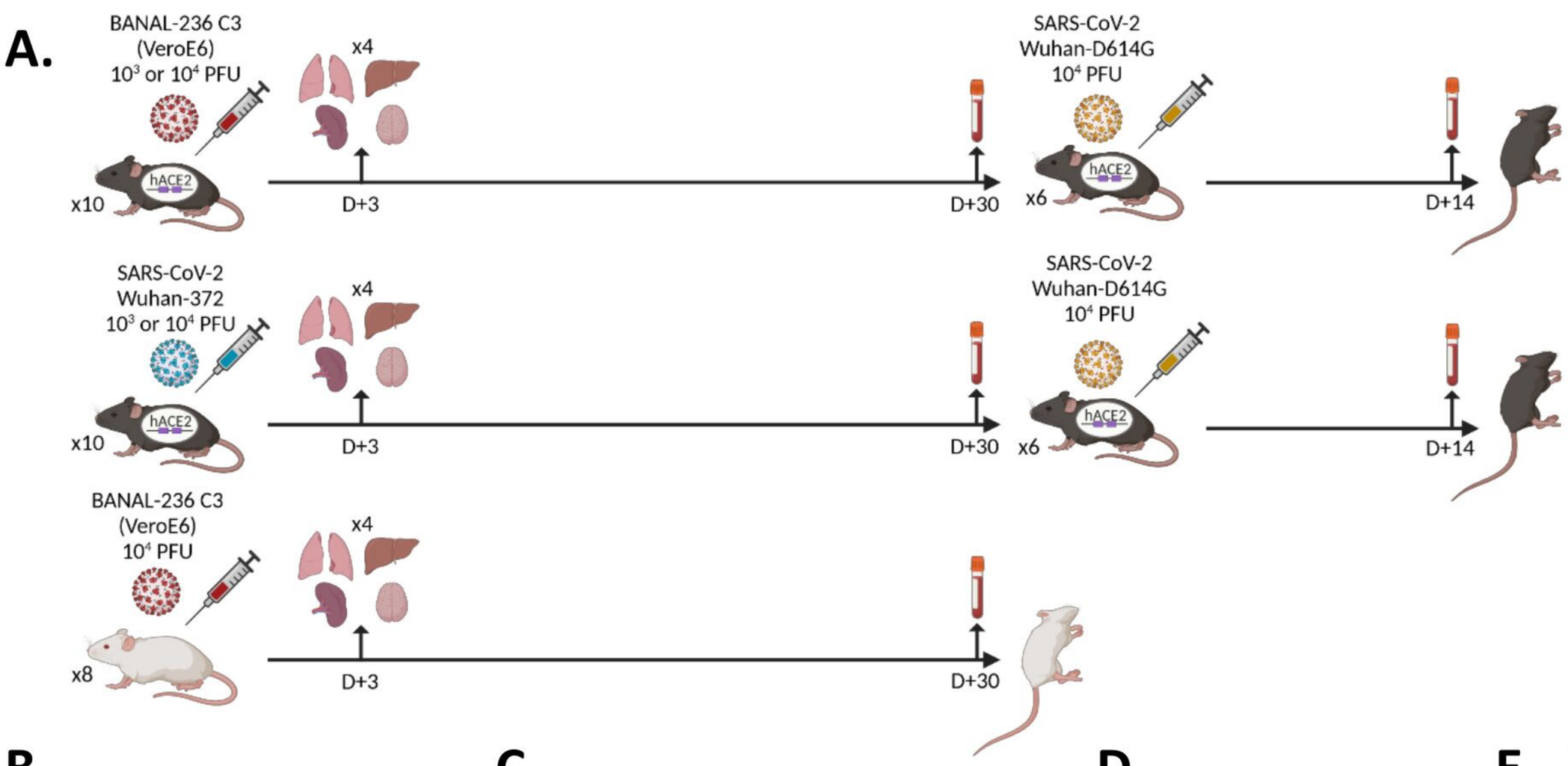
A. Schematic representation of serial passages of BANAL-236 virus in Vero E6 cells (C1 to C3) followed by 6 passages in K18-hACE2 mice or by 6 passages in Caco-2 cells at a constant multiplicity of infection (MOI) (C4 to C9).

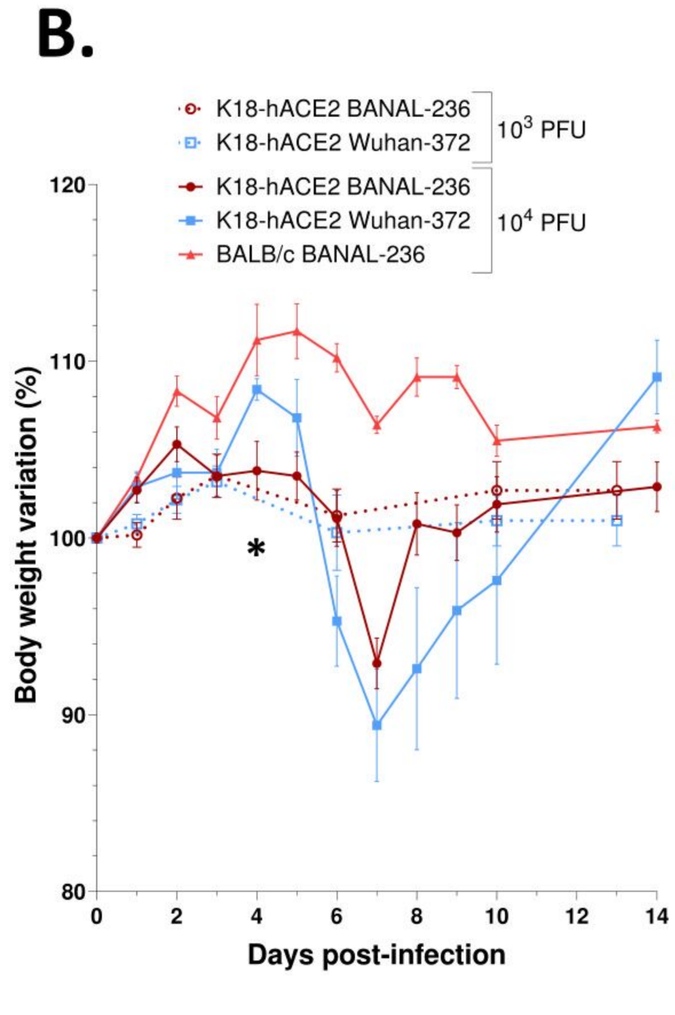
B. Mutation profiles of BANAL-236 after passages in VeroE6 cells, Caco-2 cells, K18-hACE2 mice and in the feces of macaque during the course of infection. Mutations reported in the heatmap were detected in at least three samples at a frequency above 10% in at least one sample. Rows are ordered according to the different passages and columns are clustered by SNVs frequencies.

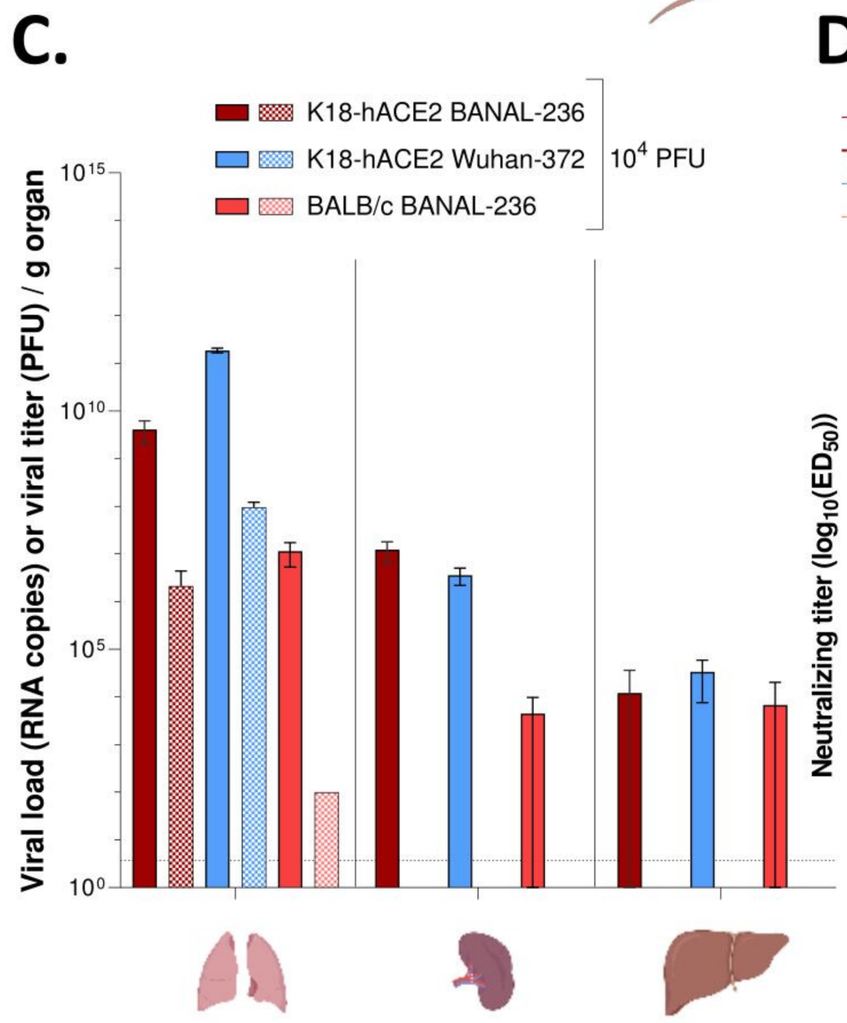
C. Stacked histogram to compare the SNV frequencies of BANAL-236 from the Caco-2 and mice serial passages along the MZ9370003.2 genome. Mutations reported were detected in at least three serial samples. K18-hACE2 lung passages are color coded in red in the upper panel and Caco-2 passages are color coded in green in the lower panel. Each colored square height summarizes the allelic frequency (AF).

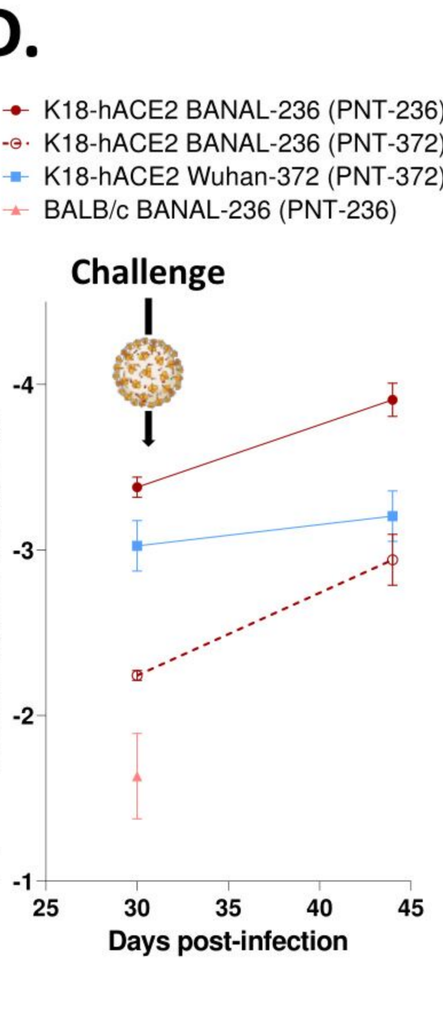
Fig. 5. Antibody testing against bat sarbecoviruses in Lao human populations.

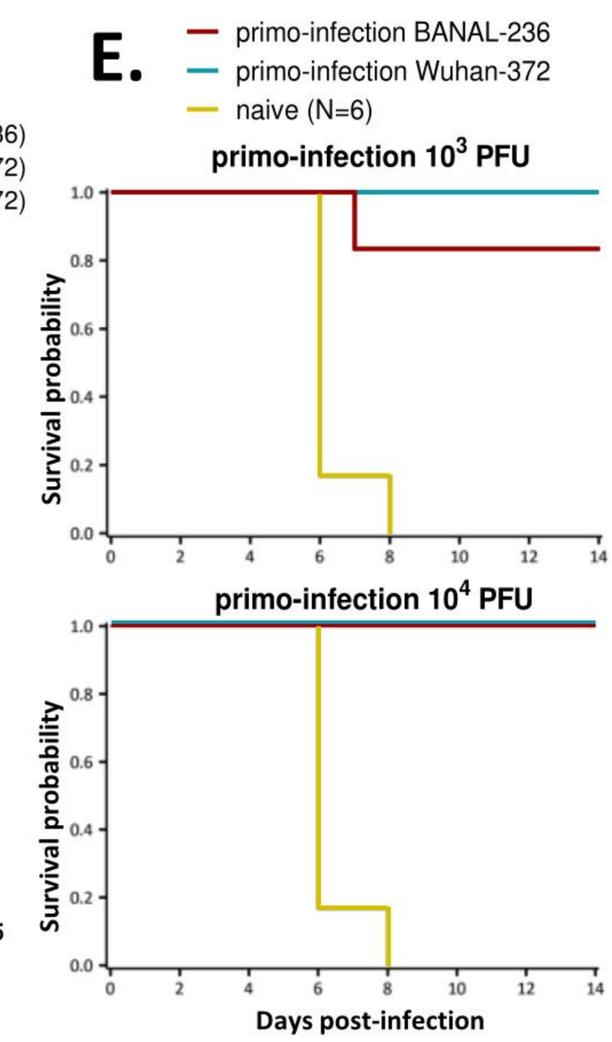
(A-C) Results of Pseudoneutralisation (PNT), Luciferase immunoprecipitation (LIPS) and LuLISA tests are shown. Pre-pandemic French sera (black) were used as negative controls and Laotian sera samples from confirmed SARS-CoV-2 infection (green) were used as cross-reacting positive controls. Laotian sera samples collected in the general population before 2019 (blue, n=100) or late 2020 (gray, n=100) or in people exposed to bats (purple, n=74) were tested for BANAL-236 (**A**), BANAL-52 (**B**) and BANAL-103 (**C**) antibody responses. The ANOVA non-parametric Kruskal-Wallis test was conducted to compare each sub-population to the reference French population (* P<0.05; **P<0.005; ***P<0.0005; ****P<0.0001). The red bars represent the median.

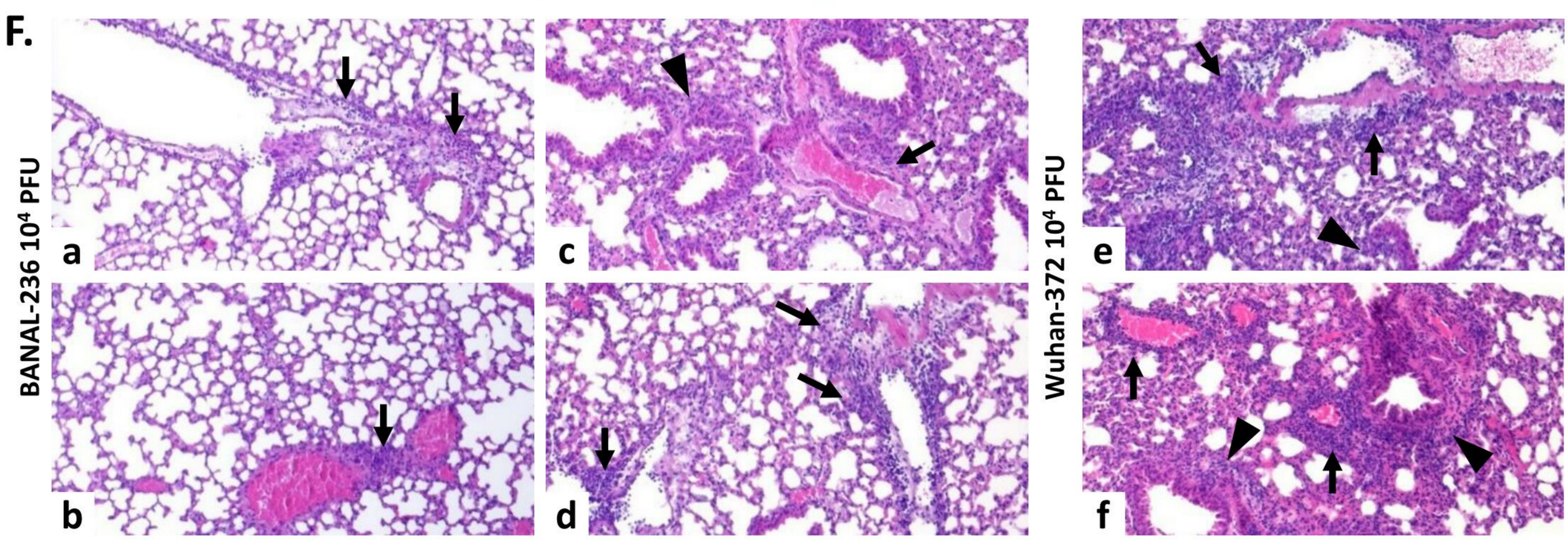

A.
BANAL-236 C3
(VeroE6)
10³ or 10⁴ PFU
x10
hACE2
x4
D+3
D+30
SARS-CoV-2
Wuhan-D614G
10⁴ PFU
x6
D+14
SARS-CoV-2
Wuhan-372
10³ or 10⁴ PFU
BANAL-236 C3
(VeroE6)
10⁴ PFU
x8

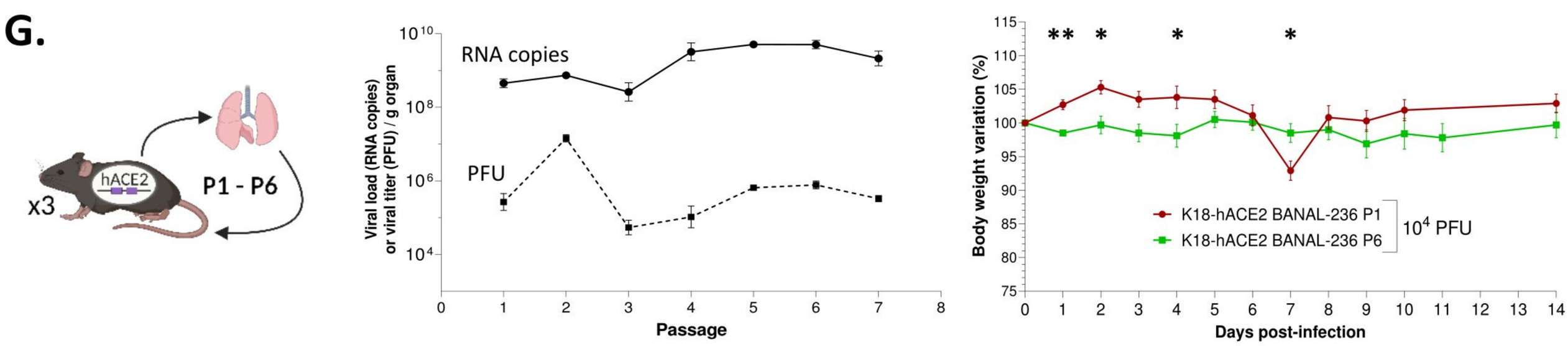

B.
K18-hACE2 BANAL-236
K18-hACE2 Wuhan-372
10³ PFU
K18-hACE2 BANAL-236
K18-hACE2 Wuhan-372
BALB/c BANAL-236
10⁴ PFU
Body weight variation (%)
Days post-infection


C.
K18-hACE2 BANAL-236
K18-hACE2 Wuhan-372
BALB/c BANAL-236
10⁴ PFU
Viral load (RNA copies) or viral titer (PFU) / g organ


D.
K18-hACE2 BANAL-236 (PNT-236)
K18-hACE2 BANAL-236 (PNT-372)
K18-hACE2 Wuhan-372 (PNT-372)
BALB/c BANAL-236 (PNT-236)
Challenge
Neutralizing titer (log10(ED50))
Days post-infection

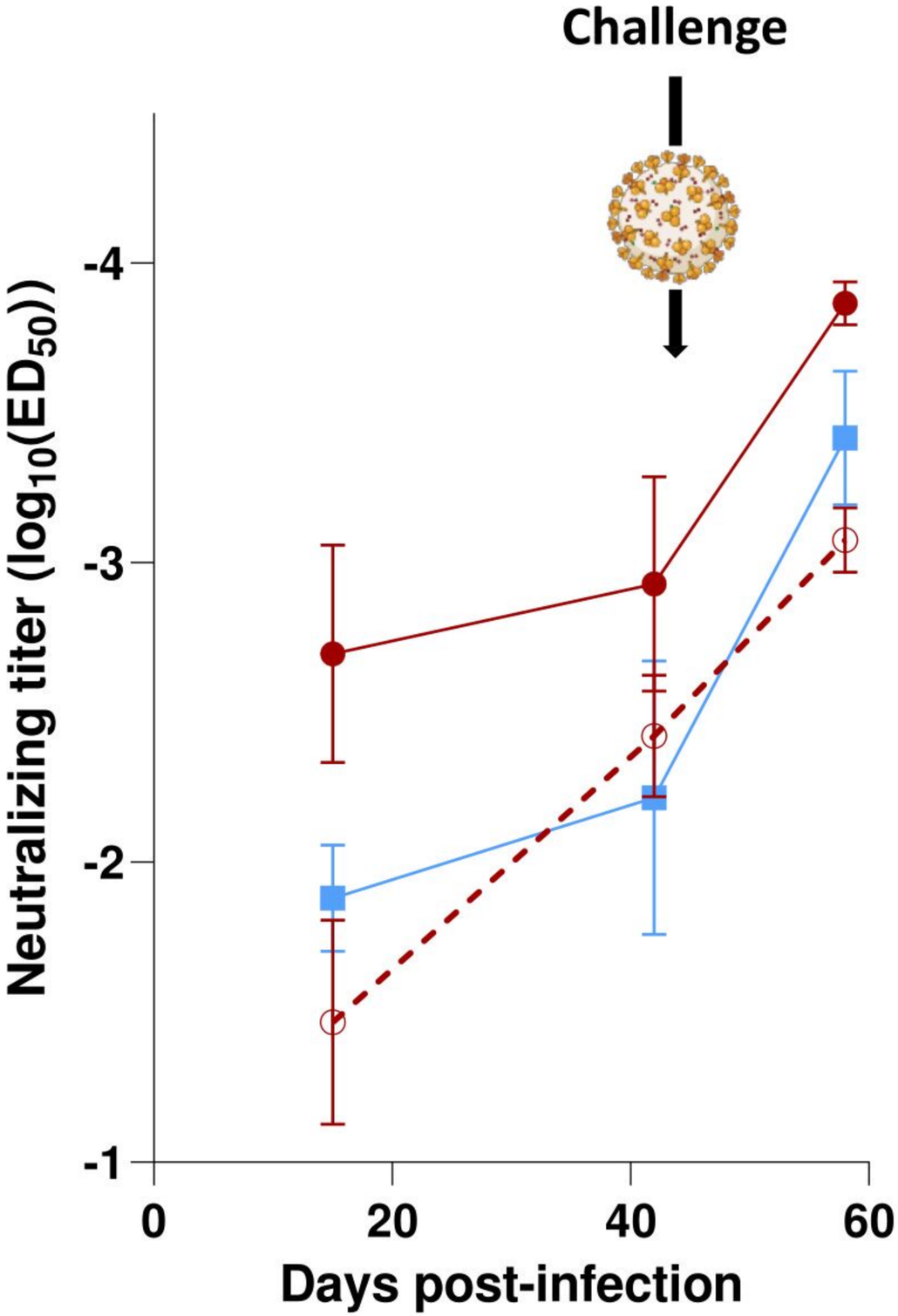

E.
primo-infection BANAL-236
primo-infection Wuhan-372
naive (N=6)
primo-infection 10³ PFU
primo-infection 10⁴ PFU
Survival probability
Days post-infection


F.
BANAL-236 10⁴ PFU
Wuhan-372 10⁴ PFU
a
b
c
d
e
f


G.
hACE2
x3
P1 - P6
RNA copies
PFU
Viral load (RNA copies) or viral titer (PFU) / g organ
Passage
K18-hACE2 BANAL-236 P1
K18-hACE2 BANAL-236 P6
10⁴ PFU
Body weight variation (%)
Days post-infection

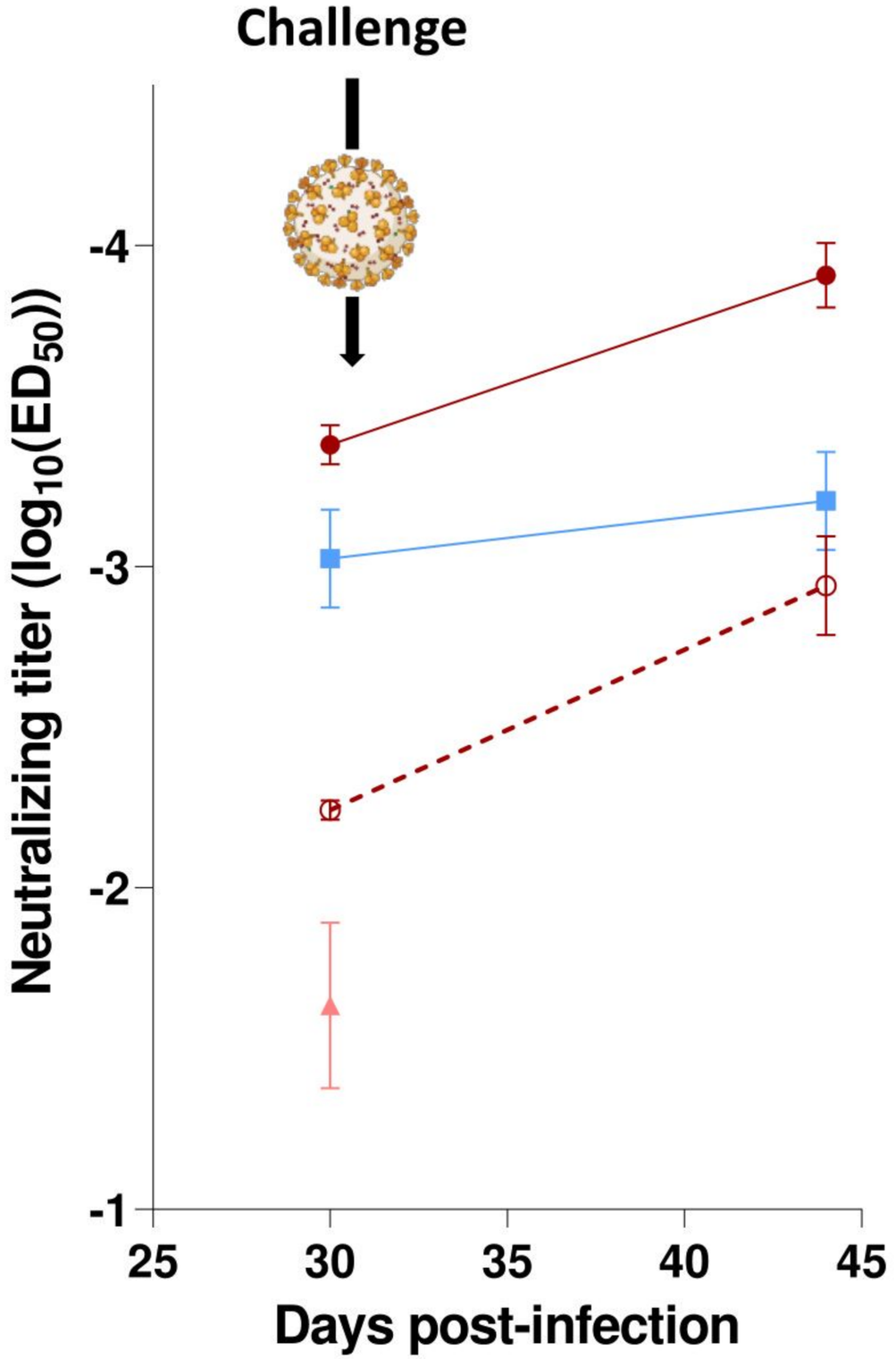
Infection at 10^3 PFU

- K18-hACE2 BANAL-236 (PNT-236)
- K18-hACE2 BANAL-236 (PNT-372)
- K18-hACE2 Wuhan-372 (PNT-372)

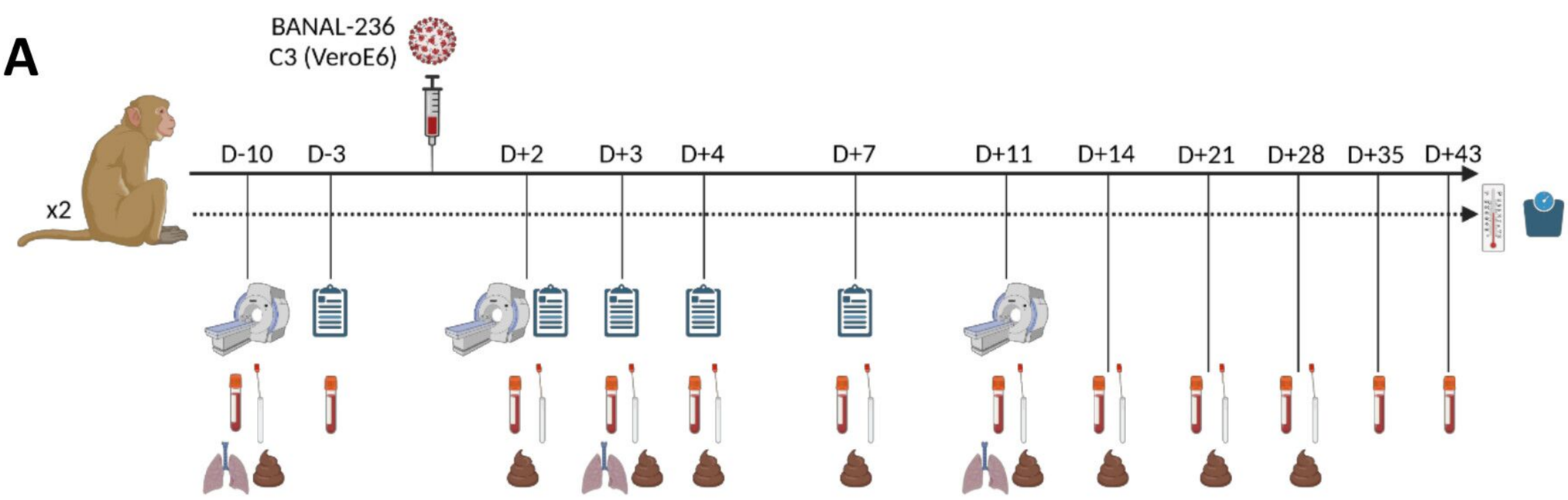


Infection at 10^4 PFU

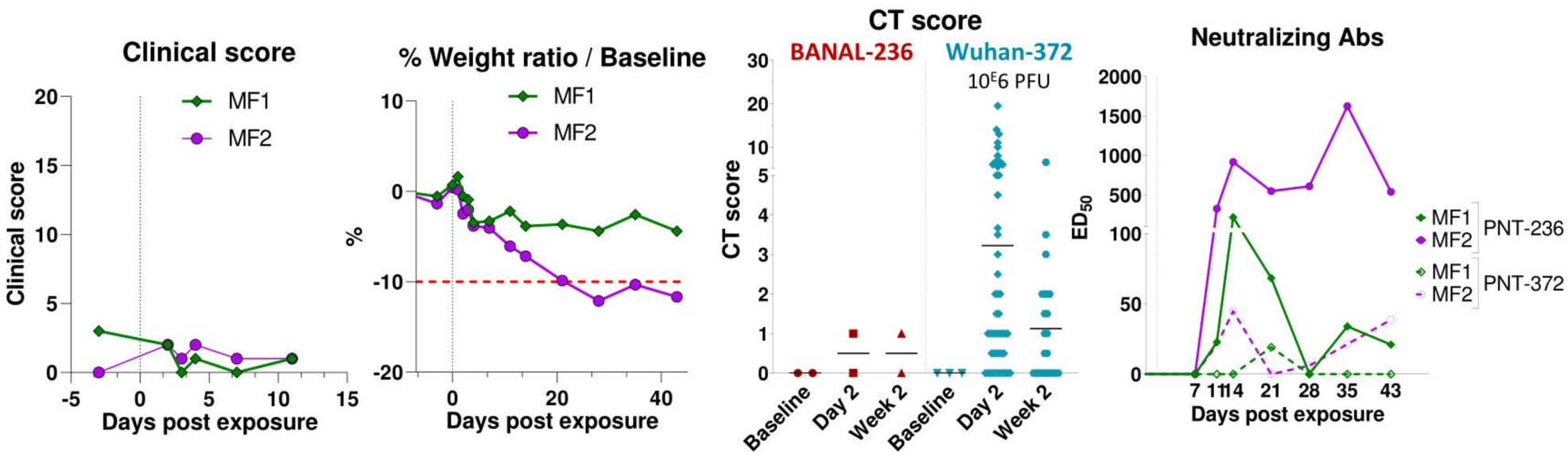
- K18-hACE2 BANAL-236 (PNT-236)
- K18-hACE2 BANAL-236 (PNT-372)
- K18-hACE2 Wuhan-372 (PNT-372)
- BALB/c BANAL-236 (PNT-236)



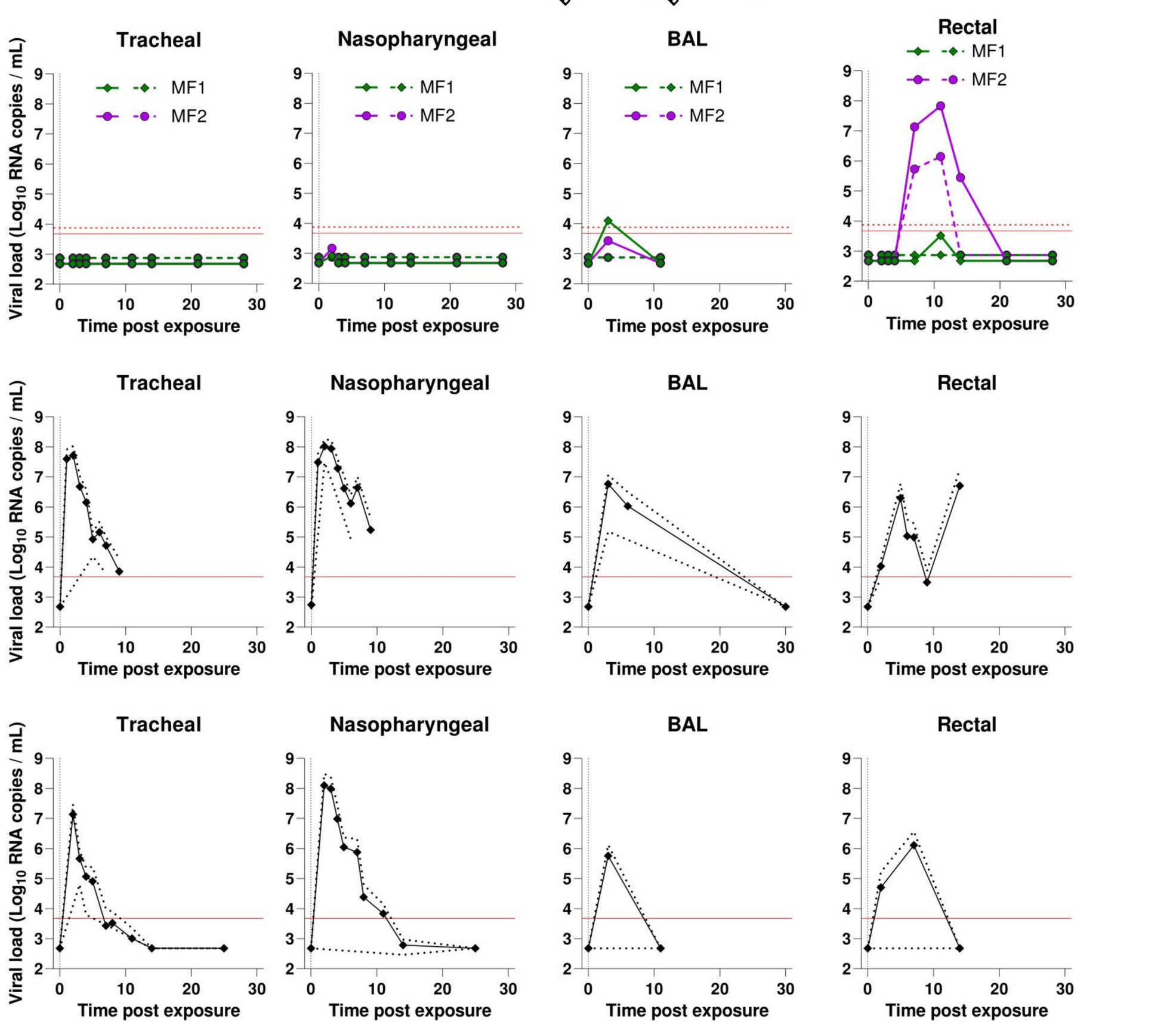
A


BANAL-236
C3 (VeroE6)
x2
D-10
D-3
D+2
D+3
D+4
D+7
D+11
D+14
D+21
D+28
D+35
D+43

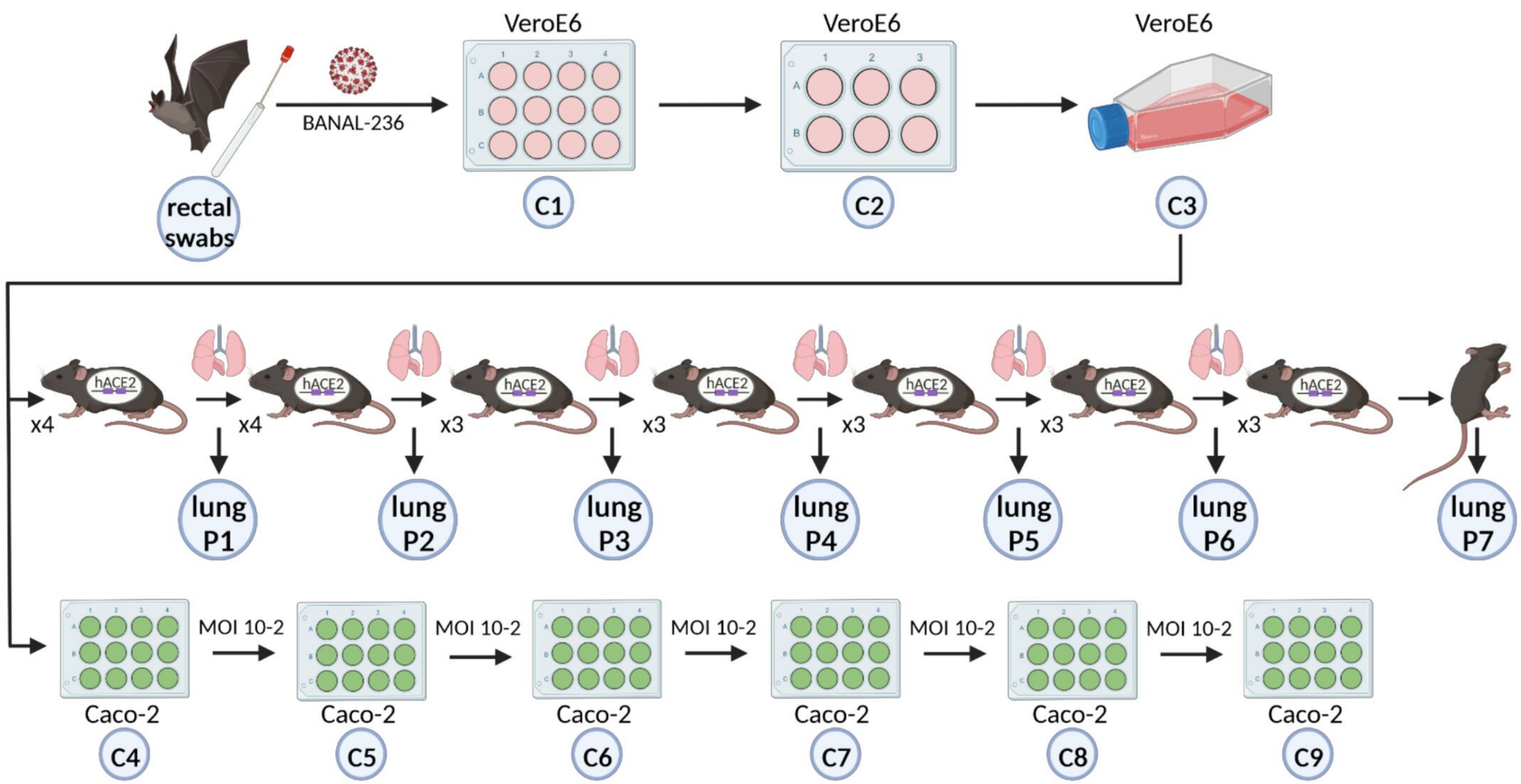
B


Clinical score
MF1
MF2
Clinical score
Days post exposure
% Weight ratio / Baseline
%
CT score
BANAL-236
Wuhan-372
10E6 PFU
Baseline
Day 2
Week 2
Neutralizing Abs
ED50
PNT-236
PNT-372
Days post exposure

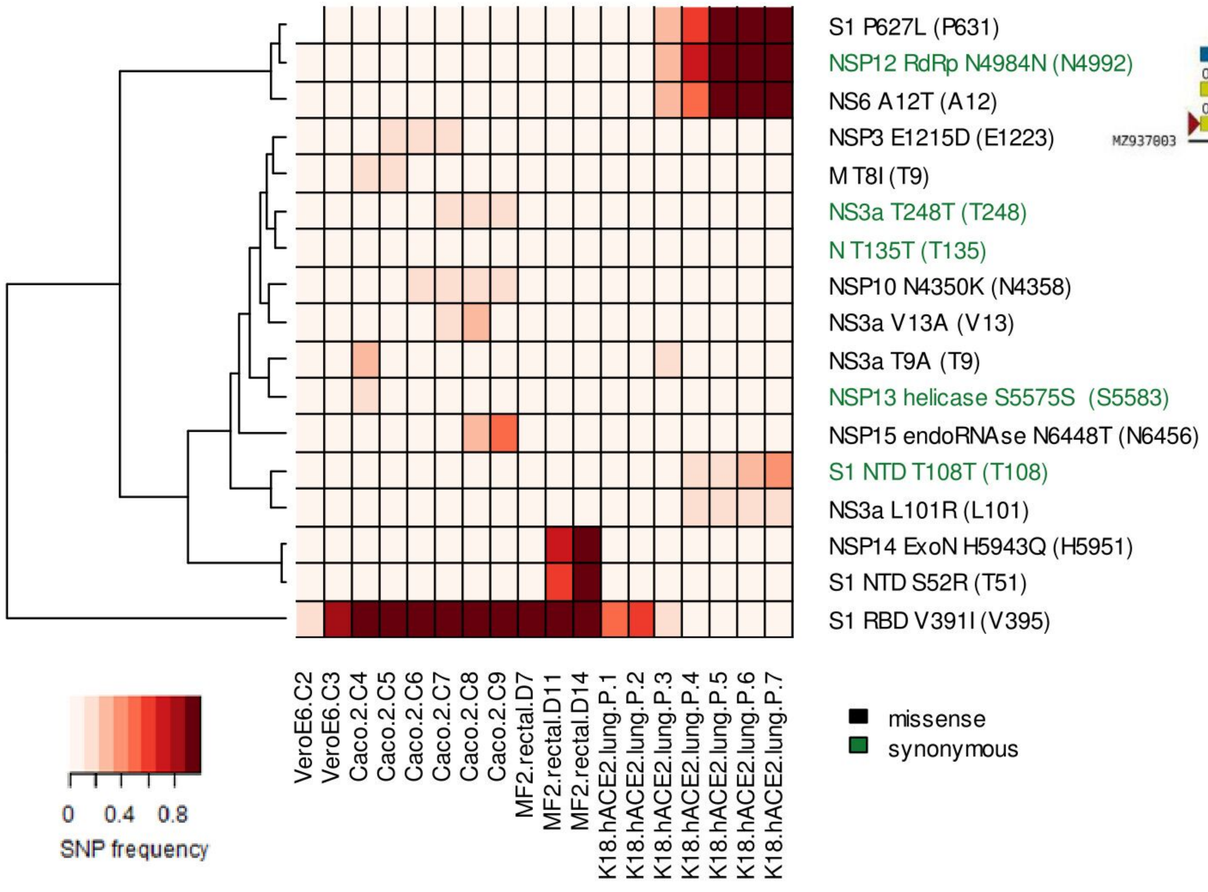
C


BANAL-236
0,6.10E6 PFU
Wuhan-372
10E6 PFU
Wuhan-372
10E5 PFU
Tracheal
Nasopharyngeal
BAL
Rectal
MF1
MF2
Viral load (Log10 RNA copies / mL)
Time post exposure

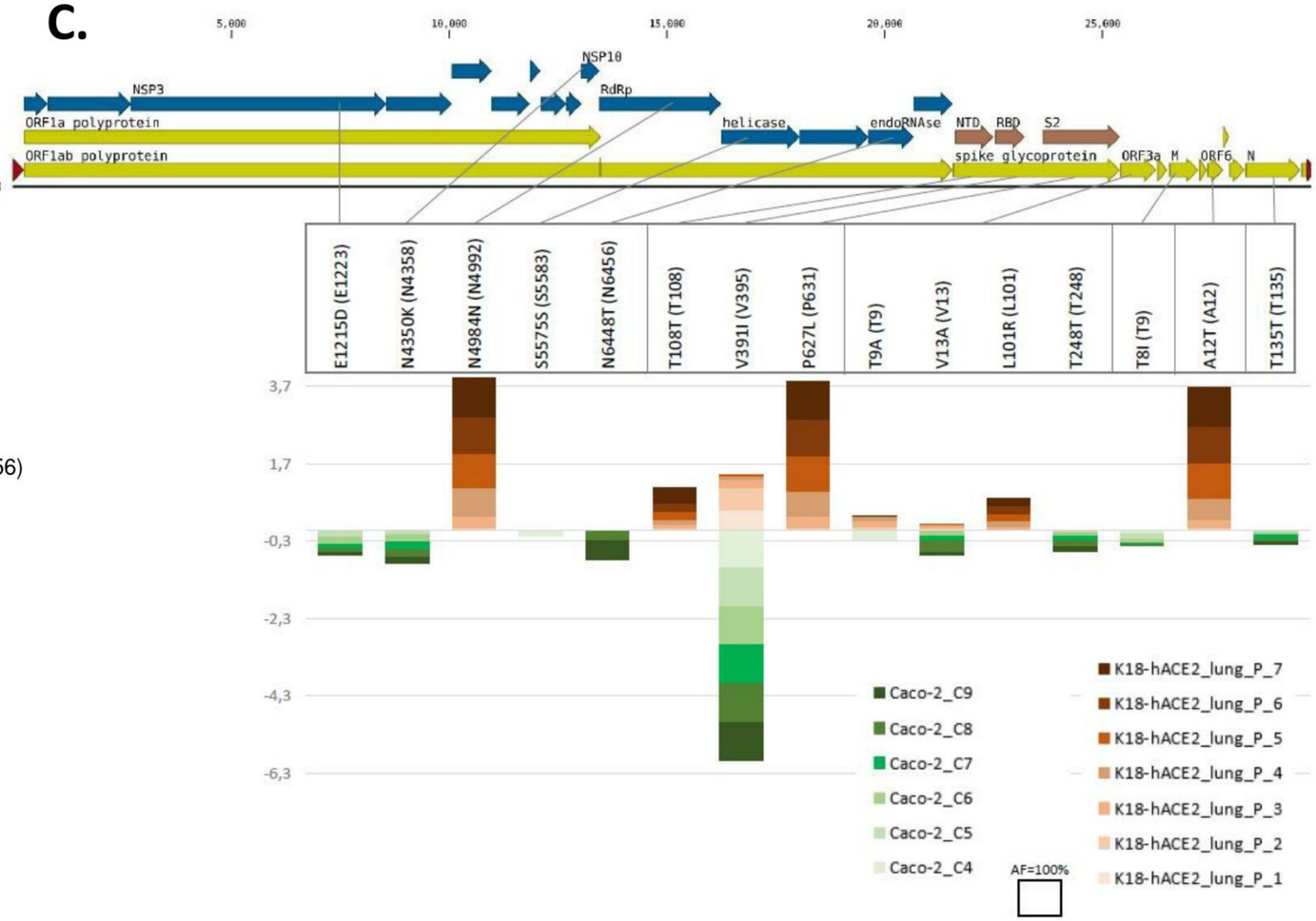
A.


VeroE6
VeroE6
VeroE6
BANAL-236
rectal swabs
C1
C2
C3
hACE2
x4
x4
x3
x3
x3
x3
x3
lung P1
lung P2
lung P3
lung P4
lung P5
lung P6
lung P7
MOI 10-2
MOI 10-2
MOI 10-2
MOI 10-2
MOI 10-2
Caco-2
Caco-2
Caco-2
Caco-2
Caco-2
Caco-2
C4
C5
C6
C7
C8
C9

B.

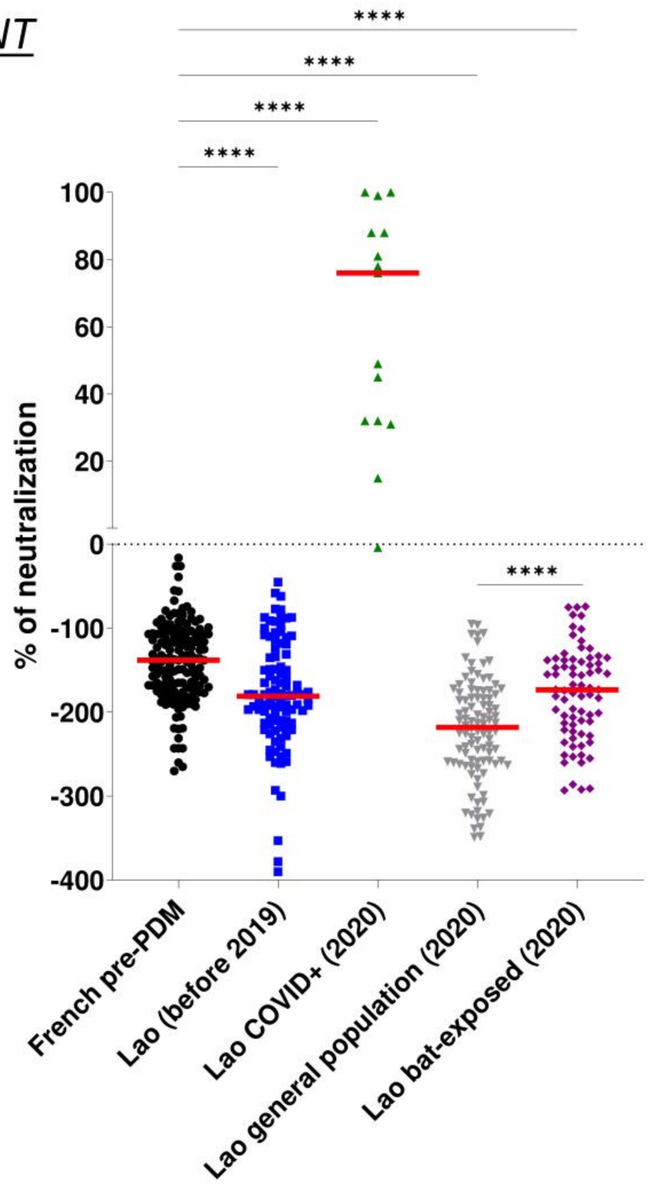

S1 P627L (P631)
NSP12 RdRp N4984N (N4992)
NS6 A12T (A12)
NSP3 E1215D (E1223)
M T8I (T9)
NS3a T248T (T248)
N T135T (T135)
NSP10 N4350K (N4358)
NS3a V13A (V13)
NS3a T9A (T9)
NSP13 helicase S5575S (S5583)
NSP15 endoRNAse N6448T (N6456)
S1 NTD T108T (T108)
NS3a L101R (L101)
NSP14 ExoN H5943Q (H5951)
S1 NTD S52R (T51)
S1 RBD V391I (V395)
VeroE6.C2
VeroE6.C3
Caco.2.C4
Caco.2.C5
Caco.2.C6
Caco.2.C7
Caco.2.C8
Caco.2.C9
MF2.rectal.D7
MF2.rectal.D11
MF2.rectal.D14
K18.hACE2.lung.P.1
K18.hACE2.lung.P.2
K18.hACE2.lung.P.3
K18.hACE2.lung.P.4
K18.hACE2.lung.P.5
K18.hACE2.lung.P.6
K18.hACE2.lung.P.7
0
0.4
0.8
SNP frequency
missense
synonymous

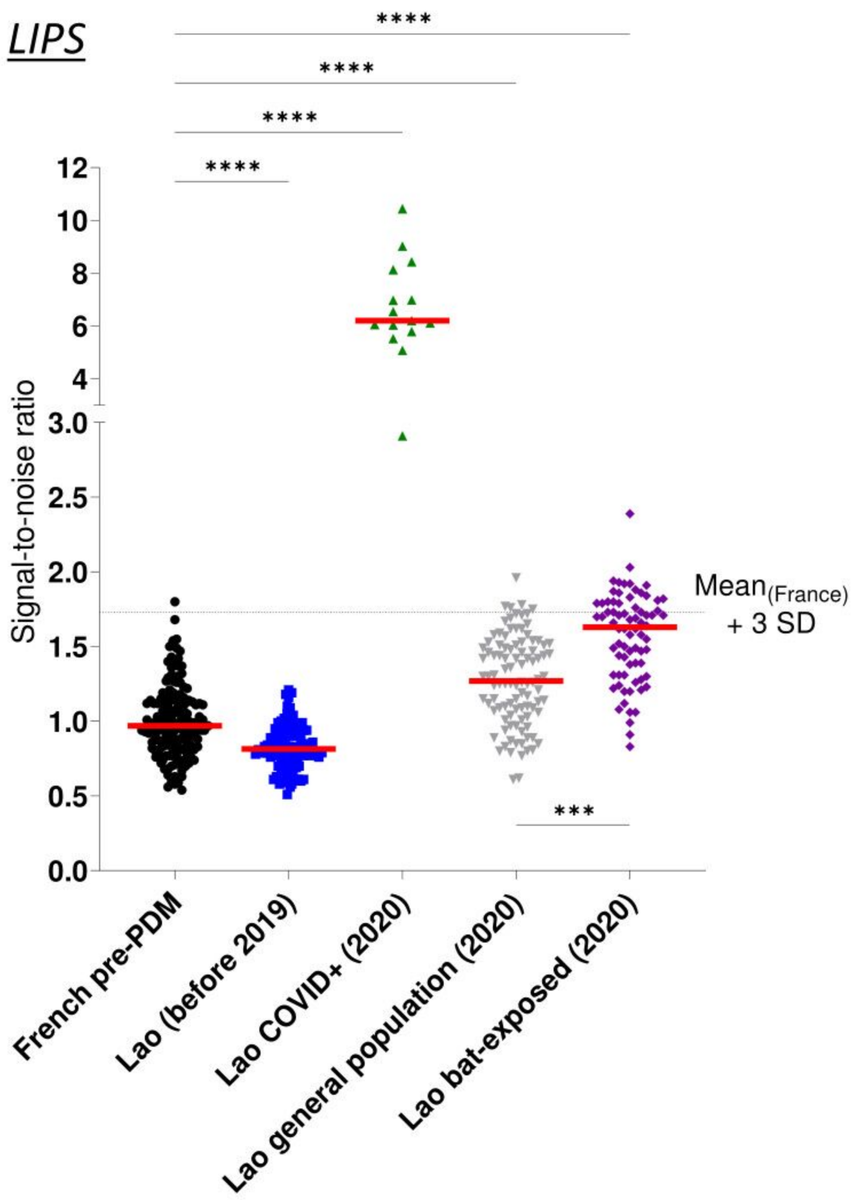
C.

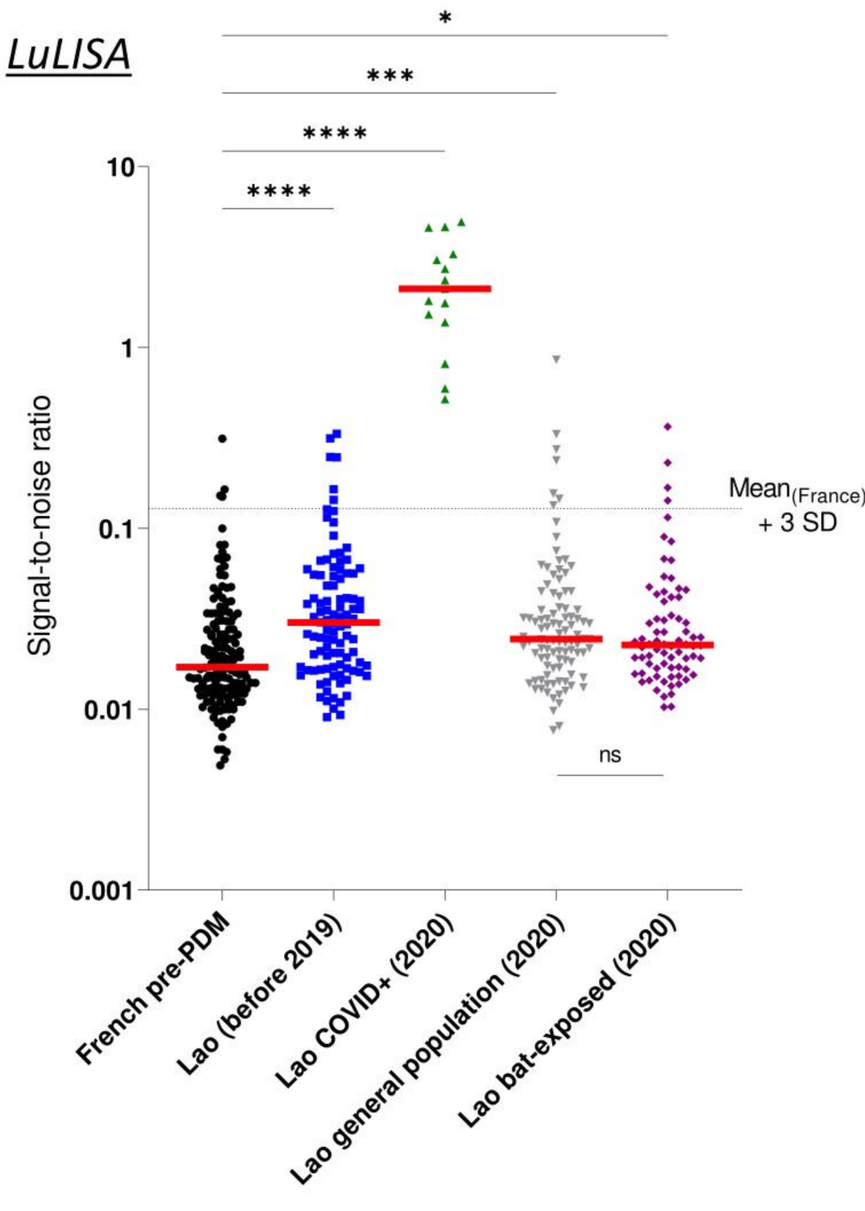

5,000
10,000
15,000
20,000
25,000
NSP10
NSP3
RdRp
ORF1a polyprotein
helicase
endoRNAse
NTD
RBD
S2
ORF1ab polyprotein
spike glycoprotein
ORF3a
M
ORF6
N
MZ937003
E1215D (E1223)
N4350K (N4358)
N4984N (N4992)
S5575S (S5583)
N6448T (N6456)
T108T (T108)
V391I (V395)
P627L (P631)
T9A (T9)
V13A (V13)
L101R (L101)
T248T (T248)
T8I (T9)
A12T (A12)
T135T (T135)
3,7
1,7
-0,3
-2,3
-4,3
-6,3
Caco-2_C9
Caco-2_C8
Caco-2_C7
Caco-2_C6
Caco-2_C5
Caco-2_C4
K18-hACE2_lung_P_7
K18-hACE2_lung_P_6
K18-hACE2_lung_P_5
K18-hACE2_lung_P_4
K18-hACE2_lung_P_3
K18-hACE2_lung_P_2
K18-hACE2_lung_P_1
AF=100%

A.

BANAL-236

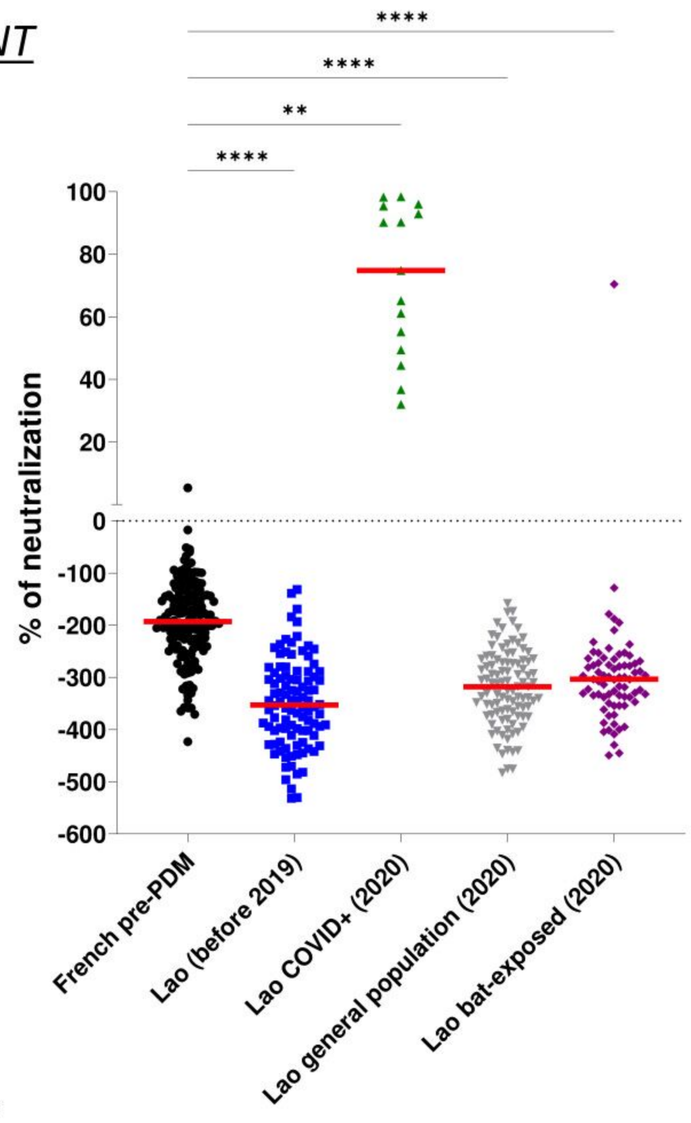


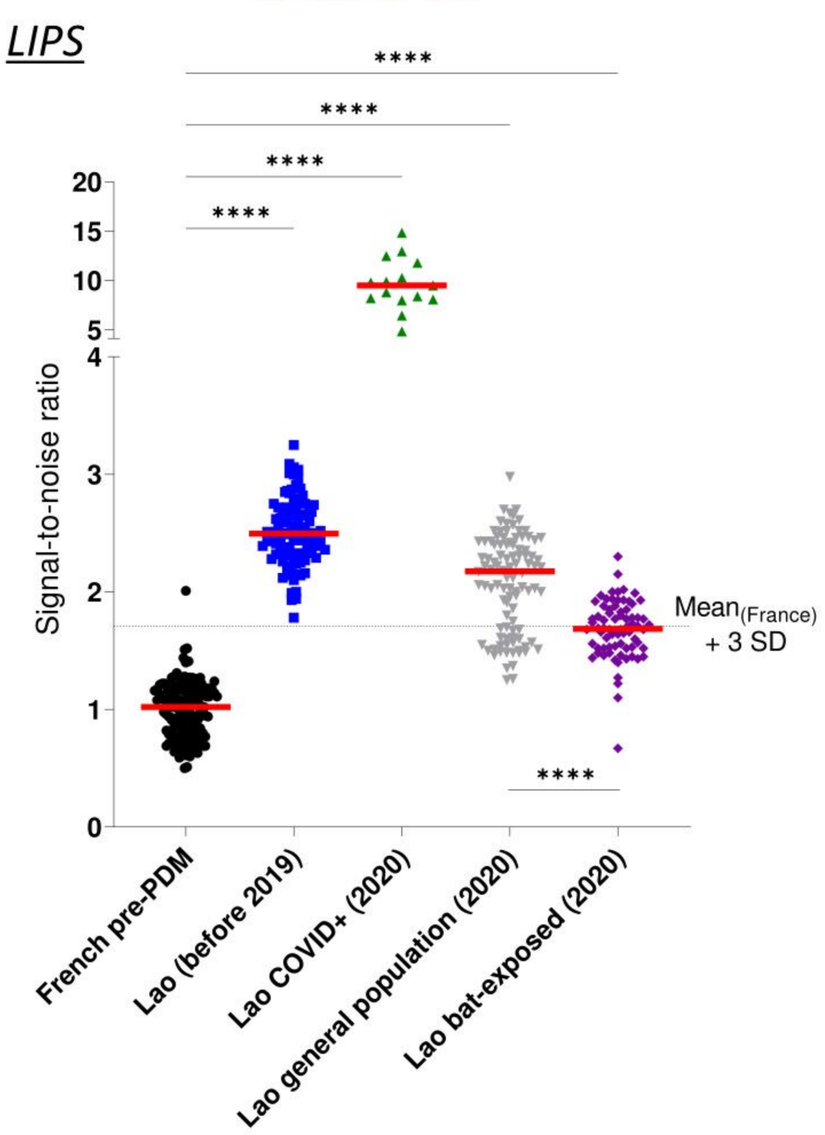


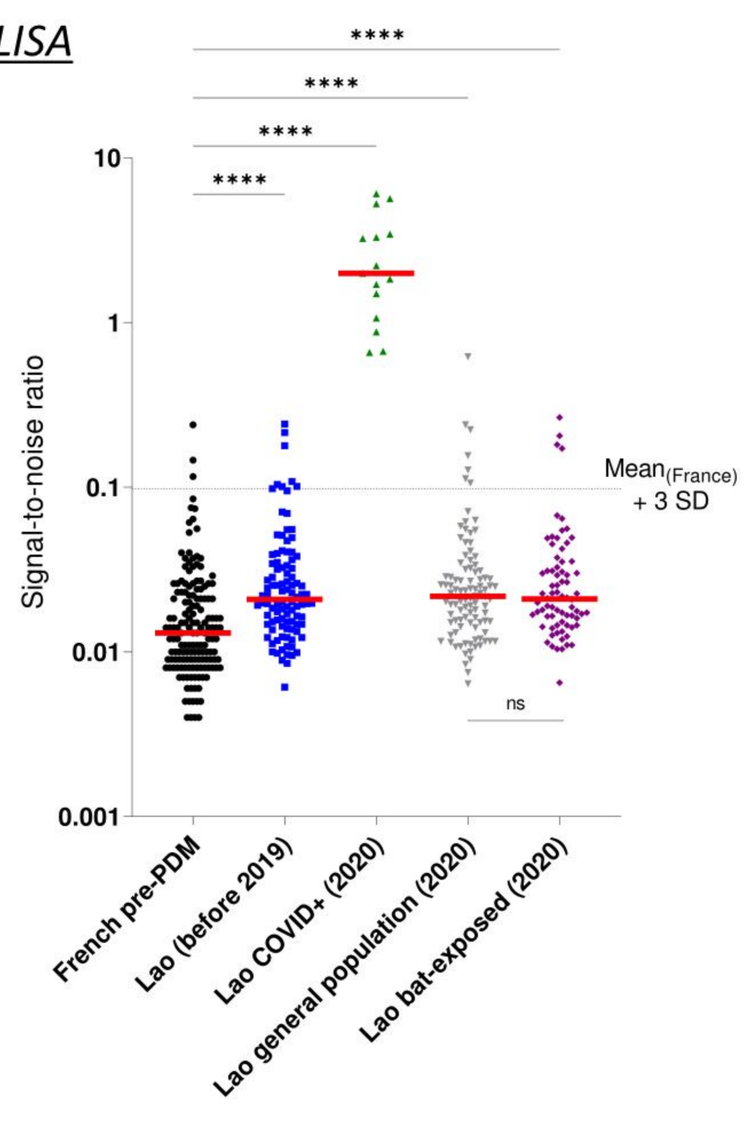


B.

BANAL-52







C.

BANAL-103