A sarbecovirus found in Russian bats uses human ACE2

Stephanie N. Seifert¹ and Michael C. Letko¹#

1. Paul G. Allen School for Global Health, Washington State University, Pullman, WA, U.S.

# Corresponding authors:
Michael Letko, Ph.D.
Paul G. Allen School for Global Health
Washington State University
PO Box 647090
Pullman, WA 99164-7090
Email: michael.letko@wsu.edu
ABSTRACT

Spillover of sarbecoviruses from animals to humans resulted in outbreaks of severe acute respiratory syndrome SARS-CoVs and the ongoing COVID-19 pandemic. Efforts to identify the origins of SARS-CoV-1 and -2 has resulted in the discovery of numerous animal sarbecoviruses – the majority of which are only distantly related to known human pathogens and do not infect human cells. The receptor binding domain (RBD) on sarbecoviruses engages receptor molecules on the host cell and mediates cell invasion. Here, we tested the receptor tropism for RBDs from two sarbecoviruses found in Russian horseshoe bats to screen cell entry. While these two viruses are in a viral lineage distinct from SARS-CoV-1 and -2, one virus was capable of using human ACE2 to facilitate cell entry. Our findings demonstrate that sarbecoviruses circulating in wildlife outside of Asia also exhibit compatibility with human ACE2 and should be taken into consideration for future universal sarbecovirus vaccine candidates.
Zoonotic spillover of sarbecoviruses from animals to humans has led to the emergence of highly pathogenic human viruses, SARS-CoV-1 and -2, with the later leading to the largest global pandemic in modern history. Researchers around the world are ramping up the pace of viral discovery efforts, expanding the sequence databases with new animal sarbecoviruses in circulation. While some laboratory experiments have been performed with these new viruses, demonstrating a range of host tropisms, several viruses remain untested, and thus their ability to transmit to humans is unknown.

Coronaviruses are covered with a “spike” glycoprotein that engages with receptor molecules on the surface of host cells and mediates viral infection of the cell. A small region within the spike proteins of sarbecoviruses, known as the receptor binding domain (RBD), contains all of the structural information necessary to engage with the host receptor. We and others have experimentally classified the majority of published sarbecovirus RBDs into different clades based on sequence and functional data: clade 1, identified in Asian bats, contains no deletions and binds to host receptor, Angiotensin-Converting Enzyme 2 (ACE2), whereas clade 2, also identified in Asian bats, contains 2 deletions and does not use ACE2 and clade 3 viruses, found more widely in Africa and Europe, contain 1 deletion and also infect using an unknown receptor\(^1-3\). Recently, several viruses were identified in China that comprise a fourth clade that also interact with human ACE2\(^4\).

In late 2020, two clade 3 sarbecoviruses were identified in Rhinolophus bats in Russia: Khosta-1 was found in *Rhinolophus ferrumequinum* and Khosta-2 in *R. hipposideros*\(^5\). Similar to other European and African clade 3 viruses, the Khosta viruses
are divergent from the RBD found in SARS-CoV-1 and -2, and were thus presumed not to possess any tropism for human cells. However, a recent study demonstrated that a new clade (clade 4) of sarbecoviruses found in China containing a deletion similar to clade 3 sarbecoviruses can bind to human ACE2\(^4\). Here, we tested the RBDs from Russian sarbecoviruses, along with other clade 3 viruses and report that the virus, Khosta-2, is capable of infecting cells using the human orthologue of ACE2.

Results

Khosta virus receptor binding domains are distinct from human viruses

Khosta-1 and -2 were identified by Alkhovsky and colleagues in bat samples collected between March-October 2020 near Sochi National Park\(^5\). Phylogenetic analysis of the conserved viral gene, Orf1ab, revealed these viruses were most closely related to another sarbecovirus found in Bulgaria in 2008 (known as BM48-31 or Bg08), and form a lineage sarbecoviruses distinct from human pathogens, SARS-CoV-1 and -2\(^5\). A list of viruses and accession numbers used in this study can be found in table 1. Phylogenetic analysis of the spike RBD further reflected the close relatedness between Khosta -1 and -2 with BM48-31 and other clade 3 RBD viruses we have previously tested from Uganda and Rwanda\(^1,6\) (Fig. 1a). Clade 3 RBDs, including the Khosta viruses, contain a truncated surface-exposed loop, as compared to the ACE2-dependent, clade 1 viruses such as SARS-CoV and additionally vary in most of the residues known for clade 1 viruses to interact with ACE2\(^1,2,6,7\).
RBD from Khosta viruses mediate entry into human cells

Using our scalable Sarbecovirus RBD entry platform, we replaced the RBD from SARS-CoV-1 spike with the Khosta RBDs and generated chimeric spike expression plasmids (Fig. 1b). For comparison, we also included chimeric RBD spikes for other clade 3 RBDs we have previously tested (BM48-31, Uganda, Rwanda) as well as SARS-CoV-2 and related RatG13 viruses. These chimeric spike expression constructs were used to produce BSL2-compatible viral reporter pseudotypes with Vesicular Stomatitis Virus expressing a dual GFP-luciferase reporter. All of the chimeric spike proteins expressed to similar levels in mammalian cells and incorporated in Vesicular Stomatitis Virus (VSV).

Chimeric spike with the RBD from BM48-31 and RatG13 showed reduced incorporation but this did not correlate with viral entry phenotypes observed in later experiments (Fig. 1c, e).

To test human cell compatibility, we first infected the human liver cell line, Huh-7, with pseudotypes bearing the chimeric Khosta RBD spikes. In the absence of the addition of an exogenous protease, trypsin, the pseudotypes exhibited almost no entry in these cells, which has been observed for other sarbecoviruses and is attributed to low endogenous expression of ACE2. However, when trypsin was included during the infection, entry signal strongly increased for SARS-CoV-1 and -2 RBDs as well as the Khosta RBDs (Fig. 1d). As we and others have shown previously, trypsin enhancement of sarbecovirus entry is still receptor dependent, suggesting that the Khosta virus RBDs were using a receptor present in human cells to mediate infection.
The RBD from Khosta-2 infects cells using human ACE2

To characterize potential receptors that the Khosta viruses may be using, we performed a classic receptor tropism test, where we transfected Baby Hamster Kidney (BHK) cells with human orthologues of known coronavirus receptors and then infected with our pseudotype panel. Unlike 293T cells, which express low levels of human ACE2 and potentially other coronavirus receptors and have been shown to have low but measurable permissivity to SARS-CoV infection, BHK cells are generally considered completely non-permissive for sarbecoviruses unless a suitable receptor is supplemented. The Khosta-1 RBD failed to infect cells expressing any of the receptors, while Khosta 2 RBD clearly infected cells expressing human ACE2 (figure 1e). The level of cell entry mediated by the Khosta 2 RBD was similar to RatG13, a bat sarbecovirus with high similarity to SARS-CoV-2 in the RBD that also binds human ACE2, albeit with lower efficiency than the human pathogen (Fig. 1e). In contrast to the ACE2 results, only HCoV-229E could infect cells expressing Aminopeptidase N (APN), and MERS-CoV spike could only infect cells expressing dipeptidyl peptidase IV (DPP4) – the known receptors for these viruses, demonstrating these receptors were expressed to functional levels (Fig. 1e).

Discussion

To date, all animal derived sarbecoviruses capable of infecting cells through human ACE2 have been found in Asia. Here, we provide evidence that animal-derived sarbecoviruses circulating in wildlife outside of Asia can also possess the capacity to infect human cells and use human ACE2 (Fig. 1). Khosta 1 and 2 viruses are most closely related to other clade 3 RBD viruses, which have been found across a much wider
geographic range than the clade 1 viruses\textsuperscript{1,6,16,17}. Thus, our findings show that sarbecovirus host receptor compatibility is not geographically locked and that other sarbecoviruses, which may only be distantly related to human SARS-CoVs, may also stochastically possess the capacity to infect humans. Importantly, as the researchers who initially discovered the Khosta viruses note with their findings: the Khosta bat sarbecoviruses are genetically distinct from human SARS-CoVs in that they lack genetic information encoding for some of the genes thought to antagonize the immune system and contribute to pathogenicity, such as Orf8\textsuperscript{5}. Unfortunately, because coronaviruses are known to recombine in co-infected hosts, the recent identification of SARS-CoV-2 spillover from humans back in wildlife populations opens the possibility of new human-compatible sarbecoviruses\textsuperscript{18-23}.

In the presence of trypsin, both Khosta-1 and -2 RBDs were capable of infecting human cells, with Khosta-1 performing notably stronger than Khosta-2, however in our receptor-specific assay, only Khosta-2 could infect cells expressing human ACE2 without exogenous protease (Fig. 1d, e). We have previously shown that a small number of clade 2 RBDs, such as As6526, also exhibit protease-mediated, ACE2-independent entry, and similar phenotypes have been described for other bat coronaviruses\textsuperscript{1,24}. Because not all of the clade 2 and 3 viruses exhibit this phenotype, these findings collectively suggest that other coronaviruses can infect human cells through a presently unknown receptor. Sarbecoviruses have been shown to co-circulate in their reservoirs, so this variation in receptor usage among closely related viruses may even represent an evolutionary strategy for viral persistence\textsuperscript{2}. 
Although the Khosta-2 RBD exhibited clear compatibility with human ACE2, pseudotype entry was approximately 10-fold less than for SARS-CoV-1 and -2, and was to levels closer to RatG13 RBD (Fig. 1e). The RBD for RatG13, while also compatible with human ACE2, has been shown by other groups to exhibit poor binding to human ACE2. A single point mutation, T403R/K, has been shown to improve RatG13-RBD binding and entry with human ACE2, although this mutation is already present in Khosta-2 (K403). We have previously demonstrated that MERS-CoV can rapidly adapt to overcome a semi-permissive receptor, suggesting that limited entry phenotypes are easily overcome by betacoronaviruses. As new sarbecoviruses are discovered in wildlife, there is a propensity in the research community to compare single amino acids based on known human SARS-CoVs as an indicator for compatibility with humans. Our work underscores the challenges in predicting viral function from sequence data alone and sheds light on another group of sarbecoviruses with zoonotic potential. Current universal sarbecovirus vaccines in development include mostly clade 1 viruses and one of the clade 2 viruses but do not include any members from clade 3. Taken together, our findings with the Khosta viruses show that the sarbecoviruses may pose an even broader geographic threat than previously appreciated.

Methods

Phylogenetic analysis

Genbank accession numbers for all sarbecovirus spike sequences used in this study are available in table 1. Amino acid sequences for the receptor binding domain of the spike glycoprotein were aligned using ClustalW multiple sequence alignment with default
parameters. A maximum likelihood phylogenetic tree was inferred with PhyML v. 3.0\textsuperscript{31} using the ‘WAG’ matrix +G model of amino acid substitution as selected by Smart Model Selection method with 1000 bootstrap replicates\textsuperscript{32}. The final tree was then visualized as a cladogram with FigTree v1.4.4 (https://github.com/rambaut/figtree).

Plasmids and sequences

Untagged human orthologues of ACE2 (Q9BYF1.2), APN (NP_001141.2), and DPP4 (XM_005246371.3) were described previously\textsuperscript{1}. Spike sequences from HCoV-229E (AB691763.1), MERS-CoV (JX869059.2), and SARS-CoV-1 (AY278741) were codon-optimized, appended with a carboxy-terminal FLAG tag sequence separated by a flexible poly-glycine linker and cloned into pcDNA3.1+ as previously described\textsuperscript{1}. The SARS-CoV-1 RBD was removed with KpnI and Xhol and replaced with codon-optimized, synthesized RBD fragments as previously described\textsuperscript{1}.

Cells and pseudotype assay

293T, Huh-7 (human liver cells), and BHKs were maintained under standard cell culture conditions in DMEM with L-glutamine, antibiotics, and 10% Fetal Bovine Serum. Single-cycle, Vesicular Stomatitis Virus (VSV) pseudotype assays were performed as previously described\textsuperscript{1}. Briefly, 293T “producer cells” were transfected with spike plasmids or empty vector as a “no spike” control and infected 24-hours later with VSV-ΔG-luc/GFP particles, generating chimeric-spike pseudotyped particles that were harvested 72 hours post-transfection and stored at -80°C. Target cells were plated in 96-well format, and spin-infected in quadruplicate with equivalent volumes of viral pseudotypes at 1200xG for 1
hour at 4°C. Infected cells were incubated for approximately 18-20 hours and luciferase was measured using the Promega BrightGlo luciferase kit following manufacturers’ instructions (Promega). Entry signal was normalized to the average signal for the “no spike” control. Plates were measured and analyzed in triplicate. Data are representative of four complete biological replicates.

Westernblot

Viral pseudotypes were concentrated and 293T producer cells were lysed in 1% SDS and clarified as described previously. Lysates were analyzed on 10% Bis-Tris PAGE-gel (Thermo Fisher Scientific) and probed for FLAG (Sigma-Aldrich; A8592; 1:10000); GAPDH (Sigma-Aldrich, G8795, 1:10000); and/or VSV-M (Kerafast, 23H12, 1:5000). Signal was detected using SuperSignal West substrate (Thermo-Fisher).

AUTHOR CONTRIBUTIONS STATEMENT
S.N.S. and M.L. performed experiments and wrote the manuscript. M.L. designed the study, analyzed the data, and assembled the figures.

ADDITIONAL INFORMATION
The authors declare no competing interests.
**Figure 1.** Khosta-2 uses human ACE2 to infect cells. (a) Sarbecovirus Receptor Binding Domain Cladogram based on amino acid sequenced. Countries of origin and known host receptors are indicated to the right. Clade 1 viruses are shown in red and orange, clade 2 in grey, clade 3 in blue and clade 4 in purple. (b) Diagram of spike constructs used for this study. The SARS-CoV-1 RBD was replaced with RBDs from other sarbecoviruses. (c) Expression and incorporation of viral pseudotypes by western blot. (d) Huh-7 cells were infected with pseudotypes in the presence of absence of trypsin. Cells were infected in triplicate. (e) BHK cells were transfected with receptors and infected in the absence of trypsin. Cells were infected in quadruplicate.
### Table 1. Sarbecovirus sequences used in this study

| Virus       | Accession | Clade | Host species          | Location                                           |
|-------------|-----------|-------|-----------------------|---------------------------------------------------|
| SARS Urbani | AY278741  | 1     | human                 | Guangdong, China (origin of outbreak)             |
| WRV1        | KF367457  | 1     | Rhinolophus sinicus  | Yunnan, China                                    |
| LYRa11      | KF508996  | 1     | Rhinolophus affinis  | Baoshan, Yunnan, China                           |
| Re7327      | KY417191  | 1     | Rhinolophus sinicus  | Kunming, Yunnan Province, China                  |
| Rs4231      | KY417146  | 1     | Rhinolophus sinicus  | Kunming, Yunnan Province, China                  |
| Rs4064      | KY417144  | 1     | Rhinolophus sinicus  | Kunming, Yunnan Province, China                  |
| RaSHC014    | KC881005  | 1     | Rhinolophus sinicus  | Yunnan, China                                    |
| SARS-CoV-2  | MN908947  | 1     | human                 | Wuhan, Hubei, China                              |
| Ra1013      | MN996532  | 1     | Rhinolophus affinis  | Yunnan, China                                    |
| Ra1615      | from preprint | 4 | Rhinolophus affinis | Mojiang County, Yunnan Province, China |
| As6526      | KY417142  | 2     | Aselliscus stoliczkanus | Kunming, Yunnan Province, China              |
| Yunnan2011  | JX935988  | 2     | Chaerophon plicata   | Yunnan, China                                    |
| Shaanxi2011 | JX935987  | 2     | Rhinolophus pusillus | Shaanxi, China                                   |
| 279-2005    | DQ488557  | 2     | Rhinolophus macrotris | Hubei, China                                     |
| Rs4237      | KY417147  | 2     | Rhinolophus sinicus  | Kunming, Yunnan Province, China                  |
| Rs4061      | KY417143  | 2     | Rhinolophus sinicus  | Kunming, Yunnan Province, China                  |
| Rp3         | DQ01615   | 2     | Rhinolophus pearsoni | Nanning, Guangxi, China                          |
| Rs4247      | KY417148  | 2     | Rhinolophus sinicus  | Kunming, Yunnan Province, China                  |
| HKU3-8      | GQ153543  | 2     | Rhinolophus sinicus  | Guangdong, China                                 |
| HKU3-13     | GQ153548  | 2     | Rhinolophus sinicus  | Guangdong, China                                 |
| GX2013      | KJ473615  | 2     | Rhinolophus sinicus  | Guangxi, China                                   |
| Longquan-140| KF294457  | 2     | Rhinolophus monoceros | Longquan, Zhejiang, China                      |
| YN2013      | KJ473616  | 2     | Rhinolophus sinicus  | Yunnan, China                                    |
| R4092       | KY417145  | 2     | Rhinolophus ferrumequinum | Kunming, Yunnan Province, China              |
| ZXC21       | MG772934  | 2     | Rhinolophus sinicus  | Zhoushan Cty, Zhejiang, China                    |
| ZC45        | MG772933  | 2     | Rhinolophus sinicus  | Zhoushan Cty, Zhejiang, China                    |
| JL2012      | KJ473611  | 2     | Rhinolophus ferrumequinum | JiLin, China                                |
| HuB2013     | KJ473614  | 2     | Rhinolophus sinicus  | Hubei, China                                     |
| R41         | DQ412042  | 2     | Rhinolophus ferrumequinum | Yichang, Hubei, China                          |
| HeB2013     | KJ473612  | 2     | Rhinolophus ferrumequinum | HebeI, China                                  |
| 273-2005    | DQ488556  | 2     | Rhinolophus ferrumequinum | Hubei, China                                   |
| BM48-31     | NC014470  | 3     | Rhinolophus blasii   | Strandja Nature Park, Bulgaria                  |
| Uganda      | MT226044  | 3     | Rhinolophus ferrumequinum | Uganda                                    |
| Rwanda      | MT226045  | 3     | Rhinolophus ferrumequinum | Rwanda                                   |
| Khosta 1    | MZ190137  | 3     | Rhinolophus ferrumequinum | Greater Caucasus, Russia                      |
| Khosta 2    | MZ190138  | 3     | Rhinolophus hipposideros | Greater Caucasus, Russia                      |
REFERENCES

1. Letko M, Marzi A, Munster V. Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nature Microbiology* 2020; (4).

2. Hu B, Zeng LP, Yang XL, et al. Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS coronavirus. *PLoS Pathog* 2017; 13(11): e1006698.

3. Starr TN, Greaney AJ, Hilton SK, et al. Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding. *Cell* 2020; 182(5): 1295-310 e20.

4. Guo H, Hu B, Si HR, et al. Identification of a novel lineage bat SARS-related coronaviruses that use bat ACE2 receptor. *Emerg Microbes Infect* 2021; 10(1): e1006698.

5. Wells HL, Letko M, Ssebide B, et al. The evolutionary history of ACE2 usage within the subgenus Sarbecovirus. *Virus Evol* 2021.

6. Hoffmann M, Muller MA, Drexler JF, et al. Differential sensitivity of bat cells to infection by enveloped RNA viruses: coronaviruses, paramyxoviruses, filoviruses, and influenza viruses. *PLoS One* 2013; 8(8): e72942.

7. Shulla A, Heald-Sargent T, Subramanya G, Zhao J, Perlman S, Gallagher T. A transmembrane serine protease is linked to the severe acute respiratory syndrome coronavirus receptor and activates virus entry. *J Virol* 2011; 85(2): 873-82.

8. Peck KM, Cockrell AS, Yount BL, Scobey T, Baric RS, Heise MT. Glycosylation of mouse DPP4 plays a role in inhibiting Middle East respiratory syndrome coronavirus infection. *J Virol* 2015; 89(8): 4696-9.

9. Liu K, Tan S, Niu S, et al. Cross-species recognition of SARS-CoV-2 to bat ACE2. *Proc Natl Acad Sci U S A* 2021; 118(1).

10. Ge XY, Li JL, Yang XL, et al. Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. *Nature* 2013; 503(7477): 535-8.

11. Menachery VD, Yount BL, Jr., Debbink K, et al. A SARS-like cluster of circulating bat coronaviruses shows potential for human emergence. *Nat Med* 2015; 21(12): 1508-13.

12. Wang Y, Shang J, Graham R, Baric RS, Li F. Receptor Recognition by the Novel Coronavirus from Wuhan: an Analysis Based on Decade-Long Structural Studies of SARS Coronavirus. *J Virol* 2020; 94(7).

13. Zhou P, Yang X, Wang X, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 2020.

14. Drexler JF, Gloza-Rausch F, Glende J, et al. Genomic characterization of severe acute respiratory syndrome-related coronavirus in European bats and classification of coronaviruses based on partial RNA-dependent RNA polymerase gene sequences. *J Virol* 2010; 84(21): 11336-49.

15. Tao Y, Tong S. Complete Genome Sequence of a Severe Acute Respiratory Syndrome-Related Coronavirus from Kenyan Bats. *Microbiol Resour Announc* 2019; 8(28).
18. Zhang Q, Zhang H, Gao J, et al. A serological survey of SARS-CoV-2 in cat in Wuhan. *Emerg Microbes Infect* 2020; 9(1): 2013-9.

19. Palmer MV, Martins M, Falkenberg S, et al. Susceptibility of white-tailed deer (Odocoileus virginianus) to SARS-CoV-2. *J Virol* 2021.

20. Oude Munnink BB, Sikkema RS, Nieuwenhuijse DF, et al. Transmission of SARS-CoV-2 on mink farms between humans and mink and back to humans. *Science* 2021; 371(6525): 172-7.

21. Fagre A, Lewis J, Eckley M, et al. SARS-CoV-2 infection, neuropathogenesis and transmission among deer mice: Implications for spillback to New World rodents. *PLoS Pathog* 2021; 17(5): e1009585.

22. Di Guardo G. Susceptibility of white-tailed deer to SARS-CoV-2. *Vet Rec* 2021; 189(10): 408-9.

23. Chandler JC, Bevins SN, Ellis JW, et al. SARS-CoV-2 exposure in wild white-tailed deer (Odocoileus virginianus). *Proc Natl Acad Sci U S A* 2021; 118(47).

24. Menachery VD, Dinnon KH, 3rd, Yount BL, Jr., et al. Trypsin treatment unlocks barrier for zoonotic bat coronaviruses infection. *J Virol* 2019.

25. Zech F, Schniertshauer D, Jung C, et al. Spike residue 403 affects binding of coronavirus spikes to human ACE2. *Nat Commun* 2021; 12(1): 6855.

26. Shang J, Ye G, Shi K, et al. Structural basis of receptor recognition by SARS-CoV-2. *Nature* 2020; 581(7807): 221-4.

27. Liu K, Pan X, Li L, et al. Binding and molecular basis of the bat coronavirus RaTG13 virus to ACE2 in humans and other species. *Cell* 2021; 184(13): 3438-51 e10.

28. Letko M, Miazgowicz K, McMinn R, et al. Adaptive Evolution of MERS-CoV to Species Variation in DPP4. *Cell Rep* 2018; 24(7): 1730-7.

29. Martinez DR, Schafer A, Leist SR, et al. Chimeric spike mRNA vaccines protect against Sarbecovirus challenge in mice. *Science* 2021; 373(6558): 991-8.

30. Tortorici MA, Czudnochowski N, Starr TN, et al. Broad sarbecovirus neutralization by a human monoclonal antibody. *Nature* 2021; 597(7874): 103-8.

31. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* 2010; 59(3): 307-21.

32. Lefort V, Longueville JE, Gascuel O. SMS: Smart Model Selection in PhyML. *Mol Biol Evol* 2017; 34(9): 2422-4.