AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Virology

Isolation and Characterization of a Novel Bat Coronavirus Closely Related to the Direct Progenitor of Severe Acute Respiratory Syndrome Coronavirus

Xing-Lou Yang,[a] Ben Hu,[a] Bo Wang,[a] Mei-Niang Wang,[a] Qian Zhang,[a] Wei Zhang,[a] Li-Jun Wu,[a] Xing-Yi Ge,[a] Yun-Zhi Zhang,[b] Peter Daszak,[c] Lin-Fa Wang,[d] Zheng-Li Shi[a]

Key Laboratory of Special Pathogens and Center for Emerging Infectious Diseases, Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan, China[a]; Yunnan Institute of Endemic Diseases Control and Prevention, Dali, China[b]; EcoHealth Alliance, New York, New York, USA[c]; Program in Emerging Infectious Diseases, Duke-NUS Graduate Medical School, Singapore, Singapore[d]

We report the isolation and characterization of a novel bat coronavirus which is much closer to the severe acute respiratory syndrome coronavirus (SARS-CoV) in genomic sequence than others previously reported, particularly in its S gene. Cell entry and susceptibility studies indicated that this virus can use ACE2 as a receptor and infect animal and human cell lines. Our results provide further evidence of the bat origin of the SARS-CoV and highlight the likelihood of future bat coronavirus emergence in humans.

The 2002-2003 outbreak of severe acute respiratory syndrome coronavirus (SARS-CoV) was a significant public health threat at the beginning of the 21st century (1). Initial evidences showed that the masked palm civet (*Paguma larvata*) was the primary suspect in the animal origin of SARS-CoV (2, 3). Later studies suggested that Chinese horseshoe bats are natural reservoirs and that the masked palm civet most likely served as an intermediate amplification host for SARS-CoV (4, 5). From our longitudinal surveillance of bat SARS-like coronavirus (SL-CoV) in a single bat colony of the species *Rhinolophus sinicus* in Kunming, Yunnan Province, China, we found a high prevalence of diverse SL-CoVs (6). Whole-genome sequence comparison revealed that these SL-CoVs have 78% to 95% nucleotide sequence identities to SARS-CoV, with the major differences located in the spike protein (S) genes and the region of open reading frame 8 (ORF8). We recently isolated a bat SL-CoV strain (WIV1) and constructed an infectious clone of another strain (SHC014); significantly, these strains are closely related to SARS-CoV and capable of using the same cellular receptor (angiotensin-converting enzyme 2 [ACE2]) as SARS-CoV (6, 7). Despite the high similarity in genomic sequences and receptor usage of these two strains, there is still some difference between the N-terminal domains of the S proteins of SARS-CoV and other SL-CoVs, indicating that other unknown SL-CoVs are circulating in bats.

Here we report the isolation of a new SL-CoV strain, named bat SL-CoV WIV16. SL-CoV WIV16 was isolated from a single fecal sample of *Rhinolophus sinicus*, which was collected in Kunming, Yunnan Province, in July 2013. The full genomic sequence of SL-CoV WIV16 (GenBank accession number KT444582) was determined and contained 30,290 nucleotides (nt) and a poly(A) tail which is slightly larger than those of SARS-CoVs and other bat SL-CoVs (6, 8–13). The WIV16 genome has a 40.9% G+C content and short untranslated regions (UTRs) of 264 and 339 nt at the 5′ and 3′ termini, respectively. Its gene organization is identical to that of WIV1 and slightly different from that of the civet SARS-CoV and other bat SL-CoVs due to an additional ORF (name ORFx) detected between the ORF6 and ORF7 genes of the WIV1 and WIV16 genomes (data not shown). The conserved transcriptional regulatory sequence was identified upstream of ORFx, indicating that this is likely to be a potential functional gene. The overall nucleotide sequence of WIV16 has 96% identity (higher than that of any previously reported bat SL-CoVs) to human and civet SARS-CoVs (Table 1) (4–6, 8–13). A detailed comparison of protein sequences between SARS-CoV GZ02, a strain from an early-phase patient, and all reported bat SL-CoVs indicated that WIV16 is the closet progenitor of the SARS-CoV in most proteins, particularly in the S protein (Table 1).

The S protein is responsible for virus entry and is functionally divided into two domains, denoted S1 and S2. The S1 domain is involved in receptor binding, and the S2 domain is involved in cellular membrane fusion (14). S1 is functionally subdivided into two domains, an N-terminal domain (S1-NTD) and a C-terminal domain (S1-CTD), both of which can bind to host receptors and hence function as receptor-binding domain (RBDs) (15). All isolates of SARS-CoV and SL-CoV have high identity in both their nucleotide and their amino acid sequences in the S2 region but are highly diverse in their S1 regions. The WIV16 S gene has 95% sequence identity at the nucleotide level and 97% identity at the amino acid level to SARS-CoVs, much higher than those of WIV1, which has 88% identity at the nucleotide level and 90% identity at the amino acid level. Unlike with other bat SL-CoVs, the S1-NTD of WIV16 is very similar to that of SARS-CoV (Fig. 1). The S1-NTD of WIV16 has an amino acid sequence identity to SARS-CoVs of 94% but of only 50% to 75% to other bat SL-CoVs. It is worth noting that the WIV16 RBD (amino acids [aa] 318 to 510) has 95% sequence identity to the SARS-CoV RBD but is almost identical to

Received 13 October 2015 **Accepted** 21 December 2015

Accepted manuscript posted online 30 December 2015

Citation Yang X-L, Hu B, Wang B, Wang M-N, Zhang Q, Zhang W, Wu L-J, Ge X-Y, Zhang Y-Z, Daszak P, Wang L-F, Shi Z-L. 2016. Isolation and characterization of a novel bat coronavirus closely related to the direct progenitor of severe acute respiratory syndrome coronavirus. J Virol 90:3253–3256. doi:10.1128/JVI.02582-15.

Editor: S. Perlman

Address correspondence to Zheng-Li Shi, zlshi@wh.iov.cn.

TABLE 1 Genomic comparison of SARS-CoV GZ02 with civet SARS-CoV and other bat SL-CoVs

FL or ORF	No. of nt	No. of aa	% nt identity/% aa identity											
			SZ3	WIV16	WIV1	Rs3367	RsSHC014	Rs672	Rp3	Rf1	Rm1	LYRa11	HKU3-1	YNLF_31C	BM48-31
FL			99.8	96.0	95.6	95.7	95.4	93.4	92.6	87.8	88.2	90.9	87.9	93.5	78.8
P1a	13,134	4,377	99.9/99.9	96.6/98.1	96.9/98.0	96.9/98.0	96.8/98.1	96.4/98.1	94.9/96.6	88.0/94.3	88.0/93.6	91.0/95.9	88.2/94.3	96.0/97.3	76.9/81.6
P1b	8088	2,695	99.9/99.9	96.1/99.1	96.3/99.4	96.3/99.4	96.4/99.5	96.0/99.3	96.2/99.1	90.9/98.3	91.4/98.6	93.8/98.9	90.9/98.6	96.8/99.2	85.5/96.0
S	3,768	1,255	99.6/99.0	95.4/97.3	90.2/92.4	90.2/92.5	88.4/90.2	77.6/80.1	78.1/80.2	75.5/78.4	78.0/80.6	83.3/89.9	77.0/79.4	76.1/79.2	70.9/76.0
S1	2,040	680	99.5/98.8	92.6/95.4	83.3/86.5	83.4/86.8	79.9/82.4	68.8/67.0	69.1/66.7	66.7/66.1	69.0/67.4	80.3/84.4	69.2/67.2	67.5/66.7	65.8/64.5
S2	1,728	575	99.8/99.3	98.3/99.5	98.3/99.5	98.2/99.3	98.3/99.5	88.0/95.5	88.4/96.2	85.5/92.7	88.3/96.0	87.3/96.3	85.9/93.9	86.0/93.7	76.7/89.6
ORF3a	825	274	99.0/97.8	99.2/98.2	99.0/97.8	99.2/98.2	99.3/98.2	90.4/90.8	84.0/84.3	88.6/86.9	83.5/84.3	89.7/91.6	83.0/82.5	89.0/88.3	73.1/71.5
E	231	76	100.0/100.0	99.1/100.0	99.1/100.0	99.1/100.0	98.7/98.7	99.6/100.0	97.8/100.0	96.5/96.1	96.1/98.7	98.3/98.7	97.4/100.0	99.6/100.0	90.0/92.1
M	666	221	99.8/99.5	97.4/98.2	97.4/98.2	97.4/98.2	97.4/97.7	97.7/98.6	93.4/97.3	95.5/97.7	94.7/97.3	94.7/97.7	95.0/98.6	95.9/98.6	81.5/91.4
ORF6	192	63	100.0/100.0	95.3/92.1	95.8/93.7	97.9/96.8	97.4/96.8	97.4/98.4	94.8/92.1	94.8/93.7	94.8/92.1	94.3/95.2	94.8/93.7	92.7/88.9	65.1/50.0
ORF7a	369	122	100.0/100.0	94.3/95.1	94.9/95.1	94.9/95.9	94.6/95.9	94.3/95.9	93.8/95.1	92.1/91.8	93.0/93.4	93.2/94.3	93.0/94.3	96.7/96.7	63.9/58.5
ORF7b	135	44	100.0/100.0	96.3/93.2	95.6/93.2	95.6/93.2	96.3/93.2	95.6/93.2	96.3/93.2	94.1/90.9	95.6/93.2	86.7/90.9	92.6/93.2	97.0/93.2	65.0/70.0
ORF8	369	122	99.5/98.4	50.1/38.6	50.7/39.5	50.7/39.5	50.7/40.4	51.6/39.5	53.3/39.5	82.1/81.8	52.1/39.5	51.0/38.3	52.1/37.7	82.1/82.6	NA
N	1,269	422	99.9/100.0	98.4/99.5	98.4/99.8	98.7/100.0	98.3/99.5	97.6/98.6	96.7/98.1	94.2/95.7	96.4/97.9	96.9/97.9	96.2/96.7	97.2/98.3	78.5/88.2

SARS-CoV GZ02 was isolated from patients in the early phase of the SARS outbreak in 2003. SARS-CoV SZ3 was identified from *Paguma larvata* in 2003 collected in Guangdong, China. SL-CoV WIV16, WIV1, Rs3367, and RsSHC014 were identified from *Rhinolophus sinicus* collected in Yunnan, China, during 2011 to 2013. SL-CoV YNLF_31C was identified from *R. ferrumequinum* collected in Yunnan, China, in 2013. SL-CoV LYRa11 was identified from *R. affinis* collected in Yunnan, China, in 2011. SL-CoV Rs672, Rp3, and HKU3-1 were identified from *R. sinicus* collected in China (respectively, from Guangxi in 2004, Guizhou in 2006, and Hong Kong in 2005). Rf1 and Rm1 were identified from *R. ferrumequinum* and *R. macrotis*, respectively, collected in Hubei, China, in 2003. Bat SARS-related CoV BM48-31 was identified from *R. blasii* collected in Bulgaria in 2008. FL, full-length genome; S1, the N-terminal domain of the S protein (aa 1 to 680); S2, the C-terminal domain of the S protein (aa 681 to 1255); NA, not available. A pairwise comparison was conducted for all ORFs at the nucleotide and amino acid levels. The full-length genomes were compared at the nucleotide level.

FIG 1 Similarity plot based on the nucleotide sequence of the S gene of bat SL-CoV WIV16. S genes of human/civet SARS-CoVs and bat SL-CoV WIV1 were used as reference sequences, with a window of 200 bp and a step size of 20 bp under the Kimura model. CenterPos, center position.

WIV1's. Thus, the WIV16 S gene is likely a recombinant of WIV1's gene and a recent ancestor of SARS-CoV's gene.

High sequence conservation of the WIV16 RBD with that of SARS-CoVs predicts that WIV16 is likely also to use ACE2 as a cellular entry receptor. This was confirmed by infection of HeLa cells expressing ACE2 from human, civet, and Chinese horseshoe bat, respectively (Fig. 2A). Cell susceptibility testing using different cell lines further indicated that WIV16 has the same host range as WIV1 (Fig. 2B) (6).

To assess whether the major sequence difference of the S1-NTD will have an effect on virus entry and/or replication, the growth kinetics of the two viruses were comparatively studied. Vero E6 cells were infected with WIV1 or WIV16 at a multiplicity of infection (MOI) of 1, and virus production in the medium supernatant was determined at four time points postinfection by quantification of viral RNA (Fig. 3; see the figure legend for more technical detail). The two viruses grew at very similar rates, with WIV16's rate being slightly lower than WIV1's rate during the 48-h period examined in this study. It is hard to conclude whether this subtle difference is significant and related to the S1-NTD sequence difference. Further investigation with more cell lines is required to confirm this preliminary observation.

In conclusion, we isolated and characterized a novel bat SL-CoV isolate, WIV16, which is the closest ancestor to date of the SARS-CoV. Our results provide further evidence that Chinese horseshoe bats are natural reservoirs of SARS-CoVs. It should be noted that WIV16 is not the closest strain to the human SARS-CoVs with regard to ORF8. Full-length ORF8 is present in several SARS-CoV genomes of early-phase patients, all civet SARS-CoVs, and bat SL-CoVs. It is split into two ORFs (ORF8a and -b) in most human SARS-CoVs from late-phase patients due to a deletion event in this part of the genome (3). Recently, two papers reported that they found a full-length ORF8 which has higher similarities to SARS-CoV GZ02 and civet SARS-CoV SZ3, suggesting that SARS-CoV derived from a complicated recombination and genetic evolution among different bat SL-CoVs (10, 12). Considering everything together, we predict that there are diverse SL-CoVs to be discovered

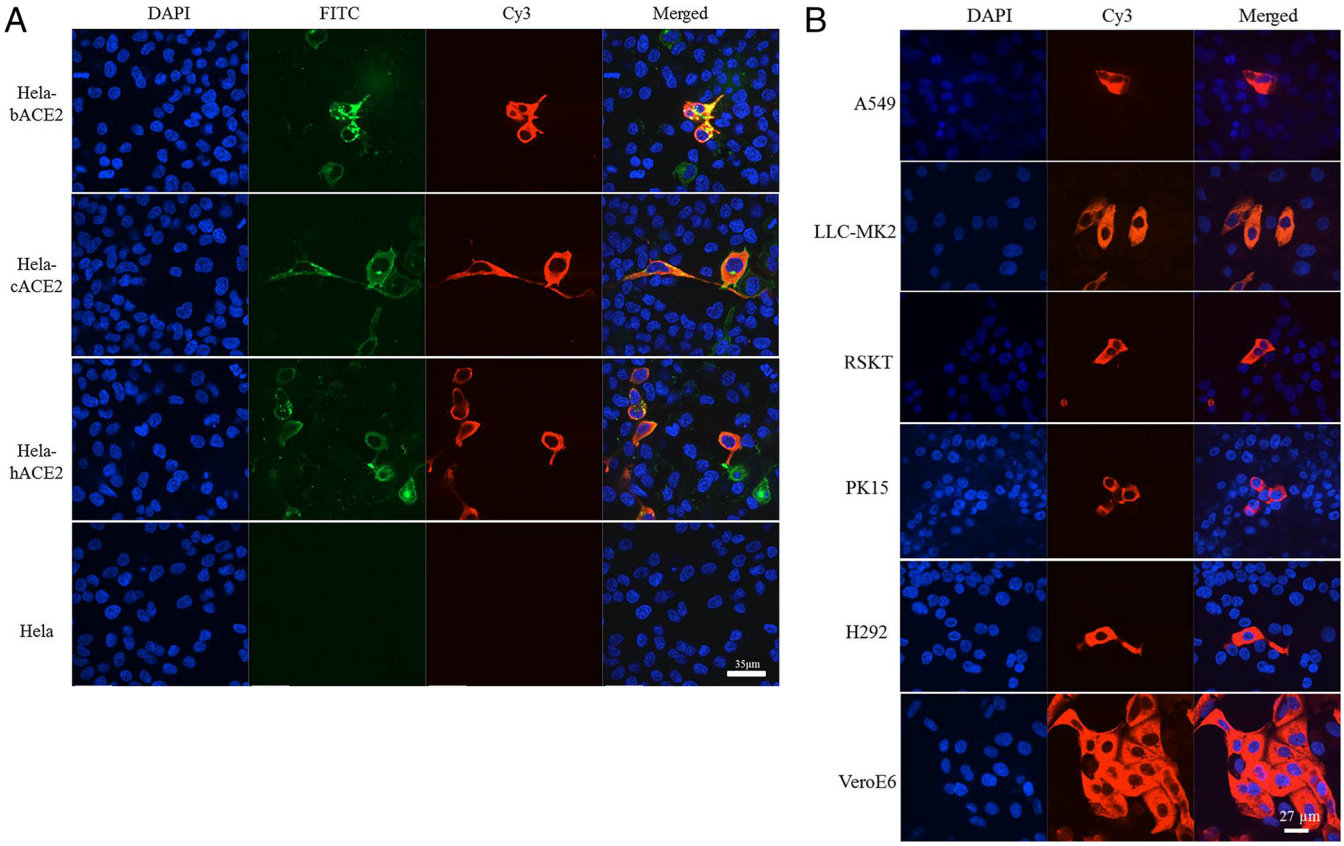

FIG 2 Receptor analysis (A) and susceptibility test (B) results for bat SL-CoV WIV16. (A) HeLa cells with and without the expression of ACE2. ACE2 expression was detected with goat anti-human ACE2 antibody, followed by fluorescein isothiocyanate (FITC)-conjugated donkey anti-goat IgG. Virus replication was detected with rabbit antibody against the SL-CoV Rp3 nucleocapsid protein, followed by cyanine 3 (Cy3)-conjugated mouse anti-rabbit IgG. Nuclei were stained with DAPI (4′,6-diamidino-2-phenylindole). The columns (from left to right) show staining of nuclei (blue), ACE2 expression (green), virus replication (red), and all three (merged triple-stain images). b, bat; c, civet; h, human. (B) Virus infection in A549, LLC-MK2, RSKT, PK15, H292, and Vero-E6 cells. The columns (from left to right) show staining of nuclei (blue), virus replication (red), and both nuclei and virus replication (merged double-stain images). A549 and H292, human lung cells; LLC-MK2, macaque kidney cells; RSKT, Chinese horseshoe bat kidney cells; PK15, pig kidney cells; VeroE6, African green monkey kidney cells.

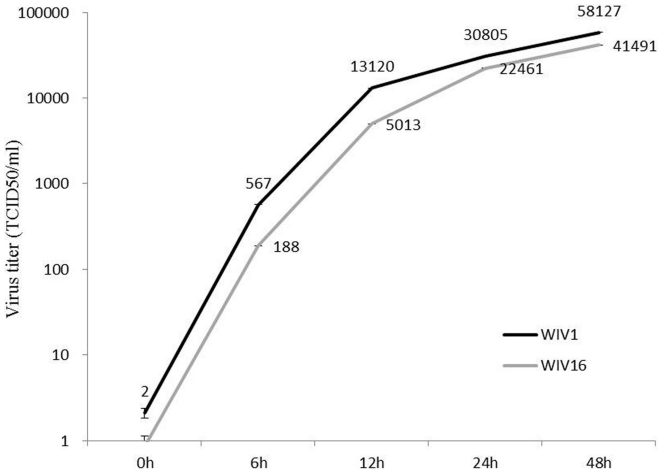

FIG 3 One-step growth curve of bat SL-CoV WIV16 compared with that of WIV1. Vero E6 cells were infected with WIV16 or WIV1 at an MOI of 1. Supernatants were collected at 0, 12, 24, and 48 h postinfection. The viruses in the supernatant were determined by one-step reverse real-time PCR ($n = 3$). Virus RNA that had been extracted from a virus with a known titer was used to set up the standard curve. Error bars represent standard deviations. TCID50, 50% tissue culture infective doses.

in bats. Continued surveillance of this group of viruses in bats will be necessary and important not only for a better understanding of the spillover mechanism but also for more effective risk assessment and prevention of future SARS-like disease outbreaks.

Nucleotide sequence accession number. The full genomic sequence of SL-CoV WIV16 was deposited in GenBank under accession number KT444582.

FUNDING INFORMATION

National Natural Science Foundation of China (NSFC) provided funding to Zheng-Li Shi under grant numbers 81290341 and 31321001. Minister of Science and Technology of the People's Republic of China provided funding to Zheng-Li Shi under grant numbers 2014ZX10004001-003 and 2013FY113500. HHS | National Institutes of Health (NIH) provided funding to Peter Daszak and Zheng-Li Shi under grant number NIAID R01AI110964. National Research Foundation Singapore (NRF) provided funding to Lin-Fa Wang under grant number NRF2012NRF-CRP001-056.

REFERENCES

1. **Peiris JSM, Guan Y, Yuen KY.** 2004. Severe acute respiratory syndrome. Nat Med **10:**S88–S97. http://dx.doi.org/10.1038/nm1143.
2. **Guan Y, Zheng BJ, He YQ, Liu XL, Zhuang ZX, Cheung CL, Luo SW, Li PH, Zhang LJ, Guan YJ, Butt KM, Wong KL, Chan KW, Lim W,**

Shortridge KF, Yuen KY, Peiris JS, Poon LL. 2003. Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China. Science **302:**276–278. http://dx.doi.org/10.1126/science.1087139.

3. **Song HD, Tu CC, Zhang GW, Wang SY, Zheng K, Lei LC, Chen QX, Gao YW, Zhou HQ, Xiang H, Zheng HJ, Chern SW, Cheng F, Pan CM, Xuan H, Chen SJ, Luo HM, Zhou DH, Liu YF, He JF, Qin PZ, Li LH, Ren YQ, Liang WJ, Yu YD, Anderson L, Wang M, Xu RH, Wu XW, Zheng HY, Chen JD, Liang G, Gao Y, Liao M, Fang L, Jiang LY, Li H, Chen F, Di B, He LJ, Lin JY, Tong S, Kong X, Du L, Hao P, Tang H, Bernini A, Yu XJ, Spiga O, Guo ZM, Pan HY, He WZ, Manuguerra JC, Fontanet A, Danchin A, Niccolai N, Li YX, Wu CI, Zhao GP.** 2005. Cross-host evolution of severe acute respiratory syndrome coronavirus in palm civet and human. Proc Natl Acad Sci U S A **102:**2430–2435. http://dx.doi.org/10.1073/pnas.0409608102.
4. **Lau SKP, Woo PCY, Li KSM, Huang Y, Tsoi HW, Wong BHL, Wong SSY, Leung SY, Chan KH, Yuen KY.** 2005. Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats. Proc Natl Acad Sci U S A **102:**14040–14045. http://dx.doi.org/10.1073/pnas.0506735102.
5. **Li WD, Shi ZL, Yu M, Ren WZ, Smith C, Epstein JH, Wang HZ, Crameri G, Hu ZH, Zhang HJ, Zhang JH, McEachern J, Field H, Daszak P, Eaton BT, Zhang SY, Wang LF.** 2005. Bats are natural reservoirs of SARS-like coronaviruses. Science **310:**676–679. http://dx.doi.org/10.1126/science.1118391.
6. **Ge XY, Li JL, Yang XL, Chmura AA, Zhu G, Epstein JH, Mazet JK, Hu B, Zhang W, Peng C, Zhang YJ, Luo CM, Tan B, Wang N, Zhu Y, Crameri G, Zhang SY, Wang LF, Daszak P, Shi ZL.** 2013. Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. Nature **503:**535–538. http://dx.doi.org/10.1038/nature12711.
7. **Menachery VD, Yount BL, Jr, Debbink K, Agnihothram S, Gralinski LE, Plante JA, Graham RL, Scobey T, Ge XY, Donaldson EF, Randell SH, Lanzavecchia A, Marasco WA, Shi ZL, Baric RS.** 2015. A SARS-like cluster of circulating bat coronaviruses shows potential for human emergence. Nat Med **21:**1508–1513. http://dx.doi.org/10.1038/nm.3985.
8. **Ren W, Li W, Yu M, Hao P, Zhang Y, Zhou P, Zhang S, Zhao G, Zhong Y, Wang S, Wang LF, Shi Z.** 2006. Full-length genome sequences of two SARS-like coronaviruses in horseshoe bats and genetic variation analysis. J Gen Virol **87:**3355–3359. http://dx.doi.org/10.1099/vir.0.82220-0.
9. **Yuan J, Hon CC, Li Y, Wang D, Xu G, Zhang H, Zhou P, Poon LL, Lam TT, Leung FC, Shi Z.** 2010. Intraspecies diversity of SARS-like coronaviruses in Rhinolophus sinicus and its implications for the origin of SARS coronaviruses in humans. J Gen Virol **91:**1058–1062. http://dx.doi.org/10.1099/vir.0.016378-0.
10. **Wu Z, Yang L, Ren X, Zhang J, Yang F, Zhang S, Jin Q.** 3 October 2015. ORF8-related genetic evidence for Chinese horseshoe bats as the source of human severe acute respiratory syndrome coronavirus. J Infect Dis http://dx.doi.org/10.1093/infdis/jiv476.
11. **He B, Zhang Y, Xu L, Yang W, Yang F, Feng Y, Xia L, Zhou J, Zhen W, Feng Y, Guo H, Zhang H, Tu C.** 2014. Identification of diverse alphacoronaviruses and genomic characterization of a novel severe acute respiratory syndrome-like coronavirus from bats in China. J Virol **88:**7070–7082. http://dx.doi.org/10.1128/JVI.00631-14.
12. **Lau SKP, Feng Y, Chen HL, Luk HKH, Yang WH, Li KSM, Zhang YZ, Huang Y, Song ZZ, Chow WN, Fan RYY, Ahmed SS, Yeung HC, Lam CSF, Cai JP, Wong SSY, Chan JFW, Yuen KY, Zhang HL, Woo PCY.** 2015. Severe acute respiratory syndrome (SARS) coronavirus ORF8 protein is acquired from SARS-related coronavirus from greater horseshoe bats through recombination. J Virol **89:**10532–10547. http://dx.doi.org/10.1128/JVI.01048-15.
13. **Drexler JF, Gloza-Rausch F, Glende J, Corman VM, Muth D, Goettsche M, Seebens A, Niedrig M, Pfefferle S, Yordanov S, Zhelyazkov L, Hermanns U, Vallo P, Lukashev A, Muller MA, Deng H, Herrler G, Drosten C.** 2010. 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 **84:**11336–11349. http://dx.doi.org/10.1128/JVI.00650-10.
14. **Paul Masters S, Perlman S.** 2013. Coronaviridae, p 825–858. *In* Knipe DM, Howley PM, Cohen JI, Griffin DE, Lamb RA, Martin MA, Racaniello VR, Roizman B (ed), Fields virology, vol I. Lippincott Williams & Wilkins, Philadelphia, PA.
15. **Li F.** 2012. Evidence for a common evolutionary origin of coronavirus spike protein receptor-binding subunits. J Virol **86:**2856–2858. http://dx.doi.org/10.1128/JVI.06882-11.