JOURNAL OF VIROLOGY, May 2008, p. 5084–5088 Vol. 82, No. 10
0022-538X/08/$08.00+0 doi:10.1128/JVI.02722-07

Identification of a Novel Coronavirus from a Beluga Whale by Using a Panviral Microarray[▽]†

Kathie A. Mihindukulasuriya,[1] Guang Wu,[1] Judy St. Leger[2] Robert W. Nordhausen,[3] and David Wang[1]*

Departments of Molecular Microbiology and Pathology & Immunology, Washington University School of Medicine, St. Louis, Missouri[1]; SeaWorld, San Diego, 500 SeaWorld Dr., San Diego, California 92109[2]; and California Animal Health and Food Safety Laboratory, University of California at Davis, West Health Science Drive, Davis, California 95616[3]

Received 21 December 2007/Accepted 5 March 2008

The emergence of viruses such as severe acute respiratory syndrome coronavirus and Nipah virus has underscored the role of animal reservoirs in human disease and the need for reservoir surveillance. Here, we used a panviral DNA microarray to investigate the death of a captive beluga whale in an aquatic park. A highly divergent coronavirus, tentatively named coronavirus SW1, was identified in liver tissue from the deceased whale. Subsequently, the entire genome of SW1 was sequenced, yielding a genome of 31,686 nucleotides. Phylogenetic analysis revealed SW1 to be a novel virus distantly related to but most similar to group III coronaviruses.

An estimated 75% of emerging diseases arise from zoonotic sources (30). Zoological parks and aquariums provide a unique opportunity for emerging virus surveillance. For example, in 1999, the first harbinger of West Nile virus emergence in North America was the mysterious death of birds at the Bronx Zoo/Wildlife Conservation Park (25). Thus, zoo populations may serve as sentinels for emerging viruses.

Panviral DNA microarrays represent one approach for massively parallel viral surveillance. We have previously described a panviral DNA microarray (ViroChip) capable of detecting thousands of known viruses as well as novel viruses related to known viral families in a single assay (35). ViroChip has previously been used to identify severe acute respiratory syndrome (SARS) coronavirus (19, 35); xenotropic murine leukemia virus-related virus, a novel human retrovirus, in patients with familial prostate cancer (32); and a novel clade of human rhinoviruses (16).

In this paper, a ViroChip was used to interrogate primary liver tissue from a recently deceased beluga whale for the presence of viruses. Microarray hybridization strongly suggested that a coronavirus was present in the liver tissue. Subsequent complete genome sequencing and phylogenic analysis revealed that the virus was a novel, highly divergent coronavirus most similar overall to group 3 coronaviruses. We have tentatively named this virus coronavirus SW1.

Clinical history and necropsy results. A 13-year-old, male, captive-born beluga whale died after a short medical illness characterized by generalized pulmonary disease and terminal acute liver failure. The liver demonstrated a diffuse increased friability with multifocal, red-yellow mottling and irregularly shaped areas of obvious necrosis (Fig. 1A). Histological examination of liver tissue demonstrated a severe, multifocal, and coalescing centrilobular-to-massive acute hepatic necrosis (data not shown). To study the liver in more detail, conventional transmission electron microscopy was performed as previously described (12). Abundant nondescript round viral particles measuring ~60 to 80 nm with cores of approximately 45 to 50 nm were identified in the cytoplasm, but this was insufficient to identify the virus (Fig. 1B). We note that while the observed particles were smaller than those typically associated with coronaviruses, coronavirus particles as small as 50 nm have been reported (26).

Virus isolation attempts. Liver tissue homogenate was inoculated into bovine turbinate, Vero, MARC 145, primary fetal porcine kidney, rabbit kidney (RK-13Ky), BHK, bovine embryonic testicle, MDCK, bovine pulmonary arterial endothelium, and human rectal tumor 18 cells and embryonating chicken eggs. No evidence of viral growth was observed.

Panviral DNA microarray analysis. RNA was extracted from liver tissue samples of the infected and two control, uninfected whales. Two hundred nanograms of RNA was randomly amplified and hybridized to the panviral microarray as previously described (35). Multiple oligonucleotides derived from various coronaviruses gave strong hybridization intensity in the infected liver, suggesting the presence of a coronavirus in the infected liver.

Consensus coronavirus PCR and complete genome sequencing. To confirm the microarray findings, reverse transcription-PCR (RT-PCR) was performed with published consensus coronavirus primers (9). A PCR product of 454 bp that possessed 70% amino acid identity with the 1ab replicase polyprotein of avian infectious bronchitis virus as determined by tBLASTx was obtained (2). The entire viral genome was subsequently sequenced using shotgun sequencing, RT-PCR, and 5′ and 3′ rapid amplification of cDNA ends. The initial assembly was confirmed by sequencing a series of overlapping RT-

* Corresponding author. Mailing address: Washington University School of Medicine, Campus Box 8230, 660 S. Euclid Ave., St. Louis, MO 63110. Phone: (314) 286-1123. Fax: (314) 362-1232. E-mail: davewang@borcim.wustl.edu.

† Supplemental material for this article may be found at http://jvi.asm.org/.

▽ Published ahead of print on 19 March 2008.

A

B

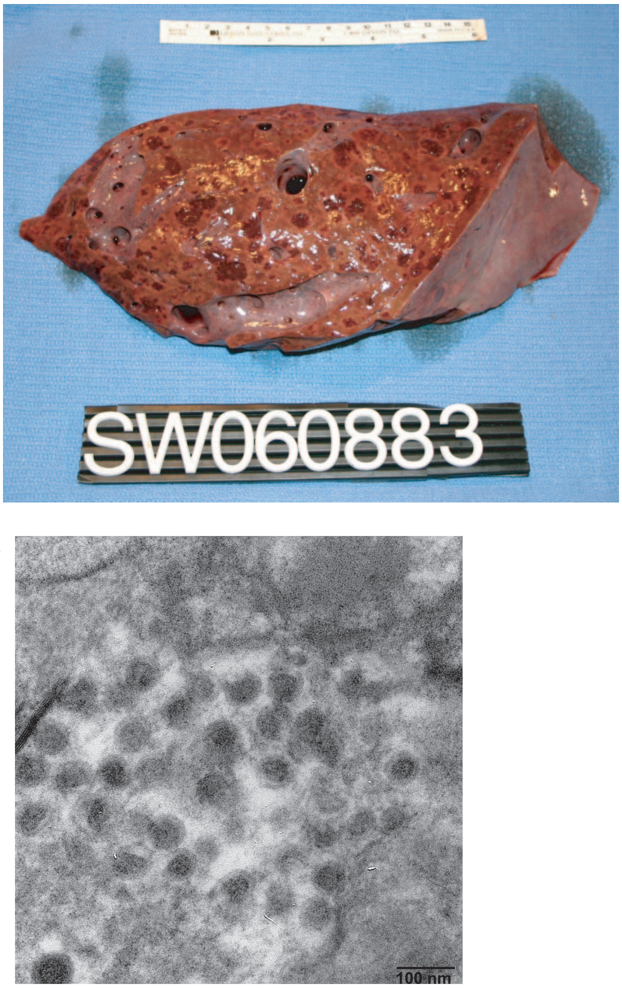

FIG. 1. (A) Explanted liver of the dead Beluga whale. (B) Electron microscopy of whale liver at ×129,300 magnification. Bar, 100 nm.

PCR products, yielding the finished genome of 31,686 nucleotides (nt).

Analysis of viral ORFs. Putative open reading frames (ORFs) were predicted using NCBI's ORF Finder (37) and the results refined using information about the noncoding sequence of SW1 to determine the most likely start sites. SW1 contained 14 putative ORFs (Table 1), including ORFs with similarity to the five major ORFs conserved in all known coronaviruses (Table 1 and Fig. 2). SW1 encoded eight putative accessory proteins whose genes were located between the M and N genes. None of these proteins had any detectable sequence similarity to proteins in other known coronaviruses, and their functional roles are currently unknown. A number of the ORFs had noteworthy features. The ORF 6 protein possessed amino acid similarity (BLAST 1e−06) to human astrovirus capsid proteins. Astrovirus capsid proteins have recently been demonstrated to disrupt tight junctions and thereby increase the barrier permeability of polarized cell monolayers, resulting in increased viral dissemination (27). The ORF 10 protein had significant amino acid similarity to a number of uridine kinases (BLAST 2e−26); no virus described to date encodes a uridine kinase (11). In addition, since some viruses encode secreted proteins that interfere with the host immune response (1, 33), the accessory proteins were analyzed for the presence of signal sequences by using SignalP (4). The ORF 7 and 8 proteins contained putative signal sequences, suggesting that they may be secreted. Finally, analysis with PolyPhobius (15) suggested that among the accessory genes, ORFs 5b and 9 contained transmembrane domains.

Analysis of noncoding viral RNA. Known coronavirus 5′ untranslated regions (UTRs) range from 209 to 528 nt (6), including a leader sequence of 65 to 98 nt. In SW1, the 5′ UTR was 523 nt, with a leader sequence of 79 nt. Typically, the final 7 to 18 nt of the leader form the transcription-regulating sequence (TRS) motif, which defines the 5′ end of each subgenomic RNA. Using MEME (3), a putative TRS motif was identified. This was experimentally confirmed by amplifying the 5′ ends of the mRNAs for the N and S genes, using a primer in the putative leader sequence and a second primer within each gene. Comparison of this amplified sequence to the genomic sequence confirmed that the TRS motif was 5′A AACA. Ten of the ORFs were immediately preceded by a TRS consensus sequence (Table 1 and Fig. 3). While ORFs 3, 5b, and 5c were not preceded by a TRS sequence, internal translation from subgenomic RNAs has been described for coronaviruses (18, 20). The SW1 3′ UTR of 369 nt fell within the known size range for other coronaviruses (288 to 506 nt) (6, 29).

Phylogenetic analysis reveals that SW1 is a novel, highly divergent coronavirus. Coronaviruses are classified based on genomic organization and phylogenetic analysis of full-length genomes (13). Phylogenetic analysis of the five major ORFs by use of ClustalX V1.83 (31) (Fig. 2A to E) demonstrated that overall SW1 was most closely related to group III coronaviruses. In addition, its genomic organization was also most similar to that of known group III coronaviruses (Fig. 3).

The emergence of SARS in 2003 marked a renaissance in the field of coronavirology. Since then, new members of the family *Coronaviridae* have been identified in birds (14, 24), humans (34, 38), bats (21, 22, 28, 39), and wild mammals from Chinese live-animal markets (8). In this study, we identified a novel coronavirus in the liver tissue of a deceased beluga

TABLE 1. Predicted ORFs

ORF	Predicted size (aa) of protein	Presence of TRS sequence	Distance (nt) from TRS to ATG	% Identity with avian infectious bronchitis virus[a]
1a	3,955	Yes	447	25
1ab	6,664			42
2 (S)	1,473	Yes	30	21
3 (E)	96			27
4 (M)	261	Yes	62	32
5a	139	Yes	0	NA
5b	173			NA
5c	176			NA
6	229	Yes	1	NA
7	162	Yes	0	NA
8	60	Yes	7	NA
9	153	Yes	0	NA
10	211	Yes	0	NA
11 (N)	380	Yes	105	35

[a] NA, not applicable.

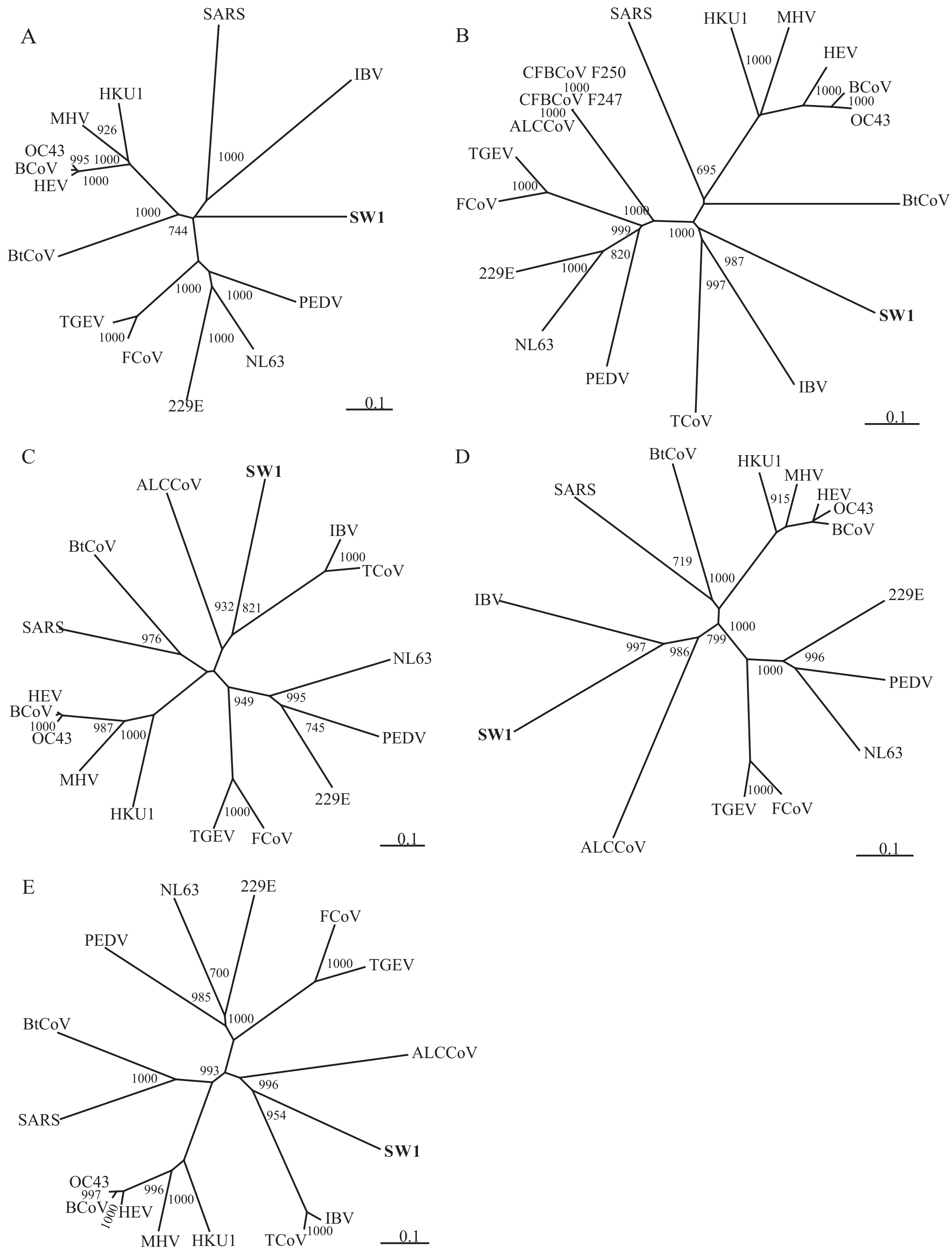

FIG. 2. Phylogenetic analysis of SW1. Phylogenetic trees were constructed from protein sequences by using the neighbor-joining method with 1,000 bootstrap replicates. Abbreviations: PEDV, porcine epidemic diarrhea virus; NL63, human coronavirus NL63; FCoV, feline coronavirus; TGEV, transmissible gastroenteritis virus; 229E, human coronavirus 229E; BCoV, bovine coronavirus; MHV, murine hepatitis virus strain JHM; SARS, SARS coronavirus; BtCoV, bat coronavirus (BtCoV/133/2005); HEV, porcine hemagglutinating encephalomyelitis virus; HKU1, human coronavirus HKU1; OC43, human coronavirus OC43; IBV, infectious bronchitis virus; TCoV, turkey coronavirus; CFBCoV F250, Chinese ferret badger coronavirus Guangxi/F250/2006; CFBCoV F247, Chinese ferret badger coronavirus Guangxi/F247/2006; ALCCoV, Asian leopard cat coronavirus Guangxi/F230/2006. The accession numbers of the sequences used are found in Table S1 in the supplemental material. (A) ORF 1ab; (B) spike; (C) envelope; (D) membrane; (E) nucleocapsid.

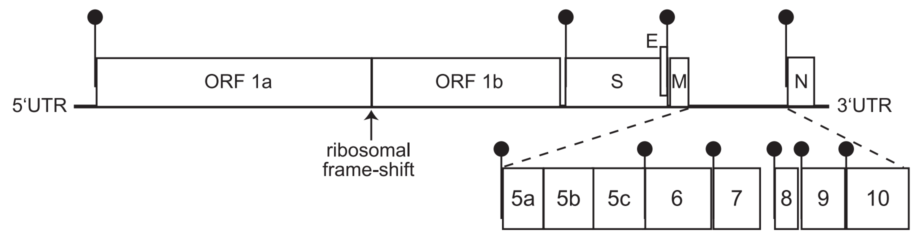

FIG. 3. Genome organization of SW1. Diagrammatic representation of the 5′-to-3′ arrangement of ORFs of the genome of SW1. The TRSs are indicated by black circles.

whale. While there is a report of immunohistochemical staining of the small intestine tissue from harbor seals (*Phoca vitulina*) with acute necrotizing enteritis with antisera to group I coronaviruses (5), this is the first description of the complete genome sequence of a coronavirus found in a marine mammal.

The detection of a novel coronavirus in a deceased beluga whale raises a number of questions, including whether beluga whales are the natural host for this virus and whether the virus was pathogenic to the whale. There is precedence for animal coronaviruses causing hepatic pathology (23, 36). In addition, SARS and HKU1 may be associated with liver disease and hepatitis (7, 10). Thus, the liver damage seen during the beluga whale necropsy (Fig. 1A) may have been caused by SW1 infection, although this remains to be experimentally verified. Furthermore, it is not yet clear whether beluga whales are the natural host, an amplifying host, or a dead-end host for SW1.

In conclusion, we have used a ViroChip to identify a novel coronavirus directly from primary animal tissues. Furthermore, the identification of a previously unrecognized virus in a captive animal underscores the vast diversity of viruses that remains unexplored in animals. These viruses have the potential to be transmitted to humans or other animals, with significant implications for human and animal health. Continued systematic surveillance of animal populations in zoos and aquaria is key for public health preparedness for future outbreaks.

Accession numbers. Primary microarray data have been deposited in NCBI GEO under accession number GSE9238. The nucleotide sequence for the SW1 genome was deposited in GenBank under accession number EU111742.

We thank Nicky Branich (Wyoming Veterinary Diagnostic Laboratory in Laramie, WY) and the National Veterinary Services Laboratory in Ames, IA, for their efforts to culture SW1, Allison Case and Holly Reed for assistance with the necropsy examination and tissue collection, and Fred Murphy for advice with electron micrographs.

REFERENCES

1. **Alcami, A.** 2003. Structural basis of the herpesvirus M3-chemokine interaction. Trends Microbiol. **11:**191–192.
2. **Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389–3402.
3. **Bailey, T. L., and M. Gribskov.** 1998. Combining evidence using p-values: application to sequence homology searches. Bioinformatics **14:**48–54.
4. **Bendtsen, J. D., H. Nielsen, G. von Heijne, and S. Brunak.** 2004. Improved prediction of signal peptides: SignalP 3.0. J. Mol. Biol. **340:**783–795.
5. **Bossart, G. D., and J. C. Schwartz.** 1990. Acute necrotizing enteritis associated with suspected coronavirus infection in three harbour seals (*Phoca vitulina*). J. Zoo Wildl. Med. **21:**84–87.
6. **Brian, D. A., and R. S. Baric.** 2005. Coronavirus genome structure and replication. Curr. Top. Microbiol. Immunol. **287:**1–30.
7. **Chiu, C. Y., A. A. Alizadeh, S. Rouskin, J. D. Merker, E. Yeh, S. Yagi, D. Schnurr, B. K. Patterson, D. Ganem, and J. L. DeRisi.** 2007. Diagnosis of a critical respiratory illness caused by human metapneumovirus by use of a pan-virus microarray. J. Clin. Microbiol. **45:**2340–2343.
8. **Dong, B. Q., W. Liu, X. H. Fan, D. Vijaykrishna, X. C. Tang, F. Gao, L. F. Li, G. J. Li, J. X. Zhang, L. Q. Yang, L. L. Poon, S. Y. Zhang, J. S. Peiris, G. J. Smith, H. Chen, and Y. Guan.** 2007. Detection of a novel and highly divergent coronavirus from Asian leopard cats and Chinese ferret badgers in southern China. J. Virol. **81:**6920–6926.
9. **Drosten, C., S. Gunther, W. Preiser, S. van der Werf, H.-R. Brodt, S. Becker, H. Rabenau, M. Panning, L. Kolesnikova, R. A. M. Fouchier, A. Berger, A.-M. Burguiere, J. Cinatl, M. Eickmann, N. Escriou, K. Grywna, S. Kramme, J.-C. Manuguerra, S. Muller, V. Rickerts, M. Sturmer, S. Vieth, H.-D. Klenk, A. D. M. E. Osterhaus, H. Schmitz, and H. W. Doerr.** 2003. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. N. Engl. J. Med. **348:**1967–1976.
10. **Esper, F., C. Weibel, D. Ferguson, M. L. Landry, and J. S. Kahn.** 2006. Coronavirus HKU1 infection in the United States. Emerg. Infect. Dis. **12:**775–779.
11. **Finn, R. D., J. Mistry, B. Schuster-Bockler, S. Griffiths-Jones, V. Hollich, T. Lassmann, S. Moxon, M. Marshall, A. Khanna, R. Durbin, S. R. Eddy, E. L. Sonnhammer, and A. Bateman.** 2006. Pfam: clans, web tools and services. Nucleic Acids Res. **34:**D247–D251.
12. **Garner, M. M., S. D. Atkinson, S. L. Hallett, J. L. Bartholomew, R. W. Nordhausen, H. Reed, L. Adams, and B. Whitaker.** 2008. Renal myxozoanosis in weedy sea dragons, Phyllopteryx taeniolatus (Lacepede), caused by Sinuolinea phyllopteryxa n. sp. J. Fish Dis. **31:**27–35.
13. **Gorbalenya, A. E., E. J. Snijder, and W. J. Spaan.** 2004. Severe acute respiratory syndrome coronavirus phylogeny: toward consensus. J. Virol. **78:**7863–7866.
14. **Jonassen, C. M., T. Kofstad, I. L. Larsen, A. Lovland, K. Handeland, A. Follestad, and A. Lillehaug.** 2005. Molecular identification and characterization of novel coronaviruses infecting graylag geese (*Anser anser*), feral pigeons (*Columbia livia*) and mallards (*Anas platyrhynchos*). J. Gen. Virol. **86:**1597–1607.
15. **Kall, L., A. Krogh, and E. L. Sonnhammer.** 2005. An HMM posterior decoder for sequence feature prediction that includes homology information. Bioinformatics **21**(Suppl. 1)**:**i251–i257.
16. **Kistler, A., P. C. Avila, S. Rouskin, D. Wang, T. Ward, S. Yagi, D. Schnurr, D. Ganem, J. L. DeRisi, and H. A. Boushey.** 2007. Pan-viral screening of respiratory tract infections in adults with and without asthma reveals unexpected human coronavirus and human rhinovirus diversity. J. Infect. Dis. **196:**817–825.
17. Reference deleted.
18. **Krishnan, R., R. Y. Chang, and D. A. Brian.** 1996. Tandem placement of a coronavirus promoter results in enhanced mRNA synthesis from the downstream-most initiation site. Virology **218:**400–405.
19. **Ksiazek, T. G., D. Erdman, C. S. Goldsmith, S. R. Zaki, T. Peret, S. Emery, S. Tong, C. Urbani, J. A. Comer, W. Lim, P. E. Rollin, S. F. Dowell, A. E. Ling, C. D. Humphrey, W. J. Shieh, J. Guarner, C. D. Paddock, P. Rota, B. Fields, J. DeRisi, J. Y. Yang, N. Cox, J. M. Hughes, J. W. LeDuc, W. J. Bellini, and L. J. Anderson.** 2003. A novel coronavirus associated with severe acute respiratory syndrome. N. Engl. J. Med. **348:**1953–1966.
20. **Lapps, W., B. G. Hogue, and D. A. Brian.** 1987. Sequence analysis of the bovine coronavirus nucleocapsid and matrix protein genes. Virology **157:**47–57.
21. **Lau, S. K., P. C. Woo, K. S. Li, Y. Huang, H. W. Tsoi, B. H. Wong, S. S. Wong, S. Y. Leung, K. H. Chan, and K. Y. Yuen.** 2005. Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats. Proc. Natl. Acad. Sci. USA **102:**14040–14045.
22. **Li, W., Z. Shi, M. Yu, W. Ren, C. Smith, J. H. Epstein, H. Wang, G. Crameri, Z. Hu, H. Zhang, J. Zhang, J. McEachern, H. Field, P. Daszak, B. T. Eaton, S. Zhang, and L. F. Wang.** 2005. Bats are natural reservoirs of SARS-like coronaviruses. Science **310:**676–679.
23. **Liu, M., C. W. Chan, I. McGilvray, Q. Ning, and G. A. Levy.** 2001. Fulminant viral hepatitis: molecular and cellular basis, and clinical implications. Expert Rev. Mol. Med. **3(10):**1–19.
24. **Liu, S., J. Chen, J. Chen, X. Kong, Y. Shao, Z. Han, L. Feng, X. Cai, S. Gu,**

and M. Liu. 2005. Isolation of avian infectious bronchitis coronavirus from domestic peafowl (*Pavo cristatus*) and teal (*Anas*). J. Gen. Virol. **86:**719–725.
25. **Ludwig, G. V., P. P. Calle, J. A. Mangiafico, B. L. Raphael, D. K. Danner, J. A. Hile, T. L. Clippinger, J. F. Smith, R. A. Cook, and T. McNamara.** 2002. An outbreak of West Nile virus in a New York City captive wildlife population. Am. J. Trop. Med. Hyg. **67:**67–75.
26. **Masters, P. S.** 2006. The molecular biology of coronaviruses. Adv. Virus Res. **66:**193–292.
27. **Moser, L. A., M. Carter, and S. Schultz-Cherry.** 2007. Astrovirus increases epithelial barrier permeability independently of viral replication. J. Virol. **81:**11937–11945.
28. **Poon, L. L., D. K. Chu, K. H. Chan, O. K. Wong, T. M. Ellis, Y. H. Leung, S. K. Lau, P. C. Woo, K. Y. Suen, K. Y. Yuen, Y. Guan, and J. S. Peiris.** 2005. Identification of a novel coronavirus in bats. J. Virol. **79:**2001–2009.
29. **Sawicki, S. G., D. L. Sawicki, and S. G. Siddell.** 2007. A contemporary view of coronavirus transcription. J. Virol. **81:**20–29.
30. **Taylor, L. H., S. M. Latham, and M. E. Woolhouse.** 2001. Risk factors for human disease emergence. Philos. Trans. R. Soc. Lond. B **356:**983–989.
31. **Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins.** 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **25:**4876–4882.
32. **Urisman, A., R. J. Molinaro, N. Fischer, S. J. Plummer, G. Casey, E. A. Klein, K. Malathi, C. Magi-Galluzzi, R. R. Tubbs, D. Ganem, R. H. Silverman, and J. L. DeRisi.** 2006. Identification of a novel Gammaretrovirus in prostate tumors of patients homozygous for R462Q RNASEL variant. PLoS Pathog. **2:**e25.
33. **van Berkel, V., J. Barrett, H. L. Tiffany, D. H. Fremont, P. M. Murphy, G. McFadden, S. H. Speck, and H. W. I. V. Virgin.** 2000. Identification of a gammaherpesvirus selective chemokine binding protein that inhibits chemokine action. J. Virol. **74:**6741–6747.
34. **van der Hoek, L., K. Pyrc, M. F. Jebbink, W. Vermeulen-Oost, R. J. Berkhout, K. C. Wolthers, P. M. Wertheim-van Dillen, J. Kaandorp, J. Spaargaren, and B. Berkhout.** 2004. Identification of a new human coronavirus. Nat. Med. **10:**368–373.
35. **Wang, D., A. Urisman, Y. T. Liu, M. Springer, T. G. Ksiazek, D. D. Erdman, E. R. Mardis, M. Hickenbotham, V. Magrini, J. Eldred, J. P. Latreille, R. K. Wilson, D. Ganem, and J. L. DeRisi.** 2003. Viral discovery and sequence recovery using DNA microarrays. PLoS Biol. **1:**E2.
36. **Weiss, S. R., and S. Navas-Martin.** 2005. Coronavirus pathogenesis and the emerging pathogen severe acute respiratory syndrome coronavirus. Microbiol. Mol. Biol. Rev. **69:**635–664.
37. **Wheeler, D. L., D. M. Church, S. Federhen, A. E. Lash, T. L. Madden, J. U. Pontius, G. D. Schuler, L. M. Schriml, E. Sequeira, T. A. Tatusova, and L. Wagner.** 2003. Database resources of the National Center for Biotechnology. Nucleic Acids Res. **31:**28–33.
38. **Woo, P. C., S. K. Lau, C. M. Chu, K. H. Chan, H. W. Tsoi, Y. Huang, B. H. Wong, R. W. Poon, J. J. Cai, W. K. Luk, L. L. Poon, S. S. Wong, Y. Guan, J. S. Peiris, and K. Y. Yuen.** 2005. Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. J. Virol. **79:**884–895.
39. **Woo, P. C., S. K. Lau, K. S. Li, R. W. Poon, B. H. Wong, H. W. Tsoi, B. C. Yip, Y. Huang, K. H. Chan, and K. Y. Yuen.** 2006. Molecular diversity of coronaviruses in bats. Virology **351:**180–187.