JOURNAL OF VIROLOGY, Mar. 2005, p. 3223–3225 Vol. 79, No. 5
0022-538X/05/$08.00+0 doi:10.1128/JVI.79.5.3223–3225.2005

Letter to the Editor

Genetic Variability of Human Respiratory Coronavirus OC43

A recent paper by J. R. St-Jean and colleagues reported the complete genome sequences of the human coronavirus OC43 (HCoV-OC43) laboratory strain from the American Type Culture Collection (ATCC) and an HCoV-OC43 clinical isolate, designated Paris (7). The ATCC HCoV-OC43 (VR759) strain originated in 1967 and was passaged several times in suckling mouse brain and in cell culture. The contemporary HCoV-OC43 Paris strain was isolated in 2001 and was subsequently cultured in an HRT-18 cell line by the authors. A high degree of genetic stability was stated for HCoV-OC43 since only six nucleotide variations in the whole genome could be observed between both HCoV-OC43 strains, with isolation dates 34 years apart. There are, however, some arguments to suggest that the HCoV-OC43 genome is not as stable as suggested by the authors and that the HCoV-OC43 Paris strain might not be a contemporary strain but might be a result of cross-contamination with the ATCC HCoV-OC43 strain.

A first argument is based on the reported evolutionary rates of RNA viruses in general and coronaviruses in particular. An evolutionary rate of 4.0×10^{-4} nucleotide substitutions per site per year was estimated for severe acute respiratory syndrome (SARS) coronavirus by using ORF1ab sequence data (5), and an evolutionary rate of 7.5×10^{-4} nucleotide substitutions per site per year was estimated for porcine transmissible gastroenteritis virus by using S gene sequence data (6). Most RNA viruses have been reported to have evolutionary rates in the range of 10^{-4} to 10^{-3} substitutions per site per year, although slightly slower mutation rates have also been described (1, 2). Here, we estimated the evolutionary rate for the complete genome sequence data of HCoV-OC43 (one strain; GenBank accession number NC_005147 [isolated in 1967]) and BCoV (four strains; GenBank accession numbers U00735 and AF220295 [isolated in 1972] and NC_003045 and AF391542 [isolated in 1998]) using a maximum likelihood (ML) approach. A ML phylogenetic tree for the complete genomes was reconstructed with PAUP version 4.10 (8), and the rate of nucleotide substitution was estimated by using Rhino software version 1.2 (http://evolve.zoo.ox.ac.uk/), which implements a molecular clock model accommodating serially sampled sequences (4). This approach resulted in an estimate of 1.54×10^{-4} nucleotide substitutions per site per year (95% confidence interval, 0.97×10^{-4} to 2.12×10^{-4}) (4), which is 30-fold higher then the mutation rate (5.7×10^{-6} nucleotide substitutions per site per year) corresponding to the data presented by St-Jean et al. The maximum likelihood evolutionary rate was used to assess the probability that the contemporary HCoV-OC43 strain described by St-Jean and colleagues was effectively sampled in 2001. Assuming that nucleotide substitutions follow a Poisson process (3), the probability of observing only six mutations ($n = 6$) between two isolates sampled 34 years apart ($t = 34$) can be calculated by using the following model:

$$P(n, t|\lambda) = \frac{e^{-\lambda t}}{n!}(\lambda t)^n$$

where λ denotes the evolutionary rate in nucleotide substitutions per year.

For the maximum likelihood estimate of the evolutionary rate, the expected number of mutations after 34 years of evolution is 163. The probability of observing only six mutations after 34 years of evolution is 4.98×10^{-61} (the probability of observing six or fewer mutations is 5.17×10^{-61}). Even for the lower 95% confidence interval limit of the evolutionary rate, the probability of observing only six mutations or fewer during this time is extremely low ($P = 4.92 \times 10^{-36}$) and therefore we believe that it seems unlikely that the Paris HCoV-OC43 strain is truly a circulating strain from 2001. It should be noted that this approach is conservative since it ignores any shared ancestry of the two viral strains.

Furthermore, in a paper by Vabret and colleagues from the same laboratories as St-Jean et al. describing an HCoV-OC43 outbreak in Normandy, France, in 2001, M gene sequence data analysis revealed an estimated genetic distance of up to approximately 32 nucleotide changes per 1,000 nucleotides between some of the clinical isolates (9). For the whole M gene, this observation would apply to up to approximately 22 nucleotide variations between some of the 2001 isolates, whereas St-Jean and coworkers found only one nucleotide change in the M gene between two HCoV-OC43 strains isolated 34 years apart. Also, St-Jean and colleagues found only one variation in the S gene, the most variable coronavirus gene. Comparison of this observation to the data of the Normandy outbreak would imply that the S gene of two HCoV-OC43 strains isolated 34 years apart is more conserved than the M gene of several clinical HCoV-OC43 isolates, all from the same year (2001), which seems improbable.

In conclusion, we believe that the findings of St-Jean et al. regarding the genetic stability of human coronavirus OC43 are unlikely. We doubt that the HCoV-OC43 contemporary strain presented by the authors is a circulating strain from 2001, which is supported by an extremely low probability value. We propose that this might be due to contamination with the HCoV-OC43 ATCC strain of the cell line used to propagate the clinical isolate.

This work was supported by a fellowship from the Flemish Fonds voor Wetenschappelijk Onderzoek (FWO) to Leen Vijgen and by FWO grant G.0288.01.

REFERENCES

1. **Domingo, E., and J. J. Holland.** 1988. High error rates, population equilibrium and evolution of RNA replication systems, p. 3–36. *In* E. Domingo, J. J. Holland, and P. Ahlquist (ed.), RNA genetics, vol. 3. CRC Press, Boca Raton, Fla.
2. **Jenkins, G. M., A. Rambaut, O. G. Pybus, and E. C. Holmes.** 2002. Rates of molecular evolution in RNA viruses: a quantitative phylogenetic analysis. J. Mol. Evol. **54:**156–165.
3. **Kimura, M.** 1983. The neutral allele theory of molecular evolution. Cambridge University Press, Cambridge, United Kingdom.
4. **Rambaut, A.** 2000. Estimating the rate of molecular evolution: incorporating non-contemporaneous sequences into maximum likelihood phylogenies. Bioinformatics **16:**395–399.
5. **Salemi, M., W. M. Fitch, M. Ciccozzi, M. J. Ruiz-Alvarez, G. Rezza, and M. J. Lewis.** 2004. Severe acute respiratory syndrome coronavirus sequence characteristics and evolutionary rate estimate from maximum likelihood analysis. J. Virol. **78:**1602–1603.
6. **Sanchez, C. M., F. Gebauer, C. Sune, A. Mendez, J. Dopazo, and L. Enjuanes.**

Downloaded from http://jvi.asm.org/ on April 16, 2015 by ST ANDREWS UNIV

1992. Genetic evolution and tropism of transmissible gastroenteritis coronaviruses. Virology **190:**92–105.
7. **St-Jean, J. R., H. Jacomy, M. Desforges, A. Vabret, F. Freymuth, and P. J. Talbot.** 2004. Human respiratory coronavirus OC43: genetic stability and neuroinvasion. J. Virol. **78:**8824–8834.
8. **Swofford, D. L.** 1998. PAUP* 4.0: phylogenetic analysis using parsimony (*and other methods). Sinauer Associates, Sunderland, Massachusetts.
9. **Vabret, A., T. Mourez, S. Gouarin, J. Petitjean, and F. Freymuth.** 2003. An outbreak of coronavirus OC43 respiratory infection in Normandy, France. Clin. Infect. Dis. **36:**985–989.

Leen Vijgen
Philippe Lemey
Els Keyaerts
Marc Van Ranst*
Laboratory of Clinical and Epidemiological Virology
Department of Microbiology and Immunology
Rega Institute for Medical Research
University of Leuven, Leuven, Belgium
Minderbroedersstraat 10
B-3000 Leuven, Belgium

*Phone: (32 16) 347908
Fax: (32 16) 347900
E-mail: marc.vanranst@uz.kuleuven.ac.be

Authors' Reply

In their comment regarding our recently published paper (11), Vijgen et al. present different arguments to underline their hypothesis that the genetic stability of the human coronavirus HCoV-OC43 that we report should be seen as an highly improbable event.

According to Vijgen et al., statistical analysis based on RNA virus evolution models reveals that the mutation rate per site per year for HCoV-OC43 should be between 0.97×10^{-4} and 2.12×10^{-4} (95% confidence interval), with a maximum likelihood of 1.54×10^{-4}.

Based on their arguments, the six-nucleotide difference we found between the HCoV-OC43 ATCC strain and the HCoV-OC43 Paris isolate obtained in March 2001 (11) represents a mutation rate of 5.7×10^{-6}, which is 30-fold lower than expected from theoretical statistical calculations. Furthermore, using the Poisson process model, Vijgen et al. calculated that the probability of this result (only six nucleotide differences between the genome of the HCoV-ATCC strain and the Paris isolate) is only 1.9×10^{-16} (4.92×10^{-36} for six nucleotides or fewer).

The statistical arguments presented by Vijgen et al. appear very interesting and convincing in terms of statistical analysis based on some previously reported evolutionary rates of RNA viruses. Furthermore, as the authors of the paper comparing these two HCoV-OC43 variants (11), we do indeed acknowledge that our published data of apparent genetic stability is surprising. However, we are putting forward herein some hypotheses that may contribute to explain the similarity of the genome of the two HCoV-OC43 variants described in our paper (11), which may explain the apparent genetic stability of HCoV-OC43.

Before the sequencing of its genome, the HCoV-OC43 Paris isolate was passaged on the HRT-18 cell line six times. The human coronavirus contained in the original clinical isolate, which comes from the upper respiratory tract of a 68-year-old immunocompromised male who was not related whatsoever to laboratory work and was not in contact with any laboratory workers who had manipulated the HCoV-OC43 ATCC virus (citation from the Materials and Methods section of our paper), consisted, as a typical RNA virus, of a complex mixture of variants, or quasispecies (2, 3, 6). As reported by Elena (3), some sort of evolutionary constraint may exist for adaptation in a particular environment. According to this assumption, by having been cultured within the HRT-18 cell line for six passages, the HCoV-OC43 Paris strain may have been subject to this type of constraint, which would have highly favored its replication in the HRT-18 cells. In other words, the six passages on the HRT-18 cells would have selected a particular variant (with only a six-nucleotide difference from HCoV-OC43 ATCC) that was present in the quasispecies mixture in the clinical isolate. Furthermore, we can now add the supplemental data that the Paris variant was the only one among several clinical isolates that was able to replicate in HRT-18 cells, meaning that this particular isolate could have already been very similar to HCoV-OC43 ATCC at the onset of the viral isolation procedure (A. Vabret and F. Freymuth, unpublished results).

As already described in our paper (11), highly stringent laboratory precautions were used in order to eliminate a possible cross-contamination between the two viruses (citation from the Materials and Methods section of our paper: "The HCoV-OC43 ATCC strain and the Paris isolate were never cultured at the same time, and stringent laboratory precautions were used in order to eliminate possible cross-contamination."). We are still highly confident that all of the manipulations that served to either isolate or amplify the HCoV-OC43-Paris variant were performed properly, as we were always acutely aware of the potential problems of cross-contamination.

As Vijgen et al. are suggesting a possible contamination of the HRT-18 cell line used to propagate the HCoV-OC43-ATCC variant, we have performed RT-PCR assays on the HRT-18 cells that were used to replicate the HCoV-OC43 and were not able to detect any HCoV-OC43 RNA even after molecular hybridization using a probe specific for the M gene of the virus (Vabret and Freymuth, unpublished). A preexisting undetected infection of the HRT-18 cell line used to propagate HCoV-OC43 therefore appears highly unlikely.

The point mutation located in the M gene of the HCoV-OC43 Paris variant (T432C) is quite interesting. Indeed, while it is absent in the M gene of the HCoV-OC43 ATCC variant, this precise mutation is found in all 20 samples of HCoV-OC43 that were isolated during the Normandy outbreak of 2001 (13) and directly sequenced by Vabret et al. without any passage on any cell line (Vabret and Freymuth, unpublished). Furthermore, the T432C mutation is found neither in the first reported sequence of the M gene, which was reported by our laboratory (8), nor in the complete genome sequences of the HCoV-OC43 so far available (11; Vijgen et al., GenBank accession no. NC_005147). Therefore, this mutation appears to be a marker for the clinical isolates from the outbreak of HCoV-OC43 in Normandy during winter 2001 (13) and is not found in HCoV-OC43 ATCC.

As reported by Vijgen et al., the evolutionary rates of RNA viruses in general (6), and of coronaviruses in particular (10), are usually higher than the results presented in our paper (11). However, some recent publications reported a higher-than-expected genetic stability for RNA viruses. For example, it has been reported that the usually highly variable regions of the VP proteins of hepatitis A virus (4) and the highly variable *env* gene of feline immunodeficiency virus (FIV) (7) are shown experimentally to be more genetically stable than predicted statistically. Indeed, even though the latter is a retrovirus that could establish a latent infection, its mutation rate is usually comparable to that of human immunodeficiency virus type 1, at about 0.34% per site per year (5). However, Ikeda et al. (7)

Downloaded from http://jvi.asm.org/ on April 16, 2015 by ST ANDREWS UNIV

showed that FIV could remain very stable despite a 10-year infection in cats, as the mutation rate within the *env* gene was only 0.015% per year (7). This result represents an unexpected genetic stability, as the rate of 0.015% is about 23-fold lower than what was previously reported (5). This finding is in the same order of magnitude as the 30-fold less-than-expected variability calculated by Vijgen et al. on the basis of our published experimental data (11). Moreover, even though the SARS-HCoV was pointed out by Vijgen et al. as an example to illustrate the extreme genetic variability of coronaviruses (9), another very interesting paper on SARS-HCoV indicates instead that this virus shows a relative genetic stability, with only 2 mutations in the S gene and 1 mutation in the N gene found in 10 different clinical isolates over a 3-month period (12). Finally, another recent study on a coronavirus, the feline coronavirus, also indicated that this virus could be more genetically stable than theoretically expected. Indeed, the sequencing of a small portion of the S gene from samples isolated from five persistently infected cats revealed low rates of mutations, ranging from none over a period of 17 months for one cat to only nine nucleotide changes over a 5-year period for another animal (1).

In conclusion, given the above comments and supportive arguments from the published literature, and even though they are, as underlined by the comment of Vijgen et al., unexpected, we do stand by our results concerning the apparent genetic stability of the HCoV-OC43 virus.

REFERENCES

1. **Addie, D. D., I. A. T. Schaap, L. Nicolson, and O. Jarrett.** 2003. Persistence and transmission of natural type I feline coronavirus infection. J. Gen. Virol. **84:**2735–2744.
2. **Domingo, E.** 1997. Rapid evolution of viral RNA genomes. J. Nutr. **127**(suppl. 5)**:**958S–961S.
3. **Elena, S. F.** 2002. Restrictions to RNA virus adaptation: an experimental approach. Antonie Leeuwenhoek **81:**135–142.
4. **Gabrieli, R., G. Sanchez, A. Macaluso, F. Cenko, S. Bino, L. Palombi, E. Buonomo, R. M. Pinto, A. Bosch, and M. Divizia.** 2004. Hepatitis in Albanian children: molecular analysis of hepatitis A virus isolates. J. Med. Virol. **72:**533–537.
5. **Green, W. K., J. Meers, G. del Fierro, P. R. Carnegie, and W. F. Robinson.** 1993. Extensive sequence variation of feline immunodeficiency virus env genes in isolates from naturally infected cats. Arch. Virol. **133:**51–62.
6. **Holland, J., and E. Domingo.** 1998. Origin and evolution of viruses. Virus Genes **16:**13–21.
7. **Ikeda, Y., T. Miyazawa, Y. Nishimura, K. Nakamura, Y. Tohya, and T. Mikami.** 2004. High genetic stability of TM1 and TM2 strains of subtype B feline immunodeficiency virus in long-term infection. J. Vet. Med. Sci. **66:** 287–289.
8. **Mounir, S., and P. J. Talbot.** 1992. Sequence analysis of the membrane protein gene of human coronavirus OC43 and evidence for O-glycosylation. J. Gen. Virol. **73:**2731–2736.
9. **Salemi, M., W. M. Fitch, M. Cicozzi, M. J. Ruiz-Alvarez, G. Rezza, and M. J. Lewis.** 2004. Severe acute respiratory syndrome coronavirus sequence characteristics and evolutionary rate estimate from maximum likelihood analysis. J. Virol. **78:**1602–1603.
10. **Sanchez, C. M., F. Gebauer, C. Sune, A. Mendez, J. Dopazo, L. Enjuanes.** 1992. Genetic evolution and tropism of transmissible gastroenteritis coronaviruses. Virology **109:**92–105.
11. **St-Jean, J. R., H. Jacomy, M. Desforges, A. Vabret, F. Freymuth, and P. J. Talbot.** 2004. Human respiratory coronavirus OC43: genetic stability and neuroinvasion. J. Virol. **78:**8824–8834.
12. **Tong, S., J. R. Lingappa, Q. Chen, B. Shu, A. C. LaMonte, B. T. Cook, C. Birge, S. W. Chern, X. Liu, R. Galloway, L. Q. Mai, W. F. Ng, J.-Y. Yang, J. Butany, J. A. Comer, S. S. Monroe, S. R. Beard, T. G. Ksiazek, D. Erdman, P. A. Rota, M. A. Pallansch, and L. J. Anderson.** 2004. Direct sequencing of SARS-coronavirus S and N genes from clinical specimens shows limited variation. J. Infect. Dis. **190:**1127–1131.
13. **Vabret, A., T. Mourez, S. Gouarin, J. Petitjean, and F. Freymuth.** 2003. An outbreak of coronavirus OC43 respiratory infection in Normandy, France. Clin. Infect. Dis. **36:**985–989.

Julien R. St-Jean
Hélène Jacomy
Marc Desforges
Pierre J. Talbot*
Laboratory of Neuroimmunovirology
INRS-Institut Armand-Frappier
531 boulevard des Prairies
Laval, Québec H7V 1B7, Canada

Astrid Vabret
François Freymuth
Laboratoire de Virologie Humaine et Moléculaire
Centre Hospitalier Régional et Universitaire de Caen
Avenue Georges Clémenceau
14033 Caen, France

*Phone: (450) 686–5515
Fax: (450) 686–5566
E-mail: pierre.talbot@inrs-iaf.uquebec.ca

Downloaded from http://jvi.asm.org/ on April 16, 2015 by ST ANDREWS UNIV