Article published online: 2020-12-02

Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2): Codon Usage and Replicative Fitness

Darja Kanduc[1]

[1] Department of Biosciences, Biotechnologies, and Biopharmaceutics, University of Bari, Bari, Italy

Global Med Genet 2020;7:92–94.

Address for correspondence Darja Kanduc, PhD, Department of Biosciences, Biotechnologies, and Biopharmaceutics, University of Bari, Via Orabona 4, Bari 70125, Italy (e-mail: dkanduc@gmail.com).

Abstract

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) codon usage, as shown by the polyprotein coding sequence, shows better translation potential in the human host when compared with human coronavirus OC43 (HCoV-OC43) codon usage. Such translational advantage might facilitate SARS-CoV-2 replication, immunogenicity, and pathogenicity, thus also accounting for the less harmful character of HCoV-OC43 infection.

Keywords

- codon usage
- HCoV-OC43
- SARS-CoV-2
- viral replication fitness

Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection causes a respiratory syndrome with altered pulmonary and alveolar function that can evolve into acute respiratory insufficiency and death.[1] Progressive immune-associated injury is a hallmark of SARS,[2] and alteration of the lung functions is possibly due to specific autoimmune cross-reactions[3,4] against alveolar surfactant-related proteins,[5] with a higher titer of antibodies independently associated with a worse clinical classification.[6] In conflict, the human coronavirus OC43 (HCoV-OC43) generally relates to less serious disturbances as common cold.[7] Currently, the molecular determinants and the mechanisms that underlie such a different pathogenic load are unknown.

Based on previous reports[8,9] suggesting that rare host codons can inhibit viral protein expression and favor viral latency, this study investigated the codon usage in HCoV-OC43 and SARS-CoV-2. Specifically, usage of the 61 amino acid (aa) specifying codons was analyzed in the HCoV-OC43 and SARS-CoV-2 polyprotein (aka orf1ab) ORF (open reading frame) and then was compared with the codon usage of the human ORFeome.[10]

Main results are reported in ►**Table 1**, which shows the different usage of a set of eight codons, whereas full data for the 61 codons in *Homo sapiens*, HCoV-OC43, SARS-CoV-2, SARS-CoV, and Middle East respiratory syndrome coronavirus (MERS-CoV) are detailed in ►**Supplementary Table S1** (online only).

►**Table 1** describes the following:

- Eight codons are often used in HCoV-OC43 polyprotein ORF but occur at a lesser extent in the *H. sapiens* ORFeome, with a human-to-viral usage ratio smaller than 1, that is, from the translational point of view, the human-to-viral usage ratio is unfavorable to HCoV-OC43 since the optimal ratio value for HCoV-OC43 polyprotein synthesis in the human host is approximately 1.[8,9,11,12]
- The human-to-viral usage ratio remains suboptimal for translational expression in the three HCoV-OC43 isolates collected in 1987, 1990, and 2011, respectively.
- Usage of the eight codons is lower in SARS-CoV-2 polyprotein ORF so that the human-to-viral usage ratio reaches values closer to approximately 1 and is more suitable for the viral polyprotein translation in the human host.

In summary, in the context of CoV polyprotein expression, ►**Table 1** documents that eight codons are more often used in HCoV-OC43 polyprotein ORF than in the human ORFeome and might represent a translational constraint for HCoV-OC43 polyprotein expression, thus limiting HCoV-OC43 replication, diffusion, and pathogenicity, given the essential role of coronavirus polyprotein for generation of viral progeny.[13,14]

published online
December 2, 2020

DOI https://doi.org/10.1055/s-0040-1721080.
ISSN 2699-9404.

© 2020. The Author(s).
This is an open access article published by Thieme under the terms of the Creative Commons Attribution License, permitting unrestricted use, distribution, and reproduction so long as the original work is properly cited. (https://creativecommons.org/licenses/by/4.0/)

Georg Thieme Verlag KG, Rüdigerstraße 14, 70469 Stuttgart, Germany

Table 1 Codon usage bias between human ORFeome and polyprotein ORFs from HCoV-OC43 and SARS-CoV-2: a representative analysis of eight codons[a,b,c]

Aa	Codon	*Homo sapiens*	HCoV-OC43[d] Isolate: 1987	HCoV-OC43[e] Isolate: 1990	HCoV-OC43[f] Isolate: 2011	SARS-CoV-2[g] Isolate: 2019	SARS-CoV-2[h] Isolate:2020
Asp	GAT	21.8	54.5	54.5	54.8	35.4	35.5
Phe	TTT	17.6	50.2	49.8	50.6	35.8	35.9
Ile	ATT	16.0	30.7	30.6	31.1	23.8	23.9
Leu	TTG	12.9	32.4	32.4	32.4	16.3	16.5
Asn	AAT	17.0	41.7	41.7	41.9	37.8	37.9
Arg	CGT	4.5	12.0	11.8	11.7	8.3	8.2
Ser	AGT	12.1	25.2	25.7	25.8	16.8	16.7
Tyr	TAT	12.2	45.5	45.5	45.9	29.3	29.2

Abbreviations: HCoV-OC43, human coronavirus OC43; ORF, open reading frame; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.
[a]Codon frequency expressed per thousand.
[b]Codon usage of *Homo sapiens* obtained at www.kazusa.or.jp/codon.[10]
[c]Codon usage of polyprotein ORFs obtained at www.geneinfinity.org.
[d]HCoV-OC43/human/USA/1987; GenBank: KF530083.1; ID = AGT51637.1; taxon:31631.
[e]HCoV-OC43/human/USA/1990; GenBank: KF530088.1; ID = AGT51687.1; taxon:31631.
[f]HCoV-OC43/UK/London/2011; GenBank: KU131570.1; ID = AMK59674.1; taxon:31631.
[g]SARS-CoV-2/Wuhan-Hu-1/2019; GenBank:MN908947.3; ID = QHD43415.1; taxon:2697049.
[h]SARS-CoV-2/UW239/2020/USA; GenBank: MT251975.1; ID = QIQ68493.1; taxon:2697049.

The possibility of translational block/delay emerges when considering the iteration of the eight suboptimal codons along the HCoV-OC43 polyprotein ORFs. Among many, two examples from the HCoV-OC43 polyprotein (isolate 2011) are the following aa sequences: (1) YDDVNASLFVDYSNL that is coded by a row of codons (given capital) abundant in the viral polyprotein ORF but not in the human ORFeome, TAT-GAT-GAT-gtt-AAT-gct-AGT-TTG-TTT-gtg-GAT-TAT-AGT-AAT-TTG, and (2) IDDHRITSITSDKFDFII that is coded by the viral nucleotide sequence ATT-GAT-GAT-cat-CGT-atc-act-AGT-ATT-act-AGT-GAT-aag-TTT-GAT-TTT-ATT-ATT, where 12 codons out of 15 are suboptimal for translation in the human host. Clearly, the potential to be translated of such HCoV-OC43 coding sequences progressively and severely diminishes along the succession of the suboptimal codons.

To conclude, the data suggest a link between CoV codon usage and CoV pathogenicity in humans. Usage of relatively rare human codons and their clustering along viral sequences can represent a major translational block at the basis of HCoV-OC43 low expression, low immunogenicity, and low pathogenicity. Vice versa, the possibility to be translated is rescued in SARS-CoV-2 polyprotein ORF by the minor usage of rare human codons (►**Table 1**). When compared with HCoV-OC43, the higher translatability of SARS-CoV-2 correlates to a higher viral protein expression and a higher capability of evoking immune responses. Consequently, also, anamnestic immune responses of increased avidity and affinity can arise and lead to autoimmune pathologies in case of repeated exposures to SARS viruses,[15–18] according to the recently clarified phenomenon of immunologic memory imprinting, also known as Original Antigenic Sin.[19]

More in general, the data further support the translational control of viral protein expression as a mechanism by which the human host can silence and tolerize viral invasion.[8,9] Hence, this study warns that microbiology methodologies such as codon optimization, insertion/modification of translational enhancers, or addition of viral vectors *inter alia*, can increase the expression, replicative fitness, diffusion, and pathogenicity of the infectious agents under study, thus altering the finely tuned equilibrium between immunogenicity and immunotolerance.[20,21]

Funding
None.

Conflict of Interest
None declared.

References

1 Wang Z, Yang B, Li Q, Wen L, Zhang R. Clinical features of 69 cases with Coronavirus Disease 2019 in Wuhan, China. Clin Infect Dis 2020;71(15):769–777
2 Cameron MJ, Bermejo-Martin JF, Danesh A, Muller MP, Kelvin DJ. Human immunopathogenesis of severe acute respiratory syndrome (SARS). Virus Res 2008;133(01):13–19
3 Kanduc D. From anti-SARS-CoV-2 immune responses to COVID-19 via molecular mimicry. Antibodies (Basel) 2020;9(03):E33
4 Lin YS, Lin CF, Fang YT, et al. Antibody to severe acute respiratory syndrome (SARS)-associated coronavirus spike protein domain 2 cross-reacts with lung epithelial cells and causes cytotoxicity. Clin Exp Immunol 2005;141(03):500–508
5 Kanduc D, Shoenfeld Y. On the molecular determinants of the SARS-CoV-2 attack. Clin Immunol 2020;215:108426
6 Zhao J, Yuan Q, Wang H, et al. Antibody responses to SARS-CoV-2 in patients of novel coronavirus disease 2019. Clin Infect Dis 2020 (e-pub ahead of print). Doi: 10.1093/cid/ciaa344

7 Lim YX, Ng YL, Tam JP, Liu DX. Human coronaviruses: a review of virus–host interactions. Diseases 2016;4(03):26
8 Kanduc D. Role of codon usage and tRNA changes in rat cytomegalovirus latency and (re)activation. J Basic Microbiol 2016;56(06): 617–626
9 Kanduc D. Rare human codons and HCMV translational regulation. J Mol Microbiol Biotechnol 2017;27(04):213–216
10 Nakamura Y, Gojobori T, Ikemura T. Codon usage tabulated from international DNA sequence databases: status for the year 2000. Nucleic Acids Res 2000;28(01):292
11 Quax TE, Claassens NJ, Söll D, van der Oost J. Codon bias as a means to fine-tune gene expression. Mol Cell 2015;59(02):149–161
12 Supek F. The code of silence: widespread associations between synonymous codon biases and gene function. J Mol Evol 2016;82 (01):65–73
13 Tijms MA, Nedialkova DD, Zevenhoven-Dobbe JC, Gorbalenya AE, Snijder EJ. Arterivirus subgenomic mRNA synthesis and virion biogenesis depend on the multifunctional nsp1 autoprotease. J Virol 2007;81(19):10496–10505
14 Krichel B, Falke S, Hilgenfeld R, Redecke L, Uetrecht C. Processing of the SARS-CoV pp1a/ab nsp7-10 region. Biochem J 2020;477 (05):1009–1019
15 Agrawal AS, Tao X, Algaissi A, et al. Immunization with inactivated Middle East Respiratory Syndrome coronavirus vaccine leads to lung immunopathology on challenge with live virus. Hum Vaccin Immunother 2016;12(09):2351–2356
16 Cameron MJ, Kelvin AA, Leon AJ, et al. Lack of innate interferon responses during SARS coronavirus infection in a vaccination and reinfection ferret model. PLoS One 2012;7(09):e45842
17 Tseng CT, Sbrana E, Iwata-Yoshikawa N, et al. Immunization with SARS coronavirus vaccines leads to pulmonary immunopathology on challenge with the SARS virus. PLoS One 2012;7(04):e35421
18 Yasui F, Kai C, Kitabatake M, et al. Prior immunization with severe acute respiratory syndrome (SARS)-associated coronavirus (SARS-CoV) nucleocapsid protein causes severe pneumonia in mice infected with SARS-CoV. J Immunol 2008;181(09):6337–6348
19 Kanduc D, Shoenfeld Y. Inter-pathogen peptide sharing and the original antigenic sin: solving a paradox. Open Immunol J 2018; 8:16–27
20 Kanduc D. Immunogenicity, Immunopathogenicity, and Immunotolerance in One Graph. Anticancer Agents Med Chem 2015; 15:1264–1268. Doi: 10.2174/1871520615666150716105543
21 Kanduc D. Hydrophobicity and the Physico-Chemical Basis of Immunotolerance. Pathobiology 2020;87:268–276