

# SARS-CoV-2 codon usage bias downregulates host expressed genes with similar codon usage

Andres Mariano Alonso<sup>2,3</sup> and Luis Diambra<sup>1,3,\*</sup>

<sup>1</sup> *CREG, Universidad Nacional de La Plata. La Plata, Argentina.*

<sup>2</sup> *InTech, Universidad Nacional de San Martin. Chascomus, Argentina.*

<sup>3</sup> *CONICET, Argentina.*

Correspondence\*:

Luis Diambra

ldiambra@gmail.com

## ABSTRACT

Severe acute respiratory syndrome is quickly spreading throughout the world and was declared as a pandemic by the World Health Organisation (WHO). The pathogenic agent is a new coronavirus (SARS-CoV-2) that infects pulmonary cells with great effectiveness. In this study we focus on the codon composition for the viral proteins synthesis and its relationship with the proteins synthesis of the host. Our analysis reveals that SARS-CoV-2 preferred codons have poor representation of G or C nucleotides in the third position, a characteristic which could conduct to an unbalance in the tRNAs pools of the infected cells with serious implications in host protein synthesis. By integrating this observation with proteomic data from infected cells, we observe a reduced translation rate of host proteins associated with highly expressed genes, and that they share the codon usage bias of the virus. The functional analysis of these genes suggests that this mechanism of epistasis contributes to understand some deleterious collateral effect as result of the viral replication. In this manner, our finding contribute to the understanding of the SARS-CoV-2 pathogeny and could be useful for the design of a vaccine based on the live attenuated strategy.

**Keywords:** SARS-CoV-2, codon usage bias, translational control, pathogeny, vaccine design.

## 1 INTRODUCTION

The new SARS-CoV-2 coronavirus is the causative agent of the current pandemic of COVID-19. This highly pathogenic virus has quickly become the latest threat to the modern human lifestyle. Since the end of 2019 up to the redaction of this paper, this virus has infected over 3 million people, leading to mild symptoms from fever, lung function reduction, to severe acute respiratory syndrome (SARS) and even death. Today there are over 215.000 death all over the world. However, we are still far from determining the final mortality figure. In the absence of vaccines or effective antiviral treatments against SARS-CoV-2, it is important to understand how this virus appropriates the host translation apparatus and subverts the immune defences of infected cells. This can be the first step in the development of novel therapeutics.

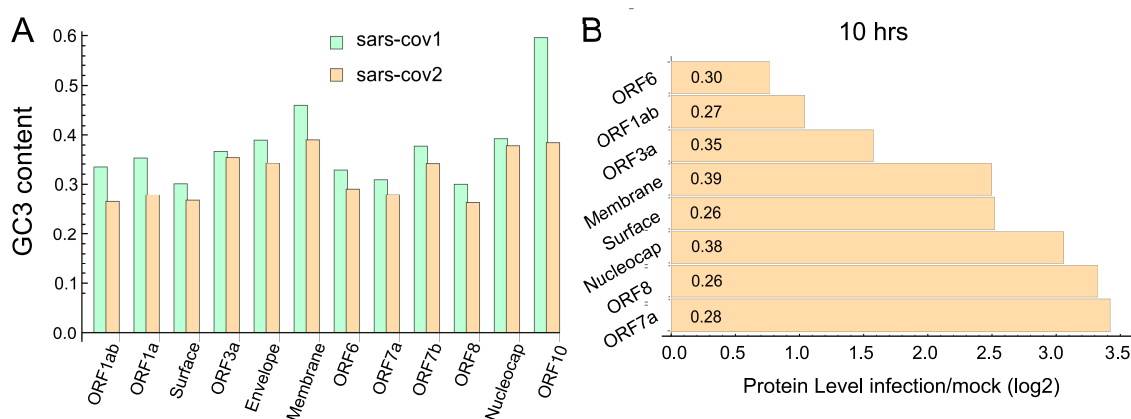
As intracellular parasites, viruses replication mandatory depends on the translational machinery of their cellular hosts to translate viral transcripts. Thus, virus replication requires ribosomes, tRNA and translation factors from the cell host. On the other hand, codon usage bias is a feature subject to natural selection and affects the genomes from all domains of life. It is known that more frequently used codons

are used for coding highly abundant proteins (30, 9, 35, 11). Virus genomes have also preference in the codon usage, but in this case the bias is constrained by the host translational machinery (38). The effects of codon composition of a transcript on its translation have been reported in literature (34, 13, 26, 39), and is considered an important determinant of gene expression (41, 42, 46). However, the codon usage of a gene also affects the translation of other genes (12). In fact the virus replication demands not only ribosomes, but also a lot tRNA resource for the most demanded codons. Thus, the consumption of specific tRNAs for the virus replication could be an alternative to control host protein synthesis machinery, as well as generating deleterious collateral effects on the function of the host cells.

Here, we study recent proteomic data from SARS-CoV-2 infected cells (4) from a novel point of view. By considering the codon usage of the virus ORFeome, we characterise a set of genes whose expression could be affected by the massive demand of the tRNA which implies the virus replication. We find that those host genes encoding proteins with similar codons to the virus ORFeome have lower translational rate. Extrapolating this finding to highly expressed genes in lung, we find a small set of genes that can be downregulated. These genes are involved in translation, immune systems, cell calcification, etc., and their roles on the SARS-CoV-2 cell pathogenesis should be the target of further studies.

## 2 RESULTS

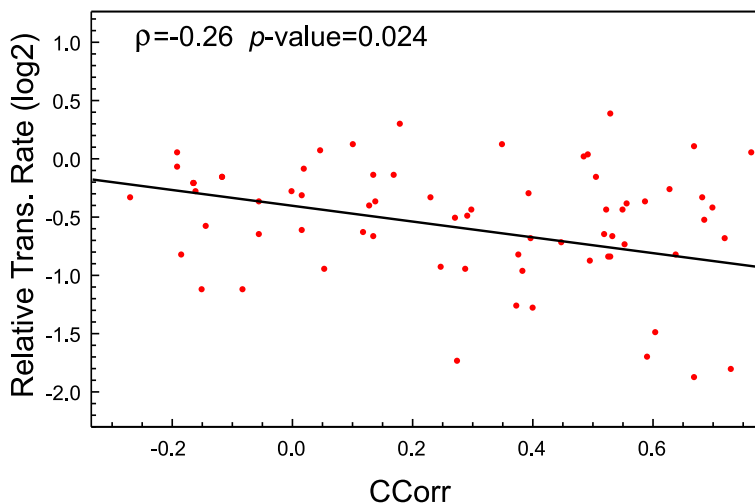
We study codon composition of two coronavirus complete genomes by measuring the percentage of G and C nucleotides at the wobble position of the codons (GC3). Figure 1A depicts the GC3 content of each annotated coding sequence of the SARS-CoV and SARS-CoV-2. The comparison of the GC3 content reveals that codons preferentially used by the new virus have lower content of G and C at the third position than SARS-CoV. In particular, the ORFs corresponding to the proteins ORF1ab, ORF1a, surface, ORF6, ORF7a and ORF8 exhibit a very low content of GC3. In particular, the last two proteins are highly expressed at 10 hrs PI, as it can be seen in Figure 1B.



**Figure 1.** SARS-CoV-2 ORFeome have lower content of GC3. **(A)** Bar chart comparing the G and C nucleotides on the third nucleotide (GC3) at every codon in the ORFeome of SARS-CoV and SARS-CoV-2. **(B)** Bar chart showing the proteins level of nine viral proteins at 10 hrs post infection (data obtained from (4)). The GC3 content for each viral ORF is showed in the corresponding bar.

Most of these proteins are highly expressed at 10h post infection (PI). This implies that the replication of SARS-CoV-2 would demand a higher supply for specific tRNAs than his ancestor. In turn, this could lead to an imbalance in the tRNA pool needed for the normal synthesis of the proteins of the host cell, altering

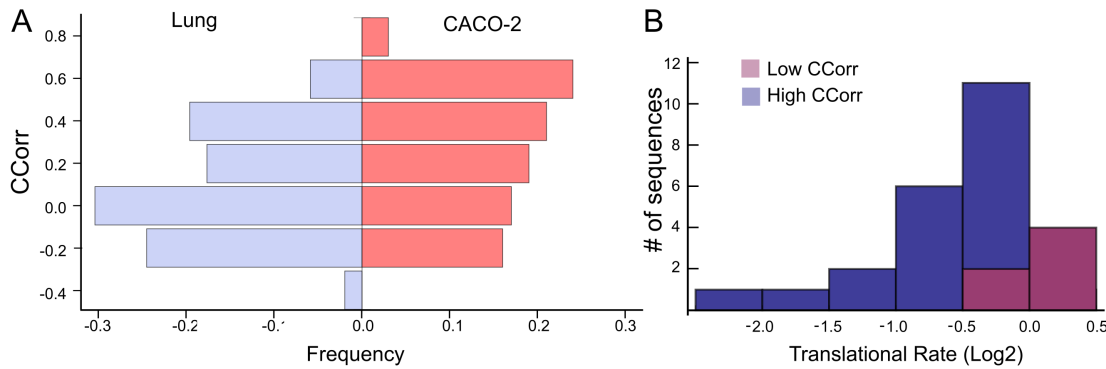
its proteostasis. Of course, as tRNA pools depend on the cell types, the postulated imbalance could be a tissue-dependent feature. To check this hypothesis, we make use of the recently available data about the proteome profile and translational rate in SARS-CoV-2-infected cells (4). In this study CACO-2 cell line was infected with SARS-CoV-2 and mocked, the last one was used as control. From this proteome we select the coding sequences corresponding to the 100 most abundant proteins in the mock infected cells at 10h PI. Then, we compute the codon correlation (CCorr) between the codon usage of each one of these sequences and the codon usage of all coding sequences in SARS-CoV-2. Figure 2 depicts a raster plot of the translational rate of these genes from SARS-CoV-2-infected cells *versus* the corresponding CCorr (red dots). The black line is the adjusted linear model, which shows a significant and negative correlation between the translation rate and the codon composition of each sequence. This means that the translational rate of coding sequence in SARS-CoV-2-infected cells are lower than in mock-infected cells. Although the correlation ( $\rho = 0.26$ ) with codon composition is not high, reflecting the fact that other factors could be regulating the translation rate, is significant at a of 0.05 ( $p$ -value= 0.024 < 0.05).



**Figure 2.** The translational rate of genes with high CCorr in SARS-CoV-2-infected cells is lower than the mock-infected cells. Raster plot showing negative correlation between translational rate of the 100 most expressed proteins in SARS-Cov-2 infected CACO-2 cells versus virus codon usage.

This analysis suggests then that the codon composition of highly expressed host-genes is a determinant of its own translation rate, and that viral replication induces a particular case of epistasis which could affect host cell proteostasis. Consequently, we find it relevant to extrapolate this phenomenon to the cells that conform lung tissue, the main target of the SARS-CoV-2 infection. To this aim, we select the 100 most expressed genes in lung from the GTEx database, and compute the CCorr for each sequence. Figure 3A shows that the codon composition profile of the most expressed genes from both cells types are different, it is evident that the lung cells share fewer codons in common with the SARS-CoV-2 than the CACO-2 cells. Consequently, the tRNAs pool used for the virus could be a more scarce supply in lung cells than in the CACO-2 cells.

We search for those genes, highly expressed in the lung tissue, which could be affected by the depletion of the tRNAs consumed by virus replication. They are those whose codon composition is similar to SARS-CoV-2 (CCorr > 0.25), they are listed in Table 1. Before analysing these 27 genes of interest, we compare the translation rate of these gene with the translation rate associated with highly expressed genes, but which differs in their codon usage bias (CCorr < -0.075). In this sense, we used the Mann-Whitney U test



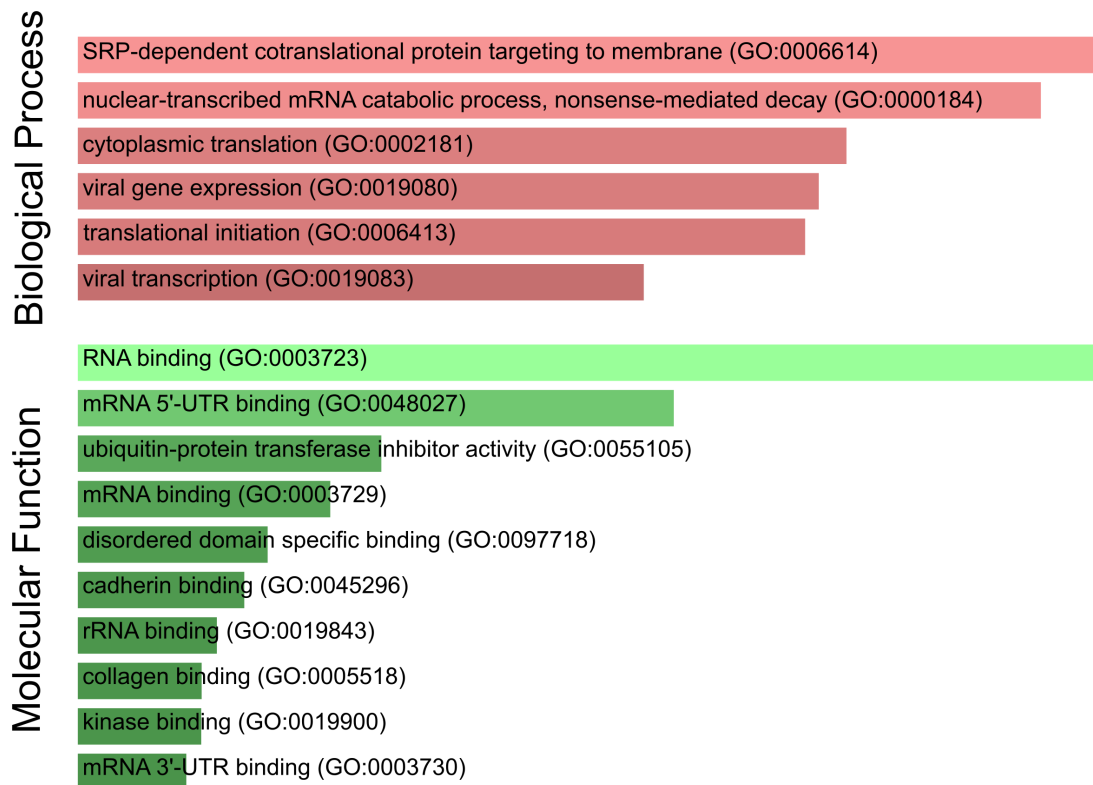
**Figure 3.** High codon correlation between SARS-CoV-2 and lung cells is associated with a translational rate decay over specific transcripts. **(A)** Frequency distribution of CCorr on the 100 highest expressed genes in lung cells and CACO-2 cells. **(B)** Histogram of the translational rate of 100 highest expressed genes in lung cells. Transcripts with high and low codon correlation (CCorr) with SARS-CoV-2 are highlighted.

for median differences of independent samples to analyse the difference in translational rate between these two groups of genes. This test identified significant differences between these groups ( $p$ -value=0.0007), as it can be seen in Fig. 3B. We want to remark that, since there is still not available data on translational rate from lung cells exist, the last analysis was performed by using translational rate data available from CACO-2 cells, even knowing that this data would be underestimating the difference between these two groups of genes.

Further, we perform a GO term enrichment analysis over the 27 genes listed in Table 1, whose translation rate is decreased, by means of the Enrichr online software (7). Figure 4 illustrates the main enrichment pathways. This analysis reveals that codon usage could promote extensive changes in the translation machinery of the host in agreement with previous report in CACO-2 cells (4). It is known that when canonical translation is impaired, as part of the host defence program, specific 40S ribosomal subunits are needed to support uncapped viral mRNA translation (22).

Our list of genes is enriched with several ribosomal proteins that constitute both 40S (RPS6, RPS3A, RPS7, RPSA, RPS25, RPS13, RPS12, RPS27A) and 60S (RPL4, RPL5, RPL21, RPLP1, RPL34, RPL9, RPL24, RPL17) ribosomal subunits. All of them belong to the nonsense-mediated decay pathway that control the mRNAs with abnormal termination. Termination codon can be recognised if the 3' untranslated region is short, or if it does not have an exon junction complex downstream of the termination codon (28). These genes are also part of the process of SRP-dependent cotranslational protein destined for the endoplasmic reticulum.

It is known that these proteins are involved in translation, in particular the gene *EEF1A1*, an isoform of the alpha subunit of the elongation factor-1 expressed in lungs, which is responsible for the GTP-dependent binding of aminoacyl-tRNA to the ribosome. Beyond its function in translation, *EEF1A1* takes part in the innate immune system, by activating directly the transcription of IFN-gamma (25). IFN-gamma activates immune cells, such as macrophages and natural killer cells and stimulates the major histocompatibility complex (MHC) II-dependent presentation of antigens to the immune system(37). This process could be downregulated due to a decrease in the translation rate of *EEF1A1*. Furthermore, *B2M*, other gene in the list of decreased translational rate, encodes one of the proteins that conform MHC class I found on the cell surface of all nucleated cells (2). Together with *TXNIP*, *RPS3A* and *RPS27A*,



**Figure 4.** SARS-CoV-2 codon usage may have an impact on host translational machinery. GO enrichment analysis performed over genes of interest. The most enriched biological processes and molecular functions are highlighted.

it is involved in the antigen processing and presentation. Moreover, this gene, along with S100A11, HSP90AB1 and also EEF1A1, regulates the exocytosis of granules containing inflammatory mediators in neutrophils (24). In addition, our analysis shows a translation decrease in A2M, a plasmatic protein that inhibits a broad spectrum of proteases, including trypsin, thrombin and collagenase (6). Finally, we highlight the presence in our analysis of the HSP90AB1, a known chaperone that facilitates the maturation of a wide range of proteins and its attenuation has been related to idiopathic pulmonary fibrosis and cystic fibrosis (16, 43).

In addition, we have also identified two genes (SAT1 and MGP) belonging to the pathway of endothelial cell calcification regulated by NOTCH1 (44). This finding acquires special relevance in the context of the acute lung injury (ALI) observed in many infected patients (18). The first gene SAT1 catalyses the acetylation of polyamines (spermidine and spermine) and carries it out of the cell. The polyamine excess is a prominent source of oxidative stress which can increase inflammatory response (20). Polyamines have also been connected with the immune system (32). On the other hand, the MGP gene encodes the matrix gla protein which is also highly expressed in all vasculatures. Recent studies suggest that MGP downregulates the tissue calcification by sequestering bone morphogenetic proteins (44). Mutations in this gene cause Keutel syndrome, which is characterised by peripheral pulmonary stenosis and abnormal cartilage calcification (27).

Other interesting protein in our study is SPARC-like 1 (SPARCL-1), also known as Hevin, commonly associated with regulation of cell migration, modulation of extracellular matrix proteins (15) and is has

Gene Name	Codon corr	T. rate 6hs	T. rate 10hs
RPS3A	0.582	0.686	0.980
RPL5	0.559	0.585	0.271
RPL9	0.546	0.723	0.738
SPARCL1	0.536	NA	NA
EEF1A1	0.531	0.948	0.627
RPS12	0.443	0.155	0.649
RPL17	0.438	NA	0.372
TPT1	0.415	1.048	0.981
RPL21	0.411	NA	NA
RPS6	0.398	1.206	0.590
RPL34	0.376	0.981	0.727
RPS13	0.372	1.386	0.209
RPL4	0.370	0.919	0.799
RPS27A	0.365	NA	NA
RPSA	0.352	0.902	0.643
SAT1	0.350	NA	NA
A2M	0.340	NA	NA
TXNIP	0.317	NA	NA
RPS25	0.315	0.728	0.753
FN1	0.308	NA	NA
RPL24	0.306	0.576	0.890
RPLP1	0.303	NA	NA
RPS7	0.297	0.965	1.036
HSP90AB1	0.296	0.989	0.738
B2M	0.277	0.926	0.825
S100A11	0.273	1.620	0.520
MGP	0.272	NA	NA

**Table 1.** Set of most expressed genes in lung (GTEx) and positive correlation with virus codon usage. col 1: gene name, col 2: CCorr, cols 4 and 5 translational rate, infected CACO-2 cells, measured at 6 and 10 hr PI, respectively. Translational rate are relative to mock cells.

been shown to be involved in lymphocyte transendothelial migration through high endothelial venules (14). In this context, it is important to mention that clinical studies over patients that suffer several cases of COVID19 document a dysregulation of immune response related particularly with a lower lymphocytes count (18). In addition, our analysis reveals a translational rate decay of fibronectin (FN1), a master organiser of extracellular matrices that mediates cellular interactions playing important roles in cell adhesion, hemostasis and thrombosis (31, 43). According to our observation previously data show that SARS-CoV infected cells experiment a downregulation in fibronectin expression (40); as in other human virus like cytomegalovirus this last observation could be associated with pathogenesis of tissue inflammation (36).

### 3 DISCUSSION

The viral infection of human cells triggers an ensemble of host processes based on the interferons that interfere with viral replication. These processes have coevolved with the viral response to the host defence, thus virus counteract these process by a diversity of immuno-modulatory mechanisms. For example, NS1 protein plays a central role in the influenza infection by suppressing the host IFNs response. Further, the N protein of porcine reproductive and respiratory syndrome virus impairs the IFN transcription by acting over the TXK, which together with the EEF1A1 and PARP1 form the trimolecular complex that binds to the IFN- $\gamma$  gene promoter (25, 21). Hemagglutinin of IAVs has been shown to facilitate IFNAR

ubiquitination and degradation, reducing the levels of IFNAR, and thus suppressing the expression of IFN-stimulated antiviral proteins (45).

These examples illustrate the action of viral dedicated factors that downregulate the transcription of IFNs. Up to the present, no dedicated factor with analogue function have been identified in SARS-CoV and SARS-CoV-2. However, a recent report found a significant lack of IFN type I and III at the transcriptional level in human alveolar adenocarcinoma cells (3). On the other hand, a marked up-regulation of inflammatory mediators at the protein level (CXCL10, CCL2, IFN- $\alpha$  and  $\gamma$ ) has been observed in patients with SARS-CoV, without a significant amount of specific antibody (5). Several case of absence of protective immunity due to previous infection seem to indicate a similar landscape for COVID-19. Until now, the manner in which some patients fail to develop adaptive immunity is yet to be elucidated. As mentioned in the Results section, the decreased translate of B2M could be related with this last observation since is a crucial factor for the stable presentation of antigens derived from virus or tumours proteins; this antigens are recognised by cytotoxic T cells that eventually eliminate the target cell stimulating apoptosis to prevent systemic dissemination of the disease (19).

MGP is also expressed at high levels in heart, kidney, and lung which is particular interesting in the context of several comorbidities and collateral effects observed in the COVID-19 patients (29). MGP is activated by a Vitamin K dependent carboxylase (10). Interestingly, an epidemiological study with hemodialysis patients showed that Vitamin K<sub>2</sub> could prevent vascular calcification diseases (1). In contrast, the anticoagulant warfarin, which inhibits Vitamin K-dependent enzymes, was recently associated with an increased incidence of aortic valve calcification (17).

Summing up, if the depletion of a selected set of tRNA, induced by virus replication, affects the expression level or the co-translation folding of this proteins, one could expect the emergence of several systemic disorders.

## **4 CONCLUSION**

Codon usage bias is thought to have significant effects on translation rate, where rare codons are assumed to be translated more slowly than common codons (33). It is assumed that rare and common codons are defined by usage rates of highly expressed genes. On the other hand, it is known that synonymous variations in a coding sequence can affects its folding and/or resulting expression level (26). However, whether the codon composition of viral ORFome can affect the translation rate of host genes has not been thoroughly explored. Here, we have shown that the synthesis of the proteins associated with highly expressed genes, and with similar codon usage to the one of the virus, appears to be downregulated. Following this idea we determine which genes in lung could be affected by the viral replication. A functional analysis of these genes reveal that they could be related to collateral effects observed in COVID-19 patients (18). Further studies are mandatory to corroborated or discard the putative relationship established here.

One of the main obstacles in the recent development of vaccines has been the finding of increased infectivity observed to occur after immunisations with whole virus vaccines or complete spike protein vaccines. This phenomenon has been observed both in vaccines against SARS coronavirus as in respiratory syncytial virus. On the other hand, other vaccine strategy has been recently assayed, focusing on altering the codon-pair usage without affecting protein sequence. This codon deoptimization strategy has been reduced virus replication (8, 23). We believe that our results shed light on how codon use could affect virus

attenuation, help decrease the damaging side effect, providing an exciting opportunity for live-attenuated vaccine development.

## 5 MATERIALS AND METHODS

The coding sequences associated with the genome of SARS-CoV and SARS-CoV-2 were obtained from the NCBI (NC\_004718.3 and NC\_045512.2, respectively). Highly expressed genes in the lung tissue were retrieved from the GTEx portal ([gtexportal.org](http://gtexportal.org)). Proteomic and translational data from infected CACO-2 cells were retrieved from supplementary material of preprint (4), which are publically available. These data contain protein levels from 6381 proteins and translational rate from 2715 proteins, consequently the proteomic information is not available for many proteins.

Given the coding sequence  $s$ , we compute the codon usage frequency as  $f_s(c) = \frac{N_c}{L_s}$ , where  $L_s$  is total number of codons in the sequence  $s$ , and  $N_c$  is the number of times that codon  $c$  is present in  $s$ . In a similar manner, we define codon usage frequency relative to the viral ORFome:  $f_O(c) = \frac{N_c}{L_O}$ , where  $L_O$  is the total number of codons in the ORFome, and  $N_c$  is the number of occurrences observed codon  $c$  in the ORFome. Thus,  $f_s(c)$  and  $f_O(c)$  are vector of 64 elements. To compute the CCorr of a given sequence  $s$  with viral ORFome we consider the Pearson's correlation coefficient between the vector  $f_s(c)$  and  $f_O(c)$ , as follows:

$$\text{CCorr} = \frac{\sum_c (f_s(c) - \bar{f}_s) \cdot (f_O(c) - \bar{f}_O)}{\sqrt{\sum_c (f_s(c) - \bar{f}_s)^2} \cdot \sqrt{\sum_c (f_O(c) - \bar{f}_O)^2}}. \quad (1)$$

## CONFLICT OF INTEREST STATEMENT

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## FUNDING

AA is a postdoctoral fellow of the CONICET (Argentina) and LD is researcher member of the CONICET (Argentina).

## ACKNOWLEDGMENTS

We would like to thank Nara Guisoni, Alejandra Carrea and Veronica C ceres for their insightful suggestions and discussion. We are thankful to Catalina Trousdell for her judicious comments on the manuscript.

## DATA AVAILABILITY STATEMENT

The datasets analysed for this study can be found in the ProteomeXchange repository (ID=PXD017710) <http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PXD017710>



## REFERENCES

1. Aoun, M., Makki, M., Azar, H., Matta, H., and Chelala, D. N. (2017). High dephosphorylated-uncarboxylated MGP in hemodialysis patients: risk factors and response to vitamin K2, A pre-post intervention clinical trial. *BMC Nephrology* doi:10.1186/s12882-017-0609-3
2. Bernier, G. M. (1980).  $\beta$ 2-microglobulin: structure, function and significance. *Vox Sanguinis* 38, 323–327. doi:10.1111/j.1423-0410.1980.tb04500.x
3. Blanco-Melo, D., Nilsson-Payant, B. E., Liu, W.-C., Møller, R., Panis, M., Sachs, D., et al. (2020). Sars-cov-2 launches a unique transcriptional signature from in vitro, ex vivo, and in vivo systems. *bioRxiv* doi:10.1101/2020.03.24.004655
4. Bojkova, D., Klann, K., Koch, B., Widera, M., Krause, D., Ciesek, S., et al. (2020). Sars-cov-2 infected host cell proteomics reveal potential therapy targets. *Preprint* doi:10.21203/rs.3.rs-17218/v1
5. Cameron, M. J., Ran, L., Xu, L., Danesh, A., Bermejo-Martin, J. F., Cameron, C. M., et al. (2007). Interferon-mediated immunopathological events are associated with atypical innate and adaptive immune responses in patients with severe acute respiratory syndrome. *Journal of Virology* doi:10.1128/jvi.00527-07
6. Cater, J. H., Wilson, M. R., and Wyatt, A. R. (2019). Alpha-2-macroglobulin, a hypochlorite-regulated chaperone and immune system modulator. *Oxidative medicine and cellular longevity* doi:10.1155/2019/5410657
7. Chen, E. Y., Tan, C. M., Kou, Y., Duan, Q., Wang, Z., Meirelles, G. V., et al. (2013). Enrichr: Interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* doi:10.1186/1471-2105-14-128
8. Coleman, J. R., Papamichail, D., Skiena, S., Futcher, B., Wimmer, E., and Mueller, S. (2008). Virus attenuation by genome-scale changes in codon pair bias. *Science* 320, 1784–7. doi:10.1126/science.1155761
9. Dana, A. and Tuller, T. (2014). The effect of tRNA levels on decoding times of mRNA codons. *Nucleic Acids Research* doi:10.1093/nar/gku646
10. Danziger, J. (2008). Vitamin K-dependent proteins, warfarin, and vascular calcification. *Clinical Journal of the American Society of Nephrology* doi:10.2215/CJN.00770208
11. Diambra, L. A. (2017). Differential bicodon usage in lowly and highly abundant proteins. *PeerJ* doi:10.7717/peerj.3081
12. Frumkin, I., Lajoie, M. J., Gregg, C. J., Hornung, G., Church, G. M., and Pilpel, Y. (2018). Codon usage of highly expressed genes affects proteome-wide translation efficiency. *Proceedings of the National Academy of Sciences of the United States of America* 115, E4940–E4949. doi:10.1073/pnas.1719375115
13. Gingold, H. and Pilpel, Y. (2011). Determinants of translation efficiency and accuracy. *Molecular Systems Biology* doi:10.1038/msb.2011.14
14. Girard, J. P. and Springer, T. A. (1995). Cloning from purified high endothelial venule cells of hevin, a close relative of the antiadhesive extracellular matrix protein SPARC. *Immunity* doi:10.1016/1074-7613(95)90083-7
15. Girard, J. P. and Springer, T. A. (1996). Modulation of endothelial cell adhesion by hevin, an acidic protein associated with high endothelial venules. *Journal of Biological Chemistry* doi:10.1074/jbc.271.8.4511
16. Haase, M. and Fitze, G. (2015). Hsp90abl: Helping the good and the bad. *Gene* 575. doi:10.1016/j.gene.2015.08.063

- 17 .Holden, R. M., Sanfilippo, A. S., Hopman, W. M., Zimmerman, D., Garland, J. S., and Morton, A. R. (2007). Warfarin and aortic valve calcification in hemodialysis patients. *Journal of Nephrology*
- 18 .Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., et al. (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet* doi:10.1016/S0140-6736(20)30183-5
- 19 .Hulpke, S. and Tampé, R. (2013). The MHC I loading complex: A multitasking machinery in adaptive immunity. *Trends in Biochemical Sciences* doi:10.1016/j.tibs.2013.06.003
- 20 .Hussain, T., Tan, B., Ren, W., Rahu, N., Dad, R., Kalhor, D. H., et al. (2017). Polyamines: therapeutic perspectives in oxidative stress and inflammatory diseases. *Amino Acids* doi:10.1007/s00726-017-2447-9
- 21 .Kenney, S. P. and Meng, X. J. (2015). An SH3 binding motif within the nucleocapsid protein of porcine reproductive and respiratory syndrome virus interacts with the host cellular signaling proteins STAM1, TXK, Fyn, Hck, and cortactin. *Virus Research* doi:10.1016/j.virusres.2015.04.004
- 22 .Kwan, T. and Thompson, S. R. (2019). Noncanonical translation initiation in Eukaryotes. *Cold Spring Harbor perspectives in biology* 11. doi:10.1101/cshperspect.a032672
- 23 .Le Nouen, C., Brock, L. G., Luongo, C., McCarty, T., Yang, L., Mehedi, M., et al. (2014). Attenuation of human respiratory syncytial virus by genome-scale codon-pair deoptimization. *Proceedings of the National Academy of Sciences of the United States of America* 111, 13169–13174. doi:10.1073/pnas.1411290111
- 24 .Lee, K., Kang, H.-S., Jeon, J., Kim, E.-M., Yoon, S.-R., Song, H., et al. (2005). Vdup1 is required for the development of natural killer cells. *Immunity* 22, 195–208. doi:10.1016/j.immuni.2004.12.012
- 25 .Maruyama, T., Nara, K., Yoshikawa, H., and Suzuki, N. (2007). Txk, a member of the non-receptor tyrosine kinase of the Tec family, forms a complex with poly(ADP-ribose) polymerase 1 and elongation factor 1 $\alpha$  and regulates interferon- $\gamma$  gene transcription in Th1 cells. *Clinical and Experimental Immunology* doi:10.1111/j.1365-2249.2006.03249.x
- 26 .McCarthy, C., Carrea, A., and Diambra, L. (2017). Bicondon bias can determine the role of synonymous SNPs in human diseases. *BMC Genomics* doi:10.1186/s12864-017-3609-6
- 27 .Munroe, P. B., Olgunturk, R. O., Fryns, J. P., Van Maldergem, L., Ziereisen, F., Yuksel, B., et al. (1999). Mutations in the gene encoding the human matrix Gla protein cause Keutel syndrome. *Nature Genetics* doi:10.1038/5102
- 28 .Nicholson, P., Yepiskoposyan, H., Metze, S., Orozco, R. Z., Kleinschmidt, N., Mühlemann, O., et al. (2010). Nonsense-mediated mRNA decay in human cells: Mechanistic insights, functions beyond quality control and the double-life of NMD factors. *Cellular and Molecular Life Sciences* doi:10.1007/s00018-009-0177-1
- 29 .Nikolich-Zugich, J., Knox, K. S., Rios, C. T., Natt, B., Bhattacharya, D., and Fain, M. J. (2020). SARS-CoV-2 and COVID-19 in older adults: what we may expect regarding pathogenesis, immune responses, and outcomes. *GeroScience* doi:10.1007/s11357-020-00186-0
- 30 .Pan, A., Dutta, C., and Das, J. (1998). Codon usage in highly expressed genes of Haemophilus influenzae and Mycobacterium tuberculosis: Translational selection versus mutational bias. *Gene* doi:10.1016/S0378-1119(98)00257-1
- 31 .Pankov, R. and Yamada, K. M. (2002). Fibronectin at a glance. *Journal of Cell Science* doi:10.1242/jcs.00059
- 32 .Pérez-Cano, F. J., Franch, À., Castellote, C., and Castell, M. (2003). Immunomodulatory action of spermine and spermidine on NR8383 macrophage line in various culture conditions. *Cellular Immunology* doi:10.1016/j.cellimm.2003.09.009

- 33 .Piovesan, A., Vitale, L., Pelleri, M. C., and Strippoli, P. (2013). Universal tight correlation of codon bias and pool of RNA codons (codonome): The genome is optimized to allow any distribution of gene expression values in the transcriptome from bacteria to humans. *Genomics* doi:10.1016/j.ygeno.2013.02.009
- 34 .Plotkin, J. B. and Kudla, G. (2011). Synonymous but not the same: The causes and consequences of codon bias. *Nature Reviews Genetics* 12, 32–42. doi:10.1038/nrg2899
- 35 .Quax, T. E., Claassens, N. J., Söll, D., and van der Oost, J. (2015). Codon bias as a means to fine-tune gene expression. *Molecular Cell* doi:10.1016/j.molcel.2015.05.035
- 36 .Reinhardt, B., Winkler, M., Schaarschmidt, P., Pretsch, R., Zhou, S., Vaida, B., et al. (2006). Human cytomegalovirus-induced reduction of extracellular matrix proteins in vascular smooth muscle cell cultures: a pathomechanism in vasculopathies? *Journal of General Virology* 87, 2849–2858
- 37 .Schroder, K., Hertzog, P. J., Ravasi, T., and Hume, D. A. (2004). Interferon- $\gamma$ : an overview of signals, mechanisms and functions. *Journal of Leukocyte Biology* doi:10.1189/jlb.0603252
- 38 .Shackelton, L. A., Parrish, C. R., and Holmes, E. C. (2006). Evolutionary basis of codon usage and nucleotide composition bias in vertebrate DNA viruses. *Journal of Molecular Evolution* doi:10.1007/s00239-005-0221-1
- 39 .Shah, P. and Gilchrist, M. A. (2011). Explaining complex codon usage patterns with selection for translational efficiency, mutation bias, and genetic drift. *Proceedings of the National Academy of Sciences of the United States of America* doi:10.1073/pnas.1016719108
- 40 .Surjit, M., Liu, B., Jameel, S., Chow, V. T., and Lal, S. K. (2004). The SARS coronavirus nucleocapsid protein induces actin reorganization and apoptosis in COS-1 cells in the absence of growth factors. *Biochemical Journal* doi:10.1042/BJ20040984
- 41 .Tuller, T., Carmi, A., Vestsigian, K., Navon, S., Dorfan, Y., Zaborse, J., et al. (2010). An evolutionarily conserved mechanism for controlling the efficiency of protein translation. *Cell* doi:10.1016/j.cell.2010.03.031
- 42 .Tuller, T., Waldman, Y. Y., Kupiec, M., and Ruppin, E. (2010). Translation efficiency is determined by both codon bias and folding energy. *Proceedings of the National Academy of Sciences of the United States of America* doi:10.1073/pnas.0909910107
- 43 .Wang, Y. and Ni, H. (2016). Fibronectin maintains the balance between hemostasis and thrombosis. *Cellular and Molecular Life Sciences* doi:10.1007/s00018-016-2225-y
- 44 .White, M. P., Theodoris, C. V., Liu, L., Collins, W. J., Blue, K. W., Lee, J. H., et al. (2015). NOTCH1 regulates matrix gla protein and calcification gene networks in human valve endothelium. *Journal of Molecular and Cellular Cardiology* doi:10.1016/j.yjmcc.2015.04.006
- 45 .Xia, C., Vijayan, M., Pritzl, C. J., Fuchs, S. Y., McDermott, A. B., and Hahn, B. (2016). Hemagglutinin of influenza A virus antagonizes type I interferon (IFN) responses by inducing degradation of Type I IFN receptor 1. *Journal of Virology* doi:10.1128/jvi.02749-15
- 46 .Zhou, T., Weems, M., and Wilke, C. O. (2009). Translationally optimal codons associate with structurally sensitive sites in proteins. *Molecular Biology and Evolution* doi:10.1093/molbev/msp070