Abstract

Dietary protein intake influences the development of the virus infection cycle. This work aimed to elucidate the role of essential amino acids (EAAs) in the kinetics of SARS-CoV-2 virus replication. Studying the biochemistry of the metabolic interaction of the host and the pathogen should help to create an adequate model for control of the COVID-19 infection process.

Statistical samples were formed from epidemiological databases and systematic information. Amino acid sequences were obtained from protein databases (https://www.ncbi.nlm.nih.gov/ or https://www.uniprot.org/uniprot).

The analysis of statistical data on the spread of the COVID-19 pandemic carried out on large representative samples showed that the prevalence of SARS-CoV-2 infections or COVID-severity increased with excessive consumption of protein, sugar, and fat. A statistically significant correlation (r=0.83) was found between the productive pathogenicity of SARS-CoV-2 and the mass of animal protein consumed.

Consumed animal proteins contain more dietary EAAs (mean 44%) than plant-based ones (mean 35%). Many SARS-CoV-2 proteins contain very high amounts of EAAs (mean 50%) and need for their rapid translation high amounts of lysine, leucine, valine, and threonine (KLTV). The number of free EAAs can be a crucial metabolic factor in the development of COVID-19. Limiting KLTV can reduce the synthesis rate of the SARS-CoV-2 proteins and, subsequently, virions replication.

Many proteins involved in immune response also contain high levels of EAAs. How the synthesis of antibodies or acute-phase proteins competes with the need for EAAs during viral polypeptides elongation remains unclear.

EAAs are vital for both host/pathogen interaction and SARS-CoV-2 productive replication. Further research and detailed biochemistry of host/pathogen metabolic reactions should help to elucidate the deficit of the EAAs in the SARS-CoV-2 replication kinetics and in the development of the COVID-19 disease.

***Materials and methods**. The scientific data and information was found in publications and media available on the Internet, as well as taken from statistical databases for 5 continents and 48 countries. Statistics were acquired from the well known and reliable databanks: https://www.who.int; https://www.worldometers.info; https://ourworldindata.org; https://databank.worldbank.org; http://www.fao.org;*

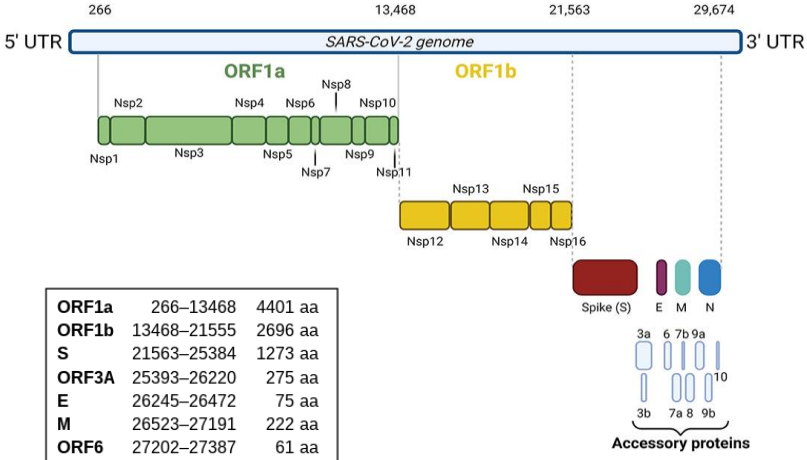
Rate of prevalence (RPr) or infection fatality rate (IFR) of the virus were calculated as a ratio between quantities of total cases and population. Amino acid sequences of proteins were obtained from databases (https://www.ncbi.nlm.nih.gov/, https://www.uniprot.org/uniprot/). The relationship between statistical data was estimated as a Pearson correlation coefficient (r).

Abbreviations: AAs – Amino Acids, EAAs – Essential Amino Acids, gRNA – genomic RNA, HCoV – Human Coronaviruses, IFR - Infection Fatality Rate, KLTV – Lysine+Leucine+Threonine+Valine, NSP - non-structural proteins, RPr - Rate of Prevalence, WHO - World Health Organization.

Introduction

Human Coronavirus SARS-CoV-2 (Fig. 1-2) was discovered over three years ago. It caused the infectious disease COVID-19. According to official statistics, the virus infected more than 765 Mio, and about 7 Mio patients died. The COVID-19 pandemic blocked almost all aspects of human activity. WHO announced the COVID-19 pandemic from 11.03.2020 till 03.05.2023. Epidemiologists, virologists, and clinicians warn of a real danger from new Coronavirus or mutants. Therefore, it is necessary to continue investigating all aspects of host-pathogen interactions during the development of this viral infection. It is still actually to find highly effective drug to prevent the replication cycle of the SARS-CoV-2 virus.

SARS-CoV-2 genome structure, translation and reproduction kinetics



ORF1a	266–13468	4401 aa
ORF1b	13468–21555	2696 aa
S	21563–25384	1273 aa
ORF3A	25393–26220	275 aa
E	26245–26472	75 aa
M	26523–27191	222 aa
ORF6	27202–27387	61 aa
ORF7a	27394–27759	121 aa
ORF8	27894–28259	121 aa
N	28274–29533	419 aa
ORF10	29558–29674	38 aa

Fig. 1. Virus SARS-CoV-2 genome organization
[Pizzato et al. FV2022]

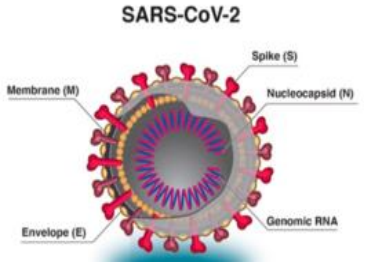


Fig. 2. Schema of the SARS-CoV-2 virus
[Ghaffari at all., 2020]

The SARS-CoV-2 virus belongs to the Betacoronavirus genus. Betacoronaviruses are the largest (about 30 Kbp) and most developed positive-sense RNA viruses (Fig. 1). Coronaviruses that cause disease in humans are referred to as HCoV. Here were the data of the null variant (Wuhan-Hu-1, Genbank MN 908947.3) used. The virus has a genome size of 29.9 thousand nucleotides. As an obligatory parasite, the virus uses an intracellular apparatus for the synthesis and post-translational modifications of all the necessary polypeptides and RNA molecules. The genome of Coronaviruses encodes the synthesis of four structural and 22-26 functional polypeptides (Fig. 1-3).

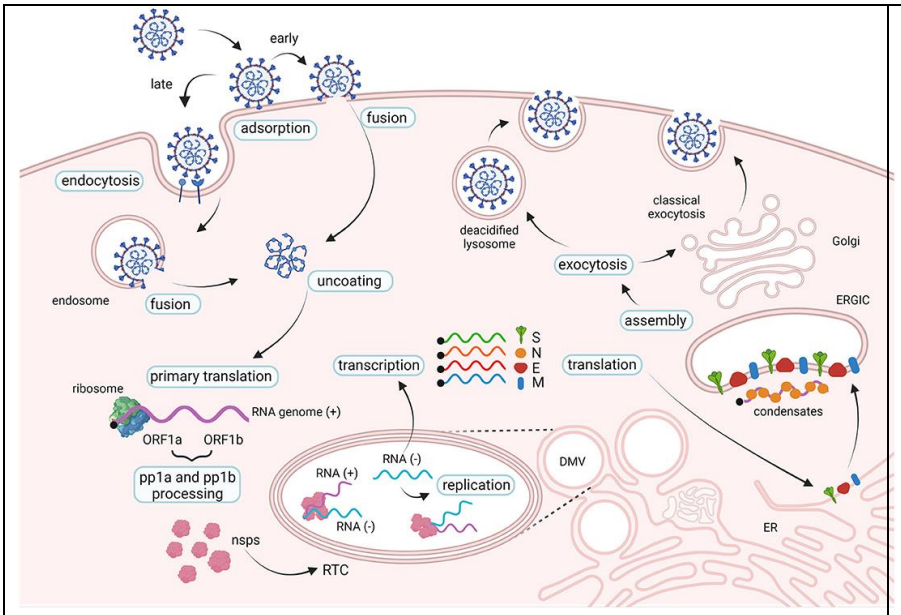


Fig. 3. Life cycle of SARS-CoV-2 virus [Pizzato et al. FV2022].

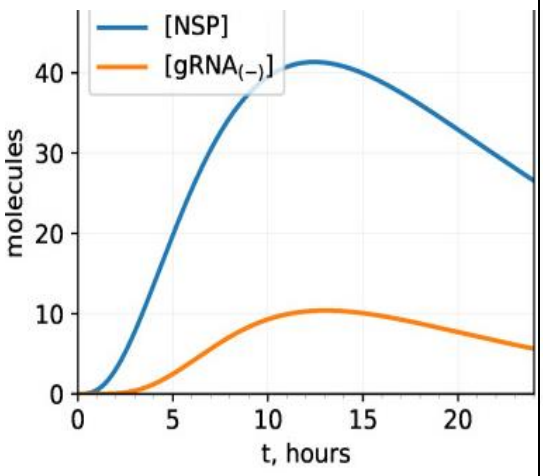


Fig. 4. Kinetics of –RNA transcription and translation of NSP [Grebennikov et al. 2021]

The open reading frame ORF1ab is located in the 5'-genomic region of the genomic RNA (gRNA) (Fig. 1). The ORF1ab makes up approximately two-thirds of the 30 Kbp viral genome. It encodes the production of two, primarily translated, polyproteins Pp1a and Pp1ab, processed in 16 non-structural proteins (NSPs). These NSPs are very essential for the virus replication cycle. The polyproteins are synthesized directly upon the virus particle invasion and uncoating of gRNA using it as a template (Fig. 3-4). The Pp1a polypeptide is 1.4–2.2 times faster than Pp1ab expressed. That is a co-translational result of a programmed -1 ribosome frameshift common to Coronaviruses.
As in other HCoV, four genes encoding structural proteins: spike (S), envelope (E), membrane (M), and nucleocapsid (N) are located in the 5' to 3' direction of the remaining third of the SARS-CoV-2 genome. These genes border on a variable number of 6-10 accessory proteins ORFs (Fig. 2-3).

Characteristics of the development of the COVID-19 pandemic

Epidemiological data of the infectious disease COVID-19 varied significantly across different continents and states (Fig. 5-6). The infection rate of prevalence (RPr) in the USA or Europe was a thousand times higher than in some countries in Asia or Africa (Fig. 1). In countries with low income, the infection fatality rate (IFR) was 100-200 times lower than the world average.
Comparing epidemic statistics in regions with high and low COVID rates should help assess risk factors for the pandemic. Due to such incomparable differences in statistical data, it was suggested in some publications that methods of recording and documenting confirmed cases may differ in several countries or that not all organizations provided real epidemiological facts. Large samples of statistical data are more representative.

The direct relationship between nutritional factors and the impact of the COVID-19 pandemic was found using available statistics for continents and 48 countries (Fig. 5-6).

Comparison of the amino acid composition

Consumed animal products contain more protein compared with plant ones (Fig. 7).
The HCoV proteins contain much more EAAs than consumed plants (Fig. 8-9). The polyprotein Pp1a contains more EAA than the structural S or N proteins. Edible plant proteins contain significantly less leucine, lysine, and especially threonine and valine than SARS-CoV-2 polypeptides (Fig. 10). Optimal synthesis of the SARS-CoV-2 virus Pp1a polyprotein requires a timely of saturated intracellular concentration of these four EAAs. EAAs are not synthesized *de novo* in the human body. Most of the EAAs enter all tissues of the body from food. With a predominantly plant-based diet (by vegans), a less amount of EAAs enters the bloodstream. Such a deficiency of four limiting EAAs can block the transition to the exponential phase of the rate of early viral protein synthesis, which leads to a slowdown in the reproduction of the pathogen or an abortive SARS-CoV-2 infection.

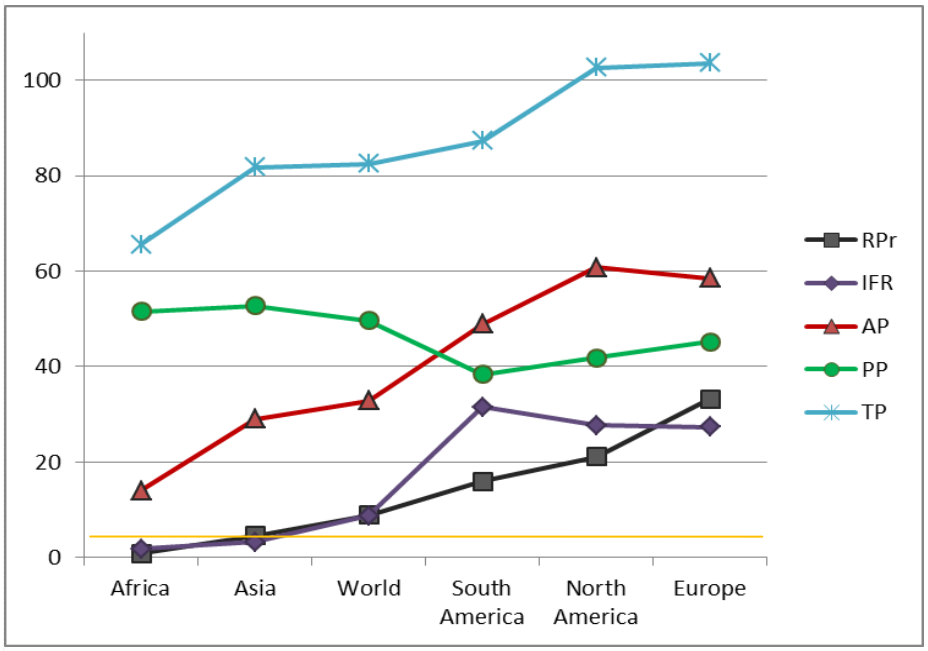


Fig. 5. Relation between outcomes of COVID-19 pandemic and protein intake
Correlation between RPr, IFR of the SARS-CoV-2 and amount of consumed: animal protein (AP); plant protein (PP) or total protein (TP). Along the y-axis are given the values of the factors, AP, PP, TP, RPr and IFR correspondingly: RPr=total amount infected/100 people, IFR=amount COVID-19 deaths/10K patients (30.05.2023). AP, PP, TP=g/day/person. Yellow line shows the epidemic threshold.

Table 1. Correlation coefficients (r) between COVID-19 outcomes and consumed proteins.
Calculation was done for data from Fig. 5.

	RPr	IFR
IFR	0,836	
AP	0,911	0,920
PP	-0,629	-0,951
TP	0,918	0,802

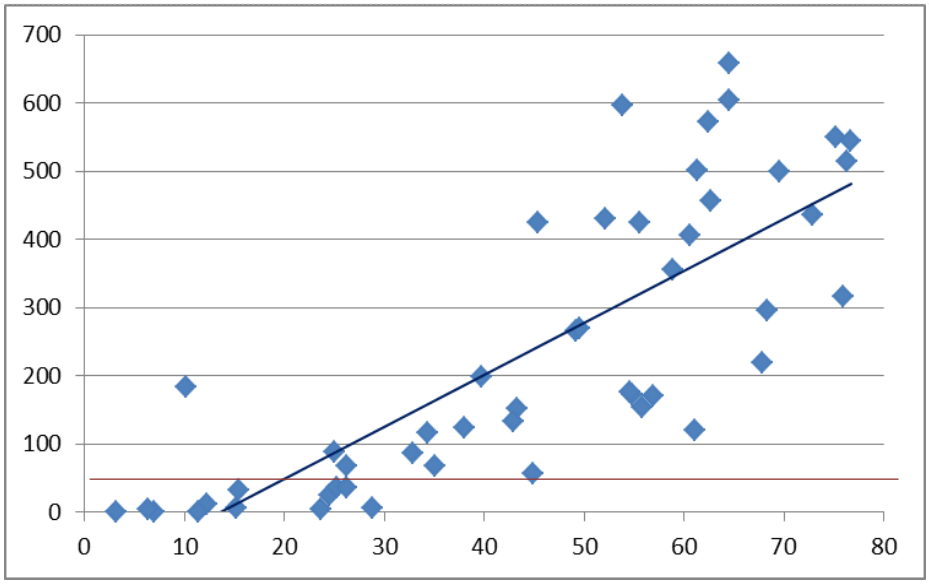


Table 2. Correlation coefficients (r) between COVID-19 outcomes and consumed AP in 50 countries

Factors	RPr	IFR
AP	0,830	0,519
PP	-0,212	-0,101

Fig. 6. Relation between RPr and animal protein intake

Correlation between RPr of the SARS-CoV-2 infection and amount of consumed animal protein (AP) in top 48 countries. Along the y-axis are given the values of the factor RPr. RPr=total amount infected/1000 people (30.05.2023), X-axis AP in g/day/person. Red line shows the epidemic threshold

The limited AAs in food can be a competitive metabolic factor that reduces or inhibits the rate of intracellular synthesis of nonstructural polyproteins of SARS-CoV-2 virions (1, 3-4). Deficiencies of EAAs, mainly free valine and threonine (Fig. 10), can suppress the early translation of SARS-CoV-2 virus proteins. A diet low in EAAs and KLTV may prevent rapid, highly productive viral replication, pathogenic development of COVID-19.

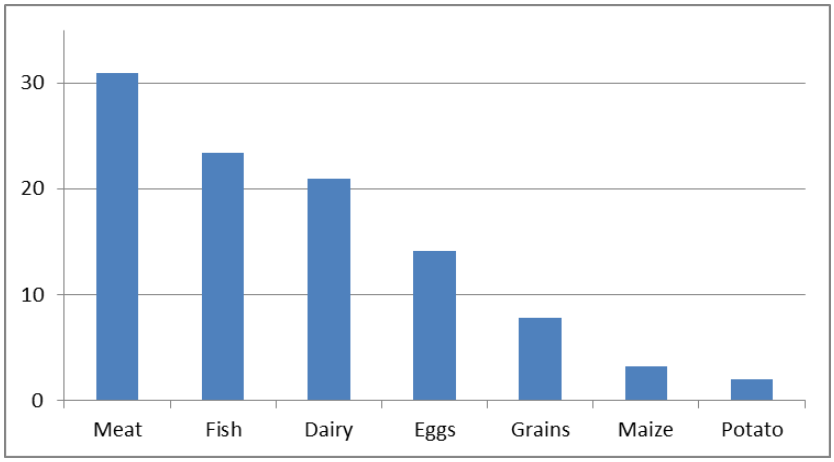


Fig. 7. Amount (g) of proteins in 100 g animal or plant staple food

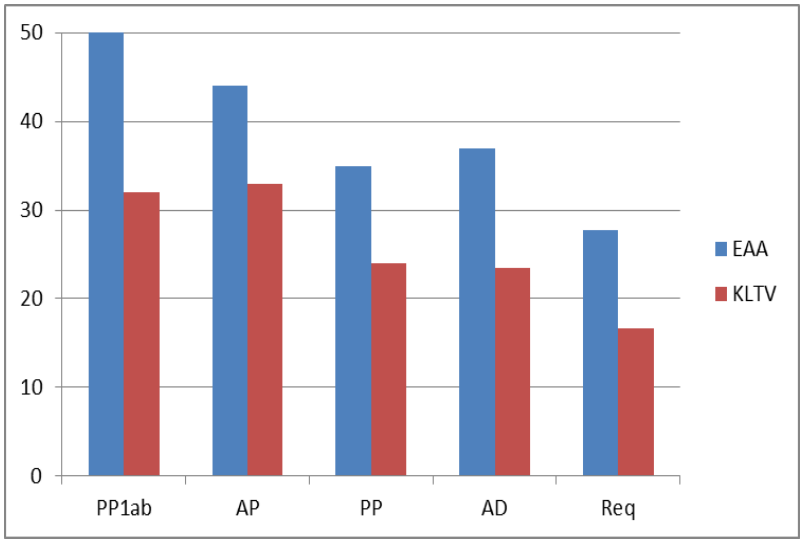


Fig. 8. Amount of EAAs or KLTV in proteins from HCoV and food
Red: KLTV, light blue: total EAAs. AD: standard American Diet [Paul et al., N2019]. AP: mean from meat, fish, milk and egg products and PP: mean from rice, wheat, potato and soya [Gardner et al., NR19]. Req: amount of EAAs required per day pro capita [WHO, FAO]. Pp1ab mean polyproteins from HCoV: SARS-CoV-2, SARS-CoV-1, MERS, HCoV-OC43 and HCoV-HRU1(UniProtKB). On the ordinate axis indicated amount of EAAs or KLTV as percentage of the total HCoV proteins, for food g/100 g.

The concentration of amino acids in the blood

Amino acids are absorbed from hydrolyzed dietary proteins in the intestine and transported into the bloodstream. The concentration of AAs in the blood plasma varies in both normal conditions and in various pathologies. Their blood concentrations may depend on many factors, but in tissues of a healthy body are successfully regulated. The content of free AAs in the bloodstream changes during the day and can significantly increase after a meal. Free amino acid homeostasis is maintained through the synthesis/degradation of polypeptides and AAs, protein proteolysis, AAs uptake from the gut into the bloodstream, and their utilization in various tissues. Fasting plasma concentrations of some EAAs are lower in vegans than in those who consume animal products. This difference is higher in the postprandial phase.

The SARS-CoV-2 virus modifies the consumption profile and distribution of proteins and AAs in the body and host cells. The pathogen alters central cellular pathways such as translation, splicing, carbon metabolism, proteostasis, and nucleic acid metabolism. The serum metabolome of COVID-19 patients is distinctive and works as an important sign in investigating pathogenesis, determining a diagnosis, predicting severe cases, and improving treatment.

The development of a viral infection leads to an imbalance in the metabolism of AAs in the body. The composition of AAs in patients' plasma can vary significantly depending on the severity of COVID-19.

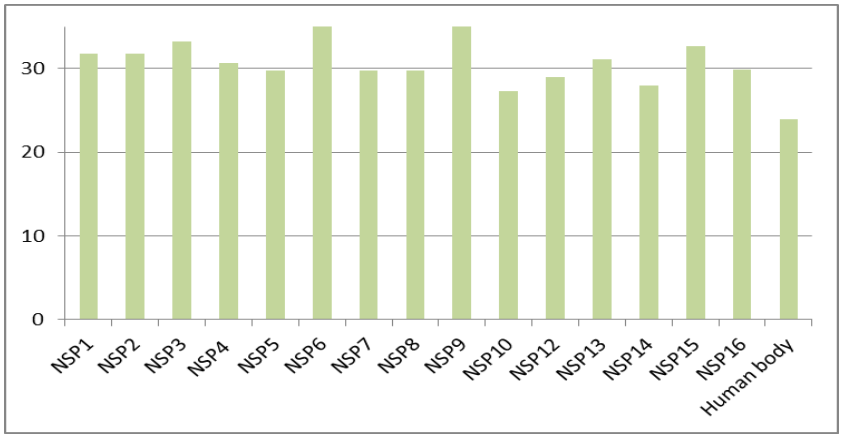


Fig. 9. Amount of KLTV in NSP from SARS-CoV-2

Ratio of KLTV as % from NSP molecules of the SARS-CoV-2 virus from (UniProtKB). Human body: ratio KLTV, % of total protein in human body [FAO, 2007]. On the ordinate axis indicated amount of KLTV as percentage of the total protein.

SARS-CoV-2 reproduction and translation kinetics

A saturated concentration of translation substrates should be sufficient for the optimal rate of protein synthesis. Virion replication passes through the lag phase into the exponential phase (Fig. 11-12). The doubling time of the SARS-CoV-2 virus is 6-10 hours. The polypeptides of the SARS-CoV-2 virus are elongated six amino acid residues per second.

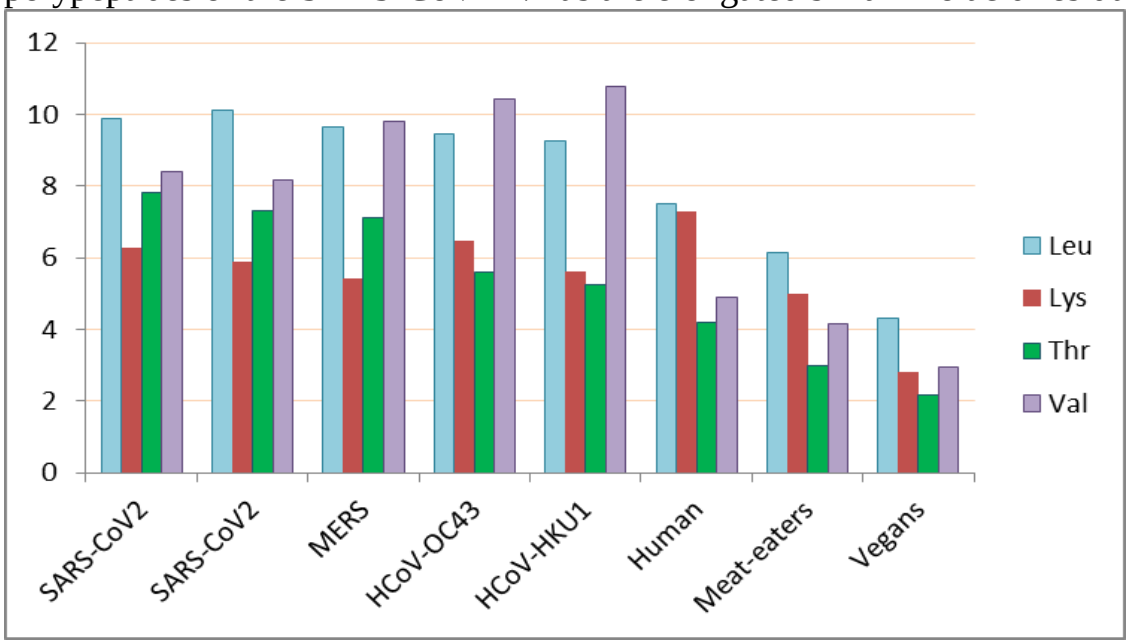


Fig. 10. Content of selected EAAs as percentage of total protein

Human: % of EAAs in proteins of whole body [FAO, 2007]. Meat-eaters and Vegans: composition (%) of EAA in ingested 100 g food proteins by meat-eaters or vegans [Schmidt et al., 2016]. For HCoV used sequence of polyprotein Pp1a from UniProtKB. On the ordinate axis indicated amount of AAs as percentage of the total protein [Gardner et al., 2019].

At the peak of pathogenesis, the total number of viral particles reaches in the body of an infected patient 10^9-10^{11} with a mass greater than 100 g. As a result of successful reproduction of the SARS-CoV-2 virus, lung weight increased, but the total body weight of animals or humans decreased.

Factors affecting the rate of translation of SARS-CoV-2 polypeptides

The deficiency of EAAs will lead to the interruption of the Pp1a translation, which is known to be synthesized as the first in HCoV. Inhibition of translation of functional Pp1a and Pp1ab polypeptides can consequently lead to a radical decrease the synthesis rate of Coronavirus gRNA, subgenomic RNAs and the polypeptides they encode. Such a decrease of synthesis rate of Pp1a molecules or abortive translation can prevent the development of the log phase of replication, reduce the pathogenic effect of the virus, and avoid cytokine storm in COVID-19 patients.

A high rate of viral protein synthesis is necessary for a rapid replication cycle and efficient production of a large population of progeny virions (Fig. 11-12). The amount of KLTV may be the immediate primary limiting factor in the rate of translation of polypeptides, and hence the replication of the SARS-CoV-2 virus. The high requirement of Pp1a for EAAs makes their replication very sensitive, especially to the shortage of KLTV. Such lack of limited EAAs can lead to inhibition of SARS-CoV-2 virus replication (Fig. 12) or even to unproductive infection.

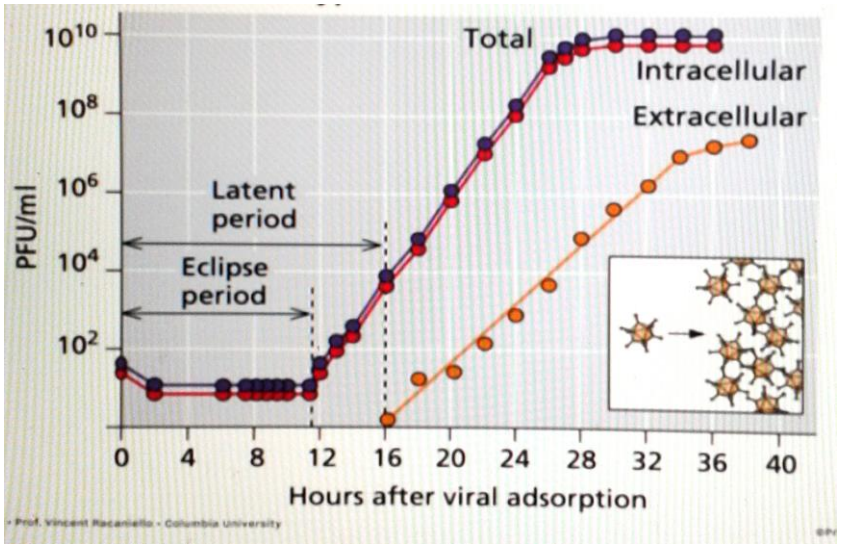


Fig. 11. Kinetics of viral replication (Racaniello, 2023)

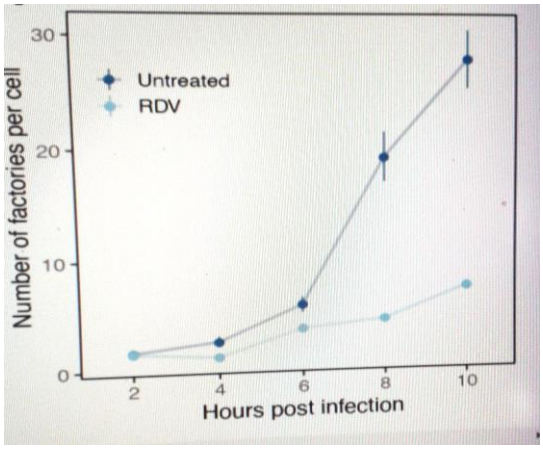


Fig. 12. Efficient and inhibited production of virions in the host cell (Lee at al. 2022)

Conclusions

The presented work shows that:

- The values of RPr and IFR in the pathogenesis of the SARS-CoV-2 virus differed significantly and could vary by two orders of magnitude in different regions.
- RPr and IFR depend on the type of diet. These epidemiological factors reflect the dynamics and severity of COVID-19.
- In populations consuming WHO-recommended or lower amounts of animal protein, RPr, and IFR were below the epidemic threshold, in contrast to countries with very high consumption of animal products. Animal proteins were a critical factor in the severity of the COVID-19 disease
- Diet animal proteins contain much more EAA than plant proteins. For the synthesis of most polypeptides of the SARS-CoV-2 virus, a large amount of free EAA is required. Especially important is the high concentration of KLTV for continuous translation of the long polyproteins Ppla and Pp1ab.
- Deficiency of EAA, especially KLTV, can reduce the rate of translation of SARS-CoV-2 virus polypeptides.
- A decrease in the rate of viral protein synthesis can slow down or block productive viral replication

The study of the biochemistry of the host/SARS-CoV-2 metabolic interaction helped our team to create a prototype antiviral drug.
Our team is going to prepare a patent application.

Author's related publications

- Ponomarenko S.V. Statistical Analysis of Critical Socioeconomic Factors in the Development of COVID-19 Disease. *Voprosy statistiki*. 2023; 30(1):90-100. https://doi.org/10.34023/2313-6383-2023-30-1-90-100
- Ponomarenko S.V. Dietary factors influencing the COVID-19 epidemic process. *FARMAKOEKONOMIKA. Modern Pharmacoeconomics and Pharmacoepidemiology*. 2022; 15 (4): 463–471. https://doi.org/10.17749/2070-4909/farmakoekonomika.2022.135.
- Ponomarenko S. V. Influence of Essential Amino Acids on the Synthesis of Polyproteins of the SARS-CoV-2 Virus in the COVID-19 Pathogenesis. *Health, Food & Biotechnology*. 2023; 5(1), P.19-34. https://doi.org/10.36107/hfb.2023.i1.s162