



Article

Unique tRNA Fragment Upregulation with SARS-CoV-2 but Not with SARS-CoV Infection

Isabella Imirowicz ^{1,†}, Azeem Saifee ^{1,†}, Leanne Henry ¹, Leo Tunkle ¹, Alexander Popescu ² , Philip Huang ¹ , Jibiana Jakpor ^{1,2}, Ava Barbano ^{1,2}, Rohit Goru ¹, Audrey Gunawan ¹ , Maria Sicilia ², Mori Ono ¹, Xiaoyong Bao ³ and Inhan Lee ^{1,*}

¹ Outreach Division, miRcore, Ann Arbor, MI 48104, USA

² miRcore Volunteer Program, miRcore, Ann Arbor, MI 48104, USA

³ Department of Pediatrics, The University of Texas Medical Branch, Galveston, TX 77555, USA; xibao@utmb.edu

* Correspondence: inhan@mircore.org

† These authors contributed equally to this work.

Abstract: Unlike other coronaviruses, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has rapidly infected the global population, with some suffering long-term effects. Thanks to extensive data on SARS-CoV-2 made available through global, multi-level collaborative research, investigators are getting closer to understanding the mechanisms of SARS-CoV-2 infection. Here, using publicly available total and small RNAseq data of Calu3 cell lines, we conducted a comparative analysis of the changes in tRNA fragments (tRFs; regulatory small noncoding RNAs) in the context of severe acute respiratory syndrome coronavirus (SARS-CoV) and SARS-CoV-2 infections. We found extensive upregulation of multiple tRFs in SARS-CoV-2 infection that was not present in SARS-CoV or other virus infections our group has studied. By comparing the total RNA changes in matching samples, we identified significant downregulation of TRDMT1 (tRNA methyltransferase), only in SARS-CoV-2 infection, a potential upstream event. We further found enriched neural functions among downregulated genes with SARS-CoV-2 infection. Interestingly, theoretically predicted targets of the upregulated tRFs without considering mRNA expression data are also enriched in neural functions such as axon guidance. Based on a combination of expression data and theoretical calculations, we propose potential targets for tRFs. For example, among the mRNAs downregulated with SARS-CoV-2 infection (but not with SARS-CoV infection), SEMA3C is a theoretically calculated target of multiple upregulated tRFs and a ligand of NRP1, a SARS-CoV-2 receptor. Our analysis suggests that tRFs contribute to distinct neurological features seen in SARS-CoV-2.



Citation: Imirowicz, I.; Saifee, A.; Henry, L.; Tunkle, L.; Popescu, A.; Huang, P.; Jakpor, J.; Barbano, A.; Goru, R.; Gunawan, A.; et al. Unique tRNA Fragment Upregulation with SARS-CoV-2 but Not with SARS-CoV Infection. *Int. J. Mol. Sci.* **2024**, *25*, 399. <https://doi.org/10.3390/ijms25010399>

Academic Editors: Hee-Jae Cha and József Tőzsér

Received: 6 October 2023

Revised: 18 December 2023

Accepted: 22 December 2023

Published: 28 December 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Coronaviruses have been responsible for several notable viral outbreaks in the 21st century, including the severe acute respiratory syndrome (SARS) epidemic between 2002 and 2004, caused by SARS-CoV (CoV), and the coronavirus disease 2019 (COVID-19) pandemic, caused by SARS-CoV-2 (CoV2), which has persisted over 3 years and is ongoing. These two viruses, which share 79% of their genome sequences [1,2], primarily affect the respiratory system and are transmitted via droplets and particles emitted from an infected person's nose or mouth. Spike proteins of both viruses bind the ACE2 receptor to enter human cells. Besides ACE2, the most prominent receptor, other receptors that mediate CoV2 infection of human cells include neuropilin-1 (NRP1), Eph receptors, transmembrane serine protease 2 (TMPRSS2), P2X7, and CD147 (BSG) [3,4].

Both CoV and CoV2 are neuroinvasive viruses able to directly enter brain tissue via olfactory bulb neurons, the blood-brain barrier, and other peripheral neural pathways.

Neurological symptoms associated with CoV infection include epilepsy, muscle weakness, and chronic symptoms such as persistent fatigue, diffuse myalgia, depression, and sleep cycle disruption [5]. CoV2 infection has been linked to acute neurological symptoms including stroke, encephalopathies, anosmia, and ageusia [6]. Direct viral infection of neurons or glial and endothelial cells can lead to neuronal cell death while systemic infection can induce inflammation, vascular damage, and autoimmune responses that damage neural tissues [7].

Though both viruses can enter neural cells, the nature and timing of COVID-19 and SARS symptoms differ. While the neurological symptoms of SARS are mostly observed in the late disease stage, COVID-19 presents unique neurological symptoms such as anosmia during early CoV2 infection, even preceding respiratory symptoms [6]. Thus, CoV2 probably has additional ways of effectively interacting with neural cells, including olfactory receptor cells. Among CoV2 receptors, NRP1 is abundantly expressed in the respiratory and olfactory epithelium and is upregulated with CoV2 infection [8]. Blocking interactions between the CoV2 spike protein and NRP1 reduces viral infection [9].

Meanwhile, though sequencing technology advances have helped identify many types of human non-coding RNAs, functional studies are still in their infancy with regard to viral infection. Among the small non-coding RNAs, tRNA-derived fragments (tRFs) have specific functions related to external stresses such as viral infection or toxins. For example, overexpressed 5'-ends of tRNA-Gln-CTG fragments (tRF5-Gln-CTG) enhance respiratory syncytial virus replication [10]. Moreover, our group reported that certain tRF5s are significantly increased in COVID-19 patient nasal samples [11].

In this study, we utilized publicly available RNAseq data of cell lines infected by either CoV or CoV2 to identify differences in their mechanisms of infection. We found tRF5 expressions to be distinctively upregulated in CoV2-infected cell lines and nervous system-related KEGG pathways to be enriched among downregulated genes specific to CoV2. When we computationally predicted potential targets of these upregulated tRF5s with CoV2 infection without considering mRNA expressions, our posited targets showed enriched neurological functions. Considering the theoretical targets and genes only downregulated with CoV2 infection, we propose an NRP1-CoV2 model which may account for early-onset neurological symptoms involving CoV2 infection with tRF upregulation. We thus propose measuring tRFs in long COVID patients to study long-term neurological symptoms.

2. Results

2.1. Uniquely Upregulated tRF5s with CoV2 Infection

To determine the molecular functions of CoV2 infection, we chose well-controlled Calu-3 (human epithelial cells from lung tissue) cell line data for our analysis. The GSE148729 dataset contains various samples of gene expression profile data, including human cell lines infected with CoV and CoV2 [12]. Since both small and total RNAseq data from the same cell lines were obtained for three different infection time points, we found these data useful in identifying regulatory small RNA functions in CoV2 infection. Though drawn from only two samples per condition, the time series data exhibits a clear trend. To reduce other variables, we used Calu-3 cell lines after 4 h of mock infection as a control with additional validation using 24 h of mock data only for the 24 h infection cases. We first analyzed small RNAseq data to understand expression changes after CoV2 infection. Similar to our previous report using patient nasal data [11], here in cell line experimental data we again found that tRFs, particularly tRF5s, were most distinctively upregulated with CoV2 among all small RNAs. Figure 1B shows a volcano plot focusing only on tRF5s after 24 h infection with CoV2, clearly demonstrating the upregulation of many tRF5s not seen with CoV infection (Figure 1A). Figure 1C,D shows the fold changes at different time points for tRF5s, with the top 10 most differentially expressed tRF5s (all upregulated and also found upregulated in patient nasal data) together with tRF5-Glu-CTC (11th, adjusted p -value $< 10^{-6}$, upregulated and most abundant) at the 24 h infection time point, these changes increasing as the infection progressed. Among these upregulated tRF5s with CoV2

infection, tRF5-Gln-CTG and tRF5-Glu-CTC were selected for qRT-PCR validation and confirmed as upregulated in nasal samples of COVID-19 patients [11].

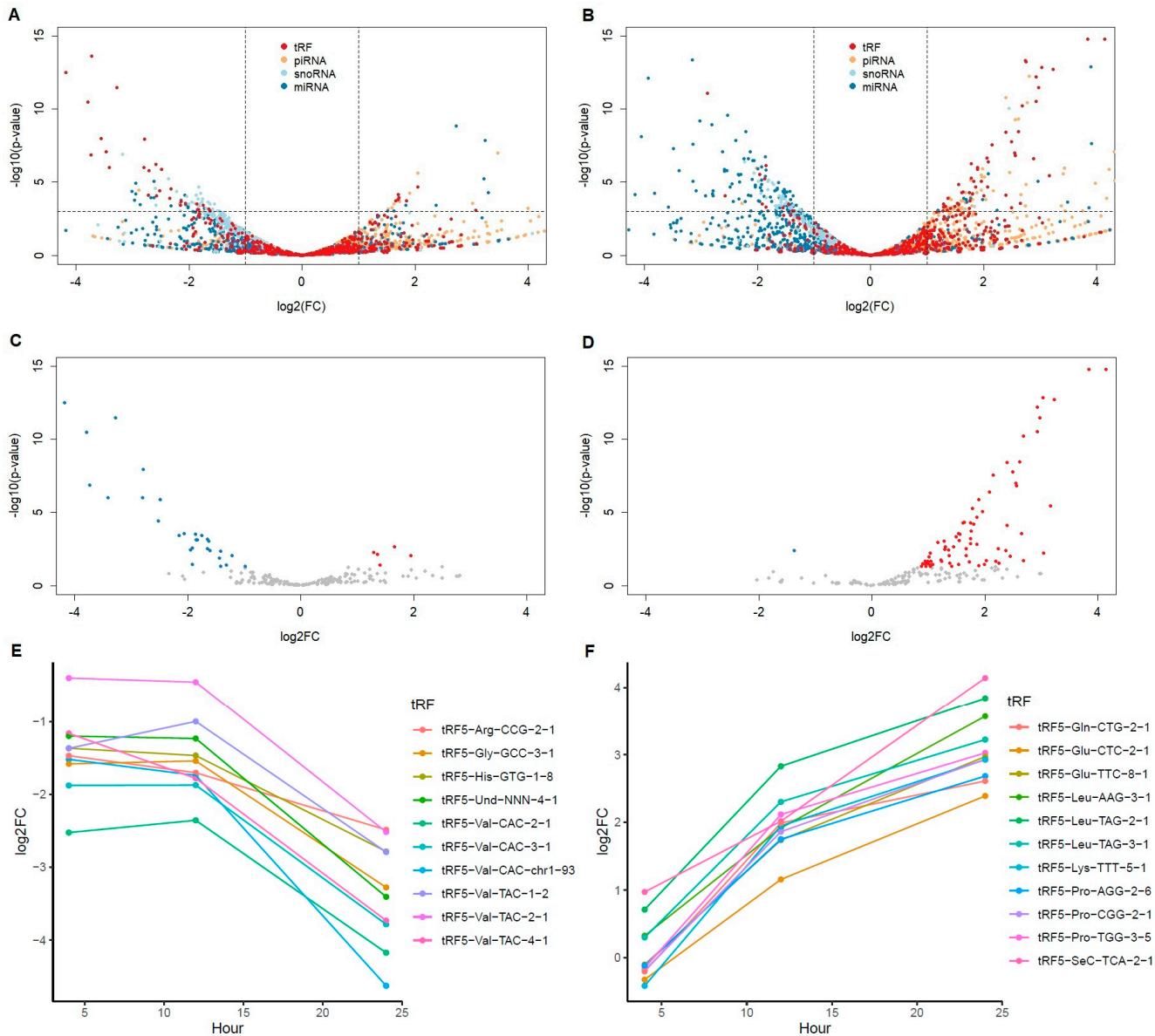


Figure 1. Differentially expressed tRF5s after SARS virus infection. (A,B) Volcano plots of all small RNAs with tRF (including all types of tRFs, red), piRNA (orange), snoRNA (light blue), miRNA (blue) at 24 h of SARS-CoV (A) and SARS-CoV-2 (B) infection in comparison to mock 4 h. Horizontal line is at $p = 0.001$. Vertical lines are at $\log_{2}FC = \pm 1$. (C,D) Volcano plots of only tRF5 species with read count more than 5 at any samples at 24 h of SARS-CoV (C) and SARS-CoV-2 (D) infection in comparison to mock 4 h. Colored dots are $p < 0.05$; red indicates upregulation ($\log_{2}FC > 1$) and blue downregulation ($\log_{2}FC < -1$). (E,F) Log₂FC changes in comparison to mock 4 h over time for the top 10 most significant tRF5s for SARS-CoV (E) and SARS-CoV-2 (F) infection at 24 h. SARS-CoV-2 has one more tRF5, Glu-CTC (11th rank), included due to its importance and abundance.

2.2. More Downregulated Genes with CoV2 Infection Than with CoV Infection

To investigate gene expression changes with CoV and CoV2 infection, we analyzed total RNAseq data and mapped them against gene-level expression values, combining all transcripts from the bowtie2 analysis. While more genes are upregulated than downregulated with both virus infections, CoV2 infection led to more genes being downregulated (Figure 2B) in comparison to CoV infection (Figure 2A). Such ratio differences (more down-

regulated genes in CoV2 than in CoV) decrease as the cutoff *p*-value increases, implying that the greater downregulation is specific to CoV2 infection. For example, with the adjusted *p*-value < 0.001 cutoff, 35.7% and 41.7% of differentially expressed genes were downregulated with CoV and CoV2 infection at 24 h, respectively (16.8% more downregulation in CoV2 infection). With the adjusted *p*-value < 0.05 cutoff, 45.2% and 48.6% of differentially expressed genes were downregulated with CoV and CoV2 infection at 24 h, respectively (7.5% more downregulation in CoV2 infection).

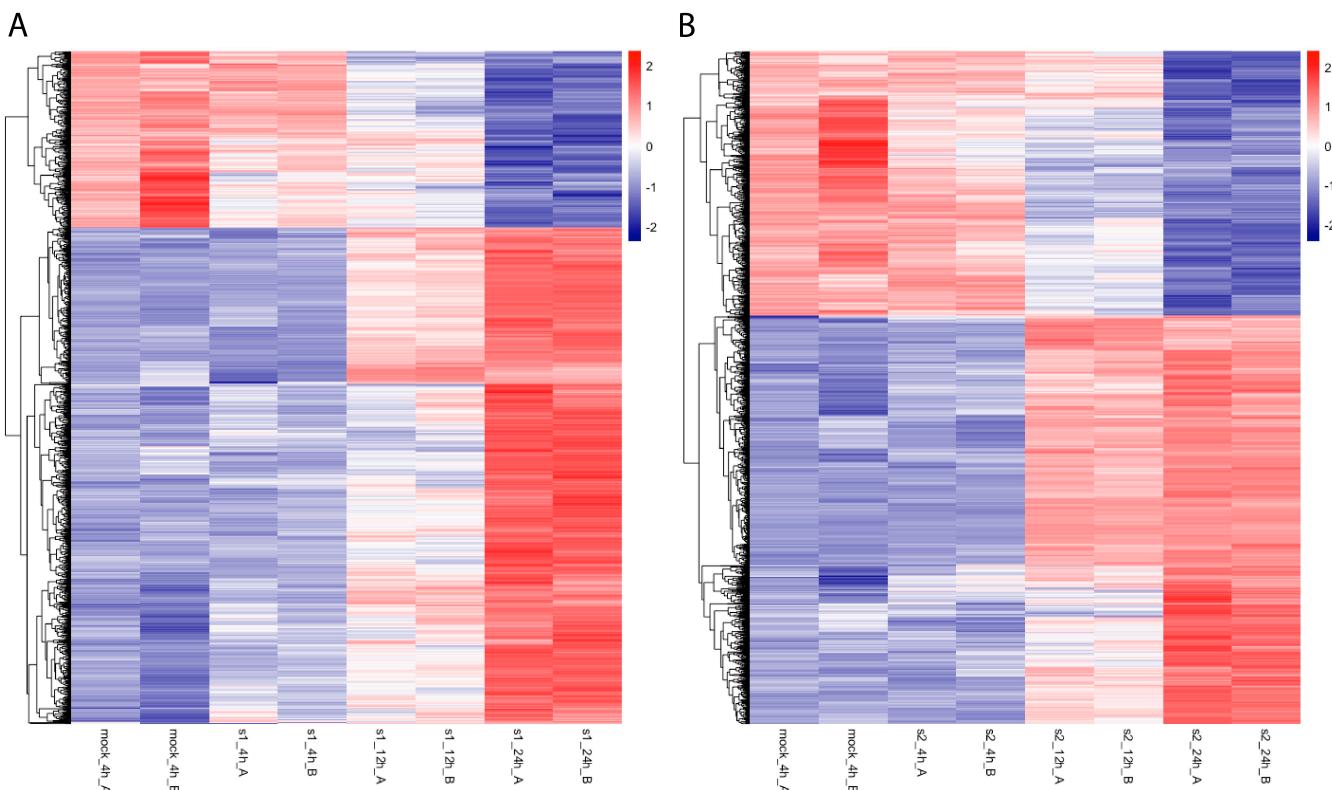


Figure 2. Heatmap of differentially expressed mRNAs (adjusted *p*-value < 10^{-5} at 24 h in comparison to mock 4 h) after SARS-CoV (A) and SARS-CoV-2 (B) infection. From left to right, mock 4 h and then infection time points at 4 h, 12 h, and 24 h of these gene expression levels are shown for each replicate A and B. After all SARS-CoV or SARS-CoV-2 infected samples were normalized using DESeq2, heatmaps were generated with normalized values of corresponding mRNAs using the R package heatmap with a scale-by-row option.

To shed light on the abundant tRF expressions with CoV2 infection, we compared expression patterns of genes potentially associated with tRF generation. Since tRNA is well known for its wide variety of modifications that respond to cellular environments [13] and tRNA stability is affected by methylation status, we checked tRNA methyltransferase and tRNA demethylase expression values for both virus infections. We also checked ANG, ELAC2, and DICER1, enzymes known to be associated with tRF biogenesis. For tRNA methyltransferases, we obtained the protein names from UniProt [14] and for tRNA demethylase, we checked *ALKBH1*, *ALKBH3*, and *FTO* gene expressions [15]. Among these, only differentially expressed genes with either CoV or CoV2 are shown in Table 1. Notably, while *DICER1* expressions were enhanced with both CoV and CoV2 infections, CoV2 infection increased *DICER1* expression 48% more than did CoV infection. Among all tRNA methyltransferases, *TRDMT1* (also known as *DNMT2*) was significantly downregulated only with CoV2 infection and not with CoV infection (Table 1).

Table 1. Differentially expressed genes among methyltransferases, demethyltransferases, and ribonucleases at 24 h after infections.

Gene Symbol	SARS-CoV Infection		SARS-CoV-2 Infection	
	Log2FC	p-Value	Log2FC	p-Value
tRNA Methyltransferase				
<i>FTSJ1</i>	−0.41	0.058	−0.70	2.3×10^{-4}
<i>METTL1</i>	−0.54	0.026	−0.17	0.53
<i>METTL8</i>	−0.53	0.052	−0.81	7.7×10^{-4}
<i>THUMPD3</i>	0.24	0.17	0.36	0.022
<i>TRDMT1</i>	−0.30	0.41	−1.43	3.2×10^{-5}
<i>TRMT2B</i>	−0.57	0.026	−0.37	0.13
<i>TRMT44</i>	0.59	0.036	0.67	0.60
<i>TRMT61A</i>	−0.56	0.0039	−0.15	0.49
<i>TRMT9B</i>	0.07	0.89	0.63	0.049
<i>TYW3</i>	−0.22	0.36	−0.42	0.033
Demethylase				
<i>ALKBH3</i>	−0.77	0.0022	−0.56	0.019
Ribonuclease¹				
<i>DICER1</i>	1.18	3.0×10^{-29}	1.75	2.8×10^{-69}

¹ We checked only ANG, ELAC2, and DICER1, known to be involved in tRF biogenesis.

To compare the functions of these downregulated genes using more comparable numbers to reduce statistical bias, we analyzed genes with an adjusted *p*-value < 0.05. Among the total 2292 and 2573 genes downregulated with CoV and CoV2 infection at 24 h, respectively, 1608 genes were commonly downregulated. Table 2 shows the enriched KEGG pathways for (1) commonly downregulated, (2) downregulated only in CoV infection, and (3) downregulated only in CoV2 infection at 24 h post-infection with adjusted *p* < 0.05. Commonly downregulated genes are involved in metabolism, neurodegenerative diseases, infectious diseases (bacterial), and translation processes. For CoV infection, proteasome (hsa03050) is the only pathway not found in the enriched pathways from commonly downregulated genes. On the other hand, CoV2 infection showed several specific enriched KEGG pathways, including the nervous system (hsa04723, hsa04728, hsa04730) and signal transduction (hsa04071, hsa04015), implying that additional molecular functions are disrupted with CoV2 infection.

Table 2. Enriched KEGG pathways for downregulated genes (adjusted *p*-value < 0.05) after 24 h infection with SARS-CoV and SARS-CoV-2 compared to 4 h mock infection.

KEGG ID	KEGG Description	Observed Gene Count	Background Gene Count	False Discovery Rate
Commonly downregulated genes				
hsa03010	Ribosome	83	130	9.43×10^{-36}
hsa01100	Metabolic pathways	209	1447	5.57×10^{-12}
hsa00190	Oxidative phosphorylation	45	130	1.83×10^{-11}
hsa04714	Thermogenesis	61	229	1.83×10^{-11}
hsa05012	Parkinson disease	57	240	3.47×10^{-9}
hsa05014	Amyotrophic lateral sclerosis	68	352	9.68×10^{-8}

Table 2. Cont.

KEGG ID	KEGG Description	Observed Gene Count	Background Gene Count	False Discovery Rate
hsa05010	Alzheimer disease	68	355	1.12×10^{-7}
hsa05016	Huntington disease	60	298	1.75×10^{-7}
hsa05020	Prion disease	52	265	3.40×10^{-6}
hsa04142	Lysosome	29	126	0.00015
hsa04510	Focal adhesion	36	198	0.0011
hsa01212	Fatty acid metabolism	16	54	0.0015
hsa04932	Non-alcoholic fatty liver disease	29	148	0.0016
hsa03060	Protein export	10	23	0.0026
hsa04260	Cardiac muscle contraction	20	87	0.003
hsa00062	Fatty acid elongation	10	25	0.0039
hsa05110	Vibrio cholerae infection	13	48	0.0105
hsa05100	Bacterial invasion of epithelial cells	16	70	0.012
hsa01200	Carbon metabolism	22	117	0.013
hsa04810	Regulation of actin cytoskeleton	32	209	0.0219
hsa03008	Ribosome biogenesis in eukaryotes	16	77	0.0246
hsa05205	Proteoglycans in cancer	30	196	0.0279
hsa04146	Peroxisome	16	79	0.0282
hsa00010	Glycolysis/Gluconeogenesis	14	65	0.031
hsa04910	Insulin signaling pathway	22	133	0.0415
Downregulated only in SARS-CoV				
hsa03010	Ribosome	19	130	0.00021
hsa05012	Parkinson disease	26	240	0.00021
hsa05016	Huntington disease	30	298	0.00021
hsa05020	Prion disease	28	265	0.00021
hsa05014	Amyotrophic lateral sclerosis	30	352	0.0011
hsa01100	Metabolic pathways	81	1447	0.0015
hsa05010	Alzheimer disease	28	355	0.0051
hsa03050	Proteasome	8	43	0.0116
hsa00190	Oxidative phosphorylation	14	130	0.0134
hsa04932	Non-alcoholic fatty liver disease	15	148	0.0137
Downregulated only in SARS-CoV-2				
hsa01100	Metabolic pathways	149	1447	4.78×10^{-14}
hsa00280	Valine, leucine, and isoleucine degradation	12	46	0.0021
hsa05132	Salmonella infection	27	209	0.0021
hsa00100	Steroid biosynthesis	8	20	0.0025
hsa00640	Propanoate metabolism	9	34	0.0094
hsa01200	Carbon metabolism	17	117	0.0099
hsa04723	Retrograde endocannabinoid signaling	19	145	0.012
hsa05031	Amphetamine addiction	12	66	0.012

Table 2. Cont.

KEGG ID	KEGG Description	Observed Gene Count	Background Gene Count	False Discovery Rate
hsa05110	Vibrio cholerae infection	10	48	0.0121
hsa04071	Sphingolipid signaling pathway	16	116	0.0153
hsa00020	Citrate cycle (TCA cycle)	7	29	0.036
hsa00190	Oxidative phosphorylation	16	130	0.036
hsa00340	Histidine metabolism	6	21	0.036
hsa00620	Pyruvate metabolism	8	38	0.036
hsa04728	Dopaminergic synapse	16	128	0.036
hsa04730	Long-term depression	10	59	0.036
hsa00564	Glycerophospholipid metabolism	13	97	0.0375
hsa04015	Rap1 signaling pathway	21	202	0.0375
hsa04962	Vasopressin-regulated water reabsorption	8	44	0.0478

2.3. Enriched Neural Functions of Theoretical Targets of Upregulated tRF5s

Given that many tRFs were uniquely upregulated and the ratio of the downregulated mRNAs increased with SARS-CoV-2 infection compared to SARS-CoV infection, we utilized tRF and mRNA expressions to identify tRF5 potential targets. We first performed theoretical hybridization calculations of uniquely upregulated tRF5s with entire untranslated regions (UTRs) of all coding transcripts.

Since most tRFs are longer than 30 nt, many mRNA UTRs have potential interaction sites for tRFs, making genome-wide analysis practically impossible. Furthermore, unlike the case for miRNA target prediction, there are minimal studies available for tRF target prediction. While tRF target research is limited, we here implemented two features in target prediction to reduce the number of potential theoretical interactions. First, we had previously identified that the 3'-end (not the 5'-end) of tRF5-Glu-CTC is functional in translational repression through targeting 3' UTR [16]. Since we found that tRF5-Glu-CTC was also upregulated with SARS-CoV-2 infection, we used this feature as the first filter among all potential interactions between tRF5 and 3' UTRs: 3' UTR interacting with the 3'-ends of tRF5s. Second, among these tRF5-target pairs, we further filtered targets having simultaneous 5' UTR interacting with the 5'-ends of tRF5s. In our earlier study, we observed a stronger translation repression when there were simultaneous interaction sites of miRNA with both the 5' and 3' UTR [17]. Since these sites are a subset of the tRF5s-3' UTR interaction pairs, our theoretically predicted targets may have certain biology functions enriched, which would be meaningful in itself.

We then investigated the enriched functions of these theoretically predicted targets. To reduce redundant target gene lists from similar sequence tRF5s, we chose representative tRF5s as they were the most abundantly expressed among tRNAs translating the same amino acids. All enriched functions identified by String-db of the theoretical targets of these 6 representative tRF5s are shown in Table S1. All 6 tRF5 targets showed enriched gene ontology (GO) biological processes and 51 GO terms were found to be enriched in more than three tRF5 targets. Among these terms, we decided to focus on *neuro* or *nerv* terms, since others are general terms like metabolic, localization, signal, development, modification, and communication. Table 3 shows the enriched GO biological processes containing *neuro* or *nerv* terms among the enriched functions in Table S1. Targets of four tRF5s (tRF5-Glu-CTC-2-1, tRF5-Gln-CTG-2-1, tRF5-Leu-AAG-3-1, and tRF5-SeC-TCA-2-1) had enriched neural GO terms. To understand the significance of enriched GO biological processes, we randomly chose 1000 genes (average target number of the 6 tRFs = 995) among expressed mRNAs (total number 20,924) in Calu-3 cell lines and checked their

enriched functions. Among the 10 random 1000-gene sets, five sets showed GO biology process terms enriched but none had neural functions.

Table 3. Enriched GO biological process terms including *neuro* or *nerv* for tRF5-predicted target genes.

Term ID	Term Description	tRF5 Name
GO:0001764	Neuron migration	tRF5-Glu-CTC-2-1
GO:0007399	Nervous system development	tRF5-Gln-CTG-2-1 tRF5-Glu-CTC-2-1 tRF5-Leu-AAG-3-1
GO:0007417	Central nervous system development	tRF5-Glu-CTC-2-1 tRF5-Leu-AAG-3-1
GO:0010975	Regulation of neuron projection development	tRF5-SeC-TCA-2-1
GO:0010976	Positive regulation of neuron projection development	tRF5-Glu-CTC-2-1
GO:0022008	Neurogenesis	tRF5-Gln-CTG-2-1 tRF5-Glu-CTC-2-1
GO:0030182	Neuron differentiation	tRF5-Gln-CTG-2-1 tRF5-Glu-CTC-2-1 tRF5-Leu-AAG-3-1
GO:0031175	Neuron projection development	tRF5-Leu-AAG-3-1
GO:0045664	Regulation of neuron differentiation	tRF5-Glu-CTC-2-1 tRF5-SeC-TCA-2-1
GO:0045665	Negative regulation of neuron differentiation	tRF5-Glu-CTC-2-1
GO:0045666	Positive regulation of neuron differentiation	tRF5-Glu-CTC-2-1
GO:0048666	Neuron development	tRF5-Gln-CTG-2-1 tRF5-Glu-CTC-2-1 tRF5-Leu-AAG-3-1
GO:0048699	Generation of neurons	tRF5-Gln-CTG-2-1 tRF5-Glu-CTC-2-1 tRF5-Leu-AAG-3-1
GO:0048812	Neuron projection morphogenesis	tRF5-Leu-AAG-3-1
GO:0050767	Regulation of neurogenesis	tRF5-Glu-CTC-2-1 tRF5-Leu-AAG-3-1
GO:0050769	Positive regulation of neurogenesis	tRF5-Glu-CTC-2-1
GO:0051960	Regulation of nervous system development	tRF5-Gln-CTG-2-1 tRF5-Glu-CTC-2-1 tRF5-Leu-AAG-3-1
GO:0051962	Positive regulation of nervous system development	tRF5-Glu-CTC-2-1

Interestingly, the remaining two tRF targets are enriched in GO biological terms related to signal transduction. Among other enriched biological functions, targets of tRF5-Lys-TTT-5-1 are enriched in the GO component Synaptic vesicle (GO:0008021), and those of tRF5-Pro-TGG-3-5 are enriched in the KEGG Neurotrophin signaling pathway (hsa04772). These theoretical targets are general ones not confined to specific cell types.

We then determined which predicted tRF5 targets (having theoretical interaction sites) were also downregulated in Calu-3 cell lines with CoV2 infection but not with CoV infection. Among the target genes involved in Table 3 GO terms (thus focused on neural functions), the predicted targets from two or more of these 4 tRF5s which were also identified as downregulated only in CoV2 infection (adjusted $p < 0.05$) were ABAT, BCR, DHFR, DPYSL3, EEF2K, GLUD1, GRIN2A, HMGB1, MAPT, NDST1, NFIA, RIMS1, SEMA3C, STK25, and

TACC1. While all other proteins are located mostly in the cell membrane or internal compartment, SEMA3C is mostly located in the extracellular compartment and is one of the ligands for NRP1 [18], which CoV2 uses as a receptor [8]. SEMA3C was predicted as a target of tRF5-Leu-AAG-3-1 and tRF5-SeC-TCA-2-1.

3. Discussion

Our cell line data analysis shows unique upregulation of the tRF5 family with CoV2 infection, not seen in other viral infection studies. Consistent with our nasal sample data from COVID-19 patients, tRF5-Gln-CTG and tRF5-Glu-CTC were also identified as upregulated in lung epithelial cell lines (Calu-3). Since cell line data have fewer variables, we identified additional upregulated tRF5s at 24 h after CoV2 infection. Most of these tRF5s only increased with CoV2 infection while additional mRNAs were downregulated in CoV2 infection compared with CoV infection, suggesting tRF5 functions merit close scrutiny with regard to COVID-19 research.

We greatly appreciate this kind of clearly defined paired sequencing data being publicly available so that citizen scientists can do small RNA research to uncover additional features with validation at each step. The unexpectedly highly upregulated tRF5s with CoV2 infection aligned well with expression value changes of genes currently known to be involved in tRF biogenesis. Only CoV2 significantly reduced expressions of *TRDMT1*, one of the well-known tRNA methyltransferases blocking tRF generation, while CoV2 increased expressions of *DICER1* more than did CoV, possibly helping tRF generation. Having confirmed the close alignment of mRNA data with tRF biogenesis, we further investigated mRNA expressions with regard to tRF functions to derive useful guidelines in future molecular functional studies of tRFs, as well as to provide insight into COVID-19 symptoms.

Since our study focus is tRF5s, we investigated downregulated mRNAs in more detail. General trends in mRNA changes due to both SARS virus infections were similar—cellular responses that help viral replication. Enriched KEGG pathways for commonly downregulated mRNAs in both CoV and CoV2 included multiple metabolism-related pathways. They also included functions to reduce various defenses such as lysosomes, peroxisomes, and thermogenesis. Bacterial invasion of epithelial cells GO-term was also enriched among these commonly downregulated genes, aligning well with the reduced co-infection among COVID-19 patients compared with those with influenza [19], presumably due to reduced competition with other pathogens during early infection. Enriched neurodegenerative disease pathways among commonly downregulated genes may be related to keeping cells alive for viral replication. CoV2 was reported to continuously replicate in Calu-3 cells even 72 h post-infection [20] and our analyzed data points were up to 24 h.

The enriched KEGG pathways of downregulated genes from CoV infection alone all fell under these same categories. However, enriched KEGG pathways of downregulated genes solely with CoV2 infection departed from these common features in ways such as including the nervous system and signal transduction, mostly involving neural cells. These changes were seen as early as 24 h post-infection of lung epithelial cells. While both CoV and CoV2 infections clearly affect neurological symptoms and we observed upregulated *NRP1* in both CoV and CoV2 infections, the key difference between them is the vast number of upregulated tRF5s. We propose that these upregulated tRF5s promote further enhancement of CoV2 interactions with neural cells.

In our earlier study, we detailed the interactions of tRF5-Glu-CTC with target gene *APOER2* and identified the 3'-end of tRF5-Glu-CTC as responsible for protein reduction through the *APOER2* 3' UTR. Because only a few tRF5-target studies are available, we decided to expand tRF5-Glu-CTC interaction patterns with *APOER2* to all other tRF5 target predictions in this study. Additionally, we filtered the predicted targets to include additional interaction with the 5' UTR of the same gene. Therefore, our predicted targets may miss other potential interactions. However, given the excess of possible targets for a given tRF5, we believe that our filters serve as a good starting point for tRF5 target analysis.

To our surprise, many target genes of upregulated tRF5 showed enriched functions related to the nervous system, features that were seen in downregulated mRNAs only with CoV2 infection. Since NRP1 has multiple functions in the immune system [21], *NRP1* expression may increase to enhance CoV2 replication. In addition, given our finding that the upregulation happens at an early infection time point, other mechanisms such as compensating for reduced neural functions may also be at play. When we focused on smaller groups of genes with neural functions, which are potential targets of multiple upregulated tRF5s and downregulated in infection with CoV2 but not with CoV, we were intrigued by SEMA3C. When we relaxed the distance criteria described in Methods from 19 nt to 16 nt, tRF5-Pro-AGG-2-6 also could target SEMA3C, which is expressed in most tissue types.

We therefore suggest a model involving tRF5-SEMA3C-NRP1 with CoV2 infection centered on neural functions. In Figure 3, the left panel represents the very initial CoV2 interaction, wherein most NRP1 receptors are occupied by other ligands such as SEMA3C (green squares), reducing the ability of CoV2 to bind to NRP1. The right panel shows CoV2 infection downregulating SEMA3C expression. With less SEMA3C, more NRP1 receptors are available for CoV2 binding, thus increasing the infection of other NRP1-presenting cells, including respiratory and olfactory epithelium. This is one example of highly upregulated tRFs possibly contributing to additional neurological symptoms not seen in other viral infections.

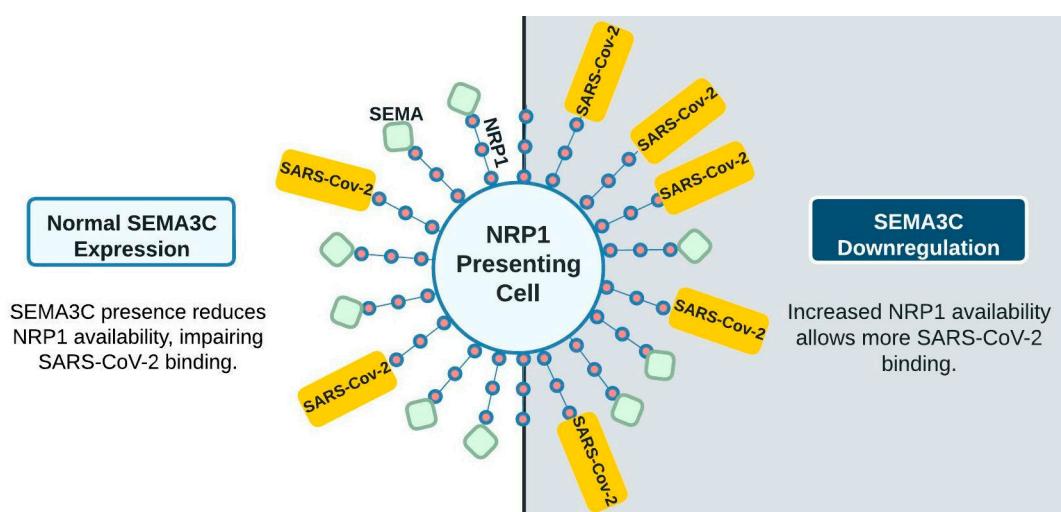


Figure 3. Proposed model of SARS-CoV-2 and SEMA competition for NRP1 interaction. With fewer ligands, more receptors are available for SARS-CoV-2 binding. Upregulated tRF5s are not shown in this diagram for clarity.

Some who recover from COVID-19 can experience long-term health effects (long COVID). Long COVID includes a variety of ongoing symptoms, lasting months or years and affecting different organs depending on the individual, and is thus not considered one illness. There are clearly multiple mechanisms leading to long COVID, including mitochondrial depletion [22]. Our findings point to a specific mechanism underlying neurological symptoms associated with long COVID. Recently, disability-adjusted life years (DALYs) were calculated based on 140,000 people who had COVID-19 compared with 6 million non-infected controls at the 2-year mark post-COVID. Neurologic DALYs per 1000 persons were found to be the most abundant, followed by cardiovascular cases [23]. While there is no current cure for long COVID, it would be interesting to see whether the tRF expressions are persistently enhanced in patients who have recovered from COVID-19.

Based on our paired data analysis, we suggest a model based on tRF5 upregulation specific to CoV2 infection and enriched neural functions both in mRNAs specifically downregulated in CoV2 and in theoretically predicted target mRNAs of the upregulated

tRF5s. We have previously identified specific tRF5s upregulated in Alzheimer's disease [24]. It seems that tRF5 functions are related to both infection and neurodegenerative diseases and may provide further insight into unique neurological symptoms such as anosmia with COVID-19.

4. Materials and Methods

4.1. Downloading Fastq Files from a GEO Dataset

We used gene expression omnibus (GEO) dataset GSE148729, which includes data from bulk and single-cell polyA RNA sequencing, small RNA sequencing, and total RNA sequencing of different human cell lines infected with CoV or CoV2 viruses. We selected Calu-3 (human epithelial cells from lung tissue) cell line samples of small RNA (GSM4477932 through GSM4477949) and total RNA sequencing data (GSM4477950 through GSM4477967). These samples included two replicates per condition at 4, 12, and 24 h, with appropriate controls at the 4 and 24 h time points. We used the fastq-dump command in the SRA toolkit [25,26] to download fastq files.

4.2. Aligning Sequencing to Transcripts

We used cutadapt [27] to trim 3'-adaptor sequences and discarded read sequences of size less than 15. The sequences trimmed from the small RNAseq data were AAAAAAAA and TGGAATTCTCGGGTGCCAAGG, and the sequences trimmed from the total RNAseq data were AGATCGGAAGAGCACACGTCTGAACCTCCAGTCA for pass 1 and AGATCGGAA-GAGCGTCGTAGGGAAAGAGTGT for pass 2, as referenced from Illumina. We used bowtie2 [28] to align reads to reference sequences, allowing for one nucleotide mismatch for small RNAseq analysis. The reference for small RNAseq was built from combined small RNA sequences of (1) downloaded sequences of sno/miRNA tract using the UCSC Table Browser [29] and (2) prepared tRFs from 33 unique nucleotides of 5'-end of tRNAs (tRF5) and those of 3'-end (tRF3) as well as upstream pre-tRNA sequences (tRF1) of mature tRNAs. To build tRF references in more detail, we downloaded all tRNA identified from UCSC genome browsers and used the tRNA names as their root names. 33 nt is 1-4 nucleotides longer than most tRF5s and, based on years of observation, enough to effectively identify the tRFs using bowtie2 software (version 2.4.1). 53.5% of tRF5s were identified as redundant sequences and were thus removed, keeping only representative ones. For tRF3 references, we did not include ACC, the mature tRNA-added sequence, at the 3'-end. To build the reference for total RNA mapping, we downloaded NCBI RefSeq transcripts using the UCSC Table Browser. The mapped sequences were then processed using SAMtools [30] commands view, sort, index, and idxstats to obtain read counts for each transcript. Figure 4 shows the schematics of the process.

4.3. Differential Expression Analysis

We used DESeq2 [31] to identify differentially expressed small RNAs and mRNAs. Expression data at time points of 4, 12, and 24 h after infection were compared with the expression data of mock 4 h samples to keep the same reference point due to missing 12 h mock data. We separately used 24 h mock as a reference for the 24 h infection samples and compared the results to confirm that the results using the 4 h mock were valid (the top 12 most significantly upregulated tRF5s were the same except for one per each analysis; all were top-ranking upregulated tRF5s). Normalized expression values were also obtained using the DESeq2 process to generate heatmaps.

4.4. Involving Student Scientists

We invited high school students who had performed transcriptome analysis for over a year and reached conclusions on their research to join weekly RNAseq data analysis sessions with undergraduate mentors. After the students understood the research project, they divided the samples and performed specific steps of RNAseq data analysis each week, reporting their progress in a shared document. To ensure the quality of data analysis,

students compared their results with those of counterparts during the weekly sessions after independently analyzing data. The results from the students' individual analyses were reproduced using shell scripts for validation. All the confirmed results were used here and those who had processed data following all steps were invited to finalize this research as authors.

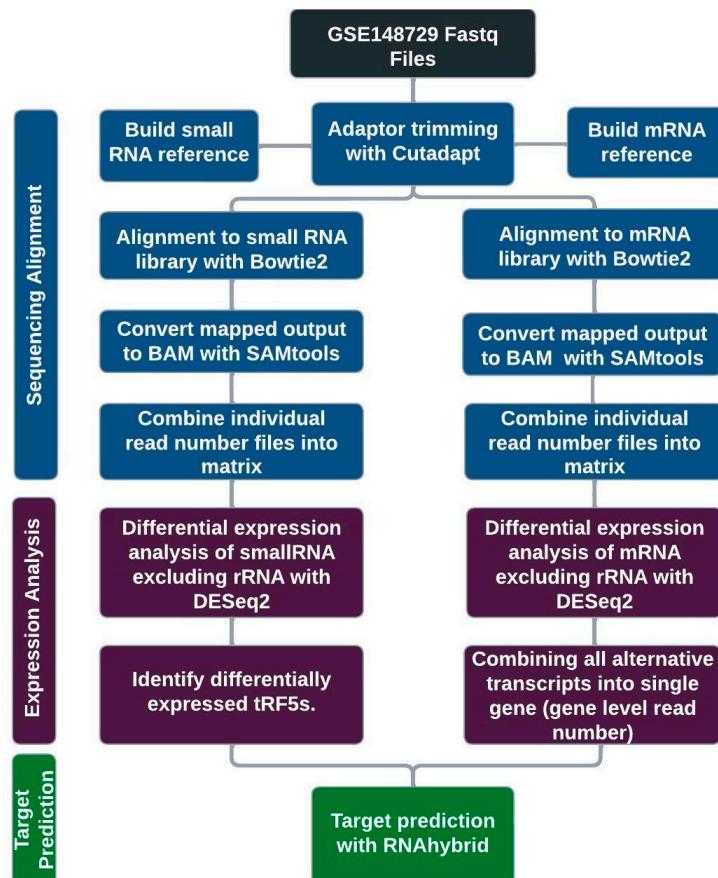


Figure 4. Schematics of sequencing data analysis. After sequence alignment, mRNA transcript read numbers were combined for gene-level analysis. Target predictions for upregulated tRF5s were calculated independently and then compared with their gene level expression changes.

4.5. tRNA Methyltransferase Collection

We used UniProt [14] to identify tRNA methyltransferase. To specify the search terms, we used advanced search with function -> Enzyme classification [EC] value of Methyltransferases [2.1.1.-] and Keyword [KW] of tRNA* for human, yielding these 25 proteins: C14orf172, FTSJ1, H_YH95C04.1, LCMT2, METTL1, METTL2A, METTL2B, METTL6, METTL8, NSUN2, NSUN3, THUMPD3, TRDMT1, TRMO, TRMT1, TRMT10C, TRMT11, TRMT13, TRMT1L, TRMT2A, TRMT2B, TRMT44, TRMT5, TRMT61A, TRMT61B, TRMT9B, and TYW3.

4.6. Candidate Target Prediction

We downloaded 5' and 3' UTR of the NCBI RefSeq using the UCSC Table Browser. We then selected tRF5s differently expressed with SARS-CoV-2 infection and determined tRF5 sequences from the most abundant sequence reads in the bowtie2 analysis among the reads mapped to the initial 33-mer tRF5 database. We first used RNAhybrid [32] to identify all 5' or 3' UTR sites interacting with these tRF5 sequences with a hybridization energy of less than -25 kcal/mol. Next, we selected pairs with consecutive sequence matches of 8-mer or more. These tRF5-mRNA pairs were further narrowed down in two ways: (1) 3'-end sides of tRF5s interacting with 3' UTRs [16] and (2) among these, 5' UTR of the

same mRNA interacting with the same tRF5 [17], but from the other side of the tRF5 and at least 19 nt away the 3' UTR-interacting 8-mers. Note that our target prediction process yields a tRF5-target subset of conventional tRF5 interactions with 3' UTR.

4.7. Enrichment Testing of tRF5 Targets

Since tRNAs carrying the same amino acid can have highly similar sequences, some tRF5s also have similar sequences, potentially leading to similar predicted targets. To avoid over-counting enriched functions among these potential targets, we chose representative tRF5s as they were the most abundantly expressed among tRNAs translating the same amino acids. We then analyzed enriched functions of these representative tRFs for the top 10 significant tRF5s using String-db [33].

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms25010399/s1>.

Author Contributions: X.B. and I.L. conceived the project, I.I., A.S., L.H., A.P., P.H., J.J., A.B., L.T., R.G., A.G., M.S. and M.O. analyzed sequencing data and biological functions, L.T. and I.L. developed target prediction methods and analyzed targets, I.I., A.S., L.H., A.P., P.H., J.J., A.B., A.G. and I.L. wrote the paper. All authors have read and agreed to the published version of the manuscript.

Funding: This work was partly supported by grants from NIH R01 AI116812, R21 AG069226, R21 AI166543, R21 AI166543-02S1, and a Research Pilot Grant from the Institute for Human Infections and Immunity (IHII) at UTMB to XB.

Data Availability Statement: The data presented in this study are available in supplementary materials.

Acknowledgments: We thank the Michigan Institute for Computational Discovery and Engineering (MICDE) and the University of Michigan student organization Genes In Disease And Symptoms (GIDAS) for their support in computational resources. We thank the miRcore volunteer program (MVP) for enabling students to participate in advanced computational research.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Lu, R.; Zhao, X.; Li, J.; Niu, P.; Yang, B.; Wu, H.; Wang, W.; Song, H.; Huang, B.; Zhu, N.; et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: Implications for virus origins and receptor binding. *Lancet* **2020**, *395*, 565–574. [[CrossRef](#)] [[PubMed](#)]
2. Hu, B.; Guo, H.; Zhou, P.; Shi, Z.L. Characteristics of SARS-CoV-2 and COVID-19. *Nat. Rev. Microbiol.* **2021**, *19*, 141–154. [[CrossRef](#)] [[PubMed](#)]
3. Masre, S.F.; Jufri, N.F.; Ibrahim, F.W.; Abdul Raub, S.H. Classical and alternative receptors for SARS-CoV-2 therapeutic strategy. *Rev. Med. Virol.* **2021**, *31*, e2207. [[CrossRef](#)] [[PubMed](#)]
4. Zalpoor, H.; Akbari, A.; Samei, A.; Forghaniesfidvajani, R.; Kamali, M.; Afzalnia, A.; Manshouri, S.; Heidari, F.; Pornour, M.; Khoshmirsafa, M.; et al. The roles of Eph receptors, neuropilin-1, P2X7, and CD147 in COVID-19-associated neurodegenerative diseases: Inflamasome and JaK inhibitors as potential promising therapies. *Cell Mol. Biol. Lett.* **2022**, *27*, 10. [[CrossRef](#)]
5. Moldofsky, H.; Patcai, J. Chronic widespread musculoskeletal pain, fatigue, depression and disordered sleep in chronic post-SARS syndrome; a case-controlled study. *BMC Neurol.* **2011**, *11*, 37. [[CrossRef](#)] [[PubMed](#)]
6. Ahmed, M.U.; Hanif, M.; Ali, M.J.; Haider, M.A.; Kherani, D.; Memon, G.M.; Karim, A.H.; Sattar, A. Neurological Manifestations of COVID-19 (SARS-CoV-2): A Review. *Front. Neurol.* **2020**, *11*, 518. [[CrossRef](#)]
7. Troyer, E.A.; Kohn, J.N.; Hong, S. Are we facing a crashing wave of neuropsychiatric sequelae of COVID-19? Neuropsychiatric symptoms and potential immunologic mechanisms. *Brain Behav. Immun.* **2020**, *87*, 34–39. [[CrossRef](#)]
8. Cantuti-Castelvetri, L.; Ojha, R.; Pedro, L.D.; Djannatian, M.; Franz, J.; Kuivinen, S.; van der Meer, F.; Kallio, K.; Kaya, T.; Anastasina, M.; et al. Neuropilin-1 facilitates SARS-CoV-2 cell entry and infectivity. *Science* **2020**, *370*, 856–860. [[CrossRef](#)]
9. Daly, J.L.; Simonetti, B.; Klein, K.; Chen, K.E.; Williamson, M.K.; Anton-Plagaro, C.; Shoemark, D.K.; Simon-Gracia, L.; Bauer, M.; Hollandi, R.; et al. Neuropilin-1 is a host factor for SARS-CoV-2 infection. *Science* **2020**, *370*, 861–865. [[CrossRef](#)]
10. Choi, E.J.; Wu, W.; Zhang, K.; Lee, I.; Kim, I.H.; Lee, Y.S.; Bao, X. ELAC2, an Enzyme for tRNA Maturation, Plays a Role in the Cleavage of a Mature tRNA to Produce a tRNA-Derived RNA Fragment During Respiratory Syncytial Virus Infection. *Front. Mol. Biosci.* **2020**, *7*, 609732. [[CrossRef](#)]
11. Wu, W.; Choi, E.J.; Wang, B.; Zhang, K.; Adam, A.; Huang, G.; Tunkle, L.; Huang, P.; Goru, R.; Imirowicz, I.; et al. Changes of Small Non-coding RNAs by Severe Acute Respiratory Syndrome Coronavirus 2 Infection. *Front. Mol. Biosci.* **2022**, *9*, 821137. [[CrossRef](#)] [[PubMed](#)]

12. Wyler, E.; Mosbauer, K.; Franke, V.; Diag, A.; Gottula, L.T.; Arsie, R.; Klironomos, F.; Koppstein, D.; Honzke, K.; Ayoub, S.; et al. Transcriptomic profiling of SARS-CoV-2 infected human cell lines identifies HSP90 as target for COVID-19 therapy. *iScience* **2021**, *24*, 102151. [CrossRef] [PubMed]
13. Suzuki, T. The expanding world of tRNA modifications and their disease relevance. *Nat. Rev. Mol. Cell Biol.* **2021**, *22*, 375–392. [CrossRef]
14. UniProt, C. UniProt: The Universal Protein Knowledgebase in 2023. *Nucleic Acids Res.* **2023**, *51*, D523–D531. [CrossRef]
15. Zhang, W.; Foo, M.; Eren, A.M.; Pan, T. tRNA modification dynamics from individual organisms to metaepitranscriptomics of microbiomes. *Mol. Cell* **2022**, *82*, 891–906. [CrossRef] [PubMed]
16. Deng, J.; Ptashkin, R.N.; Chen, Y.; Cheng, Z.; Liu, G.; Phan, T.; Deng, X.; Zhou, J.; Lee, I.; Lee, Y.S.; et al. Respiratory Syncytial Virus Utilizes a tRNA Fragment to Suppress Antiviral Responses Through a Novel Targeting Mechanism. *Mol. Ther.* **2015**, *23*, 1622–1629. [CrossRef]
17. Lee, I.; Ajay, S.S.; Yook, J.I.; Kim, H.S.; Hong, S.H.; Kim, N.H.; Dhanasekaran, S.M.; Chinnaiyan, A.M.; Athey, B.D. New class of microRNA targets containing simultaneous 5'-UTR and 3'-UTR interaction sites. *Genome Res.* **2009**, *19*, 1175–1183. [CrossRef]
18. Yin, L.; Li, J.; Wang, J.; Pu, T.; Wei, J.; Li, Q.; Wu, B.J. MAOA promotes prostate cancer cell perineural invasion through SEMA3C/PlexinA2/NRP1-cMET signaling. *Oncogene* **2021**, *40*, 1362–1374. [CrossRef]
19. Patton, M.J.; Orihuela, C.J.; Harrod, K.S.; Bhuiyan, M.A.N.; Dominic, P.; Kevil, C.G.; Fort, D.; Liu, V.X.; Farhat, M.; Koff, J.L.; et al. COVID-19 bacteremic co-infection is a major risk factor for mortality, ICU admission, and mechanical ventilation. *Crit. Care* **2023**, *27*, 34. [CrossRef]
20. Mautner, L.; Hoyos, M.; Dangel, A.; Berger, C.; Ehrhardt, A.; Baiker, A. Replication kinetics and infectivity of SARS-CoV-2 variants of concern in common cell culture models. *Virol. J.* **2022**, *19*, 76. [CrossRef]
21. Chuckran, C.A.; Liu, C.; Bruno, T.C.; Workman, C.J.; Vignali, D.A. Neuropilin-1: A checkpoint target with unique implications for cancer immunology and immunotherapy. *J. Immunother. Cancer* **2020**, *8*, e000967. [CrossRef] [PubMed]
22. Chen, T.H.; Chang, C.J.; Hung, P.H. Possible Pathogenesis and Prevention of Long COVID: SARS-CoV-2-Induced Mitochondrial Disorder. *Int. J. Mol. Sci.* **2023**, *24*, 8034. [CrossRef] [PubMed]
23. Bowe, B.; Xie, Y.; Al-Aly, Z. Postacute sequelae of COVID-19 at 2 years. *Nat. Med.* **2023**, *29*, 2347–2357. [CrossRef] [PubMed]
24. Wu, W.; Lee, I.; Spratt, H.; Fang, X.; Bao, X. tRNA-Derived Fragments in Alzheimer's Disease: Implications for New Disease Biomarkers and Neuropathological Mechanisms. *J. Alzheimers Dis.* **2021**, *79*, 793–806. [CrossRef] [PubMed]
25. Kodama, Y.; Shumway, M.; Leinonen, R.; International Nucleotide Sequence Database, C. The Sequence Read Archive: Explosive growth of sequencing data. *Nucleic Acids Res.* **2012**, *40*, D54–D56. [CrossRef]
26. Leinonen, R.; Sugawara, H.; Shumway, M.; International Nucleotide Sequence Database, C. The sequence read archive. *Nucleic Acids Res.* **2011**, *39*, D19–D21. [CrossRef]
27. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* **2011**, *17*, 3. [CrossRef]
28. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357–359. [CrossRef]
29. Karolchik, D.; Hinrichs, A.S.; Furey, T.S.; Roskin, K.M.; Sugnet, C.W.; Haussler, D.; Kent, W.J. The UCSC Table Browser data retrieval tool. *Nucleic Acids Res.* **2004**, *32*, D493–D496. [CrossRef]
30. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R.; Genome Project Data Processing, S. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079. [CrossRef]
31. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **2014**, *15*, 550. [CrossRef] [PubMed]
32. Rehmsmeier, M.; Steffen, P.; Hochsmann, M.; Giegerich, R. Fast and effective prediction of microRNA/target duplexes. *RNA* **2004**, *10*, 1507–1517. [CrossRef] [PubMed]
33. Szklarczyk, D.; Gable, A.L.; Nastou, K.C.; Lyon, D.; Kirsch, R.; Pyysalo, S.; Doncheva, N.T.; Legeay, M.; Fang, T.; Bork, P.; et al. The STRING database in 2021: Customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res.* **2021**, *49*, D605–D612. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.