

<b>Introduction</b>	<b>1</b>
<b>Methods and Results</b>	<b>2</b>
Protein–protein Interaction	2
Proteomics Analysis	4
Translateomics Analysis	7
<b>Discussion</b>	<b>12</b>
Key Findings	12
Proteomics vs. Translateomics	12
<b>Conclusion</b>	<b>13</b>
Implication	13
Limitation	13
<b>Reference</b>	<b>14</b>
<b>Appendix</b>	<b>14</b>
Appendix 1: Links to jupyter notebook for each part of the analysis	14
Appendix 2: Panther result of PPI specific to NSP2	14
Appendix 3: Panther result of PPI specific to EGFP	16

Document instruction

- *Click the table of content to jump to a specific section of the report.*
- *Click the link under Protein–protein Interaction manhattan plot to access complete analysis.*
- *Click the link in the Discussion section to jump to bookmarked visualisation under Methods and Results.*
- *Lengthy result screenshots were noted in the report and listed in Appendix.*
- *The word count for this document is 1590 excluding reference and appendix, the work count for method workflow screenshot is 481. Total word count is 2071, did not exceed 10% of word limit.*

# The Comprehensive Protein–Protein Interaction, Proteomics and Translateomics Analysis on SARS-CoV-2

## Introduction

SARS-CoV-2 is a strain of coronavirus that caused COVID-19, resulting in a pandemic that lasted for more than three years, incurred significant costs socially and economically in our society.

At the time, urgent solution for possible treatment of SARS-CoV-2, which requires our understanding of the molecule details of how SARS-CoV-2 infects cells, how does it modulates the host cell pathways, those knowledge will allow us inhibit these pathways to prevents viral replication, developing treatments accordingly, and reduce the risk of future pandemic from similar coronavirus.

Building on the relevant studies (Gordon et al., 2020) and (Bojkova et al., 2020), this report focuses on two separate studies on SARS-CoV-2:

1. **Protein–protein interaction:** when using SARS-CoV-2 proteins as “baits”, what are the human proteins that NSP2 specifically interacts with in comparison with the EGFP control group.

Assigned in group 11, the viral and control groups used in this study are:  
NSP2\_1, NSP2\_2, NSP2\_3, EGFP\_1, EGFP\_2, EGFP\_4.

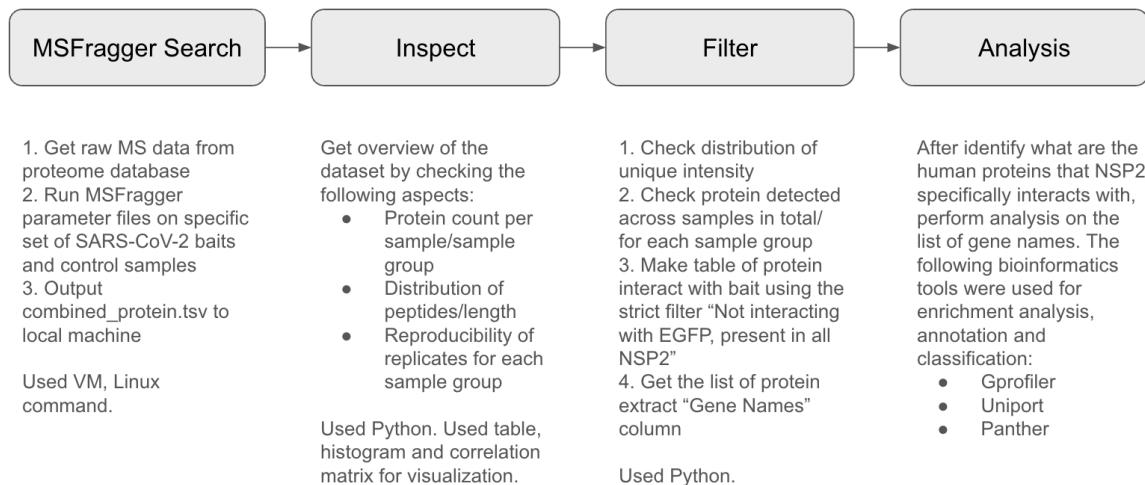
We used raw MS data from a proteome database and MSFagger parameter files to generate *combined\_protein.tsv* for this analysis.

2. **Proteomics and translateomics analysis** during infection by monitoring host and viral protein at different time points, the datasets used were summarised below:
  - a. For proteomics analysis:
    - *Timecourse of proteome measurements.xlsx* - the primary dataset
    - *SARS-CoV-2 proteins.xlsx* - used to differentiate human proteins from viral proteins due to the lack of label in timecourse dataset
  - b. For translateomics analysis:
    - *Translatome\_measurements\_use.xlsx*

## Methods and Results

The methods used for each part of study were summarised as workflow and included as screenshots. The complete jupyter notebook for each part of analysis can be found in Appendix 1.

### Protein–protein Interaction

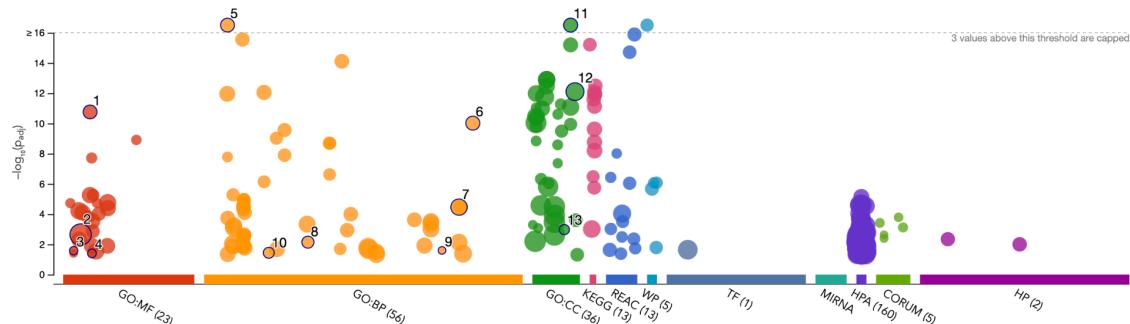


By using EGFP as a control group, we filter out the list of proteins that were detected from none of the EGFP samples but all of the NSP2 samples, so that we are confident they interact with NSP2.

Here's the gene names of that list of proteins:

PPP2R3B GIGYF2 EIF4E2 VDAC3 RAP1GDS1 PGRMC1 VDAC1 MT-CO2 FKBP15  
COX4I1 COX5B TMPO KIF26A ATP5PD SLC27A2 ELOC COX7A2 BSG CANX  
UQCRC2 RPN1 COX5A ATP5PB WASHC4 COX7C WASHC2A UQCRC1 POR  
RPN2 GNB1 RNF2 WASHC5 AMOT MT-ATP6 RRM1 CASP2 DDOST ATP5MF  
NDUFA9 ATP5F1D CORO1C ATP5MG MIEF1 KIF2A ISCA2 ARPC3 PSMC2  
THAP11 IPO5

Here's the manhattan plot based on the list of gene names:



[Click to access complete result](#)

The tree-like structure from panther tool corresponds to the results generated when using Gprofiler, which can be accessed in Appendix 2.

The key terms regarding biological process are summarised below:

1. Oxidative phosphorylation:  
This is the process where cells generate ATP, essential for energy production in aerobic organisms. Drives the proton transmembrane transport (term 2).
2. Proton transmembrane transport:  
This is about movement of protons across membranes, which is supported by oxidative phosphorylation (term 1).
3. Carbohydrate derivative biosynthetic process  
It's about the synthesis of carbohydrate derivatives, which can influence and be influenced by cellular energy states managed through processes like oxidative phosphorylation (term 1).
4. Arp2/3 complex-mediated actin nucleation  
Affect endosome fission (term 5).
5. Endosome fission  
Regarding the division of endosomes, it can affect how cells process and recycle materials.
6. protein N-linked glycosylation via asparagine  
Influence various cellular processes including those involved in material trafficking and energy production. Involve post Translational modification with sugars (glycans).

In addition, we also filtered the list of proteins only interacting with EGFP:

ZNRD2 TBC1D4 SF3B1 EPRS1 SNRPA1 PDHB ACP1 PEBP1 CPOX RPIA SUCLG1  
MRNIP PKN3 KLHDC4 COA7 PPP1R9B CCDC59 TIMM13 LDHB HDAC1 FLYWCH2  
FAM136A RPS27L GSK3B DARS1

Using the panther tool, there were no statistically significant results, which is listed as Appendix 3.

Interestingly, when analysing the reproducibility of replicates across samples, we get the following table:

	NSP2_1 Unique Intensity	NSP2_2 Unique Intensity	NSP2_3 Unique Intensity
NSP2_1 Unique Intensity	1.000000	0.990646	0.989126
NSP2_2 Unique Intensity	0.990646	1.000000	0.996416
NSP2_3 Unique Intensity	0.989126	0.996416	1.000000

	EGFP_1 Unique Intensity	EGFP_2 Unique Intensity	EGFP_4 Unique Intensity
EGFP_1 Unique Intensity	1.000000	0.318192	0.692082
EGFP_2 Unique Intensity	0.318192	1.000000	0.313779
EGFP_4 Unique Intensity	0.692082	0.313779	1.000000

## Proteomics Analysis



1. Inspect dataset
2. Perform log 10 transformation due to the large range of dataset values  
→ When performing transformation, add a small epsilon to the values, since columns like 'Control\_6h\_1' have zero values
3. Distinguish viral protein from human protein using both datasets

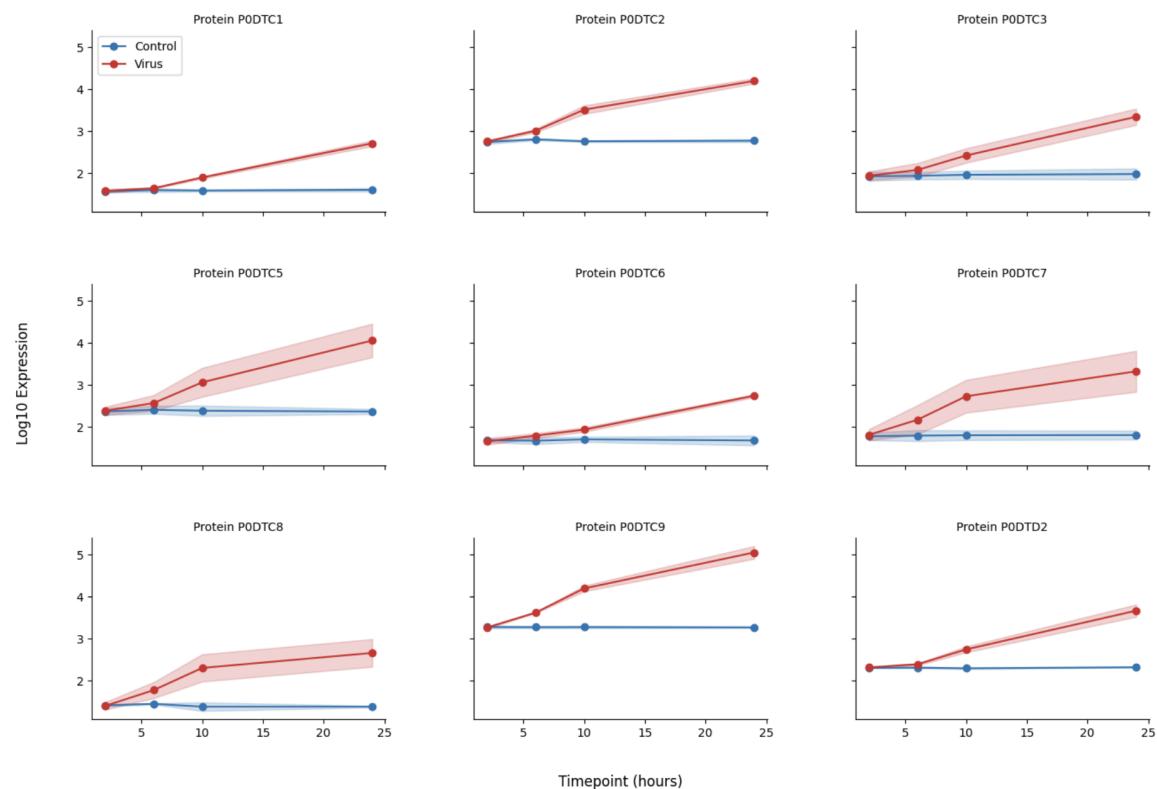
Used Python.

1. Check distribution of expression values for each timepoint/sample/replicate
2. Use relative intensity value/ratio to check which viral proteins are detected at which timepoint
3. Compare average ratio across viral and human protein
4. Show the overall change in expression for the proteome
5. Filter the data to get significantly up or down regulated proteins at each timepoint  
→ The ratio is already log 2 transformed  
→ Perform - log 10 transformation on p-values  
→ Use 5% as the threshold given common standard
6. Interpret up or down regulated proteins at each timepoint using Gprofiler, Uniport, Panther

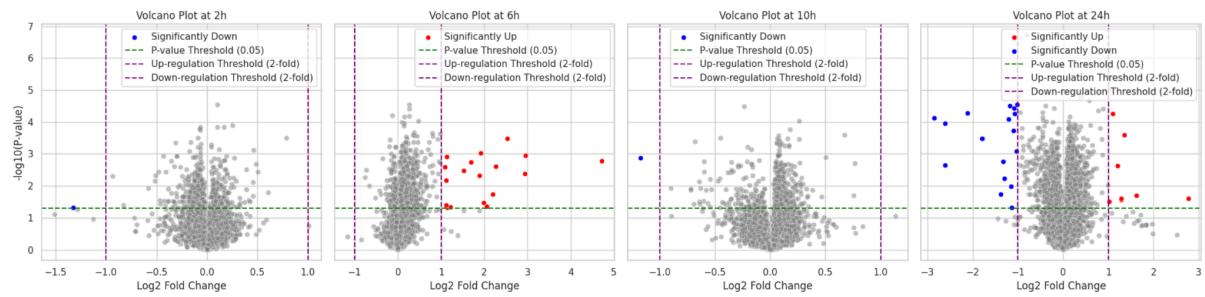
Used Python. Used line graph, scatter plot, volcano plot

This dataset contains 9 viral proteins and 6371 human proteins.

Here's the visualisation of their expression profile:



Here's the volcano plot showing differentially expressed proteins:



To summarise, here's the differentially expressed proteins in list format:

- Up-regulated Genes
  - 2h: None
  - 6h: SLC16A7, FMNL1, KRT14;KRT16;KRT17, GSTM1, HLA-C, CLU, ENO2;ENO3, CPM, ITGA3, CAV1, NEXN, UHRF1;UHRF2, GNB4, COA4, TMOD2, MTRR
  - 10h: None
  - 24h: LYPD3, KRT14;KRT16;KRT17;KRT20, H2AFX;HIST1H2AC;HIST2H2AC, LY6K, HKDC1, STAG1, CKAP4;ZFYVE19, PPP6R1

Note: For results like 'KRT14;KRT16;KRT17', we would keep 'KRT14' for analysis as the first gene name has the highest confidence.

- Down-regulated Genes
  - 2h: FOXJ3
  - 6h: None
  - 10h: LRRCC1
  - 24h: SCD, APOA2, TTR, AFP, CCK, DBI, PPIA;PPID;PPIF;PPIH;RANBP2, TMBIM6, SEPW1, APLP2;APP, INSL4, SHH, AFP, GPR64, IDI2, TSPYL1, ATP6V0A4, COA4

Using panther tools, we found the following results:

- Up\_6h: no statistically significant results.
- Up\_24h:

	Homo sapiens (REF)		upload_1 (▼ Hierarchy NEW! ⓘ)				
	#	#	expected	Fold Enrichment	+/-	raw P value	FDR
<a href="#">GO biological process complete</a>	<a href="#">74</a>	<a href="#">4</a>	.04	92.70	+	7.46E-08	1.13E-03
<a href="#">intermediate filament organization</a>							
<a href="#">↳ intermediate filament cytoskeleton organization</a>	<a href="#">94</a>	<a href="#">4</a>	.05	72.98	+	1.96E-07	1.49E-03
<a href="#">↳ intermediate filament-based process</a>	<a href="#">95</a>	<a href="#">4</a>	.06	72.21	+	2.05E-07	1.03E-03

- Down\_24h:

	Homo sapiens (REF)		upload_1 (▼ Hierarchy NEW! ⓘ)				
	#	#	expected	Fold Enrichment	+/-	raw P value	FDR
<a href="#">GO biological process complete</a>	<a href="#">8</a>	<a href="#">2</a>	.01	> 100	+	2.77E-05	4.18E-02
<a href="#">positive regulation of CoA-transferase activity</a>							
<a href="#">↳ regulation of CoA-transferase activity</a>	<a href="#">8</a>	<a href="#">2</a>	.01	> 100	+	2.77E-05	3.80E-02
<a href="#">skeletal muscle fiber differentiation</a>	<a href="#">8</a>	<a href="#">2</a>	.01	> 100	+	2.77E-05	3.49E-02
<a href="#">protein peptidyl-prolyl isomerization</a>	<a href="#">25</a>	<a href="#">4</a>	.03	> 100	+	9.99E-09	1.51E-04
<a href="#">↳ peptidyl-proline modification</a>	<a href="#">37</a>	<a href="#">4</a>	.04	> 100	+	5.18E-08	3.91E-04
<a href="#">↳ peptidyl-amino acid modification</a>	<a href="#">523</a>	<a href="#">6</a>	.53	11.24	+	1.03E-05	2.22E-02
<a href="#">positive regulation of viral genome replication</a>	<a href="#">31</a>	<a href="#">3</a>	.03	94.84	+	4.04E-06	1.53E-02
<a href="#">↳ positive regulation of viral process</a>	<a href="#">64</a>	<a href="#">3</a>	.07	45.94	+	3.67E-05	4.26E-02
<a href="#">↳ regulation of viral process</a>	<a href="#">165</a>	<a href="#">4</a>	.17	23.76	+	2.14E-05	3.60E-02
<a href="#">↳ regulation of viral life cycle</a>	<a href="#">143</a>	<a href="#">4</a>	.15	27.41	+	1.22E-05	2.31E-02
<a href="#">protein folding</a>	<a href="#">223</a>	<a href="#">5</a>	.23	21.97	+	2.52E-06	1.27E-02
<a href="#">↳ protein maturation</a>	<a href="#">494</a>	<a href="#">6</a>	.50	11.90	+	7.41E-06	1.87E-02
<a href="#">steroid metabolic process</a>	<a href="#">259</a>	<a href="#">5</a>	.26	18.92	+	5.24E-06	1.58E-02

In addition, we compared the differentially expressed human proteins with the proteins that specifically interact with NSP2 from the PPI step, they do not overlap.

## Translateomics Analysis



1. Inspect dataset
2. Perform log 10 transformation due to the large range of dataset values  
→ When performing transformation, add a small epsilon to address zero values
3. Distinguish viral protein from human protein using 'Organism' column

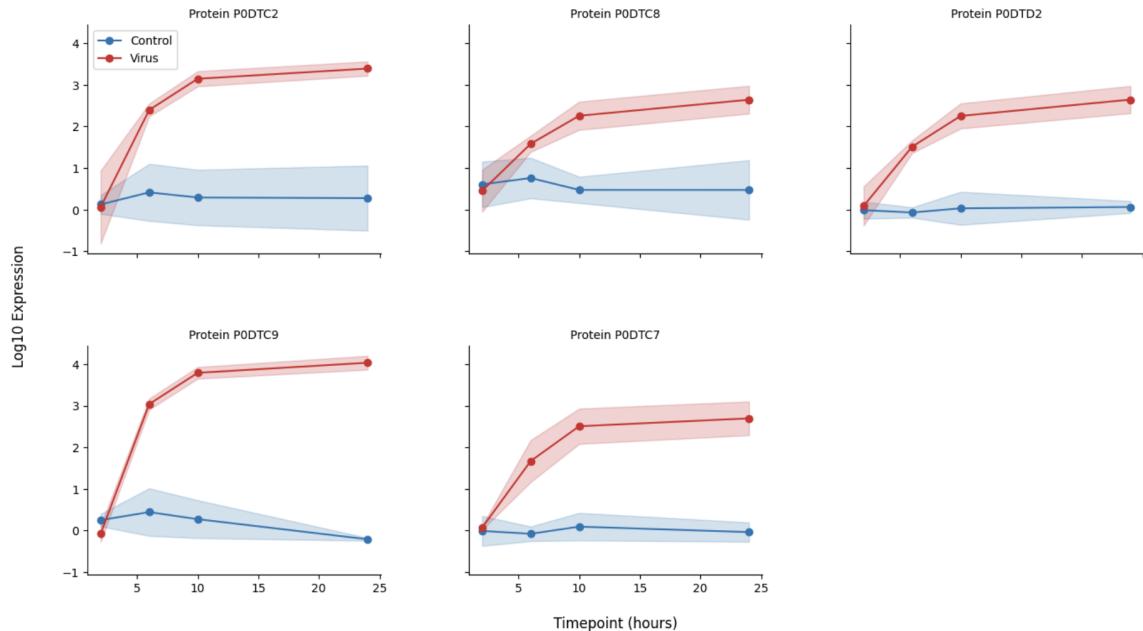
Used Python.

1. Skipped checking distribution of expression values for each timepoint/sample/replicate due to high percentage of missing values which were imputed.
2. Compare proteome dataset with translateome dataset for protein quantity
3. Use relative intensity value/ratio to check which viral proteins are translated at which timepoint
4. Compare average ratio across viral and human protein
5. Show the overall translateome change in protein expression
6. Filter the data to get significantly up or down regulated proteins at 24h, use the same method as proteome analysis, but the p-value threshold was changed to 0.1%
7. Compare fold-change expression at proteome and translateome level and get their correlation
8. Interpret up or down regulated proteins at each timepoint using Gprofiler, Uniport, Panther

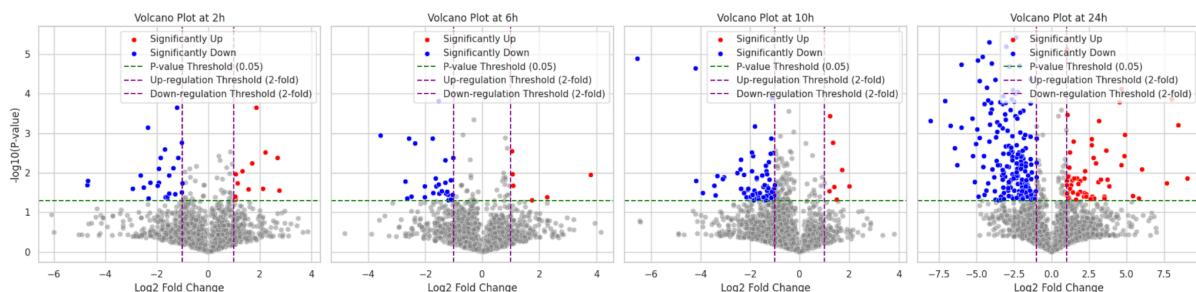
Used Python. Used line graph, scatter plot, volcano plot

This dataset contains 5 viral proteins and 2707 human proteins.

Here's the visualisation of their expression profile:



Here's the volcano plot showing differentially expressed proteins:



To simplify the analysis, we used very restrictive p value cutoff (0.001), and only focus on the 24h time point:

- Up regulated proteins: C16orf62 CEBPZ GDI1 ENO1 KRT17 ATP5A1 SLK
- Down regulated proteins:  
 AFP FABP1 VIL1 RRBP1 ANXA4 CDH17 LGALS3 SERPINA1 IDH1 TTR JUP  
 SLC9A3R1 ALDH1A1 H1F0 APOA1 OAT HIST1H1C CYP51A1 DPP4 HIST1H1E  
 CTSB RPL35 TMBIM6 ACAA2 APOE CTNNA1 APOA2 SEPTIN9 EIF1 H3F3A  
 H3F3AP4 H3F3B SEC61B EHD1 NDUFA4 ATP6V1G1 NDUFAF7 DHCR7 EPCAM  
 UBE2C DYNLRB1 TECR CIRBP MYO1D UQCR10 CUL3 TSPYL1 ABHD10 UBE2Z  
 KPNA2 NOL7

Using panther tools, we found the following results:

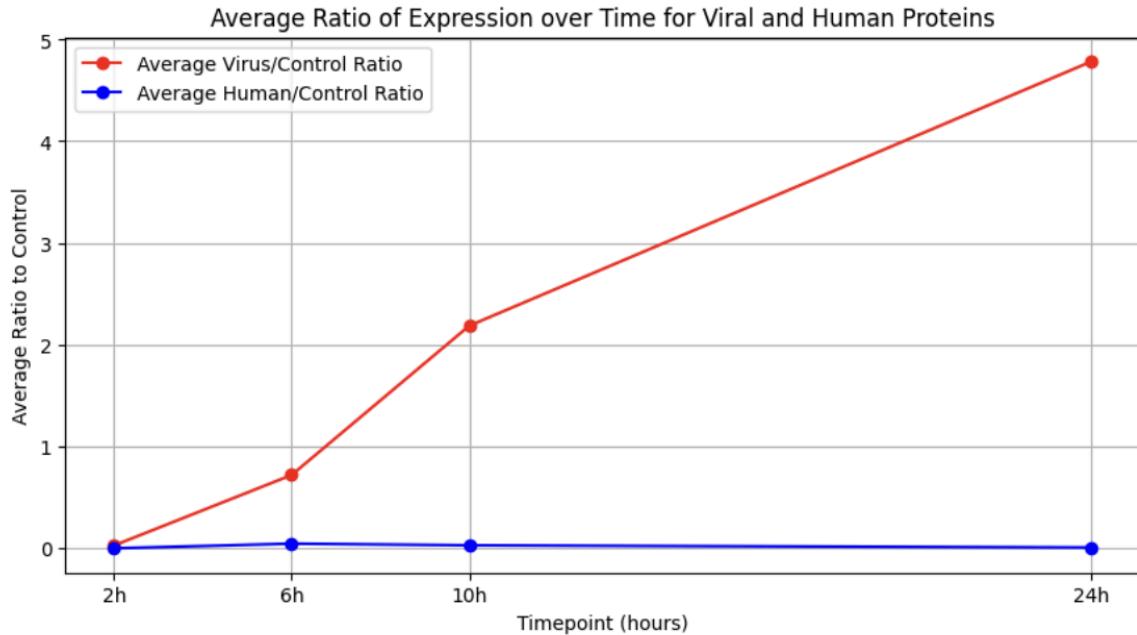
- Up\_24h: no statistically significant results.
- Down\_24h:

GO biological process complete	Homo sapiens (REF)		upload_1 (▼ Hierarchy NEW! ⓘ)				
	#	#	expected	Fold Enrichment	+/-	raw P value	FDR
<a href="#">negative regulation of very-low-density lipoprotein particle remodeling</a>	4	2	.01	> 100	+	3.05E-05	2.31E-02
↳ <a href="#">negative regulation of multicellular organismal process</a>	1092	10	2.49	4.01	+	1.46E-04	5.01E-02
↳ <a href="#">negative regulation of cellular component organization</a>	683	8	1.56	5.13	+	1.41E-04	5.09E-02
↳ <a href="#">regulation of very-low-density lipoprotein particle remodeling</a>	6	2	.01	> 100	+	7.61E-05	3.11E-02
<a href="#">positive regulation of phospholipid efflux</a>	4	2	.01	> 100	+	3.05E-05	2.20E-02
↳ <a href="#">regulation of phospholipid efflux</a>	4	2	.01	> 100	+	3.05E-05	2.10E-02
<a href="#">high-density lipoprotein particle clearance</a>	7	3	.02	> 100	+	3.88E-07	1.47E-03
↳ <a href="#">plasma lipoprotein particle clearance</a>	24	4	.05	72.98	+	2.45E-07	1.85E-03
↳ <a href="#">regulation of plasma lipoprotein particle levels</a>	59	4	.13	29.69	+	9.91E-06	8.32E-03
<a href="#">cellular response to indole-3-methanol</a>	5	2	.01	> 100	+	5.08E-05	2.74E-02
↳ <a href="#">response to indole-3-methanol</a>	6	2	.01	> 100	+	7.61E-05	3.20E-02
<a href="#">positive regulation of CoA-transferase activity</a>	8	3	.02	> 100	+	6.20E-07	1.56E-03
↳ <a href="#">regulation of CoA-transferase activity</a>	8	3	.02	> 100	+	6.20E-07	1.34E-03
<a href="#">high-density lipoprotein particle assembly</a>	10	3	.02	> 100	+	1.32E-06	2.00E-03
↳ <a href="#">plasma lipoprotein particle assembly</a>	24	3	.05	54.73	+	2.18E-05	1.74E-02
↳ <a href="#">protein-lipid complex assembly</a>	27	3	.06	48.65	+	3.14E-05	2.07E-02
↳ <a href="#">protein-containing complex assembly</a>	1269	11	2.90	3.80	+	1.04E-04	4.04E-02
↳ <a href="#">cellular component biogenesis</a>	2705	17	6.18	2.75	+	5.40E-05	2.63E-02
<a href="#">regulation of Cdc42 protein signal transduction</a>	8	2	.02	> 100	+	1.42E-04	4.98E-02
<a href="#">phospholipid efflux</a>	13	3	.03	> 100	+	3.14E-06	3.17E-03
<a href="#">reverse cholesterol transport</a>	17	3	.04	77.27	+	7.42E-06	7.01E-03

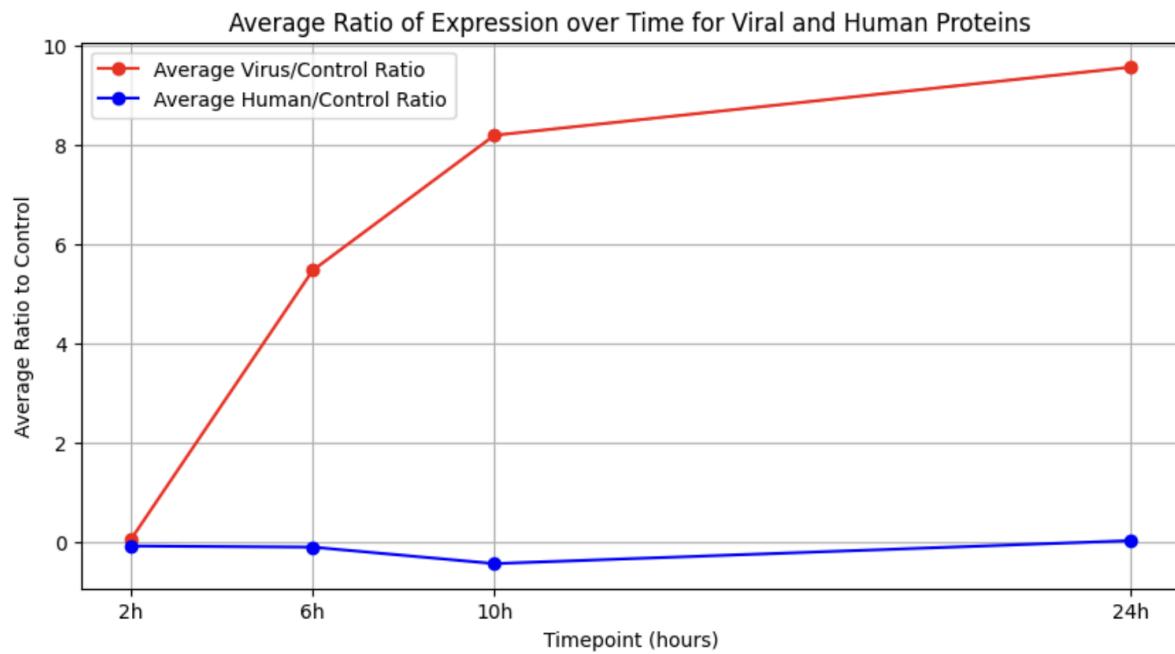
<a href="#">high-density lipoprotein particle remodeling</a>	<u>17</u>	<u>3</u>	.04	77.27	+	7.42E-06	6.60E-03
↳ <a href="#">plasma lipoprotein particle remodeling</a>	<u>32</u>	<u>3</u>	.07	41.05	+	5.29E-05	2.76E-02
↳ <a href="#">protein-lipid complex remodeling</a>	<u>32</u>	<u>3</u>	.07	41.05	+	5.29E-05	2.66E-02
↳ <a href="#">protein-containing complex remodeling</a>	<u>35</u>	<u>3</u>	.08	37.53	+	6.94E-05	3.00E-02
<a href="#">cholesterol biosynthetic process</a>	<u>37</u>	<u>4</u>	.08	47.34	+	1.49E-06	1.88E-03
↳ <a href="#">sterol biosynthetic process</a>	<u>44</u>	<u>4</u>	.10	39.81	+	3.03E-06	3.27E-03
↳ <a href="#">organic hydroxy compound metabolic process</a>	<u>486</u>	<u>9</u>	1.11	8.11	+	1.30E-06	2.19E-03
↳ <a href="#">sterol metabolic process</a>	<u>133</u>	<u>6</u>	.30	19.75	+	5.62E-07	1.70E-03
↳ <a href="#">steroid metabolic process</a>	<u>259</u>	<u>7</u>	.59	11.83	+	1.89E-06	2.19E-03
↳ <a href="#">steroid biosynthetic process</a>	<u>115</u>	<u>4</u>	.26	15.23	+	1.37E-04	5.06E-02
↳ <a href="#">cholesterol metabolic process</a>	<u>119</u>	<u>6</u>	.27	22.08	+	2.91E-07	1.47E-03
↳ <a href="#">secondary alcohol metabolic process</a>	<u>130</u>	<u>7</u>	.30	23.58	+	1.74E-08	2.63E-04
↳ <a href="#">alcohol metabolic process</a>	<u>326</u>	<u>8</u>	.74	10.75	+	6.68E-07	1.26E-03
↳ <a href="#">small molecule metabolic process</a>	<u>1629</u>	<u>13</u>	3.72	3.49	+	4.99E-05	2.90E-02
↳ <a href="#">secondary alcohol biosynthetic process</a>	<u>37</u>	<u>4</u>	.08	47.34	+	1.49E-06	2.05E-03
↳ <a href="#">alcohol biosynthetic process</a>	<u>92</u>	<u>4</u>	.21	19.04	+	5.76E-05	2.72E-02
↳ <a href="#">small molecule biosynthetic process</a>	<u>461</u>	<u>7</u>	1.05	6.65	+	7.84E-05	3.12E-02
<a href="#">cholesterol efflux</a>	<u>29</u>	<u>3</u>	.07	45.30	+	3.91E-05	2.46E-02
<a href="#">negative regulation of production of molecular mediator of immune response</a>	<u>45</u>	<u>3</u>	.10	29.19	+	1.48E-04	4.98E-02
<a href="#">cholesterol homeostasis</a>	<u>88</u>	<u>4</u>	.20	19.90	+	4.84E-05	2.93E-02
↳ <a href="#">sterol homeostasis</a>	<u>89</u>	<u>4</u>	.20	19.68	+	5.06E-05	2.83E-02
<a href="#">integrin-mediated signaling pathway</a>	<u>94</u>	<u>4</u>	.21	18.63	+	6.26E-05	2.87E-02
<a href="#">negative regulation of hydrolase activity</a>	<u>185</u>	<u>5</u>	.42	11.83	+	6.28E-05	2.79E-02
<a href="#">hormone metabolic process</a>	<u>206</u>	<u>5</u>	.47	10.63	+	1.04E-04	3.94E-02

Next, we will list results relating to the comparison between proteome and translateome datasets.

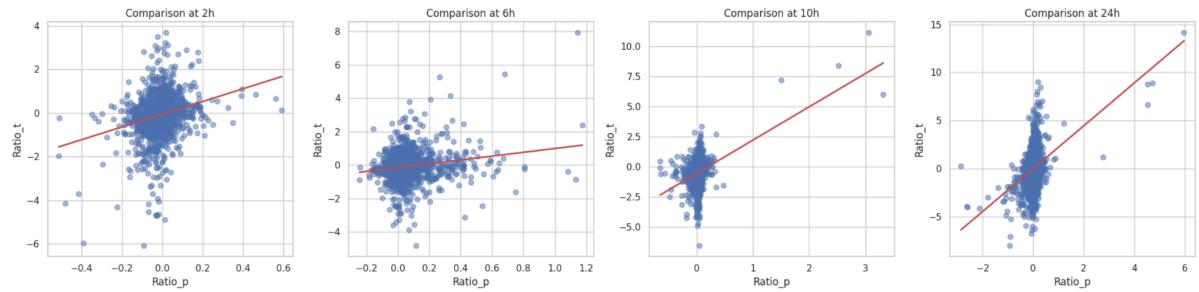
Here are comparison of average ratio for viral and human proteins:  
(Proteome)



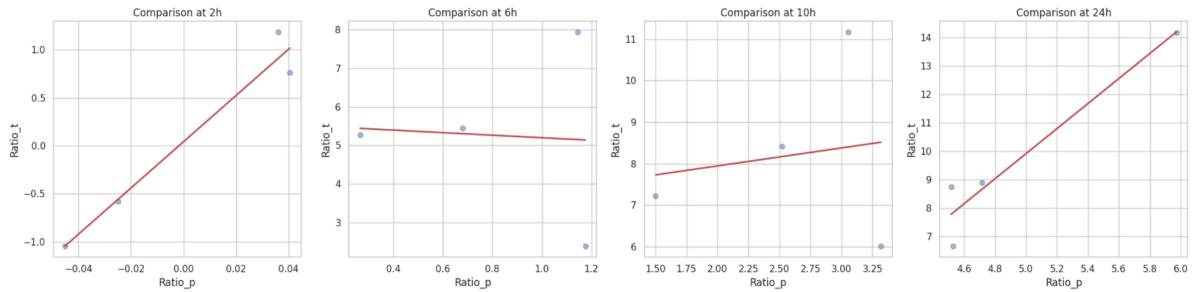
(Translateome)



Here's the comparison of correlation about the fold-change expression:  
 (both human and viral protein)



(viral protein only)



## Discussion

### Key Findings

1. Using NSP2 as bait, [the confident list of proteins](#) affect energy production and transportation, material trafficking, as well as maintaining structural organisation. This means, the selected viral group may disrupt energy production, affect the host's metabolism, and affect transportation of molecules in the cell.  
We also noticed that there are [more similarities for protein among covid group than the control group](#), which justifies our approach.
2. Regarding [the significantly up or down regulated proteins in proteome dataset](#), biological processes like 'Intermediate filament organisation' are up regulated, indicating enhanced support to cell structure. Biological processes like 'Positive Regulation of CoA-Transferase Activity' and 'Positive Regulation of Viral Genome Replication' were down regulated, indicating cells redirecting energy from growth or repairing to immune responses.  
We also compared the list of proteins that specifically interacts with NSP2 and found they do not overlap with this list of proteins. In addition, the list of NSP2 proteins were related to ATP and metabolic process, while in the human protein, metabolic process related proteins were down regulated.
3. Regarding [the significantly up or down regulated proteins in the translateome dataset](#), for the up regulated proteins, there were no significant results. For the down regulated proteins, biological processes like 'Extracellular exosome' indicate the cell to cell communication may have been obstructed.
4. For the expression profile change of viral proteins, compared with human proteins, they become more expressed overtime, where the increase for [proteome data](#) was gradually within the 24 h window, but for [translateome data](#), the increase was drastic in the first 10 h but saturates afterwards.

### Proteomics vs. Translateomics

To compare the proteome and translatome dataset, we can look at the following two aspects.

From [the average expression ratio over time](#), we can see that viral proteins become more expressed overtime for both proteome and translatome dataset. For proteome at 6h, there was some suppression that changed the trend of the graph. Similarly, at 10h for translatome data, more human proteins were translated after this time point. However, it is important to note that the change in viral proteins was drastic while the human protein expression ratio only fluctuates on a small scale.

From [the comparison of correlation about the fold-change expression](#), we can see the viral proteins were more pronounced as outliers as time went by, while the human proteins from unclear association moved to stronger positive correlation. If we compare the range of the scatter plot, the horizontal range becomes more concentrated over time, while the vertical range increases to almost twice at the 24 h timepoint than it started with. We should also

note that whatever happened at 24h were not just newly translated proteins, but also accumulated effects from the past 24h.

## Conclusion

### Implication

From this study, we understand SARS-CoV-2 primarily affects the host's metabolic process, energy production and structural maintenance. When infected, host cells may experience reduced energy production and shift focus in the metabolic process, changing the host cell's priority from cell growth to immune response. The proteome expression changes were stably increased throughout the 24 h time window, while the translateome change slows down after the 10 h timepoint.

Those findings would help us understand how SARS-CoV-2 infect cells, which pathways were influenced during the process, what symptoms are associated with the infection, and this knowledge will allow us to have better targets when developing treatments.

### Limitation

However, there are potential flaws in the experiment design that could be improved.

For the proteome dataset, there are 9 viral proteins, and 6371 human proteins, but for the translatome dataset, there are 5 viral proteins, and 2710 human proteins. The reason could be that not all proteins will be translated since the MS process is not very sensitive, given we only add the heavy arginine and heavy lysine at each time point and measure it immediately, the newly synthesised protein may not be enough to be detected.

Due to limited time and resources (that we were only allocated to 3 control and viral proteins at group 11), we only explored a subset of proteins which interacts with NSP2, and our observation is only partial and based on our allocated samples.

By addressing the limitation, the future study will gain a more comprehensive understanding of SARS-CoV-2 and the potential treatment for it.

## Reference

[1] Gordon, D. E., Jang, G. M., Bouhaddou, M., Xu, J., Obernier, K., O'Meara, M. J., ... Krogan, N. J. (2020). A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature*, 583(7816), 459–468. <https://doi.org/10.1038/s41586-020-2286-9>

[2] Bojkova, D., Klann, K., Koch, B., Widera, M., Krause, D., Ciesek, S., ... Cinatl, J. (2020). Proteomics of SARS-CoV-2-infected host cells reveals therapy targets. *Nature*, 583(7816), 469–472. <https://doi.org/10.1038/s41586-020-2332-7>

## Appendix

Appendix 1: Links to jupyter notebook for each part of the analysis

- Protein–protein Interaction:  
<https://colab.research.google.com/drive/1iOeL-byAKAJbQL-SIZmpznSvxjdQn3wT?usp=sharing>
- Proteomics Analysis:  
<https://colab.research.google.com/drive/1cdccdwBPgJePV0sz52iRJvNiwcCOoXgr?usp=sharing>
- Translateomics Analysis:  
[https://colab.research.google.com/drive/1ozb8L\\_jM07-Vkv5fn74Cnxb65fCvAeWD?usp=sharing](https://colab.research.google.com/drive/1ozb8L_jM07-Vkv5fn74Cnxb65fCvAeWD?usp=sharing)

Appendix 2: Panther result of PPI specific to NSP2

	Homo sapiens (REF)		upload_1 (▼ Hierarchy NEW! ⓘ)				
	#	#	expected	Fold Enrichment	+/-	raw P value	FDR
<a href="#">GO biological process complete</a>							
<a href="#">endosome fission</a>	3	2	.01	> 100	+	1.66E-05	4.66E-03
↳ <a href="#">cellular process</a>	14769	47	35.16	1.34	+	3.81E-05	9.94E-03
<a href="#">mitochondrial electron transport, cytochrome c to oxygen</a>	23	6	.05	> 100	+	1.30E-11	1.64E-08
↳ <a href="#">mitochondrial ATP synthesis coupled electron transport</a>	92	9	.22	41.09	+	8.47E-13	1.42E-09
↳ <a href="#">ATP synthesis coupled electron transport</a>	92	9	.22	41.09	+	8.47E-13	1.28E-09
↳ <a href="#">respiratory electron transport chain</a>	117	9	.28	32.31	+	7.71E-12	1.06E-08
↳ <a href="#">electron transport chain</a>	127	11	.30	36.38	+	7.59E-15	1.91E-11
↳ <a href="#">generation of precursor metabolites and energy</a>	373	17	.89	19.14	+	6.55E-18	2.48E-14
↳ <a href="#">cellular metabolic process</a>	5659	32	13.47	2.37	+	3.72E-08	2.01E-05
↳ <a href="#">metabolic process</a>	7829	34	18.64	1.82	+	1.06E-05	3.15E-03
↳ <a href="#">cellular respiration</a>	194	15	.46	32.47	+	2.84E-19	1.43E-15
↳ <a href="#">energy derivation by oxidation of organic compounds</a>	273	15	.65	23.08	+	4.96E-17	1.50E-13
↳ <a href="#">oxidative phosphorylation</a>	119	15	.28	52.94	+	1.45E-22	2.19E-18
↳ <a href="#">aerobic respiration</a>	163	15	.39	38.65	+	1.96E-20	1.48E-16
↳ <a href="#">aerobic electron transport chain</a>	87	9	.21	43.45	+	5.05E-13	9.54E-10
<a href="#">neuron-neuron synaptic transmission</a>	8	2	.02	> 100	+	1.54E-04	3.76E-02
<a href="#">protein N-linked glycosylation via asparagine</a>	24	3	.06	52.50	+	2.48E-05	6.69E-03
↳ <a href="#">organonitrogen compound biosynthetic process</a>	1337	14	3.18	4.40	+	1.75E-06	6.45E-04
↳ <a href="#">organonitrogen compound metabolic process</a>	4641	24	11.05	2.17	+	4.69E-05	1.20E-02
↳ <a href="#">biosynthetic process</a>	4002	21	9.53	2.20	+	1.73E-04	4.16E-02
↳ <a href="#">carbohydrate derivative metabolic process</a>	997	12	2.37	5.06	+	2.74E-06	9.86E-04
↳ <a href="#">carbohydrate derivative biosynthetic process</a>	604	12	1.44	8.34	+	1.25E-08	9.02E-06

<a href="#">proton motive force-driven mitochondrial ATP synthesis</a>	64	7	.15	45.94	+	1.55E-10	1.81E-07
↳ <a href="#">proton motive force-driven ATP synthesis</a>	73	7	.17	40.27	+	4.01E-10	4.33E-07
↳ <a href="#">ATP biosynthetic process</a>	84	7	.20	35.00	+	1.09E-09	1.10E-06
↳ <a href="#">purine ribonucleoside triphosphate biosynthetic process</a>	95	7	.23	30.95	+	2.62E-09	2.47E-06
↳ <a href="#">purine nucleoside triphosphate biosynthetic process</a>	96	7	.23	30.62	+	2.82E-09	2.51E-06
↳ <a href="#">nucleoside triphosphate biosynthetic process</a>	109	7	.26	26.97	+	6.88E-09	5.48E-06
↳ <a href="#">nucleoside triphosphate metabolic process</a>	209	7	.50	14.07	+	6.03E-07	2.34E-04
↳ <a href="#">nucleoside phosphate metabolic process</a>	536	9	1.28	7.05	+	4.19E-06	1.41E-03
↳ <a href="#">nucleobase-containing small molecule metabolic process</a>	605	9	1.44	6.25	+	1.11E-05	3.23E-03
↳ <a href="#">nucleoside phosphate biosynthetic process</a>	269	9	.64	14.05	+	1.27E-08	8.71E-06
↳ <a href="#">organophosphate biosynthetic process</a>	549	9	1.31	6.89	+	5.09E-06	1.64E-03
↳ <a href="#">purine nucleoside triphosphate metabolic process</a>	193	7	.46	15.23	+	3.52E-07	1.40E-04
↳ <a href="#">ribonucleoside triphosphate biosynthetic process</a>	101	7	.24	29.11	+	4.03E-09	3.38E-06
↳ <a href="#">ribonucleoside triphosphate metabolic process</a>	193	7	.46	15.23	+	3.52E-07	1.44E-04
↳ <a href="#">purine ribonucleoside triphosphate metabolic process</a>	186	7	.44	15.81	+	2.74E-07	1.15E-04
↳ <a href="#">purine ribonucleotide biosynthetic process</a>	195	8	.46	17.23	+	1.82E-08	1.20E-05
↳ <a href="#">purine nucleotide biosynthetic process</a>	231	8	.55	14.55	+	6.77E-08	3.20E-05
↳ <a href="#">purine-containing compound biosynthetic process</a>	240	8	.57	14.00	+	9.09E-08	4.17E-05
↳ <a href="#">purine-containing compound metabolic process</a>	498	8	1.19	6.75	+	2.10E-05	5.76E-03
↳ <a href="#">nucleotide biosynthetic process</a>	267	9	.64	14.16	+	1.19E-08	8.98E-06
↳ <a href="#">nucleotide metabolic process</a>	528	9	1.26	7.16	+	3.71E-06	1.27E-03
↳ <a href="#">purine nucleotide metabolic process</a>	468	8	1.11	7.18	+	1.34E-05	3.82E-03
↳ <a href="#">purine ribonucleotide metabolic process</a>	386	8	.92	8.70	+	3.28E-06	1.15E-03
↳ <a href="#">ribonucleotide metabolic process</a>	406	8	.97	8.28	+	4.76E-06	1.56E-03
↳ <a href="#">ribose phosphate metabolic process</a>	414	8	.99	8.12	+	5.49E-06	1.73E-03
↳ <a href="#">ribonucleotide biosynthetic process</a>	210	8	.50	16.00	+	3.24E-08	1.81E-05
↳ <a href="#">ribose phosphate biosynthetic process</a>	217	8	.52	15.48	+	4.18E-08	2.04E-05
↳ <a href="#">ATP metabolic process</a>	162	7	.39	18.15	+	1.07E-07	4.62E-05
 <a href="#">regulation of Arp2/3 complex-mediated actin nucleation</a>	33	3	.08	38.18	+	6.58E-05	1.66E-02
↳ <a href="#">regulation of actin nucleation</a>	42	3	.10	30.00	+	1.36E-04	3.38E-02
 <a href="#">proton transmembrane transport</a>	146	11	.35	31.64	+	3.62E-14	7.81E-11
↳ <a href="#">inorganic cation transmembrane transport</a>	638	12	1.52	7.90	+	2.30E-08	1.45E-05
↳ <a href="#">inorganic ion transmembrane transport</a>	733	12	1.75	6.88	+	1.05E-07	4.66E-05
↳ <a href="#">transmembrane transport</a>	1283	14	3.05	4.58	+	1.07E-06	4.05E-04
↳ <a href="#">transport</a>	3665	22	8.73	2.52	+	1.05E-05	3.18E-03
↳ <a href="#">establishment of localization</a>	3935	22	9.37	2.35	+	3.37E-05	8.93E-03
↳ <a href="#">localization</a>	4498	25	10.71	2.33	+	6.81E-06	2.10E-03
↳ <a href="#">monoatomic cation transmembrane transport</a>	653	12	1.55	7.72	+	2.97E-08	1.72E-05
↳ <a href="#">monoatomic ion transmembrane transport</a>	815	13	1.94	6.70	+	3.75E-08	1.96E-05
↳ <a href="#">monoatomic ion transport</a>	980	14	2.33	6.00	+	3.98E-08	2.01E-05
↳ <a href="#">monoatomic cation transport</a>	787	13	1.87	6.94	+	2.49E-08	1.50E-05
 <a href="#">post-transcriptional gene silencing</a>	46	3	.11	27.39	+	1.79E-04	4.23E-02

### Appendix 3: Panther result of PPI specific to EGFP

**Analysis Summary:** Please report in publication [?](#)

**Analysis Type:** PANTHER Overrepresentation Test (Released 20240807)

**Annotation Version and Release Date:** GO Ontology database DOI: 10.5281/zenodo.12173881 Released 2024-06-17

**Analyzed List:** upload\_1 (Homo sapiens) [Change](#)

**Reference List:** Homo sapiens (all genes in database) [Change](#)

**Annotation Data Set:**  [?](#)

**Test Type:**  Fisher's Exact  Binomial

**Correction:**  Calculate False Discovery Rate  Use the Bonferroni correction for multiple testing [?](#)  No correction

**Results** [?](#)

	Reference list	upload_1
Uniquely Mapped IDs:	<a href="#">20580</a> out of 20580	<a href="#">25</a> out of 25
Unmapped IDs:	<a href="#">0</a>	<a href="#">0</a>
Multiple mapping information:	<a href="#">0</a>	<a href="#">0</a>

Export [Table](#) [XML with user input ids](#) [JSON with user input ids](#)

No statistically significant results. [Click to see all results.](#)