livers

MDPI

Article

Understanding the SARS-CoV-2–Human Liver Interactome Using a Comprehensive Analysis of the Individual Virus–Host Interactions

Giovanni Colonna

Unit of Medical Informatics—AOU-L. Vanvitelli, University of Campania, 80138 Naples, Italy; giovanni.colonna@unicampania.it

Abstract: Many metabolic processes at the molecular level support both viral attack strategies and human defenses during COVID-19. This knowledge is of vital importance in the design of antiviral drugs. In this study, we extracted 18 articles (2021–2023) from PubMed reporting the discovery of hub nodes specific for the liver during COVID-19, identifying 142 hub nodes. They are highly connected proteins from which to obtain deep functional information on viral strategies when used as functional seeds. Therefore, we evaluated the functional and structural significance of each of them to endorse their reliable use as seeds. After filtering, the remaining 111 hubs were used to obtain by STRING an enriched interactome of 1111 nodes (13,494 interactions). It shows the viral strategy in the liver is to attack the entire cytoplasmic translational system, including ribosomes, to take control of protein biosynthesis. We used the SARS2-Human Proteome Interaction Database (33,791 interactions), designed by us with BioGRID data to implement a reverse engineering process that identified human proteins actively interacting with viral proteins. The results show 57% of human liver proteins are directly involved in COVID-19, a strong impairment of the ribosome and spliceosome, an antiviral defense mechanism against cellular stress of the p53 system, and, surprisingly, a viral capacity for multiple protein attacks against single human proteins that reveal underlying evolutionary–topological molecular mechanisms. Viral behavior over time suggests different molecular strategies for different organs.

Keywords: COVID-19; COVID-19 molecular mechanisms; SARS-CoV-2; liver interactome; ribosome; liver proteome during COVID-19; viral strategy

check for updates

Citation: Colonna, G. Understanding the SARS-CoV-2–Human Liver Interactome Using a Comprehensive Analysis of the Individual Virus–Host Interactions. *Livers* **2024**, *4*, 209–239. https://doi.org/10.3390/livers4020016

Academic Editor: Melanie Deutsch

Received: 11 February 2024
Revised: 25 March 2024
Accepted: 10 April 2024
Published: 30 April 2024

Copyright: © 2024 by the author. 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

The impact of COVID-19 on various organs is under intense investigation, as clinicians have identified this infection as a systemic disease, leading to significant research efforts in this area. Liver manifestations of COVID-19 have garnered attention because of their clinical significance in vulnerable patient populations. Studies have reported diverse outcomes, with liver damage ranging from mild and self-limiting in healthy individuals to severe and potentially fatal in the elderly, obese, and those with pre-existing liver conditions [1].

Despite the challenges, researchers have made progress in elucidating the pathophysiological mechanisms associated with liver involvement in COVID-19 [2]. Various tissues have shown viral RNA, suggesting potential direct viral involvement [3–6]. However, histological analysis has revealed non-specific findings [2], showing the need for further investigation into the mechanisms of liver damage [7,8]. However, the precise molecular mechanisms underlying liver injury are still not understood.

Computational approaches, such as gene expression analysis, have emerged as valuable tools in understanding COVID-19 pathogenesis and its effects on liver metabolism [9,10]. By evaluating changes in gene expression, researchers have identified hub genes associated with COVID-19 [11], offering insights into disease progression and potential therapeutic tar-

gets. These hub genes play crucial roles in coordinating metabolic processes, although there is disagreement in identifying, characterizing, and classifying these types of nodes [12].

Many studies on PubMed [11,13–22] from 2021 to 2023 reported hub genes linked to COVID-19. They aim to determine regulatory processes from datasets based on microarray or transcriptome sequencing technology. Despite this progress, challenges persist in studying COVID-19-associated liver damage because these studies rely on static and probabilistic models, lacking spatial and temporal resolution, which limits our understanding of disease progression [23–25]. The lack of a unified definition of liver damage complicates data interpretation and comparison across studies.

To address these challenges, we have turned to computational tools for analyzing protein–protein interactions and functional pathways associated with SARS-CoV-2. Here, we apply a biological reverse engineering protocol that involves deriving a model of the biological relationships established between the nodes implementing the networks, with no a priori knowledge of their computational protocols [26,27]. This approach can provide valuable insights into the molecular mechanisms underlying COVID-19 pathogenesis, decreasing biased conclusions from low-resolution data [28] with a more systematic understanding of the complex regulatory networks that govern disease [29,30]. The reverse engineering is based on the direct validation of the biological message exchanged between two nodes by validating it with an external tool.

However, the concept of degeneracy in biological systems [29,31] adds another layer of complexity to our understanding of COVID-19 pathogenesis. Degeneracy refers to the situation where distinct processes within a biological system can perform similar functions or roles, making it challenging to pinpoint exact cause–effect relationships [32–34]. This complexity underscores prizing rigorous experimental validation in elucidating the molecular mechanisms underlying COVID-19-associated liver damage.

Experimental validation of computational findings through methods of biophysics and biochemical tests is crucial for confirming the relevance of identified hub genes and biological pathways in COVID-19 pathogenesis. By integrating computational and experimental approaches, researchers can overcome the limitations of individual methodologies and gain a more comprehensive understanding of the disease (more details in Appendix A).

In conclusion, although researchers have made significant progress in understanding the liver manifestations of COVID-19, many challenges still exist. By integrating computational and experimental approaches and leveraging bioinformatics tools, researchers can gain deeper insights into the molecular mechanisms underlying COVID-19 pathogenesis in the liver, leading to developing more effective therapeutic strategies.

2. Materials and Methods

2.1. BioGRID

BioGRID [35] is the source of experimental interactions of SARS-CoV-2 (BioGRID Version 4.4.223 as of July 2023): https://thebiogrid.org/search.php?search=SARS-CoV-2*&organism=2697049, (accessed on 23 July 2023).

BioGRID is a curated biological database of protein–protein interactions, genetic interactions, chemical interactions, and post-translational modifications. It also collects all the experimentally proven data on the interactions between the 31 SARS-CoV-2 proteins and the human proteome. The quantitative SAINT analysis [36] was used to identify SARS-CoV-2 viral–host proximity interactions in human or model system cells [11,13–22] and those with a Bayesian FDR =< 0.01 were high confidence. Scores are the sum of peptide counts from four mass spec runs with a higher score indicating a higher degree of connectivity between proteins. This statistical model assigns the number of peptide identifications for each interactor to a probability distribution, which is then used to estimate the likelihood of a true interaction.

2.2. STRING

STRING [37,38] (https://version-11-5.string-db.org/, accessed on 1 July 2023) is a proteomic database focusing on the networks and interactions of proteins in an array of species. The curated interactions are direct (physical) and indirect (functional) associations. In this paper, we establish the PPI network according to version 11.5 of the STRING database. We constructed PPI networks by mapping proteins to the STRING database with a confidence score of 0.900 and with all interaction source active (see also note in the Supplementary Materials).

Regarding cluster analysis, STRING also provides the most reliable clusters in terms of compactness, metabolic functionality, and *p*-value, calculated on the network data. The cluster analysis uses the K-means clustering method [39] where K-means clustering is an unsupervised centroid-based learning algorithm.

2.3. Protein Enrichment

It is to some extent based on prior knowledge, and the statistical enrichment of the annotated features may not be an intrinsic property of the input. To obtain an enrichment test from STRING that is statistically valid, we must insert the entire set of enriched proteins into STRING ensuring that "first shell" and "second shell" are both set to "none". To confirm the procedure's correctness, we also checked the STRING notes to the network for a specific notice that disappears when the analysis is performed correctly. By adding new interaction partners to the network, we can extend the interaction neighborhood according to the required confidence score. We used 0.9 as a confidence score.

2.4. Cytoscape and Network Topology Analysis

Cytoscape [40,41] through Network Analyzer was used to analyze the topological parameters of networks. Using Cytoscape software (Version 3.10.1), we visualized and analyzed PPI networks, which offer diverse plugins for multiple analyses. Cytoscape represents PPI networks as graphs with nodes illustrating proteins and edges depicting associated interactions. We examined network architecture for topological parameters such as clustering coefficient, centralization, density, network diameter, and so on. Our analysis included undirected edges for every network. We termed the number of connected neighbors of a node in a network as the degree of a node. P(k) is used to describe distributing node degrees, which counts the number of nodes with degree k where k = 0, 1, 2, ... We calculated the power law of distribution of node degrees, which is one of the most crucial network topological characteristics. The coefficient R-squared value (R^2), also known as the coefficient of determination, gives the proportion of variability in the dataset. We also examined other network parameters, including the distribution of various topological features. We performed a calculation of hub and bottleneck nodes based on relevant topological parameters. By examining the PPI network, we found the top 7 hub nodes. These nodes had higher degree values than the others and were in two central modules that were connected and compact.

2.5. CentiScaPe

Regarding centralities for undirected, directed, and weighted networks, CentiScaPe [42] computes specific centrality parameters describing the network topology. These parameters facilitate users in locating the most important nodes within a complex network. The computation of the plugin produces both numerical and graphical results, facilitating identifying key nodes even in extensive networks. Integrating network topological quantification with other numerical node attributes can provide relevant node identification and functional classification.

2.6. GO and KEGG Pathway Analyses

To better research and show the biological function of interacting proteins, we performed GO analysis, which included biological process (BP), cellular component (CC),

molecular function (MF), and many other evaluations using the specific tools present in STRING. All functions shown by STRING are significant, having a *p*-value of <0.05.

2.7. SARS2-Human Proteome Interaction Database (SHPID)

We have collected in a single database all the files made available online by BioGRID, containing all the curated physical interactions of the 31 SARS-CoV-2 proteins gained through experiments in human cellular systems with viral baits, followed by purification and characterization with mass spectrometry. These data are available as a zip file containing multiple zip files (32 zip files) each comprising interactions and post-translational modifications for each single SARS-CoV-2 protein for 33,823 interactions (as of July 2023). The database therefore contains the set of all real interactions existing between the SARS-CoV-2 proteome and all the proteins of the human proteome. We highlight that not all interactions are real and some could derive from artifacts of the method, such as non-biological interactions, only because of the random encounter between proteins in the system used, representing an encounter that would never happen in reality during an infection. However, the interactions derived from BioGRID all, even those with the lowest score, have a significant statistic with an FDR =< 0.01. This allows us to identify as many significant comparisons as possible while maintaining a low false positive rate, i.e., the probability of a false positive is less than 1%, so only 338 interactions among all are truly null.

This database is the comprehensive repository of all interactions acknowledged as biologically possible between the virus and its human host. The database also contains interactions between individual viral proteins, where known. As part of database search actions, one can ask who interacts with whom, with queries that use single human or viral proteins. The search can include multiple sets of proteins.

2.8. Comparison between GO Pairs in Enriched Networks

In modeled networks, STRING analytically defines the enriched biological terms using two parameters. Strength is the measure of how large an enrichment is, expressed as Log10 [Log10 (observed/expected)], while false discovery rate (FDR) is the measure of the statistical significance of an enrichment given as a *p*-value after the Benjamini–Hochberg procedure. The higher the strength value, the greater the biological effect because of genetic enrichment, indicating increased gene expression, which suggests a higher likelihood of the event occurring. Since STRING characterizes biological functions as pairs in which strength and FDR often show very different numerical values from each other, we use the product P [P = strength x $-$ log10 *p*-value] to carry out a quantitative evaluation. When "strength" has a very high value and *p* has a low value, this product is enhanced (the extremes of their numerical values, very high and low, represent the most favorable situation for evaluating an effect). This facilitates us to compare and evaluate different pairs. Two pairs, one characterized by S = 0.35 and FDR = 1.0×10^{-11} and another characterized by S = 1.9 and FDR = 1.0×10^{-6}, could lead one to think that the first is more significant. If we analyze the P value, we have 3.85 and 11.4. This tells us that the increase in gene expression in the second case is prevalent. The higher the value of the product, the more reliable the result of one pair will be over the other. We consider that strength = 1 means a 10-fold genetic enrichment. However, it is important to remember that all FDR values reported by STRING in its biological functionality characterizations (GO, KEGG, etc.) are always significant and never greater than 0.05.

2.9. Highlighting the Nodes of a STRING Network Involved in the Same Biological Process (GO)

STRING makes visible all the nodes involved in the same biological process evidenced through its databases mapped onto the proteins (GO, KEGG, REACTOME, and so on) by activating the process itself with a click of the cursor on the process line. Activation means that all nodes involved in the same metabolic process are colored. Nodes involved in multiple processes receive multiple colors. This tool is very useful when one wants to

analyze involving multiple nodes in many metabolic processes, distinguishing the effect of different processes between nodes and identifying which nodes represent the crossing points. If individual nodes do not show any coloration after clicking, this identifies certain components of a path, or group, that a specific activated process does not influence. The relationships that determine the coloring of the nodes depend on the knowledge base that STRING organizes for a specific network by extracting data and information from the scientific literature in PubMed.

3. Results

3.1. Hub Data of Human Liver during COVID-19

As mentioned in the Introduction, we carefully selected 11 projects [11,13–22] out of the 18 projects identified in the scientific literature between 2021 and 2023. These papers deal with characterizing hepatic metabolic processes that are viral targets in patients affected by COVID-19. The distinguishing feature of these projects is using different techniques to conduct bioinformatic analyses on profiled patient genes. In particular, the author studied the hub genes that coordinated the metabolic activities of the human liver during COVID-19 infection. They were considered as potential drug targets for this liver pathology. Owing to their high significant rank, hub nodes can also serve as functional seeds to extract related functions from the human proteome. By enriching the nodes that express these functions, it is possible to broaden the functional spectrum of action of the virus, accessing the mechanisms used by SARS-CoV-2 to manipulate human proteins and metabolic processes, as well as information on the molecular strategy adopted. The surprising discovery is that the hub nodes highlighted by these projects are too numerous and different from each other (Table 1). Since they concern the same disease and the same virus, we should have a set of similar hub genes that control the viral strategy by inducing dysregulations in metabolic processes, but we could also come across hub nodes that coordinate normal metabolic activities (housekeeping activities).

Table 1. Hub genes found in the liver by different scientific projects during COVID-19 (2021–2023).

Article Title	HUB Genes
Demonstration of the impact of COVID-19 on metabolic associated fatty liver disease by bioinformatics and system biology approach [13].	SERPINE1, IL1RN, THBS1, TNFAIP6, GADD45B, TNFRSF12A, PLA2G7, PTGES, PTX3, and GADD45G.
Comprehensive DNA methylation profiling of COVID-19 and hepatocellular carcinoma to identify common pathogenesis and potential therapeutic targets [14].	MYLK2, FAM83D, STC2, CCDC112, EPHX4, and **MMP1**.
Exploration and verification of COVID-19-related hub genes in liver physiological and pathological regeneration [11].	ASPM, **BUB1B**, CDC20, CENPF, CEP55, KIF11, KIF4, NCAPG, NUF2, NUSAP1, PBK, PTTG1, **RRM2**, TPX2, and UBE2C.
Systems biology approach reveals a common molecular basis for COVID-19 and non-alcoholic fatty liver disease [NAFLD] [15].	**IL6**, **IL1B**, PTGS2, **JUN**, **FOS**, ATF3, **SOCS3**, CSF3, NFKB2, and HBEGF.
To investigate the internal association between SARS-CoV-2 infections and cancer through bioinformatics [16].	MMP9, **FOS**, COL1A2, COL2A1, DKK3, IHH, CYP3A4, PPARGC1A, MMP11, and APOD.
Target and drug predictions for SARS-CoV-2 infection in hepatocellular carcinoma patients [17].	Upregulated, *PDGFRB*, MMP*14*, VWF, CD*34*, NES, MCAM, CSPG*4*, **MMP**1, *SPARCL1*, and *MMP10*. Downregulated, ***IL1B***, *S100A12*, *FCGR3B*, *CCR1*, *S100A8*, *CCL3*, ***CCL2***, *CCL4*, *CLEC4D*, and *LILRA1*.

Table 1. *Cont.*

Article Title	HUB Genes
Bioinformatics analysis reveals molecular connections between non-alcoholic fatty liver disease [NAFLD] and COVID-19 [18].	ACE, ADAM17, DPP4, TMPRSS2 and NAFLD-related genes such as **TNF**, **AKT1**, MAPK14, **HIF1A**, SP1, and **IL10**.
Organ-specific or personalized treatment for COVID-19: rationale, evidence, and potential candidates [19].	**CCL2**, **CCL5**, CXCL10, HAO2, BAAT, and *SLC27A2*.
Differential Co-Expression Network Analysis Reveals Key Hub-High Traffic Genes as Potential Therapeutic Targets for COVID-19 Pandemic [20].	**IL6**, IL18, **IL10**, **TNF**, SOCS1, **SOCS3**, ICAM1, PTEN, RHOA, GDI2, SUMO1, CASP1, IRAK3, ADRB2, PRF1, GZMB, OASL, **CCL5**, HSP90AA1, HSPD1, IFNG, MAPK1, RAB5A, and TNFRSF1A.
A systems biology approach for investigating significantly expressed genes among COVID-19, hepatocellular carcinoma, and chronic hepatitis B [21].	ACTB, ATM, CDC42, DHX15, EPRS, GAPDH, **HIF1A**, HNRNPA1, HRAS, HSP90AB1, HSPA8, **IL1B**, **JUN**, POLR2B, PTPRC, RPS27A, SFRS1, SMARCA4, SRC, **TNF**, UBE2I, and VEGFA.
Identification of Key Pathways and Genes in SARS-CoV-2 Infecting Human Intestines by Bioinformatics Analysis [22]	**AKT1**, TIMP1, NOTCH, CCNA2, **RRM2**, TTK, **BUB1B**, KIF20A, and PLK1.

Note: In bold red, hub genes found in common between different projects.

From these papers, we have collected 142 hub nodes of the liver cell landscape found connected to COVID-19, of which 21.12% comprise a group of 30 genes in common between different projects, while all the others are different. One hundred and twenty-six hub genes remain after removing those in common. We show this gene list in the Supplementary Materials as Table S1. Barabasi's research showed that biological networks exhibit scale-free properties, with a few genes controlling multiple connections within different functional modules, while most genes have only a few connections [43,44]. It is rather suspicious that the same tissue has a metabolic network operated by such a disproportionate number of hub genes during viral aggression. This suggests heterogeneity of networks. The differences in databases used to extract relationships are a common cause of conflicting results [45,46]. The relationships between the virus and the host occur at the molecular level, through protein interactions. These interactions occur between viral proteins and human proteins and are determined by both human defensive strategies and viral attack strategies. Therefore, it is likely that hub nodes unrelated to the pathology have also been identified. To understand how and why, we applied a biological protocol that involves identifying the real physical relationships established between the nodes that implement the liver network, with no a priori knowledge of the computational protocols. The fundamental biological events between virus and host drive these interactions, thus necessitating a biological evaluation of each individual interaction (see Materials and Methods for details).

Considering the ongoing SARS-CoV-2 pandemic, BioGRID implemented a project called the BioGRID COVID-19 Coronavirus Curation Project (https://thebiogrid.org/project/3) accessed on 1 July 2023. BioGRID is a biomedical interaction repository with experimental data compiled through curation [35]. BioGRID has accumulated fundamental experimental data supporting the role of SARS-CoV-2 in human infection. This project collected comprehensive datasets of all known physical interactions between the proteins of the human proteome and those of SARS-CoV-2. In the purification processes of these proteins, researchers used physical methods such as affinity capture–MS and proximity label–MS and curators of BioGRID have selected and classified both interactors and physical interactions into various levels of statistical significance. This is because some interactions may be random because the laboratory method does not reproduce the cellular environment. Indeed, breaking cells to favor bait–prey interaction also allows for random encounters that do not happen. Today we have over 30,000 interactions (as of July 2023) from the human proteome when its proteins interact one-to-one with the 31 viral proteins of SARS-CoV-2. These interactions possess an unparalleled quality, characterized by their non-redundancy

and high-confidence interactions occurring at a rapid rate, showed by score values obtained through statistical filtering, as determined by Significance Analysis of INTeractome (SAINT) express version 3.6.0 [36].

We have obtained a dataset comprising the entire viral genome (31 proteins) and its interactions with the human proteome. With it, we have created a unique database of the human–virus relationships to search for physical/functional interaction between a viral protein and a human protein. Using our proposed conceptual application framework, we can gain a good understanding of the molecular mechanism of a viral infection. A similar approach has already helped researchers recognize targeted viral complexes of five common human viruses [47]. This recognition is based on biological information. Because of its small genome, a virus must reach maximum performance in interfering with the functional processes determined by human cellular proteins aimed at ensuring normal organic homeostasis. The virus learns over time to implement its attack strategy on specific animal targets by evolutionarily studying the structure of the target proteins. Many viruses use proteins containing large segments of intrinsic disorder [48]. The key to interaction lies in each interaction having specific and well-defined structural foundations, no matter how transient they are. To obtain this knowledge, the virus employs lengthy periods of co-evolution, parasitizing humans or similar species [49]. Therefore, if an interaction is present in this peculiar archive, it means that it has a strategic value of attack or defense, for the virus and for humans. The database also searches for multiple interactions of a human protein with different viral proteins.

Therefore, prioritizing the characterization of the 126 hub genes is an important issue. They should represent the highest-ranking genes, most affected by the virus, and that are therefore optimal to use as functional seeds. This should facilitate identifying genes associated with the pathology and genes involved in normal metabolic regulation, but also uncertified genes included in networks with no experimental certainty. STRING uses many standardized databases [45] as a source of data and information for calculating network models. It produces a detailed analysis of all the scientific articles underlying each single interaction and also corroborates the models calculated with biological analyses, such as GO or KEGG, and with structural analyses using systems such as UniProt. Using STRING, we can manage six data channels that parametrize the network calculation differently and are influenced by various confidence levels. In this way, we can modulate results with very different parameters of reliability, origin, and statistical significance.

On STRING, we inputted the 126 hub genes as functional seeds to extract their relationships from the entire human proteome. These genes, decoded by STRING, should interact to form a protein–protein network model showing compact sub-graphs. Therefore, we left the six channels open to make the most out of each source, but we set the interaction score to 0.900. As STRING networks have a lot of low-scoring interactions, if we want to limit their number per protein, we should use a filter. We used the highest confidence score cut-off to limit the number of interactions to those that have the highest confidence and then are more likely to be true positives. By implementing this strategy, we can narrow down the information only to our input proteins and their network pattern.

3.2. Comprehensive Liver Interactome during COVID-19

The graph in Figure S1 of the Supplementary Materials shows numerous nodes that are not connected (31%). A significant number of the remaining elements do not form a compact and connected graph, with only a portion exhibiting connectivity. This is an indicator of poor functional connectivity, but it also says that many of these hubs may not possess the basis of significant experimental certainty. Manipulating genomic data in the pipeline, from input to extracting functional properties of the network, suffers from a lack of accurate data and an indifference for control over know-how. This makes it impossible to carry out any robust analysis, because the disconnected nodes make any topological analysis or functional consideration unreliable [50–53]. To overcome these shortcomings, we can extend the interactions by setting an enrichment of our network

with new interaction partners (seeds), always depending on confidence value. This allows us to know whether the input shows evidence of statistical enrichment for any known biological function or pathway. The various external databases, including Gene Ontology, KEGG pathways, UniProt keywords, PubMed publications, and others, which annotate the STRING maps, can provide considerable help. The STRING enrichment method retrieves functional enrichment for the set of input proteins. This will show which input protein has enriched terms and describe each term with all its annotations, providing only answers with FDR =< 0.05. Regarding publications, STRING extracts all available scientific texts from PubMed to cover the maximum knowledge about each interaction, also including full-text articles. Figure S2 (see Supplementary) shows the network of Figure S1 implemented with 500 first-order (direct) nodes and 500 second-order (indirect) nodes. Despite its compactness and size, the resulting graph still shows some unconnected nodes. We removed the 15 unconnected nodes (APOD, BAAT, CCDC112, CSPG4, CYP3A4, DKK3, EPHX4, HAO2, MMP11, NES, PLA2G7, SLC27A2, SPARCL1, STC2, and UGT2B7) using an appropriate tool present in STRING to ensure a connected network. Pruning also has the aim of minimizing non-informative enrichment. As a result, we still have 111 residual original hub proteins within the final network, which suggests that there are enrichments consistent with the functional seeds used. In Table S2, we report the list of the 111 remaining hub nodes. It is also important to note that STRING in all the calculated networks has always used data and information extracted from no less than 10,000 scientific articles from PubMed (downloadable), which have generated a specific knowledge base for interactions used in the calculation. By employing a sequential cleaning approach, we can gain precise information and data, which is ensured by the exceptional dependability of each individual interaction among nodes, unveiling their authentic biological credibility.

The enrichment produced a network that includes all principal human proteins in liver tissues during COVID-19. According to STRING, the network shows 7313 functional associations with biological processes spanning 14 categories. A set of 2344 biological processes (GO), 195 KEGG pathways, and 960 reactome pathways characterize the breadth of functional activities. This network appears very well organized and contains all those functional relationships that also involve the original hub proteins. The compact groupings of certain nodes suggest molecular complexes, even very large ones. We can see these molecular complexes in the peripheral areas of the network. They operate as metabolic nano-machines that carry out specific molecular processes [54,55]. For example, the sub-graph at the bottom left is rich in proteins of the Splicing Factor 3B complex that, together with other 17S U2 small nuclear ribonucleoprotein particle (snRNP) components, may play a role in the spliceosome during the selective processing of microRNAs (miRNAs) [56]. This sub-graph also collects many of the proteins involved in transforming molecules of precursor messenger RNA (pre-mRNA) into mature mRNA. Including this complex is not random because RNA splicing is among the major downregulated proteomic signatures in COVID-19 patients [57]. Certainly, the virus needs to manipulate the host splicing machinery to its advantage to control the production of its proteome [58]. In fact, going back along the periphery of the network, we encounter compact sets of genes involved in all phases of cellular translational processes and the entire ribosomal complex, just to mention the most important. At least in the liver, these appear to be the most obvious targets of SARS-CoV-2. EXCEL FILE S1 reports all the nodes of the interactome in Figure 1 with their degrees. These nodes also include all the remaining original hubs (111 nodes). In EXCEL FILE S1, we can also note a few dozen high-ranking genes, all specific for the various phases of the cytoplasmic translation processes. However, before proceeding with other observations, we have reported in EXCEL FILE S2 all 26,990 interactions relating to the interactome in Figure 1. The file also reports the sources of each single binary interaction and the combined score. The interest in this file produced by STRING is because it shows (in red) the quantitative impact of the component deriving from the experimental data alone on the combined value of the score. Thus, this file is useful as a reference in evaluating each individual interaction for the score of 0.900 (highest confidence) we have always used.

As these results show, even for a binary relationship with a score of 0.900, the experimental certification that makes it certain can often be missing, thus introducing serious and invisible anomalies into the graph. We then processed in our SARS2-Human Proteome Interaction Database (SHPID) each single protein of the entire interactome (1111 nodes) to find out which viral proteins had interacted with the network proteins, as well as with the remaining original hub proteins. Some of these proteins no longer exhibit the high-connectivity characteristics that were crucial when they were designated as hubs in the original papers. For example, hub nodes like MCAM, LILRA1, GDI2, COL2A1, TNFAIP6, or PTX3 now have low ranks. What happened reveals that their COVID-19-associated high functional rank disappeared because their value was likely highly inflated by the frequency of studies because of their relevance in diseases or their functional importance in the cell or because they are poorly characterized. A quick check using EXCEL FILE S2 highlighted the widespread lack of valid biochemical and biophysical experimental data for these proteins, meaning that they did not provide adequate evidence for the functional hypotheses in which they had been implicated. Despite the experimental difficulty, we observe in this interactome that proteins localize to specific molecular complexes within a defined range of modules.

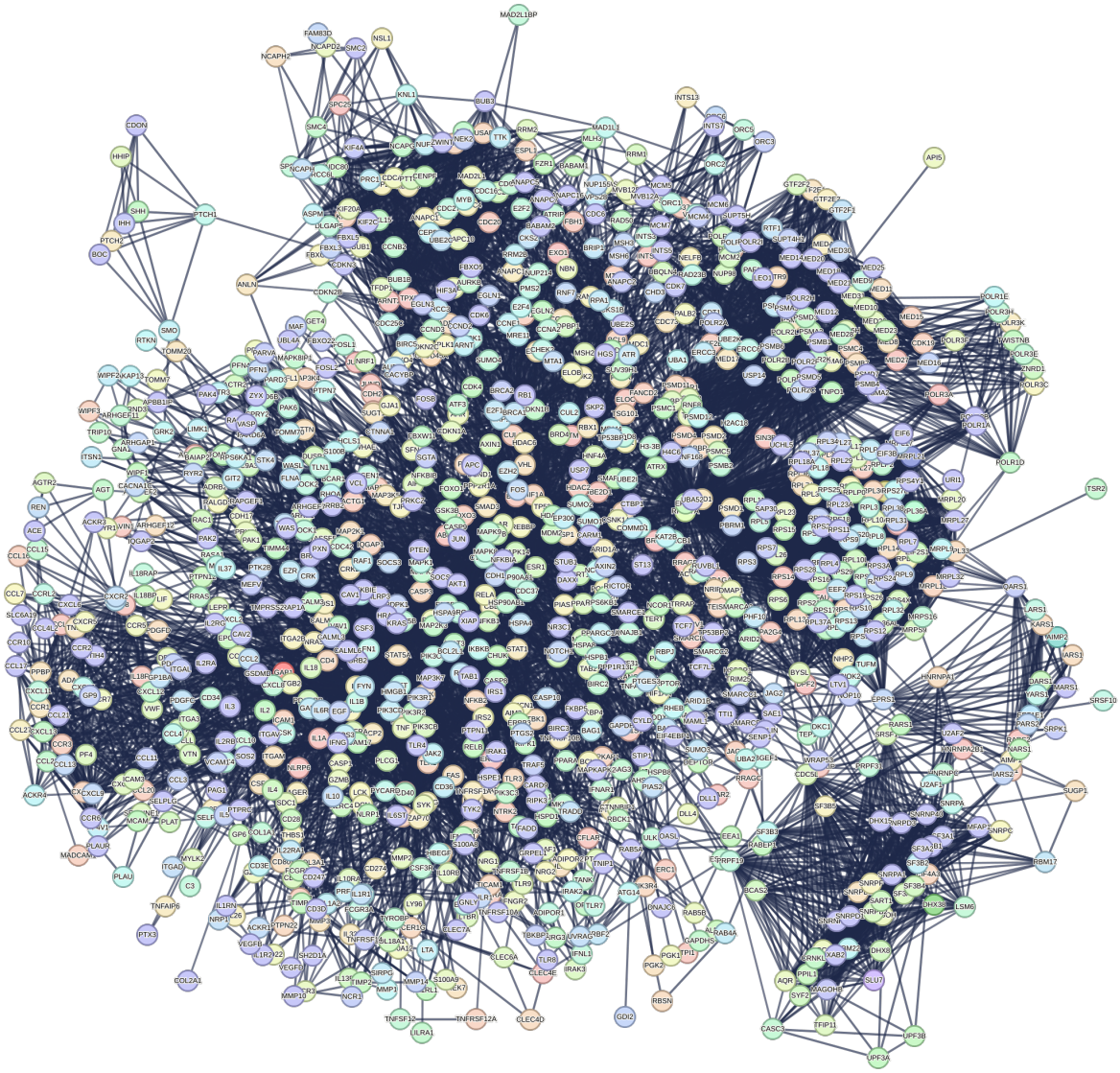

Figure 1. Comprehensive interactome of liver tissue proteins during COVID-19. STRING calculated the graph through enrichment, using as seeds the set of 111 hub proteins obtained after pruning.

We enriched this network with 500 first-order (direct) nodes and 500 second-order (indirect) nodes. Settings: interaction score of 0.900 (highest confidence); all six channels open. Network parameters: number of nodes, 1111; number of edges, 13,494, while expected statistical number is 8838; average node degree, 24.3; avg. local clustering coefficient, 0.623; PPI *p*-value, $<1.0 \times 10^{-16}$; network diameter, 7; network density, 0.022; network heterogeneity, 1.030; network centralizations, 0.128; connected components, 1. (Topological parameters calculated by Cytoscape).

3.3. Metabolic Stress Related to COVID-19 in the Liver

EXCEL FILE S1 shows that protein RPS27A, with a degree of 161, serves as the primary hub. The original hub node list (refer to Table 1) also contained RPS27A. One alias of RPS27A, Ubiquitin-40S Ribosomal Protein S27a, suggests its function as a conserved protein responsible for directing cellular proteins toward degradation by the 26S proteasome [59]. Thus, its role in the liver holds significance. We also know RPS27A plays a significant role in the progression of various human cancers, including HCC [60]. Its landscape of action during viral infection of the liver is interesting. Investigations of SARS-CoV-2 infection have shown large-scale chromatin structural changes because of metabolic stress [61,62]. In situations of oxidative stress [63], induced by phases of the viral cycle [64,65], both oxidizing agents and the need to signal this stress, as well as variations in sensitivity to oxygen, have highlighted the importance of HIF in signaling [66]. These effects are a common feature of both tumors and COVID-19 [67,68]. The shift from the TCA cycle to glycolysis requires cells to upregulate multiple glycolytic enzymes, which are less energetically efficient. One of the transcriptional regulators involved in the response to oxidative stress is HIF1A [69], which remains inactive in normoxic conditions because of its interaction with HIF1AN, an oxygen sensor that hinders interactions with other transcriptional co-activators. SIRT1 serves as an energetic sensor [70], connecting transcriptional regulation to intracellular energetic demands, while TP53BP1 acts as a p53-binding protein, participating in the response to DNA damage.

In tumor progression, the stressful events described affect the p53 protein. The role of p53 (gene TP53) is to inhibit proliferating cancer cells through cell cycle arrest [71]. Therefore, it normally performs a protective cellular action. The main cellular antagonist of p53 is MDM2, as it triggers the degradation of p53 [72] and supports cancerous growth. MDM2 and p53 establish a feedback loop to preserve balance, complemented by involving RPL11, a ribosomal protein that inhibits MDM2 and enhances p53 stabilization and activation in normal conditions [73]. Therefore, RPL11–MDM2–p53 form an axis regulated by RPS27A [73]. When activated by cellular stress phenomena, RPS27A hinders the interaction between RPL11 and MDM2, promoting the degradation activity of p53 through the catalytic activity of free MDM2, thus starting the oncogenic process. Hence, this system of proteins works as a sensor and regulator of cellular stress, acting on p53 and RPS27A to regulate their specific activity.

Figure 2 demonstrates the influence of DNA damage and oxidative stress on these same metabolic players during COVID-19. By highlighting the proteins involved in these processes through a tool that colors the nodes specifically involved (refer to Materials and Methods for further information) we can identify them within the liver protein interactome, also visualizing their role and functional relationships. Table 2 shows the activated biological processes, their statistical value, and the colors of the nodes in the network.

The analysis of Figure 2 reveals that RPL11 and RPS27A are not implicated in the pathways through which cellular stress is detected and transmitted to TP53 and MDM2. These two proteins are not colored; thus, they do not display any stimulation from their interconnected nodes. The non-involvement of RPS27A also suggests that RPL11 continues its activity of blocking the biological function of MDM2 towards TP53. This analysis hypothesizes an activity of TP53 in protecting liver cells by interfering in viral action. Only data from laboratory experiments can offer certainties, even though clinical observations of mild liver damage appear to corroborate this hypothesis. However, EXCEL FILE S2 shows that the experimental component of all the interactions highlighted in Figure 2 and used

to evaluate the hypothesis on the functional activity of TP53 during infection is very high for each protein, so the interactions all rely on a solid experimental basis, which strongly supports this conclusion.

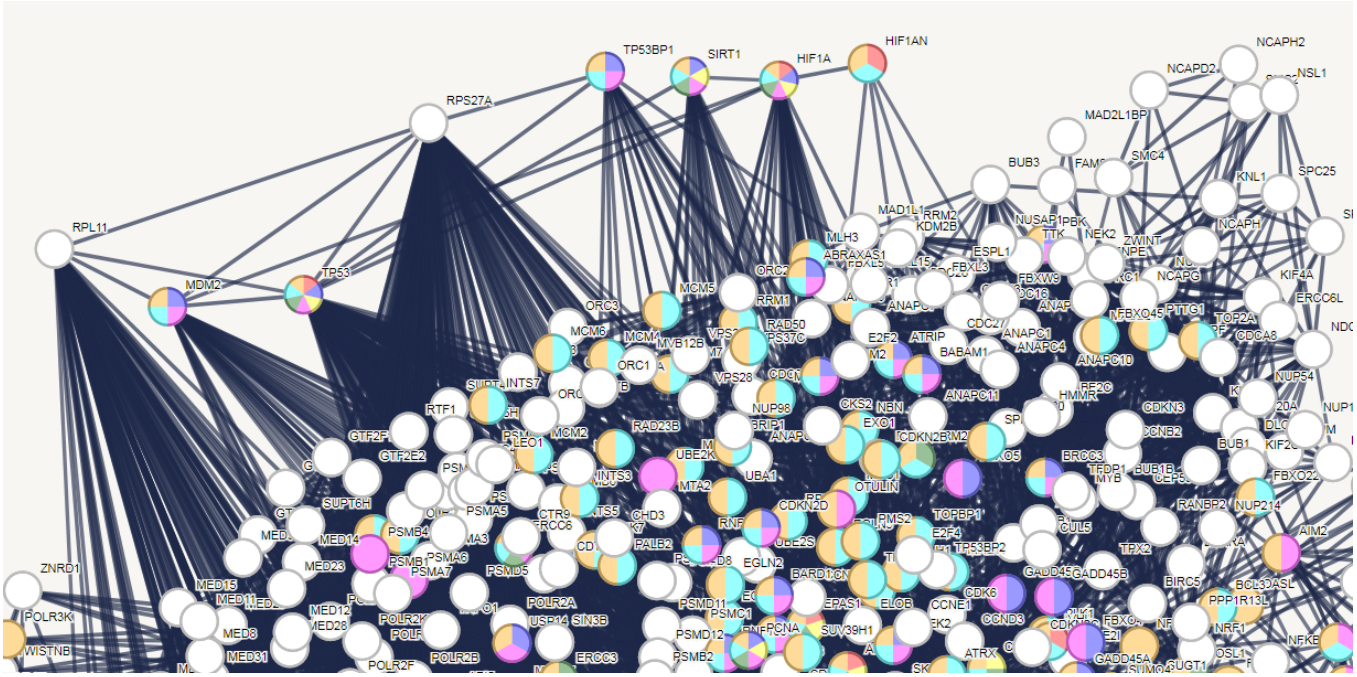

Figure 2. Role of TP53 (p53) and RPS27A in liver infection by SARS-CoV-2. The network is that of Figure 1 and the nodes at the top left have been carefully extrapolated to highlight both the mutual relationships and the abundance of functional connections with the central core of the network. The degree for each single node is RPL11, 104; MDM2, 45; TP53, 133; RPS27A, 161; TP53BP1, 23; SIRT1, 26; HIF1A, 35; HIF1AN, 5. The colors of the individual nodes show the type of metabolic stress (DNA damage and/or hypoxia) induced by COVID-19 in the liver. The biological stress processes (GO) activated are those shown in Table 2.

Table 2. Biological processes related to COVID metabolic stress in the liver.

GO-Term Biological Process	Description	*p*-Value	Node Color
GO:0043620	Regulation of DNA-template transcription in response to stress	1.90×10^{-3}	
GO:0080135	Regulation to cellular response stress	9.77×10^{-37}	
GO:1900407	Regulation of cellular response to oxidative stress	8.69×10^{-5}	
GO:0034599	Cellular response to oxidative stress	1.93×10^{-11}	
GO:0080134	Regulation of response to stress	7.11×10^{-67}	
GO:0006979	Response to oxidative stress	2.99×10^{-12}	
GO:0033554	Cellular response to stress	2.98×10^{-45}	
GO:0006950	Response to stress	5.86×10^{-85}	

3.4. The Reverse Engineering Actions

EXCEL FILE S3 reports all the liver proteins that interact with the viral proteins. Only 51 proteins (in red) of the original hubs interact with the virus. In our experimental conditions, the human proteins interacting with the 31 viral proteins are only 626 out of 1111 proteins (56%). They originate 2680 SARS-CoV-2–host interactions (roughly 20% of the total) of which only 134 can actually be null. These interactions include most of the proteins involved in the translational processes that control protein biosynthesis. In particular, the virus takes possession of the ribosomal system and all the supporting protein complexes to control and promote the biosynthesis of its proteins. This result supports the idea that viruses target high-ranked proteins and proteins crucial in certain biological processes [74]. Several authors have already noted this remarkable ability of individual SARS-CoV-2 proteins to interact with many human proteins, making therapeutic and pathobiological observations [75–77].

There is a notable difference in action between DNA and RNA viruses. Scientists classify viruses according to their DNA or RNA genome. DNA viruses replicate using DNA-dependent DNA polymerase. RNA viruses exhibit greater heterogeneity, especially with ssRNA (+) viruses like coronaviruses. The genetic material of these viruses is very similar to a mRNA. Compared to the genomes of DNA viruses, RNA viruses have smaller genomes that encode fewer proteins and can undergo rapid and direct translation within the host cell. The proteins of RNA viruses have developed a strategy by interacting with host proteins through specific protein-binding motifs. In fact, RNA viruses attacking with few proteins need them to have as multifunctional a capacity as possible. Therefore, we expect RNA virus proteins to possess the capacity to interact with multiple molecular partners. This ability to multitask implies quite specific evolutionary structural adjustments. Indeed, RNA viruses encode proteins characterized by many binding interfaces, but physically with smaller binding surfaces, to hit a greater number of cellular targets [78,79]. Another structural feature to achieve efficient multitasking is to have various segments of intrinsically disordered structure along the protein sequences that are very suitable for expressing multiple, even uncorrelated, activities [80,81]. We could say that the proteins of RNA viruses have had a specialized evolution to develop very peculiar biophysical characteristics. It is widely acknowledged that viral non-structural proteins engage in interactions with host cell proteins, resulting in the formation of replication complexes [82].

Asserting that viral proteins attack human proteins needs quantitative validation and specific information regarding the proteins involved. This question has a particular meaning. In all protein databases, as we have already pointed out, the spatio-temporal characteristics of the archived proteins are missing. Multiple participants hinder the reconstruction of events. While the interaction between many molecules is a recognized concept, the precise mechanisms, meeting sites, timing, and frequency remain elusive. We have limited knowledge in providing mechanistic information about the targeted complex.

3.5. Individual Human Proteins Interacting with Many Viral Proteins and Their Distribution Graph

In EXCEL FILE S3, we can see that some human liver proteins interact with many viral proteins. It is a known fact that multiple viral proteins can target specific human proteins [83]. These interactions described in EXCEL FILE S3 could be a resource for researchers aiming to identify important specific host–virus interactions in the dynamics of disease transmission [84], in particular, to describe the viral diversity associated with different hosts and different tissues, as well as detect shared associations useful for identifying with whom, where, and how they are shared [83,84]. However, some authors report that, in viral infections, the most common ratio of protein–protein interactions between virus and host is 1:1 [85]. Viral proteins, as well as human proteins, are integrated and interact in a specific functional context. This explains much of the binding specificity between proteins. However, even in the best-case scenario, only a handful of viral proteins could interact with a single human protein. This limitation arises from the physical impossibility of locating

suitable binding surfaces on a single molecule and the potential electrostatic repulsions and structural constraints caused by proximity on a crowded structure. In the absence of temporal data on the frequency and specificity of these attacks, we can reasonably think that this massive attack is likely directed towards the entire ribosome and its ancillary complexes, of which the targeted protein is a component, given that the most targeted proteins are the ribosomal ones. But this hypothesis also has another side. It shows the total lack, even in the best databases, of the spatio-temporal characteristics relating to individual human proteins. Given the unlikelihood of crowding on a single protein, the attack is more likely to be sequential, i.e., at different times. A comprehensive understanding of human biology, and that of other living beings, requires acknowledging the dynamic nature of metabolism.

Table 3 shows the human proteins most attacked by viral proteins in the range 12–20. Its main purpose is to showcase the different levels of affected human proteins, both high and low. The degree of each protein (see EXCEL FILE S1) is in brackets. A high degree is because the majority are proteins organized into complexes.

Table 3. Human proteins subjected to multiple attacks by SARS-CoV-2 proteins.

Human Protein	Number of Interacting Viral Proteins **
RPL18A * (84)	20
RPL13 (84)	19
ALDOA (4), CDC42 (52), EIF2S1 (45)	18
RRM2B (3)	17
RPL13A (98), RPL21 * (87), RPL30 * (85)	16
PSMC1 (30), RPL26 * (96), RPL7A (85), RPL (9)	15
BUB3 (19), RPL7 (95), RPL8 (95), RPS24 (90), RPS6 (93), RPS9 * (102), SNRPD1 (38), SRC (97), STIP1 (12)	14
BAG2 (7), RAC1 (11), RPL12 (93), RPL27A (85), RPS27L (82).	13
EIF6 (46), MCM7 (20), HYOU1, PTGES3 (23), RPL27 (84), RPL13 (84), RPL35A (84), RPS10 (87), RPS11 * (108), RPSA (99).	12

Note: * Proteins marked with an asterisk also interact with ORF1ab. ** For more extensive details about interactions, see EXCEL FILE S3.

That some human proteins interact with many viral proteins presupposes many shared structural motifs. But this also suggests that viral motifs in their evolution must gain host-like mechanisms to be successful in invasion. This supports the observations that conformational flexibility, spatial diversity, abundance, and slow evolution are the characteristic features of the human proteins targeted by viral proteins [74]. Viral proteins mimic host-binding surfaces of domains to interact with human proteins, which occur through domain–motif interactions. In EXCEL FILE S3, we can also observe that the interacting viral proteins are not only non-structural proteins (NSPs) and there is also a significant presence of accessory proteins. However, viral proteins intervene in large numbers, targeting mostly the proteins of the ribosomal system. This allows the virus to take control of protein biosynthesis and redirect it towards the synthesis of the viral genome and its own proteins. That many viral proteins attack one host protein also means that many of them have mimicked the same human motif. In addition, we must consider an average of around 47% of disordered segments in coronavirus proteins [86,87]. This favors attacks on specific cellular targets of the host. An interesting discovery is that among the viral proteins that interact with ribosomal proteins (RPL18A, RPL21, RPL30, RPL26, RPS9, and RPS11) there is also the long viral polypeptide ORF1ab. Since ORF1ab is certainly not a target to be blocked but is the viral polypeptide that must be translated, the asterisked proteins mentioned above could represent points of structural contact of the viral protein

ORF1ab with the human ribosome. In fact, some of them (RPL18A, RPL21, RPL30, and RPL26) are specific components of the large ribosomal subunit, the complex responsible for peptide chain elongation and the synthesis of proteins in the cell, while RPS9 and RPS11 are components of the small ribosomal subunit as part of ribosomal process, which couples processing steps of RNA folding and RNA cleavage [88,89]. Most ribosomes end translation at a stop codon present in the first stem of the pseudo-knot. Meanwhile, coronavirus protein synthesis employs regulatory mechanisms, such as ribosomal frameshifting, promoted by a conserved stem-loop of RNA that forms a promoting pseudo-knot structure [90]. Ribosomes stall at the pseudo-knot and undergo a -1 frameshift at the slippery sequence, leading to translating ORF1ab fusion polypeptide [91,92]. In coronavirus, this phenomenon allows the virus to encode multiple types of proteins from a single mRNA, compacting the information. In this way, virus translation dominates host translation because of high levels of virus transcripts.

In Table 3, we also find the involvement of lower-degree human proteins that are not ribosomal proteins. Some of them are key because they are involved in crucial metabolic functions of the liver. We report, as examples, ALDOA, RRM2B, BAG2, and HGS. ALDOA is the tetramer of hepatic-type aldolase B that binds to the hepatic cytoskeleton and to actin-containing stress fibers. The presence of disordered segments in the C-terminals favors the possibility of scaffolding and suggests that aldolase can regulate cell contraction [93,94]. RRM2B forms a complex with RRM1 where it plays a key catalytic role in repairing damaged DNA together with p53 and provides deoxyribonucleoides in G1/G2-locked cells [95,96]. BAG2 is a co-chaperone regulator of the HSP70 and HSC70 chaperones. It acts as a nucleotide exchange factor by promoting the release of ADP from HSP70 and HSC70 proteins, triggering the release of the client/substrate protein [97,98]. Finally, hepatocyte growth factor (HGS) is involved in intracellular signal transduction mediated by cytokines and growth factors. It regulates endosomal sorting and plays a critical role in the recycling and degradation of membrane receptors [99–101]. The liver serves as the site of localization for many of these proteins, emphasizing their tissue specificity.

3.6. Distribution of Viral Proteins Interacting with Single Human Proteins

Figure 3 shows the distribution graph of the entire set of human liver proteins (626 proteins) interacting with viral proteins (see also EXCEL FILE S3). Each point on the curve reports the set of human proteins that have the same number of interacting viral proteins. The fit shows that the distribution conforms to a power law, albeit with an R^2 value of 0.5278, suggesting an acceptable fit. This value is at the low limits of reliability and may imply existing heterogeneities in the distribution, which makes the results difficult to explain. This should not be surprising because the distribution reflects the overall structural and functional behavior of the entire set of human proteins with different roles from each other and subjected to sequential functional stress by viral proteins in complex and metabolically differentiated cellular environments. Hundreds of interactions are one-to-one (those on the left side of the curve), while others involve multiple interactions (multi-to-one), to up to 20 viral proteins per single human protein (in the tail). The connectivity distribution in Figure 3 is quite consistent with the power law's prediction of preferential attachment [102]. Thus, our model should show the emergence of a scale-free topology [103] from interaction results. So, if the connectivity distribution follows a power law, then new nodes will have a better chance of connecting to those with already many neighbors because of the preferential attachment rule.

Comparative and evolutionary genomic analyses support the birth of complex structures in the cell that make up organized and complicated cellular nano-machines [104]. Genomics has also shown that parts associate with each other to form integrated systems with modular and hierarchical structures [105]. This organizational process should also be intrinsic in the modeling of liver metabolic reactions that arise from protein–protein interactions. In accordance, complex networks exhibit higher-order organization in connectivity, showing links that can be modulated and modeled using sub-graphs of the network [106].

Some authors have also shown that networks contain within themselves information about the organization of these compact modules (sub-graphs) such as emergence of the protein complexes [106,107]. From the peculiarity of these models emerges an important intrinsic structural characteristic of biological networks, namely hierarchical modularity, i.e., a higher level of organization, the growing mechanisms of which, unfortunately, remain unknown. Researchers have never quantitatively tested these qualitative and observational relationships in real biological interaction networks. Our network model, related to liver tissue, shows human protein complexes strongly involved in viral infection. We believe that the preferences of viral proteins toward the interior of these complexes should reflect the mechanisms used by viruses to manipulate host protein complexes.

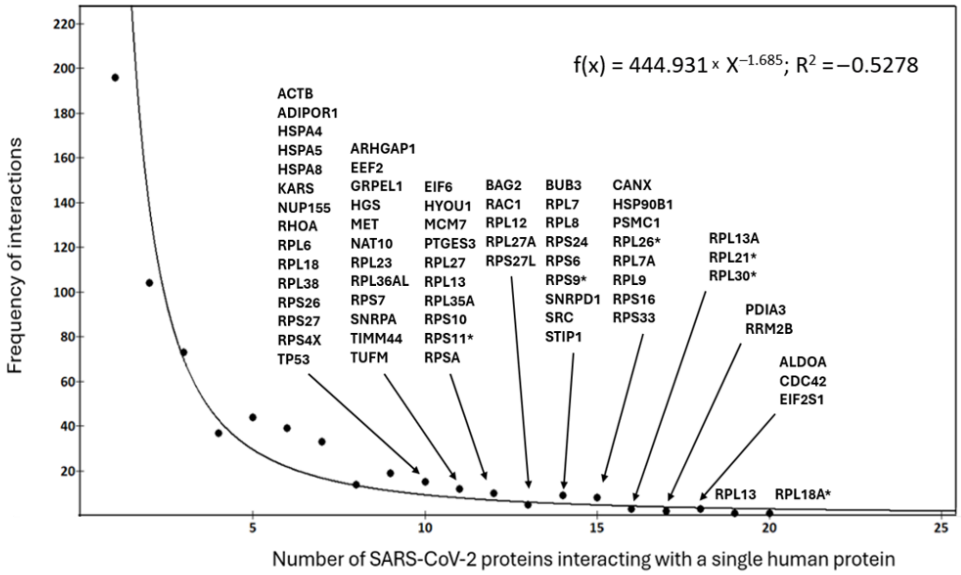

Figure 3. Distribution of viral proteins interacting with single human proteins. The curve is the exponential fit (displayed at the top right). Data calculated from EXCEL FILE S3. The figure also shows the most targeted human proteins (from 10 onwards). The asterisked proteins are those that also interact with ORF1ab.

Based on our collective data, it is evident that the evaluation of virus action should be conducted within the framework of viral preferential attack strategies on intricate protein organizations. However, how viruses manipulate sub-graphs of local host networks, such as human protein complexes, has never been addressed from a topological–computational perspective, preferring to focus on the preferential targeting of viral proteins with hub or bottleneck nodes, despite that no formal definition exists to separate hub proteins from non-hub proteins [12,108].

A systematic analysis of the protein complexes, identified as direct protein–protein targets, has been carried out to discover new drugs [109] or even through bioinformatic approaches [47], almost never considering a topological point of view. In this type of analysis, both local topological aspects of the network and evolutionary ones should contribute, but, to date, discrimination of the topological and functional properties of complex viral targets during an infection is lacking. Our analysis identified compact sub-networks of human proteins targeted by multiple viral pathogen proteins. But what is perplexing is that during the infection, the targeting process of a complex protein system, such as the ribosome, seems to depend on the connectivity of neighboring proteins in the network (because of the preferential attachment, which is a topological parameter). Conversely, the interaction of a viral protein ought to be primarily determined by the likelihood of a physical encounter associated with the decrease in free energy because of binding, exploiting chemical–physical parameters from evolutionary laws.

We can hypothesize, from the analysis presented in Figure 3, that multiple types of interaction activities could compete concurrently. If this is the case, upon closer analysis, we should be able to discern more exponential decays that would better characterize the distribution. In Figure 4 (top), we observe that the degree distribution seems to follow a single power law. However, the fit in the log–log scale indicates that the single power law distribution is at the lower limit to adequately meet or explain the data characteristics. One-to-one and one-to-many interactions behave differently and make the analytical representation heterogeneous when considered together. The bottom graph shows that the distribution, always in the log–log scale, displays two different slopes, unlike what happens when fitting with a single power law. In both fits, the values of R^2 are very good, suggesting a combination of two solutions (or two decays) that are linearly independent. The biphasic distribution suggests the hypothesis that there may be at least two dominant classes of co-existing proteins with differentiated functional responses. One class (in black) should contain human proteins essential for metabolic adaptations following viral infection. These proteins can be under-expressed or lost when pathophysiological conditions induce profound metabolic changes. Proteins belonging to the other class (depicted in red) are essential for critical physiological processes of viruses and hosts but are also essential for the virus to gain energy. Thus, these human proteins, highly expressed, exhibit enhanced resistance to pathological processes that induce functional variability. Depending on the characteristics of the local context, it is possible for all proteins to transmigrate between both classes. In the lower graph of Figure 4, there is on the x axis the transition degree, TD. Its value breaks the distribution into two parts and identifies the boundary between nodes with an interaction degree of less than 12 (in black, made up of proteins that are on average poorly connected) and nodes having a degree greater than TD (in red, composed of evolutionarily older proteins that are on average much more connected). In our analysis, each of these sub-networks follows a single power law degree distribution well, while differing in the value of power law exponents.

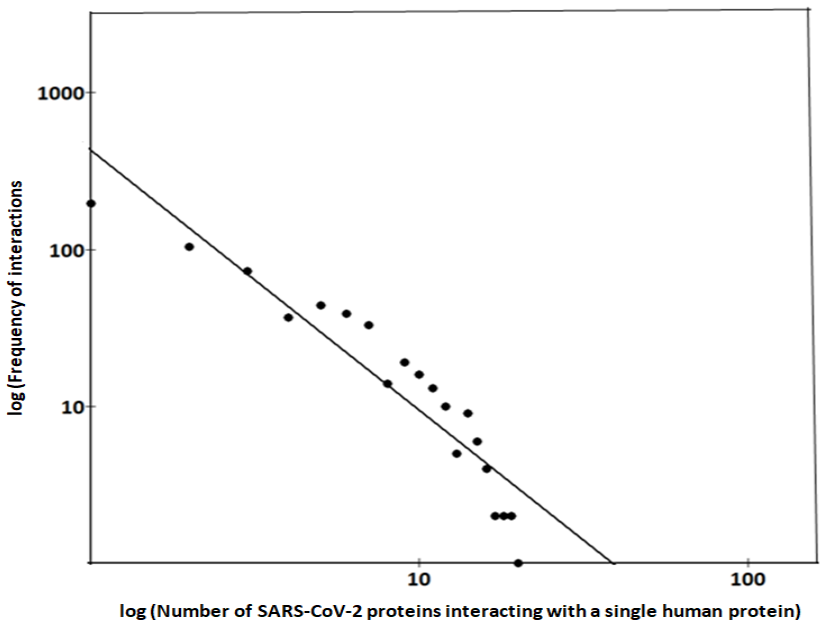

Figure 4. *Cont.*

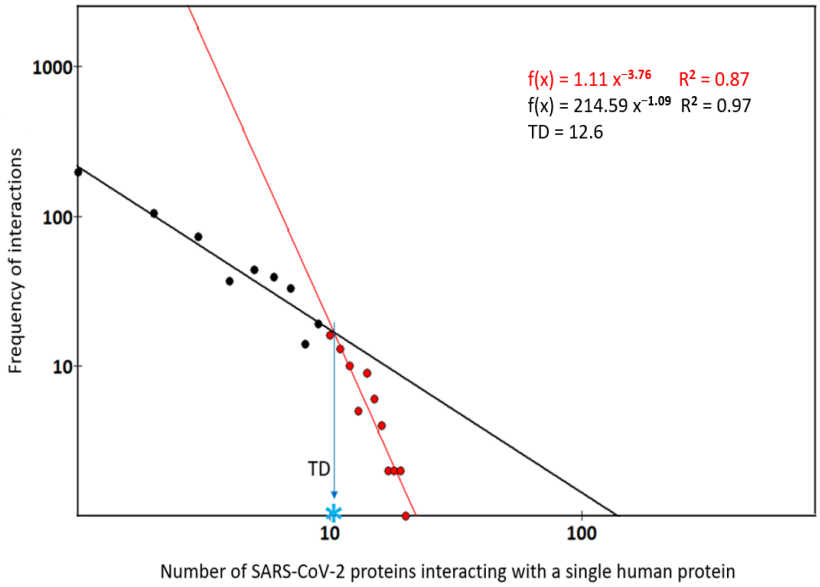

Figure 4. Linear distributions of interacting viral proteins with a single human protein (log–log scales). **Upper figure**—Distribution graph considered as a single power law. Fitting: $f(x) = 431.26\ x^{-1.66}$ and R^2 is 0.3675. **Lower figure**—Biphasic representation of the power law. The graph displays the fitting equations. TD is the transition degree, the estimated point (marked by blue star) at which the slope of the distribution sharply changes. Its value is around 12.

This biphasic model suggests all proteins can gain new interactions with rate (greater slope) and number of interactions (the rich get richer) always increasing, as happens for older proteins (red ones). Proteins can also lose their interactions, both with and without the loss of their connecting partners. It is a kinetic model which through the different slopes reflects the evolutionary behavior of proteins, considering two classes of proteins, one with a rapid action but also with a fast residence time and the second with opposite properties of greater resilience. Both classes adequately describe, both in topological and evolutionary terms, the nature of the biexponential model. The model, in fact, shows a situation in which the oldest proteins, the most conserved by evolution, increase their interactions because of the establishment of new and specific kinetic conditions. Although our results are built on solid foundations of statistics and experimentation, it is important to interpret them with caution due to all the limitations previously described.

3.7. Comprehensive Analysis of Liver Metabolic Activities during COVID-19

To support the structural and functional organizational events previously found for these proteins and the complexes involved, we analyze the data using the many specific databases that STRING maps onto the protein data of calculated networks. Table 4 reports some analyses of biological processes made by STRING on the interactome data shown in Figure 1. The table shows the most statistically reliable results. Although all data used in this study have a high intrinsic significance, analyses on extensive sets, where gene expression variability could also play a fundamental role, must be carefully evaluated. Therefore, in their evaluation, the value of the intensity of the expression of the genes that code for the proteins of a process, contained in the strength parameter (see Section 2.8), was also considered. The results show that the *p*-value (FDR) is important, but the level of gene expression influences its significance. Then, the intensity of the biological action also depends on the intensity of gene expression.

The gene expression depends on cellular signals, but the biological results depend on the phenotype "interpretation" of that information, which is displayed by the synthesis of proteins (and non-coding RNA). Thus, this parameter allows for the definition of a similarity metric between gene expressions, which we can use to reposition and compare biological processes [110,111].

Table 4. Liver Biological processes during COVID-19 infection.

1—Normal biological processes related to nodes certified by reverse engineering in the liver infected by COVID-19				
GO Term Biological Process	**Description**	**P**	***p*-Value**	**Strength**
GO:0019221	Cytokine-mediated signaling pathway	47.50	8.51×10^{-57}	0.82
GO:0002181	Cytoplasmic translation	46.53	2.05×10^{-44}	1.05
GO:0071345	Cellular response to cytokine stimulus	42.97	1.59×10^{-63}	0.68
GO:0033044	Regulation of chromosome separation	37.73	9.62×10^{-36}	1.02
GO:0010965	Regulation of mitotic sister chromatid separation	36.30	8.03×10^{-34}	1.04
GO:0033045	Regulation of sister chromatid segregation	36.20	6.46×10^{-34}	1.02
GO:0051983	Regulation of chromosome segregation	34.60	4.60×10^{-35}	0.97
GO:0030071	Regulation of mitotic metaphase/anaphase transition	33.87	3.68×10^{-32}	1.04
GO:0033044	Regulation of chromosome organization	32.37	3.03×10^{-39}	0.82
GO:0007346	Regulation of mitotic cell cycle	32.25	1.18×10^{-46}	0.70
GO:1901987	Regulation of cell cycle phase transition	30.16	2.98×10^{-42}	0.71
GO:0006412	Translation	29.30	4.59×10^{-40}	0.72
GO:1901990	Regulation of mitotic cell cycle phase transition	27.66	2.42×10^{-37}	0.74
GO:1990869	Cellular response to chemokine	23.92	8.36×10^{-24}	0.96
GO:0034243	Regulation of transcript. elongat. from RNA polym. II	17.94	5.25×10^{-19}	0.91
GO:0007088	Regulation of mitotic nuclear division	17.50	3.89×10^{-20}	0.85
2—Negative regulation of biological processes related to nodes certified by reverse engineering in the liver infected by COVID-19				
GO Term Biological Process	**Description**	**P**	***p*-Value**	**Strength**
GO:0043069	Negative regulation of programmed cell death	18.94	2.65×10^{-36}	0.52
GO:0043066	Negative regulation of apoptotic process	18.31	7.95×10^{-35}	0.51
GO:1901988	Negative regulation of cell cycle phase transition	15.97	3.11×10^{-22}	0.71
GO:0045786	Negative regulation of cell cycle	15.25	1.63×10^{-24}	0.63
GO:0010948	Negative regulation of cell cycle process	14.98	1.08×10^{-22}	0.68
GO:0009892	Negative regulation of metabolic process	14.36	3.19×10^{-43}	0.33
GO:0010605	Neg. regulation of macromolecule metabolic process	14.22	6.61×10^{-41}	0.34
GO:1901991	Neg. regulation of mitotic cell cycle phase transition	13.83	8.82×10^{-18}	0.73
GO:0045930	Negative regulation of mitotic cell cycle	13.33	2.12×10^{-19}	0.69
GO:0031324	Negative regulation of cellular metabolic process	12.03	2.37×10^{-34}	0.35
GO:0060548	Negative regulation of cell death	11.95	1.43×10^{-34}	0.35
GO:2000816	Neg. regulation of mitotic sister chromatid separation	11.88	7.56×10^{-11}	1.0
GO:0045841	Neg. regulation mitotic metaphase/anaphase transition	10.46	2.29×10^{-10}	1.01
GO:2001237	Neg. regulation of extrinsic apoptotic signaling pathway	9.67	5.60×10^{-12}	0.76
GO:0051348	Negative regulation of transferase activity	8.90	1.17×10^{-15}	0.59

Table 4. *Cont.*

3—Dysregulated biological processes related to nodes certified by reverse engineering in the liver infected by COVID-19				
3A—Local Network Clustering (STRING)	**Description**	**P**	***p*-Value**	**Strength**
CL.152	Viral mRNA translation	89.03	7.21×10^{-46}	1.19
CL:159	Viral mRNA translation	55.38	1.06×10^{-45}	1.23
CL:162	Cytoplasmic ribosomal proteins	54.16	1.41×10^{-43}	1.23
CL.143	Viral mRNA transl. and Sec61 translocon complex	53.10	6.93×10^{-47}	1.11
3B—Reactome Pathways	**Description**	**P**	***p*-Value**	**Strength**
HSA-192823	Viral mRNA translation	64.09	2.56×10^{-53}	1.2
HSA-72764	Eukaryotic translational termination	61.79	2.32×10^{-52}	1.18
HSA-72689	Formation of a pool of free 40S subunits	58.97	1.91×10^{-51}	1.15
HSA-72737	CAP-dependent translation initiation	53.73	1.98×10^{-49}	1.09
HSA-1799339	SRP-dependent co-translational prot. targeting to member	53.17	2.20×10^{-48}	1.1
HSA-9679506	SARS-CoV-1 infections	38.58	5.77×10^{-50}	0.76
HSA-9754678	SARS-CoV-2 modulation of host translational machinery	26.18	2.39×10^{-23}	1.12
HSA-9692914	SARS-CoV-1 host interactions	32.98	1.06×10^{-32}	1.03
HSA-9705683	SARS-CoV-2 host interactions	31.14	1.61×10^{-36}	0.86
HSA-9678108	SARS-CoV-1 infection	30.73	1.12×10^{-33}	0.93
HSA-9735869	SARS-CoV-1 modulates host translational machinery	28.19	1.28×10^{-23}	1.22
HAS-9754678	SARS-CoV-2 modulation of host translational machinery	26.18	2.39×10^{-23}	1.12
HSA-9694516	SARS-CoV-2 infections	25.52	1.07×10^{-34}	0.75
HSA-9705671	SARS-CoV-2 activates/modulates innate/adaptative immune responses	11.06	5.57×10^{-14}	0.75
HSA-597592	Post-translational protein modification	7.95	1.28×10^{-22}	0.36
HSA-9772572	Early SARS-CoV-2 infection events	3.68	1.3×10^{-5}	0.72

4—Protein domain characteristics in the liver infected by COVID-19					
4A—Prot. Domains (InterPro)	**Description**	**Count in Network**	**P**	***p*-Value**	**Strength**
IPR036048	Chemokine interleukin-8-like superfamily	29 of 44	15.03	1.11×10^{-14}	1.07
IPR039809	Chemokine beta/gamma/delta	15 of 26	8.03	8.90×10^{-7}	1.01
IPR033899	CXC chemokine domain	12 of 14	7.30	1.54×10^{-6}	1.18
IPR011332	Zinc-binding ribosomal protein	9 of 10	7.01	6.92×10^{-5}	1.2
IPR011029	Death-like domain superfamily	29 of 97	6.01	2.23×10^{-8}	0.72
IPR008271	Serine/threonine-protein kinase, active site	52 of 310	4.21	9.00×10^{-8}	0.47
IPR001875	Death effector domain	5 of 7	3.84	3.10×10^{-3}	1.1
IPR0000488	Death domain	11 of 35	2.81	6.3×10^{-3}	0.74

Table 4. *Cont.*

4—Protein domain characteristics in the liver infected by COVID-19					
4B—Prot. Domains (SMART)	**Description**	**Count in Network**	**P**	***p*-Value**	**Strength**
SM00199	Intercrine alpha family (small cyt/chem CXC)	28 of 42	16.80	5.07×10^{-15}	1.07
SM00252	Src homology 2 domains	22 of 104	3.24	2.5×10^{-5}	0.6
SM00219	Tyrosine kinase, catalytic domain	20 of 88	2.64	2.5×10^{-4}	0.6
4C—Annotated Keywords (UniProt)	**Description**	**Count in Network**	**P**	***p*-Value**	**Strength**
KW-0689	Ribosomal protein	90 of 175	44.83	5.05×10^{-46}	0.96
KW-0687	Ribonucleoprotein	112/278	42.17	4.14×10^{-49}	0.85
KW-0945	Host–virus interaction	148/540	33.03	3.81×10^{-48}	0.68
KW-0747	Spliceosome	50 of 138	16.77	5.14×10^{-20}	0.81
KW-0395	Inflammatory response	56 of 163	16.56	1.73×10^{-21}	0.78
KW-0132	Cell division	88 of 384	14.25	2.31×10^{-23}	0.61
KW-0498	Mitosis	69 of 75	13.43	4.53×10^{-20}	0.65
KW-0131	Cell cycle	137/651	13.13	1.09×10^{-23}	0.57
KW-0647	Proteasome	25 of 52	11.57	2.74×10^{-12}	0.93

The networks representing the clusters are reported in the Supplements as Figures, from Figures S3–S6. While the functional characteristics as Tables, from Tables S4–S7.

The table is split into four sections that show the primary aspects of the metabolic context encountered by the liver during COVID-19. The data are shown in decreasing order determined by the P value. As we note, some *p*-values, despite being remarkably low, are repositioned because of variability in the intensity of gene expression. In the first part of the table (Part 1) we can see that cellular activity is mainly involved in promoting cytokine signaling processes, cellular translation, and the cell cycle. In the second part (Part 2), we have the negative regulations resulting from the viral attack. Surprisingly, one of the main viral activities is to alter the programmed processes of cell death, followed by strong interference to alter the processes of the cell cycle in its various phases. These data suggest a viral activity that aims to implement a systemic spread of intact but infected cells, very similar in result to the spread of cancerous metastases. If we observe the interaction data in EXCEL FILE S3, we can see that the virus attacks proteins of the cellular matrix and cytoskeleton, such as ACTB, ACTR3, FN1, CDC42, COL2A1, COL18A1, ITGA3, ITGA5, ITGAV, FLNA, ACTL6A, ACTR2/3, and others, similar to what the cancer cell does to spread metastasis. Other researchers have noticed similar strategies [112], such as extending particular stages of the cell cycle and managing programmed cell death. Part 3A shows some of the clusters calculated by STRING which show the involvement of the virus in mRNA translation and in ribosomal cytoplasmic proteins. Local STRING network clusters are pre-computed protein clusters derived by hierarchically clustering the full STRING network.

The Supplementary Materials (under Clustering) provide a comprehensive overview of all four clusters of Part 3A, featuring their topological parameters and a GO analysis for each, to facilitate the identification of the metabolic framework of action. Extremely low FDR values characterize all these contexts, demonstrating that the cytoplasmic translational system, including ribosomes, is the most statistically significant virus target.

Part 3B (Reactome) shows the most reliable metabolic pathways that involve extensive virus–host interactions and identifies sets of proteins that also perform the same action as SARS-CoV-1. Part 4 highlights the specific human protein domains targeted by viruses. One interesting aspect is that the presence and incidence (count in the net) of these proteins have been quantized. Many of these domains (Parts 4A and 4B) are involved in the molecular mechanisms of chemokine/cytokine signaling and in the reprogramming processes of programmed cell death. The last part, 4C, shows in which downregulated biological processes we find these domains and in what abundance, including spliceosome-mediated RNA processing. The set of this information is in excellent agreement with that discussed earlier and also opens up other observations. Although our results are built on solid foundations of statistics and experimentation, it is important to interpret them with caution because of all the limitations described. In this study, we did not discuss one-to-one interactions of the proteins of this viral pathogen with other human proteins. The most surprising of these observations (see EXCEL FILE S3), is the large number of one-to-one interactions, that, for instance, characterize the S1 viral protein (spike), which interacts with many individual human proteins involved in different metabolic processes [113].

4. Discussion

COVID-19 involves many cellular biochemical adaptations affecting specific biochemical and physiological pathways that generate profound systemic alterations which are reflected in specific organ adaptations. This justifies a specific study of the alterations generated in the liver by SARS-CoV-2. The study shows the interactions between viral and human proteins involved in molecular and/or biological processes and their consequences because of the infection. To the best of our knowledge, we have presented here the most comprehensive and in-depth analysis of SARS-CoV-2–human PPIs within liver infection by COVID-19.

Our analysis revealed that viral targets are enriched in human protein complexes, such as ribosomes or proteasomes, and results confirm that viral infection affects large protein complexes involved in the human translational system. During the attack, we observed a significant presence of scaffolding and housekeeping proteins among the viral targets. In this way, the virus takes possession of and controls the entire apparatus that manages mRNA translation, blocking similar activities of the host. The strategy is to encourage viral replication. Therefore, understanding the host molecular mechanisms involved in protein–protein interactions (PPIs) controlled by SARS-CoV-2 is crucial for the design of new antiviral strategies, as well as because there are human proteins that could be better targets than viral ones. However, the results show the interactions that are crucial factors for regulating cellular metabolism and survival during stressful times, which have relevance in viral infections for disease progression.

Many pathological features of SARS-CoV-2 in the liver have remained unclear because the underlying molecular mechanisms are unknown (1). Although many host proteins can interact with viral proteins, only some of them are essential for a full infection in a virus-specific manner. The results also show that the biological control exerted by the various human hubs, as reported in the literature, was not always confirmed, nor was it shown which of them physically interacted with viral proteins. The results presented in our reverse engineering approach are all experimentally based because the proteins involved and their specific interactions come from BioGRID. Through a comprehensive collection of all BioGRID one-to-one interaction data, we could filter these proteins, revealing the functional characteristics of those involved in virus–host interactions. Although many host proteins can interact with multiple viral proteins, only some of them were crucial for infection in a virus-specific manner, after filtering out the less significant ones to reduce noise.

The limit of this approach does not lie in the methods used but in the acquisition and representation of tissue information on a spatial and temporal scale, which remains a limit to be overcome technologically. This is the real challenge. Considering the intricacy in

representing the spatio-temporal organization of cells and tissues as metabolic scenarios, our aim has been to choose specific biological processes applicable in real-world scenarios. We extracted from the literature an extensive set of heterogeneous hub data of the liver of infected patients by comparing them with the biological data set of our database and pruning those of low significance. We have shown the accuracy and biological robustness of our conclusions. Next, we evaluated these liver datasets and showed they could detect metabolic patterns of hepatic tissues within COVID-19. Our data showed that inverse engineering can map and reconstruct the metabolic distribution of various biomolecules, providing valuable multimodal insights into coronavirus disease.

From the distribution analysis of the human proteins, used as targets by the viral proteins, we have highlighted that the best fit of the data is the one that provides a biphasic power law. This allowed us to highlight at least two classes of proteins related to two different distributions that consider two operational kinetics of the two classes. The main connection of evolutionarily consolidated proteins is to a resilient class that quickly enhances its connectivity. The second group consists of proteins that are already loosely linked, primarily concerned with pathological aspects and exhibiting slower connectivity growth. Thus, the forces driving this protein behavior are both evolutionary and topological, albeit to varying degrees.

A set of over 33,000 experimental human–virus interactions curated by BioGRID provided the biological basis for each individual interaction. Added to this is that for every single interaction to model the STRING network, we used a score of 0.900. In evaluating key interactions, we have considered the quantitative incidence of the experimental contribution to the value of the combined score using the parametric data reported in EXCEL FILE S2. It is worth considering that only a solid experimental basis can make a protein–protein interaction certain and reliable in the real metabolic world. Recent results show that biases of the experimental procedures used to infer networks can affect the resulting topology [114]. Additionally, we can expect that study bias may affect the sensitivity of experiments, considering that proteins that have been excessively studied are tested more frequently than others [115].

Today, a network can capture functional modules and cellular connectivity processes because proteomic data contain a relational and informational component connected to protein–protein interactions. But the biological events that distinguish a cell, whether normal or infected, represent how the genetic code is executed that triggers one of the many metabolic processes of which a hub node is part or can manage. Therefore, it is very difficult, if not impossible, to distinguish when, how, and with whom a hub node is involved in an altered or normal process. As mentioned previously, the actual activity of a node does not derive from understanding the human metabolic activities in which it seems involved, but from knowledge of the specific spatio-temporal events that involve it. This is because a key node, be it a hub or a bottleneck, is a crossroads through which many pieces of information can pass, even if we do not know which ones and in what order. This constraint currently limits human knowledge, but we will overcome it to enable drawing real conclusions.

This study analyzes in depth some protein–protein interactions between virus and host involving molecular complexes in the cellular system represented by liver tissue during COVID-19. The results allow us to provide an account, albeit approximate, of the mapping of these interactions. SARS-CoV-2 identifies multiprotein complexes with which high biological functions are associated as optimal targets for attack. An advantage for this virus is that, being an ssRNA (+) virus, it has a very rapid cytoplasmic production of viral proteins. The affected multiprotein complexes are RNA splicing, transcription, and translation machineries, but also cell signaling proteins, which function as part of complexes on the order of mega-Daltons and are made of dozens of proteins [114]. With ribosomes and spliceosomes, these complexes reach an even greater molecular weight, because, on average, they comprise 100–300 different proteins, including structural and regulatory RNAs [115]. We should also consider that these complexes, which function as scaffolds

for viral proteins, are also subject to regulation of their function through mediating post-translational modifications. As already noted, we have little knowledge of the dynamics of the information flows that drive events that give rise to molecular phenomena, such as signaling or translation. We do not know PTMs of subunits or information about the structure/function relationships to organize the architecture of these complexes. All this makes any proposal of a dynamic hypothesis on viral strategy murky. However, although we still have a static understanding of metabolic actions, knowing the details involving some key human proteins in these complexes could open a new era in antiviral pharmacology.

One last observation deserves to be noted to conclude this discussion. We found a smaller quantity of important ribosomal interactions associated with RPLs and RPSs, as opposed to the information documented in the BioGRID file concerning the ORF1ab protein. This result, together with the fact that, of the 1111 human proteins of the interactome, only 626 interact with viral proteins, opens considerations on the systemic activity of the virus in various human organs. These results suggest a different viral strategy in different tissues/organs. Many researchers speak of a process of evolutionary adaptation of the virus to humans, favored by its successful propensity to mutate. The mutation rate of the virus genome has been estimated at 1×10^{-3} substitutions per base (30 nucleotides/genome) per year under neutral genetic drift conditions or 1×10^{-5}–1×10^{-4} substitutions per base in each transmission event [116], but, tracking a systematic gene-by-gene comparison analysis with a reference genome (i.e., the first sequence data of a patient from Wuhan in the National Center for Biotechnology Information (NCBI), annotation NC_045512.2), only six of mutations had over 50% frequency in global SARS-CoV-2 up to 2023 (NSP12, S, NSP4, N, ORF9b, and NSP3) [116].

Viral evolution occurs on time scales comparable to virus transmission events and to dynamics that involve many factors [117]. These factors encompass the fluctuation of infected individuals over time, the varying percentages of immune profiles in populations, human mobility, the effectiveness of transmission between individuals, as well as the interplay between viral strains and lineage extinction [117]. The complexity of all this makes it challenging, if not outright impossible, to establish global evolutionary theories through experimental evidence, although it is still workable to have coherent discussions about individual factors of variability. Consequently, numerous hypotheses have emerged regarding the evolution of SARS-CoV-2, including the notion that the virus becomes less virulent [118]. Without going into the merits of these observations and the many existing hypotheses, we note that the sampling of data we collected covers patients scattered around the world who became infected between 2021 and 2023. The genomic profiling focuses on the liver. Thus, our data cover a wide window of the evolution of SARS-CoV-2 in relation to liver tissue [119] and regarding high-ranking proteins (hubs), known to be the preferential target of the virus. Although 22% of them did not meet the experimental requirements to be reliable, we discovered that only 51 of these proteins (refer to EXCEL FILE S3) ultimately played a role in the infection, although many had reduced connectivity. They, through functional enrichment, showed us how remarkable the viral activity was against specific proteins of the entire hepatic cellular translation system. This strategy never changed over 3 years. Checking BioGRID, the interaction data show that ORF1ab also interacts with many other proteins of the human translational system but not in the liver. This suggests a unique and specific viral behavior, i.e., over time, viral methods and proteins attacking the liver showed no significant changes in strategy. It is logical to speculate that a different strategy should be considered in relation to the protein–protein interactions of SARS-CoV-2 in diverse human tissues/organs. The complexity arises when attempting to illustrate this hypothesis, as the data used are sourced from deceased patients, rendering it impossible to distinguish between the systemic response of the patient's phenotype and the effects specifically tied to the organ being examined. We could also find this information in those poorly interacting hub nodes that we often discard, which could represent unstable ongoing variations of molecular strategy, but it is not yet consolidated. So, although this

result may already exist in another context, where different design objectives obscure it, in this investigation, we present precise molecular data that support a different way to approach the distribution of nodes in an interactome, suggesting new design hypotheses. The scientific community should verify these data.

5. Conclusions

The aim of this study was to give an overall view of the molecular mechanisms involved in SARS-CoV-2 liver infection. Our research shows that COVID-19 affects only 50% of liver proteins, but it triggers a vast network of interactions among them. Based on this observation, we can infer that the virus does not attack the molecular mechanisms that are vital for cellular metabolism. Instead, it seems to affect the protein complexes governed by influential human proteins, employing a variety of types of viral proteins. The ability of these proteins to interact with many human proteins, each with distinct structural characteristics, is essential for controlling specific biological processes, such as translation. The virus attacks the entire ribosomal system, demonstrating the importance of controlling protein biosynthesis. All this also suggests that specific human proteins can serve as targets for antiviral drugs.

Two things appeared important from the set of multiple analyses that characterized this study. Researchers hunt for hub proteins because they may be ideal drug targets. Many nodes, unfortunately, turn out not to be hubs, but they support inappropriate functional hypotheses that are widespread in the literature. This shows how necessary it is to use only validated interaction data for computational analyses [120,121]. To this we add that metabolism is degenerate. The complexity of establishing cause–effect relationships in biological processes [122–126] cannot be addressed by probing a few specific proteins, like with Western blotting. A protein believed to be involved in a biological process can often be found in various forms of aggregation in multiple functional sub-networks [122–126]. Therefore, without determining the specific molecular process within that context, we cannot make any conclusions regarding cause or effect. This also generates inappropriate functional hypotheses.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/livers4020016/s1, Figure S1: Network of the pruned set of 126 nodes; Figure S2: Enriched network of the 126 original hub genes; Figure S3: CLUSTER CL:143; Figure S4: CLUSTER CL:152; Figure S5: CLUSTER CL:159; Figure S6: CLUSTER CL:162; Table S1: List of original hub genes from the literature, including those shared by multiple articles (142 hub genes); Table S2: Original set stripped of shared genes (126 hub genes); Table S3: Comprehensive set of enriched functions of the interactome in Figure 1 of the article; Table S4: Functional characteristics of Cluster-CL:143; Table S5: Functional characteristics of Cluster-CL:152; Table S6: Functional characteristics of Cluster-159; Table S7: Functional characteristics of Cluster-CL:162; Clustering: Cluster CL:143, Viral mRNA Translation, and Sec61 translocon complex; Cluster CL:152, Viral mRNA Translation; Cluster CL:159, Viral mRNA Translation; Cluster CL162, Cytoplasmic ribosomal proteins. EXCEL FILE S1: Node Degree of figure 1; EXCEL FILE S2: Percentage composition of interaction sources between human proteins and those of SARS-CoV-2; EXCEL FILE S3: Comprehensive Interactions Data in Liver.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The SARS 2-Human Proteome Interaction Database (SHPID) was assembled with online data from the COVID-19 Coronavirus Project from BioGRID. These data are 100% freely available to both academic and commercial users under the MIT License and are provided with no warranty at the following address: https://downloads.thebiogrid.org/File/BioGRID/Latest-Release/BIOGRID-PROJECT-covid19_coronavirus_project-LATEST.zip, accessed on 1 July 2023. The zip file contains multiple zip files (32 zip files) each comprising interactions and post-translational modifications for each single viral protein for a total of 33,823 interactions (as of June 2023).

Conflicts of Interest: The author declares no conflicts of interest.

Appendix A

An appendix is necessary to frame the reason for a reverse engineering approach and explain why everything must be based on reliable data. When we think of a biological network, we think it comprises a one-to-one set of interactions of its nodes. What the nodes exchange is functional information, therefore, a biological network is an information system that manages the metabolic information of the cell and the entire organism. The more precise the metabolic information, the greater the homeostatic capacity of the entire organism. Therefore, when we refer to two nodes that have a functional relationship, i.e., exchange information, we must be very sure that the interaction exists. We can only get high certainty experimentally, for example, through the methods of biochemistry and biophysics. The mutual information between two variables, i.e., the nodes, measures the amount of information that one variable contains about another. The higher the certainty, the greater the reduction in the functional uncertainty of one variable following the knowledge of another. Mutual information between two variables is a fundamental concept of information theory as defined by Shannon [127]. To apply this concept to one-to-one biological interactions, we should define it in terms of entropy, as an uncertainty of the information that is transmitted [128].

Reciprocal information is a measure of dependency between variables, which can be analyzed by interaction networks [37]: if two components have strong interactions, their reciprocal information will be high, increasing the certainty of the event. The intrinsic information of an event, also called self-information, is the amount of intrinsic uncertainty associated with it. The more certain the event, the lower the amount of uncertainty associated with it. From which, the more certain the information one possesses, knowing that the event has occurred, the lower the associated uncertainty will be, but the lower the self-information or intrinsic information, i.e., its total entropy, will also be.

In entropic terms:

$$I(X, Y) = H(X) - H(X/Y) \tag{A1}$$

where I(X, Y) is the uncertainty existing in the relationship between the two nodes X and Y and it depends on the level of informational uncertainty relating to each variable. H(X) is the entropy of the information system, or self-information or intrinsic information, which is the amount of uncertainty associated with the interaction between the two nodes. H(X/Y) is the conditional entropy, i.e., the entropy of a variable Y conditioned from the knowledge we have of the other variable, and, the higher the knowledge, the lower the associated uncertainty.

Relation (A1) tells us that the uncertainty between two metabolic nodes, which have a physical and/or functional relationship [I(X, Y)], depends on the intrinsic uncertainty [H(X)] associated with the relationship itself. We can reduce this uncertainty by increasing our knowledge of the metabolic behavior of the interacting nodes [H(X/Y)]. However, by confirming a metabolic event between two nodes through experimentation, we obtain secure and certain information, thus eliminating the conditional uncertainty and reducing the self-information, intrinsic information, or entropy of the system.

It is crucial to acknowledge that information fully controls biological networks. The greater the certainty of the information we possess about the physical/functional relationships existing in the network, the more certain and real the metabolic or pathological previsions of our computational model are. Therefore, since the relationships between two metabolic nodes are physical/functional, we gain the greatest certainty of all only through conducting experiments, with the methods of biophysics, with which we measure the type and strength of the interaction, and of biochemistry, with which we measure the levels of function. So, the relationships between computational models and experimental data are one cornerstone of systems biology. Reverse engineering aims to understand which functional processes are real and which are dysregulated through external control of the certainty of the biological event of virus–host interaction. The goal is the ultimate biological

determination of existing interactions, not the detailed characterization of these interactions, knowing that difficulties increase because we deal with non-linear interactions.

The existence of many errors undermines these principles very often because uncertainty is intrinsic in the multiple contexts that provide data and information relating to the biomolecules necessary to calculate biological networks. Although next-generation sequencing studies provide extensive sequence information, the precise knowledge of virus–host one-to-one protein interactions and potential targets for antiviral therapies remains limited, partial, and incomplete. Typically, metadata for PPIs [129] should include experimental details of tens of thousands of virus–human interactions. Some databases, such as BioGRID, STRING, or INTACT, have used standardized procedures, but many others, generalists, have collected virus–host interactions in different ways and contexts [45] and do not have a standard format.

These platforms are online and useful for checking results. The fundamental reason lies in the crucial distinction between reproducibility (repeating an experiment to obtain the same result) and replicability (interpreting the same data in different contexts). It is important to recognize that interpretations of data may vary depending on context, data quality, or analysis method. Standardization of data and protocols is necessary to obtain a univocal understanding and interpretation of research results. The vast differences between databases make it extremely challenging to compare their data when the lack of experimental details obscures the nature of an interaction. What we often observe in interactomics papers is an abnormal bloom of hub genes/proteins far beyond the needs of any biological network [46]. Therefore, STRING, a platform that for each calculated interaction in a graph creates a specific knowledge base by querying thousands of scientific articles on PubMed, and BioGRID, a platform that archives only curated experimental data of the one-to-one interactions of SARS-CoV-2 proteins with the human proteome, are two indispensable tools to guarantee the best possible certainty of the data under analysis. The liver is a very complex organ with a highly dynamic metabolism, where the sequential regulation of cellular processes plays a crucial role [119,130]. Therefore, studying its metabolic behavior during COVID-19 requires knowledge of the control systems and areas [131], which are not always found out in liver diseases [46].

References

1. Kariyawasam, J.C.; Jayarajah, U.; Abeysuriya, V.; Riza, R.; Seneviratne, S.L. Involvement of the Liver in COVID-19: A Systematic Review. *Am. J. Trop. Med. Hyg.* **2022**, *106*, 1026–1041. [CrossRef]
2. Beigmohammadi, M.T.; Jahanbin, B.; Safaei, M.; Amoozadeh, L.; Khoshavi, M.; Mehrtash, V.; Jafarzadeh, B.; Abdollahi, A. Pathological findings of postmortem biopsies from lung, heart, and liver of 7 deceased COVID-19 patients. *Int. J. Surg. Pathol.* **2021**, *29*, 135–145. [CrossRef]
3. Ryan, P.M.; Caplice, N.M. Is Adipose Tissue a Reservoir for Viral Spread, Immune Activation, and Cytokine Amplification in Coronavirus Disease 2019? *Obesity* **2020**, *28*, 1191–1194. [CrossRef]
4. Hamming, I.; Timens, W.; Bulthuis, M.L.C.; Lely, A.T.; Navis, G.J.; van Goor, H. Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first step in understanding SARS pathogenesis. *J. Pathol.* **2004**, *203*, 631–637. [CrossRef]
5. Ding, Y.; He, L.; Zhang, Q.; Huang, Z.; Che, X.; Hou, J.; Wang, H.; Shen, H.; Qiu, L.; Li, Z.; et al. Organ distribution of severe acute respiratory syndrome (SARS) associated coronavirus (SARS-CoV) in SARS patients: Implications for pathogenesis and virus transmission pathways. *J. Pathol.* **2004**, *203*, 622–630. [CrossRef]
6. Birman, D. Investigation of the Effects of COVID-19 on Different Organs of the Body. *Eurasian J. Chem. Med. Pet. Res.* **2023**, *2*, 24–36.
7. Paolini, A.; Borella, R.; De Biasi, S.; Neroni, A.; Mattioli, M.; Tartaro, D.L.; Simonini, C.; Franceschini, L.; Cicco, G.; Piparo, A.M.; et al. Cell Death in Coronavirus Infections: Uncovering Its Role during COVID-19. *Cells* **2021**, *10*, 1585. [CrossRef]
8. Yuan, C.; Ma, Z.; Xie, J.; Li, W.; Su, L.; Zhang, G.; Xu, J.; Wu, Y.; Zhang, M.; Liu, W. The role of cell death in SARS-CoV-2 infection. *Signal Transduct. Target. Ther.* **2023**, *8*, 357. [CrossRef]
9. Jothimani, D.; Venugopal, R.; Abedin, M.F.; Kaliamoorthy, I.; Rela, M. COVID-19 and the liver. *J. Hepatol.* **2020**, *73*, 1231–1240. [CrossRef]
10. Guan, G.W.; Gao, L.; Wang, J.W.; Wen, X.J.; Mao, T.H.; Peng, S.W.; Zhang, T.; Chen, X.M.; Lu, F.M. Exploring the mechanism of liver enzyme abnormalities in patients with novel coronavirus-infected pneumonia. *Chin. J. Hepatol.* **2020**, *28*, 100–106.
11. Shi, J.; Li, G.; Yuan, X.; Wang, Y.; Gong, M.; Li, C.; Ge, X.; Lu, S. Exploration and verification of COVID-19-related hub genes in liver physiological and pathological regeneration. *Front. Bioeng. Biotechnol.* **2023**, *11*, 1135997. [CrossRef]

12. Vandereyken, K.; Van Leene, J.; De Coninck, B.; Cammue, B.P.A. Hub Protein Controversy: Taking a Closer Look at Plant Stress Response Hubs. *Front. Plant Sci.* **2018**, *9*, 694. [CrossRef]
13. Huang, T.; Zheng, D.B.; Song, Y.B.; Pan, H.B.; Qiu, G.; Xiang, Y.B.; Wang, Z.B.; Wang, F. Demonstration of the impact of COVID-19 on metabolic associated fatty liver disease by bioinformatics and system biology approach. *Medicine* **2023**, *102*, e34570. [CrossRef]
14. Luo, H.; Chen, J.; Jiang, Q.; Yu, Y.; Yang, M.; Luo, Y.; Wang, X. Comprehensive DNA methylation profiling of COVID-19 and hepatocellular carcinoma to identify common pathogenesis and potential therapeutic targets. *Clin. Epigenetics* **2023**, *15*, 100. [CrossRef]
15. Jiang, S.-T.; Liu, Y.-G.; Zhang, L.; Sang, X.-T.; Xu, Y.-Y.; Lu, X. Systems biology approach reveals a common molecular basis for COVID-19 and non-alcoholic fatty liver disease (NAFLD). *Eur. J. Med. Res.* **2022**, *27*, 251. [CrossRef]
16. Shen, Q.; Wang, J.; Zhao, L. To investigate the internal association between SARS-CoV-2 infections and cancer through bioinformatics. *Math. Biosci. Eng.* **2022**, *19*, 11172–11194. [CrossRef]
17. Wang, L.; Ding, Y.; Zhang, C.; Chen, R. Target and drug predictions for SARS-CoV-2 infection in hepatocellular carcinoma patients. *PLoS ONE* **2022**, *17*, e0269249. [CrossRef]
18. Abolfazli, P.; Aghajanzadeh, T.; Ghaderinasrabad, M.; Nchama, C.N.A.; Mokhlesi, A.; Talkhabi, M. Bioinformatics analysis reveals molecular connections between non-alcoholic fatty liver disease (NAFLD) and COVID-19. *J. Cell Commun. Signal.* **2022**, *16*, 609–619. [CrossRef]
19. Mousavi, S.Z.; Rahmanian, M.; Sami, A. Organ-specific or personalized treatment for COVID-19: Rationale, evidence, and potential candidates. *Funct. Integr. Genom.* **2022**, *22*, 429–433. [CrossRef]
20. Hasankhani, A.; Bahrami, A.; Sheybani, N.; Aria, B.; Hemati, B.; Fatehi, F.; Farahani, H.G.M.; Javanmard, G.; Rezaee, M.; Kastelic, J.P.; et al. Differential Co-Expression Network Analysis Reveals Key Hub-High Traffic Genes as Potential Therapeutic Targets for COVID-19 Pandemic. *Front. Immunol.* **2021**, *12*, 789317. [CrossRef]
21. Sokouti, B. A systems biology approach for investigating significantly expressed genes among COVID-19, hepatocellular carcinoma, and chronic hepatitis B. *Egypt. J. Med. Hum. Genet.* **2022**, *23*, 146. [CrossRef]
22. Chen, J.C.; Xie, T.A.; Lin, Z.Z.; Li, Y.Q.; Xie, Y.F.; Li, Z.W.; Guo, X.G. Identification of key pathways and genes in SARS-CoV-2 infecting human intestines by bioinformatics analysis. *Biochem. Genet.* **2022**, *60*, 1076–1094. [CrossRef]
23. Steuer, R. Computational approaches to the topology, stability and dynamics of metabolic networks. *Phytochemistry* **2007**, *68*, 2139–2151. [CrossRef]
24. Hartman, E.; Scott, A.M.; Karlsson, C.; Mohanty, T.; Vaara, S.T.; Linder, A.; Malmström, L.; Malmström, J. Interpreting biologically informed neural networks for enhanced proteomic biomarker discovery and pathway analysis. *Nat. Commun.* **2023**, *14*, 5359. [CrossRef]
25. Wu, S.; Liu, X.; Dong, A.; Gragnoli, C.; Griffin, C.; Wu, J.; Yau, S.-T.; Wu, R. The metabolomic physics of complex diseases. *Proc. Natl. Acad. Sci. USA* **2023**, *120*, e2308496120. [CrossRef]
26. Yang, Y.; Fang, Q.; Shen, H.-B. Predicting gene regulatory interactions based on spatial gene expression data and deep learning. *PLoS Comput. Biol.* **2019**, *15*, e1007324. [CrossRef]
27. Chikofsky, E.; Cross, J. Reverse engineering and design recovery: A taxonomy. *IEEE Softw.* **1990**, *7*, 13–17. [CrossRef]
28. Fornito, A.; Zalesky, A.; Breakspear, M. Graph analysis of the human connectome: Promise, progress, and pitfalls. *Neuroimage* **2013**, *80*, 426–444. [CrossRef]
29. Green, S. Can biological complexity be reverse engineered? *Stud. Hist. Philos. Sci. Part C Stud. Hist. Philos. Biol. Biomed. Sci.* **2015**, *53*, 73–83. [CrossRef]
30. Natale, J.L.; Hofmann, D.; Hernández, D.G.; Nemenman, I. Reverse-engineering biological networks from large data sets. *arXiv* **2017**, arXiv:1705.06370.
31. de Camargo, R.S.; de Miranda, G.; Løkketangen, A. A new formulation and an exact approach for the many-to-many hub location-routing problem. *Appl. Math. Model.* **2013**, *37*, 7465–7480. [CrossRef]
32. Qu, Y.; Jiang, J.; Liu, X.; Yang, X.; Tang, C. Non-epigenetic mechanisms enable short memories of the environment for cell cycle commitment. *BioRxiv* **2020**. [CrossRef]
33. Pisco, A.O.; D'hérouël, A.F.; Huang, S. Conceptual Confusion: The case of Epigenetics. *BioRxiv* **2016**, 053009. [CrossRef]
34. Squire, L.R.; Genzel, L.; Wixted, J.T.; Morris, R.G. Memory consolidation. *Cold Spring Harb. Perspect. Biol.* **2015**, *7*, a021766. [CrossRef]
35. Oughtred, R.; Rust, J.; Chang, C.; Breitkreutz, B.; Stark, C.; Willems, A.; Boucher, L.; Leung, G.; Kolas, N.; Zhang, F.; et al. The BioGRID database: A comprehensive biomedical resource of curated protein, genetic, and chemical interactions. *Protein Sci.* **2020**, *30*, 187–200. [CrossRef]
36. Teo, G.; Liu, G.; Zhang, J.; Nesvizhskii, A.I.; Gingras, A.-C.; Choi, H. SAINTexpress: Improvements and additional features in Significance Analysis of INTeractome software. *J. Proteom.* **2013**, *100*, 37–43. [CrossRef]
37. 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.* **2020**, *49*, D605–D612, Erratum in: *Nucleic Acids Res.* **2021**, *49*, 10800. [CrossRef]
38. Szklarczyk, D.; Kirsch, R.; Koutrouli, M.; Nastou, K.; Mehryary, F.; Hachilif, R.; Gable, A.L.; Fang, T.; Doncheva, N.T.; Pyysalo, S.; et al. The STRING database in 2023: Protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res.* **2022**, *51*, D638–D646. [CrossRef]

39. Kumar, A.; Ingle, Y.S.; Pande, A.; Dhule, P. Canopy clustering: A review on pre-clustering approach to k-means clustering. *Int. J. Innov. Adv. Comput. Sci. (IJIACS)* **2014**, *3*, 22–29.
40. Doncheva, N.T.; Morris, J.H.; Gorodkin, J.; Jensen, L.J. Cytoscape StringApp: Network Analysis and Visualization of Proteomics Data. *J. Proteome Res.* **2018**, *18*, 623–632. [CrossRef]
41. Chung, F.; Lu, L.; Dewey, T.G.; Galas, D.J. Duplication Models for Biological Networks. *J. Comput. Biol.* **2003**, *10*, 677–687. [CrossRef] [PubMed]
42. Scardoni, G.; Tosadori, G.; Faizan, M.; Spoto, F.; Fabbri, F.; Laudanna, C. Biological network analysis with CentiScaPe: Centralities and experimental dataset integration. *F1000Research* **2015**, *3*, 139. [CrossRef]
43. Wuchty, S.; Ravasz, E.; Barabási, A.-L. The architecture of biological networks. In *Complex Systems Science in Biomedicine*; Springer: Boston, MA, USA, 2006; pp. 165–181.
44. Almaas, E.; Vázquez, A.; Barabási, A.-L. Scale-free networks in biology. In *Biological. Networks. Complex Systems and Interdisciplinary Science*; Képès, F., Ed.; Word Scientific Publishing Co.: Hackensack, NJ, USA, 2007; Chapter 1; Volume 3, pp. 1–20; ISBN 978-981-270-695-9.
45. Szklarczyk, D.; Jensen, L.J. Protein-protein interaction databases. *Protein-Protein Interact. Methods Appl.* **2015**, *1278*, 39–56. [CrossRef]
46. Sharma, A.; Colonna, G. System-Wide Pollution of Biomedical Data: Consequence of the Search for Hub Genes of Hepatocellular Carcinoma Without Spatiotemporal Consideration. *Mol. Diagn. Ther.* **2021**, *25*, 9–27. [CrossRef]
47. Yang, S.; Fu, C.; Lian, X.; Dong, X.; Zhang, Z. Understanding Human-Virus Protein-Protein Interactions Using a Human Protein Complex-Based Analysis Framework. *mSystems* **2019**, *4*. [CrossRef]
48. Mishra, P.M.; Verma, N.C.; Rao, C.; Uversky, V.N.; Nandi, C.K. Intrinsically disordered proteins of viruses: Involvement in the mechanism of cell regulation and pathogenesis. *Prog. Mol. Biol. Transl. Sci.* **2020**, *174*, 1–78.
49. Villarreal, L.P. The widespread evolutionary significance of viruses. In *Origin and Evolution of Viruses*; Elsevier Science Direct.: Amsterdam, The Netherlands, 2008; Chapter 21; pp. 477–516. [CrossRef]
50. Guidotti, R.; Gardoni, P.; Chen, Y. Network reliability analysis with link and nodal weights and auxiliary nodes. *Struct. Saf.* **2017**, *65*, 12–26. [CrossRef]
51. De Vico Fallani, F.; Richiardi, J.; Chavez, M.; Achard, S. Graph analysis of functional brain networks: Practical issues in translational neuroscience. *Philos. Trans. R. Soc. B Biol. Sci.* **2014**, *369*, 20130521. [CrossRef]
52. Li, V.; Silvester, J. Performance Analysis of Networks with Unreliable Components. *IEEE Trans. Commun.* **1984**, *32*, 1105–1110. [CrossRef]
53. Knight, S.; Nguyen, H.X.; Falkner, N.; Bowden, R.; Roughan, M. The Internet Topology Zoo. *IEEE J. Sel. Areas Commun.* **2011**, *29*, 1765–1775. [CrossRef]
54. Militello, G.; Álvaro, M. Structural and organisational conditions for being a machine. *Biol. Philos.* **2018**, *33*, 35. [CrossRef]
55. Akyildiz, I.F.; Brunetti, F.; Blázquez, C. Nanonetworks: A new communication paradigm. *Comput. Netw.* **2008**, *52*, 2260–2279. [CrossRef]
56. Will, C.L.; Urlaub, H.; Achsel, T.; Gentzel, M.; Wilm, M.; Lührmann, R. Characterization of novel SF3b and 17S U2 snRNP proteins, including a human Prp5p homologue and an SF3b DEAD-box protein. *EMBO J.* **2002**, *21*, 4978–4988. [CrossRef]
57. Wang, C.; Chen, L.; Chen, Y.; Jia, W.; Cai, X.; Liu, Y.; Ji, F.; Xiong, P.; Liang, A.; Liu, R.; et al. Abnormal global alternative RNA splicing in COVID-19 patients. *PLoS Genet.* **2022**, *18*, e1010137. [CrossRef]
58. Wang, E.T.; Sandberg, R.; Luo, S.; Khrebtukova, I.; Zhang, L.; Mayr, C.; Kingsmore, S.F.; Schroth, G.P.; Burge, C.B. Alternative isoform regulation in human tissue transcriptomes. *Nature* **2008**, *456*, 470–476. [CrossRef]
59. Luo, J.; Zhao, H.; Chen, L.; Liu, M. Multifaceted functions of RPS27a: An unconventional ribosomal protein. *J. Cell. Physiol.* **2022**, *238*, 485–497. [CrossRef]
60. Zhang, Y.; Zhang, J.; Chen, X.; Yang, Z. Polymeric immunoglobulin receptor (PIGR) exerts oncogenic functions via activating ribosome pathway in hepatocellular carcinoma. *Int. J. Med. Sci.* **2021**, *18*, 364–371. [CrossRef]
61. Vandelli, A.; Monti, M.; Milanetti, E.; Armaos, A.; Rupert, J.; Zacco, E.; Bechara, E.; Delli Ponti, P.; Tartaglia, G.G. Structural analysis of SARS-CoV-2 genome and predictions of the human interactome. *Nucleic Acids Res.* **2020**, *48*, 11270–11283. [CrossRef]
62. Chiariello, A.M.; Abraham, A.; Bianco, S.; Esposito, A.; Vercellone, F.; Conte, M.; Fontana, A.; Nicodemi, M. Multiscale modelling of chromatin 4D organization in SARS-CoV-2 infected cells. *bioRxiv* **2023**. [CrossRef] [PubMed] [PubMed Central]
63. Chernyak, B.V.; Popova, E.N.; Prikhodko, A.S.; Grebenchikov, O.A.; Zinovkina, L.A.; Zinovkin, R.A. COVID-19 and oxidative stress. *Biochemistry* **2020**, *85*, 1543–1553. [CrossRef]
64. Jana, S.; Heaven, M.R.; Stauft, C.B.; Wang, T.T.; Williams, M.C.; D'Agnillo, F.; Alayash, A.I. HIF-1α-Dependent Metabolic Reprogramming, Oxidative Stress, and Bioenergetic Dysfunction in SARS-CoV-2-Infected Hamsters. *Int. J. Mol. Sci.* **2022**, *24*, 558. [CrossRef]
65. Serebrovska, Z.O.; Chong, E.Y.; Serebrovska, T.V.; Tumanovska, L.V.; Xi, L. Hypoxia, HIF-1α, and COVID-19: From pathogenic factors to potential therapeutic targets. *Acta Pharmacol. Sin.* **2020**, *41*, 1539–1546. [CrossRef]
66. Wing, P.A.; Keeley, T.P.; Zhuang, X.; Lee, J.Y.; Prange-Barczynska, M.; Tsukuda, S.; Morgan, S.B.; Harding, A.C.; Argles, I.L.A.; Kurlekar, S.; et al. Hypoxic and pharmacological activation of HIF inhibits SARS-CoV-2 infection of lung epithelial cells. *Cell Rep.* **2021**, *35*, 109020. [CrossRef]

67. Zhu, Z.; Zheng, Z.; Liu, J. Comparison of COVID-19 and Lung Cancer via Reactive Oxygen Species Signaling. *Front. Oncol.* **2021**, *11*, 708263. [CrossRef]
68. Bhandari, V.; Hoey, C.; Liu, L.Y.; Lalonde, E.; Ray, J.; Livingstone, J.; Lesurf, R.; Shiah, Y.-J.; Vujcic, T.; Huang, X.; et al. Molecular landmarks of tumor hypoxia across cancer types. *Nat. Genet.* **2019**, *51*, 308–318. [CrossRef]
69. Cimmino, F.; Avitabile, M.; Lasorsa, V.A.; Montella, A.; Pezone, L.; Cantalupo, S.; Visconte, F.; Corrias, M.V.; Iolascon, A.; Capasso, M. HIF-1 transcription activity: HIF1A driven response in normoxia and in hypoxia. *BMC Med. Genet.* **2019**, *20*, 37. [CrossRef]
70. Varghese, B.; Chianese, U.; Capasso, L.; Sian, V.; Bontempo, P.; Conte, M.; Benedetti, R.; Altucci, L.; Carafa, V.; Nebbioso, A. SIRT1 activation promotes energy homeostasis and reprograms liver cancer metabolism. *J. Transl. Med.* **2023**, *21*, 627. [CrossRef]
71. Wang, X.; Simpson, E.R.; Brown, K.A. p53: Protection against Tumor Growth beyond Effects on Cell Cycle and Apoptosis. *Cancer Res.* **2015**, *75*, 5001–5007. [CrossRef]
72. Moll, U.M.; Petrenko, O. The MDM2-p53 interaction. *Mol. Cancer Res.* **2003**, *1*, 1001–1008.
73. Liu, Y.; Deisenroth, C.; Zhang, Y. RP–MDM2–p53 pathway: Linking ribosomal biogenesis and tumor surveillance. *Trends Cancer* **2016**, *2*, 191–204. [CrossRef]
74. Halehalli, R.R.; Nagarajaram, H.A. Molecular principles of human virus protein–protein interactions. *Bioinformatics* **2014**, *31*, 1025–1033. [CrossRef]
75. Gordon, D.E.; Jang, G.M.; Bouhaddou, M.; Xu, J.; Obernier, K.; White, K.M.; O'Meara, M.J.; Rezelj, V.V.; Guo, J.Z.; Swaney, D.L.; et al. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* **2020**, *583*, 459–468. [CrossRef]
76. Gordon, D.E.; Hiatt, J.; Bouhaddou, M.; Rezelj, V.V.; Ulferts, S.; Braberg, H.; Jureka, A.S.; Obernier, K.; Guo, J.Z.; Batra, J.; et al. Comparative host-coronavirus protein interaction networks reveal pan-viral disease mechanisms. *Science* **2020**, *370*, eabe9403. [CrossRef]
77. Komarova, A.V.; Combredet, C.; Sismeiro, O.; Dillies, M.-A.; Jagla, B.; David, R.Y.S.; Vabret, N.; Coppée, J.-Y.; Vidalain, P.-O.; Tangy, F. Identification of RNA partners of viral proteins in infected cells. *RNA Biol.* **2013**, *10*, 943–956. [CrossRef]
78. Li, J.; Guo, M.; Tian, X.; Wang, X.; Yang, X.; Wu, P.; Liu, C.; Xiao, Z.; Qu, Y.; Yin, Y.; et al. Virus–host interactome and proteomic survey reveal potential virulence factors influencing SARS-CoV-2 pathogenesis. *Med* **2021**, *2*, 99–112. [CrossRef]
79. Stukalov, A.; Girault, V.; Grass, V.; Karayel, O.; Bergant, V.; Urban, C.; Haas, D.A.; Huang, Y.; Oubraham, L.; Wang, A.; et al. Multilevel proteomics reveals host perturbations by SARS-CoV-2 and SARS-CoV. *Nature* **2021**, *594*, 246–252. [CrossRef]
80. Zhou, Y.; Liu, Y.; Gupta, S.; Paramo, M.I.; Hou, Y.; Mao, C.; Luo, Y.; Judd, J.; Wierbowski, S.; Bertolotti, M.; et al. A comprehensive SARS-CoV-2–human protein–protein interactome reveals COVID-19 pathobiology and potential host therapeutic targets. *Nat. Biotechnol.* **2022**, *41*, 128–139. [CrossRef]
81. Khorsand, B.; Savadi, A.; Naghibzadeh, M. SARS-CoV-2-human protein-protein interaction network. *Inform. Med. Unlocked* **2020**, *20*, 100413. [CrossRef]
82. Ghosh, N.; Saha, I.; Sharma, N. Interactome of human and SARS-CoV-2 proteins to identify human hub proteins associated with comorbidities. *Comput. Biol. Med.* **2021**, *138*, 104889. [CrossRef]
83. Srinivasan, S.; Cui, H.; Gao, Z.; Liu, M.; Lu, S.; Mkandawire, W.; Narykov, O.; Sun, M.; Korkin, D. Structural Genomics of SARS-CoV-2 Indicates Evolutionary Conserved Functional Regions of Viral Proteins. *Viruses* **2020**, *12*, 360. [CrossRef]
84. Shuler, G.; Hagai, T. Rapidly evolving viral motifs mostly target biophysically constrained binding pockets of host proteins. *Cell Rep.* **2022**, *40*, 111212. [CrossRef] [PubMed]
85. Mendez-Rios, J.; Uetz, P.; Luo, Y.; Muesing, M.A.; E Ballestas, M.; Kaye, K.M.; Moyano, D.F.; Rotello, V.M.; Gazzé, G.; Rodland, K.D.; et al. Global approaches to study protein–protein interactions among viruses and hosts. *Futur. Microbiol.* **2010**, *5*, 289–301. [CrossRef] [PubMed]
86. Goh, G.K.-M.; Dunker, A.K.; Uversky, V.N. Understanding Viral Transmission Behavior via Protein Intrinsic Disorder Prediction: Coronaviruses. *J. Pathog.* **2012**, *2012*, 738590. [CrossRef] [PubMed]
87. Anjum, F.; Mohammad, T.; Asrani, P.; Shafie, A.; Singh, S.; Yadav, D.K.; Uversky, V.N.; Imtaiyaz Hassan, C.M. Identification of intrinsically disorder regions in non-structural proteins of SARS-CoV-2: New insights into drug and vaccine resistance. *Mol. Cell. Biochem.* **2022**, *477*, 1607–1619. [CrossRef]
88. Anger, A.M.; Armache, J.-P.; Berninghausen, O.; Habeck, M.; Subklewe, M.; Wilson, D.N.; Beckmann, R. Structures of the human and Drosophila 80S ribosome. *Nature* **2013**, *497*, 80–85. [CrossRef]
89. Singh, S.; Broeck, A.V.; Miller, L.; Chaker-Margot, M.; Klinge, S. Nucleolar maturation of the human small subunit processome. *Science* **2021**, *373*, eabj5338. [CrossRef]
90. Baranov, P.V.; Henderson, C.M.; Anderson, C.B.; Gesteland, R.F.; Atkins, J.F.; Howard, M.T. Programmed ribosomal frameshifting in decoding the SARS-CoV genome. *Virology* **2005**, *332*, 498–510. [CrossRef]
91. Rehfeld, F.; Eitson, J.L.; Ohlson, M.B.; Chang, T.C.; Schoggins, J.W.; Mendell, J.T. CRISPR screening reveals a dependency on ribosome recycling for efficient SARS-CoV-2 programmed ribosomal frameshifting and viral replication. *Cell Rep.* **2023**, *42*, 112076. [CrossRef]
92. Khrustalev, V.V.; Giri, R.; Khrustaleva, T.A.; Kapuganti, S.K.; Stojarov, A.N.; Poboinev, V.V. Translation-associated mutational U-pressure in the first ORF of SARS-CoV-2 and other coronaviruses. *Front. Microbiol.* **2020**, *11*, 559165. [CrossRef]
93. Kusakabe, T.; Motoki, K.; Hori, K. Mode of Interactions of Human Aldolase Isozymes with Cytoskeletons. *Arch. Biochem. Biophys.* **1997**, *344*, 184–193. [CrossRef]

94. Esposito, G.; Vitagliano, L.; Costanzo, P.; Borrelli, L.; Barone, R.; Pavone, L.; Izzo, P.; Zagari, A.; Salvatore, F. Human aldolase A natural mutants: Relationship between flexibility of the C-terminal region and enzyme function. *Biochem. J.* **2004**, *380 Pt 1*, 51–56. [CrossRef] [PubMed]
95. Guittet, O.; Håkansson, P.; Voevodskaya, N.; Fridd, S.; Gräslund, A.; Arakawa, H.; Nakamura, Y.; Thelander, L. Mammalian p53R2 Protein Forms an Active Ribonucleotide Reductasein Vitro with the R1 Protein, Which Is Expressed Both in Resting Cells in Response to DNA Damage and in Proliferating Cells. *J. Biol. Chem.* **2001**, *276*, 40647–40651. [CrossRef] [PubMed]
96. Yamaguchi, T.; Matsuda, K.; Sagiya, Y.; Iwadate, M.; Fujino, M.A.; Nakamura, Y.; Arakawa, H. p53R2-dependent pathway for DNA synthesis in a p53-regulated cell cycle checkpoint. *Cancer Res.* **2001**, *61*, 8256–8262. [PubMed]
97. Rauch, J.N.; Gestwicki, J.E. Binding of Human Nucleotide Exchange Factors to Heat Shock Protein 70 (Hsp70) Generates Functionally Distinct Complexes in Vitro. *J. Biol. Chem.* **2014**, *289*, 1402–1414. [CrossRef] [PubMed]
98. Takayama, S.; Xie, Z.; Reed, J.C. An Evolutionarily Conserved Family of Hsp70/Hsc70 Molecular Chaperone Regulators. *J. Biol. Chem.* **1999**, *274*, 781–786. [CrossRef] [PubMed]
99. Yu, Z.; Zeng, J.; Wang, J.; Cui, Y.; Song, X.; Zhang, Y.; Cheng, X.; Hou, N.; Teng, Y.; Lan, Y.; et al. Hepatocyte growth factor-regulated tyrosine kinase substrate is essential for endothelial cell polarity and cerebrovascular stability. *Cardiovasc. Res.* **2020**, *117*, 533–546. [CrossRef] [PubMed]
100. Wu, L.; Cheng, Y.; Geng, D.; Fan, Z.; Lin, B.; Zhu, Q.; Li, J.; Qin, W.; Yi, W. O-GlcNAcylation regulates epidermal growth factor receptor intracellular trafficking and signaling. *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2107453119. [CrossRef]
101. Han, J.; Goldstein, L.A.; Hou, W.; Watkins, S.C.; Rabinowich, H. Involvement of CASP9 (caspase 9) in IGF2R/CI-MPR endosomal transport. *Autophagy* **2020**, *17*, 1393–1409. [CrossRef]
102. Vázquez, A. Growing network with local rules: Preferential attachment, clustering hierarchy, and degree correlations. *Phys. Rev. E* **2003**, *67*, 056104. [CrossRef]
103. Giuraniuc, C.V.; Hatchett, J.P.L.; Indekeu, J.O.; Leone, M.; Castillo, I.P.; Van Schaeybroeck, B.; Vanderzande, C. Trading Interactions for Topology in Scale-Free Networks. *Phys. Rev. Lett.* **2005**, *95*, 098701. [CrossRef]
104. Caetano-Anollés, D.; Caetano-Anollés, K.; Caetano-Anollés, G. Evolution of Macromolecular Structure: A 'Double Tale' of Biological Accretion and Diversification. *Sci. Prog.* **2018**, *101*, 360–383. [CrossRef] [PubMed]
105. Caetano-Anollés, G.; Aziz, M.F.; Mughal, F.; Gräter, F.; Koç, I.; Caetano-Anollés, K.; Caetano-Anollés, D. Emergence of Hierarchical Modularity in Evolving Networks Uncovered by Phylogenomic Analysis. *Evol. Bioinform.* **2019**, *15*. [CrossRef]
106. Benson, A.R.; Gleich, D.F.; Leskovec, J. Higher-order organization of complex networks. *Science* **2016**, *353*, 163–166. [CrossRef] [PubMed]
107. Michoel, T.; Joshi, A.; Nachtergaele, B.; Van de Peer, Y. Enrichment and aggregation of topological motifs are independent organizational principles of integrated interaction networks. *Mol. Biosyst.* **2011**, *7*, 2769–2778. [CrossRef]
108. Almaas, E. Biological impacts and context of network theory. *J. Exp. Biol.* **2007**, *210*, 1548–1558. [CrossRef]
109. Modell, A.E.; Blosser, S.L.; Arora, P.S. Systematic targeting of protein–protein interactions. *Trends Pharmacol. Sci.* **2016**, *37*, 702–713. [CrossRef]
110. Subramanian, S.; Kumar, S. Gene Expression Intensity Shapes Evolutionary Rates of the Proteins Encoded by the Vertebrate Genome. *Genetics* **2004**, *168*, 373–381. [CrossRef]
111. Szaflik, T.; Smolarz, B.; Mroczkowska, B.; Kulig, B.; Soja, M.; Romanowicz, H.; Bryś, M.; Forma, E.; Szyłło, K. An Analysis of *ESR2* and *CYP19A1* Gene Expression Levels in Women with Endometriosis. *Vivo* **2020**, *34*, 1765–1771. [CrossRef] [PubMed]
112. Zeng, C.; Evans, J.P.; King, T.; Zheng, Y.-M.; Oltz, E.M.; Whelan, S.P.J.; Saif, L.J.; Peeples, M.E.; Liu, S.-L. SARS-CoV-2 spreads through cell-to-cell transmission. *Proc. Natl. Acad. Sci. USA* **2021**, *119*, e2111400119. [CrossRef]
113. Colonna, G. Molecular mechanisms driving the action of the Spike S1 subunit of the SARS-CoV-2 virus in human metabolism by interactomic analysis. (manuscript in preparation).
114. Reményi, A.; Good, M.C.; Bhattacharyya, R.P.; Lim, W.A. The Role of Docking Interactions in Mediating Signaling Input, Output, and Discrimination in the Yeast MAPK Network. *Mol. Cell* **2005**, *20*, 951–962. [CrossRef]
115. Staley, J.P.; Woolford, J.L. Assembly of ribosomes and spliceosomes: Complex ribonucleoprotein machines. *Curr. Opin. Cell Biol.* **2009**, *21*, 109–118. [CrossRef]
116. Abbasian, M.H.; Mahmanzar, M.; Rahimian, K.; Mahdavi, B.; Tokhanbigli, S.; Moradi, B.; Sisakht, M.M.; Deng, Y. Global landscape of SARS-CoV-2 mutations and conserved regions. *J. Transl. Med.* **2023**, *21*, 152. [CrossRef]
117. Markov, P.V.; Ghafari, M.; Beer, M.; Lythgoe, K.; Simmonds, P.; Stilianakis, N.I.; Katzourakis, A. The evolution of SARS-CoV-2. *Nat. Rev. Microbiol.* **2023**, *21*, 361–379. [CrossRef]
118. Telenti, A.; Hodcroft, E.B.; Robertson, D.L. The evolution and biology of SARS-CoV-2 variants. *Cold Spring Harb. Perspect. Med.* **2022**, *12*, a041390. [CrossRef]
119. Gebhardt, R.; MAatz-Soja, M. Liver zonation: Novel aspects of its regulation and its impact on homeostasis. *World J. Gastroenterol. WJG* **2014**, *20*, 8491. [CrossRef]
120. Burley, S.K.; Bhikadiya, C.; Bi, C.; Bittrich, S.; Chen, L.; Crichlow, G.V.; Christie, C.H.; Dalenberg, K.; Di Costanzo, L.; Duarte, J.M.; et al. RCSB Protein Data Bank: Powerful new tools for exploring 3D structures of biological macromolecules for basic and applied research and education in fundamental biology, biomedicine, biotechnology, bioengineering and energy sciences. *Nucleic Acids Res.* **2020**, *49*, D437–D451. [CrossRef]

121. Burke, D.F.; Bryant, P.; Barrio-Hernandez, I.; Memon, D.; Pozzati, G.; Shenoy, A.; Zhu, W.; Dunham, A.S.; Albanese, P.; Keller, A.; et al. Towards a structurally resolved human protein interaction network. *Nat. Struct. Mol. Biol.* **2023**, *30*, 216–225. [CrossRef]
122. Baek, M.; DiMaio, F.; Anishchenko, I.; Dauparas, J.; Ovchinnikov, S.; Lee, G.R.; Wang, J.; Cong, Q.; Kinch, L.N.; Schaeffer, R.D.; et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **2021**, *373*, 871–876. [CrossRef]
123. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596*, 583–589. [CrossRef]
124. Bryant, P.; Pozzati, G.; Elofsson, A. Improved prediction of protein-protein interactions using AlphaFold2. *Nat. Commun.* **2022**, *13*, 1265. [CrossRef] [PubMed]
125. Evans, R.; O'Neill, M.; Pritzel, A.; Antropova, N.; Senior, A.; Green, T.; Žídek, A.; Bates, R.; Blackwell, S.; Yim, J.; et al. Protein complex prediction with AlphaFold-Multimer. *bioRxiv* **2021**. [CrossRef]
126. Humphreys, I.R.; Pei, J.; Baek, M.; Krishnakumar, A.; Anishchenko, I.; Ovchinnikov, S.; Zhang, J.; Ness, T.J.; Banjade, S.; Bagde, S.R.; et al. Computed structures of core eukaryotic protein complexes. *Science* **2021**, *374*, eabm4805. [CrossRef]
127. Shannon, C. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [CrossRef]
128. Cover, T.; Thomas, J. *Elements of Information Theory*; Wiley: New York, NY, USA, 1991.
129. Orchard, S.; Kerrien, S.; Abbani, S.; Aranda, B.; Bhate, J.; Bidwell, S.; Bridge, A.; Briganti, L.; Brinkman, F.S.L.; Cesareni, G.; et al. Protein interaction data curation: The International Molecular Exchange (IMEx) consortium. *Nat. Methods* **2012**, *9*, 345–350. [CrossRef]
130. Tyson, J.J.; Chen, K.C.; Novak, B. Sniffers, buzzers, toggles and blinkers: Dynamics of regulatory and signaling pathways in the cell. *Curr. Opin. Cell Biol.* **2003**, *15*, 221–231. [CrossRef]
131. Kremling, A.; Saez-Rodriguez, J. Systems biology—An engineering perspective. *J. Biotechnol.* **2007**, *129*, 329–351. [CrossRef]

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.