Unveiling the Biomarkers of Cancer and COVID-19 and Their Regulations in Different Organs by Integrating RNA-Seq Expression and Protein–Protein Interactions

Nimisha Ghosh,*# Indrajit Saha,*# and Dariusz Plewczynski

ABSTRACT: Cancer and COVID-19 have killed millions of people worldwide. COVID-19 is even more dangerous to people with comorbidities such as cancer. Thus, it is imperative to identify the key human genes or biomarkers that can be targeted to develop novel prognosis and therapeutic strategies. The transcriptomic data provided by the next-generation sequencing technique makes this identification very convenient. Hence, mRNA (messenger ribonucleic acid) expression data of 2265 cancer and 282 normal patients were considered, while for COVID-19 assessment, 784 and 425 COVID-19 and normal patients were taken, respectively. Initially, volcano plots were used to identify the up- and down-regulated genes for both cancer and COVID-19. Thereafter, protein–protein interaction (PPI) networks were prepared by combining all the up- and down-regulated genes for each of cancer and COVID-19. Subsequently, such networks were analyzed to identify the top 10 genes with the highest degree of connection to provide the biomarkers. Interestingly, these genes were all up-regulated for cancer, while they were down-regulated for COVID-19. This study had also identified common genes between cancer and COVID-19, all of which were up-regulated for cancer, while they were down-regulated for COVID-19. This analysis revealed that FN1 was highly up-regulated in different organs for cancer, while EEF2 was dysregulated in most organs affected by COVID-19. Then, functional enrichment analysis was performed to identify significant biological processes. Finally, the drugs for cancer and COVID-19 biomarkers and the common genes between them were identified using the Enrichr online web tool. These drugs include lucanthone, etoposide, and methotrexate, targeting the biomarkers for cancer, while paclitaxel is an important drug for COVID-19.

INTRODUCTION

According to the W.H.O. information, cancer is the second leading cause of death worldwide. However, an early detection can result in a lower mortality rate. When a person is afflicted with cancer, the normal cells change to tumorous ones; therefore, identification of the genes that are involved in this transformation is very crucial for effective treatment. Though there are many methods that have been proposed for the investigation of cancer, human genome sequencing developed by the Human Genome Project was a landmark in the development of cheaper techniques but with a higher throughput to obtain a comprehensive knowledge of the entire genomes. The new techniques in the Human Genome Project have been included in next-generation sequencing (NGS) as well. In this regard, gene expression from microarray data has been used for a long time for the identification of biomarkers in cancer. However, it is usually very noisy. On the other hand, RNA sequencing (RNA-seq) can find a very low level expression of genes; thus, its development as part of the NGS technology is very important as it can produce expression data with very low noise. There are many studies in the literature such as refs 5–67, which have used RNA-seq data for the analysis of various types of cancer. RNA-seq can rapidly sequence and analyze transcriptomic data to identify potential biomarkers in cancer, a possible biomarker type being mRNA. mRNA is a type of RNA molecule that is able to carry genetic information from the DNA nucleus to the ribosome, where the amino acid sequence of the protein products of gene expression is specified by the mRNA’s sequence. However, due to the complex and high dimensional nature of the data, it is a non-trivial task to analyze the same. In this regard, tools such as BioExpress and Oncomine are proposed in order to search for biomarkers.

Meanwhile, COVID-19, the disease caused by SARS-CoV-2, has been disrupting our lives for more than 2 years now. By August 2022, almost 6.5 million people have died worldwide due to this virus. This can be attributed to the fact that SARS-CoV-2 has several variants due to its genetic mutations. Moreover, COVID-19 is especially detrimental to people suffering from comorbidities such as cancer and neurological diseases. Thus, identifying biomarkers in COVID-19 is a very significant step in finding an effective treatment for the same, and the contribution of transcriptome analysis is noteworthy here as well. In ref 14, Hasan et al. considered two RNA-seq data sets and one

Received: July 12, 2022
Accepted: October 13, 2022
Published: November 18, 2022

Metrics & More | Article Recommendations | Supporting Information

© 2022 The Authors. Published by American Chemical Society

https://doi.org/10.1021/acsomega.2c04389
ACS Omega 2022, 7, 43589–43602

43589
microarray data set of COVID-19 for peripheral blood mononuclear cells (PBMCs) to identify the biomarkers. Through their analysis, they concluded that blood cells can be used to diagnose COVID-19 as well as to develop corresponding drugs. Auwul et al.\textsuperscript{15} have analyzed PBMC GSE152418 and CRA002390 data sets using gene co-expression analysis. Through their analysis, they identified four key gene modules and a hub gene signature using module membership statistics and PPI networks. They have also identified drugs by considering drug–gene interaction analysis. Hasankhani et al.\textsuperscript{16} have performed weighted gene co-expression network analysis (WGCNA) on RNA-seq data from PBMC for healthy persons and 17 mild and severe COVID-19 patients. The co-expression analysis revealed that 72% of modules that were identified in healthy samples were altered by SARS-CoV-2. In their work, many transcriptional regulatory factors with important immunoregulatory roles in SARS-CoV-2 infection were also identified, which included NFKB1, HIF1A, AHR, and TP53. Sagulkoo et al.\textsuperscript{17} proposed a multi-level biological network analysis framework to provide candidate drugs targeting the key genes using the drug–gene interaction network and structural analysis as well as key gene identification via protein–protein interaction (PPI) network analysis and survival analysis based on differentially expressed genes (DEGs) in leukocyte transcriptomic profiles. Their analysis revealed CDC25A, GUSB, MYB1L2, and SDAD1 as key genes in severe COVID-19. Medini et al.\textsuperscript{18} have analyzed RNA-seq data sets for three blood data sets with 48 healthy and 119 afflicted patients and two respiratory tract data sets with 157 healthy and 524 affected patients. In their work, they have found reduced mtDNA (mitochondrial DNA) gene expression in blood. To assess the impact of SARS-CoV-2, Park et al.\textsuperscript{19} have used shotgun metatranscriptomics (total RNA-seq) to profile human tissues in 39 patients who have succumbed to COVID-19. Their study revealed a marked disruption of cellular and transcriptional programs among COVID-19 and normal patients.

Motivated by the literature, in this work, we considered mRNA expression data of 2265 cancer and 282 normal patients for 16088 human genes common among different organs such as bladder, breast, head and neck, esophagus, liver, kidney, lung, and stomach, while for COVID-19, mRNA expression data of 784 COVID-19 and 425 normal patients across 14412 human genes for nasopharynx, blood, respiratory tract, lung, heart, liver, intestine, stomach, eye, kidney, brain, pancreas, and uterus were taken into consideration. In order to identify the corresponding up- and down-regulated genes, respective volcano plots were created for both cancer versus normal and COVID-19 versus normal cases. Thereafter, the top 10 genes with the highest degrees (biomarkers) were identified using PPIs by combining the respective up- and down-regulated genes. For both cancer and COVID-19 afflicted patients, the identification of these biomarkers is important for early detection as well as effective treatment for the diseases. The identified biomarkers for cancer were FN1, UBE2C, CCNB1, CDK1, MAD2L1, AURKA, TOP2A, TPX2, NUSAP1, and KIF11, while for COVID-19, such biomarkers were EEF2, NDUFB7, NHP2, RPL9, MRPL15, RP55, RPS15, UQCRQ, RPL35, and RPS9. Furthermore, the regulation of the different biomarkers in the organs affected by cancer and COVID-19 are reported through heatmaps. Moreover, the pathway analysis of these biomarkers is reported using KEGG, while their biological significance is shown through gene ontology (GO) enrichment analysis. Several enriched pathways for the biomarkers of cancer include the p53 signaling pathway and human immunodeficiency virus 1 infection, while for COVID-19, the induced pathways include coronavirus disease, diabetic cardiomyopathy, Parkinson disease, and Alzheimer disease. This study has also identified up-regulated genes such as CXCL10, CXCL9, and IDO1 to be the common genes between cancer and COVID-19. Finally, drugs targeting the identified biomarkers as well as the common genes are reported for both cancer and COVID-19 using the Enrichr online web tool.\textsuperscript{20,21} In this regard, luncanthone, etoposide, and methotrexate were some of the drugs targeting the biomarkers for cancer, while paclitaxel targets the biomarkers for COVID-19. Thus, this work summarizes the biomarkers that can be investigated further to combat both cancer and COVID-19.

\section*{MATERIALS AND METHODS}

In this section, the data preparation is elaborated at first, which is then followed by the discussion on the pipeline of the proposed work.

\textbf{Data Preparation.} In this work, initially, mRNA expression data of 2265 cancer and 282 normal patients were collected from the respective up- and down-regulated genes. For both cancer and COVID-19 afflicted patients, the identification of these

\begin{table}[h!]
\centering
\caption{Statistics for the Cancer Data Set}
\begin{tabular}{|l|l|l|l|l|l|l|}
\hline
\textbf{cancer type} & \textbf{abbreviation of cancer type} & \textbf{organ} & \textbf{number of genes} & \textbf{tumor/cancer} & \textbf{normal} \\
\hline
bladder urothelial carcinoma & BLCA & bladder & 18539 & 56 & 11 \\
breast invasive carcinoma & BRCA & breast & 18637 & 778 & 100 \\
head and neck squamous cell carcinoma & HNSC & head and neck & 18692 & 263 & 31 \\
esophageal carcinoma & ESCA & esophagus & 21657 & 185 & 13 \\
liver hepatocellular carcinoma & LIHC & liver & 18015 & 17 & 9 \\
kidney renal clear cell carcinoma & KIRC & kidney & 18691 & 469 & 68 \\
lung squamous cell carcinoma & LUSC & lung & 18936 & 223 & 17 \\

\end{tabular}
\end{table}

\begin{table}[h!]
\centering
\caption{Statistics for the COVID-19 Data Set}
\begin{tabular}{|l|l|l|l|l|l|}
\hline
\textbf{GEOID} & \textbf{organ} & \textbf{number of genes} & \textbf{COVID-19} & \textbf{normal} \\
\hline
GSE163151 & blood & 21952 & 145 & 113 \\
GSE164332 & brain & 57996 & 9 & 7 \\
GSE164073 & eye & 25222 & 9 & 9 \\
GSE162736 & heart & 23194 & 24 & 24 \\
GSE159201 & intestine & 33550 & 12 & 12 \\
GSE173707 & kidney & 24975 & 9 & 9 \\
GSE151803 & liver & 22316 & 12 & 9 \\
GSE147507 & lung & 20748 & 21 & 29 \\
GSE152075 & nasopharynx & 19744 & 430 & 54 \\
GSE165880 & pancreas & 22057 & 6 & 6 \\
GSE156063 & respiratory tract & 15811 & 93 & 141 \\
GSE153684 & stomach & 26501 & 9 & 9 \\
GSE171995 & uterus & 19715 & 5 & 3 \\
\hline
\end{tabular}
\end{table}
the cancer Genome Atlas (TCGA), while mRNA expression data for 784 COVID-19 and 425 normal patients were downloaded from gene expression omnibus (GEO) of NCBI. The eight organs considered for cancer were bladder, breast, head and neck, esophagus, liver, kidney, lung, and stomach, while for COVID-19, the 13 considered organs were
nasopharynx, blood, respiratory tract, lung, heart, liver, intestine, stomach, eye, kidney, brain, pancreas, and uterus. This resulted in 16088 and 14412 genes for cancer and COVID-19, respectively. The statistics for cancer data are reported in Table 1, while that for COVID-19 are presented in Table 2. The expression data of COVID-19 were also used to prepare COVID19db by Zhang et al.24 All the expression data are provided in the Supporting Information as excel files. Please note that in this work, the malignant tumors were considered and referred to as cancer throughout the manuscript.

**Pipeline of the Work.** The primary motivation of this work is to identify the human genes as biomarkers that are mostly affected by different types of cancer and COVID-19 diseases. The pipeline of the work is provided in Figure 1. The pipeline provides the different steps involved in the identification of biomarkers. In this regard, mRNA expression data of 2265 cancer and 282 normal patients were considered with 16088 genes, while mRNA expression data for 784 COVID-19 and 425 normal patients were considered with 14412 genes (Step 1). Then, in Step 2, the corresponding up- and down-regulated genes were identified using volcano plots on these expression data of cancer and COVID-19. Thereafter, the corresponding PPIs were identified using STRING database3 by combining the up- and down-regulated genes for each of cancer and COVID-19.

![Figure 3](image-url)
cases (Step 3). Based on the degree, the top 10 genes were then selected as biomarkers for each of the case in Step 4. Once the biomarkers were identified, their regulations were analyzed in the different organs for both cancer and COVID-19. Moreover, their corresponding pathways were explored using KEGG, while their biological significance had been reported through GO enrichment analysis. Finally, FDA-approved drugs targeting the biomarkers for both cancer and COVID-19 were identified using the Enrichr4 tool. The aforementioned analysis was explored in Step 6.

**RESULTS**

This work was executed according to the pipeline shown in Figure 1. To carry out the experiments in this work, MATLAB R2021a was used on an Intel Core i5-8250U CPU @ 1.80 GHz machine with 8 GB RAM and Windows 10 operating system. Initially, we had used mRNA expression data of cancer and COVID-19 patients along with the corresponding normal individuals to identify the respective up- and down-regulated genes using the volcano plot, where the p-value was less than 0.05 and the log fold change (FC) value was greater than 3. This resulted in 107 and 151 up- and down-regulated genes for cancer, respectively, while for COVID-19, the statistics was 28 and 256, respectively. The volcano plots for cancer versus normal and COVID-19 versus normal cases are shown in Figure 2a,b, respectively. Based on the FC values derived from the volcano plot, the top 10 up- and down-regulated genes for cancer versus normal and COVID-19 versus normal cases are reported in Supporting Information, Tables S1 and S2, respectively. As can be seen from Table S1, MMP11 with a p-value of $1.30 \times 10^{-212}$ and an FC value of 24.82 is the top ranked up-regulated gene for the cancer versus normal case, while the corresponding top ranked down-regulated gene is ADH1B with a p-value of 2.21E-110 and an FC value of $-52.24$. On the other hand, MTRNR2L8 with a p-value of $6.49 \times 10^{-65}$ and an FC value of 134.59 and RPS21 with a p-value of $6.18 \times 10^{-59}$ and an FC value of $-7.06$ are the top ranked up- and down-regulated genes for the COVID-19 versus normal case, respectively. The total list of up- and down-regulated genes for both cancer and COVID-19 are provided in the Supporting Information, Table S3.

Once the up- and down-regulated genes were identified, they were combined together, and their PPIs were evaluated for both cancer and COVID-19. This led to the identification of 219 genes for cancer and 257 genes for COVID-19. The top most biomarker for cancer is $\text{UBC}$, with a p-value of $2.84 \times 10^{-53}$ and an FC value of 134.59, while for COVID-19, the top most biomarker is $\text{CDK1}$, with a p-value of $4.13 \times 10^{-54}$ and an FC value of 10.40. The top 10 biomarkers for cancer are $\text{CDK1}$, $\text{TOP2A}$, $\text{TPX2}$, $\text{CCNB1}$, $\text{UB2L1}$, $\text{CCNB1}$, $\text{CDK1}$, $\text{TOP2A}$, $\text{TPX2}$, and $\text{CCNB1}$, respectively, reported in Tables 3 and 4. The top 10 biomarkers for COVID-19 are $\text{UBC}$, $\text{NUSAP1}$, $\text{UBC}$, $\text{UBC}$, $\text{UBC}$, $\text{UBC}$, $\text{UBC}$, $\text{UBC}$, $\text{UBC}$, and $\text{UBC}$, respectively.

| Table 3. Top 10 Genes as Biomarkers for Cancer |
|-----------------------------------------------|
| gene  | degree | regulation | p-value  | FC value |
|------|--------|------------|---------|----------|
| FN1  | 54     | up         | $1.00 \times 10^{-46}$ | 3.94     |
| UBE2C| 48     | up         | $9.21 \times 10^{-93}$ | 11.76    |
| CCNB1| 47     | up         | $7.05 \times 10^{-57}$ | 4.55     |
| CDK1 | 47     | up         | $5.58 \times 10^{-91}$ | 5.48     |
| MAD2L1| 47     | up         | $4.50 \times 10^{-101}$ | 4.03     |
| AURKA| 45     | up         | $9.60 \times 10^{-120}$ | 4.93     |
| TOP2A| 45     | up         | $1.28 \times 10^{-90}$ | 10.40    |
| TPX2 | 45     | up         | $4.19 \times 10^{-97}$ | 10.10    |
| NUSAP1| 44     | up         | $1.89 \times 10^{-95}$ | 6.99     |
| KIF11| 44     | up         | $1.00 \times 10^{-90}$ | 4.38     |

| Table 4. Top 10 Genes as Biomarkers for COVID-19 |
|-----------------------------------------------|
| gene  | degree | regulation | p-value  | FC value |
|------|--------|------------|---------|----------|
| EEF2 | 47     | down       | $1.83 \times 10^{-44}$ | -3.58    |
| NDUFB7| 47     | down       | $1.37 \times 10^{-56}$ | -6.73    |
| NHP2 | 45     | down       | $9.52 \times 10^{-41}$ | -3.63    |
| RPL9 | 44     | down       | $2.93 \times 10^{-29}$ | -3.16    |
| MRPL15| 43     | down       | $1.24 \times 10^{-39}$ | -3.05    |
| RPS5 | 41     | down       | $6.98 \times 10^{-38}$ | -4.04    |
| RPS15| 40     | down       | $5.78 \times 10^{-33}$ | -3.24    |
| UQCRQ| 40     | down       | $2.29 \times 10^{-33}$ | -3.37    |
| RPL35| 39     | down       | $2.35 \times 10^{-32}$ | -5.10    |
| RPS9 | 39     | down       | $2.54 \times 10^{-34}$ | -3.81    |

![Figure 4](https://example.com/figure4.png)

Figure 4. Venn diagram to show the number of common genes between cancer and COVID-19.

| Table 5. Common Genes between Cancer and COVID-19 |
|-----------------------------------------------|
| common genes | degree | regulation | p-value  | FC value | degree | regulation | p-value  | FC value |
|---------------|--------|------------|---------|----------|--------|------------|---------|----------|
| CXCL9         | 13     | Up         | $3.83 \times 10^{-23}$ | 5.85     | 15     | Up         | $8.06 \times 10^{-47}$ | 4.80     |
| CXCL10        | 15     | Up         | $1.56 \times 10^{-32}$ | 6.48     | 24     | Up         | $1.63 \times 10^{-71}$ | 10.13    |
| IDO1          | 6      | Up         | $4.46 \times 10^{-46}$ | 3.95     | 4      | Up         | $4.21 \times 10^{-28}$ | 3.16     |
Figure 5. Heatmap to represent the regulation of the biomarkers in different organs for (a) cancer, (b) corresponding normal, (c) COVID-19, and (d) corresponding normal and regulation of common genes for (e) cancer, (f) corresponding normal, (g) COVID-19, and (h) Corresponding normal.
COVID-19 is EEF2, which is a down-regulated gene and has a degree of 47, a p-value of $1.83 \times 10^{-44}$, and an FC value of $-3.58$. Figure 4 shows the number of common genes between cancer (219) and COVID-19 (257), while Table 5 reports these common genes; they being CXCL9, CXCL10, and IDO1. All of these three genes are up-regulated in both cancer and COVID-19.

Based on the average expression values, the regulation of the biomarkers for cancer and corresponding normal patients for the different afflicted organs is represented in Figure 5a,b, respectively, where the colors represent the regulation; a darker color shows that a gene is highly regulated in a particular organ. For example, FN1 is highly regulated in the breast, head and neck, esophagus, liver, kidney, and stomach of a cancer patient as opposed to a normal individual. On the other hand, such regulation for COVID-19 and corresponding normal patients is shown in Figure 5c,d, respectively. For example, EEF2 is down-regulated in organs such as the respiratory tract, nasopharynx, liver, brain, and blood of a COVID-19 patient. These observations support our previous discussions on FN1 and EEF2. Moreover, the regulation of the common genes pertaining to cancer and COVID-19 are reported through Figure 5e,h. As can be seen from Figure 5e, CXCL10 is highly expressed in the lung of cancer patients. This is in line with the observation made by Mahmood et al. Further analysis for the average expression values for the biomarkers of cancer and COVID-19 are provided in Figure 6a,b, respectively, while such an analysis for the common genes is shown in Figure 6c,d. Figure 6a also corroborates the fact that all the biomarkers are up-regulated in cancer and down-regulated in COVID-19, while the three common genes are highly expressed in both cancer and COVID-19.

**DISCUSSION**

According to ref, DNA damage may be caused by the dysregulation of EEF2. Also, the possible association of down regulation of EEF2 with COVID-19 severity has been mentioned in ref. Please note that biomarkers for COVID-19 such as EEF2, NHP2, RPL9, MRPL15, RPSS, RPS15, RPL35, and RPS9 are all directly interacting with various SARS-CoV-2 proteins such as NSP7, Spike, Envelope, ORF6, ORF7a, ORF7b, ORF9b, and ORF10. As reported in the Results section, the three common genes that are up-regulated in both cancer and COVID-19 are CXCL9, CXCL10, and IDO1. The intratumoral accumulation of CXCL9 and CXCL10, which promotes tumor-infiltrating lymphocytes (TIL) chemotactic recruitment, may enhance TIL-dependent immune intervention in cancer. Following immune checkpoint blockade, CXCL9 has been
demonstrated to be necessary for antitumor immune responses. On the other hand, patients dying from SARS-CoV-2 have shown higher plasma levels of CXCL9. CXCL10 is responsible for immune responses in the lung and COVID-19 infection. On the other hand, ID O 1 can be considered to be an ideal target for cancer immunotherapy.

**KEGG Pathway Analysis.** Some important pathways for the different biomarkers of cancer and COVID-19 and the common genes between them are shown in Figure 7a–c, respectively. These results were collected from the Enrichr tool. The bubbles in the plots represent the number of genes (biomarkers) associated with each pathway; a smaller bubble represents a lower number of biomarkers, while a larger bubble indicates the opposite, and the colors are based on the corresponding adjusted p-value. As can be seen from Figure 7a, biomarkers for cancer are enriched in pathways, which include the p53 signaling pathway (CCNB1 and CDK1), human immunodeficiency virus 1 infection (CCNB1 and CDK1), and so on, with the corresponding adjusted p-values of $3.33 \times 10^{-3}$ and $1.82 \times 10^{-2}$, respectively. Biomarkers for COVID-19 are enriched in pathways such as coronavirus disease (RPS15, RPS9, RP55, RPL35, and RPL9), diabetic cardiomyopathy (NDUFB7 and...
UQCRQ), Parkinson disease (NDUFB7 and UQCRQ), Huntington disease (NDUFB7 and UQCRQ), Alzheimer disease (NDUFB7 and UQCRQ), pathways of neurodegeneration (NDUFB7 and UQCRQ), and so on with the corresponding adjusted p-values of $4.11 \times 10^{-7}$, $1.48 \times 10^{-2}$, $1.57 \times 10^{-2}$, $1.82 \times 10^{-2}$, $2.14 \times 10^{-2}$, and $3.16 \times 10^{-2}$, respectively. These results highlight that COVID-19 can aggravate the pathways for other diseases as well, leading to comorbidity. On the other hand, the common genes target pathways such as the chemokine signaling pathway (CXCL10, CXCL9), African trypanosomiasis (ID O 1), IL-17 signaling pathway, TNF signaling pathway, Hepatitis C, Influenza A, Epstein–Barr virus infection, coronavirus disease, and so on (all triggered by CXCL10) with the corresponding adjusted p-values of $1.27 \times 10^{-3}$, $1.46 \times 10^{-2}$, $2.18 \times 10^{-2}$, $2.33 \times 10^{-2}$, $2.97 \times 10^{-2}$, $2.98 \times 10^{-2}$, $3.23 \times 10^{-2}$, and $3.43 \times 10^{-2}$, respectively. Among these, chemokine, IL-17, and TNF signaling pathways are all pathways in cancer, while coronavirus disease is a pathway in COVID-19.

**GO Enrichment Analysis.** GO enrichment analysis was performed to understand the significance of the different interacting human genes in biological activities. Similar to KEGG pathways, the GO enrichment results were collected.

---

**Figure 8.** Significant GO biological processes corresponding to biomarkers for (a) cancer and (b) COVID-19 and (c) common genes between cancer and COVID-19.
from the Enrichr tool as well and considered only for the biomarkers. The results of the analysis for the different biomarkers of cancer and COVID-19 and the common genes in terms of biological processes are shown in Figure 8a−c, respectively, while those for cellular and molecular processes are shown in Supporting Information Figures S1 and S2. The results of the analysis for the different biological processes are shown in Figure 8a with adjusted p-value of 2.93 × 10−8, 3.24 × 10−9, and 1.84 × 10−7, respectively.

For COVID-19, such pathways for biomarkers such as RPS15, RPS9, RPP5, RPL35, EEF2, and RPL9 are as follows: translation (GO:0006412), SRP-dependent cotranslational protein targeting to membrane (GO:0006614), and cytoplasmic translation (GO:0002181) with an adjusted p-value of 8.02 × 10−9 for all and some significant pathways for the common genes are positive regulation of calcium ion transmembrane transport (GO:1904427), positive regulation of release of sequestered calcium ion into cytosol (GO:0051281), and positive regulation of calcium ion transport into cytosol (GO:0010524) with the adjusted p-value of 1.73 × 10−4.

**Drug-Targeting Biomarkers.** There are many drugs that are prescribed for cancer patients, while drugs to combat COVID-19 is still under development. In this regard, drug repurposing can be a viable alternative for effective identification

| disease | human genes | drugs | drug bank ID | treatment |
|---------|-------------|-------|--------------|-----------|
| cancer | FN1, UBE2C, CCNB1, CDK1, MAD2L1, AURKA, TOP2A, TPX2, NUSAP1, KIF11 | lucanthone | DB04967 | inhibits post-radiation DNA repair in tumor cells. |
| | | etoposide | DB00773 | it is also shown to be toxic to glioma cells by inhibiting autophagy. |
| | | methotrexate | DB00563 | a wide variety of cancers. |
| | | trifluvidine | DB00432 | chemotherapy for certain types of metastatic gastrointestinal cancers. |
| | | resveratrol | DB02709 | herpes simplex virus types 1 and 2, found to have potential anticancer properties. |
| | | belinostat | DB05015 | relapsed or refractory peripheral T-cell lymphoma (PTCL). |
| | | thalidomide | DB01041 | newly diagnosed multiple myeloma and erythema nodosum leprosum. |
| | | ciclopirox | DB01188 | mild to moderate onychomycosis of fingernails and toenails. It can be considered to be a novel chemotherapeutic for the treatment of colorectal cancer. |
| | | vinblastine | DB00570 | breast cancer, testicular cancer, neuroblastoma, Hodgkin’s and non-Hodgkins lymphoma, mycosis fungoides, histiocytosis, and Kaposi’s sarcoma. |
| COVID-19 | EEF2, NDUFB7, RPL9, MRPL15, RPS5, RPS15, UQCRQ, RPL35, RPS9 | disodium selenite | DB11127 | potential therapy in the prevention or management of atherosclerosis, reduces COVID-19. |
| | | midecamycin | DB13456 | a variety of infections caused by susceptible bacteria. |
| | | amikacin | DB00479 | infections caused by more resistant strains of Gram negative bacteria and some Gram positive bacteria. |
| | | paclitaxel | DB01229 | a potent inhibitor of main protease of SARS-CoV-2. |
| | | metformin | DB00331 | glycemic control in type 2 diabetes mellitus |
| | | hydrochloride | DB00634 | | |
| | | ambroxol | DB00742 | airway secretion clearance therapy, has anti SARS-CoV-2 activity. |
| | | clindamycin | DB01190 | serious infections caused by susceptible anaerobic, streptococcal, staphylococcal, and pneumococcal bacteria. |
| Common | CXCL10, CXCL9, IDO1 | imatinib | DB00619 | leukemia, myelodysplastic/myeloproliferative disease, systemic mastocytosis, hypereosinophilic syndrome, dermatofibrosarcoma protuberans, and gastrointestinal stromal tumors. |
| | | lucanthone | DB04967 | inhibits post-radiation DNA repair in tumor cells. |
| | | etoposide | DB00773 | it is also shown to be toxic to glioma cells by inhibiting autophagy. |
| | | methotrexate | DB00563 | a wide variety of cancers. |
| | | trifluvidine | DB00432 | chemotherapy for certain types of metastatic gastrointestinal cancers. |
| | | resveratrol | DB02709 | herpes simplex virus types 1 and 2, found to have potential anticancer properties. |
| | | belinostat | DB05015 | relapsed or refractory peripheral T-cell lymphoma (PTCL). |
| | | thalidomide | DB01041 | newly diagnosed multiple myeloma and erythema nodosum leprosum. |
| | | ciclopirox | DB01188 | mild to moderate onychomycosis of fingernails and toenails. It can be considered to be a novel chemotherapeutic for the treatment of colorectal cancer. |
| | | vinblastine | DB00570 | breast cancer, testicular cancer, neuroblastoma, Hodgkin’s and non-Hodgkins lymphoma, mycosis fungoides, histiocytosis, and Kaposi’s sarcoma. |
of a drug for COVID-19, and as such, the biomarkers can be considered as good target candidates. For both cancer and COVID-19, the drugs that interact with the biomarkers were identified using DSigDB in the Enrichr tool. The results for the corresponding drugs are reported in Table 6 along with the relevant drug IDs as collected from Drug Bank and the possible treatments. The corresponding drug–protein interactions are shown in Figure 9.

As can be seen from Table 6, lucanthone, etoposide, methotrexate, trifluridine, resveratrol, belinostat, ciclopirox, and vinblastine are few of the drugs that are used for treating several types of cancer and related to the identified biomarkers as well. On the other hand, a drug such as paclitaxel, which targets the biomarkers of COVID-19 such as EEF2 and NDUFB7, is not only used for cancer but is also used for the treatment of COVID-19. Al-Motawa et al. have also provided a scientific rationale for repurposing paclitaxel for the treatment of COVID-19.

In the past, paclitaxel has also been judged for antiviral activity, especially for viral helicase.

Drugs such as imatinib and decitabine targeting the common proteins (genes) of both cancer and COVID-19, which are used for the treatment of cancer, are also under trial for the treatment of COVID-19 as well. Roflumilast, which is another drug targeting the common proteins (genes), is used for decreasing the risk of exacerbations in patients with severe chronic obstructive pulmonary disease (COPD) and is also under trial for treating cancer and COVID-19.

**CONCLUSIONS**

Both cancer and COVID-19 have been responsible for many deaths around the globe. Though various types of cancer have been prevalent for a very long time, SARS-CoV-2, the virus causing COVID-19, has been around for more than 2 years now. In this work, we have identified the human genes as biomarkers that can be targeted for cancer and COVID-19 diseases. Though we conducted the experiments with mRNA expression data of cancer and COVID-19 to identify the up- and down-regulated genes, our main focus in this work was to identify the biomarkers...
using the PPI network as they are the ones that are connected to most of the other human proteins and consequently aid in the progression of a disease. Thus, we identified 10 biomarkers for both cancer and COVID-19 with the highest degrees and have also shown the regulation of each biomarker in the corresponding human organs affected by cancer and COVID-19. We have also reported the corresponding KEGG pathways and the results of GO enrichment analysis for the biomarkers. Finally, the different drugs targeting the biomarkers of cancer and COVID-19 are also reported in this work. Such drugs including lucanthone, etoposide, methotrexate, trifluridine, resveratrol, belinostat, ciclopirox, and vinblastine are targeting the biomarkers of cancer. As per the literature, these drugs are already in use for several types of cancer. Similarly, for the biomarkers of COVID-19, the identified drug paclitaxel is under trial as per the literature. Therefore, we hope that the identified biomarkers may help the ongoing research in cancer and COVID-19.

伦理审查和知情同意

The ethical approval or individual consent was not applicable.

可用性数据和材料

The Supporting Information of this work is available at “http://www.nitttrkol.ac.in/indrajit/projects/Cancer-COVID-19-Biomarkers/”.

版权许可

Not applicable.

相关联系信息

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.2c04389.

Top 10 up- and down-regulated genes for cancer versus normal, top 10 up- and down-regulated genes for COVID-19 versus normal, link and description of Supporting Information files, significant GO cellular processes corresponding to biomarkers for (a) cancer and (b) COVID-19, significant GO molecular processes corresponding to biomarkers for (a) cancer and (b) COVID-19 and (c) common genes between cancer and COVID-19, up- and down-regulated genes for cancer, up- and down-regulated genes for COVID-19, genes identified from the PPI of cancer data, and genes identified from the PPI of COVID-19 data (PDF).

作者信息

Nimisha Ghosh — Faculty of Mathematics, Informatics and Mechanics, University of Warsaw, Warsaw 02-097, Poland; Department of Computer Science and Information Technology, Institute of Technical Education and Research, Siksha ‘O’ Anusandhan (Deemed to Be University), Bhubaneswar 751030 Odisha, India; orcid.org/0000-0002-0697-6368; Email: ghosh.nimisha@gmail.com

Indrajit Saha — Department of Computer Science and Engineering, National Institute of Technical Teachers’ Training and Research, Kolkata 700106 West Bengal, India; Email: indrajit@nitttrkol.ac.in

作者贡献

N.G., I.S. and D.P. contributed equally.

资金

This work was carried out during the tenure of an ERCIM ‘Alain Bensoussan’ Fellowship Program awarded to Dr. Nimisha Ghosh. This work has also been partially supported by the CRG short-term research grant on COVID-19 (CVD/2020/009991) from the Science and Engineering Research Board (SERB), Department of Science and Technology, Govt. of India. This work was co-funded by the Warsaw University of Technology within the Excellence Initiative: Research University (IDUB) programme. This work has also been co-supported by the Polish National Science Centre (2019/35/O/ST6/02484 and 2020/37/B/NZ2/03757).

笔记

The authors declare no competing financial interest.

参考文献

(1) Yobe, S.; Thyagarajan, B. Review of Clinical Next-Generation Sequencing. Arch Pathol Lab Med 2017, 141, 1544–1557.
(2) Herzyk, P. Next-Generation Sequencing. Handbook of Pharmacogenomics and Stratified Medicine, 2014; Elsevier, pp 125–145.
(3) Fahrudin, T. M.; Syarif, I.; Barabkha, A. R. Ant colony algorithm for feature selection on microarray datasets. International Electronics Symposium (IES), 2016, pp 351–356.
(4) Coleto-Alcudia, V.; Vega-Rodriguez, M. A. A multi-objective optimization approach for the identification of cancer biomarkers from RNA-seq data. Expert Systems with Applications 2022, 193, 116480.
(5) Chen, H.; Li, C.; Peng, X.; Zhou, Z.; Weinstein, JN; au, fnm; Liang, H. A Pan-Cancer Analysis of Enhancer Expression in Nearly 9000 Patient Samples. Cell 2018, 173, 386–e12.
(6) Hoadley, K.A.; Yau, C.; Hinoue, T.; Wolf, DM; Lazar, AJ; Dril, E.; Shen, R.; Taylor, AM; Cherniack, AD; Thorsson, V.; Abkarian, R.; Bowlby, R.; Wong, CK; Wisnerowicz, M.; Sanchez-Vega, F.; Robertson, AG; Schneider, BG; Lawrence, MS; Noushmehr, H.; Malta, TM; au, fnm; Stuart, JM; Benz, CC; Laird, PW Cell-Of-Origin Patterns Dominate the Molecular Classification of 10,000 Tumors from 33 Types of Cancer. Cell 2018, 173, 291–e6.
(7) Denkiewicz, M.; Saha, I.; Rakshit, S.; Sarkar, J. P.; Plewczynski, D. Identification of Breast Cancer Subtype Specific MicroRNAs Using Survival Analysis to Find Their Role in Transcriptomic Regulation. Front. Genet. 2019, 10, 01047.
(8) Saha, I.; Rakshit, S.; Denkiewicz, M.; Sarkar, J. P.; Maity, D.; Maulik, U.; Plewczynski, D. Survival Analysis with the Integration of RNA-Seq and Clinical Data to Identify Breast Cancer Subtype Specific Genes; Pattern Recognition and Machine Intelligence, 2019, pp 139–146.
(9) Wan, Q.; Dingerdissen, H.; Fan, Y.; Gulzar, N.; Pan, Y.; Wu, T.-J.; Yan, C.; Zhang, H.; Mazumder, R. BioXpress: an integrated RNA-seq-derived gene expression database for pan-cancer analysis. Database 2015, 2015, bay019.
(10) Rhodes, D. R.; Kalyana-Sundaram, S.; Mahavisno, V.; Varambally, N.; Yu, Y.; Briggs, T.-J.; Barrette, C.; Anset, H.; Kincead-Beal, R.; Kulkarni, P.; Varambally, S.; Ghosh, D.; Chinnaiyan, A. M. Oncomine 3.0: Genes, Pathways, and Networks in
a Collection of 18,000 Cancer Gene Expression Profiles. 

Ghosh, N.; Nandi, S.; Saha, I. A review on evolution of emerging SARS-CoV-2 variants based on spike glycoprotein. 

Int. Immunopharmacol. 2022, 105, 108565.

Saha, I.; Ghosh, N.; Sharma, N.; Nandi, S. Hotspot Mutations in SARS-CoV-2. 

Front. Genet. 2021, 12, 753440.

Saha, I.; Ghosh, N.; Pradhan, A.; Sharma, N.; Maity, D.; Mitra, K. Whole genome analysis of more than 10,000 SARS-CoV-2 virus unveils global genetic diversity and target region of NSP6. 

Briefings in Bioinformatics 2021, 22, 1106–1121.

Hasan, M. I.; Rahman, M. H.; Islam, M. B.; Islam, M. Z.; Hossain, M. A.; Moni, M. Systems Biology and Bioinformatics approach to Identify blood based signatures molecules and drug targets of patient with COVID-19. 

Informatics in Medicine Unlocked 2022, 28, 100840.

Awuwel, M. R.; Rahman, M. R.; Gov, E.; Shahjaman, M.; Moni, M. A. Bioinformatics and machine learning approach identifies potential drug targets and pathways in COVID-19. 

Briefings in Bioinformatics 2021, 22, 22.
(45) Kieliszek, M. Selenium in the Prevention of SARS-CoV-2 and Other Viruses. *Biol. Trace Elem. Res.* 2022, 19, 1–8.
(46) Ahmed, M. Z.; Zia, Q.; Haque, A.; Alqahtani, A. S.; Almarfadi, O. M.; Banawas, S.; Alqahtani, M. S.; Ameta, K. L.; Haque, S. Aminoglycosides as potential inhibitors of SARS-CoV-2 main protease: an in silico drug repurposing study on FDA-approved antiviral and anti-infection agents. *J Infect Public Health* 2021, 14, 611–619.
(47) Bocci, G.; Bradfute, S. B.; Ye, C.; Garcia, M. J.; Parvathareddy, J.; Reichard, W.; Surendranathan, S.; Bansal, S.; Bologa, C. G.; Perkins, D. J.; et al. Virtual and In Vitro Antiviral Screening Revive Therapeutic Drugs for COVID-19. *ACS Pharmacol. Transl. Sci.* 2020, 3, 1278–1292.
(48) Al-Horani, R. A.; Kar, S.; Aliter, K. F. Potential Anti-COVID-19 Therapeutics that Block the Early Stage of the Viral Life Cycle: Structures, Mechanisms, and Clinical Trial. *Int. J. Mol. Sci.* 2020, 21, 21.
(49) Li, G.; Ruan, S.; Zhao, X.; Liu, Q.; Dou, Y.; Mao, F. Transcriptomic signatures and repurposing drugs for COVID-19 patients: findings of bioinformatics analyses. *Comput Struct Biotechnol J* 2021, 19, 1–15.
(50) S, S. L. J.; V, R. Scope of adjuvant therapy using roflumilast, a PDE-4 inhibitor against COVID-19. *Pulm. Pharmacol. Ther.* 2021, 66, 101978.