Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Transcriptional profiles and common genes link lung cancer with the development and severity of COVID-19

S.S. Cury a,1, J.S. Oliveira a,1, C.A.O. Biagi-Júnior b,c,d, W.A. Silva Jr b,c,d, P.P. Reis e, O. Cabral-Marques f,g,h,i, E.N. Hasimoto e, P.P. Freire a,b,i, R.F. Carvalho a,b,i,

a Department of Structural and Functional Biology, Institute of Biosciences, São Paulo State University (UNESP), Botucatu, SP, Brazil
b Department of Genetics, Ribeirão Preto Medical School, University of São Paulo (USP), Ribeirão Preto, SP, Brazil
c Center for Cell-Based Therapy (CEPID/FAPESP), National Institute of Science and Technology in Stem Cell and Cell Therapy (INCT/CNPq), Regional Blood Center of Ribeirão Preto, Ribeirão Preto, SP, Brazil
d Institute for Cancer Research (IPIC), Guarapuava, PR, Brazil
e Department of Surgery and Orthopedics, Faculty of Medicine, São Paulo State University (UNESP), Botucatu, SP, Brazil
f Department of Immunology, Institute of Biomedical Sciences, University of São Paulo (USP), São Paulo, SP, Brazil
g Department of Clinical and Toxicological Analyses, School of Pharmaceutical Sciences, University of São Paulo (USP), São Paulo, SP, Brazil
h Network of Immunity in Infection, Malignancy, and Autoimmunity (NIIMA), Universal Scientific Education and Research Network (USERN), São Paulo, SP, Brazil
i Department of Pharmacy and Postgraduate Program of Health and Science, Federal University of Rio Grande do Norte (UFRN), Natal, RN, Brazil

ARTICLE INFO

Keywords:
SARS-CoV-2
NSCLC
SCLC
AS49
Calu-3
Host-virus interactome

ABSTRACT

Lung cancer patients with COVID-19 present an increased risk of developing severe disease and, consequently, have poor outcomes. Determining SARS-CoV-2-host interactome in lung cancer cells and tissues, infected or uninfected with SARS-CoV-2, may reveal molecular mechanisms associated with COVID-19 development and severity in lung cancer patients. Here, we integrated transcriptome data of lung tumors from patients with small- or non-small cell lung cancer (SCLC and NSCLC) and normal lung and lung cancer cells infected with SARS-CoV-2. We aimed to characterize molecular mechanisms potentially associated with COVID-19 development and severity in lung cancer patients and to predict the SARS-CoV-2-host cell interactome. We found that the gene expression profiles of lung cell lines infected with SARS-CoV-2 resemble more primary lung tumors than non-malignant lung tissues. In addition, the transcriptomic-based interactome analysis of SCLC and NSCLC revealed increased expression of cancer genes BRCA1 and CENPF, whose proteins are known or predicted to interact with the SARS-CoV-2 spike glycoprotein and helicase, respectively. We also found that TRIB3, a gene coding a putative host-SARS-CoV-2 interacting protein associated with COVID-19 infection, is co-expressed with the up-regulated genes MTHFD2, ADM2, and GPT2 in all tested conditions. Our analysis identified biological processes such as amino acid metabolism and angiogenesis and 22 host mediators of SARS-CoV-2 infection and replication that may contribute to the development and severity of COVID-19 in lung cancers.

Abbreviations: SARS-CoV-2, Severe acute respiratory syndrome coronavirus 2; COVID-19, Coronavirus disease 2019; ACE2, Angiotensin-converting enzyme 2; LUAD, Lung adenocarcinoma; LUSC, Squamous cell carcinoma; DEGs, Differentially expressed genes; NSCLC, Non-small cell lung cancer; SCLC, Small-cell lung cancer; GEO, Gene Expression Omnibus; GSEA, Gene Set Enrichment Analysis; P-HIPSTer, Pathogen–Host Interactome Prediction using Structure Similarity; TCGA, The Cancer Genome Atlas; GTEx, Genotype-Tissue Expression; PCA, Principal component analysis; Nps13, Helicase; COPD, Chronic obstructive pulmonary disease; FC, Fold change; P.adj, adjusted P-value.

* Corresponding authors at: Department of Immunology, Institute of Biomedical Sciences - University of São Paulo, Lineu Prestes Avenue, 1730 São Paulo, Brazil (P.P. Freire), Department of Structural and Functional Biology, Institute of Biosciences, São Paulo State University (UNESP), CEP: 18.618-689, Botucatu, São Paulo, Brazil (R.F. Carvalho).

E-mail addresses: paula.freire@unesp.br (P.P. Freire), robson.carvalho@unesp.br (R.F. Carvalho).

1 These authors contributed equally.
1. Introduction

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) causes the coronavirus disease 2019 (COVID-19) (Wu et al., 2021; Zhou et al., 2020a). The first case was identified at the end of 2019 in Wuhan, China, and COVID-19 rapidly became a worldwide pandemic (Dong et al., 2020). Although most patients with SARS-CoV-2 infections are asymptomatic or experience mild symptoms, others may develop a severe illness requiring hospitalization (Gu et al., 2020; Kedor et al., 2020; Wu et al., 2020). Severe disease is characterized by pneumonia, acute respiratory distress syndrome, cytokine storm, lymphopenia, venous thromboembolism, multiple organ failure, and death (Mehta et al., 2020; Tan et al., 2020; Wu et al., 2020). The main determinants of COVID-19 severity are gender (men are more vulnerable than women), older age, laboratory abnormalities (e.g., lymphopenia and neutrophilia), and comorbidities such as cardiovascular diseases, chronic respiratory diseases, diabetes, and hypertension (Gu et al., 2020; Huang et al., 2020; Williamson et al., 2020; Zhang et al., 2020b; Zhou et al., 2020a). Cancer patients with COVID-19 also have high rates of hospitalization and more severe outcomes due to poorly understood mechanisms (Albiges et al., 2020; Dai et al., 2020; Robilotti et al., 2020). This increased susceptibility may be due to specific demographic characteristics and associated comorbidities, such as advanced age, smoking, diabetes, and hypertension (Derosa et al., 2020; Kuderer et al., 2020; Lee et al., 2020). In addition, side effects of chemotherapy, such as cardiomyopathy, arterial hypertension, systemic immunosuppression, and accelerated cellular senescence may also contribute to the worsening of COVID-19 in cancer patients (Albiges et al., 2020; Derosa et al., 2020).

The highest frequency of severe events occurs in COVID-19 patients with hematologic and lung cancers and/or metastasis (Dai et al., 2020). Specifically, patients with lung cancers have worse COVID-19 outcomes (Luo et al., 2020). Interestingly, the SARS-CoV-2 virus is more likely to infect lung cancer patients due to poorly understood reasons, leading to increased mortality (Wu et al., 2021). Increased expression of the angiotensin-converting enzyme 2 (ACE2) has been postulated as the main reason, as demonstrated for other lung chronic diseases (Bui et al., 2021; Hoffmann et al., 2020; Muus et al., 2021). SARS-CoV-2 cell entry depends on binding the viral spike (S) proteins to ACE2 and the serine protease TMPRSS2 for S protein priming (Hoffmann et al., 2020). ACE2 is highly expressed in lung adenocarcinoma (LUAD) and squamous cell carcinoma (LUSC) compared to normal tissues (Pinto et al., 2020; Zhang et al., 2020a). The transcriptional profile induced by SARS-CoV-2 infection shares 308 differentially expressed genes (DEGs) with lung adenocarcinoma cell lines, normal human bronchial epithelial cells, and lung tumor samples (Chen et al., 2021). The mapped network of SARS-CoV-2-host protein–protein interactions in human cancers revealed 46 proteins originated from genes that are relevant or with indications of a role in cancer (Tutuncuoglu et al., 2020). COVID-19 and lung adenocarcinomas share gene signatures associated with immune infiltration and checkpoints (Jiang et al., 2022). Twenty-three compounds currently used or investigated in clinical trials for cancer treatment could be repurposed to inhibit these newly identified virus-host protein–protein interactions (Tutuncuoglu et al., 2020). Therefore, comparisons of the transcriptional profiles of proteins associated with SARS-CoV-2 entry, replication, and host-pathogen interactions in lung cancer cells infected with SARS-CoV-2 with lung tumors may reveal new mechanisms related to the increased susceptibility to SARS-CoV-2 infections and COVID-19 severity in lung cancer patients. Moreover, combined transcriptomic analyses from the two histological subtypes - non-small cell lung cancer (NSCLC) and small-cell lung cancer (SCLC) - are also lacking and may help identify molecular mechanisms associated with SARS-CoV-2.

Here, we combined lung cancer and non-cancer transcriptomic data from cells and tissues infected or uninfected with SARS-CoV-2. This strategy indicated candidate molecular mechanisms that may explain COVID-19 development and severity in lung cancer patients and predicted virus-host protein–protein interactions. First, we found increased expression of cancer-associated genes BRCA1 and CENPF in both SCLC and NSCLC, which encode host proteins interacting with SARS-CoV-2. By integrating the altered gene expression of lung carcinomas and cell lines infected with SARS-CoV-2, we found molecular mechanisms of SARS-CoV-2 infection in lung cancer. These results provide a plausible explanation for the increased susceptibility to severe COVID-19 in lung cancer patients. Ultimately, these results highlight targets for developing therapeutic strategies for these patients.

2. Material and methods

2.1. Differential expression analysis of lung cancer tissue and lung cancer cell lines infected with SARS-CoV-2

We compared RNA sequencing (RNA-Seq) data of 199 NSCLC tissues with 19 paired normal lung tissues and 79 SCLC tissues with seven healthy normal lung tissues (Fig. 1A and Table S1). These datasets are available in Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo) under the accession number GSE81089 (Mezhuyetski et al., 2018) and GSE60052 (Ilijan et al., 2016), and accessible via Biojupies platform. We next compared the RNA-Seq data of the cancer cell lines A549 and Calu-3 infected with SARS-CoV-2 for 24 h with mock (controls) from GSE147507 (Blanco-Melo et al., 2020) (Fig. 1B). Finally, we used the BioJupies platform (https://amp.pharm.mssm.edu/biojupies/) (Torre et al., 2018) to identify the DEGs (Log2 of fold change ≥ |1| and false discovery rate (FDR) < 0.05).

2.2. Enrichment analysis

We used the online tool EnrichR (https://maayanlab.cloud/Enrichr/) (Chen et al., 2013; Kuleshov et al., 2016) to perform functional enrichment analysis of the gene lists generated in the clustering and differential expression analyses. We selected the Top5 or Top10 most enriched biological processes using the gene ontology (GO) – 2018 and ARCH54, considering a P-value < 0.01.

2.3. Tissue-specific gene networks

The HumanBase web-tool (https://hb.flatironinstitute.org) (Greene et al., 2015) was used to generate the lung tissue-specific gene network of up-regulated genes altered in lung cancer and cell lines infected with SARS-CoV-2. We considered coexpression, interaction, transcriptional factor binding, Gene Set Enrichment Analysis (GSEA) perturbations as active interaction sources, applying minimum interaction confidence of 0.1 and a maximum number of 15 genes.

2.4. Lung cancer proteins that potentially interact with coronaviruses

We performed a transcriptomic-based interactome analysis to identify DEGs coding for proteins predicted to interact with SARS-CoV and SARS-CoV-2. SARS-CoV-2 cell entry mediators were selected for our first screening by using the human proteins described in the COVID-19 Cell Atlas (https://www.covid19cellatlas.org) and the literature (HCA Lung Biological Network et al., 2020; Hoffmann et al., 2020; Li et al., 2020) (Table S2).

We next analyzed the gene expression of host proteins that potentially interact with human coronaviruses from the Pathogen–Host Interactome Prediction using Structure Similarity (P-HIPSter, http://phister.org) database (Lasso et al., 2019) (Table S3). SARS-CoV was included in the analysis considering the evolutionary relationship between the novel SARS-CoV-2 (Zhou et al., 2020b). We also complemented our list of host proteins interacting with SARS-CoV-2 by using the list of human proteins potentially interacting with SARS-CoV-2 developed by Gordon et al. 2020 (Gordon et al., 2020) (Table S4), which includes cancer proteins targeted by SARS-CoV-2 (Table S5) (Tutuncuoglu et al., 2020).
2.5. Validation of the identified differentially expressed genes in independent lung cancer datasets from TCGA

We used the web-tool XenaBrowser (https://xenabrowser.net/) (Goldman et al., 2020) to identify DEGs (P-value < 0.05) between NSCLC samples (n = 1013, The Cancer Genome Atlas - TCGA) and normal lung samples (n = 288) from Genotype-Tissue Expression (GTEx) project.

2.6. Data representation and analysis

Morpheus (https://software.broadinstitute.org/morpheus/) was used to generate the heatmaps and perform the clustering analyses. The Venn diagram was constructed using the Van de Peer Lab software (http://bioinformatics.psb.ugent.be/webofs/Venn/). Principal component analysis (PCA) was performed using NetworkAnalyst 3.0 online software (Xia et al., 2015; Zhou et al., 2019).

3. Results

3.1. Cell lines infected with SARS-CoV-2 are suitable in vitro models to study COVID-19 in lung cancer

We included three RNA-Seq datasets publicly available for re-analysis, comprising 199 NSCLC vs. 19 normal lung tissues (GSE81089), 79 SCLC vs. seven normal lung tissues (GSE60052) (Fig. 1A), A549 and Calu-3 lung cancer cell lines infected with SARS-CoV-2 vs. mock controls (GSE147507) (Fig. 1B). We first employed principal component analysis (PCA) of transcriptome data to explore similarity patterns among normal and cancer cells and tissues. As expected, this unsupervised clustering confirmed that cancer cell lines A549 and Calu-3 are more similar to cancer tissues than normal lung tissues (Fig. 1C), which was confirmed by the hierarchical clustering analysis (Fig. 1D). We statistically demonstrated the obvious; however, these results support the concept that these lung cancer cell lines should be considered to understand the biology of SARS-CoV-2 in lung cancer patients. Also, warrant caution in translating findings from these in vitro models to COVID-19 patients without cancer.

By analyzing the gene expression profile of NSCLC and SCLC compared to the respective normal lung tissues, we identified 4411 DEGs in NSCLC (2212 up- and 2199 down-regulated, respectively) and 3840 DEGs in SCLC (1901 up- and 1939 down-regulated; Table S6). The DEGs enrich biological processes related to mitotic processes and DNA replication for NSCLC and SCLC (Figs. S1 and S2). Of note, the down-regulated genes in SCLC tumors significantly enriched terms such as focal adhesion, extracellular matrix organization, and cytokine-mediated signaling pathway (Table S7). In addition, we identified 1770 DEGs in A549 cells infected with SARS-CoV-2 (871 up- and
3.2. Lung cancer tumor tissues and cell lines have variability in the expression of SARS-CoV-2 entry-associated genes

We next asked whether DEGs in lung cancer tissues or cells infected with SARS-CoV-2 are translated into proteins that potentially interact with 24 SARS-CoV-2 entry-associated proteins previously described in the literature (Table S2). We detected that a few of these genes were differentially expressed and shared by all conditions (Fig. S3, Fig. S4, and Table S9).

3.3. Transcriptomics analyses reveal candidates for host-directed therapies that potentially interfere with SARS-CoV-2 infection and replication in lung cancer

We used the DEGs shared among all conditions studied as sources of evidence to select proteins potentially interacting with SARS-CoV-2 using the P-HIPSter database (Lasso et al., 2019) and proteins identified by Gordon et al. (Gordon et al., 2020) as physically associated with SARS-CoV-2 proteins. Lung cancer tissues overexpressed the genes HOOK1, GGHI, CIT, and CENPF, coding for proteins physically associated with SARS-CoV-2 proteins, and WASF1, BRCA1, SPTBN2, FRMD5, UHRF1, and IGSF9 coding for proteins potentially interacting with SARS-CoV as predicted by P-HIPSter (Fig. 2B). In addition, seven DEGs (INHBE, STC2, SEPSECS, NPTXI, WFSI, DCTPP1, and SLC9A3R1) coding for proteins interacting with SARS-CoV-2 were found in the Calu-3 and A549 cells infected with SARS-CoV-2 (Fig. 2B). Among these genes, the overexpressed genes INHBE, STC2, and SEPSECS code for proteins physically associated with SARS-CoV-2 proteins, as predicted by Gordon et al., 2020. These results reveal candidates for host-directed therapies that may interfere with SARS-CoV-2 infection and replication in lung cancer.

We evaluated the expression profile and enrichment of biological processes for the 280 DEGs shared by all conditions tested (Fig. 2C and Table S10). This analysis defined four clusters based on gene expression patterns. Cluster 1 is composed of genes associated with cellular division. These genes are down-regulated in infected cell lines and up-regulated in tumor tissues of lung cancer patients. Cluster 2 is composed of up-regulated genes in all conditions. These genes are involved in amino acid transport and metabolic processes. Cluster 3 mainly consists of genes related to cell migration and angiogenesis, which are downregulated in all conditions. Finally, Cluster 4 contains genes associated with inflammation and leukocyte chemotaxis. These genes are up-regulated genes in infected cell lines and down-regulated in tumor tissues of lung cancer patients (Fig. 2D and Table S10).

Interestingly, we identified that TRIB3 potentially interacts with the up-regulated genes in all conditions. TRIB3 was recently found, by our research group, as a potential host-SARS-CoV-2 interacting protein associated with COVID-19 infection during lung aging (de Moraes et al., 2021). The genes MTHFD2, ADAM2, and GPT2 were predicted to be co-expressed with TRIB3 using ARCHS4 kinases prediction from the Enrichr tool (P-value = 0.0002, Table S11), while the tissue-specific gene network performed in HumanBase tool using the lung tissue as background indicated TRIB3 directly interacting with MTHFD2, GPT2, and SLC1A4 (Fig. 2E).

We also analyzed the expression profile of all DEGs found in lung cancer cells infected with SARS-CoV-2 in the NSCLC and SCLC tissues. To detect all subtle alterations in the expression of these DEGs in tumor tissues, we performed a clustering analysis without considering any statistical filter for fold change and adjusted P-value (Fig. 3A). Five clusters of genes were identified according to their expression profile (k-means clustering). The top10 enriched biological processes for each cluster are demonstrated in Fig. 3B. Cluster A represents overexpressed genes in SARS-CoV-2 infected cell lines. This cluster A has an opposite expression pattern in lung cancer tissues. This set of genes is associated with negative regulation of the viral life cycle, RNA transcription, and inflammatory response related to type I interferon. Cluster B is composed of genes with relatively low levels of positive regulation in tumor tissues. The genes of Cluster B were associated with the regulation of viral replication and defense response to viruses by the host. Notably, these genes were not up-regulated in lung cancer tissues (when applying the statistical filters) compared to normal lung tissues. Among the Cluster B genes, we found ZNF245 and APOBEC3F with relatively high expression in NSCLC or SCLC that are associated with the host response to the virus for further investigation (ZNF245, logFC = 0.68 (SCLC) and logFC = 0.05 (NSCLC; not altered), and APOBEC3F, logFC = 0.21 (NSCLC) and logFC = 0.04 (SCLC; not altered)). Cluster C consists of genes with the same expression pattern (up, but not statistically significant according to our cutoff threshold) in all conditions. These genes are associated with amino acid transport, metabolic processes, and extra-cellular matrix. Cluster D represents genes related to apoptosis and cellular processes with the same direction of change (down, but not statistically significant according to our cutoff threshold) in all conditions. Cluster E contains genes involved in mitosis and cell replication under-expressed in the cell lines infected with SARS-CoV-2 and in opposite directions in tumor tissues of patients with NSCLC and SCLC. All enriched terms for each cluster are described in Table S12.

3.4. The differential gene expression profile found in independent lung cancer datasets is similar to those from TCGA NSCLC

To confirm the differential expression of our candidates in independent lung cancer datasets, we further validated our findings by comparing the expression profiles of selected genes in The Cancer Genome Atlas (TCGA) NSCLC tissues with GTEx normal lung tissues (Table S13). The list of selected genes included SARS-CoV-2 entry-associated genes (ACE2 and TMPRSS2), genes associated with host response to the virus (APOBEC3F and ZNF245), and the genes that were up- or down-regulated in all conditions (considering statistical cutoffs thresholds). Different from what we have seen in the first cohort of lung cancer patients (GSE81089 and GSE60052), we observed an up-regulation of ACE2 in TCGA tumor samples. In addition, we noticed that the ACE2 gene has a wide range of variation in expression in NSCLC cases from TCGA, while it is expressed at lower levels in GTEx normal lung tissues (Fig. S5). However, we confirmed TMPRSS2 as significantly under-expressed in the TCGA dataset (Fig. 4). Except for the genes ITGA2, LAMB3, DHCR24, and ZNF254, most selected genes were confirmed to be dysregulated in TCGA lung cancer samples (Fig. 4).

4. Discussion

Previous studies have demonstrated a higher risk and vulnerability of lung cancer patients to SARS-CoV-2 infection (Dai et al., 2020; Luo et al., 2020; Wu et al., 2021). Therefore, we analyzed publicly available lung transcriptomic data to compare transcriptional alterations caused by SARS-CoV-2 in lung cancer cell lines with tumor transcriptional profiles of NSCLC and SCLC lung cancer patients. While lung cancer cell lines, A549 and Calu-3, have been broadly used to understand SARS-CoV-2 infection and replication in non-cancer COVID-19 patients (Blanco-Melo et al., 2020; Wyler et al., 2021), we demonstrate that these cell lines, when infected with SARS-CoV-2, are valid models to study COVID-
Fig. 2. The expression landscape of the SARS-CoV-2 interaction proteins in lung cancer tissue and cells. A) Heat-scatter plot of genes coding for SARS-CoV-2 interacting proteins in NSCLC, SCLC, Calu-3 (SARS-CoV-2 infected), and A549 (SARS-CoV-2 infected) compared to their respective controls. The color and size of the circles correspond to the log2FC and -log10 transformed adjusted P-value, respectively. B) Heatmap illustrating differentially expressed genes (DEGs) that code for proteins that interact with SARS-CoV as predicted by the P-HIPSTer database and Gordon et al., 2020. Thirty DEGs are overlapped between NSCLC and SCLC (green Venn diagram) and seven between Calu-3 and A549 infected with SARS-CoV-2 (blue Venn diagram). The K-means analysis generated clusters associated with increased and decreased expression. C) Heatmap illustrating differentially expressed genes among all conditions evaluated. K-means analysis yielded four Clusters related to increased and decreased expression in all conditions tested. D) Top5 biological processes enriched by the genes from each cluster identified in Fig. 2C (considering P-value < 0.05). Cluster1: genes down-regulated in lung cancer cells infected with SARS-CoV-2 and up-regulated in lung cancers; Cluster2: genes up-regulated in all conditions; Cluster3: genes down-regulated in all conditions; Cluster4: genes up-regulated in infected lung cancer cells and down-regulated in non-infected lung cancers tissues. E) Tissue-specific gene network of up-regulated genes identified in all conditions tested (lung cancer and lung cancer cell lines infected with SARS-CoV-2). The network was generated using lung tissue as background in the HumanBase online tool. FC: fold change; p.adj: adjusted P-value; NSCLC: non-small cell lung cancer; SCLC: small-cell lung cancer; DEGs: differentially expressed genes. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
In lung cancer. Importantly, we found that NSCLC and SCLC over-express the cancer-associated genes *BRCA1* and *CENPF*, which are translated into proteins that are known (Gordon et al., 2020) or predicted (Lasso et al., 2019) to interact with the SARS-CoV-2 spike glycoprotein and helicase (nps13), respectively.

We selected A549 and Calu-3 cell lines to evaluate SARS-CoV-2 transcriptional alterations in cell lines with high and low transfection rates (Calu-3 and A549, respectively) (Blanco-Melo et al., 2020). In contrast to Calu-3 cells, A549 cells are relatively non-permissive to SARS-CoV-2 replication, possibly due to low expression levels of ACE2 (Blanco-Melo et al., 2020; Harcourt et al., 2020; Letko et al., 2020). However, we found significant changes in gene expression profiles between infected and non-infected A549 cells (Fig. S3). This data suggest that SARS-CoV-2 may induce intracellular signaling and modify the transcription of genes regardless of a low viral load or reduced expression of the known virus entry-associated receptor ACE2. Therefore,
although few cells in lung adenocarcinomas express high levels of ACE2 (Han et al., 2021), they might induce relevant molecular changes in lung cancer patients with COVID-19 (Wu et al., 2021).

We also identified that SARS-CoV-2 infection does not induce alterations in ACE2 in the Calu-3 and A549 cell lines. Evidence suggests that ACE2 upregulation is induced by type I interferons during host immune responses (Ziegler et al., 2020), which may explain why no changes were observed in the expression of ACE2 identified in vitro after SARS-CoV-2 infection. The expression levels of ACE2 and TMPRSS2 may indicate a higher susceptibility to SARS-CoV-2 infection in patients with LUAD and LUSC (Kong et al., 2020; Wang et al., 2021). In our results, the expression of ACE2 was not significantly different in NSCLC and SCLC compared to non-cancer lungs. One possible explanation is that few LUAD malignant cells highly express ACE2, as shown recently by scRNA sequencing (Han et al., 2021). Additionally, the expression of ACE2 is highly variable, and the expression levels of ACE2 for a group of TCGA NSCLC patients are similar to levels found in GTEx normal lungs (Fig. S5). This finding suggests a high variability of ACE2 expression in lung cancer tissues. Furthermore, TMPRSS2 - classical SARS-CoV-2 entry-associated protease - was down-regulated in lung carcinomas, including the TCGA NSCLC, and was found as down-regulated in LUAD at a scRNA-seq resolution (Han et al., 2021).

Even presenting a discrete change in the expression of known entry-associated receptors and proteases, lung carcinomas can also express other host-virus interacting genes. Therefore, we analyzed the gene expression profile of genes that encode proteins potentially interacting with SARS-CoV-2 in lung carcinomas. We found ten up-regulated genes in NSCLC and SCLC that were predicted to interact with coronaviruses. Among them, we identified the tumor suppressor gene BRCA1. Computational modeling revealed that SARS-CoV-2 potentially interacts with BRCA1 via spike protein subunit 2 (Singh and Bharara Singh, 2020). In addition, BRCA1 was suggested to be involved in other viral infections, such as HIV-1, indicating that BRCA1 enhances HIV-1 transcription by acting as a transcriptional cofactor (Guendel et al., 2015). We also identified the cancer gene CENPF, a host centrosome component, predicted to interact with the SARS-CoV helicase (nps13) protein (Gordon et al., 2020; Muus et al., 2021). The nps13 protein belongs to the superfamily 1 of the six helicase superfamilies responsible for opening up ribo- and deoxyribonucleic acid duplexes 5′ to 3′ direction. Therefore, nps13 is a potential drug target for COVID-19 therapy (Kousar et al., 2020).

We identified 22 genes with the same expression pattern in all conditions evaluated (NSCLC, SCLC, A549 infected, and Calu-3 infected). It was recently demonstrated that lung samples from COVID-19 and lung cancer patients have DEGs shared by both diseases, and most overlapping genes are found in opposite directions (Kuchi et al., 2022). However, these authors confirmed a similar molecular perturbation between COVID-19 and lung cancer, corroborating our findings. By carefully analyzing this list of 22 genes, we observed that the nine shared up-regulated genes (Cluster 2) were associated with amino acid metabolism processes. Evidence suggests that patients with COVID-19 have dysregulation of amino acid metabolism and differential serum levels of amino acids (Ansone et al., 2021; Thomas et al., 2020). The list also has 13 down-regulated genes (Cluster 3) associated with angiogenesis. As a result of vascular inflammation, endothelial dysfunction is also a known characteristic of COVID-19 patients (Huertas et al., 2020). These patients present pulmonary endothelial cell dysfunction (Won et al., 2022) and a decreased vascular density in the sublingual circulation (Rovas et al., 2021; Won et al., 2022). However, new vessel growth in the affected lungs is predominant in COVID-19 patients (Ackermann et al., 2020). Thus, future studies should consider evaluating the biological processes described herein in lung cancer patients with COVID-19.

We also found TRIB3 potentially interacting with the nine up-regulated genes identified in all tested conditions. Interestingly, our research group recently found TRIB3 - predicted to interact with viral nucleocapsid protein and RNA-dependent RNA polymerase - decreased in males during aging, asthma, and COPD, possibly contributing to the severity of COVID-19 in these patients (de Moraes et al., 2021). These results demonstrate the relevance of TRIB3 in lung tissues and its association with COVID-19 in other comorbidities, as lung cancer.

Despite integrating robust bulk RNA-Seq data from high-quality datasets, this study has several limitations. This exploratory research...
was based on publicly available transcriptome data, only using computational predictions, which may not correspond to what is found in lung cancer patients with COVID-19. Thus, the results discussed here need to be experimentally validated. The analysis of the interaction between human and viral proteins did not consider different variants of SARS-CoV-2, such as omicron (Islam et al., 2022c) or SARS-CoV-2 co-infection with other viruses (Islam et al., 2022a, 2022b), which may reflect different transcriptomic profiles in lung cancer cells. Moreover, cell lines fail to recapitulate all aspects of the bulk tumor tissues; therefore, data regarding the tumor microenvironment’s implication should be further investigated. Nonetheless, considering the urgent need to better understand COVID-19, especially in high-risk patients, such as lung cancer patients, this study opens new exploratory opportunities to address the genetic background associated with SARS-CoV-2 infection.

5. Conclusions

In conclusion: (1) Transcriptional profiles indicate that A549 and Calu-3 cell lines are valid models to investigate the SARS-CoV-2 infection in lung cancer rather than in normal lung tissues; (2) Our analyses demonstrate that lung carcinomas express high levels of cancer genes BRCAl and CENPF that encode proteins potentially interacting with SARS-CoV-2 spike protein subunit 2 and nsp13, respectively; and (3) we also pointed out 22 genes with the same expression pattern in lung cancer and COVID-19, which are associated with amino acid metabolism and angiogenesis.

CRediT authorship contribution statement

S.S. Cury: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing, Visualization. J.S. Oliveira: Methodology, Formal analysis, Investigation, Writing – original draft, Writing – review & editing, Visualization. C.A.O. Biagi-Júnior: Methodology, Formal analysis, Investigation, Writing – review & editing, Visualization. W.A. Silva: Investigation, Data curation, Writing – review & editing, Visualization. P.P. Reis: Investigation, Data curation, Writing – review & editing, Visualization. O. Cabral-Marques: Investigation, Data curation, Writing – review & editing, Visualization. E.N. Hasimoto: Investigation, Data curation, Writing – review & editing, Visualization. P.P. Freire: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing – review & editing, Visualization, Supervision. R. F. Carvalho: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing – review & editing, Visualization, Supervision, Project administration, Funding acquisition, Resources, Writing – original draft.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

This paper analyzes existing publicly available datasets.

Acknowledgement

We thank Nadia El Khawanky (Department of Hematology and Oncology, Faculty of Medicine) from the University of Freiburg, Freiburg, Germany, for the English corrections. The results shown here are in part based upon data generated by the TCGA Research Network (http://cancergenome.nih.gov/) and by the Genotype-Tissue Expression project (GTEx) (https://gtexportal.org/).

Funding

This study was supported by Santander Universities (Banco Santander Brasil S/A) for Covid-19 research, by the Sao Paulo Research Foundation – FAPESP (grant 13/50343-1 to RFC and grant 20/09146-1 to PPF) and by the National Council for Scientific and Technological Development, CNPq (Process 311530/2019-2 to RFC, scholarship #141601/2019-1 to SSC), by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, CAPES, Brasil, Finance Code 001 (scholarship to JSO and scholarship #88882.378695/2019-01 to CAOBB).

Data Availability Statement

Publicly archived datasets analyzed herein are available at Gene Expression Omnibus ( GEO) under the accession numbers: GSE81089, GSE60052, and GSE147507.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi. org/10.1016/j.gene.2022.147047.

References

Ackermann, M., Verleden, S.E., Kuehnel, M., Haverich, A., Welte, T., Laenger, F., Vanstapel, A., Werlein, C., Stark, H., Tsankov, A., Li, W.W., Li, W.V., Mentzer, S.J., Jonik, D., 2020. Pulmonary vascular endothelialitis, thrombosis, and angiogenesis in Covid-19. N. Engl. J. Med. 383, 120–128. https://doi.org/10.1056/NEJMoa15432.

Albiges, L., Foulon, S., Bayle, A., Gachot, B., Pommeter, F., Willekens, C., Stoclin, A., Merafd, M., Griscelli, F., Lacroix, L., Netzer, F., Hueso, T., Bälleygueui, C., Ammari, S., Colomba, E., Baziarello, G., Perret, A., Hollebecque, A., Hadoux, J., Michot, J.-M., Chaput, N., Saada, V., Hauchecorne, M., Micol, J.-B., Sun, R., Valteau-Couzanet, D., André, F., Scotte, F., Besse, B., Soria, J.-C., Barlesi, F., 2020. Determinants of the outcomes of patients with cancer infected with SARS-CoV-2: results from the Gustave-Roussy cohort. Nat. Cancer 1, 965–973. https://doi.org/10.1038/s43018-020-00120-5.

Ansonne, L., Briviba, M., Silamikielis, I., Terentjeva, A., Perkons, I., Birzniece, L., Rovite, V., Rozentale, B., Viksna, L., Klavins, K., Klovins, J., 2021. Amino acid metabolism is significantly altered at the time of admission in hospital for severe COVID-19 patients: findings from longitudinal targeted metabolomics analysis. Microbiol. Spectr. 9, e00338-e21. https://doi.org/10.1128/ micspec.00338-21.

Blanco-Melo, D., Nilsson-Payant, B.E., Liu, W.-C., Uhl, S., Hoagland, D., Maller, R., Jordan, T.X., Oishi, K., Parisi, M., Sachs, D., Wang, T.T., Schwartz, R.E., Lim, J.K., Albrecht, R.A., tenOever, B.R., 2020. Imbalanced host response to SARS-CoV-2 drives development of COVID-19. Cell 181, 1036–1045.e9. https://doi.org/10.1016/j.cell.2020.04.026.

Bui, L.T., Winters, N.I., Chng, M.-I., Joseph, C., Gutierrez, A.J., Habermann, A.C., Adams, T.S., Schupp, J.C., Poli, S., Peter, L.M., Taylor, C.J., Blackburn, J.B., Richmond, R.W., Nicholson, A.G., Roat, D., Wallace, W.A., Ronan, I.O., Jenkins, R.G., Kaminska, N., Kropsk, J.A., Banovich, N.E., Misharin, A.V., Tsankov, A.M., Spira, A., Barbry, P., Brazuma, A., Samakolovis, C., Shepherd, D.P., Rawlins, E.L., Theis, F.J., Griffonnet, J., Lee, H., Schiller, H.B., Hofman, P., Powell, J.E., Schultz, J.I., White, J.L., Whitsett, J., Choi, J., Lundeberg, J., Kaminski, N., Kropsk, J.A., Banovich, N.E., Orozlos-Montes, J., Rajagopal, J., Meyer, K.B., Krauonow, M.A., Szehr-Farry, K., Zhang, K., Lafayatis, R., Leroy, S., Hanifa, M., Nawn, M.C., Nikolici, M.Z., van den Berge, M., Kuhnmund, M., Marquette, C.-H., Von Papen, M., Eickelberg, O., Rosenblatt-Rosen, O., Reyfman, P.A., Pe'er, D., Horvath, P., Tata, P.R., Regev, A., Rojas, M., Seibold, M.A., Shalek, A.K., Spence, J.R., Teichmann, S.A., Quake, S., Duong, T.E., Biancalani, T., Desai, T., Sun, X., Zaragosi, L.E., 2021. Chronic lung diseases are associated with gene expression programs favoring SARS-CoV-2 entry and severity. Nat. Commun. 12 (1) https://doi.org/10.1038/s41467-021-24467-0.

Chen, E.Y., Tan, C.M., Kou, Y., Duan, Q., Wang, Z., Meirelles, G.V., Clark, N.R., Ma, A., 2013. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics 14, 128. doi: 10.1186/1471-2105-14-128.

Chen, F., Zhang, Y., Surgang, R., Ramani, S., Corry, D., Kheradmand, F., Creighton, C.J., 2021. Meta-analysis of host transcriptional responses to SARS-CoV-2 infection reveals their manifestation in human tumors. Sci. Rep. 11, 791. doi: 10.1038/s41598-021-22221-4.

Dai, M., Liu, D., Liu, M., Zhou, F., Li, G., Chen, Z., Zhang, Z., You, H., Wu, M., Zheng, Q., Xiong, Y., Xiong, H., Wang, C., Chen, C., Xiong, F., Zhang, Y., Peng, Y., Ge, S., Zhou, B., Yu, T., Wang, L., Wang, H., Liu, Y., Chen, Y., Mei, J., Gao, X., Li, Zhuyan, Gan, L., He, C., Li, Zhen, Shi, Y., Qi, Y., Yang, J., Tenen, D.G., Chai, L., Mucci, L.A., Ackermann, M., Verleden, S.E., Kuehnel, M., Haverich, A., Welte, T., Laenger, F., 2020. Pulmonary vascular endothelialitis, thrombosis, and angiogenesis in Covid-19. N. Engl. J. Med. 383, 120–128. https://doi.org/10.1056/NEJMoa15432.

Dow, T.E., Biancalani, T., Desai, T., Sun, X., Zaragosi, L.E., 2021. Chronic lung diseases are associated with gene expression programs favoring SARS-CoV-2 entry and severity. Nat. Commun. 12 (1) https://doi.org/10.1038/s41467-021-24467-0.
Islam, F., Dhawan, M., Nafady, M.H., Emran, T.B., Mitra, S., Choudhary, O.P., Akter, A., 2022e. Understanding the omicron variant (B.1.1.529) of SARS-CoV-2: mutational impacts, concerns, and lessons. Aims. J. Med. Genet. 59, 101663. https://doi.org/10.1016/j.ajmg.2022.101663.
Kaminski, N., Krasnow, M., Kropski, J.A., Kuhnemund, M., Lafyatis, R., Lee, H., Leroy, S., Linnarson, S., Lundeberg, J., Meyer, K., Misharin, A., Nawijn, M., Nikolic, M.Z., Ordovas-Montanes, J., Pe'er, D., Powell, J., Quake, S., Rajagopal, J., Tata, P.B., Rawlins, E.L., Regev, A., Reyfman, P.A., Rojas, M., Rosen, O., Saeb-Parsy, K., Samakovlis, C., Schiller, H., Schultze, J.L., Seibold, M.A., Shalek, A.K., Shepherd, D., Spence, J., Spira, A., Sun, X., Teichmann, S., Theis, F., Tsankov, A., van den Berge, M., von Papen, M., Whitsett, J., Xavier, R., Xu, Y., Zarrago, L.-E., Zhang, K., 2020. SARS-CoV-2 Receptor ACE2 Is an Interferon-Stimulated Gene in Human Airway Epithelial Cells and Is Detected in Specific Cell Subsets across Tissues. Cell 181, 1016–1035.e19. doi: 10.1016/j.cell.2020.04.035.