



# Genome-Wide Profiling Reveals Alternative Polyadenylation of Innate Immune-Related mRNA in Patients With COVID-19

Sanqi An<sup>1†</sup>, Yueqi Li<sup>1†</sup>, Yao Lin<sup>1†</sup>, Jiemei Chu<sup>1</sup>, Jinming Su<sup>1</sup>, Qiuli Chen<sup>1</sup>, Hailong Wang<sup>1</sup>, Peijiang Pan<sup>1</sup>, Ruili Zheng<sup>1</sup>, Jingyi Li<sup>1</sup>, Junjun Jiang<sup>1,2\*</sup>, Li Ye<sup>1,2\*</sup> and Hao Liang<sup>1,2\*</sup>

## OPEN ACCESS

### Edited by:

William J. Liu,  
National Institute for Viral Disease  
Control and Prevention (China CDC),  
China

### Reviewed by:

Guoli Shi,  
National Institutes of Health (NIH),  
United States  
Ulf Kahlert,  
Heinrich Heine University of  
Düsseldorf, Germany

### \*Correspondence:

Junjun Jiang  
jiangjunjun@gxmu.edu.cn  
Li Ye  
yeli@gxmu.edu.cn  
Hao Liang  
lianghao@gxmu.edu.cn

<sup>†</sup>These authors have contributed  
equally to this work and share  
first authorship

### Specialty section:

This article was submitted to  
Vaccines and Molecular Therapeutics,  
a section of the journal  
Frontiers in Immunology

**Received:** 10 August 2021

**Accepted:** 13 October 2021

**Published:** 27 October 2021

### Citation:

An S, Li Y, Lin Y, Chu J, Su J, Chen Q,  
Wang H, Pan P, Zheng R, Li J,  
Jiang J, Ye L and Liang H (2021)  
Genome-Wide Profiling Reveals  
Alternative Polyadenylation of Innate  
Immune-Related mRNA in  
Patients With COVID-19.  
*Front. Immunol.* 12:756288.  
doi: 10.3389/fimmu.2021.756288

<sup>1</sup> Biosafety Level-3 Laboratory, Life Sciences Institute & Guangxi Collaborative Innovation Center for Biomedicine, Guangxi Medical University, Nanning, China, <sup>2</sup> Guangxi Key Laboratory of AIDS Prevention and Treatment, School of Public Health, Guangxi Medical University, Nanning, China

The coronavirus disease 2019 (COVID-19) pandemic has caused many deaths worldwide. To date, the mechanism of viral immune escape remains unclear, which is a great obstacle to developing effective clinical treatment. RNA processing mechanisms, including alternative polyadenylation (APA) and alternative splicing (AS), are crucial in the regulation of most human genes in many types of infectious diseases. Because the role of APA and AS in response to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection remains unknown, we performed *de novo* identification of dynamic APA sites using a public dataset of human peripheral blood mononuclear cell (PBMC) RNA-Seq data in COVID-19 patients. We found that genes with APA were enriched in innate immunity-related gene ontology categories such as neutrophil activation, regulation of the MAPK cascade and cytokine production, response to interferon-gamma and the innate immune response. We also reported genome-wide AS events and enriched viral transcription-related categories upon SARS-CoV-2 infection. Interestingly, we found that APA events may give better predictions than AS in COVID-19 patients, suggesting that APA could act as a potential therapeutic target and novel biomarker in those patients. Our study is the first to annotate genes with APA and AS in COVID-19 patients and highlights the roles of APA variation in SARS-CoV-2 infection.

**Keywords:** COVID-19, alternative polyadenylation, immunity, alternative splicing, APA regulator

## INTRODUCTION

SARS-CoV-2 causes the respiratory disease known as COVID-19. By the start of August 2021, there had been more than 200 million cases and at least 4,200,000 deaths caused by COVID-19 around the world.

mRNA-processing events, including APA and AS, play key roles in various diseases (1, 2). APA is a widespread mechanism of gene regulation that generates different 3' ends of transcripts (3). APA leads to the production of distinct protein isoforms and repressing gene expression (4, 5). APA regulators control APA by binding to the APA site during mRNA processing (6). AS enables an mRNA to

differentiate into isoforms that may have different biological functions (3). Therefore, APA and AS are involved in transcriptional gene regulation. Overall, the diversity of the transcriptome and proteome is enhanced by APA and AS, which are two important regulatory mechanisms for many biological processes. Indeed, through APA, a single gene can encode multiple different 3' ends of transcripts. Polyadenylation affects numerous aspects of mRNA metabolism, including transcription termination by RNAP II, mRNA stability, and the efficiency of translation (7). APA upregulates target genes through miRNA repression escape (6). Shortening of the average 3'UTR length and widespread APA in response to virus infection have been reported (8). The expression level of genes with APA is altered and enriched in immune-related categories including interferon (8). Interferons have shown *in vitro* and *in vivo* antiviral effect against SARS-CoV-2, and they have been suggested as a potential treatment for COVID-19 patients (9, 10). Moreover, APA has an effect on viral replication and plays a crucial role in the antiviral innate immune response (8). The mechanisms of changes in APA are considered to be regulated in *cis* through genetic aberrations (11) and in *trans* by regulatory proteins in response to dynamic environmental changes (12). To date, there has been no research on whether APA plays an essential role in patients infected with SARS-CoV-2.

Similar to alternative polyadenylation, hundreds of host genes showed AS upon viral infection (13). For instance, infection of human cells with influenza A virus induced a broad program of alternative splicing of host genes (14). However, the in-depth molecular basis of infection and pathogenesis of SARS-CoV-2 in human cells has not been further explained. It has also been reported that SARS-CoV-2 disrupts mRNA splicing *in vitro* (1). There is an urgent need to determine whether these AS anomalies correlate with clinical features.

Based on the above information, we aimed to reveal overall dynamic changes in APA and AS in COVID-19 patients. DaPars (6) and rMATS (15) were used to directly detect APA and AS events from RNA-Seq data. Our findings will contribute to our understanding of the pathogenesis, potential molecular targets, and development of new APA-based therapeutic targets of COVID-19.

## MATERIALS AND METHODS

### Data Extraction

The sequencing data used in this study were retrieved from two large-scale multiomic studies of COVID-19 (16, 17), in which 128 patients admitted to Albany Medical Center in Albany, NY were collected for moderate to severe respiratory issues presumably related to SARS-CoV-2 infection from April 6, 2020 to May 1, 2020. Patients who were positive ( $n = 102$ ) and negative ( $n = 26$ ) for the virus were assigned to the COVID-19 and non-COVID-19 groups, respectively. Due to missing sequencing data for 2 positive patients, a total of 126 cases were included in this study. RNA-Seq data of the COVID-19 cohort mentioned above were available at the SRA database

SRP279280 (<https://www.ncbi.nlm.nih.gov/sra/>) (16). To validate our results, independent RNA-Seq data for PBMCs from the non-COVID-19 and COVID-19 groups were downloaded from the GSA database CRA002390 (<https://bigd.big.ac.cn/gsa/>) (17). The reads were mapped to the hg38 human genome using HISTA2 (v2.1.0) (18). We used StringTie (v1.3.4d) to calculate the TPMs (transcripts per million) of Ensembl annotated genes (19). Differential gene expression analyses were performed using DESeq2 according to the read counts of each gene determined by HTSeq (18). Genes with FDR (false discovery rate)  $\leq 0.05$  and mean CPM (Counts per Million)  $> 100$  were determined to be differentially expressed genes, as we described previously (20) (Table S1).

### Alternative Polyadenylation Analysis

We used APA to link genetic variation to variations in gene expression and disease risk. DaPars is a bioinformatic algorithm dedicated to *de novo* identification and quantification of dynamic APA events using standard RNA-Seq (6). In this study, we used DaPars (0.9.1) to infer, identify and quantify APA in RNA-Seq data for COVID-19 patients and non-COVID-19 cohorts. The APA with  $FDR \leq 0.05$  and  $\Delta|PDUI| > 0.1$  were determined as the significantly different APA between COVID-19 and non-COVID-19 as described previously (21). The percentage of distal polyA site usage index (PDUI) was defined as follows:

$$PDUI = \frac{L}{L + S}$$

S and L are the abundances of transcripts with proximal distal polyA sites in each sample. The following is a detailed description of the options used with DaPars.  $FDR\_cutoff=0.05$ ,  $PDUI\_cutoff=0.1$  (Table S2).

### Differential Alternative Splicing Analysis

The main types of AS include exon skipping (SE), intron retention (IR), alternative 5' splice site (A5SS), alternative 3' splice site (A3SS) and mutually exclusive exon (MXE). In total, 126 samples, including 100 COVID-19 and 26 non-COVID-19 patient samples, were used for analysis. We used hisat2 (v2.1.0) for sequence alignment and rMATS for AS analysis (15). Using  $FDR \leq 0.05$  as the threshold, the AS analysis results were screened to obtain a matrix of differential AS genes and expression levels. We ran rMATS on bam as input data. The following is a description of the options used with rMATS:  $FDR \leq 0.05$ , Read length = 51, and Thread=32 (Table S3). The data were visualized using the Integrative Genomics Viewer (IGV) tool (22).

### GO Function Enrichment and KEGG Pathway Analyses

To describe potential different mechanisms between those with and without COVID-19, differential APA genes, AS genes, and differentially expressed genes screened in the previous step were used for GO function enrichment analysis and KEGG pathway enrichment analysis using The Database for Annotation, Visualization and Integrated Discovery (DAVID, <https://david.ncicrf.gov/>) and Metascape

(<https://metascape.org/gp/index.html#/main/step1>) (23, 24). The Benjamini-Hochberg (BH) method was used to obtain the FDR value, and  $FDR \leq 0.05$  was taken as the threshold. Pathway/GO terms satisfying this condition were defined as those significantly enriched among differentially expressed genes. The top hits with the most significant enrichment (the lowest p-value) are shown in a histogram.

## Principal Component Analysis

Principal component analysis (PCA) is used to transform original correlating variables into two uncorrelated principal components by linear transformation and to characterize the contribution rate by variance (also called the eigenvalue); we selected the first two principal components for analysis according to the contribution rate. After standardizing the input alternative splicing PSI (percent spliced in index) and PDUI values, the R packages factominer (V2.4) and factoextra (v1.0.7) were applied to reduce the PCA dimension of AS and APA genes, and the first and second principal components were extracted for two-dimensional visualization.

## Analyses of the Relationship of APA With DEG

We performed cumulative distribution analysis of log<sub>2</sub> fold changes in mRNA expression in R as we described previously (20). The empirical cumulative distribution functions of the log<sub>2</sub>-fold mRNA expression values were computed using the ecdf function, and the corresponding p-value was included in the plot. miRNA-binding sites were predicted by TargetScanHuman (25), and Hypergeometric Optimization of Motif Enrichment (HOMER) software was used for motif enrichment analysis (26).

## RESULTS

### Dynamic APA Events in COVID-19 Patients

First, we developed a computational framework to systematically analyze the dynamic changes in APA, gene expression and AS in two independent databases. The computational framework is illustrated in **Figure 1**. The detailed method is given in the *Materials and Methods* section.

We performed *de novo* identification and quantification of dynamic APA events using existing RNA-Seq data for PBMCs from COVID-19 patients. Based on data for 100 COVID-19 patients and 26 non-COVID-19 subjects collected in GEO, we detected 145 sites that were significant dynamic APA events ( $FDR \leq 0.05$  and  $\Delta|PDUI| > 0.1$ ) in COVID-19 (**Figures 2A, B**) (**Table S2**). We quantified the degree of difference of APA in COVID-19 as a change in the percentage of distal polyA site usage index ( $\Delta PDUI$ ), which could identify shortening (negative index) or lengthening (positive index) of 3'UTRs. As shown in **Figures 2A, B**, in COVID-19 samples, most dynamic APA events presented a shorter 3'UTR. The mean PDUI value was significantly lower in the COVID-19 group than in the non-COVID-19 group (**Figure 2C**). Additionally, COVID-19 patients with shorter 3'UTRs had a decreasing number

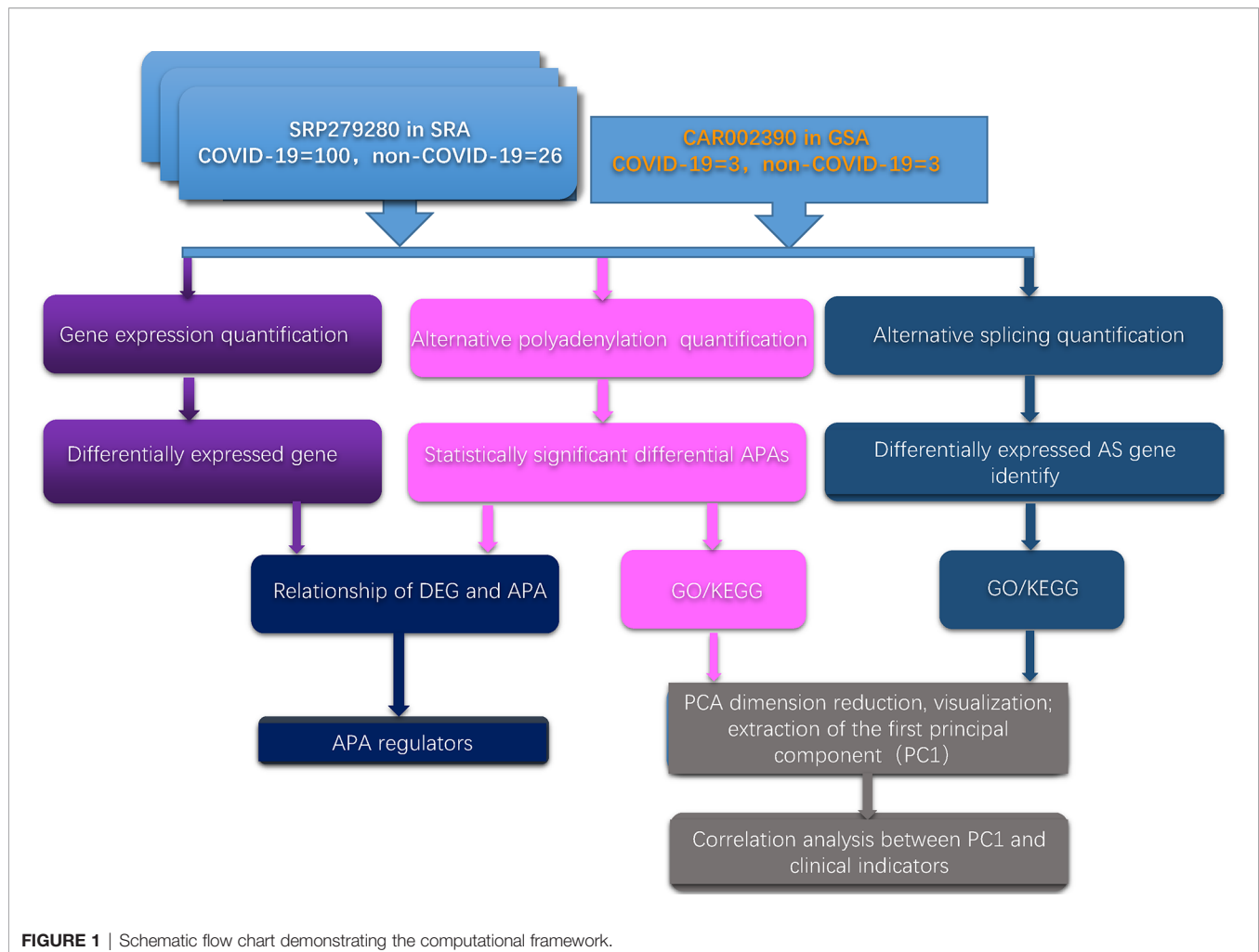
of ventilator-free days compared with those with longer 3'UTRs. COVID-19 patients with shorter 3'UTRs also had higher intensive care unit (ICU) admission rates. We concluded that APA is related to the clinical characteristics of COVID-19. Furthermore, the canonical polyA signal AATAAA was successfully identified by HOMER motif enrichment analysis of dynamic APA sites (**Figure S1A**) (27), AATAAA was found to be the most predominant motifs in human APA events in previous study (8). One example of a changed APA event in COVID-19 compared with non-COVID-19 is given for the IGSF6 gene (Immunoglobulin Superfamily Member 6) (**Figure 2D**), with a shorter 3'UTR predominating in COVID-19 samples compared with matched non-COVID-19 samples in GEO (16). Furthermore, we validated this result with another independent data in GSA (17) (**Figure S1B**). Collectively, these analyses revealed global changes in APA in COVID-19 patients.

### Functional Analysis of APA in COVID-19

According to previous studies, the type I interferon (IFN) response characterizes the local immune response phase due to SARS-CoV-2 attack (28). The IFN system caused by TNF- $\alpha$  and IFN- $\gamma$  mirrors the tissue damage and inflammation that occur in COVID-19 samples (29), and most GO and pathway enrichment analyses based on gene expression data have revealed that interferon terms are enriched in COVID-19 patients (17). We performed GO and pathway enrichment analyses to explore changes based on APA in response to SARS-CoV-2 infection and found that APA is associated with key biological processes and pathways of COVID-19 (**Figure S2**). **Figure 3A** shows significant enrichment of significantly different APA in biological processes associated with innate immune responses, such as neutrophil activation, regulation of the MAPK cascade and cytokine production, response to IFN- $\gamma$  and the innate immune response. In fact, viruses infecting vertebrate hosts must overcome the IFN-mediated antiviral response before replicating and propagating to new hosts (30). As shown in **Figure 3B**, the GSEA results showed that genes with APA were significantly enriched in response to IFN- $\gamma$ . Interestingly, this result is similar to previous GO results based on differentially expressed genes in COVID-19 (17). To further investigate the specific innate immune regulatory APA in COVID-19, we selected representative samples to track the IFN-related genes CD14, IL6 and IFNGR1 (**Figure S3**). In addition, we performed KEGG pathway enrichment analysis using genes with significantly different APA (**Figure 3C**). This analysis indicated that genes expressing APA were significantly enriched in pathogenic *Escherichia coli* infection, endocytosis, phagosome human cytomegalovirus infection, Epstein-Barr virus infection and human T-cell leukemia virus 1 infection. We concluded that APA was involved in the IFN signaling pathway in the antiviral innate immune response of COVID-19 patients.

### APA and DEGs in COVID-19

It is widely accepted that 3'UTR shortening through APA may upregulate target genes *via* miRNA repression escape (6). Because the results of immune-related pathway enrichment for APA were similar to the results of pathway enrichment for the differentially expressed genes, we speculated that APA affected the expression level of immune-related genes in COVID-19



patients. To address this possibility, we examined differentially (DEGs) and nondifferentially expressed genes to assess whether the APA sites were overrepresented in COVID-19 associated DEGs. As expected, almost 6% of significantly DEGs underwent APA, whereas only 1% of nondifferentially expressed genes underwent APA, indicating a significantly enriched occurrence of APA target sites among DEGs ( $p=2.8 \times 10^{-33}$ , two-tailed chi-square test; **Figure 4A**), suggesting that APA affects host gene expression level. Moreover, the fold change ratios of the genes with APA were significantly upregulated compared to genes without APA target sites in COVID-19 ( $p=1.2 \times 10^{-36}$ , two-tailed Wilcoxon test; **Figure 4B**). We also observed that 77% of genes with shorter 3'UTRs in COVID-19 samples lost at least 1 predicted miRNA-binding site, indicating that APA might upregulate parental genes by escaping miRNA repression (**Figure S4A**).

It has been reported that the STAT1/STAT3 axis is required for TNF- $\alpha$ - and IFN- $\gamma$ -induced inflammatory cell death-PANoptosis in COVID-19 (29). Thus, we manually compared APA in multiple IFN-related genes, especially STAT1 in two independent datasets to verify our results (**Figures 4C, E**). We found a shorter 3'UTR for STAT1 and significantly higher gene expression levels in two

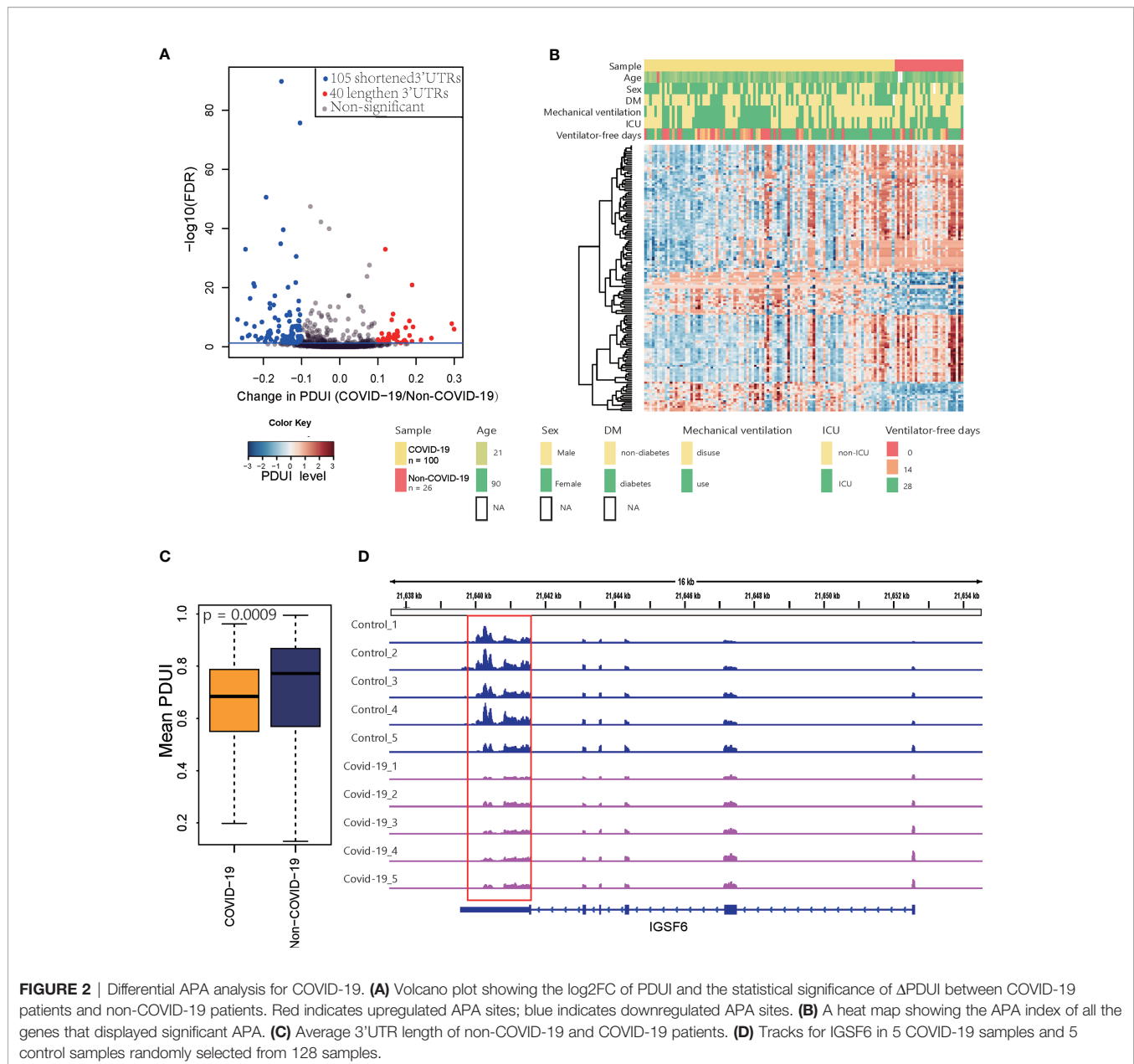
independent datasets (**Figures 4D, F**). Since there were too many samples, we randomly selected five representative samples to show the trends. The findings are consistent with the theoretical hypothesis that 3'UTR shortening through APA may upregulate target genes through miRNA repression escape. Collectively, these analyses revealed that APA play an important role in the regulation of gene expression in COVID-19 patients.

### Regulatory Factors of APA

The core polyadenylation trans factors include the activity of 3' processing factors. The 3' processing complex consists of over 20 core proteins (27), including cleavage and polyadenylation specificity factor (CPSF), cleavage factors (CFim and CFIm), Two cytoplasmic poly(A)-binding proteins (PABPC1 and PABPC4), cleavage stimulatory factor or cleavage stimulation factor (CSTF), poly(A)-binding protein nuclear 1 (PABPN1), FIP1L1, PCF11 and SYMP. Previous studies have found that alteration in the expression of some 3' processing factors caused significant changes in poly (A) site choice (31).

To investigate whether APA changes with the expression level of 3' processing factors in COVID-19 patients, we analyzed changes in expression of 3' processing factors in PBMCs from



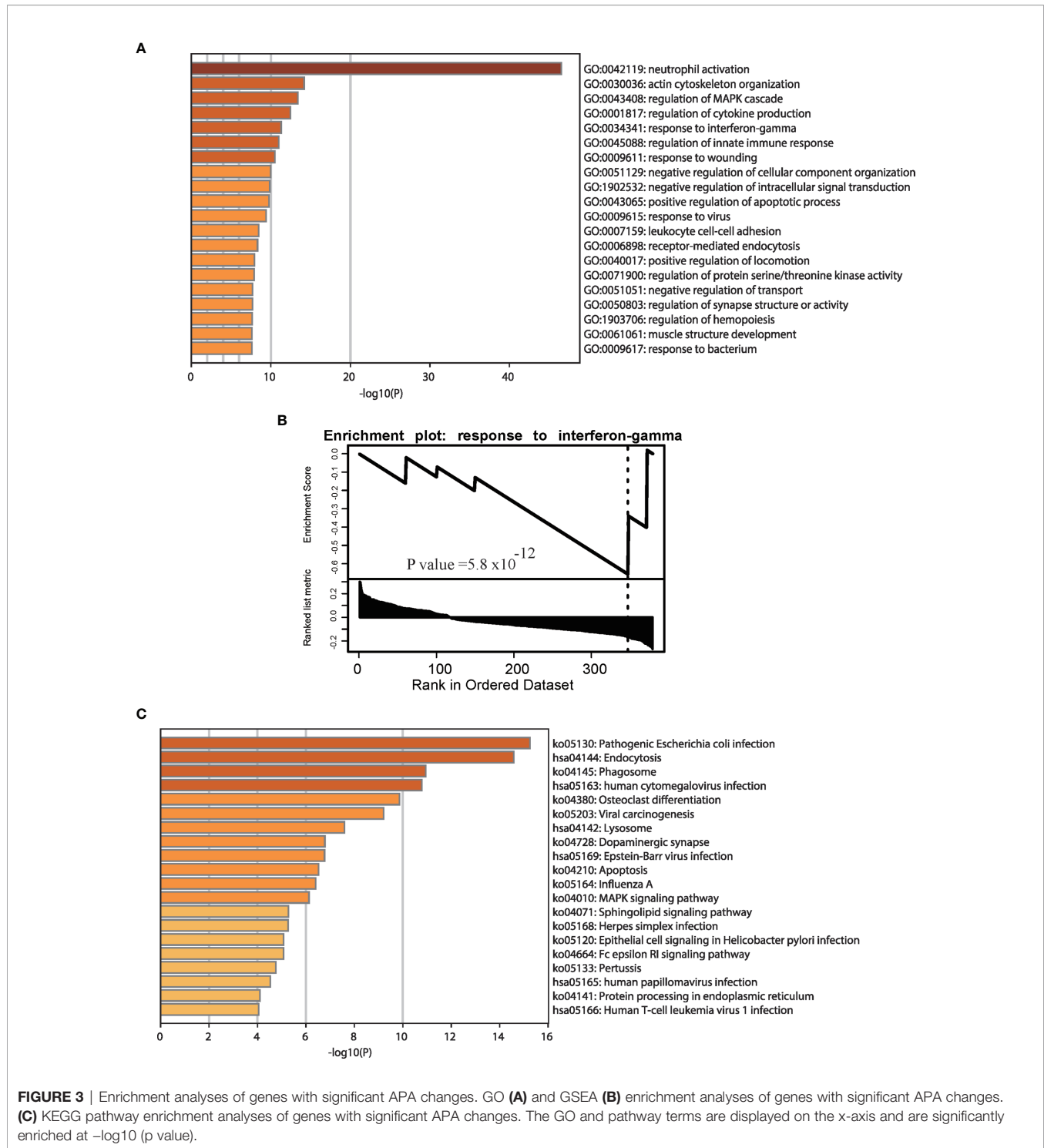


COVID-19 patients. **Figures 5A, B** show that the expression level of most 3' processing factors was significantly higher in COVID-19 patients ( $P < 0.05$ ), especially CPSF2, PAPOLG, FIP1L1, and PCF11. To validate our results, we analyzed RNA-Seq data from other independent COVID-19 datasets and observed the same results (**Figure S4B**), suggesting change of APA is associated with expression level of 3' processing factors in COVID-19 patients.

In brief, we concluded that high expression level of 3' processing factors might be one of the reasons underlying genome-wide APA when patients are infected with SARS-CoV-2. Our results were consistent with previous studies that 3'UTR shortening correlated positively with high expression of polyadenylation factors (32).

## Performance Evaluation of APA and AS

To further determine the power of APA in COVID-19 patients, we compared APA with AS, and AS has been proven to play an important role in COVID-19. We performed alternative splicing analysis for COVID-19 patients by using rMARTs (33), a computational tool employed to detect differential AS events based on RNA-Seq data. We detected a total of 806 events in the COVID-19 sample, including 640 unique AS events, which are 292 SE events, 188 RI events, 47 A5SS events, 81 A3SS events, and 32 MXE events (**Figure 6A**). Moreover, the pathways enriched in AS were immune related (**Figure 6B**) (**Table S3**). Considering that SE was found to be the most prevalent AS event, we conducted principal component

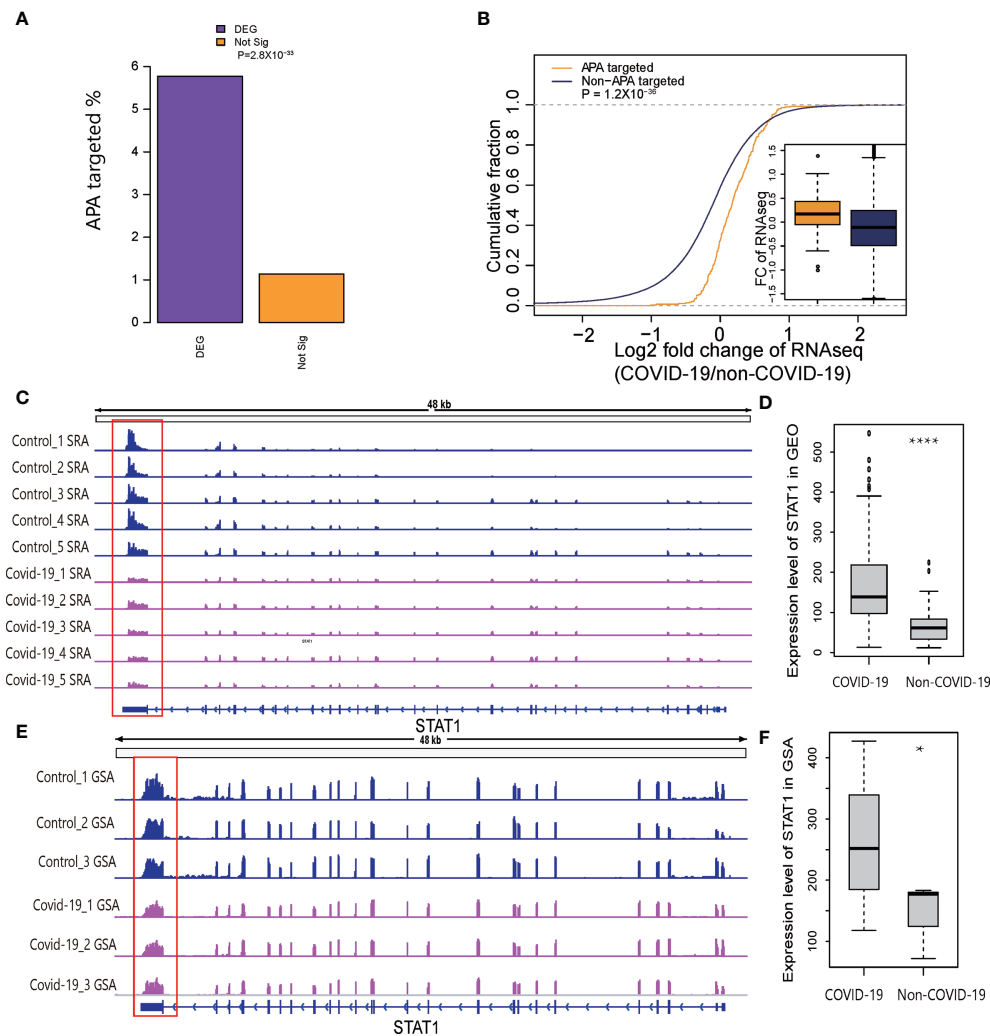


**FIGURE 3** | Enrichment analyses of genes with significant APA changes. GO **(A)** and GSEA **(B)** enrichment analyses of genes with significant APA changes. **(C)** KEGG pathway enrichment analyses of genes with significant APA changes. The GO and pathway terms are displayed on the x-axis and are significantly enriched at  $-\log_{10}(p \text{ value})$ .

analysis (PCA) to construct APA and SE signatures as described previously (34).

We identified two significant components that explained 61% and 12% of the APA variation. They also explained 10.6% and 9.4% of the SE variation, respectively. The first PC (principal component) mainly separated COVID-19 patients from non-COVID-19 patients with regard to APA and SE (**Figures 6C, D**). Then, principal component 1

was selected to act as a signature score predicting COVID-19 status. To further explore the diagnostic potential of the above dynamic APA and SE, we performed receiver operating characteristic (ROC) curve analysis. As illustrated in **Figure 6C**, the area under the curve (AUC) for discriminating COVID-19 from non-COVID-19 was 0.83 for APA, suggesting its high diagnostic potential in COVID-19. Furthermore, the AUC for SE was 0.76 (**Figure 6E**), which



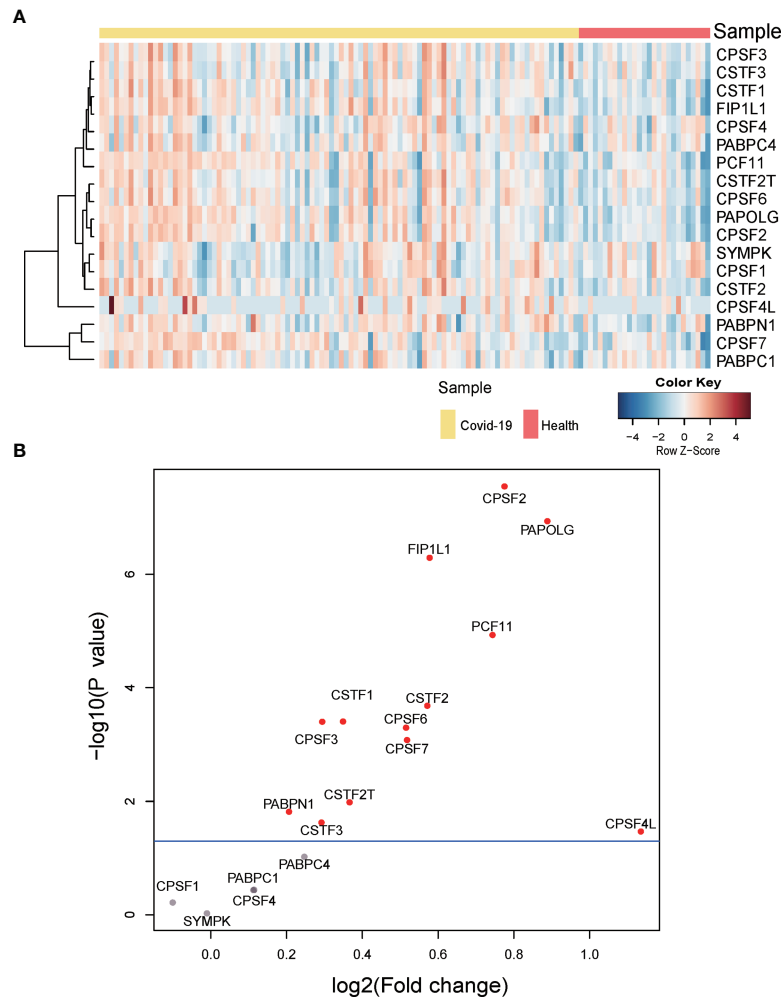
**FIGURE 4** | Relationship of APA and DEGs in COVID-19. **(A)** Barplot representing the percentages of the DEGs or nonsignificant genes enriched for significant APA sites. **(B)** Plot of the cumulative fraction of log<sub>2</sub>-fold change of gene expression ratios in COVID-19 patients comparing genes with significant APA sites versus those overlapping without significant APA sites. The p value of a two-tailed Wilcoxon test is indicated. **(C)** Tracks displaying the read coverage for the STAT1 gene in samples randomly selected from SRP279280. **(D)** Expression level of STAT1 in COVID-19 and non-COVID-19 patients in GEO. **(E)** Tracks displaying the read coverage of validation RNA-Seq data (CRA002390) for the STAT1 gene. **(F)** Expression level of STAT1 in COVID-19 and non-COVID-19 patients in GSA. \*P < 0.05, \*\*\*\*P < 0.05\*10<sup>-4</sup>.

indicated that APA provided greater accuracy than SE in COVID-19. Furthermore, we investigated the association between the clinical outcomes of COVID-19 patients and APA conditions. We found that the APA index was significantly changed in the different hospital free days groups (**Figure S5A**). However, there were not significantly differences in the APA index between different ventilator free day groups and male vs female COVID-19 patients (**Figures S5B, C**). Hence, we fill the knowledge gap of regarding the role of APA in the clinical outcomes of COVID-19 patients.

## DISCUSSION

In brief, we performed *de novo* identification of dynamic APA and AS using two independent datasets of PBMC RNA-Seq data

from COVID-19 patients. We found that APA and gene expression of APA regulators in the hosts was perturbed in 100 COVID-19 patients. This study contributes to a better understanding of the activation and evasion of interferon responses by SARS-CoV-2, and this will be beneficial to the study of new treatment methods for COVID-19. Even though all of the in silico works were performed in two independent datasets, the limitations of our in silico work cannot be ignored. For instance, we could not demonstrate which APA regulator mainly regulates abnormal APA in COVID-19 patients. Wet experiments need to be conducted to demonstrate the detailed regulatory mechanism in future studies. At the same time, the effects of abnormal APA on phenotypes in human immune cell lines were also need to be conducted using wet experiments.



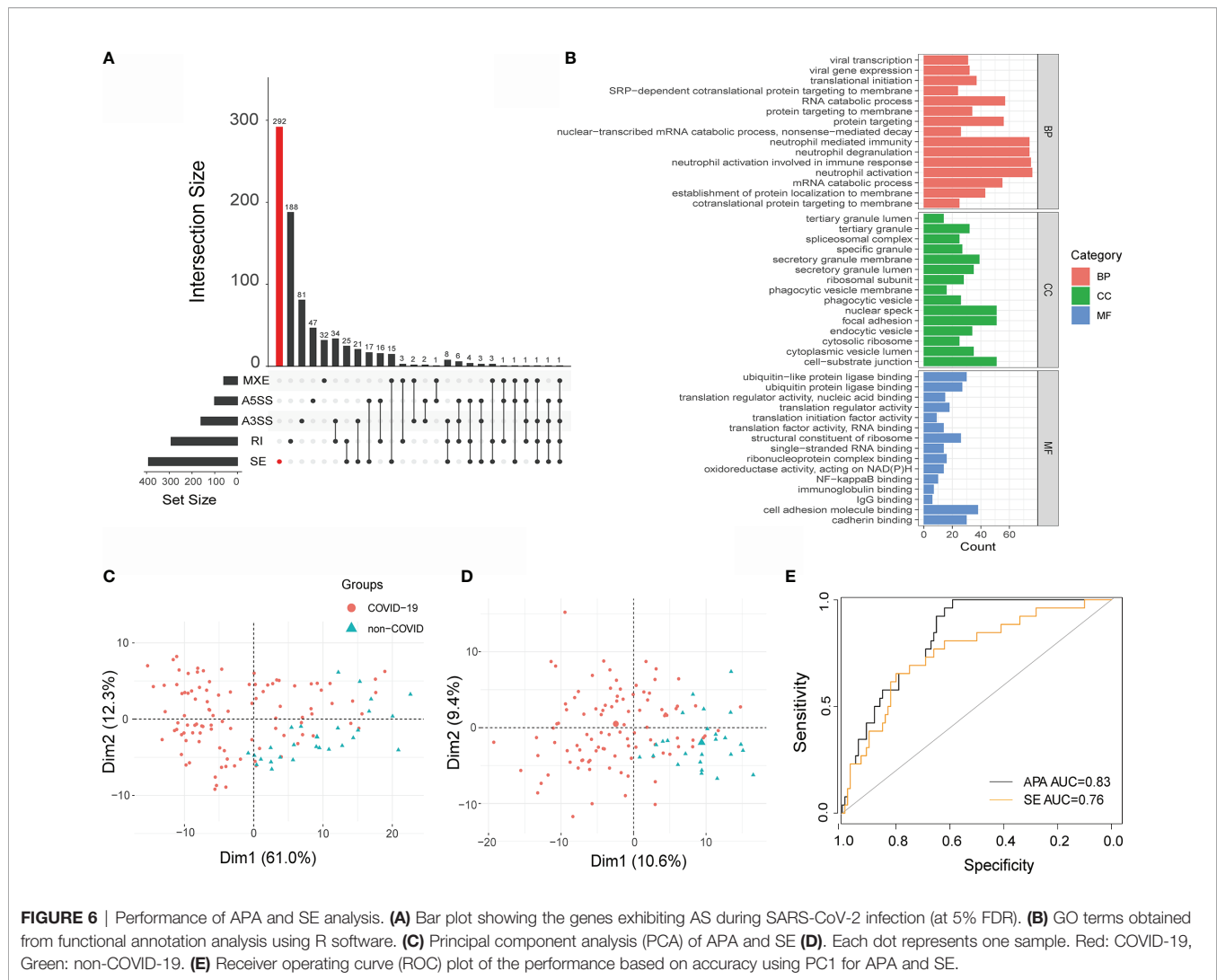
**FIGURE 5** | Gene expression of APA regulators. **(A)** Heatmap of gene expression of regulators of APA. **(B)** Volcano plot of differentially expressed APA regulators. Significant regulators ( $p \text{ value} \leq 0.05$ ) are highlighted in red.

Transcriptome analysis of PBMCs can indicate numerous RNA mis-APA and mis-AS events, which may serve as disease-specific biomarkers (35). Although PBMCs are peripheral blood cells comprising T cells, B cells, NK cells, and monocytes, they are widely used to identify potential biomarkers for COVID-19 (36–38). For instance, RNA-Seq data for COVID-19 PBMCs indicated immune-related transcriptomic profiles (1). However, PBMCs contain many cell types, which is difficult to verify by wet experiments.

It was reported that antiviral pathways and interferon pathways play important roles in COVID-19 patients (38–41), but the exact mechanism remained unclear. Simultaneously, previous studies have found that APA and AS regulated multiple viral immune processes (7, 8, 13, 42–44). However, there is a lack of studies on whether APA affects SARS-CoV-2 infection. Innate lymphoid cells including neutrophils and macrophages secrete signaling factors that regulate innate and

adaptive immune responses. Recent studies have reported that innate immunity may play a more central role in combating SARS-CoV-2 rather than adaptive immunity (45). Based on our results, APA may cause neutrophil activation by affecting the release of inflammatory factors from macrophages. APA may play a central role in innate immunity in COVID-19 patients. In addition, neutrophils can express costimulatory molecules and MHC-II after exposure to cytokines, such as IFN- $\gamma$  (46). APA may also affect the release of antiviral factors such as interferon and exert antiviral effects. It has been reported that IFN-induced STAT1 phosphorylation remains intact in the presence of SARS-CoV-2 ORF6 (47) and that alteration of STAT1 increases susceptibility to virus infections because it is involved in various signaling pathways both upstream and downstream of IFN production. Interestingly, in COVID-19 patients, APA (and AS) genes were enriched in IFN related categories, especially pathogenic *Escherichia coli* infection, endocytosis, phagosome





human cytomegalovirus infection, and Epstein-Barr virus infection signaling pathways.

We revealed that 3'UTR shortening through APA might play an essential role in COVID-19 and that APA upregulated target genes by facilitating miRNA repression escape in SARS-CoV-2 infection, which was consistent with previous report that 3'UTR shortening through APA might upregulate target genes by miRNA repression escape in other infectious diseases. Although previous studies showed that global 3'UTR shortening affects protein abundance (8), our study revealed that the impact of the 3'UTR on protein production may depend on the gene. However, APA of the genes confers different functions and needs further investigation. Most of the *trans* factors assessed were highly expressed in COVID-19 patients, in accordance with previous results that 3'UTR shortening was associated with highly expressed *trans* factors. For instance, the key poly-A trans-factor CSTF64 was significantly upregulated, leading to preferential 3' UTR shortening in tumors (6). The expression of 3' processing

factors was down-regulated when cells were infected by vesicular stomatitis virus, which might be one of the reasons underlying genome-wide APA when cells were infected with viruses (8). Meanwhile, we found that the expression level of 3' processing factors was also altered in COVID-19 patients. Therefore, the question here is how SARS-CoV-2 disrupts APA and AS. According to the research conclusion that SARS-CoV-2 proteins could bind to transcription factor and splicing factors (1), we proposed the following hypothesis: SARS-CoV-2 proteins can bind to APA factors affecting the gene expression level of APA factors to regulate APA. Although SE has been reported to play a key role in SARS-CoV-2 infection (1), our PCA results of dynamic APA and SE showed that APA provided greater accuracy than SE (**Figure 6C**). This conclusion will contribute to a better understanding of mRNA-processing mechanisms in COVID-19 samples. Furthermore, APA may disrupt antigen presentation by MHC in infected cells, and interference with APA and AS might further aid SARS-CoV-2 in evading the host immune response.

Our article contributes to the development of new APA-based therapeutic targets of COVID-19.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding authors.

## AUTHOR CONTRIBUTIONS

SA and YQL collected the omics data and conducted the analysis. YL and JC produced the diagrams. SA, LY, and JS wrote the draft of the manuscript. RZ, JL, QC, HW, PP, and HL revised the draft. JJ supervised the project. All authors contributed to the article and approved the submitted version.

## REFERENCES

- Banerjee AK, Blanco MR, Bruce EA, Honson DD, Chen LM, Chow A, et al. SARS-CoV-2 Disrupts Splicing, Translation, and Protein Trafficking to Suppress Host Defenses. *Cell* (2020) 183:1325–39.e1321. doi: 10.1016/j.cell.2020.10.004
- Shulman ED, Elkon R. Cell-Type-Specific Analysis of Alternative Polyadenylation Using Single-Cell Transcriptomics Data. *Nucleic Acids Res* (2019) 47:10027–39. doi: 10.1093/nar/gkz781
- Zhang Y, Liu L, Qiu Q, Zhou Q, Ding J, Lu Y, et al. Alternative Polyadenylation: Methods, Mechanism, Function, and Role in Cancer. *J Exp Clin Cancer Res* (2021) 40:51. doi: 10.1186/s13046-021-01852-7
- Bartel DP. MicroRNAs: Target Recognition and Regulatory Functions. *Cell* (2009) 136:215–33. doi: 10.1016/j.cell.2009.01.002
- Wang ET, Sandberg R, Luo S, Khrebtkova I, Zhang L, Mayr C, et al. Alternative Isoform Regulation in Human Tissue Transcriptomes. *Nature* (2008) 456:470–6. doi: 10.1038/nature07509
- Xia Z, Donehower LA, Cooper TA, Neilson JR, Wheeler DA, Wagner EJ, et al. Dynamic Analyses of Alternative Polyadenylation From RNA-Seq Reveal a 3'-UTR Landscape Across Seven Tumour Types. *Nat Commun* (2014) 5:5274. doi: 10.1038/ncomms6274
- Gruber AJ, Zavolan M. Alternative Cleavage and Polyadenylation in Health and Disease. *Nat Rev Genet* (2019) 20:599–614. doi: 10.1038/s41576-019-0145-z
- Jia X, Yuan S, Wang Y, Fu Y, Ge Y, Ge Y, et al. The Role of Alternative Polyadenylation in the Antiviral Innate Immune Response. *Nat Commun* (2017) 8:14605. doi: 10.1038/ncomms14605
- Fatoba AJ, Maharaj L, Adeleke VT, Okpeku M, Adeniyi AA, Adeleke MA. Immunoinformatics Prediction of Overlapping CD8(+) T-Cell, IFN-Gamma and IL-4 Inducer CD4(+) T-Cell and Linear B-Cell Epitopes Based Vaccines Against COVID-19 (SARS-CoV-2). *Vaccine* (2021) 39:1111–21. doi: 10.1016/j.vaccine.2021.01.003
- Zhang F, Mears JR, Shakib L, Beynor JI, Shanaj S, Korsunsky I, et al. IFN-Gamma and TNF-Alpha Drive a CXCL10+ CCL2+ Macrophage Phenotype Expanded in Severe COVID-19 Lungs and Inflammatory Diseases With Tissue Inflammation. *Genome Med* (2021) 13:64. doi: 10.1186/s13073-021-00881-3
- Stacey SN, Sulem P, Jonasdottir A, Masson G, Gudmundsson J, Gudbjartsson DF, et al. A Germline Variant in the TP53 Polyadenylation Signal Confers Cancer Susceptibility. *Nat Genet* (2011) 43:1098–103. doi: 10.1038/ng.926
- Di Giammartino DC, Nishida K, Manley JL. Mechanisms and Consequences of Alternative Polyadenylation. *Mol Cell* (2011) 43:853–66. doi: 10.1016/j.molcel.2011.08.017

## FUNDING

China Postdoctoral Science Foundation (2020M683623XB), The 64th Batch of the China Postdoctoral Science Foundation (2018M643382), Young Scientists Fund of the Guangxi Natural Science Foundation (2018GXNSFBA281014), Guangxi Bagui Scholar (to JJ), Guangxi Medical University Training Program for Distinguished Young Scholars (to JJ), Guangxi Science Fund for Distinguished Young Scholars (2018GXNSFFA281001), National Natural Science Foundation of China (82160389).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2021.756288/full#supplementary-material>

- Ashraf U, Benoit-Pilven C, Lacroix V, Navratil V, Naffakh N. Advances in Analyzing Virus-Induced Alterations of Host Cell Splicing. *Trends Microbiol* (2019) 27:268–81. doi: 10.1016/j.tim.2018.11.004
- Thompson MG, Dittmar M, Mallory MJ, Bhat P, Ferretti MB, Fontoura BM, et al. Viral-Induced Alternative Splicing of Host Genes Promotes Influenza Replication. *Elife* (2020) 9:e55500. doi: 10.7554/eLife.55500
- Wang J, Pan Y, Shen S, Lin L, Xing Y. rMATS-DVR: rMATS Discovery of Differential Variants in RNA. *Bioinformatics* (2017) 33:2216–7. doi: 10.1093/bioinformatics/btx128
- Overmyer KA, Shishkova E, Miller IJ, Balnis J, Bernstein MN, Peters-Clarke TM, et al. Large-Scale Multi-Omic Analysis of COVID-19 Severity. *Cell Syst* (2021) 12:23–40.e27. doi: 10.1016/j.cels.2020.10.003
- Xiong Y, Liu Y, Cao L, Wang D, Guo M, Jiang A, et al. Transcriptomic Characteristics of Bronchoalveolar Lavage Fluid and Peripheral Blood Mononuclear Cells in COVID-19 Patients. *Emerg Microbes Infect* (2020) 9:761–70. doi: 10.1080/22221751.2020.1747363
- Kim D, Langmead B, Salzberg SL. HISAT: A Fast Spliced Aligner With Low Memory Requirements. *Nat Methods* (2015) 12:357–60. doi: 10.1038/nmeth.3317
- Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. StringTie Enables Improved Reconstruction of a Transcriptome From RNA-Seq Reads. *Nat Biotechnol* (2015) 33:290–5. doi: 10.1038/nbt.3122
- An S, Huang W, Huang X, Cun Y, Cheng W, Sun X, et al. Integrative Network Analysis Identifies Cell-Specific Trans Regulators of M6a. *Nucleic Acids Res* (2020) 48:1715–29. doi: 10.1093/nar/gkz1206
- Zhang Y, Shen L, Shi Q, Zhao G, Wang F. Comprehensive Analysis of APA Events and Their Association With Tumor Microenvironment in Lung Adenocarcinoma. *Front Genet* (2021) 12:645360. doi: 10.3389/fgene.2021.645360
- Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative Genomics Viewer. *Nat Biotechnol* (2011) 29:24–6. doi: 10.1038/nbt.1754
- Huang da W, Sherman BT, Lempicki RA. Bioinformatics Enrichment Tools: Paths Toward the Comprehensive Functional Analysis of Large Gene Lists. *Nucleic Acids Res* (2009) 37:1–13. doi: 10.1093/nar/gkn923
- Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape Provides a Biologist-Oriented Resource for the Analysis of Systems-Level Datasets. *Nat Commun* (2019) 10:1523. doi: 10.1038/s41467-019-09234-6
- Agarwal V, Bell GW, Nam JW, Bartel DP. Predicting Effective microRNA Target Sites in Mammalian mRNAs. *Elife* (2015) 4:e05005. doi: 10.7554/eLife.05005
- Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, et al. Simple Combinations of Lineage-Determining Transcription Factors Prime Cis-

- Regulatory Elements Required for Macrophage and B Cell Identities. *Mol Cell* (2010) 38:576–89. doi: 10.1016/j.molcel.2010.05.004
27. Tian B, Hu J, Zhang H, Lutz CS. A Large-Scale Analysis of mRNA Polyadenylation of Human and Mouse Genes. *Nucleic Acids Res* (2005) 33:201–12. doi: 10.1093/nar/gki158
  28. Schultze JL, Aschenbrenner AC. COVID-19 and the Human Innate Immune System. *Cell* (2021) 184:1671–92. doi: 10.1016/j.cell.2021.02.029
  29. Karki R, Sharma BR, Tuladhar S, Williams EP, Zalduondo L, Samir P, et al. Synergism of TNF-Alpha and IFN-Gamma Triggers Inflammatory Cell Death, Tissue Damage, and Mortality in SARS-CoV-2 Infection and Cytokine Shock Syndromes. *Cell* (2021) 184:149–68.e117. doi: 10.1016/j.cell.2020.11.025
  30. Garcia-Sastre A. Ten Strategies of Interferon Evasion by Viruses. *Cell Host Microbe* (2017) 22:176–84. doi: 10.1016/j.chom.2017.07.012
  31. Thomas PE, Wu X, Liu M, Gaffney B, Ji G, Li QQ, et al. Genome-Wide Control of Polyadenylation Site Choice by CPSF30 in Arabidopsis. *Plant Cell* (2012) 24:4376–88. doi: 10.1105/tpc.112.096107
  32. Shepard PJ, Choi EA, Lu J, Flanagan LA, Hertel KJ, Shi Y. Complex and Dynamic Landscape of RNA Polyadenylation Revealed by PAS-Seq. *RNA* (2011) 17:761–72. doi: 10.1261/rna.2581711
  33. Shen S, Park JW, Huang J, Dittmar KA, Lu ZX, Zhou Q, et al. MATS: A Bayesian Framework for Flexible Detection of Differential Alternative Splicing From RNA-Seq Data. *Nucleic Acids Res* (2012) 40:e61. doi: 10.1093/nar/gkr1291
  34. Zhang B, Wu Q, Li B, Wang D, Wang L, Zhou YL. M(6)A Regulator-Mediated Methylation Modification Patterns and Tumor Microenvironment Infiltration Characterization in Gastric Cancer. *Mol Cancer* (2020) 19:53. doi: 10.1186/s12943-020-01170-0
  35. Sznajder LJ, Scotti MM, Shin J, Taylor K, Ivankovic F, Nutter CA, et al. Loss of MBNL1 Induces RNA Misprocessing in the Thymus and Peripheral Blood. *Nat Commun* (2020) 11:2022. doi: 10.1038/s41467-020-15962-x
  36. Sadanandam A, Bopp T, Dixit S, Knapp D, Emperumal CP, Vergidis P, et al. A Blood Transcriptome-Based Analysis of Disease Progression, Immune Regulation, and Symptoms in Coronavirus-Infected Patients. *Cell Death Discovery* (2020) 6:141. doi: 10.1038/s41420-020-00376-x
  37. Caruso FP, Scala G, Cerulo L, Ceccarelli M. A Review of COVID-19 Biomarkers and Drug Targets: Resources and Tools. *Brief Bioinform* (2021) 22:701–13. doi: 10.1093/bib/bbaa328
  38. Mathew D, Giles JR, Baxter AE, Oldridge DA, Greenplate AR, Wu JE, et al. Deep Immune Profiling of COVID-19 Patients Reveals Distinct Immunotypes With Therapeutic Implications. *Science* (2020) 369. doi: 10.1126/science.abc8511
  39. Wise J. Covid-19: Rare Immune Response may Cause Clots After AstraZeneca Vaccine, Say Researchers. *BMJ* (2021) 373:n954. doi: 10.1136/bmj.n954
  40. Putri DU, Wang CH, Tseng PC, Lee WS, Chen FL, Kuo HP, et al. Profiles of Peripheral Immune Cells of Uncomplicated COVID-19 Cases With Distinct Viral RNA Shedding Periods. *Viruses* (2021) 13(3):514. doi: 10.3390/v13030514
  41. Lee EE, Song KH, Hwang W, Ham SY, Jeong H, Kim JH, et al. Pattern of Inflammatory Immune Response Determines the Clinical Course and Outcome of COVID-19: Unbiased Clustering Analysis. *Sci Rep* (2021) 11:8080. doi: 10.1038/s41598-021-87668-z
  42. Xu X, Mann M, Qiao D, Brasier AR. Alternative mRNA Processing of Innate Response Pathways in Respiratory Syncytial Virus (RSV) Infection. *Viruses* (2021) 13(2):218. doi: 10.3390/v13020218
  43. Lin J, Hu Y, Nunez S, Foulkes AS, Cieply B, Xue C, et al. Transcriptome-Wide Analysis Reveals Modulation of Human Macrophage Inflammatory Phenotype Through Alternative Splicing. *Arteriosclerosis Thrombosis Vasc Biol* (2016) 36:1434–47. doi: 10.1161/ATVBAHA.116.307573
  44. Fabozzi G, Oler AJ, Liu P, Chen Y, Mindaye S, Dolan MA, et al. Strand-Specific Dual RNA Sequencing of Bronchial Epithelial Cells Infected With Influenza A/H3N2 Viruses Reveals Splicing of Gene Segment 6 and Novel Host-Virus Interactions. *J Virol* (2018) 92:e00518–18. doi: 10.1128/JVI.00518-18
  45. Mallapaty S. Kids and COVID: Why Young Immune Systems are Still on Top. *Nature* (2021) 597:166–8. doi: 10.1038/d41586-021-02423-8
  46. Vono M, Lin A, Norrby-Teglund A, Koup RA, Liang F, Lore K. Neutrophils Acquire the Capacity for Antigen Presentation to Memory CD4(+) T Cells *In Vitro* and *Ex Vivo*. *Blood* (2017) 129:1991–2001. doi: 10.1182/blood-2016-10-744441
  47. Lei X, Dong X, Ma R, Wang W, Xiao X, Tian Z, et al. Activation and Evasion of Type I Interferon Responses by SARS-CoV-2. *Nat Commun* (2020) 11:3810. doi: 10.1038/s41467-020-17665-9
- Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 An, Li, Lin, Chu, Su, Chen, Wang, Pan, Zheng, Li, Jiang, Ye and Liang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.