Single Extracellular Vesicles (EV) Proteomic Profiling Altered and Identifies Co-Localization of SARS-CoV-2 Nucleocapsid Protein with CD81/Integrin-Rich EV Subpopulation in Sputum Samples of COVID-19 Patients

Ruiting Sun
First Affiliated Hospital of Guangzhou Medical University https://orcid.org/0000-0002-4562-9991
Yanling Cai
Shenzhen Second People's Hospital
Yumin Zhou
the first affiliated hospital of guangzhou medical university
Ge Bai
the first affiliated hospital of Guangzhou medical university
Airu Zhu
The First Affiliated Hospital of Guangzhou Medical University
Panyue Kong
The First Affiliated Hospital of Guangzhou Medical University
Jing Sun
The First Affiliated Hospital of Guangzhou Medical University
Yimin Li
The First Affiliated Hospital of Guangzhou Medical University
Yuefei Liu
Shenzhen secretech Co.Ltd
Wenting Liao
Shenzhen secretech Co.Ltd
Jiye Liu
Shenzhen Secretech Co. Ltd
Nan Cui
Shenzhen secretech Co.Ltd
Jinsheng Xiang
Shenzhen secretech Co.Ltd
Bing Li
Guangzhou Medical University

Jincun Zhao
The First Affiliated Hospital of Guangzhou Medical University
Di Wu
Shenzhen secretaria Co.Ltd
Pixin Ran (pixinRan@hotmail.com)
The First Affiliated Hospital of Guangzhou Medical University https://orcid.org/0000-0003-2221-8986

Research Article

Keywords: COVID-19, SARS-CoV-2, Single extracellular vesicles, Inflammatory response, Extracellular vesicles e subpopulation

Posted Date: December 30th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1154862/v1

Title:

Full title: Single extracellular vesicles (EV) proteomic profiling altered and identifies co-localization of SARS-CoV-2 nucleocapsid protein with CD81/integrin-rich EV subpopulation in sputum samples of COVID-19 patients

Running title: Single EVs proteomic profiling in sputum samples of COVID-19 patients

Ruiting Sun[1#], *Yanling Cai* [2,3#], *Yumin Zhou*[1#], *Ge Bai*[1#], *Airu Zhu*[1#], *Panyue Kong*[1], *Jing Sun*[1], *Yimin Li*[1], *Yuefei Liu*[3], *Wenting Liao*[3], *Jiye Liu*[3], *Nan Cui*[3], *Jinsheng Xiang*[3], *Bing Li*[1], *Jincun Zhao*[1], *Di Wu*[3,4*], ***Pixin Ran***[1*]

[1]*National Center for Respiratory Medicine, State Key Laboratory of Respiratory Disease, National Clinical Researcher Center for Respiratory Diseases, Guangzhou Institute of Respiratory Health, the First Affiliated Hospital of Guangzhou Medical University, Guangzhou, Guangdong 510120, China*
[2]*Shenzhen Second People's hospital, Postdoctoral Workstation of Zhongshan School of Medicine, Sun Yat-Sen University, Shenzhen, Guangdong 518000, China*
[3] *Shenzhen Secretech Co. Ltd, Shenzhen, Guangdong 518000, China*
[4] *Vesicode AB, Solna, Sweden*

#These authors contributed equally to this work.
*Corresponding author:
Pixin Ran, MD, PhD. Email: pixinRan@hotmail.com / pxran@gzhmu.edu.cn
Di Wu, PhD. Email：di.wu@vesicode.se

Keywords

COVID-19, SARS-CoV-2, Single extracellular vesicles, Inflammatory response, Extracellular vesicles e subpopulation

Highlights

1. **Single-EV protein sequencing profiling (proximity barcoding assay) was applied to analyze sputum EVs from 20 hospitalized COVID-19 patients and 20 healthy individuals.**
2. **Sputum EVs from COVID-19 patients carried SARS-CoV-2 N protein，which was correlated with IL-6 and TGF-β expression levels.**
3. **Viral protein was found to co-localize with a specific EV subpopulation featured expression of multiple proteins, including CD81, SNAI2 and integrins.**
4. **Single extracellular vesicles proteomic profile in the sputum of COVID-19 patients reveals altered host and virus-derived proteins in relation to infection**

Abstract

Understanding the pathogenesis of SARS-CoV-2 is crucial to respond to the current coronavirus disease 2019（COVID-19）pandemic. Sputum samples from 20 COVID-19 patients and healthy controls were collected, respectively. During the isolation of infectious SARS-CoV-2 virus, EV-like vesicles were associated with virions under a transmission electron microscope. Next, the expression of IL6 and TGF-β increased in EVs derived from the sputum of patients, and these were highly correlated with the expression of the SARS-CoV-2 N protein. Further, proximity barcoding assay (PBA) was used to investigate the immune-related proteins in the EVs, and the relationship between EVs and SARS-CoV-2 N protein in COVID-19 patients' samples. Particularly, to investigate the differential contribution of the specific EV subsets, the protein expression of a single EV was detected and analyzed for the first time. Among the 40 EV subpopulations, 18 were found to have significant differences. The EV subpopulation regulated by CD81 were most likely to correlate with the changes in the pulmonary microenvironment after SARS-CoV-2 infection. This study provides evidence on

the association between EVs and the SARS-CoV-2 virus, give a deep insight into the possible pathogenesis of SARS-CoV-2 infection and the possibility of nanoparticles drug intervention in viral infection.

Introduction

Globally, there are more than 179.27 million confirmed coronavirus disease 2019 (COVID-19) cases, with more than 5,089,384 deaths reported to WorldOmeters as of Nov. 11，2021. The number of countries having more than 10 million confirmed cases worldwide increased to 36. With the global outbreak of COVID-19 and the limited availability of clinical treatments, researchers around the world are looking for potential drugs to treat COVID-19 [1, 2]. Due to the severe clinical progression of COVID-19 patients with increased inflammation and immune response disorders, a considerable number of patients had severe pneumonia, and even developed acute respiratory distress syndrome (ARDS) [3]. Cytokine storm-related syndrome (IL6, IL1, TNFα, etc.) has been proposed as the trigger for ARDS. Accordingly, treatments (such as corticosteroids) to control the inflammatory cytokine signaling are being used to reduce the mortality of patients with COVID-19 [4, 5]. Extracellular vesicles (EV), especially exosomes, have emerged as key mediators in various physio-pathological processes related to virus infection and are actively involved in response to virus-induced injury [6], mediating inflammatory response and inflammation-related protection, since they display dual beneficial and detrimental roles [7].

Exosomes, which are functional vehicles secreted by various types of cells, possess a diameter of 30–130 nm and carry a complex cargo of proteins, lipids, and nucleic acids. It has been proven that the hepatitis A virus can hijack exosome membranes and transport virus pathogenesis-related proteins [8], genomes, and even virus particles to all parts of the body, using the characteristics of the free shuttle, and the accurate localization of exosomes between

the host and target cells [9]. Blocking the vesicle release in HCV-positive cells increased the intracellular dsRNA levels, and the activation of toll-like receptor 3, inhibited HCV replication [10]. Therefore, it was speculated that EVs might also be involved in transmitting severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2). The membrane hijacking by SARS-CoV-2 likely promotes the virus spread within the lung through EV-like vesicles. However, little is known about the role of EVs during SARS-CoV-2 infection and subsequent immune response. Furthermore, the EV subsets in sputum samples and subsequent changes in their proteomic features during SARS-CoV-2 infection remains poorly understood. Therefore, it is of great significance to investigate the relationship between EVs and SARS-CoV-2 infection. In particular, the identification of proteins transmitted by EVs would help identify potential drug targets, develop vaccines, or reasonably reuse existing drugs according to the specific protein expression.

The present study aims to identify the EV and virus co-expression proteins using a library of DNA tagged antibodies. EVs obtained from healthy donors and COVID-19 patients were examined. Single EVs were classified into subpopulations according to their proteomic features via an unsupervised machine learning method. The subpopulations were visualized. Meanwhile, the proportion of each subpopulation was quantified, and the proteomic fingerprints for each subpopulation were profiled. Then, the co-expression of the N protein of SARS-CoV-2 (SARS-CoV-2 N) with other EV proteins was selected from the protein combination dataset and analyzed to predict virus-EV association. Thus, the present study may increase the knowledge on the EV subset involved in the pathogenesis of the COVID-19 disease, which could be useful for the design of therapeutic strategies to fight SARS-CoV-2 infection, and the screening of potential targets and drugs for COVID-19 treatment.

Results

Demographic information and Characterization of EVs from the sputum samples of COVID-19 patients

Twenty severe COVID-19 patients (nCOV group, n=20) and healthy controls (HC group, n=20) were enrolled in the current study. Most of them were over 50 years old, with an average age of 56.6 and 56.7 years old, respectively. The sputum of these patients was collected when they were still in the ICU. Twelve of the 20 COVID-19 required ventilator, who was diagnosed with acute respiratory distress syndrome (ARDS). The patient information is provided in Figure 1A.

The sputum samples were collected as described in the Materials and Methods section. The SARS-CoV-2 virus was successfully isolated from the sputum supernatant of a COVID-19 patient, demonstrating an infectious virus in the sputum. (**Figure 1B**, red arrow). Surprisingly, EV-like vesicles were found close to the virions (**Figure 1B**, blue arrow). Next, the EVs were isolated and identified from the sputum by differential ultracentrifugation. The EVs were cup-shaped and had a lipid bilayer membrane vesicle structure by electron microscopy (**Figure 1C**). To further characterize the nature of these released vesicles, particle tracking was performed using a NanoSight instrument. The nanoparticle tracking analysis (NTA) provided the particle size distribution profiles and concentration measurements. As shown in **Figure 1D**, the primary peak was observed at approximately 86 nm, consistent with the size of most EVs（30-200 nm). The size distribution was quite monodispersed. The particle from sputum sample from the COVID-19 patients had a concentration of $2.35\times10^{8} \pm 2\times10^{6}$ particles/ml.

Patients with COVID-19 secreted more proteins in individual EV, and EVs participated in the immune response

The EV and virus co-expression proteins were identified using the PBA method. The scheme of the workflow is illustrated in **Figure 2A**. The antibody-conjugated oligonucleotides were brought into the proximity on the same EV due to the protein-antibody interaction, thereby obtaining the same EV tag barcoding [13]. EVs obtained from different sources were characterized by the presence of specific combinations of surface proteins and their abundance, allowing each EV to be quantified in the mixed samples, to serve as markers for specific engagement in the disease. After library construction and sequencing, the original data were obtained in fastQ file format. After quality control and tag extraction, the file of the identified individual EVs and detected proteins are summarized for each sample.

From the 5μl sputum samples, the mean number of EVs detected in the PBA exhibited an increasing trend after SARS-CoV-2 infection (**Figure 2B**, left), and the number of proteins detected in EVs obtained from the nCOV group were doubled (**Figure 2B**, middle) than the HC group. Among the proteins under investigation, sputum EVs in the nCOV group had a higher number of detected proteins (3.6 proteins/EV) than the 1.9 proteins/EV in the HC group (**Figure 2B**, right). The samples under investigation included the sputum of patients with COVID-19 infection (nCOV, n=20), healthy controls (HC, n=20) and PBS negative controls (PBS, n=4). As shown in **Figure 2C**, SARS-CoV-2 N protein signals could be detected in EVs obtained from COVID-19 patients (**Figure 2C**). In the control group, the protein signal of some individuals was slightly higher than that of PBS, which was considered as an acceptable systematic error and antibody nonspecific binding. In the subsequent analysis, the SARS-CoV-2 N protein was centralized, and the data for SARS-CoV-2 N expression of the control group was set as 0 for the data processing of the nCOV group. By the summation of

signals of each detected protein on all EVs of the sample, the EV associated protein expression of each sample was obtained.

Consistent with previous reports on the expression of cytokines in serum [14], the expression levels of IL6 and TGF-β were also increased in EVs of COVID-19 patients. Furthermore, this elevation was highly correlated with the SARS-CoV-2 N protein expression (**Figure 2D**). We also identified other proteins that were significantly increased after SARS-CoV-2 infection, including T-cell activation marker CD26, human leukocyte antigen HLA-A, and adhesion molecule MAdCAM-1 (mucosal addressing cell adhesion molecule-1), which were overexpressed in inflammatory mucosal tissues (**Fig. S1**). These results show that EVs were involved in the immune response to COVID-19. However, although immunoglobulin A (IgA) might be higher than IgM (consistent with the findings in serum [15]), there was no significant difference in the total expression of IgA and IgM in sputum EVs of healthy controls and patients with COVID-19 (**Figure 2E**). After TMM normalization, protein expression heatmap (**Figure 2F**) showed that nCOV patients have a general shift of EV proteomic profile compared to HC samples, although with exceptions. Differentially expression proteins were analyzed in a volcano plot after normalizing TMM protein expression data and then generated the dot plot (**Fig. S2**). Compared to HC group, the abundance of TROP-2, CD36, EGFR, IgA and IgM decreased in nCOV group while CDH1 (E-cadherin1), ZEB-1 and ZO-1 increased.

EV subpopulations atlas and the featured change in patients with COVID-19.

The algorithm FlowSOM was used to analyze the behavior of all markers on all individual EVs, and the clusters of EVs were generated using a self-organizing map. The clusters, which were the EV subpopulations, were determined according to proteomic fingerprints of each EV. The investigators detected 9,377,119 EVs with an average of 234,428 EVs per sample (**Figure**

2B). The dimensionality reduction indicated the substantial phenotypic similarity and differences between patients with COVID-19 and controls. The t-distributed stochastic neighbor embedding (tSNE) plot for each sample is shown in Figure 3A, which identified 40 clusters. The clustering of individual EVs obtained from all samples was displayed in the tSNE plot, in which 40 clusters were color labelled (**Figure 3B**).

Next, a modeling approach was employed to detect the features that distinguish healthy individuals from infected individuals. **Figure 4A** shows the similarity and differences in EV proteomics between the sputum sample of the nCOV and HC groups, in which all EVs in the nCOV group were colored green, while EVs in the HC group were colored red. Through these red and green markers, the different subpopulations of EVs between these two groups can be more intuitively distinguished (**Figure 4A**). The proteomic similarity of EVs was observed in the tSNE plot. The proportion of each subpopulation was quantified. Among the 40 subpopulations, 18 clusters were found with significant differences. These were cluster 2, 3, 4, 6, 7, 9, 10, 12, 13, 14, 32 and 34, which had a significantly elevated ratio of EV subpopulations in the nCOV group, while the ratio for the subpopulation decrease in cluster 16, 21, 22, 26, 27 and 33 (**Figure 4B** and **S3**).

Among these, cluster 2, 3, 4, 12, 13, 34 and 33 accounted more in the distribution of the subpopulation with differences, and the difference was particularly significant. We further analyzed the seven groups, and the proteomic fingerprints for each subpopulation were profiled. For each differentially expressed EV subpopulation, the location in the total EVs and the top 7 featured proteins are shown in Figure 5. First, we were concerned about cluster 2, which constituted 4.92% of all EVs in the nCOV group, and only 0.55% of all EVs in the HC group (**Figure 5A** and **S3**). The EVs in cluster 2 contained a large amount of SARS-CoV-2 N protein. These should be the EVs directly contacted with SARS-CoV-2 or secreted by the cells responsible for viral replication. These EVs highly expressed exosome biomarker CD81, and

the following cell adhesion molecules: epithelial cell adhesion molecule (EpCAM), CDH1, ITGB4, ITGA5, SNAI2, etc. **Figure 5B** shows the proteomic characteristics of cluster 2 (100 proteins). Clusters 3, 4, 12, 13 and 34 are EVs that increased in the nCOV group, with highly expressed protein CLEC2A, CD81, ITGB3, CD151 and ITGB2, respectively (**Figure 5C**). In contrast, cluster 33 was reduced, and comprised 3.37% of the EVs in the nCOV group, while this comprised of 14.12% of EVs in HC group (**Fig. S3**). Cluster 33 featured a higher expression of EGFR and IgA (**Figure 5D**). The proteomic profiles are shown in Figure 5D.

EVs regulated by CD81 are more likely to carry SARS-CoV-2 proteins

Protein combinations are defined as the co-localization of two proteins on the same individual EV, which can be considered as the fingerprints of individual EVs. The investigators employed the correlation analysis heat map triangular modeling, and profiled the protein co-expression on individual EVs obtained from sputum samples of COVD-19 infected individuals. The quantity of each possible pair of co-expressed proteins was obtained as the protein combination dataset, and this was used as input variables for the abundance and differential analysis (**Figure 6A**). The protein combinations exhibited a universal increasing trend in the nCOV group, except for the combinations of EGFR and IgA. The co-expression between the integrin subgroups significantly increased.

To investigate the co-localization of viral protein with EVs, we further analyzed the combinations of SARS-CoV-2 N protein with other proteins on individual EVs. Among the markers that regulate EVs, the co-expression of CD9, CD63, CD81 and Alix with SARS-CoV-2 N were calculated (**Figure 6B**), and found that the EVs regulated by CD81 were more likely to bind to the SARS-CoV-2 N protein (**Figure 6C**). In addition, we found that cluster 2 (**Figure 5B**), cluster 4 (**Figure 5C**), cluster 6, cluster 7, cluster 12 and cluster 34 were the EVs that were highly expressed after SARS-CoV-2 infection, while CD81 was highly

expressed in both clusters. In particular, the protein matrix of cluster 4 shows that the expression of CD81 is the most and abnormally high (**Figure 6D**). These results suggest that the EVs regulated by CD81 are the most likely subpopulations of EVs that cause the changes in the pulmonary microenvironment after SARS-CoV-2 infection. The protein expression distribution of CD81(red) and SARS-CoV-2 N (blue) in all EV and the green part represented the co-expression region (**Figure 6E**).

Discussion

In the current study we isolated and identified EVs from the sputum of COVID-19 patient to investigate EV inflammatory and immune responses in COVID-19 patients. We found EV-like vesicles that coexisted alongside virions (**Figure 1A**), and the mean number of EVs showed an increasing trend after the SARS-CoV-2 infection (**Figure 2B**). The nucleocapsid protein of SARS-CoV-2 (SARS-CoV-2 N) is an important structural protein, which is located in the core part of the virus particle, and binds to the viral RNA, playing an important role in the process of virus packaging and other process [16, 17]. As expected, the SARS-CoV-2 N protein can be detected in EVs obtained from patients with COVID-19 infection: when the mean signal value in the control group was taken as the baseline (the blue dotted line in **Figure 2C**), 19 of 20 individuals were detected for SARS-CoV-2 N protein in EVs; when the maximum signal value in the control group was taken as the baseline (the red dotted line in **Figure 2C**), 12 of 20 individuals were detected. In any case, the above results prove that EVs might have some roles in SARS-CoV-2 transmission. That maybe membrane hijacking by SARS-CoV-2 likely promotes the virus spread through Exosome-like vesicles, which requires further studies.

We used the systems biology approach (Proximity Barcoding Assay, PBA) to determine the expression of membrane proteins in EVs with and without COVID-19 infection. The

amount of protein encapsulated in EVs obtained from sputum significantly increased in patients with COVID-19 (**Figure 2B**), and the viral infection stimulated the EVs secretion. The cell and animal models of SARS-CoV-2 infection [2], in addition to the serum profiling of COVID-19 patients, consistently revealed the unique and inappropriate inflammatory response [18]. The investigators detected the increased expression of IL-6 and TGF-β in EVs obtained from COVID-19 patients (**Figure 2D**), which is consistent with the previous results in peripheral blood [19]. Furthermore, the expression of IL-6 and TGF-β was highly consistent with that of the SARS-CoV-2 N protein in EVs. Meanwhile, we found that most integrins and other adhesion molecules were also upregulated (**Fig. S1**), which could jointly influence the interaction of immune cells with the local microenvironment [20, 21]. All these indicate that EVs are engaged in the immune response to COVID-19 infection. Secretory immunoglobulin A (IgA) play an important role in the protection and homeostatic regulation of the respiratory mucosal epithelium, which is referred to as "immune exclusion" [15]. However, there was no significant difference in the total expression of IgA in sputum EVs before and after the infection with SARS-CoV-2 (**Figure 2E**). It was considered that there should be differences in IgA in some EV subpopulations, but this could not be reflected in the analysis of the total protein.

Toobtain the protein expression of a single EV, the algorithm FlowSOM was applied to analyze the behavior of all markers on all individual EVs, and generate the clusters of EVs using a self-organizing map. The clustering of individual EVs obtained from all samples was displayed in the tSNE plot, in which 40 clusters were color labelled (**Figure 3B**). After quantifying the proportion of each subpopulation, we found that there were significant differences in 18 clusters (**Fig. S3**).

The EVs in cluster 2, which constituted 4.92% of all EVs in the nCOV group and only 0.55% of all EVs in the HC group (**Figure 5A and S3**), contained a large amount of protein

247 SARS-CoV-2 N. It was considered that these are EVs that directly transport SARS-CoV-2. In
248 cluster 2, epithelial cell adhesion molecule (EpCAM), CDH1, ITGB4, ITGA5, SNAI2, CD81,
249 ITGB2, ZEB1, CD151, mucosal addressin cell adhesion molecule-1 (MAdCAM-1), and other
250 adhesion molecules, were highly expressed. A previous study has shown that CDH1 is required
251 for HCV (Hepatitis C virus) infection, while CDH1 silencing significantly inhibits the HCV
252 infection in primary human hepatocytes at the post-binding entry step [22]. Furthermore,
253 ITGB2, ITGB3 and CD151 were involved in the process of vesicle internalization and
254 recycling to the cell membrane [23-25]. ITGB3 plays a central role in intracellular
255 communication through extracellular vesicles [26]. Meanwhile, CD151 plays a critical role in
256 influenza A virus signaling [27], and ITGB4 participates in cell recognition through CD81 [21].
257 The present results merely confirm that these were all highly expressed in cluster 2. Therefore,
258 we boldly speculate that the high expression of adhesion proteins, such as EpCAM and
259 CDH1(**Figure 2F and S2**), on EVs may make these more susceptible to SARS-CoV-2
260 infection. These adhesion factors (ITGB4, ITGB2, ITGB3 and CD151) improve the
261 recognition function of EVs, and eventually, these EVs were more likely to carry virus
262 particles, and absorbed by the recipient cells. Once these EVs are ingested by epithelial cells,
263 the expression of CDH1 in cells would be reduced due to the virus-induced epithelial to
264 mesenchymal transition (EMT) [28]. Hence, it was found that ZEB1 and Snai2 were both
265 highly expressed. In addition, the abnormal expression of integrins plays an important role in
266 fibrosis formation [30]. The histological examination of biopsy samples obtained from
267 COVID-19 patients revealed the bilateral diffuse alveolar damage with cellular fibromyxoid
268 exudates [31], and the processes mentioned above may be partly responsible.
269 Except for cluster 2, cluster 3, 4, 6, 7, 9, 10,12, 13, 14, 32 and 34 were the EVs that
270 increased in the nCOV group, with highly express protein CLEC2A, CD81, ITGA1, ITGA5,
271 ITGB6, ITGB3, SNAI1, CD151, GPC1, RETN and ITGB2, respectively (**Figure 5C**). These

are all correlated to the protein expression of cluster 2. Adhesion molecules are involved in various important physiological functions and pathological processes, including leukocyte adhesion to vascular endothelial cells, and lymphocyte homing during the process of inflammation [20]. This process is controlled through the modulation of integrin binding to endothelial and mucosal ligands (e.g. integrin α4β7 and MAdCAM-1) [32]. After the SARS-CoV-2 infection, there may be some new types of multistep adhesion cascade in EVs, leading to inflammation. Indeed, all this needs to be further verified by more follow-up experiments.

Back to the EVs themselves, it is known that EVs are regulated by surface markers, such as CD9, CD63, CD81 and ALIX [33], while different factors regulate different EV functions [34, 35]. We found that the EVs regulated by CD81 were more likely to bind to the SARS-CoV-2 N protein (**Figure 6C**). In addition, it is known that cluster 2 (**Figure 5B**), cluster 4 (**Figure 5C**), cluster 6, cluster 7, cluster 12 and cluster 34 (**Fig. S3**) are EVs with an upregulated protein expression after SARS-CoV-2 infection, and that CD81 is highly expressed in all these clusters. These results suggest that the EVs regulated by CD81 are the most likely subpopulations of EVs that cause the changes in the pulmonary microenvironment after SARS-CoV-2 infection.

Furthermore, it was found that hepatitis C virus (HCV), which have been extensively studied, enters the host cell through interactions with a cascade of cellular factors，such as CDH1, claudin-1(CLDN1), and occludin (OCLN) [22]. Subsequently, it can be be recognized and absorbed by recipient cells. Unexpectedly, these transmission processes may be similar to those of SARS-CoV-2, the EVs regulated by CD81 also highly expressed CDH1. The difference is that EGFR is not required for the transmission of the SARS-CoV-2 (**Figure 6A**). HCV uses a dynamic and multi-step process to engage and enter host cells, in which EGFR is necessary for internalization [36].

In conclusion, we found that EVs (mostly regulated by CD81) can carry the SARS-CoV-2 N protein, and its expression is highly correlated with that of inflammatory factors in EVs. These results demonstrate that EVs derivedthe from sputum of patients may participate in the infection and immune response of COVID-19. The mechanism of the HCV infection, and subsequently internalizing these into recipient cells, might have some similarities to the relationship between EVs and SARS-CoV-2 infection, giving us many hints. This can provide some information for the further study of COVID-19 and promote our understanding of the possible pathogenesis of SARS-CoV-2 infection and the possibility of nanoparticles drug intervention in viral infection.

Materials and methods

Patient and healthy donor selection and inclusion criteria

The present study was approved by the Ethics Committee of the First Affiliated Hospital of Guangzhou Medical University (Guangzhou, China). Written informed consent was obtained from all study participants. A sample ID was applied to ensure sample tracking with confidentiality on sample donor identity. The healthy control (HC) group included 20 healthy donors without symptoms, such as cough, allergy, respiratory tract discomfort, and so on. The nCOV group included 20 patients with RT-PCR confirmed infection of SARS-CoV-2. The clinic-pathological conditions of patients included in the present study are shown in Figure 1A. The HC group consisted of 15 males and 5 females, with an average age of 56.6 years old. The nCOV group consisted of 14 males and 6 females, with an average age of 56.7 years old.

Detection of SARS-CoV-2 for the diagnosis of COVID-19:

The presence of the SARS-CoV-2 was detected by real-time RT-PCR methods [11]. Nucleic acid was extracted from respiratory samples and sputum using a Viral RNA extraction kit obtained from Daan Gene Co., Ltd. (Guangzhou, China). The RNA extraction from sputum and blood was performed using a total RNA extraction kit obtained from Sangon Biotech (Shanghai, China). The real-time PCR assay kit for targeting the SARS-CoV-2 RdRp and N gene regions was provided by Daan Gene Co., Ltd.

Sputum sample collection and pretreatment

The sputum of patients was collected during a pulmonary exacerbation, and in a stable condition. Similarly, the sputum of normal people (HC group) was induced by the inhalation of hypertonic (NaCl 5%) or isotonic (NaCl 0.9%) saline. The sputum samples were observed under a microscope to ensure the qualified samples. The standard is that the ratio of white blood cells to squamous cells is greater than 2.5. The sputum samples were dispersed with PBS at a ratio of 1:3, and centrifuged at 500 g to remove the cells, cell debris, and aggregates. All operations were performed in biosafety laboratories, and the samples were aliquoted and stored at -80°C until analysis.

Virus Isolation and Transmission Electron Microscopy

Vero E6 cells were used for virus isolation. A quantitative reverse transcription PCR (qRT-PCR)–positive sputum swab specimen was saved in viral transport media (DMEM containing 1% bovine serum albumin, 15 μg/mL amphotericin, 100 units/mL penicillin G, and 100 μg/mL streptomycin). Before virus isolation, the sample was filtered with a 0.45-μm strainer and diluted 1:10 with DMEM containing 2% FBS and antimicrobial drugs. Cells were infected at 37°C for 1 h. The inoculum was removed and replaced with a fresh culture medium.

Cytopathic effect (CPE) was observed in Vero E6 cells infected with SARS-CoV-2 isolate after 72 h but not in mock-infected cells. Culture supernatant was negatively stained and visualized by transmission electron microscopy.

Characterization of extracellular vesicles in the sputum sample

After collecting the induced sputum, the EVs were isolated from the sputum supernatant using a standard ultracentrifugation protocol after initial extraction using the EV extraction kit (ExoQuick-TC, EXOTC50A-1, SBI, USA). The concentration and size distribution of particles in the sputum were investigated via a nanoparticle tracking system (Nanosight NS300, Malvern Panalytical Ltd., UK) [12]. The morphology of the EVs were recorded via transmission electron microscopy (H-7650, Hitachi, Japan).

EV capture, fixation and permeabilization

The streptavidin-coated PCR plates (PCR0STF-SA5/100, Biomat, Italy) were incubated with 2.5µg/ml of biotinylated cholera toxin subunit B (C34779, Thermo-Fisher Scientific, USA) in PBS (C10010500BT, Gibco, USA) at room temperature for two hours. Then, the plate wells were rinsed for three times with PBST washing buffer, 0.05% Tween-20 (003005, Thermo-Fisher Scientific, USA) in PBS. Afterwards, 20µl of sputum/PBS samples were added to the wells of the plate. Wells were rinsed with PBST after incubation at room temperature for two hours. The fixation step was performed by adding 20µl 4% paraformaldehyde in PBS (BL539A, Biosharp, China) into each well. Thereafter, 0.2% Triton-X (T8787, Sigma-Aldrich, USA) in PBS were added for permeabilization to facilitate the detection of the inner proteins of EVs. Then, the wells of the plate were rinsed for three times with PBST before further tests.

EV proteomics analysis via Proximity Barcoding Assay (PBA)

The experimental method for the proximity barcoding assay (PBA) was performed adopting the previously published protocol [13]. For the EV proteomic analysis, 100 antibodies were conjugated with DNA oligonucleotides comprising 8-nucleotide (nt) protein tag, 8-nucleotide (nt) molecule tag and universal sequences as adapters. The proteins under investigation included typical EV biomarkers, biomarkers related in lung diseases, and a panel of cell adhesion molecules. The PBA tests were designed according to the protocols in Vesicode AB (Solna, Sweden), and performed in Secretech (Shenzhen, China).

Data processing

After the DNA sequencing, the raw data was obtained in bcl file format. After running the bcl2fastq program (Illumina, USA), the fastq files of each sample were generated according to the sample indexes. Using fastx_toolkit, low-quality reads (Phred quality score Q<20) were removed before further analysis. The clean data files for each sample constituted the DNA reads of 75 bp, and the EV tag, protein tag, and molecule tag were extracted. The molecule tags were used to deduplicate the amplified sequences due to the PCR reaction for library construction, and the unique reads were used in the subsequent assays. The protein tags were translated to the protein name by matching the antibody-DNA tag conjugation list (**SI. Table 1**). The EV-protein matrix contained columns for protein expression and rows of single EVs for each sample, as indicated by the detected EV tags (**SI. Table 2**).

The EV-associated protein expression levels were obtained by summating the quantity of a certain protein detected on all EVs. The data was normalized using the count per million (CPM) method, accounting for the library size, and the trimmed mean (TMM) method, accounting for composition bias. The protein combinations are the information of the protein co-expression on the same EV. An unsupervised machine learning algorithm, FlowSOM, was applied to cluster

EVs, according to the proteomic features of EVs. The number of clusters was determined according to the consensus matrix, in which the lowest number of clusters for optimal separability was selected. The proteomic similarity of EVs was observed in the T-distributed Stochastic Neighbor Embedding (t-SNE) plot. The proportion of each subpopulation was quantified. The proteomic fingerprints for each subpopulation were profiled.

The protein combinations were summarized in the format of the EV tag-(p1, p2, p3…). The quantity of each possible pair of co-expressed protein was obtained as the protein combination dataset, and used as input variables for the abundance and differential analysis. The differential analysis between the nCOV and HC groups was performed and visualized. The co-expression of SARS-CoV-2 protein SARS-CoV-2 N with EV proteins was selected from the protein combination dataset and analyzed to predict virus-EV association.

Quantification and Statistical Analysis

ANOVA and Student's t-test were performed to analyze the differences in mean values between groups using GraphPad Prism 7. All results were expressed as mean ± standard error of the mean (SEM), and were corrected for multiple comparisons. P-values <0.05 were considered statistically significant. The significance was indicated by asterisks and * stands for a P-values ≤0.05, ** for P-values ≤0.005, *** for P-values ≤0.001 and **** for P-values ≤0.0001.

Declarations

Acknowledgements and Role of the funding source

This work is supported by the emergency grants for prevention and control of SARS-CoV-2 of Ministry of Guangdong province (2020B111133001), the China Postdoctoral Science Project (2020T130025ZX and 2019M652860), the independent project of the State Key

Laboratory of Respiratory Diseases (SKLRD-QN-201913), National Natural Science Foundation of China (NSFC, 82000044), the Local Innovative and Research Teams Project of Guangdong Pearl River Talents Program (2017BT01S155), and Zhongnanshan Medical Foundation of Guangdong Province (ZNSA-2020013).

Conflict of interest and Consent for publication

The authors have declared that no conflict of interest exists. This work described has not been submitted elsewhere for publication, in whole or in part, and all the authors listed have approved the manuscript that is enclosed.

Ethics approval

All procedures followed were in accordance with the ethical standards of the Medical Ethical Council of the First Affiliated Hospital of Guangzhou Medical University and with the Helsinki Declaration of 1975, as revised in 2000. Informed consent was obtained from all patients for being included in the study. This article does not contain any studies with animal subjects performed by any of the authors.

Author contributions

PR, RS, DW, and JZ conceived the study; RS, YZ, GB, JS, PK, LY and AZ collected clinical specimen and executed the experiments; RS, YC, YZ, GB, PK, YL, WL, JL, NC, JX and DW analyzed the data; JZ, BL and YZ contributed to critical revision of the manuscript; PR, RS, YC and DW wrote the manuscript. All authors revised and approved the final version.

Availability of Data and Material

All the high-throughput sequencing data and experimental materials generated in this study are available from the corresponding authors upon reasonable request.

Figure legends

Figure 1. Demographic information and Characterization of EVs from the sputum samples of COVID-19 patients. **(A)** Donor information and the clinical complications of COVID-19 infected patients. **(B-C)** The morphology of SARS-CoV-2 and EVs recorded by Transmission Electron Microscopy. **(D)** The Nanoparticle Tracking Analysis of particles in the sputum samples of COVID-19 patients (50μl of EVs were extracted from 100μl of sputum, and detected after 200-fold dilution). In the inset, the size distribution (black lines) of EVs is depicted (no magnet was employed). The error bars (in red) indicate the standard error of the mean.

Figure 2. Patients with COVID-19 secreted more proteins in individual EV and EVs participated in the immune response. (**A**) The Proximity Barcoding Assay (PBA) for the analysis of the protein profile on a single EV level. A lipid membrane binding layer captured the EVs of each sample; that is, streptavidin-biotin-CTB, coated on the wall of a 0.2ml well. After fixation with formaldehyde and permeabilization with Triton X, the 100-antibody panel of the barcoded antibodies (SI. Table 1) were applied to interact with the proteins on individual EVs. The PBA templates endow each EV with a specific 16-bp EV tag through extension reaction. The DNA sequence results were prepared into the sequencing library, and read out through high throughput sequencing using Illumina NextSeq 500. (**B**) The quantification of EVs and proteins detected in the PBA, and the ratio of proteins per EV from COVID-19 patients (nCOV group) and healthy controls (HC group) were plotted. (**C**) SARS-CoV-2 N protein signals can be detected in EVs obtained from nCOV, HC and PBS negative controls. **(D)** The quantification of IL6 and TGF-β proteins, and their correlation with SARS-CoV-2 N protein, detected from the nCOV and HC groups. COVID-19 patients (nCOV, n = 20), healthy controls (HC, n = 20), and PBS negative controls (PBS, n=4). **(E)**

The quantification of IgA and IgM proteins. ****Indicates P<0.0001 in the t-test. **(F)** Total expression of proteins in sputum EV samples of nCOV and HC groups, the heatmap shows the proteomic profile of samples.

Figure 3. The EV subpopulations from 40 different samples were determined using unsupervised machine learning process, FlowSOM. **(A)** Forty subpopulations were displayed in the t-distributed stochastic neighbor embedding (tSNE) plot. The tSNE representation of EVs from all analyzed samples (n = 40) were colored by manually annotating the EV type. **(B)** The tSNE plot for each sample obtained from healthy individuals (HC, n = 20) and COVID-19 infected patients (nCoV, n= 20), and was conducted for each cluster.

Figure 4. EV subpopulations atlas and the featured change in patients with COVID-19. **(A)**The similarity and differences in EV proteomics between the sputum samples in the COVID-19 and HC groups are shown, in which all EVs in the nCoV group were colored green, while the EVs in the HC group were colored red. **(B)** The quantification of the EV subpopulation for the nCOV and HC groups. Compared to the HC group, the nCOV group had a significantly high ratio of EV subpopulations, which was indicated as cluster 2, 3, 4, 6, 7, 9, 10, 12, 13, 14, 32 and 34, while there was a decreased ratio of subpopulations in cluster 16, 21, 22, 26, 27 and 33. (*indicates P<0.05, **indicates P<0.01, ***indicates P<0.001, and ****indicates P<0.0001 in the t-test).

Figure 5. The differential distribution of EVs obtained from the nCOV and HC groups, the EV cluster distribution for the nCOV (red) and HC (blue) groups. **(B)** For each differentially expressed EV subpopulation, the location in the total EVs and the top 7 featured proteins are shown. **(A-B)** cluster 2 constitutes 4.92% of all EVs in the nCOV group, but this

was only 0.55% in the HC group. The EVs of cluster 2 show the high expression of SARS-CoV-2 N protein with EV biomarker CD81 and cell adhesion molecules, including EpCAM, CDH1, ITGB4, ITGA5, SNAI2 and so on. **(C-E)** The proteomic profiles of cluster 3, 4,13 and 33. Cluster 33 featured the higher expression of EGFR and IgA, comprising 14.12% of EVs in the HC group and 3.37% of EVs in the nCOV group.

Figure 6. EVs regulated by CD81 are more likely to carry SARS-CoV-2. **(A)** The protein combinations were demonstrated in the matrix. The fold change in expression was color coded, and the significance was indicated with the size of the dots. **(B-C)** The co-localization of the SARS-CoV-2 N protein with EV biomarkers. **(B)** The Venn diagram analysis for multiple groups of data. **(C)** The levels of enriched co-localization. The number of associated query proteins in a term are shown on the right side of each term bar. **(D)** The proteomic profiles of cluster 4. **(E)** The protein expression distribution of CD81(red) and SARS-CoV-2 N (blue) in all EV, and the green part represented the co-expression region.

References

1. Abbott, T.R., et al., *Development of CRISPR as an Antiviral Strategy to Combat SARS-CoV-2 and Influenza.* Cell, 2020. **181**(4): p. 865-876 e12.
2. Sun, J., et al., *Generation of a Broadly Useful Model for COVID-19 Pathogenesis, Vaccination, and Treatment.* Cell, 2020. **182**(3): p. 734-743 e5.
3. Wu, F., et al., *Clinical characteristics of COVID-19 infection in chronic obstructive pulmonary disease: a multicenter, retrospective, observational study.* J Thorac Dis, 2020. **12**(5): p. 1811-1823.
4. Sanchez-Cerrillo, I., et al., *COVID-19 severity associates with pulmonary redistribution of CD1c+ DC and inflammatory transitional and nonclassical monocytes.* J Clin Invest, 2020.
5. Huang, C., et al., *Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China.* Lancet, 2020. **395**(10223): p. 497-506.
6. Sung, P.S., T.F. Huang, and S.L. Hsieh, *Extracellular vesicles from CLEC2-activated platelets enhance dengue virus-induced lethality via CLEC5A/TLR2.* Nat Commun, 2019. **10**(1): p. 2402.

7. Delpech, J.C., et al., *Neuroimmune Crosstalk through Extracellular Vesicles in Health and Disease.* Trends in Neurosciences, 2019. **42**(5): p. 361-372.
8. Feng, Z., et al., *A pathogenic picornavirus acquires an envelope by hijacking cellular membranes.* Nature, 2013. **496**(7445): p. 367-71.
9. Jiang, W., et al., *Hepatitis A virus structural protein pX interacts with ALIX and promotes the secretion of virions and foreign proteins through exosome-like vesicles.* J Extracell Vesicles, 2020. **9**(1): p. 1716513.
10. Grunvogel, O., et al., *Secretion of Hepatitis C Virus Replication Intermediates Reduces Activation of Toll-Like Receptor 3 in Hepatocytes.* Gastroenterology, 2018. **154**(8): p. 2237-2251 e16.
11. Wang, Y., et al., *Kinetics of viral load and antibody response in relation to COVID-19 severity.* J Clin Invest, 2020. **130**(10): p. 5235-5244.
12. Sun, R., et al., *ALIX increases protein content and protective function of iPSC-derived exosomes.* J Mol Med (Berl), 2019. **97**(6): p. 829-844.
13. Wu, D., et al., *Profiling surface proteins on individual exosomes using a proximity barcoding assay.* Nat Commun, 2019. **10**(1): p. 3854.
14. Arunachalam, P.S., et al., *Systems biological assessment of immunity to mild versus severe COVID-19 infection in humans.* Science, 2020. **369**(6508): p. 1210-1220.
15. Yu, H.Q., et al., *Distinct features of SARS-CoV-2-specific IgA response in COVID-19 patients.* Eur Respir J, 2020. **56**(2).
16. Ferretti, A.P., et al., *Unbiased Screens Show CD8(+) T Cells of COVID-19 Patients Recognize Shared Epitopes in SARS-CoV-2 that Largely Reside outside the Spike Protein.* Immunity, 2020.
17. Le Bert, N., et al., *SARS-CoV-2-specific T cell immunity in cases of COVID-19 and SARS, and uninfected controls.* Nature, 2020. **584**(7821): p. 457-462.
18. Wilk, A.J., et al., *A single-cell atlas of the peripheral immune response in patients with severe COVID-19.* Nat Med, 2020. **26**(7): p. 1070-1076.
19. Agrati, C., et al., *Expansion of myeloid-derived suppressor cells in patients with severe coronavirus disease (COVID-19).* Cell Death Differ, 2020. **27**(11): p. 3196-3207.
20. Weninger, W., M. Biro, and R. Jain, *Leukocyte migration in the interstitial space of non-lymphoid organs.* Nat Rev Immunol, 2014. **14**(4): p. 232-46.
21. Sun, H., et al., *Distinct chemokine signaling regulates integrin ligand specificity to dictate tissue-specific lymphocyte homing.* Dev Cell, 2014. **30**(1): p. 61-70.
22. Li, Q., et al., *Hepatitis C virus depends on E-cadherin as an entry factor and regulates its expression in epithelial-to-mesenchymal transition.* Proc Natl Acad Sci U S A, 2016. **113**(27): p. 7620-5.
23. Yamada, M., et al., *The tetraspanin CD151 regulates cell morphology and intracellular signaling on laminin-511.* FEBS J, 2008. **275**(13): p. 3335-51.
24. Rosetti, F., et al., *A Lupus-Associated Mac-1 Variant Has Defects in Integrin Allostery and Interaction with Ligands under Force.* Cell Rep, 2015. **10**(10): p. 1655-1664.

25. Kovacheva, M., et al., *Conditional knockdown of integrin beta-3 reveals its involvement in osteolytic and soft tissue lesions of breast cancer skeletal metastasis.* J Cancer Res Clin Oncol, 2020.
26. Fuentes, P., et al., *ITGB3-mediated uptake of small extracellular vesicles facilitates intercellular communication in breast cancer cells.* Nat Commun, 2020. **11**(1): p. 4261.
27. Qiao, Y., et al., *CD151, a novel host factor of nuclear export signaling in influenza virus infection.* J Allergy Clin Immunol, 2018. **141**(5): p. 1799-1817.
28. Harris, T.J. and U. Tepass, *Adherens junctions: from molecules to morphogenesis.* Nat Rev Mol Cell Biol, 2010. **11**(7): p. 502-14.
29. Liao, H.D., Y. Mao, and Y.G. Ying, *The involvement of the laminin-integrin alpha7beta1 signaling pathway in mechanical ventilation-induced pulmonary fibrosis.* J Thorac Dis, 2017. **9**(10): p. 3961-3972.
30. Reed, N.I., et al., *The alphavbeta1 integrin plays a critical in vivo role in tissue fibrosis.* Sci Transl Med, 2015. **7**(288): p. 288ra79.
31. Xu, Z., et al., *Pathological findings of COVID-19 associated with acute respiratory distress syndrome.* Lancet Respir Med, 2020. **8**(4): p. 420-422.
32. Sun, H., et al., *beta7 Integrin Inhibition Can Increase Intestinal Inflammation by Impairing Homing of CD25(hi)FoxP3(+) Regulatory T Cells.* Cell Mol Gastroenterol Hepatol, 2020. **9**(3): p. 369-385.
33. Roucourt, B., et al., *Heparanase activates the syndecan-syntenin-ALIX exosome pathway.* Cell Res, 2015. **25**(4): p. 412-28.
34. Ghossoub, R., et al., *Syntenin-ALIX exosome biogenesis and budding into multivesicular bodies are controlled by ARF6 and PLD2.* Nature Communications, 2014. **5**(3): p. 3477.
35. van Niel, G., et al., *The tetraspanin CD63 regulates ESCRT-independent and -dependent endosomal sorting during melanogenesis.* Dev Cell, 2011. **21**(4): p. 708-21.
36. Baktash, Y., et al., *Single Particle Imaging of Polarized Hepatoma Organoids upon Hepatitis C Virus Infection Reveals an Ordered and Sequential Entry Process.* Cell Host Microbe, 2018. **23**(3): p. 382-394 e5.

Figures

Figure 1

A

Group HC: Healthy control (n=20, average age=56.6)			Group nCOV: COVID-19 patients (n=20, average age=56.7)			
Patient number	Gender[1]	Age	Patient number	Gender[1]	Age	Clinical complications[2]
HC1	M	67	nCOV1	F	25	N
HC2	M	56	nCOV2	F	65	N
HC3	F	48	nCOV3	M	64	A(severe), B(I type), C, D, E
HC4	M	54	nCOV4	M	49	A(severe), D, E, F
HC5	M	58	nCOV5	M	41	A(severe), B, F
HC6	M	61	nCOV6	F	72	A(severe), B, F
HC7	M	68	nCOV7	M	58	A(severe), D, E, G
HC8	F	54	nCOV8	F	82	N
HC9	M	40	nCOV9	M	49	C, D, G
HC10	M	52	nCOV10	M	53	A, B
HC11	M	53	nCOV11	M	61	A, B(II type), C, D, F
HC12	F	60	nCOV12	M	42	A(severe), D
HC13	F	56	nCOV13	M	72	A(severe), D, E, G, H
HC14	M	65	nCOV14	M	26	I
HC15	M	59	nCOV15	M	79	A(severe), D, E, F, G, H
HC16	M	50	nCOV16	M	50	N
HC17	F	41	nCOV17	F	55	A(severe), B, F
HC18	M	64	nCOV18	M	66	A(severe), B, J
HC19	M	69	nCOV19	M	68	A(severe), B, F, G
HC20	M	57	nCOV20	F	57	N

[1]Gender:
M, male
F, female

[2]Clinical complications:
A, Acute respiratory distress syndrome (severe)
B, Respiratory failure
C, Septic shock
D, Myocardial damage
E, Abnormal coagulation function
F, Sepsis
G, Acute kidney injury
H, Multiple organ dysfunction syndrome
I, Hepatic insufficiency
J, Hypoproteinemia
N, None

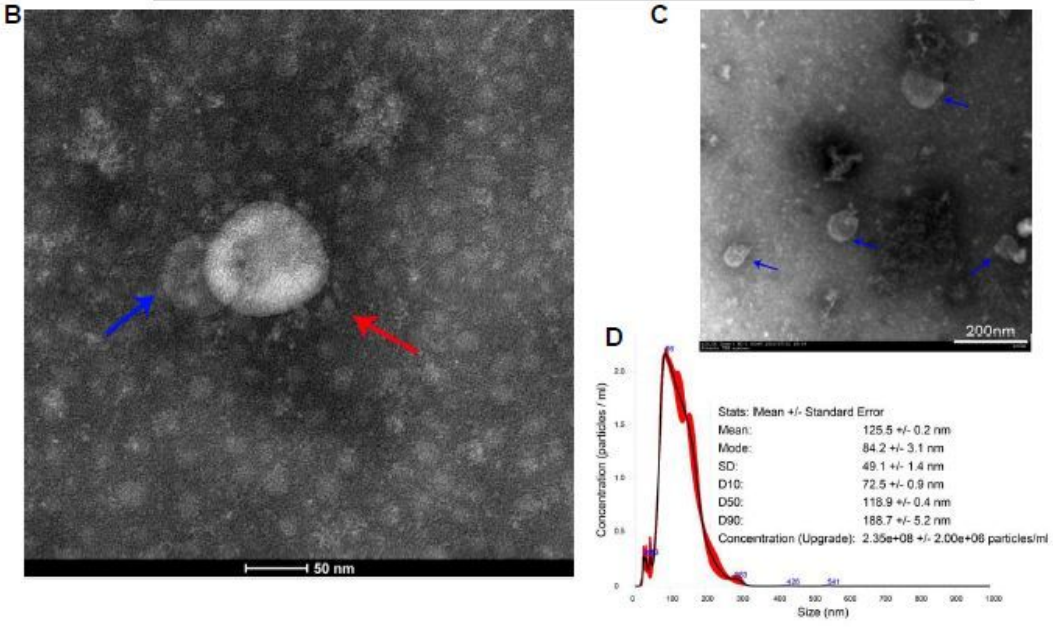

Figure 1

Demographic information and Characterization of EVs from the sputum samples of COVID-19 patients. (A) Donor information and the clinical complications of COVID-19 infected patients. (B-C) The morphology of SARS-CoV-2 and EVs recorded by Transmission Electron Microscopy. (D) The

Nanoparticle Tracking Analysis of particles in the sputum samples of COVID-19 patients (50μl of EVs were extracted from 100μl of sputum, and detected after 200-fold dilution). In the inset, the size distribution (black lines) of EVs is depicted (no magnet was employed). The error bars (in red) indicate the standard error of the mean.

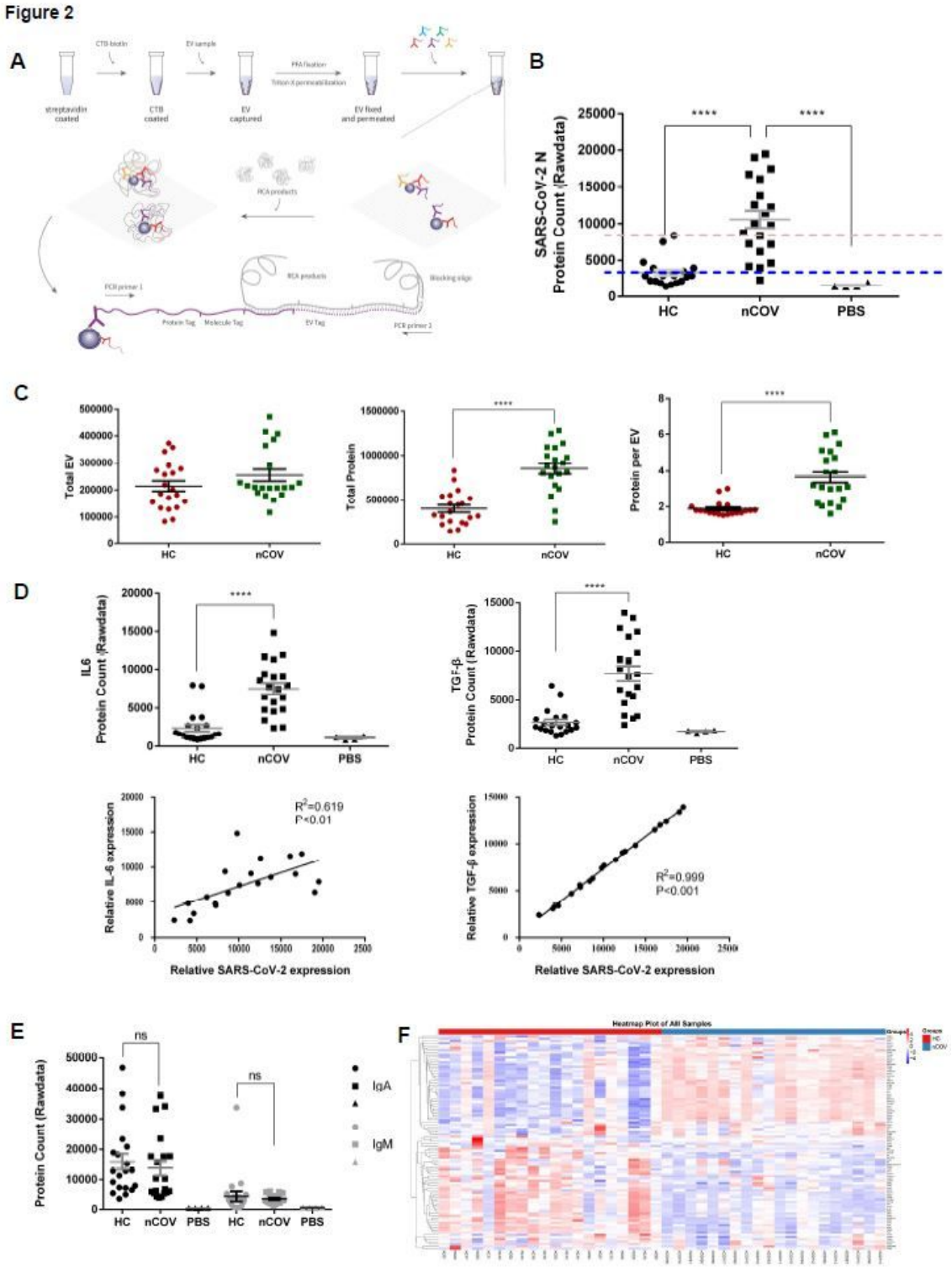

Figure 2

Patients with COVID-19 secreted more proteins in individual EV and EVs participated in the immune response. (A) The Proximity Barcoding Assay (PBA) for the analysis of the protein profile on a single EV level. A lipid membrane binding layer captured the EVs of each sample; that is, streptavidin-biotin-CTB, coated on the wall of a 0.2ml well. After fixation with formaldehyde and permeabilization with Triton X, the 100-antibody panel of the barcoded antibodies (SI. Table 1) were applied to interact with the proteins on individual EVs. The PBA templates endow each EV with a specific 16-bp EV tag through extension reaction. The DNA sequence results were prepared into the sequencing library, and read out through high throughput sequencing using Illumina NextSeq 500. (B) The quantification of EVs and proteins detected in the PBA, and the ratio of proteins per EV from COVID-19 patients (nCOV group) and healthy controls (HC group) were plotted. (C) SARS-CoV-2 N protein signals can be detected in EVs obtained from nCOV, HC and PBS negative controls. (D) The quantification of IL6 and TGF-β proteins and their correlation with SARS-CoV-2 N protein, detected from the nCOV and HC groups. COVID-19 patients (nCOV, n = 20), healthy controls (HC, n = 20), and PBS negative controls (PBS, n=4). (E) The quantification of IgA and IgM proteins. ****Indicates $P<0.0001$ in the t-test. (F) Total expression of proteins in sputum EV samples of nCOV and HC groups, the heatmap shows the proteomic profile of samples.

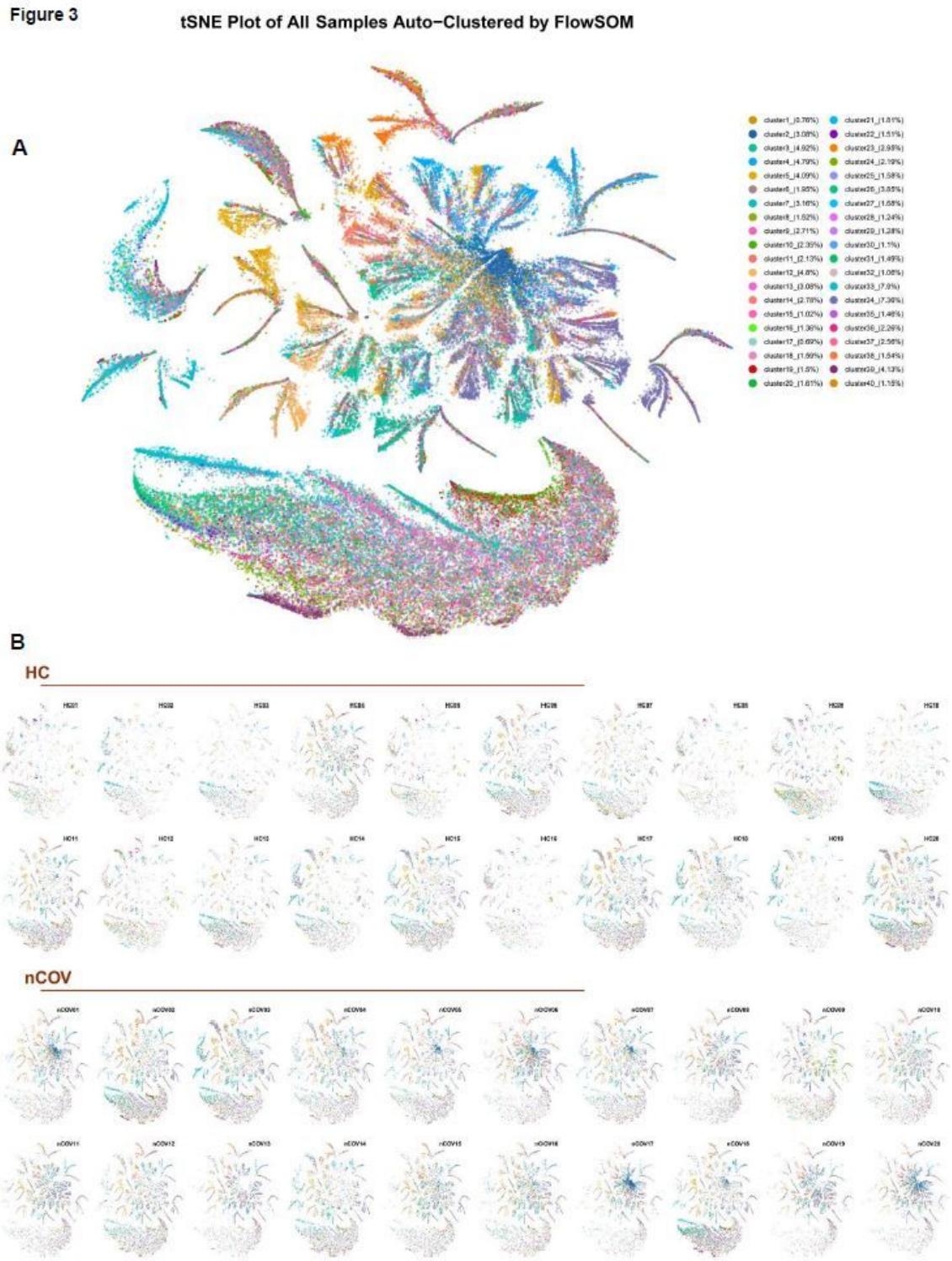

Figure 3

The EV subpopulations from 40 different samples were determined using unsupervised machine learning process, FlowSOM. (A) Forty subpopulations were displayed in the t-distributed stochastic neighbor embedding (tSNE) plot. The tSNE representation of EVs from all analyzed samples (n = 40) were colored by manually annotating the EV type. (B) The tSNE plot for each sample obtained from healthy individuals (HC, n = 20) and COVID-19 infected patients (nCoV, n= 20), and was conducted for each cluster.

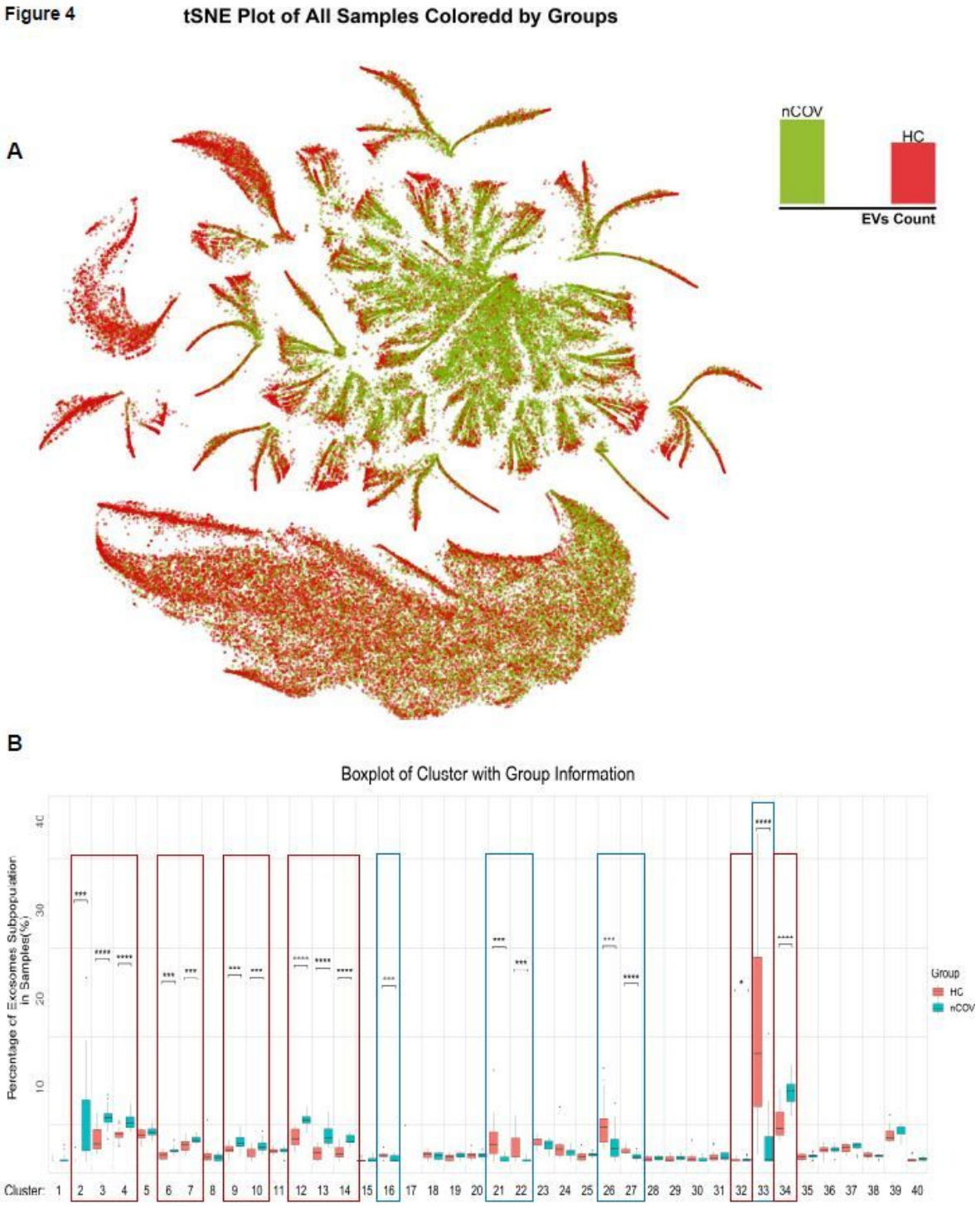

Figure 4

EV subpopulations atlas and the featured change in patients with COVID-19. (A)The similarity and differences in EV proteomics between the sputum samples in the COVID-19 and HC groups are shown, in which all EVs in the nCoV group were colored green, while the EVs in the HC group were colored red. (B) The quantification of the EV subpopulation for the nCOV and HC groups. Compared to the HC group, the nCOV group had a significantly high ratio of EV subpopulations, which was indicated as cluster 2, 3, 4, 6,

7, 9, 10, 12, 13, 14, 32 and 34, while there was a decreased ratio of subpopulations in cluster 16, 21, 22, 26, 27 and 33. (*indicates P<0.05, **indicates P<0.01, ***indicates P<0.001, and ****indicates P<0.0001 in the t-test).

Figure 5

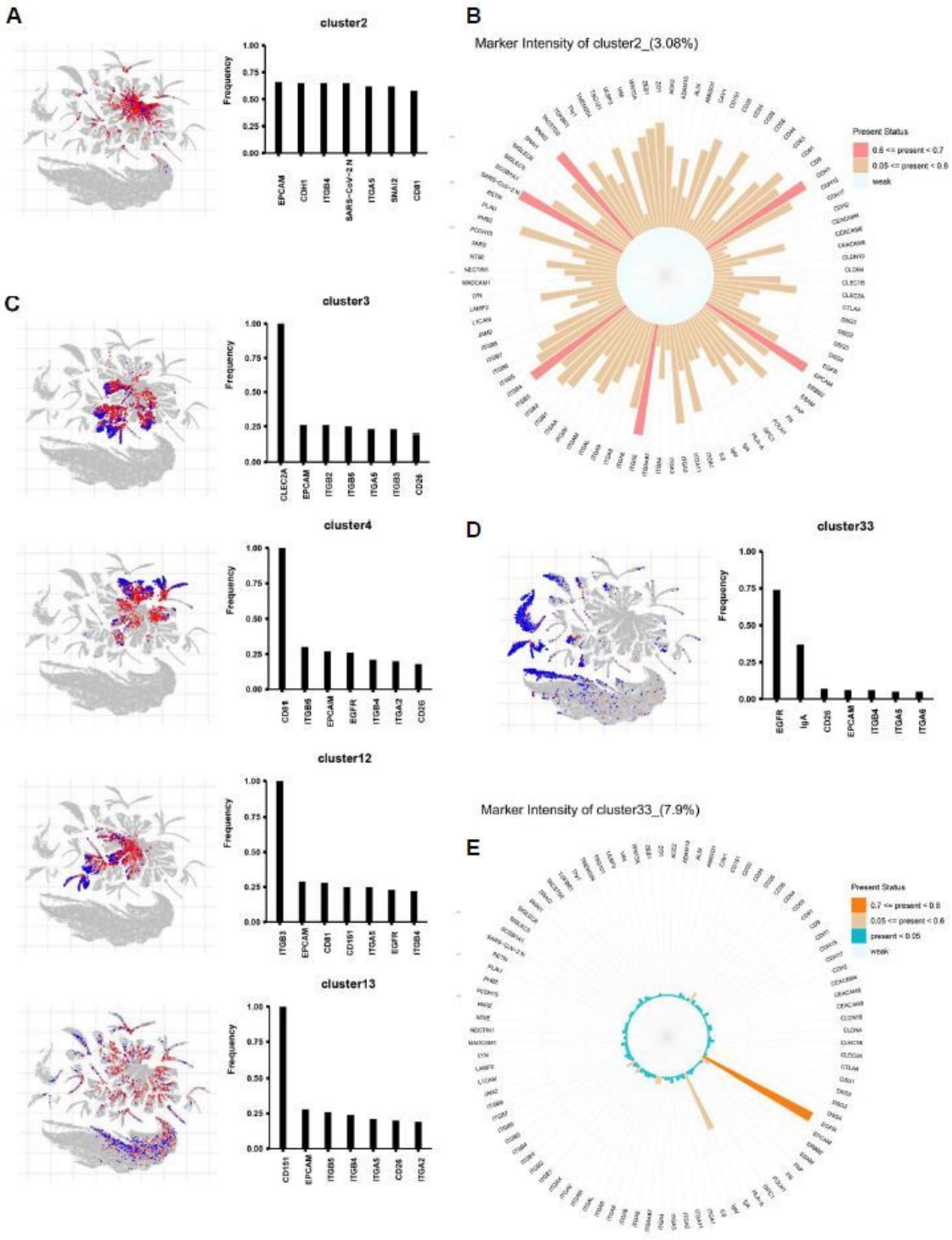

Figure 5

The differential distribution of EVs obtained from the nCOV and HC groups, the EV cluster distribution for the nCOV (red) and HC (blue) groups. (B) For each differentially expressed EV subpopulation, the location in the total EVs and the top 7 featured proteins are shown. (A-B) cluster 2 constitutes 4.92% of all EVs in the nCOV group, but this was only 0.55% in the HC group. The EVs of cluster 2 show the high expression of SARS-CoV-2 N protein with EV biomarker CD81 and cell adhesion molecules, including EpCAM, CDH1, ITGB4, ITGA5, SNAI2 and so on. (C-E) The proteomic profiles of cluster 3, 4,13 and 33. Cluster 33 featured the higher expression of EGFR and IgA, comprising 14.12% of EVs in the HC group and 3.37% of EVs in the nCOV group.

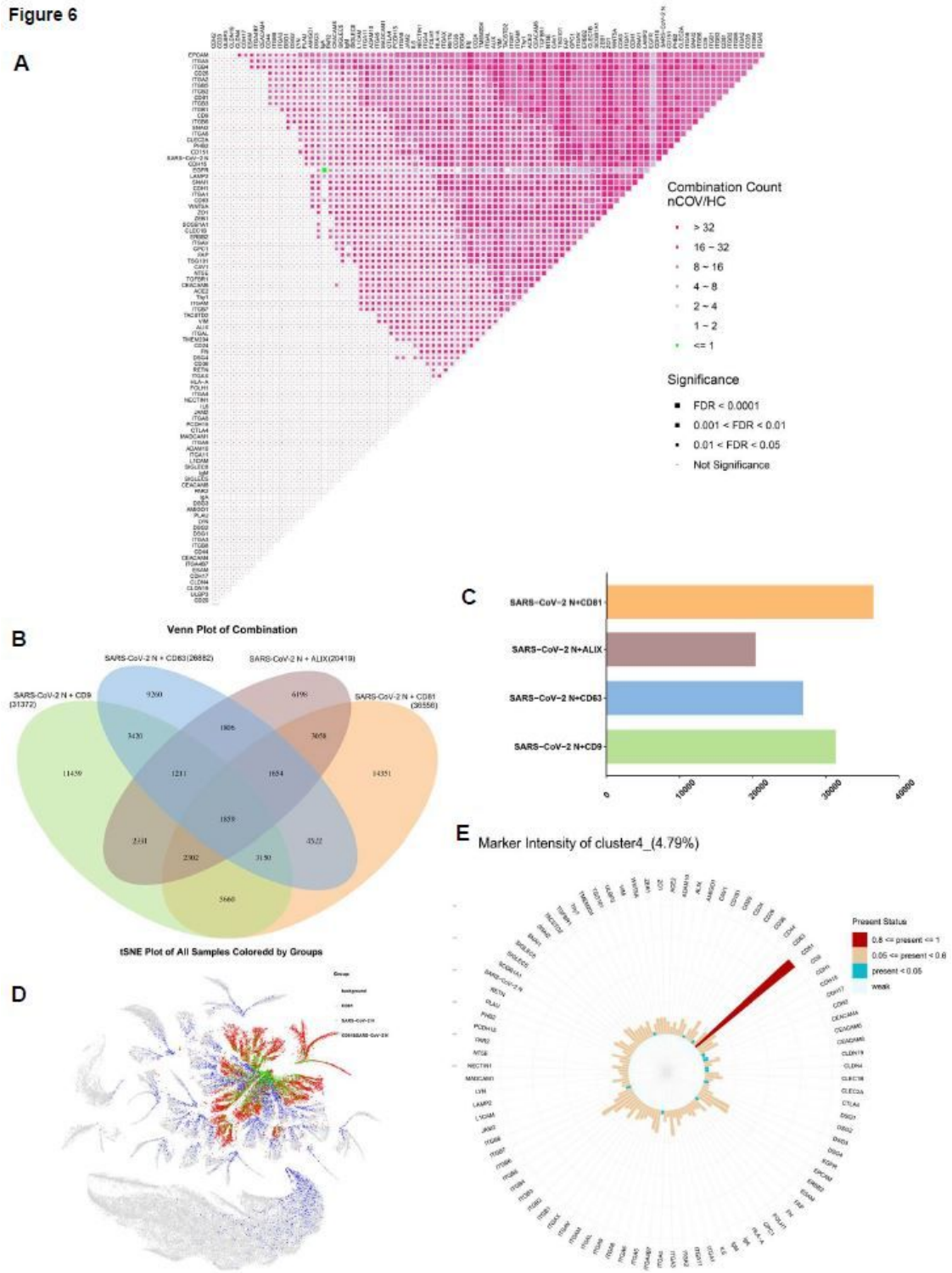

Figure 6

EVs regulated by CD81 are more likely to carry SARS-CoV-2. (A) The protein combinations were demonstrated in the matrix. The fold change in expression was color coded, and the significance was indicated with the size of the dots. (B-C) The co-localization of the SARS-CoV-2 N protein with EV biomarkers. (B) The Venn diagram analysis for multiple groups of data. (C) The levels of enriched co-localization. The number of associated query proteins in a term are shown on the right side of each term

bar. (D) The proteomic profiles of cluster 4. (E) The protein expression distribution of CD81(red) and SARS-CoV-2 N (blue) in all EV, and the green part represented the co-expression region.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Graphicalabstract.pdf
- Supplementaryinformation.pdf
- SFig2.jpg
- SFig3.jpg
- SFig1.jpg