The Nucleocapsid Protein of SARS-CoV-2 Abolished Pluripotency in Human Induced Pluripotent Stem Cells

Zebin Lin[1,2#], Zhiming Wu[3#], Jinlian Mai[2], Lishi Zhou[2], Yu Qian[1], Tian Cai[4], Zhenhua Chen[4], Ping Wang[5*] and Bin Lin[2*]

[1]School of Pharmaceutical Sciences, Sun Yat-Sen University, Guangzhou, China

[2]Guangdong Beating Origin Regenerative Medicine Co. Ltd., Foshan, China

[3]Department of Urology, Sun Yat-sen University Cancer Center, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Guangzhou, China

[4]Nanhai District People's Hospital of Foshan, Foshan, China

[5]School of Medical Imaging, Tianjin Medical University, Tianjin, China

[#]These authors contributed equally to the work.

*** Correspondence:**
Ping Wang

pingwang@tmu.edu.cn;

Bin Lin

linbin@beatingorigin.com.

Keywords: SARS-CoV-2, nucleocapsid protein, human induced pluripotent stem cell, pluripotency, fibroblast.

Abstract

The COVID-19 pandemic caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is raging across the world, leading to a global mortality rate of 3.4% (estimated by World Health Organization in March 2020). As a potential vaccine and therapeutic target, the nucleocapsid protein of SARS-CoV-2 (nCoVN) functions in packaging the viral genome and viral self-assembly. To investigate the biological effects of nCoVN to human stem cells, genetically engineered human induced pluripotent stem cells (iPSC) expressing nCoVN (iPSC-nCoVN) were generated by lentiviral expression systems, in which the expression of nCoVN could be induced by the doxycycline. The proliferation rate of iPSC-nCoVN was decreased. Unexpectedly, the morphology of iPSC started to change after nCoVN expression for 7 days. The pluripotency marker TRA-1-81 were not detectable in iPSC-nCoVN after a four-day induction. Meanwhile, iPSC-nCoVN lost the ability for differentiation into cardiomyocytes with a routine differentiation protocol. The RNA-seq data of iPSC-nCoVN (induction for 30 days) and immunofluorescence assays illustrated that iPSC-nCoVN were turning to fibroblast-like cells. Our data suggested that nCoVN disrupted the pluripotent properties of iPSC and turned them into other types of cells, which provided a new insight to the pathogenic mechanism of SARS-CoV-2.

1 Introduction

Right now, the COVID-19 pandemic is sweeping the world, causing a huge crisis in public health and economics globally. According to the continuously updated data from World Health Organization, to date, nearly three million infected cases were confirmed, while more than 200,000 individuals died because of COVID-19 (https://www.who.int/emergencies/diseases/novel-coronavirus-2019). Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which was proved to be the pathogen of COVID-19, has 79% identity in genomes with severe acute respiratory syndrome coronavirus (SARS-CoV) (Lu et al., 2020). Twelve coding regions were predicted in SARS-CoV-2, including spike protein, nucleocapsid protein, envelope protein, and membrane protein (Lu et al., 2020; Wu et al., 2020; Xu et al., 2020). The Cryo-EM structure of spike protein had been determined (Wrapp et al., 2020), and more and more evidences showed that the spike protein binds human ACE2 to entry into host cells (Hoffmann et al., 2020; Wrapp et al., 2020), which indicated that SARS-CoV-2 might share similar pathogenic mechanisms with SARS-CoV. Because of the very limited knowledge of SARS-CoV-2, we sought to understand the biology of SARS-CoV-2 based on the previous studies about SARS-CoV.

As one of the most studied proteins in SARS-CoV, the nucleocapsid protein binds to viral RNA to package the genome in a ribonucleoprotein particle (Chang et al., 2014). Unlike the spike protein with a certain mutation frequency, the sequence of nucleocapsid protein was more stable (Chinese, 2004), which meant it was an ideal target for diagnostic tools (Severance et al., 2008; Suresh et al., 2008; Das et al., 2010) and antiviral therapy (Cheung et al., 2008; Chang et al., 2016). The pathogenic effects in host cells caused by the nucleocapsid protein were also studied. It was reported that the nucleocapsid protein inhibited type I interferon production after virion infected the host cells (Hu et al., 2017), which was considered as a possible mechanism of immune escape. The nucleocapsid protein inhibited cell cytokinesis and proliferation (Zhou et al., 2008), and regulated several pathways, such as transforming growth factor-beta signaling (Zhao et al., 2008), AP-1 signal transduction pathway (He et al., 2003), and NF-KappaB pathway (Zhang et al., 2007b). Besides, the nucleocapsid protein was reported as an apoptosis inducer in COS-1 cells (Surjit et al., 2004; Zhang et al., 2007a) and HPF cells (Zhao et al., 2006).

As the nucleocapsid protein of SARS-CoV-2 (nCoVN) has 88.1% identity with the nucleocapsid protein of SARS-CoV (Lu et al., 2020), it is reasonable to speculate that they share a same pathogenic pathway in host cells. The original goal of this study is to determine the physiological malfunctions, such as cardiac fibrosis, in human cardiomyocytes expressing nCoVN by using human induced pluripotent stem cells (iPSC) and direct cardiac-differentiation methods. However, the morphology of iPSC altered obviously when nCoVN had been expressed for 7 days. This unexpected observation inspired us some new thoughts: (i) It was likely that the adult stem cells could be infected by SARS-CoV-2, in spite of lacking of the clinical data; (ii) iPSC were appropriate study materials for stem cell research because of avoiding many ethical issues; (iii) The preliminary data indicated that nCoVN seemed to be deleterious to iPSC, which meant it might also cause damages in other stem cells. Therefore, we turned to investigate whether nCoVN obstructed the pluripotency maintenance in iPSC. We believed that this study could help us to understand the deleterious effects of nCoVN to human adult stem cells and embryonic stem cells.

2 Method

2.1 Cell culture and differentiation assay

Human induced pluripotent stem cells (iPSC) DYR0100 (from The American Type Culture Collection, ATCC) were plated on Matrigel matrix (hESC-Qualified, LDEV-Free, Corning, 354277)-coated plates, and then were cultured in DMEM/F-12 medium (Gibco, 11320033) supplemented with STEMUP® ES/iPS cell culture medium supplement (Nissan Chemical Corporation). STEMUP medium was changed every two days. iPSC were passaged every three to four days or when the cell culture was 80-90% confluent. During passages, iPSC were rinsed with 1× DPBS (Gibco, 14040133) for one time then were treated with 0.5 mM EDTA (Invitrogen, 15575020) in 1× DPBS (Gibco, 14190144) for 10 minutes at room temperature. The split ratio was 1:3-1:6. The detailed differentiation protocol was described in the previous published reports (Lin et al., 2017; Shekhar et al., 2018). Briefly, iPSC were treated with the small molecule CHIR99021 (Tocris, 4423, final concentration 10 μM) in the RPMI-BSA medium [RPMI 1640 Medium (HyClone, SH30027.01) supplemented with 213 μg/ml AA2P (l-ascorbic acid 2-phosphate magnesium) (Sigma, A8960) and 0.1% bovine serum albumin (BSA) (Sigma, A1470)] for 24 hours, then were incubated with RPMI-BSA medium for 48 hours. On differentiation day 4, cells were treated with the small molecule IWP2 (Tocris, 3533, final concentration 5 μM) in RPMI-BSA medium. After 48 hours, the medium was changed to RPMI-BSA medium. Then, RPMI 1640 medium supplemented with 3% KnockOut Serum Replacement (Gibco, 10828-028) was used to culture the cardiomyocytes in the following experiments. All the cells in this study (except iPSC-derived cardiomyocytes) were kept culturing in the STEMUP medium until they were applied to other assays.

2.2 Generation of iPSC-nCoVN

The cDNA of nCoVN with a N-terminal 6× His Tag coding sequence (GeneMedi) and puromycin resistance gene were sub-cloned into the plasmid pCW-Cas9-Blast (Addgene, 83481) to replace Cas9 and Blast cDNA, respectively. Lentivirus preparation using a third generation lentivirus packaging system were referred to the previous report (Jiang et al., 2015). We followed and modified the protocol from Zhang lab to detect MOI of the lentivirus and perform transduction (Shalem et al., 2014). After 24 hours of transduction, medium was changed to fresh STEMUP medium supplemented with doxycycline hyclate (Sigma, D9891) for induction. Two days later, puromycin (InvivoGen, ant-pr-1, final concentration 2 μg/mL) was added into the STEMUP medium supplemented with doxycycline hyclate. After 2-3 days' selection, which resulted in a transduction efficiency of ~30%, single cell clones were manually picked and re-seeded in separated wells. For

nCoVN expression induction, doxycycline hyclate (Sigma, D9891) was supplemented in the stem cell culture medium at a final concentration of 2 μg/mL, and the same amount of DMSO was added to the stem cell culture medium for controls.

2.3 Reverse transcription-PCR and Quantitative Real-time PCR

Total RNA was extracted by using the UNlQ-10 Column Trizol Total RNA Isolation Kit (Sangon Biotech, B511321-0100) prior to the treatment with DNase I (Sangon Biotech, B618252) for 30 minutes. mRNA was reverse transcribed by using iScript Reverse Transcription Supermix (Bio-Rad, 1708841). Quantitative Real-time PCR was performed by using a PikoReal Real-Time PCR System (Thermo Fisher) with SsoAdvanced™ Universal SYBR® Green Supermix (Bio-Rad, 1725271). The primers for Reverse transcription-PCR and Quantitative Real-time PCR are as followed (from 5′ to 3′):

ACE2-RT-F: GGTCTTCTGTCACCCGATTT;

ACE2-RT-R: ACCACCCCAACTATCTCTCG;

nCoVN-RT-F: CATTGGCATGGAAGTCACAC;

nCoVN-RT-R: TCTGCGGTAAGGCTTGAGTT;

GAPDH-RT-F: TGGGTGTGAACCATGAGAAG;

GAPDH-RT-R: GTGTCGCTGTTGAAGTCAGA.

2.4 The proliferation assay

iPSC, iPSC-GFP and iPSC-nCoVN were seeded in 96-well plates with the same cell number. After 24 hours, CCK-8 reagent was added in the medium to monitor the proliferation rate (Beyotime, C0038). The absorbance at 450 nm was measured at 24 hours, 42 hours, 48 hours, 60 hours and 72 hours by using Varioskan Flash Multimode Reader (Thermo Scientific). The cell-free medium with CCK-8 reagent were used as blank control sets. The data were analyzed and plotted using GraphPad Prism 6.

2.5 Immunofluorescence Staining

Cells were fixed with 4% paraformaldehyde at room temperature for 20 minutes and washed three times with 1× PBS. Cells were then permeabilized with PBS containing 0.25% Triton X-100 at room temperature for 10 minutes. After incubating in the blocking buffer (1× PBS with 10% goat serum), cells were stained with different primary antibodies at 4°C overnight. These primary antibodies were [target, dilution, species, company, product number]: Troponin T Cardiac Isoform, 1:100, mouse, Thermo Fisher, MA5-12960; alpha-smooth muscle actin, 1:100, mouse, Bioss, bsm-33187M; S100A4, 1:100, rabbit, Bioss, bs-3759R; SSEA4, 1:250, mouse, Invitrogen, 14-8843-80; TRA-1-81, 1:250, mouse, Invitrogen, 14-8883-80; 6× His Tag, 1:250, mouse, Sangon Biotech, D191001; OCT4, 1:200, mouse, Abcam, ab184665; ACE2, 1:200, rabbit, Bioss, bs-1004R; vimentin, 1:250, rabbit, Bioss, bs-0756R. Cells were washed three times with PBS containing 0.1% Triton X-100, then incubated with the Alexa Fluor 488 goat anti-mouse or Alexa Fluor 555 goat anti-rabbit IgG secondary antibodies at 37°C for 1 hour. Nuclei were labeled with DAPI (4′,6-diamidino-2-phenylindole, 1 μg/ml) for 5 minutes. Images were obtained by using the DMi6000 B inverted

microscope (Leica) or the FV1000 confocal laser scanning microscope (Olympus), then were analyzed by using ImageJ software.

2.6 RNA-seq analysis method

RNA-seq was performed by Novogene Co., Ltd. We obtained 151 bp paired-end RNA-seq reads from an Illumina Novaseq instrument, average 23 million read pairs for 3 iPSC-GFP and 3 iPSC-nCoVN samples. Adapters and low-quality bases in reads were trimmed by trim_galore (v0.6.5; http://www.bioinformatics.-babraham.ac.uk/projects/trim_galore/). We employed Kallisto (v0.46.0) (Bray et al., 2016) to determine the read count for each transcript and quantified transcript abundance as transcripts per kilobase per million reads mapped (TPM), using gene annotation in the GENCODE database (v32, GRCh38) (Frankish et al., 2019). Then we summed the read counts and TPM of all alternative splicing transcripts of a gene to obtain gene expression levels. We restricted our analysis to 22,201 expressed genes with an average TPM >=1 in either iPSC-GFP or iPSC-nCoVN samples. DESeq2 (v1.26.0) (Love et al., 2014) was used to identify differentially expressed genes (DEGs) (false discovery rate (FDR) <0.05 and abs(log2FoldChange) >3). The pathway analysis was performed by ToppGene Suite (Chen et al., 2009). The raw RNA-seq data have been deposited in GEO, and the accession number is pending.

We took gene expression values (i.e. $\log_2$(TPM)) in iPSC/ESC and fibroblast from ENCODE (Consortium, 2012; Davis et al., 2018). Combining with our RNA-seq data, quantile normalization (Bolstad et al., 2003) was performed, and then ComBat (Johnson et al., 2007) was used to remove the batch correction. We selected the top 1000 genes with the largest variance to calculate the correlation coefficients between samples. The heatmap was generated by pheatmap function in R.

2.7 Statistic

Values were expressed as mean ± SD (standard deviation). Statistical significances were evaluated using one-way ANOVA with Bonferroni correction or Student's T-Test. P<0.05 was considered statistically significant.

3 Results

3.1 *ACE2* was expressed in various of stem cells

As ACE2 was the major receptor of SARS-CoV-2 on the cell membrane (Hoffmann et al., 2020; Wrapp et al., 2020), we first examined whether *ACE2* was expressed in the stem cells. Thanks to the gene expression data collection in Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/), it was convenient to analyze the *ACE2* expression profiles in sorts of stem cells. Figure 1A showed the *ACE2* expression values in different stem cells from different projects, including human embryonic stem cells (Kim et al., 2014), iPSC (Yang et al., 2014), human epithelial stem cells (Yang et al., 2014), human adipose stem cells (Onate et al., 2013), human hematopoietic stem cells (Pang et al., 2011), and human mesenchymal stem cells (Bernstein et al., 2010). The expression values of a housekeeping gene *GAPDH* were simultaneously collected as controls. *ACE2* was expressed in each kind of stem cells, though the expression values were relatively low compared with *GAPDH*. The reverse transcription-PCR (RT-PCR) results showed that *ACE2* was expressed in iPSC, iPSC-derived cardiomyocytes (iPSC-CM) and human coronary artery endothelial cells (HCAEC) (Figure 1B). The images from immunofluorescence assays clearly showed that ACE2 protein was located on the cell membrane of iPSC (Figure 1C), suggesting that the pluripotent stem cells were the potential targets of SARS-CoV-2.

3.2 Expression of nCoVN changed the morphology of iPSC

To study whether physiological activities in iPSC were disturbed by nCoVN, a human induced pluripotent stem cell line (iPSC-nCoVN) in which the expression of nCoVN could be modulated by a Tet-On system was generated by a lentiviral expression system. In this system, nCoVN cDNA sequence (with a 6× His Tag coding sequence) was conjugated to puromycin resistance gene through a T2A peptide encoding sequence, and the transcription was relied on the induction of tetracycline or doxycycline (Dox). After puromycin selection, two single cell clones were seeded in separated wells by manual colony-picking. Sequentially, iPSC-nCoVN were divided into two groups: one was induced by Dox for nCoVN expression (Dox), the other was added with DMSO as a control set (DMSO), meanwhile, a GFP-expressed iPS cell line (iPSC-GFP), in which the expression of GFP was modulated by the same Tet-On system, was used as another control set in the following assays (Figure 2A).

The expression of nCoVN was confirmed at the mRNA and protein levels. The transcriptional level of nCoVN was measured by Real-time PCR in iPSC, Dox, and DMSO groups, and nCoVN expression increased about 267-fold in the Dox group compared with the DMSO group (Figure 2B). The nCoVN protein was detected by using an anti-6× His Tag antibody in cells from the Dox group (Supplementary Figure 1). The proliferation rate was compared among iPSC, iPSC-GFP and Dox groups by using a cell counting kit. The absorbance at 450 nm (A450) was measured at 24 hours, 42 hours, 48 hours, 60 hours and 72 hours after cell seeding. After three days of cell seeding, Dox group showed a decreased proliferation rate than both of iPSC and iPSC-GFP groups, indicating that nCoVN might hamper the growth and division of iPSC (Figure 2C). This observation was consistent with the previous finding about the nucleocapsid protein of SARS-CoV (Zhou et al., 2008).

We continued to induce nCoVN expression in iPSC. The phase-contrast images of iPSC-nCoVN with a 7-day, a 9-day, and an 11-day inductions (Dox) and counterpart controls (DMSO) were shown in Figure 2D. In the DMSO group, a typical morphology of stem cells with high nucleus/cytoplasm ratio and close cell membrane contacts was observed, while iPSC-nCoVN after a 7-day induction started to exhibit endothelial cell morphological features and lower nucleus/cytoplasm ratio. After a 14-day induction, most of the cells exhibited distinct shapes from wild-type iPSC, such as neuron-like cells, endothelial-like cells and fibroblast-like cells (Figure 2E). These data showed that continuous expression of nCoVN caused obvious morphological changes in iPSC.

3.3 Expression of nCoVN disabled the pluripotent properties of iPSC

Next, we examined the pluripotency markers in iPSC and iPSC-nCoVN. The pluripotency markers SSEA4 and TRA-1-81, which were expressed in human embryonic stem cells and iPSC, were widely applied in identification of pluripotent stem cells (Abujarour et al., 2013; Trusler et al., 2018). The immunofluorescence staining images illustrated that iPSC-nCoVN completely lost the expression of SSEA4 and TRA-1-81, namely, iPSC-nCoVN lost the pluripotency in the presence of nCoVN (Supplementary Figure 2A, B). We traced the expression of TRA-1-81 in iPSC-nCoVN with a 2-day, a 4-day, a 6-day, and an 8-day inductions (Figure 3A). On Day 2, TRA-1-81 was still expressed in iPSC-nCoVN; however, from Day 4, TRA-1-81 was not detectable in most of the cells, suggesting that the pluripotent fate of iPSC-nCoVN was determined in the first 4 days. To further test the pluripotency in iPSC and iPSC-nCoVN, we directly differentiated these cells to cardiomyocytes by using a routine protocol, and the differentiation assays were performed under the same conditions. As expected, the differentiation efficiency could reach 60% in iPSC; however, on differentiation day 12, only a very small portion of cells from iPSC-nCoVN were expressed cardiac Troponin T,

accompanied by many cell deaths (Figure 3B, C). This differentiation assay provided solid evidence that the pluripotency maintenance of iPSC-nCoVN was disrupted by nCoVN.

3.4 Long-term expression of nCoVN drove iPSC to fibroblast

Since the pluripotency lost due to short-term expression of nCoVN, we are extremely interested in the cell fate of iPSC-nCoVN under long-term expression of nCoVN. After a 28-day induction in the stem cell culture medium, some spindle-shaped iPSC-nCoVN, which exhibited a typical fibroblast morphological feature, were observed (Supplementary Figure 2C). The antibodies against fibroblast markers vimentin, alpha-smooth muscle actin (α-SMA) and S100A4 were used to verify the cell type of these fibroblast-like cells. The results from immunofluorescence assays confirmed that these markers were expressed in nCoVN-expressing cells (Figure 4A, B, C; Supplementary Figure 2D). To further investigate the transcriptomic profiles of iPSC-nCoVN under the long-term nCoVN expression, doxycycline-induced iPSC-nCoVN and iPSC-GFP for 30 days were applied to RNA-seq. Through differentially express analysis, iPSC-nCoVN showed a dramatic gene expression change comparing with iPSC-GFP (Supplementary Table). Totally, 3,080 genes were significantly differentially expressed (FDR<0.05, |log2FoldChange|>3). Among them, the down-regulated genes in iPSC-nCoVN were most significantly enriched with proliferation and stem cell related pathways, including the Yamanaka factors-associated genes, such as *POU5F1*, *LIN28A*, *NANOG*, and *SOX2* (with a 790-fold, a 2306-fold, a 253-fold, and an 18-fold decrease, respectively) (Figure 4E); while the extracellular matrix and extracellular matrix-associated pathway was the most significantly enriched pathway in the up-regulated genes (Figure 4F). Next, we used RNA-seq data from the ENCODE project to evaluate the cell type of iPSC-nCoVN. Comparing with the transcriptome of the pluripotent stem cells and fibroblast in the ENCODE project (Consortium, 2012; Davis et al., 2018), iPSC-GFP samples were clustered with H7 and GM23338, which were the embryonic stem cells (ESC) and iPSC, respectively; while iPSC-nCoVN samples were clustered with multiple kinds of fibroblast (Figure 4G). Furthermore, iPSC-nCoVN with a 40-day induction, which were kept culturing in the stem cell medium, were totally differentiated to fibroblast (Figure 4D).

4 Discussion

According to the current knowledge about the life cycle of SARS-CoV, the nucleocapsid protein was translated by the host cell translation protein synthesis machinery (McBride et al., 2014; Song et al., 2019), and was localized mainly in the cytoplasm (Rowland et al., 2005). The primary function of nucleocapsid protein was to package the viral genome into nucleocapsids to protect the genomic RNA (McBride et al., 2014). During the formation of nucleocapsids, numerous nucleocapsid proteins bound to the viral RNA and started oligomerization. The viral reproductive strategies would synthesize nucleocapsid proteins as many as possible to meet the requirements of viral assembly, which meant the nucleocapsid proteins were overproduced. The findings that redundant nucleocapsid proteins interfered with the normal physiology of host cells were reported (Surjit et al., 2004; Zhao et al., 2006; Zhang et al., 2007a; Zhao et al., 2008; Zhou et al., 2008; Hu et al., 2017). In this study, we first presented that nCoVN abolished pluripotency and reduced the proliferation rate in human induced pluripotent stem cells. Long-term expression of nCoVN drove iPSC to fibroblast in spite of using the stem cell culture conditions. It was reported that the nucleocapsid protein of SARS-CoV facilitated TGF-β-induced PAI-1 expression to promote lung fibrosis (Zhao et al., 2008), which was also the possible pathway that nCoVN turned iPSC to fibroblast.

The time-course assays showed that the pluripotency marker disappeared in four days after nCoVN expression. This finding might be applied to a cell-based chemical screening model, in which the

candidate chemicals with potential ability to halt the iPSC differentiation caused by nCoVN are easily identified. More importantly, SARS-CoV-2 is not necessary in this model, which means it could be used in the routine laboratories and applied to high-throughput equipment with less risk.

In addition, how nCoVN breaks the pluripotency maintenance of iPSC is still a riddle. The pluripotency maintenance in stem cells requires delicate regulations to maintain the balance of pluripotency gene expression in a complicated network. Since nCoVN can bind RNAs, it is possible that nCoVN suppresses the key pluripotency gene's translation through occupying the particular sites of RNAs. Although the mechanism is unknown, the toxic effects of nCoVN are clear, which reminds us that SARS-CoV-2 might impair the reproductive system and hematopoietic system. In conclusion, we first reported expressing nCoVN could totally change the cell fate of iPSC, which provided new clues to help people fighting against the virus.

5 Conflict of Interest

Author Zebin Lin, Jinlian Mai, Lishi Zhou, and Bin Lin were employed by the company Guangdong Beating Origin Regenerative Medicine Co. Ltd. The remaining 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.

6 Author Contributions

ZL, ZW, PW, and BL had substantial contributions to the design of the paper; ZL, ZW, JM, LZ, YQ, and TC performed the experiments and analysed the data; ZC provided critical suggestions to improve the paper; ZL, ZW, PW, and BL wrote the manuscript. All authors (ZL, ZW, JM, LZ, YQ, TC, ZC, PW, and BL) had read and approved the final manuscript.

7 Funding

This work was supported by grants from the National Natural Science Foundation of China (NSFC) to Ping Wang (31900812).

8 Acknowledgments

This work is dedicated to all the medical staff who are still fighting against COVID-19 in China. Your efforts make us safer.

9 Figure legends

Figure 1. *ACE2* was expressed in human stem cells. (A) Expression values of *ACE2* and *GAPDH* derived from the Gene Expression Omnibus database. (B) Images from agarose gel electrophoresis for analyzing the Reverse transcription-PCR products. *ACE2* was expressed in iPSC, iPSC-CM and HCAEC. iPSC, human induced pluripotent stem cell; iPSC-CM, human induced pluripotent stem cell-derived cardiomyocyte; HCAEC, human coronary artery endothelial cell; NC, negative control. (C) Representative immunofluorescent staining images of ACE2 (red) and the pluripotency marker OCT4 (green) in iPSC. The cell nuclei were stained by DAPI (blue). The scale bar represents 10 μm.

Figure 2. nCoVN affected the proliferation and morphology of iPSC. (A) Schematic diagram illustrating the generation of iPSC-nCoVN and controls. Purple cells indicated iPSC without nCoVN expression, while yellow cells indicated iPSC with nCoVN expression. Green cells were iPSC stably

expressing GFP under the doxycycline induction. (B) The mRNA expression level of nCoVN was significantly elevated in iPSC-nCoVN for a long-term induction (n=3). **, $p<0.001$. (C) The time course of cellular proliferation from iPSC, iPSC-GFP and iPSC-nCoVN (n=6). At 72 hours after cell seeding, the values of A450 were significantly increased in iPSC and iPSC-GFP groups. **, $p<0.001$. (D) Representative phase-contrast images from nCoVN-positive cells (Dox group) and control cells (DMSO group) after a 7-day, a 9-day, and an 11-day inductions. The scale bar is 50 μm. (E) Representative phase-contrast images of iPSC-nCoVN after a 14-day induction show detailed morphological alterations. Images were taken under objectives with 10×, 20×, and 40× magnifications (from left to right panels). White dashed line boxes indicate the regions that are magnified in the right panel. The scale bar represents 50 μm.

Figure 3. iPSC-nCoVN lost the pluripotency. (A) Representative immunofluorescent staining images of pluripotency marker TRA-1-81 (green) in iPSC-nCoVN for a 2-day, a 4-day, a 6-day, and an 8-day inductions. The cell nuclei were stained by DAPI (blue). The scale bar represents 50 μm. (B) Representative immunofluorescent staining images of cardiomyocyte marker cardiac Troponin T (Red) in iPSC- and iPSC-nCoVN-derived cardiomyocytes. The cell nuclei were stained by DAPI (blue). The scale bar represents 50 μm. (C) The cardiac differentiation efficiency of iPSC and iPSC-nCoVN. Images taken from (B) were analyzed by using ImageJ software. The efficiency was calculated as the portion of cardiac Troponin T positive cells in all the cells. Approximately 6,000 cells were counted in each group. **, $p<0.001$.

Figure 4. Long-term expression of nCoVN turned iPSC to fibroblast. (A-C) Representative immunofluorescent staining images of vimentin (green), S100A4 (green), and α-SMA (green) in iPSC-nCoVN after a 10-day induction. The cell nuclei were stained by DAPI (blue). The scale bars represent 10 μm. (D) The representative bright field image of cells after a 40-day nCoVN expression. The morphology exhibits typical fibroblast features. The scale bar represents 50 μm. (E) The two most significantly enriched pathways in the down-regulated genes. The histogram shows the significance of the two pathways by using $-\log_{10}$(q values with Bonferroni correction). The Log 2-fold change values of a total of 14 genes in these two pathways were exhibited by the heatmap. (F) The three most significantly enriched pathways in the up-regulated genes. The histogram shows the significance of the three pathways by using $-\log_{10}$(q values with Bonferroni correction). The Log 2-fold change values of the top 20 up-regulated genes in these pathways were exhibited by the heatmap. (G) The heatmap of the correlation coefficients among iPSC-GFP, iPSC-nCoVN, ESC/iPSC and fibroblast from the ENCODE project. The 1000 most variable genes in the samples were used to calculate the correlation coefficients. iPSC-GFP and iPSC-nCoVN clustered with pluripotent stem cell and fibroblast, respectively.

10 References

Abujarour, R., Valamehr, B., Robinson, M., Rezner, B., Vranceanu, F., and Flynn, P. (2013). Optimized surface markers for the prospective isolation of high-quality hiPSCs using flow cytometry selection. *Sci Rep* 3**,** 1179. doi: 10.1038/srep01179.

Bernstein, P., Sticht, C., Jacobi, A., Liebers, C., Manthey, S., and Stiehler, M. (2010). Expression pattern differences between osteoarthritic chondrocytes and mesenchymal stem cells during chondrogenic differentiation. *Osteoarthritis Cartilage* 18(12)**,** 1596-1607. doi: 10.1016/j.joca.2010.09.007.

365 Bolstad, B.M., Irizarry, R.A., Astrand, M., and Speed, T.P. (2003). A comparison of normalization
366 methods for high density oligonucleotide array data based on variance and bias.
367 *Bioinformatics* 19(2)**,** 185-193. doi: 10.1093/bioinformatics/19.2.185.

368 Bray, N.L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq
369 quantification. *Nat Biotechnol* 34(5)**,** 525-527. doi: 10.1038/nbt.3519.

370 Chang, C.K., Hou, M.H., Chang, C.F., Hsiao, C.D., and Huang, T.H. (2014). The SARS coronavirus
371 nucleocapsid protein--forms and functions. *Antiviral Res* 103**,** 39-50. doi:
372 10.1016/j.antiviral.2013.12.009.

373 Chang, C.K., Lo, S.C., Wang, Y.S., and Hou, M.H. (2016). Recent insights into the development of
374 therapeutics against coronavirus diseases by targeting N protein. *Drug Discov Today* 21(4)**,**
375 562-572. doi: 10.1016/j.drudis.2015.11.015.

376 Chen, J., Bardes, E.E., Aronow, B.J., and Jegga, A.G. (2009). ToppGene Suite for gene list
377 enrichment analysis and candidate gene prioritization. *Nucleic Acids Res* 37(Web Server
378 issue)**,** W305-311. doi: 10.1093/nar/gkp427.

379 Cheung, Y.K., Cheng, S.C., Sin, F.W., Chan, K.T., and Xie, Y. (2008). Investigation of
380 immunogenic T-cell epitopes in SARS virus nucleocapsid protein and their role in the
381 prevention and treatment of SARS infection. *Hong Kong Med J* 14 Suppl 4**,** 27-30.

382 Chinese, S.M.E.C. (2004). Molecular evolution of the SARS coronavirus during the course of the
383 SARS epidemic in China. *Science* 303(5664)**,** 1666-1669. doi: 10.1126/science.1092002.

384 Consortium, E.P. (2012). An integrated encyclopedia of DNA elements in the human genome.
385 *Nature* 489(7414)**,** 57-74. doi: 10.1038/nature11247.

386 Das, D., Kammila, S., and Suresh, M.R. (2010). Development, characterization, and application of
387 monoclonal antibodies against severe acute respiratory syndrome coronavirus nucleocapsid
388 protein. *Clin Vaccine Immunol* 17(12)**,** 2033-2036. doi: 10.1128/CVI.00293-10.

389 Davis, C.A., Hitz, B.C., Sloan, C.A., Chan, E.T., Davidson, J.M., Gabdank, I., et al. (2018). The
390 Encyclopedia of DNA elements (ENCODE): data portal update. *Nucleic Acids Res* 46(D1)**,**
391 D794-D801. doi: 10.1093/nar/gkx1081.

392 Frankish, A., Diekhans, M., Ferreira, A.M., Johnson, R., Jungreis, I., Loveland, J., et al. (2019).
393 GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res*
394 47(D1)**,** D766-D773. doi: 10.1093/nar/gky955.

395 He, R., Leeson, A., Andonov, A., Li, Y., Bastien, N., Cao, J., et al. (2003). Activation of AP-1 signal
396 transduction pathway by SARS coronavirus nucleocapsid protein. *Biochem Biophys Res*
397 *Commun* 311(4)**,** 870-876. doi: 10.1016/j.bbrc.2003.10.075.

398 Hoffmann, M., Kleine-Weber, H., Schroeder, S., Kruger, N., Herrler, T., Erichsen, S., et al. (2020).
399 SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically
400 Proven Protease Inhibitor. *Cell*. doi: 10.1016/j.cell.2020.02.052.

401 Hu, Y., Li, W., Gao, T., Cui, Y., Jin, Y., Li, P., et al. (2017). The Severe Acute Respiratory
402 Syndrome Coronavirus Nucleocapsid Inhibits Type I Interferon Production by Interfering
403 with TRIM25-Mediated RIG-I Ubiquitination. *J Virol* 91(8). doi: 10.1128/JVI.02143-16.

404 Jiang, W., Hua, R., Wei, M., Li, C., Qiu, Z., Yang, X., et al. (2015). An optimized method for high-
405 titer lentivirus preparations without ultracentrifugation. *Sci Rep* 5**,** 13875. doi:
406 10.1038/srep13875.

Johnson, W.E., Li, C., and Rabinovic, A. (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 8(1)**,** 118-127. doi: 10.1093/biostatistics/kxj037.

Kim, J.J., Khalid, O., Namazi, A., Tu, T.G., Elie, O., Lee, C., et al. (2014). Discovery of consensus gene signature and intermodular connectivity defining self-renewal of human embryonic stem cells. *Stem Cells* 32(6)**,** 1468-1479. doi: 10.1002/stem.1675.

Lin, B., Lin, X., Stachel, M., Wang, E., Luo, Y., Lader, J., et al. (2017). Culture in Glucose-Depleted Medium Supplemented with Fatty Acid and 3,3',5-Triiodo-l-Thyronine Facilitates Purification and Maturation of Human Pluripotent Stem Cell-Derived Cardiomyocytes. *Front Endocrinol (Lausanne)* 8**,** 253. doi: 10.3389/fendo.2017.00253.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15(12)**,** 550. doi: 10.1186/s13059-014-0550-8.

Lu, R., Zhao, X., Li, J., Niu, P., Yang, B., Wu, H., et al. (2020). Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* 395(10224)**,** 565-574. doi: 10.1016/S0140-6736(20)30251-8.

McBride, R., van Zyl, M., and Fielding, B.C. (2014). The coronavirus nucleocapsid is a multifunctional protein. *Viruses* 6(8)**,** 2991-3018. doi: 10.3390/v6082991.

Onate, B., Vilahur, G., Camino-Lopez, S., Diez-Caballero, A., Ballesta-Lopez, C., Ybarra, J., et al. (2013). Stem cells isolated from adipose tissue of obese patients show changes in their transcriptomic profile that indicate loss in stemcellness and increased commitment to an adipocyte-like phenotype. *BMC Genomics* 14**,** 625. doi: 10.1186/1471-2164-14-625.

Pang, W.W., Price, E.A., Sahoo, D., Beerman, I., Maloney, W.J., Rossi, D.J., et al. (2011). Human bone marrow hematopoietic stem cells are increased in frequency and myeloid-biased with age. *Proc Natl Acad Sci U S A* 108(50)**,** 20012-20017. doi: 10.1073/pnas.1116110108.

Rowland, R.R., Chauhan, V., Fang, Y., Pekosz, A., Kerrigan, M., and Burton, M.D. (2005). Intracellular localization of the severe acute respiratory syndrome coronavirus nucleocapsid protein: absence of nucleolar accumulation during infection and after expression as a recombinant protein in vero cells. *J Virol* 79(17)**,** 11507-11512. doi: 10.1128/JVI.79.17.11507-11512.2005.

Severance, E.G., Bossis, I., Dickerson, F.B., Stallings, C.R., Origoni, A.E., Sullens, A., et al. (2008). Development of a nucleocapsid-based human coronavirus immunoassay and estimates of individuals exposed to coronavirus in a U.S. metropolitan population. *Clin Vaccine Immunol* 15(12)**,** 1805-1810. doi: 10.1128/CVI.00124-08.

Shalem, O., Sanjana, N.E., Hartenian, E., Shi, X., Scott, D.A., Mikkelson, T., et al. (2014). Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* 343(6166)**,** 84-87. doi: 10.1126/science.1247005.

Shekhar, A., Lin, X., Lin, B., Liu, F.Y., Zhang, J., Khodadadi-Jamayran, A., et al. (2018). ETV1 activates a rapid conduction transcriptional program in rodent and human cardiomyocytes. *Sci Rep* 8(1)**,** 9944. doi: 10.1038/s41598-018-28239-7.

Song, Z., Xu, Y., Bao, L., Zhang, L., Yu, P., Qu, Y., et al. (2019). From SARS to MERS, Thrusting Coronaviruses into the Spotlight. *Viruses* 11(1). doi: 10.3390/v11010059.

448 Suresh, M.R., Bhatnagar, P.K., and Das, D. (2008). Molecular targets for diagnostics and
449 therapeutics of severe acute respiratory syndrome (SARS-CoV). *J Pharm Pharm Sci* 11(2)**,**
450 1s-13s. doi: 10.18433/j3j019.

451 Surjit, M., Liu, B., Jameel, S., Chow, V.T., and Lal, S.K. (2004). The SARS coronavirus
452 nucleocapsid protein induces actin reorganization and apoptosis in COS-1 cells in the absence
453 of growth factors. *Biochem J* 383(Pt 1)**,** 13-18. doi: 10.1042/BJ20040984.

454 Trusler, O., Huang, Z., Goodwin, J., and Laslett, A.L. (2018). Cell surface markers for the
455 identification and study of human naive pluripotent stem cells. *Stem Cell Res* 26**,** 36-43. doi:
456 10.1016/j.scr.2017.11.017.

457 Wrapp, D., Wang, N., Corbett, K.S., Goldsmith, J.A., Hsieh, C.L., Abiona, O., et al. (2020). Cryo-
458 EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* 367(6483)**,**
459 1260-1263. doi: 10.1126/science.abb2507.

460 Wu, A., Peng, Y., Huang, B., Ding, X., Wang, X., Niu, P., et al. (2020). Genome Composition and
461 Divergence of the Novel Coronavirus (2019-nCoV) Originating in China. *Cell Host Microbe*
462 27(3)**,** 325-328. doi: 10.1016/j.chom.2020.02.001.

463 Xu, X., P., C., J., W., J., F., H., Z., X., L., et al. (2020). Evolution of the novel coronavirus from the
464 ongoing Wuhan outbreak and modeling of its spike protein for risk of human transmission.
465 *Sci China Life Sci* 63. doi: https://doi.org/10.1007/s11427-020-1637-5.

466 Yang, R., Zheng, Y., Burrows, M., Liu, S., Wei, Z., Nace, A., et al. (2014). Generation of
467 folliculogenic human epithelial stem cells from induced pluripotent stem cells. *Nat Commun*
468 5**,** 3071. doi: 10.1038/ncomms4071.

469 Zhang, L., Wei, L., Jiang, D., Wang, J., Cong, X., and Fei, R. (2007a). SARS-CoV nucleocapsid
470 protein induced apoptosis of COS-1 mediated by the mitochondrial pathway. *Artif Cells*
471 *Blood Substit Immobil Biotechnol* 35(2)**,** 237-253. doi: 10.1080/10731190601188422.

472 Zhang, X., Wu, K., Wang, D., Yue, X., Song, D., Zhu, Y., et al. (2007b). Nucleocapsid protein of
473 SARS-CoV activates interleukin-6 expression through cellular transcription factor NF-
474 kappaB. *Virology* 365(2)**,** 324-335. doi: 10.1016/j.virol.2007.04.009.

475 Zhao, G., Shi, S.Q., Yang, Y., and Peng, J.P. (2006). M and N proteins of SARS coronavirus induce
476 apoptosis in HPF cells. *Cell Biol Toxicol* 22(5)**,** 313-322. doi: 10.1007/s10565-006-0077-1.

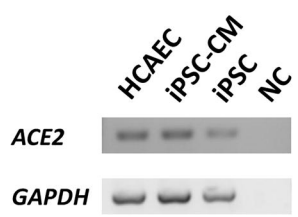
477 Zhao, X., Nicholls, J.M., and Chen, Y.G. (2008). Severe acute respiratory syndrome-associated
478 coronavirus nucleocapsid protein interacts with Smad3 and modulates transforming growth
479 factor-beta signaling. *J Biol Chem* 283(6)**,** 3272-3280. doi: 10.1074/jbc.M708033200.

480 Zhou, B., Liu, J., Wang, Q., Liu, X., Li, X., Li, P., et al. (2008). The nucleocapsid protein of severe
481 acute respiratory syndrome coronavirus inhibits cell cytokinesis and proliferation by
482 interacting with translation elongation factor 1alpha. *J Virol* 82(14)**,** 6962-6971. doi:
483 10.1128/JVI.00133-08.

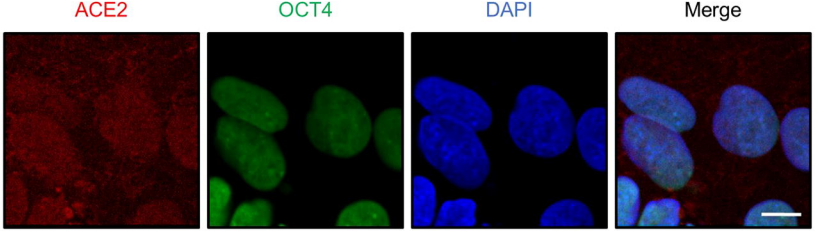
A

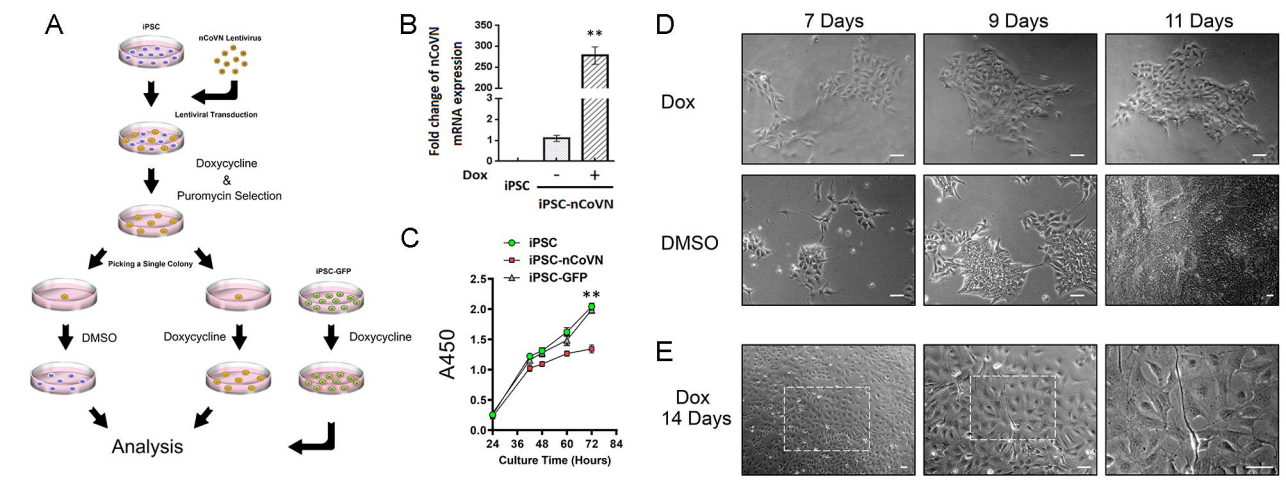
Dataset	Sample	Cell Line	ACE2 Ave Value	GAPDH Ave Value
GDS5408	GSM1309417	human embryonic stem cell H1 1	25.61	24162.72
	GSM1309418	human embryonic stem cell H1 2	30.16	26123.47
	GSM1309421	human embryonic stem cell H9 1	39.93	24325.00
	GSM1309422	human embryonic stem cell H9 2	27.90	23177.18
GDS5638	GSM1235179	human induced pluripotent stem cell 1	62.92	29904.83
	GSM1235180	human induced pluripotent stem cell 2	76.38	30926.80
	GSM1235184	human epithelial stem cell 1	67.86	15493.84
	GSM1235185	human epithelial stem cell 2	70.71	23540.97
GDS5056	GSM1187676	human adipose stem cell 1	5.77	14.29
	GSM1187677	human adipose stem cell 2	5.72	14.31
	GSM1187678	human adipose stem cell 3	5.87	14.34
GDS3942	GSM812988	human hematopoietic stem cell 1	3.77	12.47
	GSM812989	human hematopoietic stem cell 2	4.71	12.54
	GSM812990	human hematopoietic stem cell 3	4.20	12.61
GDS3785	GSM490983	human mesenchymal stem cell	5.47	12.29

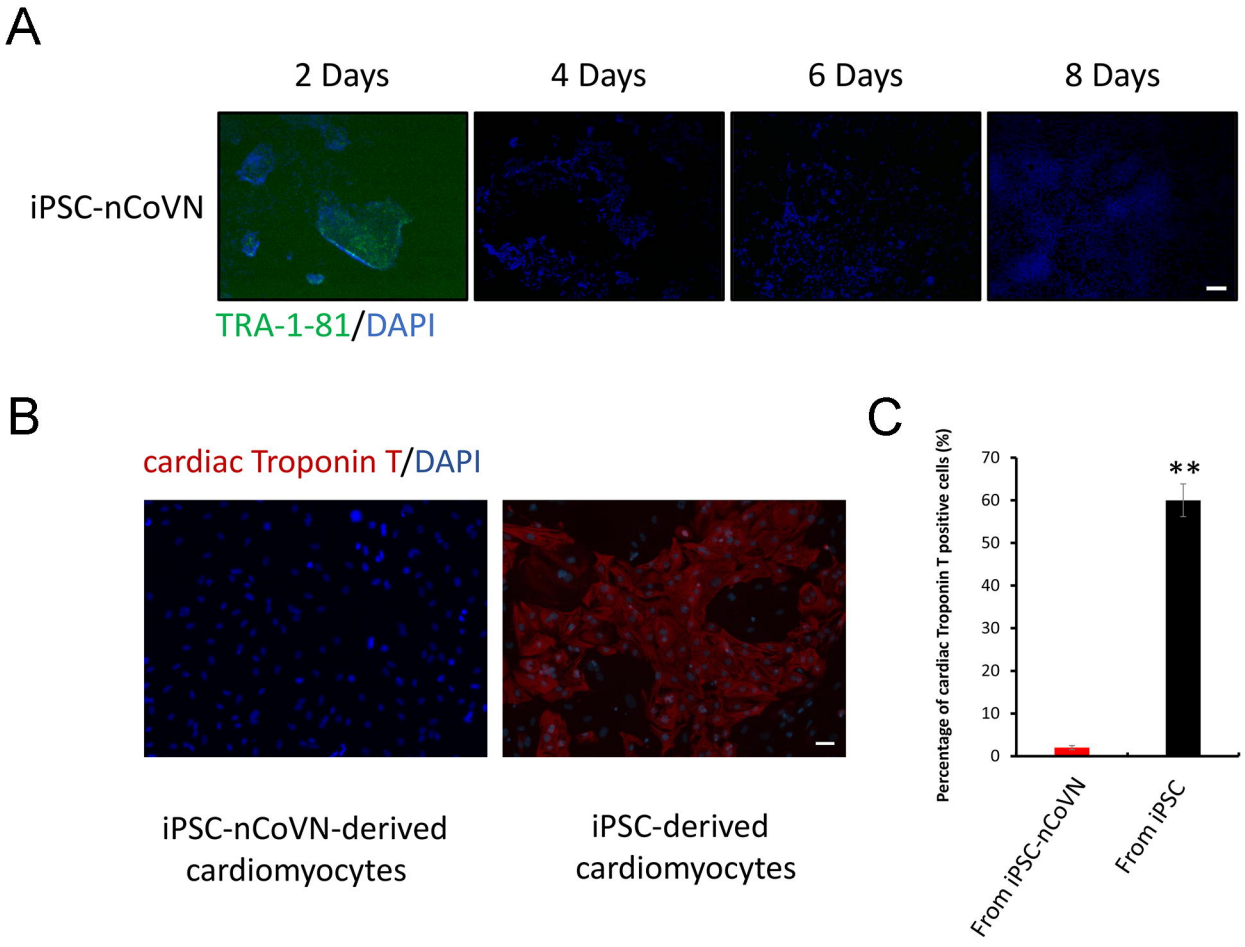
B

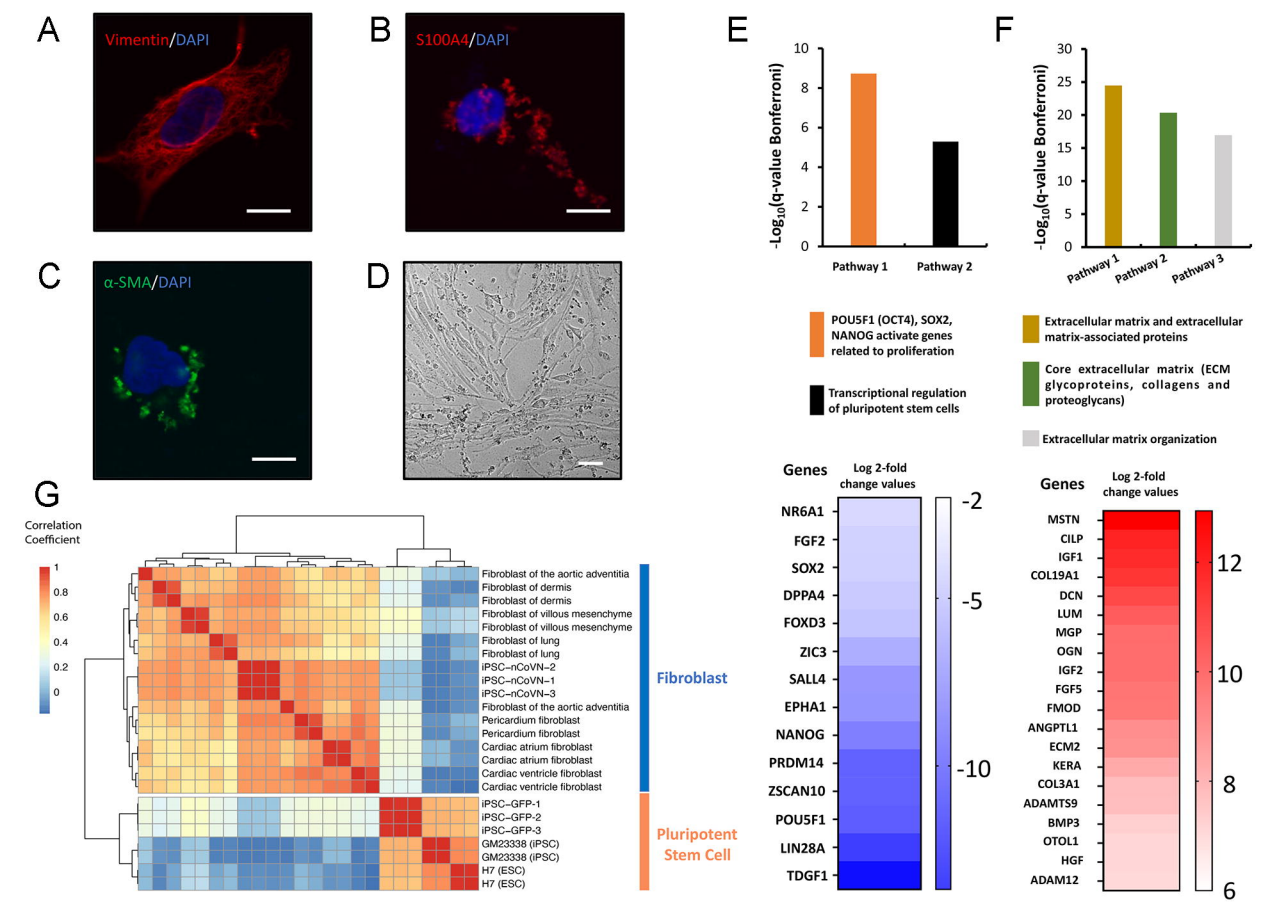


C




A
iPSC
nCoVN Lentivirus
Lentiviral Transduction
Doxycycline
&
Puromycin Selection
Picking a Single Colony
iPSC-GFP
DMSO
Doxycycline
Doxycycline
Analysis
B
Fold change of nCoVN mRNA expression
350
300
250
200
3
2
1
0
**
Dox
iPSC
-
+
iPSC-nCoVN
C
iPSC
iPSC-nCoVN
iPSC-GFP
A450
2.5
2.0
1.5
1.0
0.5
0.0
**
24
36
48
60
72
84
Culture Time (Hours)
D
7 Days
9 Days
11 Days
Dox
DMSO
E
Dox
14 Days


A
2 Days
4 Days
6 Days
8 Days
iPSC-nCoVN
TRA-1-81/DAPI
B
cardiac Troponin T/DAPI
iPSC-nCoVN-derived cardiomyocytes
iPSC-derived cardiomyocytes
C
Percentage of cardiac Troponin T positive cells (%)
70
60
50
40
30
20
10
0
**
From iPSC-nCoVN
From iPSC


A
Vimentin/DAPI
B
S100A4/DAPI
C
α-SMA/DAPI
D
E
-Log10(q-value Bonferroni)
Pathway 1
Pathway 2
POU5F1 (OCT4), SOX2, NANOG activate genes related to proliferation
Transcriptional regulation of pluripotent stem cells
F
-Log10(q-value Bonferroni)
Pathway 1
Pathway 2
Pathway 3
Extracellular matrix and extracellular matrix-associated proteins
Core extracellular matrix (ECM glycoproteins, collagens and proteoglycans)
Extracellular matrix organization
G
Correlation Coefficient
Fibroblast of the aortic adventitia
Fibroblast of dermis
Fibroblast of dermis
Fibroblast of villous mesenchyme
Fibroblast of villous mesenchyme
Fibroblast of lung
Fibroblast of lung
iPSC–nCoVN–2
iPSC–nCoVN–1
iPSC–nCoVN–3
Fibroblast of the aortic adventitia
Pericardium fibroblast
Pericardium fibroblast
Cardiac atrium fibroblast
Cardiac atrium fibroblast
Cardiac ventricle fibroblast
Cardiac ventricle fibroblast
iPSC–GFP-1
iPSC–GFP-2
iPSC–GFP-3
GM23338 (iPSC)
GM23338 (iPSC)
H7 (ESC)
H7 (ESC)
Fibroblast
Pluripotent Stem Cell
Genes
Log 2-fold change values
NR6A1
FGF2
SOX2
DPPA4
FOXD3
ZIC3
SALL4
EPHA1
NANOG
PRDM14
ZSCAN10
POU5F1
LIN28A
TDGF1
Genes
Log 2-fold change values
MSTN
CILP
IGF1
COL19A1
DCN
LUM
MGP
OGN
IGF2
FGF5
FMOD
ANGPTL1
ECM2
KERA
COL3A1
ADAMTS9
BMP3
OTOL1
HGF
ADAM12