Systematic temporal analysis of transcriptomes using *TrendCatcher* identifies early and persistent neutrophil activation as an early hallmark of severe COVID-19

Xinge Wang[1,2,3], Mark Sanborn[1, 2, 3], Yang Dai[1] and Jalees Rehman[1, 2, 3*]

[1]Department of Biomedical Engineering, University of Illinois Colleges of Engineering and Medicine, Chicago, IL, USA

[2]Department of Pharmacology and Regenerative Medicine, University of Illinois College of Medicine, Chicago, IL, USA

[3]Division of Cardiology, Department of Medicine, University of Illinois College of Medicine, Chicago, IL, USA

*Please address correspondence to Jalees Rehman (jalees@uic.edu)

The University of Illinois at Chicago, College of Medicine (M/C 868)
835 S. Wolcott Ave. Rm. E403, Chicago, IL 60612, USA
Office phone: +1 312-996-5552; Fax: +1 312-996-1225

Keywords:
RNA-sequencing
Single cell RNA-sequencing
Temporal analysis
Gene expression dynamics
Differential expression
Genomics
Transcriptomics
Pathway analysis
COVID-19
Immunity
Inflammation

Abstract

Studying the dynamic shifts in gene expression during disease progression may provide important insights into the biological mechanisms that distinguish adaptive and maladaptive responses. Existing computational tools for the analysis of transcriptomic data are not designed to optimally identify distinct temporal patterns when analyzing dynamic differentially expressed genes (DDEGs). Moreover, there is also a lack of methods to assess and visualize the temporal progression of biological pathways or gene sets mapped from time course transcriptomic datasets. In this study, we developed the open-source R package *TrendCatcher* (https://github.com/jaleesr/TrendCatcher), which applies the smoothing spline ANOVA model and break point searching strategy to estimate dynamic signals. *TrendCatcher* identifies and visualizes distinct dynamic transcriptional gene signatures and biological processes from sequential datasets. We used *TrendCatcher* to perform a systematic temporal analysis of COVID-19 peripheral blood transcriptomic datasets, including both whole blood bulk RNA-seq and PBMC scRNA-seq time course data. *TrendCatcher* uncovered the early and persistent activation of neutrophils as well as impaired type I interferon (IFN-I) signaling in circulating cells as early hallmark of patients who progressed to develop severe COVID-19, whereas no such patterns were identified in individuals receiving SARS-CoV-2 vaccinations or patients with mild COVID-19. These results underscore the importance of systematic temporal analysis of gene expression during disease progression to identify biomarkers and possible therapeutic targets.

Introduction

Time-course transcriptomic profiling has been widely used to study and model dynamic biological processes in cells (Bar-Joseph, Gitter, & Simon, 2012). By profiling mRNA levels during consecutive time points, researchers can infer dynamic responses to various external cues that cannot be observed by looking at only initial and terminal states. Recent improvements in high-throughput RNA sequencing (RNA-seq) technologies, including single-cell RNA-sequencing (scRNA-seq) provide viable approaches to study dynamic gene expression changes (Hwang, Lee, & Bang, 2018). Especially scRNA-seq allows for the in-depth analysis of temporal changes in distinct cell populations, thus providing insights into the heterogeneity and dynamics of responses to environmental cues or pathogenic stimuli. However, the analysis and visualization of longitudinal bulk RNA-seq data or scRNA-seq can be computationally challenging.

Currently, there are two predominant strategies for the analysis of sequential transcriptomic datasets. One strategy treats the sampling time points as categorical variables and is based on generalized linear models (GLMs). GLM-based packages include DESeq2 (Love, Huber, & Anders, 2014), edgeR (McCarthy, Chen, & Smyth, 2012) and limma (Ritchie et al., 2015). A complementary strategy is to treat time as a continuous variable and fit the time expression data into a spline-like model. These methods include DESeq2Spline (DESeq2 adopted with spline model for temporal RNA-seq datasets) fitting, ImpulseDE2 (Fischer, Theis, & Yosef, 2018) and Next maSigPro (Nueda, Tarazona, & Conesa, 2014). The former strategies focus on the magnitude of change instead of the time order of gene expression, and may also suffer from a relative loss of statistical testing power, especially if many time points are assessed (Fischer et al., 2018). The latter strategy increases the power of detecting dynamic genes, but is based on strong model assumptions which are not optimally suited for multiphasic responses, such as gene trajectory patterns that reflect initial acute stimuli followed by counter-regulatory compensatory responses. Furthermore, there is a lack of tools that leverage existing knowledge of functional pathway databases to infer and visualize pathway trajectories instead of individual gene trajectories only.

We found this is challenging in a complex disease, such as coronavirus disease 19 (COVID-19) caused by the SARS-CoV-2 virus infection (Zhou et al., 2020). COVID-19 is characterized by distinct disease progression patterns that suggest diverse host immune responses (Wu & McGoogan, 2020). Patients with severe disease exhibit profound inflammatory responses and immunopathology (Pedersen & Ho, 2020). COVID-19 immunophenotyping studies involve a

large number of time points from corresponding RNA-seq and scRNA seq datasets (Bergamaschi et al., 2021; Bernardes et al., 2020). Existing approaches for temporal transcriptomic analysis either do not take the temporal sequence into account when identifying dynamic differentially expressed genes (DDEGs) using pairwise comparison between time points, nor systematically analyze pathways that are dysregulated at defined time points.

In this study, we developed *TrendCatcher*, an open-source R package tailored for longitudinal bulk RNA-seq and scRNA-seq analysis. *TrendCatcher* uses a framework that combines the smoothing spline ANOVA model and break point searching strategy, which identifies inflection points when gene expression trends reverse. We show that *TrendCatcher* outperformed commonly used methods for longitudinal RNA-seq analysis when using simulated time course data for benchmarking. We also analyzed bulk RNA-seq and scRNA-seq gene expression profiles of peripheral blood cells in COVID-19 patients at various disease time points. *TrendCatcher* allowed us to identify and visualize dynamic gene expression signature profiles in peripheral blood that were associated with poor disease outcomes during the early phases of disease, and could thus serve as novel mechanistic targets as well as early biomarkers for patient prognostication.

Results

TrendCatcher accurately identifies DDEGs in simulated datasets

First, we tested the prediction performance of the *TrendCatcher* platform (**Figure 1A**) using a set of simulated time course RNA-seq datasets, because simulated data provides defined standards to assess the accuracy of novel analytical platforms. We considered a comprehensive collection of gene temporal trajectory patterns to simulate a set of realistic data with biological characteristics (Bar-Joseph et al., 2012). We embedded 10,000 simulated trajectories with varied temporal patterns, including 90% non-dynamic trajectories, 2.5% monotonous transition trajectories (continuously increasing or continuously decreasing gene expression levels throughout the time course), 2.5% impulse shaped single break point trajectories (only one temporal inflection point, i.e. up-peak-down or down-trough-up), 2.5% two break point trajectories and 2.5% three break point trajectories (multimodal dynamic response, e.g., a combination of 2 or more basic types of trajectories). Compared to DESeq2, DESeq2Spline and ImpulseDE2, *TrendCatcher* had the highest area under the ROC curve (AUC) in a mixed simulated dataset for time-course data with 7 time points (**Figure 1B**). We also tested each model's performance on a varying number of time points, including 3, 5, 7, 9 and 11 time points. As shown in **Figure 1C**, *TrendCatcher* had the highest prediction AUC across all time points, with the AUC values range from 0.88 to 0.90.

We next evaluated the prediction performance for each type of temporal trajectories. As shown in (**Figure 1 – Supplement 1**), although other three methods achieved slightly higher accuracy than *TrendCatcher* in monotonic trajectories, their AUCs dropped markedly when more complicated trajectories were embedded. DESeq2 only achieved AUC values of 0.49 to 0.62 for both bi-phasic trajectory and multimodal trajectory. The DESeq2Spline approach using a spline curve fitting model also dropped to an AUC of approximately 0.7 once multiphasic trajectories were introduced. These results suggests that existing approaches for longitudinal or time course analyses are well-suited for monotonic trajectories (continuously up or continuously down) but that *TrendCatcher* maybe more broadly applicable, because it identifies monotonic, biphasic (up-down, down-up) and multi-phasic shifts in gene expression, which are especially important in complex pathological setting when initial biological responses are followed by counter-regulatory adaptive or maladaptive responses.

TrendCatcher identifies rapid but transient upregulation of interferon signaling in peripheral blood following SARS-CoV-2 infection in a non-human primate model

To define the key dynamic gene signatures associated with SARS-CoV-2 infection in peripheral blood, we first analyzed the global transcriptomics profiles from a non-human primate dataset (Aid et al., 2020), in which samples were collected on Days 0 (uninfected controls), 1, 2, 4, 7, 10 and 14 following the live SARS-CoV-2 inoculation. This experiment in non-human primates had the advantage of clearly defining the timing of inoculation and following the time course of dynamic genes. *TrendCatcher* identified 962 DDEGs out of 12,754 total genes, accounting for 7.6% of total expression, suggesting that over 90% of the expressed genes in the peripheral blood remain close to the baseline expression levels even in the setting of the SARS-CoV-2 infection. We observed two major types of dynamic trajectories: (a) 167 genes followed a biphasic “0D-2D up, 2D-14D down” pattern, with their expression level peaking at Day 2 and gradually returning close to baseline levels at Day 14 (**Figure 2A**). These dynamic genes were primarily associated with host defense biological pathways, such as defense response to viruses, regulation of viral life cycle and type I interferon signaling pathways (**Figure 2B**). (b) 263 genes follow a “0D-14D down” pattern (**Figure 2A**). This set of genes followed a monotonous trajectory, with their expression gradually decreasing until Day 14. Interestingly, we found these genes were primarily associated with mitochondrial ATP synthesis and oxidative phosphorylation, suggesting that disruption of mitochondrial respiration and ATP generation may be an important feature of the circulating cells transcriptomic shift in a SARS-CoV-2 infection (**Figure 2C**). *TrendCatcher* assigned trajectory pattern types to all dynamic genes, and provided hierarchical pie charts to visualize the composition of trajectory patterns (**Figure 2-Supplement1 A**).

To systematically assess and visualize the dynamic programming of the top biological pathways associated with SARS-CoV-2 infection, *TrendCatcher* generated a *TimeHeatmap*. The *TimeHeatmap* function of *TrendCatcher* visualizes shifts in pathways by displaying the mean-fold change of individual DDEGs in a given pathway at defined time points while also depicting the number of dynamic genes and the percentage of dynamic genes within that pathway (**Figure 2D**). This quantifies the pathway level shifts over time, and allows for the magnitude of pathway change to be visualized. Together with the number and fraction of dynamic genes in a given pathway, to gauge the relative importance of the pathway during the time course.

173 During initial infection (Day 0 to Day2), pathways related to innate immune response and
174 interferon pathways were highly upregulated. Examples are upregulation of pathways such as
175 defense response to virus and regulation of innate immune response increased with an
176 average log2 fold-change of 2.57 and 1.76 within the first day. Mucosal immune response,
177 antimicrobial humoral response and killing of cells of other organisms were activated during
178 the later stage of infection (Day 4 to Day 7), with an average log2 fold-change around 2. On
179 the other hand, mitochondrial ATP synthesis coupled electron transport and protein targeting
180 to ER, on the other hand, were gradually down-regulated until Day 14. Dynamic gene
181 signatures from type I interferon signaling pathway and mitochondrial ATP synthesis coupled
182 electron transport were shown using traditional heatmaps in **Figure 2-Supplement 1 B,C**.
183 The temporal analysis of this non-human primate dataset highlights the rapidity of interferon
184 signaling activation btu also underscores that this initial burst of immune activation is transient
185 and followed by a gradual downregulation of the antiviral interferon responses during the first
186 week following infection.

187 **Increased generation of immunoglobulin synthesis in plasma B cells as early as**
188 **Day 1 of symptom onset in patients diagnosed with SARS-CoV-2 infection**
189
190 Next, we analyzed the longitudinal gene expression profiles of peripheral blood mononuclear
191 cells (PBMCs) obtained from patients diagnosed with a SARS-CoV-2 infection who were
192 admitted to the hospital, but predominantly had uncomplicated disease progression with 4 out
193 5 patients showing only mild symptoms (Zhu et al., 2020) (**Figure 3A**). The study performed
194 single cell RNA-seq analysis on PBMCs, thus allowing for a cell-type specific analysis of gene
195 expression shifts in distinct B cell subtypes, T cell subtypes and NK cells. We adopted the
196 cell-type labels from the original study and generated pseudo-bulk RNA-seq datasets for each
197 cell type in order to quantify changes in DDEGs for a specific cell type. We also adopted the
198 time annotation using stages, by binning the disease processes of COVID-19 patients from
199 symptom onset to discharge into stages 0, 1, 2, 3 and 4. We defined Stage 0 as the time point
200 of samples obtained from healthy controls and stages 1, 2, 3 and 4 represented Day 1 to Day
201 16 from symptom onset. We only applied *TrendCatcher* to cell types containing more than
202 1,000 cells at any given stage to ensure the robustness of the results.
203
204 *TrendCatcher* identified 400 DDEGs in memory B cells, 213 DDEGs in naive B cells, 1,413
205 DDEGs in plasma B cells, 398 DDEGs in $CD4^+$ T cells, 645 DDEGs in $CD8^+$ T cells, 423
206 DDEGs in MAIT, 1,161 DDEGs in naive T cells and 667 DDEGs in NK cells. The
207 *TimeHeatmap* of plasma B cells visualized the most dynamic biological pathways. Importantly,
208 this temporal analysis of single cell RNA-seq data showed how rapidly plasma B cells ramp
209 up the upregulation Fc-gamma receptor signaling and immunoglobulin synthesis as early as
210 stage 1 (which corresponds to Day 1 of symptom onset) (**Figure 3B**). However, not all
211 immunoglobulin genes are upregulated during the same temporal phase. As shown in (**Figure**
212 **3-Supplement 1**), genes involved in immunoglobulin synthesis show distinct temporal
213 patterns. Due to the comparatively lower number of plasma B cells, the prominence of such
214 increased immunoglobulin changes may be diluted in peirpheral blood bulk RNA-seq analysis
215 or PBMC bulk RNA-seq analysis. However, temporal analysis of all peripheral blood
216 mononuclear cell types, across the whole time course demonstrated increases in type I
217 interferon signaling and defense responses to viruses as the most prominent changes over
218 time (**Figure 3C**), thus mirroring the responses to SARS-CoV-2 we observed in the peripheral
219 blood of non-human primates (**Figure 2**). To define cell-type specific temporal dynamics, we
220 generated *TimeHeatmaps* for individual PBMC cell types and found significant upregulation
221 of Type I interferon signaling in T cells, NK cells and memory B cells during Stage 1 but

222 subsequent downregulation by Stage 2 of the disease in patients with mild symptoms (**Figure**
223 **3-Supplement 2A-C, 3A-D**).

224 ***TrendCatcher* identifies early and persistent neutrophil activation as a hallmark**
225 **of severe COVID-19**
226
227 We then assessed whether a systematic temporal analysis of gene expression trends could
228 be used to distinguish disease severity and prognosis of COVID-19 patients. We thus applied
229 *TrendCatcher* to a time course human whole blood bulk RNA-seq dataset (Bergamaschi et
230 al., 2021), which contained longitudinal whole blood transcriptomes from COVID-19 patients
231 with mild, moderate and severe clinical outcomes. *TrendCatcher* identified 77 DDEGs from
232 the mild group, 226 DDEGs from the moderate group, and 1,205 DDEGs from the severe
233 group. Only 42 DDEGs were shared among these three groups (**Figure 4A**), and these were
234 primarily associated with B cell mediated immunity, Fc-gamma receptor signaling pathways,
235 humoral immune response and lymphocyte mediated immunity (**Figure 4B**). However,
236 *TrendCatcher* identified 978 DDEGs uniquely shown in the severe COVID-19 patient group.
237 Importantly, these genes were strongly enriched for neutrophil-related biological pathways,
238 including neutrophil activation and neutrophil mediated immunity. These severe-disease
239 associated genes also included genes found in pathways such as myeloid cell differentiation,
240 reactive oxygen species metabolic process, and positive regulation of cytokine production
241 (**Figure 4B**).
242
243 To systematically characterize which biological processes in whole blood RNA-seq samples
244 were most dynamic in severe COVID-19 patients and how they progress over time, we applied
245 the *TrendCatcher*'s *TimeHeatmap* function. As seen in the non-human primate PBMC and the
246 human COVID-19 single cell RNA-seq datasets, we again observed the upregulation of genes
247 involved in the response to virus, humoral immune response and type I interferon signaling
248 pathway within Week 1 of disease onset (**Figure 4C**). Importantly, some dynamic biological
249 responses were only enriched in the group of patients which subsequently progressed to
250 severe COVID-19, including neutrophil activation (117 DDEGs, 23.4% of the corresponding
251 GO term), blood coagulation (36 DDEGs, 10.5% of the corresponding GO term), and
252 regulation of response to cytokine stimulus and respiratory burst (**Figure 4C, Figure 4-**
253 **Supplement 1A-B**). For instance, neutrophil activation gene expression increased by a mean
254 of 1.3 log2-units (approximately 2.5-fold increase in mean gene expression) within Week 1,
255 increased continuously until Week 4, and only very gradually decreased in surviving patients
256 by Week 7. However, the summation of the averaged log2 fold-change from the *TimeHeatmap*
257 was larger than zero, which indicates that the neutrophil activation may not have returned fully
258 to baseline levels by 7 weeks. To confirm this, we applied LOESS smooth curve fitting to all
259 neutrophil activation DDEGs identified from the three severity groups. LOESS fitting confirmed

260 that severe COVID-19 patients showed markedly higher neutrophil activation at the early
261 stage of infection (Week 1- Week 2), and also remained highly activated even after 7 weeks.
262 Such a persistent neutrophil activation was not observed in either mild or moderate COVID-
263 19 patient groups (**Figure 4D**).
264
265 We also observed severe COVID-19 patients showed evidence of greater humoral immune
266 response gene upregulation (**Figure 4E**), as well as markedly higher upregulation blood
267 coagulation genes (**Figure 4F**) and respiratory burst genes (**Figure 4G**), which remained
268 upregulated even after Week 6. These were not found to be upregulated in patients with mild
269 or moderate COVID-19 disease. These DDEGs and dynamic biological pathways may thus
270 be suited to serve as early biomarkers that distinguish severe COVID-19 from mild and
271 moderate COVID. All these dynamic gene signatures from neutrophil activation, blood
272 coagulation and respiratory burst pathways were listed using traditional heatmap (**Figure 4-**
273 **Supplement 2A-C**).

274 **Early impaired IFN-I signaling in PBMC provides a hallmark of severe COVID-19**
275
276 Next, to define the cell-type specific dynamic gene signatures and biological processes, we
277 used *TrendCatcher* to analyze a human PBMC single cell RNA-seq dataset time course
278 dataset in which patients were categorized as having either moderate or severe COVID-19
279 (Liu et al., 2021). Importantly, as this dataset only contained mRNA from peripheral blood
280 mononuclear cells, it lacked mRNA from neutrophils. *TrendCatcher* generated "pseudo-bulk"
281 mRNA profiles for each cell type in order to perform the analysis of gene expression dynamics
282 in a cell-type specific manner. *TrendCatcher* identified more dynamic shifts in almost all cell
283 types from severe COVID-19 patients than moderate groups (**Supplement Table 1, 2**). To
284 identify dynamic responsive processes unique to the severe group, we compared GO
285 enrichment for each cell type between severe and moderate. As shown in **Figure 5A**, all innate
286 immune cells and some adaptive immune cells (including B cells, $CD8^+$ T cells) from moderate
287 COVID-19 were highly enriched in type I interferon signaling, negative regulation of viral
288 process and defense to virus, whereas this was not observed significantly enriched in severe
289 COVID-19. Severe COVID-19 patients were highly dynamic in MAPK cascade, NF-κB
290 signaling, T cell receptor signaling and positive regulation of cytokine production for both NK
291 cells and $CD4^+$ T cells. For monocytes and DC (dendritic cells), no uniquely enriched dynamic
292 biological processes were observed in severe COVID-19 patients versus moderate COVID-
293 19 patients.
294
295 Furthermore, we also found the extent of the interferon signaling response to be the key
296 distinguishing feature between moderate and severe COVID-19. To quantitatively compare
297 the trajectory differentiation of the type I interferon signaling pathway, we performed the
298 LOESS smooth fitting to the DDEGs identified from this pathway. As shown in **Figure 5B**,
299 there is a profound separation between these two groups, with early strong activation of IFN-
300 I signaling in NK cells, monocytes, B cells and $CD8^+$ T cells in moderate COVID-19 patients
301 but not in severe COVID-19 patients. Patients with moderate COVID-19 showed activation of
302 type I interferon signaling pathways whereas patients with severe COVID-19 had a blunted
303 activation of type I interferon signaling. This is also shown in the cell-type specific
304 *TimeHeatmap*. As shown in **Figure C,D**, although NK cells from both moderate and severe
305 COVID-19 patients demonstrated activation of type I interferon response within the first week,
306 moderate COVID-19 exhibited a stronger activation than severe group, with average 1.62 log2
307 fold-change compared to 0.87 log2 fold-change. In $CD8^+$ T cells, only moderate COVID-19
308 groups were observed to have a strong type I interferon response within the first week. On the
309 other hand, $CD8^+$ T cells in patients who would go on to develop severe COVID-19, showed
310 upregulation of cell proliferation and cell differentiation genes instead. Together, these data

311 suggest that a robust early Type I interferon response in PBMCs is associated with reduced
312 severity of COVID-19.

313 ***TrendCatcher* identifies metabolic gene expression shifts in NK cells as a**
314 **hallmark response to COVID-19 vaccination**
315
316 We next applied *TrendCatcher* to a longitudinal human PBMC single cell RNA-seq vaccination
317 dataset (Arunachalam et al., 2021), to provide insights into how the immune system
318 physiologically responds to mRNA vaccines over time. The vaccination study collected single-
319 cell PBMCs from 56 healthy volunteers vaccinated with the Pfizer-BioNtech at Day 0, 1, 2, 7,
320 21, 22, 28, and 42 after the vaccination. We processed each patient's scRNA-seq individually.
321 We clustered cells using the Seurat algorithm (Hao et al., 2021) and annotated the cell types
322 using SingleR (Aran et al., 2019). Then we remove cell types containing less than 1,000 cells
323 for each time point across all samples. As **Figure 6A** shown, it is the UMAP of one patient's
324 PBMCs single cell data at Day 0.
325
326 *TrendCatcher* identified 650 DDEGs in NK cells, 450 in B cells, 23 in $CD8^+$ T cells, 62 in
327 monocytes and only 6 in $CD4^+$ T cells. This indicates NK cells exhibit a strong dynamic gene
328 expression shift after vaccination. After comparing the GO enrichment analysis across cell
329 types, we found NK cells gene expression shifts were enriched in metabolic processes in
330 response to the Pfizer-BioNtech SARS-CoV-2 mRNA vaccine, such as regulation of cellular
331 amino acid metabolism and ATP metabolism (**Figure 6B**). As shown in **Figure 6C**, a
332 *TimeHeatmap* of NK cells shows 54 DDEGs, which account for nearly half of the ATP
333 metabolic process pathway genes and exhibited a mean of 0.44 log2 fold upregulation after
334 the first dose of vaccination. These findings indicate the reprogramming of metabolism in NK
335 cells may be an indicator of an intact vaccine response. Another hallmark of the vaccine
336 response was the activation of interferon pathways, including type I interferon signaling
337 pathway and response to interferon-gamma (**Figure 6-Supplement 1A**). Importantly, the
338 *TimeHeatmap* demonstrates that these gene expression shifts were very transient, usually
339 decreasing by Day 2 or Day 7, and were again upregulated when the subjects received the
340 booster vaccine at Day 21. Interferon pathways showed an average increase of 0.71 and 1.42
341 log2 fold-change. In adaptive immune cells, we also observed strong type I interferon signaling
342 pathways activated in both B cells and $CD8^+$ T cells (**Figure 6-Supplement 1B,C**).
343 Additionally, B cells demonstrated upregulation of antigen processing and protein synthesis
344 (**Figure 6-Supplement 1B**), which were upregulated only for the first day, and then began
345 decreasing. It is noteworthy that early upregulation of interferon pathways in the PBMCs of
346 vaccine recipients mirrors that seen in patients with moderate COVID-19 infection, consistent
347 with the notion that early and transient interferon upregulation is a hallmark of a healthy
348 immune response to the SARS-CoV-2 infection.

Discussion

Temporal analysis of gene expression is gaining importance in the analysis of complex dynamic processes such as disease progression. Besides gene dynamic pattern characterization, time-course gene expression data are also used to infer regulatory and signaling relationships among genes (Krieger, Lupo, Levy, & Barkai, 2020; Yuan & Bar-Joseph, 2019). Integrating with other different types of measurements, such as pathology and infection over time helps disentangle the complex dynamic processes and possible underlying mediators (Graw, 2021). Thus, accurate identification of dynamic differentially expressed genes (DDEGs) in a time course RNA-seq or single cell RNA-seq study can help identify time-dependent disease mechanisms, adaptive and maladaptive molecular signatures as well as potential biomarkers that may be associated with disease severity.

In recent years, tools have been implemented to characterize time course RNA-seq data; however, these tools were focused on either bulk RNA-seq datasets (Fischer et al., 2018; Love et al., 2014; Nueda et al., 2014) and few infer trajectories from scRNA-seq (Qiu et al., 2017; Van den Berge et al., 2020). Compared to methods that analyze time course data without considering the sequential nature of time points, modeling time as a continuous function avoids a relative loss of statistical testing power, especially when many multiple time points were studied (Fischer et al., 2018). There is also a need for methods which combine DDEGs identification with the visualization of dynamic pathway shifts.

In this study, we developed an open-source R-software package to perform temporal analysis in longitudinal bulk RNA-seq and scRNA-seq datasets. *TrendCatcher* can identify DDEGs and infer the trajectories for each DDEG. By combining a time-window sliding strategy on inferred gene trajectories and leveraging annotated biological pathway databases, *TrendCatcher* can infer and visualize the most dynamic biological pathways in response to the external stimuli. Furthermore, by utilizing the *TimeHeatmap* function, *TrendCatcher* can help researchers identify the magnitude and dynamic nature of pathways shifts. Using simulated datasets for benchmarking, *TrendCatcher* achieved higher accuracy (AUC = 0.9) compared to other commonly used methods for the analysis of temporal gene expression datasets, when analyzing three or more time points. The advantage of using *TrendCatcher* was especially apparent in the setting of biphasic or multiphasic temporal trajectories, in which gene expression levels can change their trend of upregulation or downregulation during the time course.

385 Human immune responses to SARS-CoV-2 infection are highly dynamic and time dependent
386 upregulation as well as downregulation of the immune responses are key to ensure optimal
387 host defense (Chen et al., 2020; Gong et al., 2020; Mehta et al., 2020; Qin et al., 2020). A
388 recent study of B cell dynamics highlighted how clonal lineages expand along infection
389 processes (Montague et al., 2021) and a parallel study of T cell memory clones identified the
390 dynamics of T-cell memory clones (Minervina et al., 2021). Understanding the dynamics of
391 the COVID-19 immune response could form the basis of developing therapies that are
392 appropriate for a given time window (Goyal, Cardozo-Ojeda, & Schiffer, 2020). Personalized
393 medicine or precision medicine are gaining traction by developing therapies tailored to patients
394 based on their genotype and phenotype, but personalization likely also requires tailoring
395 therapies based on the temporal phase of a disease. These studies highlight the need for
396 time-course or longitudinal analyses of the host responses to the SARS-CoV-2 infection.
397
398 In this study, we applied *TrendCatcher* to systematically analyze sequential blood samples
399 from either non-human primates infected with SARS-CoV-2, human patients with COVID-19
400 of varying severity or human subjects who received a SARS-CoV-2 mRNA vaccine.
401 *TrendCatcher* identified dynamic gene expression and biological pathway signatures for (a)
402 SARS-CoV-2 infection progression over time, (b) severe COVID-19 vs. moderate or mild
403 COVID-19 and (c) COVID-19 vaccine recipients versus control. *TrendCatcher* established
404 response to virus, humoral immune response and type I interferon signaling pathway
405 activation across peripheral blood cell types in mild, moderate and severe COVID-19.
406 However, we found temporal patterns of gene expression shifts that were unique in the severe
407 COVID-19 patients. Severe COVID-19 was associated with marked activation of neutrophils
408 as well as blunted type I interferon signaling as early as week 1. Importantly, severe COVID-
409 19 was associated with persistent activation of neutrophils and genes regulating coagulation
410 for as long as 6 weeks, underscoring that the temporal analysis by *TrendCatcher* identified
411 hallmarks of severe COVID-19 in peripheral blood samples.
412
413 Although roles of monocytes and macrophages in severe COVID-19 have been well
414 demonstrated (Merad & Martin, 2020), less is known about neutrophils, which are also an
415 important element of innate immunity, in COVID-19 (Laforge et al., 2020; Tomar, Anders,
416 Desai, & Mulay, 2020). As one of the first responders to microbial threats, neutrophils infiltrate
417 tissues and produce inflammatory molecules to facilitate pathogen clearance. Neutrophils are
418 particularly active in highly vascularized organs, such as lungs and kidneys, which are prime
419 targets in COVID-19 (Libby & Luscher, 2020). However, dysregulated neutrophil response,
420 such as prolonged activation may cause damage to vessels and underlying parenchymal
421 (Bardoel, Kenny, Sollberger, & Zychlinsky, 2014; Metzemaekers et al., 2021;

422 Papayannopoulos, 2018). Recently, some studies also focused on the kinetics of circulating
423 neutrophils during COVID-19 progression, and the key role neutrophils play in COVID-19
424 pathology (Kaiser et al., 2021; Metzemaekers et al., 2021; Radermecker et al., 2020). In our
425 study, we observed strongly neutrophil activation which was unique to the severe COVID-19
426 group. Importantly, we showed that neutrophils remained highly activated in surviving patients
427 with severe COVID-19 even after seven weeks. The increased activation of coagulation genes
428 in the whole blood mRNA studies in severe COVID-19 patients also points to another feature
429 that is associated with severe COVID-19, thrombotic and embolic complications such as
430 strokes (Ackermann et al., 2020; Aid et al., 2020; Rapkiewicz et al., 2020). We believe the
431 analysis of the RNA-seq and scRNA-seq data support neutrophil activation and upregulation
432 of coagulation genes as defining features of severe COVID-19, highlighting their role as early
433 biomarkers to provide prognostic information and thus optimally treat patients who have a
434 higher likelihood of progressing to severe disease. Importantly, our results also raise questions
435 about the role of neutrophils and of coagulation in what is referred to as “long COVID” or “post-
436 acute sequelae of COVID” (Carfi, Bernabei, Landi, & Gemelli Against, 2020; Huang et al.,
437 2021; Nalbandian et al., 2021), because we observed that the activation persists for as long
438 as seven weeks after the initial infection.
439
440 Our temporal analysis showed that severe COVID-19 single-cell PBMCs were characterized
441 by impaired type I interferon signaling at the onset of infection (Week 1), compared to the
442 moderate COVID-19 group. This impaired type I interferon signaling was identified in innate
443 cells, such as NK cells and monocytes, and adaptive cells, such as B cell and $CD8^+$ T cells.
444 Type I interferons, which are essential for antiviral immunity (McNab, Mayer-Barber, Sher,
445 Wack, & O'Garra, 2015; Muller et al., 1994), have been shown to be upregulated in COVID-
446 19 (Blanco-Melo et al., 2020; Hadjadj et al., 2020). Other studies also suggested that impaired
447 type I interferon signaling may promote severe COVID-19 and that interferon therapy could
448 be used as therapy in severe COVID-19 (Darazam et al., 2021). However, since these studies
449 were not designed for longitudinal analyses, the timing of when to intervene on IFN signaling
450 was not clear. Traditional analyses of gene expression data are not suited for the identification
451 of temporal windows and thus do not address whether impaired IFN production is present at
452 the onset of infection, whether it was delayed, or whether IFN production is exhausted after
453 an initial activation (Hadjadj et al., 2020). *TrendCatcher* addressed this question by showing
454 the dynamic timing of type I IFN from both the moderate group and severe group.
455
456 In our COVID-19 vaccination single cell PBMC temporal data analysis, we also identified
457 prominent metabolic shifts in NK cells. Cellular metabolism is recognized as an important
458 factor that can determine the fate and function of lymphocytes (Loftus & Finlay, 2016; O'Neill,

459 Kishton, & Rathmell, 2016). Certain metabolic pathways have been shown to have direct
460 immunoregulatory roles (Loftus & Finlay, 2016). Activated NK cells engage in a robust
461 metabolic response, which is required for normal effector functions (O'Brien & Finlay, 2019),
462 and our data suggests that assessing NK metabolic shifts may be an intriguing alternative to
463 assessing vaccine responsiveness, which does not rely on antibody levels which can be highly
464 variable in laboratory assays.
465
466 In conclusion**,** we have developed the novel *TrendCatcher* R package platform designed for
467 time course RNA-seq data analysis which identifies and visualizes distinct dynamic
468 transcriptional programs. When applied to real whole blood bulk RNA-seq time course
469 datasets from COVID-19 patients, we observed that patients with severe COVID-19 showed
470 gene expression profiles consistent with profound neutrophil activation early during the
471 progression of the disease (starting from the first week of symptom onset). This indicates that
472 neutrophil activation may be a key early biomarker and possible mediator of severe COVID-
473 19. When applied to scRNA-seq time course vaccination data of PBMCs (peripheral blood
474 mononuclear cells), *TrendCatcher* identified specific shifts in NK cell metabolic gene
475 expression programs. This raises important questions about the role of NK cells in vaccine
476 efficacy. The distinct kinetics of gene expression shifts identified by *TrendCatcher* in COVID-
477 19 patients may allow for the identification of therapeutic targets during distinct temporal
478 phases.
479
480
481 **Acknowledgments**
482 The studies were supported by NIH grants R01-HL149300 (to JR), P01-HL151327 (to JR),
483 and T32-HL139439 (to MS).
484 **Competing interests**
485 J.R. is a reviewing editor of eLife. There are no other competing interests to report.
486
487 **Author contributions**
488 X.W. and J.R. designed the study, X.W. performed the analyses and wrote the first draft of the
489 manuscript, J.R. co-wrote the initial draft of the manuscript. M.S. and Y.D. provided critical
490 revisions of the manuscript.

491 **Materials and Methods**

492 ***TrendCatcher* framework**

493 The main components of the *TrendCatcher* framework are shown in **Figure 1**. *TrendCatcher*
494 requires two main inputs: the raw count table C of a temporal study with a dimension of $m \times n$,
495 where m denotes the number of genes and n denotes the number of samples, and a user-
496 defined baseline time variable T, such as "0 hour". Since samples may have different
497 sequencing depths and batch effect, *TrendCatcher* integrates with limma (Ritchie et al., 2015)
498 and provides preprocessing steps, such as batch correction and normalization. For single-cell
499 RNA sequencing datasets, *TrendCatcher* extracts cells for each cell type annotated in the
500 meta data slot of Seurat object and convert it into cell-type specific "pseudobulk" time course
501 RNA library. Based on a user-specified threshold, relatively low abundant genes are removed
502 from the count table, and reads are normalized and batch effects are removed.
503 *TrendCatcher*'s core algorithm is composed of five main steps: (a) baseline fluctuation
504 confidence interval estimation, (b) model dynamic longitudinal count, (c) time point dynamic
505 p-value calculation, (d) gene-wise dynamic p-value calculation, and (e) break point screening
506 and gene-wise dynamic pattern assignment. Mathematical details will be expanded in the
507 following sections. For the output of *TrendCatcher*, there are mainly two components: a master
508 table and a set of functions for versatile visualization purposes. The master table contains all
509 the dynamic details of each single gene, including its dynamic p-value, its break point location
510 time, and its dynamic trajectory pattern. In addition to the master table, *TrendCatcher*
511 produces five main types of visualizations: (a) a figure showing the observed counts and fitted
512 splines of each gene, (b) genes trajectories from each trajectory pattern group (c) a
513 hierarchical pie chart that represents trajectory pattern composition, (d) a *TimeHeatmap* to
514 infer trajectory dynamics of top dynamic biological pathways and (e) a two-sided bar plot to
515 show the top most positively and negatively changed (averaged accumulative log2FC)
516 biological pathways.

517

518 **Baseline fluctuation confidence interval estimation**

519 We assumed that the observed number of RNA-seq reads count from the baseline time
520 (e.g., t=0 hour) $X_{i,t_{baseline}}$ was generated from a negative binomial distribution (Equation 1):

$$X_{i,t_{baseline}} \sim NB(\mu_{i,t_{baseline}}, \varphi_i) \quad (1)$$

521 Where gene $i = 1, \dots, n$ and $\mu_{i,t_{baseline}}$ is the mean count of gene i at baseline time, and φ_i is
522 the dispersion factor. First, the dispersion factor φ_i was pre-estimated as a constant
523 hyperparameter for each gene with DESeq2 (Love et al., 2014), as shown in Equation 2. Here,
524 $\sigma_i(t)$ is the variance at time t.

$$\sigma_i(t)^2 = \mu_i(t) + \varphi_i * \mu_i(t)^2 \quad (2)$$

Then, $\mu_{i,t_{baseline}}$ was estimated using maximum likelihood from Equation 3.

$$\hat{\mu}_{i,t_{baseline}} = argmax\ L_{NB}(u,\varphi) = argmax\ \frac{\Gamma(\varphi + x)}{x!\,\Gamma(\varphi)}\left(\frac{\mu}{\mu+\varphi}\right)^x\left(\frac{\varphi}{\varphi+\mu}\right)^\varphi \quad (3)$$

Based on the negative binomial distribution and the estimated mean count for baseline time, we constructed the 90% confidence interval $[X_{i,t_{baseline},0.05}, X_{i,t_{baseline},0.95}]$ as the baseline count fluctuation interval (Equation 4(a) and Equation 4(b)).

$$P\left(x_{i,t_{baseline}} \leq X_{i,t_{baseline},0.05} \mid \mu_{i,t_{baseline}}, \varphi_i\right) = 0.05 \quad (4a)$$

$$P\left(x_{i,t_{baseline}} \geq X_{i,t_{baseline},0.95} \mid \mu_{i,t_{baseline}}, \varphi_i\right) = 0.95 \quad (4b)$$

Model dynamic longitudinal count estimation

To model the time-dependent gene expression value, we applied a smoothing spline ANOVA model (C, 2014; Metwally et al., 2018) with a negative binomial family constraint to fit the reads from samples across non-baseline multiple time points. The random variable $X_{i,t(t\neq t_{baseline})}$ is assumed to follow the NB distribution in Equation 5, with a positive integer α represents number of failures before the αth success in a sequential of Bernoulli trials, and $p(t) \in (0,1)$ represents the success probability.

$$X_{i,t(t\neq t_{baseline})} \sim NB(\alpha, p(t)) \quad (5)$$

The log-likelihood given a time-course observed count $x = \{x_{i,t(t\neq t_{baseline})}\}_{i=1,\ldots,n;\ t=1,\ldots,T}$ is calculated as Equation 6.

$$L = log\ L(\alpha, p(t) \mid X = x\ )$$
$$= \sum_{t=1}^{T} \sum_{i=1}^{n} [x_{i,t(t\neq t_{baseline})}\ log\ (1 - p(t) + \alpha log p(t) + log\Gamma\left(\alpha + x_{i,t(t\neq t_{baseline})}\right) - loglog\Gamma(\alpha)] + C \quad (6)$$

Taking the logit link and model time effect, we define the logit link η

$$\eta = log\ \frac{p(t)}{1 - p(t)} \quad (7)$$

To allow for flexibility in the estimation of η , and find the best trade-off between goodness of fit and the smoothness of the spline curve, soft constraints of the form $J(\eta)$ is added to the minus log-likelihood, with the smoothing parameter $\lambda > 0$.

$$-L + \lambda\ J(\eta) \quad (8)$$

The solution to the optimization of Equation 8 leads to a smoothing fitting to the reads from samples across different non-baseline time points. The estimated mean of $\mu_{i,t(t\neq t_{baseline})}$ can be estimated using Equation 9.

$$\mu_{i,t(t\neq t_{baseline})} = \hat{\alpha} e^{\hat{\eta}(t)} = \frac{\hat{\alpha}\hat{p}(t)}{1-\hat{p}(t)} \tag{9}$$

Gene's dynamic p-value calculation

To calculate gene's non-baseline dynamic signal significance, each gene's non-baseline estimated mean count $\mu_{i,t(t\neq t_{baseline})}$ was tested against the baseline fluctuation interval. Based on Equation 10(a) and Equation 10(b), for each gene at each single non-baseline time point, a dynamic time p-value was calculated.

$$If\ \mu_{i,t(t\neq t_{baseline})} \geq \mu_{i,t_{baseline}},\ then\ p_{i,t} = \int P(x \geq \mu_{i,t(t\neq t_{baseline})} \mid \mu_{i,t_{baseline}}, \varphi_i) \tag{10a}$$

$$If\ \mu_{i,t(t\neq t_{baseline})} \leq \mu_{i,t_{baseline}},\ then\ p_{i,t} = \int P(x \leq \mu_{i,t(t\neq t_{baseline})} \mid \mu_{i,t_{baseline}}, \varphi_i) \tag{10b}$$

Then, we applied Fisher's combined probability test method to calculate a gene-wise dynamic p-value (Equation 11):

$$\chi^2_{i,2t} \sim -2\sum_{t=0}^{T} ln\,(p_{i,t}) \tag{11}$$

Trajectory pattern assignment

First, we connect all the significant dynamic signal time points, with a p-value threshold less than 0.05. Then we applied a break point searching strategy to capture the gene expression change trend. The definition of break point is defined using Equation 12(a) and Equation 12(b).

$$If\ \mu_{i,t} > \mu_{i,t_{next}}\ \ AND\ \ \mu_{i,t} > \mu_{i,t_{previous}}, break\ point\ type\ I. \tag{12a}$$

$$If\ \mu_{i,t} < \mu_{i,t_{next}}\ \ AND\ \ \mu_{i,t} < \mu_{i,t_{previous}}, break\ point\ type\ II. \tag{12a}$$

There are two types of break points: type I means gene up-regulated followed by a down regulation, and type II means gene expression level down regulated and then followed by an upregulation. By screening along the break point, the master-pattern and sub-pattern were assigned to each gene.

563 ***TimeHeatmap* enrichment analysis and Two-sided bar plot**

564 To build the *TimeHeatmap* for visualizing the biological pathway enrichment change over time,
565 we designed a window-sliding strategy to capture all the up-regulated or down-regulated
566 genes within each time interval. If we denote time vector as t_j and $j \in 1, \ldots, T$, each time
567 interval is denoted as $[t_{j-1}, t_j]$. We found N_{up} up-regulated genes and N_{down} down-regulated
568 genes within the time window $[t_{j-1}, t_j]$, then Fisher's exact test was performed to obtain the
569 Gene Ontology (GO) term enrichment with the corresponding time interval for up-regulated
570 genes and down-regulated genes separately. Users can select the top most enriched
571 biological pathways (ranked by enrichment p-value) for each time interval (by default is 10).
572 Then for each selected GO term within the corresponding time window, we calculated the
573 averaged log 2 fold-change (Avg_log2FC) of all the DDEGs from this GO term. A series of
574 Avg_log2FC values over time characterize the trajectory dynamics of the corresponding
575 biological pathway, it is defined as biological pathway trajectory inference in this study. The
576 summation of the series of Avg_log2FC estimates the averaged accumulative log 2 fold-
577 change (GO_mean_logFC) for the corresponding GO term. *TrendCatcher* ranks biological
578 pathways based on their dynamic magnitude inferred from the GO_mean_logFC value. Users
579 are free to choose the top most positively and negatively changed (averaged accumulative
580 log2FC) biological pathways. Besides GO enrichment analysis (Yu, Wang, Han, & He, 2012),
581 *TrendCatcher* also packages Enrichr (Kuleshov et al., 2016) biological pathway databases.
582 Visualization is constructed using the ComplexHeatmap (Gu, Eils, & Schlesner, 2016)
583 package.
584

585 **Simulated dataset**

586 To mimic the real biological RNA-seq dataset, we only allowed ~10% of the genes to be
587 dynamic responsive genes. In this study, we embedded 5 different types of trajectories into
588 the temporal RNA-seq simulated datasets, including non-dynamic trajectory (~90%),
589 monotonic trajectory (~2.5%), bi-phase trajectory (~2.5%), and multimodal trajectory (~5%)
590 including 2 break-point and 3 break-point trajectory. Each type of trajectory was constructed
591 by adding negative binomial distribution noise to the embedded trajectory count. For example,
592 for monotonic trajectory, we defined the first and last time-points' RNA expression level
593 change to be 0.5-2 log2 fold-change, then we added negative binomial distribution noise to
594 the embedded monotonic trajectory. To check how the AUC changes as time-course study
595 extends longer, we sampled 3, 5, 7, 9 and 11 different number of time points with evenly time
596 intervals, and randomly sampled 3 replicates for each time point. To validate how the
597 prediction AUC varies across different types of trajectories, we embedded ~10% of the genes

598 to have non-dynamic trajectories, and ~90% of the genes have a specific dynamic trajectory
599 (monotonic, bi-phase or multimodal).

600

601 **Pseudobulk RNA library construction**

602 To construct “pseudo-bulk” RNA library from single cell RNA-seq datasets, all cells for each
603 cell type in a given sample were computationally “pooled” by summing all counts for a given
604 gene. Since pseudobulk libraries composed of few cells are not likely modeled properly, we
605 removed cell types containing less than 1k cells in this study. Lowly expressed genes were
606 removed for each cell type as well, using the filterByExpr function from *edgeR* R package
607 (McCarthy et al., 2012). Gene counts were transformed using the log of the counts per million
608 (cpm) and library size was normalized using calcNormFactors function with the method =
609 “RLE”.

610

611 **Gene set enrichment of DDEGs**

612 Gene sets enrichment analysis in this study was performed using clusterProfiler (Yu et al.,
613 2012) R package, and enrichment comparison from multiple groups were visualized using the
614 compareCluster function.

615

Dataset	Organism	Number of time points	Study	Figure	Dataset Source	Reference
Bulk Peripheral Blood (Aid et al.)	Macaca mulatta	7	Covid vs. Normal	Figure 2	GSE156701	(Aid et al., 2020)
scRNA PBMC (Zhu et al.)	Homo Sapien	5	Covid vs. Normal	Figure 3	https://db.cngb.org/search/project/CNP0001102/	(Zhu et al., 2020)
Bulk Peripheral Blood (Bergamaschi et al.)	Homo Sapien	8	Severe vs. Moderate	Figure 4	https://www.covid19cellatlas.org/patient/citiid/	(Bergamaschi et al., 2021)
scRNA PBMC (Liu et al.)	Homo Sapien	4	Severe vs. Moderate	Figure 5	GSE161918	(Liu et al., 2021)
scRNA PBMC (Arunachalam et al.)	Homo Sapien	8	Vaccination	Figure 6	GSE171964 and GSE169159	(Arunachalam et al., 2021)

616

617 Table 1. Datasets presented in this study

618

Innate immune cells					
NK		Monocyte		DC	
Moderate	Severe	Moderate	Severe	Moderate	Severe
708	1,179	1,479	1,845	22	17

Supplement Table 1. Number of DDEGs for each innate immune cell type identified by *TrendCatcher*. Using FDR less than 0.05.

Adaptive immune cells							
B		$CD4^+$ T		$CD8^+$ T		MAIT	
Moderate	Severe	Moderate	Severe	Moderate	Severe	Moderate	Severe
969	1,344	190	1,060	634	817	71	13

Supplement Table 2. Number of DDEGs for each adaptive immune cell type identified by *TrendCatcher*. Using FDR less than 0.05.

References

Ackermann, M., Verleden, S. E., Kuehnel, M., Haverich, A., Welte, T., Laenger, F., . . . Jonigk, D. (2020). Pulmonary Vascular Endothelialitis, Thrombosis, and Angiogenesis in Covid-19. *New England Journal of Medicine, 383*(2), 120-128. doi:10.1056/NEJMoa2015432

Aid, M., Busman-Sahay, K., Vidal, S. J., Maliga, Z., Bondoc, S., Starke, C., . . . Barouch, D. H. (2020). Vascular Disease and Thrombosis in SARS-CoV-2-Infected Rhesus Macaques. *Cell, 183*(5), 1354-+. doi:10.1016/j.cell.2020.10.005

Aran, D., Looney, A. P., Liu, L. Q., Wu, E., Fong, V., Hsu, A., . . . Bhattacharya, M. (2019). Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nature Immunology, 20*(2), 163-+. doi:10.1038/s41590-018-0276-y

Arunachalam, P. S., Scott, M. K. D., Hagan, T., Li, C. F., Feng, Y. P., Wimmers, F., . . . Pulendran, B. (2021). Systems vaccinology of the BNT162b2 mRNA vaccine in humans. *Nature, 596*(7872), 410-+. doi:10.1038/s41586-021-03791-x

Bar-Joseph, Z., Gitter, A., & Simon, I. (2012). Studying and modelling dynamic biological processes using time-series gene expression data. *Nat Rev Genet, 13*(8), 552-564. doi:10.1038/nrg3244

Bardoel, B. W., Kenny, E. F., Sollberger, G., & Zychlinsky, A. (2014). The Balancing Act of Neutrophils. *Cell Host & Microbe, 15*(5), 526-536. doi:10.1016/j.chom.2014.04.011

Bergamaschi, L., Mescia, F., Turner, L., Hanson, A. L., Kotagiri, P., Dunmore, B. J., . . . Smith, K. G. C. (2021). Longitudinal analysis reveals that delayed bystander CD8(+) T cell activation and early immune pathology distinguish severe COVID-19 from mild disease. *Immunity, 54*(6), 1257-+. doi:10.1016/j.immuni.2021.05.010

Bernardes, J. P., Mishra, N., Tran, F., Bahmer, T., Best, L., Blase, J. I., . . . Deutsch, C.-O. I. (2020). Longitudinal Multi-omics Analyses Identify Responses of Megakaryocytes, Erythroid Cells, and Plasmablasts as Hallmarks of Severe COVID-19. *Immunity, 53*(6), 1296-+. doi:10.1016/j.immuni.2020.11.017

Blanco-Melo, D., Nilsson-Payant, B. E., Liu, W. C., Uhl, S., Hoagland, D., Moller, R., . . . tenOever, B. R. (2020). Imbalanced Host Response to SARS-CoV-2 Drives Development of COVID-19. *Cell, 181*(5), 1036-+. doi:10.1016/j.cell.2020.04.026

C, G. (2014). Smoothing spline ANOVA models: R package gss. *J Stat Softw., 58*(5), 25.

Carfi, A., Bernabei, R., Landi, F., & Gemelli Against, C.-P. A. (2020). Persistent Symptoms in Patients After Acute COVID-19. *Jama-Journal of the American Medical Association, 324*(6), 603-605. doi:10.1001/jama.2020.12603

Chen, G., Wu, D., Guo, W., Cao, Y., Huang, D., Wang, H. W., . . . Ning, Q. (2020). Clinical and immunological features of severe and moderate coronavirus disease 2019. *Journal of Clinical Investigation, 130*(5), 2620-2629. doi:10.1172/jci137244

Darazam, I. A., Shokouhi, S., Pourhoseingholi, M. A., Irvani, S. S. N., Mokhtari, M., Shabani, M., . . . Khoshkar, A. (2021). Role of interferon therapy in severe COVID-19: the COVIFERON randomized controlled trial. *Scientific Reports, 11*(1). doi:10.1038/s41598-021-86859-y

Fischer, D. S., Theis, F. J., & Yosef, N. (2018). Impulse model-based differential expression analysis of time course sequencing data. *Nucleic Acids Res, 46*(20), e119. doi:10.1093/nar/gky675

Gong, J., Dong, H., Xia, Q. S., Huang, Z. Y., Wang, D. K., Zhao, Y., . . . Lu, F. E. (2020). Correlation analysis between disease severity and inflammation-related parameters in patients with COVID-19: a retrospective study. *Bmc Infectious Diseases, 20*(1). doi:10.1186/s12879-020-05681-5

Goyal, A., Cardozo-Ojeda, E. F., & Schiffer, J. T. (2020). Potency and timing of antiviral therapy as determinants of duration of SARS-CoV-2 shedding and intensity of inflammatory response. *Science Advances, 6*(47). doi:10.1126/sciadv.abc7112

Graw, F. (2021). DISEASE Deciphering the triad of infection, immunity and pathology. *Elife, 10*. doi:10.7554/eLife.72379

681 Gu, Z. G., Eils, R., & Schlesner, M. (2016). Complex heatmaps reveal patterns and
682 correlations in multidimensional genomic data. *Bioinformatics, 32*(18), 2847-2849.
683 doi:10.1093/bioinformatics/btw313
684 Hadjadj, J., Yatim, N., Barnabei, L., Corneau, A., Boussier, J., Smith, N., . . . Terrier, B.
685 (2020). Impaired type I interferon activity and inflammatory responses in severe
686 COVID-19 patients. *Science, 369*(6504), 718-+. doi:10.1126/science.abc6027
687 Hao, Y. H., Hao, S., Andersen-Nissen, E., Mauck, W. M., Zheng, S. W., Butler, A., . . . Satija,
688 R. (2021). Integrated analysis of multimodal single-cell data. *Cell, 184*(13), 3573-+.
689 doi:10.1016/j.cell.2021.04.048
690 Huang, C. L., Huang, L. X., Wang, Y. M., Li, X., Ren, L. L., Gu, X. Y., . . . Cao, B. (2021). 6-
691 month consequences of COVID-19 in patients discharged from hospital: a cohort
692 study. *Lancet, 397*(10270), 220-232. doi:10.1016/s0140-6736(20)32656-8
693 Hwang, B., Lee, J. H., & Bang, D. (2018). Single-cell RNA sequencing technologies and
694 bioinformatics pipelines. *Experimental and Molecular Medicine, 50*.
695 doi:10.1038/s12276-018-0071-8
696 Kaiser, R., Leunig, A., Pekayvaz, K., Popp, O., Joppich, M., Polewka, V., . . . Stark, K.
697 (2021). Self-sustaining IL-8 loops drive a prothrombotic neutrophil phenotype in
698 severe COVID-19. *Jci Insight, 6*(18). doi:10.1172/jci.insight.150862
699 Krieger, G., Lupo, O., Levy, A. A., & Barkai, N. (2020). Independent evolution of transcript
700 abundance and gene regulatory dynamics. *Genome Research, 30*(7), 1000-1011.
701 doi:10.1101/gr.261537.120
702 Kuleshov, M. V., Jones, M. R., Rouillard, A. D., Fernandez, N. F., Duan, Q., Wang, Z., . . .
703 Ma'ayan, A. (2016). Enrichr: a comprehensive gene set enrichment analysis web
704 server 2016 update. *Nucleic Acids Res, 44*(W1), W90-97. doi:10.1093/nar/gkw377
705 Laforge, M., Elbim, C., Frere, C., Hemadi, M., Massaad, C., Nuss, P., . . . Becker, C. (2020).
706 Tissue damage from neutrophil-induced oxidative stress in COVID-19 (Jul,
707 10.1038/s41577-020-0407-1, 2020). *Nature Reviews Immunology, 20*(9), 579-579.
708 doi:10.1038/s41577-020-00425-7
709 Libby, P., & Luscher, T. (2020). COVID-19 is, in the end, an endothelial disease. *European*
710 *Heart Journal, 41*(32), 3038-3044. doi:10.1093/eurheartj/ehaa623
711 Liu, C., Martins, A. J., Lau, W. W., Rachmaninoff, N., Chen, J. G., Imberti, L., . . . Clinicians,
712 C. (2021). Time-resolved systems immunology reveals a late juncture linked to fatal
713 COVID-19. *Cell, 184*(7), 1836-+. doi:10.1016/j.cell.2021.02.018
714 Loftus, R. M., & Finlay, D. K. (2016). Immunometabolism: Cellular Metabolism Turns
715 Immune Regulator. *Journal of Biological Chemistry, 291*(1), 1-10.
716 doi:10.1074/jbc.R115.693903
717 Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and
718 dispersion for RNA-seq data with DESeq2. *Genome Biol, 15*(12), 550.
719 doi:10.1186/s13059-014-0550-8
720 McCarthy, D. J., Chen, Y., & Smyth, G. K. (2012). Differential expression analysis of
721 multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids*
722 *Res, 40*(10), 4288-4297. doi:10.1093/nar/gks042
723 McNab, F., Mayer-Barber, K., Sher, A., Wack, A., & O'Garra, A. (2015). Type I interferons in
724 infectious disease. *Nature Reviews Immunology, 15*(2), 87-103. doi:10.1038/nri3787
725 Mehta, P., McAuley, D. F., Brown, M., Sanchez, E., Tattersall, R. S., Manson, J. J., &
726 Speciality, H. L. H. A. (2020). COVID-19: consider cytokine storm syndromes and
727 immunosuppression. *Lancet, 395*(10229), 1033-1034. doi:10.1016/s0140-
728 6736(20)30628-0
729 Merad, M., & Martin, J. C. (2020). Pathological inflammation in patients with COVID-19: a
730 key role for monocytes and macrophages. *Nature Reviews Immunology, 20*(6), 355-
731 362. doi:10.1038/s41577-020-0331-4
732 Metwally, A. A., Yang, J., Ascoli, C., Dai, Y., Finn, P. W., & Perkins, D. L. (2018).
733 MetaLonDA: a flexible R package for identifying time intervals of differentially
734 abundant features in metagenomic longitudinal studies. *Microbiome, 6*(1), 32.
735 doi:10.1186/s40168-018-0402-y

Metzemaekers, M., Cambier, S., Blanter, M., Vandooren, J., de Carvalho, A. C., Malengier-Devlies, B., . . . Proost, P. (2021). Kinetics of peripheral blood neutrophils in severe coronavirus disease 2019. *Clinical & Translational Immunology, 10*(4). doi:10.1002/cti2.1271

Minervina, A. A., Komech, E. A., Titov, A., Koraichi, M. B., Rosati, E., Mamedov, I. Z., . . . Pogorelyy, M. V. (2021). Longitudinal high-throughput TCR repertoire profiling reveals the dynamics of T-cell memory formation after mild COVID-19 infection. *Elife, 10*. doi:10.7554/eLife.63502

Montague, Z., Lv, H. B., Otwinowski, J., DeWitt, W. S., Isacchini, G., Yip, G. K., . . . Mok, C. K. P. (2021). Dynamics of B cell repertoires and emergence of cross-reactive responses in patients with different severities of COVID-19. *Cell Reports, 35*(8). doi:10.1016/j.celrep.2021.109173

Muller, U., Steinhoff, U., Reis, L. F. L., Hemmi, S., Pavlovic, J., Zinkernagel, R. M., & Aguet, M. (1994). FUNCTIONAL-ROLE OF TYPE-I AND TYPE-II INTERFERONS IN ANTIVIRAL DEFENSE. *Science, 264*(5167), 1918-1921. doi:10.1126/science.8009221

Nalbandian, A., Sehgal, K., Gupta, A., Madhavan, M. V., McGroder, C., Stevens, J. S., . . . Wan, E. E. Y. (2021). Post-acute COVID-19 syndrome. *Nature Medicine, 27*(4), 601-615. doi:10.1038/s41591-021-01283-z

Nueda, M. J., Tarazona, S., & Conesa, A. (2014). Next maSigPro: updating maSigPro bioconductor package for RNA-seq time series. *Bioinformatics, 30*(18), 2598-2602. doi:10.1093/bioinformatics/btu333

O'Brien, K. L., & Finlay, D. K. (2019). Immunometabolism and natural killer cell responses. *Nature Reviews Immunology, 19*(5), 282-290. doi:10.1038/s41577-019-0139-2

O'Neill, L. A. J., Kishton, R. J., & Rathmell, J. (2016). A guide to immunometabolism for immunologists. *Nature Reviews Immunology, 16*(9), 553-565. doi:10.1038/nri.2016.70

Papayannopoulos, V. (2018). Neutrophil extracellular traps in immunity and disease. *Nature Reviews Immunology, 18*(2), 134-147. doi:10.1038/nri.2017.105

Pedersen, S. F., & Ho, Y. C. (2020). SARS-CoV-2: a storm is raging. *Journal of Clinical Investigation, 130*(5), 2202-2205. doi:10.1172/jci137647

Qin, C., Zhou, L. Q., Hu, Z. W., Zhang, S. Q., Yang, S., Tao, Y., . . . Tian, D. S. (2020). Dysregulation of Immune Response in Patients With Coronavirus 2019 (COVID-19) in Wuhan, China. *Clinical Infectious Diseases, 71*(15), 762-768. doi:10.1093/cid/ciaa248

Qiu, X. J., Mao, Q., Tang, Y., Wang, L., Chawla, R., Pliner, H. A., & Trapnell, C. (2017). Reversed graph embedding resolves complex single-cell trajectories. *Nature Methods, 14*(10), 979-+. doi:10.1038/nmeth.4402

Radermecker, C., Detrembleur, N., Guiot, J., Cavalier, E., Henket, M., d'Emal, C., . . . Marichal, T. (2020). Neutrophil extracellular traps infiltrate the lung airway, interstitial, and vascular compartments in severe COVID-19. *Journal of Experimental Medicine, 217*(12). doi:10.1084/jem.20201012

Rapkiewicz, A. V., Mai, X. C., Carsons, S. E., Pittaluga, S., Kleiner, D. E., Berger, J. S., . . . Reynolds, H. R. (2020). Megakaryocytes and platelet-fibrin thrombi characterize multi-organ thrombosis at autopsy in COVID-19: A case series. *Eclinicalmedicine, 24*. doi:10.1016/j.eclinm.2020.100434

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., & Smyth, G. K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res, 43*(7), e47. doi:10.1093/nar/gkv007

Tomar, B., Anders, H. J., Desai, J., & Mulay, S. R. (2020). Neutrophils and Neutrophil Extracellular Traps Drive Necroinflammation in COVID-19. *Cells, 9*(6). doi:10.3390/cells9061383

Van den Berge, K., de Bezieux, H. R., Street, K., Saelens, W., Cannoodt, R., Saeys, Y., . . . Clement, L. (2020). Trajectory-based differential expression analysis for single-cell sequencing data. *Nature Communications, 11*(1). doi:10.1038/s41467-020-14766-3

Wu, Z. Y., & McGoogan, J. M. (2020). Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China Summary of a Report of 72 314 Cases From the Chinese Center for Disease Control and Prevention. *Jama-Journal of the American Medical Association, 323*(13), 1239-1242. doi:10.1001/jama.2020.2648

Yu, G., Wang, L. G., Han, Y., & He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS, 16*(5), 284-287. doi:10.1089/omi.2011.0118

Yuan, Y., & Bar-Joseph, Z. (2019). Deep learning for inferring gene relationships from single-cell expression data. *Proceedings of the National Academy of Sciences of the United States of America, 116*(52), 27151-27158. doi:10.1073/pnas.1911536116

Zhou, F., Yu, T., Du, R. H., Fan, G. H., Liu, Y., Liu, Z. B., . . . Cao, B. (2020). Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet, 395*(10229), 1054-1062. doi:10.1016/s0140-6736(20)30566-3

Zhu, L. N., Yang, P. H., Zhao, Y. Z., Zhuang, Z. K., Wang, Z. F., Son, R., . . . Liu, W. J. (2020). Single-Cell Sequencing of Peripheral Mononuclear Cells Reveals Distinct Immune Response Landscapes of COVID-19 and Influenza Patients. *Immunity, 53*(3), 685-+. doi:10.1016/j.immuni.2020.07.009

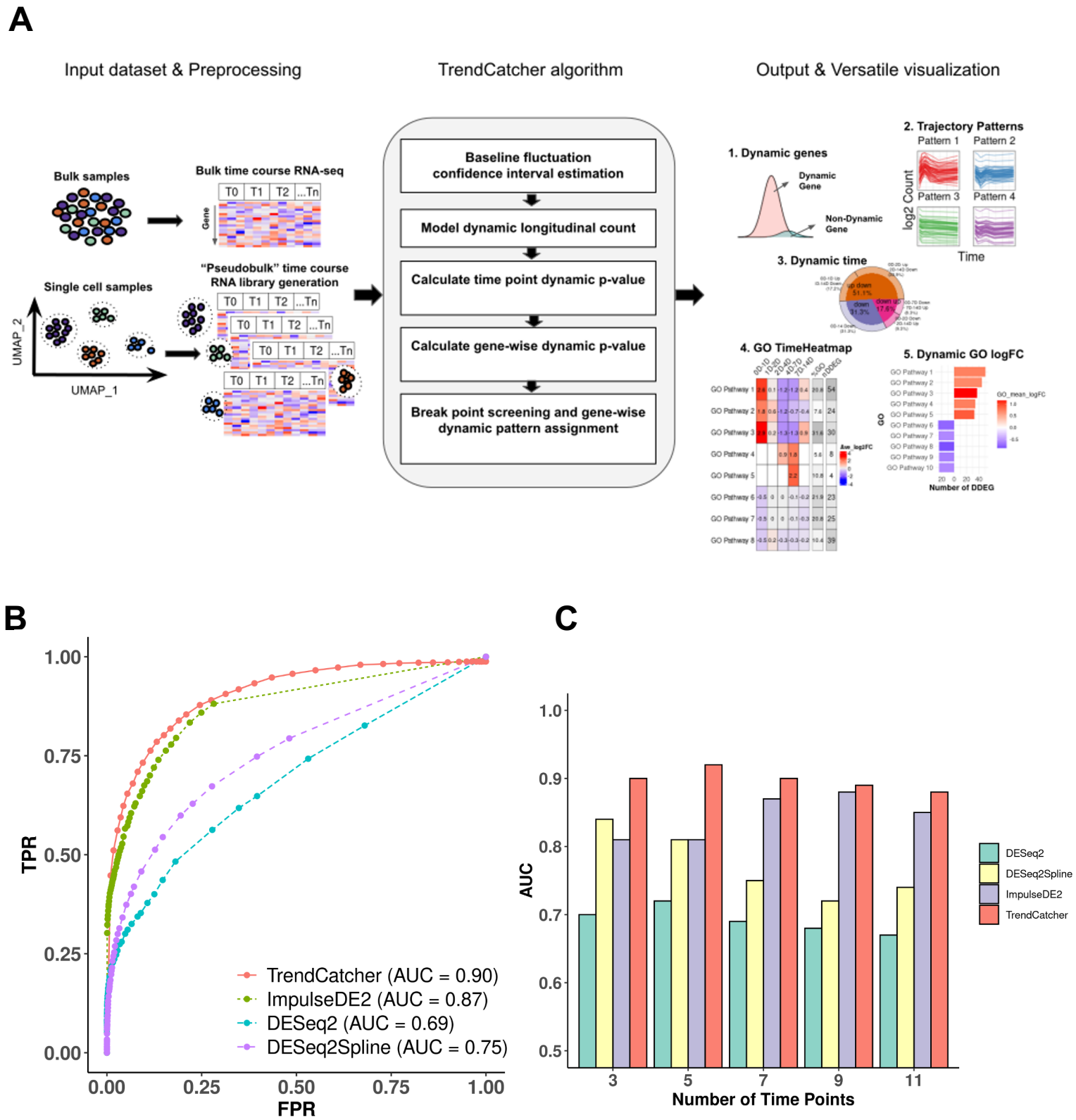


Figure 1. **Overview and benchmarking of of *TrendCatcher***. (A) *TrendCatcher* framework. *TrendCatcher* preprocess input data, which includes creating cell-type specific "pseudobulk" datasets for temporal analysis when single-cell data is used. *TrendCatcher*'s core algorithm is composed of five main steps. *TrendCatcher*'s output includes four main types of visualizations and DDEG identification (numbered 1-5). (B) *TrendCatcher* prediction ROC for a 7 time point simulated dataset compared to DESeq2, DESeq2Spline and ImpulseDE2, with mixed trajectories. (C) *TrendCatcher* prediction performance (AUC) across different number of time points, from 3 to 11 time points. *TrendCatcher*'s AUC values across time points from 3 to 11 are, 0.90, 0.92, 0.90, 0.89 and 0.88.

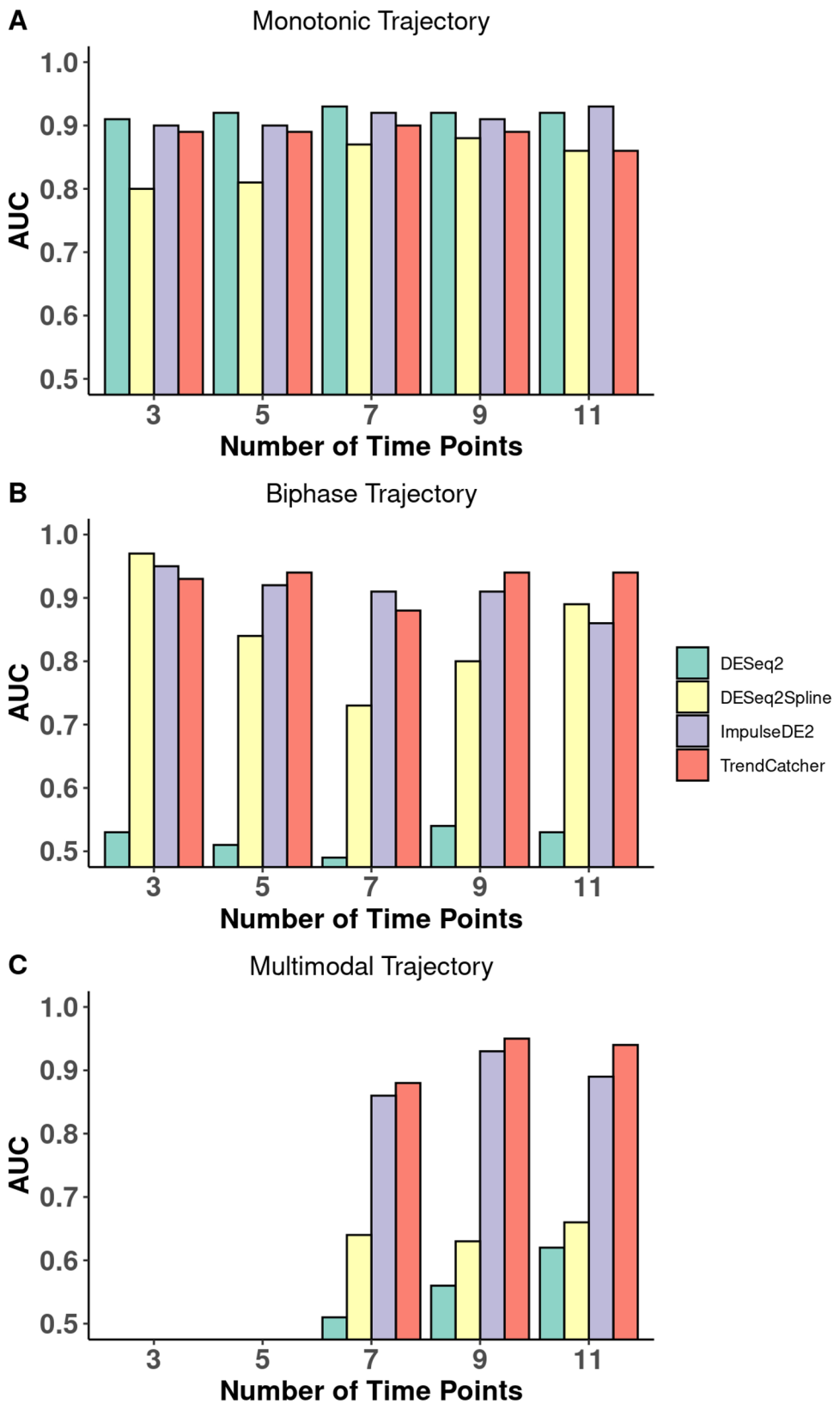
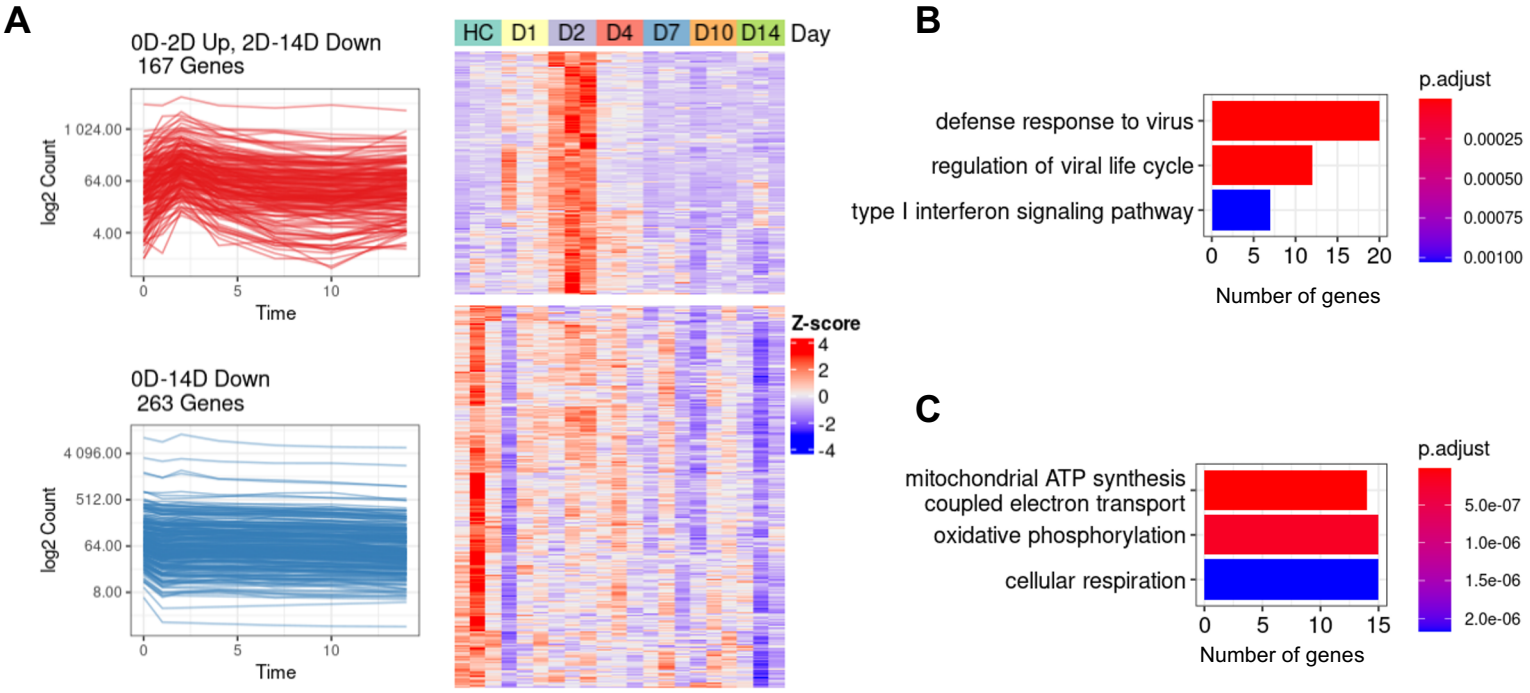


Figure 1 - Supplement 1. (A) Prediction performance of *TrendCatcher* across varying numbers of time points for monotonic trajectories. (B) Prediction performance of *TrendCatcher* across varying numbers of time points for biphasic trajectories. (C) Prediction performance of *TrendCatcher* across varying numbers of time points for multimodal trajectories. DESeq2 is shown in green, DESeq2Spline is shown in yellow, ImpulseDE2 is shown in purple, *TrendCatcher i*s shown in red.



D

	0D-1D	1D-2D	2D-4D	4D-7D	7D-10D	10D-14D	%GO	nDDEG
defense response to virus	2.57	0.1	-1.11	-1.18	-0.42	0.69	26.5	69
regulation of innate immune response	1.76	0.25	-1.03	-0.74	-0.32	-0.22	12.4	39
type I interferon signaling pathway	2.89	0.14	-1.33	-1.33	-0.48	1.81	32.6	31
regulation of viral life cycle	2.46	0.07	-1.11	-1.1	-0.34	1.3	27.7	41
regulation of viral genome replication	2.73	0.04	-1.14	-1.18	-0.35	1.42	34.9	29
interferon-gamma-mediated signaling pathway	2.26	0.01	-1.13	-1.06	-0.46	-0.14	23.1	21
response to interferon-beta	1.65	0.06	-0.82	-0.8	-0.26	-0.28	35.5	11
neutrophil mediated immunity	1.2		-0.55	0.56	-0.27	-0.5	9.6	48
regulation of response to cytokine stimulus	1.77	0.78	-1.02	-0.81	-0.34	-0.22	14.8	28
antimicrobial humoral response			0.95	1.86		-0.67	9.9	14
mucosal immune response				1.9			13.5	5
killing of cells of other organism				2.01	0.37	-0.68	13.4	9
mitochondrial ATP synthesis coupled electron transport	-0.26	-0.03	-0.14	-0.11	-0.13	-0.31	15.2	15
protein targeting to ER	-0.54	0.01	0.06	0.28		-0.49	19.2	23

Ave_log2FC: 2, 1, 0, -1, -2

Figure 2. Dynamic gene expression in peripheral blood following SARS-CoV-2 inoculation in a non-human primate model. (A) Analysis of the two predominant trajectory patterns in the non-human primate peripheral blood RNA-seq data from Day 0 to Day 14. The top-left figure represents 167 DDEGs following an up-down expression pattern, which peaked at Day 2 then slowly decreased until Day 14. The top-right figure represents their expression using a traditional z-score normalized heatmap. The bottom-left figure represents 263 DDEGs following a monotonic down-regulated trajectory pattern and their gene expression values were represented in the corresponding heatmap on the right. (B) Top 3 Gene ontology (GO) enrichment analysis pathways using 167 DDEGs from trajectory pattern “0D-2D Up, 2D-14D Down”. (C) Top 3 Gene ontology (GO) enrichment analysis pathways using 263 DDEGs from trajectory pattern “0D-14D Down”. (D) *TimeHeatmap* of the top 15 dynamic pathways and their dynamic time windows visualizes the temporal patterns. Each column represents a time window. The “%GO” column represents the percentage of DDEGs found in the corresponding pathway. The “nDDEG” column represents number of DDEGs found in the corresponding pathway. Color represents the averaged log fold-change of the DDEGs within each time window for the corresponding pathway.

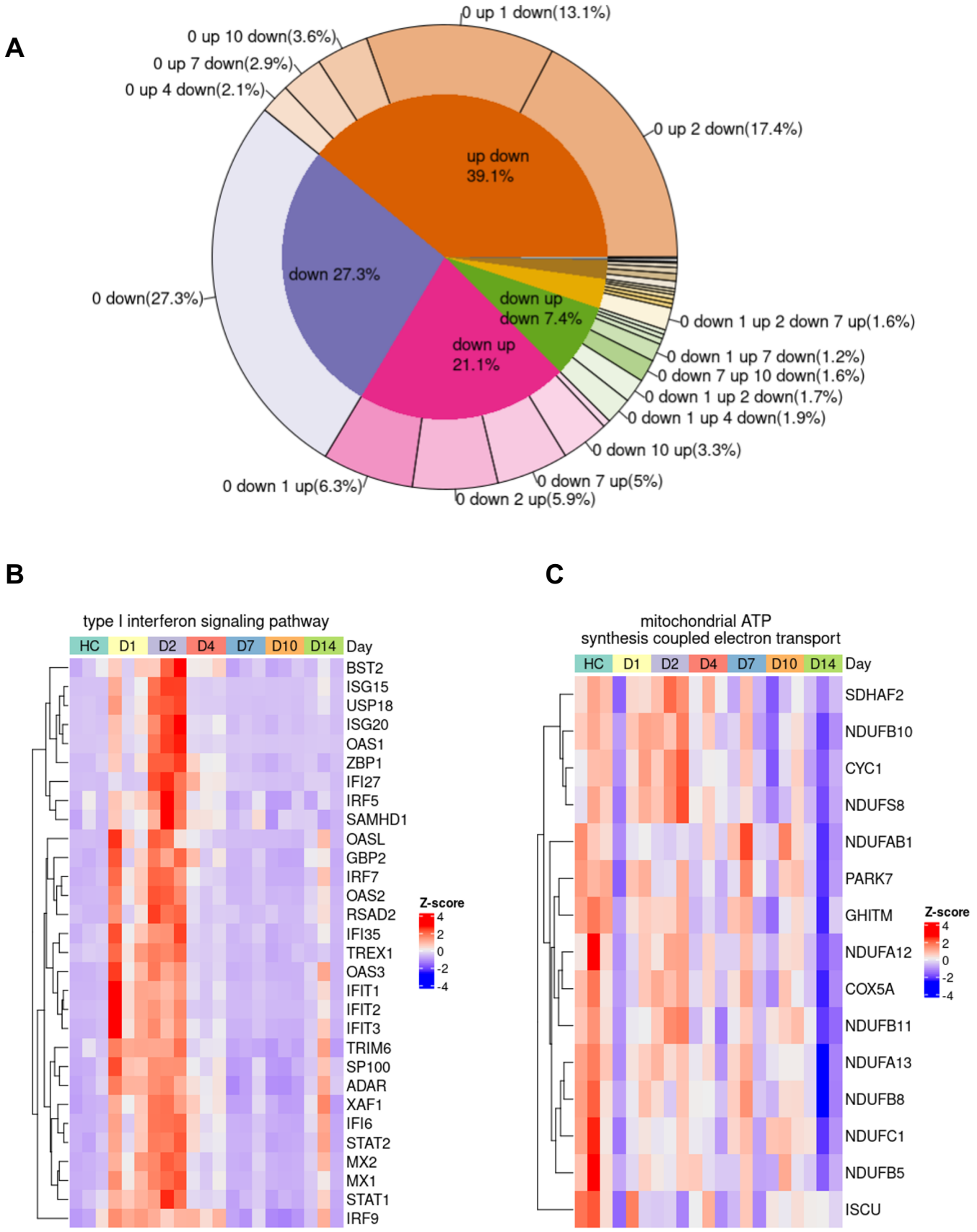


Figure 2-Supplement 1. (A) Hierarchical pie chart shows the composition of trajectory patterns of DDEGs identified in a non-human primate bulk peripheral blood mRNA (Aid et al.). (B) Dynamic gene signatures for SARS-CoV-2 infection from type I interferon signaling pathway. (C) Dynamic gene signatures response for SARS-CoV-2 infection from mitochondrial ATP synthesis coupled electron transport pathway. Color represents normalized z-score for gene expression.

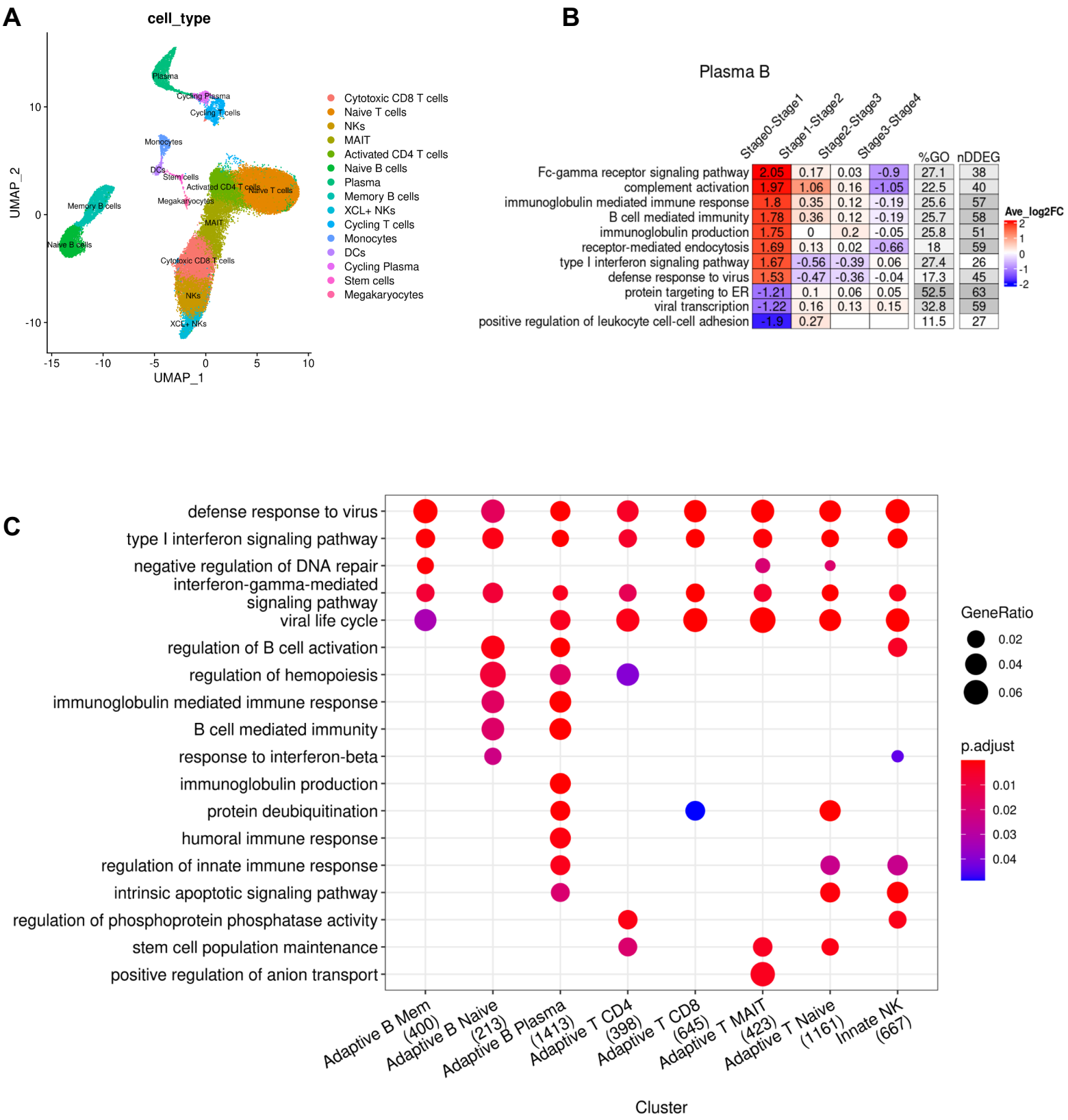


Figure 3. Cell-type specific dynamic gene expression in peripheral blood mononuclear cells following SARS-CoV-2 infection in patients. (A) UMAP visualization of single cell RNA-seq PBMC dataset (Zhu et al.) with annotated cell types from the original study. (B) *TimeHeatmap* of top dynamic biological pathway from plasma B cells. Each column represents a time window (stage). Stage 0 represents uninfected baseline. Each row represents the most dynamic biological pathways. Color represents the averaged log fold-change of the DDEGs within each time window from the corresponding pathway. (C) Top GO enrichment comparison analysis using DDEGs from each cell type. Dot size represents gene ratio. Dot color represents GO enrichment p-value. The number in the x-axis labels represent the number of DDEGs for each cell type.

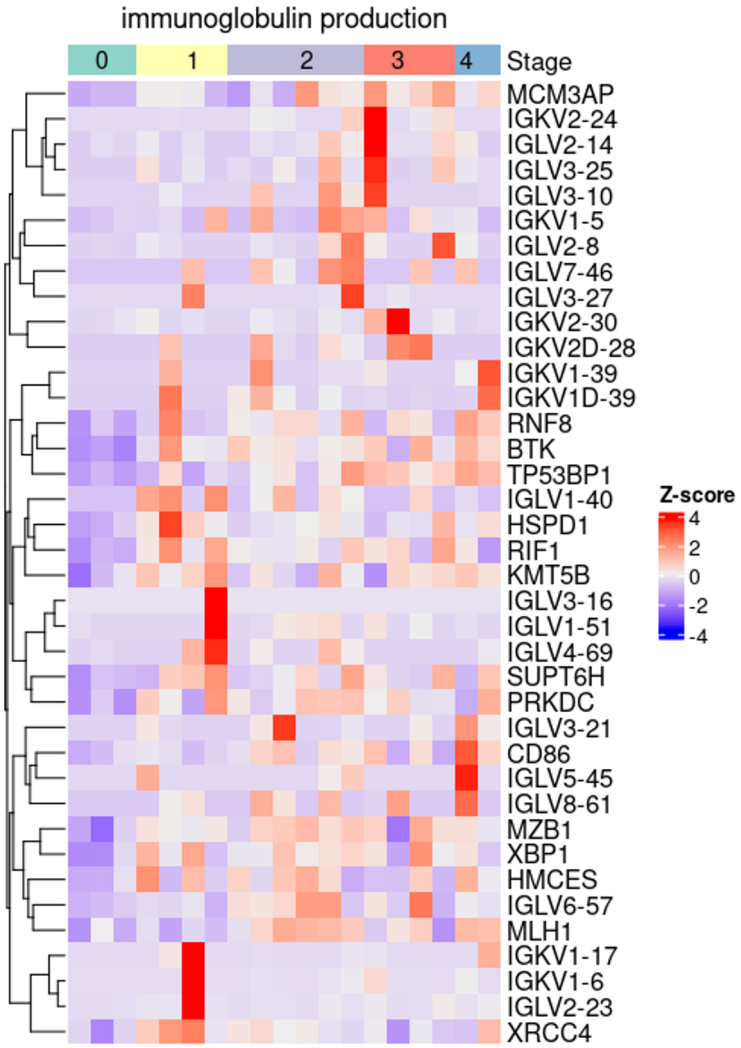


Figure 3-Supplement 1. Heatmap showing B cell's 38 DDEGs identified from immunoglobulin production process. Color represents normalized z-score of gene expression.

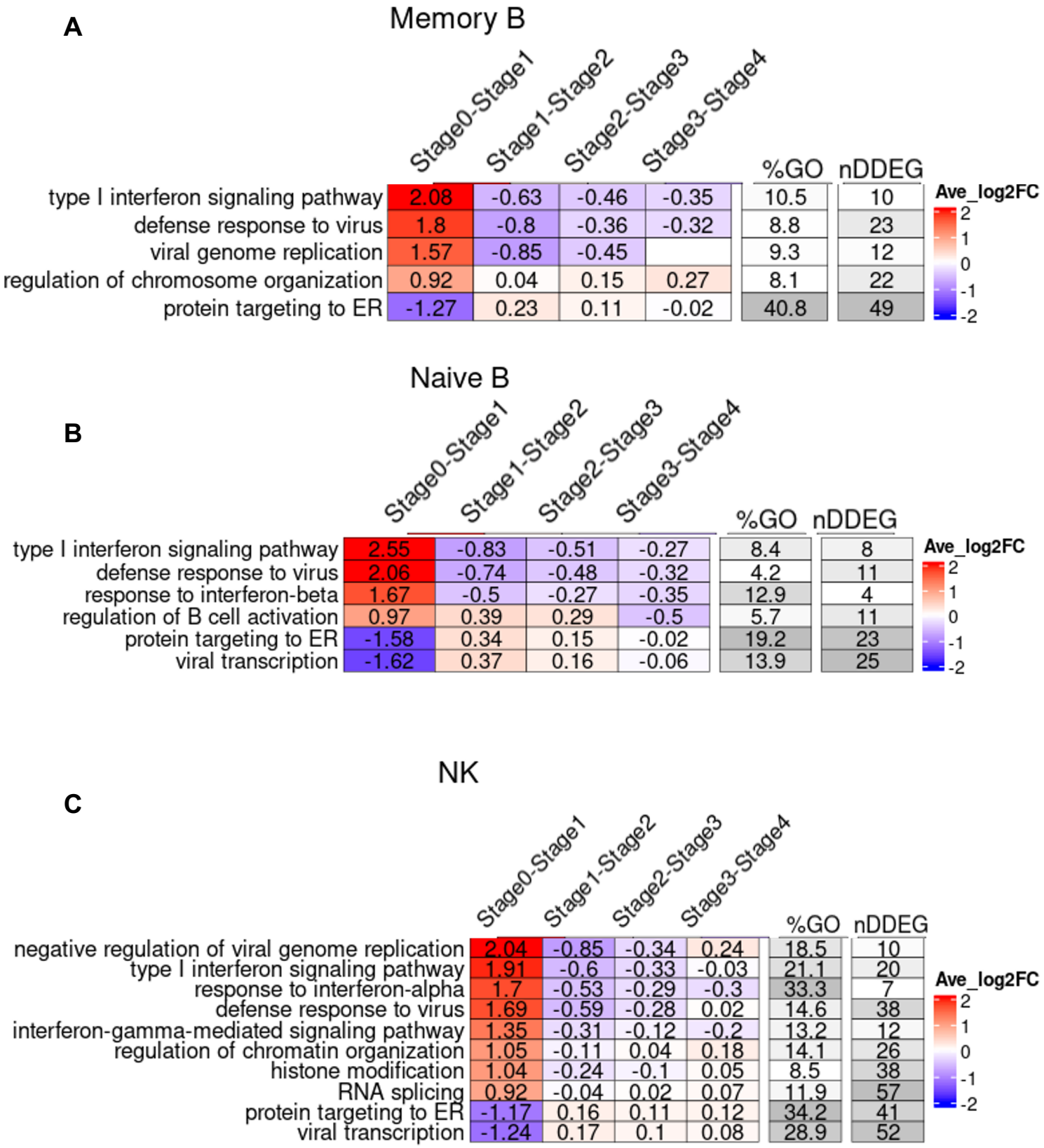


Figure 3-Supplement 2. (A) *TimeHeatmap* of Memory B cells. (B) *TimeHeatmap* of Naive B cells. (C) *TimeHeatmap* of NK cells.

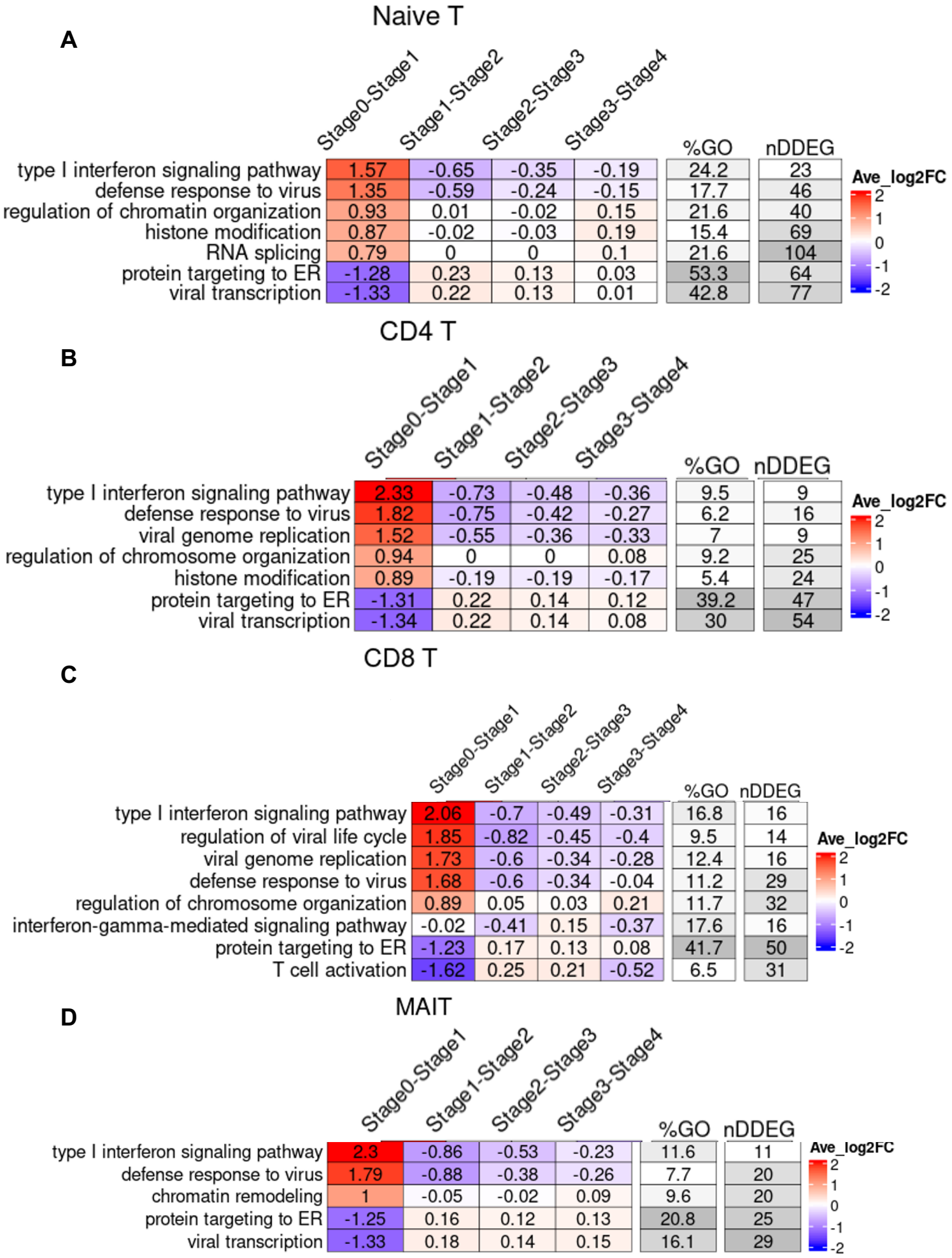
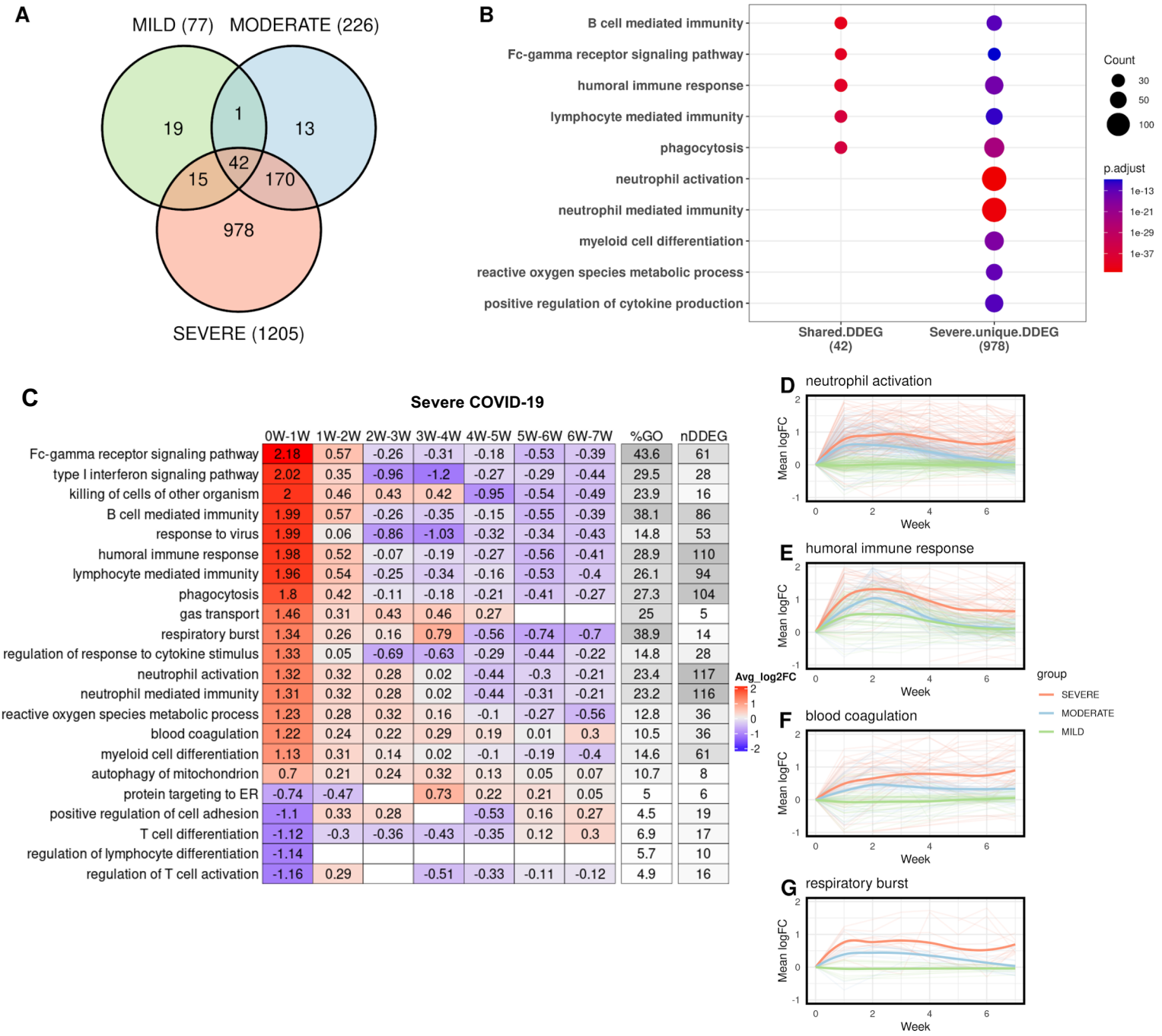


Figure 3-Supplement 3. (A) *TimeHeatmap* of Naive T cells. (B) *TimeHeatmap* of CD4⁺ T cells. (C) *TimeHeatmap* of CD8⁺ T cells. (D) *TimeHeatmap* of MAIT cells.



	0W-1W	1W-2W	2W-3W	3W-4W	4W-5W	5W-6W	6W-7W	%GO	nDDEG
Fc-gamma receptor signaling pathway	2.18	0.57	-0.26	-0.31	-0.18	-0.53	-0.39	43.6	61
type I interferon signaling pathway	2.02	0.35	-0.96	-1.2	-0.27	-0.29	-0.44	29.5	28
killing of cells of other organism	2	0.46	0.43	0.42	-0.95	-0.54	-0.49	23.9	16
B cell mediated immunity	1.99	0.57	-0.26	-0.35	-0.15	-0.55	-0.39	38.1	86
response to virus	1.99	0.06	-0.86	-1.03	-0.32	-0.34	-0.43	14.8	53
humoral immune response	1.98	0.52	-0.07	-0.19	-0.27	-0.56	-0.41	28.9	110
lymphocyte mediated immunity	1.96	0.54	-0.25	-0.34	-0.16	-0.53	-0.4	26.1	94
phagocytosis	1.8	0.42	-0.11	-0.18	-0.21	-0.41	-0.27	27.3	104
gas transport	1.46	0.31	0.43	0.46	0.27			25	5
respiratory burst	1.34	0.26	0.16	0.79	-0.56	-0.74	-0.7	38.9	14
regulation of response to cytokine stimulus	1.33	0.05	-0.69	-0.63	-0.29	-0.44	-0.22	14.8	28
neutrophil activation	1.32	0.32	0.28	0.02	-0.44	-0.3	-0.21	23.4	117
neutrophil mediated immunity	1.31	0.32	0.28	0.02	-0.44	-0.31	-0.21	23.2	116
reactive oxygen species metabolic process	1.23	0.28	0.32	0.16	-0.1	-0.27	-0.56	12.8	36
blood coagulation	1.22	0.24	0.22	0.29	0.19	0.01	0.3	10.5	36
myeloid cell differentiation	1.13	0.31	0.14	0.02	-0.1	-0.19	-0.4	14.6	61
autophagy of mitochondrion	0.7	0.21	0.24	0.32	0.13	0.05	0.07	10.7	8
protein targeting to ER	-0.74	-0.47		0.73	0.22	0.21	0.05	5	6
positive regulation of cell adhesion	-1.1	0.33	0.28		-0.53	0.16	0.27	4.5	19
T cell differentiation	-1.12	-0.3	-0.36	-0.43	-0.35	0.12	0.3	6.9	17
regulation of lymphocyte differentiation	-1.14							5.7	10
regulation of T cell activation	-1.16	0.29		-0.51	-0.33	-0.11	-0.12	4.9	16

Figure 4. Temporal analysis of whole blood RNA-seq data in patients grouped according to disease severity. (A) Venn diagram of DDEGs identified from three COVID-19 severity groups. (B) Top GO enrichment from shared DDEGs across three groups compared with top GO enrichment from severe group. Dot size represents number of DDEGs enriched from the corresponding biological pathway. (C) *TimeHeatmap* of the top dynamic pathways from the severe group. Each column represents a time window, with W depicting week. (D-E) LOESS curve fitting of DDEGs identified in the severe COVID-19 group of the (D) neutrophil activation pathway, (E) humoral immune response pathway, (F) blood coagulation pathway and (G) respiratory burst pathway. Red curves represent the severe group, blue curves represent the moderate group, and green curves represent the mild group.

A

Mild COVID-19

	0W-1W	1W-2W	2W-3W	3W-4W	4W-5W	5W-6W	6W-7W	%GO	nDDEG
Fc-gamma receptor signaling pathway	1.71	0.49	-0.03	-0.74	-0.6	-0.31	-0.25	18.6	26
B cell mediated immunity	1.68	0.6	-0.06	-0.73	-0.59	-0.3	-0.25	13.3	30
lymphocyte mediated immunity	1.68	0.6	-0.06	-0.73	-0.59	-0.3	-0.25	8.3	30
phagocytosis	1.68	0.6	-0.21	-0.65	-0.59	-0.3	-0.25	7.9	30
humoral immune response	1.63	0.56	0.25	-0.85	-0.58	-0.29	-0.24	8.4	32
protein targeting to ER			0.73	-1.82	1		0.53	4.2	5

Avg_log2FC: 2, 1, 0, -1, -2

B

Moderate COVID-19

	0W-1W	1W-2W	2W-3W	3W-4W	4W-5W	5W-6W	6W-7W	%GO	nDDEG
type I interferon signaling pathway	2.34	-0.63	-0.57	-0.52	-0.47	-0.42	-0.4	26.3	25
humoral immune response	1.53	1.14	-0.87	-0.99	-0.51	-0.27	-0.28	15.5	59
lymphocyte mediated immunity	1.48	1.29	-0.86	-1.06	-0.54	-0.24	-0.22	13.9	50
B cell mediated immunity	1.47	1.32	-0.87	-1.09	-0.54	-0.23	-0.22	20.8	47
Fc-gamma receptor signaling pathway	1.43	1.42	-0.89	-1.12	-0.54	-0.21	-0.2	28.6	40
phagocytosis	1.39	1.23	-0.87	-1.04	-0.52	-0.23	-0.2	13.4	51

Avg_log2FC: 4, 2, 0, -2, -4

Figure 4-Supplement 1. (A) *TimeHeatmap* of top dynamic GO terms found in severe COVID-19 shown in mild group. (B) *TimeHeatmap* of top dynamic GO terms found in severe COVID-19 shown in moderate group.

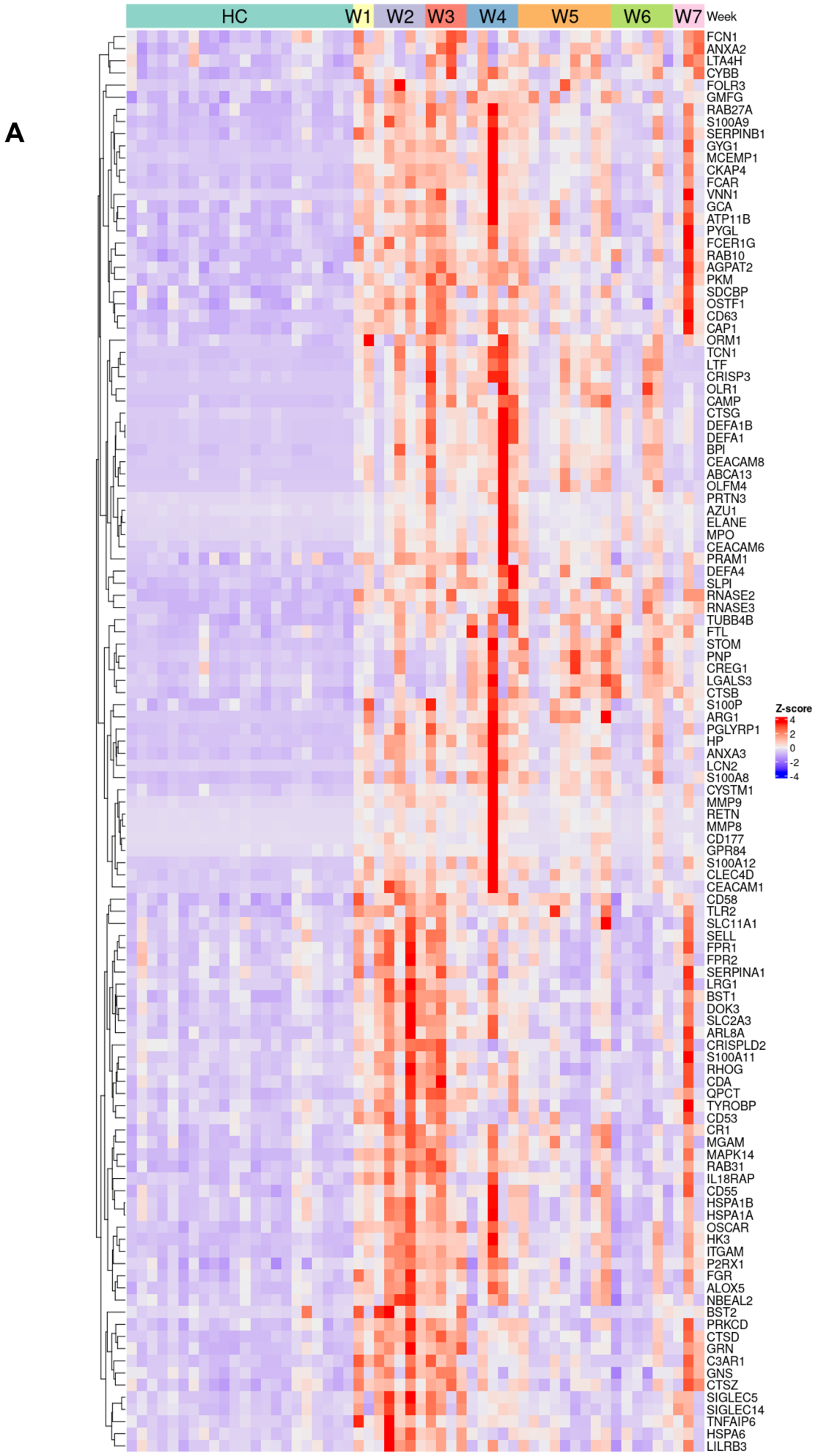


Figure 4-Supplement 2. (A) Heatmap of severe COVID-19 DDEGs identified from the neutrophil activation pathway. The heatmap is ordered by week following infection with healthy controls (HC) on the left. Color represents Z-score normalized expression values. Each column represents one sample.

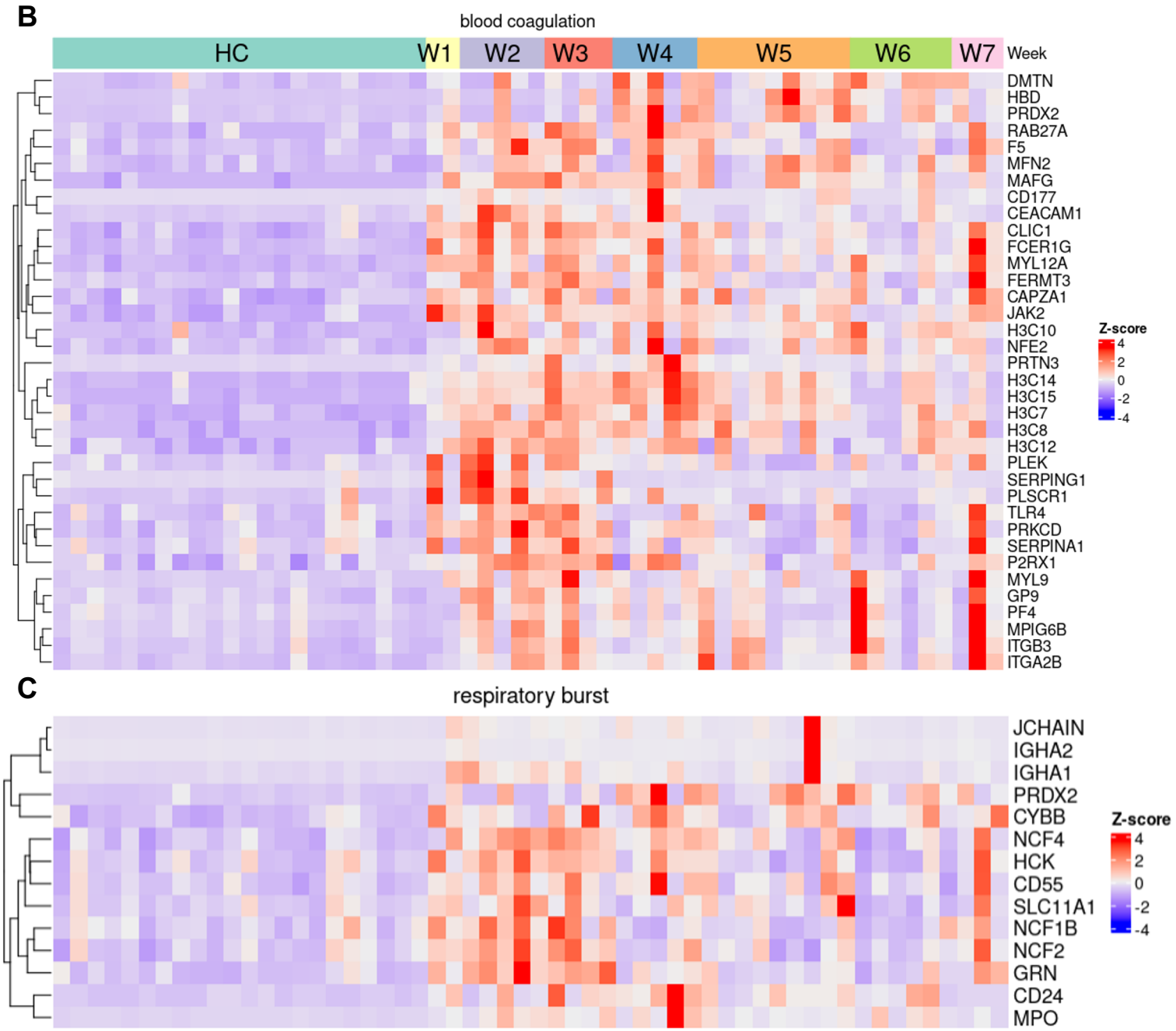


Figure 4-Supplement 2. (B) Severe COVID-19 DDEGs identified from the blood coagulation pathway. (C) Severe COVID-19 DDEGs identified from respiratory burst pathway. The heatmap is ordered by week following infection with healthy controls (HC) on the left. Color represents Z-score normalized expression values.

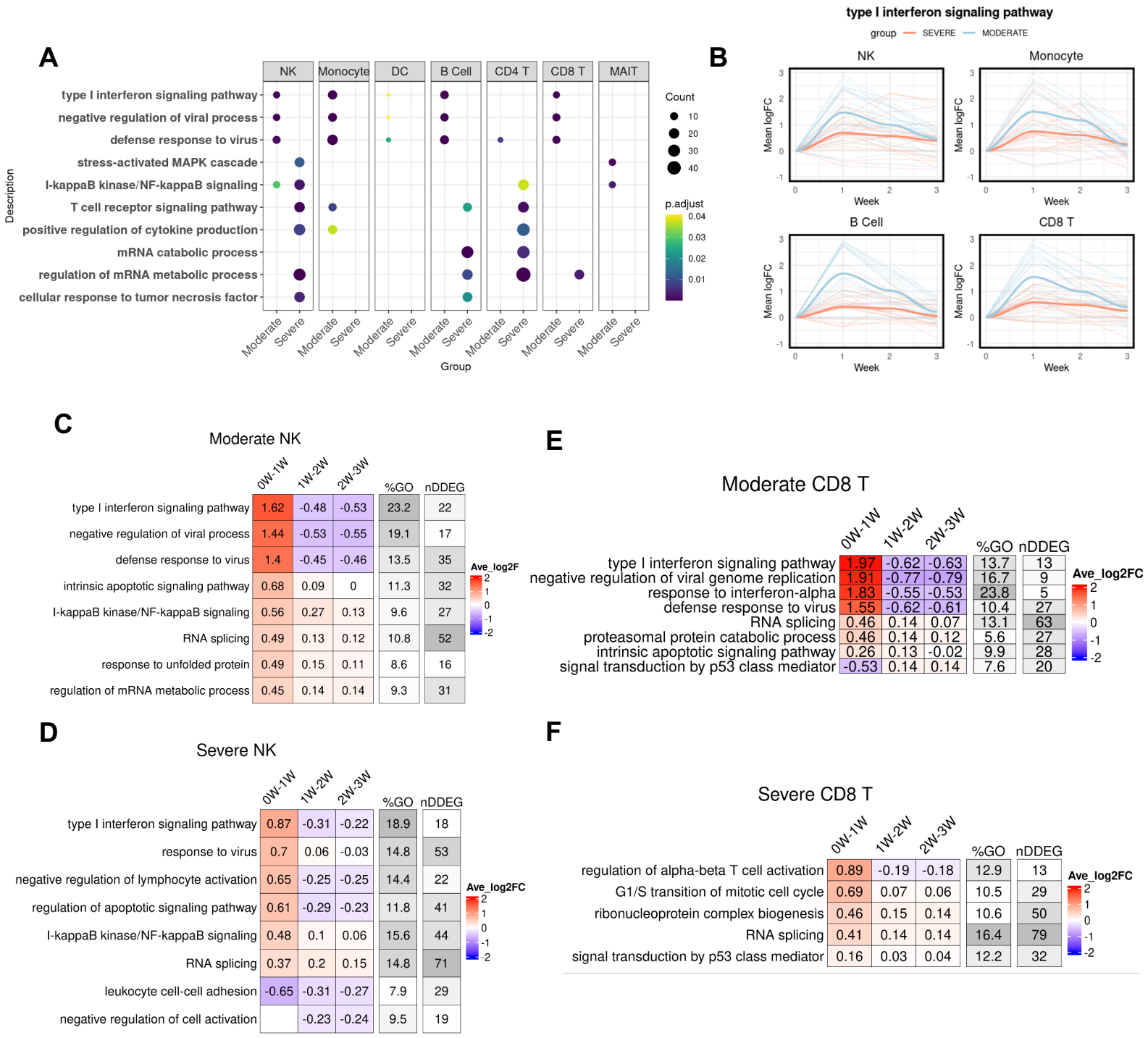


Figure 5. Temporal analysis of single cell RNA-seq data of PBMCs from patients with either moderate or severe COVID-19. (A) Dot plot showing GO enrichment comparison between severe COVID-19 and moderate COVID-19 for each cell type. The size of the dot represents gene count. (B) LOESS curve fitting on DDEGs identified from the type I interferon pathway using TrendCatcher from moderate COVID-19 and severe COVID-19. Blue color indicates moderate group and red color indicates severe group. (C) *TimeHeatmap* of NK cells from moderate COVID-19. (D) *TimeHeatmap* of NK cells from severe COVID-19. (E) *TimeHeatmap* of $CD8^+$T cells from moderate COVID-19. (F) *TimeHeatmap* of $CD8^+$T cells from severe COVID-19. Each column represents a time window with W depicting week.

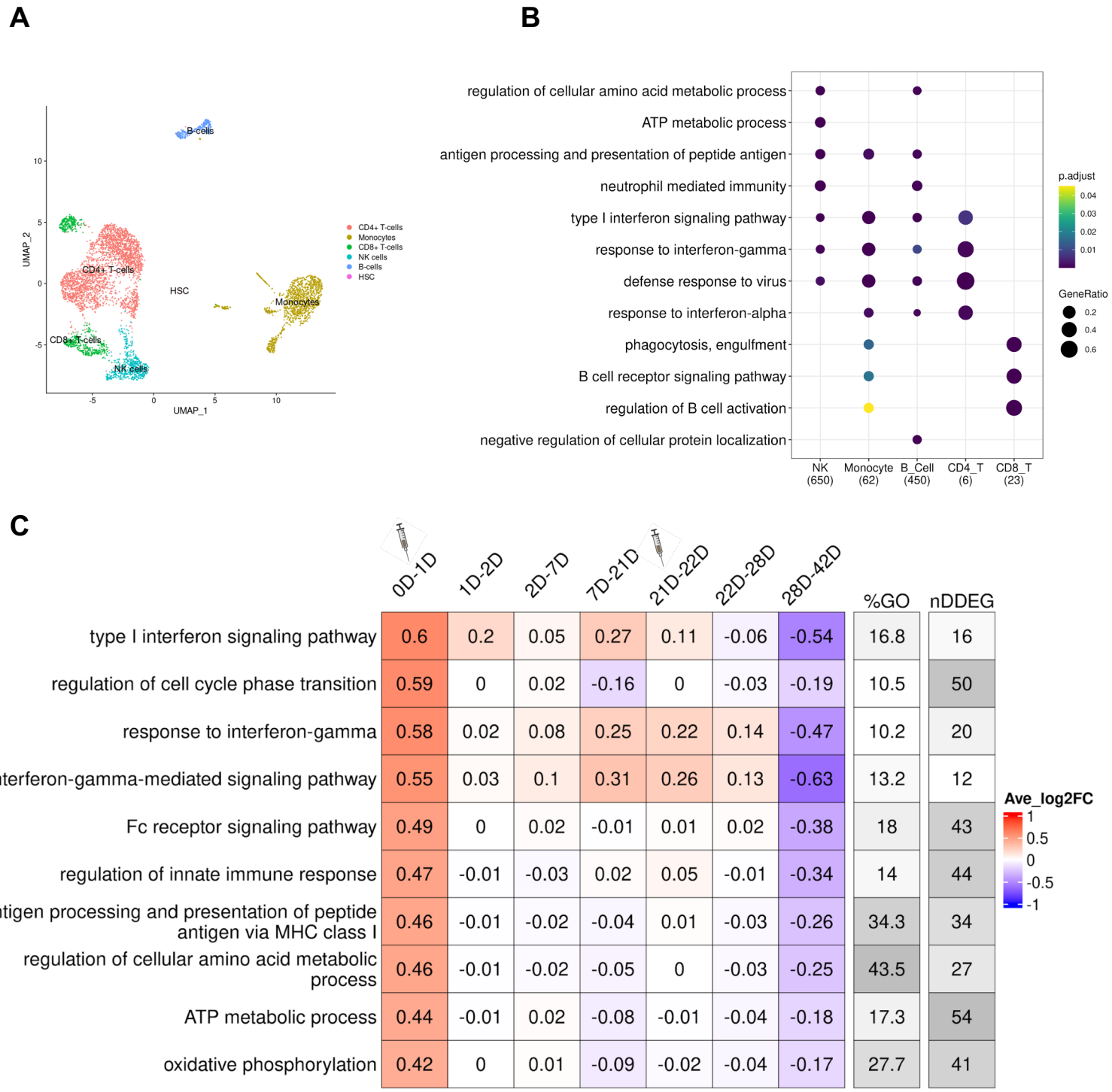


Figure 6. Temporal analysis of PBMC single cell RNA-seq data from human subjects receiving the SARS-CoV-2 mRNA vaccine. (A) UMAP of the single-cell transcriptional profile of patient 2047 at Day 0. Cell types were auto-annotated by SingleR. (B) Dot plot of comparison of the top GO terms enriched from cell-type specific DDEGs. *TrendCatcher* identified 650, 62, 450, 6 and 23 DDEGs from NK cells, monocytes, B cells, $CD4^+$ T cells and $CD8^+$ T cells. Size of the dot represents gene ratio from the enrichment analysis corresponding to each cell type. (C) T*imeHeatmap* of NK cells. Each column represents a time window measured by days. The second dose was administered at Day 21.

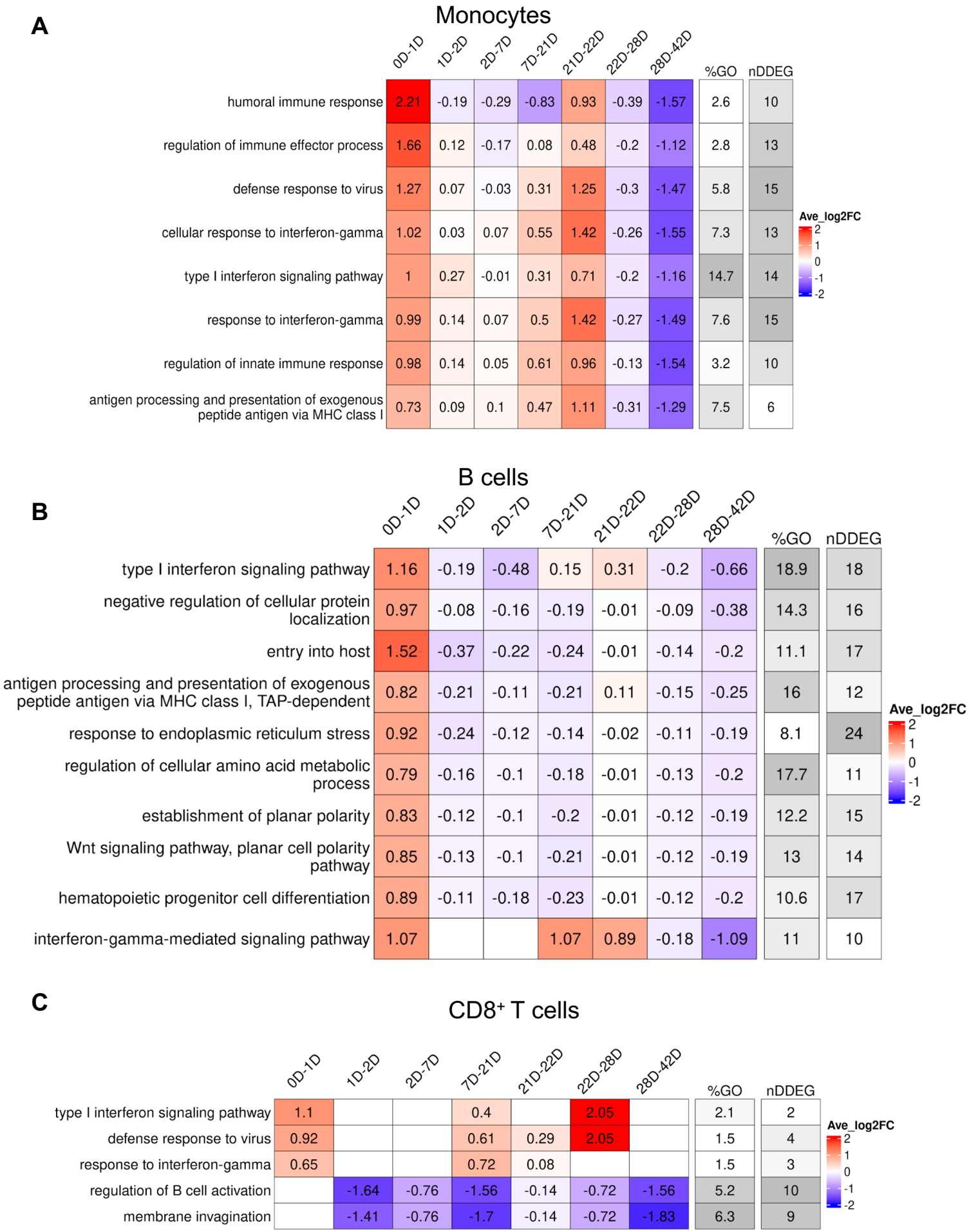


Figure 6-Supplement 1. (A) *TimeHeatmap* of monocytes. (B) *TimeHeatmap* of B cells. (C) *TimeHeatmap* of CD8⁺T cells. Day 1 is the first dose of vaccination, Day 21 is the second dose of vaccination.