

1   **Additional analyses exploring the hypothesized transdifferentiation of plasmablasts to**  
2   **developing neutrophils in severe COVID-19**

4   Aaron J. Wilk<sup>1,2\*</sup>, Arjun Rustagi<sup>3\*</sup>, Nancy Q. Zhao<sup>2\*</sup>, Beth A. Martin<sup>3</sup>, Angela J. Rogers<sup>3\*\*</sup>,  
5   Catherine A. Blish<sup>1,3,4\*\*</sup>

6   Affiliations

8   <sup>1</sup>Stanford Medical Scientist Training Program, <sup>2</sup>Stanford Immunology Program, <sup>3</sup>Department of  
9   Medicine, Stanford University School of Medicine, Stanford, CA 94305; <sup>4</sup>Chan Zuckerberg  
10   Biohub, San Francisco, CA

11   \*equal contributions

12   \*\*Corresponding authors:

14   Angela J. Rogers, MD  
15   Department of Medicine, Division of Pulmonary and Critical Care  
16   Stanford University School of Medicine  
17   [ajrogers@stanford.edu](mailto:ajrogers@stanford.edu)

19   Catherine A. Blish, MD, PhD  
20   Department of Medicine, Division of Infectious Diseases and Geographic Medicine  
21   Stanford University School of Medicine  
22   [cblish@stanford.edu](mailto:cblish@stanford.edu)

23   **Abstract**

25   We thank Alquicira-Hernandez et al. for their reanalysis of our single-cell transcriptomic dataset  
26   profiling peripheral immune responses to severe COVID-19. We agree that careful analysis of  
27   single-cell sequencing data is important for generating cogent hypotheses but find several  
28   aspects of their criticism of our analysis to be problematic. Here we respond briefly to  
29   misunderstandings and inaccuracies in their commentary that may have led to misinformed  
30   interpretation of our results.

31   **Main**

33   Alquicira-Hernandez et al.<sup>1</sup> question the plausibility of the potential lineage relationships  
34   between plasmablasts and developing neutrophils that we postulated as a part of our recent  
35   work<sup>2</sup>. We appreciate their commentary and concur that careful computational analysis of  
36   single-cell RNA sequencing (scRNA-seq) data is necessary. Our study, the first to publicly share  
37   scRNA-seq data to profile immunity in COVID-19, was by its design and execution descriptive,  
38   correlative, and hypothesis-generating, given the limitations of the dataset acknowledged in our  
39   original manuscript. Our goal was to develop a resource for the scientific community to better  
40   understand COVID-19, and to identify distinctive immune features for further study. We regret  
41   that we may have not adequately conveyed the hypothesis-generating nature of our study; if

42 any reader came away with the impression that we had claimed to have “proven” the existence  
43 of a plasmablast to neutrophil transition, this was not our intent, and we apologize.  
44

45 We chose our words very carefully when describing our findings, explicitly choosing not to say  
46 that we had “proved” or “concluded” new hypotheses with scRNA-seq data alone, particularly as  
47 it related to a potential transdifferentiation pathway. In this regard, we point to our original  
48 manuscript rather than Alquicira-Hernandez et al.’s paraphrasing, which left out important  
49 context for our stated conclusions. To wit, our final statement on this putative pathway reads,  
50 “Collectively, we observe a developing neutrophil population that may be characteristic of ARDS  
51 in severe COVID-19 infection; our data suggest that these cells may derive from plasmablasts,  
52 but they may also represent developing neutrophils derived from emergency granulopoiesis”<sup>2</sup>.  
53

54 Alquicira-Hernandez et al. argue that transdifferentiation between plasmablasts and developing  
55 neutrophils is biologically implausible and, therefore, that the association between these two cell  
56 types in uniform manifold approximation and projection (UMAP) space must represent an  
57 artifact of the computational pipeline we selected. Alquicira-Hernandez et al. thus re-analyzed  
58 our data with different preprocessing recipes to see if the phenotypic association between  
59 plasmablasts and developing neutrophils would break, implying that the relationship between  
60 these cells would be artifactual.  
61

62 Our response to this argument is, as follows:  
63

64 1) Alquicira-Hernandez et al. assert that we drew our hypothesis solely from the proximity of  
65 plasmablasts and developing neutrophils in non-linear dimensionality reduction space, which is  
66 incorrect. Our hypothesis was based primarily on a cellular trajectory analysis by RNA velocity  
67 (Figure 4 of our original manuscript). This orthogonal computational technique uses the kinetics  
68 of RNA splicing to calculate a time derivative of gene expression, which can computationally  
69 infer the trajectory of cellular differentiation<sup>3,4</sup>. Based on this analysis, we believe our hypothesis  
70 is plausible because we observed sequential downregulation of genes associated with  
71 plasmablasts and upregulation of genes associated with neutrophil development across the  
72 inferred latent time trajectory. This was coincident with upregulation of C/EBP transcription  
73 factors known to drive neutrophil development, and consistent with the pattern observed  
74 previously during B cell to macrophage transdifferentiation<sup>5-9</sup>. Thus, the results of RNA velocity-  
75 based analysis led us to postulate that the phenotypic relationship between plasmablasts and  
76 developing neutrophils could represent transition between the two cell types.  
77

78 2) Alquicira-Hernandez et al.’s argument takes two inherently contradictory positions on the  
79 sufficiency of our dataset to prove, or disprove, a plasmablast-to-neutrophil transition in COVID-  
80 19. Given that we have stated that our data is insufficient to make such a firm conclusion, we do  
81 not find that Alquicira-Hernandez et al. could definitively disprove this putative transition using  
82 the same data alone.  
83

84 3) This argument overextends the interpretability of scRNA-seq data by overemphasizing the  
85 role of parameter tuning in preprocessing. While we agree that careful selection of

86 preprocessing parameters is an essential component of scRNA-seq data analysis, there are  
87 nonetheless a plethora of reasonable ways to analyze and visualize a scRNA-seq dataset. For  
88 many of these parameters, there is no widely-accepted formal method of determining what is  
89 the “best” to use for each dataset. For example, Alquicira-Hernandez et al. argue that we may  
90 have overfit our data by regressing the number of unique molecular identifiers (UMIs) and genes  
91 detected per cell, but the impact of such potential overfitting is likely to be inconsequential given  
92 the extremely high ratio of variables to covariates.  
93

94 To explore this last idea further and extend the reanalysis performed by Alquicira-Hernandez et  
95 al., we have conducted a second analysis of our dataset using 24 different combinations of  
96 covariates and two other parameters not discussed by Alquicira-Hernandez et al. These  
97 additional parameters are the number of highly-variable genes (nHVG) used for transformation  
98 and dimensionality reduction, and the number of principal components (nPC) used for  
99 dimensionality reduction. This second analysis reveals that, in the vast majority of combinations,  
100 there is still a phenotypic association between plasmablasts and developing neutrophils (**Figure 1a**). This is similar to the UMAP projections generated by Alquicira-Hernandez et al., which still  
101 show developing neutrophils closely associated with plasmablasts, with several plasmablasts  
102 embedding closer to developing neutrophils than with other plasmablasts (**Figure 1a**, purple  
103 box). To examine this phenotypic relationship outside of the non-linear dimensionality reduction  
104 manifold space of UMAP, we additionally hierarchically clustered pseudobulk average gene  
105 expression profiles of each cell type in our dataset, which again indicates that plasmablasts and  
106 developing neutrophils are phenotypically related (**Figure 1b**). It is important to note that this  
107 does not indicate relationship in terms of cell lineage, but merely relationship of transcriptional  
108 phenotype. Taken together, these analyses re-affirm our previous decision to explore this  
109 relationship further with RNA velocity analysis.  
110

111 We performed RNA velocity analysis using preprocessing parameters employed in our original  
112 manuscript (**Figure 1c**, green box) and using preprocessing parameters used by Alquicira-  
113 Hernandez et al. (**Figure 1d**, purple box). To analyze the dynamic relationship between  
114 additional cell types that may be biologically related to plasmablasts and developing neutrophils,  
115 we embedded only plasmablasts, developing neutrophils, B cells, and a population of low-  
116 density mature neutrophils we identified. We found that, with both sets of preprocessing  
117 parameters, developing neutrophils appear to transition from plasmablasts and do not occupy  
118 similar UMAP manifold space as B cells and mature neutrophils (**Figure 1c, d**).  
119

120 Alquicira-Hernandez et al. hypothesized that plasmablasts should be more closely related to B  
121 cells than developing neutrophils, and that developing neutrophils should be phenotypically  
122 associated with mature neutrophils. Upon finding that an embedding of these four cell types  
123 alone showed a relative lack in relatedness between plasmablasts and B cells, the authors  
124 concluded that plasmablasts and developing neutrophils must be misclassified as related cell  
125 types. However, it is incorrect to assume B cells and plasmablasts should be phenotypically  
126 related in UMAP space, as these cell types are dramatically different in terms of gene module  
127 expression (eg. proliferation) that is easily detected at the transcriptional level (**Figure 1b**), and  
128 because the kinetics of B cell-to-plasmablast differentiation in these patients may not enable  
129

130 identification of intermediate cell states in the periphery. While it does remain possible that  
131 developing neutrophils and plasmablasts are related in UMAP space because they are both  
132 proliferative cell types, there are other proliferative T and NK cells in the dataset that are not  
133 phenotypically related and this argument does not have bearing on trajectories predicted by  
134 RNA velocity. We thus conclude that our selection of preprocessing parameters was reasonable  
135 and would have led to the same hypotheses had we chosen different parameters.

136

137 Alquicira-Hernandez et al. also imply that we did not fully consider the possibility that  
138 hemophagocytic lymphohistiocytosis (HLH) could have explained our findings because of the  
139 difficulty in making this diagnosis. First, we would like to reiterate that such an explanation  
140 would be expected to result in an increase in complexity (# genes detected per # UMI  
141 sequenced per cell) of developing neutrophils, which we did not observe (Extended Data Figure  
142 9 of our original manuscript<sup>2</sup>). While it is possible that internalization of one cell into another,  
143 either by emperipoleisis or hemophagocytosis, could confound our interpretation, these  
144 mechanisms of intact cell ingestions are exceptionally rare behaviors of both neutrophils and B  
145 lymphocytes at any differentiation stage, and none of the published series of peripheral smears  
146 in COVID-19 have revealed this phenomenon<sup>10–12</sup>. In addition, none of the patients in our cohort  
147 had clinical characteristics of HLH and none received granulocyte colony-stimulating factor (G-  
148 CSF) or granulocyte-macrophage colony-stimulating factor (GM-CSF) therapy on the original  
149 authors' review. To probe the possibility of HLH in our cohort more deeply, we requested an  
150 independent review of all available clinical data (including demographics, ethnicity, laboratory  
151 and imaging data, physical examinations, and clinical treatment plans) from our cohort by a  
152 hematologist and expert in adult HLH (Dr. Beth A. Martin). Standard scoring (HLH-94, H score)  
153 and expert assessment were utilized. The available clinical data was sufficient to conclude that  
154 no patient had HLH. Of note, no patient met criteria for bone marrow biopsy or other tissue  
155 evaluation for the presence of hemophagocytosis.

156

157 In conclusion, we would like to thank Alquicira-Hernandez et al. for their commentary and  
158 reanalysis of our work. While it remains possible that the phenotypic association and predicted  
159 trajectory dynamics between plasmablasts and developing neutrophils is an incidental finding,  
160 the additional analysis presented here indicates this finding is not an artifact of our analytical  
161 pipeline. We believe that a plasmblast-to-neutrophil transdifferentiation in severe COVID-19  
162 remains an intriguing and plausible hypothesis, one that we are working to validate through  
163 isolation of the correct cell population and DNA sequencing of the BCR loci to conclusively  
164 determine the developmental origins of these cells. Until we are able to generate such data, we  
165 would like to reiterate that our previous manuscript is exploratory, observational, and does not  
166 claim to have demonstrated the veracity of this transition. Finally, we would agree with the  
167 authors' title with a slight modification, that there is “**No direct, mechanistic** evidence that  
168 plasmablasts transdifferentiate into developing neutrophils in severe COVID-19”: our study was  
169 not designed to find this evidence.

170

171

172 **CONTRIBUTIONS**

173 AJW performed bioinformatic analyses; BAM performed independent review of clinical data;  
174 AJW, AR, NQZ, AJR, and CAB wrote the manuscript.

175

176 **COMPETING INTERESTS**

177 The authors declare no competing interests.

178

179 **DATA AVAILABILITY**

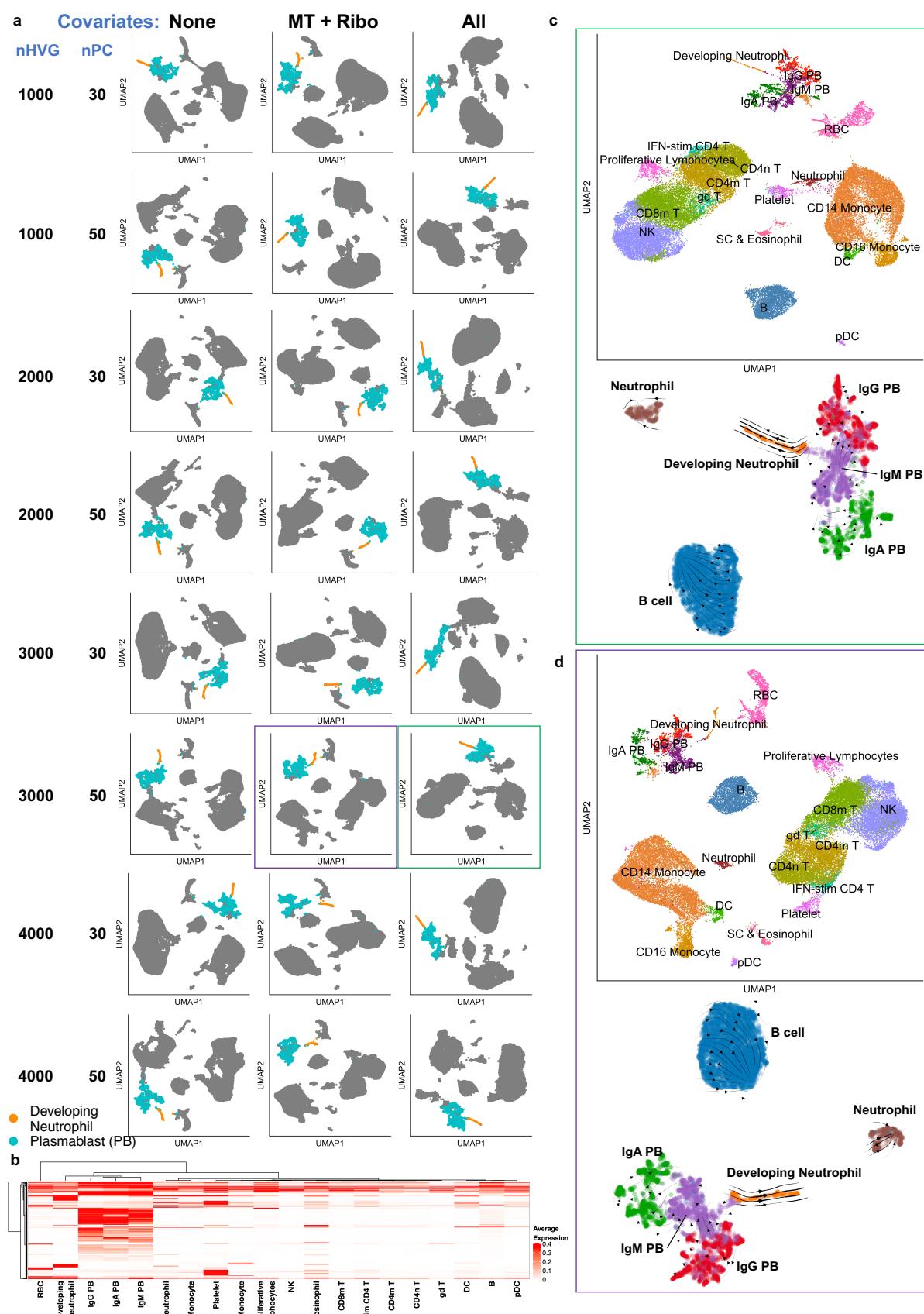
180 All data analyzed in this manuscript are publicly available, and relevant accessions and web  
181 links are provided in the original publication.

182 REFERENCES

- 183 1. Alquicira-Hernandez, J., Powell, J. E. & Phan, T. G. No evidence that plasmablasts  
184 transdifferentiate into developing neutrophils in severe COVID-19 disease.  
185 2020.09.27.312538 (2020) doi:10.1101/2020.09.27.312538.
- 186 2. Wilk, A. J. *et al.* A single-cell atlas of the peripheral immune response in patients with  
187 severe COVID-19. *Nat. Med.* **26**, 1070–1076 (2020).
- 188 3. La Manno, G. *et al.* RNA velocity of single cells. *Nature* **560**, 494–498 (2018).
- 189 4. Bergen, V., Lange, M., Peidli, S., Wolf, F. A. & Theis, F. J. Generalizing RNA velocity to  
190 transient cell states through dynamical modeling. *Nat. Biotechnol.* (2020)  
191 doi:10.1038/s41587-020-0591-3.
- 192 5. DeKoter, R. P. & Singh, H. Regulation of B lymphocyte and macrophage development by  
193 graded expression of PU. 1. *Science* **288**, 1439–1441 (2000).
- 194 6. Graf, T. Historical origins of transdifferentiation and reprogramming. *Cell Stem Cell* **9**, 504–  
195 516 (2011).
- 196 7. Rapino, F. *et al.* C/EBP $\alpha$  induces highly efficient macrophage transdifferentiation of B  
197 lymphoma and leukemia cell lines and impairs their tumorigenicity. *Cell Rep.* **3**, 1153–1163  
198 (2013).
- 199 8. Di Tullio, A. *et al.* CCAAT/enhancer binding protein  $\alpha$  (C/EBP $\alpha$ )-induced transdifferentiation  
200 of pre-B cells into macrophages involves no overt retrodifferentiation. *Proceedings of the  
201 National Academy of Sciences* **108**, 17016–17021 (2011).
- 202 9. Francesconi, M. *et al.* Single cell RNA-seq identifies the origins of heterogeneity in efficient  
203 cell transdifferentiation and reprogramming. *Elife* **8**, (2019).
- 204 10. Nazarullah, A., Liang, C., Villarreal, A., Higgins, R. A. & Mais, D. D. Peripheral Blood  
205 Examination Findings in SARS-CoV-2 Infection. *Am. J. Clin. Pathol.* **154**, 319–329 (2020).
- 206 11. Lee, C.-T. & Teo, W. Z. Y. Peripheral Blood Smear Demonstration of Lymphocyte Changes  
207 in Severe COVID-19. *Am. J. Trop. Med. Hyg.* (2020) doi:10.4269/ajtmh.20-0721.

208 12. Lüke, F. et al. Coronavirus disease 2019 induces multi-lineage, morphologic changes in  
209 peripheral blood cells. *EJHaem* (2020) doi:10.1002/jha2.44.

210



212 **Figure 1. scRNA-seq data processed with different preprocessing parameters supports**  
213 **the hypothesis of plasmablast-developing neutrophil transdifferentiation. a)** UMAP  
214 embeddings generated with 24 different combinations of regressed covariates, number of  
215 highly-variable genes (nHVG), and number of principal components for dimensionality reduction  
216 (nPC). “None”, no covariates regressed; “MT + Ribo”, only percentage of mitochondrial and  
217 ribosomal reads regressed; “All”, percentage of mitochondrial and ribosomal reads, as well as  
218 number of UMIs and number of genes detected per cell regressed. Only developing neutrophils  
219 and plasmablasts are colored; all other cell types are gray. “PB”, plasmablast. **b)** Hierarchically-  
220 clustered pseudobulk average expression profiles of the top 250 HVG for each cell type. **c)**  
221 UMAP dimensionality reduction projection of full dataset generated using original preprocessing  
222 parameters colored by cell type (top; green box corresponds to embedding shown in panel **(a)**).  
223 UMAP embedding of plasmablasts, developing neutrophils, B cells, and mature neutrophils  
224 generated using original preprocessing parameters overlaid with RNA velocity stream (bottom).  
225 **d)** Same plots as described in **(c)**, using preprocessing parameters used by Alquicira-  
226 Hernandez et al. (percentage of mitochondrial and ribosomal reads regressed, 3000 highly-  
227 variable genes, 50 principal components).