Basal Expression of Interferon-Stimulated Genes Drives Population Differences in Monocyte Susceptibility to Influenza Infection

Mary O'Neill[1], Hélène Quach[2], Julien Pothlichet[3], Yann Aquino[1], Aurélie Bisiaux[1], Nora Zidane[4], Matthieu Deschamps[1], Valentina Libri[5], Milena Hasan[5], Shen-Ying Zhang[6,7,8], Qian Zhang[6], Daniela Matuozzo[7,8], Aurélie Cobat[7,8], Laurent Abel[6,7,8], Jean-Laurent Casanova[6,7,8,9,10], Nadia Naffakh[11], Maxime Rotival[1,13,*], Lluis Quintana-Murci[1,12,13,14,*]

[1]Human Evolutionary Genetics Unit, Institut Pasteur, UMR 2000, CNRS, Paris 75015, France
[2]Muséum National d'Histoire Naturelle, UMR7206, CNRS, Université de Paris, Paris 75016, France
[3]DIACCURATE, Paris, 75014, France
[4]Biodiversity and Epidemiology of Bacterial Pathogens Unit, Institut Pasteur, Paris 75015, France
[5]Cytometry and Biomarkers UTechS, Institut Pasteur, Paris 75015, France
[6]St. Giles Laboratory of Human Genetics of Infectious Diseases, Rockefeller Branch, The Rockefeller University, New York, NY, 10065, USA
[7]Laboratory of Human Genetics of Infectious Diseases, Necker Branch, Necker Hospital for Sick Children, INSERM UMR 1163, Necker Hospital for Sick Children, Paris 75015, France
[8]Paris University, Imagine Institute, Paris 75015, France
[9] Pediatric Hematology-Immunology Unit, Necker Hospital for Sick Children, Paris, France
[10] Howard Hughes Medical Institute, New York, NY, USA
[11]RNA Biology of Influenza Virus Unit, Institut Pasteur, Paris, 75015, France
[12]Chair of Human Genomics and Evolution, Collège de France, Paris, 75005, France

[13]M.R. and L.Q.M. contributed equally to this work.
[14]Lead Contact.
[*]To whom correspondence may be addressed. Email: maxime.rotival@pasteur.fr or quintana@pasteur.fr

SUMMARY

There is considerable inter-individual immunological and clinical variability upon influenza A virus (IAV) infection in humans; yet, the factors underlying such heterogeneity remain elusive. Here, using an *ex vivo* cellular model that captures natural human variation in the transcriptional responses of monocytes to IAV, we find significant differences in viral mRNA levels between individuals of African and European ancestry. Using single cell analyses, we show that the overall number of cells that will ultimately become infected, rather than the amount of viral transcript expression per cell, is the main driver of the higher IAV mRNA levels detected in European cells. Finally, we identify 135 genes, including the interferon-stimulated genes *IFITM3*, *MX1*, and *OAS3*, for which basal mRNA expression levels associate with IAV mRNA levels post infection. Our findings reveal that basal differences in activation of IRF/STAT-induced antiviral mechanisms may contribute to individual and population disparities in susceptibility to IAV infection.

INTRODUCTION

Influenza A viruses (IAVs) are responsible for seasonal epidemics and occasional pandemics in humans; annually, flu is responsible for 3-5 million cases of severe illness and up to 650,000 deaths, being therefore a major public health concern (WHO, 2020). The severity and clinical outcomes of influenza infection are driven by a variety of intrinsic and extrinsic factors. While the viral determinants of disease are increasingly well described, the contribution of host factors is less clear (Clohisey and Baillie, 2019; Fukuyama and Kawaoka, 2011; Gounder and Boon, 2019; Horby et al., 2012; Horby et al., 2013; Schrauwen et al., 2014). Individuals with suboptimal immunity – pregnant women, the elderly and the very young, and those with pre-existing medical conditions – have long been appreciated as having elevated risk of developing severe disease (CDC, 2020; WHO, 2020). Single-gene inborn errors of *TLR3*- and *IRF7*-dependent type I and II interferon (IFN) immunity can also account for severe influenza (Ciancanelli et al., 2015; Hernandez et al., 2018; Lim et al., 2019), but common genetic variants contributing to susceptibility have not been convincingly described. Yet, in each flu season, particularly during pandemics, even individuals from demographics least at risk succumb to severe influenza, highlighting considerable inter-individual variation among otherwise healthy individuals (Beigel et al., 2005; Louie et al., 2009). Such severe cases are often associated with increased levels of inflammatory cytokines and chemokines (i.e., cytokine storm) (Gao et al., 2013; Liu et al., 2016).

During the course of infection, IAV must first breach physical barriers in the respiratory tract (e.g. mucus) before binding to sialic acid receptors on airway epithelial cells – the principal site of IAV replication. Infection of the epithelium triggers innate immune responses that result in the production of cytokines that induce antiviral gene transcription among neighboring cells, and recruit immune cells to the site of infection (Lamichhane and Samarasinghe, 2019; Vangeti et al., 2018). Monocytes are among the early milieu of recruited cells, as attested by their increased numbers in the peripheral blood and nasal mucosae of IAV-infected patients (Giamarellos-Bourboulis et al., 2009; Gill et al., 2008; Gill et al., 2005; McClain et al., 2013; Oshansky et al., 2014). They can be divided into different subsets based on surface receptor expression of CD14 and CD16: classical ($CD14^{++}CD16^{-}$), nonclassical ($CD14^{+}CD16^{++}$), and intermediate ($CD14^{++}CD16^{+}$) (Geissmann et al., 2003; Ziegler-Heitbrock et al., 2010). Whereas classical monocytes prevail in the blood of healthy individuals, intermediate monocytes dominate in their airways (Baharom et al., 2016). Upon being recruited, monocytes secrete inflammatory cytokines and chemokines, and differentiate into macrophages and dendritic cells with phagocytic properties (Kratofil et al., 2017; Lamichhane and Samarasinghe, 2019). Despite these protective roles, monocytes are themselves susceptible to IAV infection (Hoeve et al.,

2012), and their excessive infiltration can induce immunopathology (Brandes et al., 2013; Cole et al., 2017; Lin et al., 2014). Specifically, intermediate monocytes are over-represented in blood samples from severe influenza cases with no previously known risk factor (Cole et al., 2017), suggesting they have a particular role in disease etiology.

The recent advent of single-cell approaches has provided a powerful tool to explore immune cell heterogeneity (Papalexi and Satija, 2018). In the context of influenza, single-cell RNA-sequencing (scRNA-seq) allows for the simultaneous capture and characterization of both host and viral transcriptomes as IAV mRNAs are polyadenylated. Using this technology, extreme variation in viral replication has been reported for epithelial cell lines (Russell et al., 2018), a phenomenon that has been replicated in other systems including animal models and patient samples (Cao et al., 2020; Kudo et al., 2019; Ramos et al., 2019; Russell et al., 2019; Russell et al., 2018; Steuerman et al., 2018; Sun et al., 2020; Wang et al., 2020). Importantly, single-cell analyses can distinguish between bystander (exposed but uninfected) and infected cells, allowing the characterization of antiviral states. Although viral contributions to variation in cellular responses to infection have begun to be probed with this technology (Russell et al., 2019; Sun et al., 2020), how host responses vary between individuals and populations remains unexplored.

Here, we report the first comprehensive single-cell study of natural human variation in monocytes' response to influenza infection. To do so, we leveraged the data generated by the EVOIMMUNOPOP project, which has characterized the genetic architecture of transcriptional responses of primary monocytes to bacterial and viral stimuli in 200 individuals of African and European ancestry (Quach et al., 2016). Specifically, we have reported marked differences in transcriptional responses to IAV between the two populations, which are in part accounted for by host genetic factors. Herein, we document the surprising finding that primary monocytes of individuals of European ancestry are more susceptible to IAV infection than those from individuals of African descent. By applying scRNA-seq to the mononuclear phagocytes (i.e. monocytes or their derived cell-types) of individuals of both ancestries who were found to be at either extreme of this cellular phenotype (resistant to susceptible), we reveal extensive heterogeneity in viral transcription/degradation among cells challenged with IAV, even within an individual donor. Yet, we report evidence that the overall number of cells that become infected, rather than the amount of viral transcript expression per cell, drives the observed differences at the population level. Finally, we document the role of the type I IFN response in cellular susceptibility to infection, and importantly, establish that basal expression levels of IFN-stimulated genes correlate with the percentage of IAV-expressing cells following influenza infection.

RESULTS

A Mononuclear Phagocyte Model of Immune Response Variability to IAV Infection

To characterize population-level variation in IAV molecular phenotypes, we examined bulk RNA-sequencing (RNA-seq) data collected from the $CD14^+$ fraction of peripheral blood mononuclear cells (PBMCs), from 200 males residing in Belgium of African (AFB) or European (EUB) descent (100 donors of each ancestry), following *ex vivo* challenge with a live strain of IAV for 6 hours (Quach et al., 2016). After mapping the data to a combined human-IAV reference, and excluding 1 sample (EUB) with low quality RNAs, we found that 0.02-13.5% of RNA-seq reads from each sample were of viral origin (**Figure 1A**). This indicates marked inter-individual variation in the ability of the IAV strain to infect and/or replicate in mononuclear phagocytes. Such heterogeneity was found to be highly reproducible, as attested by the strong correlation between viral mRNA levels from replicate infection experiments in cells from six individuals, three of each ancestry, who span the phenotypic spectrum (Pearson r = 0.96, *p*-value = 1.3×10^{-3}). Interestingly, AFB and EUB samples presented overlapping but significantly shifted distributions of total IAV-mapping reads (4.9% vs. 6.8% of reads, respectively, Wilcoxon *p*-value = 5.3×10^{-8}), and of each of the 10 primary viral transcripts (**Figure 1B**, Wilcoxon *p*-value < 5.5×10^{-4}), suggesting a role of ancestry-related host factors in driving viral mRNA levels. However, no genome-wide significant hits were detected when treating viral mRNA levels as a quantitative trait among ~7M Single Nucleotide Polymorphisms (genotyping/exome sequencing), dismissing the likelihood of large-effect, common genetic variants explaining the phenotype. These analyses suggest that polygenic and/or nongenetic factors drive the strong variability in the IAV molecular phenotypes observed between the populations of different ancestry.

Marked Cell-to-Cell Variation in Infection Potential

We next investigated whether cellular heterogeneity may help explain inter-individual differences in IAV mRNA levels, as measured by the percentage of IAV-mapping reads in the bulk RNA-seq. Focusing again on the mononuclear phagocytes from six individuals (three of each ancestry) who span the phenotypic spectrum, we found that only a fraction of each donors' cells harbored the IAV viral nucleoprotein (NP) at 6 hours post infection (hpi), illustrating probable differences in cell-to-cell infection potential (**Figure 1C**). The proportion of NP-expressing cells (defined by protein fluorescence > 2.35) from these 6 donors was highly correlated with IAV mRNA levels (r^2=0.82, *p*-value = 7.74×10^{-3}). Importantly, AFB and EUB individuals differed with respect to proportions of canonical monocyte subsets - classical ($CD14^{++}CD16^-$), intermediate ($CD14^{++}CD16^+$), and nonclassical ($CD14^+CD16^{++}$) - defined via flow cytometry at basal state

(prior to IAV-challenge), which make up the majority of $CD14^+$ cells in PBMCs (**Figures 1D and S1**). We observed that the basal proportion of intermediate monocytes, but not of non-classical or classical, was negatively correlated with IAV mRNA levels measured 6hpi (Pearson r = -0.30, *p*-value = 2.0×10^{-4}), and explained ~8% of differences in IAV mRNA levels between populations. Our results suggest that within an individual's circulating monocyte population, there exists cell-to-cell variation in infection potential, and between donors, the proportion of monocytes that become infected likely drives the observed patterns of population variability in viral mRNA levels.

Single-Cell RNA-Sequencing Mirrors Inter-Individual Variability in Immune Phenotypes

To investigate more precisely the role of cellular heterogeneity in driving population immune variability, we performed a time-course experiment, both in the presence and absence of IAV, on the $CD14^+$ fraction of PBMCs for a subset of eight donors. These individuals were selected on the basis of ancestry and their associated IAV mRNA levels in the bulk RNA-seq experiment (**Figure 1A**). We deliberately chose individuals whose cells previously demonstrated varying levels of *ex vivo* susceptibility, from more resistant (lower than average IAV mRNA levels; 2 of each ancestry) to more susceptible (higher than average IAV mRNA levels; 2 of each ancestry). Non-stimulated (control) and IAV-challenge conditions were then carried out in parallel on positively-selected monocytes from the eight donors, with droplet-based scRNA-seq being performed immediately before infection initiation (T_0), and at 2 (T_2), 4 (T_4), 6 (T_6), and 8 (T_8) hours post infection initiation. After removal of doublets and low-quality droplets (**Figures S2A-C**), we characterized 96,386 high-quality single cells (Lun et al., 2016a; Lun et al., 2016b; McCarthy et al., 2017). As additional quality control, we clustered cells based on their proximity in a *k*-nearest-neighbor (*knn*) graph (Xu and Su, 2015), and used SingleR to predict cell types based on transcriptional profiles (Aran et al., 2019). We found that cells were divided into 21 communities (**Figure S2D**), and, as expected, the majority were of myeloid origin (>98%). Negligible fractions (<2%) of other cell types were also detected and were mostly assigned to four clusters (clusters 2, 17, 19, 20) (**Figure S2E**), revealing minor contaminants of the monocyte isolation. Four clusters, making up <7% of the cells, displayed hallmarks of dying cells such as high mitochondrial content and lower UMI counts (clusters 3, 8, 10, 11; **Figure S2F**). We excluded these contaminant- and dying-cell clusters, resulting in a final dataset of 88,559 cells.

At T_0, >99% of cells were predicted to be monocytes, which we distinguish into classical, intermediate, and nonclassical subsets based on their mRNA expression profiles (**Supplementary Note; Tables S1 and S2; Figures S3 and S4**). We reproduced, in the eight

donors, the negative association between the estimated proportion of intermediate monocytes prior to infection and their viral mRNA levels post infection (Pearson r = -0.83, p-value = 0.01). At later time-points, however, a substantial fraction of cells (18% in total) were predicted to be macrophages, indicating monocyte differentiation over the course of the experiment. We thus distinguished cells from T_2-T_8 into two subsets, *CD16*$^-$ and *CD16*$^+$, based on *FCGR3A* (i.e. *CD16*) mRNA expression (**Supplementary Note; Tables S1 and S2; Figures S3 and S4**). For clarity, we refer to cells as monocytes at T_0 and as mononuclear phagocytes from T_2-T_8.

To explore the extent to which scRNA-seq captures the observed inter-individual variability in viral mRNA levels, we quantified IAV transcripts in the mononuclear phagocytes of each individual (**Figure 2A**). We created pseudo-bulk estimates by averaging the percent of viral mRNAs per-cell across all cells from each donor at each time point (**Figure 2B**). Even as early as T_2, the differences in viral mRNAs levels across donors based on scRNA-seq data strongly correlated with those observed at 6hpi in the bulk RNA-seq study (Quach et al., 2016) (Pearson r = 0.95, *p*-value=3.7×10^{-3}). The correlation was maximized at T_4 when viral transcription was at its peak (Pearson r = 0.97, *p*-value=5.1×10^{-5}), and remained significant at T_6 (Pearson r = 0.95, *p*-value=3.0×10^{-4}) and T_8 (Pearson r = 0.84, *p*-value=8.9×10^{-3}). These analyses collectively indicate that the *ex vivo* susceptibility phenotype (resistant to susceptible) is highly reproducible among individuals, even across different experimental protocols and technologies.

Inter-Individual Variation Is Determined by the Proportion of Infected Cells

To investigate the contribution of cellular heterogeneity to inter-individual differences in viral mRNA levels, we next sought to distinguish cells that were successfully infected by the virus from those that were not, based on their viral mRNA content. Among mononuclear phagocytes exposed to IAV, the proportions of viral mRNAs were bimodally distributed (**Figure 2C**), mirroring observations at the protein level (**Figure 1C**). As mononuclear phagocytes not exposed to virus (T_2-T_8) were deliberately multiplexed with IAV-challenged cells, low amounts of viral mRNAs present in the cellular suspension (i.e. ambient mRNAs) were captured by the cell-containing droplets during encapsulation (Young and Behjati, 2020), leading to 53.1% of these droplets containing viral transcripts. Yet, reassuringly, the distribution of viral mRNAs levels was heavily skewed towards zero among droplets containing non-stimulated cells. We thus used droplets of non-stimulated cells to estimate the maximal contamination fraction of cellular reads contributed by such ambient mRNAs (**Figure 2D).** When allowing up to 10% of a cell's transcriptome to originate from ambient mRNAs, we found that 50.3% of IAV-challenged cells significantly expressed IAV transcripts above ambient levels; using this as a proxy for infection,

we deemed these cells as 'infected'. The remaining cells were considered as 'bystanders', which either did not come into contact with the virus or were able to fully repress the transcription of viral mRNAs.

Using a Gaussian mixture model, we further sub-classified infected cells into two different states: those with lower IAV mRNA levels (<1-6%) and those with higher IAV mRNA levels (6-83%). While viral mRNA levels are dictated by both the rate of transcription and degradation, for simplicity we refer to these infected cell states as 'low IAV-transcribers' and 'high IAV-transcribers', respectively. When comparing the relative proportions of each cell state across individuals and time points, we found that the proportion of high IAV-transcribing cells strongly differed between time points, but displayed only moderate variation across individuals (R^2=63% vs 23%, respectively, Anova *p*-value < 1.5×10^{-3}) (**Figure 2E**). Conversely, the percentage of infected cells (both low and high IAV-transcribing) was stable over time (R^2<1%), but highly variable between individuals (R^2=97%, Anova *p*-value = 1.5×10^{-17}). Moreover, the proportion of infected cells correlated with that of viral mRNA levels recovered 6hpi from the independent bulk RNA-seq study (Quach et al., 2016) (Pearson r = 0.94, *p*-value = 4.0×10^{-4}). These results show that infected cells can be divided into two distinct subpopulations that differ with respect to levels of viral mRNAs at later time points and, importantly, that inter-individual variability in viral mRNA levels can be explained by the percentage of cells that become infected.

Cellular Susceptibility to Infection Drives Population Differences in Response to IAV

To extrapolate these findings to the population level, we used the transcriptional profiles obtained from the scRNA-seq data at T_6 to estimate the proportions of reads coming from each inferred cell state among the bulk RNA-seq profiles of EVOIMMUNOPOP samples (Quach et al., 2016). We first assessed the quality of our deconvolution method by cross-validation, comparing across our eight donors the estimated cell fractions obtained from deconvolution of the bulk RNA-seq profiles with the true proportions assessed through scRNA-seq (**Figure S5A**). We reliably captured inter-individual variation in the percentage of reads that originate from bystander (Pearson r = 0.97, *p*-value=$4.2\text{x}10^{-5}$) and infected cells (Pearson r = 0.98, *p*-value=$3.7\text{x}10^{-5}$), as well as, to a lesser extent, from dying/dead cells (Pearson r = 0.74, *p*-value=$3.5\text{x}10^{-2}$). Applying the method to all 199 samples passing QC, we found that the percentage of reads from infected cells (both low and high IAV-transcribers) ranged from 0 to 70%, with this percentage being between 38 and 52% in more than half of the donors (**Figure 3A**). Our deconvolution method further allowed the separation of infected cells into low and high IAV-transcribers, although it was slightly less accurate at doing so (Pearson r = 0.80 and 0.73,

respectively, *p*-values < 0.05), and revealed that the percentage of high IAV-transcribers among infected cells was stable across populations (**Figure 3B**).

We found that differences in the estimated percentage of infected cells alone could explain 63% of the inter-individual variability in viral mRNA levels (**Figure 3C**), while variation in the percentage of high/low transcribers among infected cells accounted for an additional 19% of variance in viral mRNA expression. Interestingly, variation in high/low transcribers had the strongest impact on viral mRNAs in EUB individuals (Pearson r = 0.66 and 0.09, for EUB and AFB, respectively, interaction *p*-value = 9.5×10^{-8}, **Figure S5B**). This could be explained by the overall higher rate of infection that we observed in EUB relative to AFB (+9.6 % of infected cells, Wilcoxon *p*-value= 5.2×10^{-10}), together with higher cell death (+5% of dying cells, Wilcoxon *p*-value = 2.6×10^{-10}). Nonetheless, differences in the proportions of infected cells were sufficient to account for the observed difference in viral mRNA levels between AFB and EUB individuals (*p*-value=0.16 after adjusting on infected cells, compared to $p<5.3\times10^{-8}$ without adjustment). Taken together, these results show that while variation in the proportion of high IAV-transcribing infected cells has a sizeable impact on viral mRNA levels, population differences in viral mRNA levels are primarily driven by the overall proportion of cells that will ultimately become infected.

Interplay of Cytokine and Ribosome Networks Drive Cell States upon Infection

Given the potential role of monocyte heterogeneity in influenza severity (Cole et al., 2017), and the distinctive transcriptional profiles of *CD16*$^-$ and *CD16*$^+$ mononuclear phagocyte subsets in our scRNA-seq data (**Supplementary Note; Tables S1 and S2; Figures S3 and S4**), we next combined this subset classification with the cell states inferred above (unexposed, bystander, and low and high-IAV transcribing infected cells) to characterize cellular responses to infection over time (**Figure 4A**). To account for differences in sample sizes, we subsampled each subset (*CD16*$^-$/*CD16*$^+$), cell state, and time point in our scRNA-seq data to 100 cells (corresponding to the number of cells in the least abundant category), while ensuring representation of all donors. We then focused on the 6,669 host genes with average $\log_2$ normalized count >0.1 in at least one subgroup. We found that, regardless of the condition and mononuclear phagocyte subset, changes in gene expression increased over time but quickly plateaued (**Figure 4B, Table S3A**). While similar absolute levels of gene expression changes were observed among all cells at T_2, we detected marked changes among cells exposed to IAV (bystander and infected) relative to those unexposed at later time points. Correlation analysis revealed a substantial divergence, from T_4 onwards, in transcriptional responses of cells exposed to IAV versus unexposed. The transcriptome of each cell state remained stable over the rest of the experiment; between 73

and 90% of genes up-regulated at T_4 in a given cell state (relative to T_0, FDR<1% & $\log_2$FC >0.2) remained so at later time points (**Figure 4C**).

CD16$^-$ and *CD16*$^+$ subsets behaved similarly upon stimulation, with changes in gene expression between cell states being strongly correlated among subsets (Pearson r = 0.83-0.95, *p*-values < 2.2×10^{-16}, **Figures S6A-D)**. GO term enrichment analyses of shared responses (FDR<1% & $\log_2$FC >0.2 in same direction in both subsets) uncovered several functional categories interacting to shape the activation state of cells (**Figures 4D and 4E**; **Table S3B**). Both bystander and infected cells showed increased mRNA expression of genes involved in antigen processing and presentation via class I MHC (GO:0019885, OR=53.7, FDR < 2.0×10^{-6}) and the response to type I IFN (GO:0034340, OR >14.8, FDR < 3.3×10^{-20}). Yet, bystander cells showed increased mRNA expression of type I IFN response and defense response to virus pathways relative to infected cells (GO:0034340: OR=13.4, FDR=4.4×10^{-7}; GO:0051607: OR=9.0, FDR=2.1×10^{-7}), while infected cells displayed higher mRNA expression of mitochondrial (GO:0005743, OR=4.7, FDR=3.3×10^{-3}) and ribosomal genes (GO: 0005840, OR=117, FDR=1.0×10^{-78}) (**Figure 4D and 4F**).

Furthermore, among infected cells, ribosomal genes showed higher activity among high IAV-transcribing cells relative to low IAV-transcribing cells (**Figures 4E and 4F**, comparison only made at T_4 due to sample size constraints, e.g. GO:0019083: OR=137, FDR=6.1×10^{-65}). This observation is consistent with the notion that the expression of viral proteins is completely dependent on cellular ribosomes, with recent data suggesting that IAVs do not induce a global shut-off of cellular translation but rather a reshaping of the translation landscape (Bercovich-Kinori et al., 2016; Li, 2019; Machkovech et al., 2019). Likewise, among bystander cells, numerous ribosomal genes were downregulated at later time points relative to unexposed cells (**Figure 4D and 4F**, GO:0019083, OR=5.3, FDR=4.2×10^{-6}), suggesting that repression of ribosomal subunits plays an active role in limiting viral replication. Collectively these results suggest that IFN signaling and ribosomal expression interact to shape cell states upon IAV-challenge.

Increased IRFs and STATs Activity Drives Stronger Antiviral Response of *CD16*$^+$ Monocytes upon IAV Infection

Despite the similar transcriptional responses of *CD16*$^-$ and *CD16*$^+$ subsets to *ex vivo* stimulation, *CD16*$^+$ cells were less likely to be infected by IAV than *CD16*$^-$ cells (42.3% sd: 4.0% for *CD16*$^+$ relative to 49.4% sd: 5.4% for *CD16*$^-$, generalized linear model with CD16$^{+/-}$ status, donor, and time point as covariates, *p*-value=0.006). This may be attributable to the higher expression of

type I IFN-stimulated genes in *CD16*$^+$ cells, which we observe both in the presence and absence of IAV (**Supplementary Note; Tables S2A and S2B**), but could also result from subset-specific responses to IAV. To investigate this, we performed an interaction test on the subsampled scRNA-seq data, and searched for genes whose transcriptional response to IAV-challenge differed between *CD16*$^-$ and *CD16*$^+$ subsets in either infected and/or bystander cells (**Figures S6A and S6B**; **Table S3A**). At FDR≤1%, we identified a total of 335 such genes, of which 98 differed between subsets only in bystander cells, 144 only in infected cells, and 93 in both. Hierarchical clustering highlighted eight major patterns of transcriptional responses (modules) among the 335 genes, several of which were associated with specific biological functions (**Figure 5A; Tables S3C and S3D**). Of particular interest, module 1 (green) was enriched for genes in the antiviral response pathway (GO:0051607, OR=23.2, FDR<5.43×10^{-7}) and displayed a stronger response in infected *CD16*$^+$ cells relative to *CD16*$^-$ infected cells. Of additional interest was the transient *CD16*$^+$ specific transcription of the inflammatory cytokine genes *IL6* and *TNF* after stimulation by IAV (**Figure 5B**), consistent with the observation that genes involved in the regulation and production of IL-6 and TNFα are over-expressed in *CD16*$^+$ subsets between T_2 and T_8, both in the presence and absence of IAV (**Table S2B**). These results reveal the strong antiviral and inflammatory potential of *CD16*$^+$ relative to *CD16*$^-$ mononuclear phagocytes in response to viral infection (Cros et al., 2010).

We next sought to characterize the regulatory architecture underlying the 335 genes whose transcriptional response to IAV-challenge differed between monocyte subsets. Using SCENIC (Aibar et al., 2017), we identified 113 high-confidence gene regulatory networks, or 'regulons', that are active in non-stimulated and/or IAV-challenged cells, each composed of a transcription factor (TF) and a set of predicted targets (genes). We then used these 113 regulons to search for an enrichment/depletion of TF targets among the eight modules of genes displaying subset-specific response to infection (**Table S3E**). Among those modules associated with an increased expression in cells exposed to virus (modules 1-5), we observed a widespread over-representation of targets of IFN regulatory factors (IRFs) and signal transducing and activators of transcription (STATs) (**Figure 5C**), highlighting the central role of the IFN response upon IAV challenge. Interestingly, several of these factors displayed subset-specific activity themselves in response to IAV (*IRF1/2/7* and *STAT1/2/3*, FDR<1%), mirroring the expression patterns of module 1 (Pearson r >0.92). These results highlight mononuclear phagocyte subset differences in the strength of antiviral responses upon IAV infection, and reveal the contribution of IRF and STAT TFs to subset-specific responses upon infection.

Basal Differences in Activation of Antiviral Responses Correlate with IAV Levels

Finally, we turned our attention from the transcriptional effects of IAV-challenge to the causal determinants of such outcomes. To identify specific genes that might underlie infection potential, we focused on the 4,589 genes that were expressed at > 0.1 $\log_2$ normalized counts in at least one of the canonical monocyte subsets defined at T_0 (**Supplementary Note; Table S4A**). We identified a total of 3,136 genes that differed among our eight donors in either classical, intermediate, and/or nonclassical monocyte subsets (Kruskal-Wallis Rank Test, FDR 1%). Within each subset, focusing on genes that significantly differ between donors, we searched for those for which expression at basal state was nominally (*p*-value < 0.01) correlated with viral mRNA levels among our eight donors. We found that, depending on the monocyte subset, between 2.5 to 6.8% of genes matched these criteria, resulting in a set of 135 genes displaying correlation with IAV in at least one subset (**Figure 6A**). These genes were collectively enriched for several related biological processes such as defense response to virus (GO:0051607, OR=8.4, FDR=6.8×10^{-7}) and type I IFN signaling (GO:0060337, OR=10.7, FDR=1.0×10^{-4}) (**Table S4B**).

Overlap with the TF targets identified by SCENIC revealed strong enrichments of several IRFs and STATs among the 135 genes, including IRF7 and STAT1, STAT2 and IRF9 that form the tripartite IFN-stimulated gene factor 3 (ISGF3) (**Figures 6B and 6C**; **Table S4C**). Interestingly, among the candidate genes identified, we found the well-known host viral restriction factors encoded by *IFITM3*, *OAS3*, and *MX1*, as well as critical TFs (*IRF7*, *IRF9*) involved in the severity of IAV-infection both in mice and humans (Allen et al., 2017; Ciancanelli et al., 2016; Zhang, 2020). Notably, *IFITM3* expression was higher in nonclassical and intermediate monocytes relative to classical monocytes, and presented higher expression in *ex vivo* resistant versus susceptible individuals across all subsets (**Figure 6D**). Together, our results provide evidence that the basal mRNA expression of genes related to IFN-induced and antiviral responses are indicative of the proportion of cells that will become infected in the first cycle of IAV infection.

DISCUSSION

The role monocytes play in initiating and coordinating the immune response against viral infections is increasingly appreciated, as well as the pathogenic effects of their excessive recruitment (Lamichhane and Samarasinghe, 2019; Stegelmeier et al., 2019). Here, we use $CD14^+$ positively-selected primary monocytes as a model to explore their specific contribution to influenza infection. Such an *ex vivo* system using healthy donors allows us to normalize for viral and environmental determinants of disease and focus specifically on host-related factors. The combination of transcriptomic and cytometric analyses in a large group of individuals of African and European ancestry, together with ad hoc single-cell analyses, has increased our understanding of the cellular factors driving population variation in responses to IAV at several levels.

At the single-cell level, we document extreme heterogeneity in viral mRNA and protein expression, extending the observation of high cell-to-cell variation in IAV replication to circulating human mononuclear phagocytes (Cao et al., 2020; Kudo et al., 2019; Ramos et al., 2019; Russell et al., 2019; Russell et al., 2018; Steuerman et al., 2018; Sun et al., 2020; Wang et al., 2020). We also show, as documented in other cell types, that bystander cells display increased mRNA expression of IFN-stimulated genes relative to infected cells; yet, we additionally observe both an induction of ribosomal gene mRNA expression in IAV-transcribing cells and a down regulation of these genes in bystander cells at later time points. While the former is likely induced by the virus to enhance mRNA translation (Panthu et al., 2017), the repression of ribosomal expression observed in bystanders may reflect a host mechanism to contain infection by shutting down the translational machinery of neighboring cells. Besides these general patterns, we uncover cellular subset-specific responses to IAV. For example, after 4 hours of stimulation, infected *CD16*$^+$ mononuclear phagocytes display stronger antiviral responses relative to *CD16*$^-$ cells, driven by stronger IRF TFs activity. Moreover, *CD16*$^+$ cells display transient mRNA expression of both *IL6* and *TNF*, the secretion of which has been shown to play a key role in regulating inflammatory responses to IAV in the lung (Damjanovic et al., 2011; Matikainen et al., 2006). These observations collectively suggest a distinct role of *CD16*$^+$ mononuclear phagocytes in regulating the response to IAV infection.

At the individual level, our data reveal extensive variation in the degree of viral transcription and/or degradation across donors. Single cell analyses showed that while mononuclear phagocytes co-exist in all three cell states in each donor (bystander, low IAV-transcribing, and high IAV-transcribing infected cells), individuals strongly differ with respect to the proportion of infected cells. We find that the percentage of cells that become infected upon IAV challenge

negatively correlates with the proportion of intermediate monocytes; yet, this correlation accounts for only a fraction of inter-individual variation in *ex vivo* susceptibility to IAV. Focusing on the transcriptional profiles of individual donors' monocyte populations prior to viral exposure, we found that basal levels of an IRF/STAT-driven antiviral program are a key determinant of the proportion of cells that will ultimately become infected. Further studies are needed to determine whether circulating intermediate monocytes play a causal role in reducing permissiveness to IAV infection *in vivo*, or simply act as a marker of a heightened antiviral state.

At the population level, we uncovered differences in the levels of viral mRNAs between individuals of African and European ancestry, all residing in Belgium, following *ex vivo* IAV challenge. Through deconvolution analysis, we provide evidence that IAV is more efficient at infecting the monocytes of European versus African ancestry individuals on average. Given the importance of basal activation of antiviral defense genes in the outcome of infection, our results suggest an ancestry-related difference in the steady state level of monocytes' antiviral activation. While genetic factors might contribute to such population differences, we did not detect any large-effect loci associated with IAV mRNA levels, but other subtler forms of genetic control, such as the contribution of rare variants or small-effect *cis*- and *trans*-acting loci, remain a possibility. Thus, our observations highlight the need of systems immunology approaches (Davis et al., 2017) in large cohorts where environmental exposures (e.g. past infections) are surveyed in a systematic fashion to quantify the respective contribution of genetic and environmental variables to the observed population differences in cellular susceptibility to IAV infection.

An outstanding question is whether and how the population differences we report here *ex vivo* translate into disparities in severity of influenza infection *in natura*. While some vulnerable individuals may present symptoms with relatively low levels of IAV replication, lower viral replication is generally associated with asymptomatic forms of disease (Huang et al., 2011), suggesting that *in vitro* replication may be indicative of susceptibility to IAV infection. In support of this idea, several of the 135 genes that we find here to be predictive of monocyte permissiveness to IAV infection have been associated with severe forms of influenza in mice and/or humans (e.g. *IFITM3*, *MX1*, *IRF7*, *IRF9*) (Allen et al., 2017; Ciancanelli et al., 2016; Zhang, 2020). This reinforces the notion that some of the cellular determinants of susceptibility to IAV infection that we identify here may translate to other cell types, including epithelial cells.

Altogether, our findings indicate that basal activation of an IRF/STAT driven antiviral program contribute to intra- and inter-population variation in monocytes' responses to IAV, with possible implications for the disparities in disease susceptibility observed both between individuals and populations.

DECLARATION OF INTERESTS

The authors declare no competing interests.

ACKNOWLEDGMENTS

This work was supported by the *Institut Pasteur*, the *Collège de France*, the French Government's Investissement d'Avenir program, *Laboratoires d'Excellence* "Integrative Biology of Emerging Infectious Diseases" (ANR-10- LABX-62-IBEID) and "Milieu Intérieur" (ANR-10-LABX-69-01), the *Fondation de France* (n°00106080), and the *Fondation pour la Recherche Médicale* (Equipe FRM DEQ20180339214). M.B.O was supported by a *European Molecular Biology Organization* long-term fellowship (ALTF 229-2017). We also wish to thank Tzachi Hagai and John Marioni for input and feedback, and Sylvie van Der Werf and Vincent Enouf for providing resources and protocols relating to influenza viruses.

AUTHORS CONTRIBUTIONS

M.B.O, M.R. and L.Q.M. conceived and designed the study. M.B.O., H.Q., J.P., Y.A., A.B., N.Z., and M.D. conducted the experiments at the Human Evolutionary Genetics Unit. M.B.O., M.R., Y.A. and D.M. developed computational methods and performed bioinformatic analyses. J.P., V.L., M.H., S.Z.Y., Q.Z., A.C. L.A., J.L.C. and N.N. provided resources, expertise and feedback. M.R. and L.Q.M supervised the study. L.Q.M secured funding. M.B.O., M.R. and L.Q.M. wrote the manuscript with input from all authors.

REFERENCES

10xGenomics (2020). Chromium Single Cell 3' Reagent Kits User Guide (v3 Chemistry).

Aibar, S., González-Blas, C.B., Moerman, T., Huynh-Thu, V.A., Imrichova, H., Hulselmans, G., Rambow, F., Marine, J.C., Geurts, P., Aerts, J.*, et al.* (2017). SCENIC: single-cell regulatory network inference and clustering. Nat Methods *14*, 1083-1086.

Allen, E.K., Randolph, A.G., Bhangale, T., Dogra, P., Ohlson, M., Oshansky, C.M., Zamora, A.E., Shannon, J.P., Finkelstein, D., Dressen, A.*, et al.* (2017). SNP-mediated disruption of CTCF binding at the *IFITM3* promoter is associated with risk of severe influenza in humans. Nat Med *23*, 975.

Aran, D., Looney, A.P., Liu, L., Wu, E., Fong, V., Hsu, A., Chak, S., Naikawadi, R.P., Wolters, P.J., Abate, A.R.*, et al.* (2019). Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. Nat Immunol *20*, 163-172.

Baharom, F., Thomas, S., Rankin, G., Lepzien, R., Pourazar, J., Behndig, A.F., Ahlm, C., Blomberg, A., and Smed-Sorensen, A. (2016). Dendritic Cells and Monocytes with Distinct Inflammatory Responses Reside in Lung Mucosa of Healthy Humans. J Immunol *196*, 4498-4509.

Becton, Dickinson and Company. (2019) FlowJo™ Software for Mac [software application] Version v10.6.1. Ashland, OR.

Beigel, J.H., Farrar, J., Han, A.M., Hayden, F.G., Hyer, R., de Jong, M.D., Lochindarat, S., Nguyen, T.K., Nguyen, T.H., Tran, T.H.*, et al.* (2005). Avian influenza A (H5N1) infection in humans. N Engl J Med *353*, 1374-1385.

Benaglia, T., Chauveau, D., Hunter, D.R., and Young, D.S. (2009). mixtools: An R Package for Analyzing Mixture Models. J Stat Soft *32*, 1-29.

Bercovich-Kinori, A., Tai, J., Gelbart, I.A., Shitrit, A., Ben-Moshe, S., Drori, Y., Itzkovitz, S., Mandelboim, M., and Stern-Ginossar, N. (2016). A systematic view on influenza induced host shutoff. eLife *5, e18311.*

Brandes, M., Klauschen, F., Kuchen, S., and Germain, R.N. (2013). A systems analysis identifies a feedforward inflammatory circuit leading to lethal influenza infection. Cell *154*, 197-212.

Cao, Y., Guo, Z., Vangala, P., Donnard, E., Liu, P., McDonel, P., Ordovas-Montanes, J., Shalek, A.K., Finberg, R.W., Wang, J.P.*, et al.* (2020). Single-cell analysis of upper airway cells reveals host-viral dynamics in influenza infected adults. bioRxiv, 2020. http://www.biorxiv.org/content/10.1101/2020.04.15.042978v1

Center For Disease Control and Prevention (2020). People at High Risk For Flu Complications (accessed 01/12/2020). https://www.cdc.gov/flu/highrisk/

Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., and Lee, J.J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience *4*, 7.

Ciancanelli, M.J., Abel, L., Zhang, S.-Y., and Casanova, J.-L. (2016). Host genetics of severe influenza: from mouse Mx1 to human IRF7. Curr Opin Immunol *38*, 109-120.

Ciancanelli, M.J., Huang, S.X., Luthra, P., Garner, H., Itan, Y., Volpi, S., Lafaille, F.G., Trouillet, C., Schmolke, M., Albrecht, R.A.*, et al.* (2015). Infectious disease. Life-threatening influenza and impaired interferon amplification in human IRF7 deficiency. Science *348*, 448-453.

Clohisey, S., and Baillie, J.K. (2019). Host susceptibility to severe influenza A virus infection. Crit Care *23*, 303.

Cole, S.L., Dunning, J., Kok, W.L., Benam, K.H., Benlahrech, A., Repapi, E., Martinez, F.O., Drumright, L., Powell, T.J., Bennett, M.*, et al.* (2017). M1-like monocytes are a major immunological determinant of severity in previously healthy adults with life-threatening influenza. JCI Insight *2*, e91868.

Cros, J., Cagnard, N., Woollard, K., Patey, N., Zhang, S.Y., Senechal, B., Puel, A., Biswas, S.K., Moshous, D., Picard, C.*, et al.* (2010). Human CD14dim monocytes patrol and sense nucleic acids and viruses via TLR7 and TLR8 receptors. Immunity *33*, 375-386.

Damjanovic, D., Divangahi, M., Kugathasan, K., Small, C.L., Zganiacz, A., Brown, E.G., Hogaboam, C.M., Gauldie, J., and Xing, Z. (2011). Negative regulation of lung inflammation and immunopathology by TNF-α during acute influenza infection. Am J Pathol *179*, 2963-2976.

Davis, M.M., Tato, C.M., and Furman, D. (2017). Systems immunology: just getting started. Nat Immunol *18*, 725-732.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21.

Fukuyama, S., and Kawaoka, Y. (2011). The pathogenesis of influenza virus infections: the contributions of virus and host factors. Curr Opin Immunol *23*, 481-486.

Gao, R., Bhatnagar, J., Blau, D.M., Greer, P., Rollin, D.C., Denison, A.M., Deleon-Carnes, M., Shieh, W.J., Sambhara, S., Tumpey, T.M.*, et al.* (2013). Cytokine and chemokine profiles in lung tissues from fatal cases of 2009 pandemic influenza A (H1N1): role of the host immune response in pathogenesis. Am J Pathol *183*, 1258-1268.

García-Alcalde, F., Okonechnikov, K., Carbonell, J., Cruz, L.M., Götz, S., Tarazona, S., Dopazo, J., Meyer, T.F., and Conesa, A. (2012). Qualimap: evaluating next-generation sequencing alignment data. Bioinformatics *28*, 2678-2679.

Geissmann, F., Jung, S., and Littman, D.R. (2003). Blood monocytes consist of two principal subsets with distinct migratory properties. Immunity *19*, 71-82.

1000 Genomes Project Consortium (2012). An integrated map of genetic variation from 1,092 human genomes. Nature *491*, 56-65.

Giamarellos-Bourboulis, E.J., Raftogiannis, M., Antonopoulou, A., Baziaka, F., Koutoukas, P., Savva, A., Kanni, T., Georgitsi, M., Pistiki, A., Tsaganos, T.*, et al.* (2009). Effect of the novel influenza A (H1N1) virus in the human immune system. PLoS ONE *4*, e8393.

Gill, M.A., Long, K., Kwon, T., Muniz, L., Mejias, A., Connolly, J., Roy, L., Banchereau, J., and Ramilo, O. (2008). Differential recruitment of dendritic cells and monocytes to respiratory mucosal sites in children with influenza virus or respiratory syncytial virus infection. J Infect Dis *198*, 1667-1676.

Gill, M.A., Palucka, A.K., Barton, T., Ghaffar, F., Jafri, H., Banchereau, J., and Ramilo, O. (2005). Mobilization of plasmacytoid and myeloid dendritic cells to mucosal sites in children with respiratory syncytial virus and other viral respiratory infections. J Infect Dis *191*, 1105-1115.

Gong, T., and Szustakowski, J.D. (2013). DeconRNASeq: a statistical framework for deconvolution of heterogeneous tissue samples based on mRNA-Seq data. Bioinformatics *29*, 1083-1085.

Gounder, A.P., and Boon, A.C.M. (2019). Influenza Pathogenesis: The Effect of Host Factors on Severity of Disease. J Immunol *202*, 341-350.

Hart, T., Komori, H.K., LaMere, S., Podshivalova, K., and Salomon, D.R. (2013). Finding the active genes in deep RNA-seq gene expression studies. BMC Genomics *14*, 778.

Heaton, H., Talman, A.M., Knights, A., Imaz, M., Gaffney, D.J., Durbin, R., Hemberg, M., and Lawniczak, M.K.N. (2020). Souporcell: robust clustering of single-cell RNA-seq data by genotype without reference genotypes. Nature methods *17*, 615-620.

Hernandez, N., Melki, I., Jing, H., Habib, T., Huang, S.S.Y., Danielson, J., Kula, T., Drutman, S., Belkaya, S., Rattina, V.*, et al.* (2018). Life-threatening influenza pneumonitis in a child with inherited IRF9 deficiency. J Exp Med *215*, 2567-2585.

Hoeve, M.A., Nash, A.A., Jackson, D., Randall, R.E., and Dransfield, I. (2012). Influenza virus A infection of human monocyte and macrophage subpopulations reveals increased susceptibility associated with cell differentiation. PLoS ONE *7*, e29443.

Horby, P., Nguyen, N.Y., Dunstan, S.J., and Baillie, J.K. (2012). The Role of Host Genetics in Susceptibility to Influenza: A Systematic Review. PLoS ONE *7*, e33180.

Horby, P., Nguyen, N.Y., Dunstan, S.J., and Kenneth Baillie, J. (2013). An updated systematic review of the role of host genetics in susceptibility to influenza. Influenza Other Respi Viruses *7*, 37-41.

Huang, Y., Zaas, A.K., Rao, A., Dobigeon, N., Woolf, P.J., Veldman, T., Øien, N.C., McClain, M.T., Varkey, J.B., Nicholson, B.*, et al.* (2011). Temporal dynamics of host molecular responses differentiate symptomatic and asymptomatic influenza a infection. PLoS Genet *7*, e1002234.

Johnson, W.E., Li, C., and Rabinovic, A. (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics (Oxford, England) *8*, 118-127.

Kang, H.M., Subramaniam, M., Targ, S., Nguyen, M., Maliskova, L., McCarthy, E., Wan, E., Wong, S., Byrnes, L., Lanata, C.M.*, et al.* (2018). Multiplexed droplet single-cell RNA-sequencing using natural genetic variation. Nat Biotechnol *36*, 89-94.

Kawakami, E., Watanabe, T., Fujii, K., Goto, H., Watanabe, S., Noda, T., and Kawaoka, Y. (2011). Strand-specific real-time RT-PCR for distinguishing influenza vRNA, cRNA, and mRNA. J Vir Methods *173*, 1-6.

Kratofil, R.M., Kubes, P., and Deniset, J.F. (2017). Monocyte Conversion During Inflammation and Injury. Arterioscler Thromb Vasc Biol *37*, 35-42.

Kudo, E., Song, E., Yockey, L.J., Rakib, T., Wong, P.W., Homer, R.J., and Iwasaki, A. (2019). Low ambient humidity impairs barrier function and innate resistance against influenza infection. Proc Natl Acad Sci U S A *116*, 10905-10910.

Lamichhane, P.P., and Samarasinghe, A.E. (2019). The Role of Innate Leukocytes during Influenza Virus Infection. J Immunol Res *2019*, 8028725.

Langfelder, P., Zhang, B., and Horvath, S. (2008). Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut package for R. Bioinformatics *24*, 719-720.

Li, S. (2019). Regulation of Ribosomal Proteins on Viral Infection. Cells *8*, 508.

Lim, H.K., Huang, S.X.L., Chen, J., Kerner, G., Gilliaux, O., Bastard, P., Dobbs, K., Hernandez, N., Goudin, N., Hasek, M.L.*, et al.* (2019). Severe influenza pneumonitis in children with inherited TLR3 deficiency. J Exp Med *216*, 2038-2056.

Lin, S.J., Lo, M., Kuo, R.L., Shih, S.R., Ojcius, D.M., Lu, J., Lee, C.K., Chen, H.C., Lin, M.Y., Leu, C.M.*, et al.* (2014). The pathological effects of CCR2+ inflammatory monocytes are

amplified by an IFNAR1-triggered chemokine feedback loop in highly pathogenic influenza infection. J Biomed Sci *21*, 99.

Liu, Q., Zhou, Y.H., and Yang, Z.Q. (2016). The cytokine storm of severe influenza and development of immunomodulatory therapy. Cell Mol Immunol *13*, 3-10.

Louie, J.K., Acosta, M., Winter, K., Jean, C., Gavali, S., Schechter, R., Vugia, D., Harriman, K., Matyas, B., Glaser, C.A.*, et al.* (2009). Factors associated with death or hospitalization due to pandemic 2009 influenza A(H1N1) infection in California. JAMA *302*, 1896-1902.

Lun, A.T., Bach, K., and Marioni, J.C. (2016a). Pooling across cells to normalize single-cell RNA sequencing data with many zero counts. Genome Biol *17*, 75.

Lun, A.T., McCarthy, D.J., and Marioni, J.C. (2016b). A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor. F1000Res *5*, 2122.

Machkovech, H.M., Bloom, J.D., and Subramaniam, A.R. (2019). Comprehensive profiling of translation initiation in influenza virus infected cells. PLoS Pathog *15*, e1007518.

Matikainen, S., Sirén, J., Tissari, J., Veckman, V., Pirhonen, J., Severa, M., Sun, Q., Lin, R., Meri, S., Uzé, G.*, et al.* (2006). Tumor necrosis factor alpha enhances influenza A virus-induced expression of antiviral cytokines by activating RIG-I gene expression. J Virol *80*, 3515-3522.

McCarthy, D.J., Campbell, K.R., Lun, A.T., and Wills, Q.F. (2017). Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. Bioinformatics *33*, 1179-1186.

McClain, M.T., Park, L.P., Nicholson, B., Veldman, T., Zaas, A.K., Turner, R., Lambkin-Williams, R., Gilbert, A.S., Ginsburg, G.S., and Woods, C.W. (2013). Longitudinal analysis of leukocyte differentials in peripheral blood of patients with acute respiratory viral infections. J Clin Virol *58*, 689-695.

Okonechnikov, K., Conesa, A., and García-Alcalde, F. (2016). Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. Bioinformatics *32*, 292-294.

Oshansky, C.M., Gartland, A.J., Wong, S.S., Jeevan, T., Wang, D., Roddam, P.L., Caniza, M.A., Hertz, T., Devincenzo, J.P., Webby, R.J.*, et al.* (2014). Mucosal immune responses predict clinical outcomes during influenza infection independently of age and viral load. Am J Respir Crit Care Med *189*, 449-462.

Panthu, B., Terrier, O., Carron, C., Traversier, A., Corbin, A., Balvay, L., Lina, B., Rosa-Calatrava, M., and Ohlmann, T. (2017). The NS1 Protein from Influenza Virus Stimulates Translation Initiation by Enhancing Ribosome Recruitment to mRNAs. J Mol Biol *429*, 3334-3352.

Papalexi, E., and Satija, R. (2018). Single-cell RNA sequencing to explore immune cell heterogeneity. Nat Rev Immunol *18*, 35-45.

Pertea, M., Pertea, G.M., Antonescu, C.M., Chang, T.C., Mendell, J.T., and Salzberg, S.L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol *33*, 290-295.

Pons, P., and Latapy, M. (2005). Computing Communities in Large Networks Using Random Walks. Paper presented at: Computer and Information Sciences - ISCIS 2005, LNCS *3373*, 284-293.

Quach, H., Rotival, M., Pothlichet, J., Loh, Y.E., Dannemann, M., Zidane, N., Laval, G., Patin, E., Harmant, C., Lopez, M.*, et al.* (2016). Genetic Adaptation and Neandertal Admixture Shaped the Immune System of Human Populations. Cell *167*, 643-656 e617.

Ramos, I., Smith, G., Ruf-Zamojski, F., Martínez-Romero, C., Fribourg, M., Carbajal, E.A., Hartmann, B.M., Nair, V.D., Marjanovic, N., Monteagudo, P.L.*, et al.* (2019). Innate Immune Response to Influenza Virus at Single-Cell Resolution in Human Epithelial Cells Revealed Paracrine Induction of Interferon Lambda 1. J Virol *93*, e00559-19.

Russell, A.B., Elshina, E., Kowalsky, J.R., Te Velthuis, A.J.W., and Bloom, J.D. (2019). Single-Cell Virus Sequencing of Influenza Infections That Trigger Innate Immunity. J Virol *93*, e00500-19.

Russell, A.B., Trapnell, C., and Bloom, J.D. (2018). Extreme heterogeneity of influenza virus infection in single cells. eLife *7,* e32303.

Schrauwen, E.J., de Graaf, M., Herfst, S., Rimmelzwaan, G.F., Osterhaus, A.D., and Fouchier, R.A. (2014). Determinants of virulence of influenza A virus. Eur J Clin Microbiol Infect Dis *33*, 479-490.

Stegelmeier, A.A., van Vloten, J.P., Mould, R.C., Klafuric, E.M., Minott, J.A., Wootton, S.K., Bridle, B.W., and Karimi, K. (2019). Myeloid Cells during Viral Infections and Inflammation. Viruses *11,* 168.

Steuerman, Y., Cohen, M., Peshes-Yaloz, N., Valadarsky, L., Cohn, O., David, E., Frishberg, A., Mayo, L., Bacharach, E., Amit, I.*, et al.* (2018). Dissection of Influenza Infection In Vivo by Single-Cell RNA Sequencing. Cell syst. *6*, 679-691.e674.

Sun, J., Vera, J.C., Drnevich, J., Lin, Y.T., Ke, R., and Brooke, C.B. (2020). Single cell heterogeneity in influenza A virus gene expression shapes the innate antiviral response to infection. PLoS Pathog *16*, e1008671.

Vangeti, S., Yu, M., and Smed-Sorensen, A. (2018). Respiratory Mononuclear Phagocytes in Human Influenza A Virus Infection: Their Role in Immune Protection and As Targets of the Virus. Front Immunol *9*, 1521.

Vester, D., Lagoda, A., Hoffmann, D., Seitz, C., Heldt, S., Bettenbrock, K., Genzel, Y., and Reichl, U. (2010). Real-time RT-qPCR assay for the analysis of human influenza A virus transcription and replication dynamics. J Virol Methods *168*, 63-71.

Wang, C., Forst, C.V., Chou, T.W., Geber, A., Wang, M., Hamou, W., Smith, M., Sebra, R., Zhang, B., Zhou, B.*, et al.* (2020). Cell-to-Cell Variation in Defective Virus Expression and Effects on Host Responses during Influenza Virus Infection. mBio *11*, e02880-19.

World Health Organization (2020). Influenza (Seasonal) Fact sheet (accessed 01/12/2020). https://www.who.int/news-room/fact-sheets/detail/influenza-(seasonal)

Xu, C., and Su, Z. (2015). Identification of cell types from single-cell transcriptomes using a novel clustering method. Bioinformatics *31*, 1974-1980.

Young, M.D., and Behjati, S. (2020). SoupX removes ambient RNA contamination from droplet based single-cell RNA sequencing data. bioRxiv, https://www.biorxiv.org/content/10.1101/303727v2

Young, M.D., Wakefield, M.J., Smyth, G.K., and Oshlack, A. (2010). Gene ontology analysis for RNA-seq: accounting for selection bias. Genome Biol *11*, R14.

Zhang, Q. (2020). Human genetics of life-threatening influenza pneumonitis. Hum Genet *139*, 941-948.

Zheng, G.X., Terry, J.M., Belgrader, P., Ryvkin, P., Bent, Z.W., Wilson, R., Ziraldo, S.B., Wheeler, T.D., McDermott, G.P., Zhu, J.*, et al.* (2017). Massively parallel digital transcriptional profiling of single cells. Nat Commun *8*, 14049.

Ziegler-Heitbrock, L., Ancuta, P., Crowe, S., Dalod, M., Grau, V., Hart, D.N., Leenen, P.J., Liu, Y.J., MacPherson, G., Randolph, G.J.*, et al.* (2010). Nomenclature of monocytes and dendritic cells in blood. Blood *116*, e74-80.

Figures

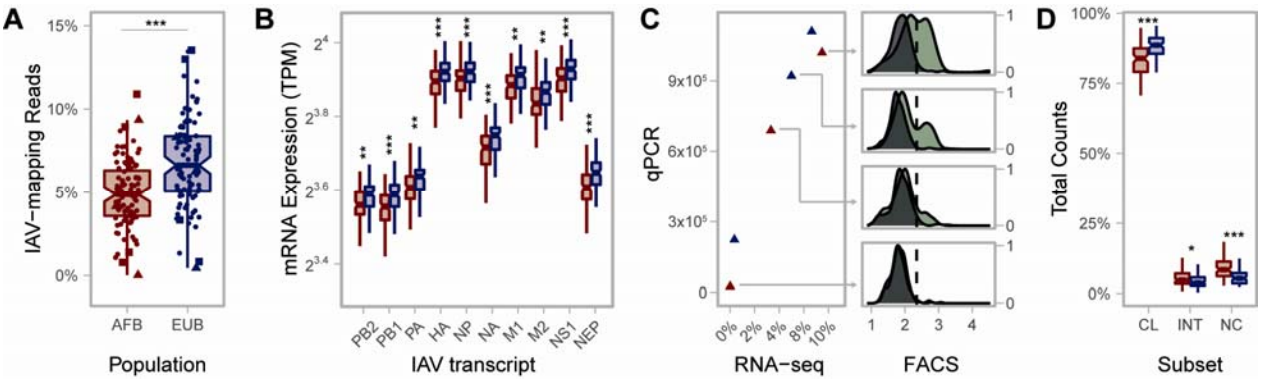

Figure 1. Marked Individual and Population Variation in Influenza and Cellular Phenotypes
(A) Inter- and intra-population variation in the percentage of RNA-seq reads mapping to the IAV genome (n_{AFB}=100, n_{EUB}=99). Colors reflect population (AFB in red and EUB in blue). The percentage of RNA-seq reads mapping to the IAV genome is significantly different between populations (p-value=5.3×10^{-8}). Donors used in the replication experiment (n_{AFB}=3, n_{EUB}=3) are designated with triangles, donors used in the scRNA-seq experiment (n_{AFB}=4, n_{EUB}=4) are designated with squares, and all other donors are designated with circles.
(B) Inter- and intra-population variation in viral mRNA expression at 6hpi (n_{AFB}=100, n_{EUB}=99). Expression levels for each of the 10 primary transcripts of the IAV are plotted. Colors reflect population (AFB in red and EUB in blue). All IAV transcripts are significantly different between populations (p-value<0.001). Outlier points are not displayed.
(C) Replication of interindividual variation in the amount of viral mRNA recovered from cells at 6hpi, and demonstration of cellular heterogeneity in viral protein expression. For six individuals spanning the phenotypic spectrum of IAV-mapping reads (RNA-seq), IAV nucleoprotein (NP) expression was measured (qPCR) in an independent experiment. Results from the two experiments are highly correlated (Pearson r = 0.96, p-value = 1.3×10^{-3}). In parallel, flow cytometry analysis for NP protein expression (FACS) was performed and four representative density plots are show. Grey shows the background fluorescence from control samples, while green shows the fluorescence intensity of NP from IAV-challenged samples. Dotted line depicts the threshold used to classify a cell as expressing NP.
(D) Inter- and intra-population variation in the number of classical (CD14^{++}CD16^{-}), intermediate (CD14^{++}CD16^{+}), and nonclassical (CD14^{+}CD16^{++}) monocytes following CD14^{+} selection from frozen PBMCs (n_{AFB}=89, n_{EUB}=85). Colors reflect population (AFB in red and EUB in blue). All three subsets are significantly different between populations (p-value<0.01). Outlier points not displayed. Abbreviations: African ancestry individuals from Belgium (AFB), European ancestry individuals from Belgium (EUB), transcripts per million (TPM), influenza A virus (IAV), quantitative polymerase chain reaction (qPCR), fluorescence-activated cell sorting (FACS), classical (CL), intermediate (INT), and nonclassical (NC).
* p-value < 0.01
** p-value < 0.001
*** p-value < 0.0001

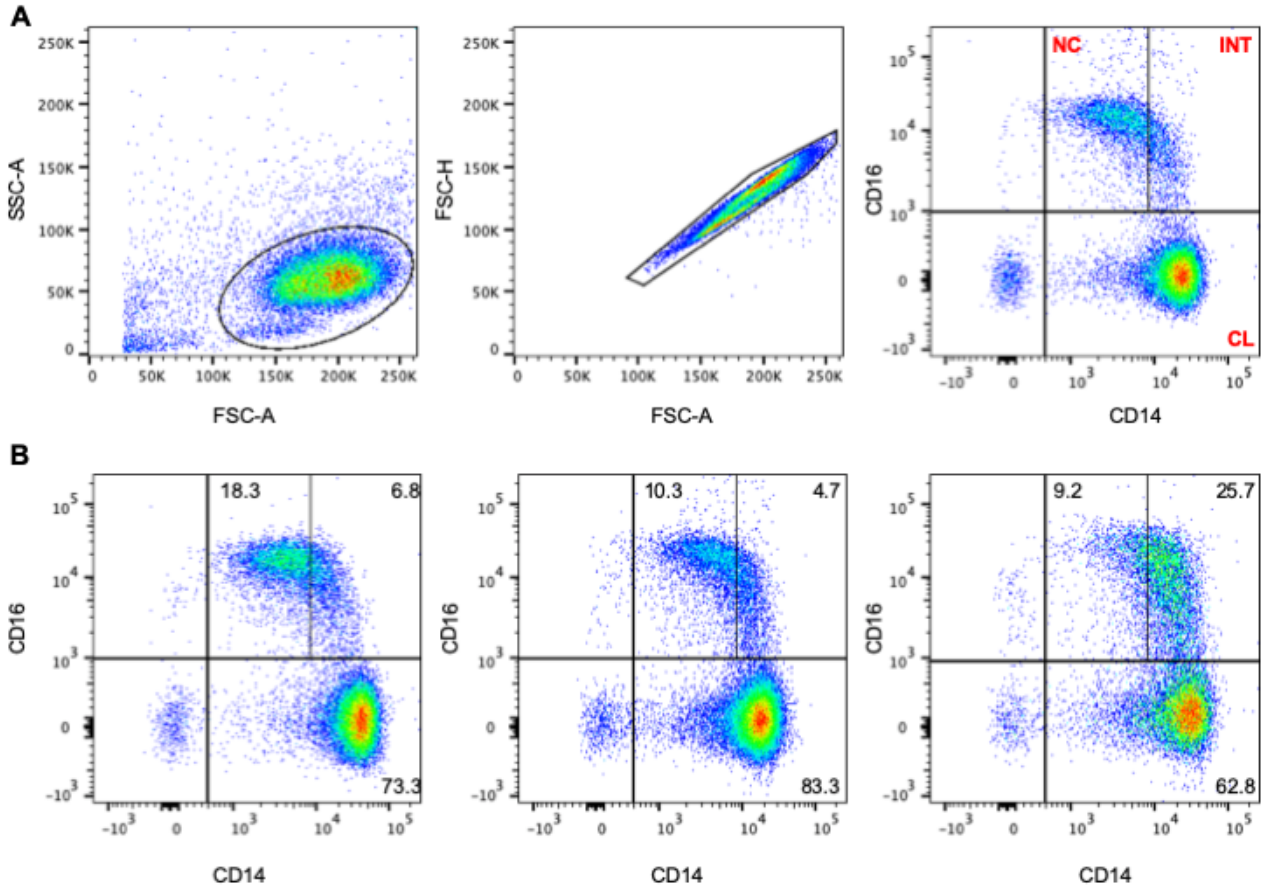

Figure S1. Flow Cytometry Analysis and Gating Strategy of EVOIMMUNOPOP Samples
(A) Gating strategy to discriminate distinct monocyte subsets based on CD14 and CD16 expression, after doublet exclusion with forward scatter-height and forward scatter-area.
(B) Sample FACS profiles for three donors of the EVOIMMUNOPOP cohort with varying proportions of monocyte subsets.
Abbreviations: classical (CL), intermediate (INT), and nonclassical (NC).

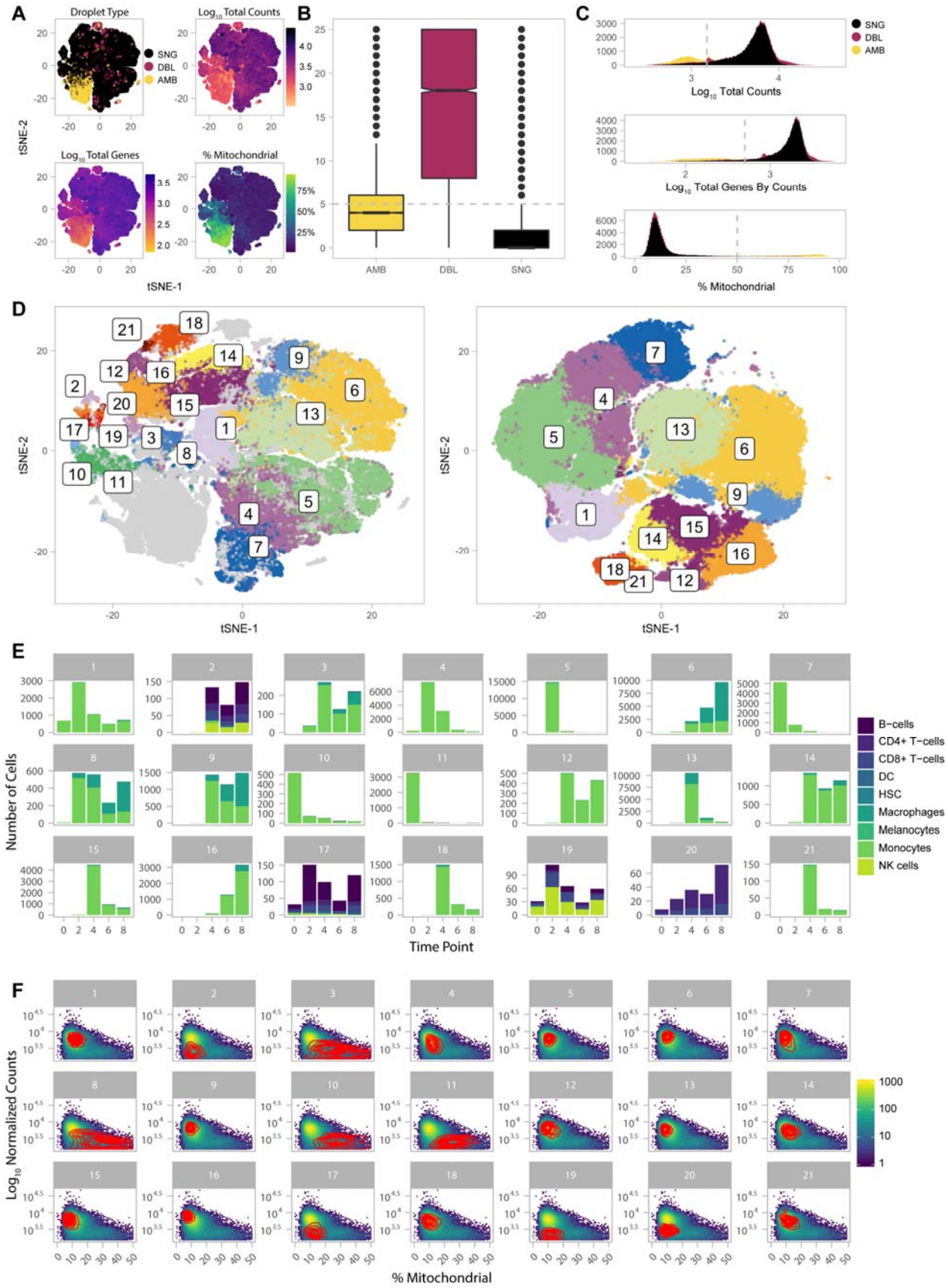
A
Droplet Type
SNG
DBL
AMB
Log10 Total Counts
Log10 Total Genes
% Mitochondrial
tSNE-1
tSNE-2
B
AMB
DBL
SNG
C
Log10 Total Counts
Log10 Total Genes By Counts
% Mitochondrial
D
tSNE-1
tSNE-2
E
Number of Cells
Time Point
B-cells
CD4+ T-cells
CD8+ T-cells
DC
HSC
Macrophages
Melanocytes
Monocytes
NK cells
F
Log10 Normalized Counts
% Mitochondrial

Figure S2. Quality Control of Droplet-Based scRNA-seq Data Excludes Doublets, Contaminants, and Dying Cells

(A) t-distributed stochastic neighbor embeddings (tSNEs) of 132,130 cell-containing droplets colored by various QC metrics.

(B) Doublet detection based on genetics and nearest-neighbors. Cell-containing droplets were traced back to donors using two independent methods - Demuxlet and SoupOrCell - both of which capitalize on genetic variation in the sequencing reads. Comparison of genetically deemed doublets to high-confidence singlets (concordant identification of the donor across both programs) revealed that doublets are more likely to share nearest-neighbors in a *knn*-graph with other doublets, and we used this feature to identify droplets presumed to contain two or more cells originating from the same donor. Barcodes with > 5 genetic doublets as nearest-neighbors were considered doublets and excluded from post-QC analyses.

(C) Distribution of standard QC metrics across all 132,130 barcodes, colored by droplet type. Droplets with < 1500 total counts (top), < 500 genes (middle), or > 50% mitochondrial gene content (bottom) were excluded from further analysis and these thresholds are designated with grey dotted lines.

(D) Pre-QC tSNE (left) and post-QC tSNE (right) colored by graph-based clusters. Cells that were excluded in the first stages of QC (e.g. doublets, of cells with low read counts and high mitochondrial content) are colored in grey on the pre-QC tSNE.

(C) SingleR cell type predictions for 96,386 high-quality single cells stratified by cluster membership and time point are displayed. We excluded clusters 2, 17, 19 and 20 from further analyses due to the predominance of contaminant cell types.

(D) Identification of dying cell clusters. 2D kernel density estimation (red contour) is plotted for each cluster against the background of the total dataset represented by the hexagonal heatmap showing the distribution for two quality control metrics for 96,386 high-quality single cells. We exclude clusters 3, 8, 10, and 11. Clusters 2, 17, 19, and 20 are also outliers, but were already deduced to be contaminant cell populations.

Abbreviations: SNG, singlets; DBL, doublets; AMB, ambiguous

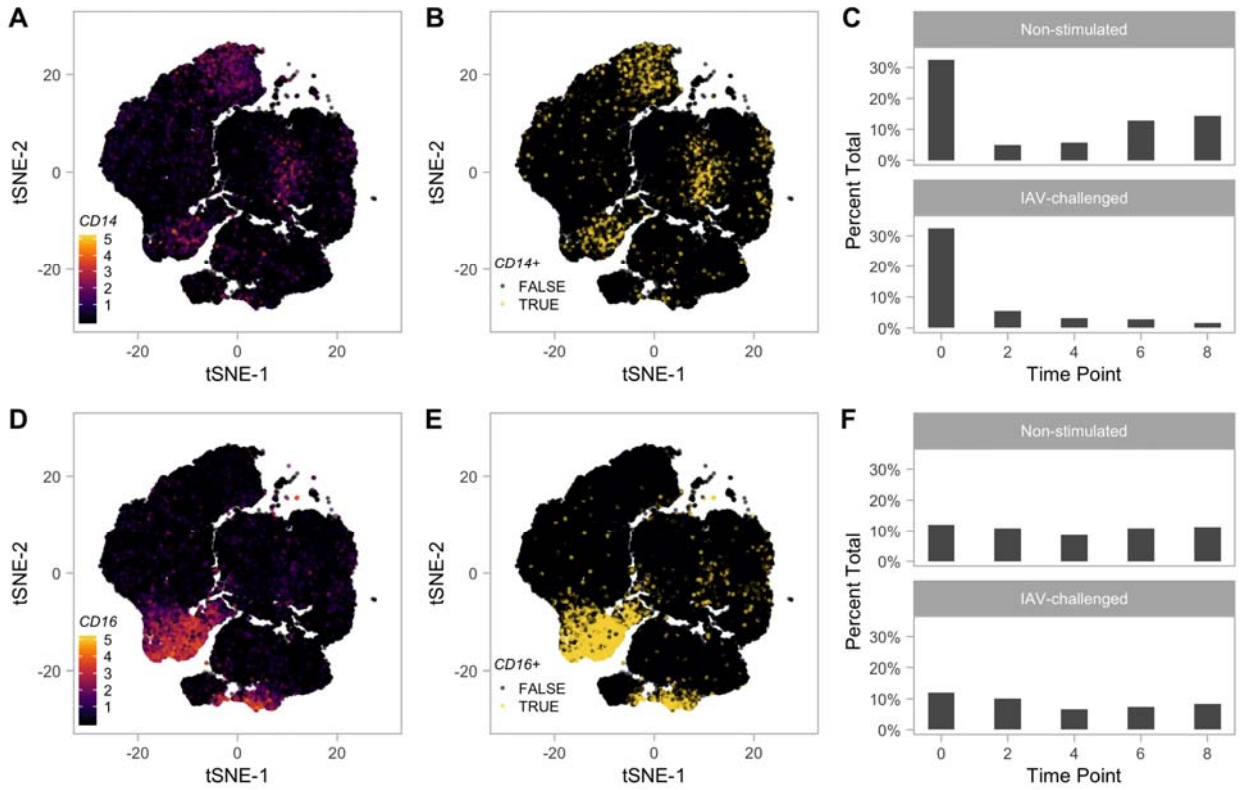

Figure S3. mRNA Expression of *CD14* and *CD16* (*FCGR3A*) Highlight Distinct Cell Populations

(A) tSNE plot colored by $\log_2$ normalized counts for *CD14*.
(B) tSNE plot colored based on significance of *CD14* expression when accounting for ambient RNA (*CD14*$^+$, FDR < 0.01).
(C) Barplot showing the percentage of *CD14*$^+$ cells across time points and conditions.
(D) tSNE plot colored by $\log_2$ normalized counts for *FCGR3A* (aka *CD16*).
(E) tSNE plot colored based on significance of *FCGR3A* expression when accounting for ambient RNA (*CD16*$^+$, FDR < 0.01).
(F) Barplot showing the percentage of *CD16*$^+$ cells across time points and conditions.

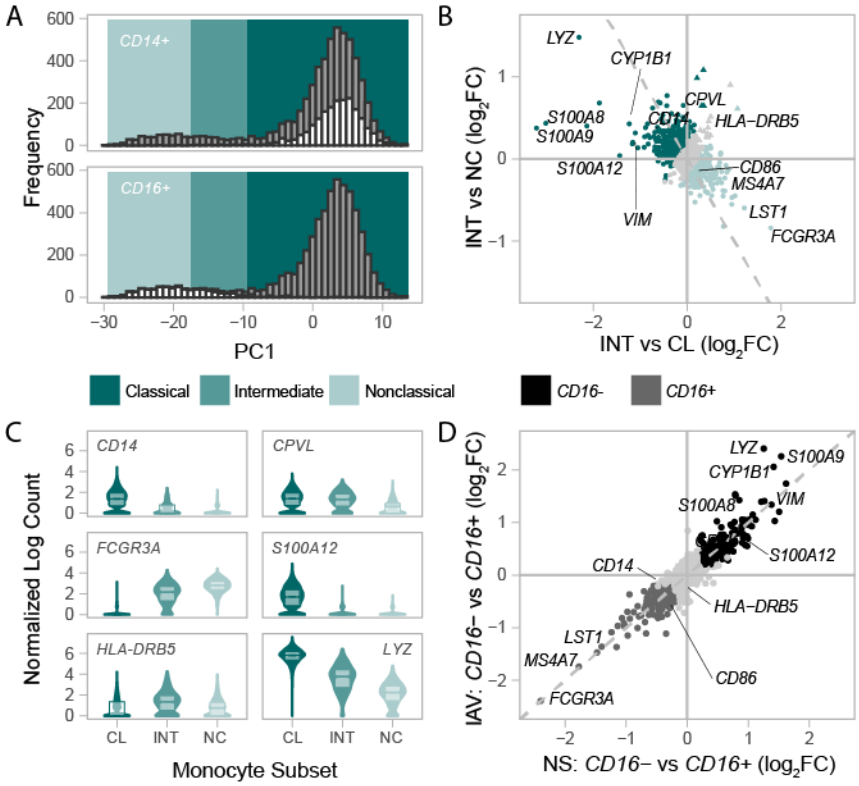

Figure S4. scRNA-seq Accurately Recovers Monocyte and Mononuclear Phagocyte Subsets and their Functional Roles

(A) Distribution of 6,601 monocytes along PC1 at T_0, highlighting *CD14*$^+$ (top panel) and *CD16*$^+$ (bottom panel) cells. Differing shades of green distinguish nonclassical (light green), intermediate (medium green), and classical (dark green) monocyte subsets.

(B) Average gene expression difference among the three canonical monocyte subsets for 4,589 genes at T_0. The average $\log_2$FC change in expression between intermediate and classical subsets are plotted on the x-axis, while the average $\log_2$FC between intermediate and nonclassical subsets are plotted on the y-axis. Triangles designate the 18 genes upregulated in the intermediate subset relative to both classical and nonclassical subsets ($\log_2$FC>0.2, FDR≤1%).

(C) Example distributions of gene expression for monocyte subset markers at T_0.

(D) Average gene expression differences between *CD16*$^-$ and *CD16*$^+$ monocyte subsets for 5,368 genes across T_2 to T_8 in two conditions. Differences between the two subsets in the non-stimulated condition are plotted on the x-axis, while differences between the two subsets in the IAV-challenged condition are plotted on the y-axis. Abbreviations: classical (CL), intermediate (INT), nonclassical (NC), non-stimulated (NS), and IAV-challenged (IAV).

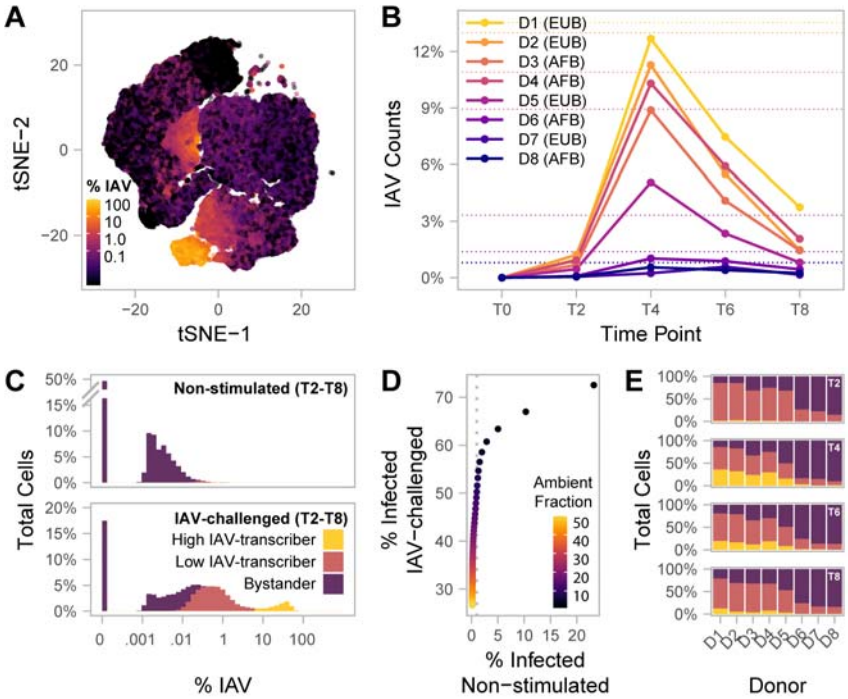

Figure 2. *Ex vivo* Susceptible and Resistant Individuals Differ in Viral-Related Phenotypes
(A) t-SNE representation of single-cell RNA-sequencing data, colored by the percentage of counts of viral origin across all high-quality monocytes.
(B) Pseudo-bulk estimates of the percentage of counts of viral origin in IAV-challenged condition (T_2-T_8). Donors are colored based on the rank of their prior phenotype, with light colors reflecting those who were previously highly susceptible and darker colors reflecting those who were previously resistant to *ex vivo* challenge; the dotted colored lines mark the percentage of reads mapping to IAV at 6hpi from the bulk RNA-seq study.
(C) Distribution of counts of viral origin across all donors, from T_2 to T_8. Cells are shown separately for non-stimulated (top) and IAV-challenged (bottom) conditions. Fill color reflects the cell state assignments: unexposed/bystander cells (purple), low IAV-transcribing infected cells (peach), and high IAV-transcribing infected cells (yellow). Note that the threshold used to define infected cells is dependent on the number of viral mRNAs in the ambient pool, and varies across libraries.
(D) Determination of the maximum contamination fraction by ambient RNA. The number of non-stimulated cells deemed to significantly express IAV transcripts (presumed false positives) versus the number of IAV-challenged cells deemed to significantly express IAV transcripts across a range of maximum contamination fractions from 1-50%. Dotted grey line is drawn at 1% on the x-axis. A maximum contamination fraction of 10% results in 1% of non-stimulated cells being classified as infected cells (false positive proxy), and half of IAV-challenged cells showing evidence of viral transcription.
(E) Proportions of cell states from the IAV-challenged condition at T_2, T_4, T_6, and T_8, in each of the eight donors. Donors are ordered across the x-axis based on the percentage of reads mapping to IAV at 6hpi from the bulk RNA-seq study, from most susceptible to most resistant.

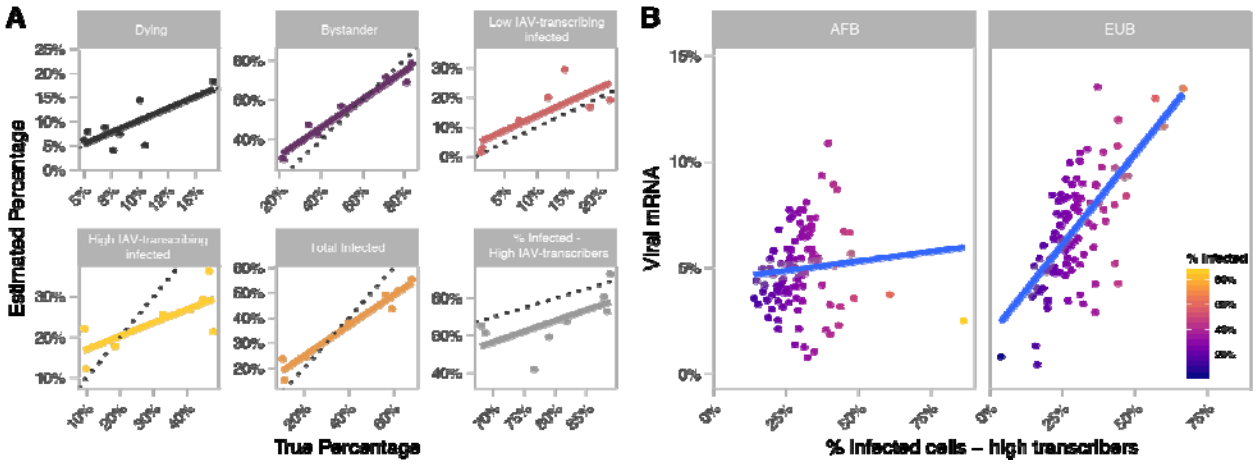

Figure S5. Deconvolution of Bulk RNA-Seq Profiles Correctly Infers Proportion of Infected Cells and Highlights Greater Infectivity among European-Ancestry Individuals

(A) Cross-validation of the deconvolution method. For each cell state, comparison between the percentage of UMIs that originate from that cell state in the scRNA-seq data and the percentages estimated from their bulk RNA-seq profiles are plotted. A comparison is also provided for the total percentage of infected cells (orange) and the percentage of high IAV-transcribers among infected cells (grey). For each comparison, a regression line is depicted with a 95% confidence interval. Dotted lines indicate when estimates are identical to the observed values.

(B) Percentage of IAV reads as a function of the estimated percentage of high IAV transcribers among infected cells, for African and European ancestry individuals (AFB and EUB, respectively). Colors reflect the estimated percentage of infected cells for the sample. One individual with no infected cells was excluded (n_{AFB}=99, n_{EUB}=99).

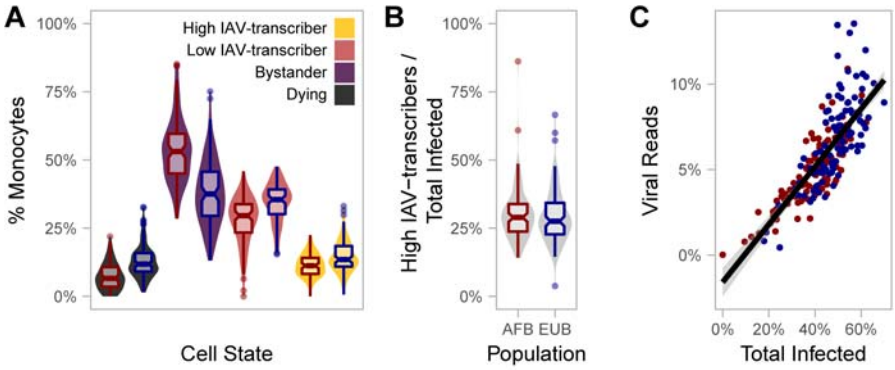

Figure 3. Deconvolution of Bulk RNA-seq Reveals Population Differences in the Proportions of Infected Cells

(A) Estimated distribution of the percentage of cells from each cell state in the bulk RNA-seq data (n_{AFB}=100, n_{EUB}=99). Fill colors reflect cell state assignments, while outlines of boxplots reflect population (AFB in red and EUB in blue).

(B) Distribution of the percentage of high IAV-transcribers among infected cells, stratified by population. One individual with no infected cell was excluded (n_{AFB}=99, n_{EUB}=99).

(C) Percentage of RNA-seq reads of viral origin as a function of the estimated proportion of infected cells (n_{AFB}=100, n_{EUB}=99), colored by population (AFB in red and EUB in blue).

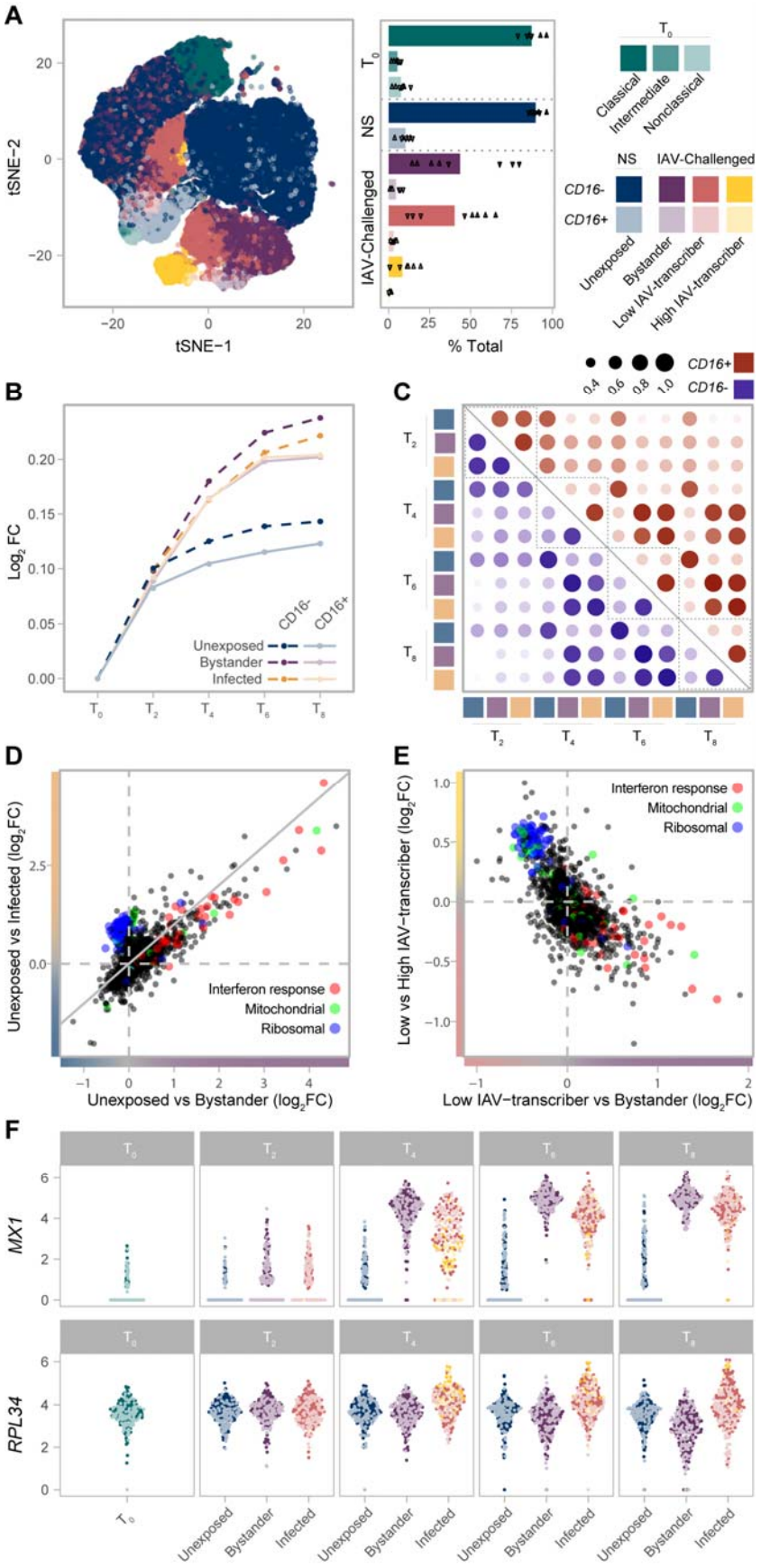
A
tSNE−2
tSNE−1
% Total
T0
NS
IAV-Challenged
Classical
Intermediate
Nonclassical
CD16-
CD16+
Unexposed
Bystander
Low IAV-transcriber
High IAV-transcriber
B
Log2 FC
Unexposed
Bystander
Infected
C
D
Unexposed vs Infected (log2FC)
Unexposed vs Bystander (log2FC)
Interferon response
Mitochondrial
Ribosomal
E
Low vs High IAV−transcriber (log2FC)
Low IAV−transcriber vs Bystander (log2FC)
F
MX1
RPL34

Figure 4. Gradient of mRNA Expression from Ribosomal and IFN-Stimulated Genes Separates Bystander and Infected Cells

(A) Two-dimensional tSNE of transcriptional profiles of cells collected across all five different time-points, two conditions, and eight donors. The relative proportions of cells deemed as classical, intermediate, and non-classical monocytes at T_0, those deemed *CD16*$^-$ and *CD16*$^+$ in the non-stimulated condition from T_2-T_8, and those deemed bystander, low IAV-transcribers (infected), and high IAV-transcribers (infected) stratified by CD16$^{-/+}$ status in the IAV-challenged condition from T_2-T_8 are plotted to the right and colored according to the legend. The bar plot represents the mean across the eight donors with individual triangles plotted for each donor (right-side-up triangles represent donors from the more susceptible extreme of the cellular susceptibility phenotype, while upside-down triangles represent donors from the more resistant extreme).

(B) Average $\log_2$FC in expression across time (relative to T_0) in each monocyte subset and cell state, based on a subsampled data sets with balanced representation of all donors, cell states, and subsets.

(C) Correlation between fold changes in expression relative to T_0 across cell states and consecutive time-points, based on a subsampled data set with balanced representation of all donors, cell states, and monocyte subsets. Correlations are shown separately for *CD16*$^+$ and *CD16*$^-$ monocyte subsets (upper and lower triangle, respectively). The size (and color) of dots reflect the correlation coefficient as indicated in the key.

(D). Transcriptional responses of cells upon IAV-challenge (T_2-T_8) highlight the interplay between IFN response (GO:0034340), ribosomal (GO:0005840), and mitochondrial (GO:0005743) genes. The $\log_2$FC change in gene expression between unexposed and bystander cells is plotted on the x-axis, while the $\log_2$FC change in gene expression between unexposed and infected cells is plotted on the y-axis. Values are plotted based on a meta-analysis across time points and subsets, based on a subsampled data set with balanced representation of all donors.

(E) The interplay between IFN response (GO:0034340), ribosomal (GO:0005840), and mitochondrial (GO:0005743) genes among cells exposed to IAV. The $\log_2$FC change in gene expression between low IAV-transcribing infected and bystander cells is plotted on the x-axis, while the $\log_2$FC change in gene expression between low IAV-transcribing infected and high IAV-transcribing infected cells is plotted on the y-axis. Values are plotted based on a meta-analysis across monocyte subsets at T_4.

(F) mRNA expression levels of representative IFN-stimulated (*MX1*) and ribosomal (*RPL34*) genes across the subsampled dataset. Colors reflect the cell state and subset assignment depicted at the top right of the figure, with the exception of cells at T_0, which are colored based on the two-subset designation (*CD16*$^-$ vs *CD16*$^+$).

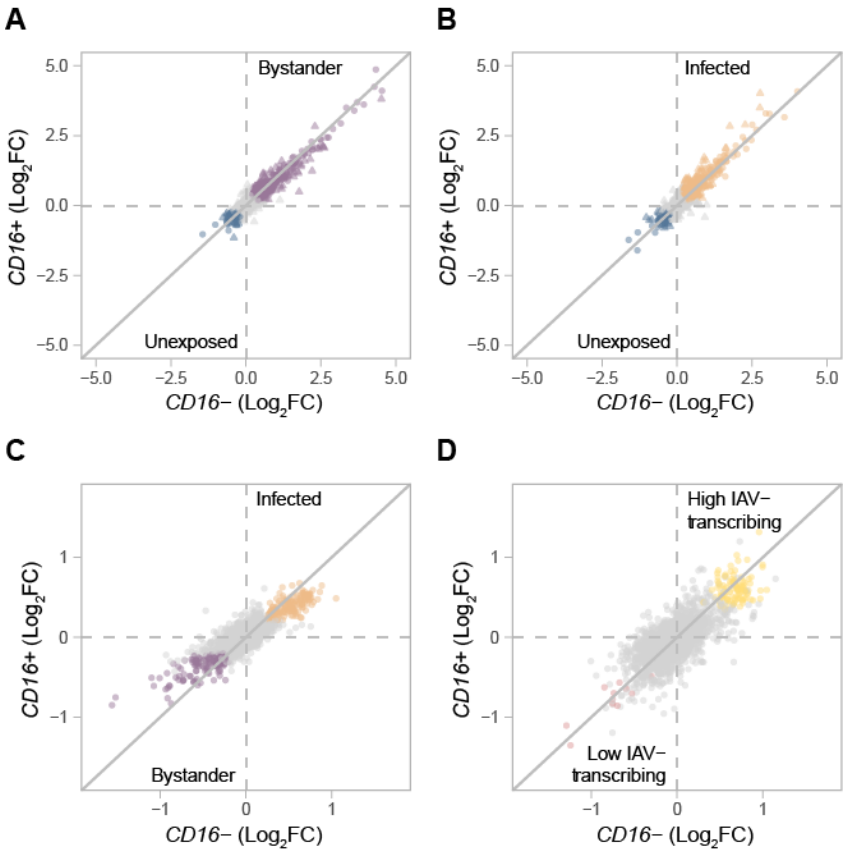

Figure S6. Monocyte Subsets Display Similar Transcriptional Responses of Monocytes to IAV.

For each cell state comparison, the average $\log_2$FC in expression between (A) unexposed and bystander, (B) unexposed and infected, (C) bystander and infected, and (D) low and high IAV-transcribing infected cells is plotted for *CD16*$^-$ cells on the x-axis and *CD16*$^+$ cells on the y-axis. Colors reflect common transcriptional responses among cell states (>0.2 $\log_2$FC in both subsets, FDR≤1%). Triangles in *A* & *B* represent genes which demonstrate a stronger response in a particular monocyte subset (interaction test, FDR≤1%).

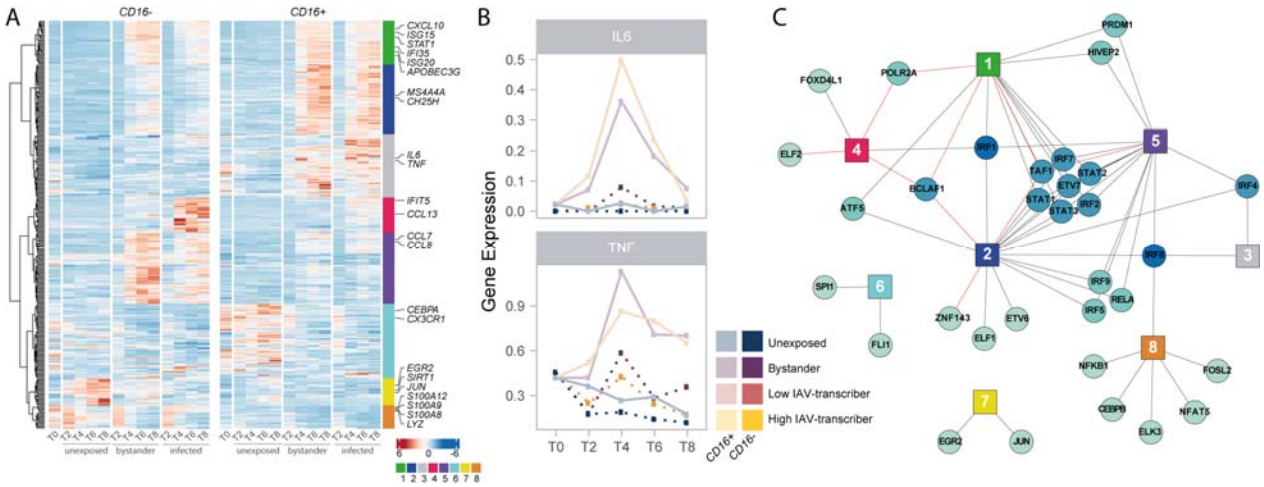

Figure 5. IRFs and STATs Have a Central Role in the Subset-Specific Responses to IAV Infection

(A) Heatmap of scaled gene expression from 335 genes displaying a subset-specific response to infection. Genes are grouped into 8 modules based on hierarchical clustering of their expression patterns. Representative genes from each module are highlighted.

(B) Mean expression over time from *IL6* and *TNF*, across the different monocyte subsets and cell states.

(C) Network of transcription factors (round nodes) associated with each module (square nodes). TF nodes are colored according to the number of modules they are associated with.

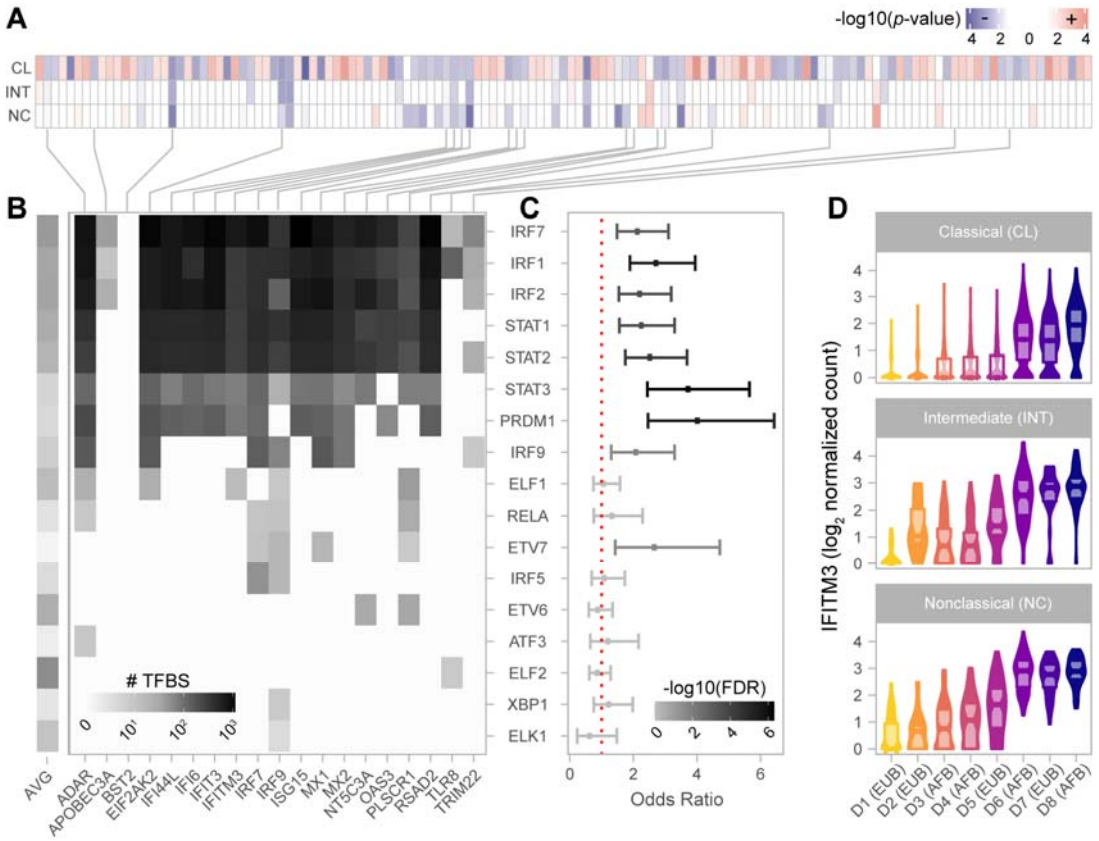

Figure 6. **Basal Expression of IFN-Stimulated Genes, Targets of IRF and STAT TFs, Predicts the Proportion of Infected Cells**

(A) Association of viral mRNA levels estimated from bulk with expression of 135 candidate genes, across the 3 monocytes subsets. Associations are shown in blues or red depending on the sign of the correlation.

(B) Heatmap of the number of TF motifs (as reported by SCENIC) located around the promoter of 19 IFN-stimulated/defense response to virus genes that correlate with IAV levels. Only 17 TFs with at least one binding motif are shown. For the same 17 TFs, the average number of binding motifs across all 4,578 genes expressed at T_0 (AVG) are shown on the left-hand side.

(C) For the same 17 TFs, enrichment of SCENIC-predicted targets among the 135 genes associated with bulk IAV levels at T_0 (odds ratio and 95% confidence interval). Red dotted line designates an odds ratio equal to 1. For each TF, color code indicates significance of enrichment among the 135 candidate genes.

(D) Expression of *IFITM3* across the 8 donors stratified by monocyte subset. Colors reflect the different donors which are ordered by decreasing viral mRNA levels from the bulk RNA experiment. For each donor and monocyte subset, the violin plots show the full distribution of *IFITM3* expression across individual cells and boxplots highlight the median and interquartile range.

STAR Methods

RESOURCE AVAILABILITY

Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Lluis Quintana-Murci (lluis.quintana-murci@pasteur.fr).

Materials Availability

This study did not generate new unique reagents.

Data and Code Availability

The bulk RNA-seq data used in this study are available at the European Genome Phenome archive under accession number [EGAS00001001895]. The single cell RNA-seq data generated during this study are available at the European Genome Phenome archive under accession number [EGAS0000XXXXXXX]. Code generated as part of this study is available on Github (https://github.com/ONeillMB1/PopDiff_MonocyteIAV)

EXPERIMENTAL MODEL AND SUBJECT DETAILS

The EVOIMMUNOPOP cohort has been previously described (Quach et al., 2016). Briefly, 200 healthy male donors living in Belgium of self-reported European descent (EUB) or African descent (AFB) were recruited. Inclusion was restricted to nominally healthy individuals between 19 and 50 years of age at the time of sample collection. Serological testing was performed for all donors to exclude those with serological signs of past or ongoing infection with human immunodeficiency virus (HIV), hepatitis B virus (HBV) or hepatitis C virus (HCV).

METHOD DETAILS

Monocyte Subset Characterization of EVOIMMUNOPOP Samples via FACS

For 174 of the 200 EVOIMMUNOPOP donors, proportions of classical, intermediate and nonclassical monocytes were determined based on a fraction of 10^5 $CD14^+$ positively-selected monocytes, stained according to the manufacturer's instructions, with fluorescent APC-conjugated anti-CD14 and PE-conjugated anti-CD16 antibodies (Cat#130-091-243 and Cat #130-091-245, respectively, Miltenyi Biotec). Samples were then analyzed on a MACSQuant Analyzer 10 benchtop flow cytometer (Miltenyi Biotec).

Replication of Viral Phenotypes

To establish reproducibility of viral phenotypes and explore cellular heterogeneity, we challenged 6×10^6 primary monocytes from 6 representative donors with IAV as previously described (Quach et al., 2016). Following 6 hours of infection, cells were split in two fractions to measure two viral phenotypes: (1) intracellular NP levels via FACS and (2) viral mRNA levels of NP via qPCR. Five-hundred thousand monocytes from each donor were collected by centrifugation and washed once with 5 ml PBS + 1% fetal bovine serum (FBS), before being fixed with 1 ml of PBS + 4% paraformaldehyde for 20 minutes. This fixed fraction was then used for a Flow cytometry-based assay of NP protein expression. The remaining fraction of monocytes were collected by centrifugation and lysed in a guanidinium thiocyanate solution, and stored at –80°C until RT-qPCR quantification.

To quantify the intracellular content of viral NP in monocytes from each donor, the fixed fraction of cells was washed with 5 ml PBS + 1% FBS. Cells were then permeabilized in 1 ml PBS + 0.1% triton for 30 minutes and washed with 5 ml PBS + 1% FBS. Cells were resuspended in 1 ml cold PBS + 1% FBS and distributed into a 96-well plate (200 µl of cell suspension/well). After centrifugation, cells were incubated in 100 µl staining solution containing a fluorescent FITC-conjugated anti-IAV NP antibody (Cat#MA1-7322, ThermoFisher Scientific) at a dilution of 1:50 for 45 minutes at 4°C. Cells were then washed twice and resuspended in cold PBS + 1% FBS. Samples were analyzed on a MACSQuant Analyzer 10 benchtop flow cytometer (Miltenyi Biotec).

To replicate the viral phenotype at the mRNA level, total RNA was extracted with the Nucleospin miRNA kit (Cat#740971.250, Macherey Nagel), which includes an enzymatic digestion of genomic DNA. RNA quality and quantity were assessed with a Nanodrop spectrometer and the Agilent Bioanalyzer RNA 6000 nano kit (Cat#5067-1511, Agilent Technologies). To quantify mRNA levels of viral NP, we developed a real-time RT-PCR assay for the analysis of viral messenger RNAs (mRNAs) of the segment 5, which codes for the IAV NP, as previously described (Kawakami et al., 2011; Vester et al., 2010). Briefly, viral RNAs from strain A/USSR/90/1977(H1N1) were extracted from supernatant of infected MDCK cells using the QIAamp viral RNA mini kit (Cat# 52904, Qiagen). After reverse transcription using an oligo d(T)16 primer (Cat# N8080128, ThermoFisher Scientific), cDNAs were used as template for full length PCR amplification using an extended 5'-primer with a T7 promoter sequence (**Table S5A**). The PCR product was then *in vitro* transcribed using the T7 Ribomax Large Scale RNA production system (Cat#P1300, Promega), and synthetic viral mRNAs were purified with the Nucleospin miRNA kit (Cat#740971.50, Macherey-Nagel). The concentration of purified

synthetic viral mRNA transcripts was assessed by spectrophotometry, and converted into molecular copies by considering the molecular weight of the corresponding sequence. The resulting synthetic viral mRNAs were considered as a reference standard for the quantification of viral mRNA copies in our monocyte samples challenged with IAV.

To assess the number of viral mRNA copies in these samples, total RNAs were reverse transcribed, together with the reference standard, using a primer specific to the IAV NP, which was extended at the 5' end with Tag 3 sequence (**Table S5B**). qPCR reactions were performed on a QuantStudio 7 Flex Real-time PCR system (Applied Biosystems), in triplicates using the Power Sybr Green Master Mix (Cat# 4367659, ThermoFisher Scientific) according to the manufacturer's instructions. A negative water sample was included in each run and the specificity of the PCR product was monitored by analyzing the melting curve. The amount of NP transcripts was determined using the corresponding synthetic viral mRNA standard curve (10-fold dilution series). IAV mRNA levels estimated based on RNA-seq were compared with either RT-qPCR-based estimates of mRNA levels or the percentage of NP-expressing cells estimated by FACS, using a Pearson's correlation test.

Single-Cell Analyses and RNA Sequencing

For eight selected donors (4 susceptible and 4 resistant to *ex-vivo* infection, equally split by ancestries), 100×10^6 PBMCs were thawed, washed twice and resuspended in complete medium: pre-warmed RPMI-1640 Glutamax medium, supplemented with 10% FCS and 1% penicillin/streptomycin (Cat# 15140-122, Life Technologies). Monocytes were then positively selected with magnetic CD14 microbeads, according to the manufacturer's instructions (Cat#130-050-201, Miltenyi Biotec). The number of monocytes was determined with the Countless2 automated cell counter system (Cat# AMQAX1000, ThermoFisher Scientific) in the presence of trypan blue. For each donor, monocytes were seeded at 0.5×10^6 monocytes per well on 24-well NUNC plates in 500 µL of complete media and allowed to rest for one hour at 37°C under 5% CO_2. Five-hundred microliters of complete media (non-stimulated) or A/USSR/90/1977(H1N1) at a concentration of 1×10^6 pfu/mL in complete media (IAV-challenged, MOI=1) were added to each sample. Following one hour of staging at 4°C, plates were centrifuged at 1300 rpm for 10 minutes at 4°C, media was removed by pipette, and each well was washed with 1mL complete media. The spin was repeated, media removed by pipette, and samples were resuspended in 1mL pre-warmed complete media before being transferred to an incubator at 37°C under 5% CO_2 to initiate infection (T_0).

At each time point (T_0, T_2, T_4, T_6, and T_8), samples were mixed by pipetting and transferred to Eppendorf tubes. Wells were washed with 300uL of PBS + 0.04% BSA and transferred to the same tubes. Collection tubes were centrifuged at 1300 rpm for 10 minutes, media was removed and replaced with 1mL PBS + 0.04% BSA and an aliquot of 10µL was taken to count each sample on a Countless2 automated cell counter system, before repeating the centrifugations. Individual samples were adjusted to 2×10^6 live cells/mL.

Samples were multiplexed for running on the 10X Chromium (Cat# 120223 & 1000074, 10X Genomics) by mixing equal proportions from 6-8 samples in a manner that balanced conditions and allowed us to assess for batch effects across lanes (**Table S5C**). Multiplexed samples were counted with the Countless2 automated cell counter system and adjusted to target recovery of 10,000 cells per reaction of the Chromium Single Cell 3' Reagent Kits v3 (Cat# 1000092 & 1000078, 10X Genomics) assuming a recovery rate of 50%. GEM Generation & Barcoding, Post GEM-RT Cleanup & cDNA Amplification, and 3' Gene Expression Library Construction were performed as per manufacturer's instructions (10xGenomics, 2020). All 13 libraries were mixed prior to sequencing across 13 different lanes from an Illumina HiSeq X (28bp barcode + 91bp insert – target 400 M reads pairs per lane), leading to a total of 5.3 billon reads.

QUANTIFICATION AND STATISTICAL ANALYSIS

Analysis of Bulk RNA-Seq Profiles from the EVOIMMUNOPOP Cohort

A combined human-IAV reference was generated by concatenation of the primary human genome assembly (GRCh38) with the 8 segments of the human influenza A virus (IAV) A/USSR/90/1977(H1N1) genome (accession numbers CY010372-CY010379). Comprehensive human gene annotation was obtained from GENCODE (release 27) and merged with the 12 known transcripts of A/USSR/90/1977(H1N1). RNA-seq reads (FASTQs) for all 970 samples that passed quality control in our previous study (Quach et al., 2016) [accession EGAS00001001895] were mapped to the combined reference with the STAR aligner (v.2.5.0a) (Dobin et al., 2013) and assessed for quality with QualiMap 'bamqc' and 'rnaseq' (García-Alcalde et al., 2012; Okonechnikov et al., 2016). Expression of viral mRNAs was measured as the percentage of uniquely-mapped reads aligning to the IAV genome. Reassuringly, the mean percentage of RNA-seq reads among samples from the IAV-challenged condition was 5.86% versus <0.01% in the other four conditions. Comparison of the percentage of IAV reads between populations was done using a Wilcoxon rank test. StringTie (v.1.3.3) (Pertea et al., 2015) was used to quantify expression levels in transcripts per million mapped reads (TPM) for each annotated transcript. Gene expression data were filtered to remove genes with little evidence of

activation (mean zTPM score < -3) (Hart et al., 2013) in any of the 5 conditions, and their quality was checked by principal component analysis (PCA). As GC content, 5′/3′ bias, date of the experiment and library batch were previously determined to be the strongest confounding factors on transcript expression (Quach et al., 2016), we corrected the data for these factors. First, we adjusted the data for GC content and 5′/3′ bias using linear models. Then, we imputed missing values by k-nearest neighbor imputation and adjusted for experiment date and library batch by sequentially running ComBat (Johnson et al., 2007) for each batch effect, with condition and population as covariates. After batch effect correction, only IAV-stimulated samples were kept for downstream analyses.

Genotyping data [accession EGAS00001001895] were obtained for all individuals based on both Illumina HumanOmni5-Quad BeadChips and whole-exome sequencing with the Nextera Rapid Capture Expanded Exome kit (Quach et al., 2016). The 3,782,260 SNPs obtained after stringent quality control were then used for imputation, based on the 1,000 Genomes Project imputation reference panel (Phase 1 v3.2010/11/23) (1000 Genomes Project Consortium, 2012), leading to a final set of 19,619,457 high-quality SNPs, of which 7,766,248 SNPs had a MAF ≥5% in our cohort. GWAS was performed using PLINK v1.90b5.4 (Chang et al., 2015), treating inverse-normal rank-transformed viral mRNA levels as a quantitative trait, with population included as a covariate.

Quantification of Canonical Monocyte Subsets in EVOIMMUNOPOP Samples

FlowJo v10.6.1 software (Becton, Dickinson and Company. 2019) was used with the gating strategy depicted in **Figure S1** to quantify monocyte subsets for 174 EVOIMMUNOPOP donors. Population-level differences in proportion of canonical monocyte subsets were assessed using Wilcoxon Rank tests. Correlations of all 3 monocyte subsets with IAV mRNA levels were assessed using a linear model of the form

(1) $IAV \sim X + Pop$,

where '*IAV'* are IAV mRNA levels, '*X*' is the percentage of monocyte in the subset of interest, and '*Pop'* is and indicator variable separating AFB from EUB individuals. The contribution (η) of the intermediate monocyte subset to population differences in IAV levels was assessed by comparing the effect of population on IAV mRNA levels between model (1) and a model of the form

(2) $\underline{IAV \sim Pop}$

With η given by $\eta = 100 \times \left(1 - \frac{\beta_2}{\beta_1}\right)$, with β_i the effect of population on IAV levels in model (i).

Processing of Single-Cell RNA Sequencing Data

Basic pre-processing of the sequencing data was performed with CellRanger v3.0.2 (Zheng et al., 2017), including the *mkfastq*, *count*, and *aggr* commands. Default parameters and our combined human-IAV reference were used, and batch correction was disabled in the *aggr* command. Cell-containing droplets (*n*=132,130) were traced back to individual donors using two independent methods, Demuxlet and SoupOrCell, which capitalize on genetic variation in the sequencing reads (Heaton et al., 2020; Kang et al., 2018). Barcodes with ambiguous and/or non-concordant calls between the two programs were used to establish suitable QC metrics. We found that barcodes deemed as *doublets* (i.e. the droplet contained two or more cells originating from different donors) were more likely to be nearest-neighbors in a *knn*-graph with other *doublets* than assigned *singlets*. We used this feature to identify droplets presumed to contain two or more cells originating from the same donor; barcodes with > 5 *doublets* as nearest-neighbors were excluded from further analysis (**Figures S2A and S2B**). Additionally, droplets containing low-quality cells (i.e. damaged, dying) were excluded using the following thresholds: < 1500 total counts, < 500 genes, or > 50% mitochondrial gene content (**Figure S2C**). This QC resulted in a final data set of 96,386 single cells.

Transcriptomes (i.e. counts) were adjusted for the presence of ambient RNA with SoupX, (https://github.com/constantAmateur/SoupX/, accessed November 28, 2019) (Young and Behjati, 2020), using estimated contamination fractions (per 10X library) from SoupOrCell (Heaton et al., 2020). SoupX-adjusted counts were normalized using pool-based size factors followed by deconvolution as implemented in the scran R package (Lun et al., 2016b). Feature selection was performed by (i) constructing a mean-variance trend in the log-counts and retaining genes found to exhibit more variation than expected assuming Poisson-distributed technical noise, as implemented in the makeTechTrend and TrendVar functions from package scran (Lun et al., 2016b), and (ii) selecting genes expressed in at least 25 cells (*n*=22,603). The first 10 PCs of the data were retained for data visualization and clustering analyses. Graph-based clustering was performed by building the shared nearest-neighbor graph with the buildSNNGraph function from scran (Lun et al., 2016b) using a series of *k* values, and cell clusters were defined with the igraph Walktrap algorithm (Pons and Latapy, 2005). Similar clustering results were obtained based on the *knn*-graphs generated using *k*=25, 50, 75, and 100, and *k*=25 was used for all downstream analyses (**Figure S2D**). Cell types were predicted using SingleR and the built-in BlueprintEncodeData reference (Aran et al., 2019). Based on the clustering and cell-type predictions, we removed cells belonging to clusters associated with lymphoid cell types or low QC metrics from downstream analyses (**Figures S2E and S2F**).

Pseudo-bulk Estimates from scRNA-seq Data

Pseudo bulk estimates of IAV mRNA levels were computed by measuring, for each donor and time point, the mean percentage of reads of viral origin across all cells from the sample. At each time point, we then used a Pearson's correlation test to compare pseudo-bulk estimates for the 8 donors with IAV mRNA levels obtained in bulk data at 6hpi.

Accounting for Ambient RNA Contamination in scRNA-Seq Data and Assigning Cell States.

Droplet-based scRNA-seq methods capture ambient mRNAs present in the cell suspension in addition to cell specific mRNAs. To estimate which cells in our experiment were genuinely expressing mRNAs for *CD14*, *FCGR3A (CD16)*, and those originating from the virus, we implemented a two-step strategy utilizing the estimateNonExpressionCells function of the SoupX package (Young and Behjati, 2020). This function estimates whether each cell contains significantly more counts of a provided gene-set than would be expected under a Poisson model, given the estimated ambient RNA from its library of origin and the maximum contamination fraction. First, we used the viral genes to estimate the true maximum contamination fraction, based on the assumption that cells from the non-stimulated state should only contain viral reads from ambient mRNA captured in their droplets. To do so, we modified the estimateNonExpressionCells function to return *p*-values, and performed the test on each of our 13 libraries with a range of maximum contamination values from 1-50% (step of 1%) using the viral genes. We then computed FDR adjusted *p*-values for each maximum contamination value on the 88,559 high-quality, single monocytes. The number of non-simulated cells deemed to significantly express IAV transcripts (FDR<0.01) was used as a proxy for false positives. In examining the relationship between this number and the number of IAV-challenged cells found to significantly express viral transcripts at FDR<0.01 (**Figure 2D**), we found that a maximum contamination fraction of 10% resulted in a 1% false positive rate (defined as the percentage of non-stimulated cells from T_2-T_8 that were deemed to significantly express IAV transcripts). At this threshold, half of IAV-challenged cells showed evidence of viral transcription.

Assigning Cell States and Investigating Sources of Variability in IAV Levels

We used a maximum contamination fraction of 10% to test for significant expression of IAV transcripts in each cell (**Figures 2C** and **2E**). IAV-challenged cells that contained a significant amount of IAV transcripts were considered as infected, while the others were deemed

bystanders. To distinguish low from high IAV-transcribing cells, a Gaussian mixture model was fitted to the total percentage of viral mRNAs per cell across all infected cells, using the normalmixEM function from mixtools R package with *k*=2 (Benaglia et al., 2009). Each cell was assigned to the cluster with the highest posterior probability, and the cluster of cells with higher IAV content was annotated as high IAV-transcribing. We then studied how the relative proportions of each cell state varied across donors and time points by modeling either the percentage of infected cells (*INF*) or the percentage of high IAV-transcribing cells among infected cells (*HI*) with a linear model of the form

(3) INF ~ TP + Ind

and

(4) HI ~ TP + Ind

where *TP* and *Ind* correspond to factor variables reflecting the time point, and individual respectively. Type II Anova was used to infer the percentage of variance explained by each factor, and the significance of the effect.

Cell States Deconvolution from Bulk RNA Sequencing

To assess the percentage of total transcripts that originate from each cell state across the 199 IAV-challenged samples, we pooled cells from T_6 into 3 groups, based on their assigned cell-state (bystander, infected: high and low IAV-transcribing) and to which we added a 4th group containing all singlets that were either (i) assigned to cluster numbers 3, 8, 10, and 11 (dying cells) or (ii) discarded based on their high mitochondrial content or low read counts (dead cells). We then estimated pseudo-bulk profiles for each group by summing UMIs across all cells and computing the number of UMIs associated to each gene per million of sequenced UMIs. TPM profiles obtained from bulk data were then normalized to improve comparison with pseudo-bulk. Specifically, we first computed a global pseudo-bulk profile of the entire single cell dataset as the average of the pseudo bulk profiles from the 4 cell states (bystander, infected: high and low IAV-transcribing, or dying/dead), weighted by the percentage of UMIs they contribute to the overall pool of cells. To account for the difference in how gene expression is quantified between the two methods (3' end counts for scRNA-seq and full-length gene coverage for bulk RNA-seq), we computed for each gene i a normalization factor s_i given by

(5) $s_i = \log(\overline{TPM_i}) - \log(PB_i)$

where PB_i is the number of UMI per million for gene *i* in the global pseudo-bulk profile, and $\overline{TPM_i}$ is the average expression of the gene *i* in the 199 IAV-stimulated samples from the bulk RNA-seq data. For each gene, s_i was then subtracted from the log transformed TPM to yield a

normalized TPM profiles. We next applied DeconRNAseq (Gong and Szustakowski, 2013) to the normalized log TPM profiles from all individuals, using the log-transformed pseudo bulk profiles from the 4 cell states as a basis for deconvolution. Quality of the deconvolution was assessed using leave-one-out cross validation, based on the eight individuals for whom we had scRNA-seq data. Specifically, for each of these eight individuals, bulk mRNAs were decomposed using pseudo-bulk profiles recomputed based on the seven other individuals. The resulting proportions were then compared with the percentage of UMIs that originate in each cell-state in the scRNA-seq to assess the quality of the deconvolution. Comparisons between populations were performed using Wilcoxon rank tests.

The effect of the percentage of infected cells and percentage of high IAV-transcribing cells among infected cells on the total IAV mRNA levels were assessed by modeling

(6) IAV ~ INF + POP

and

(7) IAV ~ HI + POP

where *IAV* are the IAV mRNA levels across the 199 bulk mRNA samples, *INF* and *HI* are respectively the percentage of infected cells and the percentage of high IAV-transcribing cells among infected cells that we estimated from the deconvolution, and *POP* is a factor variable reflection the population (EUB or AFB). The fraction η of population differences attributable to difference in rate of infection was estimating by comparing model (6) with model (8) below

(8) IAV ~ POP

and computing $\eta = 100 \times \left(1 - \frac{\beta_6}{\beta_8}\right)$, with β_i the effect of population on IAV levels in model (i). To assess how the contribution of the percentage of high IAV transcribing cells to total IAV mRNA levels differed between populations, we used a linear model of the form

(9) IAV ~ HI + POP + HI:POP

and tested for significant effect of the interaction term *HI:POP* on *IAV* mRNA levels.

Characterizing Monocyte Subsets and Transcriptional Profiles from Single-Cell Data

Principal components analysis of 6,601 cells at T_0 was used to order monocytes along a differentiation axis separating *CD14*$^+$ cells from *CD16*$^+$ cells. We then computed the average percentage of classical and nonclassical monocytes obtained by FACS across the eight donors, weighting each individual by the number of high-quality cells in the scRNA-seq data at T_0. Based on these percentages (87.1% for classical and 7.6% for nonclassical), we annotated the monocytes on each side of the differentiation axis as classical and nonclassical, respectively, with the remaining 5.3% of monocytes being annotated as intermediates. Validity of our

approach was confirmed by correlating the proportion of monocytes assigned to each subset across the eight donors, with the percentage of classical, intermediate and nonclassical monocytes estimated by FACS.

Differential expression between subsets was assessed for the 4,859 genes expressed at an average logCount > 0.1 in any of the 3 subsets. Specifically, Wilcoxon rank tests were implemented in the scran package (Lun et al., 2016b), using the findMarkers function and blocking on donor. We considered genes to be differentially expressed (DE) between monocyte subsets when gene expression was significant at an FDR≤1% and $\log_2$FC>0.2. The 848 genes that differed between classical (CL) and nonclassical (NC) monocyte subsets were classified according to their behavior in intermediate monocytes (INT). They were either deemed 'similar to classical' (DE between INT and NC, but not between INT and CL), 'similar to nonclassical' (DE between INT and CL, but not between INT and NC), or 'intermediate' (all other cases).

At later time points, comparisons between *CD16*$^+$ and *CD16*$^-$ monocytes subsets were done based on 5,368 genes expressed with a normalized log count > 0.1 on average in either subset, in at least one condition and time point. For each subset, $\log_2$ fold change in gene expression relative to T_0 were correlated across times points. For each condition, differential expression between *CD16*$^+$ and *CD16*$^-$ cells was assessed with findMarkers (Lun et al., 2016b), based on Wilcoxon rank tests and blocking on donors and time points. Again, an FDR≤1% and $\log_2$FC>0.2 were required to define differentially expressed genes. To assess how CD16$^+$ status alters the infection of monocytes by IAV, we used logistic regression to model bystander/infected status as a function of *CD16*$^{+/-}$ status, while adjusting on donor, and time point (as factors).

Characterizing Subset-Specific Responses to IAV Challenge

To allow comparison between responses of *CD16*$^+$ and *CD16*$^-$ monocytes, 100 cells were subsampled from each subset and cell state, and at each time point. When subsampling, we ensured balanced representation of all donors across each monocyte subset and cell state, by using sampling weights that were inversely proportional to each donor representation in the original dataset. After sampling, a total of 6,669 genes with normalized log counts >0.1 on average in at least one group (cell-state x subset x time point) was selected for further analyses. For each monocyte subset, differences in expression between cell states (unexposed, bystander, infected) as well as between high- and low-IAV transcribing infected cells were performed using the findMarkers function from the scran package (Lun et al., 2016b) and blocked on time-point. For each comparison, genes were considered to be differentially expressed between cell states when gene expression was significant at an FDR=1% (Wilcoxon

rank tests) and the $\log_2$ fold change was > 0.2. In addition, for each comparison between cell states, we tested for differences in response between subsets using a linear model of the form:

$$(10)\ Expr_i \sim State_i + subset_i + State_i \,.\, subset_i$$

where $Expr_i$ is the expression of the gene being tested in cell i, $State_i$ is an indicator variable that distinguishes the two cell states being compared (e.g. unexposed and bystander), and $subset_i$ is an indicator variable that reflects the *CD16*$^{+/-}$ status of cell i. The 335 genes with significant interactions at a 1% FDR (for unexposed-bystander and unexposed-infected comparisons) were clustered using the hclust R function with method 'Ward.D2'. DynamicTreeCut algorithm (Langfelder et al., 2008) was used to identify eight major patterns of response to IAV.

Transcription Factor Enrichment Analyses

To estimate Transcription Factor (TF) activity and define TF-targets relationships, we ran the R SCENIC pipeline (Aibar et al., 2017) on the expression matrix (pre-normalization) on a random subsample of 4800 cells (100 cells from each donor at each time point and each condition, pre-exclusion of dying and contaminant cells) with default parameters. For each gene, motif-enrichment was considered for either *cis*-regulatory regions located <10kb from the TSS (distal regulatory elements), or between 500 bp upstream and 100 bp downstream of the promoter (proximal regulatory elements). To do so, motif-enrichment scores for all human genes (hg38 build, refseq_r80), were retrieved from https://resources.aertslab.org/cistarget and used as input for the Rcistarget package (Aibar et al., 2017).

Sets of high-confidence targets for the 113 TFs whose activity could be quantified by SCENIC were then extracted and used for enrichment analysis. For each gene module, TF enrichment was assessed using a Fisher's exact test with the 6,669 expressed genes as background (**Table S3**). Resulting *p*-values were adjusted using a global Benjamini-Hochberg correction for all eight modules and 113 TFs.

For each TF, with its targets enriched among one of the eight modules, TF activity inferred by SCENIC was used to test for subset-specific activity using a linear model of the form:

$$(11)\ TF_i \sim State_i + subset_i + State_i \,.\, subset_i$$

where TF_i is the activity of the TF being tested in cell i, $State_i$ is an indicator variable that distinguishes the two cell states being compared (e.g. unexposed and bystander), and $subset_i$ is an indicator variable that reflects the *CD16*$^{+/-}$ status of cell i. Average TF activity was then

computed for each cell state, subset and time point, and correlated with gene expression of the associated module, to assess the link between TF activation and the TF-target enriched modules.

Prediction of the Outcome of IAV Infection Using Basal Gene Expression

For each of the three monocyte subsets detected at basal sate, Kruskall-wallis test was used to search for genes whose expression levels significantly differ across donors. Within each monocyte subset, we then computed the average expression of each gene for all eight donors and correlated it with IAV mRNA levels obtained from the bulk RNA-seq data at 6hpi. Genes that differed in expression between donors (FDR≤1%), and passed a nominal *p*-value threshold of 0.01 for association with IAV levels in any of the 3 subsets, were selected for downstream enrichment analyses. For genes nominally correlated with viral mRNA levels, TF enrichment was assessed as previously using a Fisher's exact test with all 4,859 genes expressed at T_0 as background (**Table S4**), and Benjamini-Hochberg correction for all 113 TFs was applied.

Gene Ontology Enrichment Analyses

All Gene Ontology (GO) enrichment analyses were performed with GOSeq package using default settings (Young et al., 2010). Background gene sets consisted of all genes that had average log-normalized expression value > 0.1 in at least one of the groupings being examined, and are described in the text. Only enrichments significant at FDR≤5% are reported.