Title

Single-cell profiling of Ebola virus infection *in vivo* reveals viral and host transcriptional dynamics

Author list

Dylan Kotliar[1,2,3,4,16,*], Aaron E. Lin[2,3,5,16,18,*], James Logue[6], Travis K. Hughes[3,4,7,8], Nadine M. Khoury[2,3], Siddharth S. Raju[1,2,3], Marc H. Wadsworth II[3,7,8], Han Chen[9], Jonathan R. Kurtz[6], Bonnie Dighero-Kemp[6], Zach B. Bjornson[9], Nilanjan Mukherjee[9], Brian A. Sellers[10], Nancy Tran[3,8], Matthew R. Bauer[2,3], Gordon C. Adams[2,3], Ricky Adams[6], John L. Rinn[3,11], Marta Melé[3,12], Garry P. Nolan[9,17], Kayla G. Barnes[3,13,14,17], Lisa E. Hensley[6,17,*], David R. McIlwain[9,17,*], Alex K. Shalek[3,4,7,8,17], Pardis C. Sabeti[2,3,15,17], Richard S. Bennett[6,17]

[1]Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA. [2]FAS Center for Systems Biology, Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA. [3]Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA. [4]Harvard-MIT Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, MA 02142, USA. [5]Harvard Program in Virology, Harvard Medical School, Boston, MA 02115, USA. [6]Integrated Research Facility, Division of Clinical Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Frederick, MD 21702, USA. [7]Department of Chemistry, Institute for Medical Engineering and Sciences (IMES), Koch Institute for Integrative Cancer Research, MIT, Cambridge, MA 02142, USA. [8]Ragon Institute of MGH, Harvard, and MIT, Cambridge, MA 02139, USA. [9]Department of Pathology, Stanford University, Stanford, CA 94305, USA. [10]Trans-NIH Center for Human Immunology, Autoimmunity, and Inflammation, National Institutes of Health, Bethesda, MD 20814, USA. [11]BioFrontiers Institute, University of Colorado Boulder, Boulder, CO 80303, USA. [12]Life Sciences Department, Barcelona Supercomputing Center, Barcelona, Catalonia 08034, Spain. [13]Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health, Harvard University, Boston, MA, USA. [14]MRC-University of Glasgow Centre for Virus Research, Glasgow, UK. [15]Howard Hughes Medical Institute, Chevy Chase, MD 20815, USA. [16]These authors contributed equally. Authors listed alphabetically by last name. [17]These authors contributed equally. [18]Lead contact. *Correspondence: Dylan_Kotliar@hms.harvard.edu (D.K.), ALin@broadinstitute.org (A.E.L.), Lisa.Hensley@nih.gov (L.E.H.), McIlwain@stanford.edu (D.R.M.)

Summary

Ebola virus (EBOV) causes epidemics with high case fatality rates, yet remains understudied due to the challenge of experimentation in high-containment and outbreak settings. To better understand EBOV infection *in vivo*, we used single-cell transcriptomics and CyTOF-based single-cell protein quantification to characterize peripheral immune cell activity during EBOV infection in rhesus monkeys. We obtained 100,000 transcriptomes and 15,000,000 protein profiles, providing insight into pathogenesis. We find that immature, proliferative monocyte-lineage cells with reduced antigen presentation capacity replace conventional circulating monocyte subsets within days of infection, while lymphocytes upregulate apoptosis genes and decline in abundance. By quantifying viral RNA abundance in individual cells, we identify molecular determinants of tropism and examine temporal dynamics in viral and host gene expression. Within infected cells, we observe that EBOV down-regulates *STAT1* mRNA and interferon signaling, and up-regulates putative pro-viral genes (e.g., *DYNLL1* and *HSPA5*), nominating cellular pathways the virus manipulates for its replication. Overall, this study sheds light on EBOV tropism, replication dynamics, and elicited immune response, and provides a framework for characterizing interactions between hosts and emerging viruses in a maximum containment setting.

Introduction

Ebola virus (EBOV) is among the world's most lethal pathogens, with an estimated case fatality rate of 66% in the recent epidemic in the Democratic Republic of the Congo (Ilunga Kalenga et al., 2019; World Health Organization, 2019) and 40% in the 2013–2016 epidemic in West Africa (Lo et al., 2017). EBOV infection in humans causes Ebola virus disease (EVD), characterized by fever, malaise, muscle aches, and gastrointestinal distress, rapidly progressing to coagulopathy, shock, and multi-organ failure (Malvy et al., 2019). While recently developed vaccines (Kennedy et al., 2017) and monoclonal antibody therapeutics (Mulangu et al., 2019) have shown great promise for preventing and treating EVD, case fatality rates in treated patients still exceed 30%, highlighting the need for further research into disease pathogenesis.

Studies of EVD pathogenesis, while paramount, face numerous logistical challenges which have limited their scope relative to studies of other pathogens. Experiments involving live EBOV require maximum containment (e.g., biosafety level 4 [BSL-4]) and therefore are restricted to a small number of highly specialized research facilities. *In vivo* studies are especially challenging: human EVD is difficult to study in the midst of deadly outbreaks in resource-limited settings, necessitating animal models. However, commonly used laboratory mouse lines are resistant to naturally occurring EBOV isolates, limiting their utility for research (Bray, 2001; Rasmussen et al., 2014). Moreover, rodents and other small animal models such as ferrets lack the primate-specific NPC1 genotype (Diehl et al., 2016), the key cellular receptor for EBOV entry (Carette et al., 2011; Côté et al., 2011), and do not always recapitulate human EVD-like disease (Bray et al., 2001; Geisbert et al., 2002). EVD in nonhuman primates (NHPs) most closely resembles human EVD (Bennett et al., 2017; Geisbert et al., 2015; St Claire et al., 2017), but NHP studies are often limited to small sample sizes, which reduces power to identify statistically significant trends and to discern meaningful inter-individual variability.

The two predominant approaches to studying EVD – analyzing infected cells in culture and infected animals *in vivo* – have revealed important, if somewhat contradictory, aspects of how EBOV impacts the host immune system. In cell culture, EBOV infects myeloid cells, potently inhibiting both production of type 1 interferon (Basler et al., 2003; Gupta et al., 2001; Harcourt et al., 1998) and signal transduction downstream of interferon receptors (Harcourt et al., 1999; Kash et al., 2006; Leung et al., 2006; Reid et al., 2008). Under-activation of this key innate antiviral response hinders the ability of antigen-presenting cells to activate the adaptive immune system to combat infection (Bosio et al., 2003; Lubaki et al., 2013; Mahanty et al., 2003), a key determinant of fatal outcomes (Baize et al., 1999) and could be due to reduced presentation of viral proteins by antigen-presenting cells (Lüdtke et al., 2016). In contrast to these culture-based findings, EVD *in vivo* is characterized by high fever and dramatic up-regulation of hundreds of interferon stimulated genes (Caballero et al., 2016; Liu et al., 2017; Reynard et al., 2019; Speranza et al., 2018), correlating with the release of dozens of inflammatory cytokines (Caballero et al., 2016; Reynard et al., 2019; Wauquier et al., 2010), suggesting that an aberrant

over-activation of innate and adaptive immunity underlies much of EVD pathology, rather than solely virus-mediated cytotoxicity (Geisbert et al., 2003a, 2003b).

High-throughput single-cell technologies, such as single-cell RNA-sequencing (scRNA-Seq) and protein quantification by CyTOF (Bendall et al., 2011), have made it possible to analyze the response of individual cells to viral infection at unprecedented resolution (Hamlin et al., 2017; Hein and Weissman, 2019; Newell et al., 2012; O'Neal et al., 2019; Russell et al., 2018, 2019; Steuerman et al., 2018; Zanini et al., 2018a; Zhao et al., 2020). By generating mRNA or protein profiles for thousands of cells in a sample, these methods can quantify the cell-type composition and expression programs of individual cells -- signals that are obscured in bulk measurements. By quantifying viral RNA within individual cells, scRNA-Seq allows comparison of gene expression between infected and uninfected cells in a diseased host (*i.e.*, bystander cells), which can yield a far more nuanced view of host (Steuerman et al., 2018) and viral (Hein and Weissman, 2019) gene expression within infected cells. Further, this approach can be used to disentangle the direct effects of EBOV infection within a cell from the effect of the inflammatory cytokine milieu. However, many scRNA-Seq technologies require droplet generator devices and inactivation protocols that can be challenging to establish in a maximum containment facility. As a result, such approaches have yet to be applied to a risk group 4 (RG-4) pathogen such as EBOV. Furthermore, high-volume exhaust, superheated components, and other aspects of CyTOF instrumentation make these devices incompatible with installation in maximum containment facilities (Logue et al., 2019), necessitating development of new protocols compatible with sample inactivation to study RG-4 pathogens.

Here, we describe the first investigation of an RG-4 agent under maximum containment with high-dimensional single-cell technologies. We apply CyTOF and Seq-Well--a portable single-cell RNA-seq platform (Gierahn et al., 2017; Hughes et al., 2019)--to a combined total of 90 peripheral blood mononuclear cell (PBMC) samples (90 by CyTOF, 28 by Seq-Well) collected from 21 rhesus monkeys prior to infection or at multiple timepoints following lethal EBOV challenge *in vivo*. We further inoculated PBMCs with EBOV *ex vivo*, using pre-defined experimental parameters, and profiled their gene expression with Seq-Well. These data allow us to dissect host-virus interactions and comprehensively catalog changes in cell-type abundance and cell state over the course of EVD. Moreover, as EBOV harbors an RNA genome and transcribes polyadenylated mRNAs, we were able to detect viral RNA within individual cells, allowing us to define EBOV tropism with high resolution and identify EBOV-associated transcriptional changes in putative pro- and antiviral genes.

We find that EVD leads to widespread changes in the circulating monocyte populations *in vivo*, both in NHPs as well as in acute human infections, with replacement of conventional monocyte subsets with a highly proliferative monocyte precursor population and a macrophage-like population that is enriched for EBOV-infected cells. Furthermore, by comparing infected and uninfected bystander monocytes, we resolve the apparent contradiction between *in vivo* and *in vitro* studies of EBOV. We find that bystander cells of all major immune cell types express an interferon response program, but that this response is suppressed specifically within infected

monocytes *in vivo*, consistent with previous studies in culture. In addition to down-regulating host antiviral genes, we observe that EBOV drives the up-regulation of candidate pro-viral genes, such as *DYNLL1* and *HSPA5*, within infected cells. Taken together, this dataset constitutes a unique resource for the study of EBOV, enabling the study of host immune response in infected and bystander cells across cell types, in natural (*in vivo*) and experimentally controlled (*ex vivo*) contexts.

Results

Single-cell characterization of RNA and protein expression in circulating immune cells from Ebola virus infected rhesus monkeys

To comprehensively profile EBOV-induced immune dysfunction *in vivo*, we collected peripheral immune cells from rhesus monkeys prior to infection and at multiple days post-infection (DPI), corresponding to several stages of acute EVD (**Figure 1**). Cohorts of ≥3 nonhuman primates (NHPs) were sacrificed as baseline uninfected controls (B), at pre-defined DPI, or upon reaching pre-determined humane euthanasia criteria. These cohorts were recently characterized for viral load, clinical score, blood chemistry (Bennett et al. in submission) and liver pathology (Greenberg et al., 2020). Viral load first became detectable in all animals on DPI 3, preceding detectable clinical signs (*e.g.*, fever) by 1–2 days (**Figure 2A**). Clinical signs of EVD progressed until humane euthanasia criteria were uniformly reached between DPI 6–8 **(Figure S1A**). For each NHP, cells collected multiple times throughout disease were used for CyTOF, while cells collected prior to infection and at sacrifice were used for CyOF and Seq-Well (**Figure S1B**).

After standard quality control filters (**Materials and Methods**), we obtained single-cell transcriptomes and 42-protein CyTOF profiles from ~58,000 and ~15,000,000 PBMCs, respectively. We visualized these data with uniform manifold approximation and projection (UMAP) non-linear dimensionality reduction (Becht et al., 2018; McInnes et al., 2018) (**Figures 2B, 2C, 2E, 2F, and 2H**). Unsupervised clustering of either the transcriptomes or a down-sampled set of 1,100,000 protein profiles (**Materials and Methods**) yielded clusters that could be readily identified as the major circulating immune cell types using well-known RNA and protein markers (**Figures 2B, 2C, S2A, and S2B**). The CyTOF-based clustering gave concordant results with an approach based on manual gating of conventional cell-type marker genes (**Figures S3A and S3B, Materials and Methods**). After batch correction of the CyTOF data and integration of the transcriptomes to adjust for technical sources of variation **(Materials and Methods**), samples were well-distributed across cell-type clusters (**Figures S2C–S2F**) but separated by DPI (**Figures 2E and 2F**), suggestive of dynamic cell states over the course of disease progression. By sub-clustering within broad cell-type categories, we further identified cell subtypes based on the expression of identifying marker genes (**Figure S4)**.

Cell-type abundance, proliferation, and EBOV infection rates vary throughout EVD

In addition to the major PBMC cell types, a cluster of immature neutrophils emerged during EVD, marked by high gene expression of *CD177* and *SOD2*, and protein expression of CD66 and CD11b. Though neutrophils are typically removed during density-based PBMC isolations, immature neutrophils (*i.e.*, band cells) – which are less dense than mature polymorphonuclear neutrophils – can be released from the bone marrow and co-isolate with PBMCs in infectious and autoimmune inflammatory conditions (Carmona-Rivera and Kaplan, 2013; Darcy et al., 2014; Deng et al., 2016), including during EVD (Eisfeld et al., 2017). Neutrophils were almost entirely absent from baseline samples in our data but constituted a high proportion of cells in late EVD samples (scRNA-Seq: 0.2% of baseline cells compared to 65.1% of late EVD cells; CyTOF: 9.3% of baseline compared to 49.8% of late EVD; **Figures S5A and S5B**), which supports the hypothesis that band cells are released into the periphery from the bone marrow in response to cytokines elicited during EVD.

Next, we quantified absolute abundance of each cell type over the course of EVD by combining CyTOF data with complete blood counts (CBC) (**Materials and Methods**) (Bennett et al. in submission). CBC provided direct neutrophil, lymphocyte, and monocyte abundances, and we integrated CBC and CyTOF data to obtain differential abundances of the lymphocyte cell types. Cell-type percentage estimates based on CyTOF were in general agreement with those based on scRNA-Seq (**Figure S5A**).

In agreement with previous NHP studies (Ebihara et al., 2011; Fisher-Hoch et al., 1985), there was a >5-fold increase in neutrophil abundance by DPI 4 relative to baseline, before levels returned to baseline in late EVD ($p < .05$ for DPI 3–4, $p = 0.059$ for DPI 5, Wilcoxon signed-rank test, **Figures 2D and S5B**). Also consistent with previous studies, we observed a marked decrease in lymphocyte abundance with NK cells declining one day before the other cell types ($p < 0.05$ on DPI 3–6 for B, NK, CD8+ T, and CD4+ T; except for CD4+ T on DPI 5, Wilcoxon signed-rank test). Interestingly, all lymphocyte populations slowly recovered after DPI 4 (**Figure S5B**). Monocyte abundance initially increased >2-fold before declining precipitously between DPI 4 and 5.

Changes in circulating cell-type abundance could reflect cell proliferation and/or death, as well as movement of cells into and out of bone marrow, lymph, and tissues. While we were unable to directly quantify rates of death or movement between different compartments, we estimated the fraction of actively dividing cells using the proliferation marker Ki67 (encoded by the gene *MKI67*) in both the CyTOF and scRNA-Seq data, and found good agreement between the two modalities (**Figure S5C**).

The fraction of Ki67+ monocytes increased dramatically from 17% at baseline to 56% at DPI 5 and remained >40% for the remainder of disease ($p = 1.1\text{x}10^{-5}$ rank-sum test of DPI 5-8 vs. baseline samples), suggesting an increase in proliferation (**Figures 2G and S5D**). By contrast,

neutrophil proliferation remained roughly constant (**Figures 2G and S5D**) despite the dramatic changes in abundance (**Figures 2D and S5B**), further evidence that immature neutrophils were released from the bone marrow during disease (Summers et al., 2010). Intriguingly, the fraction of dividing T and NK cells stayed relatively constant for most of the time course but increased dramatically on DPI 8 for both of the NHPs (out of 6 total) that survived until then, both by RNA (**Figure S5C**) and protein levels (p = 0.022 rank-sum test of DPI 8 vs. baseline for NK, CD8+ T, and CD4+ T, **Figure S5D**). Proliferation is a core component of effective T-cell mediated viral clearance, but requires time for activated T cells to accumulate; the observation that significant proliferation of circulating T cells only occurred in the 2 animals that survived until the latest DPI suggests that those animals may have begun to mount a T-cell response.

Not all cell types support EBOV entry and replication; here, we were able to identify which cells were infected *in vivo* using scRNA-Seq because EBOV has an RNA genome and produces poly-adenylated mRNA transcripts (**Figure 2H**). However, uninfected cells may also contain EBOV reads due to ambient RNAs that contaminate single-cell profiles (Fleming et al., 2019; Young and Behjati, 2018). We therefore developed a statistical approach to identify infected cells as those that contain more EBOV transcripts than would be expected by chance, based on the relative abundance of EBOV-mapped transcripts in a cell and the amount of ambient RNA contamination (**Materials and Methods**). This allowed us to control the false positive rate (FPR) at a pre-specified level while maximizing power to call infected cells. At a FPR of 1%, we were well-powered to identify an infected cell when ≥1% of its transcripts mapped to EBOV, and estimated an average sensitivity of 51% when ≥0.1% of cellular transcripts derived from EBOV (**Figure S2G**) though the sensitivity for a given cell depends on read depth and other parameters (**Materials and Methods**). In addition, we spiked uninfected Madin-Darby canine kidney (MDCK) cells into a subset of PBMC samples to serve as a negative control (**Table S1, Materials and Methods**).

Monocytes comprised the main infected cell population *in vivo*, first detectable at DPI 4, with an increasing fraction of infected monocytes thereafter (**Figure 2I**). Consistent with previous studies, T cells, B cells, and neutrophils were not identified as infected more often than would be expected by chance (1% FPR threshold), nor more often than MDCK control cells. We did not observe any infected plasmacytoid (pDC) or conventional dendritic cells (cDC) in circulation, though infected DCs have been observed in culture and in lymph nodes (Geisbert et al., 2003c) (see **Discussion**).

Interferon response drives gene expression programs across multiple cell types

Having examined temporal shifts in the frequency of each immune cell type, we next sought to comprehensively catalog changes in their respective gene expression profiles throughout EVD. To increase statistical power to detect differentially expressed genes, we grouped cells into EVD stages based on clinically relevant phenomena: "incubation" which precedes detectable viral load or clinical signs (DPI 1 and 2; CyTOF only), "early" when there is detectable viral load but no clinical signs (DPI 3), "middle" when there is detectable viral load and clinical signs (DPI 4

and 5), and "late" when animals uniformly reached human euthanasia criteria (DPI 6–8) (**Figure 1**).

We compared transcriptomes of cells from each EVD stage to baseline for each cell type individually (**Materials and Methods**). This identified 1,437 differentially expressed genes with an FDR corrected *q*-value < 0.05 and a fold-change of greater than 30% in at least one cell type and stage (**Table S2**). To identify patterns of gene expression associated with cell type and time, we performed unsupervised clustering of the differential expression signatures and identified 11 modules of genes sharing similar patterns of changes (**Figures 3A and 3B**, **Table S3**, **Materials and Methods**). We excluded neutrophils, pDCs, cDCs, and plasmablasts because of small sample sizes. Three modules which we term "Global" were broadly up or down-regulated across cell types, and the remaining modules were cell-type specific.

The “Global up” module contained 136 genes, consisting mostly of regulators and targets of the interferon (IFN) alpha (α) and gamma (γ) signal transduction cascade such as *STAT1*, *IRF7*, *MX1*, and *ISG15*. Gene sets labeled “response to interferon alpha”, and “response to interferon gamma” were significantly enriched in this module (IFNα: OR = 69.5, $q = 8\text{x}10^{-39}$; IFNγ: OR = 45.9, $q = 1\text{x}10^{-39}$; Fisher's exact test; **Table S3**). The emergence of an IFN response was consistent with another observation: *IFNγ* mRNA rises >10-fold in CD8+ T-cells from an average of 0.4 transcripts per ten thousand (TP10K) at baseline to 5.0 at mid stage EVD. Concurrently, type 1 IFN (α/β) mRNAs rose from undetectable at baseline to 0.03–0.05 TP10K at late EVD in monocytes, along with a large number of other cytokines (**Figure S6A**). However, the increase in type 1 IFN mRNAs was not statistically significant, as IFN mRNAs are expressed transiently (Lin et al., 2011). To further characterize the dynamics of the IFN response, we assigned an interferon stimulated gene (ISG) score to each cell, reflecting the average expression of literature-annotated ISGs that overlap with the "Global up" module (**Table S3, Materials and Methods**). The median ISG score increased substantially throughout EVD across each cell type ($p < 1\text{x}10^{-5}$ for all cell types and periods, rank-sum test, **Figure 3C**).

As there is substantial overlap between the genes stimulated by IFNα and IFNγ, we sought to determine if one cytokine predominated, or if both acted independently. We therefore identified genes that were annotated as regulated by IFNα but not IFNγ (*i.e.*, uniquely IFNα-regulated) and vice versa. Both uniquely IFNα- and uniquely IFNγ-regulated genes were significantly enriched in the “Global up” module ($q < 0.01$, **Table S3**), with a larger fold-change for the uniquely IFNα-regulated genes (IFNα OR = 20.9, IFNγ OR = 16.6). This pattern held true for each cell type and EVD stage separately (**Figure S6B**). These results suggest that both IFNα and IFNγ substantially and independently influenced the gene expression profiles of circulating cells during EVD.

The "Global late down" module contained 144 genes that were predominantly down-regulated across cell types during late EVD. It contained numerous regulators of translation initiation and elongation (e.g. *EEF2*, *EEF1D*, *EIF3E*, and *PABPC1*; REACTOME_TRANSLATION gene set enrichment $q = 5.2\text{x}10^{-7}$, **Table S3**), which is consistent with a core antiviral function of IFN

being to down-regulate translation (Li et al., 2015). The “Global down” module contained 156 genes, and similar to "Global late down", included several other genes involved in translation (e.g. *QARS*, *NOP53*, and *EIF3D*). In addition, this module was most significantly enriched for the HALLMARK_MITOTIC_SPINDLE gene set (q = 8.5x10^{-7}) suggesting a global down-regulation of cell cycling upon activation.

Cell-type and temporally-specific modules underlie cell states related to clinical phenomena in EVD

After elucidating the global effects of EBOV infection on immune cells, we next investigated the transcriptional responses specific to each cell population.

The 2 modules “B/T early up” and “Lymph late up” reflect changing gene expression states of lymphocytes at different stages of acute EVD. “B/T early up” is strongly associated with the gene set HALLMARK_TNFA_SIGNALING_VIA_NFKB (q = 1.3x10^{-9}) and is characterized by many lymphoid activation genes including the canonical marker *CD69* (Testi et al., 1994), *CD48* (McArdel et al., 2016), and the transcription factor *FOS* (Foletta et al., 1998). This module is unlikely to represent antigen-dependent activation via the BCR/TCR as it occurs in most lymphocytes and does not coincide with proliferation. Indeed, several of the top up-regulated genes, such as *GADD45B* and *ZFP36L2*, are associated with growth arrest. In addition, the 5th most enriched gene set in the “B down” module is “REACTOME_ANTIGEN_ACTIVATES_B_CELL_RECEPTOR_LEADING_TO_GENERATION_OF_SECOND_MESSENGERS” (q = 0.00017) suggesting a reduction in BCR activation generally. Thus, the “B/T early up” module likely represents a cytokine-mediated, non-antigen-dependent activation of lymphocytes.

The “Lymph late up” module is up-regulated in late EVD across all lymphocyte cell types. The top associated gene sets implicate DNA repair (PUJANA_ATM_PCC_NETWORK, q = 0.00031) and apoptosis via TRAIL (HAMAI_APOPTOSIS_VIA_TRAIL_UP, q = 0.00032). This latter gene set is potentially consistent with previous reports of T-cell apoptosis in EVD (Geisbert et al., 2000; Iampietro et al., 2017; Wauquier et al., 2010) and with the lymphopenia in our dataset (**Figure 2D**).

The “NK mid up” module is highly specific to NK cells during early EVD (**Figure 3A**) and is most enriched for “MARSON_BOUND_BY_FOXP3_STIMULATED” (q = *0.07*1) and “BIOCARTA_CDC42RAC_PATHWAY” (q = 0.08, **Table S3**). *FOXP3* expression is characteristic of invariant NK cells (Engelmann et al., 2011) that secrete a wide variety of cytokines in response to stimulation (Krovi and Gapin, 2018). The CDC42/RAC pathway is essential for the polarization of cytolytic granules in NK cells, a requirement for effective cytotoxicity (Sinai et al., 2010; Tybulewicz and Henderson, 2009). This suggests that NK cells became activated with increased cytotoxicity in the mid EVD stage, but not in late EVD.

Monocytes express reduced MHC class II mRNAs and proteins independent of infection status

Monocytes were of particular interest. In addition to being the preferred target of EBOV (**Figure 2I**), they had far more significant gene expression changes during EVD than the other cell types. 1,020 genes (11.6% of total genes tested) were differentially expressed in monocytes at one or more EVD stages versus baseline, compared to 505 genes (6.6%) for B cells, the cell type with the second most differentially expressed genes. We therefore focused our attention on characterizing monocytes in detail.

One prominent feature of the monocyte differential expression profile was the striking down-regulation of several MHC class II (MHC-II) genes by mid and late EVD (**Figure 3D**). Monocytes and professional antigen-presenting cells display viral antigens on MHC-II proteins at the cell surface to stimulate the adaptive immune response. While IFNγ typically up-regulates MHC-II gene and protein expression (Steimle et al., 1994), we observed decreased MHC-II on monocytes despite elevated *IFNγ* mRNA levels in T cells (**Figure S6A**) and widespread IFNγ transcriptional response in monocytes (**Figure S6B**). Previous reports have described loss of HLA-DR, one of the 4 MHC-II proteins, during EBOV infection of monocytes *ex vivo* (Hensley et al., 2002), in experimentally infected NHPs (Menicucci et al., 2017) and in human EVD cases (Lüdtke et al., 2016), similar to observations of reduced HLA-DR on monocytes in patients experiencing septic shock (Wolk et al., 2000). However, the specific MHC genes affected, the cell-type specificity, temporal dynamics, and relationship with EBOV infection status, have not been previously described.

We observed widespread changes in levels of MHC genes throughout EVD (**Figure 4A**). The most striking decreases occurred in MHC-II genes of monocytes (>5-fold for *DPA*, *DPB*, and *DRA* by late EVD, $q < 1\text{x}10^{-21}$ for all MHC-II genes), with smaller effect-size changes in MHC-I genes (<1.7-fold increase for *A*, *A3*, and *B* at late disease, $q < 5\text{x}10^{-4}$). B cells displayed modest reductions in MHC-II genes as well (>1.9-fold for *DPA*, *DPB*, *DRA*, and *DQA1* at late disease, $q < 1\text{x}10^{-22}$). pDCs and cDCs showed no statistically significant reduction of any MHC-II gene ($q > 0.05$) but our dataset contained few DCs (**Table S1**), so we had less power to detect these effects. We observed a corresponding pattern in the protein levels by CyTOF: in monocytes, HLA-DR protein levels decreased to a greater extent than in the other cell types ($p < 1\text{x}10^{-61}$ for monocytes in early, mid, and late stages, rank-sum test, **Figure 4B**), with a more modest reduction of HLA-DR in B cells at DPI 5–8 ($p = 0.0012$, rank-sum test) (**Figure S6C**). This phenomenon held true for each individual NHP; even as monocytes became activated, demonstrated by up-regulation of the canonical activation marker CD38 (Amici et al., 2018) (**Figures 4C and S3D**), they showed dramatic down-regulation of average HLA-DR protein expression in monocytes at DPI 5–8 versus baseline ($p = 9.5\text{x}10^{-7}$, rank-sum test) (**Figure 4D**).

Reduced MHC-II expression in monocytes was not a direct consequence of EBOV infection. Only a small (~5%) percentage of monocytes were infected at mid EVD (**Figure 2I)**, suggesting that the striking decreased levels of MHC-II genes was unlikely to be specific to infected cells. Moreover, we confirmed that the average expression of MHC-II genes was comparable or even higher in infected cells relative to uninfected cells in NHPs with EVD (*i.e.* bystanders) (**Figure 4E**). Thus we conclude that the MHC-II decrease in monocytes observed in EVD is independent of direct viral infection.

To identify co-regulated genes as well as possible drivers of MHC-II down-regulation, we looked for other genes with expression correlated with MHC-II in monocytes (**Materials and Methods**). Many of the most correlated genes were functionally involved in the antigen presentation pathway, such as *CD74* (Spearman $\rho = 0.42$, $p = 1.8 \times 10^{-296}$), which chaperones MHC-II to the endosome and prevents premature binding of antigen (Schröder, 2016); *LGMN* (Spearman $\rho = 0.41$, $p = 8.1 \times 10^{-286}$), a protease that cleaves proteins to facilitate peptide presentation on MHC-II (Dall and Brandstetter, 2016); and, *B2M* (Spearman $\rho = 0.33$, 2.6×10^{-175}), a component of the MHC class I complex (**Figure 4A**). In addition, one of the most associated genes was *ZFP36* (Spearman $\rho = 0.43$, $p < 1 \times 10^{-296}$), a protein that directly regulates mRNA stability and turnover of MHC-II and other immune-related RNAs (Pisapia et al., 2019). These findings suggest that MHC-II and other genes involved in antigen presentation may be part of a single transcriptional module, co-regulated by *ZFP36* and/or other genes.

Characterization of differentially expressed genes between infected and bystander monocytes

Next we characterized genes that were differentially expressed between infected and bystander monocytes, as these could represent host entry factors, restriction factors, or genes that are regulated by infection within a cell. For this and all subsequent differential expression analyses, EBOV transcripts were excluded from the denominator when normalizing cells by library size, to avoid a bias in the estimated expression levels of host genes in infected cells (**Materials and Methods**). We identified 505 genes that were differentially expressed between infected and bystander monocytes ($q < 0.05$) of which 276 changed by more than 30% (**Figure 5A, Table S4**). 181 (18%) of the 1,020 genes that were differentially expressed in monocytes at one or more stages of EVD were also differentially expressed in infected monocytes relative to bystanders.

We observed that the differentially expressed genes fell into 3 broad categories--genes associated with monocyte subtypes, genes associated with monocyte-to-macrophage differentiation, and interferon stimulated genes (ISGs)--which we explore in the sub-sections below.

Emergence of CD14- CD16- immature monocyte precursors suggests emergency myelopoiesis

A notable feature of the differential expression profile was that *CD14* and *FCGR3* (which codes for *CD16)* were over-expressed in infected monocytes relative to bystanders. These two genes define classical (CD14+) and non-classical (CD16+) monocytes respectively (Kapellos et al., 2019), which comprise the dominant monocyte subsets in the blood of healthy individuals. Classical monocytes are highly phagocytic scavenger cells, while non-classical monocytes are involved in complement and antibody-mediated phagocytosis. To understand the role of monocyte subsets and other genes affecting monocyte heterogeneity across EVD, we visualized the transcriptome profiles of just the monocytes in two dimensions using UMAP. The monocytes separated by DPI and by EBOV infection status, consistent with changing cell states over the disease course (**Figure 5B**). At baseline, monocytes separated into 2 distinct clusters, marked by high expression of *CD14* (**Figure 5B**, bottom half of the blue lobe of DPI panel) or *CD16* (**Figure 5B**, top half of the blue lobe of DPI panel), consistent with the conventional subtyping.

However, monocyte subsets changed dramatically during EVD, with the decline of single-positive CD14+ and CD16+ monocytes, and the corresponding rise of 2 unusual populations: a large population of CD14- CD16- cells (double negatives [DNs]) and a smaller population of CD14+ CD16+ cells (double positives [DPs])**.** To visualize the dynamics of these populations, we plotted smoothed gene expression of *CD14* and *CD16* for each EVD stage (**Figure 5C, Materials and Methods**). While 87.6% of cells fell into single CD14+ or CD16+ bins at baseline, this dropped to 33.8%, 35.6%, and 15.2% in the early, middle, and late stages of EVD, respectively. We confirmed a corresponding loss of single-positive CD14+ or CD16+ monocytes and a gain of DNs and DPs at the protein level by CyTOF (**Figure 5D**). In both RNA and protein measurements, these two populations began declining on DPI 3.

At late EVD, the most frequent monocyte population was CD14- CD16- DN cells, which rose to make up 70.4% (by scRNA-Seq) or 56.7% (by CyTOF) of the monocytes. Given that they expressed neither of the canonical (CD14/CD16) monocyte marker genes, we confirmed that their overall gene expression profiles were most correlated with single-positive monocyte populations in late EVD (Pearson correlation R=0.80 CD14+, R=0.58 CD16+) and CD14+ monocytes at baseline (R=0.40), and were less correlated with neutrophils, DCs, and lymphocytes in late EVD (R=0.28, R=0.39, R=0.14, respectively) (**Figure S7A**). DNs first emerged on DPI 3, coinciding with the 2-fold increase in monocytes we observed on that day (**Figure 2D**). The DN population is highly proliferative; while 0% of monocytes at baseline expressed moderate levels of Ki67 (smoothed log TP10K > 1), over 37% of DN monocytes expressed Ki67 beyond this threshold by late EVD (**Figures S7B and S7C**). Therefore, this population underlies the increased monocyte replication rates observed in the middle stage of EVD (**Figure 2G**).

To determine if a corresponding increase in DNs occurs in acute EVD in humans, we re-analyzed published CyTOF data from the 2013–2015 outbreak of 4 acute EVD cases that were treated at the Emory University Hospital Serious Communicable Diseases Unit (McElroy et al., 2020) (**Materials and Methods**). All of the human cases showed a congruent pattern to the

NHP data, with loss of conventional CD14+ and CD16+ single-positive monocytes and an emergence of proliferative (Ki67^{hi}) DN monocytes (**Figures 5E and S7D**). Then, the DNs disappeared and were replaced by the conventional CD14+ and CD16+ single positive monocytes at later days post symptom onset, as the human cases entered the convalescent period. Thus, the emergence of circulating DNs and the loss of conventional circulating monocyte subsets is a feature of human clinical cases as well as our NHP model of lethal EVD.

The presence of proliferating DN monocytes was surprising because mature monocytes in circulation are believed to be non-dividing (van Furth et al., 1979). However, infectious and neoplastic diseases produce cytokines such as M-CSF that induce the release of proliferating immature myeloid cells from the bone marrow, a process known as emergency myelopoiesis (Chiba et al., 2018; Cuenca et al., 2015; Sayed et al., 2019). We therefore hypothesized that the DN population may reflect immature myeloid cells released from the bone marrow by emergency myelopoiesis.

If DNs represent the product of emergency myelopoiesis, we might expect their gene expression profiles to be more similar to bone marrow resident monocyte precursors than circulating monocytes. To test this, we compared our monocyte populations against a reference scRNASeq dataset of bone marrow monocyte precursors (BM-MPs) from healthy human bone marrow (Hay et al., 2018) and mature monocytes from human PBMCs (**Figures S7E–H, Materials and Methods**). Notably, the BM-MPs showed lower expression of *CD14* and *FCGR3A* (the human *CD16* gene) than mature monocytes, consistent with the diminished expression of these genes in DNs relative to baseline single-positive monocytes in the NHP data (**Figure S7I**). In addition, the BM-MPs showed higher expression of *MPO*, *AZU1*, *S100A8*, and *S100A9* than mature monocytes (**Figures S7I and S7J**), consistent with the observation that these genes are expressed at higher levels in DNs relative to baseline monocytes (**Figure S7K**). To formally test whether our monocyte populations were more similar to the mature PBMCs or the BM-MPs, we identified the nearest neighbor of each NHP monocyte in the reference dataset. As expected, DNs from mid and late EVD were significantly more likely to be matched with BM-MPs (32% at mid, 23% late), than single positive CD14+ or CD16+ monocytes, which almost exclusively were assigned to the corresponding circulating monocyte populations in the reference data (<10% assignment to BM-MPs assignment to BM-MPs for all non-DN populations and stages, except for CD16+ at late EVD which was 15%) (**Figure 5F**).These findings suggest that DNs represent immature monocytes released from the bone marrow in response to the EVD cytokine milieu.

Monocytes expressing markers of macrophage differentiation are enriched for EBOV infection

In addition to DNs, we also observed CD14+ CD16+ DP cells, which rose to make up 20.4% of the monocytes by mid EVD (**Figure 5C**). A similar increase in DP monocytes has been observed in sepsis (Fingerle et al., 1993; Nockher and Scherberich, 1998) and other viral infections (Michlmayr et al., 2018; Zanini et al., 2018a). We found that the DP population harbored a disproportionately high percentage of EBOV-infected cells (**Figure 5G**), consistent

with the fact that *CD14* and *CD16* were both independently higher in EBOV infected monocytes than in bystanders (**Figure 5A**). At the late infection timepoints, 22.1% of DPs were infected compared to only 1.74% of DNs. Thus, the differential expression of *CD14* and *CD16* in infected cells results from increased infection of the DP cells, rather than increased expression of *CD14* on classical and *CD16* on non-classical monocytes.

We noticed that the differentially expressed genes between infected and bystander monocytes (**Figure 5A**, **Table S4**) *and* between DP and DN monocytes (**Figure S8A**, **Table S5**), were enriched for monocyte-to-macrophage differentiation associated genes, including known EBOV entry factors. It has been previously observed that freshly isolated monocytes are largely refractory to EBOV infection in cell culture, but that EBOV entry factors are up-regulated during *in vitro* macrophage differentiation, allowing increased infection (Martinez et al., 2013). *In vivo*, we observed higher levels of macrophage differentiation markers such as *NR1H3*, *ADAMDEC1*, and several cathepsins in infected cells relative to bystanders. Among these genes, the known EBOV entry factors cathepsin L (*CTSL*) and B (*CTSB*), and *GNPTAB* (Carette et al., 2011; Gnirß et al., 2012) were all expressed at significantly higher levels in infected cells than bystanders ($q = 6.7\text{x}10^{-9}$, $3.8\text{x}10^{-7}$, and $2.1\text{x}10^{-3}$, respectively). By contrast, the cellular receptor *NPC1* was not significantly differentially expressed, suggesting that natural variability in the mRNA abundance of *NPC1* likely does not influence EBOV infectivity within circulating monocytes in rhesus monkeys.

We suspected that up-regulation of the entry factors *CTSL*, *CTSB*, and *GNPTAB* might be occurring as part of a general macrophage differentiation program. We tested this hypothesis using gene sets derived from published bulk RNA-Seq data of primary blood monocytes before and after differentiation into macrophages, *in vitro* (Dong et al., 2013) (**Table S4**). We found that genes that are up-regulated during *in vitro* differentiation were significantly enriched in infected cells (OR = 3.5, $p = 3.1\text{x}10^{-11}$, Fisher's exact test) and genes that were down-regulated during differentiation were significantly enriched in bystanders (OR = 3.7, $p = 4.2\text{x}10^{-8}$, Fisher's exact test; combined chi-squared goodness of fit test $p = 2.2\text{x}10^{-30}$). Using gene set annotations from two other RNA-Seq studies of macrophage *in vitro* differentiation resulted in similar findings (chi-squared goodness of fit test $p = 2.6\text{x}10^{-9}$ (Saeed et al., 2014), Fisher's exact test $p = 5.7\text{x}10^{-12}$ (Italiani et al., 2014)). Genes associated with differentiation into M2-polarized macrophages were more enriched among EBOV-infected cells than those of M1-polarized macrophages (OR = 7.8, $p = 1.3\text{x}10^{-10}$ compared to OR = 3.3, $p = 1.6\text{x}10^{-3}$ (Italiani et al., 2014), **Table S4**).

Next, we quantified the proportion of infected cells as a function of macrophage-differentiation, and found that infectivity increased along with expression of the macrophage differentiation program. To determine the relative activity of the macrophage program in each cell, we computed a “macrophage differentiation score” consisting of a weighted sum of the 618 genes that were significantly positively or negatively correlated with *in vitro* differentiation in the (Dong et al., 2013) dataset (**Materials and Methods**). Ranking cells from lowest to highest macrophage score, we observed that the percentage of infected cells rose more than four-fold

from 3.0% to 15.0% (**Figure 5H**). Thus, our data strongly suggests that of all circulating cells, EBOV predominantly infects monocytes with the highest expression of the macrophage differentiation program.

Given that macrophage differentiation genes were over-expressed in DPs relative to DNs (**Figure S8A**), we sought to understand the relationship between the CD14 and CD16 defined monocyte subsets and expression of the macrophage differentiation program. Comparing the overall macrophage score of the different CD14/CD16-marked subpopulations confirmed that DPs generally had the highest macrophage scores while DNs and single CD14+ cells had the lowest (**Figure S8B**). Thus, some of the enrichment of infected cells among the DP subset could potentially be attributed to their more 'macrophage-like' gene expression.

However, there was substantial heterogeneity in macrophage scores within DPs and the other CD14/CD16-marked subsets, and we found that macrophage score and CD14/CD16 subset were independently predictive of infectivity. To demonstrate this, we stratified cells in each subset by macrophage score (above or below the median value across all subsets combined). This showed that less macrophage-like DPs were still more likely to be infected than more macrophage-like DNs, even though all cells in the former category had a lower macrophage score than all cells in the latter ($p = 1.9 \times 10^{-25}$, Fisher's exact Test, **Figure S8C**). However, within the DPs, more macrophage-like cells were more likely to be infected than less macrophage-like cells ($p = 0.0003$, Fisher's exact Test, **Figure S8C**). This suggests that infection could not be explained by either CD14/CD16 subset or macrophage score alone. As a further confirmation, we fit a logistic regression predicting the infection status of each cell using macrophage score, smoothed *CD14* and *CD16* expression values, and a *CD14*x*CD16* interaction term (**Materials and Methods**). As expected, the *CD14*x*CD16* interaction term (which is highest in DPs) and macrophage score were positively associated with infection status, and the *CD14* and *CD16* terms were negatively associated ($p < 0.01$ for all coefficients). These findings demonstrate that monocyte CD14/CD16 subset and differentiation status independently impact the probability of a cell being infected with EBOV, *in vivo*.

Interferon stimulated genes are down-regulated in infected monocytes relative to bystanders

Finally, we noticed that several key ISGs such as *MX1* were expressed at lower levels in infected cells than in bystanders (MAST $q = 7.7 \times 10^{-14}$, rank-sum test $p = 2.0 \times 10^{-5}$, **Figures 5A and 5I)**. To determine if infection had a suppressive effect on overall ISG expression, we compared the magnitude of the interferon response (defined previously as the ISG score, **Materials and Methods**) between infected and bystander cells at late EVD. While both bystander and infected monocytes at late EVD had higher ISG scores than monocytes at baseline, ISG scores were lower in infected cells than bystanders (not statistically significant by rank-sum test, **Figure S8D**). More strikingly, there was a significant negative correlation between ISG score and the percentage of cellular transcripts derived from EBOV (*i.e.*, the intracellular viral load) (Spearman $\rho = -0.62$, $p = 1.1 \times 10^{-11}$, **Figure 5J**). This suggests that ISGs are down-regulated during viral replication within infected cells (see **Figure 7** and **Discussion**).

Single-cell transcriptomics of *ex vivo* infected PBMCs reveals temporal dynamics in viral gene expression

In order to more thoroughly probe viral and host gene expression changes during the viral life cycle, we sought to obtain transcriptomes from a greater number of infected cells. Thus, we isolated PBMCs from 2 healthy rhesus monkeys (NHP1 and NHP2) and inoculated them *ex vivo* with either live EBOV, EBOV rendered replication-incompetent by gamma irradiation (Feldmann et al., 2019), or media only as a control (**Figure 6A**). We selected a multiplicity of infection (MOI) of 0.1 plaque forming units (pfu, titrated on Vero E6 cells)/cell to ensure that a large proportion of cells would be infected. We performed scRNA-Seq using Seq-Well at 4 hours or 24 hours post-infection (HPI), corresponding to very early (start of viral transcription) and middle-to-late stages (viral genome replication, virion assembly) of the viral life cycle. Inoculation with gamma-irradiated EBOV allowed us to characterize the host response in the absence of effective viral transcription and translation.

We obtained single-cell transcriptomes from 50,646 PBMCs inoculated *ex vivo*, and observed similar cell-type representation, clustering by treatment condition, and distribution of EBOV-infected cells as with the *in vivo* collections (**Figures S9A–S9C**), with a few notable exceptions. First, we observed that cells from NHP1 and NHP2 separated in UMAP embeddings (**Figure S9D**) and that this separation was associated with higher expression of ISGs such as *MX1* in cells derived from NHP1 compared to NHP2 (**Figures S9E and S9F**). The ISG signal was most predominant in cells from NHP1 at 24 HPI treated with either irradiated or live virus (**Figure S9G**). We therefore analyzed cells from each animal both separately and jointly to avoid potential artifacts. A second difference is that while we did not observe infected DCs *in vivo*, we found that 16% of DCs (16.0% in NHP1, 15.7% in NHP2) inoculated with live virus *ex vivo* were infected by 24 HPI (**Figure S9H**). This difference could be due to increased density of cells in culture, the higher effective MOI we used in the *ex vivo* experiment, or even changes to the expression states of DCs associated with culture conditions.

Consistent with the *in vivo* data, monocytes were the predominant infected cell type, with over 65% infected by 24 HPI after inoculation with live virus (76% in NHP1, 61% in NHP2) (**Figures S9H and S9I**). 11.8% of the monocytes treated with irradiated virus also contained a statistically significant number of viral reads by 24 HPI, despite the fact that gamma irradiation induces damage to the viral genome that eliminates productive viral replication (Feldmann et al., 2019). As expected, cells treated with irradiated virus had a significantly lower fraction of EBOV reads per cell than those treated with live virus (**Figure S10A)**. Moreover, viral RNAs from the cells treated with irradiated virus were substantially less likely to be coding-sense mRNA transcripts than anti-sense viral genomic RNA. For live-virus treated cells, 78% and 92% of detected RNAs were mRNA at 4 and 24 HPI, compared to only 37% and 44% in cells treated with irradiated virus (**Figure S10B**). This suggests that our method can detect fragments of viral genomic RNA from irradiated virus that have entered cells, but as expected, these do not reflect productive infections and do not generate significant quantities of poly-adenylated mRNAs.

Exploiting the increased resolution of the *ex vivo* dataset, we characterized the heterogeneity in viral transcript abundance per cell (*i.e.*, the intracellular viral load). We observed that the intracellular viral load varies over several orders of magnitude in infected cells, both *in vivo* and *ex vivo* (**Figures 6B and 6C**)**.** While most cells harbored viral loads below 0.1%, a substantial minority had loads of >10%, with maximum detected loads of 57.5% and 52.3% for cells *in vivo* and *ex vivo* respectively. The observed heterogeneity in viral load was not due to different numbers of transcripts detected per cell, because cells with low and high viral load had a similar range of total transcripts detected (**Figures S10C and S10D**).

We next analyzed the dynamics of EBOV gene expression to determine if it matched the predicted pattern based on established models of EBOV transcription. Transcription of EBOV's 7 genes by the viral RNA-directed RNA polymerase L follows the canonical stop-start mechanism described for filoviruses and other non-segmented negative-strand RNA viruses (Brauburger et al., 2014, 2016). L initiates transcription *de novo* (Deflubé et al., 2019) at the 3' end of the genome, and processes from 5' to 3'; at each gene transcription termination signal, L pauses and either falls off the genomic RNA template or reinitiates transcription of a new mRNA for the subsequent gene (**Figure 6D**, (Mühlberger, 2007)). As a consequence, *NP* is the first gene to be transcribed and is transcribed at the highest level, proceeding down the genome to the polymerase gene *L* being transcribed last and at the lowest level.

When we quantified the relative expression levels of EBOV genes as a function of viral load, we observed an unexpected accumulation of *GP* mRNA (**Figures 6E and 6F**) that was consistent between NHP1 and NHP2 (**Figures S10E and S10F**). At low viral loads, both *in vivo* and *ex vivo*, the gene expression distribution roughly matched the expected pattern, with most of the transcripts derived from the 3' end of the genome, in particular *NP*, and the fewest transcripts derived from the 5' genes *VP30*, *VP24*, and *L*. In agreement with this pattern, cells inoculated with irradiated virus, which has impaired transcription due to RNA cleavage or crosslinking (Feng et al., 2011; Ginoza, 1967; Ward, 1980), were highly enriched in *NP* mRNA (**Figure S10B**), suggestive of RNA fragment transcription. However, as viral load increased in cells infected with live EBOV, *GP* was the most highly expressed viral transcript. This finding is unexpected based on the start-stop mechanism where *NP* should be transcribed at a strictly higher rate than *GP*. This observation suggests a life-cycle dependent regulation of viral gene expression that has not previously been observed for EBOV (see **Discussion**).

EBOV infection down-regulates host antiviral genes and up-regulates putative pro-viral genes

Next, we exploited natural variability in viral load across infected cells to identify host gene expression changes correlated with viral replication, which may therefore represent pathways directly regulated by infection. Instead of testing for differential expression between infected and bystander cells as we did previously to define tropism-associated genes, we looked for continuous association between viral abundance and host transcript levels in infected

monocytes (**Materials and Methods**). This identified 264 genes that were negatively correlated and 211 genes that were positively correlated with viral load *ex vivo* with $q < 0.05$, of which 34 changed by more than 30% per 10-fold increase in viral load (**Figure 7A**, **Table S6**).

Consistent with our previous observation that ISG score decreased with viral load within a cell, ISGs decreased dramatically with EBOV levels, both *in vivo* and *ex vivo* (e.g., *MX1*, $q = 1.5\text{x}10^{-24}$ *ex vivo*, $q = 7.4\text{x}10^{-9}$ *in vivo*, **Figures 7A and 7B**, **Table S6**). *Ex vivo*, the most negatively associated gene was *STAT1*, the master transcriptional regulator of the IFN response ($q = 4.9\text{x}10^{-41}$ *ex vivo*, $p = 0.0072$ *in vivo* but *not significant by FDR*). Previous experiments have shown that the EBOV protein VP24 inhibits STAT1 activity by blocking its translocation to the nucleus (Leung et al., 2006) through direct binding (Zhang et al., 2012). However, this is the first observation that *STAT1* mRNA levels decrease with viral replication within infected cells *in vivo*.

We visualized the expression of *STAT1* and several other negatively regulated ISGs over the course of infection within a cell (**Figure 7C**). The expression of these genes remained relatively constant as EBOV levels rose to 1% of cellular transcripts; however as EBOV levels continued to rise beyond 1%, the expression of *STAT1* and many of its target genes declined precipitously. This suggests that there is a delay before EBOV can transcriptionally down-regulate host antiviral genes since it must transcribe and translate VP24 and other immunomodulatory proteins before they can act. The trajectories of these host antiviral genes were consistent between donor animals *ex vivo* (**Figure S10G**) despite the fact that more cells from NHP1 mounted an IFN response than NHP2.

Only a handful of host genes increased in expression level alongside viral load, but their trajectories were consistent between the two animals (**Figure S10H**). The most dramatically up-regulated gene was *DYNLL1* ($q = 2.5\text{x}10^{-27}$ *ex vivo*, $q = 1.5\text{x}10^{-5}$ *in vivo*, **Table S6**), which increased significantly both *in vivo* and *ex vivo* (**Figures 7A and 7B**). DYNLL1 is a multi-functional protein involved in intracellular transport (Barbar, 2008). Intriguingly, DYNLL1 was previously shown to increase EBOV replication in a minigenome reporter assay (Luthra et al., 2015). Our data show that *DYNLL1* mRNA is up-regulated *in vivo* and *ex vivo*, starting when EBOV RNA constitutes between 0.1–1% of the transcripts in a cell, before ISGs begin decreasing (**Figure 7D**).

Several other genes that we identified as up-regulated alongside viral replication have known or speculative pro-viral functions in protein folding and synthesis. For example, *HSPA5* ($q = 4.5\text{x}10^{-22}$ *ex vivo*, $p = 0.04$ *in vivo* but *not significant* by FDR) encodes a chaperone protein that was previously shown to be an essential host factor for EBOV (Reid et al., 2014). However, this is the first observation that *HSPA5* mRNA is up-regulated in EBOV infection. Other hits such as *DDIT3* and *NFE2L1* ($q = 1.3\text{x}10^{-10}$ and $q = 1.8\text{x}10^{-8}$ respectively, *ex vivo*) are sensors of ER and oxidative stress (Kim et al., 2016) and have been implicated in cell lines infected with Marburg virus (Hölzer et al., 2016) and in monocytes in EBOV-infected NHPs (Menicucci et al., 2017). We observe a corresponding enrichment of gene sets associated with ER stress

response among genes up-regulated alongside intracellular viral load such as BUYTAERT_PHOTODYNAMIC_THERAPY_STRESS_UP, PODAR_RESPONSE_TO_ADAPHOSTIN_UP, and HALLMARK_UNFOLDED_PROTEIN_RESPONSE (q = 4.7x10^{-14}, q = 5.3 x 10^{-5}, q = 0.004 respectively, Fisher's exact test, **Table S6**). This suggests that oxidative and/or ER stress may play an important role in the pathophysiology of infected cells, particularly at a late stage of viral replication.

Increased translation can also deplete tRNAs and free amino acids (Albers and Czech, 2016). We observed statistically significant up-regulation of *IARS* - isoleucine tRNA synthetase (q = 8.4x10^{-9} *ex vivo*) - which may reflect the cellular response to increased translational demand due to overwhelming production of viral proteins. In addition to *IARS*, several gene sets associated with depletion of amino acids were significantly up-regulated such as KRIGE_RESPONSE_TO_TOSEDOSTAT_24HR_UP (Tosedostat is an aminopeptidase inhibitor), KRIGE_AMINO_ACID_DEPRIVATION and PENG_LEUCINE_DEPRIVATION_UP (q = 2.3x10^{-6}, q = 0.0030, q = 0.0040, respectively, Fisher's exact test, **Table S6**), suggesting that viral replication exhausts cellular amino acid stores. This hypothesis is consistent with prior observations of depleted amino acids in the plasma of fatal human EVD cases (Eisfeld et al., 2017).

Discussion

Despite recurrent outbreaks, the molecular basis of EVD pathogenesis in humans remains understudied due to the biosafety and logistical challenges associated with performing clinical research during outbreaks in resource-poor settings, and experimentally investigating EBOV in maximum containment facilities. By adapting CyTOF and scRNA-Seq approaches for use in BSL-4 containment, we comprehensively surveyed the molecular correlates of disease progression and viral replication in circulating immune cells in a nonhuman primate model of EVD. This study, which is the first high-parameter, single-cell investigation of a RG-4 agent, shed new light on changes in cell-type abundance throughout lethal EVD, defined the preferred targets of EBOV amongst circulating cells, and identified genes regulated by the cytokine milieu or by direct EBOV infection.

We characterized transcriptional- and protein-level changes in monocytes during EVD in NHPs, some of which reflect disruption of their physiological antiviral function. Monocytes had over twice as many differentially expressed genes as other cell types, including genes involved in IFN response, cytokine production, myeloid differentiation, and antigen presentation (**Figures 3 and 4**). Monocytes became activated by IFN during EVD, which normally triggers an increase in expression of multiple MHC-II genes for antigen presentation (Steimle et al., 1994). Surprisingly, almost all of the MHC-II genes and other genes in the antigen presentation pathway were strikingly down-regulated in monocytes, with modest changes in B cells, and no significant changes in cDCs. Moreover, MHC-II genes decreased in both infected and uninfected

monocytes, suggesting that the cytokine milieu led to decreased MHC-II expression (**Figure 4**). These findings suggest a failure of monocytes to present antigens, which might explain why a failed or delayed adaptive immune response is a hallmark of fatal EVD in humans (Baize et al., 1999; Lüdtke et al., 2016).

As EVD progressed, conventional CD14+ and CD16+ monocyte subsets disappeared and were replaced by two unusual populations: CD14+ CD16+ (DP) monocytes, which are known to increase in other infections (Fingerle et al., 1993; Michlmayr et al., 2018; Nockher and Scherberich, 1998; Zanini et al., 2018a), and an unexpected CD14- CD16- (DN) population, that, to our knowledge, has not been previously described in viral infections (**Figures 5C–5F**). The DN monocytes were highly proliferative and their transcriptomes were more similar to bone marrow resident monocyte precursors than circulating monocytes. This suggests that they may be the product of emergency myelopoiesis, a process whereby cytokines stimulate the bone marrow to release immature myeloid lineage cells that are sometimes referred to as immunosuppressive myeloid cells (Chiba et al., 2018; Hérault et al., 2017; Sayed et al., 2019). DNs had high expression of neutrophil granule genes such as *MPO* and *AZU1*, suggesting that they may represent immature cells prior to the branching of neutrophil and monocyte lineages (*i.e.*, common myeloid progenitors). We identified the emergence of an analogous DN population in human acute EVD cases as well, which later disappeared as the patients recovered, confirming that this population becomes the dominant circulating monocyte population in human disease as well. This finding highlights the power of high-parameter methods such as scRNA-Seq and CyTOF over previous approaches like FACS; despite little-to-no detection of CD14 or CD16, the conventional markers for phenotyping monocytes, there were enough other RNA and protein markers to reliably detect these DN cells as monocyte-lineage cells.

Our data refines the picture of EBOV's tropism in NHPs, demonstrating that the predominant EBOV infected population in circulation are the DP monocytes which expand during the infection, and monocytes expressing markers of macrophage differentiation (**Figures 5C, 5G, and 5H**). Existing literature has already demonstrated that myeloid cells are major targets of EBOV (Geisbert et al., 2003b, 2003c; Greenberg et al., 2020), including DCs *ex vivo* and in lymph nodes *in vivo* (Geisbert et al., 2003c). While we observe infected monocytes and DCs *ex vivo*, only monocytes were infected more often than expected due to chance among circulating immune cells *in vivo*. This might reflect relevant biological phenomena required for DC infection, such as cell density, cell-to-cell contact, or MOI. Among monocytes, EBOV-infected cells were significantly enriched in CD14+ CD16+ DP monocytes, which were the most macrophage-like and expressed increased differentiation genes, including known EBOV entry factors like cathepsin B (Chandran et al., 2005; Martinez et al., 2013; Schornberg et al., 2006). It has been shown that cultured monocytes are only susceptible to EBOV infection upon differentiation (Martinez et al., 2013), and our data further shows that infectivity *in vivo* strongly correlates with physiological variability of monocyte differentiation state in the context of an active immune system. Furthermore, our data support the hypothesis that the relative abundance of DP

monocytes, the preferred circulating cell targets of EBOV, increases over the course of infection, perhaps driven by the cytokine milieu of EVD.

Among infected monocytes, we observed substantial heterogeneity in intracellular viral load, which we exploited as a proxy for staging EBOV's progression through its life cycle. We confidently identified infected cells with as low as 0.01% EBOV RNA out of total RNA (intracellular viral load), while some cells had up to 57.5% EBOV RNA (**Figures 6B and 6C)**. Similar heterogeneity has been observed for the segmented -ssRNA virus influenza (Russell et al., 2018), the +ssRNA virus dengue (Zanini et al., 2018a, 2018b), and the dsDNA virus HCMV (Hein and Weissman, 2019), but this is the first demonstration for a non-segmented negative-sense (NNS) virus.

By analyzing patterns of gene expression among cells at different stages of the viral infection, we identified an unexpected over-representation of *GP* mRNA accumulating late in the viral life cycle**.** NNS viruses, including EBOV, have strict transcriptional gene regulation based on the viral genome organization (**Figure 6D**) (Brauburger et al., 2014, 2016). In our data, cells with low viral load (reflecting early stages of the viral life cycle) had relative EBOV transcript abundances that mirrored the genome organization as expected -- *i.e.*, they had the highest expression of *NP* and the lowest expression of *L*. But cells with high viral load (later during the viral life cycle) had higher abundance of *GP* mRNA than *NP* mRNA (**Figures 6E and 6F)**. Given the NNS virus dogma that *NP* is transcribed more frequently than *GP* mRNA, our data suggest that alternate transcription or post-transcriptional regulatory mechanisms, such as increased stability of *GP* mRNA, may account for accumulation of *GP* mRNA late during the viral life cycle. Many viruses increase structural protein production late in the viral life cycle (Honess and Roizman, 1974; Irigoyen et al., 2016; King et al., 2018; Shin et al., 2015); increased EBOV *GP* late in the viral life cycle likely increases the formation of infectious virions.

Many host ISGs negatively correlated with intracellular viral load, suggesting that viral infection down-regulates ISG expression *in vivo* and *ex vivo* (**Figures 7A–C**). These findings strongly suggest that EBOV is specifically down-regulating ISGs rather than preferentially infecting cells with low ISG levels. First, there are multiple well-established mechanisms by which EBOV down-regulates transcription of ISGs, such as by preventing the master antiviral transcription factor, STAT1, from translocating to the nucleus (Harcourt et al., 1999; Kash et al., 2006; Leung et al., 2006; Reid et al., 2008). Second, the percentage of EBOV-infected cells increases gradually from DPI 4 onward (**Figure 2I**) despite the fact that monocytes are mounting a strong ISG response by then, which suggests that EBOV is able to overcome the inhibitory activities of ISGs and continue replicating. Third, when EBOV infects a cell, viral load must start at a low level before increasing with replication and transcription. If the causal direction was reversed -- *i.e.*, low ISG levels supported EBOV progression and high ISG levels inhibited EBOV progression -- we would expect to see some cells with low ISG and low viral load, but we do not observe any cells with low (0.01–0.1%) viral load and low ISG levels (**Figure 5J**). Thus our data suggests that EBOV is able to infect monocytes that are mounting a full IFN response,

overcome the inhibitory activities of ISGs, and transcribe viral mRNA to high levels during disease, *in vivo*.

On the other hand, several putative pro-viral host genes were positively correlated with intracellular viral load, suggesting that they are directly responding to, or are regulated by, the presence of virus in infected cells (**Figures 7A, 7B, and 7D**). *DYNLL1* was the top associated gene with viral load, both *in vivo* and *ex vivo*. Previous work identified an interaction between DYNLL1 and EBOV VP35 (Kubota et al., 2009) that increased EBOV RNA synthesis in a minigenome assay (Luthra et al., 2015). DYNLL1 associates with proteins of many viruses (Merino-Gracia et al., 2011), and increases replication of rabies virus (Tan et al., 2007) an NNS virus similar to EBOV. Here, we showed that *DYNLL1* expression is up-regulated within infected cells *in vivo*, suggesting that EBOV manipulates cellular pathways to encourage a pro-viral cellular environment. Nuclear DYNLL1 typically represses its own transcription factor ATMIN (Jurado et al., 2012); we hypothesize that EBOV VP35 may sequester DYNLL1 protein in the cytoplasm, relieving repression of ATMIN, thus up-regulating *DYNLL1* mRNA.

Additional intriguing genes that are positively correlated with intracellular viral load include *HSPA5*, *NFE2L1*, *DDIT3*, *GTF2A1*, and *IARS*, though their effect sizes are larger *ex vivo* than *in vivo*. Many of these genes are involved in the cellular stress response, which can be triggered when infection overwhelms host translation. As a chaperone protein, *HSPA5* in particular is essential for EBOV replication (Reid et al., 2014); *NFE2L1* and *DDIT3* sense ER and oxidative stress (Kim et al., 2016) and are upregulated upon filovirus infection (Hölzer et al., 2016; Menicucci et al., 2017), and our data suggest that these genes are up-regulated throughout the viral life cycle. Viral transcription and translation can also deplete tRNAs and free amino acids; as a result, viruses have evolved mechanisms to maintain translational capacity (Albers and Czech, 2016). The tRNA synthetase *IARS* could be up-regulated in response to this depleted host environment, also reflected in fatal EVD cases (Eisfeld et al., 2017). By computationally staging cells by their phase in the infection cycle, we nominated several putative pro-viral genes for further study, highlighting the utility of single-cell profiling to study host-virus interactions.

The accumulation of additional host-pathogen single-cell datasets promises to greatly enhance our understanding of infection by allowing us to determine which features of pathogenesis are shared between, or specific to, individual pathogens. For example, our scRNA-Seq and CyTOF data identified several molecular commonalities between EVD and immunosuppressive septic shock (Bray and Mahanty, 2003), which is also characterized by loss of MHC-II expression in monocytes (Monneret and Venet, 2014; Reyes et al., 2020), increased DP monocytes (Fingerle et al., 1993), and emergency myelopoiesis (Bomans et al., 2018; Cuenca et al., 2015; Reyes et al., 2020). Soluble mediators, including cytokines and glucocorticoids, could be key drivers of both EVD and sepsis pathophysiology. TNFα signaling has been extensively implicated as a driver of systemic loss of vascular resistance and shock during EVD. Glucocorticoids have been less well studied in EVD, but decrease MHC-II *ex vivo* (Hawrylowicz et al., 1994) and *in vivo* during sepsis (Tulzo et al., 2004) and reduce *CD14* expression (Nockher and Scherberich, 1997) while increasing the abundance of DP monocytes (Liu et al., 2015). Indeed, the

connection between EVD and sepsis may be direct: studies have found bacterial invasion during EVD in NHPs (Reisler et al., 2018) and in humans (Carroll et al., 2017), with immune signatures that resemble sepsis (Eisfeld et al., 2017).

In summary, this work expands our understanding of EVD, and provides a general paradigm for exploring molecular features of host-pathogen interactions, such as tropism and dysregulation of cell circuitry in infected cells, that can be applied to other emerging pathogens.

Materials and Methods

Resource Availability

Lead Contact. Further information and requests for resources and reagents should be directed to Aaron Lin (alin@broadinstitute.org).

Materials Availability. This study did not generate new unique reagents.

Data and Code Availability

The analysis scripts used in this study are available at https://github.com/dylkot/SC-Ebola.

Datasets generated during this study will be available on GEO and raw sequence data will be available on SRA.

CyTOF .fcs files will be made available via FlowRepository.

Experimental Model and Subject Details

This study included a subset (21 of 27) outbred rhesus monkeys (*Macaca mulatta*) of Chinese origin described recently (Bennett et al. in submission) (Greenberg et al., 2020). These 27 nonhuman primates (NHPs) were randomized into cohorts (**Figure S1B**, (Bennett et al. in submission)), balancing age, weight, and sex across 7 groups. All work was approved and performed in accordance with the Guide for the Care and Use of Laboratory Animals of the National Institute of Health, the Office of Animal Welfare, and the US Department of Agriculture (Bennett et al. in submission).

Method Details

Serial sampling study

This study utilized the Ebola virus/H. sapiens-tc/COD/1995/Kikwit-9510621 (EBOV/Kikwit; GenBank accession MG572235.1; *Filoviridae: Zaire ebolavirus*) isolate for the *in vivo* and *ex*

vivo challenges, obtained from the Biodefense and Emerging Infections Research Resources Repository (BEI Resources, Manassas, VA, USA). It is the standard challenge stock defined by the filovirus animal non-clinical group (FANG) for testing product efficacy for FDA approval and is well characterized.

For all 21 outbred rhesus monkeys, two baseline blood samples were collected between 0–14 and 14–30 days prior to infection (**Figure S1B**). 18 NHPs were exposed to the EBOV/Kikwit isolate diluted to a target concentration of 1,000 plaque forming units (PFU) in a volume of 1 mL/dose. All NHPs were inoculated within a 5 month period. This same cohort has already been described recently (Bennett et al. in submission) (Greenberg et al., 2020).

Clinical observations and scoring

Beginning on day post-infection (DPI) 0, NHPs were observed 1–3 times daily and given a clinical score based on five criteria: overall clinical appearance and signs of hemorrhage; respiratory rate, mucous membrane color, and dyspnea; recumbency; non-responsiveness; and core temperature (Bennett et al. in submission). Each criterium was assigned a score between 1 and 10, and all scores were added together. Once an NHP reached a combined score of >10, the animal was humanely euthanized.

Whole blood collection

Blood was drawn from anesthetized animals into BD vacutainer plastic serum separator tubes (SST) for serum viral load quantification, or in BD vacutainer plastic blood collection tubes with K_3EDTA for hematology and peripheral blood mononuclear cell (PBMC) purification (Becton Dickinson, Franklin Lakes, NJ, USA) (Bennett et al. in submission). SST tubes were centrifuged at room temperature for 10 minutes (min) at 1800 x *g* to isolate serum. K_3EDTA tubes were mixed by gentle inversion prior to hematology and PBMC purification.

Hematology and complete blood counts (CBC)

250 µL of each whole blood sample was analyzed on a Sysmex 2000i XT (Sysmex Corporation, Kobe, Hyogo Prefecture, Japan) (Bennett et al. in submission). Parameters analyzed by this instrument were: counts of basophils, eosinophils, lymphocytes, monocytes, neutrophils, white blood cell count; percentages of each cell type; and mean platelet volume.

To estimate the abundance of lymphocyte cell types, we multiplied the CBC lymphocyte count by the proportion of lymphocytes of each cell type (CD8 T cells, CD4 T cells, NK cells, and B cells) which was calculated from the unsupervised clustering of the CyTOF data (see 'CyTOF' below).

EBOV serum viral load by RT-qPCR

70 µL of sample inactivated by TRIzol LS was added to 280 µL of Buffer AVL (Qiagen, Hilden, Germany) with carrier RNA (Bennett et al. in submission). Samples were then extracted using the QIAamp Viral RNA Mini Kit (Qiagen) in accordance with the manufacturer's instructions, eluted in 70 µL of Buffer AVE, aliquoted, and frozen. Viral load was determined using the BEI Resources Critical Reagents Program experimental EZ1 RT-qPCR kit assay in accordance with the manufacturer's instructions.

PBMC purification

We centrifuged whole blood in K_3EDTA tubes at 1800 x *g* for 10 min at room temperature, removed EDTA plasma, and added phosphate buffered saline (PBS, Thermo Fisher Scientific, Waltham, MA, USA) to the pelleted cells to double the original whole blood volume. We gently poured the PBS-blood cell mixture into an Accuspin tube containing Histopaque (Sigma-Aldrich, St. Louis, MO, USA) and centrifuged at 1000 x *g* for 10 min at room temperature with the brake set to 1. Following centrifugation, we removed the top, clear supernatant layer to within 0.5 cm of the cloudy white layer containing PBMCs.

We transferred the cloudy PBMC layer to a clean 15 mL conical tube and increased the volume to 10 mL using PBS supplemented with 2% heat-inactivated fetal bovine serum (PBS/2%HI-FBS) and mixed by inversion. We then centrifuged at 300 x *g* for 10 min at 4 °C with the brake set to 1. Following centrifugation, we removed the supernatant, resuspended the cell pellet with PBS/2%HI-FBS to a final volume of 10 mL, and mixed using gentle raking to wash the cells. We repeated the wash step 2 more times with the centrifuge set to 200 x *g* for 10 min at 4 °C with the brake set to 1. We then resuspended the cell pellet in 9.5mL PBS/2%HI-FBS for counting using the Countess Cell Counting system (Thermo Fisher Scientific). We aliquoted 0.5 mL for Seq-Well, and used the remaining 9 mL volume for CyTOF.

Seq-Well

We performed Seq-Well as described previously (Gierahn et al., 2017), with the S^3 protocol (Hughes et al., 2019) and some controls and modifications to adhere to the BSL-4 environment.

As an experimental negative control to test our statistical model, we spiked Madin-Darby canine kidney (MDCK) cells, constituting ~5% of the total sample, into a subset of PBMC samples (**Table S1**) from EVD NHPs immediately before scRNA-Seq. As MDCKs were not exposed to EBOV, viral reads in these transcriptomes should be due to ambient RNA contamination (Russell et al., 2019).

After loading and sealing beads and cells in Seq-Well arrays, we placed them in a -80 °C freezer until further processing – this step was required due to time constraints in the BSL-4. Later, we removed sealed Seq-Well arrays from the -80 °C freezer, placed them in 4-well dishes, and allowed them to equilibrate to room temperature for at least 30 min. We then covered arrays in 5 mL Seq-Well Lysis Buffer per protocol.

We performed RNA hybridization and RT as specified in the protocol (Hughes et al., 2019). After RT, we collected beads by centrifugation at 1000 x *g* for 1 min at room temperature. We resuspended beads with GeneXpert Lysis Buffer (Cepheid, Sunnyvale, CA, USA) for inactivation, which was required prior to removal from the BSL-4 laboratory according to standard operating procedures. After removal, we washed beads thrice with TE buffer containing 0.01% Tween 20 and shipped at 4 °C for further library construction and sequencing, which was performed with the S^3 protocol (Hughes et al., 2019). We sequenced all libraries on either NextSeq 550 High Output or NovaSeq 6000 S2 flowcells (Illumina, San Diego, CA, USA), with 20 cycles for Read 1 (cell barcode and unique molecular index [UMI]) and 88 cycles for Read 2 (cDNA of interest). In some cases, we merged fastq files from multiple sequencing runs for increased coverage.

Depletion of abundant sequences by hybridization (DASH) of select Seq-Well libraries

We observed long concatemers of the common scRNA-Seq adaptor sequence (SeqB, 5'-AAGCAGTGGTATCAACGCAGAGTAC-3') at high frequency in some Seq-Well libraries, likely owing to low RNA input and the challenging environment of processing samples in the BSL-4 suite. These concatemers disrupted Illumina sequencing runs because the Read 1 sequencing primer annealed to multiple SeqB sequences on a single template, allowing multiple sequencing-by-synthesis reactions simultaneously. We therefore devised a strategy to remove SeqB concatemers using depletion of abundant sequences by hybridization (DASH) (Gu et al., 2016), a CRISPR-based method to degrade target DNA sequences prior to sequencing. SeqB lacked a 'NGG' protospacer adjacent motif (PAM) for the common *S. pyogenes* Cas9 (SpyCas9) for which DASH was originally described; therefore, we modified DASH to use *S. aureus* Cas9 (SauCas9) (Ran et al., 2015). Moreover, in contrast to SpyCas9, SauCas9 is a multi-turnover enzyme (Yourik et al., 2019), suggesting that SauCas9 would have higher cleavage efficiency, which was important since SeqB was present in multiple copies within a concatemer.

First, we designed a guide RNA (gRNA) to target SeqB. Based on the position of the SauCas9 PAM and the length of SeqB, only 17 nucleotides of the gRNA protospacer could anneal to SeqB. Because gRNA length is critical to SauCas9 cleavage efficiency (Friedland et al., 2015; Ran et al., 2015), we prepended 4 random bases as a 5' overhang (**Key Resources Table**). We *in vitro* transcribed this gRNA using the MEGAshortscript T7 Transcription Kit (Thermo Fisher Scientific), purified it using the RNA Clean & Concentrate Kit (Zymo Research, Irvine, CA, USA), and verified the correct RNA length on a 15% TBE-urea gel (Bio-Rad Laboratories, Hercules, CA, USA).

We performed SauCas9 DASH according to reaction conditions laid out for *in vitro* SauCas9 cleavage assays (Yourik et al., 2019). We incubated 10 pmol gRNA with 5 pmol SauCas9 (New England Biolabs [NEB], Ipswich, MA, USA) at 25 °C for 10 min, and then added up to 5 fmol DNA (2000:1000:1 RNA:protein:DNA ratio) and NEBuffer 3.1 (NEB) to 1X. We incubated this reaction at 37 °C for 2 hours (h), and quenched by adding EDTA to 50 μM, SDS to 1%, and 4

total U of Proteinase K (NEB) at room temp for 10 min. We removed degraded concatemers with two consecutive 0.8X SPRI purifications using Ampure XP beads (Beckman Coulter, Brea, CA, USA), eluted, and performed 6–9 cycles of PCR with the NEBNext Ultra II Q5 Master Mix (NEB) using Illumina P7 and the Seq-Well P5-TSO hybrid primer (Gierahn et al., 2017).

CyTOF

We added 1 mL of 16% paraformaldehyde (PFA, Electron Microscopy Sciences, Hatfield, PA, USA) to 9 mL PBMCs to fix the cells. We incubated samples at room temperature for 10 min followed by a final centrifugation at 600 x *g* for 5 min at 4 °C with the brake set to 9. We removed the supernatant, added 1 mL of PBS/5%HI-FBS for every 3 x 10^6 cells (*e.g.*, 2 mL for a sample containing 6 x 10^6 cells), aliquoted into 1 mL aliquots in cryovials, and stored in a -80 °C freezer.

We equilibrated fixed PBMCs to come to room temperature before transferring approximately 2 x 10^6 cells per sample into 1.2 mL cluster tubes in a 96-tube rack. We barcoded samples and multiplexed them into 6 batches of 16 samples using a previously described 16-plex palladium-based mass-tag cell barcoding scheme (Zunder et al., 2015). We aspirated pelleted barcoded cells to a volume of 50 µL and incubated them with 15 µL of Human TruStain FcX (Biolegend, San Diego, CA) for 10 minutes. We stained cells for 30 min with 175 µL of a reconstituted lyophilized cocktail of metal-tagged cell-surface antibodies described previously (Bjornson-Hooper et al., 2019a, 2019b) supplemented with two additional antibody channels (**Table S7**). We washed and permeabilized surface-stained cells with methanol before aspirating down to 50 µL and staining for 60 min with 190 µL of reconstituted lyophilized intracellular staining antibody cocktail (Bjornson-Hooper et al., 2019a, 2019b) (**Table S7**). We washed and resuspended fully-stained cells in a volume of 750 µL.

Following staining, we inactivated samples by adding 250 µL of 16% PFA to 750 µL of each sample, for a final concentration of 4% PFA, and incubated at 4 °C overnight. The next day, we centrifuged samples 600 x *g* for 5 min at 4°C and aspirated down to 100 µL. We resuspended samples in 1 mL 4% PFA in PBS and transferred to a clean 2 mL cryovial. We then removed samples from the BSL-4 using a dunk tank and froze at -80 °C within 30 min of PFA addition.

Following inactivation, we thawed and processed inactivated samples within a BSL-2 lab for iridium intercalation, then mixed with 1xEQ beads (Fluidigm, South San Francisco, CA, USA) and run on a CyTOF Helios (Fluidigm) instrument using a Super-Sampler introduction system (Victorian Airship & Scientific Apparatus LLC, Alamo, CA, USA).

FCS data files were normalized across all runs using the data normalization software (Finck et al., 2013) and debarcoded using the single-cell debarcoder tool (Zunder et al., 2015) as previously described. Data were uploaded and analyzed using CellEngine software (https://cellengine.com, Primity Bio, Fremont, CA, US), and a gating strategy was applied to identify cell populations using canonical markers. Frequencies for each population were

determined as a function of total CD66-CD45+ cell events and reported marker intensities are expressed as medians. Also see 'CyTOF data preprocessing, clustering, and dimensionality reduction' section below for details of unsupervised analysis of CyTOF data.

Ex vivo inoculation of PBMCs

PBMCs were isolated from healthy NHPs as previously described. We diluted cells to 3.3 x 10^6 cells/mL RPMI/10%HI-FBS and transferred 900 µL each to 2 mL external thread cryogenic vials (Corning, Corning, NY) for each experimental condition. We inoculated cells with 100 µL live virus (EBOV/Kikwit, the same stock used for *in vivo* NHP inoculation) for a final MOI of 0.1 PFU/cell, an equivalent dose of irradiated virus (EBOV/Kikwit treated with 5 mRads gamma irradiation), or media only, and incubated for 4 or 24 hours with slow rocking prior to Seq-Well processing.

Single-cell RNA-Seq raw data processing

Raw sequencing files were demultiplexed and converted to fastq using bcl2fastq version 2.20. Reads were then trimmed, aligned to a reference transcriptome, and parsed into a digital gene expression matrix using the previously published Dropseq-tools pipeline (Macosko et al., 2015) version 2.0. We used a Snakemake wrapper around Dropseq-tools that is available in the open source Github repository https://github.com/Hoohm/dropSeqPipe. In brief, we trimmed adapter sequences using Cutadapt (Martin, 2011) version 1.16, performed spliced alignment of trimmed reads using STAR aligner (Dobin et al., 2013) version 2.6.1b, identified core barcodes using the whitelist function in umi_tools (Smith et al.) version 0.5.5, and used Dropseq-tools to correct barcodes and extract digital count matrices. A frozen version of the pipeline used to process the Seq-Well data is available at the Github repository https://github.com/dylkot/dropSeqPipe-dak.

Sequencing reads were aligned to a hybrid genome/genebuild of *Macaca mulatta* (genome assembly Mmul_8.01, Ensembl gene build 92) and EBOV/Kikwit (Genbank accession KU182905.1).

Single-cell RNA-Seq data preprocessing, clustering, dimensionality reduction, and smoothing

The scRNA-Seq data was preprocessed, clustered, and visualized using Scanpy (Wolf et al., 2018). We removed cells with <300 genes detected, >10% of their UMIs derived from mitochondrial genes, or >95% of UMIs mapped to non-genic regions. We excluded ribosomal genes, genes correlated with the percentage of UMIs assigned to mitochondrial genes (Pearson R > 0.1), and *HBB* as these were largely driven by the technical covariate of whether cells had loaded into Seq-Well arrays fresh or had undergone a freeze-thaw cycle with cryoprotectant. We also excluded EBOV genes and cell-cycle genes (defined by correlation with *TOP2A,* Pearson R > 0.1) prior to clustering so that these signals would not influence identification of cell types.

We performed multiple iterations of clustering to detect and exclude doublets and to identify distinct cell populations at multiple levels of granularity. In each iteration, we filtered genes detected in <10 of the cells being clustered. We then transformed raw UMI counts by normalizing the sum of counts of each cell to 10,000 (TP10K), adding 1 to each expression value, and taking the natural logarithm. In each clustering iteration, we identified and subsetted the data to highly variable genes using the highly_variable_genes function in Scanpy (Satija et al., 2015) with the default parameters. We Z-normalized each gene and set transformed values exceeding 10 to 10. Z-normalized data was used as input to principal component analysis. We determined the number of principal components to use for downstream analysis by identifying an elbow on the Skree plot of the eigenvalues associated with each principal component. For the *in vivo* Seq-Well data, we used the Harmony algorithm (Korsunsky et al., 2019) to remove variation due to whether a PBMC sample had been processed fresh, or following a freeze-thaw. No Harmony adjustment was used for the *ex vivo* EBOV dataset. The Harmony-adjusted or raw principal components were then used to construct a nearest neighbor graph with the number of neighbors set to the maximum of 30 or 0.001 x the number of cells. Lastly, we clustered cells using the Leiden community detection algorithm (Traag et al., 2019).

We annotated broad PBMC clusters in the *in vivo* and *ex vivo* EBOV datasets based on the following marker genes: CD8+ T-cells (*CD3D*, *GZMB*, *GNLY*), CD4+ T-cells (*CD3D*, *IL7R*), B-cells (*MS4A1*, *IGHM*), Monocytes (*CFD*, *LYZ*), cDCs (*FLT3*, *IRF8*), pDCs (*IRF8*, *GZMB*), Neutrophils (*CD177*, *LCN2*), Platelets (*PF4*, *CAVIN2*), Plasmablasts (*MZB1*, *IGHM*, *IGHA*), and spike-in control MDCK cells (*COL5A2*, *SLC20A1*).

For the *in vivo* EBOV dataset, the first clustering iteration was used to identify and filter a cluster of multiplets (expressing high levels of B-cell, T-cell, Neutrophil, and Monocyte genes) and MDCK control cells. A second clustering iteration was run on the filtered data to identify broad cell type clusters of T/NK-cells, B-cells, and myeloid cells (Monocytes, cDCs, pDCs, and neutrophils). Sub-clustering was performed on each broad cell type in two iterations, the first to identify remaining doublets to exclude, and the second to cluster cells into final cell-types and sub-types based on annotation of marker genes (**Figure S4**).

An analogous sequence of clustering iterations was used for the *ex vivo* EBOV dataset. However, as there were no MDCK cells spiked in, we proceeded straight to sub-clustering of the T/NK, B, Monocyte/DC, and multiplet populations from the initial clustering iteration.

Doublets and multiplets identified during any clustering iteration were excluded. Then visualization in 2 dimensions was accomplished by computing the nearest neighbor graph with 0.001 x the number of cells as nearest neighbors, followed by Uniform Manifold Approximation and Projection (Becht et al., 2018).

Gene expression values were smoothed to facilitate direct visualization of *CD14*, *FCGR3* (which codes for CD16), and *MKI67* (which codes for Ki67) by running MAGIC (van Dijk et al., 2018) on log TP10K expression values with the 'cosine' distance metric and 3 diffusion steps.

Differential expression testing

We performed all differential expression tests using MAST (Finak et al., 2015) on log(TP10K + 1) normalized data. For all differential expression tests, we included (1) the percentage of mitochondrial reads and (2) the number of genes detected in a cell as covariates. For tests in the *in vivo* dataset, we additionally included a binary indicator covariate of whether the cell was derived from a sample that had been processed fresh, or had undergone freeze-thaw. For tests in the *ex vivo* dataset, we additionally included a binary indicator covariate of which NHP donor the sample was derived from.

For viral load comparisons, we used the log10 viral load as a continuous exogenous variable and only considered cells with ≥1 viral read. For all other comparisons, we used a binary exogenous variable indicating the reference and the query group. "Viral load" and "bystander vs. EBOV infected cells" in the *in vivo* dataset were conducted only considering cells from DPI 5–8. "Viral load" comparisons in the *ex vivo* dataset only considered cells from the 24 HPI timepoint.

Differential expression *p*-values were corrected for multiple hypothesis testing using the method of Benjamini and Hochberg (Benjamini and Hochberg, 1995).

Identifying differential gene expression modules

We clustered the log fold-change profiles of 1,437 differentially expressed genes (rows of **Figure 3A**) considering B, CD4+ T, CD8+ T, NK, and monocyte populations at the three EVD stages. Prior to clustering, insignificant values ($p > 0.2$) were first set to 0 and genes were normalized to unit L2 norm. We performed KMeans clustering with K=11 and the default parameters in Scikit-learn (Pedregosa et al., 2011) version 0.22.2.post1. Varying K above and below 11 led to highly consistent results with splitting or merging individual clusters at the margin, and we picked K=11 as the lowest value that yielded the "B down" module.

Detection of EBOV infected cells

Our infection detection model assumes that EBOV reads assigned to a cell are either due to true viral RNAs inside that cell, or due to 'ambient' extracellular EBOV RNA in the sample that, by chance, are captured in a well of the Seq-Well array along with the cell. Our null model for how many EBOV reads would be expected in a cell by chance therefore depends on two main parameters which we estimate from the data: (1) what proportion of ambient RNAs in a sample are due to EBOV, and (2) what proportion of a cell's expression profile is due to ambient RNA. Our method therefore precedes through the following steps which we describe below:

1. Estimate an ambient RNA profile for each Seq-Well array

2. Estimate the proportion of each cell's transcripts that are due to ambient RNA
3. Determine if there are more EBOV transcripts in each cell than would be expected based on the proportion of EBOV in the ambient RNA and the cell's estimated level of ambient RNA contamination.

1. Estimate an ambient RNA profile for each Seq-Well array

We assume that cell barcodes with few transcripts detected correspond to wells in the Seq-Well array that lack a cell, and therefore, that any transcripts assigned to those cell barcodes derive from ambient RNA. We consider all cell barcodes with fewer than 50 transcripts detected to be empty wells and compute an ambient RNA profile of the proportion of transcripts from these barcodes assigned to each gene. This is analogous to the approach used in several ambient RNA correction approaches such as (Fleming et al., 2019).

We denote the number of genes in the dataset as *G* and define the ambient RNA profile for a given Seq-Well array, *a*, as a *G*-dimensional vector $P^{(a)}$. $P_g^{(a)}$ denotes the estimated proportion of ambient RNAs that are assigned to gene *g* based on transcripts from all cell barcodes with fewer than 50 UMIs. Since $P_g^{(a)}$ is a proportion, the following constraints hold:

$$0 < P_g^{(a)} < 1; \sum_{g=1}^{G} P_g^{(a)} = 1$$

2. Estimate the proportion of each cell's transcripts that are due to ambient RNA

We adapt the previously published Consensus Non-negative Matrix Factorization (cNMF) method (Kotliar et al., 2019) to estimate the ambient RNA contamination level of each cell. In brief, cNMF learns a user-specified number of gene expression programs (GEPs), each a non-negative *G*-dimensional vector representing the average expression profile of an individual cell type or cellular activity (e.g. cell-cycle or interferon response) that are present in the data. In addition, it learns a "Usage" matrix reflecting the % contribution of each GEP in each cell (*i.e.*, the % of each cell's transcripts derived from each GEP). For the following notation, matrices are indicated in bold. Denoting the number of cells in the dataset as *C* and the user-specified number of GEPs as *K*, cNMF is given an input *C*x*G* matrix of transcript counts $\boldsymbol{X}_{\mathrm{cg}}$ and returns $\boldsymbol{G}_{\mathrm{kg}}$ a non-negative *K*x*G* matrix of GEPs reflecting the relative contribution of gene *g* in GEP *k*, as well as a non-negative *C*x*K* matrix $\boldsymbol{U}_{\mathrm{ck}}$ reflecting the percentage of transcripts in cell *c* that are due to GEP *k*.

We adapt this approach to return an updated, *C*x(*K+1*) dimensional usage matrix that includes an additional column reflecting the usage of the ambient RNA profile. We run cNMF as published to obtain the $\boldsymbol{G}_{\mathrm{kg}}$ GEP matrix. We then append the ambient RNA profile for a given array as an additional row to $\boldsymbol{G}_{\mathrm{kg}}$ and normalize each row to sum to 1 like so:

$$\tilde{G}_i = \frac{G_i}{|G_i|_1}$$

$$\tilde{\mathbf{G}}^{(a)} = \begin{bmatrix} \tilde{G}_1 \\ \tilde{G}_2 \\ \vdots \\ \tilde{G}_K \\ \tilde{P}^{(a)} \end{bmatrix}$$

where $||_1$ denotes the L1 norm, a tilde is used to denote an L1 normalized vector, and G_i denotes the ith row of $\boldsymbol{G}_{\mathrm{kg}}$. We then run a final iteration of NMF with the GEP matrix fixed to $\tilde{\mathbf{G}}^{(\mathrm{a})}$. This jointly estimates a usage of the ambient RNA profile and the other GEPS using non-negative least squares. For a given cell *c* from an array *a*, this amounts to solving the following non-negative least squares optimization:

$$\min_{U^{(a)}_{c1} \ldots U^{(a)}_{c(K+1)} \geq 0} || C_i - \sum_{k=1}^{K+1} U^{(a)}_{ck} \tilde{G}^{(a)}_k ||_2$$

We combine these coefficients for all cells and programs into a single matrix and L1 normalize the usages to sum to 1:

$$\boldsymbol{U}^{(\mathrm{a})} = \begin{bmatrix} U^{(a)}_{11} & \cdots & U^{(a)}_{1(K+1)} \\ \vdots & \ddots & \vdots \\ U^{(a)}_{C(a)1} & \cdots & U^{(a)}_{C(a)(K+11)} \end{bmatrix}$$

$$\tilde{\mathbf{U}}^{(\mathrm{a})} = \begin{bmatrix} \frac{U^{(a)}_1}{|U^{(a)}_1|_1} \\ \vdots \\ \frac{U^{(a)}_{C(a)}}{|U^{(a)}_{C(a)}|_1} \end{bmatrix}$$

where $C(a)$ denotes the number of cells derived from array *a*. The "K+1"th column of $\tilde{\mathbf{U}}^{(a)}$ reflects the estimated proportion of transcripts due to ambient RNA in all of the cells from array *a*. We repeat this calculation for each array separately and denote the estimated contribution of ambient RNA for a cell *c* as A_c.

<u>3. Determine if there are more EBOV transcripts in each cell than would be expected by chance</u>

For each cell, we determine a threshold, T_c, for the number of EBOV transcripts required to call that cell infected while keeping the false positive rate below *f*, a user specified threshold (*f* = 0.01 in all of our analyses). We first calculate $E^{(a)}$, the proportion of ambient RNA transcripts in array *a* that are due to EBOV as follows:

$$E^{(a)} = \sum_{g \in \{EBOV\ genes\}} P_g^{(a)}$$

Then, for a cell with N_c transcripts total; a given proportion, A_c, of its reads derived from ambient RNA; $E^{(a)}$, the proportion of ambient RNA contaminating reads expected to map to EBOV, we compute T_c using binomial statistics as follows :

$$T_c = F\left(f \mid p = E^{(a)} A_c,\ N = N_c\right)$$

Where *F* is the inverse survival function of the Binomial distribution with event probability *p* and *N* trials. We identify cells with $> T_c$ reads as infected.

Estimation of the infection receiver operator characteristic

We estimate sensitivity to call an infected cell with either 1% or 0.1% of its reads due to EBOV across a range of 60 false positive rate thresholds (*f*). We first randomly sampled 2,000 cells from the live EBOV treatment samples in the *ex vivo data,* or the non-baseline samples from the *in vivo* data, to serve as an empirical distribution for A_c and N_c. For each cell and specificity threshold *f*, we then calculate the probability of correctly calling a true positive cell as positive. We model the distribution of the number of EBOV reads in a cell as the convolution of 2 binomial distributions: $BN\left(p_{cell},\ N_{cell}\right)$ is the binomially distributed number of true cell-derived reads mapping to EBOV and $BN\left(p_{ambient},\ N_{ambient}\right)$ is the analogous distribution for ambient RNA-derived reads mapping to EBOV. p_{cell} is 0.01 or 0.001 by assumption, and $p_{ambient}$ is estimated empirically for each cell as $E^{(a)} A_c$. We calculate the expected values for N_{cell} and $N_{ambient}$ and round $N_{ambient}$ up to the nearest integer, as follows:

$$N_{ambient,c} = Ceiling\left(N_c A_c\right)$$

$$N_{cell,c} = N_c - N_{ambient,c}$$

We then directly calculate the convolution of the two binomial distributions for each cell:

$$P_{convolution}\left(X_c = x\right) = \sum_{j=0}^{x-i} \sum_{i=0}^{x} BN\left(i \mid p = p_{cell},\ N = N_{cell,c}\right) BN\left(j \mid p = p_{ambient,c},\ N = N_{ambient,c}\right)$$

We then calculate the sensitivity for each cell as the probability that the convolution distribution is greater than the empirical threshold:

$$Sensitivity_c = P_{convolution}\left(X_c > T_c\right)$$

We plot the average sensitivity across the 2000 randomly sampled cells as a function of *f* for p_{cell} equal to either 0.01 or 0.001.

Gene set enrichment testing

We downloaded gene sets from the Molecular Signatures Database (Liberzon et al., 2011) version 6.2 for gene set enrichment testing. We considered all Hallmark or C2 gene sets containing greater than 10 genes that were present in our expression data. We tested expression modules for enrichment using Fisher's exact test and corrected for multiple hypothesis testing using the method of Benjamini and Hochberg (Benjamini and Hochberg, 1995).

To test continuous expression profiles for gene set enrichment (**Figure S6B**), we used the rank-sum test comparing genes in the gene set to all genes not in the set.

Scoring cells for interferon response and macrophage differentiation

We identified 58 genes in the "Global up" module that were also included in one or more of the following gene sets from the molecular signatures database: HECKER_IFNB1_TARGETS, BROWNE_INTERFERON_RESPONSIVE_GENES, MOSERLE_IFNA_RESPONSE, HALLMARK_INTERFERON_ALPHA_RESPONSE, HALLMARK_INTERFERON_GAMMA_RESPONSE (**Table S3**). We then scored cells for the average expression of these genes using the score_genes function in Scanpy (Satija et al., 2015) with 58 control genes, as this was the number of genes in the ISG set, and otherwise default parameters.

We computed a macrophage score based on the set of 618 genes annotated as significantly up or down-regulated during *in vitro* monocyte-to-macrophage differentiation (Dong et al., 2013) (**Table S4)**. We computed each cell's macrophage score as the dot-product of its expression profile for the 618 genes (in log TP10K) with the log fold-change reported for each gene in (Dong et al., 2013). This effectively weights genes by both the direction and magnitude of their change during *in vitro* macrophage differentiation.

Comparison of EVD monocyte subsets with human bone marrow and PBMC data

We obtained all of the human PBMC datasets produced using v3 or v3.1 chemistry from the 10X website (**Key Resources Table**), aggregated them together, and processed the resulting dataset using the same pipeline as the NHP Seq-Well data. Briefly, we first filtered out genes detected in fewer than 10 cells before converting to log TP10K and performed PCA as described above. Then, we used Harmony (Korsunsky et al., 2019) to integrate out variation due to the different samples of origin and used 30 nearest neighbors for Leiden community detection (Traag et al., 2019) and UMAP dimensionality reduction (McInnes et al., 2018) (**Figure S7E**). We did not perform any sub-clustering on this dataset.

We obtained Human Cell Atlas bone marrow data from the Human Cell Atlas data portal (Hay et al., 2018) and processed it according to the same pipeline as the NHP data with a few modifications. We filtered doublets prior to clustering by running Scrublet (Wolock et al., 2019) separately within each of 8 donor batches with an expected doublet rate parameter of 6%. We identified and excluded cell-cycle associated genes, as those with a Pearson correlation > 0.3 with *TOP2A*. We integrated data from the different donor batches using Harmony and used 30 nearest neighbors for Leiden community detection and UMAP dimensionality reduction. We performed 3 rounds of sub-clustering: First we clustered all of the cells to identify monocyte and dendritic lineage cells (**Figure S7F**). Second, we clustered just hematopoietic stem cells (HSCs) and monocyte/dendritic progenitor cells to identify doublets (as those falling into a cluster characterized by T-cell marker genes such as CD3D and CD3E). Finally, we re-clustered this set with the doublets excluded to identify monocyte lineage cells, plasmacytoid dendritic cells, and conventional dendritic cells (**Figure S7G**).

We confirmed our marker gene-based annotations of the myeloid cell populations by comparing these cells to the circulating human PBMC dataset. We identified the nearest neighbor of each bone marrow myeloid progenitor cell in the PBMC dataset based on Euclidean distance of TP10K-normalized cells, considering overdispersed genes identified in the PBMC dataset based on the V-score (baseline-corrected Fano factor) (Klein et al., 2015). We then visualized the nearest-PBMC assignment of the bone marrow myeloid cells on a UMAP embedding (**Figure S7H**).

Finally, we combined the monocytes and monocyte precursor cells from the human PBMC and bone marrow datasets into a single reference. We again normalized the data to log TP10K and computed UMAP embeddings following the same procedure as for the individual datasets

(**Figure S7I**), using Harmony to remove variation due to donor sample. We then down-sampled this data so that there would be equivalent numbers of cells of each of the bone marrow and PBMC clusters (*i.e.*, 982 cells per cluster as that was the number of cells in the smallest cluster). We identified the nearest neighbor of each NHP monocyte in the down-sampled reference dataset, as described above and computed the percentage of NHP monocytes assigned to CD14+ or CD16+ clusters from either human bone marrow or PBMC (**Figure 5F).**

Logistic regression prediction of EBOV infection in vivo

We used Statsmodels (Seabold and Perktold, 2010) version 0.11.1 to fit a logistic regression predicting EBOV infection status among all monocytes from late EVD, based on the following features: macrophage score, MAGIC smoothed values of *CD14*, and *CD16* (*FCGR3*), and an interaction term for the product of the MAGIC smoothed values for *CD14* and *CD16*.

CyTOF data preprocessing, clustering, and dimensionality reduction

Our clustering and dimensionality analysis of the NHP CyTOF data was analogous to the Seq-Well pipeline with a few adaptations. We down-sampled a total of 1.1 million cells, consisting of 300,000 baseline cells and 100,000 cells from each DPI, selecting a uniform number of cells from each sample at a given DPI. We used the Arcsinh transformation of CyTOF raw intensity values divided by 5, which is standard in the field. We set a ceiling of transformed intensity values for each gene at the 99.999th percentile to reduce the effect of very rare outliers. We mean-centered the data but did not variance normalize prior to PCA. We then performed multiple clustering iterations of the data using the Leiden algorithm with the number of nearest neighbors set to the maximum of 30 or 0.01% of the number of cells in the dataset.

We annotated broad PBMC clusters in the CyTOF datasets based on the following marker proteins: CD8+ T-cells (CD3, CD8), CD4+ T-cells (CD3, CD4), NK cells (CD16, CD161), B-cells (CD19, IgM), Monocytes (CD11b, BDCA3, CD14, CD16, HLA-DR), cDCs (HLA-DR, CD11c, CD1c), pDCs (HLA-DR, CD4, CD123), Neutrophils (CD11b, CD66), Platelets (CD61, BDCA3), Plasmablasts (IgM high, but little or no CD19), Basophils (CD11b, CD123, low HLA-DR). There was also a cluster of cells characterized by high HLA-DR, Ki67, and CD38 which we annotate as “Unassigned APC”.

In the first clustering iteration, we identified and excluded clusters of doublets as those expressing markers of two or more broad cell-types. We also excluded a cluster characterized by high expression of CD3 but low expression of both CD4 and CD8 that we speculated was a technical artifact as it was nearly exclusive to a single CyTOF batch.

Next, to address batch effects between the different CyTOF runs, we grouped cells into the broad categories of Monocytes and DCs, CD4+ T, CD8+ T, NK, Neutrophils, B, Plasmablasts, and Platelets, and used COMBAT (Johnson et al., 2007) to adjust for batch effect separately

within each broad category, using the default COMBAT parameters in the Scanpy implementation.

We then ran cNMF (Kotliar et al., 2019) to identify and regress out artifact signals in the data. We ran cNMF using all 42 markers as inputs and adapted the method to not perform any variance scaling prior to NMF; this is because CyTOF intensities of different markers are already expressed on a comparable scale, unlike genes in single-cell RNA-Seq data which can vary over different orders of magnitude. We also set a floor on the input data to be non-negative (because after COMBAT, a small percentage of the values were slightly below 0). We selected K=7 as the cNMF dimensionality as the solution stability fell off dramatically at higher values. One gene expression program (GEP) was characterized by high levels of the platelet markers CD61 and BDCA3. It was mostly enriched in the platelet cluster but was also elevated in a subset of cells of all major cell types, which suggests that it reflects an artifact of platelets sticking to cells. A second GEP was characterized by high levels of all of the intracellular markers and it was also distributed throughout cells of multiple clusters. We interpreted this GEP as a cell permeabilization artifact reflecting the relative accessibility of a cell's intracellular proteins to CyTOF antibody staining. We regressed these 2 GEPs out of the data by subtracting the matrix (outer) product of the Usage and GEP matrices. We denote the number of cells as *C*, the number of genes as *G* (42 in our data), and the number of programs selected as K (7 in our data). The *C*x*G* input data matrix is denoted as X, the *K*x*G* GEP matrix returned by cNMF as G and the C*x*K usage matrix as U. Then the correction is as follows:

$$X^{corrected} = X - U^{platelet} x G^{platelet} - U^{perm} x G^{perm}$$

Where $U^{platelet}$ and U^{perm} are the *Cx*1 dimensional matrices representing the usage of the platelet and permeabilization GEPs and $G^{platelet}$ and G^{perm} are the 1x*G* dimensional matrices representing the spectra of the platelet and permeabilization GEPs, and multiplication is the outer product. We also filtered cells assigned to the platelet cluster in the initial clustering, since their predominant signal had been regressed out.

We then repeated clustering of the corrected data to identify broad cell types (Neutrophils, Monocytes, and DCs), followed by sub-clustering within each broad cluster to generate the sub-clusterings in **Figure S4**. The data was visualized using the UMAP algorithm as described for the scRNA-Seq data.

The human PBMC CyTOF data was processed with the same pipeline as the NHP data with a few modifications. The data were down-sampled to 280,000 cells total (20,000 per sample), batch correction was performed using Harmony, and 0.001 x the number of cells was used for K nearest neighbor graph construction. No cNMF or COMBAT adjustment steps were performed.

Quantification and Statistical Analysis

Details of statistical testing, sample size, center, and dispersion can be found in the figure legends, the main text, and Materials and Methods.

Additional Resources

This study did not generate new additional resources (website, forum, clinical trial).

Key resources table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
See **Table S7**	(Bjornson-Hooper et al., 2019a, 2019b)	
Bacterial and Virus Strains		
Ebola virus/H. sapiens-tc/COD/1995/Kikwit-9510621 (EBOV/Kikwit; GenBank accession MG572235.1; *Filoviridae: Zaire ebolavirus*)	BEI Resources	Cat#NR-50306
Biological Samples		
None		
Chemicals, Peptides, and Recombinant Proteins		
Xpert Ebola Lysis Reagent	Cepheid	Cat#GXEBOLA-CE-50
EnGen SauCas9, 20 μM	(Yourik et al., 2019); NEB	Cat#M0654T
16% paraformaldehyde	Electron Microscopy Sciences	Cat#15710
Critical Commercial Assays		
None		
Deposited Data		

EBOV NHP infection single-cell RNA-Seq	This paper	
EBOV NHP infection CyTOF	This paper	
Human cell atlas bone marrow scRNA-Seq	(Hay et al., 2018)	https://data.humancellatlas.org/
Human healthy PBMC scRNA-Seq	10X	https://support.10xgenomics.com/single-cell-gene-expression/datasets "Aggregate of 8 Chromium Connect channels and 8 manual channels", "5k Peripheral blood mononuclear cells (PBMCs) from a healthy donor (v3 chemistry)", "5k Peripheral blood mononuclear cells (PBMCs) from a healthy donor (Next GEM)", "5k Peripheral blood mononuclear cells (PBMCs) from a healthy donor with cell surface proteins (v3 chemistry)", "5k Peripheral blood mononuclear cells (PBMCs) from a healthy donor with cell surface proteins (Next GEM)", "10k PBMCs from a Healthy Donor - Gene Expression and Cell Surface Protein", "10k PBMCs from a Healthy Donor (v3 chemistry)"
EBOV human infection CyTOF	(McElroy et al., 2020)	Author correspondence
Experimental Models: Cell Lines		
Experimental Models: Organisms/Strains		
None		

Oligonucleotides		
Primer: dN-SMRT: AAGCAGTGGTATCAACGCAGAGTGANNNGGNNNB	(Hughes et al., 2019); IDT	N/A
gRNA sequence: DASH_SeqB: GGGNNNNAAGCAGUGGUAUCAACGGUUAUAGUACUCUGGAAACAGAAUCUACUAAAACAAGGCAAAAUGCCGUGUUUAUCUCGUCAACUUGUUGGCGAGAU	This paper	N/A
Recombinant DNA		
None		
Software and Algorithms		
DropSeqPipe	(Macosko et al., 2015)	https://github.com/Hoohm/dropSeqPipe
Scanpy	(Wolf et al., 2018)	https://github.com/theislab/scanpy
MAST	(Finak et al., 2015)	https://github.com/RGLab/MAST
HARMONY	(Korsunsky et al., 2019)	https://pypi.org/project/harmony-pytorch/, https://github.com/immunogenomics/harmony
MAGIC	(van Dijk et al., 2018)	https://github.com/KrishnaswamyLab/MAGIC
cNMF	(Kotliar et al., 2019)	https://github.com/dylkot/cNMF
Scrublet	(Wolock et al., 2019)	https://github.com/AllonKleinLab/scrublet
Other		
None		

Acknowledgements

We thank S. Schaffner, E. Normandin, K. Siddle, S. Reilly, S. Weingarten-Gabbay, C. Myhrvold, K. DeRuff, M. Rudy, N. Barkas, M. Babadi, M. Reyes, N. Hacohen, A. Regev, and E. Hodis, for helpful comments on this work. We thank S. Wolock for useful feedback and for providing scripts and processed data for investigating the human bone marrow samples. We thank D. Schwarz and New England Biolabs for generously providing SauCas9 and technical advice. We thank A. Matthews and M. Kemball for help in project management and administration. We thank S. Knemeyer and SciStories for illustrations.

This work is supported by the US Food and Drug Administration (FDA) contracts HHSF223201810172C and HHSF223201610018C, National Institute of Allergy and Infectious Diseases (NIAID) U19AI110818, and HHMI. This work was partially supported by NIAID Interagency agreement NOR15003-001-0000. The nonhuman primate work completed at the NIAID Integrated Research Facility was supported in part by the NIAID Division of Intramural Research and NIAID Division of Clinical Research and was performed under Battelle Memorial Institute Contract (No. HHSN272200700016I) and manuscript drafting was performed under Laulima Government Solutions, LLC. Contract (No. HHSN272201800013C). J.L. performed this work as an employee of Battelle. J.R.K., B. D.-K., R.A., and R.S.B. are current employees of Laulima Government Solutions, LLC. D.A.K. was supported by award Number T32GM007753 from the National Institute of General Medical Sciences (NIGMS). A.E.L. was supported by the National Science Foundation (NSF) under Grant No. DGE 1144152. M.M. was a Gilead Fellow of the Life Sciences Research Foundation. K.G.B was supported by a K01 (NIH-TW010853) and an ASTMH Shope fellowship. A.K.S. was supported by the Searle Scholars Program, the Beckman Young Investigator Program, a Sloan Fellowship in Chemistry, NIH 5U24AI118672, and the Bill and Melinda Gates Foundation. The CyTOF facility at the trans-NIH Center for Human Immunology is supported by funding from the Intramural Research Program of the NIH. The authors are solely responsible for the content of this paper, which does not necessarily represent the official views of the US Department of Health and Human Services (HHS), the NIH, the NIGMS, the FDA, or the institutions and companies affiliated with the authors.

Author Contributions

Conceptualization, D.K., A.E.L., T.K.H., G.P.N., K.G.B., L.E.H., D.R.M., A.K.S., P.C.S., R.S.B.; Methodology, D.K., A.E.L., J.L., T.K.H., N.M.K., M.H.W., D.R.M., R.S.B.; Software, D.K.; Validation, S.S.R.; Formal Analysis, D.K., A.E.L., S.S.R., H.C., M.M., D.R.M.; Investigation, J.L., N.M.K., H.C., J.R.K., B.D.-K., Z.B.B., N.M., B.A.S., M.R.B., G.C.A., R.A., K.G.B., D.R.M., R.S.B.; Resources, T.K.H., M.H.W., Z.B.B., N.T.; Data Curation, D.K., A.E.L., D.R.M., R.S.B.; Writing – Original Draft, D.K., A.E.L.; Writing – Review & Editing, D.K., A.E.L., J.L., T.K.H., N.M.K., S.S.R., M.H.W., H.C., J.R.K., B.D.-K., Z.B.B., N.M., B.A.S., N.T., M.R.B., G.C.A., R.A.,

J.L.R., M.M., G.P.N., K.G.B., L.E.H., D.R.M., A.K.S., P.C.S., R.S.B.; Visualization, D.K., A.E.L., S.S.R., H.C., D.R.M.; Supervision, J.L.R., M.M., G.P.N., K.G.B., L.E.H., D.R.M., A.K.S., P.C.S., R.S.B.; Project Administration, D.K., A.E.L., J.L., N.T., K.G.B., R.S.B.; Funding Acquisition, D.K., A.E.L., M.M., G.P.N., K.G.B., L.E.H., D.R.M., P.C.S.

Declaration of Interests

A.K.S. has received compensation for consulting and SAB membership from Honeycomb Biotechnologies, Cellarity, Orche Bio, Repertoire, and Dahlia Biosciences.

References

Albers, S., and Czech, A. (2016). Exploiting tRNAs to Boost Virulence. Life *6*.

Amici, S.A., Young, N.A., Narvaez-Miranda, J., Jablonski, K.A., Arcos, J., Rosas, L., Papenfuss, T.L., Torrelles, J.B., Jarjour, W.N., and Guerau-de-Arellano, M. (2018). CD38 Is Robustly Induced in Human Macrophages and Monocytes in Inflammatory Conditions. Front. Immunol. *9*, 1593.

Baize, S., Leroy, E.M., Georges-Courbot, M.C., Capron, M., Lansoud-Soukate, J., Debré, P., Fisher-Hoch, S.P., McCormick, J.B., and Georges, A.J. (1999). Defective humoral responses and extensive intravascular apoptosis are associated with fatal outcome in Ebola virus-infected patients. Nat. Med. *5*, 423–426.

Barbar, E. (2008). Dynein light chain LC8 is a dimerization hub essential in diverse protein networks. Biochemistry *47*, 503–508.

Basler, C.F., Mikulasova, A., Martinez-Sobrido, L., Paragas, J., Mühlberger, E., Bray, M., Klenk, H.-D., Palese, P., and García-Sastre, A. (2003). The Ebola virus VP35 protein inhibits activation of interferon regulatory factor 3. J. Virol. *77*, 7945–7956.

Becht, E., McInnes, L., Healy, J., Dutertre, C.-A., Kwok, I.W.H., Ng, L.G., Ginhoux, F., and Newell, E.W. (2018). Dimensionality reduction for visualizing single-cell data using UMAP. Nat. Biotechnol.

Bendall, S.C., Simonds, E.F., Qiu, P., Amir, E.-A.D., Krutzik, P.O., Finck, R., Bruggner, R.V., Melamed, R., Trejo, A., Ornatsky, O.I., et al. (2011). Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. Science *332*, 687–696.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J. R. Stat. Soc. Series B Stat. Methodol. *57*, 289–300.

Bennett, R.S., Huzella, L.M., Jahrling, P.B., Bollinger, L., Olinger, G.G., Jr, and Hensley, L.E. (2017). Nonhuman Primate Models of Ebola Virus Disease. Curr. Top. Microbiol. Immunol. *411*, 171–193.

Bjornson-Hooper, Z.B., Fragiadakis, G.K., Spitzer, M.H., Madhireddy, D., Hu, K., Lundsten, K., and Nolan, G.P. (2019a). Cell type-specific monoclonal antibody cross-reactivity screening in non-human primates and development of comparative immunophenotyping panels for CyTOF.

Bjornson-Hooper, Z.B., Fragiadakis, G.K., Spitzer, M.H., Madhireddy, D., McIlwain, D., and Nolan, G.P. (2019b). A comprehensive atlas of immunological differences between humans, mice and non-human primates.

Bomans, K., Schenz, J., Sztwiertnia, I., Schaack, D., Weigand, M.A., and Uhle, F. (2018). Sepsis Induces a Long-Lasting State of Trained Immunity in Bone Marrow Monocytes. Front. Immunol. *9*, 2685.

Bosio, C.M., Aman, M.J., Grogan, C., Hogan, R., Ruthel, G., Negley, D., Mohamadzadeh, M., Bavari, S., and Schmaljohn, A. (2003). Ebola and Marburg viruses replicate in monocyte-derived dendritic cells without inducing the production of cytokines and full maturation. J. Infect. Dis. *188*, 1630–1638.

Brauburger, K., Boehmann, Y., Tsuda, Y., Hoenen, T., Olejnik, J., Schümann, M., Ebihara, H., and Mühlberger, E. (2014). Analysis of the highly diverse gene borders in Ebola virus reveals a distinct mechanism of transcriptional regulation. J. Virol. *88*, 12558–12571.

Brauburger, K., Boehmann, Y., Krähling, V., and Mühlberger, E. (2016). Transcriptional regulation in Ebola virus: effects of gene border structure and regulatory elements on gene expression and polymerase scanning behavior. J. Virol. *90*, 1898–1909.

Bray, M. (2001). The role of the Type I interferon response in the resistance of mice to filovirus infection. J. Gen. Virol. *82*, 1365–1373.

Bray, M., and Mahanty, S. (2003). Ebola hemorrhagic fever and septic shock. J. Infect. Dis. *188*, 1613–1617.

Bray, M., Hatfill, S., Hensley, L., and Huggins, J.W. (2001). Haematological, biochemical and coagulation changes in mice, guinea-pigs and monkeys infected with a mouse-adapted variant of Ebola Zaire virus. J. Comp. Pathol. *125*, 243–253.

Caballero, I.S., Honko, A.N., Gire, S.K., Winnicki, S.M., Melé, M., Gerhardinger, C., Lin, A.E., Rinn, J.L., Sabeti, P.C., Hensley, L.E., et al. (2016). In vivo Ebola virus infection leads to a strong innate response in circulating immune cells. BMC Genomics *17*, 707.

Carette, J.E., Raaben, M., Wong, A.C., Herbert, A.S., Obernosterer, G., Mulherkar, N., Kuehne, A.I., Kranzusch, P.J., Griffin, A.M., Ruthel, G., et al. (2011). Ebola virus entry requires the cholesterol transporter Niemann-Pick C1. Nature *477*, 340–343.

Carmona-Rivera, C., and Kaplan, M.J. (2013). Low-density granulocytes: a distinct class of neutrophils in systemic autoimmunity. Semin. Immunopathol. *35*, 455–463.

Carroll, M.W., Haldenby, S., Rickett, N.Y., Pályi, B., Garcia-Dorival, I., Liu, X., Barker, G., Bore, J.A., Koundouno, F.R., Williamson, E.D., et al. (2017). Deep sequencing of RNA from blood and oral swab samples reveals the presence of nucleic acid from a number of pathogens in patients

with acute Ebola virus disease and is consistent with bacterial translocation across the gut. mSphere *2*.

Chandran, K., Sullivan, N.J., Felbor, U., Whelan, S.P., and Cunningham, J.M. (2005). Endosomal proteolysis of the Ebola virus glycoprotein is necessary for infection. Science *308*, 1643–1645.

Chiba, Y., Mizoguchi, I., Hasegawa, H., Ohashi, M., Orii, N., Nagai, T., Sugahara, M., Miyamoto, Y., Xu, M., Owaki, T., et al. (2018). Regulation of myelopoiesis by proinflammatory cytokines in infectious diseases. Cell. Mol. Life Sci. *75*, 1363–1376.

Côté, M., Misasi, J., Ren, T., Bruchez, A., Lee, K., Filone, C.M., Hensley, L., Li, Q., Ory, D., Chandran, K., et al. (2011). Small molecule inhibitors reveal Niemann-Pick C1 is essential for Ebola virus infection. Nature *477*, 344–348.

Cuenca, A.G., Cuenca, A.L., Gentile, L.F., Efron, P.A., Islam, S., Moldawer, L.L., Kays, D.W., and Larson, S.D. (2015). Delayed emergency myelopoiesis following polymicrobial sepsis in neonates. Innate Immun. *21*, 386–391.

Dall, E., and Brandstetter, H. (2016). Structure and function of legumain in health and disease. Biochimie *122*, 126–150.

Darcy, C.J., Minigo, G., Piera, K.A., Davis, J.S., McNeil, Y.R., Chen, Y., Volkheimer, A.D., Weinberg, J.B., Anstey, N.M., and Woodberry, T. (2014). Neutrophils with myeloid derived suppressor function deplete arginine and constrain T cell function in septic shock patients. Crit. Care *18*, R163.

Deflubé, L.R., Cressey, T.N., Hume, A.J., Olejnik, J., Haddock, E., Feldmann, F., Ebihara, H., Fearns, R., and Mühlberger, E. (2019). Ebola virus polymerase uses an unconventional genome replication mechanism. Proc. Natl. Acad. Sci. USA *116*, 8535–8543.

Deng, Y., Ye, J., Luo, Q., Huang, Z., Peng, Y., and Xiong, G. (2016). Low-density granulocytes are elevated in mycobacterial infection and associated with the severity of tuberculosis. PLoS One.

Diehl, W.E., Lin, A.E., Grubaugh, N.D., Carvalho, L.M., Kim, K., Kyawe, P.P., McCauley, S.M., Donnard, E., Kucukural, A., McDonel, P., et al. (2016). Ebola virus glycoprotein with increased infectivity dominated the 2013–2016 epidemic. Cell *167*, 1088–1098.e6.

van Dijk, D., Sharma, R., Nainys, J., Yim, K., Kathail, P., Carr, A.J., Burdziak, C., Moon, K.R., Chaffer, C.L., Pattabiraman, D., et al. (2018). Recovering Gene Interactions from Single-Cell Data Using Data Diffusion. Cell *174*, 716–729.e27.

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.

Dong, C., Zhao, G., Zhong, M., Yue, Y., Wu, L., and Xiong, S. (2013). RNA sequencing and transcriptomal analysis of human monocyte to macrophage differentiation. Gene *519*, 279–287.

Ebihara, H., Rockx, B., Marzi, A., Feldmann, F., Haddock, E., Brining, D., LaCasse, R.A., Gardner, D., and Feldmann, H. (2011). Host response dynamics following lethal infection of rhesus macaques with Zaire ebolavirus. J. Infect. Dis. *204 Suppl 3*, S991–S999.

Eisfeld, A.J., Halfmann, P.J., Wendler, J.P., Kyle, J.E., Burnum-Johnson, K.E., Peralta, Z., Maemura, T., Walters, K.B., Watanabe, T., Fukuyama, S., et al. (2017). Multi-platform'omics analysis of human Ebola virus disease pathogenesis. Cell Host Microbe *22*, 817–829.

Engelmann, P., Farkas, K., Kis, J., Richman, G., Zhang, Z., Liew, C.W., Borowiec, M., Niewczas, M.A., Jalahej, H., and Orbán, T. (2011). Characterization of human invariant natural killer T cells expressing FoxP3. Int. Immunol. *23*, 473–484.

Feldmann, F., Shupert, W.L., Haddock, E., Twardoski, B., and Feldmann, H. (2019). Gamma Irradiation as an Effective Method for Inactivation of Emerging Viral Pathogens. Am. J. Trop. Med. Hyg. *100*, 1275–1277.

Feng, K., Divers, E., Ma, Y., and Li, J. (2011). Inactivation of a human norovirus surrogate, human norovirus virus-like particles, and vesicular stomatitis virus by gamma irradiation. Appl. Environ. Microbiol. *77*, 3507–3517.

Finak, G., McDavid, A., Yajima, M., Deng, J., Gersuk, V., Shalek, A.K., Slichter, C.K., Miller, H.W., McElrath, M.J., Prlic, M., et al. (2015). MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. Genome Biol. *16*, 278.

Finck, R., Simonds, E.F., Jager, A., Krishnaswamy, S., Sachs, K., Fantl, W., Pe'er, D., Nolan, G.P., and Bendall, S.C. (2013). Normalization of mass cytometry data with bead standards. Cytometry A *83*, 483–494.

Fingerle, G., Pforte, A., Passlick, B., Blumenstein, M., Ströbel, M., and Ziegler-Heitbrock, H.W. (1993). The novel subset of CD14+/CD16+ blood monocytes is expanded in sepsis patients. Blood *82*, 3170–3176.

Fisher-Hoch, S.P., Platt, G.S., Neild, G.H., Southee, T., Baskerville, A., Raymond, R.T., Lloyd, G., and Simpson, D.I.H. (1985). Pathophysiology of Shock and Hemorrhage in a Fulminating Viral Infection (Ebola). Journal of Infectious Diseases *152*, 887–894.

Fleming, S.J., Marioni, J.C., and Babadi, M. (2019). CellBender remove-background: a deep generative model for unsupervised removal of background noise from scRNA-seq datasets. bioRxiv.

Foletta, V.C., Segal, D.H., and Cohen, D.R. (1998). Transcriptional regulation in the immune system: all roads lead to AP-1. Journal of Leukocyte Biology *63*, 139–152.

Friedland, A.E., Baral, R., Singhal, P., Loveluck, K., Shen, S., Sanchez, M., Marco, E., Gotta, G.M., Maeder, M.L., Kennedy, E.M., et al. (2015). Characterization of Staphylococcus aureus

Cas9: a smaller Cas9 for all-in-one adeno-associated virus delivery and paired nickase applications. Genome Biol. *16*, 257.

van Furth, R., Raeburn, J.A., and van Zwet, T.L. (1979). Characteristics of human mononuclear phagocytes. Blood *54*, 485–500.

Geisbert, T.W., Hensley, L.E., Gibb, T.R., Steele, K.E., Jaax, N.K., and Jahrling, P.B. (2000). Apoptosis induced in vitro and in vivo during infection by Ebola and Marburg viruses. Lab. Invest. *80*, 171–186.

Geisbert, T.W., Pushko, P., Anderson, K., Smith, J., Davis, K.J., and Jahrling, P.B. (2002). Evaluation in nonhuman primates of vaccines against Ebola virus. Emerg. Infect. Dis. *8*, 503–507.

Geisbert, T.W., Young, H.A., Jahrling, P.B., Davis, K.J., Larsen, T., Kagan, E., and Hensley, L.E. (2003a). Pathogenesis of Ebola hemorrhagic fever in primate models: evidence that hemorrhage is not a direct effect of virus-induced cytolysis of endothelial cells. Am. J. Pathol. *163*, 2371–2382.

Geisbert, T.W., Young, H.A., Jahrling, P.B., Davis, K.J., Kagan, E., and Hensley, L.E. (2003b). Mechanisms underlying coagulation abnormalities in ebola hemorrhagic fever: overexpression of tissue factor in primate monocytes/macrophages is a key event. J. Infect. Dis. *188*, 1618–1629.

Geisbert, T.W., Hensley, L.E., Larsen, T., Young, H.A., Reed, D.S., Geisbert, J.B., Scott, D.P., Kagan, E., Jahrling, P.B., and Davis, K.J. (2003c). Pathogenesis of Ebola hemorrhagic fever in cynomolgus macaques: evidence that dendritic cells are early and sustained targets of infection. Am. J. Pathol. *163*, 2347–2370.

Geisbert, T.W., Strong, J.E., and Feldmann, H. (2015). Considerations in the Use of Nonhuman Primate Models of Ebola Virus and Marburg Virus Infection. J. Infect. Dis. *212 Suppl 2*, S91–S97.

Gierahn, T.M., Wadsworth, M.H., II, Hughes, T.K., Bryson, B.D., Butler, A., Satija, R., Fortune, S., Christopher Love, J., and Shalek, A. (2017). Seq-Well: portable, low-cost RNA sequencing of single cells at high throughput. Protocol Exchange.

Ginoza, W. (1967). The Effects of Ionizing Radiation on Nucleic Acids of Bacteriophages and Bacterial Cells. Annual Review of Nuclear Science *17*, 469–512.

Gnirß, K., Kühl, A., Karsten, C., Glowacka, I., Bertram, S., Kaup, F., Hofmann, H., and Pöhlmann, S. (2012). Cathepsins B and L activate Ebola but not Marburg virus glycoproteins for efficient entry into cell lines and macrophages independent of TMPRSS2 expression. Virology *424*, 3–10.

Greenberg, A.R., Huber, B.R., Liu, D.X., Logue, J.P., Woolsey Hischak, A.M., Hart, R.J., Abbott, M., Isic, N., Hisada, Y.M., Mackman, N., et al. (2020). Quantification of Viral and Host

Biomarkers in the Liver of Rhesus Macaques: a Longitudinal Study of Zaire Ebolavirus strain Kikwit (EBOV/Kik). Am. J. Pathol.

Gu, W., Crawford, E.D., O'Donovan, B.D., Wilson, M.R., Chow, E.D., Retallack, H., and DeRisi, J.L. (2016). Depletion of abundant sequences by hybridization (DASH): using Cas9 to remove unwanted high-abundance species in sequencing libraries and molecular counting applications. Genome Biol. *17*, 41.

Gupta, M., Mahanty, S., Ahmed, R., and Rollin, P.E. (2001). Monocyte-derived human macrophages and peripheral blood mononuclear cells infected with ebola virus secrete MIP-1alpha and TNF-alpha and inhibit poly-IC-induced IFN-alpha in vitro. Virology *284*, 20–25.

Hamlin, R.E., Rahman, A., Pak, T.R., Maringer, K., Mena, I., Bernal-Rubio, D., Potla, U., Maestre, A.M., Fredericks, A.C., Amir, E.-A.D., et al. (2017). High-dimensional CyTOF analysis of dengue virus-infected human DCs reveals distinct viral signatures. JCI Insight *2*.

Harcourt, B.H., Sanchez, A., and Offermann, M.K. (1998). Ebola virus inhibits induction of genes by double-stranded RNA in endothelial cells. Virology *252*, 179–188.

Harcourt, B.H., Sanchez, A., and Offermann, M.K. (1999). Ebola virus selectively inhibits responses to interferons, but not to interleukin-1beta, in endothelial cells. J. Virol. *73*, 3491–3496.

Hawrylowicz, C.M., Guida, L., and Paleolog, E. (1994). Dexamethasone up-regulates granulocyte-macrophage colony-stimulating factor receptor expression on human monocytes. Immunology *83*, 274–280.

Hay, S.B., Ferchen, K., Chetal, K., Grimes, H.L., and Salomonis, N. (2018). The Human Cell Atlas bone marrow single-cell interactive web portal. Exp. Hematol. *68*, 51–61.

Hein, M.Y., and Weissman, J.S. (2019). Functional single-cell genomics of human cytomegalovirus infection. bioRxiv.

Hensley, L.E., Young, H.A., Jahrling, P.B., and Geisbert, T.W. (2002). Proinflammatory response during Ebola virus infection of primate models: possible involvement of the tumor necrosis factor receptor superfamily. Immunol. Lett. *80*, 169–179.

Hérault, A., Binnewies, M., Leong, S., Calero-Nieto, F.J., Zhang, S.Y., Kang, Y.-A., Wang, X., Pietras, E.M., Chu, S.H., Barry-Holson, K., et al. (2017). Myeloid progenitor cluster formation drives emergency and leukaemic myelopoiesis. Nature *544*, 53–58.

Hölzer, M., Krähling, V., Amman, F., Barth, E., Bernhart, S.H., Carmelo, V.A.O., Collatz, M., Doose, G., Eggenhofer, F., Ewald, J., et al. (2016). Differential transcriptional responses to Ebola and Marburg virus infection in bat and human cells. Scientific Reports *6*.

Honess, R.W., and Roizman, B. (1974). Regulation of herpesvirus macromolecular synthesis. I. Cascade regulation of the synthesis of three groups of viral proteins. J. Virol. *14*, 8–19.

Hughes, T.K., Wadsworth, M.H., Gierahn, T.M., and Do, T. (2019). Highly efficient, massively-parallel single-cell RNA-seq reveals cellular states and molecular features of human skin pathology. bioRxiv.

Iampietro, M., Younan, P., Nishida, A., Dutta, M., Lubaki, N.M., Santos, R.I., Koup, R.A., Katze, M.G., and Bukreyev, A. (2017). Ebola virus glycoprotein directly triggers T lymphocyte death despite of the lack of infection. PLoS Pathog. *13*, e1006397.

Ilunga Kalenga, O., Moeti, M., Sparrow, A., Nguyen, V.-K., Lucey, D., and Ghebreyesus, T.A. (2019). The Ongoing Ebola Epidemic in the Democratic Republic of Congo, 2018-2019. N. Engl. J. Med. *381*, 373–383.

Irigoyen, N., Firth, A.E., Jones, J.D., Chung, B.Y.-W., Siddell, S.G., and Brierley, I. (2016). High-Resolution Analysis of Coronavirus Gene Expression by RNA Sequencing and Ribosome Profiling. PLoS Pathog. *12*, e1005473.

Italiani, P., Mazza, E.M.C., Lucchesi, D., Cifola, I., Gemelli, C., Grande, A., Battaglia, C., Bicciato, S., and Boraschi, D. (2014). Transcriptomic profiling of the development of the inflammatory response in human monocytes in vitro. PLoS One *9*, e87680.

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

Jurado, S., Conlan, L.A., Baker, E.K., Ng, J.-L., Tenis, N., Hoch, N.C., Gleeson, K., Smeets, M., Izon, D., and Heierhorst, J. (2012). ATM substrate Chk2-interacting Zn2+ finger (ASCIZ) Is a bi-functional transcriptional activator and feedback sensor in the regulation of dynein light chain (DYNLL1) expression. J. Biol. Chem. *287*, 3156–3164.

Kapellos, T.S., Bonaguro, L., Gemünd, I., Reusch, N., Saglam, A., Hinkley, E.R., and Schultze, J.L. (2019). Human Monocyte Subsets and Phenotypes in Major Chronic Inflammatory Diseases. Front. Immunol. *10*, 2035.

Kash, J.C., Muhlberger, E., Carter, V., Grosch, M., Perwitasari, O., Proll, S.C., Thomas, M.J., Weber, F., -D. Klenk, H., and Katze, M.G. (2006). Global Suppression of the Host Antiviral Response by Ebola- and Marburgviruses: Increased Antagonism of the Type I Interferon Response Is Associated with Enhanced Virulence. Journal of Virology *80*, 3009–3020.

Kennedy, S.B., Bolay, F., Kieh, M., Grandits, G., Badio, M., Ballou, R., Eckes, R., Feinberg, M., Follmann, D., Grund, B., et al. (2017). Phase 2 Placebo-Controlled Trial of Two Vaccines to Prevent Ebola in Liberia. New England Journal of Medicine *377*, 1438–1447.

Kim, H.M., Han, J.W., and Chan, J.Y. (2016). Nuclear Factor Erythroid-2 Like 1 (NFE2L1): Structure, function and regulation. Gene *584*, 17–25.

King, B.R., Samacoits, A., Eisenhauer, P.L., Ziegler, C.M., Bruce, E.A., Zenklusen, D., Zimmer, C., Mueller, F., and Botten, J. (2018). Visualization of Arenavirus RNA Species in Individual Cells by Single-Molecule Fluorescence In Situ Hybridization Suggests a Model of Cyclical Infection and Clearance during Persistence. Journal of Virology *92*.

Klein, A.M., Mazutis, L., Akartuna, I., Tallapragada, N., Veres, A., Li, V., Peshkin, L., Weitz, D.A., and Kirschner, M.W. (2015). Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. Cell *161*, 1187–1201.

Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y., Brenner, M., Loh, P.-R., and Raychaudhuri, S. (2019). Fast, sensitive and accurate integration of single-cell data with Harmony. Nature Methods *16*, 1289–1296.

Kotliar, D., Veres, A., Nagy, M.A., Tabrizi, S., Hodis, E., Melton, D.A., and Sabeti, P.C. (2019). Identifying gene expression programs of cell-type identity and cellular activity with single-cell RNA-Seq. Elife *8*.

Krovi, S.H., and Gapin, L. (2018). Invariant Natural Killer T Cell Subsets—More Than Just Developmental Intermediates. Front. Immunol. *9*, 1393.

Kubota, T., Matsuoka, M., Chang, T.-H., Bray, M., Jones, S., Tashiro, M., Kato, A., and Ozato, K. (2009). Ebolavirus VP35 interacts with the cytoplasmic dynein light chain 8. J. Virol. *83*, 6952–6956.

Leung, L.W., Hartman, A.L., Martinez, O., Shaw, M.L., Carbonnelle, C., Volchkov, V.E., Nichol, S.T., Basler, C.F., and Others (2006). Ebola virus VP24 binds karyopherin α1 and blocks STAT1 nuclear accumulation. J. Virol. *80*, 5156–5167.

Li, M.M.H., MacDonald, M.R., and Rice, C.M. (2015). To translate, or not to translate: viral and host mRNA regulation by interferon-stimulated genes. Trends Cell Biol. *25*, 320–329.

Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdóttir, H., Tamayo, P., and Mesirov, J.P. (2011). Molecular signatures database (MSigDB) 3.0. Bioinformatics *27*, 1739–1740.

Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J.P., and Tamayo, P. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst *1*, 417–425.

Lin, W.-J., Zheng, X., Lin, C.-C., Tsao, J., Zhu, X., Cody, J.J., Coleman, J.M., Gherzi, R., Luo, M., Townes, T.M., et al. (2011). Posttranscriptional control of type I interferon genes by KSRP in the innate immune response against viral infection. Mol. Cell. Biol. *31*, 3196–3207.

Liu, B., Dhanda, A., Hirani, S., Williams, E.L., Sen, H.N., Martinez Estrada, F., Ling, D., Thompson, I., Casady, M., Li, Z., et al. (2015). CD14++CD16+ Monocytes Are Enriched by Glucocorticoid Treatment and Are Functionally Attenuated in Driving Effector T Cell Responses. J. Immunol. *194*, 5150–5160.

Liu, X., Speranza, E., Muñoz-Fontela, C., Haldenby, S., Rickett, N.Y., Garcia-Dorival, I., Fang, Y., Hall, Y., Zekeng, E.-G., Lüdtke, A., et al. (2017). Transcriptomic signatures differentiate survival from fatal outcomes in humans infected with Ebola virus. Genome Biol. *18*, 4.

Lo, T.Q., Marston, B.J., Dahl, B.A., and De Cock, K.M. (2017). Ebola: Anatomy of an Epidemic. Annu. Rev. Med. *68*, 359–370.

Logue, J., Solomon, J., Niemeyer, B.F., Benam, K.H., Lin, A.E., Bjornson, Z., Jiang, S., McIlwain, D.R., Nolan, G.P., Palacios, G., et al. (2019). Innovative technologies for advancement of WHO risk group 4 pathogens research. In Global Virology III: Virology in the 21st Century, P. Shapshak, S. Balaji, P. Kangueane, F. Chiappelli, C. Somboonwit, L.J. Menezes, and J.T. Sinnott, eds. (Cham: Springer International Publishing), pp. 437–469.

Lubaki, N.M., Ilinykh, P., Pietzsch, C., Tigabu, B., Freiberg, A.N., Koup, R.A., and Bukreyev, A. (2013). The lack of maturation of Ebola virus-infected dendritic cells results from the cooperative effect of at least two viral domains. J. Virol. *87*, 7471–7485.

Lüdtke, A., Ruibal, P., Becker-Ziaja, B., Rottstegge, M., Wozniak, D.M., Cabeza-Cabrerizo, M., Thorenz, A., Weller, R., Kerber, R., Idoyaga, J., et al. (2016). Ebola Virus Disease Is Characterized by Poor Activation and Reduced Levels of Circulating CD16+ Monocytes. J. Infect. Dis. *214*, S275–S280.

Luthra, P., Jordan, D.S., Leung, D.W., Amarasinghe, G.K., and Basler, C.F. (2015). Ebola virus VP35 interaction with dynein LC8 regulates viral RNA synthesis. J. Virol. *89*, 5148–5153.

Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A.R., Kamitaki, N., Martersteck, E.M., et al. (2015). Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell *161*, 1202–1214.

Mahanty, S., Hutchinson, K., Agarwal, S., McRae, M., Rollin, P.E., and Pulendran, B. (2003). Cutting edge: impairment of dendritic cells and adaptive immunity by Ebola and Lassa viruses. J. Immunol. *170*, 2797–2801.

Malvy, D., McElroy, A.K., de Clerck, H., Günther, S., and van Griensven, J. (2019). Ebola virus disease. Lancet *393*, 936–948.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal *17*, 10–12.

Martinez, O., Johnson, J.C., Honko, A., Yen, B., Shabman, R.S., Hensley, L.E., Olinger, G.G., and Basler, C.F. (2013). Ebola virus exploits a monocyte differentiation program to promote its entry. J. Virol. *87*, 3801–3814.

McArdel, S.L., Terhorst, C., and Sharpe, A.H. (2016). Roles of CD48 in regulating immunity and tolerance. Clin. Immunol. *164*, 10–20.

McElroy, A.K., Akondy, R.S., McIlwain, D.R., Chen, H., Bjornson-Hooper, Z., Mukherjee, N., Mehta, A.K., Nolan, G., Nichol, S.T., and Spiropoulou, C.F. (2020). Immunologic timeline of Ebola virus disease and recovery in humans. JCI Insight *5*.

McInnes, L., Healy, J., and Melville, J. (2018). UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction.

Menicucci, A.R., Versteeg, K., Woolsey, C., Mire, C.E., Geisbert, J.B., Cross, R.W., Agans, K.N., Jankeel, A., Geisbert, T.W., and Messaoudi, I. (2017). Transcriptome Analysis of

Circulating Immune Cell Subsets Highlight the Role of Monocytes in Zaire Ebola Virus Makona Pathogenesis. Front. Immunol. *8*, 1372.

Merino-Gracia, J., García-Mayoral, M.F., and Rodríguez-Crespo, I. (2011). The association of viral proteins with host cell dynein components during virus infection. FEBS J. *278*, 2997–3011.

Michlmayr, D., Pak, T.R., Rahman, A.H., Amir, E.-A.D., Kim, E.-Y., Kim-Schulze, S., Suprun, M., Stewart, M.G., Thomas, G.P., Balmaseda, A., et al. (2018). Comprehensive innate immune profiling of chikungunya virus infection in pediatric cases. Mol. Syst. Biol. *14*, e7862.

Monneret, G., and Venet, F. (2014). Monocyte HLA-DR in sepsis: shall we stop following the flow? Crit. Care *18*, 102.

Mühlberger, E. (2007). Filovirus replication and transcription. Future Virol. *2*, 205–215.

Mulangu, S., Dodd, L.E., Davey, R.T., Jr, Tshiani Mbaya, O., Proschan, M., Mukadi, D., Lusakibanza Manzo, M., Nzolo, D., Tshomba Oloma, A., Ibanda, A., et al. (2019). A Randomized, Controlled Trial of Ebola Virus Disease Therapeutics. N. Engl. J. Med. *381*, 2293–2303.

Newell, E.W., Sigal, N., Bendall, S.C., Nolan, G.P., and Davis, M.M. (2012). Cytometry by time-of-flight shows combinatorial cytokine expression and virus-specific cell niches within a continuum of CD8+ T cell phenotypes. Immunity *36*, 142–152.

Nockher, W.A., and Scherberich, J.E. (1997). Expression and release of the monocyte lipopolysaccharide receptor antigen CD14 are suppressed by glucocorticoids in vivo and in vitro. J. Immunol. *158*, 1345–1352.

Nockher, W.A., and Scherberich, J.E. (1998). Expanded CD14+ CD16+ monocyte subpopulation in patients with acute and chronic infections undergoing hemodialysis. Infect. Immun. *66*, 2782–2790.

O'Neal, J.T., Upadhyay, A.A., Wolabaugh, A., Patel, N.B., Bosinger, S.E., and Suthar, M.S. (2019). West Nile Virus-Inclusive Single-Cell RNA Sequencing Reveals Heterogeneity in the Type I Interferon Response within Single Cells. J. Virol. *93*.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al. (2011). Scikit-learn: Machine Learning in Python. J. Mach. Learn. Res. *12*, 2825–2830.

Pisapia, L., Hamilton, R.S., Farina, F., D'Agostino, V., Barba, P., Strazzullo, M., Provenzani, A., Gianfrani, C., and Del Pozzo, G. (2019). Tristetraprolin/ZFP36 Regulates the Turnover of Autoimmune-Associated HLA-DQ mRNAs. Cells *8*.

Ran, F.A., Cong, L., Yan, W.X., Scott, D.A., Gootenberg, J.S., Kriz, A.J., Zetsche, B., Shalem, O., Wu, X., Makarova, K.S., et al. (2015). In vivo genome editing using Staphylococcus aureus Cas9. Nature *520*, 186–191.

Rasmussen, A.L., Okumura, A., Ferris, M.T., Green, R., Feldmann, F., Kelly, S.M., Scott, D.P., Safronetz, D., Haddock, E., LaCasse, R., et al. (2014). Host genetic diversity enables Ebola hemorrhagic fever pathogenesis and resistance. Science *346*, 987–991.

Reid, S.P., Valmas, C., Martinez, O., Sanchez, F.M., and Basler, C.F. (2008). Ebola Virus VP24 Proteins Inhibit the Interaction of NPI-1 Subfamily Karyopherin Proteins with Activated STAT1. Journal of Virology *82*, 3163–3163.

Reid, S.P., Shurtleff, A.C., Costantino, J.A., Tritsch, S.R., Retterer, C., Spurgers, K.B., and Bavari, S. (2014). HSPA5 is an essential host factor for Ebola virus infection. Antiviral Res. *109*, 171–174.

Reisler, R.B., Zeng, X., Schellhase, C.W., Bearss, J.J., Warren, T.K., Trefry, J.C., Christopher, G.W., Kortepeter, M.G., Bavari, S., and Cardile, A.P. (2018). Ebola Virus Causes Intestinal Tract Architectural Disruption and Bacterial Invasion in Non-Human Primates. Viruses *10*.

Reyes, M., Filbin, M.R., Bhattacharyya, R.P., Billman, K., Eisenhaure, T., Hung, D.T., Levy, B.D., Baron, R.M., Blainey, P.C., Goldberg, M.B., et al. (2020). An immune-cell signature of bacterial sepsis. Nat. Med. *26*, 333–340.

Reynard, S., Journeaux, A., Gloaguen, E., Schaeffer, J., Varet, H., Pietrosemoli, N., Mateo, M., Baillet, N., Laouenan, C., Raoul, H., et al. (2019). Immune parameters and outcomes during Ebola virus disease. JCI Insight *4*.

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

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*.

Saeed, S., Quintin, J., Kerstens, H.H.D., Rao, N.A., Aghajanirefah, A., Matarese, F., Cheng, S.-C., Ratter, J., Berentsen, K., van der Ent, M.A., et al. (2014). Epigenetic programming of monocyte-to-macrophage differentiation and trained innate immunity. Science *345*, 1251086.

Satija, R., Farrell, J.A., Gennert, D., Schier, A.F., and Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. Nat. Biotechnol. *33*, 495–502.

Sayed, M.F.A., Al Sayed, M.F., Amrein, M.A., Bührer, E.D., Huguenin, A.-L., Radpour, R., Riether, C., and Ochsenbein, A.F. (2019). T-cell–Secreted TNFα Induces Emergency Myelopoiesis and Myeloid-Derived Suppressor Cell Differentiation in Cancer. Cancer Research *79*, 346–359.

Schornberg, K., Matsuyama, S., Kabsch, K., Delos, S., Bouton, A., and White, J. (2006). Role of endosomal cathepsins in entry mediated by the Ebola virus glycoprotein. J. Virol. *80*, 4174–4178.

Schröder, B. (2016). The multifaceted roles of the invariant chain CD74--More than just a chaperone. Biochim. Biophys. Acta *1863*, 1269–1281.

Seabold, S., and Perktold, J. (2010). Econometric and statistical modeling with Python skipper seabold 1 1. In Proc 9th Python Sci Conf, p. 61.

Shin, Y.C., Bischof, G.F., Lauer, W.A., and Desrosiers, R.C. (2015). Importance of codon usage for the temporal regulation of viral gene expression. Proc. Natl. Acad. Sci. U. S. A. *112*, 14030–14035.

Sinai, P., Nguyen, C., Schatzle, J.D., and Wülfing, C. (2010). Transience in polarization of cytolytic effectors is required for efficient killing and controlled by Cdc42. Proc. Natl. Acad. Sci. U. S. A. *107*, 11912–11917.

Smith, T., Heger, A., and Sudbery, I. UMI-tools: Modelling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy.

Speranza, E., Bixler, S.L., Altamura, L.A., Arnold, C.E., Pratt, W.D., Taylor-Howell, C., Burrows, C., Aguilar, W., Rossi, F., Shamblin, J.D., et al. (2018). A conserved transcriptional response to intranasal Ebola virus exposure in nonhuman primates prior to onset of fever. Sci. Transl. Med. *10*.

St Claire, M.C., Ragland, D.R., Bollinger, L., and Jahrling, P.B. (2017). Animal Models of Ebolavirus Infection. Comp. Med. *67*, 253–262.

Steimle, V., Siegrist, C.A., Mottet, A., Lisowska-Grospierre, B., and Mach, B. (1994). Regulation of MHC class II expression by interferon-gamma mediated by the transactivator gene CIITA. Science *265*, 106–109.

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.e4.

Summers, C., Rankin, S.M., Condliffe, A.M., Singh, N., Peters, A.M., and Chilvers, E.R. (2010). Neutrophil kinetics in health and disease. Trends Immunol. *31*, 318–324.

Tan, G.S., Preuss, M.A.R., Williams, J.C., and Schnell, M.J. (2007). The dynein light chain 8 binding motif of rabies virus phosphoprotein promotes efficient viral transcription. Proc. Natl. Acad. Sci. U. S. A. *104*, 7229–7234.

Testi, R., D'Ambrosio, D., De Maria, R., and Santoni, A. (1994). The CD69 receptor: a multipurpose cell-surface trigger for hematopoietic cells. Immunol. Today *15*, 479–483.

Traag, V.A., Waltman, L., and van Eck, N.J. (2019). From Louvain to Leiden: guaranteeing well-connected communities. Sci. Rep. *9*, 5233.

Tulzo, Y.L., Le Tulzo, Y., Pangault, C., Amiot, L., Guilloux, V., Tribut, O., Arvieux, C., Camus, C., Fauchet, R., Thomas, R., et al. (2004). Monocyte Human Leukocyte Antigen–DR Transcriptional Downregulation by Cortisol during Septic Shock. American Journal of Respiratory and Critical Care Medicine *169*, 1144–1151.

Tybulewicz, V.L.J., and Henderson, R.B. (2009). Rho family GTPases and their regulators in lymphocytes. Nat. Rev. Immunol. *9*, 630–644.

Ward, R.L. (1980). Mechanisms of poliovirus inactivation by the direct and indirect effects of ionizing radiation. Radiat. Res. *83*, 330–344.

Wauquier, N., Becquart, P., Padilla, C., Baize, S., and Leroy, E.M. (2010). Human fatal zaire ebola virus infection is associated with an aberrant innate immunity and with massive lymphocyte apoptosis. PLoS Negl. Trop. Dis. *4*.

Wolf, F.A., Angerer, P., and Theis, F.J. (2018). SCANPY: large-scale single-cell gene expression data analysis. Genome Biol. *19*, 15.

Wolk, K., Döcke, W.D., von Baehr, V., Volk, H.D., and Sabat, R. (2000). Impaired antigen presentation by human monocytes during endotoxin tolerance. Blood *96*, 218–223.

Wolock, S.L., Lopez, R., and Klein, A.M. (2019). Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. Cell Syst *8*, 281–291.e9.

World Health Organization (2019). Ebola virus disease – Democratic Republic of the Congo.

Young, M.D., and Behjati, S. (2018). SoupX removes ambient RNA contamination from droplet based single cell RNA sequencing data. BioRxiv.

Yourik, P., Fuchs, R.T., Mabuchi, M., Curcuru, J.L., and Robb, G.B. (2019). Staphylococcus aureus Cas9 is a multiple-turnover enzyme. RNA *25*, 35–44.

Zanini, F., Robinson, M.L., Croote, D., Sahoo, M.K., Sanz, A.M., Ortiz-Lasso, E., Albornoz, L.L., Rosso, F., Montoya, J.G., Goo, L., et al. (2018a). Virus-inclusive single-cell RNA sequencing reveals the molecular signature of progression to severe dengue. Proc. Natl. Acad. Sci. U. S. A. *115*, E12363–E12369.

Zanini, F., Pu, S.-Y., Bekerman, E., Einav, S., and Quake, S.R. (2018b). Single-cell transcriptional dynamics of flavivirus infection. Elife *7*.

Zhang, A.P.P., Bornholdt, Z.A., Liu, T., Abelson, D.M., Lee, D.E., Li, S., Woods, V.L., and Saphire, E.O. (2012). The Ebola Virus Interferon Antagonist VP24 Directly Binds STAT1 and Has a Novel, Pyramidal Fold. PLoS Pathogens *8*, e1002550.

Zhao, Y., Amodio, M., Vander Wyk, B., Gerritsen, B., Kumar, M.M., van Dijk, D., Moon, K., Wang, X., Malawista, A., Richards, M.M., et al. (2020). Single cell immune profiling of dengue virus patients reveals intact immune responses to Zika virus with enrichment of innate immune signatures. PLoS Negl. Trop. Dis. *14*, e0008112.

Zunder, E.R., Finck, R., Behbehani, G.K., Amir, E.-A.D., Krishnaswamy, S., Gonzalez, V.D., Lorang, C.G., Bjornson, Z., Spitzer, M.H., Bodenmiller, B., et al. (2015). Palladium-based mass tag cell barcoding with a doublet-filtering scheme and single-cell deconvolution algorithm. Nat. Protoc. *10*, 316–333.

Figures

Figure 1.

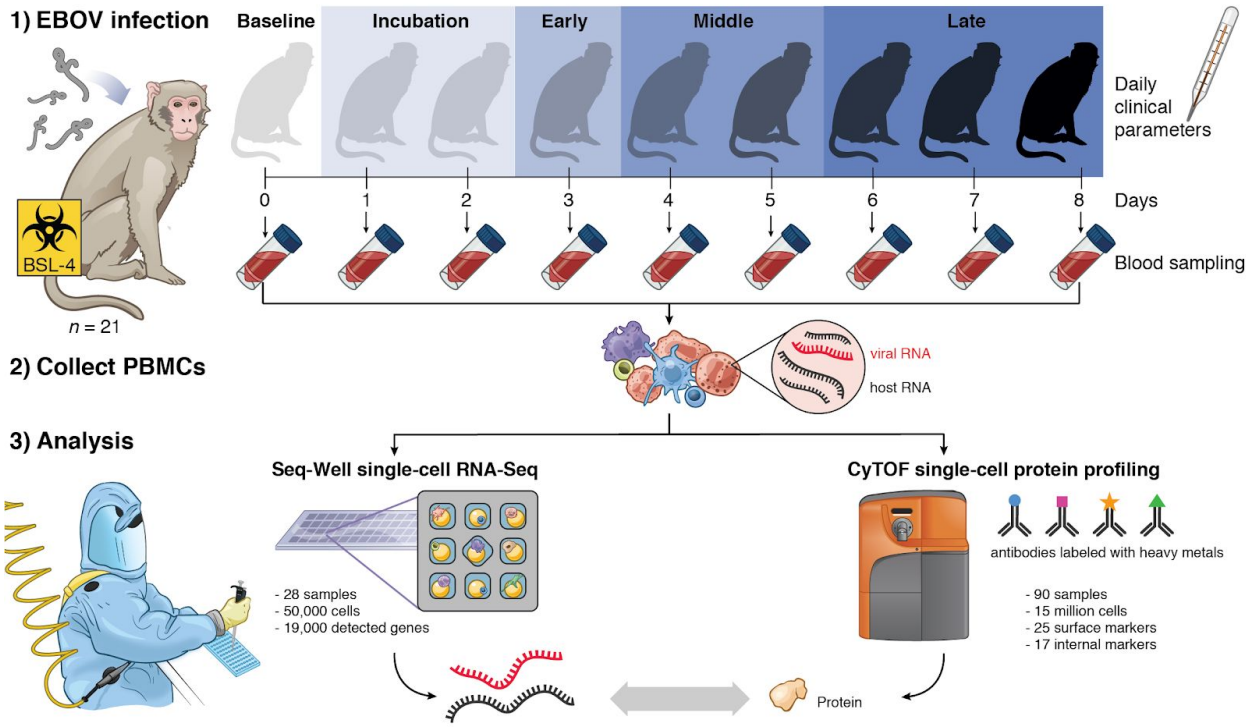


Figure 1. Studying lethal Ebola virus disease in rhesus monkeys using single-cell profiling of RNA and protein expression in peripheral immune cells

Serial sampling study design for investigating the host response to Ebola virus (EBOV) infection in rhesus monkeys. Under biosafety level 4 (BSL-4) containment, we collected blood samples from a total of 21 animals at multiple days post inoculation, extracted peripheral blood mononuclear cells (PBMCs), and profiled single-cell transcriptomes and 42 protein markers using Seq-Well and CyTOF. Seq-Well quantifies both host (black) and viral (red) RNA expression, allowing comparisons between infected and bystander cells. We also assessed daily clinical parameters for each animal, including body temperature, clinical signs, and body weight, and obtained complete blood counts for each blood draw.

See also **Figure S1** and **Table S1** for a complete listing of CyTOF and Seq-Well samples.

Figure 2.

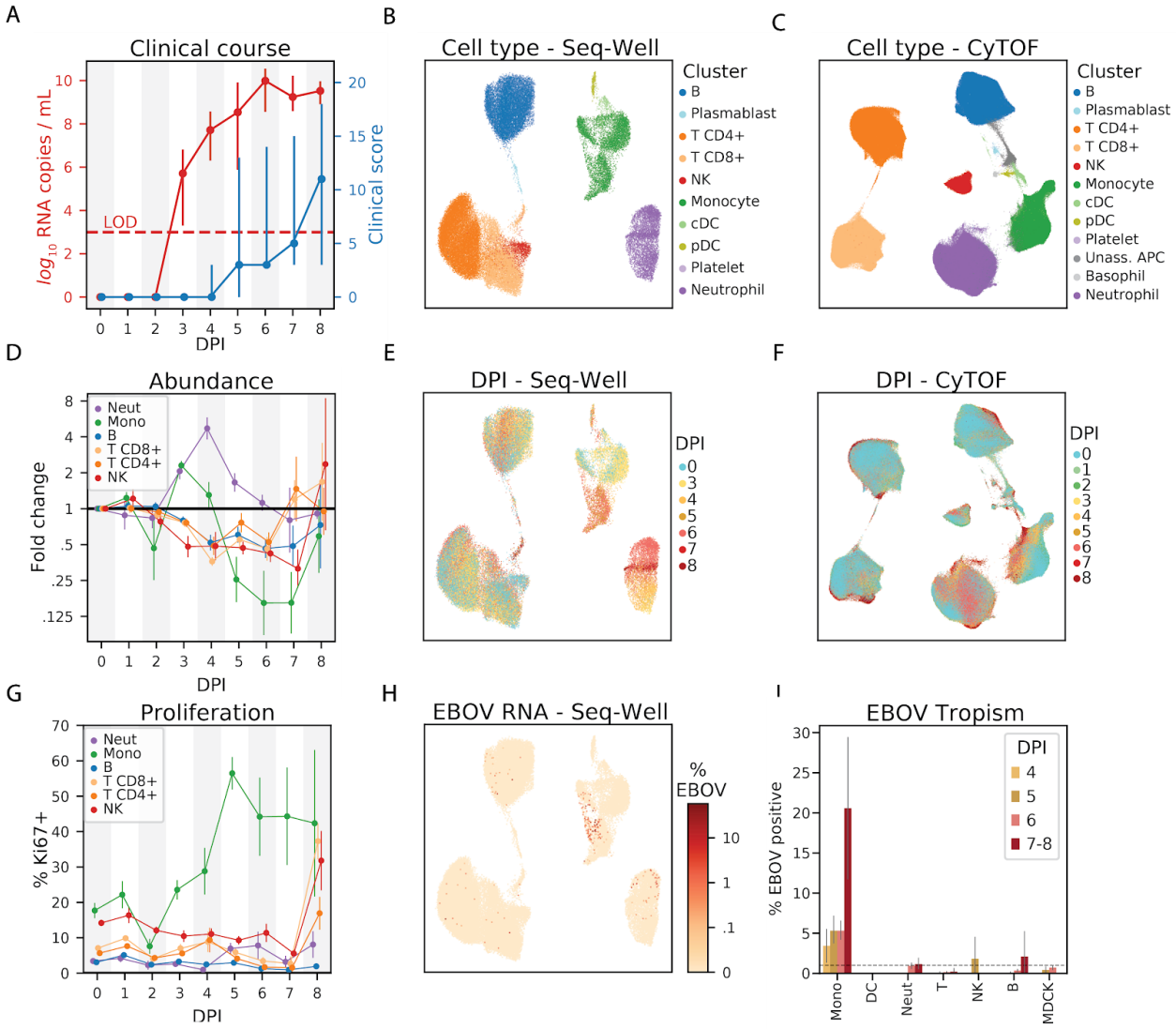


Figure 2. Single-cell profiling reveals changing cell-type abundance, proliferation rate, and infection status throughout EVD

(**A**) Time course of viral load (red, left *y*-axis, $\log_{10}$ scale) and clinical score (blue, right *y*-axis). Markers denote the mean and error bars denote the minimum and maximum values. LOD, limit of detection of viral load by RT-qPCR assay. See also **Figure S1**.

(**B and C**) Uniform Manifold Approximation and Projection (UMAP) embedding of Seq-Well (B) and CyTOF (C) data, colored by annotated cluster assignment. See also **Figures S2 and S4**.

(**D**) Fold change ($\log_2$ scale) in the absolute abundance (cells / μL of whole blood) of each cell type relative to baseline based on clustering of CyTOF data. Error bars denote the mean ± 1 standard error. See also **Figures S5A and S5B**.

(**E and F**) UMAP embedding of Seq-Well (E) and CyTOF (F) data, colored by the day post infection (DPI) on which each cell was sampled.

(**G**) Percentage of Ki67-positive cells (CyTOF intensity > 1.8, arbitrary units) of each cell type. Error bars denote the mean ± 1 standard error. See also **Figures S5C and S5D**.

(**H**) UMAP embedding of Seq-Well data, colored by the percentage of cellular transcripts mapping to EBOV.
(**I**) Percentage of infected cells by cell type based on Seq-Well expression profiles. Dashed line denotes the 1% false positive rate threshold used for calling infected cells. See also **Figure S2G**.

Figure 3.

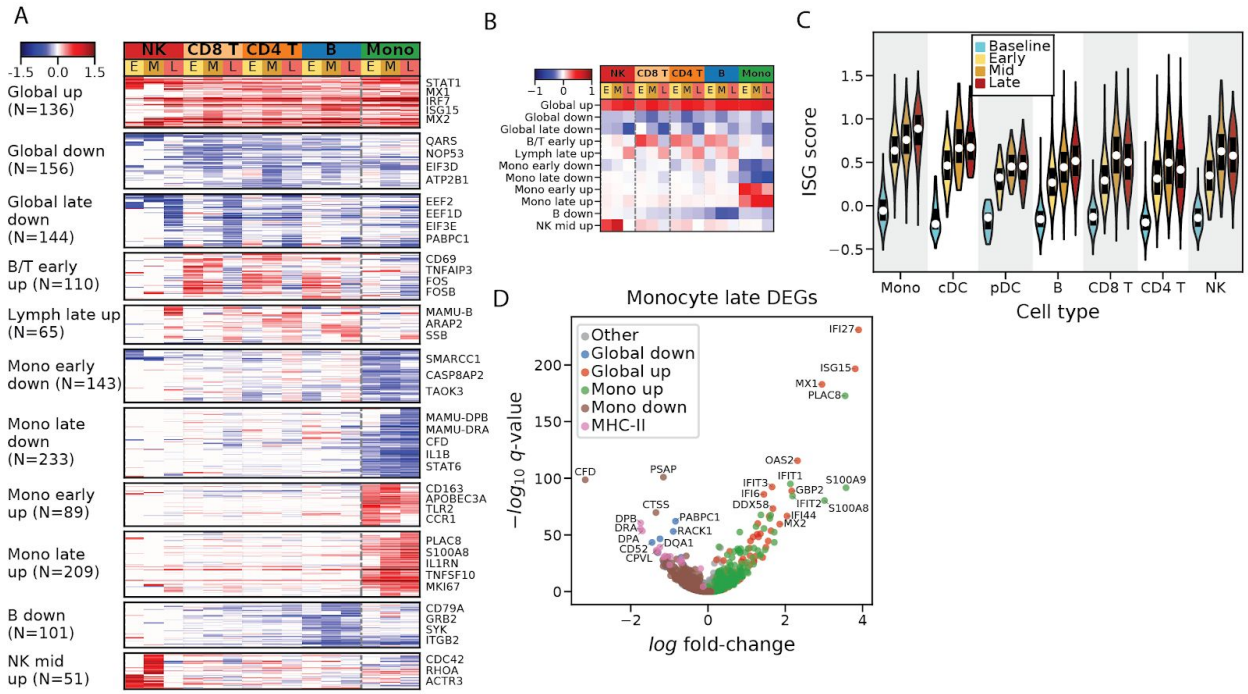


Figure 3. Patterns of differential expression across disease stages and cell types in EVD

(**A**) Fold changes (natural log scale) of 1,430 differentially expressed genes (rows) in each cell type at early (E), middle (M), and late (L) stages of disease, relative to baseline levels, with insignificant values ($p > 0.2$) set to 0. Genes were grouped into modules through unsupervised *k*-means clustering. See also **Tables S2 and S3**.

(B) Same as A but displaying the average log fold-change of each module.

(**C**) Distribution of interferon stimulated gene (ISG) scores for each cell type at baseline (blue), early (yellow), mid (orange), or late (red) EVD. White markers denote the median and bars denote the interquartile range. See also **Figures S6A and S6B**.

(**D**) Volcano plot of differential expression between monocytes in late EVD compared to baseline. Genes are colored by membership in the "Global down" or "Global late down" modules (blue), "Global up" module (red), “Mono early up” or “Mono late up” modules (green), “Mono early down” or “Mono late down” modules (brown), or if they are annotated as an MHC class II gene (purple).

Figure 4.

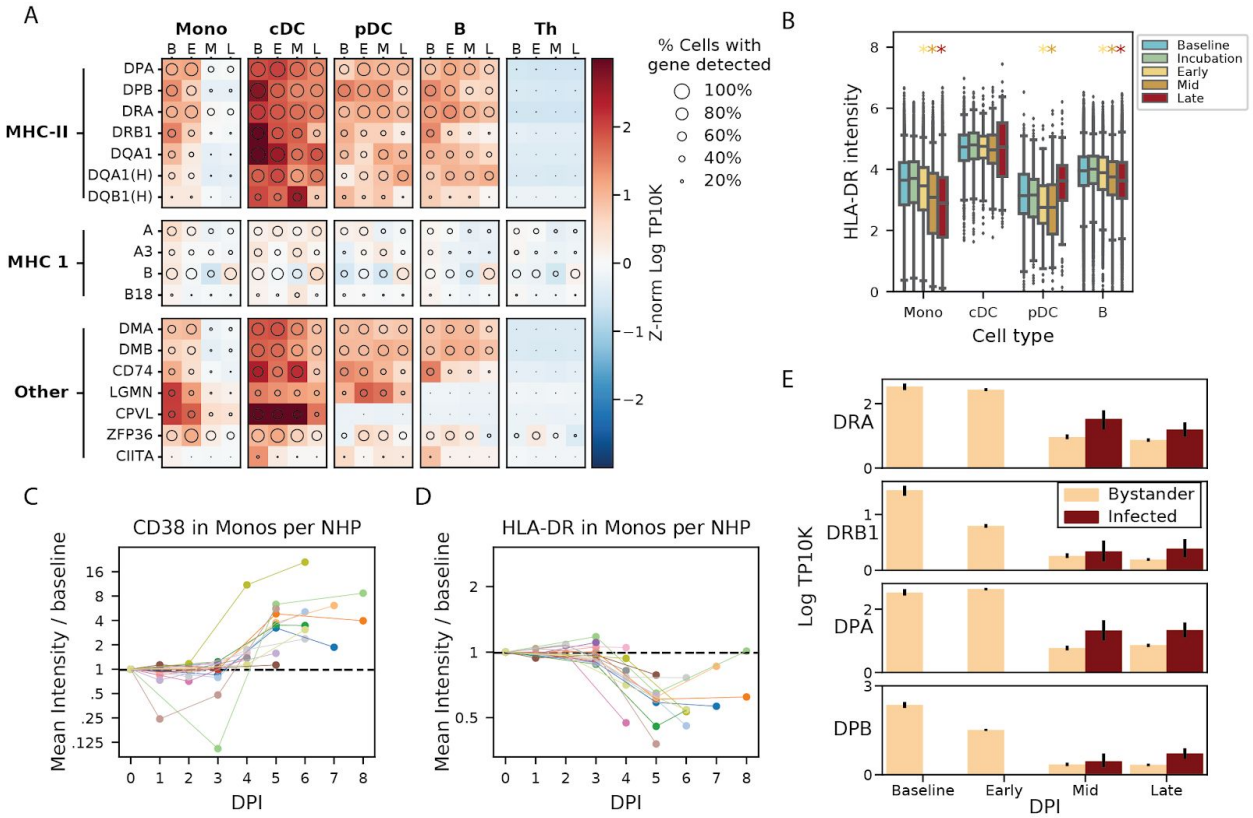


Figure 4. Monocytes dramatically reduce expression of MHC class II proteins independent of infection status

(**A**) Expression profiles of major histocompatibility (MHC) or MHC-associated genes (rows) in key cell types at baseline (B), early (E), middle (M), or late (L) stages of disease (columns). Circle size represents the percentage of cells in that group in which the gene was detected, and color denotes the average expression level, in Z-score normalized, natural log transcripts per 10,000 (TP10K). The “MAMU-” prefix, which designates major histocompatibility genes in rhesus monkeys, was removed from all MHC I and II gene symbols. Gene symbols that initially contained an “HLA-” prefix are indicated with an “(H)”.

(**B**) CyTOF intensity (arbitrary units) of HLA-DR protein in antigen-presenting cells at different stages of EVD. Boxes denote the median and interquartile range, and whiskers denote the 2.5th and 97.5th percentiles. Colored stars indicate EVD stages that decreased significantly from baseline (rank-sum test $p < 0.05$) with color corresponding to stage.

(C and D) Fold change ($\log_2$ scale) in average CD38 (C) and HLA-DR (D) CyTOF intensity on monocytes at each DPI relative to baseline for each PBMC sample. Colored lines connect serial samples from the same NHP. See also **Figure S6C**.

(**E**) Average gene expression (natural log TP10K) for four MHC class II genes in monocytes at different disease stages, stratified by cell infection status. Error bars denote 95% confidence intervals for the mean based on 200 bootstraps.

Figure 5.

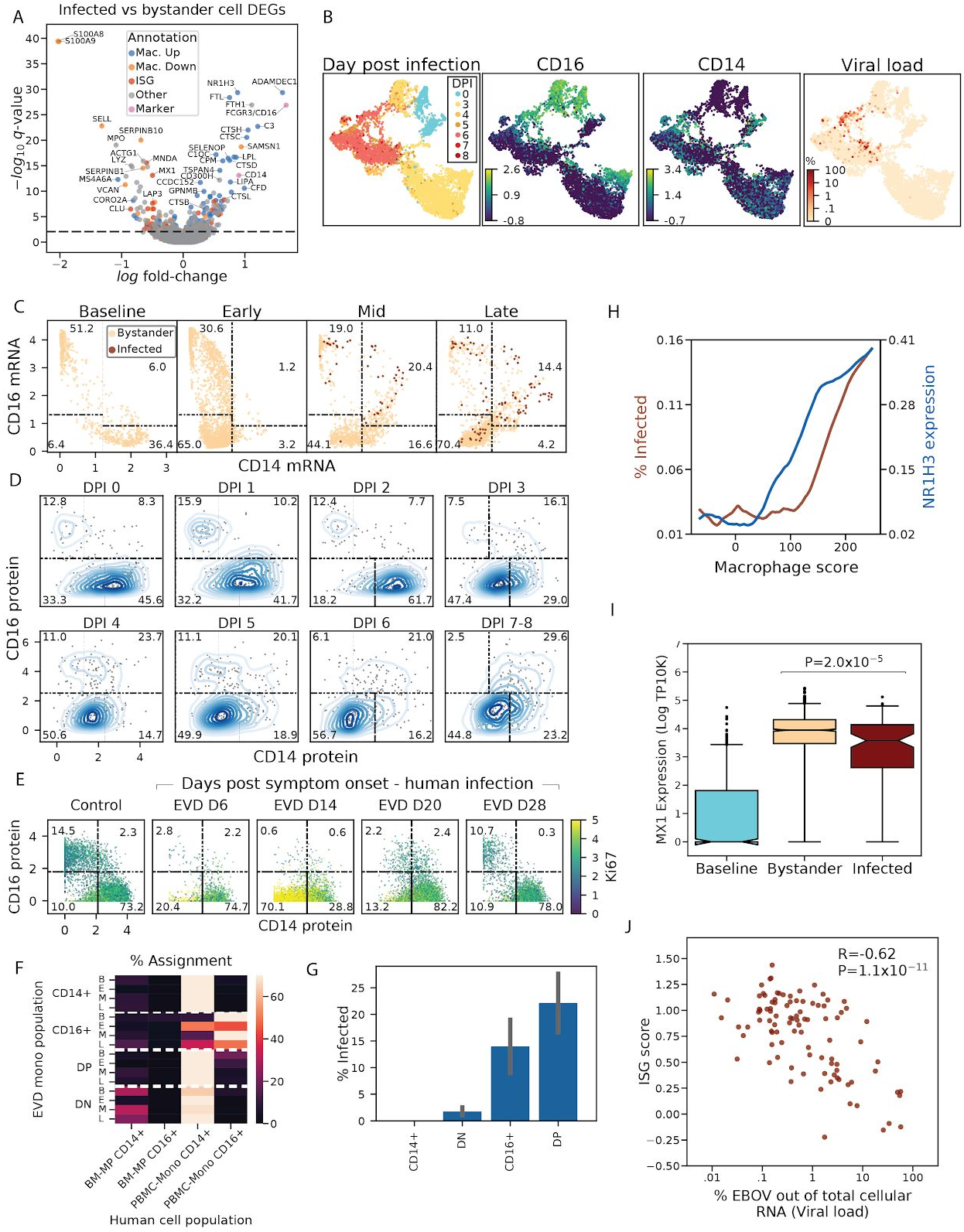

A
Infected vs bystander cell DEGs
Annotation
Mac. Up
Mac. Down
ISG
Other
Marker
−log10 q-value
log fold-change
B
Day post infection
CD16
CD14
Viral load
C
Baseline
Early
Mid
Late
Bystander
Infected
CD16 mRNA
CD14 mRNA
D
DPI 0
DPI 1
DPI 2
DPI 3
DPI 4
DPI 5
DPI 6
DPI 7-8
CD16 protein
CD14 protein
E
Days post symptom onset - human infection
Control
EVD D6
EVD D14
EVD D20
EVD D28
Ki67
F
% Assignment
EVD mono population
CD14+
CD16+
DP
DN
BM-MP CD14+
BM-MP CD16+
PBMC-Mono CD14+
PBMC-Mono CD16+
Human cell population
G
% Infected
H
% Infected
NR1H3 expression
Macrophage score
I
MX1 Expression (Log TP10K)
P=2.0x10−5
Baseline
Bystander
Infected
J
R=-0.62
P=1.1x10−11
ISG score
% EBOV out of total cellular RNA (Viral load)

Figure 5. ISG suppression, co-expression of CD14 and CD16, and expression of macrophage genes, are associated with monocyte infectivity

(**A**) Volcano plot of differential expression between infected and bystander monocytes from DPI 5–8. Genes are colored by membership in sets of genes up- (Mac. Up, blue) or down-regulated (Mac. Down, orange) during *in vitro* differentiation of monocytes into macrophages, interferon stimulated genes (ISG, red), or the marker genes *CD14* and *CD16* (Marker, pink). See also **Table S4**.

(**B**) UMAP embedding of monocyte gene expression data, colored by (left-to-right): Day post infection, *CD16* expression (log TP10K), *CD14* expression (log TP10K), and percentage of cellular transcripts mapping to EBOV.

(**C**) Scatter plot of smoothed expression (log TP10K) of *CD14* and *CD16* for monocytes in baseline, early, mid, and late disease. Cells are colored by infection status. Boxes define the CD14+, CD16+, DN, and DP subsets described in the text, and numbers denote the percentage of cells in each subset at that disease stage. See also **Figures S7A and S7B**.

(D) CD14 and CD16 protein (CyTOF intensity) expression on monocytes at each DPI. The distribution is displayed as a bivariate kernel density plot with 200 randomly sampled cells overlaid as a scatter plot. See also **Figure S7C**.

(**E**) CD14 and CD16 protein (CyTOF intensity) expression on monocytes in a case of human EVD, colored by Ki67 protein expression (CyTOF intensity) for multiple days post symptom onset. See also **Figure S7D**.

(**F**) Percent assignment of NHP CD14/CD16 subsets at each EVD stage (B - baseline, E - early, M - mid, L - Late) to human myeloid populations in the reference dataset (BM-MP - bone marrow monocyte progenitors, PBMC-CD16+ - circulating CD16+ monocytes, PBMC-CD14+ - circulating CD14+ monocytes). See also **Figures S7E–S7K**.

(G) Percentage of infected monocytes in each CD14/CD16 subset in late infection. Error bars denote 95% confidence intervals for the mean based on 1,000 bootstraps.

(H) Association between macrophage score (*x*-axis) and percentage of infected cells (left *y*-axis, red), and expression of *NR1H3* (natural log TP10K), a marker of macrophage differentiation (right *y*-axis, blue). Monocytes from late disease are ordered by macrophage score, and we averaged percent of infected cells and *NR1H3* expression within a 400-cell sliding window. See also **Figures S8A–S8C**.

(**I**) *MX1* expression (natural log TP10K) in monocytes at baseline, and uninfected bystanders or infected cells in late infection. Boxes denote the median and interquartile range, and whiskers denote the 2.5th and 97.5th percentiles. See also **Figure S8D**.

(J) Scatter plot of ISG score (*y*-axis) versus percentage of cellular transcripts mapping to EBOV (*x*-axis) for infected monocytes in late EVD (DPI 6–8).

Figure 6.

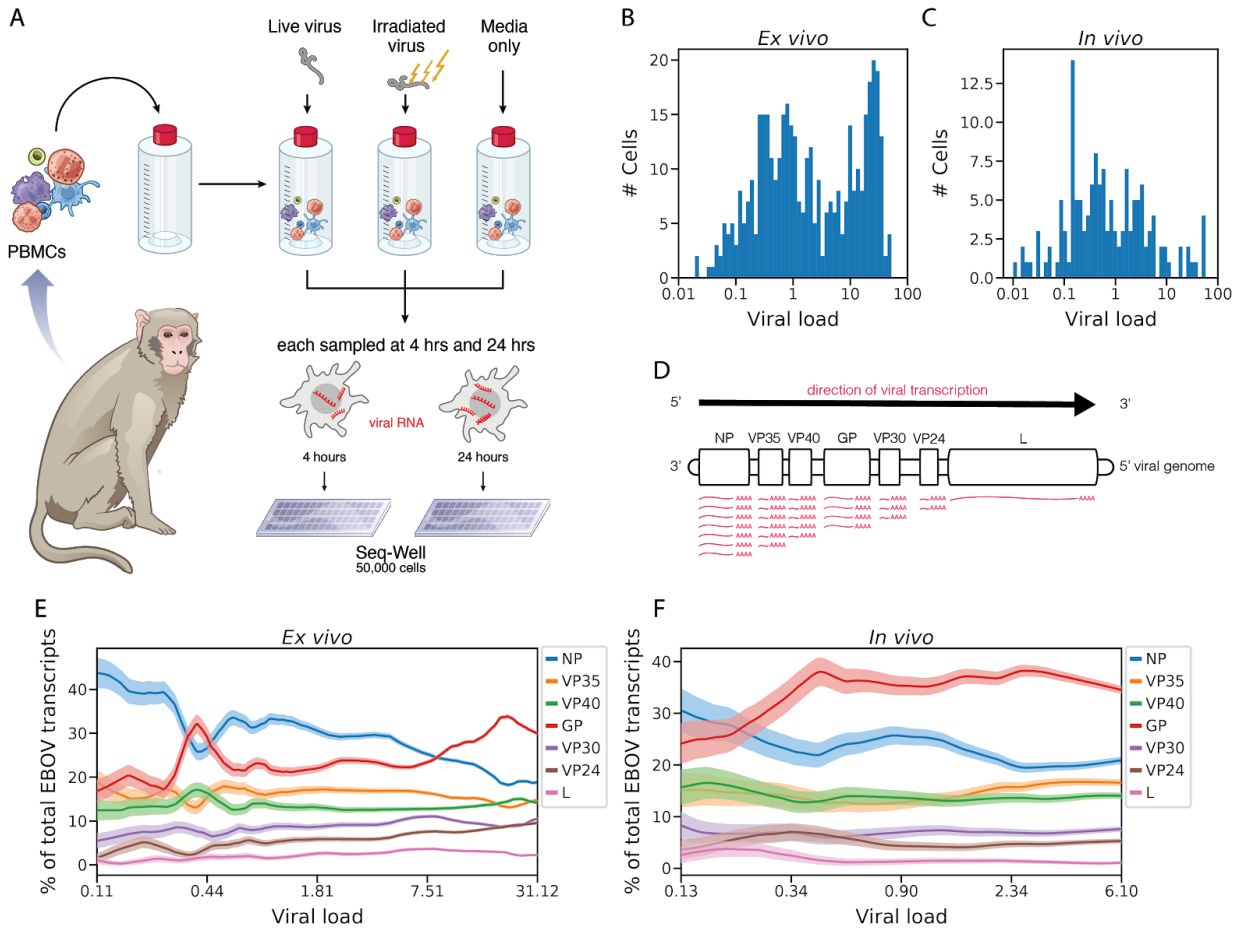


Figure 6. Viral transcriptional dynamics of infected monocytes *in vivo* and *ex vivo*

(**A**) Schematic of EBOV challenge of PBMCs *ex vivo*. See also **Figure S9**.

(**B and C**) Histogram of the percentage of cellular transcripts derived from EBOV (intracellular viral load) in monocytes from PBMCs inoculated with live virus *ex vivo* (B) or from PBMCs of NHPs infected *in vivo* (C). See also **Figures S10A–S10D**.

(**D**) Schematic of EBOV transcription. The viral RNA-directed RNA-polymerase transcribes each gene sequentially from the 19 kb genome, but occasionally releases the genomic RNA template thereby ending transcription. As a result, *NP* is transcribed most frequently, and *L* least frequently.

(**E and F**) Relative proportion of each EBOV gene versus viral load ($\log_{10}$ scale), *ex vivo* (E) or *in vivo* (F). All infected monocytes were ordered by viral load and the percentage of each viral gene was averaged over a 50-cell sliding window. Bands denote the mean ± 1 SD. See also **Figures S10E and S10F**.

Figure 7.

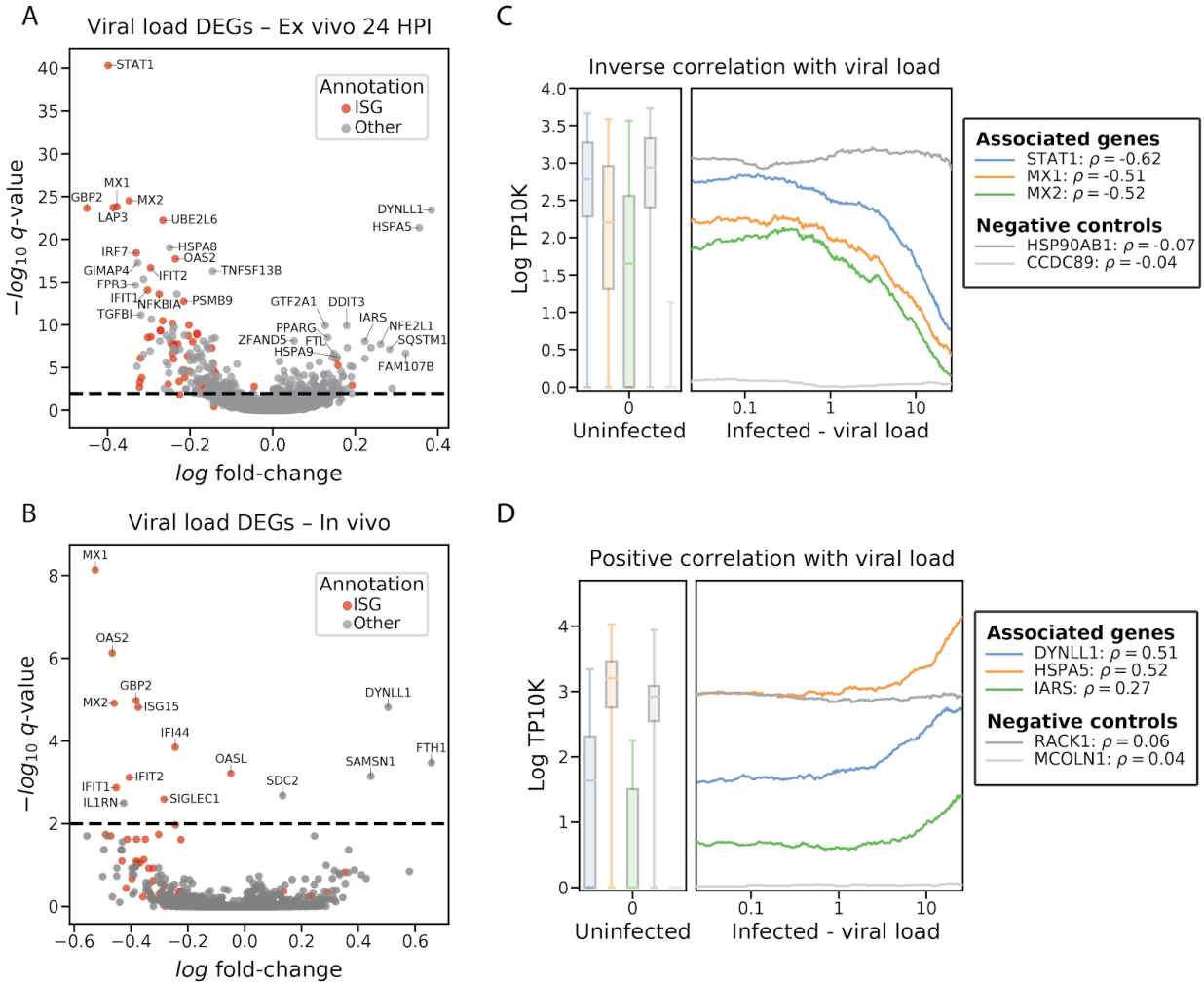


Figure 7. EBOV infection down-regulates host antiviral genes and up-regulates putative pro-viral genes

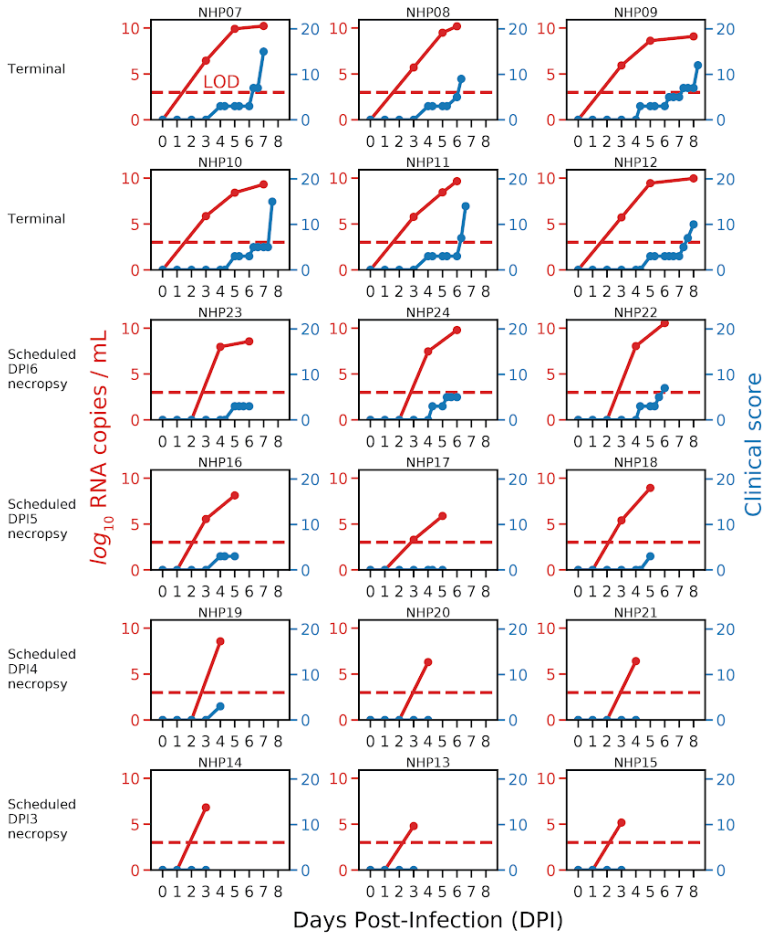
(**A and B**) Volcano plot of association between host genes expression and viral load, within infected monocytes from PBMCs 24 HPI treated with live virus *ex vivo* (A) or from PBMCs of NHPs infected *in vivo* on DPI 5-8 (B). See also **Table S6**.

(**C and D**) Association between host gene expression and viral load for selected positively (C) and negatively (D) associated genes in monocytes from *ex vivo* infections. Infected cells are ordered by viral load ($\log_{10}$ scale), and we depict gene expression (natural log TP10K) averaged over a 100-cell sliding window. Genes that are significantly associated with viral load are shown in color while unassociated negative control genes are shown in gray. Spearman correlation coefficients (ρ) for the association between viral load and gene expression are listed in the legend. Box plots represent gene expression in uninfected cells (boxes denote the median and interquartile range, and whiskers denote the 2.5th and 97.5th percentiles) See also **Figures S10G and S10H**.

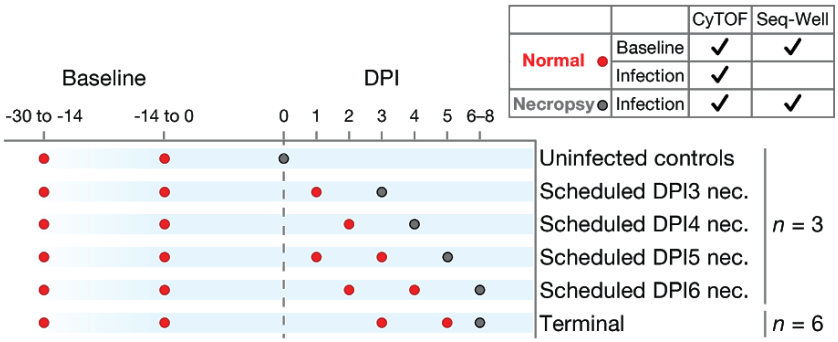
Supplementary Figures

Figure S1.

A



B



BL	1	2	3	4	5	6	7	8	DPI
42	6	6	12	6	9	5	2	2	CyTOF samples
13	0	0	3	1	3	5	1	2	Seq-Well samples

Figure S1. Blood sampling overview and clinical time course per animal replicate. Related to Figure 1.

(**A**) Each panel represents the time course of log viral load (left axis, red) and clinical score (right axis, blue) for a specific NHP. Panels are organized into rows based on the cohorts.
(**B**) Overview of study cohorts and blood draw timelines. Animals were grouped into cohorts with pre-scheduled necropsy times (at baseline, or day post infection [DPI] 3, 4, 5, 6 - *n* = 3 each), or allowed to progress until clinical score exceeded 10 (terminal), predetermined euthanasia criteria. Dots indicate scheduled blood draws for each cohort with red denoting an intermediate (non-necropsy) draw, and gray indicating a draw that coincided with euthanasia and necropsy. Necropsy and baseline normal draws were used for Seq-Well and CyTOF, while intermediate post-infection draws were available only for CyTOF.

Figure S2.

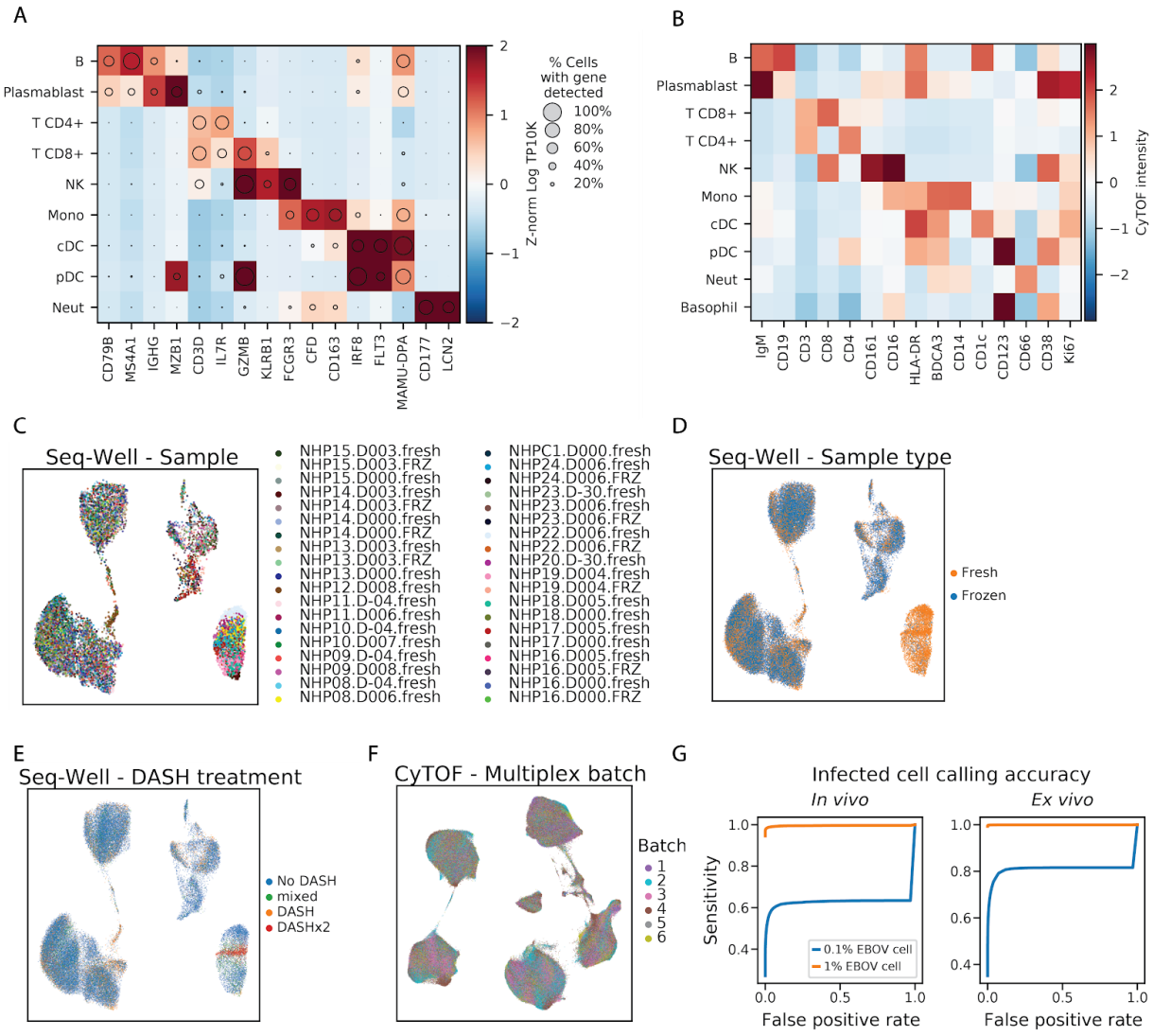


Figure S2. Cell type markers for Seq-Well and CyTOF clusters. Related to Figure 2.

(**A**) Expression profiles of cell-type marker genes (columns) for cell-type clusters (rows) based on the *in vivo* Seq-Well data. Circle area represents the percentage of cells in each group in which the gene was detected, and color denotes the average expression level (natural log TP10K).

(**B**) Average expression (Z-normalized CyTOF intensity) profiles of cell type marker genes (columns), for cell type clusters (rows), based on the CyTOF data.

(**C**) Uniform Manifold Approximation and Projection (UMAP) embedding of post-integration Seq-Well data, colored by the sample source (NHP, DPI, and whether the sample was loaded for Seq-Well without any freezing [.fresh] or was frozen with cryoprotectant and thawed prior to

Seq-Well [.FRZ]). A maximum of 500 cells per sample is plotted to increase representation across samples.
(**D**) UMAP embedding of Seq-Well data, colored by whether cells were processed fresh (orange) or after freeze/thaw (blue) prior to Seq-Well.
(**E**) UMAP embedding of Seq-Well data, colored by depletion of abundant sequences by hybridization (DASH) treatment. We developed a DASH-based method to remove a PCR adaptor artifact from some Seq-Well sequencing libraries (see Materials and Methods), and performed this 0 times (No DASH, blue), 1 time (DASH, orange), or 2 times sequentially (DASHx2, red). For a few samples, we sequenced 'No DASH' and 'DASH' libraries and merged the reads (mixed, green).
(**F**) UMAP embedding of batch-corrected CyTOF data, colored by the multiplex batch in which it was pooled, and analyzed by CyTOF.
(**G**) Receiver operating characteristic curves for identifying EBOV-infected cells. Estimates of sensitivity to detect an infected cell at various false positive rate thresholds *in vivo* (left) and *ex vivo* (right). Curves are estimated separately for a hypothetical viral load of 0.1% (blue line) and 1% (orange line).

Figure S3.

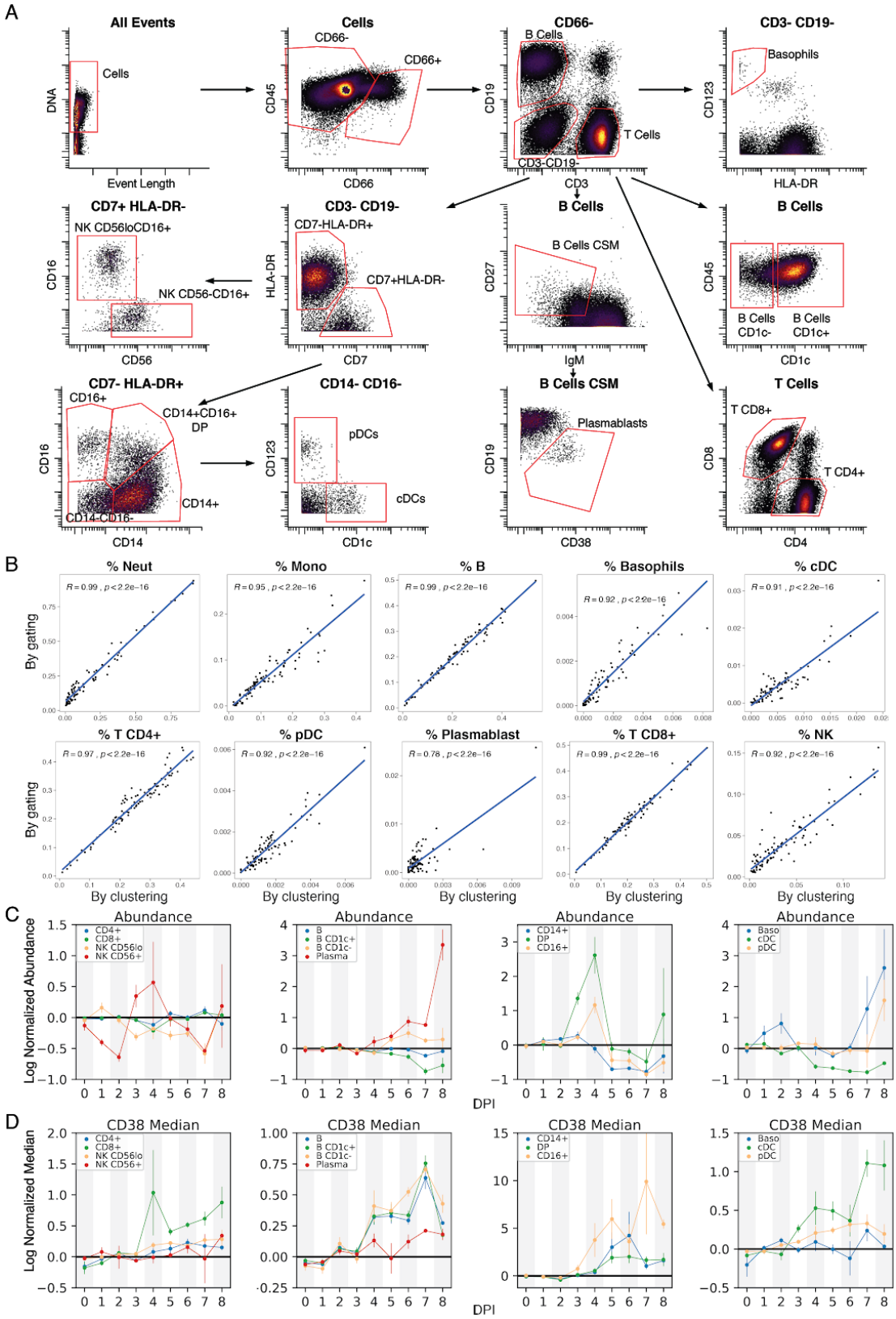

A
All Events
Cells
DNA
Event Length
Cells
CD66-
CD66+
CD45
CD66
CD66-
B Cells
T Cells
CD3-CD19-
CD19
CD3
CD3- CD19-
Basophils
CD123
HLA-DR
CD7+ HLA-DR-
NK CD56loCD16+
NK CD56-CD16+
CD16
CD56
CD3- CD19-
CD7-HLA-DR+
CD7+HLA-DR-
HLA-DR
CD7
B Cells
B Cells CSM
CD27
IgM
B Cells
B Cells CD1c-
B Cells CD1c+
CD45
CD1c
CD7- HLA-DR+
CD16+
CD14+CD16+ DP
CD14+
CD14-CD16-
CD16
CD14
CD14- CD16-
pDCs
cDCs
CD123
CD1c
B Cells CSM
Plasmablasts
CD19
CD38
T Cells
T CD8+
T CD4+
CD8
CD4
B
% Neut
% Mono
% B
% Basophils
% cDC
% T CD4+
% pDC
% Plasmablast
% T CD8+
% NK
By gating
By clustering
C
Abundance
Log Normalized Abundance
CD4+
CD8+
NK CD56lo
NK CD56+
B
B CD1c+
B CD1c-
Plasma
CD14+
DP
CD16+
Baso
cDC
pDC
DPI
D
CD38 Median
Log Normalized Median

Figure S3. Comparison of unsupervised clustering and manual gating of CyTOF data. Related to Figure 2.

(**A**) Gating strategy used to define cell populations using canonical markers. Axes indicate CyTOF ArcSinh-scaled marker intensities.

(**B**) Pearson correlation between relative cell-type abundance determined by unsupervised clustering (*x*-axis) and by manual gating (*y*-axis). Each marker represents a single PBMC sample.

(**C**) Abundances of indicated manually gated populations, normalized for each animal to pre-challenge timepoints. *x*-axis shows day post infection (DPI), lines denote mean ± 1 standard error of the mean (SEM).

(**D**) Median CD38 expression on manually gated populations shown in (C), medians are normalized to each NHP's average baseline expression. *x*-axis shows DPI, lines denote mean ± 1 SEM.

Figure S4.

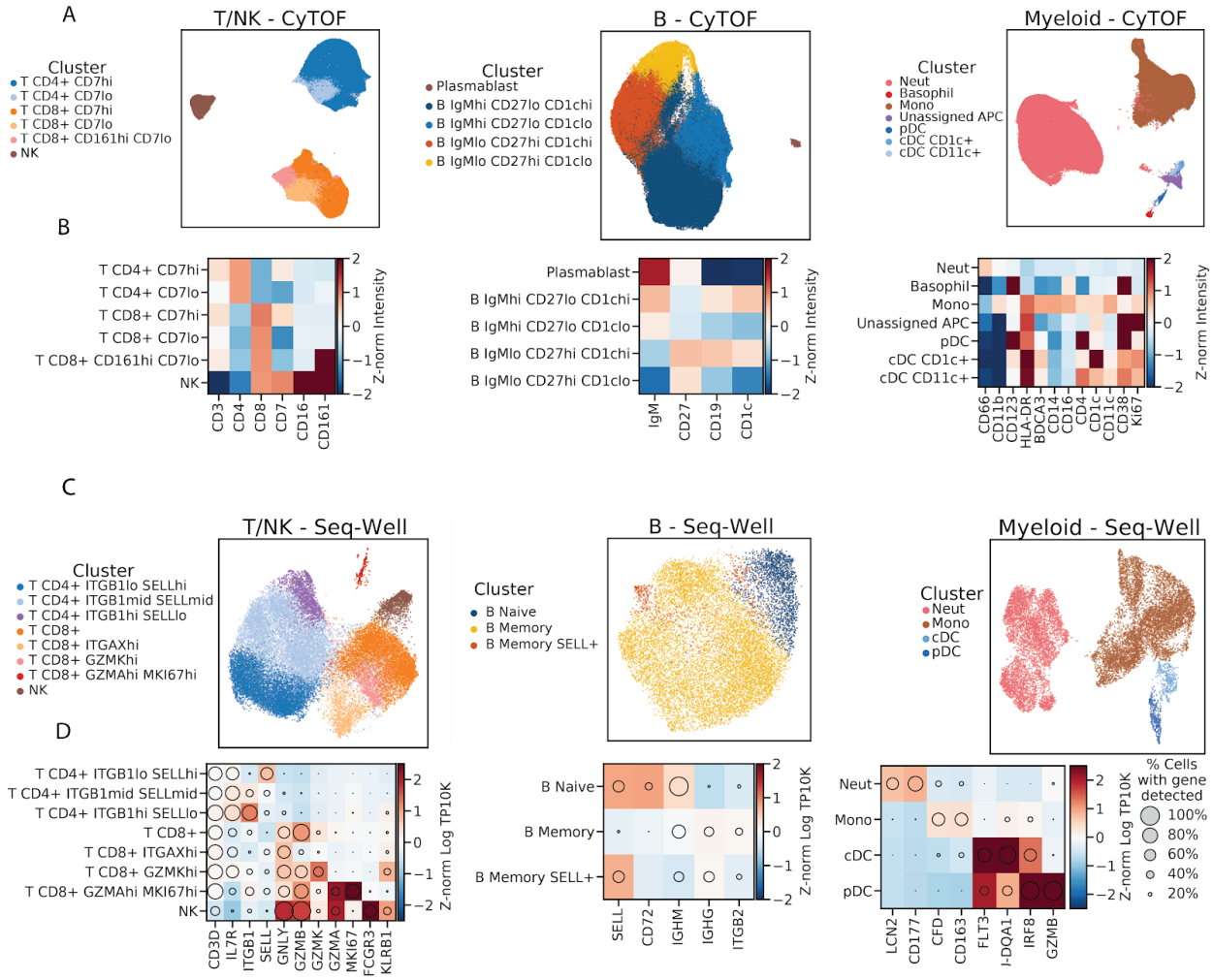


Figure S4. Identifying cell subtypes by subclustering. Related to Figure 2.

(A) UMAP embedding of broad cell-type clusters in the CyTOF data, colored by sub-cluster assignment (Neut - neutrophil, Mono - monocyte).

(B) Average expression (Z-normalized CyTOF intensity) profiles of sub-clusters for marker genes based on CyTOF data.

(C) UMAP embedding of broad cell-type clusters in the Seq-Well data, colored by sub-cluster assignment.

(D) Expression profiles of sub-clusters for marker genes based on Seq-Well data. Circle area represents the percentage of cells in which the gene was detected, and color denotes the average expression level (Z-normalized natural log TP10K).

Figure S5.

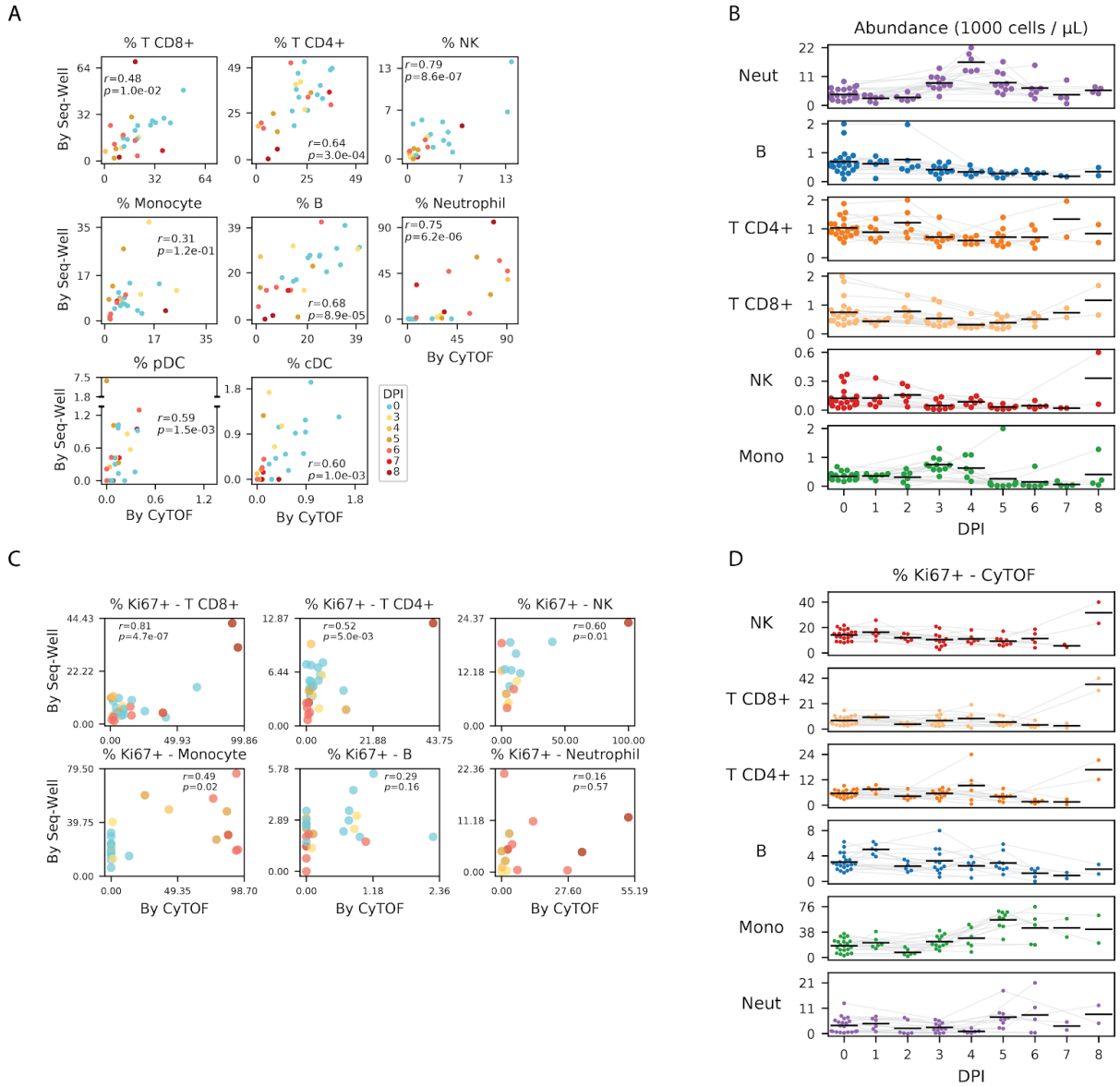


Figure S5. Estimates of cell-type abundance and proliferation over the time course. Related to Figure 2.

(**A**) Scatter plot of the percentage of cells of each cell type in a sample, inferred from CyTOF (*x*-axis) or Seq-Well (*y*-axis), for several cell types (panels). Each dot represents a sample colored by DPI. Pearson correlation coefficients (*r*) and *p*-value are provided.

(**B**) Estimates of the abundance of each cell type (rows) for each NHP (individual markers) in units of 1000 cells per μL of whole blood, based on integration of CyTOF and complete blood count (CBC) information. The mean value of each DPI is indicated with a black line. Gray lines connect serial samples from the same NHP.

(**C**) Scatter plots of the percentage of Ki67-positive cells in a sample inferred from Seq-Well (*x*-axis) or CyTOF (*y*-axis) for several cell types (panels). Each dot represents a sample colored

by DPI. Cells with smoothed expression of *MKI67* (the gene coding for Ki67) above 0.1 are called Ki67 positive by Seq-Well. Cells with CyTOF intensity above 1.8 (arbitrary units) are called Ki67 positive by CyTOF.
(**D**) Estimates of the percentage of Ki67 positive cells (CyTOF intensity > 1.8) of each cell type (rows) for each animal replicate (markers). The mean value of each DPI is indicated with a black line. Gray lines connect serial samples from the same NHP.

Figure S6.

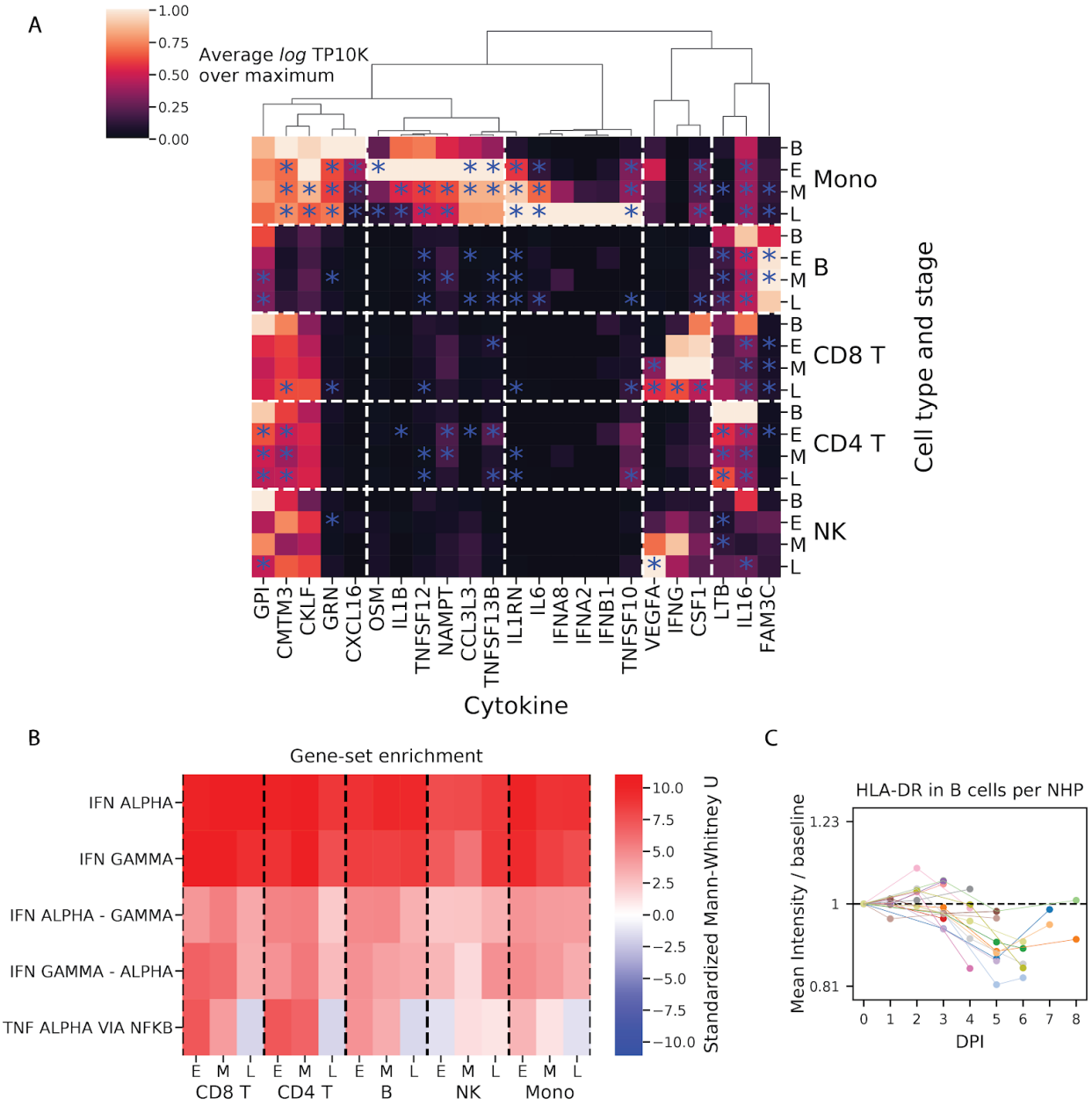


Figure S6. Quantification of cytokine expression and enrichment of response signatures. Related to Figures 3 and 4.

(**A**) Average expression values (natural log TP10K) of literature-annotated cytokines (columns) across cell types and stages of acute EVD (rows). Values are plotted as a ratio relative to the maximum across cell types and stages. Values that are statistically different from baseline ($p < .05$) are indicated with a blue star.

(**B**) Heatmap of rank-sum test statistics for comparison of differential expression log-fold changes of genes in a gene set (rows) compared to genes not in the gene set. The log fold-changes were defined from differential expression profiles of each cell type at each EVD stage (columns) relative to baseline. Five gene sets were tested -- three from the Hallmark database (IFN ALPHA, IFN GAMMA, and TNF ALPHA VIA NFKB) (Liberzon et al., 2015) and 2 constructed from the hallmark sets, as uniquely IFNα-regulated genes in “IFN ALPHA” but not “IFN GAMMA” (“IFN ALPHA - GAMMA”), and vice versa for uniquely IFNγ-regulated (“IFN GAMMA - ALPHA”). See also **Table S3**.
(**C**) Fold change ($\log_2$ scale) in average HLA-DR CyTOF intensity on B cells at each DPI relative to baseline for each PBMC sample. Colored lines connect serial samples from the same NHP.

Figure S7.

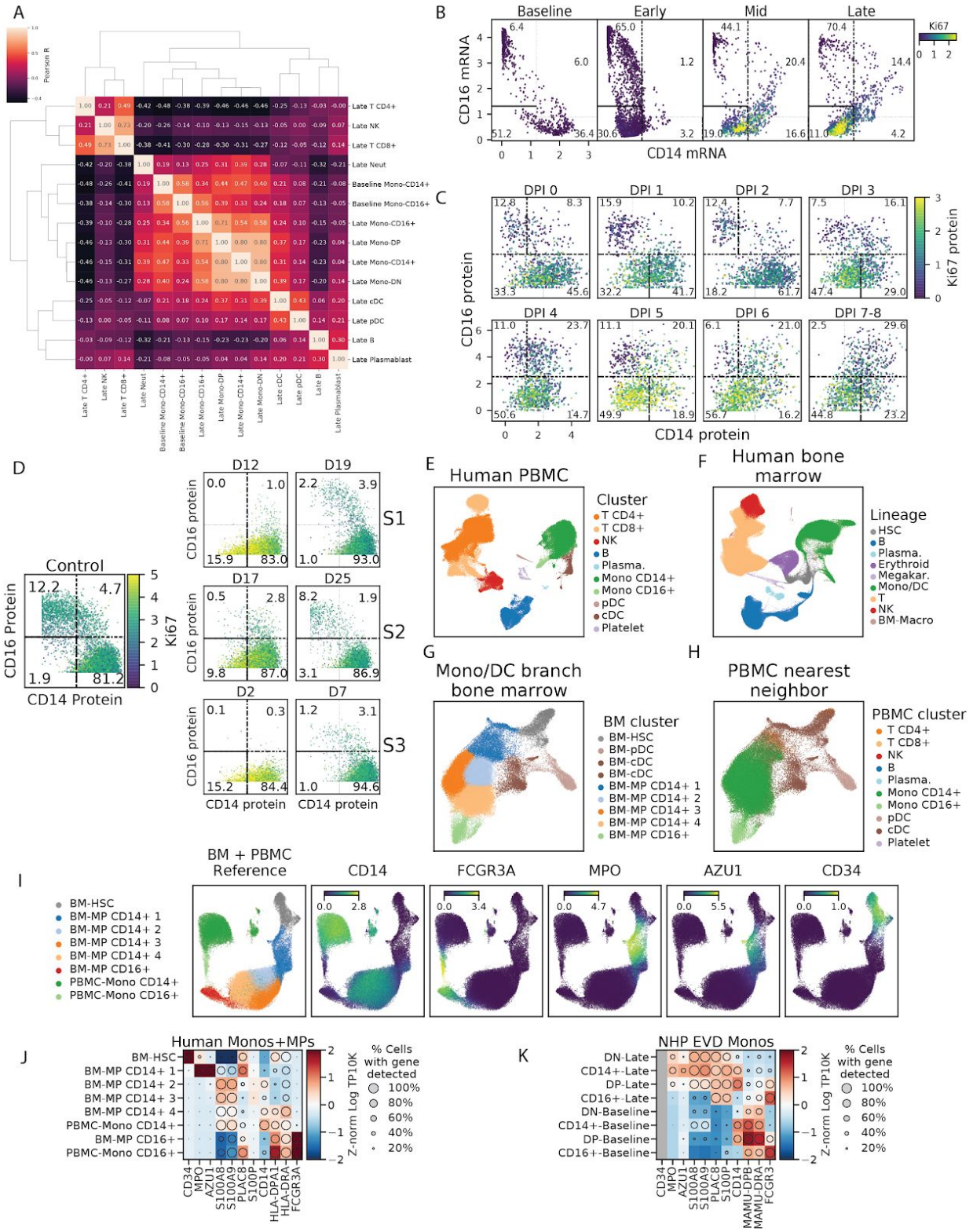

A
Pearson R
Late T CD4+
Late NK
Late T CD8+
Late Neut
Baseline Mono-CD14+
Baseline Mono-CD16+
Late Mono-CD16+
Late Mono-DP
Late Mono-CD14+
Late Mono-DN
Late cDC
Late pDC
Late B
Late Plasmablast
B
Baseline
Early
Mid
Late
Ki67
CD16 mRNA
CD14 mRNA
C
DPI 0
DPI 1
DPI 2
DPI 3
DPI 4
DPI 5
DPI 6
DPI 7-8
CD16 protein
CD14 protein
Ki67 protein
D
Control
CD16 Protein
CD14 Protein
Ki67
D12
D19
D17
D25
D2
D7
S1
S2
S3
CD16 protein
CD14 protein
E
Human PBMC
Cluster
T CD4+
T CD8+
NK
B
Plasma.
Mono CD14+
Mono CD16+
pDC
cDC
Platelet
F
Human bone marrow
Lineage
HSC
B
Plasma.
Erythroid
Megakar.
Mono/DC
T
NK
BM-Macro
G
Mono/DC branch bone marrow
BM cluster
BM-HSC
BM-pDC
BM-cDC
BM-cDC
BM-MP CD14+ 1
BM-MP CD14+ 2
BM-MP CD14+ 3
BM-MP CD14+ 4
BM-MP CD16+
H
PBMC nearest neighbor
PBMC cluster
T CD4+
T CD8+
NK
B
Plasma.
Mono CD14+
Mono CD16+
pDC
cDC
Platelet
I
BM + PBMC Reference
CD14
FCGR3A
MPO
AZU1
CD34
BM-HSC
BM-MP CD14+ 1
BM-MP CD14+ 2
BM-MP CD14+ 3
BM-MP CD14+ 4
BM-MP CD16+
PBMC-Mono CD14+
PBMC-Mono CD16+
J
Human Monos+MPs
BM-HSC
BM-MP CD14+ 1
BM-MP CD14+ 2
BM-MP CD14+ 3
BM-MP CD14+ 4
PBMC-Mono CD14+
BM-MP CD16+
PBMC-Mono CD16+
CD34
MPO
AZU1
S100A8
S100A9
PLAC8
S100P
CD14
HLA-DPA1
HLA-DRA
FCGR3A
Z-norm Log TP10K
% Cells with gene detected
100%
80%
60%
40%
20%
K
NHP EVD Monos
DN-Late
CD14+-Late
DP-Late
CD16+-Late
DN-Baseline
CD14+-Baseline
DP-Baseline
CD16+-Baseline
CD34
MPO
AZU1
S100A8
S100A9
PLAC8
S100P
CD14
MAMU-DPB
MAMU-DRA
FCGR3
Z-norm Log TP10K
% Cells with gene detected
100%
80%
60%
40%
20%

Figure S7. Extended characterization of interferon and double-negative (DN) CD14- CD16- monocytes. Related to Figure 5.

(**A**) Clustermap of pairwise Pearson correlations between cell type clusters at baseline and late EVD. Correlations are computed on average log TP10K expression values of overdispersed genes. DN and DP monocytes at late EVD are more similar to monocytes (including baseline CD14+s) than other cell types.

(**B**) Scatter plot of MAGIC-smoothed expression values (log TP10K) of *CD14* and *CD16* for monocytes in baseline, early, mid, and late disease stages. Cells are colored by smoothed expression levels of *MKI67* (the gene coding for Ki67 protein). Boxes indicate the CD14+, CD16+, DN, and DP subsets described in the text, and the numbers denote the percentage of cells falling into each subset.

(**C**) Scatter plot of protein expression (CyTOF intensity) of CD14 and CD16 for 1000 randomly sampled monocytes at each DPI. Cells are colored by Ki67 expression. Boxes indicate the CD14+, CD16+, DN, and DP subsets described in the text, and the numbers denote the percentage of cells falling into each subset.

(**D**) Scatter plot of protein expression (CyTOF intensity) of CD14 and CD16 for monocytes during human EVD. Left: monocytes from healthy human controls. Right: monocytes from 3 EVD cases (S1, S2, and S3) at various days post symptom onset. Cells are colored by Ki67 marker intensity. Boxes indicate the CD14+, CD16+, DN, and DP subsets described in the text and the numbers denote the percentage of cells falling into each subset.

(**E**) UMAP embedding of healthy human PBMCs dataset, colored by annotated cluster assignment, based on known marker genes. (Plasma. - Plasmablast).

(**F**) UMAP embedding of healthy bone marrow cells, colored by cluster assignment, based on marker genes. (HSC - hematopoietic stem cell, Plasma. - Plasmablast, Megakar. - Megakaryocyte, Mono/DC - monocyte and dendritic cell, BM-Macro - bone marrow resident macrophage).

(**G**) UMAP embedding of sub-clustered HSC and monocyte/dendritic lineage cells. (BM - bone marrow, MP - monocyte progenitor)

(**H**) Same UMAP embedding as **Figure S7G**, but cells are colored by the cluster identity of their nearest neighbor in the human PBMC dataset (**Figure S7E**).

(**I**) UMAP embedding of the merged reference dataset of healthy bone marrow HSCs and monocyte lineage cells and PBMCs. Left sub-panel is colored by cluster assignment. Right sub-panels are colored by marker gene expression (natural log TP10K).

(**J**) Expression profiles of selected genes for human bone marrow monocyte progenitors (BM-MPs) and human circulating monocytes (PBMC-Monos). Circle area represents the percentage of cells in which the gene was detected, and color denotes the average expression (Z-normalized natural log TP10K).

(**K**) Expression profiles of selected genes for NHP monocyte subsets at baseline or late EVD for orthologs of the genes in (J). Circle area represents the percentage of cells in which the gene was detected, and color denotes the average expression level (Z-normalized natural log TP10K). CD34 is grayed out because it is detected in fewer than 10 cells.

Figure S8.

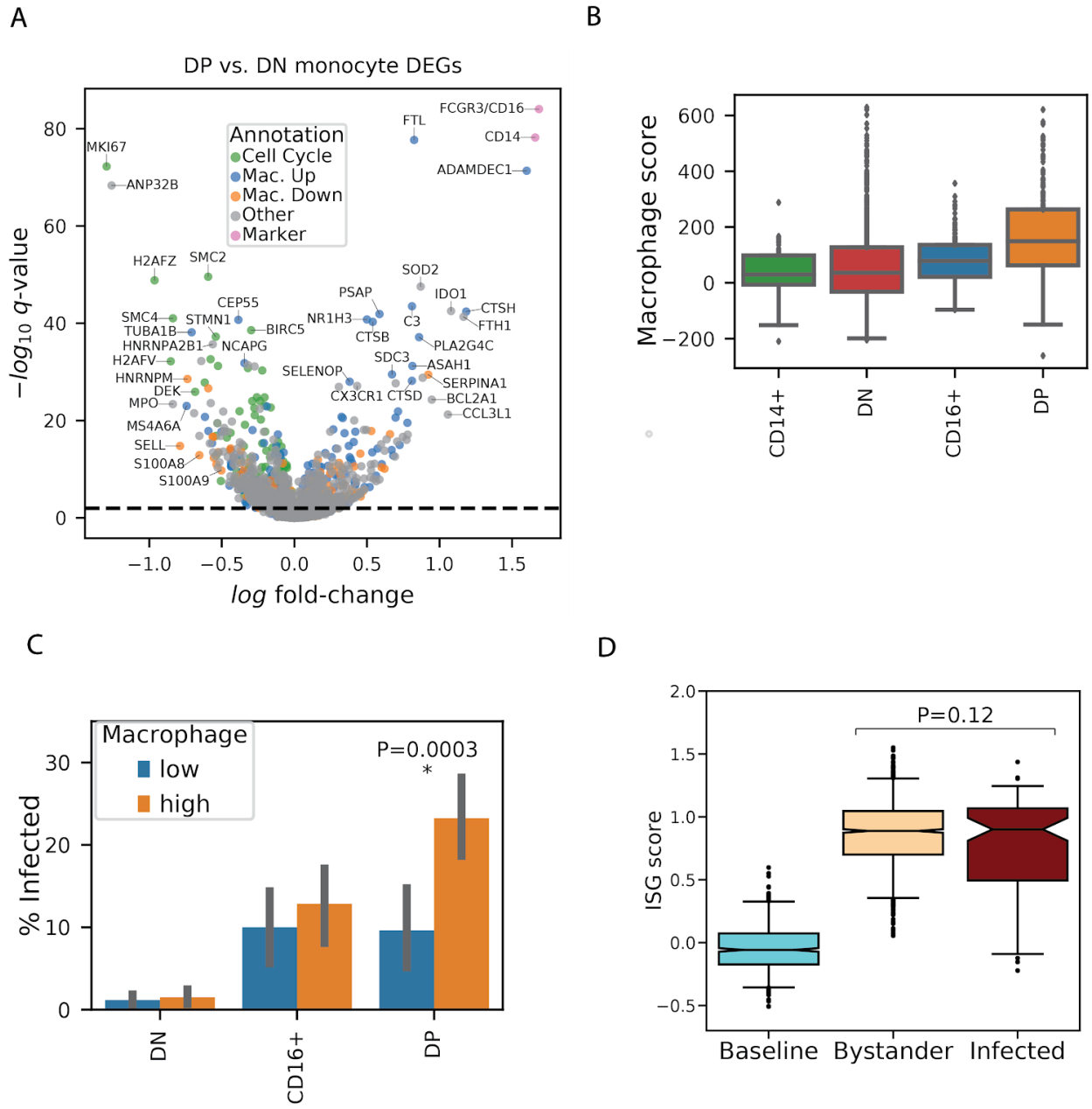


Figure S8. Extended characterization of gene expression signals associated with EBOV infection status in monocytes. Related to Figure 5.

(**A**) Volcano plot of differentially expressed genes between double positive and double negative monocyte subsets from DPI 5–8. Genes are colored by membership in cell cycle, macrophage up-regulated (Mac. Up), macrophage down-regulated (Mac. Down), or marker (*CD14*, *CD16*) gene sets. See also **Table S5**.

(**B**) Macrophage scores for monocytes in late EVD for each subset. Boxes denote the median and interquartile range, and whiskers denote the 2.5th and 97.5th percentiles.
(**C**) Percentage of infected monocytes in each subset in late disease, stratified by low or high macrophage score (below or above the median of monocytes from all subsets). Height of the bar denotes the mean, and error bars denote 95% bootstrap confidence intervals for the mean. There are no infected monocytes in the CD14+ subset.
(**D**) ISG scores of monocytes at baseline, and uninfected bystanders or infected cells in late stage EVD (DPI 6–8). Boxes denote the median and interquartile range, and whiskers denote the 2.5th and 97.5th percentiles.

Figure S9.

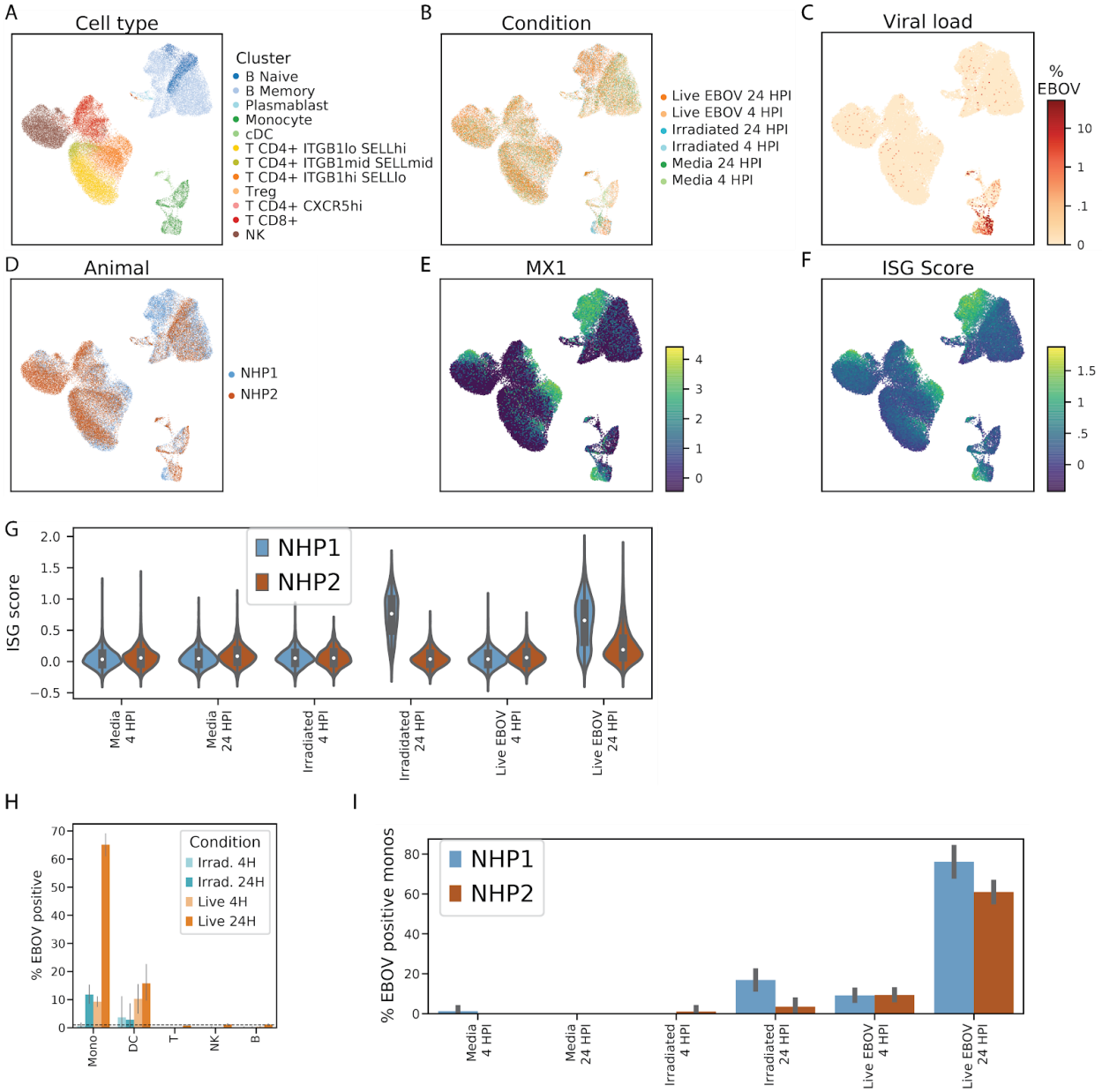


Figure S9. Overview of the *ex vivo* EBOV infection dataset. Related to Figure 6.
(**A–F**) Uniform Manifold Approximation and Projection (UMAP) embedding of Seq-Well data colored by annotated cluster assignment (A), treatment condition (B), viral load (C), NHP donor (D), *MX1* gene expression (natural log TP10K) (E), and interferon stimulated gene (ISG) score (F).
(**G**) Distributions of ISG scores across monocytes from each treatment condition, stratified by NHP donor. The central white marker denotes the median and the black bar denotes the interquartile range.
(**H**) Estimated percentage of infected cells of each cell type in the *ex vivo* dataset. The dashed line denotes the 1% false positive rate threshold used for calling infected cells.

(**I**) Percentage of EBOV-positive monocytes from each *ex vivo* treatment condition, stratified by NHP donor. Height of the bar denotes the mean, and error bars denote 95% bootstrap confidence intervals for the mean.

Figure S10.

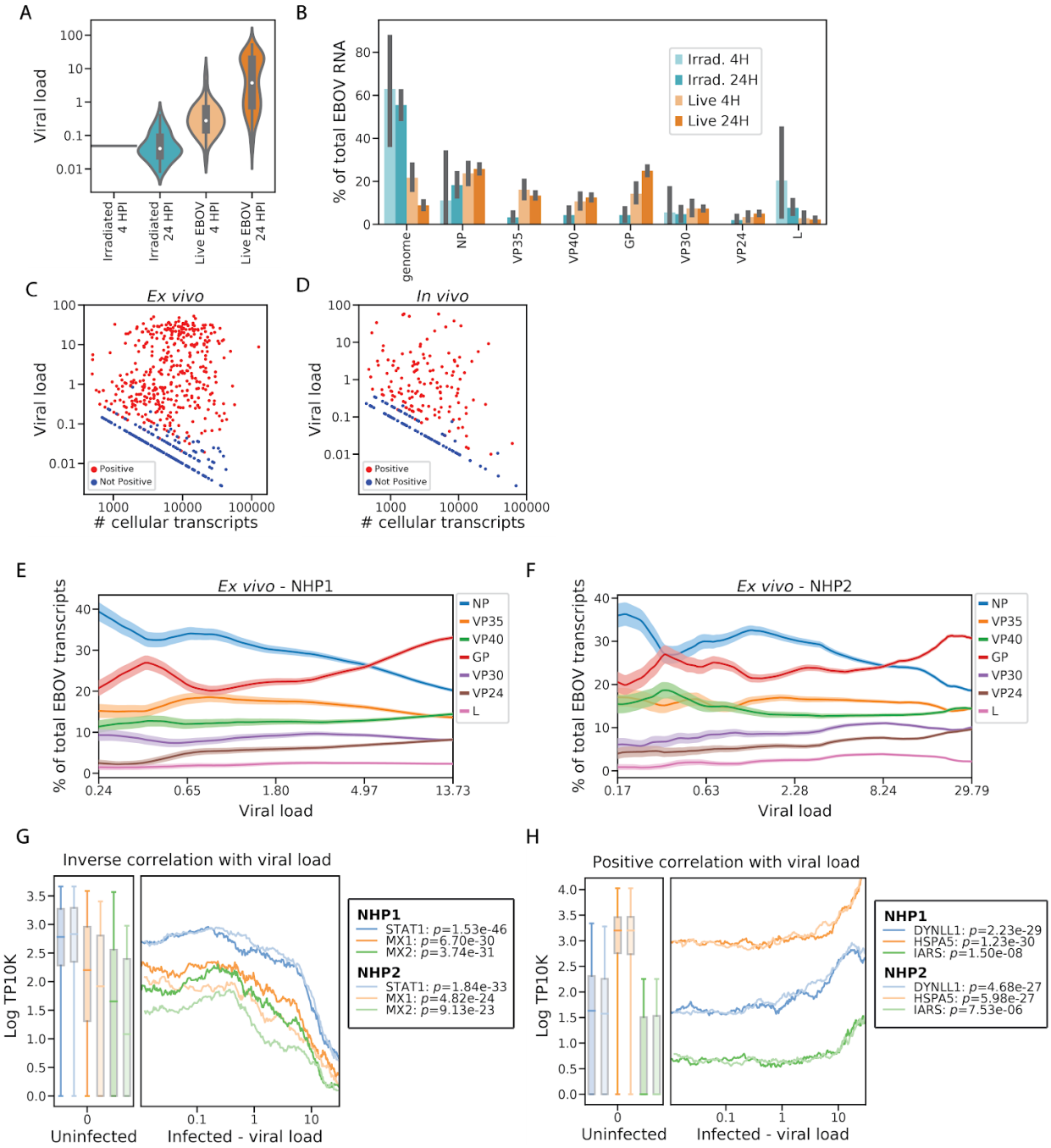


Figure S10. EBOV infection dynamics in the *ex vivo* dataset. Related to Figures 6 and 7.

(**A**) Distributions of viral loads across monocytes from different treatment conditions. The central white marker denotes the median and the black bar represents the interquartile range.

(**B**) Estimated percentage of EBOV transcripts derived from the EBOV genome or each EBOV gene, out of total viral RNA, stratified by treatment conditions. Prior to averaging, the counts of

EBOV genes for each cell was normalized to sum to one, so each cell contributes uniformly to the proportion, regardless of its total number of EBOV transcripts. Height of the bar denotes the mean, and error bars denote 95% bootstrap confidence intervals for the mean.
(**C and D**) Scatter plot of total transcripts (unique molecular identifiers) detected in a cell (*x*-axis, $\log_{10}$ scale) against viral load (*y*-axis, $\log_{10}$ scale) for cells with one or more viral reads *ex vivo* (C) or *in vivo* (D). Cells called as infected are colored in red and otherwise colored in blue.
(**E and F**) Relative proportion of each EBOV gene versus viral load ($\log_{10}$ scale) *ex vivo* for cells from donor NHP1 (E) or NHP2 (F). All monocytes were ordered by viral load and the percentage of each viral gene was averaged over a 50-cell sliding window. Color bands denote the mean ± 1 SD.
(**G and H**) Association between gene expression and viral load for selected negatively (G) and positively (G) associated host genes in monocytes, 24 HPI after inoculation with live virus *ex vivo*. In the left sub-plots, distributions of gene expression in uninfected bystander cells are shown as a boxplot with boxes denoting the interquartile range, and whiskers denoting the 2.5th and 97.5th percentiles. In the right sub-plots, we ordered infected cells by viral load ($\log_{10}$ scale) and averaged gene expression (natural log TP10K) over a 100-cell sliding window. Curves and box-plots are shown separately for the 2 donor NHPs. *p*-values for the Spearman correlation between viral load and gene expression are listed for each NHP donor in the legend.

Supplementary Tables

Table S1. Overview of samples analyzed by CyTOF and Seq-Well. Related to Figure 1.

Description of each NHP PBMC sample profiled in the study for CyTOF (tab 1) and Seq-Well (tab 2). Includes the animal ID, day post infection, number of cells passing quality control filters, the processing batch, and other technical factors such as whether the sample was processed fresh or underwent freeze-thaw, was treated with DASH, or included an MDCK cell spike-in. Subsequent tabs for the final counts of each broad cell type cluster for CyTOF (tab 3) and Seq-Well (tab 4).

Table S2. Differential expression profiles of cell types at each EVD stage relative to baseline. Related to Figure 3.

Log fold-change, *p*-values, and *q*-values for MAST differential expression tests comparing gene expression levels of cells at each EVD stage relative to baseline, for each cell type.

Table S3. Gene-set enrichment testing of differential expression modules. Related to Figure 3.

The “geneset_enrichment” tab includes the rank, gene set name, odds ratio (OR), and FDR-corrected q-value for the top 10 gene-sets associated with each differential expression module (ranked by q-value) based on Fisher’s exact tests. The “module_membership” tab includes the module assignment for all genes tested for differential expression including insignificant genes labeled as “Not significant”. The “ISG_set_forScoring” tab includes differentially expressed genes in the “global up” module that were annotated as interferon stimulated genes (ISGs) and subsequently used for computing the ISG score. The Unique_IFN_Alpha_Set species the genes included in the HALLMARK_INTERFERON_ALPHA_RESPONSE gene set but not the HALLMARK_INTERFERON_GAMMA_RESPONSE gene set, and Unique_IFN_Gamma_Set specifies the converse gene set.

Table S4. Differential expression results of infected vs. bystander monocytes, in vivo, at DPI 5–8. Related to Figure 5.

Log fold-change, *p*-values, and *q*-values for differential expression tests comparing gene expression levels between infected and bystander monocytes in late EVD. Also includes annotations of whether each gene is associated with macrophage differentiation or interferon response. Subsequent tabs specify the published gene lists associated with in vitro macrophage

differentiation that were used for gene set enrichment, as well as the contingency tables and hypothesis tests associated with each enrichment test.

Table S5. Differential expression between DP and DN monocytes in late EVD. Related to Figure S8.

Log fold-change, *p*-values, and *q*-values for differential expression tests comparing gene expression levels between DP and DN monocytes in late EVD. Also includes annotations of whether each gene is associated with macrophage differentiation, interferon response, or cell cycle.

Table S6. Association between host gene expression and viral load for infected monocytes *in vivo* and *ex vivo*. Related to Figure 7.

Log fold-change, *p*-values, and *q*-values for differential expression tests of association between log10 viral load and gene expression for cells containing ≥1 viral read. Tab 1 is for *in vivo* (cells in DPI 5-8), and tab 2 is for *ex vivo* (cells treated with live virus and collected at 24 hours post inoculation). Tabs 3 and 4 show the gene set enrichment results for separate Fisher's exact tests of genes that are significantly positively or negatively correlated with intracellular viral load ($P < .05$) including enrichment odds ratios (OR), p-values, and q-values.

Table S7. Antibody panel for cell staining prior to CyTOF. Related to Key Resources Table.