מכון ויצמן למדע
WEIZMANN INSTITUTE OF SCIENCE

Molecular characterization of human cytomegalovirus infection with single-cell transcriptomics

Document Version:
Accepted author manuscript (peer-reviewed)

Citation for published version:
Schwartz, M, Shnayder, M, Nachshon, A, Arazi, T, Kitsberg, Y, Levi Samia, R, Lavi, M, Kuint, R, Tsabari, R & Stern-Ginossar, N 2023, 'Molecular characterization of human cytomegalovirus infection with single-cell transcriptomics', *Nature Microbiology*, vol. 8, pp. 455-468. https://doi.org/10.1038/s41564-023-01325-x

Total number of authors:
10

Digital Object Identifier (DOI):
10.1038/s41564-023-01325-x

Published In:
Nature Microbiology

License:
Unspecified

General rights
@ 2020 This manuscript version is made available under the above license via The Weizmann Institute of Science Open Access Collection is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognize and abide by the legal requirements associated with these rights.

How does open access to this work benefit you?
Let us know @ library@weizmann.ac.il

Take down policy
The Weizmann Institute of Science has made every reasonable effort to ensure that Weizmann Institute of Science content complies with copyright restrictions. If you believe that the public display of this file breaches copyright please contact library@weizmann.ac.il providing details, and we will remove access to the work immediately and investigate your claim.

(article begins on next page)

1

Figure or Table # Please group Extended Data items by type, in sequential order. Total number of items (Figs. + Tables) must not exceed 10.	Figure/Table title One sentence only	Filename Whole original file name including extension. i.e.: Smith_ED_Fig1.jpg	Figure/Table Legend If you are citing a reference for the first time in these legends, please include all new references in the main text Methods References section, and carry on the numbering from the main References section of the paper. If your paper does not have a Methods section, include all new references at the end of the main Reference list.
Extended Data Fig. 1	Productive and non-productive infection in monocyte derived macrophages.	Schwartz_ED_Fig1.jpg	**a**. Gating strategy used to determine the single-cell population for all flow cytometry analyses of GFP-HCMV infected monocytes and PMA-induced MDMs. **b**. Flow cytometry analysis of primary CD14+ monocytes treated with the indicated treatments and infected with HCMV-GFP. Analysis was performed at 3dpi. Blue and red gates mark the GFP-dim and GFP-bright populations, respectively, and their percentage is noted. **c**. Gating strategy used to determine the single-cell population for flow cytometry analyses of GFP-HCMV infected M-CSF and GM-CSF/IL-4 treated monocytes. **d**. Flow cytometry analysis of CD11b surface expression in PMA and M-CSF treated monocytes and of CD11c surface expression in GM-CSF/IL-4 treated monocytes compared to primary CD14+ monocytes. Shown is a representative of two biological replicates with similar results. **e**. Flow cytometry analysis of PMA-induced monocyte-derived macrophages

			(MDMs) at different times following infection with HCMV-GFP. **f**. Measurements of infectious virus in supernatants from PMA-induced MDMs and from monocytes infected with HCMV-GFP at different times post infection. NA, not assayed. **g**. Microscopy image of PMA-induced MDMs at 3 dpi with HCMV-GFP. Scale bar is 100µm. Shown is a representative of two biological replicates. **h.** Flow cytometry analysis of PMA-induced MDMs infected with HCMV-GFP at 3 dpi showing the size, measured by forward scatter (FSC), of mock-infected, GFP-dim and GFP-bright cells. **i**. Flow cytometry plot showing the gates used to sort GFP-bright (red gate) and GFP-dim (blue gate) HCMV-GFP infected PMA-induced MDMs at 4 dpi. **j**. digital droplet PCR (ddPCR) results of GFP-dim and GFP-bright HCMV-GFP infected PMA-MDMs at 4 dpi. Upper panel shows detection of viral DNA, lower panel reflects detection of cellular genomes. The magenta line marks the threshold. **k-l**. Measurements of viral genomes by ddPCR from GFP-dim and GFP-bright FACS-sorted populations of monocytes treated with M-CSF (**i**) or GM-CSF/IL-4 (**j**). ddPCR graph is a representative of two biological replicates and shows the mean and 95% CV of poisson distribution of >10,000 PCR

			droplets. **m.** Microscopy image of HCMV-GFP infected BAL cells at 2dpi. Scale bar is 10µm. Shown is a representative of four biological replicates. **n.** Flow cytometry analysis of CD14 surface expression in BAL cells enriched for macrophages. Stained and unstained cells are in green and gray, respectively. **o.** Gating strategy used to determine the single-cell population for flow cytometry analyses of GFP-HCMV infected BAL macrophages.
Extended Data Fig. 2	Single cell transcriptome analyses of HCMV infected monocytes and MDMS	Schwartz_ED_Fig2.eps	**a.** Projection of mock-infected and HCMV-GFP infected CD14+ monocytes (n=68 and n=803, respectively) and MDMs (n=540 and n=1815, respectively) from the indicated time points, based on cellular and viral gene expression, colors indicate cell types and time post infection HCMV-infected cells are plotted in gray. **b.** Heatmap depicting the 40 most significantly increased genes between 5hpi infected and mock-infected monocytes and 12hpi-144hpi infected monocytes. ISGs are highlighted in bold. Z-scores of the log of the expression values are shown.
Extended Data Fig. 3	Features of single infected MDMs and BAL cells	Schwartz_ED_Fig3.jpg	**a**. Projection of 283 mock-infected monocytes and MDMs at 5 hpi colored by cell type. **b**. Projections as in **a**, colored by the expression level of macrophage differentiation genes. **c**. Levels of viral reads versus GFP reads in single HCMV-GFP infected MDMs. **d.** Projection of 21,839 HCMV-infected BAL cells from

			20 and 72 hpi colored by time post infection. Analysis was performed based on both cellular and viral transcripts. **e.** Projection as in d, colored by percentage of viral reads per cell. **f.** Projection as in d, colored by cell identity annotation based on reference datasets [16].
Extended Data Fig. 4	High expression of viral immediate early genes characterizes early productive MDMs.	Schwartz_ED_Fig4.jpg	**a**. Projection of HCMV-GFP infected MDMs at 5-144 hpi, analyzed based solely on cellular transcripts and colored by cluster assignment. **b**. Cluster assignment of each BAL myeloid cell according to analysis based on both viral and cellular genes (left) or on analysis based solely on cellular genes (right). Each cell is represented by a single line. The different non-productive cell clusters are colored in blue colors. Productively infected cells and clusters are colored in red. **c**. Relative expression levels, in the different MDM clusters, of viral genes from temporal classes 3-7 based on [15].
Extended Data Fig. 5	Ectopic expression of IE genes in THP1 derived macrophages drives productive infection.	Schwartz_ED_Fig5.jpg	**a**. Flow cytometry analysis of THP1 derived macrophages (DMs) infected with HCMV-GFP. Analysis was performed at 3dpi. Blue and red gates mark the GFP-dim and GFP-bright populations, respectively. **b**. ddPCR results of GFP-dim and GFP-bright HCMV-GFP infected THP1 DMs at 4 dpi. Upper panel shows detection of viral DNA, lower panel reflects detection of cellular genomes. The magenta line marks the threshold. **c**. ddPCR measurements of viral genomes from FACS-sorted GFP-bright and GFP-dim HCMV-GFP infected THP1 DMs at different times post infection. Graph is a representative of two biological replicates and shows the mean and 95% CV of poisson distribution of >10,000 PCR droplets. **d**. Measurements of infectious virus in supernatants

			from FACS-sorted GFP-bright and GFP-dim HCMV-GFP infected MDMs, at 5dpi. Points show measurements of biological replicates. p-value as calculated by one-tailed t-test on the replicates, is indicated. **e**. IE1/IE2 and mCherry protein levels at 8hpi visualized by western blot in THP1 DMs ectopically expressing UL123, UL122 or mCherry as control. GAPDH is shown as loading control. Shown is a representative of two biological replicates with similar results. **f**. pp28 (pUL99) and mCherry protein levels at 8hpi visualized by western blot in infected THP1 DMs ectopically expressing UL99 or mCherry as control. GAPDH is shown as loading control. mCherry blot is from the same samples run on a different gel. Shown is a representative of two biological replicates with similar results. **g**. Percentage of productively infected cells in infected THP1 DMs ectopically expressing mCherry (control), UL123, UL122 or UL99. Points show measurement of biological replicates. p-values as calculated by two-sided t-test on the replicates are indicated. **h**. Measurements of infectious virus in supernatants from THP1 DMs ectopically expressing UL123, UL122 or mCherry as control, at 8dpi. Points show measurements of biological replicates. p-values as calculated by one-tailed t-test on the replicates, are indicated. **i**. Percentage of productively infected cells in infected PMA-treated Kasumi-3 ectopically expressing UL123 or mCherry as control. Points show measurement of biological replicates. p-values as calculated by one-tailed t-test on the replicates are indicated. **j**. Distribution of the Spearman correlation coefficients between cellular transcript levels and viral gene expression level (left panels) or UL123 and UL122 expression levels (right panels), in

			early HCMV-GFP infected MDMs from 5hpi (top panels) or 12 hpi (bottom panels). The black line marks the matching normal distribution.
Extended Data Fig. 6	Induction of ISGs does not dictate infection outcome.	Schwartz_ED_Fig6.eps	**a**. Heatmap depicting the 40 most significantly increased genes in non-productive MDMs compared to early MDMs. ISGs are highlighted in bold. Z-scores of the log of the expression values are shown. **b**. Viral gene expression level versus ISG expression level in single HCMV-infected fibroblasts at different times post infection. Data taken from [20]. The percentage of each group of cells in the given time point is indicated. **c.** Percentage of productively infected MDMs either untreated (HCMV), or treated with interferons at the time of infection (HCMV+IFN). Points show measurements of biological replicates. p-value of the decrease in productively infected MDMs following interferon treatment, as calculated by one-tailed t-test on the replicates, is indicated. **d.** Relative expression level of representative ISGs, measured by qPCR, in infected MDMs either untreated (HCMV) or treated with interferons at the time of infection (HCMV+IFN), at 8 and 24 hpi. Points show measurements of biological replicates. p-values as calculated by one-tailed t-test on the replicates are indicated. **e.** Percentage of productively infected fibroblasts either untreated (HCMV), or treated with interferons at the time of infection (HCMV+IFN). Points show measurements of biological replicates. p-value of the decrease in productively infected fibroblasts following interferon treatment, as calculated by one-tailed t-test on the replicates, is indicated.

Extended Data Fig. 7	Intrinsic ISG levels affect HCMV infection outcome.	Schwartz_ED_Fig7.jpg	**a.** ISG expression level versus differentiation status in single bone marrow human haematopoietic stem cells. Data taken from [25] shows differentiation trajectory from hematopoietic stem cells (HSCs) to monocytes (mono, left panel) and data from [24] shows differentiation trajectory from hematopoietic stem cells (HSCs) to hematopoietic progenitor cells (HPCs, right panel). The linear regression line with confidence interval is shown. **b.** Percentage of productively infected MDMs either untreated (HCMV), or treated with interferons 24 hours prior to infection (IFN-HCMV). Points show measurements of biological replicates. p-value of the decrease in productively infected MDMs following interferon treatment, as calculated by one-tailed t-test on the replicates, is indicated. **c.** Percentage of productively infected fibroblasts either untreated (HCMV), or treated with interferons 24 hours prior to infection (IFN-HCMV). Points show measurements of biological replicates. p-value of the decrease in productively infected fibroblasts following interferon treatment, as calculated by one-tailed t-test on the replicates, is indicated. **d.** Measurements of infectious virus in supernatants from infected MDMs either untreated or treated with interferons 24 hours prior to infection, at different times post infection. **e**. qPCR measurements of viral genomes from infected MDMs either untreated or treated with interferons 24 hours prior to infection, at different times post infection. The plot shows the mean of 2 independent biological replicates. **f.** Relative expression level of representative ISGs, measured by qPCR, in infected MDMs either untreated (HCMV) or treated with ruxolitinib 24 hours prior to infection (Ruxo-HCMV), at 12 hpi.

			Points show measurements of biological replicates. p-values as calculated by one-tailed t-test on the replicates are indicated. **g**. Percentage of productively infected MDMs either untreated (HCMV), or treated with ruxolitinib 24 hours prior to infection (Ruxo-HCMV). Points show measurements of biological replicates. p-value of the increase in productively infected MDMs following Ruxo treatment, as calculated by one-tailed t-test on the replicates, is indicated. **h**. IRF9 or STAT2 protein levels as measured by western blot in THP1 DMs following CRISPR KO of IRF9 or STAT2, respectively. KO using two different gRNAs for each gene are shown. Actin was used as loading control. Shown is a representative of two biological replicates with similar results. **i.** Percentage of productively infected THP1 DMs following CRISPR KO of a control gene, IRF9 or STAT2. Two independent gRNAs were used for IRF9 and STAT2. Points show measurements of biological replicates. p-value of IRF9 or STAT2 KO compared to the control, as calculated by one-tailed t-test on the replicates is indicated. **j.** Percentage of productively infected THP1 DMs either untreated (HCMV) or treated with ruxolitinib at the time of infection (HCMV+Ruxo) following CRISPR KO of a control gene or IRF9. Points show measurements of biological replicates. p-value of IRF9 KO compared to the control, as calculated by one-tailed t-test on the replicates, is indicated.
Extended Data Fig. 8	Effects of HCMV infection on macrophage cell identity.	Schwartz_ED_Fig8.jpg	**a**. Gene set enrichment analysis (GSEA) of differentially expressed genes between GFP-dim and GFP-bright BAL infected macrophages compared to mock-infected cells, on innate immunity, immune signaling and immune identity and function pathways. **b.** Gating strategy used to determine the

			single-cell population for flow cytometry analysis of CD11b surface levels. **c.** Flow cytometry of CD11b surface levels in unstained and stained HCMV-GFP infected MDMs. **d.** Flow cytometry of CD11b surface levels in HCMV-GFP infected PMA-induced MDMs. The blue and red gates mark the non-productive and productive MDMs, respectively, as measured by GFP level. The inset in the bottom left shows the distribution of CD11b levels in the two populations. p value of the decrease in the mean fluorescence intensity in GFP-bright MDMs on the measured single cells is calculated using one-tailed t-test. Representative of two independent replicates is shown. **e.** Gating strategy used to determine the single-cell population for flow cytometry analysis of phagocytosis. **f.** Percentage of phagocytosing GFP-dim and GFP-bright HCMV-GFP infected M-CSF-induced MDMs. Points show measurements of biological replicates. p-value as calculated by t-test on the replicates is indicated.
Extended Data Fig. 9	Non-productive MDMs continuously express viral transcripts and can reactivate.	Schwartz_ED_Fig9.jpg	**a**. Heatmap depicting absolute levels of viral transcripts, in reads per million, in productive MDMs. Staging of the cells and ordering of the viral genes are as in Fig. 6a. The expression levels in the heatmap are relative to the highest level of each transcript. **b**. Percentage of viral reads in single monocytes along infection. Box plot shows standard parameters (median, Q1-Q3, 1.5 IQR). p-value as calculated by one-tailed Wilcoxon test comparing the cells in the first and last groups is indicated. n is 141, 27, 210, 348 cells for 24hpi, 48hpi, 72hpi and 144hpi, respectively. **c**. Percentage of viral reads in single non-productive PMA-induced MDMs along infection. Box plot shows standard parameters (median, Q1-Q3, 1.5 IQR). Data

			from 72 hpi and 144 hpi is from scRNA-seq of only GFP-dim cells to avoid noise coming from GFP-bright cells, expressing high levels of viral transcripts. p-value as calculated by one-tailed Wilcoxon test comparing the cells in the first and last groups is indicated. n is 413, 552, 628 cells for 24hpi, 72hpi and 144hpi, respectively. **d**. Phase-contrast and GFP microscopy images of PMA-induced MDMs infected with HCMV-GFP at different days post infection. Scale bars are 200µm. Shown are representatives of two biological replicates. **e.** Plaques formed on fibroblasts co-cultured with reactivated M-CSF-induced MDMs. Reactivation was done by either PMA or LPS treatment. Scale bars are 500µm. Shown are representatives of five and three biological replicates for PMA or LPS treatments, respectively. **f**. Plaques formed on fibroblasts co-cultured with reactivated PMA-induced MDMs. Reactivation was done by either PMA or LPS treatment. Scale bars are 500µm. Shown are representatives of two biological replicates. **g**. Plaques formed on fibroblasts co-cultured with reactivated interferon-pre treated, PMA-induced MDMs. Reactivation was done by PMA. Scale bar is 500µm. Shown is a representative of three biological replicates.
Choose an item.			

2

Item	Present?	Filename Whole original file name including extension. i.e.:	A brief, numerical description of file contents. i.e.: *Supplementary Figures 1-4, Supplementary Discussion, and Supplementary Tables 1-4.*

		Smith_SI.pdf. The extension must be .pdf	
Supplementary Information	Yes	Supplementary Table 1	Supplementary Table 1
Reporting Summary	Yes	Schwartz_reporting_summary	
Peer Review Information	No	*OFFICE USE ONLY*	

3

4

5

6

Type	Number Each type of file (Table, Video, etc.) should be numbered from 1 onwards. Multiple files of the same type should be listed in sequence, i.e.: Supplementary Video 1, Supplementary Video 2, etc.	Filename Whole original file name including extension. i.e.: *Smith_ Supplementary_Video_1.mov*	Legend or Descriptive Caption Describe the contents of the file
Choose an item.			
Choose an item.			
Choose an item.			
Choose an item.			
Choose an item.			

Choose an item.			

7

Parent Figure or Table	Filename Whole original file name including extension. i.e.: *Smith_SourceData_Fig1.xls,* or *Smith_ Unmodified_Gels_Fig1.pdf*	Data description i.e.: Unprocessed western Blots and/or gels, Statistical Source Data, etc.
Source Data Fig. 1		
Source Data Fig. 2		
Source Data Fig. 3		
Source Data Fig. 4		
Source Data Fig. 5		
Source Data Fig. 6		
Source Data Fig. 7		
Source Data Fig. 8		
Source Data Extended Data Fig./Table 1		
Source Data Extended Data Fig./Table 2		
Source Data Extended Data Fig./Table 3		

Source Data Extended Data Fig./Table 4	Schwartz_SourceData_ED_Fig4	Unprocessed western Blots
Source Data Extended Data Fig./Table 5		
Source Data Extended Data Fig./Table 6		
Source Data Extended Data Fig./Table 7	Schwartz_SourceData_ED_Fig7	Unprocessed western Blots
Source Data Extended Data Fig./Table 8		
Source Data Extended Data Fig./Table 9		
Source Data Extended Data Fig./Table 10		

8

Molecular characterization of human cytomegalovirus infection with single cell transcriptomics

Michal Schwartz[1,4*], Miri Shnayder[1,4], Aharon Nachshon[1], Tamar Arazi[1], Yaarit Kitsberg[1], Roi Levi Samia[1], Michael Lavi[1], Rottem Kuint[2], Reuven Tsabari[3], Noam Stern-Ginossar[1*]

1 Department of Molecular Genetics, Weizmann Institute of Science, Rehovot, Israel.

2 Institute of Pulmonary Medicine, Hadassah Hebrew University Medical Center, Faculty of Medicine, Hebrew University, Jerusalem, Israel.

3 Pediatric Pulmonology and CF Unit, Department of Pediatrics, Hadassah Hebrew University Medical Center, Jerusalem, Israel.

4 These authors contributed equally to this work.

* To whom correspondence should be addressed: michalsc@weizmann.ac.il or noam.stern-ginossar@weizmann.ac.il

Abstract

Human cytomegalovirus (HCMV) can result in either productive or non-productive infection, with the latter potentially leading to viral latency. The molecular factors dictating these outcomes are poorly understood. Here we used single cell transcriptomics to analyse HCMV infection progression in monocytes, which are latently infected, and macrophages, considered to be permissive for productive infection. We show that early viral gene expression levels, specifically of those encoding immediate early proteins, IE1 and IE2, are a major factor dictating productive infection. We also revealed that intrinsic, not induced, host cell interferon stimulated gene (ISG) expression level is a main determinant of infection outcome. Intrinsic ISG expression is downregulated with monocyte to macrophage differentiation, partially explaining increased macrophage susceptibility to productive HCMV infection. Furthermore, non-productive macrophages could reactivate, making them potential latent virus reservoirs. Overall, we decipher molecular features underlying HCMV infection outcomes and propose macrophages as a potential HCMV reservoir.

Introduction

Human cytomegalovirus (HCMV) is a prevalent pathogen of the beta-herpesvirus family, infecting the majority of the human population [1]. HCMV establishes latent infection which can lead to disease and mortality upon reactivation.

A wide range of cell types can be infected [2], but with variable infection outcomes, and generally differentiated cells are more permissive. Cells from the myeloid lineage are thought to play a critical role in HCMV latency and reactivation. Hematopoietic stem cells (HSCs), granulocyte-macrophage progenitor (GMPs) and blood monocytes are infected with HCMV but do not support productive infection and are the main cells in which HCMV latency has been characterized [3–5]. While bone marrow HSCs have been proposed as the long term HCMV latent reservoir, clinical data strongly implies that there are additional uncharacterized sources of latent HCMV outside the bone marrow, likely within tissues [6,7]. Unlike monocytes and earlier myeloid progenitors, terminally differentiated macrophages and dendritic cells were shown to be permissive for productive HCMV infection [8,9] yet the molecular basis for the different infection outcome between monocytes and their differentiated counterparts is not well understood.

To shed light on molecular determinants that dictate infection outcome, we scrutinized HCMV infection progression in monocytes and macrophages by single-cell RNAseq. First, we reveal that macrophages can support both productive and non-productive infection. By comparing these two infection outcomes, as well as to infected monocytes, we show that high early viral gene expression levels, particularly of immediate early genes, facilitate productive infection. On the host side, we demonstrate that intrinsic Interferon-stimulated gene (ISG) expression levels, but not their induction upon infection, have a substantial effect on infection outcome and are correlated to cellular differentiation. Finally, we show that non-productive macrophages can reactivate to produce infectious progeny. Altogether, our findings define determinants of HCMV infection outcome in the myeloid lineage and raise the possibility that macrophages serve as latent viral tissue reservoirs.

Results

Experimental system for studying HCMV infection outcome

We initially compared infection of monocytes to differentiated monocytes by using three treatments: phorbol-myristate (PMA), GM-CSF/IL-4 and M-CSF, which have all been shown to trigger monocyte differentiation [10,11]. We used an HCMV TB40-E strain containing an SV40-GFP reporter [12], which allows quantification of productive infection efficiency. As expected, 3 days post infection (dpi) in monocytes, almost all of the cells show a small increase in GFP, indicating they are infected but that viral gene expression is repressed (Fig. 1a and Extended Data Fig. 1a). Following all differentiation-inducing treatments, infection resulted in two distinct populations: GFP-dim, suggesting the cells were infected but viral gene expression is repressed, and GFP-bright, suggesting productive infection (Fig. 1a and Extended Data Fig. 1b-c). Cells within each of the treatments displayed uniform shifts in surface expression of macrophage/dendritic cell markers, negating the possibility that differences in differentiation efficiency within the cell populations can solely explain the differences in infection outcomes (Extended Data Fig. 1d). For further experiments we primarily used PMA derived macrophages because they are commonly used in HCMV studies and exhibited a large and distinct population of GFP-bright cells (Fig. 1a). The two distinct populations in infected monocyte-derived macrophages (MDMs) are not apparent at early stages of infection(Extended Data Fig. 1e), suggesting that they likely do not reflect two separate initial populations, but rather two different infection outcomes . Infection of MDMs, but not of monocytes, resulted in production of infectious virions (Extended Data Fig. 1f), confirming productive infection in these cells.

We verified that the GFP-dim and GFP-bright populations reflect different infection outcomes, by showing that GFP-bright cells are larger compared to GFP-dim cells (Extended Data Fig. 1g-h), a known feature of lytic HCMV infection. Moreover, both sorted GFP-bright and GFP-dim populations (Extended Data Fig. 1i) harbored viral DNA, however, at 4dpi viral DNA load was much higher in the GFP-bright cells (Extended Data Fig. 1j-l) and substantially increased with time in the GFP-bright but not in the GFP-dim population (Fig. 1b). This indicates that viral genome replication occurs only in the GFP-bright population. In addition, GFP-bright MDMs produce infectious virions while GFP-dim cells do not (Fig. 1c). These results indicate that the majority of MDMs were infected with the virus but a productive infection only developed in some of the cells. We further assessed infection of natural tissue myeloid cells, collected from

BronchoAlveolar Lavage (BAL) samples. The presence of macrophages in these samples was verified by size, morphology and surface marker staining (Extended Data Fig. 1m-n). Indeed, also in BAL macrophages we observed two distinct populations of GFP-dim and GFP-bright cells that differ in viral genome levels (Fig. 1d-e and Extended Data Fig. 1o).

Single-cell RNA-seq of HCMV-infected monocytes and MDMs

To shed light on the processes governing these two infection outcomes, we analyzed infected monocytes and PMA-induced MDMs from the same donor at 5, 12, 24, 48, 72 and 144 hours post infection (hpi), as well as mock-infected monocytes and MDMs at the 5 and 144 hpi time points, by single-cell RNA-sequencing (scRNA-seq, Fig. 1f). Projection of all cells according to both cellular and viral reads, revealed several distinct groups of cells (Fig. 1g, Extended Data Fig. 2a). At 5 hpi, infected monocytes were grouped with mock-infected monocytes (Fig. 1g, Extended Data Fig. 2a). Infected monocytes from 12 hpi onwards appeared as a single population (Fig.1g) and stratified from early monocytes mainly based on interferon stimulated genes (ISGs, Extended Data Fig. 2b). As expected, viral transcript load in infected monocytes remained low in all cells along infection (median=0.063%, Fig. 1h), indicating lack of productive infection [13–15]. In contrast, infected MDMs from 24 hpi and onwards, were separated into two distinct populations (Fig. 1g) differing in viral transcript load, reflecting two infection outcomes: productive, in which viral transcript load increased as infection progressed, and non-productive, in which viral transcript loads were low (median=0.226%, Fig. 1h). MDMs from earlier time points clustered with mock-infected MDMs (Fig. 1g, Extended Data 2a), suggesting milder transcriptional changes relative to later in infection.

Two distinct infection outcomes in infected macrophages

Since infection of MDMs results in non-productive and productive outcomes, we focused our analysis on infected MDMs. Notably, all mock-infected MDMs were transcriptionally distinct from the initial monocyte population (Extended Data Fig. 3a) and showed increased macrophage differentiation signature (Extended Data Fig. 3b) indicating that it is unlikely that the two distinct infection outcomes originate from two initially distinct populations. Projection of all infected MDMs according to cellular and viral transcripts, revealed that cells were divided into three

discernible populations (Fig. 2a). Most of the cells from the early time points (5-12 hpi) belong to a single group, harboring low levels of viral transcripts and indistinguishable in regard to infection outcome, while cells from later time points in infection split into two additional distinct groups (Fig. 2b), which differ greatly in the level of viral transcripts (Fig. 2c-d) and in virally encoded GFP expression levels (Extended Data Fig. 3c). The three main populations therefore correspond to early, non-productive, and productive MDMs (Fig 2e). Further sub-clustering corresponds to progression in the given trajectory with time post infection (Fig. 2b, e).

scRNA-seq analysis of infected BAL cells from 20 and 72hpi demonstrated that the cells stratify into two distinct populations and an additional small population that expresses significantly higher viral transcript levels, reflecting productively infected cells (Extended Data Fig. 3d-e). By using an annotated reference dataset [16], we found that these two populations correspond to lymphocytes, mostly T cells, and to myeloid cells, mostly macrophages, and the productive cells originate from myeloid cells (Extended Data Fig. 3f). Analysis of the myeloid cells demonstrated that infection of natural tissue macrophages also results in two outcomes: productive and non-productive, as reflected by high and low viral transcript levels, respectively (Fig. 2f).

Viral IE gene expression determines infection outcome

To delineate the contribution of cellular and viral genes to the stratification of the MDMs into distinct groups (early, non-productive, productive), we next analyzed the transcriptome solely according to cellular genes. This analysis produced the same three distinct groups of cells and very similar cluster stratification, indicating these three infection states are associated with corresponding changes in the cellular transcriptome (Fig. 3a and Extended Data Fig. 4a). There were two exceptions: Cluster 5, which includes very few non-productive cells from different time points, and cluster 7, which is the earliest productive cluster (Fig. 2a) and when clustering was based only on cellular genes, these cells clustered together with the early cells (Fig. 3a). This demonstrates that the first major transcriptional signature that defines cells as productively infected is robust viral gene expression, and this was also apparent in analysis of infected BAL macrophages (Extended Data Fig. 4b). Quantification of the expression of the different temporal classes of viral genes[15] within cells in the different clusters, reveals that in the early productive cells, compared to their non-productive counterparts, there is high relative expression of

immediate early (IE) genes, as well as specific induction of true early genes, but not of other viral temporal classes (Fig. 3b, Extended Data Fig. 4c). This suggests that expression of IE genes at early stages of infection, at levels sufficient to induce early genes, has a major role in dictating infection outcome. Indeed, the UL123 and UL122 transcripts which encode the major IE genes, IE1 and IE2, respectively, are specifically expressed in early productive cells (Fig. 3c). We further scrutinized the effect of IE gene expression on infection outcome, in macrophages derived from the monocytic cell line THP1, whose infection similarly exhibits two outcomes (Extended Data Fig. 5a-d). Induced ectopic expression of either UL123 or UL122 but not of a control gene (mCherry) or UL99, a viral late gene (Extended Data Fig. 5e-f), significantly increased the percentage of productively infected cells (Fig. 3d and Extended Data Fig. 5g). Moreover, ectopic expression of both IE genes, increased the production of infectious progeny (Extended Data Fig. 5h). The effect of IE1 was further validated in differentiated Kasumi-3 cells (Extended Fig. 5i).

Since scRNA-seq gives unprecedented detail on inter-cellular heterogeneity, we next explored whether we can identify cellular genes that correlate with the initial levels of viral gene expression but we could not detect any significant association (Extended Data Fig. 5j). This likely reflects the relatively homogenous nature of the cells we used, such that small differences in gene expression are masked by other factors, such as the amount of incoming virions [17], or variability between the content of virions [18].

Induction of ISGs does not dictate infection outcome

We next examined the features that characterize the shift from early infection to the non-productive population. The genes that were most upregulated in the non-productive cells were ISGs (Extended Data Fig. 6a, Fig. 4a-b). Analysis of ISG expression versus viral gene expression in infected MDMs revealed that at early time points, both viral gene and ISG expression are low. As infection progressed, most of the cells increased either viral gene expression or ISG expression (Fig. 4c). Notably, the cells of both trajectories were exposed to the same extracellular signals (e.g. interferons), as they were grown together in the same culture dish, indicating that the induction of viral gene expression renders the cells unresponsive to interferon signaling [19]. Since we hardly detected infected MDMs that induced both viral genes and ISGs, this analysis may imply that these events describe two competing deterministic outcomes - induction of viral

genes resulting in productive infection, or induction of ISGs prohibiting productive infection. However, analysis of the percentage of cells in the different populations at each time point indicates that productive cells likely also originate from the non-productive groups in which ISGs were already induced, as the percentage of productive cells reaches numbers that cannot be solely explained by transition from the early infected population (Fig. 4d). This therefore suggests that induction of interferon response does not forestall establishment of productive infection and the main determinant that governs infection outcome is the induction of viral gene expression above a certain threshold. Analysis of a recent single cell RNA-seq dataset along HCMV infection of fibroblasts [20] demonstrated the same phenomenon (Extended Data Fig. 6b). These observations suggest that induction of ISGs by an infected cell does not deterministically prevent the establishment of productive infection within that cell.

In accordance with these results, adding interferon at the time of infection did not affect the percentage of productively infected cells (Fig. 4e and Extended Data Fig. 6c), despite significant increase in ISG expression (Extended Data Fig 6d). Similar results were obtained in low MOI infection of fibroblasts (Fig. 4f and Extended Data Fig. 6e). Together these results demonstrate that activation of ISGs in response to infection may slow productive infection but overall does not deterministically prevent it.

Intrinsic ISG levels affect HCMV infection outcome

Analysis of monocytes compared to non-productive MDMs showed similarly low viral gene expression but higher ISG expression level even at the earliest infection stage (Fig. 4g). Further analysis of mock-infected monocytes, MDMs and fibroblasts[20] revealed that intrinsic levels of ISGs (in mock-infected cells) in fibroblasts is lower than in MDMs, which in turn is lower than in monocytes (Fig. 4h) and there is an inverse-correlation between intrinsic ISG levels and viral gene expression levels at early stages of infection (Fig. 4i). Since the permissiveness for productive infection is tightly linked to differentiation state [21,22], we examined the link between intrinsic expression of ISGs and differentiation state. Indeed, intrinsic levels of many ISGs decrease along differentiation of monocytes to MDMs [23] (Fig. 4j-k) and along differentiation of hematopoietic cells in the bone marrow [24,25] (Extended Data Fig. 7a), suggesting a tight link

between intrinsic levels of ISGs, differentiation and permissiveness to productive HCMV infection.

To explore if the levels of ISGs at the time of infection affect infection outcome, we tested the effect of interferon pre-treatment on HCMV infection. Addition of interferons 24 hours prior to infection of MDMs or fibroblasts eliminated productive infection (Fig. 4e-f, Extended Fig. 7b-d), blocked viral DNA replication (Extended Data Fig. 7e), and led to reduction in viral gene expression already at 12hpi (Fig. 4l). These results indicate that ISGs act to block the virus at the stage of immediate early viral gene expression or earlier.

We next sought to test whether reducing intrinsic ISG levels would indeed affect infection outcome. Intrinsic levels of ISGs may originate from low basal levels of interferon signaling, however pretreatment of MDMs with ruxolitinib, a potent and selective Janus kinase (JAK) 1 and 2 inhibitor did not affect the intrinsic levels of ISGs (Extended Data Fig. 7f) and correspondingly did not affect infection outcome (Extended Data Fig. 7g). It was recently suggested that basal expression of ISGs is controlled by preformed STAT2-IRF9 complexes, independent of interferon receptor signaling [26]. Indeed, depletion of IRF9 or STAT2 using CRISPR-cas9 in THP1 derived macrophages (Extended Data Fig. 7h) reduced the intrinsic expression of many ISGs (Fig. 4m) and led to an increase in the percentage of productively infected cells (Fig. 4n and Extended Data Fig. 7i). This effect was not diminished by addition of ruxolitinib at the time of infection (Extended Data Fig. 7j), illustrating it is due to reduction in intrinsic ISGs rather than reduced induction of ISGs following infection. Overall, these results indicate that intrinsic expression of ISGs, which decrease upon differentiation, plays a major role in dictating HCMV infection outcome.

Productive HCMV infection disrupts macrophage identity

We next analyzed cellular transcriptional changes and revealed that several pathways known to be modulated during HCMV infection in fibroblasts were significantly enriched in productive MDMs, including metabolism, cell cycle and translation pathways (Fig. 5a)[27,28]. As mentioned above, non-productive MDMs significantly induced, mainly and most robustly, ISGs.

Several pathways were suppressed in productive MDMs, including some which are associated with immune signaling or common functions of macrophages, such as antigen presentation and

phagocytosis as well as genes related to hematopoietic cell lineage (Fig. 5a). RNA-seq on infected BAL macrophages showed that indeed similar pathways related to immune signaling and macrophage identity are significantly suppressed in productively infected BAL macrophages and ISGs are strongly induced in the non-productively infected macrophages (Extended Data Fig. 8a). The disruption of macrophage identity and function was further supported by reduction in surface levels of CD11b, a prominent macrophage marker, in productively infected MDMs, induced by either PMA or M-CSF, (Fig. 5b and Extended Data Fig. 8b-d) as well as reduced phagocytosis activity in these cells (Fig. 5c and Extended Data Fig. 8e-f). These results show, both on the transcriptional and the functional levels, that productive HCMV infection disrupts macrophage identity and function.

Viral gene expression in HCMV infected MDMs

We next examined viral gene temporal expression along productive and non-productive infection. For the productive MDMs we binned the cells according to % viral reads to reflect progression along infection. This was followed by hierarchical clustering, revealing a temporal expression cascade in productive MDMs that is similar to that seen in infection of fibroblasts [15] (Fig. 6a, left heatmap). This analysis reflects relative expression patterns, i.e. how each viral gene is expressed relative to all other viral genes. At the absolute level, almost all viral genes increase in their expression levels as infection progresses (Extended Data Fig. 9a). In non-productive MDMs, despite viral transcript levels being low throughout infection, they slightly increase with time (Fig. 6b), unlike infected monocytes (Extended Data Fig. 9b) [13]. To exclude the possibility that this apparent increase in viral reads originates from low level cross contamination from productive cells, we conducted another experiment in which at 72 and 144 hpi only GFP-dim cells were processed, which further demonstrated a small increase in viral transcript levels (Extended Data Fig. 9c). Nevertheless, analysis of viral gene expression in non-productive cells along infection revealed broad gene expression with no clear kinetic program (Fig. 6a, right heatmap).

Non-productive MDMs reactivate and release infectious virus

The observation that viral gene expression is maintained at higher levels in non-productive macrophages compared to monocytes, indicates a less repressed state in the former.

Furthermore, in culture, these cells survived for relatively long periods (Extended Data Fig. 9d) and also *in-vivo* macrophages can be long lived cells [29]. We therefore tested if the virus residing in non-productive MDMs, differentiated either by PMA or M-CSF, could reactivate. We did not observe any spontaneous reactivation, however, additional stimulation with either PMA or LPS at 7dpi resulted in reactivation and release of infectious virus, (Fig. 6c, Extended Data Fig. 9e-f). Moreover, interferon-pretreated, infected MDMs, which, as described above, are non-productively infected, also reactivated when PMA was added at 7dpi, and released infectious virus (Fig. 6d and Extended Data Fig. 9g). It is noteworthy that in these cells there was some low level reactivation in the absence of additional stimuli, which is in line with previous findings [30].

These results indicate that in cell culture, macrophages can be infected in a latent-like manner, in which viral gene expression is repressed, but the viral genome retains the ability to reactivate upon additional stimulation. HCMV infection following exposure to interferons may further lead to a similar latent-like infection that can also result in reactivation.

Discussion

HCMV infection can result in different outcomes: productive infection leading to release of infectious virus, or non-productive infection, potentially leading to latency establishment, from which the virus can later reactivate. To identify, in an unbiased manner, determinants of HCMV infection outcome, we analyzed the progression of HCMV infection in monocytes and macrophages, which support latent and lytic infection, respectively. Our data demonstrates that in fact, HCMV infection of monocyte-derived macrophages (MDMs) can be either productive or non-productive, allowing us to analyze the decision point of infection outcome. We show that the unequivocal factor associated with the earliest split of productive cells, is viral gene expression and specifically the expression of immediate early genes and the onset of early gene expression. The importance of IE gene expression was already acknowledged in the past [31], however, ectopic expression of these genes in THP1 cells was previously shown to be insufficient to induce productive infection [32]. Here we show that overexpression of UL123 or UL122 in THP1-derived macrophages shifts infected cells towards productive infection. The notable difference between the reports is the use of differentiated versus non-differentiated THP1 cells. Further research is needed in order to uncover what are the additional barriers in non-differentiated THP1 cells and whether they extend to primary monocytes. The lack of a clear cellular gene expression signature that is associated with viral gene expression levels at the early stages of infection, indicates that in this experimental set-up, infection outcome is likely determined primarily by the amount of infectious virus entering the cell [17], which will have significant consequences on the initial induction of viral gene transcription.

The defining signature of non-productive MDMs is the expression of ISGs as was also shown in HSV-1 infected cells [33]. However, our data suggests that induction of ISGs following infection does not deterministically define infection outcome, although it does not rule out the possibility that it slows down or hampers the propagation of the virus.

Instead, we reveal that intrinsic expression levels of ISGs prior to infection dictate infection outcome, and are associated with myeloid differentiation state. This connection between intrinsic expression of ISGs, differentiation and infection was noted previously for infection of embryonic stem cells with several RNA viruses [34]. Our results suggest that this link could be a major contributing factor to the cell-type dependent infection outcome of HCMV and potentially other herpesviruses.

Beyond explaining cell type differences, this effect of intrinsic levels of ISGs could be relevant to the establishment of latency. In a setting where not all cells are infected, induction of interferons by infected cells could shift the fate of neighboring uninfected cells towards non-productive infection, potentially aiding the virus to seed latent reservoirs, and perhaps also prolonging infection in case of viral gene silencing which can be reversed [30].

Reactivation of HCMV from latency is a substantial clinical burden in immunocompromised individuals and may cause serious and often life-threatening disease [35]. The only characterized HCMV latent reservoir to date is the bone marrow HSCs [36], however, clinical data suggest that latent reservoirs exist also in tissues [7]. For many years tissue resident macrophages were considered to be short lived and be replenished regularly by circulating monocytes. However, mounting evidence indicates that resident macrophage populations reside in organs and can remain for years [29,37,38]. Our results demonstrate that macrophages have the potential to be infected in a latent-like manner, in which they carry viral genomes, repress viral gene expression and can reactivate. Natural infection of macrophages is supported by reports of HCMV reactivation in alveolar macrophages from healthy HCMV carriers [39] and of latent murine macrophages following clearance of an acute CMV infection [40], implying that tissue macrophages may serve as latent CMV reservoir.

On the other hand, our analysis of productively infected macrophages demonstrates a dramatic effect on macrophage identity and function. This analysis aligns with a recent study by Baasch et al. that showed productive MCMV infection of macrophages reprograms macrophage identity and disrupts their function, leading to higher susceptibility to secondary lung infection [41].

These findings highlight macrophages as prospective critical players during HCMV infection both in pathogenesis at the time of acute infection, as well as potential latent reservoirs in tissues.

Methods

Ethics statement

All fresh peripheral blood samples were obtained after approval of protocols by the Weizmann Institutional Review Board (IRB application 92–1) and following informed consent from the donors. The study using BronchoAlveolar Lavage (BAL) fluid samples was approved by the Hadassah Medical Organization research ethics committee in accordance with the Declaration of Helsinki (HMO-0704-20). Written informed consent to retain BAL fluid was obtained from patients undergoing bronchoscopy. Blood donors and bronchoscopy patients were not compensated.

Cell culture and virus

Primary CD14+ monocytes were isolated from fresh venous blood, obtained from healthy donors, males and females, aged 25-45, using Lymphoprep (Stemcell Technologies) density gradient followed by magnetic cell sorting with CD14+ magnetic beads (Miltenyi Biotec). Monocytes were cultured in X-Vivo15 media (Lonza) supplemented with 2.25 mM L-glutamine at 37°C in 5% CO2, at a concentration of 1-2 million cells/ml in non-stick tubes to avoid differentiation.

Where indicated, primary monocytes were treated with the following treatments immediately following isolation and plating: PMA 50ng/ml for 3 days, GM-CSF and IL-4 1000U/ml each for 6 days with cytokine replenishing after 3 days, M-CSF 20ng/ml for 6days with cytokine replenishing after 3 days, TNFα 10ng/ml for 6 days. All treatments were done in Roswell Park Memorial Institute (RPMI) with 10% heat-inactivated fetal bovine serum (FBS), 2 mM L-glutamine, and 100 units/ml penicillin and streptomycin (Beit-Haemek, Israel) and MDMs were grown in this media following treatment. 293T cells (ATCC- CRL-3216) were maintained in DMEM with 10% FBS, 2 mM L-glutamine, and 100 units/ml penicillin and streptomycin (Beit-Haemek, Israel). Primary human foreskin fibroblasts (ATCC CRL-1634) were maintained in DMEM with 10% FBS, 2 mM L-glutamine, and 100 units/ml penicillin and streptomycin (Beit-Haemek, Israel). THP1 cells, purchased from ATCC (TIB-202), were grown in RPMI media with 15% heat-inactivated fetal bovine serum (FBS), 2 mM L-glutamine, and 100 units/ml penicillin and streptomycin (Beit-Haemek, Israel). Kasumi-3, purchased from ATCC (CRL-2725) were grown in RPMI media with 20% heat-inactivated fetal bovine serum (FBS), 2 mM L-glutamine, and 100

372 units/ml penicillin and streptomycin (Beit-Haemek, Israel). Differentiation of THP1 and Kasumi-3
373 cells was done by adding PMA 50ng/ml for 2 days. Infection and growth following infection were
374 done in the media containing 10% FBS. Interferon treatment was done with a combination of
375 550 U/ml IFNα, 700 U/ml IFNβ and 500 U/ml IFNγ. Ruxolitinib treatment was done at a
376 concentration of 4μM.

377 BronchoAlveolar Lavage fluid was obtained from routine bronchoscopies of patients, males and
378 females aged 58-82. 15-40ml were obtained from each sample and diluted by adding an equal
379 volume of RPMI media with 10% heat-inactivated fetal bovine serum (FBS), 2 mM L-glutamine,
380 and 100 units/ml penicillin and streptomycin (Beit-Haemek, Israel). Diluted BAL fluid was
381 transferred through a 70uM filter and plated on a tissue culture dish. Adherent cells were
382 washed thoroughly prior to infection.

383 The TB40E strain of HCMV, containing an SV40-GFP tag, was described previously [12]. Virus was
384 propagated by adenofection of infectious bacterial artificial chromosome (BAC) DNA into
385 fibroblasts as was previously described [42]. Viral stocks were concentrated by centrifugation at
386 26000xg, 4°C for 120 minutes. Infectious virus yields were assayed on primary fibroblasts.

387 Infection and reactivation procedures

388 Infection was performed by incubating cells with the virus for 3 hr, washing twice and
389 supplementing with fresh media. Initial experiments with primary non-differentiated and
390 differentiated monocytes as well as the single cell experiments of monocytes and MDMs were
391 done at MOI=10 to ensure infection of all the cells. Notably, since this MOI is based on
392 quantification of infectious particles in fibroblasts it is likely effectively lower in monocytic cells.
393 The single cell experiments in BAL macrophages as well as follow up experiments in THP1- and
394 primary monocyte- derived macrophages were done in MOI=5. For reactivation assays, GFP-dim
395 infected MDMs were sorted at 4dpi and plated in 96-well plates (10,000 cells per well),
396 interferon-pretreated MDMs were used following infection without further sorting. At 7dpi cells
397 were either treated, as indicated, with 50ng/ml PMA (Sigma), or 500ng/ml LPS (Sigma) or left
398 untreated for 3 days and then the media was replaced with fresh untreated media. At 14dpi, 7
399 days post treatment, GFP positive cells were counted on a fluorescent microscope. At 18dpi
400 7500 fibroblasts per well were seeded with the MDMs and plaque formation was followed and
401 imaged for a period of 14 days.

Flow cytometry and sorting

Cells were analyzed on a BD Accuri C6 using the BD Accuri C6 software or a BD LSRII and sorted on a BD FACSAria III using FACSDiva software. All analyses and figures were done with FlowJo. All histograms were plotted with modal normalization.

Measurement of viral progeny

Fibroblasts were infected with known volumes of supernatants taken from infected cultures. Number of GFP positive fibroblasts were measured at 2dpi by flow cytometry or microscopy and the infectious particles per volume was calculated.

Measurement of viral genomes by digital PCR

DNA was extracted from cell pellets in a 1:1 mixture of lysis solutions A (100 mM KCl, 10 mM Tris–HCl pH 8.3, and 2.5 mM MgCl2) and B (10 mM Tris–HCl pH 8.3, 2.5 mM MgCl2, 0.25% Tween 20, 0.25% Non-idet P-40, and 0.4 mg/ml Proteinase K), for 60 min at 60˚C followed by a 10 min 95˚C incubation. Measurement of viral DNA was done using the QX200 droplet digital PCR system (Bio-Rad), using FAM labeled HCMV primer and probe: Human CMV HHV5 kit for qPCR using a glycoprotein B target (PrimerDesign); and HEX labeled RPP30 copy number assay for ddPCR (Bio-Rad), as previously described [43]. Analysis was done using the Quantasoft software.

Microscopy

Imaging was performed on an Olympus Ix73 inverted fluorescent microscope using a X4 or a X10 objective and an Olympus DP73 camera using the CellSens dimensions software or on an AxioObserver Z1 wide-field microscope using a X20 objective and Axiocam 506 mono camera using the ZEN software.

Single cell RNA sequencing library preparation

For primary monocytes and MDMs, single-cell sorting and library preparation were conducted according to the massively parallel single-cell RNA-seq (MARS-seq) protocol, as previously described [44]. In brief, single cells were FACS sorted into wells of 384-well capture plates containing 2µl of lysis buffer and reverse transcription (RT)-indexed poly(T) primers, thus generating libraries representing the 3’ end of mRNA transcripts. Eight empty wells were kept in

each 384-well plate as a no-cell control during data analysis. Immediately after sorting, each plate was spun down to ensure cell immersion into the lysis solution, snap-frozen on dry ice, and stored at -80°C until processed. Barcoded single-cell capture plates were prepared with a Bravo automated liquid handling platform (Agilent). For generation of the RNA library, mRNA from single cells was converted into cDNA in the capture plates and pooled by centrifugation to VBLOCK200 reservoir (Clickbio). The pooled sample was linearly amplified by T7 in vitro transcription, and the resulting RNA was fragmented and converted into a sequencing-ready library by tagging the samples with pool barcodes and Illumina sequences during ligation, RT, and PCR. Each pool of cells was tested for library quality, and concentration was assessed as described earlier [44]. In total we analyzed 871 monocytes and 2355 MDMs. Numbers of cells in clusters 1-10 of infected MDMs are: 347, 325, 232 229, 33, 122, 53, 274, 117, 83.

For BAL cells, pools of 10,000 cells were prepared for scRNA-seq using Chromium Single Cell 3' Gene Expression v3.1 kit according to the manufacturer's instructions (10x Genomics).

sc RNA-seq data analysis

Analysis of MARS-seq data was done with the tools described in Jaitin et al. [45]. The reference was created from the hg19 and TB40E (NCBI EF999921.1) strain of HCMV. The transcription units of the virus were based on NCBI annotations, with some changes based on the alignment results. This includes merging several transcripts (taking into account that the library maps only the 3' ends of transcripts) and adding some antisense transcripts. Read assignment to wells was based on the batch barcode (4 bp) and the well barcode (7 bp) and reads with low-quality barcodes were removed. The trimmed read (37 bp) was aligned with the reference using Bowtie 2 [46], and the counting of the reads is done based on unique molecular identifiers (UMIs, 8 bp). For each batch, the leakage noise level was estimated by comparing the number of UMIs in the eight empty wells to the total number of UMIs in the batch. The noise levels in all batches was lower than 1.5% . Wells with less than 500 or more than 5,000 expressed genes, or more than 20,000 reads were discarded. The number of wells that were used for further analysis is 3,226 .

For 10xGenomics data, we used Cell Ranger software (V6.1.2, 10x Genomics) with the default settings to process the FASTQ files with the references used in the analysis of the MARS-seq data. Cells with less than 500 or more than 10,000 expressed genes, or more than 30,000 reads were discarded. The number of wells that were used for further analysis is 21,839 .

All single cell analyses were done using Seurat V4.0.3 [16]. To visualize single-cell datasets, we performed dimension reduction by Uniform Manifold Approximation and Projection (UMAP), based on both cellular and viral genes in Fig. 1, 2 and Extended Data Fig. 3 or only cellular genes in Fig. 3 and Extended Data Fig. 4. Clusters of MDMs were numbered based on % viral reads in the cluster, from lowest to highest. Cluster 5 includes very few non-productive cells from different time points and its nature is unclear, but since it is very small, these cells were removed from all following cluster-based analyses. Annotation of cell types in BAL samples was done using the Azimuth application [16].

Calculation of viral and cellular gene expression levels in single cells

Expression of a single cellular gene or a group of cellular genes in single cells is calculated as the percentage of reads on this gene or group of genes out of the total cellular reads in a cell. % viral reads is calculated as the percentage of reads on all viral genes out of all reads in a cell. These numbers were used for coloring UMAP plots in Fig. 1h, 2c, 2f, 4a Extended Data Fig. 3b, 3e and for plots in Fig. 2d, 4b, 4c, 4g, 4h, 4i, 6b, Extended Data Fig. 6b, 7a, 9b, 9c. The list of ISGs is the combined list of the following gene sets: Hallmark interferon alpha response, Hallmark interferon gamma response and Reactome interferon signaling. The list of Macrophage differentiation genes is based on the list of genes in the GO term regulation of macrophage differentiation (GO:0045649).

Relative gene expression level of a single viral gene or a group of viral genes is calculated as the percentage of reads on this gene or group of genes out of the total viral reads in a cell. These numbers were used for coloring the UMAP plot in Fig. 3c and for plots in Fig. 3b, and Extended Data Fig. 4c.

Subtraction of cross contaminating viral reads was done in the following way: for each plate the level of noise was estimated from the ratio between the average number of UMIs in 8 empty wells to the average number of UMIs in the plate. This number was used to calculate the fraction of viral UMIs that can be attributed to leakage between cells, and was subtracted from the number of viral reads in all cells of the plate.

Bulk RNA-seq library preparation and alignment

For RNA-seq of THP1 cells, cells were collected with Tri-Reagent (Sigma-Aldrich), total RNA was extracted by phase separation and poly(A) selection was performed using Dynabeads mRNA

DIRECT Purification Kit (Invitrogen) according to the manufacturer's protocol. RNA-seq libraries were generated as previously described [47]. mRNA samples of 4 ng were subjected to DNaseI treatment and 3' dephosphorylation using FastAP Thermosensitive Alkaline Phosphatase (Thermo Scientific) and T4 PNK (NEB) followed by 3' adaptor ligation using T4 ligase (NEB). The ligated products were used for reverse transcription with SSIII (Invitrogen) for first-strand cDNA synthesis. The cDNA products were 3' ligated with a second adaptor using T4 ligase and amplified with 8 cycles of PCR for final library products of 200–300 base pairs. Raw sequences were first trimmed at their 30 end, removing the illumina adapter and polyA tail. Alignment was performed using Bowtie 1 [48] (allowing up to 2 mismatches) and reads were aligned to the human (hg19). Reads aligned to ribosomal RNA were removed. Reads that were not aligned to the genome were then aligned to the transcriptome.

For RNA-seq of BAL macrophages, pellets from cell sorting were lysed in RLT buffer (Qiagen) containing 40mM DTT. Due to the low number of cells, RNA-seq libraries were performed using a modified mcSCRB-seq [49] adapted for bulk RNA-seq. Briefly, RNA was isolated using SPRI beads, reverse transcription and template switching was performed with indexed poly(T) primers in the presence of polyethylene glycol, followed by amplification of the final library.

66-bp raw reads were aligned using STAR (V2.7.1) [50] to concatenation of the human (hg19) and the viral genomes (NCBI EF999921.1). Counting of reads per gene was done based on unique molecular identifiers (UMIs, 10 bp).

Plasmids and lentiviral transduction

UL123, UL122, UL99 or mCherry were cloned into pLVX-Puro-TetONE-SARS-CoV-2-nsp1-2XStrep (kind gift from Nevan Krogan, UCSF) in place of the SARS-CoV-2-nsp1-2XStrep cassette. UL123 was also cloned into pLVX-EF1alpha-SARS-CoV-2-nsp1-2XStrep-IRES-Puro (Addgene plasmid #141367) in place of the SARS-CoV-2-nsp1-2XStrep cassette. The UL123, UL122, and UL99 coding sequence were amplified from cDNA of HCMV infected fibroblasts with primers containing flanking regions homologous to the vector (Supplementary Table 1). The mCherry was amplified from an expression plasmid using primers containing flanking regions homologous to the vector (Supplementary Table 1). The plasmids were amplified with the appropriate primers (Supplementary Table 1). The amplified PCR fragments were cleaned using a gel extraction kit (Promega) according to the manufacturer's protocol. mCherry was cloned into the

vector by restriction-free cloning [51] and the viral genes were cloned into the vectors using a Gibson assembly protocol [52].

CRISPR KO was done using a CRISPR–cas9 system, with lentiCRISPR v2 plasmid (Addgene no. 52961). sgRNAs were cloned downstream of the U6 promoter (Supplementary Table 1).

Lentiviral particles were generated by co-transfection of the expression constructs and 2nd generation packaging plasmids (psPAX2, Addgene#12260 & pMD2.G, Addgene#12259), using jetPEI DNA transfection reagent (Polyplus transfection) into 293T cells, according to the manufacturer's instructions. 60 hours post transfection, supernatants were collected and filtered through a 0.45µm PVDF filter (Millex). THP1 cells were transduced with lentiviral particles by spinfection (800xg, 1hr) and were then puromycin-selected (1ug/mL) for 4 days. Puromycin was removed prior to subsequent differentiation. Induction of the inducible transgene was done by addition of 100ng/ml Doxycycline (Sigma) 24 hours prior to HCMV infection. Kasumi-3 cells were transduced with lentiviral particles by spinfection (800xg, 1hr). Since Doxycycline induction was not successful in our hands in Kasumi-3 cells, they have been transduced with a non-inducible version of UL123 expression plasmid and a control mCherry expression plasmid. To avoid long term effects of the transgene expression, no selection step was done and the cells were treated with PMA two days after transduction followed by HCMV infection two days later.

Quantitative real-time PCR analysis

For analysis of RNA expression, total RNA was extracted using Tri-Reagent (Sigma) according to the manufacturer's protocol. cDNA was prepared using qScript cDNA Synthesis Kit (Quanta Biosciences) according to the manufacturer's protocol. For analysis of DNA levels, cells were lysed in a 1:1 mixture of PCR solutions A (100 mM KCl, 10 mM Tris–HCl pH 8.3, and 2.5 mM MgCl2) and B (10 mM Tris–HCl pH 8.3, 2.5 mM MgCl2, 1% Tween 20, 1% Non-idet P-40, and 0.4 mg/ml Proteinase K), for 60 min at 60˚C followed by a 10 min 95˚C incubation. Real time PCR was performed using the SYBR Green PCR master-mix (ABI) on the QuantStudio 12K Flex (ABI).

Primers spanning an exon-exon junction, were used for viral transcript quantification, for ISG transcript quantification, and for the host gene Anxa5 that was used for normalization (Supplementary Table 1).

549 DNA was extracted from cell pellets or directly from adherent cells in a 1:1 mixture of lysis
550 solutions A (100 mM KCl, 10 mM Tris–HCl pH 8.3, and 2.5 mM MgCl2) and B (10 mM Tris–HCl pH
551 8.3, 2.5 mM MgCl2, 0.25% Tween 20, 0.25% Non-idet P-40, and 0.4 mg/ml Proteinase K), for 60
552 min at 60˚C followed by a 10 min 95˚C incubation. Primers for the viral UL44 gene were used to
553 quantify viral DNA and a host gene (B2M) was used for normalization (Supplementary Table 1).

554 Western blot analysis

555 Cells were lysed in RIPA buffer for 30 minutes on ice and then centrifuged at 14,000 xg for 20
556 min at 4°C. Samples were separated by 4 to 12% polyacrylamide Bis-Tris gel electrophoresis
557 (Invitrogen), blotted onto nitrocellulose membranes, and immuno- blotted with the following
558 primary antibodies: mouse anti-CMV IE1 and IE2 1:1000 (abcam), mouse anti-PP28 1:2000
559 (EastCoast Bio), mouse anti-mCherry 1:1000 (ChromoTek), rabbit anti-GAPDH 1:1000 (Cell
560 Signaling Technology), rabbit anti- IRF9 1:1000 (Cell signaling Technology), rabbit anti- STAT2
561 1:1000 (Cell signaling Technology) , mouse anti- actin 1:7500 (Sigma). The secondary antibodies
562 were goat anti-rabbit–IRDye 800CW and goat anti-mouse–IRDye 680RD (LI-COR) diluted 1:7500.
563 Membranes were visualized in a LI-COR Odyssey imaging system.

564 Correlation between cellular gene expression and viral gene expression levels

565 For calculating the Spearman correlation between cellular genes and viral expression levels, in
566 each cell, the sum of viral reads was normalized to the total number of reads in the cell, and the
567 number of reads of each cellular gene was normalized to the total number of cellular reads in
568 the cell. The correlation was calculated across 313 and 310 cells on 5007 and 5606 genes for the
569 5 hpi and 12 hpi analyses, respectively. Genes with a total number of reads less than 45 (after
570 normalization to 5,000 reads per cell) were not included.

571 Differential expression and enrichment analysis

572 Differential expression analysis on bulk RNA-seq data was performed with DESeq2 (V1.22.2) [53]
573 using default parameters, with the number of reads in each of the samples as an input.
574 Differential expression analysis of the scRNA-seq was done using Seurat V4.0.3 [16]

575 The log2 fold change values were used for enrichment analysis using GSEA (version 4.1, [54]). For
576 scRNA-seq data, only genes that were expressed in at least 20% of cells in both groups of cells in
577 a comparison were used. For bulk RNA-seq, only genes that had at least 10 reads in each sample

were used. The gene sets that were used were Hallmark and KEGG from MSigDB (version 7.4, [55]).

Viral gene expression heatmaps

Viral gene expression heatmaps depict the viral gene expression level either relative to total number of viral reads (Fig. 6a) or normalized to the total number of reads (Extended Data Fig. S9a). The expression levels in the heatmap are relative to the highest level of each gene. Productive MDMs were ordered by their % viral reads. The cells were then grouped according to the number of viral reads per bin increasing from the lowest bin to the highest bin, to ensure a similar number of cells per bin. The heatmap shows the 50 viral genes with the top relative expression at their highest expressing bin. Non-productive cells were grouped according to the time post of infection. In each bin/time group, expression level of each viral gene was calculated as the total number of reads on the gene divided by the total number of viral reads.

Marker staining and phagocytosis assay

Cells were stained in cold MACS buffer (PBS, 5% BSA, 2 mM EDTA). Cell staining was done using an APC-conjugated anti-human CD11b (#101211, Biolegend), PE-conjugated anti-human CD11c (#333149, BD), or FITC-conjugated anti-human CD14 (#130-080-701, Miltenyi) diluted 1:200. Phagocytosis was assayed using a phagocytosis kit (#601490, Cayman chemical) according to the manufacturer's instructions. The percentage of phagocytosing cells was calculated for the productive and non-productive populations (as measured by GFP level) according to the flow cytometry gate.

Graphics

Fig. 1f was created with BioRender. Flow cytometry figures were created with FlowJo. Fig. 3a and Extended Data Fig. 4b were created with the R package ggalluvial. Figure 4d was created with R package sankey.

Data availability

Sequencing data have been deposited in GEO under accession code GSE189581. Other datasets analyzed in this study: Single cell RNA-seq dataset along HCMV infection of fibroblasts is available under accession number GSE165291. RNA-seq dataset along differentiation of monocytes to MDMs is available under accession number GSE147314. Single cell RNA-seq

datasets along differentiation of hematopoietic cells in the bone marrow are available under accession number GSE75478 and through the Human Cell Atlas data portal at https://data.humancellatlas.org/explore/projects/091cf39b-01bc-42e5-9437-f419a66c8a45.

Correspondence and requests for materials should be addressed to Michal Schwartz.

Acknowledgements

We thank Nir Drayman, Roni Winkler, and the members of the Stern-Ginossar lab, for critical reading of the manuscript. We thank Dr. Keren Bahar Halpern and Dr. Shani Ben-Moshe for assistance with mcSCRB-seq and 10xGenomics library preparation. We thank Eain A Murphy for the TB40E-GFP virus strain. We thank the Weizmann flow cytometry unit for technical assistance. This study was supported by a European Research Council consolidator grant to N.S-G (CoG-2019-864012). N.S-G is a member of the European Molecular Biology Organization (EMBO) Young Investigator Program.

Author Contributions

M.Sc., M.Sh., A.N. and N.S-G. conceived and designed the project, M.Sc., M.Sh., T.A., Y.K., R.L.S. and M.L. performed the experiments. R.K. and R.T. provided BAL samples and advice, M.Sc., M.Sh., A.N. and N.S-G. analyzed and interpreted the data. M.Sc. and N.S-G. wrote the manuscript with input from all the authors.

Competing Interests

The authors declare no competing interests.

Figure 1. Experimental system for studying HCMV infection outcome.

a. Flow cytometry analysis of primary CD14+ monocytes and monocytes differentiated with PMA, infected with HCMV-GFP. Analysis was performed at 3dpi. The mock-infected GFP level is marked by a line. Blue and red gates mark the GFP-dim and GFP-bright populations, respectively, and their percentage is noted. **b**. ddPCR measurements of viral genomes from FACS-sorted GFP-bright and GFP-dim HCMV-GFP infected MDMs at different times post infection. Graph is a representative of two biological replicates and shows the mean and 95% CV of poisson distribution of >10,000 PCR droplets. **c**. Measurements of infectious virus in supernatants from FACS-sorted GFP-bright and GFP-dim HCMV-GFP infected MDMs, at different times post infection. **d**. Flow cytometry analysis of BAL macrophages infected with HCMV-GFP at 3dpi. **e**. ddPCR measurements of viral genomes from FACS-sorted GFP-bright and GFP-dim HCMV-GFP infected BAL macrophages at 3dpi. Graph is a representative of two biological replicates and shows the mean and 95% CV of poisson distribution of >10,000 PCR droplets. **f**. Experiment design. primary monocytes and PMA-induced MDMs were infected with HCMV-GFP and at 5, 12, 24, 48 ,72 and 144 hpi were processed for scRNA-seq. In parallel mock-infected cells from 5 and 144 hpi were also processed for scRNA-seq. **g.** Projection of HCMV-GFP infected CD14+ monocytes (n=803) and MDMs (n=1815) as well as mock-infected monocytes (n=68) and MDMs (n=540) from the indicated time points, based on cellular and viral gene expression, colors indicate cell types and time post infection mock-infected cells are plotted in gray. **h**. Projection as in **g**, showing all infected cells, colored by percentage of viral reads per cell.

Figure 2. Single-cell transcriptome analysis of HCMV infected macrophages.

a. Projection of 1815 HCMV-GFP infected MDMs at 5-144 hpi colored by cluster assignment. Analysis was based on both cellular and viral transcripts. **b**. Projections as in figure **a**, showing the cells from each time point along infection. **c**. Projection, as in **a**, colored by percentage of viral reads per cell. **d**. Percentage of viral reads in single MDMs from the different clusters along infection. **e**. Distribution of the single cells between the different early, non-productive and productive clusters along infection. **f**. Projection of 10,320 HCMV-GFP infected myeloid BAL cells from 20 and 72 hpi colored by percentage of viral reads per cell. Circles and triangles mark cells from 20 or 72 hpi, respectively. Analysis was based on both cellular and viral transcripts.

Figure 3. Viral IE gene expression is a major determinant of infection outcome.

a. Cluster assignment of each cell according to analysis based on both viral and cellular genes (left) or on analysis based solely on cellular genes (right). Each cell, represented by a single line, is colored by the percentage of viral reads. **b**. Relative expression levels, within cells from the different clusters, of viral genes from temporal classes 1 (immediate early) and 2 (true early) based on [15]. **c**. Projection of infected MDMs (as in Fig. 2a) colored by the log expression level of UL123 or UL122 relative to all other viral genes. **d**. Flow cytometry analysis of HCMV-GFP infected THP1-derived macrophages over expressing mCherry as control, UL123, UL122 or UL99. Replicate data and statistics are shown in Extended Data Fig. 4g.

Figure 4. Intrinsic ISG levels but not their induction affect infection outcome.

a. Projection of infected MDMs (as in Fig. 2a) colored by the expression level of ISGs. **b**. Expression level of ISGs in single MDMs from the different clusters. Box plot shows standard parameters (median, Q1-Q3, 1.5 IQR). p-values as calculated by one-tailed t-test on the cells in a group of clusters as marked are indicated. See methods for number of cells in the analysis. **c**. Viral gene expression level versus ISG expression level, in single MDMs at different times post infection. The percentage of each group of cells in the given time point is indicated. **d**. Flow chart describing the transition of cells between early, non-productive and productive groups along infection. **e**. and **f**. Flow cytometry analysis of HCMV-GFP infected MDMs (**e**) or fibroblasts (**f**) either untreated, treated with interferons at the time of infection (HCMV+IFN) or treated with interferons 24 hours prior to infection (IFN-HCMV). Mock-infected cells are shown as control. Replicate data and statistics are shown in Extended Data Fig. 5c, 5e, 5g and 5h. **g**. Viral gene expression level versus ISG expression level in single infected monocytes, early and non-productive MDMs at different times post infection. **h**. ISG expression level in mock-infected monocytes and MDMs as measured from the scRNA-seq data and in mock-infected fibroblasts [20]. Violin plot shows the median, Q1-Q3 (box), distribution of values (violin) and outliers. n is 897, 258 and 25 for fibroblasts, MDMs and monocytes, respectively. **i**. ISG expression level in mock-infected cells versus viral gene expression level at early stages of infection (12 hpi or 20 hpi, in monocytes/MDMs and fibroblasts, respectively). p-values as calculated by one-tailed t-test between cells of different types as marked are indicated. **j**. Expression levels of different

690 ISGs in cells along differentiation trajectory from monocytes to MDMs. Data taken from [23]. Violin
691 plot shows the median, Q1-Q3 (box), distribution of values (violin) and outliers. p-value as
692 calculated by one-tailed t-test comparing the cells in the first and last groups is indicated. n= 265
693 genes. **k.** Cellular transcript levels in monocytes and MDMs[23]. TPM, transcript reads per kilobase
694 per million reads mapped to cellular transcripts. **l.** Relative expression level of the viral
695 transcripts UL123 and UL22A, measured by qPCR at 12hpi, in infected MDMs either untreated
696 (HCMV) or treated with interferons 24 hours prior to infection (IFN-HCMV). Points show
697 measurement of biological replicates. p-values as calculated by one-tailed t-test on the
698 replicates are indicated. **m.** Differential expression analysis of cellular genes in uninfected THP1
699 derived macrophages (DMs) following CRISPR KO of IRF9 (left panel) or STAT2 (right panel)
700 compared to control CRISPR KO. p-values of the enrichment of ISGs in the differentially
701 expressed genes as calculated by one-tailed hypergeometric test are indicated. **n.** Flow
702 cytometry analysis of HCMV-GFP infected THP1 DMs following CRIPSR mediated knockout of a
703 control gene, IRF9 or STAT2. Two independent gRNAs were used for IRF9 and STAT2. Percentage
704 of productively infected cells are indicated for each condition. Replicate data and statistics are
705 shown in Extended Data Fig. 5n.

706

707 **Figure 5.** Effects of HCMV infection on cellular pathways and macrophage cell identity and
708 function.

709 **a**. Summary of gene set enrichment analysis (GSEA) of differentially expressed genes between
710 each of the scRNA-seq clusters and mock-infected cells from 5 hpi using annotated KEGG and
711 Hallmark pathways. Uninfected MDMs at 144 hpi are shown as control for changes occurring
712 with time in culture regardless of infection. NES- normalized enrichment score. FDR- false
713 discovery rate. **b**. Flow cytometry of CD11b surface levels in M-CSF-induced MDMs infected with
714 HCMV-GFP. The blue and red gates mark the non-productive and productive MDMs,
715 respectively, as measured by GFP level. The inset in the bottom left shows the distribution of
716 CD11b levels in the two populations. p value of the decrease in the mean fluorescence intensity
717 in GFP-bright MDMs on the measured single cells is calculated using one-tailed t-test.
718 Representative of two independent replicates is shown. **c**. Phagocytic activity measured by flow
719 cytometry of HCMV-GFP infected M-CSF-induced MDMs. The black line marks the gate for
720 identification of cells as phagocytosing. The blue and red gates mark the non-productive and

productive MDMs, respectively, as measured by GFP level. Replicate data and statistics are shown in Extended Data Fig. 6c.

Figure 6. Non-productive MDMs exhibit mild elevation of gene expression with time and can reactivate to produce viral progeny.

a. Heat map depicting normalized relative levels of the 50 top relatively expressed viral transcripts in productive MDMs (left heatmap) and in non-productive MDMs (right heatmap). Productive MDMs are binned and staged by percentage of viral reads. The expression levels in the heatmap are relative to the highest level of each transcript. Viral genes are ordered based on hierarchical clustering according to the productive MDMs. Annotations according to temporal classes [15] are shown to the left of the heatmaps. **b**. Percentage of viral reads in single non-productive MDMs along infection. Box plot shows standard parameters (median, Q1-Q3, 1.5 IQR). p-value as calculated by one-tailed Wilcoxon test comparing the cells in the first and last groups is indicated. n is 147, 189, 165, 107 cells for 24hpi, 48hpi, 72hpi and 144hpi, respectively. **c**. Number of reactivation events in HCMV infected, GFP-dim sorted, M-CSF-induced MDMs following different stimulations. Points show measurements of biological replicates. p-value of the increase in reactivation events upon stimuli, as calculated by one-tailed t-test on the replicates, is indicated. **d**. Number of reactivation events in interferon pre-treated, HCMV infected, PMA-induced MDMs either untreated or additionally stimulated with PMA. Points show measurements of biological replicates. p-value of the increase in reactivation events upon stimuli, as calculated by one-tailed t-test on the replicates, is indicated.

References

1. Staras, S. A. S. *et al.* Seroprevalence of Cytomegalovirus Infection in the United States, 1988-1994. *Clinical Infectious Diseases* **43**, 1143–1151 (2006).
2. Mocarski E, Shenk T, Griffiths P, Pass R. Cytomegaloviruses. in *Fields virology 6th ed.* (ed. Knipe DM, Howley PM, Cohen JI, Griffin DE, Lamb RA, Martin MA, Racaniello VR, Roizman B) 1960–2014 (Lippincott Williams & Wilkins, Philadelphia, PA., 2013).
3. Hahn, G., Jores, R. & Mocarski, E. S. Cytomegalovirus remains latent in a common precursor of dendritic and myeloid cells. *Proceedings of the National Academy of Sciences* **95**, 3937–3942 (1998).
4. Kondo, K., Kaneshima, H. & Mocarski, E. S. Human cytomegalovirus latent infection of granulocyte-macrophage progenitors. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 11879–11883 (1994).
5. Taylor-Wiedeman, J., Sissons, J. G., Borysiewicz, L. K. & Sinclair, J. H. Monocytes are a major site of persistence of human cytomegalovirus in peripheral blood mononuclear cells. *J. Gen. Virol.* **72 (Pt 9)**, 2059–2064 (1991).
6. Emery, V. C. Relative Importance of Cytomegalovirus Load as a Risk Factor for Cytomegalovirus Disease in the Immunocompromised Host. *CMV-Related Immunopathology* 288–301 Preprint at https://doi.org/10.1159/000061707 (1997).
7. Reddehase, M. J. & Lemmermann, N. A. W. Cellular reservoirs of latent cytomegaloviruses. *Med. Microbiol. Immunol.* **208**, 391–403 (2019).
8. Ibanez, C. E., Schrier, R., Ghazal, P., Wiley, C. & Nelson, J. A. Human cytomegalovirus productively infects primary differentiated macrophages. *J. Virol.* **65**, 6581–6588 (1991).
9. Sinzger, C., Plachter, B., Grefte, A., T. Houw The & Jahn, G. Tissue Macrophages Are Infected by Human Cytomegalovirus In Vivo. *The Journal of Infectious Diseases* **173**, 240–

245 (1996).

10. Hargett, D. & Shenk, T. E. Experimental human cytomegalovirus latency in CD14+ monocytes. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 20039–20044 (2010).
11. Poole, E., Reeves, M. & Sinclair, J. H. The use of primary human cells (fibroblasts, monocytes, and others) to assess human cytomegalovirus function. *Methods Mol. Biol.* **1119**, 81–98 (2014).
12. O'Connor, C. M. & Murphy, E. A. A Myeloid Progenitor Cell Line Capable of Supporting Human Cytomegalovirus Latency and Reactivation, Resulting in Infectious Progeny. *Journal of Virology* **86**, 9854–9865 (2012).
13. Shnayder, M. *et al.* Defining the Transcriptional Landscape during Cytomegalovirus Latency with Single-Cell RNA Sequencing. *MBio* **9**, (2018).
14. Cheng, S. *et al.* Transcriptome-wide characterization of human cytomegalovirus in natural infection and experimental latency. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E10586–E10595 (2017).
15. Rozman, B. *et al.* Temporal dynamics of HCMV gene expression in lytic and latent infections. *Cell Rep.* **39**, 110653 (2022).
16. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e29 (2021).
17. Erhard, F. *et al.* scSLAM-seq reveals core features of transcription dynamics in single cells. *Nature* **571**, 419–423 (2019).
18. Chaturvedi, S. *et al.* A molecular mechanism for probabilistic bet hedging and its role in viral latency. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 17240–17248 (2020).
19. Shnayder, M. *et al.* Single cell analysis reveals human cytomegalovirus drives latently infected cells towards an anergic-like monocyte state. *Elife* **9**, (2020).

20. Hein, M. Y. & Weissman, J. S. Functional single-cell genomics of human cytomegalovirus infection. *Nature Biotechnology* (2021) doi:10.1038/s41587-021-01059-3.

21. Reeves, M. B. Chromatin-mediated regulation of cytomegalovirus gene expression. *Virus Res.* **157**, 134–143 (2011).

22. Galinato, M. *et al.* Single-Cell Transcriptome Analysis of CD34 Stem Cell-Derived Myeloid Cells Infected With Human Cytomegalovirus. *Front. Microbiol.* **10**, 577 (2019).

23. Song, R. *et al.* IRF1 governs the differential interferon-stimulated gene responses in human monocytes and macrophages by regulating chromatin accessibility. *Cell Rep.* **34**, 108891 (2021).

24. Velten, L. *et al.* Human haematopoietic stem cell lineage commitment is a continuous process. *Nat. Cell Biol.* **19**, 271–281 (2017).

25. Setty, M. *et al.* Characterization of cell fate probabilities in single-cell data with Palantir. *Nat. Biotechnol.* **37**, 451–460 (2019).

26. Platanitis, E. *et al.* A molecular switch from STAT2-IRF9 to ISGF3 underlies interferon-induced gene transcription. *Nat. Commun.* **10**, 2921 (2019).

27. Munger, J. *et al.* Systems-level metabolic flux profiling identifies fatty acid synthesis as a target for antiviral therapy. *Nat. Biotechnol.* **26**, 1179–1186 (2008).

28. Tirosh, O. *et al.* The Transcription and Translation Landscapes during Human Cytomegalovirus Infection Reveal Novel Host-Pathogen Interactions. *PLoS Pathog.* **11**, e1005288 (2015).

29. Patel, A. A., Ginhoux, F. & Yona, S. Monocytes, macrophages, dendritic cells and neutrophils: an update on lifespan kinetics in health and disease. *Immunology* **163**, 250–261 (2021).

30. Dağ, F. *et al.* Reversible silencing of cytomegalovirus genomes by type I interferon governs

virus latency. *PLoS Pathog.* **10**, e1003962 (2014).

31. Sinclair, J. Chromatin structure regulates human cytomegalovirus gene expression during latency, reactivation and lytic infection. *Biochim. Biophys. Acta* **1799**, 286–295 (2010).

32. Yee, L.-F., Lin, P. L. & Stinski, M. F. Ectopic expression of HCMV IE72 and IE86 proteins is sufficient to induce early gene expression but not production of infectious virus in undifferentiated promonocytic THP-1 cells. *Virology* **363**, 174–188 (2007).

33. Drayman, N., Patel, P., Vistain, L. & Tay, S. HSV-1 single-cell analysis reveals the activation of anti-viral and developmental programs in distinct sub-populations. *Elife* **8**, (2019).

34. Wu, X. *et al.* Intrinsic Immunity Shapes Viral Resistance of Stem Cells. *Cell* **172**, 423–438.e25 (2018).

35. Griffiths, P. D. Cytomegalovirus in intensive care. *Rev. Med. Virol.* **20**, 1–3 (2010).

36. Goodrum, F. Human Cytomegalovirus Latency: Approaching the Gordian Knot. *Annu Rev Virol* **3**, 333–357 (2016).

37. Guilliams, M., Thierry, G. R., Bonnardel, J. & Bajenoff, M. Establishment and Maintenance of the Macrophage Niche. *Immunity* **52**, 434–451 (2020).

38. Ginhoux, F. & Jung, S. Monocytes and macrophages: developmental pathways and tissue homeostasis. *Nat. Rev. Immunol.* **14**, 392–404 (2014).

39. Poole, E. *et al.* Alveolar Macrophages Isolated Directly From Human Cytomegalovirus (HCMV)-Seropositive Individuals Are Sites of HCMV Reactivation In Vivo. *J. Infect. Dis.* **211**, 1936–1942 (2015).

40. Pollock, J. L., Presti, R. M., Paetzold, S. & Virgin, H. W., 4th. Latent murine cytomegalovirus infection in macrophages. *Virology* **227**, 168–179 (1997).

41. Baasch, S. *et al.* Cytomegalovirus subverts macrophage identity. *Cell* **184**, 3774–3793.e25 (2021).

838 42. Elbasani, E., Gabaev, I., Steinbrück, L., Messerle, M. & Borst, E. M. Analysis of essential viral
839 gene functions after highly efficient adenofection of cells with cloned human
840 cytomegalovirus genomes. *Viruses* **6**, 354–370 (2014).
841 43. Jackson, S. E. *et al.* Latent Cytomegalovirus (CMV) Infection Does Not Detrimentally Alter T
842 Cell Responses in the Healthy Old, But Increased Latent CMV Carriage Is Related to
843 Expanded CMV-Specific T Cells. *Frontiers in Immunology* **8**, (2017).
844 44. Keren-Shaul, H. *et al.* MARS-seq2.0: an experimental and analytical pipeline for indexed
845 sorting combined with single-cell RNA sequencing v1. *Nature Protocols* **14**, 1841–1862
846 (2019).
847 45. Jaitin, D. A. *et al.* Massively parallel single-cell RNA-seq for marker-free decomposition of
848 tissues into cell types. *Science* **343**, 776–779 (2014).
849 46. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**,
850 357–359 (2012).
851 47. Shishkin, A. A. *et al.* Simultaneous generation of many RNA-seq libraries in a single reaction.
852 *Nature Methods* vol. 12 323–325 Preprint at https://doi.org/10.1038/nmeth.3313 (2015).
853 48. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient
854 alignment of short DNA sequences to the human genome. *Genome Biology* vol. 10 R25
855 Preprint at https://doi.org/10.1186/gb-2009-10-3-r25 (2009).
856 49. Bagnoli, J. W. *et al.* Sensitive and powerful single-cell RNA sequencing using mcSCRB-seq.
857 *Nat. Commun.* **9**, 2937 (2018).
858 50. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
859 51. van den Ent, F. & Löwe, J. RF cloning: A restriction-free method for inserting target genes
860 into plasmids. *Journal of Biochemical and Biophysical Methods* vol. 67 67–74 Preprint at
861 https://doi.org/10.1016/j.jbbm.2005.12.008 (2006).

52. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).

53. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

54. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15545–15550 (2005).

55. Liberzon, A. *et al.* The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst* **1**, 417–425 (2015).

figure 1

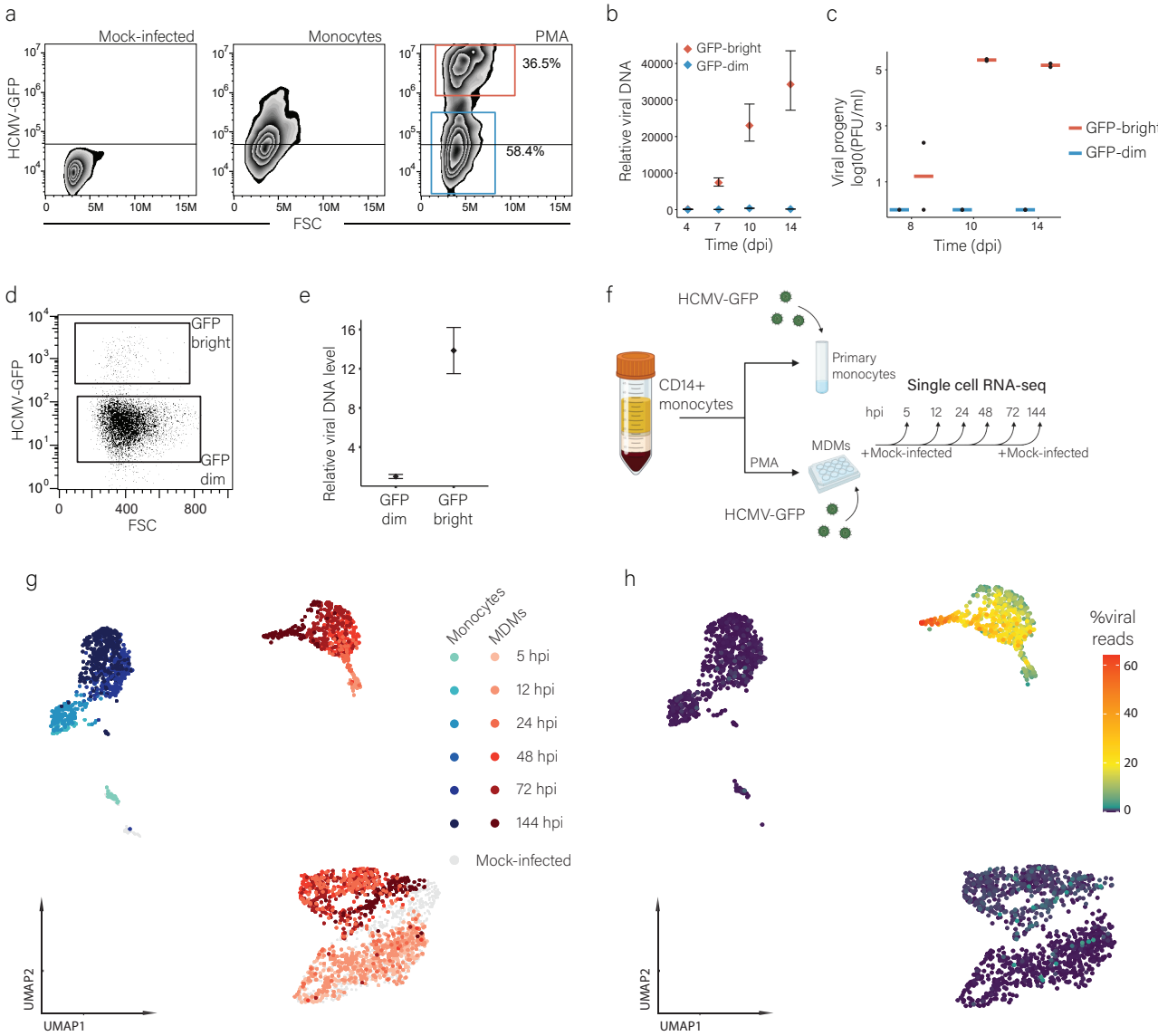

a
Mock-infected
Monocytes
PMA
HCMV-GFP
FSC
36.5%
58.4%
b
GFP-bright
GFP-dim
Relative viral DNA
Time (dpi)
c
Viral progeny log10(PFU/ml)
GFP-bright
GFP-dim
Time (dpi)
d
GFP bright
GFP dim
HCMV-GFP
FSC
e
Relative viral DNA level
GFP dim
GFP bright
f
HCMV-GFP
CD14+ monocytes
Primary monocytes
Single cell RNA-seq
hpi 5 12 24 48 72 144
MDMs
+Mock-infected
+Mock-infected
PMA
HCMV-GFP
g
Monocytes
MDMs
5 hpi
12 hpi
24 hpi
48 hpi
72 hpi
144 hpi
Mock-infected
UMAP2
UMAP1
h
%viral reads
UMAP2
UMAP1

figure 2

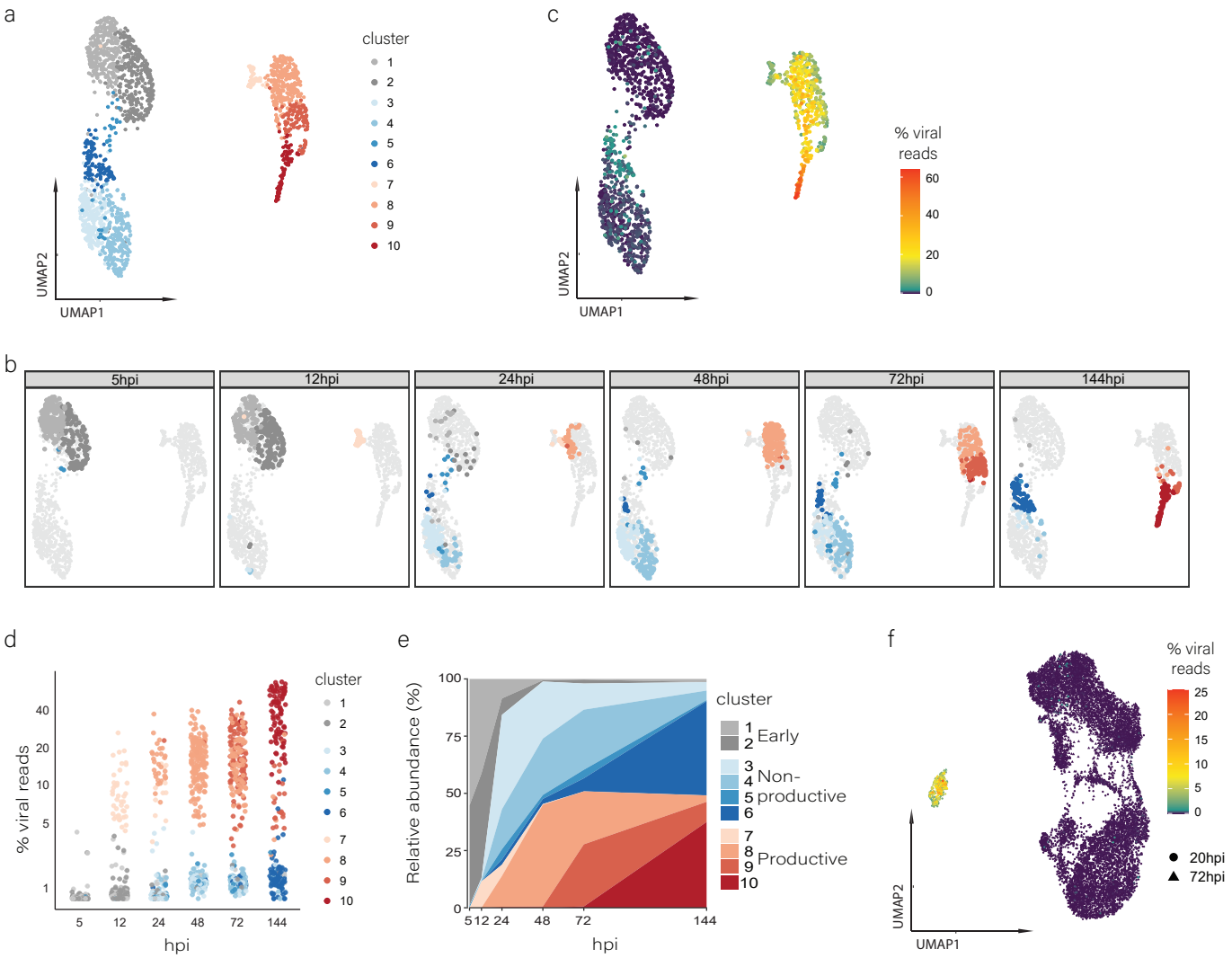

a
cluster
1
2
3
4
5
6
7
8
9
10
UMAP2
UMAP1
c
% viral reads
60
40
20
0
b
5hpi
12hpi
24hpi
48hpi
72hpi
144hpi
d
% viral reads
hpi
e
Relative abundance (%)
Early
Non-productive
Productive
f
20hpi
72hpi

figure 3

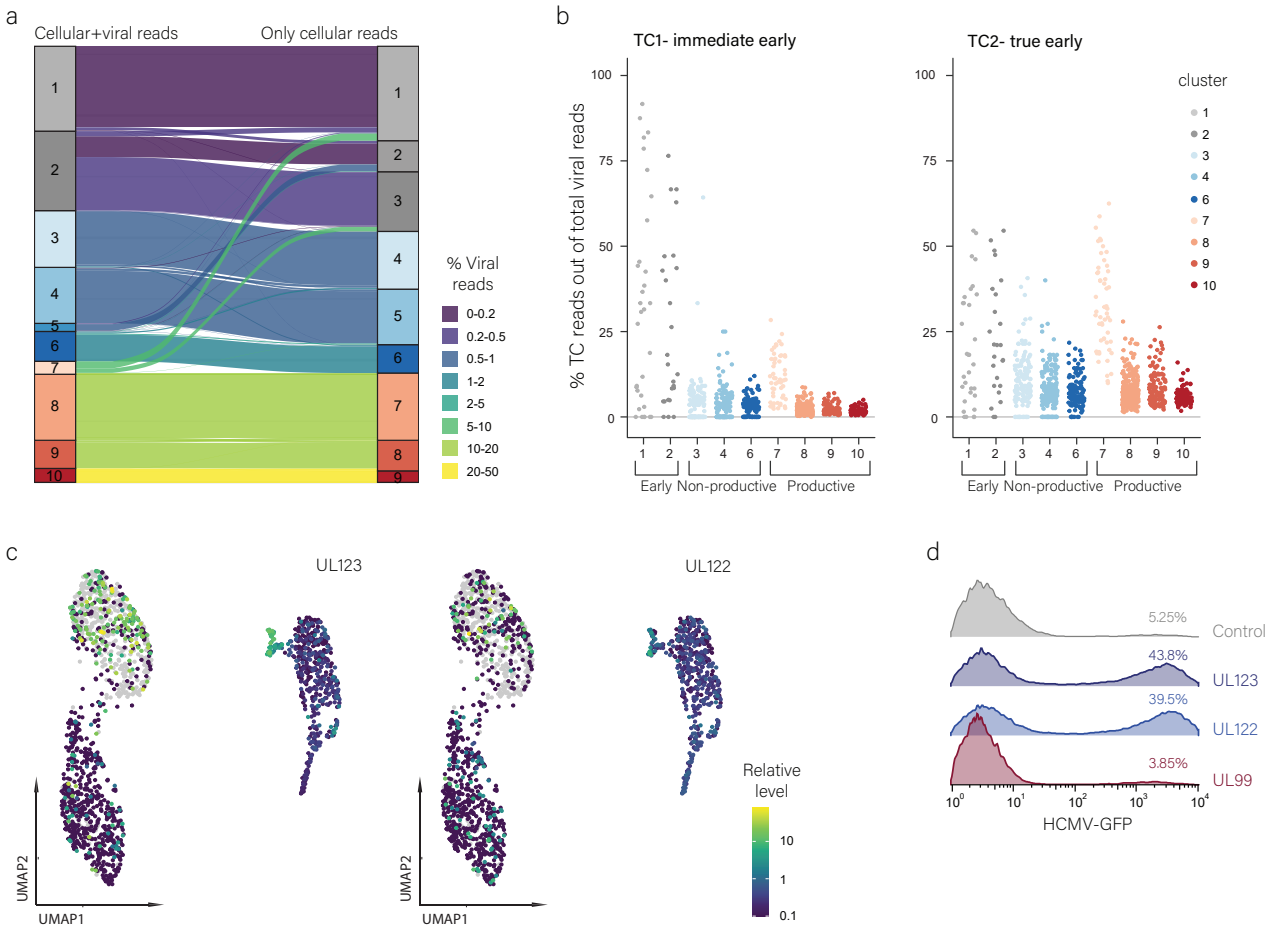

a
Cellular+viral reads
Only cellular reads
% Viral reads
0-0.2
0.2-0.5
0.5-1
1-2
2-5
5-10
10-20
20-50
b
TC1- immediate early
TC2- true early
% TC reads out of total viral reads
cluster
Early
Non-productive
Productive
c
UL123
UL122
UMAP1
UMAP2
Relative level
d
5.25%
Control
43.8%
UL123
39.5%
UL122
3.85%
UL99
HCMV-GFP

figure 4

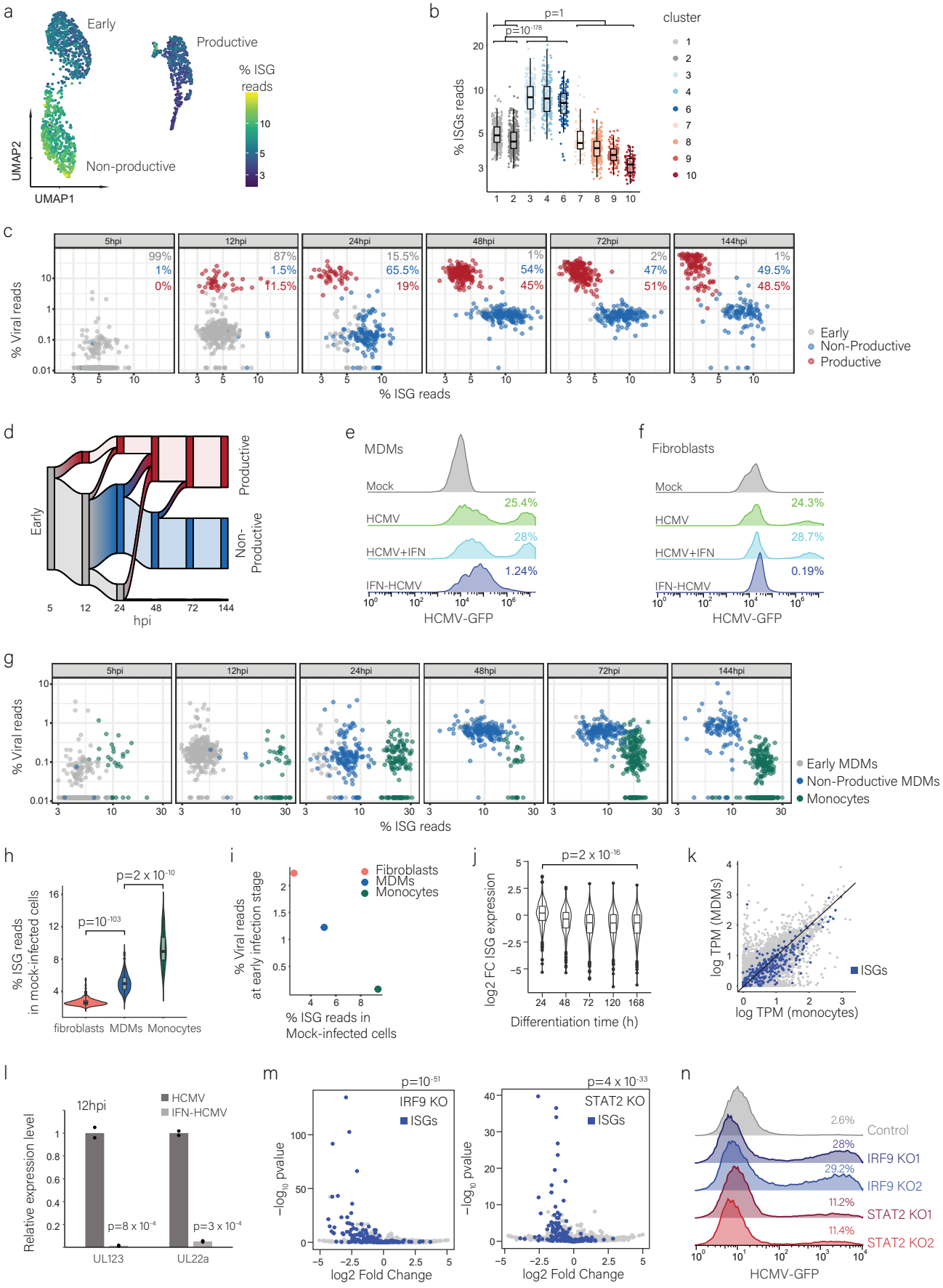
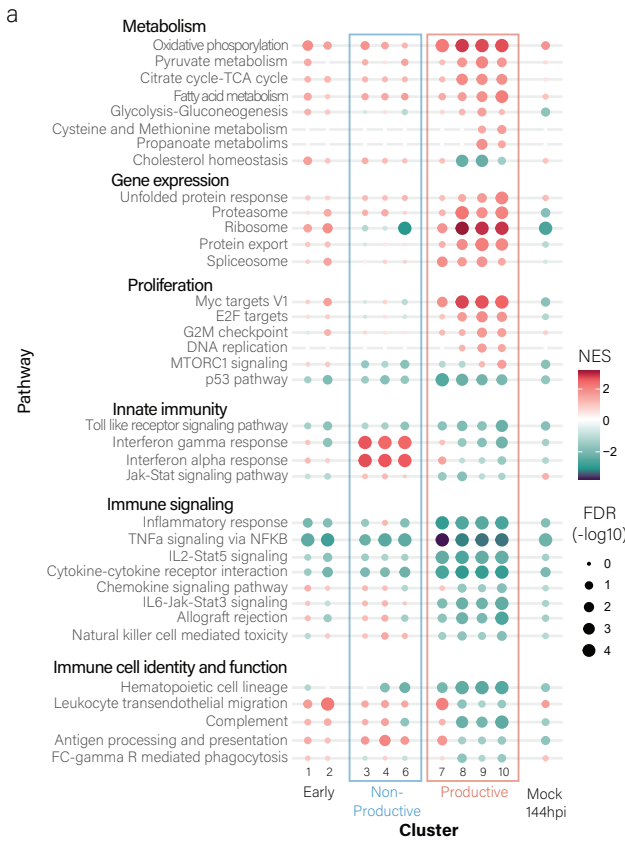
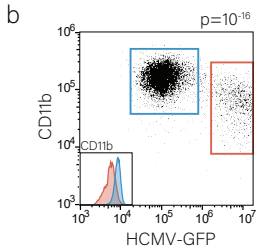

a
Early
Productive
Non-productive
% ISG reads
UMAP2
UMAP1
b
p=1
p=10-178
% ISGs reads
cluster
c
5hpi
12hpi
24hpi
48hpi
72hpi
144hpi
% Viral reads
% ISG reads
Early
Non-Productive
Productive
d
Early
Productive
Non-Productive
hpi
e
MDMs
Mock
HCMV
25.4%
HCMV+IFN
28%
IFN-HCMV
1.24%
HCMV-GFP
f
Fibroblasts
Mock
HCMV
24.3%
HCMV+IFN
28.7%
IFN-HCMV
0.19%
HCMV-GFP
g
% Viral reads
% ISG reads
Early MDMs
Non-Productive MDMs
Monocytes
h
% ISG reads in mock-infected cells
p=2 x 10-10
p=10-103
fibroblasts
MDMs
Monocytes
i
% Viral reads at early infection stage
Fibroblasts
MDMs
Monocytes
% ISG reads in Mock-infected cells
j
p=2 x 10-16
log2 FC ISG expression
Differentiation time (h)
k
log TPM (MDMs)
log TPM (monocytes)
ISGs
l
12hpi
HCMV
IFN-HCMV
Relative expression level
p=8 x 10-4
p=3 x 10-4
UL123
UL22a
m
p=10-51
IRF9 KO
ISGs
−log10 pvalue
log2 Fold Change
p=4 x 10-33
STAT2 KO
ISGs
n
2.6%
Control
28%
IRF9 KO1
29.2%
IRF9 KO2
11.2%
STAT2 KO1
11.4%
STAT2 KO2
HCMV-GFP

figure 5


a
Metabolism
Oxidative phosporylation
Pyruvate metabolism
Citrate cycle-TCA cycle
Fatty acid metabolism
Glycolysis-Gluconeogenesis
Cysteine and Methionine metabolism
Propanoate metabolims
Cholesterol homeostasis
Gene expression
Unfolded protein response
Proteasome
Ribosome
Protein export
Spliceosome
Proliferation
Myc targets V1
E2F targets
G2M checkpoint
DNA replication
MTORC1 signaling
p53 pathway
Innate immunity
Toll like receptor signaling pathway
Interferon gamma response
Interferon alpha response
Jak-Stat signaling pathway
Immune signaling
Inflammatory response
TNFa signaling via NFKB
IL2-Stat5 signaling
Cytokine-cytokine receptor interaction
Chemokine signaling pathway
IL6-Jak-Stat3 signaling
Allograft rejection
Natural killer cell mediated toxicity
Immune cell identity and function
Hematopoietic cell lineage
Leukocyte transendothelial migration
Complement
Antigen processing and presentation
FC-gamma R mediated phagocytosis
Pathway
1 2
3 4 6
7 8 9 10
Early
Non-Productive
Productive
Mock 144hpi
Cluster
NES
2
0
−2
FDR (-log10)
0
1
2
3
4


b
p=10-16
CD11b
CD11b
HCMV-GFP

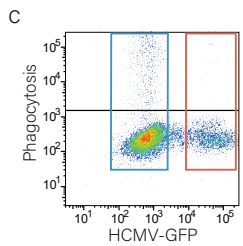
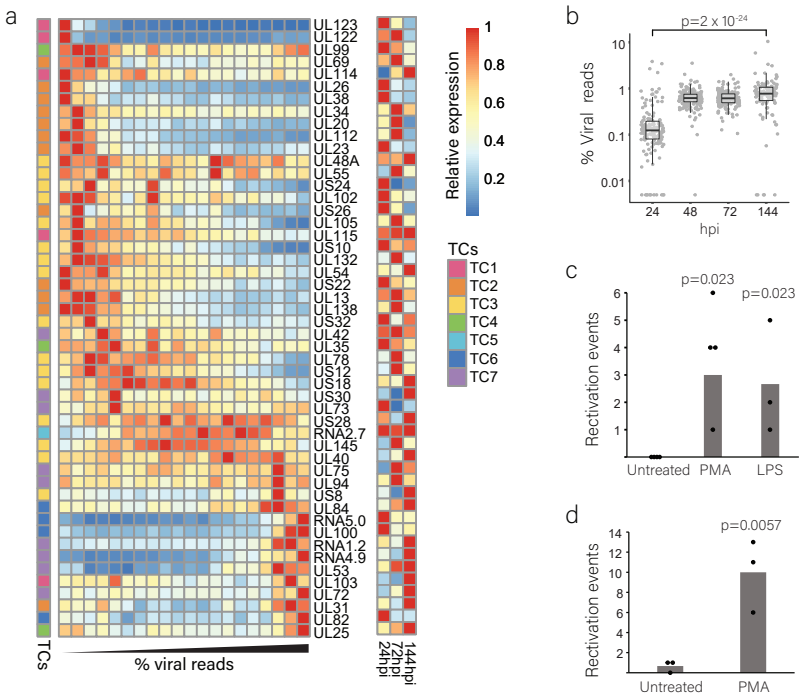

c
Phagocytosis
HCMV-GFP

figure 6

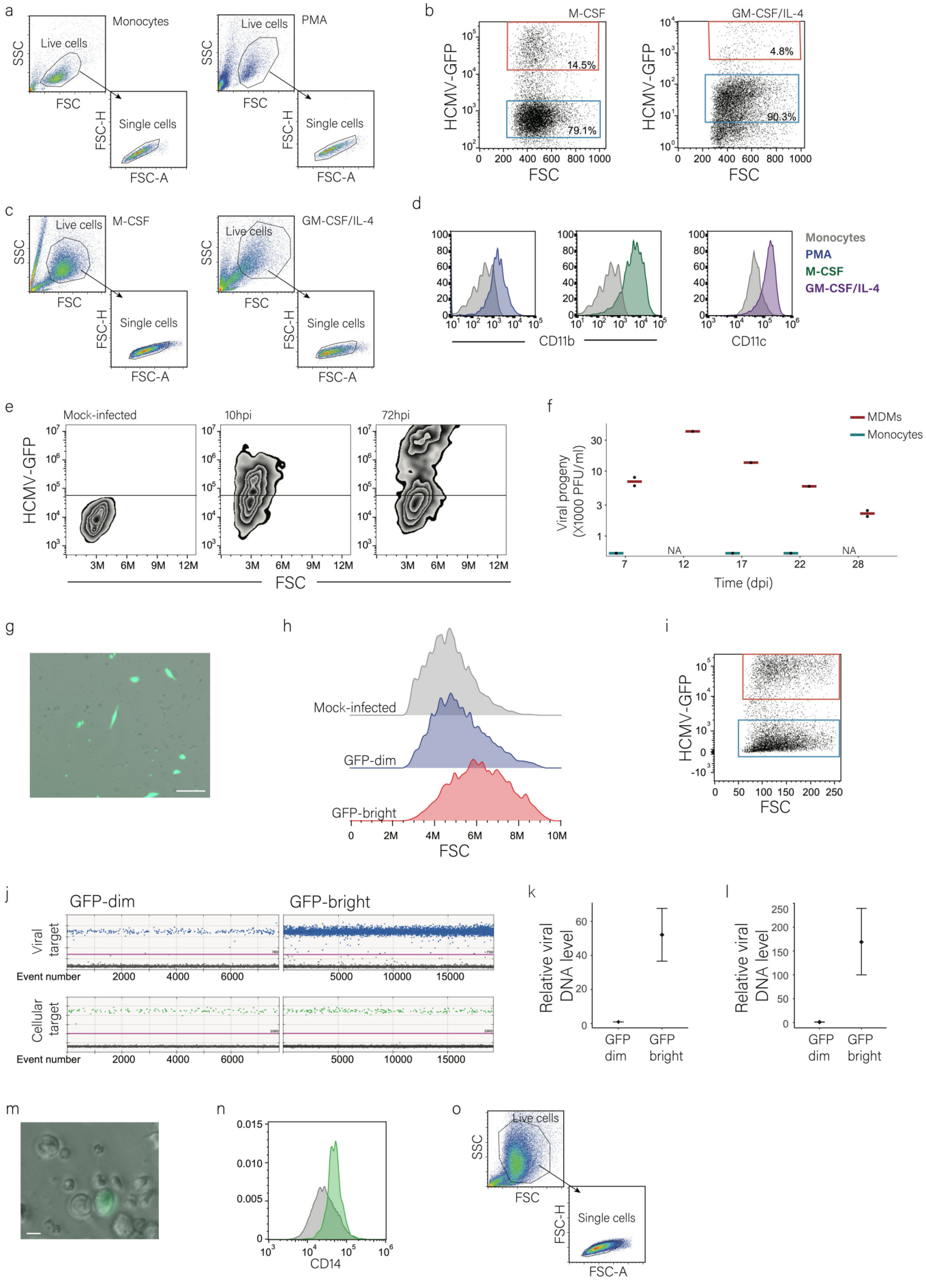

a
UL123
UL122
UL99
UL69
UL114
UL26
UL38
UL34
UL20
UL112
UL23
UL48A
UL55
US24
UL102
US26
UL105
UL115
US10
UL132
UL54
US22
UL13
UL138
US32
UL42
UL35
UL78
US12
US18
US30
UL73
US28
RNA2.7
UL145
UL40
UL75
UL94
US8
UL84
RNA5.0
UL100
RNA1.2
RNA4.9
UL53
UL103
UL72
UL31
UL82
UL25
TCs
% viral reads
144hpi
72hpi
24hpi
Relative expression
1
0.8
0.6
0.4
0.2
TCs
TC1
TC2
TC3
TC4
TC5
TC6
TC7

b
p=2 x 10-24
% Viral reads
10
1
0.1
0.01
24
48
72
144
hpi

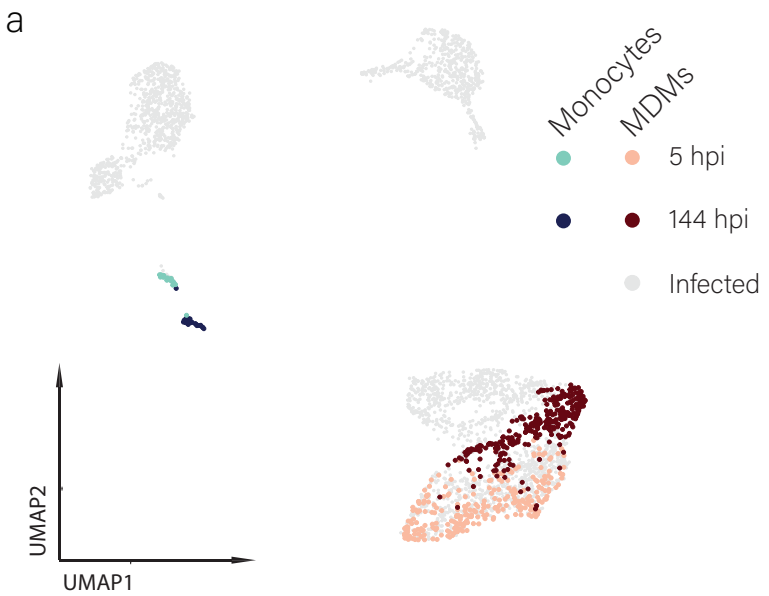
c
p=0.023
p=0.023
Rectivation events
6
5
4
3
2
1
Untreated
PMA
LPS

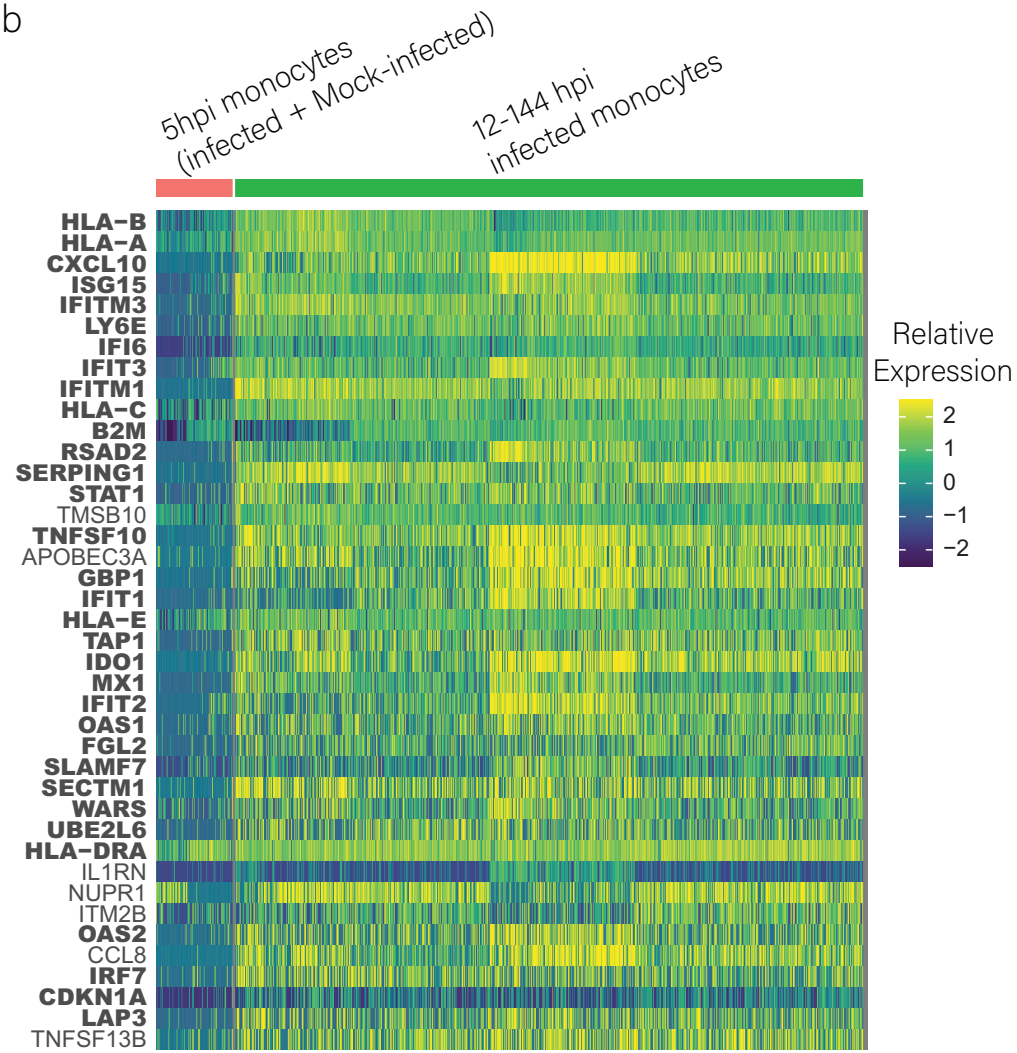
d
p=0.0057
Rectivation events
14
12
10
8
6
4
2
Untreated
PMA

Extended Data figure 1

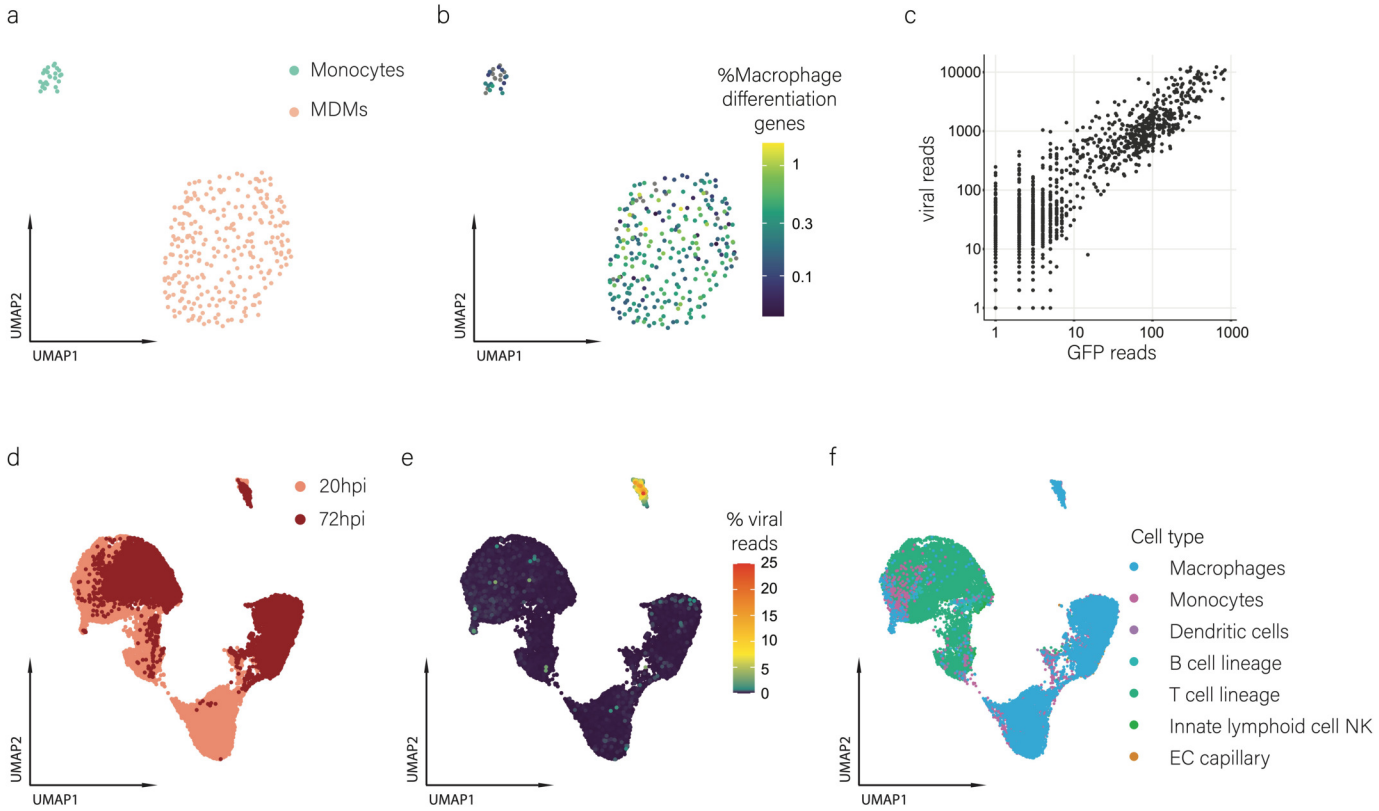


Extended Data figure 2

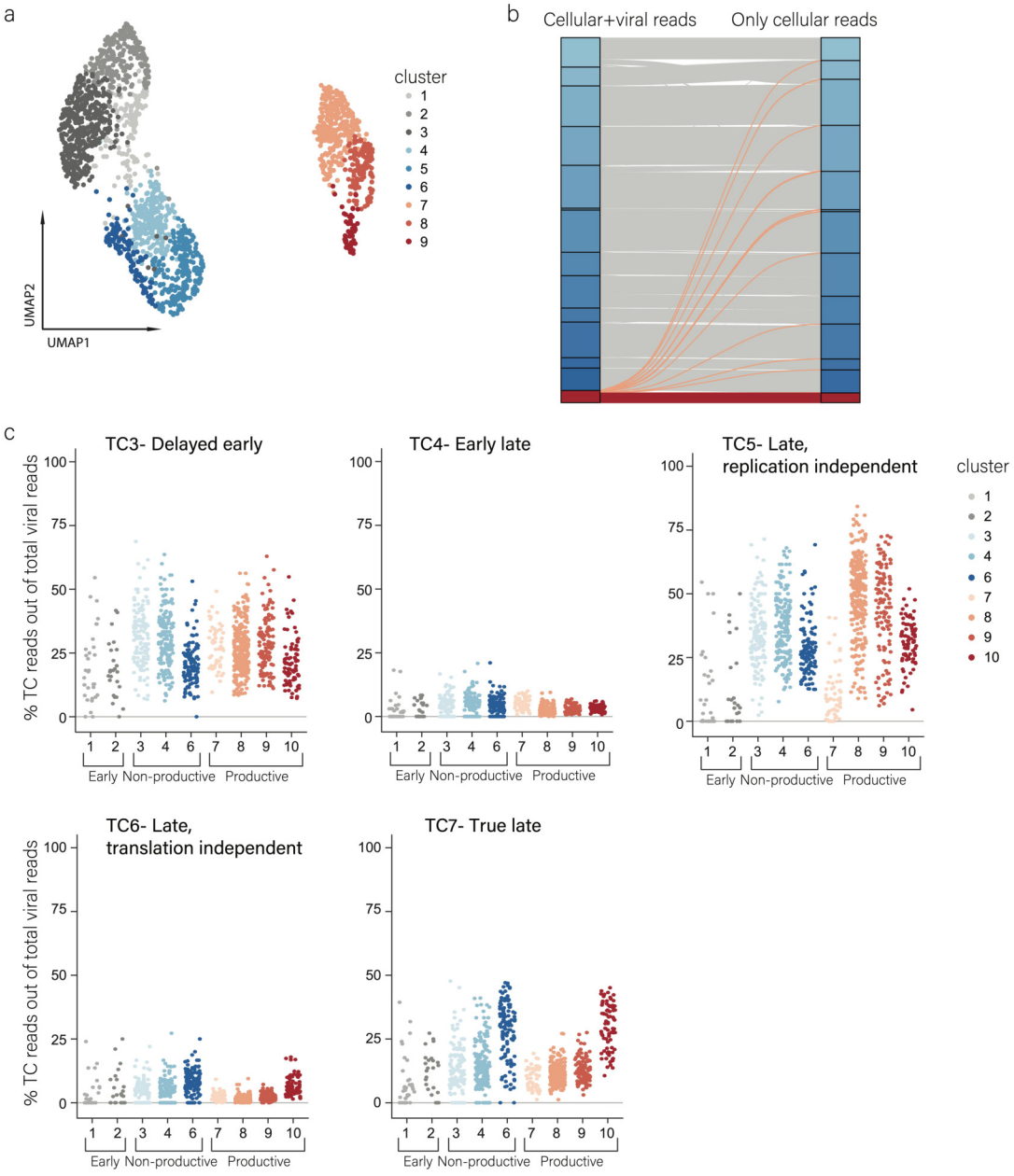




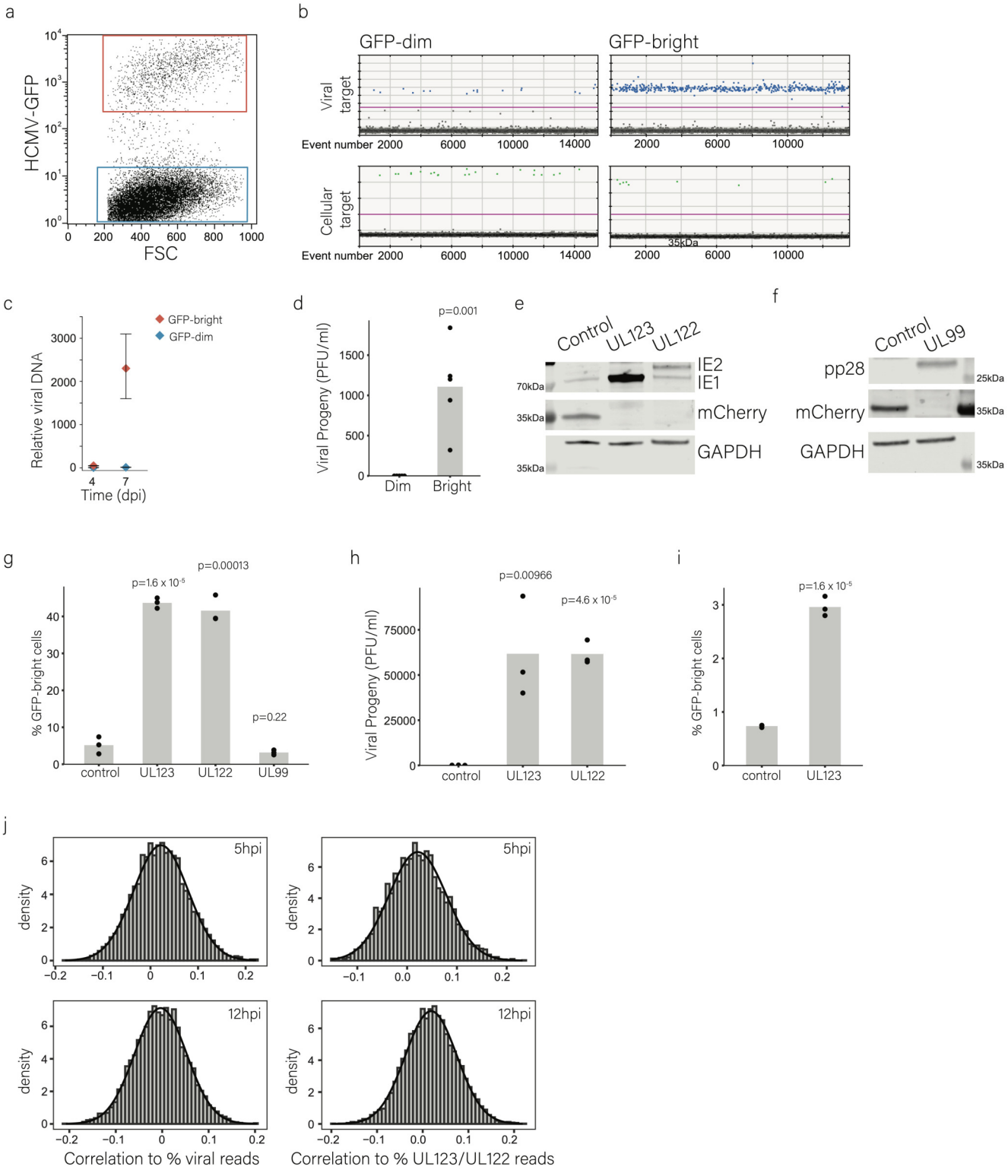
Extended Data figure 3



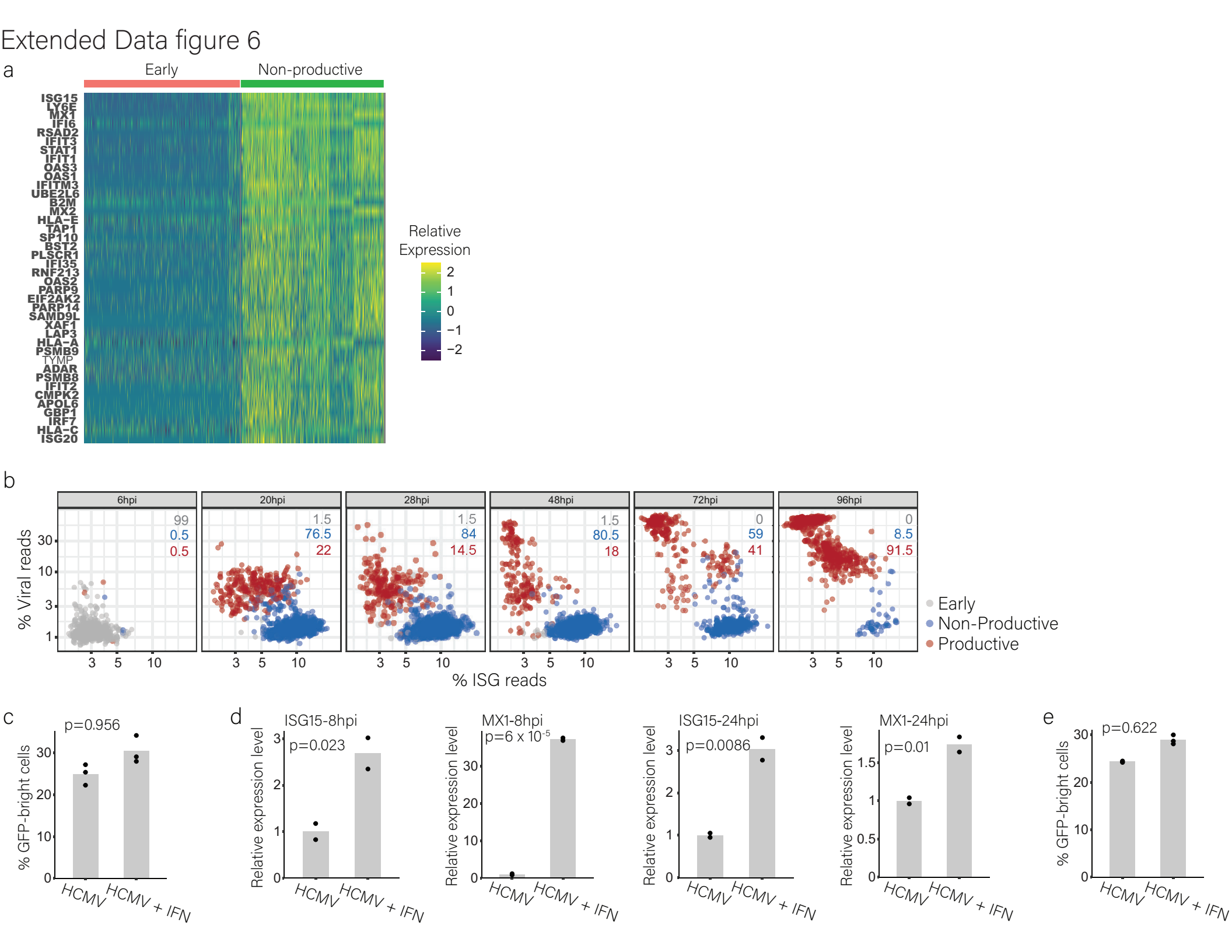
Extended Data figure 4



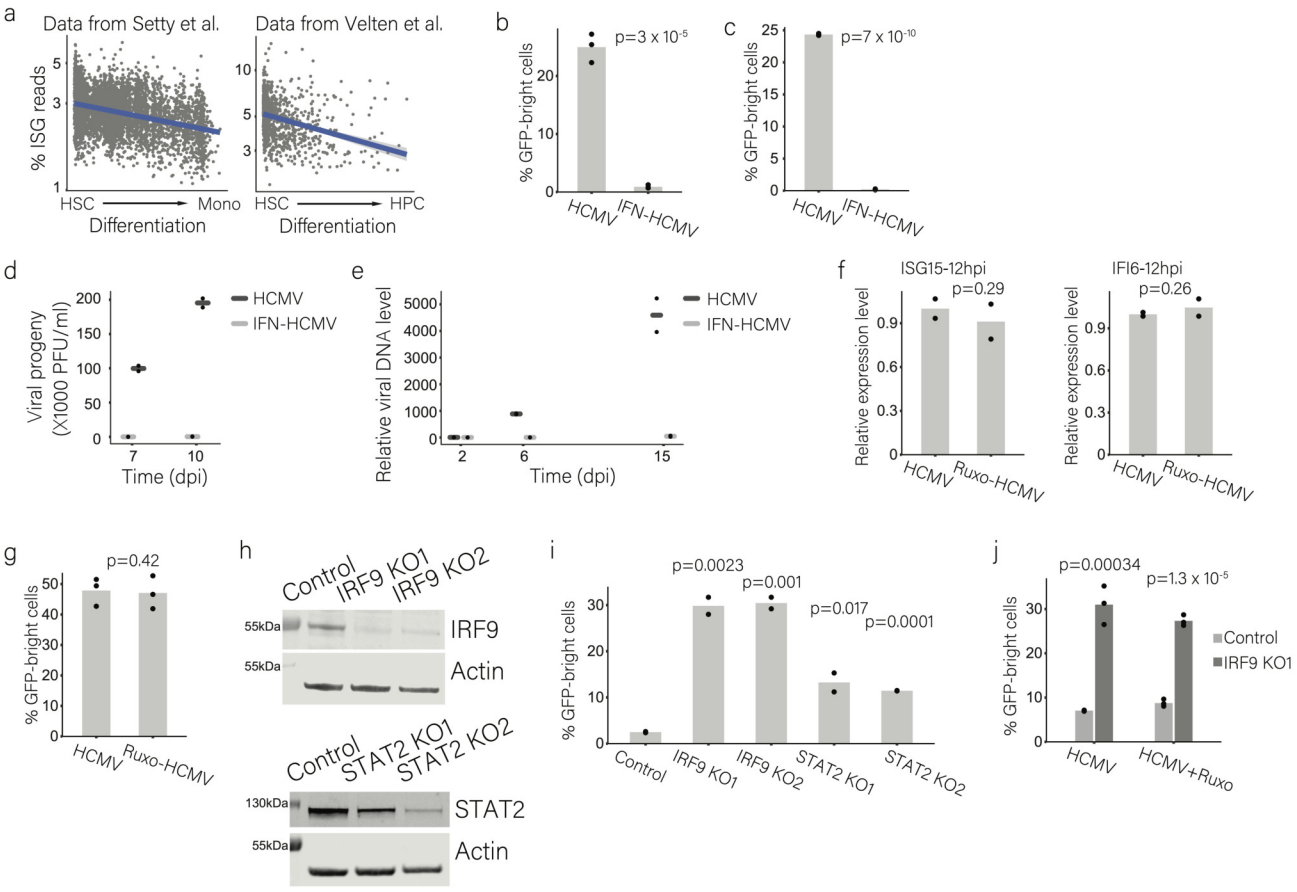
Extended Data figure 5



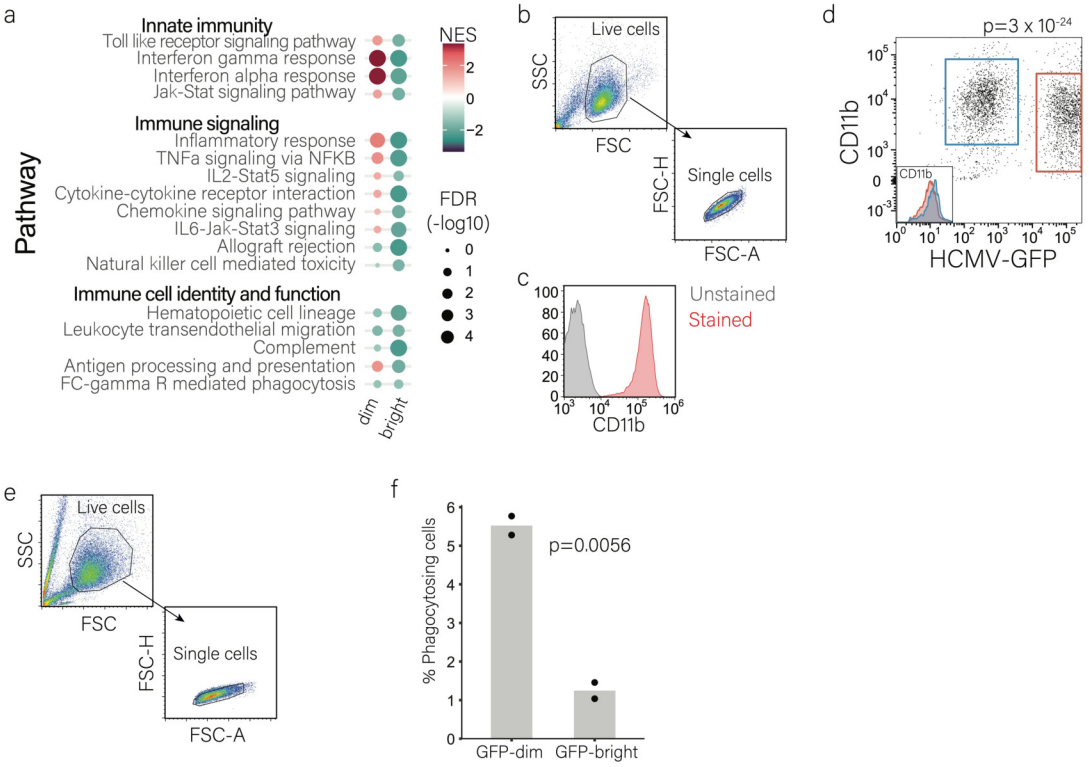
Extended Data figure 6



Extended Data figure 7



Extended Data figure 8



Extended Data figure 9