nature microbiology

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] ✉

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 expression level is a main determinant of infection outcome. Intrinsic interferon-stimulated gene 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.

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

A wide range of cell types can be infected[2], but with variable infection outcomes, and differentiated cells are generally more permissive. Cells from the myeloid lineage are thought to play a critical role in HCMV latency and reactivation. Haematopoietic stem cells (HSCs), granulocyte–macrophage progenitor 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].

Although bone marrow HSCs have been proposed as the long-term HCMV latent reservoir, clinical data strongly imply that there are additional uncharacterized sources of latent HCMV outside the bone marrow, probably 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 outcomes 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 using single-cell RNA sequencing (RNA-seq).

[1]Department of Molecular Genetics, Weizmann Institute of Science, Rehovot, Israel. [2]Institute of Pulmonary Medicine, Hadassah Hebrew University Medical Center, 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: Michal Schwartz, Miri Shnayder. ✉e-mail: michalsc@weizmann.ac.il; noam.stern-ginossar@weizmann.ac.il
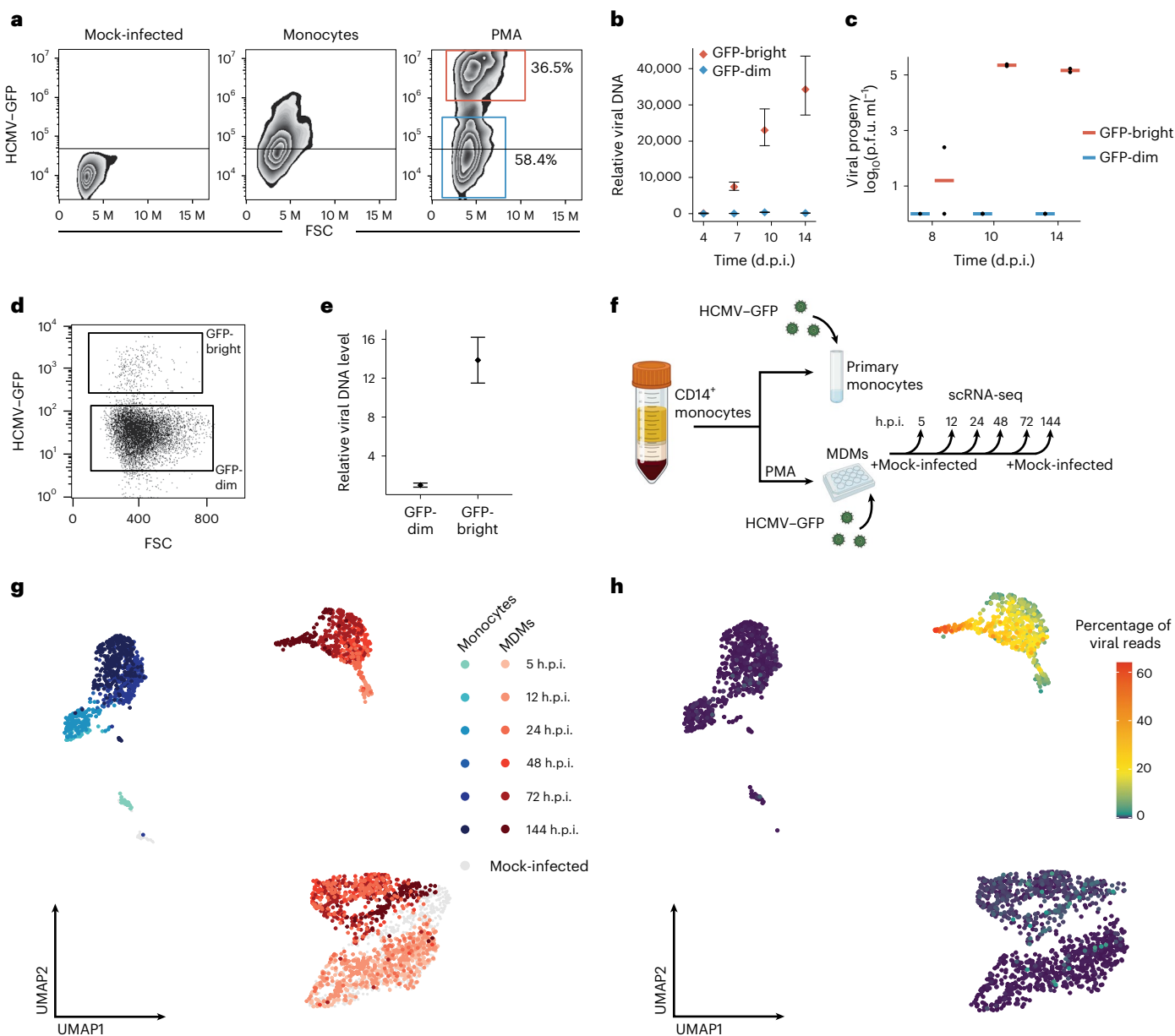
Fig. 1 | Experimental system for studying HCMV infection outcome. a, Flow cytometry analysis of primary CD14+ monocytes and monocytes differentiated with PMA, infected with HCMV–GFP. Analysis was performed at 3 d.p.i. The mock-infected GFP level is marked by a line. Blue and red gates mark the GFP-dim and GFP-bright populations, respectively, and their percentages are noted. FSC, forward scatter; M, million. **b,** droplet-digital PCR (ddPCR) measurements of viral genomes from FACS-sorted GFP-bright and GFP-dim HCMV–GFP-infected MDMs at different times post infection. The graph is representative of two biological replicates and shows the mean and 95% confidence interval 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, p.f.u., plaque forming units. **d,** Flow cytometry analysis of BAL macrophages infected with HCMV–GFP at 3 d.p.i. **e,** ddPCR measurements of viral genomes from FACS-sorted GFP-bright and GFP-dim HCMV–GFP-infected BAL macrophages at 3 d.p.i. Graph is representative of two biological replicates and shows the mean and 95% confidence interval of Poisson distribution of >10,000 PCR droplets. **f,** Experimental design. Primary monocytes and PMA-induced MDMs infected with HCMV–GFP were processed for scRNA-seq at 5, 12, 24, 48, 72 and 144 h.p.i. In parallel, mock-infected cells from 5 and 144 h.p.i. were also processed for scRNA-seq. **g,** Projection of HCMV–GFP infected CD14+ monocytes (*n* = 803) and MDMs (*n* = 1,815) as well as mock-infected monocytes (*n* = 68) and MDMs (*n* = 540) from the indicated time points, based on cellular and viral gene expression, colours indicate cell types and time post infection, mock-infected cells are plotted in grey. **h,** Projection as in **g,** showing all infected cells, coloured by percentage of viral reads per cell. Panel **f** created with BioRender.com.

First, we reveal that macrophages can support both productive and non-productive infection. By comparing these two infection outcomes, as well as infected monocytes, we show that high early viral gene expression levels, particularly of immediate early (IE) 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 with 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 can serve as latent viral tissue reservoirs.

Results

Experimental system for studying HCMV infection outcome

We initially compared infection of monocytes with differentiated monocytes using three treatments: phorbol myristate acetate (PMA), granulocyte-macrophage colony-stimulating factor and interleukin 4 (GM-CSF/IL-4) and macrophage colony stimulating factor (M-CSF), which have all been shown to trigger monocyte differentiation[10,11]. We used an HCMV TB40E strain containing a green fluorescent protein (GFP) reporter under a simian-virus 40 (SV40) promoter[12], which allows quantification of productive infection efficiency. As expected, 3 days post infection (d.p.i.) in monocytes, almost all of the cells show a small increase in GFP, indicating that they are infected but 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 in the early stages of infection (Extended Data Fig. 1e), suggesting that they probably do not reflect two separate initial populations, but rather two different infection outcomes. Infection of MDMs, but not of monocytes, resulted in the 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 than 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) harboured viral DNA; however, at 4 d.p.i. viral DNA load was much higher in the GFP-bright cells (Extended Data Fig. 1j–l) and increased substantially with time in the GFP-bright, but not 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, whereas 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 analysed infected monocytes and PMA-induced MDMs from the same donor at 5, 12, 24, 48, 72 and 144 hours post infection (h.p.i.), as well as mock-infected monocytes and MDMs at 5 and 144 h.p.i., using 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 and Extended Data Fig. 2a). At 5 h.p.i., infected monocytes were grouped with mock-infected monocytes (Fig. 1g and Extended Data Fig. 2a). Infected monocytes from 12 h.p.i. onwards appeared as a single population (Fig.1g) and were stratified from early monocytes mainly based on ISG expression (Extended Data Fig. 2b). As expected, viral transcript load in infected monocytes remained low in all cells throughout infection (median = 0.063%; Fig. 1h), indicating lack of productive infection[13–15]. By contrast, infected MDMs from

24 h.p.i. 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 and Extended Data Fig. 2a), suggesting milder transcriptional changes relative to later in infection.

Two distinct infection outcomes in infected macrophages

Because 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 an 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 h.p.i.) belong to a single group, harbouring low levels of viral transcripts and are indistinguishable with regard to infection outcome, whereas cells from the later time points split into two additional distinct groups (Fig. 2b) that 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 subclustering corresponds to progression in the given trajectory with time post infection (Fig. 2b,e).

scRNA-seq analysis of infected BAL cells from 20 and 72 h.p.i. demonstrated that the cells stratify into two distinct populations and an additional small population that expresses substantially 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 MDMs into distinct groups (early, non-productive, productive), we next analysed the transcriptome solely according to cellular genes. This analysis produced the same three distinct groups of cells and very similar cluster stratification, indicating that 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, the earliest productive cluster (Fig. 2a), when clustering was based only on cellular genes, these cells clustered together with the early cells (Fig. 3a). This demonstrates 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 relative expression of the different temporal classes of viral genes[15] within cells in different clusters reveals that in the early productive cells, compared with their non-productive counterparts, there is high relative expression of IE genes, as well as specific induction of true early genes, but not of other viral temporal classes (Fig. 3b and 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 that 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
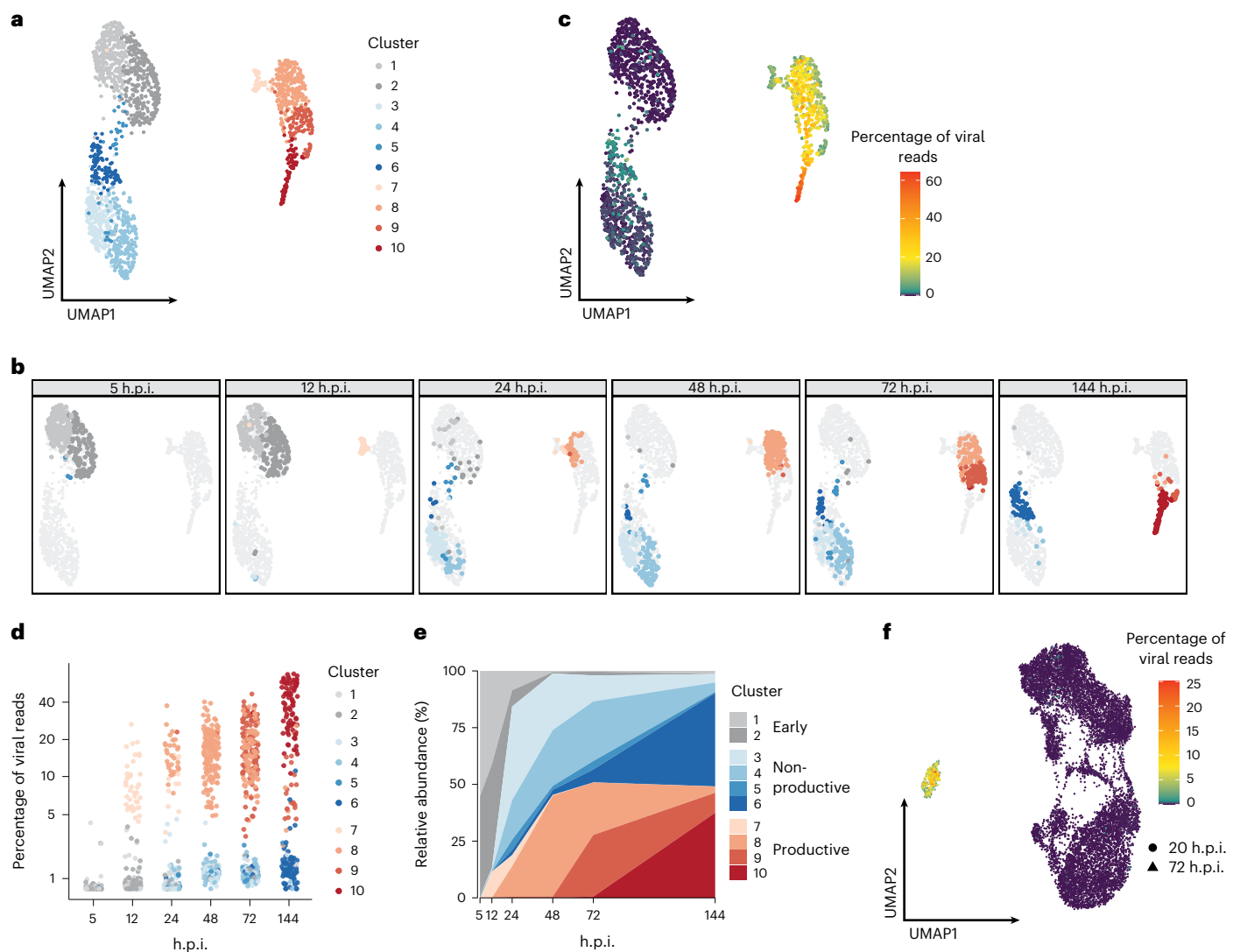
Fig. 2 | Single-cell transcriptome analysis of HCMV-infected macrophages. a, Projection of 1,815 HCMV–GFP-infected MDMs at 5–144 h.p.i. coloured by cluster assignment. Analysis was based on both cellular and viral transcripts. **b**, Projections as in **a**, showing the cells from each time point during infection. **c**, Projection, as in **a**, coloured by percentage of viral reads per cell. **d**, Percentage of viral reads in single MDMs from the different clusters throughout infection. **e**, Distribution of the single cells between the different early, non-productive and productive clusters throughout infection. **f**, Projection of 10,320 HCMV–GFP-infected myeloid BAL cells from 20 and 72 h.p.i. coloured by percentage of viral reads per cell. Circles and triangles mark cells from 20 or 72 h.p.i., respectively. Analysis was based on both cellular and viral transcripts.

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 Data Fig. 5i).

Because scRNA-seq gives unprecedented detail on intercellular 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 probably 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 and 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 (for example, interferons (IFNs)), because they were grown together in the same culture dish, indicating that the induction of viral gene expression renders the cells unresponsive to IFN signalling[19]. Because 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
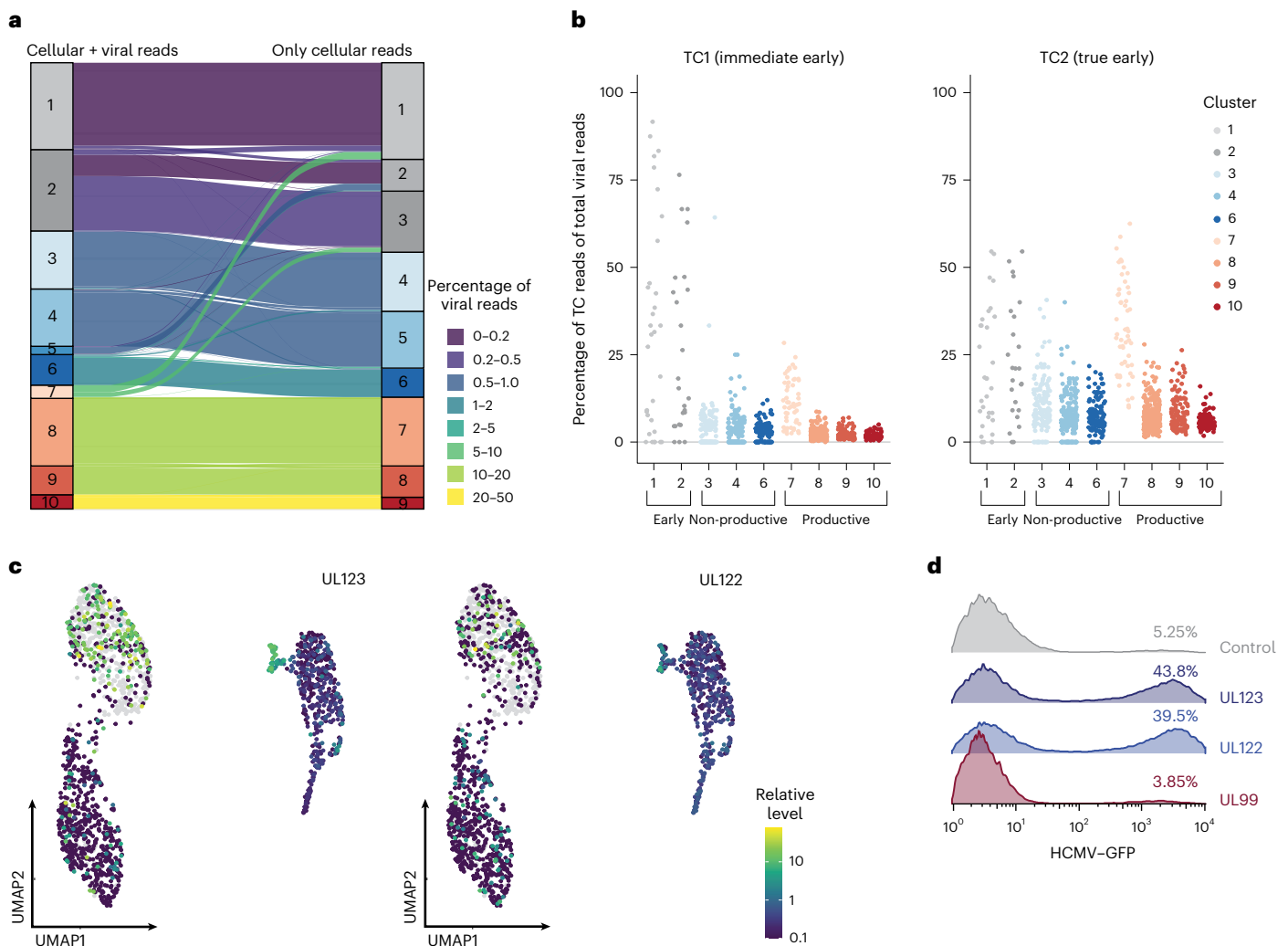
Fig. 3 | Viral IE gene expression is a major determinant of infection outcome.
a, Cluster assignment of each cell according to analysis based on both viral and cellular genes (left) or on analysis based solely on cellular genes (right). Each cell, represented by a single line, is coloured by the percentage of viral reads. **b**, Relative expression levels, within cells from the different clusters, of viral genes from temporal classes (TC) 1 (immediate early) and 2 (true early) based on

Rozman et al.[15]. **c**, Projection of infected MDMs (as in Fig. 2a) coloured 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 overexpressing mCherry as control, UL123, UL122 or UL99. Replicate data and statistics are shown in Extended Data Fig. 5g.

point indicates that productive cells probably also originate from the non-productive groups in which ISGs were already induced, because the percentage of productive cells reaches numbers that cannot be explained solely by transition from the early infected population (Fig. 4d). This therefore suggests that activation of the IFN 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 scRNA-seq dataset during 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 IFN at the time of infection did not affect the percentage of productively infected cells (Fig. 4e and Extended Data Fig. 6c), despite a significant increase in ISG expression (Extended Data Fig. 6d). Similar results were obtained in low multiplicity of infection (MOI) 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 with non-productive MDMs showed similarly low viral gene expression but revealed higher ISG expression level even at the earliest infection stage likely before IFN response initiates (Fig. 4g). Further analysis of mock-infected monocytes, MDMs and fibroblasts[20] revealed that intrinsic levels of ISGs (in mock-infected cells) in fibroblasts are lower than in MDMs, which in turn are 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). Because the permissiveness for productive infection is tightly linked to the differentiation state[21,22], we examined the link between intrinsic expression of ISGs and differentiation state. Indeed, intrinsic levels of many ISGs decrease with differentiation of monocytes to MDMs[23] (Fig. 4j,k) and with differentiation of haematopoietic cells in the bone marrow[24,25] (Extended Data Fig. 7a), suggesting a close link between intrinsic levels of ISGs, differentiation and permissiveness to productive HCMV infection.

To explore whether the levels of ISGs at the time of infection affect infection outcome we tested the effect of IFN pretreatment on HCMV infection. Addition of IFN 24 h before infection of MDMs or fibroblasts

eliminated productive infection (Fig. 4e,f and Extended Data Fig. 7b–d), blocked viral DNA replication (Extended Data Fig. 7e) and led to a reduction in viral gene expression at 12 h.p.i. (Fig. 4l). These results indicate that ISGs act to block the virus at the stage of IE 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 IFN signalling; however, pretreatment of MDMs with ruxolitinib, a potent and selective Janus kinase 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 signal transducer and activator of transcription 2–interferon regulatory factor 9 (STAT2–IRF9) complexes, independent of IFN receptor signalling[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 analysed 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 that are associated with immune signalling or common macrophage functions, such as antigen presentation and phagocytosis, as well as genes related to haematopoietic cell lineage (Fig. 5a). RNA-seq on infected BAL macrophages showed that similar pathways related to immune signalling and macrophage identity are indeed 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 a 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 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 during productive and non-productive infection. For the productive MDMs we binned the cells according to percentage of viral reads to reflect progression during infection. This was followed by hierarchical clustering, revealing a temporal expression cascade in productive MDMs that is similar to that seen in fibroblast infection[15] (Fig. 6a, left heatmap). This analysis reflects relative expression patterns; that is, 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 increase slightly 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 only GFP-dim cells were processed at 72 and 144 h.p.i. (Extended Data Fig. 9c), which further demonstrated a small increase in viral transcript levels (Extended Data Fig. 9c). Nevertheless, analysis of viral gene expression in non-productive cells throughout infection revealed broad gene expression with no clear kinetic programme (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 with monocytes indicates a less-repressed state in the former. Furthermore, in culture, these cells survived for relatively long periods (Extended Data Fig. 9d); in vivo macrophages can also be long-lived cells[29]. We therefore tested whether 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 lipopolysaccharide (LPS) at 7 d.p.i. resulted in reactivation and the release of infectious virus (Fig. 6c and Extended Data Fig. 9e,f). Moreover, IFN-pretreated, infected MDMs, which as described above are non-productively infected, also reactivated when PMA was added at 7 d.p.i., 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].

Fig. 4 | Intrinsic ISG levels, but not their induction, affect infection outcome. **a**, Projection of infected MDMs (as in Fig. 2a) coloured by the expression level of ISGs. **b**, Expression level of ISGs in single MDMs from the different clusters. The boxplot shows standard parameters (median, Q1–Q3, 1.5 interquartile range). P values calculated by one-tailed t test on the cells in a group of clusters as marked are indicated. See Methods for the 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 at the given time point is indicated. **d**, Flow chart describing the transition of cells between the early, non-productive and productive groups during infection. **e**,**f**, Flow cytometry analysis of HCMV–GFP-infected MDMs (**e**) or fibroblasts (**f**) either untreated, treated with IFNs at the time of infection (HCMV + IFN) or treated with IFNs 24 h before infection (IFN–HCMV). Mock-infected cells are shown as controls. Replicate data and statistics are shown in Extended Data Figs. 6c,e and 7b,c. **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 = 897 (fibroblasts), 258 (MDMs) and 25 (monocytes). P values calculated by one-tailed t test between cells of different types as marked are indicated. **i**, ISG expression level in mock-infected cells versus viral gene expression level at early stages of infection (12 or 20 h.p.i., in monocytes/MDMs and fibroblasts, respectively). **j**, Expression levels of different ISGs in cells during differentiation trajectory from monocytes to MDMs. Data taken from Song et al.[23]. Violin plot shows the median, Q1–Q3 (box), distribution of values (violin) and outliers. P value calculated by one-tailed t test comparing the cells in the first and last groups is indicated. n = 265 genes. **k**, Cellular transcript levels in monocytes and MDMs[23]. TPM, transcript reads per kilobase per million reads mapped to cellular transcripts. **l**, Relative expression level of the viral transcripts UL123 and UL22A, measured by quantitative polymerase chain reaction at 12 h.p.i., in infected MDMs either untreated (HCMV) or treated with IFNs 24 h before infection (IFN–HCMV). Points show measurement of biological replicates. P values calculated by one-tailed t test on the replicates are indicated. **m**, Differential expression analysis of cellular genes in uninfected THP1-derived macrophages following CRISPR KO of IRF9 (left) or STAT2 (right) compared with control CRISPR KO. P values of the enrichment of ISGs in the differentially expressed genes as calculated by one-tailed hypergeometric test are indicated. **n**, Flow cytometry analysis of HCMV–GFP-infected THP1-derived macrophages following CRISPR-mediated KO of a control gene, IRF9 or STAT2. Two independent gRNAs were used for IRF9 and STAT2. Percentage of productively infected cells is indicated for each condition. Replicate data and statistics are shown in Extended Data Fig. 7i.

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 IFNs may further lead to a similar latent-like infection that can also result in later reactivation.
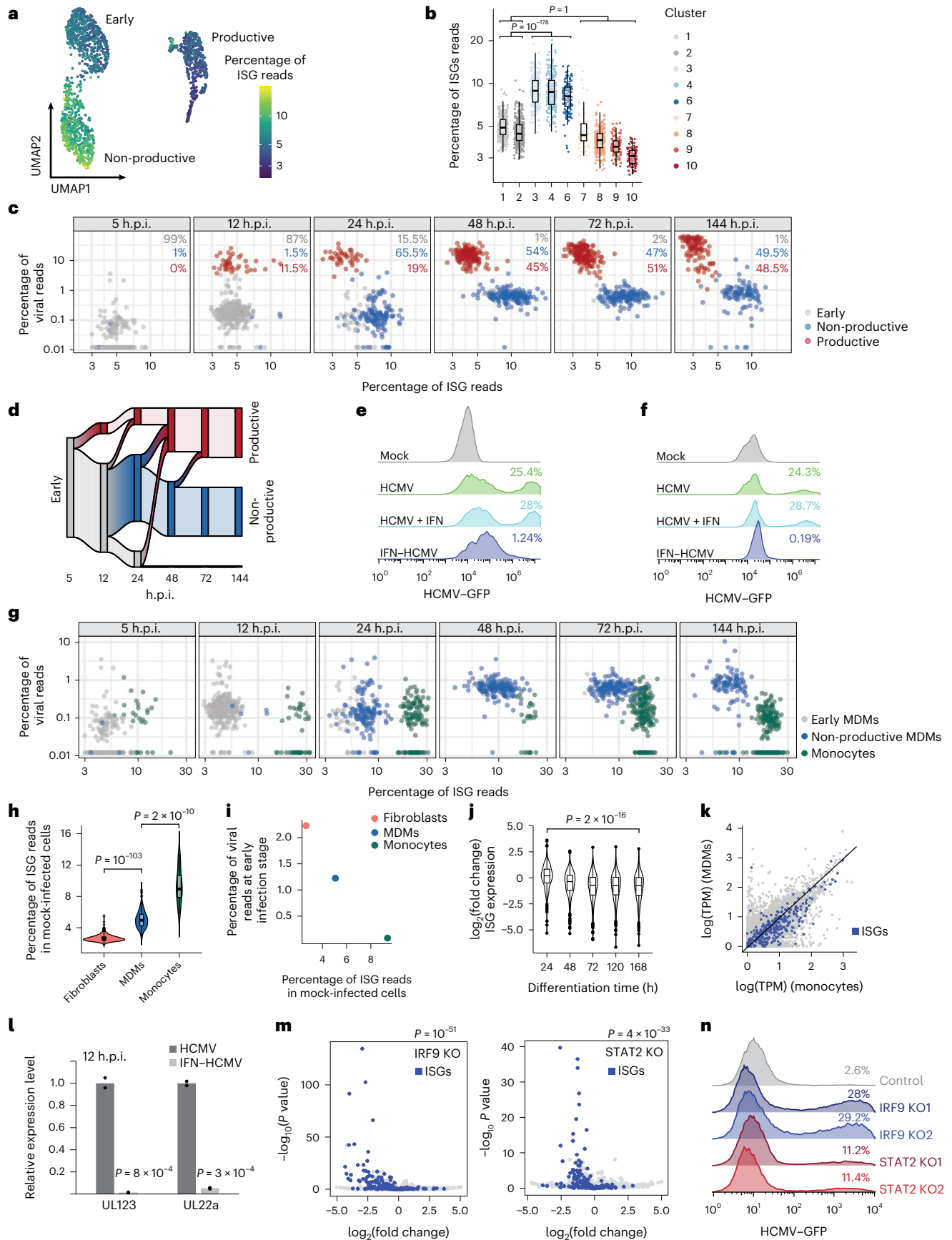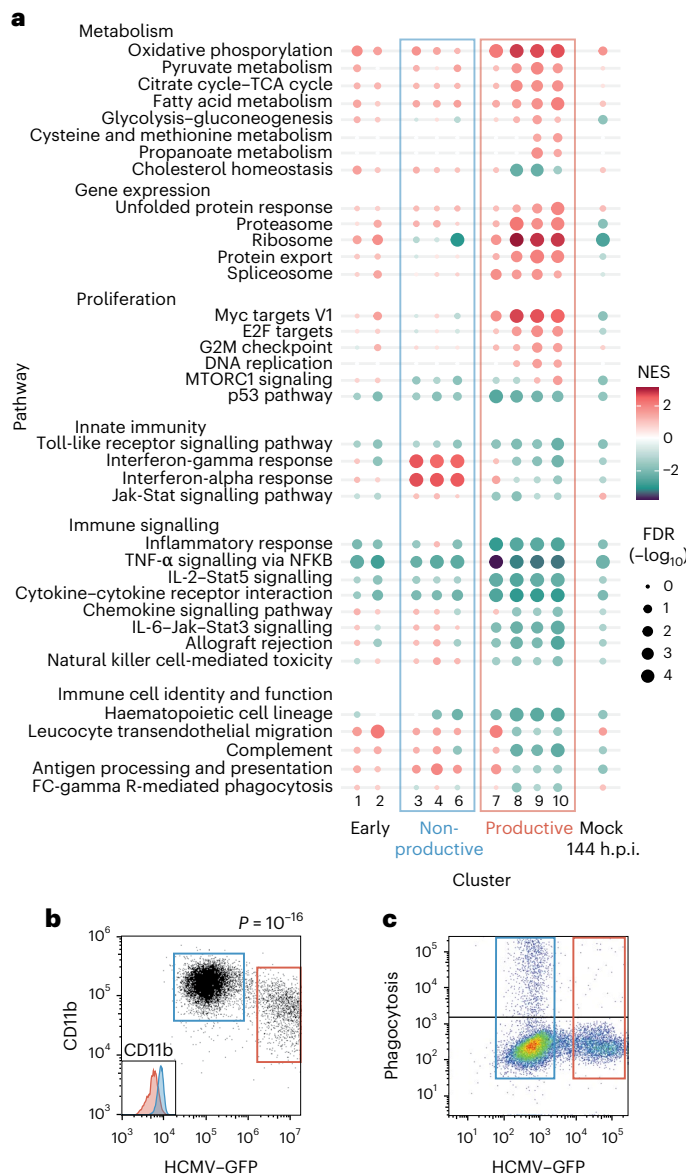
Fig. 5 | Effects of HCMV infection on cellular pathways and macrophage cell identity and function. a, Summary of gene set enrichment analysis of differentially expressed genes between each of the scRNA-seq clusters and mock-infected cells from 5 h.p.i. using annotated Kyoto Encyclopedia of Genes and Genomes and Hallmark pathways. Uninfected MDMs at 144 h.p.i. are shown as a control for changes occurring over time in culture, regardless of infection. FC, fragment crystallizable; FDR, false discovery rate; NES, normalized enrichment score; NFKB, nuclear factor kappa B; TCA, tricarboxylic acid; TNF, tumour necrosis factor. **b**, Flow cytometry of CD11b surface levels in M-CSF-induced MDMs infected with HCMV–GFP. 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 mean fluorescence intensity in GFP-bright MDMs on the measured single cells is calculated using a one-tailed *t* test. A representative of two independent replicates is shown. **c**, Phagocytic activity measured by flow cytometry of HCMV–GFP-infected M-CSF-induced MDMs. The black line marks the gate for 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. 8f.

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 analysed the progression of HCMV infection in monocytes and macrophages, which support latent and lytic infection, respectively. Our data demonstrate that HCMV infection of MDMs can be either productive or non-productive, allowing us to analyse the decision point of infection outcome. We show that the unequivocal factor associated with the earliest split of productive cells is viral gene expression, specifically the expression of IE genes and the onset of early gene expression. The importance of IE gene expression has been 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 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 associated with viral gene expression levels at the early stages of infection indicates that in this experimental set-up, infection outcome is probably determined primarily by the amount of infectious virus entering the cell[17], which will substantially affect the initial induction of viral gene transcription.

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

Instead, we reveal that intrinsic expression levels of ISGs before 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 in which not all cells are infected, induction of IFNs by infected cells could shift the fate of neighbouring 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 that 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 might exist also in tissues[7]. For many years tissue-resident macrophages were considered to be short lived and 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 cytomegalovirus infection[40], implying that tissue macrophages may also serve as latent cytomegalovirus 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 murine cytomegalovirus infection of macrophages
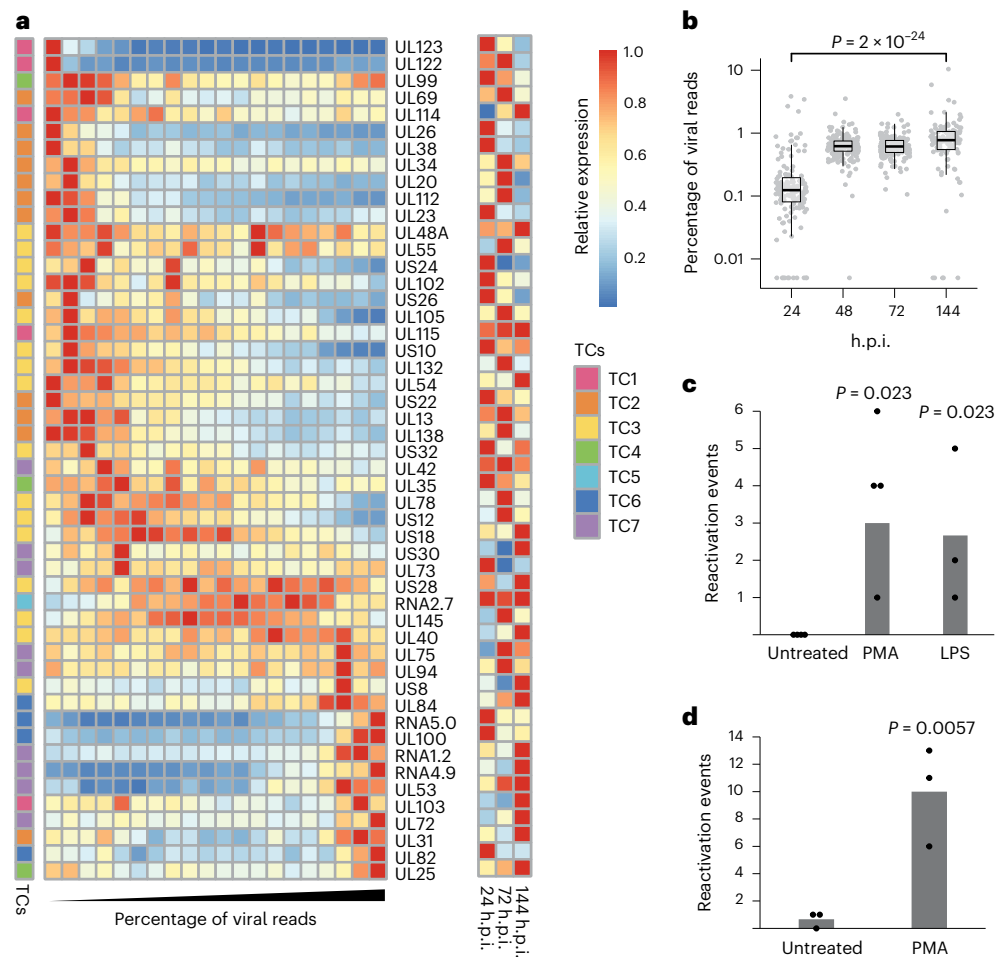
Fig. 6 | Non-productive MDMs exhibit mild elevation of gene expression with time and can reactivate to produce viral progeny. a, Heatmap depicting normalized relative levels of the 50 top relatively expressed viral transcripts in productive MDMs (left) and in non-productive MDMs (right). Productive MDMs 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 during infection. Boxplot shows standard parameters (median, Q1–Q3, 1.5 interquartile range). The P value calculated by one-tailed Wilcoxon test comparing the cells in

the first and last groups is indicated. $n = 147, 189, 165$ and 107 cells for 24, 48, 72 and 144 h.p.i., 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. The P value of the increase in reactivation events upon stimuli, calculated by one-tailed t test on the replicates, is indicated. LPS, lipopolysaccharide. **d**, Number of reactivation events in IFN-pretreated, HCMV-infected, PMA-induced MDMs either untreated or additionally stimulated with PMA. Points show measurements of biological replicates. The P value of the increase in reactivation events upon stimuli, calculated by one-tailed t test on the replicates, is indicated.

reprogrammes 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

Cell culture and virus

Primary CD14$^+$ monocytes were isolated from fresh venous blood, obtained from healthy donors, males and females, aged 25–45, using a 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% CO$_2$, at a concentration of 1–2 million cells per ml in non-stick tubes to avoid differentiation.

Where indicated, primary monocytes were treated with the following treatments immediately following isolation and plating: PMA 50 ng ml^{-1} for 3 d, GM-CSF and interleukin-4 1,000 U ml^{-1} each for 6 d with cytokine replenishing after 3 d, M-CSF 20 ng ml^{-1} for 6 d with cytokine replenishing after 3 d. All treatments were performed in RPMI with 10% heat-inactivated fetal bovine serum (FBS), 2 mM L-glutamine and 100 units ml^{-1} penicillin and streptomycin (Beit-Haemek), 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⁻¹ penicillin and streptomycin (Beit-Haemek). Primary human foreskin fibroblasts (ATCC CRL-1634) were maintained in DMEM with 10% FBS, 2 mM L-glutamine and 100 units ml⁻¹ penicillin and streptomycin (Beit-Haemek). THP1 cells, purchased from ATCC (TIB-202), were grown in RPMI media with 15% heat-inactivated FBS, 2 mM L-glutamine and 100 units ml⁻¹ penicillin and streptomycin (Beit-Haemek). Kasumi-3 cells, purchased from ATCC (CRL-2725), were grown in RPMI media with 20% heat-inactivated FBS, 2 mM L-glutamine and 100 units ml⁻¹ penicillin and streptomycin (Beit-Haemek). Differentiation of THP1 and Kasumi-3 cells was done by adding PMA 50 ng ml⁻¹ for 2 d. Infection and growth following infection were done in media containing 10% FBS. IFN treatment was done with a combination of 550 U ml⁻¹ IFN-α, 700 U ml⁻¹ IFN-β and 500 U ml⁻¹ IFN-γ. Ruxolitinib treatment was done at a concentration of 4 µM.

BAL fluid was obtained from routine bronchoscopies of patients, males and females, aged 58–82. Between 15 and 40 ml of BAL fluid was obtained from each sample and diluted by adding an equal volume of RPMI media with 10% heat-inactivated FBS, 2 mM L-glutamine and 100 units ml⁻¹ penicillin and streptomycin (Beit-Haemek). Diluted BAL fluid was transferred through a 70 µM filter and plated on a tissue culture dish. Adherent cells were washed thoroughly before infection.

The TB40E strain of HCMV, containing an SV40–GFP reporter, was described previously[12]. Virus was propagated by adenofection of infectious bacterial artificial chromosome DNA into fibroblasts as described previously[42]. Viral stocks were concentrated by centrifugation at 26,000g, 4 °C for 120 min. Infectious virus yields were assayed on primary fibroblasts.

Infection and reactivation procedures

Infection was performed by incubating cells with the virus for 3 h, washing twice and supplementing with fresh media. Initial experiments with primary non-differentiated and differentiated monocytes as well as the single-cell experiments of monocytes and MDMs were done at MOI = 10 to ensure infection of all the cells. Notably, because this MOI is based on quantification of infectious particles in fibroblasts it is probably effectively lower in monocytic cells. The single-cell experiments in BAL macrophages as well as follow-up experiments in THP1- and primary MDMs were done in MOI = 5. For reactivation assays, GFP-dim infected MDMs were sorted at 4 d.p.i. and plated in 96-well plates (10,000 cells per well), IFN-pretreated MDMs were used following infection without further sorting. At 7 d.p.i., cells were either treated, as indicated, with 50 ng ml⁻¹ PMA (Sigma), or 500 ng ml⁻¹ LPS (Sigma) or left untreated for 3 d and then the media was replaced with fresh untreated media. At 14 d.p.i., 7 d post treatment, GFP-positive cells were counted on a fluorescent microscope. At 18 d.p.i. 7,500 fibroblasts per well were seeded with the MDMs and plaque formation was followed and imaged for a period of 14 d.

Flow cytometry and sorting

Cells were analysed on a BD Accuri C6 using 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. The number of GFP-positive fibroblasts was measured at 2 d.p.i. by flow cytometry or microscopy and the number of 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% Nonidet P-40

and 0.4 mg ml⁻¹ 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-labelled HCMV primer and probe: HCMV HHV5 kit for quantitative PCR using a glycoprotein B target (PrimerDesign); and HEX-labelled RPP30 copy number assay for droplet-digital PCR (ddPCR) (Bio-Rad), as described previously[43]. Analysis was done using Quantasoft software.

Microscopy

Imaging was performed on an Olympus Ix73 inverted fluorescent microscope using a ×4 or ×10 objective and an Olympus DP73 camera using CellSens dimensions software or on an AxioObserver Z1 wide-field microscope using a ×20 objective and Axiocam 506 mono camera using ZEN software.

scRNA-seq library preparation

For primary monocytes and MDMs, single-cell sorting and library preparation were conducted according to the massively parallel scRNA-seq (MARS-seq) protocol, as described previously[44]. In brief, single cells were FACS sorted into wells of 384-well capture plates containing 2 µl of lysis buffer and reverse transcription (RT)-indexed poly(T) primers, thus generating libraries representing the 3′-end of messenger RNA 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 complementary DNA 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 analysed 871 monocytes and 2,355 MDMs. The numbers of cells in clusters 1–10 of infected MDMs are: 347, 325, 232, 229, 33, 122, 53, 274, 117 and 83.

For BAL cells, pools of 10,000 cells were prepared for scRNA-seq using the Chromium Single Cell 3′ Gene Expression v.3.1 kit according to the manufacturer's instructions (10x Genomics).

scRNA-seq data analysis

Analysis of MARS-seq data was done using the tools described in Jaitin et al.[45]. The reference was created from the hg19 and TB40E (NCBI EF999921.1) strain of HCMV. 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 base pairs (bp)) and 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 (ref. [46]), and counting of the reads was 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 with the total number of UMIs in the batch. The noise levels in all batches were lower than 1.5%. Wells with fewer than 500 or more than 5,000 expressed genes, or more than 20,000 reads were discarded. The number of wells used for further analysis was 3,226.

For 10x Genomics data, we used Cell Ranger software (v.6.1.2, 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 fewer than 500 or more than 10,000 expressed genes, or more than 30,000 reads were discarded. The number of wells used for further analysis was 21,839.

All single-cell analyses were done using Seurat v.4.0.3 (ref. [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 Figs. 1 and 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 percentage of 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 because it is very small, these cells were removed from all subsequent cluster-based analyses. Annotation of the 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. Percentage viral reads are calculated as the percentage of reads on all viral genes of all reads in a cell. These numbers were used for colouring UMAP plots in Figs. 1h, 2c,f and 4a and Extended Data Fig. 3b,e, and for plots in Figs. 2d, 4b,c,g–i and 6b and Extended Data Figs. 6b, 7a and 9b,c. The list of ISGs is the combined list of the following gene sets: Hallmark IFN-alpha response, Hallmark IFN-gamma response and Reactome IFN signalling. The list of macrophage differentiation genes is based on the list of genes in the GO term regulation of macrophage differentiation (GO:0045649).

The 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 colouring 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 eight 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 described previously[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 complementary DNA synthesis. The cDNA products were 3′ ligated with a second adaptor using T4 ligase and amplified with eight cycles of PCR for final library products of 200–300 bp. Raw sequences were first trimmed at their 30 end, removing the Illumina adapter and poly(A) tail. Alignment was performed using Bowtie 1 (ref. [48]) (allowing up to two 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 40 mM dithiothreitol. Owing to the low number of cells, RNA-seq libraries were performed using a modified molecular crowding single-cell barcoding and sequencing (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.

Sixty-six base pair raw reads were aligned using STAR (v.2.7.1)[50] to concatenation of the human (hg19) and the viral genomes (NCBI EF999921.1). Counting of reads per gene was done based on 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 N. 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, catalogue no. 141367) in place of the SARS-CoV-2-nsp1-2XStrep cassette. The UL123, UL122 and UL99 coding sequences were amplified from cDNA of HCMV-infected fibroblasts with primers containing flanking regions homologous to the vector (Supplementary Table 1). 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 knockout (KO) was done using a CRISPR–Cas9 system, with lentiCRISPR v2 plasmid (Addgene, catalogue no. 52961). Single guide RNAs were cloned downstream of the U6 promoter (Supplementary Table 1).

Lentiviral particles were generated by cotransfection of the expression constructs and second-generation packaging plasmids (psPAX2, Addgene, catalogue no. 12260 and pMD2.G, Addgene, catalogue no. 12259), using jetPEI DNA transfection reagent (Polyplus transfection) into 293T cells, according to the manufacturer's instructions. At 60 h post transfection, supernatants were collected and filtered through a 0.45-µm polyvinylidene difluoride filter (Millex). THP1 cells were transduced with lentiviral particles by spinfection (800g, 1 h) and then puromycin-selected (1 µg ml⁻¹) for 4 d. Puromycin was removed before subsequent differentiation. Induction of the inducible transgene was done by addition of 100 ng ml⁻¹ doxycycline (Sigma) 24 h before HCMV infection. Kasumi-3 cells were transduced with lentiviral particles by spinfection (800g, 1 h). Because doxycycline induction was not successful in our hands in Kasumi-3 cells, they were 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 carried out and the cells were treated with PMA 2 d after transduction followed by HCMV infection 2 d 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 the 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% Nonidet P-40 and 0.4 mg ml⁻¹ 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, which was used for normalization (Supplementary Table 1).

Primers for the viral UL44 gene were used to quantify viral DNA and a host gene (B2M) was used for normalization (Supplementary Table 1).

Western blot analysis

Cells were lysed in RIPA buffer for 30 min on ice and then centrifuged at 14,000g for 20 min at 4 °C and supernatant was collected. Samples were separated by 4% to 12% polyacrylamide Bis-Tris gel electrophoresis

(Invitrogen), blotted onto nitrocellulose membranes and immuno-blotted with the following primary antibodies: mouse anti-CMV IE1 and IE2 1:1,000 (Abcam), mouse anti-PP28 1:2,000 (EastCoast Bio), mouse anti-mCherry 1:1,000 (ChromoTek), rabbit anti-GAPDH 1:1,000 (Cell Signaling Technology), rabbit anti-IRF9 1:1,000 (Cell Signaling Technology), rabbit anti-STAT2 1:1,000 (Cell Signaling Technology), mouse anti-actin 1:7,500 (Sigma). The secondary antibodies were goat anti-rabbit–IRDye 800CW and goat anti-mouse–IRDye 680RD (LI-COR Biosciences) diluted 1:7,500. Membranes were visualized in a LI-COR Biosciences Odyssey imaging system.

Correlation between cellular gene expression and viral gene expression levels

For calculating the Spearman correlation between cellular genes and viral expression levels, in each cell, the sum of viral reads was normalized to the total number of reads in the cell, and the number of reads of each cellular gene was normalized to the total number of cellular reads in the cell. The correlation was calculated across 313 and 310 cells on 5,007 and 5,606 genes for the 5 h.p.i. and 12 h.p.i. analyses, respectively. Genes with a total number of reads below 45 (after normalization to 5,000 reads per cell) were not included.

Differential expression and enrichment analysis

Differential expression analysis on bulk RNA-seq data was performed with DESeq2 (v.1.22.2)[53] using default parameters, with the number of reads in each of the samples as an input. Differential expression analysis of the scRNA-seq was done using Seurat v.4.0.3 (ref. [16]).

The $\log_2$(fold change) values were used for enrichment analysis using GSEA (v.4.1)[54]. For scRNA-seq data, only genes that were expressed in at least 20% of cells in both groups of cells in a comparison were used. For bulk RNA-seq, only genes that had at least ten reads in each sample were used. The gene sets that were used were Hallmark and Kyoto Encyclopedia of Genes and Genomes from MSigDB (v.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. 9a). The expression levels in the heatmap are relative to the highest level of each gene. Productive MDMs were ordered by the percentage of 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 infection. In each bin/time group, the 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 allophycocyanin (APC)-conjugated anti-human CD11b (Biolegend, catalogue no. 101211), phycoerythrin (PE)-conjugated anti-human CD11c (BD, catalogue no. 333149), or fluorescein isothiocyanate (FITC)-conjugated anti-human CD14 (Miltenyi, catalogue no. 130-080-701) diluted 1:200. Phagocytosis was assayed using a phagocytosis kit (Cayman Chemical, catalogue no. 601490) 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

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

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Sequencing data have been deposited in GEO under accession code GSE189581. Other datasets analysed in this study: single-cell RNA-seq dataset during HCMV infection of fibroblasts is available under accession number GSE165291; RNA-seq dataset during differentiation of monocytes to MDMs is available under accession number GSE147314; single-cell RNA-seq datasets during differentiation of haematopoietic 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. Source data are provided with this paper. Correspondence and requests for materials should be addressed to Michal Schwartz.

References

1. Staras, S. A. S. et al. Seroprevalence of cytomegalovirus infection in the United States, 1988–1994. *Clin. Infect. Dis.* **43**, 1143–1151 (2006).
2. Mocarski, E., Shenk, T., Griffiths, P. & Pass, R. in *Fields Virology* 6th edn (eds Knipe, D. M. et al.) Ch. 62 (Lippincott Williams & Wilkins, 2013).
3. Hahn, G., Jores, R. & Mocarski, E. S. Cytomegalovirus remains latent in a common precursor of dendritic and myeloid cells. *Proc. Natl Acad. Sci. USA* **95**, 3937–3942 (1998).
4. Kondo, K., Kaneshima, H. & Mocarski, E. S. Human cytomegalovirus latent infection of granulocyte–macrophage progenitors. *Proc. Natl Acad. Sci. USA* **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**, 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. Monogr Virol. Basel* **21**, 288–301 (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, G. et al. Tissue macrophages are infected by human cytomegalovirus in vivo. *J. Infect. Dis.* **173**, 240–245 (1996).
10. Hargett, D. & Shenk, T. E. Experimental human cytomegalovirus latency in CD14+ monocytes. *Proc. Natl Acad. Sci. USA* **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. *J. Virol.* **86**, 9854–9865 (2012).
13. Shnayder, M. et al. Defining the transcriptional landscape during cytomegalovirus latency with single-cell RNA sequencing. *mBio* **9**, e00013–e00018 (2018).
14. Cheng, S. et al. Transcriptome-wide characterization of human cytomegalovirus in natural infection and experimental latency. *Proc. Natl. Acad. Sci. USA* **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. USA* **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**, e52168 (2020).

20. Hein, M. Y. & Weissman, J. S. Functional single-cell genomics of human cytomegalovirus infection. *Nat. Biotechnol.* **40**, 391–401 (2021).

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**, e46339 (2019).

34. Wu, X. et al. Intrinsic immunity shapes viral resistance of stem. *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).

42. Elbasani, E., Gabaev, I., Steinbrück, L., Messerle, M. & Borst, E. M. Analysis of essential viral gene functions after highly efficient adenofection of cells with cloned human cytomegalovirus genomes. *Viruses* **6**, 354–370 (2014).

43. Jackson, S. E. et al. Latent cytomegalovirus (CMV) infection does not detrimentally alter T cell responses in the healthy old, but increased latent CMV carriage is related to expanded CMV-specific T cells. *Front. Immunol.* **8**, 733 (2017).

44. Keren-Shaul, H. et al. MARS-seq2.0: an experimental and analytical pipeline for indexed sorting combined with single-cell RNA sequencing v1. *Nat. Protoc.* **14**, 1841–1862 (2019).

45. Jaitin, D. A. et al. Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. *Science* **343**, 776–779 (2014).

46. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).

47. Shishkin, A. A. et al. Simultaneous generation of many RNA-seq libraries in a single reaction. *Nat. Methods* **12**, 323–325 (2015).

48. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).

49. Bagnoli, J. W. et al. Sensitive and powerful single-cell RNA sequencing using mcSCRB-seq. *Nat. Commun.* **9**, 2937 (2018).

50. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

51. van den Ent, F. & Löwe, J. RF cloning: a restriction-free method for inserting target genes into plasmids. *J. Biochem. Biophys. Methods* **67**, 67–74 (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. USA* **102**, 15545–15550 (2005).

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

Acknowledgements

Author contributions

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

Competing interests

The authors declare no competing interests.

Additional information

Reprints and permissions information is available at www.nature.com/reprints.
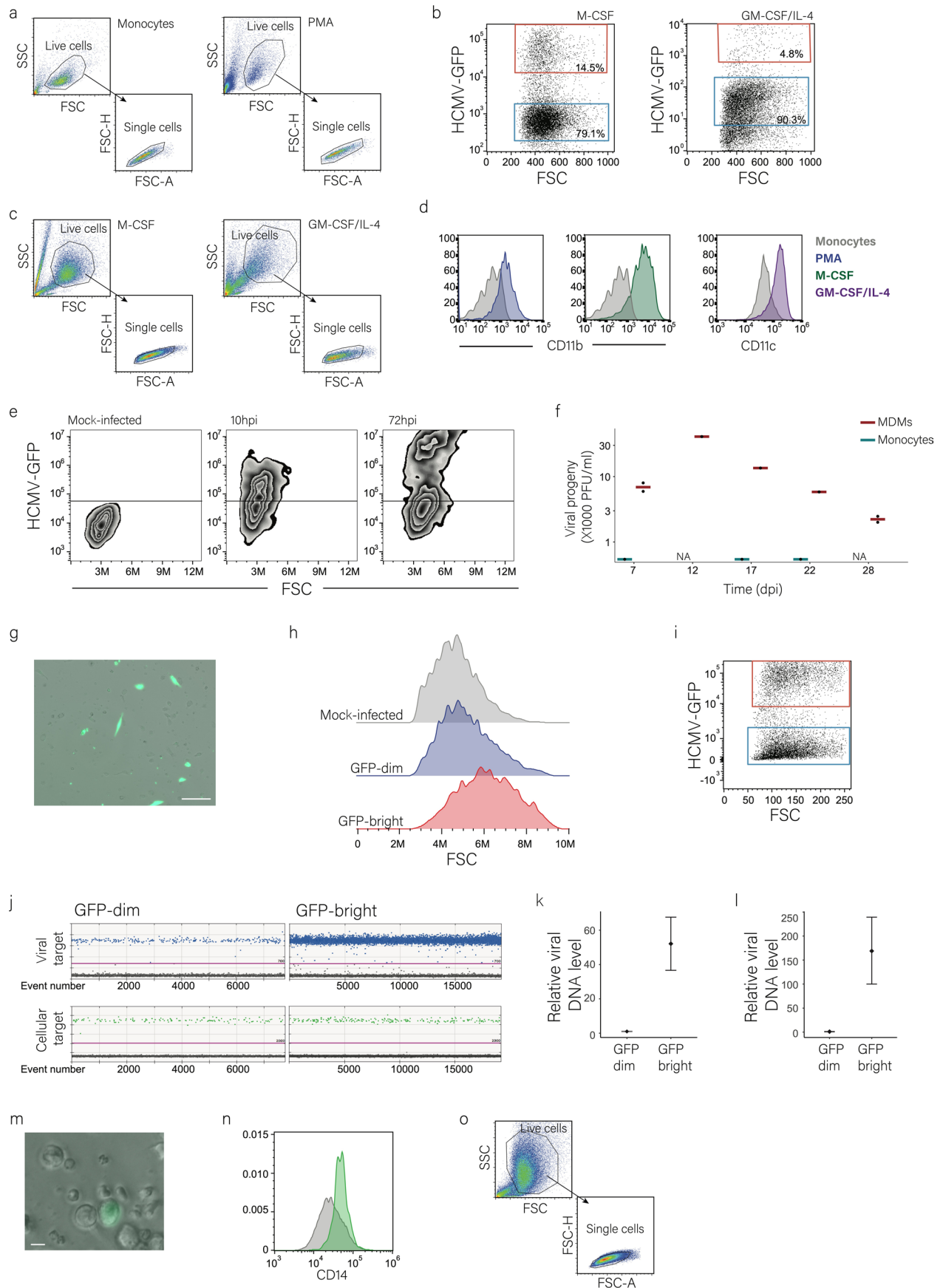
Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | Productive and non-productive infection in monocyte derived macrophages. 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 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% confidence interval 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.
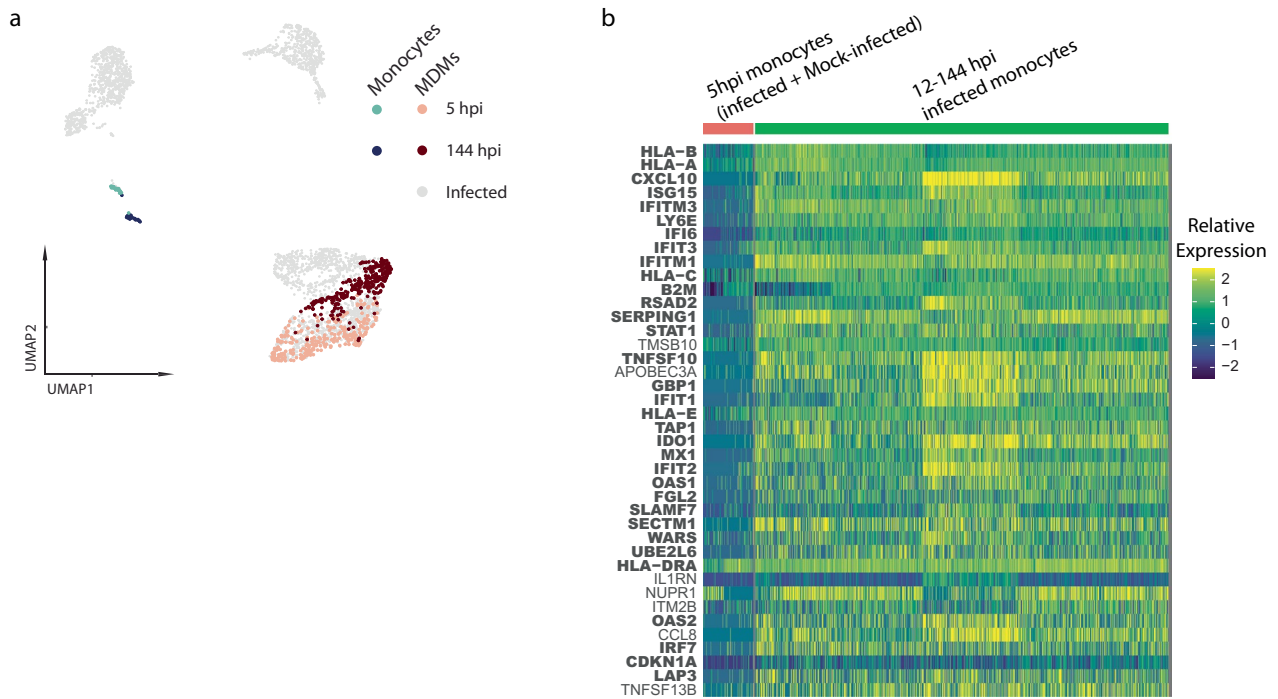
Extended Data Fig. 2 | Single cell transcriptome analyses of HCMV infected monocytes and MDMS. a. Projection of mock-infected and HCMV-GFP infected CD14 + monocytes (n = 68 and n = 803, respectively) and MDMs (n = 540 and n = 1815, respectively) from the indicated time points, based on cellular and viral 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.
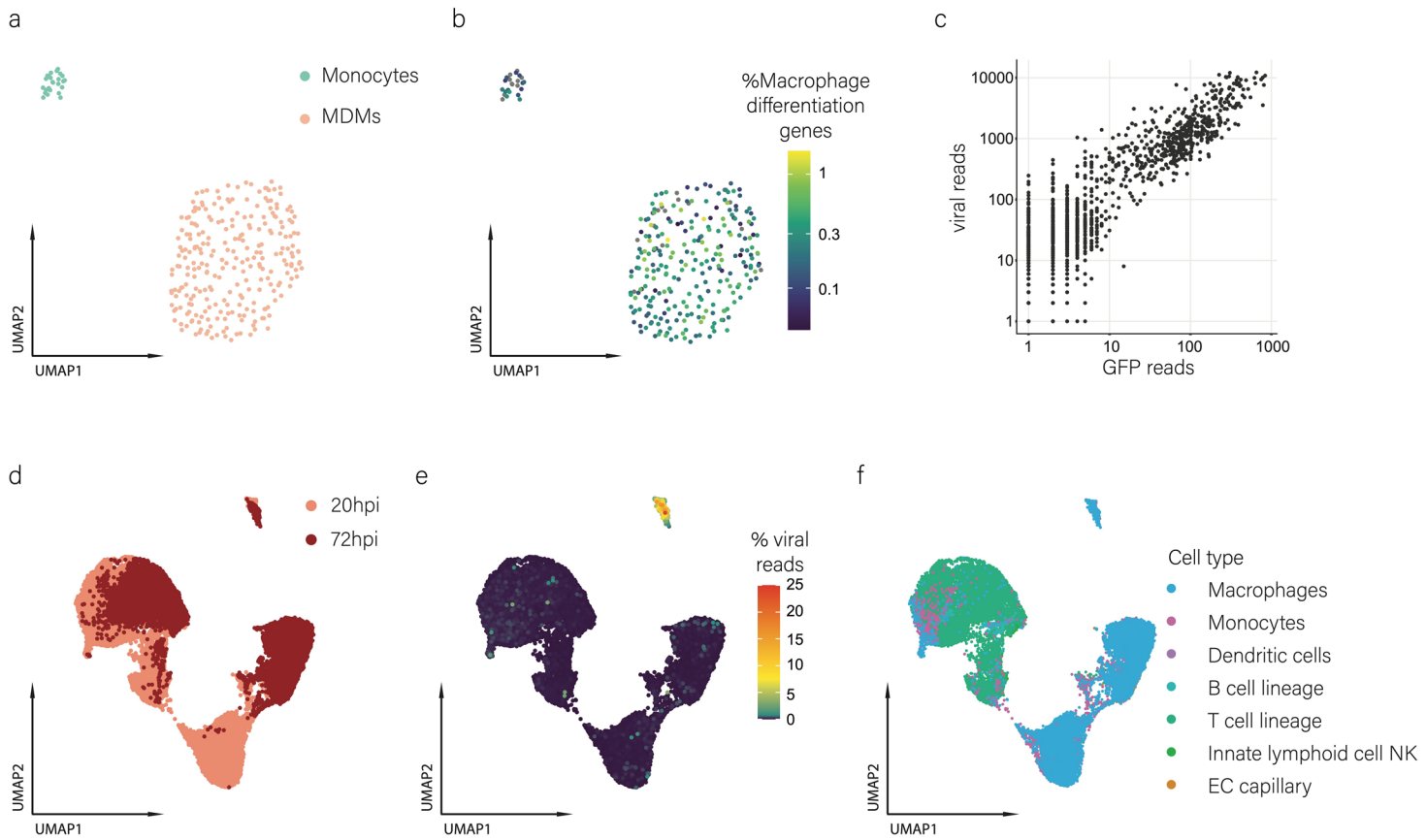
Extended Data Fig. 3 | Features of single infected MDMs and BAL cells. a. Projection of 283 mock-infected monocytes and MDMs at 5 hpi colored by cell type. **b**. Projections as in **a**, colored by the expression level of macrophage differentiation genes. **c**. Levels of viral reads versus GFP reads 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].
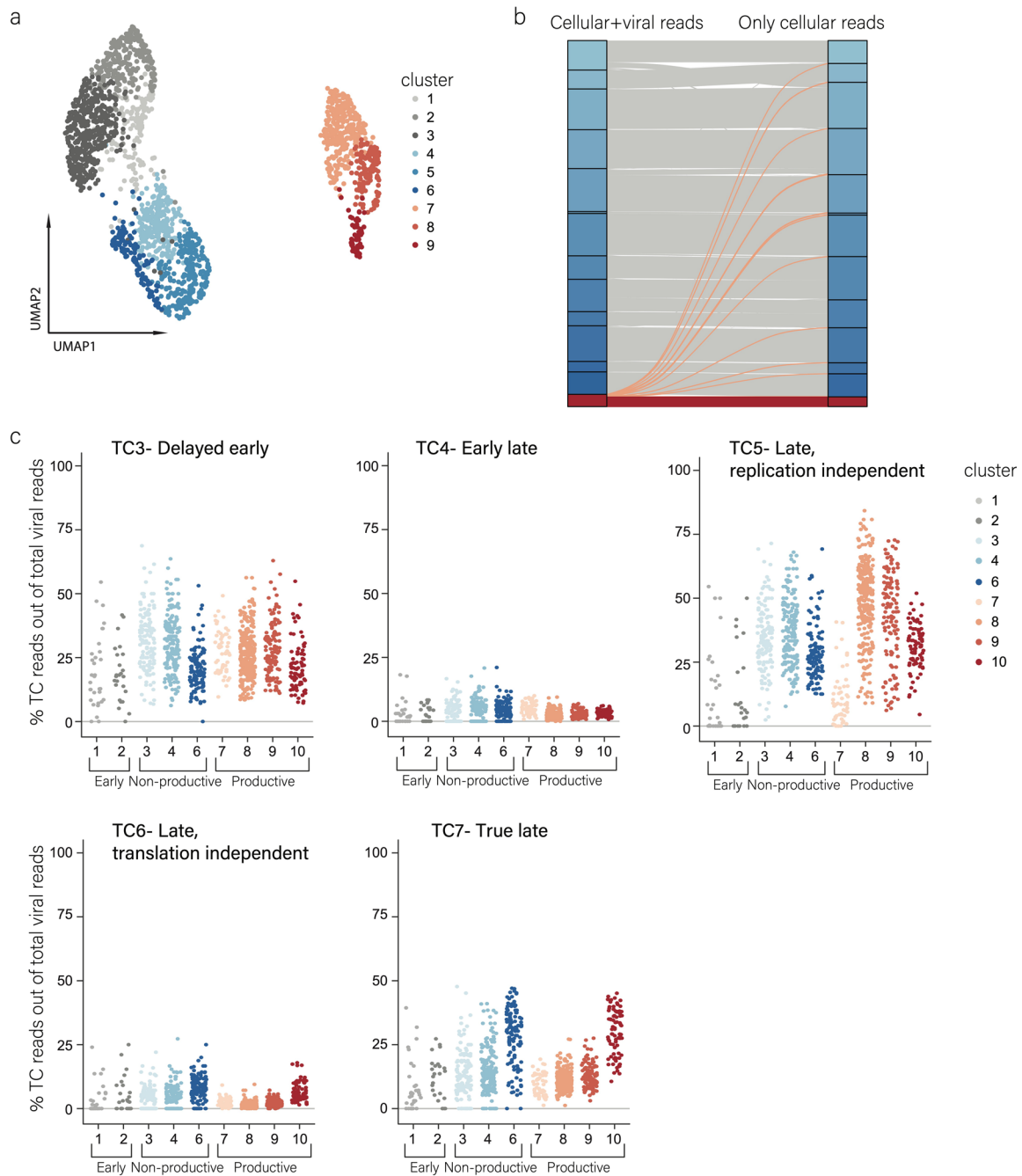
Extended Data Fig. 4 | High expression of viral immediate early genes characterizes early productive MDMs. a. Projection of HCMV-GFP infected MDMs at 5-144 hpi, analyzed based solely on cellular transcripts and colored by cluster assignment. **b**. Cluster assignment of each BAL myeloid cell according 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].
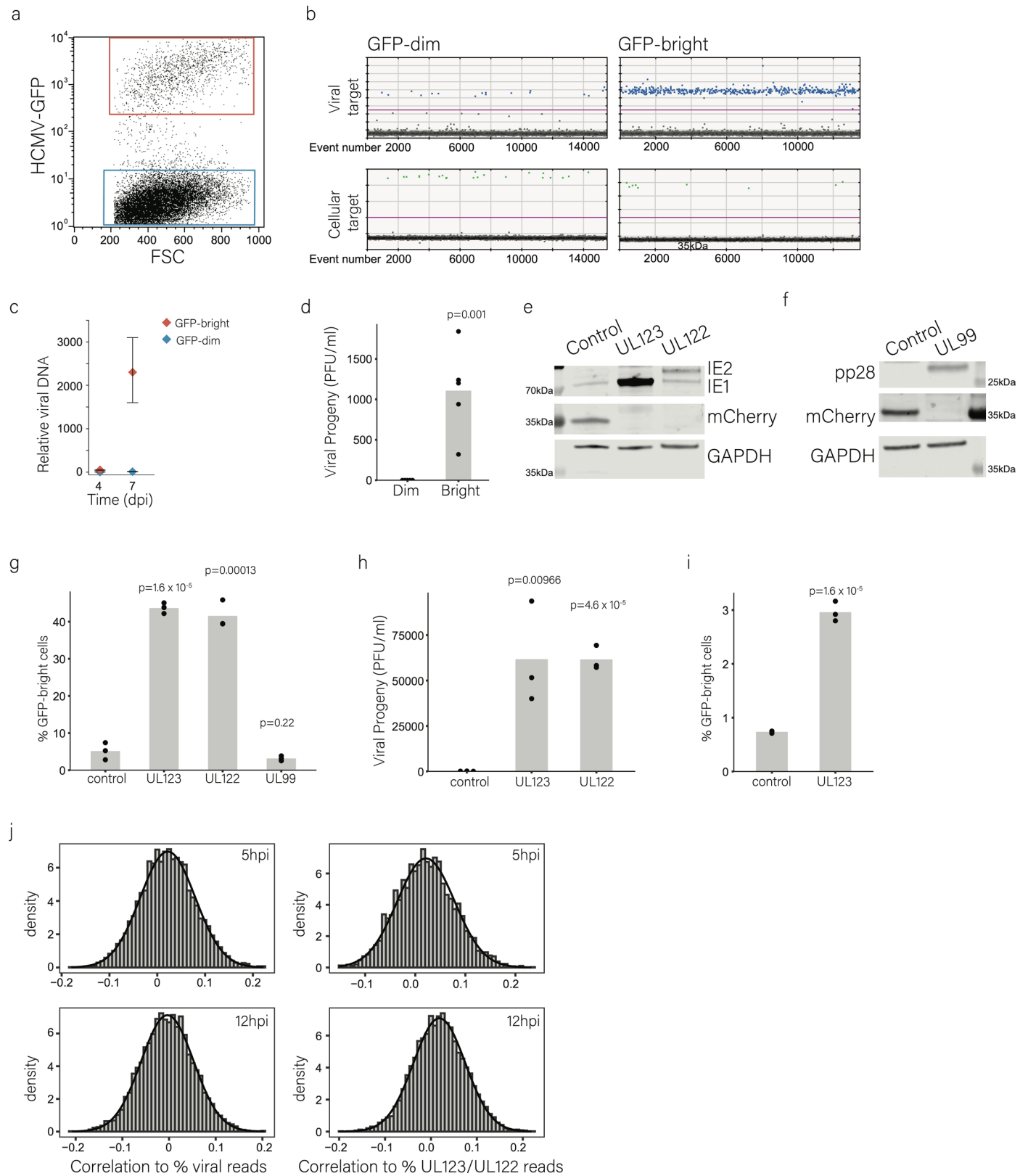
Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | Ectopic expression of IE genes in THP1 derived macrophages drives productive infection. 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% confidence interval 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.
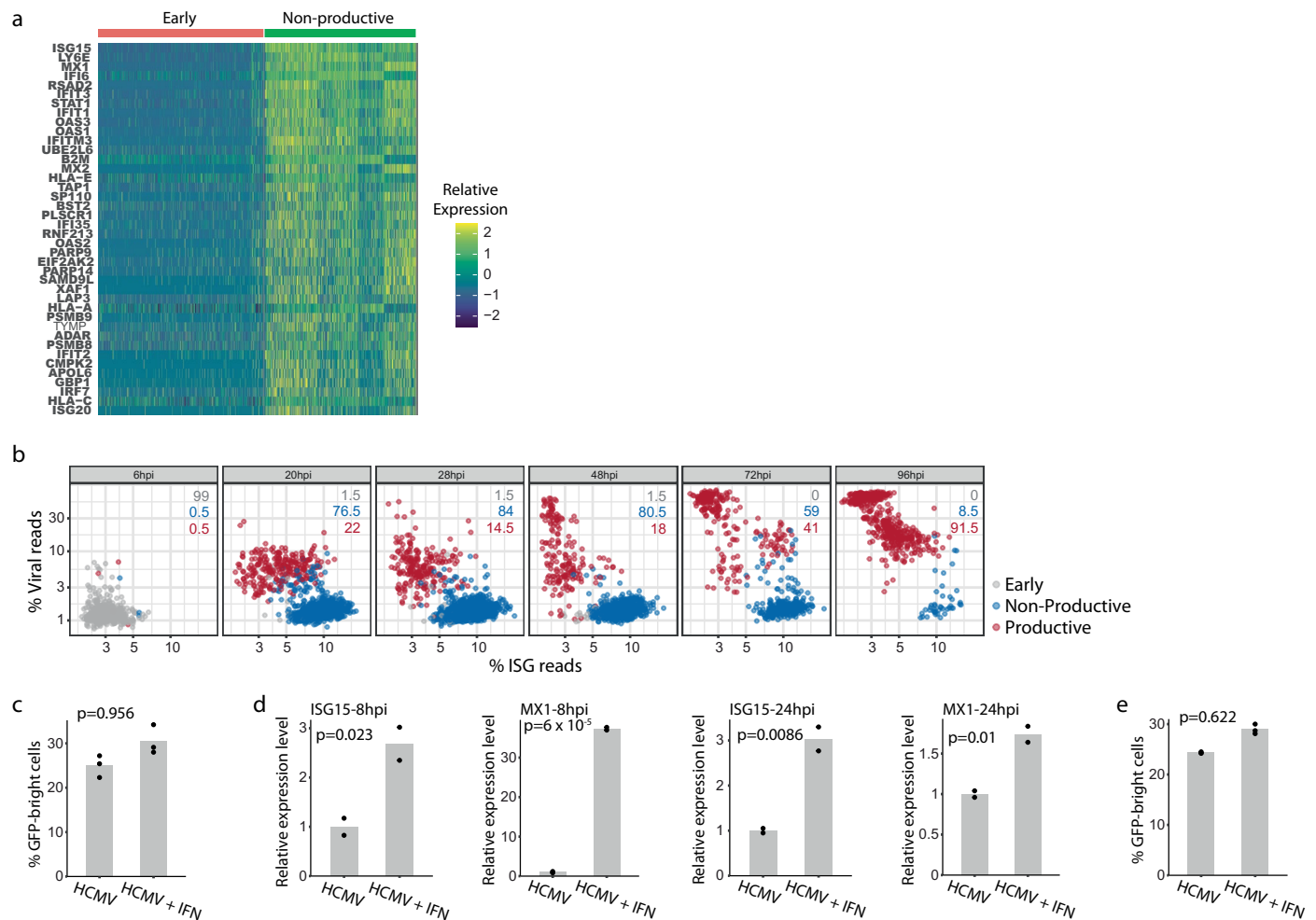
Extended Data Fig. 6 | Induction of ISGs does not dictate infection outcome. a. Heatmap depicting the 40 most significantly increased genes in non-productive MDMs compared to early MDMs. ISGs are highlighted in bold. Z-scores of the log of the expression values are shown. **b.** Viral gene expression 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 IFNs 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 IFNs 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 IFNs 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.
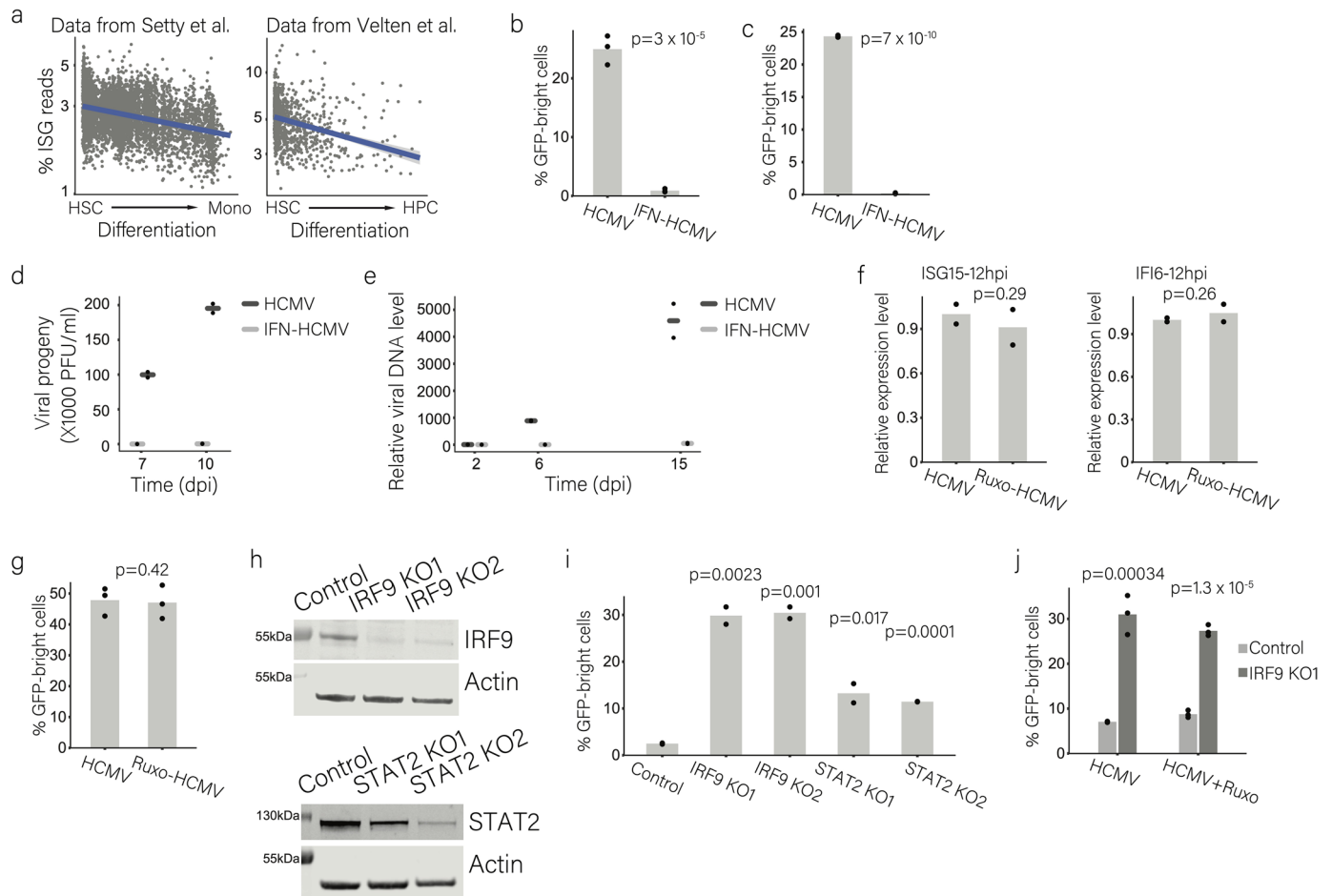
Extended Data Fig. 7 | Intrinsic ISG levels affect HCMV infection outcome.
a. ISG expression level versus differentiation status in single bone marrow human haematopoietic stem cells. Data taken from[25] shows differentiation trajectory from hematopoietic stem cells (HSCs) to monocytes (mono, left panel) 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 IFNs 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 IFNs 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 IFNs 24 hours prior to infection, at different times post infection. **e.** qPCR measurements of viral genomes from infected MDMs either untreated or treated with IFNs 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.
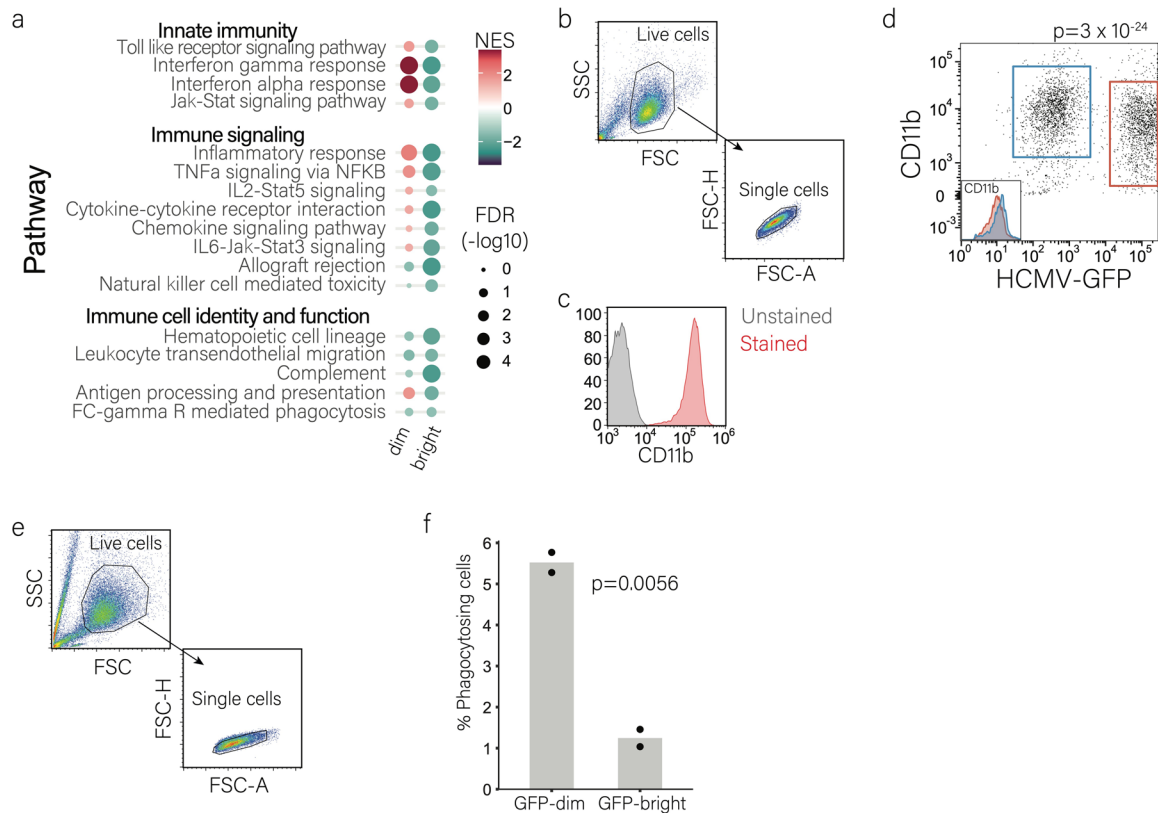
Extended Data Fig. 8 | Effects of HCMV infection on macrophage cell identity.
a. Gene set enrichment analysis (GSEA) of differentially expressed genes between GFP-dim and GFP-bright BAL infected macrophages compared to mock-infected cells, on innate immunity, immune signaling and immune identity and 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.
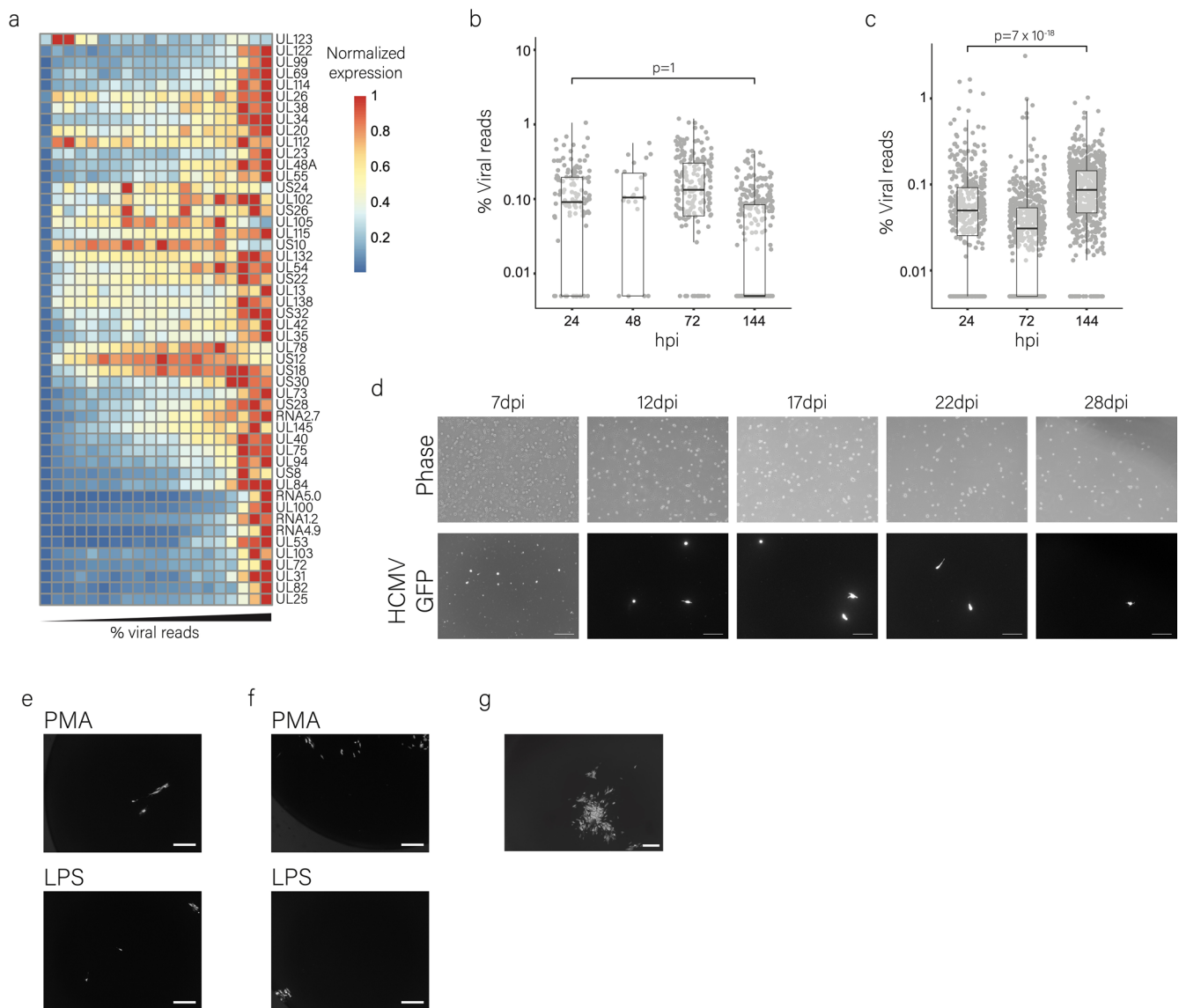
Extended Data Fig. 9 | Non-productive MDMs continuously express viral transcripts and can reactivate. a. Heatmap depicting absolute levels of viral transcripts, in reads per million, in productive MDMs. Staging of the cells and ordering of the viral genes are as in Fig. 6a. The expression levels in the heatmap are relative to the highest level of each transcript. **b**. Percentage of viral reads in single monocytes during 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 during 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 IFN-pre treated, PMA-induced MDMs. Reactivation was done by PMA. Scale bar is 500μm. Shown is a representative of three biological replicates.

nature portfolio

Corresponding author(s):	Michal Schwartz Noam Stern-Ginossar
Last updated by author(s):	Aug 16, 2022

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	For flow cytometry data acquisition we used FACSDiva (v8.0.1) or BD Accuri C6 (V1.0.264.21). For microscopy image aquisition we used CellSens dimensions (V1.12) or Zen (V2.0).
Data analysis	Bowtie 2, Bowtie 1 and STAR (V2.7.1) were used for alignment of different RNA-seq libraries. All single cell analyses were done using Seurat (V4.0.3). DESeq 2 (V1.22.2) was used for differential expression analysis. GSEA v4.1 was used for enrichment analysis. For 10xGenomics data, we used Cell Ranger software (V6.1.2, 10x Genomics). Data visualiztion was done using R (V4.1.0). For flow cytometry data analysis we used FlowJo (V10.8.0). ddPCR aquisition and analysis were done using QuantaSoft (V1.7.4).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All sequencing data from this manuscript have been deposited in GEO under accession code GSE189581. 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

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	For all experiments we used at least biological duplicates, as is customary in the field. qPCR was performed with technical triplicates, as is customary in the field. Specific sample sizes were chosen considering the variability of the measurement.
Data exclusions	We did not exclude data
Replication	We confirmed strong reproducibility between duplicates
Randomization	Samples in all experiments were allocated randomly.
Blinding	Blinding was not relevant to the study since no subjective variables were measured

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	APC-conjugated anti-human CD11b (#101211, Biolegend), PE-conjugated anti-human CD11c (#333149, BD), FITC-conjugated anti-human CD14 (#130-080-701, Miltenyi). Mouse anti-CMV IE1 and IE2 (#ab53495, abcam), mouse anti-PP28 (#CA004, EastCoast Bio), rabbit anti-GAPDH (#2118S, Cell Signaling Technology), rabbit anti- IRF9 (#76684S, Cell signaling Technology), rabbit anti- STAT2 (#72604S, Cell signaling Technology) , mouse anti-mCherry (#6G6, Chromotek), mouse anti- beta-actin (#A4700, Sigma), goat anti-rabbit–IRDye 800CW (#926-32211, LI-COR), goat anti-mouse–IRDye 680RD (#926-68070, LI-COR).
Validation	CD11c and CD11b Antibodies were validated by observing a shift in staining in differentiated cells (Extended Data figure 1d). CD14 antibodies were validated by positive staining of CD14+ selected cells and lack of staining in CD14- cells (not shown). All flow cytometry antibodies were validate for the application by the manufacturer. Western blot antibodies were validated by detection of a band in the correct size. Antibodies of viral proteins were also validated by detection of a band in samples of infected cells but not of uninfected cells (not shown). Antibodies for human proteins and the m-cherry antibody were validated for the application and the species by the manufacturer.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	THP1, Kasumi-3, 293T and Human foreskin fibroblasts were purchased from ATCC
Authentication	Cell lines were authenticated by ATCC by STR profiling

Mycoplasma contamination	All cell lines were tested negative for Mycoplasma
Commonly misidentified lines (See ICLAC register)	None

Human research participants

Policy information about studies involving human research participants

Population characteristics	For blood monocytes donors were healthy volunteers, male and female ages 25-45. BronchoAlveolar Lavage fluid was obtained from adults undergoing routine bronchoscopies, male and female ages 58-82.
Recruitment	Blood donors were recruited through the Weizmann institute. Patients undergoing routine bronchoscopies, performed in the Institute of Pulmonary Medicine, Hadassah Hebrew University Medical Center were asked for their willingness to donate some of their BAL fluid and signed an informed consent.
Ethics oversight	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 a 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.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	For GFP analysis or single cell sorting, cells were trypsinized if needed. For surface marker staining, cells were stained in cold MACS buffer (PBS, 5% BSA, 2 mM EDTA). Phagocytosis was assayed a phagocytosis kit (#601490, Cayman chemical) according to the manufacturer's instructions.
Instrument	Cells were analyzed on a BD Accuri C6 or a BD LSRII and sorted on a BD FACSAria III
Software	BD FACSDiva software was used for collection on LSRII and BD FACSAria III. BD Accuri C6 plus software was used for collection on BD Accuri C6. FlowJo was used for analysis of the data.
Cell population abundance	The cell population abundance is shown in the relevant figures. Sorting was based on GFP and was verified immediately after sorting by flow cytometry and microscopy.
Gating strategy	Live cells were determined according to FSC/SSC gating, doublets exclusion was performed by plotting the height against the area of the forward or side scatter. When sorting according to GFP levels, two separate and distinct populations were seen as show in the relevant figures.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.