Temporal Dynamics of HCMV Gene Expression in Lytic and Latent Infections

Batsheva Rozman[1], Yaarit Kitsberg[1], Aharon Nachshon[1], Michal Schwartz[1*] and Noam Stern-Ginossar[1*]

[1] Department of Molecular Genetics, Weizmann Institute of Science, Rehovot 76100, Israel.

[*] To whom correspondence should be addressed: michalsc@weizmann.ac.il

or noam.stern-ginossar@weizmann.ac.il

Abstract

Primary infection with Human cytomegalovirus (HCMV) results in a persistent lifelong infection due to its ability to establish latent infection. During productive HCMV infection, viral genes are expressed in a coordinated cascade that is characteristic of all herpesviruses and traditionally relies on the dependencies of viral genes on protein synthesis and viral DNA replication. In contrast, the transcriptional landscape associated with HCMV latency is still disputed and poorly understood. Here, we examine viral transcriptomic dynamics during the establishment of both productive and latent HCMV infections. These temporal measurements reveal that viral gene expression dynamics along productive infection and their dependencies on protein synthesis and viral DNA replication, do not fully align. This illustrates that the regulation of herpesvirus genes does not represent a simple sequential transcriptional cascade and surprisingly many viral genes are regulated by multiple independent modules. Using our improved classification of viral gene expression kinetics in conjunction with transcriptome-wide measurements of the effects of a wide array of chromatin modifiers, we unbiasedly show that a defining characteristic of latent cells is the unique repression of immediate early (IE) genes. In particular, we demonstrate that IE1 (a central IE protein) expression is the principal barrier for achieving a full productive cycle. Altogether, our findings provide an unbiased and elaborate definition of HCMV gene expression in lytic and latent infection states.

Introduction

Human cytomegalovirus (HCMV) is a pervasive pathogen of the beta herpesvirus family, infecting the majority of the world population[1]. Like all herpesviruses, HCMV persists throughout the lifetime of the host by establishing latent infection from which the virus can later reactivate, causing life-threatening disease in immunocompromised individuals such as transplant recipients and HIV patients [2].

Members of the family *Herpesviridae* are large double stranded DNA viruses, with the HCMV genome encompassing 235 kb, making it one of the largest known viruses to infect humans[3]. During productive infection, transcription from herpesvirus genomes is accomplished by the host RNA polymerase II and regulated by host and viral proteins, leading to a coordinated viral gene

expression cascade and resulting in the production of infectious progeny. Traditionally, viral genes are divided into three distinct and sequential expression groups, differing with respect to their regulation and kinetics. The transcriptional cascade begins with the immediate-early (IE) genes, which require no new cellular or viral protein synthesis for their expression. Next, early (E) genes, including those encoding proteins necessary for DNA replication, are transcribed. These genes are dependent on the presence of viral IE proteins and independent of viral DNA synthesis. Lastly, late (L) genes are those whose expression is dependent on viral DNA synthesis and generally code for structural proteins necessary for the assembly of new virions[4]. Thus, herpesvirus gene expression during productive infection is assumed to represent a classic sequential regulatory cascade.

Historically, metabolic inhibitors have been used to classify viral genes into IE, E and L temporal expression profiles. Inhibitors of protein synthesis, such as cycloheximide (CHX), lead to the specific accumulation of IE transcripts and these genes were assumed to be expressed at the onset of viral gene expression. Inhibitors of viral DNA replication, such as phosphonoformate (PFA), lead to the specific depletion of transcripts dependent on or enhanced by the onset of viral DNA synthesis and these genes were assumed to be expressed with late kinetics. Early studies provided the framework for mapping viral gene expression kinetics using hybridization methodologies [5–7] and later on by using microarrays [8]. The reannotation of the HCMV transcriptional landscape has elucidated the crowded nature of its genome and its ability to encode many overlapping RNAs [3,9]. These findings highlight the necessity for unbiased transcriptomic methods to accurately determine HCMV gene expression dynamics.

Viral gene expression during latent infection has been extremely hard to define and is still a matter of controversy[11]. In latently infected cells, the viral genome is maintained in a repressed state where no viral DNA synthesis occurs and there is no production of new infectious virions[12,13]. Since the major immediate-early promoter (MIEP) drives the expression of IE1 and IE2, two potent transactivators of viral gene expression, its regulation has been extensively studied [14-15] and it is well established that it is chromatinized and repressed in hematopoietic cells, in which latency is established [13,16,17]. Thus, repression of the MIEP is considered a major mechanism dictating latent infection. However, viral gene expression is generally repressed in latent infection, and whether the MIEP is uniquely regulated has not been addressed in an

unbiased manner. It has further been demonstrated that ectopic expression of IE1 and IE2 in non-permissive cells does not result in production of viral progeny[19], indicating that there are additional barriers which need to be crossed for successful viral reactivation. In parallel to studies that focused on deciphering the repression of the MIEP, there were extensive efforts to decipher the latent HCMV transcriptome. The leading idea for many years was that in latency, viral gene expression is generally silenced and only a limited number of genes—the presumed latency genes—are expressed[20,21]. However, recent transcriptomic studies on latent hematopoietic cells, including research from our lab, revealed that viral genes are broadly expressed but at low levels[23, 24]. These findings have complicated the characterization of latent viral gene expression and its distinction from lytic viral gene expression as well as the understanding of how the transcriptome reflects latency regulation.

Using transcriptome wide sequencing, we perform a dense time course along both lytic and latent HCMV infections. We define kinetic classes of HCMV transcripts along productive infection, revealing novel patterns of viral gene expression. We uncover that the sensitivity of viral transcripts to translation and replication inhibitors and their temporal dynamics along infection, are often two independent properties, implying that expression of many viral genes is dictated by multiple regulation modules. In latent infection, we show that at early time points post infection the latent transcriptome in experimental models is highly dominated by virion-associated input RNA. By unbiasedly measuring the effects of a wide array of chromatin modifiers that enhance viral gene expression, we reveal that the characteristic of latent cells, in terms of viral gene expression, is distinctive repression of IE genes. We further demonstrate that when HCMV entry is efficient, over-expression of IE1 is sufficient to drive a full replicative cycle in latent cells. These findings suggest that the historical terms which defined the basic herpesvirus expression cascade do not capture the full complexity of herpesvirus gene expression dynamics. Furthermore, in latency, IE genes are distinctively repressed, and this repression is the main barrier that dictates latency and prevents reactivation.

Results

Temporal analysis of HCMV gene expression along lytic and latent infections

In order to comprehensively define viral gene expression kinetics along productive infection, and to resolve the molecular events which precede the broad but repressed transcriptional state in latency, we aimed to comprehensively determine the HCMV temporal gene expression cascade along lytic and latent infection in human foreskin fibroblasts and monocytes, respectively.

Fibroblasts and primary CD14+ monocytes were infected simultaneously with the same TB40E-GFP tagged HCMV strain and were harvested for RNA-sequencing (RNA-seq) over the time course of 4, 8, 12, 24, 48, and 72 hours post infection (hpi). In order to determine immediate early genes, fibroblasts and CD14+ monocytes were treated with CHX, an inhibitor of protein synthesis, and these samples were harvested for RNA-seq at 8 hpi. Additionally, to define true late genes, we treated infected fibroblasts and CD14+ monocytes with PFA, a viral DNA replication inhibitor, and harvested for RNA-seq at 24/48 and 72 hpi, respectively (Figure 1A).

To evaluate the reproducibility of our experiment, we prepared two independent biological replicates for all time points, treatments, and cell types. Biological replicates grouped well together when applying hierarchical clustering analysis on the Pearson correlations of host and viral transcripts, between all samples (Supplementary Fig. S1A). Interestingly, replicates from all time points and treatments in latently infected CD14+ monocytes strongly correlate, indicating there are few changes in the transcriptome. In contrast, for infected fibroblasts each time point, or treatment represents a more distinct state along productive infection (Figure 1B, Supplementary Fig. S1A and B).

We calculated the percent of HCMV transcripts along infection in the lytic and latent models. In lytic infection, viral gene expression steadily increases between 4 and 72 hours, where viral transcripts compose up to half of the total mRNA by 72 hpi. As expected, in the 72 hpi PFA treated fibroblast samples viral transcripts make up only around 10% of the total mRNA, likely due to less viral DNA template (Figure 1C). In stark contrast, viral transcripts compose less than 0.25% of the total polyadenylated RNA in latently infected monocytes and the percentage of viral transcripts continuously decreases from the earliest points of monocyte infection through 72

hpi. Surprisingly, PFA treated latently infected monocytes expressed slightly higher levels of viral genes likely due to changes in the host upon drug treatment (Figure 1D).

HCMV gene expression kinetics along lytic infection reveals complex, multi-faced regulation

We first analyzed viral gene expression kinetics in lytically infected fibroblasts. After filtering for minimal expression, we were able to reliably quantify the majority of viral genes (134 genes). For each viral gene we calculated its relative expression pattern out of total viral reads in a given time point or treatment. We performed hierarchical clustering based on the normalized viral gene expression, dividing viral genes into seven distinct temporal classes (Figure 2A). We term these clusters TC1 to TC7 (Figure 2B, Table S1). There was generally good correspondence between classical temporal classes and our measurements. For example, the two canonical immediate early genes UL123 and UL122[25] were classified as TC1 and were already highly expressed at 4 hpi and their expression was not affected by translation inhibition with CHX. The genes in clusters TC2 and TC3 have early and delayed early kinetics, and contain many genes classically defined as early. The late clusters (TC5-TC7) contain mostly genes classically defined as late. The known late genes UL94 [26], UL75 [27,28], and UL32 [29] were classified as TC7 which showed late expression kinetics (peak expression 48-72 hpi) and their expression was diminished when viral DNA replication was inhibited by PFA (Figure 2A). In general, for many of the viral genes there was agreement between the mRNA temporal class we defined and previously determined protein temporal classifications[10] (Figure 2A, lower panel).

It is important to note that the seven TCs reflect how each viral gene is expressed relative to all other viral genes in the same sample. Looking at the normalized expression levels of viral genes in RPKM (reads per kilobase of transcript, per million mapped reads), which more closely reflects their absolute quantities, almost all viral genes are most highly expressed at 72 hpi regardless of the TC in which they were classified (Supplementary Fig. S2). This is expected due to viral DNA replication that leads to high levels of template DNA and a corresponding increase in viral transcription levels at later stages of infection. A unique outlier is UL123 (encoding for IE1) whose highest expression was measured at 8 hpi, and in absolute terms it was hardly expressed at 72 hpi (Supplementary Fig. S2). This unique expression pattern points to the critical

role IE1 plays at the early stages of the HCMV life cycle and to the distinct regulation it is subjected to.

Importantly, we noticed that for numerous viral genes, the temporal expression along the course of infection and the expression under CHX and PFA treatments do not fully align, revealing more than one regulatory mode of expression. For example, TC4 is composed of genes that show early expression kinetics as they are already highly expressed at 8 hpi but these genes also show dependency on DNA replication as PFA treatment greatly diminishes their expression and consequently, they demonstrate a second wave of expression at 72 hpi. Interestingly viral genes that are considered canonical late genes like UL99 (encoding for pp28) and UL80 (encoding for capsid scaffold protein) are part of this group. TC5 represents a group of genes with late expression kinetics which resembles the kinetics of TC6 and TC7, but the relative expression of genes in TC5 did not depend on DNA replication and was unaffected by PFA treatment (Figure 2A, upper panel). Furthermore, genes in TC6 which includes UL83 (encoding for the major tegument protein pp65)[30,31] and UL86 (encoding for the major capsid protein)[32] exhibit late kinetics and dependency on DNA replication, with their highest relative expression at 48-72 hpi and this expression is drastically reduced upon PFA treatment. Surprisingly, however, these genes also exhibit expression that is independent of de novo protein synthesis upon CHX treatment (Figure 2A, upper panel). These results reveal a clear deviation from the classical definition of herpesvirus gene expression cascade which traditionally relies on the dependence on protein synthesis and sensitivity to viral DNA replication inhibitors. This indicates that the canonical definition and naming is oversimplified and even misleading; for many viral genes there are multiple modules that regulate their expression. We therefore propose a division in to seven temporal classes that incorporate the canonical nomenclature but reflects better the complexity of viral gene expression: TC1: Immediate Early, TC2 -Early, TC3-Delayed early, TC4: Early-late, TC5-Late-replication independent, TC6-Late-translation independent, TC7-Late (Figure 2B).

A group of late HCMV transcripts is also expressed in a protein synthesis–independent manner

The analysis of immediate early viral gene expression, especially in the presence of CHX, can be hindered by the possible existence of virion-associated input RNA nonspecifically bound by or packaged within the virus particles and delivered to newly infected cells[33]. Indeed, many of the transcripts that presented late expression kinetics (TC5-TC7) are also relatively highly expressed at 4 hpi. Since TC6 presented such an unexpected expression profile composed of transcripts with late kinetics as well as expression that is independent of protein synthesis, we wanted to rule out the possibility that these transcripts might represent unique, stable RNA species which entered the infected cell with the virus as "input" RNA and remained intact for 8 hours, the time point at which we harvested the CHX samples. In order to assess if the expression of these genes at 8 hpi is independent of transcription, HCMV infected fibroblasts were treated with actinomycin D (ActD), an RNA polymerase inhibitor, at 2 hpi and RNA was harvested at 4 and 8 hpi (2 and 6 hours post ActD treatment). Using real time quantitative PCR analysis for TC6 transcripts, we show that the expression of these transcripts, such as UL83 and UL84, is abolished by 8 hpi under ActD treatment, indicating that these transcripts are not uniquely stable and have fully degraded by 8 hpi and hence our measurements at 8 hpi represent true viral RNA transcription taking place in the newly infected cells (Figure 3A). We also repeated this experiment and analyzed the expression of these genes relative to total viral gene expression at different time points following CHX treatment. Indeed, we detected a higher relative expression of TC6 transcripts (UL83 and UL84) upon CHX treatment, whereas the late transcripts UL32 and UL99 did not show a relative increase in their expression (Figure 3B). Overall, these results illustrate de novo expression of traditional late genes in a protein synthesis independent manner and provide further support to the validity of the new temporal classes we have identified.

High levels of input RNA at 4hpi

The relatively high level of late viral transcripts in the RNA-seq samples at 4 hpi suggested a portion of the transcriptome at 4 hpi likely originates from virion-derived RNA, rather than RNA that had been produced following infection. In order to better distinguish virion-associated input RNA from genuine transcription, we performed RNA-seq of both HCMV infected human

fibroblasts and CD14+ monocytes at 1-hour post infection, a time point which likely precedes the onset of expression for most viral genes. The expression profiles between 1 and 4 hours post infection in both systems highly correlate with each other (Figure 4A), indicating the significant presence of input RNA even at 4 hours post infection. However, in lytic infection (Figure 4A, left panel) TC1 genes, which are expressed with immediate early kinetics, show significantly higher expression at 4 hpi than at 1 hpi, illustrating clear transcription of this set of genes by 4 hours post infection. In contrast, in CD14+ monocytes we see no increased expression of specific transcripts at 4 hpi compared to 1 hpi, illustrating that if transcription of viral IE genes occurs at 4 hpi, it is low and masked by a relatively high level of input RNA (Figure 4A, right panel). We examined whether the lack of increased transcript levels at early time points post infection, together with the gradual decrease in the percent of HCMV transcripts along infection in monocytes (Figure 1D), might indicate large quantities of input RNA and no de-novo viral transcription in this system. Calculation of the decay rate of viral mRNA levels across infection in the infected CD14+ monocytes demonstrates that the decay rate of viral mRNAs is drastically reduced between 8 and 12 hpi (Figure 4B). This suggests that the source of viral reads in CD14+ monocytes is likely initially dominated by the decaying input RNA but that early in infection there is also onset of low-level RNA synthesis, which eventually dominates the measured viral transcripts.

Latently infected CD14+ monocytes express genes that resemble the late lytic profile throughout infection

Although the latency transcriptome profile at early time points post infection was masked by input RNA, we examined if we could capture differential expression of specific temporal classes of genes along the later time points in latent infection. We compared the expression profiles measured at 12, 24, 48 and 72 hpi in CD14+ monocytes to the expression we measured in late lytic infection. We obtained a significant correlation between viral gene expression in infected monocytes with the late lytic expression in fibroblasts and the correlation coefficient dropped over time due to the gradual reduction in the number of viral reads (Figure 4C). The only genes whose expression was significantly higher in monocytes compared to their relative abundance in fibroblasts were RNA2.7 and UL22a, both of which are highly expressed transcripts, and their high levels are likely related to their high mRNA stability. We further analyzed the expression

kinetics of the seven temporal classes of gene groups we defined in lytic infection in the CD14+ data set. As expected from the general decline in viral reads in latent monocytes, the expression of all TC classes declined with time (Figure 4D, upper panel). In contrast to the temporal dynamic which is clearly evident in lytic cells (Supplementary Fig. S3), the relative proportion of reads from each TC remains relatively constant throughout the course of latent infection (Figure 4D, lower panel) implying little to no changes in the overall expression profile. This relatively fixed profile likely results from residual input RNA in combination with low-level transcription which may reflect inherent susceptibilities of different loci of the viral genome to the host transcription machinery.

Diverse epigenetic inhibitors induce viral gene expression in infected monocytes

The cellular environment seems to be a key factor in determining the outcome of HCMV infection. Epigenetic regulation plays an important role in latent infection, where the viral genome is chromatinized and maintained as a repressed episome [34,18]. To verify that the expression we captured in latently infected CD14+ monocytes at 72 hpi indeed reflects low-level transcription, we asked if we could enhance viral gene expression in infected CD14+ monocytes by interfering with host chromatin modifying factors. If so, we would expect to see increased expression of specific viral genes, with variable effects between drugs that target different gene silencing pathways.

To probe a large range of molecules we performed an unbiased and high-throughput inhibitor screen for epigenetic regulators that affect HCMV gene expression. Primary CD14+ monocytes were infected with an HCMV strain TB40E virus encoding a green fluorescent protein (GFP) tag under a simian virus 40 (SV40) promoter (TB40-GFP) [35,36]. 48 hours post infection, at a time when much of the overload of input RNA has degraded and viral gene expression largely reflects low-level transcription, we incubated the latently infected cells with 140 different small molecule epigenetic inhibitors, in biological replicates. The inhibitors we tested fell within a broad range of categories, including inhibitors of bromodomains, histone deacetylases, histone methyltransferases, histone demethylases, and others (Figure 5A). The final concentration we used in this screen was set to 1uM based on calibration of six different drugs targeting different functional categories with the aim to reduce drug cytotoxicity (Supplementary Fig. S4). 24 hours

following treatment with the inhibitors (a time point we presumed is likely to reflect the immediate changes in viral gene expression), we screened by flow cytometry for changes in GFP expression as a proxy for changes in viral gene expression compared to a DMSO treated control. For the vast majority of epigenetic inhibitors, biological replicates showed little variation (Figure 5B) and treatments induced little to no changes in the GFP expression. The 15 drugs that fell within the top ten percent for fold change in GFP expression, compared to the control (Figure 5C and 5D) were enriched in HDACs, such as Trichostatin A (TSA) [37-38], Sodium Butyrate [39,40] and MC1568 [41], and also included the Sirtuin inhibitor AGK2 [42-43] all of which have been reported to enhance herpesvirus gene expression. Additionally, these drugs included UNC0631, histone methyltransferase inhibitor, and Etoposide and Mirin, inhibitors of the DNA damage pathway[44] (Figure 5E).

In order to verify that these drugs indeed elevate viral gene expression and to examine the gene expression profile that is induced, we performed transcriptome analysis of infected CD14+ monocytes following treatment with 20 drugs from the screen, most with biological replicates. The drugs that were used included the top 10% hits of our flow cytometry based screen as well as several drugs from the top 20% hits that fall into additional functional categories such as poly-ADP ribose polymerase inhibitors. Reassuringly, treatment with most of the inhibitors elicited some degree of increase in viral gene expression compared to the DMSO treated control (Figure 5F), with Trichostatin A and Cay10398, two HDAC inhibitors and Mirin, an ATM pathway inhibitor inducing the most significant increase in viral gene expression (13.6-fold, 5.5-fold and 10.3-fold correspondingly).

Transcription induction in infected monocytes reveals that the hallmark of HCMV latency is distinctive repression of IE genes

To examine which viral genes were affected by each of these drugs we compared the gene expression profiles of HCMV infected monocytes treated with epigenetic drugs to gene expression in the control sample. Importantly, viral gene expression following all the epigenetic drug treatments, regardless of their specific target and the levels of viral gene induction, were highly correlated with the control sample (Figure 6A). Statistical analysis showed almost no viral genes that were distinctively induced by each of these drugs. This overall resemblance in viral

gene expression between untreated monocytes and monocytes treated with diverse drugs that work by different mechanisms points to two critical conclusions: 1. Although there are substantial amounts of input RNA upon infection, viral RNA expression at 72 hpi mostly reflects low level transcription that can be further enhanced. 2. Upon infection of monocytes, viral gene expression is repressed, and this repression is consistent along the viral genome, such that interference with different chromatin modifiers results in overall uniformly enhanced expression of most viral genes.

Since viral gene expression is overall very low (between 0.4 to 4% of expressed mRNAs are viral), relatively small changes in a limited number of genes might be difficult to detect and quantify. Therefore, we also analyzed the changes in expression of the temporal classes across the different drugs. This pooled analysis revealed an invariable increase of the different TCs (Figure 6B). Interestingly, a common signature of most drugs was the relatively low induction of TC1 expression (IE genes) (Supplementary Fig. S5), and this effect was statistically significant when analyzed across all drugs (Figure 6B). The relative uniform increase of early and late genes 24 hours after drug treatments indicates that this profile still does not reflect reactivation, rather enhancement of transcription in latent monocytes. This finding indicates that a defining property of the latent transcriptome in monocytes is the repression of IE transcripts that differs from the general repression of viral gene expression, thus alluding to unique regulation of IE genes in latent monocytes.

HCMV gene expression induction in latent monocytes leads to reduction in the expression of immune related genes.

To probe the effects of these drugs more accurately, we analyzed the host gene expression landscape following treatment with the diverse epigenetic modifiers. Gene set enrichment analysis (GSEA) shows that compared to untreated infected monocytes, drug-treated monocytes in which viral gene expression was induced by more than 2-fold, show significant reduction in many immune response related pathways including TNFa signaling, IFN response and inflammatory response as well as reduction in pathways related to cell cycle progression and apoptosis (Figure 7A). Remarkably, although the drugs we have used target diverse chromatin modifiers, the changes in host gene expression were common to all drugs regardless of the

specific protein that was targeted. These changes coincided with the increase in viral gene expression and monocytes that were treated with BSI201, which did not lead to a significant increase in viral gene expression, showed only marginal reduction in a limited number of these pathways (Figure 7A). These common changes in cellular gene expression suggest that most of these do not stem from the direct effects of the drugs but rather likely reflect secondary effects that relate to the increase in viral RNA or protein expression. We also observed a correlation between the increase in viral gene expression and the magnitude of reduction in genes related to immune pathways such as IFNa, IFNg responses and TNFa signaling (Figure 7B). This inverse correlation further supports that the enhancement in viral gene expression accounts for a major portion of these changes in cellular gene expression.

To further examine the kinetics of viral gene expression induction, and specifically to verify that the low induction of TC1 genes is not related to the time point post drug treatment, we analyzed gene expression at different time points after drug treatment. In this experiment, we focused on TSA. TSA treatment led to the most significant increase in viral gene expression, providing us with a maximal dynamic range. 48 hpi, CD14+ monocytes were treated with TSA or DMSO as a control. Cells were harvested in replicates for RNA-seq at 0, 4, 12 and 24h post TSA treatment. We observed a small (1.2-fold) but significant increase in viral expression already at 4h post TSA treatment and the induction in viral gene expression increased with time reaching 14-fold at 24h post treatment (Figure 8A). In agreement with our previous observations, viral gene expression patterns in infected monocytes treated with TSA were highly correlated with the untreated samples at all time points post treatment (Figure 8B). These measurements illustrate that in the early time points post treatment, TSA does not lead to dynamic changes in viral gene expression but rather to enhancement of the same initial viral transcription that exists in infected monocytes. We next grouped viral transcripts according to their TC and analyzed the induction of different TC gene groups across the time points. Already at 4 hr and also at 12 hr, it was evident that TC1 genes are not induced by TSA treatment, while the expression of viral genes of other TCs already begins to increase. At 24h post TSA treatment, when viral gene induction was the strongest, expression of TC1 genes was induced but to a significantly less extent than genes from other TCs (Pval<0.001, Figure 8C). This analysis shows that the induction of viral gene expression in infected monocytes by epigenetic modifiers is associated with the relative unresponsiveness of IE genes indicating an inherent difference in the repression mechanism

between IE genes and other genes in latent monocytes. We analyzed the changes in cellular gene expression as well, revealing again significant reduction in similar immune response related pathways including TNFa signaling, IFN response and inflammatory response in TSA treated samples over time (Supplementary Fig. S6). When we examined the kinetics of host genes, we observed that down regulation of IFNa and IFNg pathway coincides with the increase in viral gene expression (Figure 8D). Overall, these results demonstrate that diverse epigenetic drug treatments lead to a uniform increase in viral gene expression with the exception of IE genes, which appear to be under additional repressive regulation.

Overexpression of UL123 in THP1 monocytic cells is sufficient to drive productive infection

The specific repression we observed for IE transcripts, together with the unique expression pattern of UL123 (Supplementary Fig. S2), prompted us to examine whether the repression of IE transcripts and specifically of UL123 serves as a functional barrier for the establishment of lytic infection. To this end, we used THP1, a monocytic cell line that serves as a model for HCMV latency and transduced these cells with lenti-virus encoding UL123 (encoding for IE1 protein) or mCherry that was used as a control (Supplementary Figure S7). For these experiments we also utilized THP1 cells that co-express PDGFRα (THP1-PDGFRα), as PDGFRα expression was shown to greatly enhance HCMV entry and infection of THP1 cells[45]. THP1 and THP1-PDGFRα cells expressing mCherry (control) or IE1 were infected with TB40-GFP and infected cells were analyzed by flow cytometry at 72 hpi. Whereas infection of THP1 that expresses mCherry (control) or IE1 showed mostly dim expression of GFP representing repressed expression of viral transcripts, infection of THP1-PDGFRα cells resulted also in a small population of cells that express a bright and distinct GFP signal, and this population rose to 55% in cells overexpressing IE1 (Figure 9A). These results indicate that in the presence of PDGFRα, IE1 over-expression is enough to support productive infection. Indeed, digital droplet PCR analysis of viral DNA confirmed that ectopic expression of IE1 led to viral DNA replication compared to control THP1-PDGFRα cells or THP1 cells not expressing PDGFRα (Figure 9B). Altogether these results indicate that, following entry, IE1 expression in THP1 cells is the principal barrier that needs to be crossed to achieve a full productive cycle.

Discussion

Using transcriptome wide sequencing, we performed a dense time course along both lytic and latent HCMV infection. We defined kinetic classes for the majority of HCMV transcripts along productive infection, revealing novel patterns of viral gene expression. We discovered that in contrast to the traditional view of herpesvirus gene expression, the dependency of viral transcripts on metabolic drugs and viral transcripts' temporal dynamics are often decoupled, implying that expression kinetics and the dependency on protein synthesis or viral DNA replication are frequently two independent properties in viral gene expression regulation. A recent study on alpha herpesvirus, varicella-zoster virus (VZV), which mapped VZV transcripts also examined their expression kinetics and observed unexpected patterns of viral gene expression[46]. In addition to the traditional IE, Early and L kinetic classes, they defined groups of genes as Early-Late and Transactivated/True-Late (TA/TL)[46] . In their analysis, Early-Late genes show two waves of expression, one of which is dependent on de novo protein synthesis and a second, later wave of expression which depends on the onset of DNA-replication. These genes are comparable to the HCMV genes characterized here in TC4 which we also name Early-Late. TA/TL genes correspond in expression dynamics to our TC6: Late-translation independent gene cluster. These genes show low level expression in the absence of de novo protein production, but their predominant expression occurs at late time points post infection and is highly dependent on viral DNA replication. The authors hypothesized that the production of these transcripts in CHX-treated samples might be due to low-level transactivation by viral tegument proteins delivered with incoming virions. Whether our TC6 genes can be expressed independently of viral proteins or whether they in fact depend on incoming viral tegument proteins is an important question which warrants further research. Regardless, the ability of these genes to be expressed early in infection suggests they play an additional important role in the early stages of infection.

More globally, the wide use in the herpesvirus literature of the terms "leaky late" and "true late" which refer to late genes whose expression is augmented by or totally dependent on DNA replication respectively, hints towards a more intricate expression regulation than initially thought. This together with the resemblance between the expression patterns we uncover here to the one described for VZV [46] strongly points to a more complex expression regulation across all herpesviruses. We therefore propose that the patterns we describe here, which are based on both

regulation and temporal expression, may better capture the multiple modules herpesviruses utilize to tightly regulate their gene expression. One clear implication of these more complex expression patterns is that they do not represent a simple sequential cascade as the traditional IE, E, L model. Therefore, a deeper understanding of herpesvirus gene regulation will be needed to decipher the connectivity between these different expression modules and how they come into play throughout productive infection.

During HCMV latency there is massive repression of viral gene expression. We revealed that at early time points post infection the latent transcriptome, in experimental models, is highly dominated by virion-associated input RNA. Looking at viral mRNA decay rate, we can infer that the percent of viral reads in infected CD14+ monocytes is set by the decay of virion associated input RNA together with the onset of low-level RNA synthesis. In order to clearly distinguish newly synthesized mRNAs from background input RNA noise and to unambiguously determine transcription patterns at early time points during latency establishment, it would be necessary to perform metabolic labeling, which allows the detection and quantification of newly synthesized RNA species [47–49]. However, since the levels of viral gene expression are incredibly low and metabolic labeling approaches facilitate labeling of only a small portion of the newly synthesized RNA pool, conducting these experiments will be challenging. When we analyze the relative proportion of reads from each TC in monocytes from 12 hpi and onwards, it remains largely constant. We hypothesize this latency transcriptome represents overall uniform repression with default transcription tendencies of the viral genome dictating the low-level transcription.

To convincingly show this low-level viral gene expression represents true transcription, we screened a wide array of chromatin modifiers and tested their effects on viral gene expression. This screen revealed compounds from several categories that enhanced viral gene expression, including inhibitors of HDACs, Sirtuins and DNA damage response, which were previously implicated in regulation of viral gene expression during herpesvirus latency [44,50–53]. Although these drugs work by diverse mechanisms, activating different changes to the chromatin, characterization of the viral gene expression enhanced by these drugs reveals that they all lead to induction of the same transcriptional program. Importantly, the only distinctive group was TC1-IE genes, as induction of IE genes was significantly less prominent than induction of viral genes from all other TC gene clusters. This implies that repression along the HCMV genome in latent

cells is uniform with an additional and unique repression of IE genes. When analyzing host gene expression due to treatment with this diverse set of drugs, we observed recurrent down regulation of the host innate immune response pathway. In agreement with our previous findings[54], this indicates that viral gene expression does have functional consequences and suggests a constant arms race between viral mRNA or viral protein expression and expression of host innate immune response genes during infection and latency.

Inspired by our observations which unbiasedly points to the unique regulation of IE genes, we show that the ectopic expression of the central IE transcript UL123, encoding for IE1, is indeed sufficient to drive a full productive cycle in non-permissive cells which serve as a model for HCMV latency. Previously, ectopic expression of IE1 and IE2 in non-permissive cells led to an increase in early gene expression, however, it did not result in viral DNA replication and viral progeny[19]. These results led to the perception that there are additional barriers needed to be crossed for successful productive infection. This, together with the controversies about the nature of the latency transcriptome, made HCMV latency and reactivation extremely enigmatic processes. Our work suggests that the other main barrier is simply the number of viral genomes which effectively enter into infected THP1 cells. We show that PDGFRα expression, which is a central HCMV entry receptor [45], is enough to induce viral gene expression and viral DNA replication in infected THP1 cells. Once this entry threshold is lowered, sufficient IE1 expression is the main impediment to achieving HCMV lytic expression in these monocytic cells. Given the established necessity of the MIEP and IE proteins in initiating productive infection, our results support a parsimonious explanation; IE1 expression represents the main barrier for productive infection and for reactivation. These findings also illustrate that HCMV latency resembles other herpesviruses in which it was shown that the switch between latent and lytic expression is primarily regulated by IE transactivators. For herpes simplex virus (HSV), ectopic expression of either IE protein ICP0 or ICP4 induced production of infectious virus in a murine trigeminal ganglia latency model [55]. As well for Kaposi's Sarcoma associated herpesvirus (KSHV) deregulated expression of ORF50, which encodes a major transactivator of delayed-early viral genes, suffices to induce the lytic gene cascade in latently infected B cells [55,56]. Interestingly, generation of HCMV mutants previously demonstrated that deletion of exon 5 which defines the UL122 gene, encoding for IE2, results in a nonviable virus[57,58], whereas IE1 deletion viruses exhibit a growth defect only at a low MOI but are able to efficiently replicate at high MOIs [59].

Consequently, IE2 was thought to be the key viral transactivator, as it is absolutely required for lytic HCMV infection. Our data points to a critical role for IE1 as the principal transactivator and further experiments addressing the role of IE2 in latency models will provide insight into the unique effects of each of these two splice variants. Another intriguing question is whether the distinctive repression we observed for IE genes is a unique property of monocytes, or does the same phenomenon exist in other cell types. Deciphering the molecular determinants that drive the differential repression in diverse cell types will help to shed light on the role of the cell environment in determining infection outcome.

In summary, our findings suggest that herpesvirus temporal gene expression cascade is dictated by a more intricate set of dependencies than was previously appreciated and that as envisioned for many years, the hallmark of HCMV latency is the unique repression of IE gene expression.

Acknowledgments

We thank Stern-Ginossar lab members, Oren Kobiler and Igor Ulitsky for providing valuable feedback. We thank Thomas Shenk for the PDGFRa expressing THP1 cells, and the Weizmann flow cytometry units for technical assistance. This study was supported by a European Research Council consolidator grant to N.S-G (CoG-2019-864012). N.S-G is an incumbent of the Skirball Career Development Chair in New Scientists and is a member of the European Molecular Biology Organization (EMBO) Young Investigator Program. The authors declare no competing interests.

Materials and Methods:

Cells and viruses

Human foreskin fibroblasts (ATCC CRL-1634) were grown in Dulbecco's Modified Eagle's Medium (DMEM) with 10% heat-inactivated fetal bovine serum (FBS), 2 mM L-glutamine, and 100 units/ml penicillin and streptomycin (Beit-Haemek, Israel) and maintained at 37°C in a 5% CO2 incubator. THP1 cells (ATCC TIB-202) were grown in Roswell Park Memorial Institute

(RPMI) with 10% heat-inactivated FBS, 2 mM L-glutamine, and 100 units/ml penicillin and streptomycin (Beit-Haemek, Israel). Primary CD14+ monocytes were isolated from fresh venous blood, obtained from healthy donors, using a Lymphoprep (StemCell Technologies) density gradient followed by magnetically activated cell sorting with CD14 magnetic beads (Miltenyi Biotec). CD14+ cells were cultured in X-Vivo 15 medium (Lonza) supplemented with 2.25 mM L-glutamine and maintained at 37°C in a 5% CO2 incubator. The bacterial artificial chromosome (BAC) containing the clinical strain TB40E[60] with an SV40-GFP tag (TB40E-GFP) was described previously [35,36]. This strain lacks the US2-US6 region, therefore, these genes were not included in our analysis. Virus was propagated by adenofection of infectious BAC DNA into fibroblasts [61]. Viral stocks were concentrated by ultracentrifugation at 30,000xg at 4°C for 120 min. Infectious virus yields were assayed on fibroblasts and THP1 cells.

Infection procedures

For the time course analysis fibroblasts and CD14+ monocytes were infected with the same stock of HCMV strain TB40E-GFP at a multiplicity of infection (MOI) of 1 and 10, respectively. For all other experiments CD14+ monocytes or THP1 cells were infected with HCMV strain TB40E-GFP at an MOI of 5. Cells were incubated with the virus for 2h, washed, and supplemented with fresh medium. Infection was monitored 2-3 dpi by measurement of GFP expression on a BD accuri flow cytometer and analysis by FlowJo. Substantial shift in GFP levels, seen in most infected fibroblasts is indicative of productive infection, while a small shift in GFP levels, as seen in all CD14+ monocytes indicates that the cells were infected but the virus is repressed, as expected in latent infection.

Cell treatments

CHX was added to infected fibroblasts and CD14+ monocytes immediately after infection at a final concentration of 100 ug/ml and 200 ug/ml or at 100 ug/ml, respectively. PFA was added to fibroblasts and CD14+ monocytes immediately after infection at a final concentration of 400 ug/ml. ActD was added to infected fibroblasts at a final concentration of 5uM.

RNA library construction

For RNA-seq time course experiments in fibroblasts and CD14+ monocytes, cells were washed with PBS and then collected with Tri-Reagent (Sigma-Aldrich), total RNA was extracted by phase separation and poly(A) selection was performed using Dynabeads mRNA DIRECT Purification Kit (Invitrogen) according to the manufacturer's protocol. RNA-seq libraries were generated as previously described [62]. Briefly, mRNA samples of ~4ng were subjected to DNaseI treatment and 3′ dephosphorylation using FastAP Thermosensitive Alkaline Phosphatase (Thermo Scientific) and T4 PNK (NEB) followed by 3′ adaptor ligation using T4 ligase (NEB). The ligated products were used for reverse transcription with SSIII (Invitrogen) for first-strand cDNA synthesis. The cDNA products were 3′ ligated with a second adaptor using T4 ligase and amplified with 8 cycles of PCR for final library products of 200–300 base pairs. For epigenetic drug-treated CD14+ monocyte samples, RNA libraries were generated from samples of ~100,000 cells according to the MARS-seq protocol[63].

Next-generation sequencing and data analysis

All RNA-Seq libraries (pooled at equimolar concentration) were sequenced using Novaseq6000 (Illumina), with read parameters: Read1: 72 cycles and Read2: 15 cycles. For the time course libraries, raw sequences were first trimmed at their 3′ end, removing the illumina adapter and polyA tail. Alignment was performed using Bowtie (allowing up to 2 mismatches) and reads were aligned to concatenation of the human (hg19) and the viral genomes (NCBI EF999921.1). Reads aligned to ribosomal RNA were removed. Reads that were not aligned to the genome were then aligned to the transcriptome.

Analysis of libraries generated with the MARS-seq protocol was done as previously described [23]. Briefly, 37-bp reads were aligned using Bowtie (allowing up to 2 mismatches) to concatenation of the human (hg19) and the viral genomes (NCBI EF999921.1). Counting of reads per gene was done based on unique molecular identifiers (UMIs) (8 bp). The transcription units of the virus were based on NCBI annotations, with some changes, including merging several transcripts (considering that the library maps only the 3' ends of transcripts) and adding some antisense transcripts. All analyses and figures were done using in-house R-scripts. Epigenetic drug treated CD14+ monocyte samples which had less than 100,000 UMIs were left out from further analysis leaving three drug treatments without duplicates- Cay10398, CPTH2, and Gemcitabine.

Inhibitor screen

The compounds used in the epigenetic inhibitor screen were purchased from Cayman Chemical as a complete Epigenetics Screening Library (96-Well, Item No. 11076). To calibrate cytotoxicity, HCMV infected CD14+ monocytes were treated with six inhibitors at increasing concentrations (500nM,1uM,10uM) for 48 hours. Cells were washed, resuspended in PBS and stained with 0.5ug/ml of PI for 10 minutes prior to analysis by FACs- LSRII. The 1μM concentration was chosen for the complete screen as it caused minimal cytotoxicity. For the inhibitor screen at 48 hpi, TB40-infected CD14+ monocytes were divided into 1.2 ml tubes containing ~100,000 cells per tube and inhibitors or DMSO for negative control were added to a final concentration of 1uM. After 24 hours of incubation, Flow cytometry was performed using an LSRII, and analysis was performed using FlowJo software.

Differential expression and enrichment analysis

Differential expression analysis was done with DESeq2 (version 1.22.2)[64] using default parameters, with the number of reads in each of the samples as an input. The normalized number of reads according to DESeq2 were used for enrichment analysis using GSEA (version 4.0.3) [65]. The MSigDB hallmark (version 7.1) gene sets were used [66]. The GSEA plots were created based on the GSEA output with in house R scripts using R package enrichplot.

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. Real time PCR was performed using the SYBR Green PCR master-mix (ABI) on the QuantStudio 12K Flex (ABI) with the following primers (forward, reverse):

UL123 (TCCCGCTTATCCTCAGGTACA, TGAGCCTTTCGAGGACATGAA)

UL84 (TCAAAGCATACGCTGAATCG, GCTTACAGTCTTGCGGTTCC)

UL83 (CCAGCGTGACGTGCATAAAG, TCGTGTTTCCCACCAAGGAC)

UL44 (AGCAAGGACCTGACCAAGTT, GCCGAGCTGAACTCCATATT)

UL99 (GGGAGGATGACGATAACGAG, TGCCGCTACTACTGTCGTTT)

UL32 (GGTTTCTGGCTCGTGGATGTCG, CACACAACACCGTCGTCCGATTAC)

RNA 2.7 (TCCTACCTACCACGAATCGC, GTTGGGAATCGTCGACTTTG)

B2M (TGCTGTCTCCATGTTTGATGTATCT, TCTCTGCTCCCCACCTCTAAGT)

Plasmids and lentiviral transduction

UL123 was cloned into the pLVX-EF1alpha-SARS-CoV-2-nsp1-2XStrep-IRES-Puro (Addgene plasmid #141367) in place of the SARS-CoV-2-nsp1-2XStrep cassette. The UL123 coding sequence was amplified from cDNA of HCMV infected fibroblasts with the primers detailed below, containing flanking regions homologous to the vector. The vector was amplified with the primers detailed below. The amplified PCR fragments were cleaned using a gel extraction kit (Promega) according to the manufacturer's protocol and the UL123 fragment was cloned into the vector using a Gibson assembly protocol. The control plasmid used was a lentiviral vector encoding mCherry.

Primer	Sequence
UL123_Fw	CGGGATCCTTACTGGTCAGCCTTGCTTCTAGT
UL123_Rv	CGCCACCATGGAGTCCTCTGCCAAGAGAAAG
Vector_Fw	ACCAGTAAGGATCCCGCCCC
Vector_Rv	GACTCCATGGTGGCGGCGAA

Lentiviral particles were generated by co-transfection of the expression constructs and 2nd generation packaging plasmids (psPAX2, Addgene#12260 & pMD2.G, Addgene#12259), using jetPEI DNA transfection reagent (Polyplus transfection) into 293T cells, according to manufacturer’s instructions. 48 hours post transfection, supernatants were collected and filtered through a 0.45µm PVDF filter (Millex). THP1 cells were transduced with lentiviral particles by

spinfection (800xg, 1hr) and were then puromycin-selected (1ug/mL) for 4 days. Puromycin was removed prior to subsequent HCMV infection. Expression of IE1 was measured by IE1 immunostaining.

IE1 immunostaining

THP1 cells were fixed with 4% formaldehyde for 15 minutes, permeabilized in 0.1% triton x100 for 15 minutes and stained with an Alexa Fluor 488-conjugated IE1 antibody (Clone 8B1.2, Millipore). Flow cytometry was performed on a BD Accuri flow cytometer and analysis was done using FlowJo.

Digital Droplet PCR

Detection of viral DNA in monocytes was done using the QX200 droplet digital PCR system (Bio-Rad), using FAM labeled HCMV primer and probe: Human CMV HHV5 kit for qPCR using a glycoprotein B target (PrimerDesign); and HEX labeled RPP30 copy number assay for ddPCR (Bio-Rad), as previously described [67]. Calibration curve was run in duplicate, using CMV positive control template (PrimerDesign). The limit of detection was 3 events per sample, with accuracy improved at 10 copies and higher. For sample preparation cells were counted and dry pellet was stored at −80°C prior to DNA extraction. DNA was extracted from the cell pellet 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, 0.25% Tween 20, 0.25% Non-idet P-40, and 0.4 mg/ml Proteinase K), for 60 min at 60°C followed by a 10 min 95°C incubation, according to the description in [68].

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. Crough, T. & Khanna, R. Immunobiology of human cytomegalovirus: from bench to bedside. *Clin. Microbiol. Rev.* **22**, 76–98, Table of Contents (2009).
3. Stern-Ginossar, N. *et al.* Decoding human cytomegalovirus. *Science* **338**, 1088–1093 (2012).
4. Knipe, D. M. *et al.* Fundamental Virology, 4th Edition; and Fields Virology, 4th Edition, Volumes I and II:Fundamental Virology, 4th Edition;Fields Virology, 4th Edition, Volumes I and II. *Clinical Infectious Diseases* vol. 34 1029–1030 (2002).
5. DeMarchi, J. M., Schmidt, C. A. & Kaplan, A. S. Patterns of transcription of human cytomegalovirus in permissively infected cells. *J. Virol.* **35**, 277–286 (1980).
6. McDonough, S. H. & Spector, D. H. Transcription in human fibroblasts permissively infected by human cytomegalovirus strain AD169. *Virology* **125**, 31–46 (1983).
7. Wathen, M. W. & Stinski, M. F. Temporal Patterns of Human Cytomegalovirus Transcription: Mapping the Viral RNAs Synthesized at Immediate Early, Early, and Late Times After Infection. *Journal of Virology* vol. 41 462–477 (1982).
8. Chambers, J. *et al.* DNA microarrays of the complex human cytomegalovirus genome: profiling kinetic class with drug sensitivity of viral gene expression. *J. Virol.* **73**, 5757–5766 (1999).
9. Gatherer, D. *et al.* High-resolution human cytomegalovirus transcriptome. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 19755–19760 (2011).
10. Weekes, M. P. *et al.* Quantitative temporal viromics: an approach to investigate host-pathogen interaction. *Cell* **157**, 1460–1472 (2014).
11. Schwartz, M. & Stern-Ginossar, N. The Transcriptome of Latent Human Cytomegalovirus. *J. Virol.* **93**, (2019).

12. Murphy, J. C., Fischle, W., Verdin, E. & Sinclair, J. H. Control of cytomegalovirus lytic gene expression by histone acetylation. *EMBO J.* **21**, 1112–1120 (2002).

13. Reeves, M. B., MacAry, P. A., Lehner, P. J., Sissons, J. G. P. & Sinclair, J. H. Latency, chromatin remodeling, and reactivation of human cytomegalovirus in the dendritic cells of healthy carriers. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 4140–4145 (2005).

14. Meier, J. L. & Stinski, M. F. Regulation of Human Cytomegalovirus Immediate-Early Gene Expression. *Intervirology* vol. 39 331–342 (1996).

15. Collins-McMillen, D. *et al.* Alternative promoters drive human cytomegalovirus reactivation from latency. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 17492–17497 (2019).

16. Nitzsche, A., Paulus, C. & Nevels, M. Temporal Dynamics of Cytomegalovirus Chromatin Assembly in Productively Infected Human Cells. *Journal of Virology* vol. 82 11167–11180 (2008).

17. Reeves, M. B. & Sinclair, J. H. Analysis of latent viral gene expression in natural and experimental latency models of human cytomegalovirus and its correlation with histone modifications at a latent promoter. *Journal of General Virology* vol. 91 599–604 (2010).

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

19. 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* vol. 363 174–188 (2007).

20. Sinclair, J. & Sissons, P. Latency and reactivation of human cytomegalovirus. *J. Gen. Virol.* **87**, 1763–1779 (2006).

21. Slobedman, B. *et al.* Human cytomegalovirus latent infection and associated viral gene expression. *Future Microbiol.* **5**, 883–900 (2010).

22. Sinclair, J. H. & Reeves, M. B. Human cytomegalovirus manipulation of latently infected cells. *Viruses* **5**, 2803–2824 (2013).

23. Shnayder, M. *et al.* Defining the Transcriptional Landscape during Cytomegalovirus Latency with Single-Cell RNA Sequencing. *MBio* **9**, (2018).

24. Cheng, S. *et al.* Transcriptome-wide characterization of human cytomegalovirus in natural infection and experimental latency. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E10586–E10595 (2017).

25. Stenberg, R. M. The Human Cytomegalovirus Major Immediate-Early Gene. *Intervirology* vol. 39 343–349 (1996).

26. Phillips, S. L. & Bresnahan, W. A. The human cytomegalovirus (HCMV) tegument protein UL94 is essential for secondary envelopment of HCMV virions. *J. Virol.* **86**, 2523–2532 (2012).

27. McWatters, B. J. P., Stenberg, R. M. & Kerry, J. A. Characterization of the human cytomegalovirus UL75 (glycoprotein H) late gene promoter. *Virology* **303**, 309–316 (2002).

28. Yurochko, A. D. *et al.* The human cytomegalovirus UL55 (gB) and UL75 (gH) glycoprotein ligands initiate the rapid activation of Sp1 and NF-kappaB during infection. *J. Virol.* **71**, 5051–5059 (1997).

29. AuCoin, D. P., Smith, G. B., Meiering, C. D. & Mocarski, E. S. Betaherpesvirus-conserved cytomegalovirus tegument protein ppUL32 (pp150) controls cytoplasmic events during virion maturation. *J. Virol.* **80**, 8199–8210 (2006).

30. Browne, E. P. & Shenk, T. Human cytomegalovirus UL83-coded pp65 virion protein inhibits antiviral gene expression in infected cells. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 11439–11444 (2003).

31. Varnum, S. M. *et al.* Identification of proteins in human cytomegalovirus (HCMV) particles: the HCMV proteome. *J. Virol.* **78**, 10960–10966 (2004).

32. Gibson, W. Structure and Assembly of the Virion. *Intervirology* vol. 39 389–400 (1996).

33. Terhune, S. S., Schröer, J. & Shenk, T. RNAs are packaged into human cytomegalovirus virions in proportion to their intracellular concentration. *J. Virol.* **78**, 10390–10398 (2004).

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

35. Sinzger, C. *et al.* Cloning and sequencing of a highly productive, endotheliotropic virus strain derived from human cytomegalovirus TB40/E. *J. Gen. Virol.* **89**, 359–368 (2008).

36. 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).

37. Park, J.-J. *et al.* Functional interaction of the human cytomegalovirus IE2 protein with histone deacetylase 2 in infected human fibroblasts. *J. Gen. Virol.* **88**, 3214–3223 (2007).

38. Bigley, T. M., Reitsma, J. M., Mirza, S. P. & Terhune, S. S. Human cytomegalovirus pUL97 regulates the viral major immediate early promoter by phosphorylation-mediated disruption of histone deacetylase 1 binding. *J. Virol.* **87**, 7393–7408 (2013).

39. Radsak, K. *et al.* Induction by sodium butyrate of cytomegalovirus replication in human endothelial cells. *Arch. Virol.* **107**, 151–158 (1989).

40. Tanaka, J., Sadanari, H., Sato, H. & Fukuda, S. Sodium butyrate-inducible replication of human cytomegalovirus in a human epithelial cell line. *Virology* vol. 185 271–280 (1991).

41. Krishna, B. A. *et al.* Transient activation of human cytomegalovirus lytic gene expression during latency allows cytotoxic T cell killing of latently infected cells. *Sci. Rep.* **6**, 24674 (2016).

42. He, M. & Gao, S.-J. A novel role of SIRT1 in gammaherpesvirus latency and replication. *Cell Cycle* **13**, 3328–3330 (2014).

43. Pei, Y. & Robertson, E. S. The Crosstalk of Epigenetics and Metabolism in Herpesvirus Infection. *Viruses* **12**, (2020).

44. Hu, H.-L. *et al.* TOP2β-Dependent Nuclear DNA Damage Shapes Extracellular Growth Factor Responses via Dynamic AKT Phosphorylation to Control Virus Latency. *Mol. Cell* **74**, 466–480.e4 (2019).

45. Wu, K., Oberstein, A., Wang, W. & Shenk, T. Role of PDGF receptor-α during human cytomegalovirus entry into fibroblasts. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E9889–E9898 (2018).

46. Braspenning, S. E. *et al.* Decoding the Architecture of the Varicella-Zoster Virus Transcriptome. *MBio* **11**, (2020).

47. Marcinowski, L. *et al.* Real-time transcriptional profiling of cellular and viral gene expression during lytic cytomegalovirus infection. *PLoS Pathog.* **8**, e1002908 (2012).

48. Erhard, F. *et al.* scSLAM-seq reveals core features of transcription dynamics in single cells. *Nature* **571**, 419–423 (2019).

49. Azarkh, Y., Dölken, L., Nagel, M., Gilden, D. & Cohrs, R. J. Synthesis and decay of varicella zoster virus transcripts. *J. Neurovirol.* **17**, 281–287 (2011).

50. Groves, I. J. *et al.* Bromodomain proteins regulate human cytomegalovirus latency and reactivation allowing epigenetic therapeutic intervention. *Proc. Natl. Acad. Sci. U. S. A.* **118**, (2021).

51. Ren, K. *et al.* An Epigenetic Compound Library Screen Identifies BET Inhibitors That Promote HSV-1 and -2 Replication by Bridging P-TEFb to Viral Gene Promoters through BRD4. *PLoS Pathog.* **12**, e1005950 (2016).

52. Hopcraft, S. E. *et al.* Chromatin remodeling controls Kaposi's sarcoma-associated herpesvirus reactivation from latency. *PLOS Pathogens* vol. 14 e1007267 (2018).

53. Koyuncu, E. *et al.* Sirtuins are evolutionarily conserved viral restriction factors. *MBio* **5**, (2014).

54. Shnayder, M. *et al.* Single cell analysis reveals human cytomegalovirus drives latently infected cells towards an anergic-like monocyte state. *Elife* **9**, (2020).

55. Halford, W. P., Kemp, C. D., Isler, J. A., Davido, D. J. & Schaffer, P. A. ICP0, ICP4, or VP16 Expressed from Adenovirus Vectors Induces Reactivation of Latent Herpes Simplex Virus Type 1 in Primary Cultures of Latently Infected Trigeminal Ganglion Cells. *Journal of Virology* vol. 75 6143–6153 (2001).

56. Lukac, D. M., Renne, R., Kirshner, J. R. & Ganem, D. Reactivation of Kaposi's sarcoma-associated herpesvirus infection from latency by expression of the ORF 50 transactivator, a homolog of the EBV R protein. *Virology* **252**, 304–312 (1998).

57. Marchini, A., Liu, H. & Zhu, H. Human cytomegalovirus with IE-2 (UL122) deleted fails to express early lytic genes. *J. Virol.* **75**, 1870–1878 (2001).

58. Sanders, R. L., Clark, C. L., Morello, C. S. & Spector, D. H. Development of cell lines that provide tightly controlled temporal translation of the human cytomegalovirus IE2 proteins for complementation and functional analyses of growth-impaired and nonviable IE2 mutant viruses. *J. Virol.* **82**, 7059–7077 (2008).

59. Greaves, R. F. & Mocarski, E. S. Defective Growth Correlates with Reduced Accumulation of a Viral DNA Replication Protein after Low-Multiplicity Infection by a Human Cytomegalovirus ie1 Mutant. *Journal of Virology* vol. 72 366–379 (1998).

60. Cobbs, C. S., Matlaf, L. & Harkins, L. E. Methods for the detection of cytomegalovirus in glioblastoma cells and tissues. *Methods Mol. Biol.* **1119**, 165–196 (2014).

61. Borst, E.-M. & Messerle, M. Analysis of human cytomegalovirus oriLyt sequence requirements in the context of the viral genome. *J. Virol.* **79**, 3615–3626 (2005).

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

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

64. 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).

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

66. Liberzon, A. *et al.* Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739–1740 (2011).

67. 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**, (2017).

68. Roback, J. D. *et al.* Multicenter evaluation of PCR methods for detecting CMV DNA in blood donors. *Transfusion* vol. 41 1249–1257 (2001).

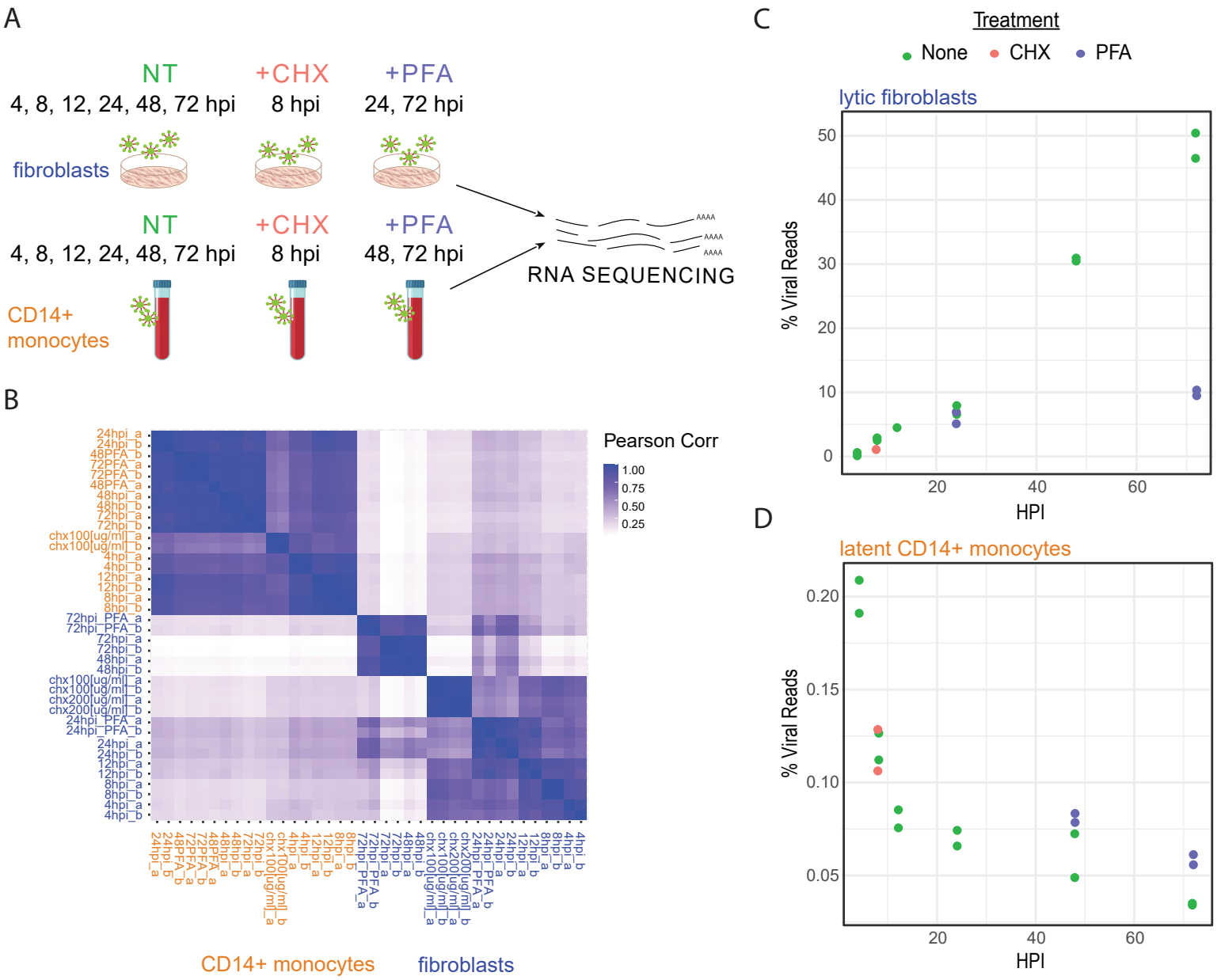

Figure 1. RNA sequencing along HCMV infection in fibroblasts and CD14+ monocytes
(A) Schematic representation of the experimental setup. Fibroblasts and CD14+ monocytes infected with HCMV strain TB40E-GFP were harvested at 4, 8, 12, 24, 48 and 72 hpi for RNA sequencing. In addition, infected fibroblasts and CD14+ monocytes were either treated with Phosphonoformic acid (PFA) and harvested at 24 and 72 or 48 and 72 hpi, respectively or treated with cycloheximide (CHX) and harvested at 8 hpi. All RNA-seq libraries were done in two biological replicates
(B) Heatmap depicting the Pearson correlations between all RNA-seq samples based on viral and host reads. Fibroblast samples are marked in blue and CD14+ monocyte samples are marked in orange. All biological replicates are denoted by a/b.
(C, D) Percentage of HCMV reads out of total mRNA reads along infection in fibroblasts (C) and CD14+ monocytes (D). The plots include the untreated (green), CHX (red), and PFA (purple) samples with both replicates of each sample.

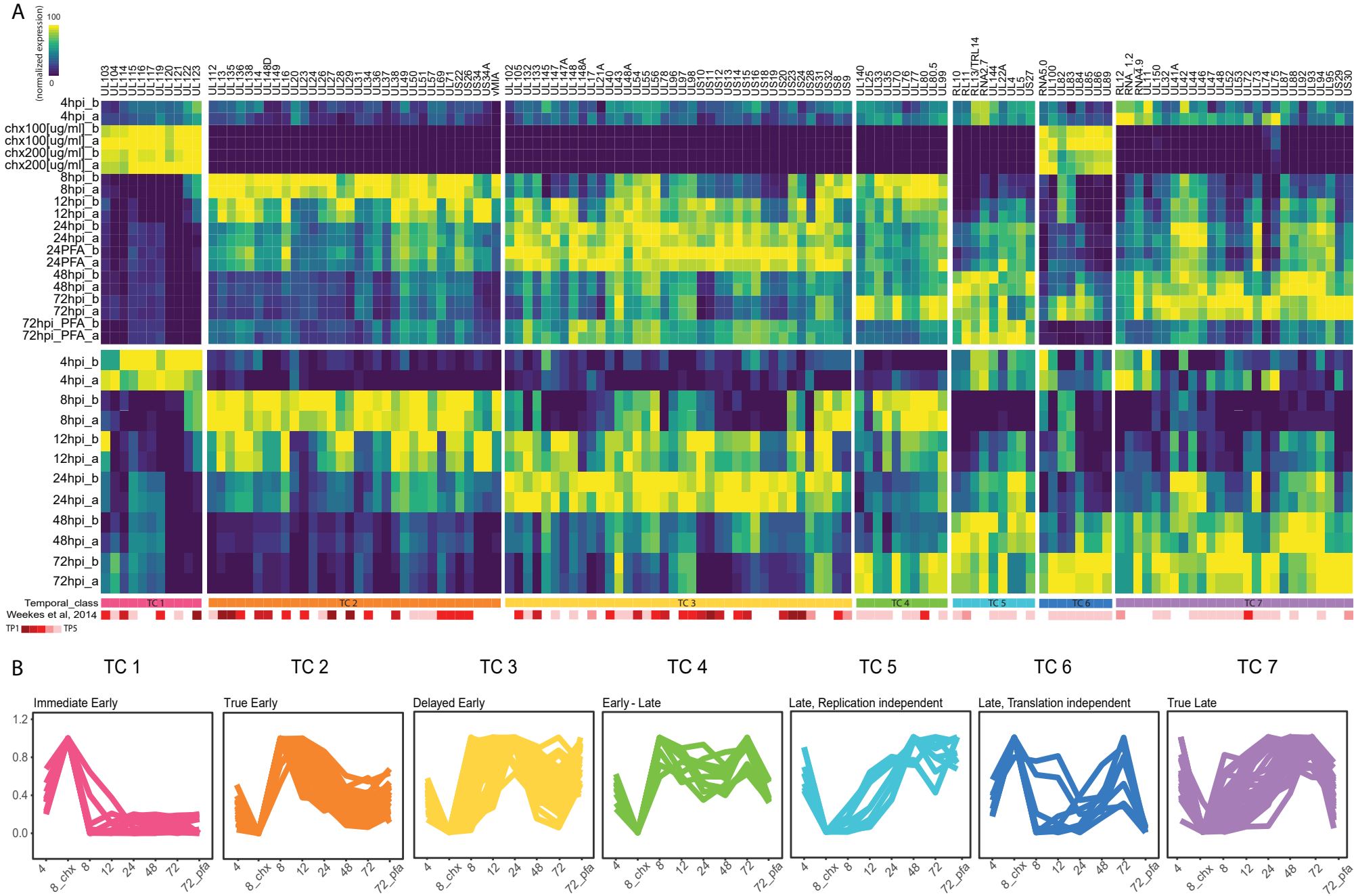

Figure 2. Temporal Classes of HCMV genes along infection of fibroblasts

(A) Heatmap depicting normalized relative expression levels of viral transcripts. Viral genes were clustered into seven temporal classes based on their expression pattern along time, as well as following CHX and PFA treatments. Scaled normalized expression patterns of viral genes are shown including CHX and PFA samples (upper panel) and without the CHX and PFA samples (lower panel). The bars below the panels show the annotations according to our defined temporal classes (top bar) and according to the protein temporal classes that were defined by Weekes et al.[10] (B) Expression profiles of all viral genes for each temporal class (TC) along infection of fibroblasts, averaged across replicates with description for each TC. Values are normalized to the maximal expression of each gene.

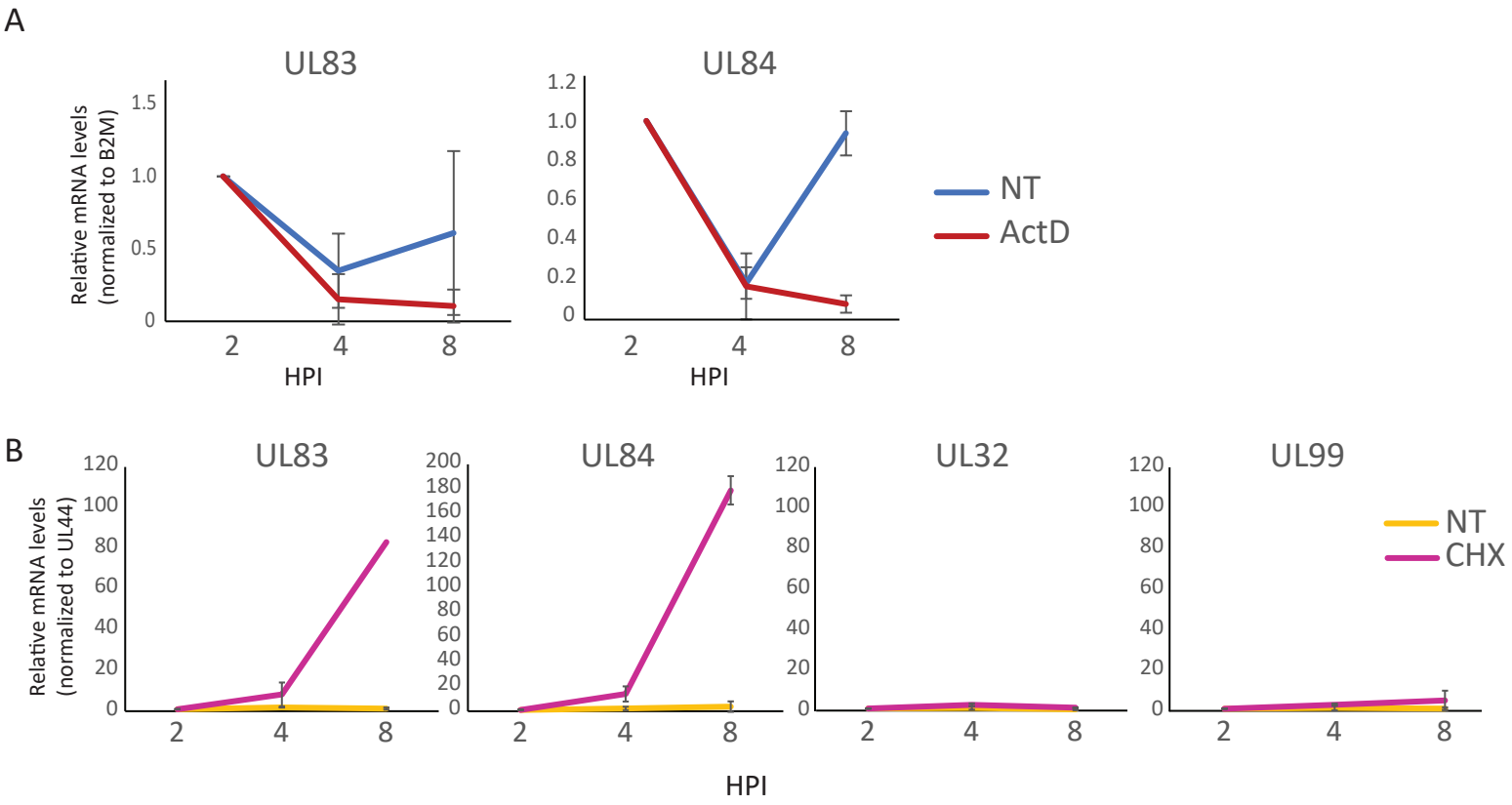

Figure 3. Translation-independent transcription of a group of late transcripts
(A) Fibroblasts were infected with HCMV and were then treated with actinomycinD (actD) or left untreated (NT) at 2 hpi. mRNA levels of the viral transcripts UL83 and UL84 (TC6) were quantified by qRT-PCR at 2, 4, and 8 hpi corresponding to 0, and 2, and 6 hours post actD treatment. RNA levels are normalized to the host B2M transcript. Mean and standard deviation are shown. (B) Fibroblasts were infected with HCMV and cells were treated with CHX or left untreated (NT) at 2 hpi. mRNA levels of viral transcripts UL83, UL84, UL32, and UL99 were quantified by qRT-PCR at 2, 4, and 8 hpi. RNA levels are normalized to viral transcript UL44. Mean and standard deviation are shown.

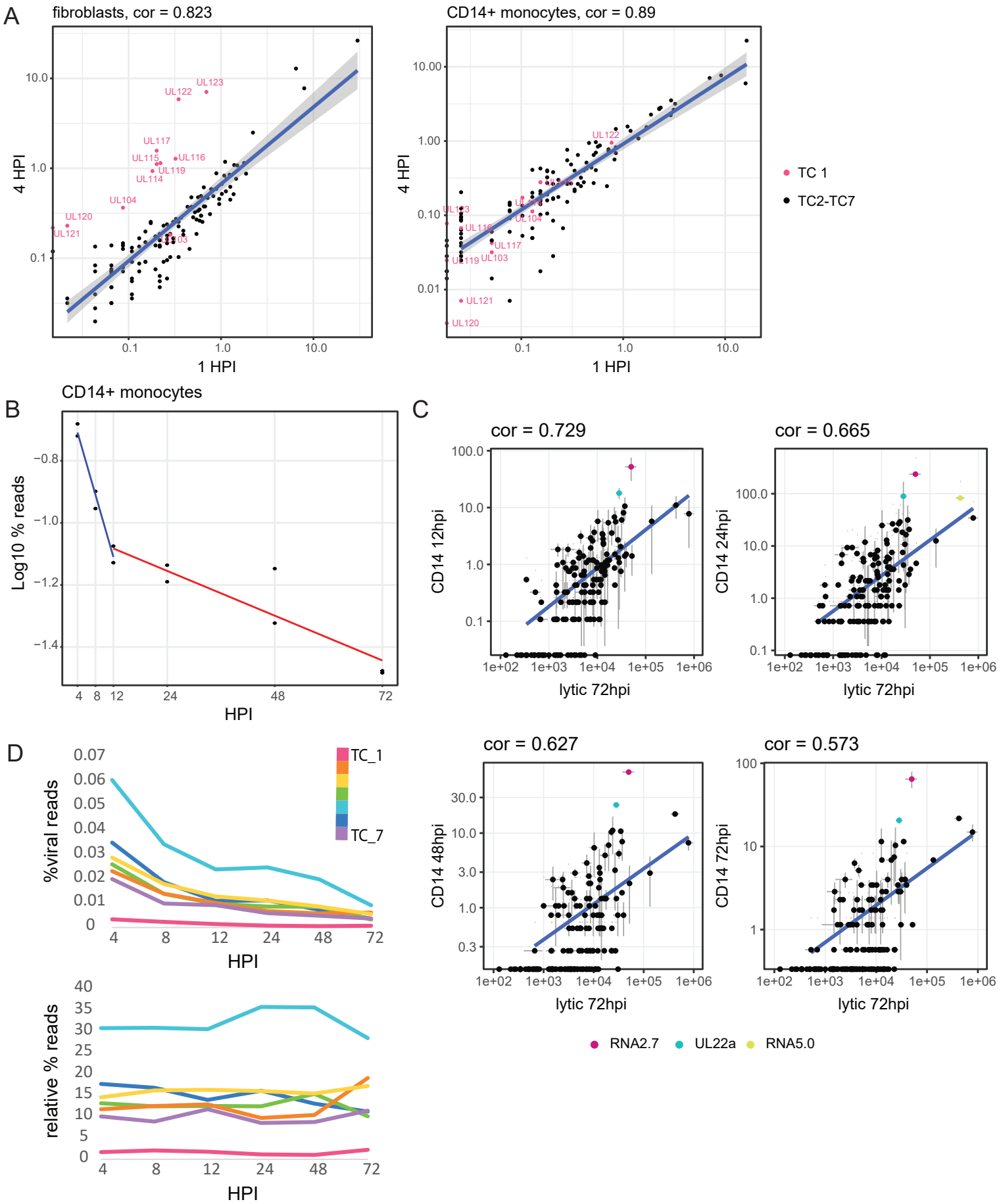

Figure 4. Low level transcription in HCMV infected CD14+ monocytes

(A) Scatter plot showing the read number for each viral gene at 4 hpi versus 1 hpi in fibroblasts (left panel) and CD14+ monocytes (right panel). Viral immediate early genes (TC 1) are marked in pink. Spearman correlations are indicated. (B) Percentage of viral reads out of total mRNA reads along HCMV infection of CD14+ monocytes. The calculated overall decay (half-life) of viral transcripts at 1−12 hpi (0.68 hours) is denoted by a blue regression line and at 12−72 hpi (8.93 hours) is denoted by a red regression line. (C) Scatter plots showing the normalized read number for viral genes at 72 hpi in fibroblasts versus 12, 24, 48, and 72 hpi in CD14+ monocytes. Spearman correlations are indicated. Labeled in color are viral transcripts whose expression significantly deviated (p.value<0.05) from the correlation (D) Expression Profile of all temporal classes (TCs) along HCMV infection of CD14+ monocytes, as calculated by percentage of reads from all viral genes in a TC out of all reads (top) or out of all viral reads (bottom). Mean values of replicates are presented.

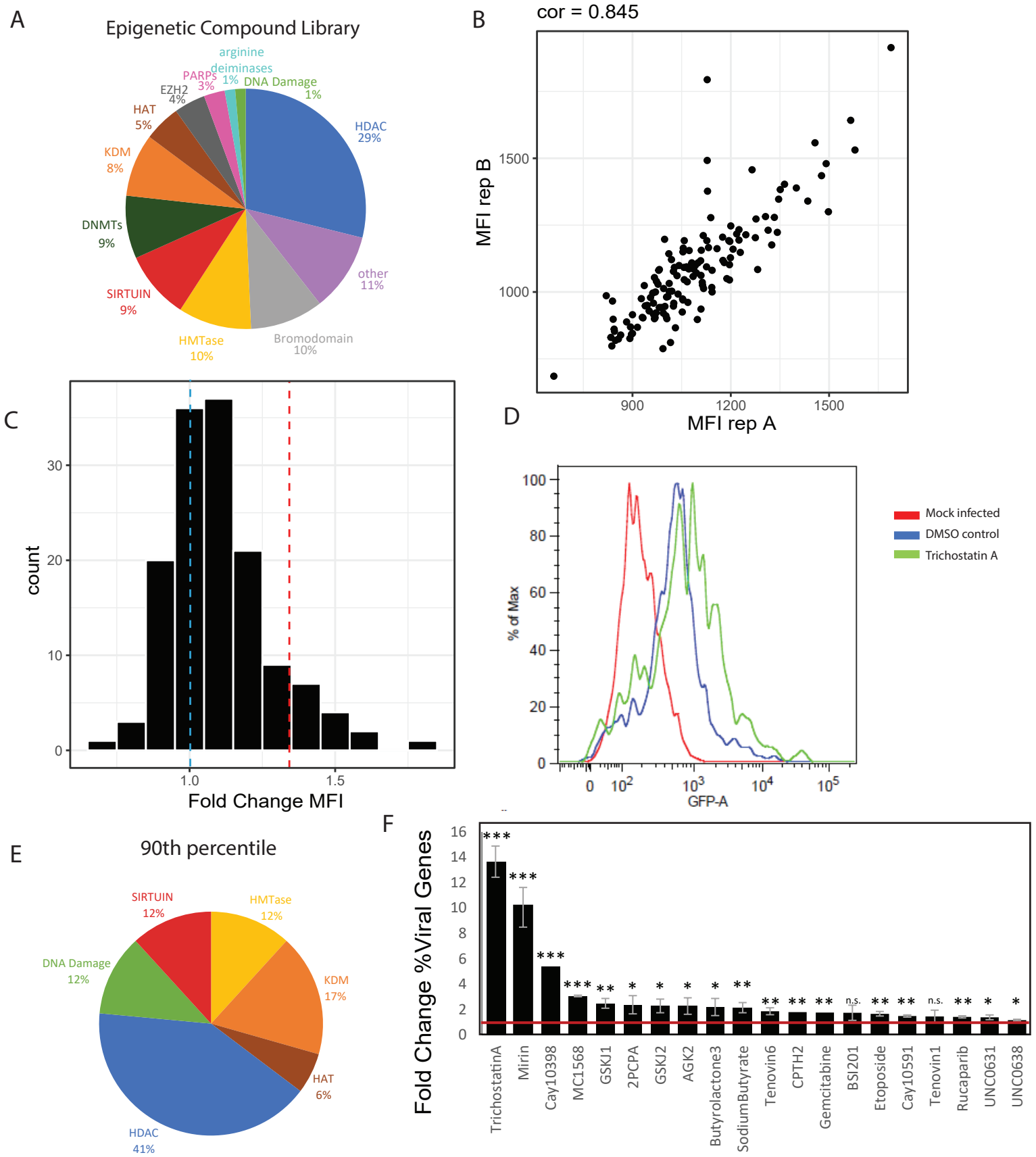

Figure 5. Epigenetic inhibitors induce viral gene transcription in HCMV infected CD14+ monocytes.
(A-F) CD14+ monocytes infected with HCMV expressing a GFP reporter were incubated with 1 μM of each of 150 compounds from an epigenetic inhibitor library, for 24 hours at 48 hpi, in biological replicates. (A) A pie chart showing the distribution of drug targets in the epigenetic inhibitor library. (B) Scatter plot showing GFP Mean fluorescence intensity (MFI) values measured by flow cytometry for each treatment of biological replicates. Spearman correlation is indicated. (C) Histogram depicting the distribution of MFI fold change, compared to a DMSO control, for all tested inhibitors. The blue dashed line shows the DMSO control sample (FC=1) and the red dashed line pinpoints the 90th percentile. (D) Flow cytometry analysis of GFP levels for a representative inhibitor, TrichostatinA, whose MFI fold change lies within the 90th percentile, compared to the DMSO sample and uninfected cells. (E) A pie chart showing the distribution of drug targets of the inhibitors in the 90th percentile of MFI fold change. (F) RNA-sequencing was performed in duplicates on HCMV infected CD14+ monocytes treated with 20 different inhibitors and DMSO control. Plot showing fold change of percent viral reads for each treatment compared to DMSO control. Mean and standard deviation of duplicates are presented. Red line marks fold change of 1. p-values representing significant increase in viral gene expression were calculated with t-test ***pval≤0.001, **pval≤0.01, *pval≤0.05, n.s. pval>0.05.

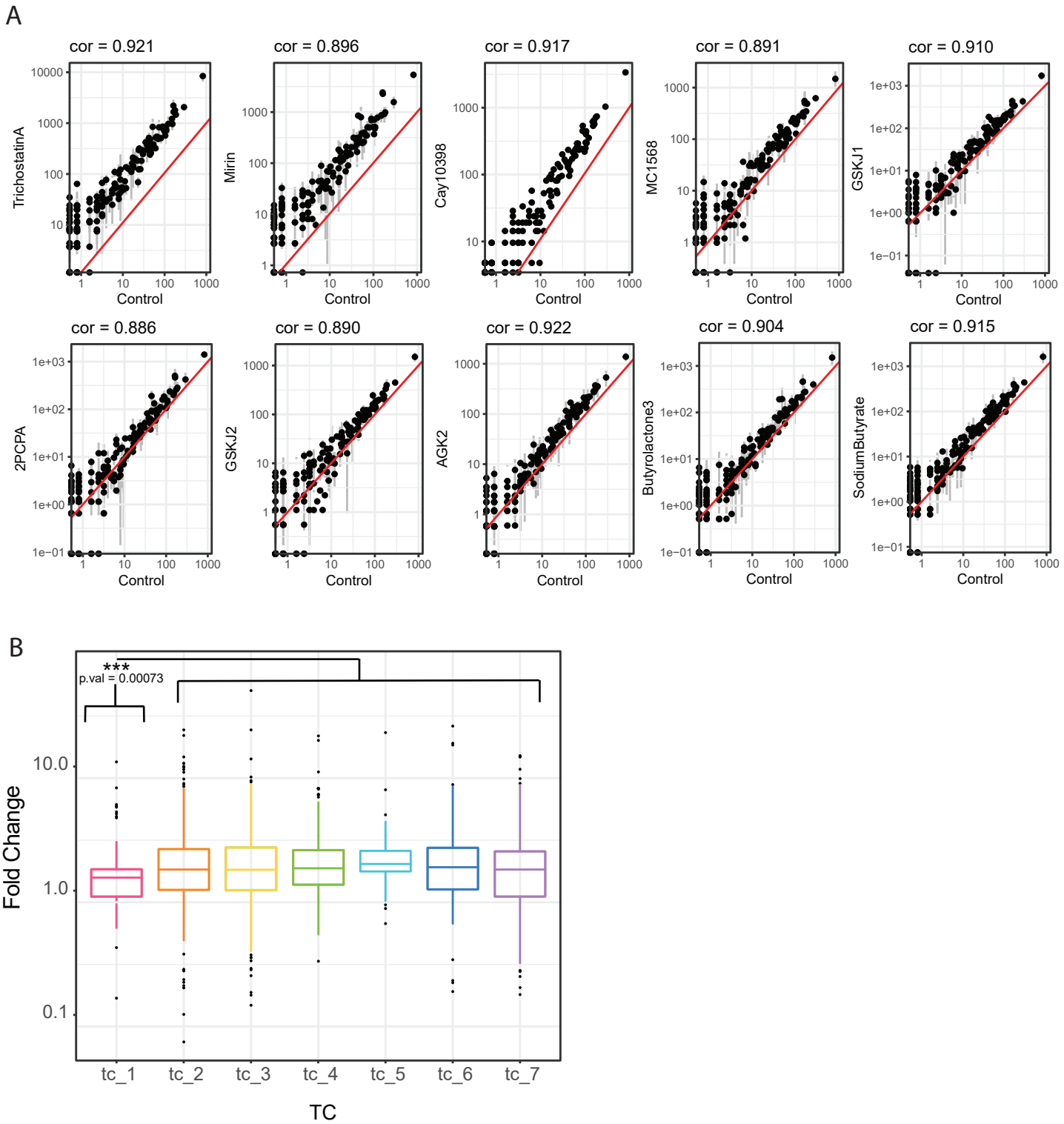

Figure 6. Unique repression of IE genes in HCMV infected CD14+ monocytes.
(A) Scatter plots showing the read number for viral genes in HCMV infected CD14+ monocytes treated with DMSO as control versus treatment with ten different inhibitors that elicited the strongest increase in viral gene expression. Spearman correlations are indicated. The red line marks 1:1 ratio. Only in the mirin sample, 2 viral genes (RNA5.0 and RL1) significantly deviated (p.value<=0.05 and FC=>2) from the correlation with the DMSO control (B) Boxplot showing the transcript level fold change of vira class between inhibitors treated and control treated HCMV infected CD14+ monocytes. p-value calculated by t-test comparing TC1 and all other TCs is indicated.

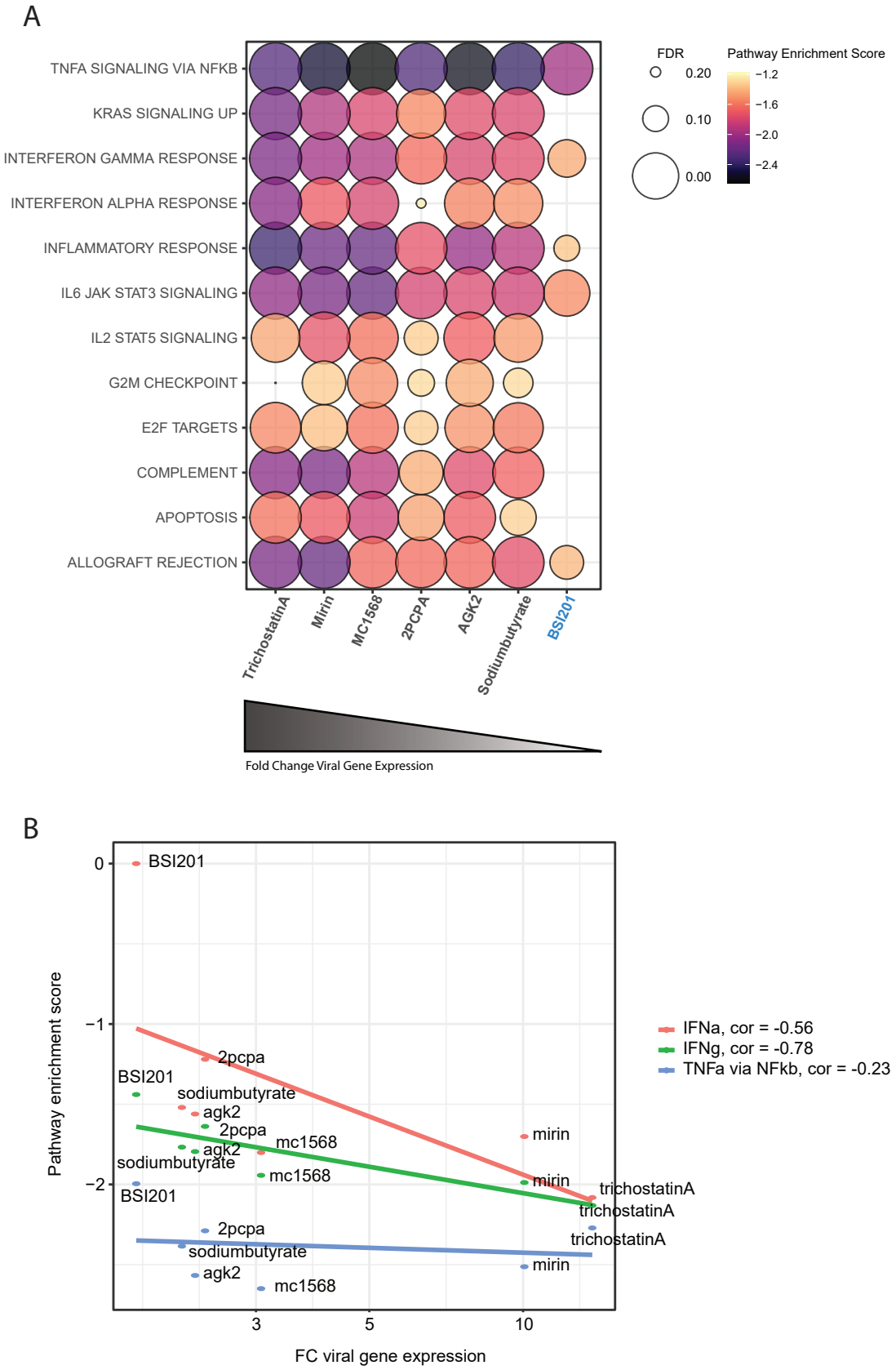

Figure 7. Induction of viral gene expression during HCMV latency leads to reduction in expression of immune related genes (A). Bubble plot showing enriched human hallmark pathways which are downregulated following inhibitor treatment of infected CD14+ monocytes for drugs which significantly induced more than 1.8- Fold Change in viral gene expression. BSI201, highlighted in blue, did not produce a significant increase in viral gene expression. (B). Scatter plot depicting enrichment scores for three host pathways: interferon alpha response (IFNa, red), interferon gamma response (IFNg, green) and TNF alpha signaling via NFKb (blue) versus viral transcript level fold change for different inhibitor treatments compared to the control. Spearman correlations are indicated.

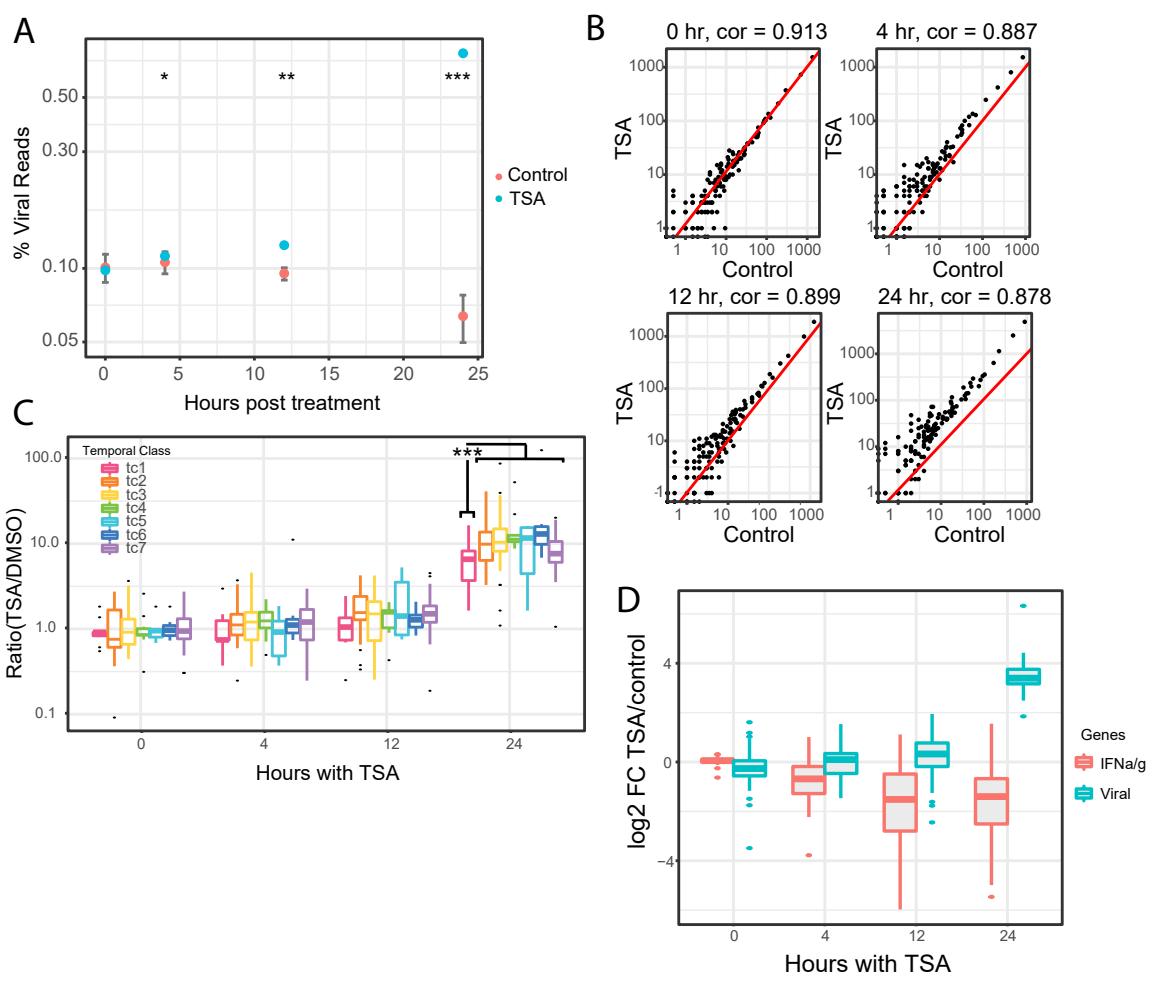

Figure 8. Unique repression of IE genes occurs at different times following TSA treatment (A-C) RNA-sequencing was performed on HCMV infected CD14+ monocytes treated with TSA and DMSO control, in biological replicates, at 0, 4, 12 and 24 hours after treatment. (A) Percentage of HCMV reads out of total mRNA reads at different time points following treatment. Error bars represent standard deviation in the control samples. p-values were calculated with t test. *pval≤0.05, **pval≤0.01, ***pval≤0.001. (B) Scatter plots showing the read number of viral genes in HCMV infected CD14+ monocytes treated with DMSO control versus TSA treatment at 0, 4, 12 and 24 hours after treatment. Spearman correlations are indicated. The red line marks 1:1 ratio. (C) Boxplot showing the transcript level fold change of viral genes from each temporal class (TC) between TSA and DMSO control treated HCMV infected CD14+ monocytes at different time points following treatment. p-value = 0.0011 calculated by t-test comparing TC1 and all other TCs at 24 hours following treatment is indicated. (D) Box plot showing the transcript level log2 Fold change of viral genes and IFNa/g induced host genes in TSA treated samples at 0, 4, 12, and 24 hours after treatment.

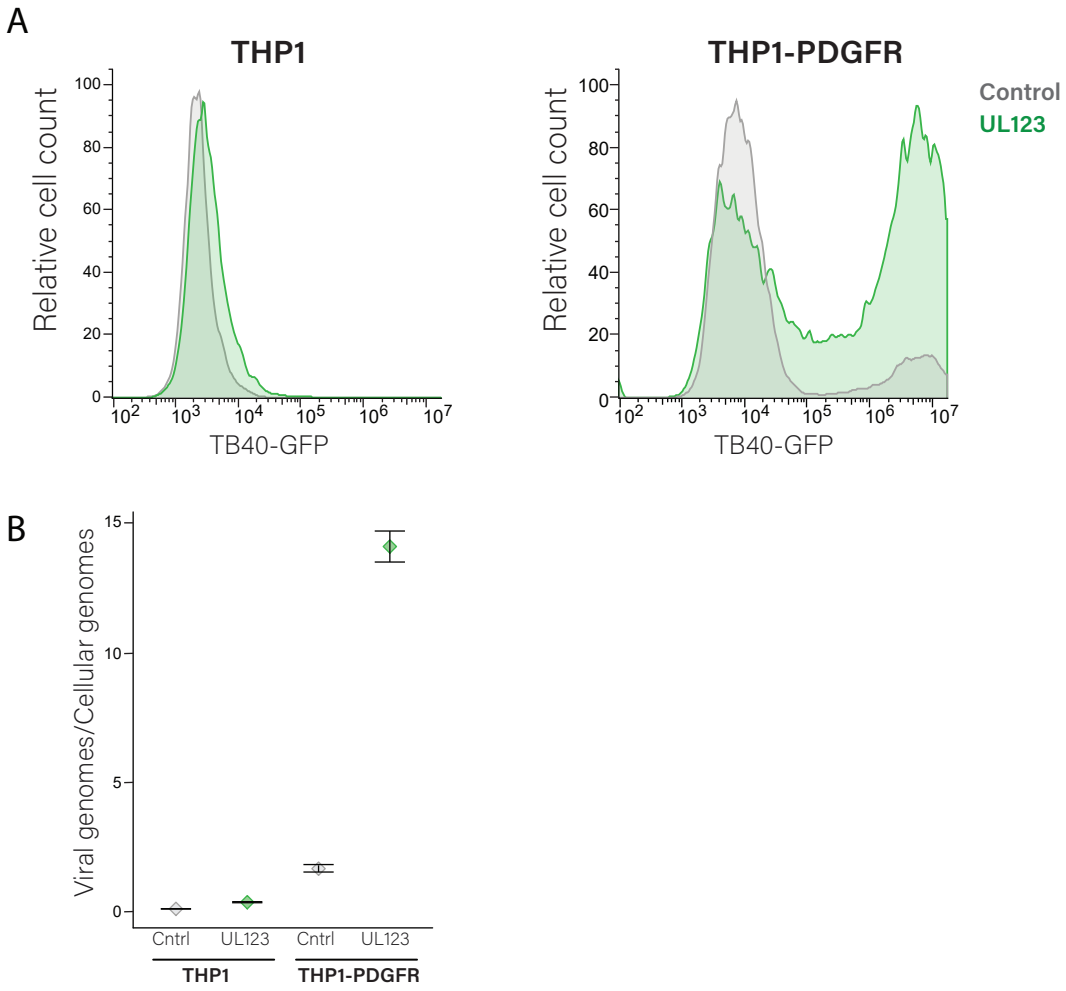

Figure 9. Overexpression of UL123 is sufficient to drive productive infection.
(A-B) THP1 cells or THP1 cells expressing PDGFRa were transduced with UL123 expression vector or mCherry as control, infected with HCMV strain TB40E-GFP. (A) GFP levels measured by flow cytometry at 72hpi. Representative experiment out of two biological replicates is presented. (B) Digital droplet PCR measurements of viral genome abundance at 72hpi. Mean and 95% CV of Poisson distribution calculated from two replicates are shown.

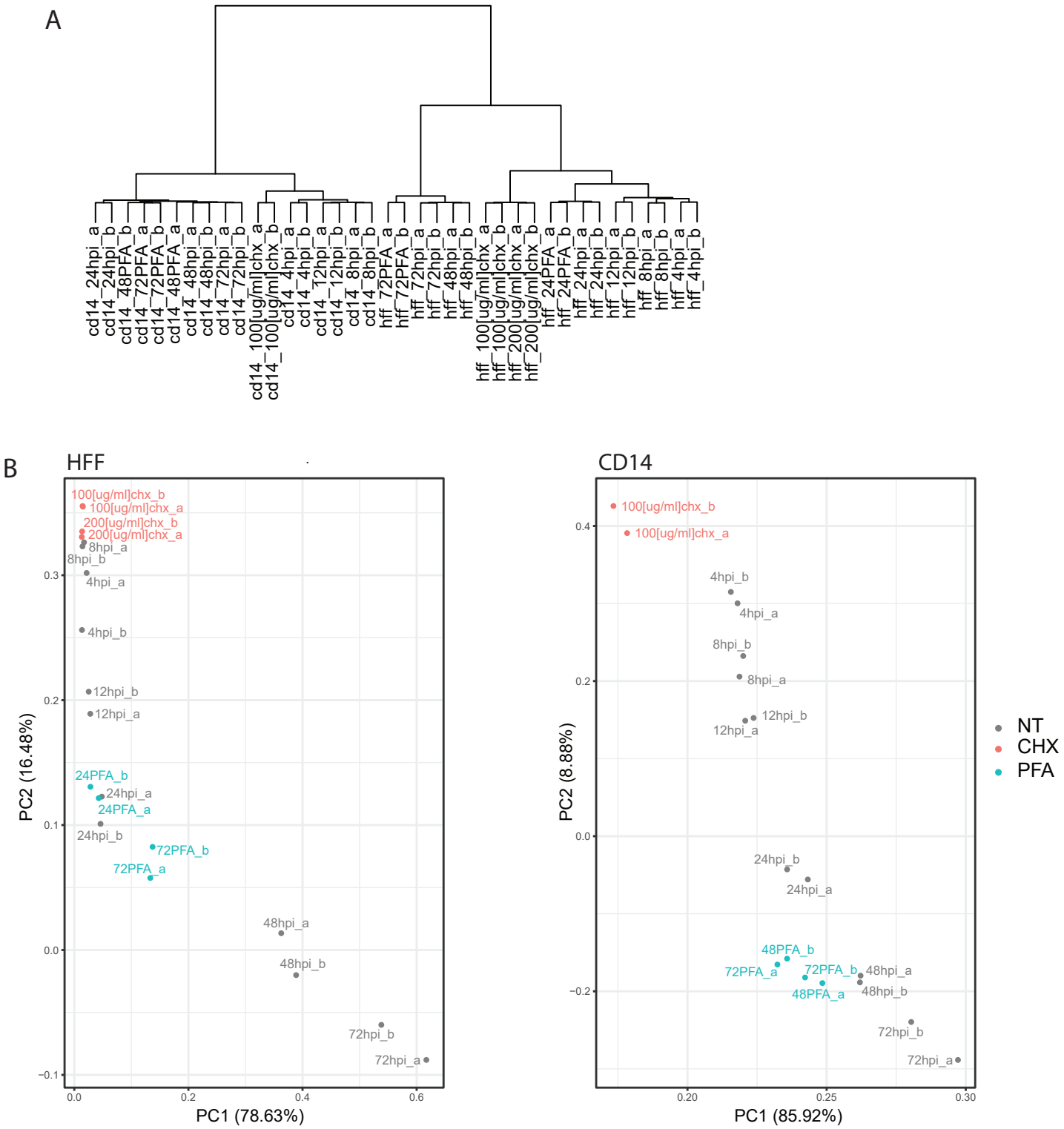

Fig S1. RNA sequencing along HCMV infection in fibroblasts and CD14+ monocytes
(A) Hierarchical clustering dendrogram of all samples. Cell lines cluster separately, and biological replicates cluster together.
(B) PCA analysis for each data set (HFF and CD14+) based on host and viral transcript reads. CHX treated samples are highlighted in red and PFA treated samples in blue.

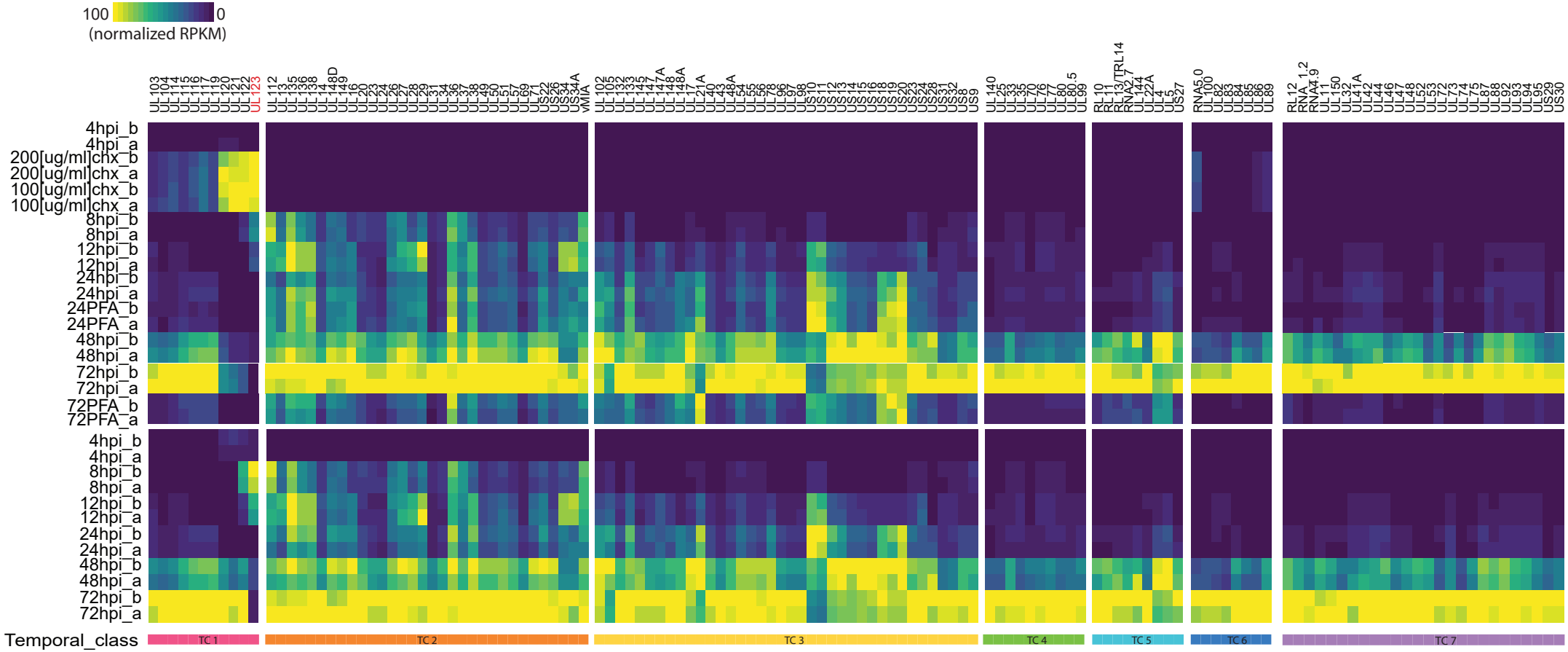

Fig S2. RPKM values of HCMV genes along infection of fibroblasts

Heatmap depicting absolute levels of viral RNA transcripts in reads per kb per million (RPKM) in HCMV infected fibroblasts. Expression patterns of viral genes are shown including CHX and PFA samples in the top panel of the heatmap and without CHX and PFA samples in the lower panel of the heatmap.

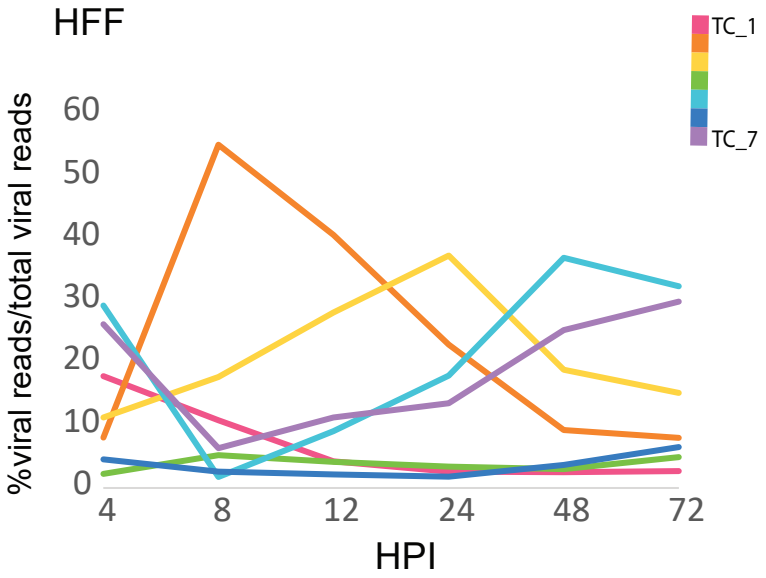

Fig S3. Temporal class dynamics along infection in fibroblasts.
Expression Profile of all temporal classes (TCs) along HCMV infection of HFFs, as calculated by percentage of reads from all viral genes in a TC out of all viral reads. Mean values from biological replicates are shown.

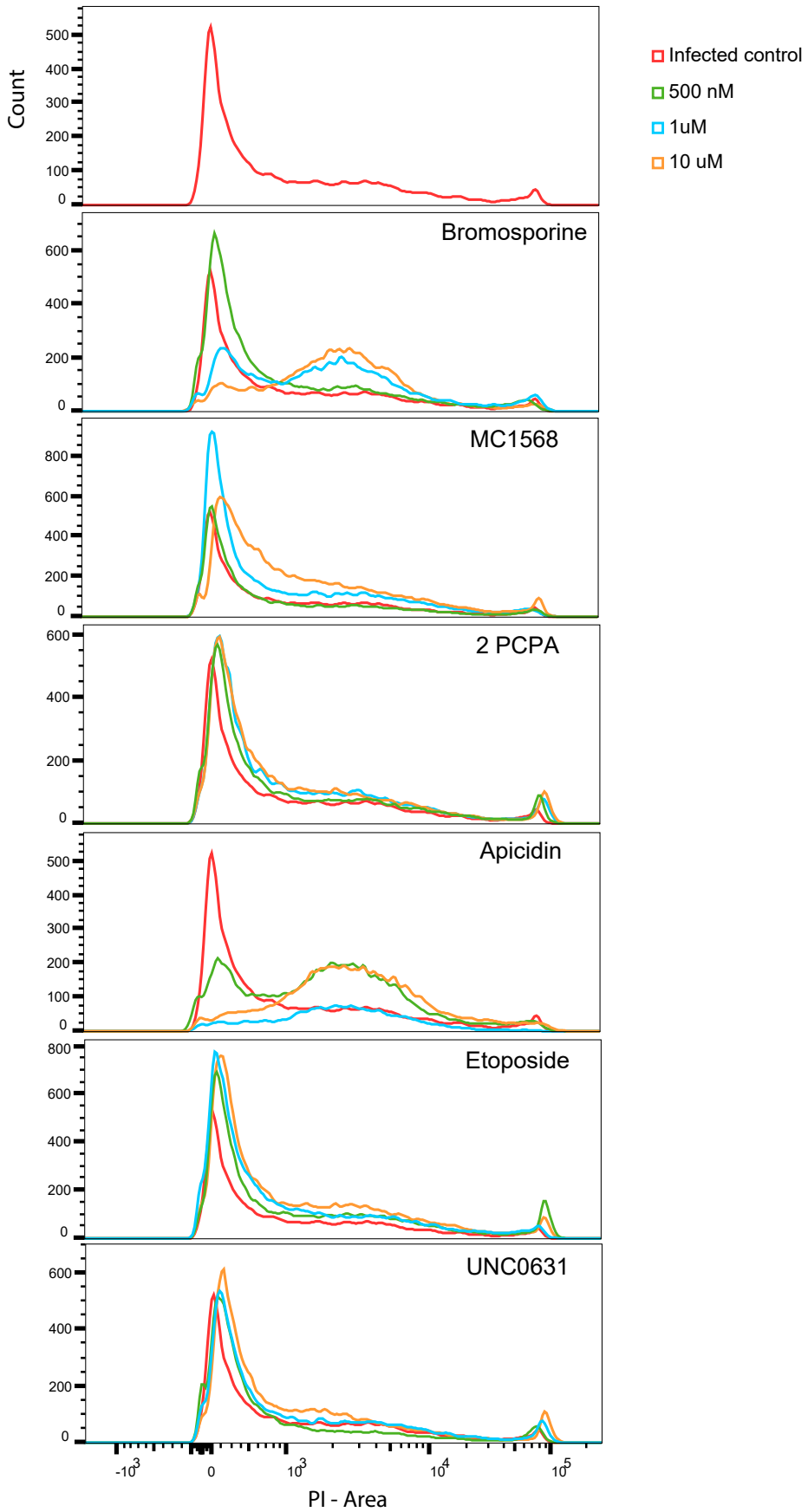

Fig S4. Dose-dependent cytotoxicity of different epigenetic inhibitors
At 24 hours post infection, HCMV infected CD14+ monocytes were incubated with six select drugs from the epigenetic inhibitor library at increasing concentrations (500nM,1uM,10uM). Cell death was measured in an infected control sample and in infected, inhibitor treated samples by propidium iodide staining and flow cytometry.

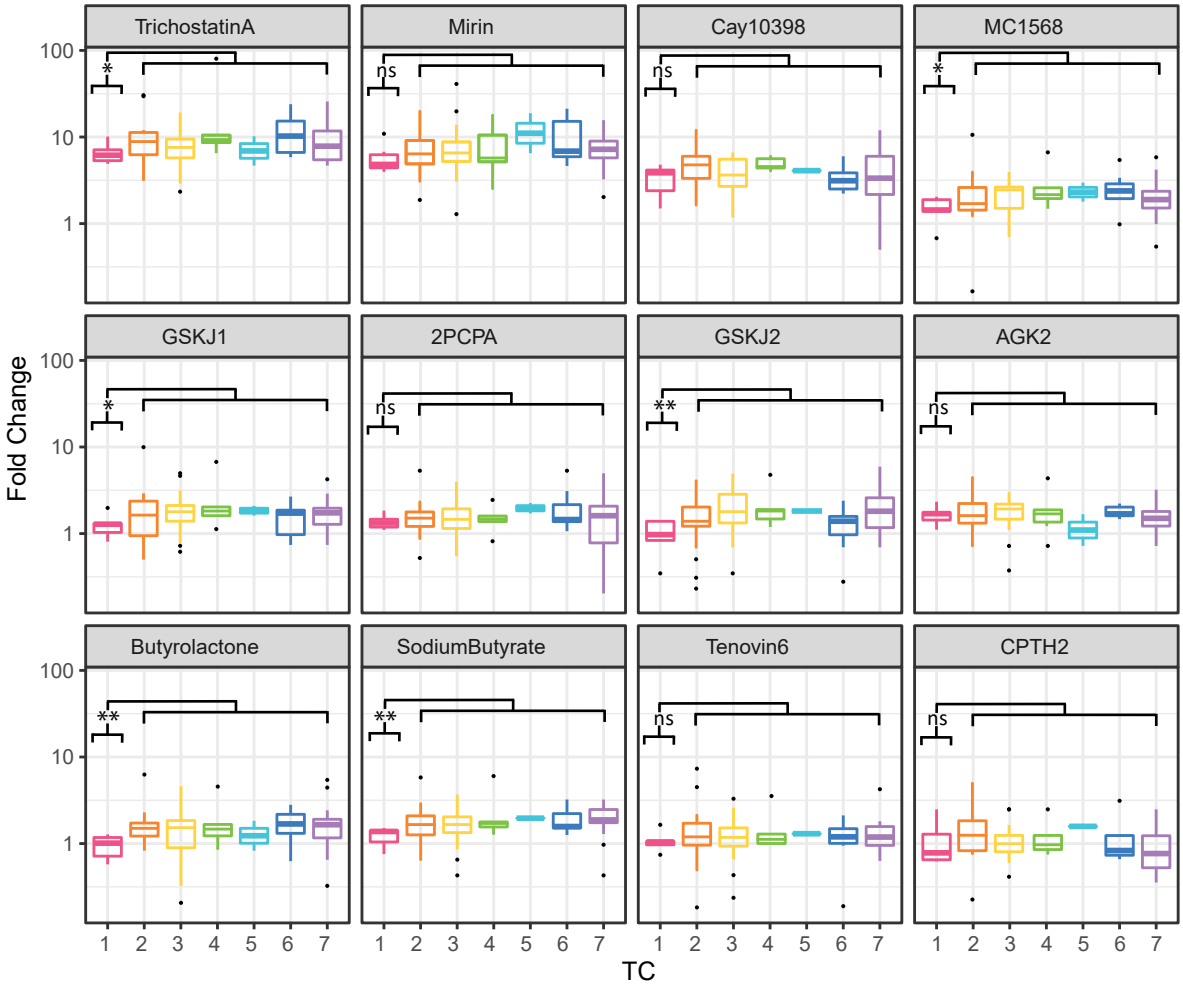

Fig S5. Epigenetic drug treatments reveal unique repression of TC1, immediate early genes
Boxplot showing the transcript level fold change of viral genes from each temporal class between inhibitor treated and control treated HCMV infected CD14+ monocytes. p-value calculated by t-test comparing TC1 and all other TCs is indicated.

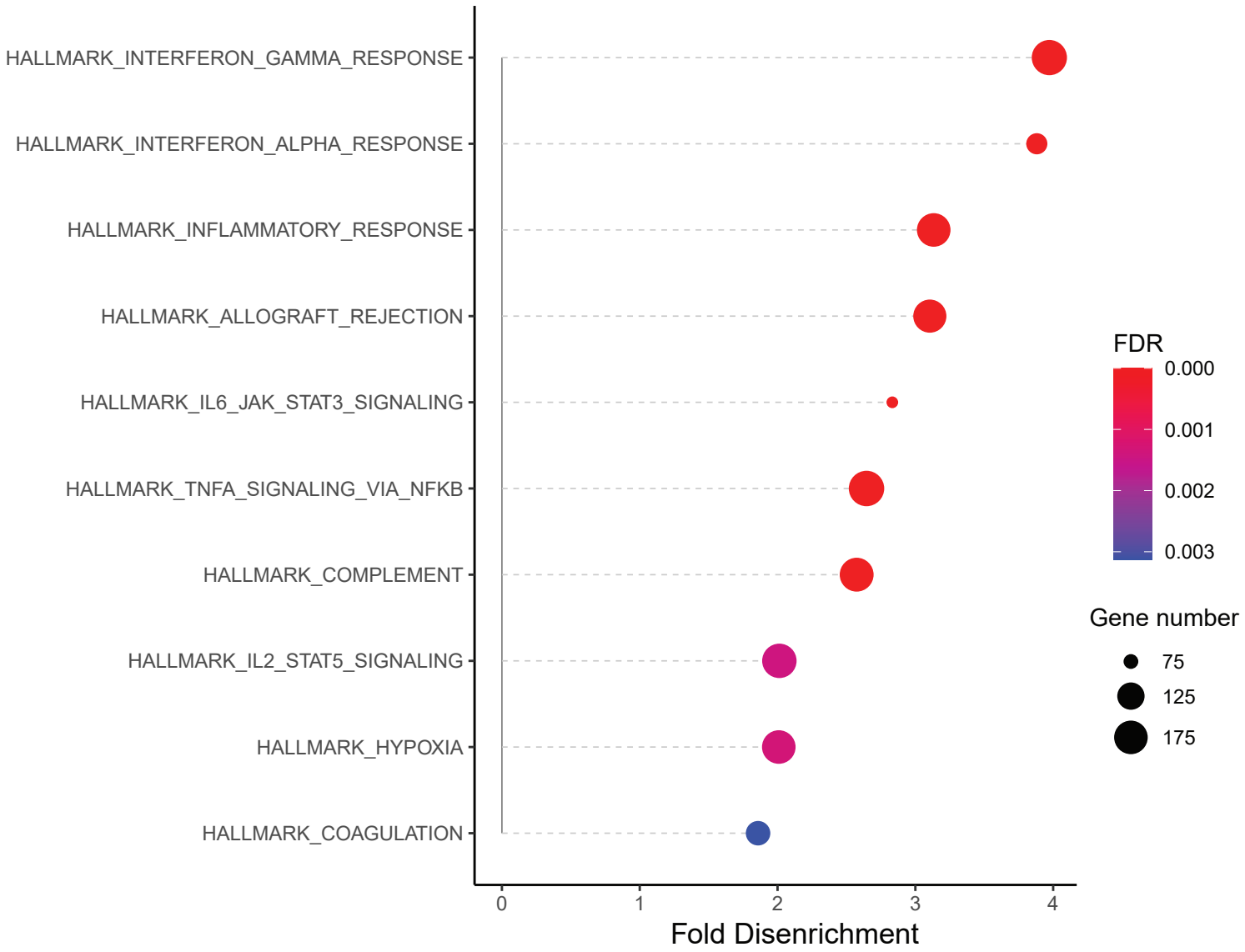

Fig S6. Induction of viral gene expression by TSA time course leads to reduction in expression of immune related genes
Bubble plot showing enriched human hallmark signature pathways which are downregulated following TSA treatment of infected CD14+ monocytes.

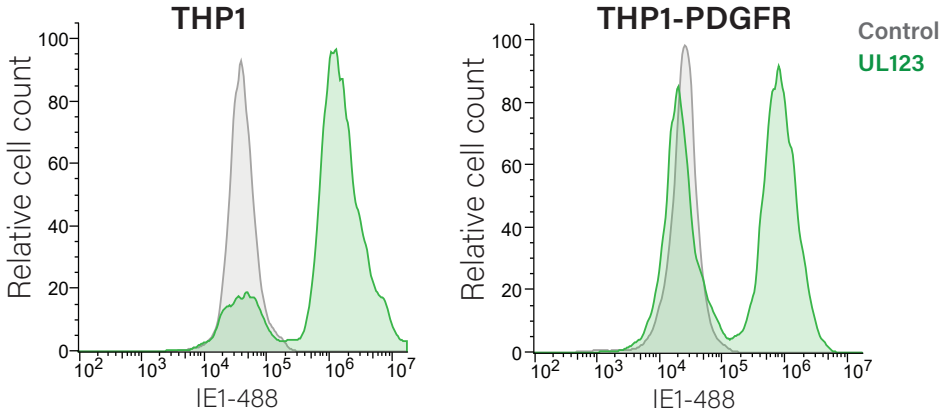

Fig S7. Over Expression of IE1

THP1 cells or THP1 cells expressing PDGFRa were transduced with UL123 expression vector or mCherry as control, infected with HCMV strain TB40E-GFP. IE1 levels were measured by flow cytometry using intracellular staining with a florescent antibody. Representative of two biological replicates are shown.