



Evidence for Multiple Subpopulations of Herpesvirus-Latently Infected Cells

Justin T. Landis,^a Ryan Tuck,^a Yue Pan,^b Carson N. Mosso,^c Anthony B. Eason,^a Razia Moorad,^a J. Stephen Marron,^{b,c}  Dirk P. Dittmer^a

^aLineberger Comprehensive Cancer Center, Department of Microbiology and Immunology, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

^bDepartment of Biostatistics, Gillings School of Global Public Health, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

^cDepartment of Statistics and Operations Research, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

ABSTRACT Kaposi's sarcoma-associated herpesvirus (KSHV)-associated primary effusion lymphomas (PEL) are traditionally viewed as homogenous regarding viral transcription and lineage of origin, but so far this contention has not been explored at the single-cell level. Single-cell RNA sequencing of latently infected PEL supports the existence of multiple subpopulations even within a single cell line. At most 1% of the cells showed evidence of near-complete lytic transcription. The majority of cells only expressed the canonical viral latent transcripts: those originating from the latency locus, the viral interferon regulatory factor locus, and the viral lncRNA nut-1/Pan/T1.1; however, a significant fraction of cells showed various degrees of more permissive transcription, and some showed no evidence of KSHV transcripts whatsoever. Levels of viral interleukin-6 (IL-6)/K2 mRNA emerged as the most distinguishing feature to subset KSHV-infected PEL. One newly uncovered phenotype is the existence of BCBL-1 cells that readily adhered to fibronectin and that displayed mesenchymal lineage-like characteristics.

IMPORTANCE Latency is the defining characteristic of the *Herpesviridae* and central to the tumorigenesis phenotype of Kaposi's sarcoma-associated herpesvirus (KSHV). KSHV-driven primary effusion lymphomas (PEL) rapidly develop resistance to therapy, suggesting tumor instability and plasticity. At any given time, a fraction of PEL cells spontaneously reactivate KSHV, suggesting transcriptional heterogeneity even within a clonal cell line under optimal growth conditions. This study employed single-cell mRNA sequencing to explore the within-population variability of KSHV transcription and how it relates to host cell transcription. Individual clonal PEL cells exhibited differing patterns of viral transcription. Most cells showed the canonical pattern of KSHV latency (LANA, vCyc, vFLIP, Kaposin, and vIRFs), but a significant fraction evidenced extended viral gene transcription, including of the viral IL-6 homolog, open reading frame K2. This study suggests new targets of intervention for PEL. It establishes a conceptual framework to design KSHV cure studies analogous to those for HIV.

KEYWORDS Kaposi's sarcoma, herpesvirus, KSHV, single cell, single-cell RNA-seq, lymphoma, primary effusion lymphoma, Castleman's disease, scRNAseq, Kaposi's sarcoma-associated herpesvirus

Latency is the default outcome after infection with Kaposi's sarcoma-associated herpesvirus (KSHV) or human herpesvirus 8. Epigenetic profiling showed that the viral genome is heavily and rapidly silenced upon uncoating of the capsid in the nucleus of any infected cell (1–6). Very few viral genes are transcribed in infected cells during the latent phase. Chief among the constitutively expressed genes is the latency-associated nuclear antigen (LANA), which is encoded by open reading frame (ORF) 73 (7–10). LANA is necessary and sufficient to replicate the viral genome in synchrony with the cell and to properly partition the viral plasmids in dividing cells (11–15). All KSHV-infected cells are uniquely marked by punctate

Editor Thomas Shenk, Princeton University

Copyright © 2022 Landis et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Dirk P. Dittmer, ddittmer@med.unc.edu.

The authors declare no conflict of interest.

This article is a direct contribution from Dirk P. Dittmer, a Fellow of the American Academy of Microbiology, who arranged for and secured reviews by Kenneth Kaye, Harvard Medical School/Brigham and Women's Hospital, and Jae Jung, Cleveland Clinic.

Received 18 November 2021

Accepted 23 November 2021

Published 4 January 2022

foci in the nucleus of infected cells that are composed of the LANA protein and viral DNA (16, 17). Since the LANA protein has a half-life of several days, even cells that initiate lytic viral replication remain positive for LANA protein. Moreover, in those cells, LANA transcription is maintained by the second, lytic promoter of LANA (18). The KSHV latency-locus promoter also directs the expression of the viral cycling homolog (vCyc; ORF 72), the ORF 71 protein, which is called vFLIP, and the K12/Kaposin protein as well as all viral microRNAs (19). K12/Kaposin and vCyc mRNAs can also independently initiate at internal promoters (20, 21). The K12/Kaposin mRNA is the most abundant RNA in KSHV-infected cells and can easily be detected by *in situ* hybridization (22). It gives rise to the Kaposin protein variants as well as the microRNA miR-K12-10a, which is embedded within the Kaposin ORF (23).

Genome-wide transcriptional profiling studies using reverse hybridization (24–30), real-time quantitative PCR arrays (31–34), and bulk RNA sequencing (RNAseq) (35, 36) have generated a complete picture of the average transcription pattern of KSHV in Kaposin's sarcoma tumor lesions, in experimentally infected cells, and in primary effusion lymphoma (PEL) cell lines. These studies uncovered that different KSHV-associated diseases, such as KS, PEL, or multicentric Castlemann's disease (MCD), exhibit different transcription programs, and that even within a disease, such as KS, different primary lesions exhibit widely varying KSHV transcription patterns (20, 33, 34).

PEL cell lines are the workhorses of KSHV research (37–39). Unlike cells derived from KS lesions, cells derived from a primary patient PEL exudate can be adapted to grow in culture indefinitely (40, 41). These long-term propagated PEL cell lines, as well as primary patient-derived PEL explants, reproducibly form tumors in immune-deficient mice (42, 43). Adaptation to culture, however, is not a very efficient process. Few robust PEL cell lines exist, and the most used PEL cell lines have been in continued culture for over 20 years. They are genetically stable but have acquired many genetic alterations in the human host genome (44, 45). In sum, PEL represents a reproducible and robust model to probe KSHV transcription at the single-cell level.

Viral latent protein expression is continuously required for PEL survival (17, 46). Since PEL tumors and cell lines grow in the presence of the KSHV polymerase inhibitors ganciclovir and Foscarnet (trisodium phosphonoformate) (43, 47), KSHV lytic DNA replication and virion formation is not required for PEL survival. Notably, polymerase inhibitor experiments make no claims as to the need for intermittent immediate-early or early protein expression or for late gene transcription that is not strictly dependent on viral DNA polymerase-mediated DNA replication.

BCBL-1 is perhaps the best-characterized PEL-derived cell line in existence. BCBL-1 was initially derived from a patient with AIDS (37, 48). Unlike other PEL, BCBL-1 does not contain Epstein-Barr virus (EBV). The BCBL-1 strain of KSHV is replication competent and infectious (37, 49, 50). The KSHV strain from BCBL-1 cells has been sequenced multiple times, and the BCBL-1 exome, transcriptome, and epigenome have been described in detail (2–4, 24, 31, 35, 36, 51). The BCBL-1 proteome, metabolome, and kinome are known as well (35, 36, 52, 53). For all intents and purposes, BCBL-1 cultures are considered clonal, overwhelmingly in a latent state of the viral life cycle, and stable regarding viral and cellular gene transcription.

Under normal growth conditions, a minority fraction of BCBL-1 reactivates KSHV spontaneously. The frequency of spontaneous reactivation varies depending on time in continuous culture and health of the culture but does not exceed 5%. The fraction of reactivating cells can be increased to 25% when exposed to saturating conditions of chemical inducers, such as phorbol esters, and/or histone deacetylase (HDAC) inhibitors, such as butyrate or valproic acid (6, 54, 55). The fraction of reactivating cells can be increased to over 50% when providing the KSHV Rta/ORF 50 protein in *trans* (56, 57). This suggests (i) that there exists a natural, host cell signaling-dependent mechanism to change viral transcription and (ii) that virus-intrinsic and cell-dependent layers of regulation render each single cell slightly different from its neighbor.

The following experiments use single-cell RNA sequencing (scRNAseq) to test the hypothesis that, prior to chemical induction, all BCBL-1 cells exhibit the same cellular

and viral transcription pattern or that, alternatively, there exists one or more subsets of cells that spontaneously initiated immediate-early, early, or late viral transcription.

scRNAseq represents different things to different people. First, scRNAseq enables high-confidence, high-throughput transcription profiling of a limited number of cells, such as those obtained by laser capture microscopy, fluorescence-activated cell sorting (FACS), or microfluidics, as adopted for the 10xGenomics or BD Rhapsody platform. FACS can be tuned to deliver at most one cell per well into a 96-well culture plate, which is the approach that was used here. Second, scRNAseq enables the analysis of complex mixtures of cells, such as those that exist within a primary tumor biopsy specimen of mixed cellularity or within a heterogeneously infected culture of cells. Recently, a number of studies have applied scRNAseq to acute virus infections, such as to human immunodeficiency virus (HIV) (58, 59), influenza virus (60, 61), or flaviviruses (62). scRNAseq has also been applied to the analysis of mixed cell populations that are latently infected with human cytomegalovirus and other herpesviruses, yielding somewhat contradictory results (63–65). Despite this progress, our understanding of the technical and biological intricacies of applying scRNAseq to virally infected cells remains incomplete.

In this study, scRNAseq is used in its original sense. Given a population of cells, which by any other method, including bulk RNAseq, appears homogenous, one can nevertheless identify novel subpopulations of cells based on their individual transcriptional activity. This is an important question, as it can lead to the discovery of cancer stem cells, of drug-resistant or virus-replication susceptible cells, and of novel biomarkers to characterize those subpopulations.

The following report is divided into three parts. Part one describes the pattern of KSHV transcription in latently infected PEL cells at the single-cell level. Two different PEL cell lines, JSC-1 and BCBL-1, could easily be distinguished based on differing viral transcription alone. This finding reaffirms the requirement to always test multiple cell lines and articulates a need to generate additional PEL cell lines to deliver insights of high rigor, robustness, and repeatability. Several viral genes, such as viral IL-6 (vIL6)/ORF K2, were affirmed as frequently transcribed genes in a fraction of cells. Thus, vIL6/ORF K2 mirrored the variable expression seen clinically in KSHV-associated multicentric Castleman's disease (16, 66, 67). Part two describes biological and technical variability in the experimental designs and how these features affect the interpretation of scRNAseq data for viral infections. scRNAseq data have less than ideal statistical properties (68, 69). The scRNAseq data for viral gene transcription had even worse statistical properties, since viral genes are transcribed at much lower levels than many cellular genes. Hence, it was important to understand nonrelevant sources of experimental noise and to assign explicit confidence levels to the different clusters of cells and genes. Part three describes the cellular profile of PEL by scRNAseq. Lastly, this study also explores a newly appreciated predilection of some BCBL-1 cells, but not of JSC-1 cells, to grow attached to fibronectin on plastic surfaces. This *in vitro* behavior mirrors the clinical findings that describe an extracavitary variant of PEL (70, 71).

RESULTS

Viral transcription patterns in latently infected clonal PEL cell lines. To understand variations in latent KSHV transcription at the single-cell level, viral mRNA levels in PEL were determined by scRNAseq. As scRNAseq is still a relatively new technology, it was important to explicitly explore sources of variability to find robust and repeatable patterns of viral gene transcription (Table 1). Two different PEL cell lines were used, BCBL-1 and JSC-1. To explore differences due to variation within a single cell line, cells were single-cell sorted at three different time points or dates from the same initial seed culture. Each time point was separated by 1 week, i.e., multiple doubling times, from the other. The cells were maintained under ideal culture conditions and collected in mid-log phase for analysis. The experiment was repeated several months later. Growth rates were equivalent across consecutive passages (Fig. 1A), as cells were seeded at the exact same density of 0.25×10^6 cells/ml at day zero and passaged before reaching stationary phase around day five.

Two types of culture conditions were compared. BCBL-1 cells were either grown in suspension in upright flasks or attached to fibronectin-coated plastic in horizontal flasks (Fig. 1B and C; see also Movie S1 in the supplemental material). Individual cells were sorted

TABLE 1 Summary of experiments^a

Plate ($N = 10$), ε_1	Data ($N = 12$)	Date (yr-mo-day) ($N = 5$), ε_2	Origin ($N = 2$), ε_3	Cell ($N = 2$), ε_4	Wells ($N = 960$), ε_5	Adherence ($N = 2$), ε_6
1	a	2018-09-04	BCBL1	One	96	No
2	b	2018-09-11	BCBL1	One	96	No
3	c	2018-09-26	BCBL1	One	96	No
4	d	2018-09-26	BCBL1	One	96	No
5	e	2018-09-26	BCBL1	One	96	No
	f	2018-09-26	BCBL1	One	0	No
6	g	2018-09-30	JSC1	One	96	No
	h	2018-09-30	JSC1	One	0	No
7	i	2018-09-30	BCBL1	One	96	Yes
8	j	2018-09-30	BCBL1	Two	96	Yes
9	k	2019-08-08	BCBL1	One	96	Yes
10	l	2019-08-08	BCBL1	One	96	Yes

^aThe cells were sorted into 10 individual 96-well plates (wells and plate columns) at 5 different dates (date column), each representing a biological replicate or different populations of cells of the same monoclonal origin (origin column). For plate 8, two cells (cell column) were sorted into each well. For plates 7 to 10, BCBL-1 cells were allowed to adhere to plastic and nonadherent cells washed off. ε refers to the factor that was used to normalize across the experiment.

into individual wells of a 96-well plate, and libraries were prepared and subjected to scRNAseq. For two biological replicates, the same library was sequenced on two different Illumina lanes. This served as a technical repeat to estimate sequencing accuracy and library depth effect. Each batch of cells produced a comparable number of reads and unique molecular indices (UMI) (Table 2), attesting to the robustness and reproducibility of this method. In addition to reads mapping to the human genome, reads mapping to the KSHV genome were recovered across all samples.

Next-generation sequencing reads were mapped to the KSHV reference genome [NM_009333](#) and counts per open reading frame (ORF) computed. The KSHV viral miRNA loci were removed prior to analysis, as these are not captured in the cDNA-generating step. The viral latent genes K12, vFLIP, vCyc, and LANA were combined, since transcripts across the latency locus are 3' coterminal and the scRNAseq assay uses poly(dT) in the initial priming step. Only cells that passed quality control (QC), based on the analysis of cellular transcripts, were used for subsequent detailed analyses. Most of the cells that failed QC and are not represented in Fig. 2 had no detectable KSHV transcripts. As the same cells also had less than the minimum permissible number of cellular mRNAs detected, these most likely represent technical failures.

The LANA/Kaposin/latency locus and nut-1/PAN RNAs were detected even in some of the cells that failed QC. These are the most abundant transcripts in any KSHV-infected cell.

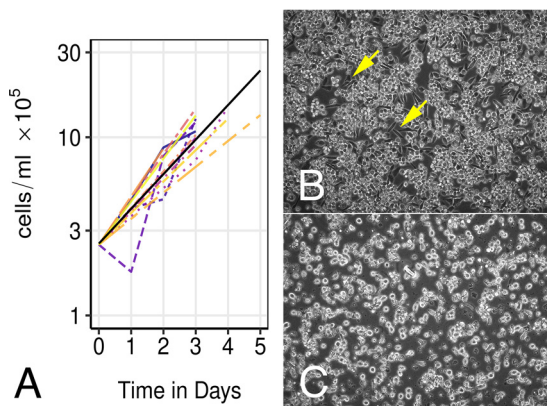


FIG 1 Cell growth characteristics and adhesion phenotype of BCBL1 cells. (A) Shown are growth curves of BCBL-1 at consecutive passages and the linear fit across all the data (black) ($n = 13$ passages). \log_{10} value of cell concentration, in cells per milliliter, is shown on the vertical axis and days postseeding on the horizontal axis. Cells were seeded at 0.25×10^6 cells/ml on day 0. (B and C) Shown is an image (200 \times) of cultures of either BCBL1 (B) or JSC1 (C) cells that were grown on horizontal fibronectin-coated flasks. The yellow arrows point to some spindle-shaped cells that are flattened out and tightly attached to the fibronectin matrix.

TABLE 2 Summary of next-generation sequencing^a

Name	Cell (no. of reads ^b)	Cell (no. of UMI)	Cell (no. of genes)	Virus (no. of reads)	Virus (no. of genes)
Plate 1	47,898,415	1,686,753	11,627	737	63
Plate 2	42,308,862	1,051,764	10,774	766	58
Plate 3	70,460,719	1,784,666	11,762	1,917	72
Plate 4	61,547,640	1,854,635	11,980	1,933	73
Plate 5a ^c	66,497,202	1,495,992	11,631	1,682	75
Plate 5b ^c	61,378,574	1,561,116	11,678	1,829	75
Plate 6a ^c	42,925,808	1,685,717	11,933	10,217	70
Plate 6b ^c	56,591,011	1,848,909	12,027	11,443	74
Plate 7	97,327,587	3,462,436	13,105	50,485	78
Plate 8	86,963,518	2,457,892	12,403	3,377	77
Plate 9	47,168,617	2,604,976	14,449	1,218	72
Plate 10	57,609,160	1,361,365	10,518	1,116	73
Median ± SEM	$6 \times 10^7 \pm 5 \times 10^6$	$1.9 \times 10^6 \pm 1.8 \times 10^5$	$11,847 \pm 296$	$1,873 \pm 4,070$	73 ± 2

^aOne Illumina library was prepared for each plate. For plates 5 and 6, the same library was sequenced twice, indicated by a and b. Shown is the total number of reads mapping to either the cell or virus genome, the number of UMIs, and the number of genes detected by at least one read.

^bData are from 2×250 paired reads at 3 libraries per lane on an Illumina HiSeq instrument (GenWiz Inc., Plainfield, NJ).

^cData sets a and b were combined for data analysis unless indicated otherwise.

Some of the cells that failed QC expressed only the latency transcripts/Kaposin, some only nut-1/PAN, and some both, suggesting that these RNAs were more abundant than any other mRNA. If each cell *a priori* contains both at equivalent levels, this combination pattern (not A and not B, only A, only B, and A and B) is consistent with random fluctuations at the limit of detection for this assay, as would be expected for cells that failed QC precisely because too few transcripts overall were recorded. Conversely, we submit that KSHV mRNA signals in cells that passed QC were not limited by sensitivity and reflect biological fluctuations. Suboptimal library depth was not the reason for differences in KSHV mRNA levels in those cells that passed QC, since sequencing the same library twice (Table 1, a and b replicates) did not change the analysis.

Figure 2 shows a heatmap representation of KSHV gene transcription for each KSHV gene and each cell in the experiment after unsupervised hierarchical clustering. A total of $n = 841$ cells were analyzed across $m = 84$ viral genes. The viral genes are listed in order of their genome location and prior classification into latent, immediate-early, early, and late genes, indicated by colors on the right and the different plates corresponding to the various technical and biological replicates colored on top. Figure 2A shows a binary representation requiring at least two reads to map to an ORF to be marked as detected (blue). Figure 2B shows the log-transformed counts. Thus, the process shown in Fig. 2A maximizes specificity as it filters out genes with only one mapped read, which could be due to misalignment, contamination, or transcriptional noise. This analysis led to the following conclusions.

- (i) The most consistently detected mRNAs were nut-1/Pan/T1.1. This was followed by the latency mRNAs mapping to ORF K12, vFLIP, vCYC, and LANA and the vIRF locus, including vIRF1, vIRF2, vIRF3 (LANA-2), and vIRF4. Some cells transcribed nut-1/PAN but not the latency transcripts, others transcribed the latency transcripts but not nut-1/PAN, and others transcribed both. Thus, the scRNAseq broadly confirmed the known transcription profiles of the previously identified viral latent genes in the majority of PEL cells in both cell lines.
- (ii) Less than 1% of cells transcribed all viral transcripts or significant amounts of late viral genes. This confirms that under optimal culture conditions, PEL are tightly latent. External stimuli are required to induce lytic reactivation and genome replication. Maintaining PEL under ideal growth conditions (Fig. 1) minimized the spontaneous reactivation rate to yield a lower limit of spontaneous reactivation.
- (iii) In some cells (plate 9 [yellow] and, to a lesser degree, plates 1 and 2), no KSHV transcription could be detected at all. In those cells, the entire KSHV genome was transcriptionally silent, or rather the majority of KSHV plasmids within a single cell were inactive even across the LANA promoter (1). This is an unexpected finding, as to date no one has been able to physically isolate a stable subclone of any PEL cell

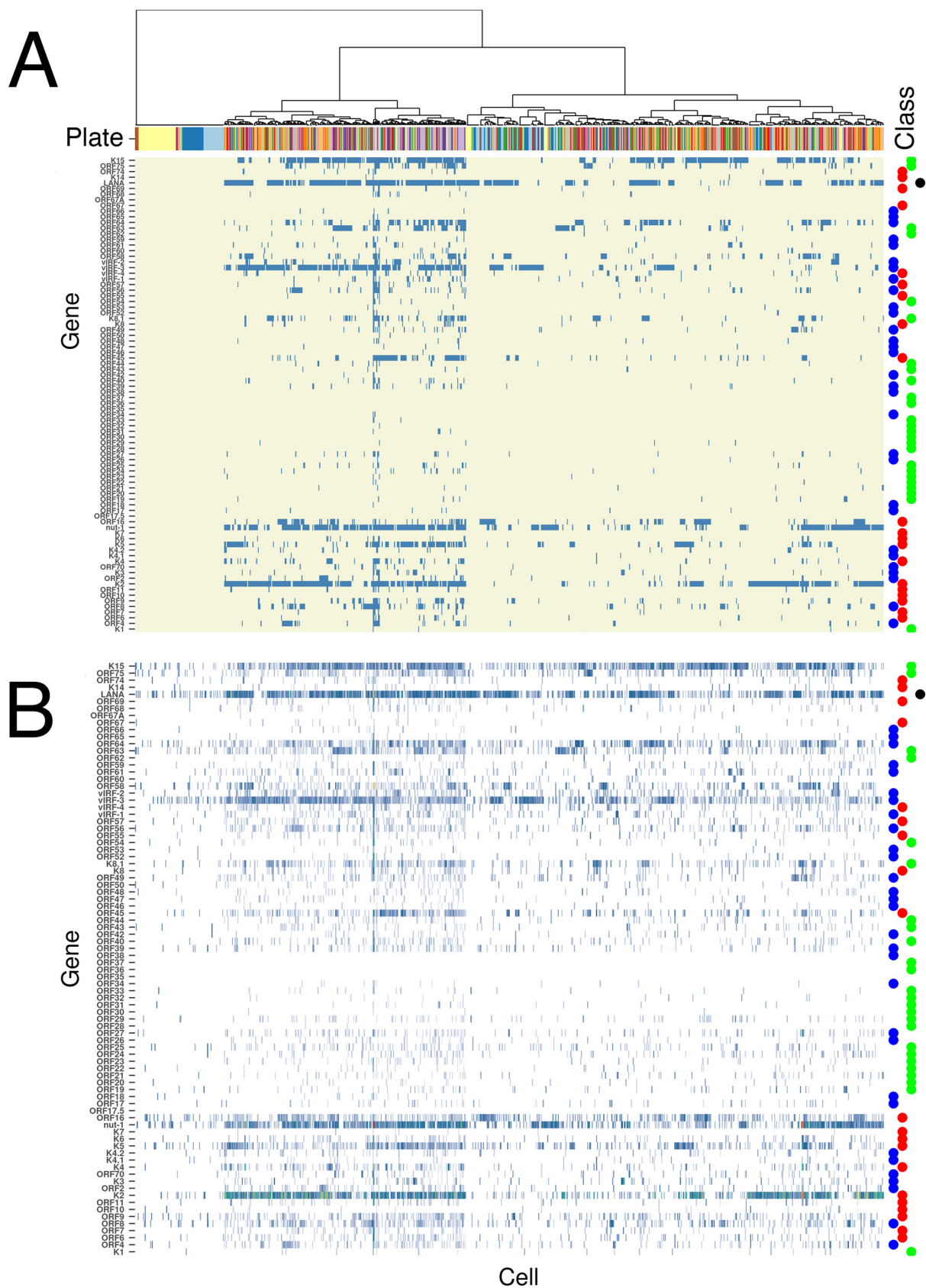


FIG 2 Heatmap of KSHV transcription in PEL. Shown are the results of unsupervised hierarchical clustering of PEL cells (columns) by KSHV genes (rows). Genes are arranged in order from 5' (bottom) to 3' (top). A dendrogram on top represents the clustering results, which were (Continued on next page)

line that is devoid of the KSHV plasmid. However, this finding does correlate with *in situ* studies of LANA (11, 17, 72), which reported variability in the number of LANA dots, a surrogate for the presence of KSHV and LANA transcription per cell. Thus, while not stable over time and subject to extreme negative selection, it seems as if into every generation a transcriptionally silent cell is born, one cell in each population.

- (iv) A significant fraction of cells transcribed vIL6/K2 or ORF 2/vDHFR. vIL6 and ORF2/vDHFR are 3' coterminal and colinear across the entire ORF of vIL6/ORF K2. Hence, this assay could not discern which transcript was measured. vIL6/ORF K2 was the discerning factor between the JSC-1 and BCBL-1 cell lines. Most JSC-1 cells transcribed ORF K2/vIL6. By comparison, few BCBL-1 cells did. It is well established that a subpopulation of PEL, MCD, and KS express ORF K2/vIL6 but also that ORF K2/vIL6 expression is heterogenous and can be regulated by cellular signaling pathways independently of KSHV lytic replication (73, 74). The ORF K2/vIL6 was among the most differentially regulated viral mRNAs among KSHV latently infected B cells.

To obtain an unbiased impression of the population structure of the experiment, each cell was assigned to one cluster with related viral gene expression (unsupervised clustering). Each gene was considered following a negative binomial distribution (Fig. 3A). Figure 3B shows a visualization using principal-component analysis (PCA) for the viral genes. Mirroring the raw data, these fall into two groups: those with broad expression across many cells, which includes nut-1, LANA, vIL6/K2, K15, and vIRF-3, each of which are separate from the others, and the remainder of the genes that have a more sporadic expression pattern particular to individual cells, cell lines, or growth conditions. Together, these 5 transcripts are responsible for 47% of the variation in the data (PC1).

The uniform manifold approximation and projection (UMAP) representation of cells' clusters are not enriched for particular plates or dates (Fig. 3C); rather, they represent cells of similar viral transcription patterns (Fig. 3D). The exception is plate 6, which represents JSC-1 cells only, whereas all other plates represent BCBL-1 cells. Hence, in this case the plate/batch separation mirrors the viral transcription clusters.

As a result of unsupervised clustering, each cell in the experiment was assigned to a cluster representing a specific, composite pattern of KSHV transcription (Table 3). For instance, plates 9 and 10 were enriched for cells with a gene expression pattern identified by clusters 1 and 5 but entirely devoid of cells with a gene expression pattern indicative of clusters 2 and 4. Plate 6 showed the JSC-1 cell-specific pattern of viral transcription.

Each of these clusters was defined by a KSHV gene signature, rarely a single gene. For instance, clusters 2 and 3 contain the cells with the highest numbers of latency transcripts (Fig. 3E). Clusters 2 and 4 show elevated levels of vIL6/ORF K2 (Fig. 3F) relative to the cells in all other clusters. Clusters 1 and 2 show high levels of nut-1/PAN (Fig. 3G). The clusters seem to be identified with the following profiles: clusters 3, 4, and 1 are high in latency locus, vIL6, and nut-1, respectively, cluster 2 is high in all the aforementioned genes, and cluster 5 is not enriched in any of the genes. The example of vIL6/ORF K2 contains an interesting lesson. Overall, JSC-1 expresses much more vIL6/ORF K2 than BCBL-1 (cluster 4). Hence, experiments aimed at understanding the role of vIL6 in PEL pathogenesis will give different results depending on which cell line is used, as would agents that target IL-6, IL-6 receptor, or IL-6-dependent signaling. However, in each BCBL-1 population there exists a sizable fraction of cells (cluster 2) that transcribe high levels of vIL6/ORF K2. Hence, there would be a signal in BCBL-1 if agents, which target vIL6, IL-6 receptor, or IL-6-dependent signaling, were tested in BCBL-1. Conversely, this suggests that independence from vIL-6 signaling can arise rapidly in PEL, as many cells do fine without vIL6/ORF K2.

This analysis provided the class assignments, also called cluster bins, used in the last part of the manuscript to test the hypothesis that the different patterns of KSHV transcription

FIG 2 Legend (Continued)

based on the binary matrix of detected/undetected genes. A gene was detected if it had at least 2 UMI counts and undetected otherwise. The top heatmap represents this binary signal, whereby detected genes are labeled in blue and nondetected genes are labeled in yellow. The lower heatmap represents the KSHV transcription pattern after log normalization. With this, we can visualize how many genes possessed exactly 1 count.

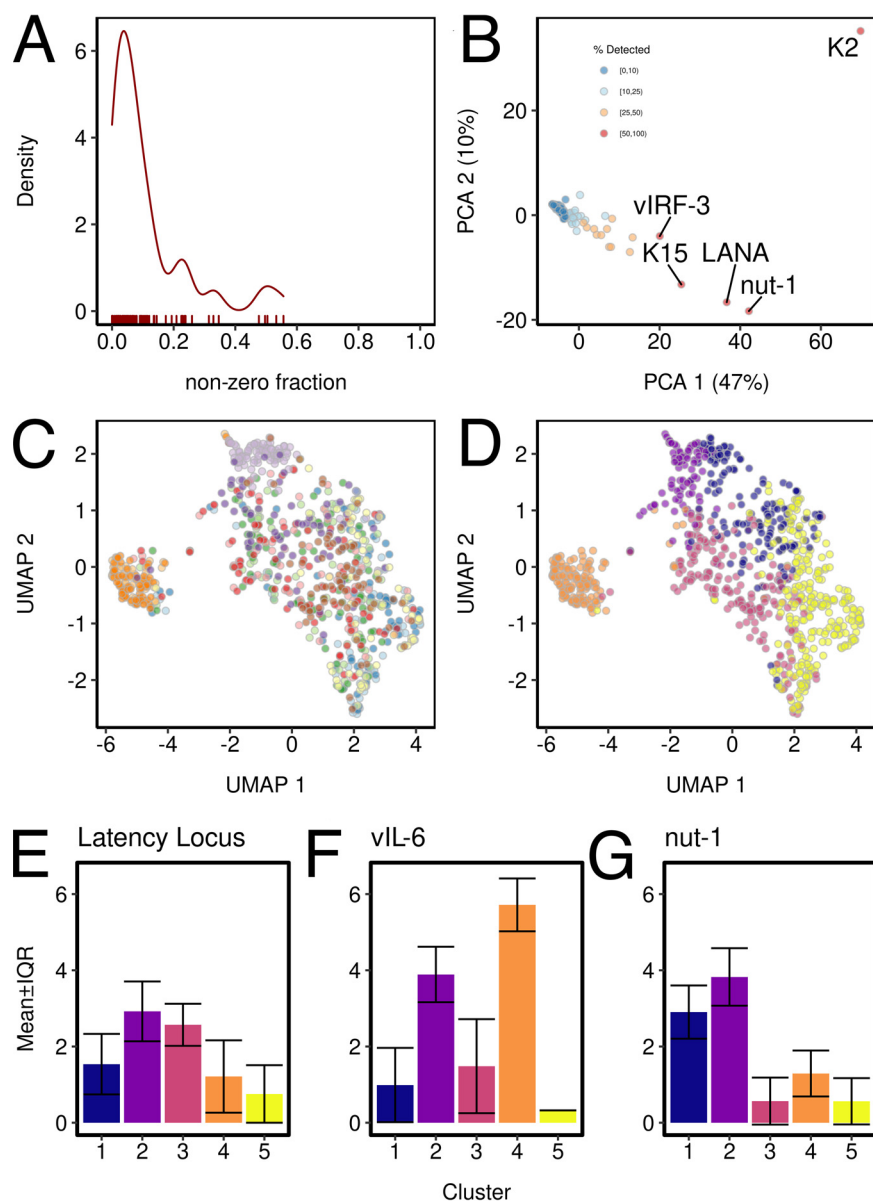


FIG 3 KSHV-based clustering of PEL. (A) Distribution of the detection rates (nonzero fraction) on the horizontal axis. This is a measure of the zero inflation for each gene. (B) Principal-component analysis (PCA) of KSHV genes. LANA, nut-1, and the vIRF have the most distinct expression patterns, and LANA and nut-1 were the most abundantly expressed genes (PC1). (C and D) The same UMAP plot of all cells. In panel C, the cells are color coded by the plate/library number (plate 1, light blue; plate 2, dark blue; plate 3, light green; plate 4, dark green; plate 5A, light red; plate 5B, dark red; plate 6A, light orange; plate 6B, dark orange; plate 7, light purple; plate 8, dark purple; plate 9, yellow; plate 10, brown), and in panel D the cells are color coded by the cluster designation as identified by unsupervised analysis using K-medoids. Cluster identification and per-plate saturation can be found in Table 3. (E) Mean \pm interquartile range (IQR)-adjusted counts for the latency locus in each of the clusters, i.e., the relative expression of LANA and the same for vIL-6 (F) and nut-1 (G).

within the same cell culture flask correlate with different patterns or pathways of host gene transcription.

The cellular single-cell transcriptome of PEL. As scRNAseq is still developing as a technology, it was important to determine the experimental factors that contribute to data variability. The overall goal was to report the most robust, reproducible, and specific results that could be garnered from this experimental data. The experimental setup reflected these goals (Table 1) and allowed for detailed analyses using pairwise comparisons. The details of these validation studies are described in Materials and

TABLE 3 Viral transcription patterns in latently infected clonal PEL cell lines^a

Plate	Cell line	Value (%) for cluster no. ^b				
		1	2	3	4	5
1	BCBL-1	1 (*)	0 (*)	12 (16)	5 (*)	56 (76)
2	BCBL-1	9 (11)	3 (*)	18 (23)	3 (*)	47 (59)
3	BCBL-1	15 (20)	5 (*)	29 (39)	2 (*)	24 (32)
4	BCBL-1	8 (11)	6 (*)	29 (42)	3 (*)	23 (33)
5A	BCBL-1	11 (16)	6 (*)	34 (48)	4 (*)	16 (23)
5B	BCBL-1	12 (17)	8 (11)	33 (47)	3 (*)	15 (21)
6A	JSC-1	0 (*)	1 (*)	0 (*)	56 (98)	0 (*)
6B	JSC-1	0 (*)	1 (*)	0 (*)	57 (98)	0 (*)
7	BCBL-1	39 (45)	46 (54)	0 (*)	0 (*)	1 (*)
8	BCBL-1	25 (39)	13 (21)	16 (25)	3 (*)	6 (10)
9	BCBL-1	6 (*)	0 (*)	2 (*)	1 (*)	60 (89)
10	BCBL-1	25 (38)	1 (*)	19 (28)	0 (*)	23 (34)

^aShown are the numbers of cells for each plate classified into each unsupervised k-means cluster. Each entry also contains an additional value indicating the percentage of cells within a plate that was classified to a given cluster.

^bShown in parentheses is the percentage of cells or an asterisk if less than 10% of cells are in a particular cluster. Indicated in boldface are the most abundant contributors to each cluster.

Methods and Fig. S2. These studies established robust, reproducible, and rigorous boundaries in which to interpret these results. The validation studies on cellular transcripts affirmed conventional wisdom: differences in transcription were greater between two different cell lines than among cells originating from the same clonal population, and biological day-to-day variability was greater than variability due to technical differences in single-cell library preparation or sequencing.

The transcription profile differed significantly between JSC-1 cells and BCBL-1 cells, as shown in Fig. 3 by UMAP analysis. Hence, BCBL-1 cells were analyzed separately to find heterogeneity within a single cell line. This data set encompasses different biological replicates, including culture enriched for adherent cells. Figure 4A shows the result of hierarchical clustering for BCBL-1 cells based on cellular transcript levels. Below the dendrogram are indicated the clusters as determined by viral gene expression and the types of biologically distinct cell types: BCBL-1 floating or BCBL-1 adhering to plastic. Note, however, that this experimental approach represents a rather crude method to select these two states or populations of BCBL-1 cells. Four clusters were identified based on cellular transcription alone and were statistically significant, with a P value of ≤ 0.05 after adjustment for false discovery rate (FDR). These “cellular” clusters did not correlate with viral transcription clusters or the adherent/suspension enrichment status of the BCBL-1 cells.

To identify genes that defined these clusters, a pairwise comparison was performed at each significant hierarchical split. As a positive control, Fig. 4B shows the result for genes that differed between JSC-1 and BCBL-1 cells in the form of a volcano plot. As expected, many transcripts were dramatically different between two cell lines. Note the scales for the fold change on the horizontal and the statistical significance on the vertical axis. The panels C, D, and E in Fig. 4 all use the same scale to allow for comparison. Figure 4C shows transcripts that were differentially regulated within BCBL-1 cells at $P \leq 0.001$, the top level of the hierarchical cluster. There were 16 transcripts. Their degree of difference was not as great as that between cell lines (compare the scale on both the horizontal axis depicting fold change and the vertical axis depicting the significance level). A complete list of differentially regulated cellular genes is provided in Table S1. Nevertheless, these transcripts pointed to a common biological pathway: PCNA, TUBA1B, TYMS, UBE2T, FEN1, TK1, DUT, STMN1, MCM7, H2AFZ, HMGB2, GMNN, CKS1B, KIAA0101, CKS2, and RPM2. Most of these are cell cycle-regulated genes, in particular genes that are involved in nucleotide synthesis and DNA replication. Many more cell cycle-regulated genes differed between the top two clusters but individually did exceed our significance threshold. Figure 4D shows the next (left-level) cluster split out cells in mitosis as ascertained by the significantly regulated genes CCNB1, CKS2, PLK1, CDC20, TPX2, CCNB2, and NUSAP1. Figure 4E, the right-level cluster split, shows no other genes were differentially regulated

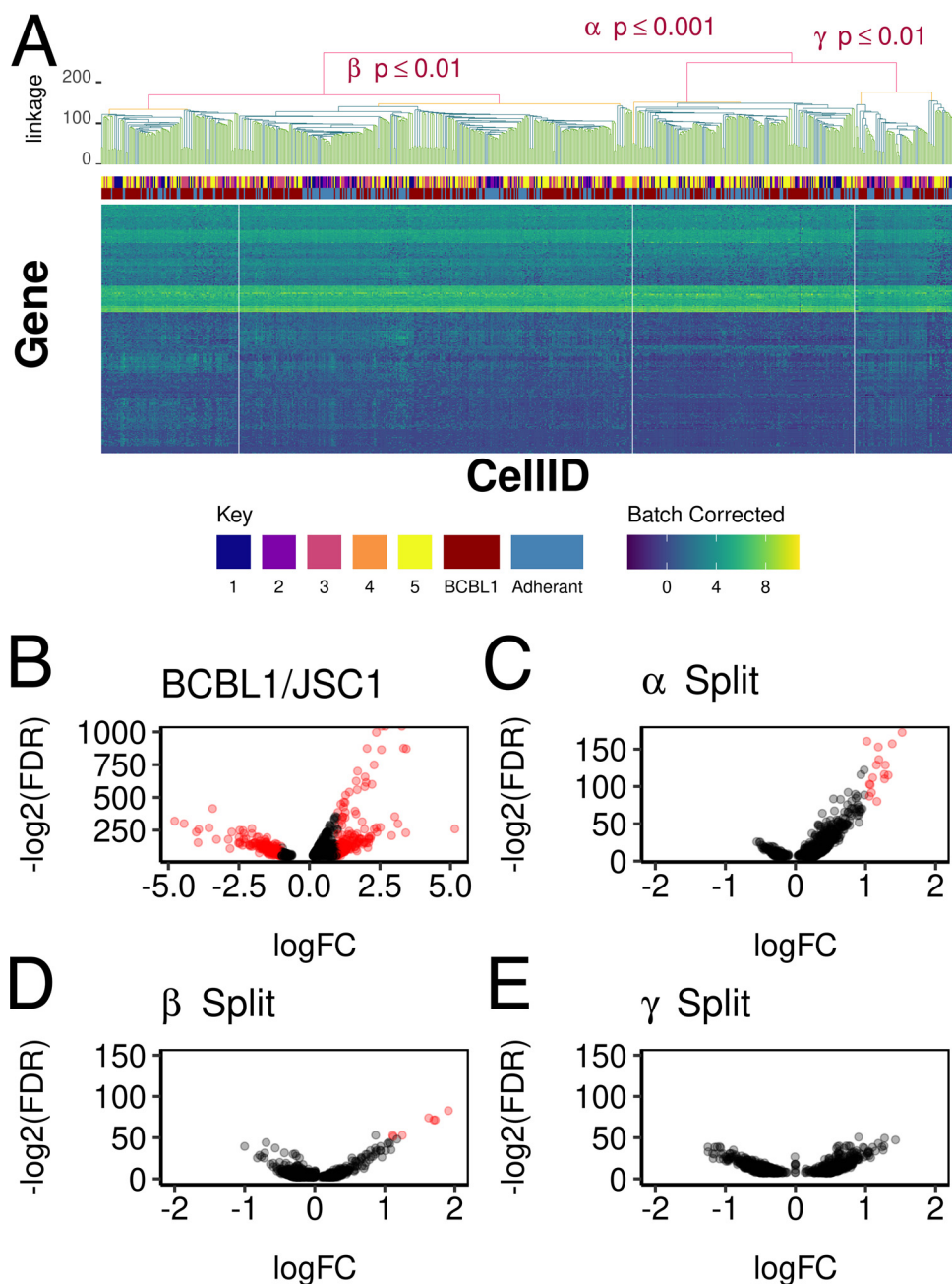


FIG 4 Association between KSHV and host cell transcription. Heatmap of cellular transcription of BCBL-1 cells after batch correction. The associated SigClust dendrogram is above with annotations of which clusters are significant. The KSHV K-means clustering assignments and cell group identity are mapped between the heatmap and the dendrogram. Panel B shows a volcano plot of genes that were differentially regulated between BCBL-1 and JSC-1 cells, panel C shows a volcano plot of genes that signify differences at the major split ($P \leq 0.001$), panel D shows a volcano plot of the left minor split, and panel E shows a volcano plot of the right minor split ($P \leq 0.01$). The individual genes are listed in Table S1.

within BCBL-1 populations. Specific Gene Ontology pathways for significantly differentially expressed genes of Fig. 4B to D can be seen in Fig. S3. This result suggests that BCBL-1 cells grown under optimal conditions in suspension represent a homogenous population with regard to cellular gene transcription.

In every culture, some BCBL-1 cells adhere to the plastic wall (Fig. 1A and B and Movie S1). This was not the phenotype of a particular subclone of BCBL-1 cells but appeared repeatedly at each new passage. It is specific for BCBL-1 cells and was not observed for JSC-1 cells. To enrich for this population, BCBL-1 cells were cultured horizontally on fibronectin-coated flasks

in the presence of 5% methyl cellulose. These experiments, however, should be considered preliminary, as they were limited by the purity of the population. No one cellular gene could be identified as a biomarker for the adherent state of BCBL-1 cells. This result supports the notion that even though BCBL-1 cells in culture can adopt two different phenotypes, one representing suspension growth as found in pleural effusions and the other representing tumor cells adhering to the body cavity walls, or even mesenchymal differentiation (75), these phenotypes are likely regulated at the posttranscriptional level.

No one cellular gene emerged as a biomarker that correlated significantly with the different clusters of viral transcription or any one viral gene in BCBL-1 cells. This suggests that, under steady-state conditions, the different patterns of viral transcription represent responses to posttranscriptional signals or even stochastic fluctuations. It may also highlight the limitations of scRNAseq, which only measures the most abundant mRNAs. It is important to also keep in mind that SigClust conducts explicit hypothesis testing to determine how big the distance between any two clusters must be to reach a specific significance cutoff (76). It represents a conservative approach to hierarchical clustering. It is quite possible that, if one had induced massive changes in host and viral transcription, e.g., via HDAC inhibitors or TLR activation across the entire culture (6, 77), correlations between host and viral genes would have emerged. This, however, was not the purpose of these experiments.

DISCUSSION

KSHV establishes latency in B cells (50, 78–82) and KSHV latent genes, including the viral miRNAs that modify the behavior of developing B cells (83–85). Consequently, KSHV has been associated with a varied spectrum of B cell abnormalities, such as localized hyperproliferation, multicentric Castleman's disease (16, 66, 67), microlymphoma (86), and solid lymph node-associated (70, 71) as well as body cavity-based PEL (87, 88), KSHV immune reconstitution syndrome (KS-IRIS; and the presumed B cell activation as measured by B cell inflammatory cytokines, such as IL-6), and KSHV-associated inflammatory syndrome (KICS) (89). Different PEL and PEL cell lines both have common and distinct cellular gene transcription profiles (90, 91). The novel insight of this study is that even within clonally selected, long-term *ex vivo*-propagated PEL cell lines, distinct subpopulations of cells exist.

By applying scRNAseq to herpesvirus latency in PEL, we uncovered subpopulations among PEL that differed from each other in terms of viral and cellular gene expression.

- (i) The latency locus-associated transcripts were present in the majority of cells, as was the Nut-1/Pan/T1.1 lncRNA. This was consistent with prior bulk transcriptional profiling studies in PEL (24, 31). The vIRF-3 and, to a lesser degree, vIRF-1, -2, and -4 transcripts were detectable in many cells, which was consistent with prior studies (92–94). Transcripts traversing the ORF K15 locus were also consistently detected. Thus, the KSHV latency status is tightly regulated in healthy PEL cells and dominated by the known latent genes, which can be considered essential for PEL survival and potential targets of therapy, including therapeutic vaccination.
- (ii) There are limitations to this interpretation, since mRNA levels are dependent on both ongoing transcription and mRNA stability. Typically, scRNAseq experiments are heavily biased towards abundant RNAs. This is of importance for interpreting the Nut-1/T.1.1/PAN results. Whereas the half-life of protein mRNAs is measured in minutes, long noncoding RNAs (lncRNAs) are much more stable. Nut-1/T.1.1/PAN is the most abundant KSHV lncRNA. It is polyadenylated but nuclear localized (26, 51, 95). Nut-1/T.1.1/PAN is also tightly associated with DNA and proteins (95–97). By comparison, the LANA mRNA, which is abundant but, unlike the LANA protein, not particularly stable, was not as consistently detectable as the Nut-1/T.1.1/PAN lncRNA. In the field, Nut-1/T.1.1/PAN lncRNA would be a better biomarker for KSHV-infected cells than LANA mRNA.
- (iii) The transcription pattern of the viral IL-6 homolog ORF K2/vIL6 was notable. It was detectable in many cells of the J5C-1 cell line but only a limited number of cells in the BCBL-1 cell line. These single-cell data mirror the biology of ORF K2/vIL6, which exhibits

highly variable expression across different KSHV diseases and has multiple modes of transcriptional activation, some that are KSHV RTA dependent and some that are independent (56, 66, 98–101). There are limitations to this interpretation, as K2/VIL6 is coterminal and overlaps ORF2 (102). These scRNAseq data are strongly biased towards the 3' end of each transcript, as it uses a poly(dT) primer in the first step. They do not capture full-length mRNAs. Hence, it was not possible to ascertain whether the count data reflect ORF 2 or ORF K2/VIL6 message.

- (iv) As expected, the two different cell lines JSC1 and BCBL1 differed from each other more so than different batches of the same PEL cell line or subpopulations that emerged within the BCBL-1 cell line. Most of these differences manifested themselves in differences in cellular gene expression, reflecting the unique patient origin and mutational profile of each PEL cell line. This suggests that host cell transcription patterns, rather than viral genes, are important determinants of PEL tumor progression and therapy responses.
- (v) The transcription profiles of different batches of the same PEL cell line were similar over time. Batch effects could be removed using a linear model (103, 104). After correction for technical variability, the transcriptionally defined subpopulations within the cell line were stable over subsequent passages, suggesting that they are generated anew at each generation with a fixed frequency. Such a mechanism is consistent with the general partition model for extrachromosomal herpesvirus genomes, which naturally creates transient imbalances in viral gene expression after individual cell division events even if the population as a whole exhibits stable genome partitioning behavior (13, 17, 72). In sum, both clonal mutational events (genetic drift) as well as individual transcription patterns contribute to the variability of each cell line.
- (vi) At any given time, a fraction of cells in PEL cell lines undergoes lytic reactivation, as measured by KSHV protein expression, e.g., ORF K2/VIL6 or ORF K8.1. Spontaneous reactivation events are rare but result in constitutive low-level shedding of infectious particles into the culture supernatant (28, 37). The experiments reported here were designed to limit spontaneous reactivation events as much as possible by maintaining the cells under ideal culture conditions. In addition, FACS eliminated cells with aberrant cellular transcription, indicative of cell death (105–107), including apoptosis induced in response to KSHV replication. Under these conditions, most late-lytic mRNAs (covering ~50% of annotated genome) were not detectable at all, suggesting that fewer than ~1/600 healthy PEL cells at any one time spontaneously enter the lytic phase or that KSHV late transcription is associated with rapid apoptosis and massive changes in overall host transcript levels. The latter hypothesis is consistent with the reported RNA degradation functions of KSHV ORF 37/SOX (108, 109).
- (vii) Many more cells transcribed early viral mRNAs (ORF50/Rta, ORF 49, ORF 45, ORF 48, ORF K5, ORF K4, and ORF K8.1) than late, ganciclovir-sensitive viral mRNAs. This suggests that one reason that the early ORF 45 and ORF K8.1 proteins are consistent markers for KSHV reactivation is that they are readily transcribed even if none of the replication-dependent, late genes are expressed (110). Even though ORF 50/Rta was transcribed in some cells, this did not correlate with complete viral reactivation. This suggests that additional safeguard mechanisms are in place to prevent widespread virus transcription in the case of Rta “blips.” One such candidate mechanism is inhibitory DNA and histone modifications, which are well documented in KSHV-infected cells (1, 2, 4, 6). Other candidate mechanisms include the need for cooperating viral immediate-early genes and/or cellular transcription factors.

The analysis of scRNAseq data is far from routine. Some of the limitations have been alluded to above, while others are more theoretical in nature. For instance, it is difficult to define the minimal cluster size of a single-cell experiment. Are single cells also single clusters? How many similar cells does it take to call a cluster, and how like each other do the transcription profiles of the cells within a cluster have to be? The SigClust algorithm addresses those questions explicitly (76). In flow cytometry, a common conservative estimate

uses $\geq 5\%$ of all events to define a new cluster (107). The five clusters defined in Fig. 3 fall above this threshold and have ≥ 70 cells each.

A recent report advocates for a mesenchymal origin for PEL (75). The authors noted that in effusions of HIV⁺ PEL patients, a population of adherent mesenchymal cells existed, which, with time, gave rise to B1 lymphoid cells. Whereas the adherent mesenchymal cells were KSHV negative, the emerging B1 lymphoid cells in some of the cultures were KSHV positive. Similar claims have been advanced the postulate of a mesenchymal origin of KS (111–117). BCBL-1 cells, but not JSC-1 cells, had an innate ability to adhere to the plastic side of the culture flask, which could be enhanced by fibronectin coating and reducing turbulence. Unfortunately, this crude method of culture was not able to obtain a pure enough population to yield a statistically different cellular or viral gene expression signature. Whether this attachment phenotype reflects a cell cycle state or is due to posttranscriptional regulatory events alone remains an open question. These studies underscore the plasticity of KSHV-infected cells by identifying five clusters of viral transcription. They are consistent with the general impression that KSHV driving endothelial reprogramming (blood endothelial to lymphatic endothelial, or BEC to LEC) as well as endothelial-to-mesenchymal transition (EMT) but should not be construed as favoring the same idea for PEL. Further studies are needed. Understanding the population composition of KSHV-associated malignancies at the single-cell level represents the next step, can be expected to yield important insights into the fundamental biology of the virus, and may uncover new avenues of intervention.

MATERIALS AND METHODS

Tissue culture. Cells were obtained from the American Type Tissue Culture Collection (ATCC) or the NIH AIDS reagent program. Cells were cultured in RPMI 1640 (Gibco) supplemented with 100 U/ml penicillin-streptomycin (Gibco), 2 mM L-glutamine (Gibco), and 10% Fetalgro bovine serum (VWR). Cells were maintained at 37°C in 5% CO₂ and passaged for no more than 3 months at a time. Specifically, cultures were seeded at 0.25×10^6 cells/ml and used at 48 h after culture initiation or diluted 1:5 at 72 h for routine passage. Cells were imaged continuously on a LEICA PAULA smart cell imager to quantitate growth and density. Cell identity was confirmed by STR typing (PowerPlex 16HS assay; Promega). All cells underwent periodical mycoplasma testing (LT07-701; Lonza).

For adherent cell enrichment, a fibronectin-coated 25-cm² plate was inoculated with cells in 10 ml of 50% culture medium (0.5×10^6 cells)–5% methylcellulose to facilitate adherence. After 48 h, regular RPMI replaced the spent methylcellulose-containing medium, and the cells were maintained until the day of scRNAseq (the time lapse recording was on adherent cells cultured without methylcellulose). At that time, the culture medium was discarded and the adherent cells washed twice with phosphate-buffered saline (PBS) without magnesium (14190-144; Gibco). Next, the cells were incubated in trypsin-PBS for 10 min at 37°C and pelleted by low-speed centrifugation. The cell pellet was suspended in 1,000 μ l of Hanks' balanced salt solution with 2% fetal bovine serum, and the cells were passed through 35- μ m cell strainers (352235; Corning Falcon). The cells were sorted directly into prepared Precise WTA single-cell encoding plates (634100; BD) on 96-well plates using a Becton, Dickinson FACSARIA II while leaving four wells empty for quality control per the manufacturer's protocol. The manufacturer's protocol settings ensure that at most 1 cell per well is deposited. For plate 8, two cells were sorted into each well.

Single-cell RNA sequencing. These studies used the Precise whole-transcriptome analysis single-cell kit (634100; BD) by following the manufacturer's protocol. The pooled library was submitted to GenWiz Inc. (Plainfield, NJ) for Illumina HiSeq sequencing using 2×250 bp paired reads at 3 pools per lane with 15% Phi-X spike as a control. The median \pm standard error of the mean per library was $6 \times 10^7 \pm 5 \times 10^6$ reads, $1 \times 10^6 \pm 2 \times 10^5$ unique molecular identifiers (UMI), and $11,847 \pm 296$ genes. Note that libraries 5 and 6 were sequenced twice on two different HiSeq lanes.

For mapping the scRNA data to the human genome, the manufacturer-supplied analysis pipeline (Precise whole transcriptome assay analysis pipeline v2.0; BD) was used as implemented at <https://www.sevenbridges.com/>. Rather than using total counts per ORF, the BD methodology calculates unique molecular identifiers (UMI), which reduces reverse transcription-PCR bias (105, 118). The manufacturer-supplied analysis pipeline yielded the UMI data matrix for each human gene in each cell for each experiment, termed data set 1.

To map the scRNA reads to the KSHV genome, a methodology similar to that for the human genome was used. This allowed for consistent comparison of UMI counts between reads mapping to the human genome versus the KSHV genome.

Prior to analysis, the cells were subjected to a series of quality control measures and filtering steps based on human gene transcription. The code and data are available on the paper bitbucket repository. The following are the filtering conditions used in the context of 20,240 protein-coding genes and 1,152 cells. Cells were required to have greater than 5,000 total UMI and greater than 1,500 uniquely detected cellular genes. If a cell did not meet both requirements, then the cell was considered apoptotic or a technical failure and removed. For filtering on the human genes, the following conditions needed to be met: the gene was expressed in greater than 0.5% of all cells, was a protein-coding gene, and was not related

to mitochondrial or ribosomal expression. If any of the previous conditions was not met, then that gene was excluded from the analysis (see Fig. S1 in the supplemental material). Other sophisticated filtering steps were explored, e.g., see reference 105, but upon experimentation, the above-described criteria and thresholds emerged as the most reliable and biologically coherent across this data set and experimental design. This analysis pipeline yielded data set 2.

scRNAseq virus data analysis. The virus scRNAseq analyses were subjected to additional normalization: (i) as many KSHV genes are 3' coterminal, it was useful to combine the counts for the multiple genes in the KSHV latency locus, comprised of K12, ORF71, ORF72, and ORF73; (ii) the KSHV miRNA genes are removed, since this scRNAseq method did not capture miRNAs, pre-miRNAs, or any RNAs that are not polyadenylated (19). The data were further processed by providing easily recognizable labels for the plates and by changing some of the names for the viral genes, from the nomenclature used in the reference sequence [NC_009333](https://www.ncbi.nlm.nih.gov/nuccore/NC_009333) to their more biologically intuitive names that have become the *de facto* standard in the field. The viral genes were log normalized in a manner similar to that for the human genes, except they were scaled by the size factors of the human cellular genes, since the viral gene population is low expressing. For heatmap representations, counts were scaled to yield a binary signal, as this aided visualization; however, all computations were performed on the transformed count data. After these processing steps, 70,644 data points, comprising 841 individual cells, were available for heatmap visualization. Unsupervised clustering was performed using either the k-means algorithm (1,000 iterations from 25 random starting points) or hierarchical clustering based on Euclidian distance using Ward's method as implemented in R v3.6.2. Code and documentation are available at the bitbucket repository for this work: https://bitbucket.org/dittmerlab/scrnaseq_bcb11/src/master/.

scRNAseq cellular data analysis. UMI were obtained, and cells and genes were filtered on raw UMI counts as described above for QC. The final data set contained 841 cells and 12,908 genes for a total of 10,855,628 data points. We lost 27% of our cells upon filtering and upwards of 36% of the genes. Most genes were lost due to low gene frequency, i.e., the gene was only detected in a few cells at very low levels. UMI was normalized by computing the log-normalized expression values. This method was utilized through the R package `scran`. Log-normalized counts have greater stability and decrease the variance between libraries.

Typically, scRNAseq data show more variation than bulk RNAseq data (68, 119). They tend to be zero-inflated and overdispersed. For experiments combining multiple libraries and/or biological samples batch effects are significant, and the data need to be adjusted for it; multiple methods have been proposed (69, 120–122). We used analysis of variance (ANOVA) on well-normalized, transformed UMI to arrive at a single correction coefficient per plate batch for nonzero entries. This model was as effective as other more complicated models, e.g., Seurat (123), perhaps since each batch did not contain mixtures of cells in various proportions to each other but only cells of a clonal cell line.

The batch and library size-adjusted data were subjected to hierarchical clustering with FDR adjustment using Ward-D as the method and a Euclidian distance matrix as implemented in the `SigClust` package (76).

First, the libraries from biological replicates of plate 5 were split across two different Illumina sequencing lanes each. A pairwise comparison of transformed UMI across all cells and genes generated an estimate of the variation that was introduced by the sequencing step alone. The two libraries were highly correlated with a regression coefficient of $r^2 = 0.997$ (Fig. S2A). The perfectly horizontal distribution on the MA plot (103) indicated that variation of normalized UMI counts was constant across the range of expression levels in the experiment (Fig. S2E). This attests to the high reproducibility of this technology. Second, two cells rather than one cell per well were sorted into plate 8 to explore sensitivity and linear range of the cDNA conversion step. Gene expression levels for the plate containing two cells per well correlated linearly with the plate containing only one cell per well with $r^2 = 0.961$ (Fig. S2B and F). This demonstrates that reagents were not limiting the cDNA reaction. A few genes were present at significantly higher counts (red and orange dots in Fig. S2B) in the 2 cells/well plate than in the 1 cell/well plate, suggesting that providing twice as much input RNA improved the cDNA conversion rate for weakly transcribed genes. Overall variability for these different library preparations was increased (Fig. S2E versus F). This was expected, as two different library preparations were compared, not two different sequencing runs on the same library. Third, two different PEL cell lines were compared, BCBL-1 and JSC-1 (37, 38). The cell lines differed dramatically regarding cellular gene expression (Fig. S2C and D). This comparison served as a positive control and to justify the level of the FDR. Any transcript that exhibited a level of variation of $M \geq 2$ (red dots in Fig. S2G) was considered differentially regulated between the two different cell lines. Extending this argument to different clusters of the same cell line, any transcript within a population of BCBL-1 cells was considered differentially regulated between single cells. Conversely, biological replicates for BCBL-1 cells served as a negative control (Fig. S2D and H). BCBL-1 cells were analyzed at three different dates (1 week apart). Plates 1 and 2 were FACS sorted on two separate dates, 1 week apart, and each was processed into a different library. Plates 3 to 5 were sorted on a different, third date, and each was processed into a different library. As expected, the biological variation between different calendar dates was larger than variation of different libraries made on the same calendar date, but not by much (Fig. S2D and H compared to Fig. S2A and E). Together, these measurements established biological plausibility for differential gene expression beyond statistical significance cutoffs alone. If the difference in mRNA transcription patterns between any two cells was less than that across biological replicates, it is unlikely that the two cells represent different subpopulations. If the difference in mRNA transcription patterns between two cells in the same biological replicate approached the level of difference between all BCBL-1 and all JSC-1 cells, it reflected biologically relevant differences between these two cells.

As scRNAseq is prone to catastrophic failures affecting only a single cell, the data sets expectedly contained outliers. These outliers could represent wells (i) that did not receive any cell during the initial

FACS procedure, (ii) that received a cell in the process of cell death, where mRNA was degraded, (iii) where mRNA extraction and reverse transcription failed, and (iv) where pipetting failed. These types of errors affect steps before library construction. They are expected to affect the gene expression profile globally, as opposed to specific changes in groups of genes, as would be observed for physiologically significant differences. These outliers were removed prior to analysis (see Materials and Methods). Afterwards, UMI count data were adjusted to remove batch by library size normalization and transformed to stabilize variation using the algorithm of the *scran* R package. All further analyses were conducted on adjusted data.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

MOVIE S1, MOV file, 17.1 MB.

FIG S1, TIF file, 0.7 MB.

FIG S2, TIF file, 1.6 MB.

FIG S3, TIF file, 1.2 MB.

TABLE S1, XLSX file, 0.3 MB.

ACKNOWLEDGMENT

This work was supported by public health service funding to D.P.D. (CA019014, CA239583, and DE018304).

REFERENCES

- Darst RP, Haecker I, Pardo CE, Renne R, Klade MP. 2013. Epigenetic diversity of Kaposi's sarcoma-associated herpesvirus. *Nucleic Acids Res* 41: 2993–3009. <https://doi.org/10.1093/nar/gkt033>.
- Gunther T, Grundhoff A. 2010. The epigenetic landscape of latent Kaposi sarcoma-associated herpesvirus genomes. *PLoS Pathog* 6:e1000935. <https://doi.org/10.1371/journal.ppat.1000935>.
- Hilton IB, Simon JM, Lieb JD, Davis IJ, Damania B, Dittmer DP. 2013. The open chromatin landscape of Kaposi's sarcoma-associated herpesvirus. *J Virol* 87:11831–11842. <https://doi.org/10.1128/JVI.01685-13>.
- Toth Z, Maglinte DT, Lee SH, Lee HR, Wong LY, Brulois KF, Lee S, Buckley JD, Laird PW, Marquez VE, Jung JU. 2010. Epigenetic analysis of KSHV latent and lytic genomes. *PLoS Pathog* 6:e1001013. <https://doi.org/10.1371/journal.ppat.1001013>.
- Krishnan HH, Naranatt PP, Smith MS, Zeng L, Bloomer C, Chandran B. 2004. Concurrent expression of latent and a limited number of lytic genes with immune modulation and antiapoptotic function by Kaposi's sarcoma-associated herpesvirus early during infection of primary endothelial and fibroblast cells and subsequent decline of lytic gene expression. *J Virol* 78:3601–3620. <https://doi.org/10.1128/jvi.78.7.3601-3620.2004>.
- Hopcraft SE, Pattenden SG, James LI, Frye S, Dittmer DP, Damania B. 2018. Chromatin remodeling controls Kaposi's sarcoma-associated herpesvirus reactivation from latency. *PLoS Pathog* 14:e1007267. <https://doi.org/10.1371/journal.ppat.1007267>.
- Dittmer DP, Lagunoff M, Renne R, Staskus K, Haase A, Ganem D. 1998. A cluster of latently expressed genes in Kaposi's sarcoma-associated herpesvirus. *J Virol* 72:8309–8315. <https://doi.org/10.1128/JVI.72.10.8309-8315.1998>.
- Kedes DH, Lagunoff M, Renne R, Ganem D. 1997. Identification of the gene encoding the major latency-associated nuclear antigen of the Kaposi's sarcoma-associated herpesvirus. *J Clin Invest* 100:2606–2610. <https://doi.org/10.1172/JCI119804>.
- Rainbow L, Platt GM, Simpson GR, Sarid R, Gao SJ, Stoiber H, Herrington CS, Moore PS, Schulz TF. 1997. The 222- to 234-kilodalton latent nuclear protein (LNA) of Kaposi's sarcoma-associated herpesvirus (human herpesvirus 8) is encoded by orf73 and is a component of the latency-associated nuclear antigen. *J Virol* 71:5915–5921. <https://doi.org/10.1128/JVI.71.8.5915-5921.1997>.
- Kellam P, Boshoff C, Whitby D, Matthews S, Weiss RA, Talbot SJ. 1997. Identification of a major latent nuclear antigen, LNA-1, in the human herpesvirus 8 genome. *J Hum Virol* 1:19–29.
- Grant MJ, Loftus MS, Stoja AP, Kedes DH, Smith MM. 2018. Superresolution microscopy reveals structural mechanisms driving the nanoarchitecture of a viral chromatin tether. *Proc Natl Acad Sci U S A* <https://doi.org/10.1073/pnas.1721638115>.
- Hu J, Garber AC, Renne R. 2002. The latency-associated nuclear antigen of Kaposi's sarcoma-associated herpesvirus supports latent DNA replication in dividing cells. *J Virol* 76:11677–11687. <https://doi.org/10.1128/jvi.76.22.11677-11687.2002>.
- Ballestas ME, Chatis PA, Kaye KM. 1999. Efficient persistence of extrachromosomal KSHV DNA mediated by latency-associated nuclear antigen. *Science* 284:641–644. <https://doi.org/10.1126/science.284.5414.641>.
- Barbera AJ, Chodaparambil JV, Kelley-Clarke B, Joukov V, Walter JC, Luger K, Kaye KM. 2006. The nucleosomal surface as a docking station for Kaposi's sarcoma herpesvirus LANA. *Science* 311:856–861. <https://doi.org/10.1126/science.1120541>.
- Tso FY, West JT, Wood C. 2019. Reduction of Kaposi's sarcoma-associated herpesvirus latency using CRISPR-Cas9 to edit the latency-associated nuclear antigen gene. *J Virol* 93:e02183-18. <https://doi.org/10.1128/JVI.02183-18>.
- Dupin N, Fisher C, Kellam P, Ariad S, Tulliez M, Franck N, van Marck E, Salmon D, Gorin I, Escande JP, Weiss RA, Alitalo K, Boshoff C. 1999. Distribution of human herpesvirus-8 latently infected cells in Kaposi's sarcoma, multicentric Castlemans disease, and primary effusion lymphoma. *Proc Natl Acad Sci U S A* 96:4546–4551. <https://doi.org/10.1073/pnas.96.8.4546>.
- Bigi R, Landis JT, An H, Caro-Vegas C, Raab-Traub N, Dittmer DP. 2018. Epstein-Barr virus enhances genome maintenance of Kaposi sarcoma-associated herpesvirus. *Proc Natl Acad Sci U S A* 115:E11379–E11387. <https://doi.org/10.1073/pnas.1810128115>.
- Staudt MR, Dittmer DP. 2006. Promoter switching allows simultaneous transcription of LANA and K14/vGPCR of Kaposi's sarcoma-associated herpesvirus. *Virology* 350:192–205. <https://doi.org/10.1016/j.virol.2006.03.006>.
- Cai X, Cullen BR. 2006. Transcriptional origin of Kaposi's sarcoma-associated herpesvirus microRNAs. *J Virol* 80:2234–2242. <https://doi.org/10.1128/JVI.80.5.2234-2242.2006>.
- Rose TM, Bruce AG, Barcy S, Fitzgibbon M, Matsumoto LR, Ikoma M, Casper C, Orem J, Phipps W. 2018. Quantitative RNAseq analysis of Ugandan KS tumors reveals KSHV gene expression dominated by transcription from the LTd downstream latency promoter. *PLoS Pathog* 14:e1007441. <https://doi.org/10.1371/journal.ppat.1007441>.
- Li H, Komatsu T, Dezube BJ, Kaye KM. 2002. The Kaposi's sarcoma-associated herpesvirus K12 transcript from a primary effusion lymphoma contains complex repeat elements, is spliced, and initiates from a novel promoter. *J Virol* 76:11880–11888. <https://doi.org/10.1128/jvi.76.23.11880-11888.2002>.
- Staskus KA, Zhong W, Gebhard K, Herndier B, Wang H, Renne R, Benke J, Pudney J, Anderson DJ, Ganem D, Haase AT. 1997. Kaposi's sarcoma-associated herpesvirus gene expression in endothelial (spindle) tumor cells. *J Virol* 71:715–719. <https://doi.org/10.1128/JVI.71.1.715-719.1997>.
- Forte E, Raja AN, Shamulailatpam P, Manzano M, Schipma MJ, Casey JL, Gottwein E. 2015. MicroRNA-mediated transformation by the Kaposi's sarcoma-associated herpesvirus Kaposin locus. *J Virol* 89:2333–2341. <https://doi.org/10.1128/JVI.03317-14>.

24. Jenner RG, Alba MM, Boshoff C, Kellam P. 2001. Kaposi's sarcoma-associated herpesvirus latent and lytic gene expression as revealed by DNA arrays. *J Virol* 75:891–902. <https://doi.org/10.1128/JVI.75.2.891-902.2001>.
25. Sarid R, Flore O, Bohenzky RA, Chang Y, Moore PS. 1998. Transcription mapping of the Kaposi's sarcoma-associated herpesvirus (human herpesvirus 8) genome in a body cavity-based lymphoma cell line (BC-1). *J Virol* 72:1005–1012. <https://doi.org/10.1128/JVI.72.2.1005-1012.1998>.
26. Zhong W, Wang H, Herndier B, Ganem D. 1996. Restricted expression of Kaposi sarcoma-associated herpesvirus (human herpesvirus 8) genes in Kaposi sarcoma. *Proc Natl Acad Sci U S A* 93:6641–6646. <https://doi.org/10.1073/pnas.93.13.6641>.
27. Chandriani S, Xu Y, Ganem D. 2010. The lytic transcriptome of Kaposi's sarcoma-associated herpesvirus reveals extensive transcription of non-coding regions, including regions antisense to important genes. *J Virol* 84:7934–7942. <https://doi.org/10.1128/JVI.00645-10>.
28. Chandriani S, Ganem D. 2010. Array-based transcript profiling and limiting-dilution reverse transcription-PCR analysis identify additional latent genes in Kaposi's sarcoma-associated herpesvirus. *J Virol* 84:5565–5573. <https://doi.org/10.1128/JVI.02723-09>.
29. Chandriani S, Ganem D. 2007. Host transcript accumulation during lytic KSHV infection reveals several classes of host responses. *PLoS One* 2: e811. <https://doi.org/10.1371/journal.pone.0000811>.
30. Zhu FX, Cusano T, Yuan Y. 1999. Identification of the immediate-early transcripts of Kaposi's sarcoma-associated herpesvirus. *J Virol* 73:5556–5567. <https://doi.org/10.1128/JVI.73.7.5556-5567.1999>.
31. Fakhari FD, Dittmer DP. 2002. Charting latency transcripts in Kaposi's sarcoma-associated herpesvirus by whole-genome real-time quantitative PCR. *J Virol* 76:6213–6223. <https://doi.org/10.1128/jvi.76.12.6213-6223.2002>.
32. Dittmer DP. 2003. Transcription profile of Kaposi's sarcoma-associated herpesvirus in primary Kaposi's sarcoma lesions as determined by real-time PCR arrays. *Cancer Res* 63:2010–2015.
33. Dittmer DP. 2011. Restricted Kaposi's sarcoma (KS) herpesvirus transcription in KS lesions from patients on successful antiretroviral therapy. *mBio* 2:e00138-11. <https://doi.org/10.1128/mBio.00138-11>.
34. Hosseiniour MC, Sweet KM, Xiong J, Namarika D, Mwafongo A, Nyirenda M, Chiwoko L, Kamwendo D, Hoffman I, Lee J, Phiri S, Vahrson W, Damania B, Dittmer DP. 2014. Viral profiling identifies multiple subtypes of Kaposi's sarcoma. *mBio* 5:e01633-14. <https://doi.org/10.1128/mBio.01633-14>.
35. Arias C, Weisburd B, Stern-Ginossar N, Mercier A, Madrid AS, Bellare P, Holdorf M, Weissman JS, Ganem D. 2014. KSHV 2.0: a comprehensive annotation of the Kaposi's sarcoma-associated herpesvirus genome using next-generation sequencing reveals novel genomic and functional features. *PLoS Pathog* 10:e1003847. <https://doi.org/10.1371/journal.ppat.1003847>.
36. Dresang LR, Teuton JR, Feng H, Jacobs JM, Camp DG, II, Purvine SO, Gritsenko MA, Li Z, Smith RD, Sugden B, Moore PS, Chang Y. 2011. Coupled transcriptome and proteome analysis of human lymphotropic tumor viruses: insights on the detection and discovery of viral genes. *BMC Genomics* 12:625. <https://doi.org/10.1186/1471-2164-12-625>.
37. Renne R, Zhong W, Herndier B, McGrath M, Abbey N, Kedes D, Ganem D. 1996. Lytic growth of Kaposi's sarcoma-associated herpesvirus (human herpesvirus 8) in culture. *Nat Med* 2:342–346. <https://doi.org/10.1038/nm0396-342>.
38. Cannon JS, Ciufu D, Hawkins AL, Griffin CA, Borowitz MJ, Hayward GS, Ambinder RF. 2000. A new primary effusion lymphoma-derived cell line yields a highly infectious Kaposi's sarcoma herpesvirus-containing supernatant. *J Virol* 74:10187–10193. <https://doi.org/10.1128/jvi.74.21.10187-10193.2000>.
39. Brulois KF, Chang H, Lee AS, Ensser A, Wong LY, Toth Z, Lee SH, Lee HR, Myoung J, Ganem D, Oh TK, Kim JF, Gao SJ, Jung JU. 2012. Construction and manipulation of a new Kaposi's sarcoma-associated herpesvirus bacterial artificial chromosome clone. *J Virol* 86:9708–9720. <https://doi.org/10.1128/JVI.01019-12>.
40. Cesarman E, Moore PS, Rao PH, Inghirami G, Knowles DM, Chang Y. 1995. In vitro establishment and characterization of two acquired immunodeficiency syndrome-related lymphoma cell lines (BC-1 and BC-2) containing Kaposi's sarcoma-associated herpesvirus-like (KSHV) DNA sequences. *Blood* 86:2708–2714. <https://doi.org/10.1182/blood.V86.7.2708.2708>.
41. Boshoff C, Gao SJ, Healy LE, Matthews S, Thomas AJ, Coignet L, Warnke RA, Strauchen JA, Matutes E, Kamel OW, Moore PS, Weiss RA, Chang Y. 1998. Establishing a KSHV+ cell line (BCP-1) from peripheral blood and characterizing its growth in Nod/SCID mice. *Blood* 91:1671–1679. <https://doi.org/10.1182/blood.V91.5.1671>.
42. Bhatt S, Ashlock BM, Toomey NL, Diaz LA, Mesri EA, Lossos IS, Ramos JC. 2013. Efficacious proteasome/HDAC inhibitor combination therapy for primary effusion lymphoma. *J Clin Invest* 123:2616–2628. <https://doi.org/10.1172/JCI64503>.
43. Staudt MR, Kanan Y, Jeong JH, Papin JF, Hines-Boykin R, Dittmer DP. 2004. The tumor microenvironment controls primary effusion lymphoma growth in vivo. *Cancer Res* 64:4790–4799. <https://doi.org/10.1158/0008-5472.CAN-03-3835>.
44. Roy D, Sin SH, Damania B, Dittmer DP. 2011. Tumor suppressor genes FHIT and WWOX are deleted in primary effusion lymphoma (PEL) cell lines. *Blood* 118:e32–e39. <https://doi.org/10.1182/blood-2010-12-323659>.
45. Luan SL, Boulanger E, Ye H, Chanudet E, Johnson N, Hamoudi RA, Bacon CM, Liu H, Huang Y, Said J, Chu P, Clemen CS, Cesarman E, Chadburn A, Isaacson PG, Du MQ. 2010. Primary effusion lymphoma: genomic profiling revealed amplification of SELPLG and CORO1C encoding for proteins important for cell migration. *J Pathol* 222:166–179. <https://doi.org/10.1002/path.2752>.
46. Godfrey A, Anderson J, Papanastasiou A, Takeuchi Y, Boshoff C. 2005. Inhibiting primary effusion lymphoma by lentiviral vectors encoding short hairpin RNA. *Blood* 105:2510–2518. <https://doi.org/10.1182/blood-2004-08-3052>.
47. Kedes DH, Ganem D. 1997. Sensitivity of Kaposi's sarcoma-associated herpesvirus replication to antiviral drugs. Implications for potential therapy. *J Clin Invest* 99:2082–2086. <https://doi.org/10.1172/JCI119380>.
48. Komanduri KV, Luce JA, McGrath MS, Herndier BG, Ng VL. 1996. The natural history and molecular heterogeneity of HIV-associated primary malignant lymphomatous effusions. *J Acquir Immune Defic Syndr Hum Retroviral* 13: 215–226. <https://doi.org/10.1097/00042560-199611010-00003>.
49. Picchio GR, Sabbe RE, Gulizia RJ, McGrath M, Herndier BG, Mosier DE. 1997. The KSHV/HHV8-infected BCBL-1 lymphoma line causes tumors in SCID mice but fails to transmit virus to a human peripheral blood mononuclear cell graft. *Virology* 238:22–29. <https://doi.org/10.1006/viro.1997.8822>.
50. Dittmer D, Stoddart C, Renne R, Linquist-Stepps V, Moreno ME, Bare C, McCune JM, Ganem D. 1999. Experimental transmission of Kaposi's sarcoma-associated herpesvirus (KSHV/HHV-8) to SCID-hu Thy/Liv mice. *J Exp Med* 190:1857–1868. <https://doi.org/10.1084/jem.190.12.1857>.
51. Sun R, Lin SF, Staskus K, Gradoville L, Grogan E, Haase A, Miller G. 1999. Kinetics of Kaposi's sarcoma-associated herpesvirus gene expression. *J Virol* 73:2232–2242. <https://doi.org/10.1128/JVI.73.3.2232-2242.1999>.
52. Wong JP, Stuhlmiller TJ, Giffin LC, Lin C, Bigi R, Zhao J, Zhang W, Bravo Cruz AG, Park SI, Earp HS, Dittmer DP, Frye SV, Wang X, Johnson GL, Damania B. 2019. Kinome profiling of non-Hodgkin lymphoma identifies Tyro3 as a therapeutic target in primary effusion lymphoma. *Proc Natl Acad Sci U S A* 116:16541–16550. <https://doi.org/10.1073/pnas.1903991116>.
53. Bhatt AP, Jacobs SR, Freermerman AJ, Makowski L, Rathmell JC, Dittmer DP, Damania B. 2012. Dysregulation of fatty acid synthesis and glycolysis in non-Hodgkin lymphoma. *Proc Natl Acad Sci U S A* 109:11818–11823. <https://doi.org/10.1073/pnas.1205995109>.
54. Shin HJ, DeCotiis J, Giron M, Palmeri D, Lukac DM. 2014. Histone deacetylase classes I and II regulate Kaposi's sarcoma-associated herpesvirus reactivation. *J Virol* 88:1281–1292. <https://doi.org/10.1128/JVI.02665-13>.
55. Shaw RN, Arbisser JL, Offermann MK. 2000. Valproic acid induces human herpesvirus 8 lytic gene expression in BCBL-1 cells. *AIDS* 14:899–902. <https://doi.org/10.1097/00002030-200005050500021>.
56. Nakamura H, Lu M, Gwack Y, Souvlis J, Zeichner SL, Jung JU. 2003. Global changes in Kaposi's sarcoma-associated virus gene expression patterns following expression of a tetracycline-inducible Rta transactivator. *J Virol* 77:4205–4220. <https://doi.org/10.1128/jvi.77.7.4205-4220.2003>.
57. Liang Y, Chang J, Lynch SJ, Lukac DM, Ganem D. 2002. The lytic switch protein of KSHV activates gene expression via functional interaction with RBP-Jkappa (CSL), the target of the Notch signaling pathway. *Genes Dev* 16:1977–1989. <https://doi.org/10.1101/gad.996502>.
58. Rato S, Rausell A, Munoz M, Telenti A, Ciuffi A. 2017. Single-cell analysis identifies cellular markers of the HIV permissive cell. *PLoS Pathog* 13:e1006678. <https://doi.org/10.1371/journal.ppat.1006678>.
59. Bradley T, Ferrari G, Haynes BF, Margolis DM, Browne EP. 2018. Single-cell analysis of quiescent HIV infection reveals host transcriptional profiles that regulate proviral latency. *Cell Rep* 25:107–117. <https://doi.org/10.1016/j.celrep.2018.09.020>.
60. Russell AB, Trapnell C, Bloom JD. 2018. Extreme heterogeneity of influenza virus infection in single cells. *Elife* 7:e032303. <https://doi.org/10.7554/eLife.32303>.
61. Wang C, Forst CV, Chou TW, Geber A, Wang M, Hamou W, Smith M, Sebra R, Zhang B, Zhou B, Ghedin E. 2020. Cell-to-cell variation in

- defective virus expression and effects on host responses during influenza virus infection. *mBio* 11:e02880-19. <https://doi.org/10.1128/mBio.02880-19>.
62. Zanini F, Pu SY, Bekerman E, Einav S, Quake SR. 2018. Single-cell transcriptional dynamics of flavivirus infection. *Elife* 7:e32942. <https://doi.org/10.7554/eLife.32942>.
 63. Goodrum F, McWeeney S. 2018. A single-cell approach to the elusive latent human cytomegalovirus transcriptome. *mBio* 9:e01001-18. <https://doi.org/10.1128/mBio.01001-18>.
 64. Shnayder M, Nachshon A, Rozman B, Bernshtein B, Lavi M, Fein N, Poole E, Avdic S, Blythe E, Gottlieb D, Abendroth A, Slobedman B, Sinclair J, Stern-Ginossar N, Schwartz M. 2020. Single cell analysis reveals human cytomegalovirus drives latently infected cells towards an anergic-like monocyte state. *Elife* 9:e52168. <https://doi.org/10.7554/eLife.52168>.
 65. Shnayder M, Nachshon A, Krishna B, Poole E, Boshkov A, Binyamin A, Maza I, Sinclair J, Schwartz M, Stern-Ginossar N. 2018. Defining the transcriptional landscape during cytomegalovirus latency with single-cell RNA sequencing. *mBio* 9:e00013-18. <https://doi.org/10.1128/mBio.00013-18>.
 66. Parravicini C, Chandran B, Corbellino M, Berti E, Paulli M, Moore PS, Chang Y. 2000. Differential viral protein expression in Kaposi's sarcoma-associated herpesvirus-infected diseases: Kaposi's sarcoma, primary effusion lymphoma, and multicentric Castlemans disease. *Am J Pathol* 156:743–749. [https://doi.org/10.1016/S0002-9440\(10\)64940-1](https://doi.org/10.1016/S0002-9440(10)64940-1).
 67. Du MQ, Liu H, Diss TC, Ye H, Hamoudi RA, Dupin N, Meignin V, Oksenhendler E, Boshoff C, Isaacson PG. 2001. Kaposi sarcoma-associated herpesvirus infects monotypic (IgM lambda) but polyclonal naive B cells in Castlemans disease and associated lymphoproliferative disorders. *Blood* 97:2130–2136. <https://doi.org/10.1182/blood.v97.7.2130>.
 68. Hicks SC, Townes FW, Teng M, Irizarry RA. 2018. Missing data and technical variability in single-cell RNA-sequencing experiments. *Biostatistics* 19:562–578. <https://doi.org/10.1093/biostatistics/kxx053>.
 69. Tung PY, Blischak JD, Hsiao CJ, Knowles DA, Burnett JE, Pritchard JK, Gilad Y. 2017. Batch effects and the effective design of single-cell gene expression studies. *Sci Rep* 7:39921. <https://doi.org/10.1038/srep39921>.
 70. Chadburn A, Hyjek E, Mathew S, Cesarman E, Said J, Knowles DM. 2004. KSHV-positive solid lymphomas represent an extra-cavitary variant of primary effusion lymphoma. *Am J Surg Pathol* 28:1401–1416. <https://doi.org/10.1097/01.pas.0000138177.10829.5c>.
 71. Katano H, Suda T, Morishita Y, Yamamoto K, Hoshino Y, Nakamura K, Tachikawa N, Sata T, Hamaguchi H, Iwamoto A, Mori S. 2000. Human herpesvirus 8-associated solid lymphomas that occur in AIDS patients take anaplastic large cell morphology. *Mod Pathol* 13:77–85. <https://doi.org/10.1038/modpathol.3880012>.
 72. Chiu YF, Sugden AU, Fox K, Hayes M, Sugden B. 2017. Kaposi's sarcoma-associated herpesvirus stably clusters its genomes across generations to maintain itself extrachromosomally. *J Cell Biol* 216:2745–2758. <https://doi.org/10.1083/jcb.201702013>.
 73. Moore PS, Boshoff C, Weiss RA, Chang Y. 1996. Molecular mimicry of human cytokine and cytokine response pathway genes by KSHV. *Science* 274:1739–1744. <https://doi.org/10.1126/science.274.5293.1739>.
 74. Cannon JS, Nicholas J, Orenstein JM, Mann RB, Murray PG, Browning PJ, DiGiuseppe JA, Cesarman E, Hayward GS, Ambinder RF. 1999. Heterogeneity of viral IL-6 expression in HHV-8-associated diseases. *J Infect Dis* 180:824–828. <https://doi.org/10.1086/314956>.
 75. Sanchez-Martin D, Uldrick TS, Kwak H, Ohnuki H, Polizzotto MN, Annunziata CM, Raffeld M, Wyvill KM, Aleman K, Wang V, Marshall VA, Whitby D, Yarchoan R, Tosato G. 2017. Evidence for a mesothelial origin of body cavity effusion lymphomas. *J Natl Cancer Inst* 109:djx016. <https://doi.org/10.1093/jnci/djx016>.
 76. Huang H, Liu Y, Yuan M, Marron JS. 2015. Statistical significance of clustering using soft thresholding. *J Comput Graph Stat* 24:975–993. <https://doi.org/10.1080/10618600.2014.948179>.
 77. Gregory SM, West JA, Dillon PJ, Hilscher C, Dittmer DP, Damania B. 2009. Toll-like receptor signaling controls reactivation of KSHV from latency. *Proc Natl Acad Sci U S A* 106:11725–11730. <https://doi.org/10.1073/pnas.0905316106>.
 78. Hassman LM, Ellison TJ, Kedes DH. 2011. KSHV infects a subset of human tonsillar B cells, driving proliferation and plasmablast differentiation. *J Clin Invest* 121:752–768. <https://doi.org/10.1172/JCI44185>.
 79. Decker LL, Shankar P, Khan G, Freeman RB, Dezube BJ, Lieberman J, Thorley-Lawson DA. 1996. The Kaposi sarcoma-associated herpesvirus (KSHV) is present as an intact latent genome in KS tissue but replicates in the peripheral blood mononuclear cells of KS patients. *J Exp Med* 184:283–288. <https://doi.org/10.1084/jem.184.1.283>.
 80. McHugh D, Caduff N, Barros MHM, Ramer PC, Raykova A, Murer A, Landtwing V, Quast I, Styles CT, Spohn M, Fowotade A, Delecluse HJ, Papoudou-Bai A, Lee YM, Kim JM, Middeldorp J, Schulz TF, Cesarman E, Zbinden A, Capaul R, White RE, Allday MJ, Niedobitek G, Blackbourn DJ, Grundhoff A, Munz C. 2017. Persistent KSHV infection increases EBV-associated tumor formation in vivo via enhanced EBV lytic gene expression. *Cell Host Microbe* 22:61–73. <https://doi.org/10.1016/j.chom.2017.06.009>.
 81. Faure A, Hayes M, Sugden B. 2019. How Kaposi's sarcoma-associated herpesvirus stably transforms peripheral B cells towards lymphomagenesis. *Proc Natl Acad Sci U S A* 116:16519–16528. <https://doi.org/10.1073/pnas.1905025116>.
 82. Aalam F, Nabiee R, Castano JR, Totonchy J. 2020. Analysis of KSHV B lymphocyte lineage tropism in human tonsil reveals efficient infection of CD138+ plasma cells. *PLoS Pathog* 16:e1008968. <https://doi.org/10.1371/journal.ppat.1008968>.
 83. Sin SH, Eason AB, Bigi R, Kim Y, Kang S, Tan K, Seltzer TA, Venkataraman R, An H, Dittmer DP. 2018. Kaposi's sarcoma-associated herpesvirus latency locus renders B cells hyperresponsive to secondary infections. *J Virol* 92:e01138-18. <https://doi.org/10.1128/JVI.01138-18>.
 84. Boss IW, Nadeau PE, Abbott JR, Yang Y, Mergia A, Renne R. 2011. A Kaposi's sarcoma-associated herpesvirus-encoded ortholog of microRNA miR-155 induces human splenic B-cell expansion in NOD/LtSz-scid IL2Rgammanull mice. *J Virol* 85:9877–9886. <https://doi.org/10.1128/JVI.05558-11>.
 85. Totonchy J, Osborn JM, Chadburn A, Nabiee R, Argueta L, Mikita G, Cesarman E. 2018. KSHV induces immunoglobulin rearrangements in mature B lymphocytes. *PLoS Pathog* 14:e1006967. <https://doi.org/10.1371/journal.ppat.1006967>.
 86. Carbone A, De Paoli P, Ghoghini A, Vaccher E. 2015. KSHV-associated multicentric Castlemans disease: a tangle of different entities requiring multi-target treatment strategies. *Int J Cancer* 137:251–261. <https://doi.org/10.1002/ijc.28923>.
 87. Cesarman E, Chang Y, Moore PS, Said JW, Knowles DM. 1995. Kaposi's sarcoma-associated herpesvirus-like DNA sequences in AIDS-related body-cavity-based lymphomas. *N Engl J Med* 332:1186–1191. <https://doi.org/10.1056/NEJM199505043321802>.
 88. Gaidano G, Ghoghini A, Gattei V, Rossi MF, Cilia AM, Godeas C, Degan M, Perin T, Canzonieri V, Aldinucci D, Saglio G, Carbone A, Pinto A. 1997. Association of Kaposi's sarcoma-associated herpesvirus-positive primary effusion lymphoma with expression of the CD138/syndecan-1 antigen. *Blood* 90:4894–4900.
 89. Uldrick TS, Wang V, O'Mahony D, Aleman K, Wyvill KM, Marshall V, Steinberg SM, Pittaluga S, Maric I, Whitby D, Tosato G, Little RF, Yarchoan R. 2010. An interleukin-6-related systemic inflammatory syndrome in patients co-infected with Kaposi sarcoma-associated herpesvirus and HIV but without multicentric Castlemans disease. *Clin Infect Dis* 51:350–358. <https://doi.org/10.1086/654798>.
 90. Fan W, Bubman D, Chadburn A, Harrington WJ, Jr, Cesarman E, Knowles DM. 2005. Distinct subsets of primary effusion lymphoma can be identified based on their cellular gene expression profile and viral association. *J Virol* 79:1244–1251. <https://doi.org/10.1128/JVI.79.2.1244-1251.2005>.
 91. Yanagisawa Y, Sato Y, Asahi-Ozaki Y, Ito E, Honma R, Imai J, Kanno T, Kano M, Akiyama H, Sata T, Shinkai-Ouchi F, Yamakawa Y, Watanabe S, Katano H. 2006. Effusion and solid lymphomas have distinctive gene and protein expression profiles in an animal model of primary effusion lymphoma. *J Pathol* 209:464–473. <https://doi.org/10.1002/path.2012>.
 92. Wies E, Mori Y, Hahn A, Kremmer E, Sturzl M, Fleckenstein B, Neipel F. 2008. The viral interferon-regulatory factor-3 is required for the survival of KSHV-infected primary effusion lymphoma cells. *Blood* 111:320–327. <https://doi.org/10.1182/blood-2007-05-092288>.
 93. Rivas C, Thlick AE, Parravicini C, Moore PS, Chang Y. 2001. Kaposi's sarcoma-associated herpesvirus LANA2 is a B-cell-specific latent viral protein that inhibits p53. *J Virol* 75:429–438. <https://doi.org/10.1128/JVI.75.1.429-438.2001>.
 94. Pozharskaya VP, Weakland LL, Zimring JC, Krug LT, Unger ER, Neisch A, Joshi H, Inoue N, Offermann MK. 2004. Short duration of elevated vRF-1 expression during lytic replication of human herpesvirus 8 limits its ability to block antiviral responses induced by alpha interferon in BCBL-1 cells. *J Virol* 78:6621–6635. <https://doi.org/10.1128/JVI.78.12.6621-6635.2004>.
 95. Sztuba-Solinska J, Rausch JW, Smith R, Miller JT, Whitby D, Le Grice SFJ. 2017. Kaposi's sarcoma-associated herpesvirus polyadenylated nuclear RNA: a structural scaffold for nuclear, cytoplasmic and viral proteins. *Nucleic Acids Res* 45:6805–6821. <https://doi.org/10.1093/nar/gkx241>.

96. Sun R, Lin SF, Gradoville L, Miller G. 1996. Polyadenylated nuclear RNA encoded by Kaposi sarcoma-associated herpesvirus. *Proc Natl Acad Sci U S A* 93:11883–11888. <https://doi.org/10.1073/pnas.93.21.11883>.
97. Rossetto CC, Pari GS. 2011. Kaposi's sarcoma-associated herpesvirus non-coding polyadenylated nuclear RNA interacts with virus- and host cell-encoded proteins and suppresses expression of genes involved in immune modulation. *J Virol* 85:13290–13297. <https://doi.org/10.1128/JVI.05886-11>.
98. Gasperini P, Sakakibara S, Tosato G. 2008. Contribution of viral and cellular cytokines to Kaposi's sarcoma-associated herpesvirus pathogenesis. *J Leukoc Biol* 84:994–1000. <https://doi.org/10.1189/jlb.1107777>.
99. Polizzotto MN, Uldrick TS, Wang V, Aleman K, Wyvill KM, Marshall V, Pittaluga S, O'Mahony D, Whitby D, Tosato G, Steinberg SM, Little RF, Yarchoan R. 2013. Human and viral interleukin-6 and other cytokines in Kaposi sarcoma herpesvirus-associated multicentric Castlemann disease. *Blood* 122:4189–4198. <https://doi.org/10.1182/blood-2013-08-519959>.
100. Foster CB, Lehrnbecher T, Samuels S, Stein S, Mol F, Metcalf JA, Wyvill K, Steinberg SM, Kovacs J, Blauvelt A, Yarchoan R, Chanock SJ. 2000. An IL6 promoter polymorphism is associated with a lifetime risk of development of Kaposi sarcoma in men infected with human immunodeficiency virus. *Blood* 96:2562–2567. <https://doi.org/10.1182/blood.V96.7.2562>.
101. Damania B, Jeong JH, Bowser BS, DeWire SM, Staudt MR, Dittmer DP. 2004. Comparison of the Rta/Orf50 transactivator proteins of gamma-2-herpesviruses. *J Virol* 78:5491–5499. <https://doi.org/10.1128/jvi.78.10.5491-5499.2004>.
102. Bruce AG, Barcy S, DiMaio T, Gan E, Garrigues HJ, Lagunoff M, Rose TM. 2017. Quantitative analysis of the KSHV transcriptome following primary infection of blood and lymphatic endothelial cells. *Pathogens* 6:11. <https://doi.org/10.3390/pathogens6010011>.
103. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. 2015. limma powers differential expression analyses for RNA-seq and microarray studies. *Nucleic Acids Res* 43:e47. <https://doi.org/10.1093/nar/gkv007>.
104. Satija R, Farrell JA, Gennert D, Schier AF, Regev A. 2015. Spatial reconstruction of single-cell gene expression data. *Nat Biotechnol* 33:495–502. <https://doi.org/10.1038/nbt.3192>.
105. Islam S, Zeisel A, Joost S, La Manno G, Zajac P, Kasper M, Lonnerberg P, Linnarsson S. 2014. Quantitative single-cell RNA-seq with unique molecular identifiers. *Nat Methods* 11:163–166. <https://doi.org/10.1038/nmeth.2772>.
106. Ordonez-Rueda D, Baying B, Pavlinic D, Alessandri L, Yeboah Y, Landry JJM, Calogero R, Benes V, Paulsen M. 2020. Apoptotic cell exclusion and bias-free single-cell selection are important quality control requirements for successful single-cell sequencing applications. *Cytometry A* 97:156–167. <https://doi.org/10.1002/cyto.a.23898>.
107. Bruggner RV, Bodenmiller B, Dill DL, Tibshirani RJ, Nolan GP. 2014. Automated identification of stratifying signatures in cellular subpopulations. *Proc Natl Acad Sci U S A* 111:E2770–7. <https://doi.org/10.1073/pnas.1408792111>.
108. Glaunsinger B, Ganem D. 2004. Lytic KSHV infection inhibits host gene expression by accelerating global mRNA turnover. *Mol Cell* 13:713–723. [https://doi.org/10.1016/s1097-2765\(04\)00091-7](https://doi.org/10.1016/s1097-2765(04)00091-7).
109. Abernathy E, Gilbertson S, Alla R, Glaunsinger B. 2015. Viral nucleases induce an mRNA degradation-transcription feedback loop in mammalian cells. *Cell Host Microbe* 18:243–253. <https://doi.org/10.1016/j.chom.2015.06.019>.
110. Davis ZH, Hesser CR, Park J, Glaunsinger BA. 2016. Interaction between ORF24 and ORF34 in the Kaposi's sarcoma-associated herpesvirus late gene transcription factor complex is essential for viral late gene expression. *J Virol* 90:599–604. <https://doi.org/10.1128/JVI.02157-15>.
111. Parsons CH, Szomju B, Kedes DH. 2004. Susceptibility of human fetal mesenchymal stem cells to Kaposi sarcoma-associated herpesvirus. *Blood* 104:2736–2738. <https://doi.org/10.1182/blood-2004-02-0693>.
112. Gasperini P, Espigol-Frigole G, McCormick PJ, Salvucci O, Maric D, Uldrick TS, Polizzotto MN, Yarchoan R, Tosato G. 2012. Kaposi sarcoma herpesvirus promotes endothelial-to-mesenchymal transition through Notch-dependent signaling. *Cancer Res* 72:1157–1169. <https://doi.org/10.1158/0008-5472.CAN-11-3067>.
113. Jones T, Ye F, Bedolla R, Huang Y, Meng J, Qian L, Pan H, Zhou F, Moody R, Wagner B, Arar M, Gao SJ. 2012. Direct and efficient cellular transformation of primary rat mesenchymal precursor cells by KSHV. *J Clin Invest* 122:1076–1081. <https://doi.org/10.1172/JCI58530>.
114. Lee MS, Yuan H, Jeon H, Zhu Y, Yoo S, Shi S, Krueger B, Renne R, Lu C, Jung JU, Gao SJ. 2016. Human mesenchymal stem cells of diverse origins support persistent infection with Kaposi's sarcoma-associated herpesvirus and manifest distinct angiogenic, invasive, and transforming phenotypes. *mBio* 7:e02109-15. <https://doi.org/10.1128/mBio.02109-15>.
115. Li Y, Zhong C, Liu D, Yu W, Chen W, Wang Y, Shi S, Yuan Y. 2018. Evidence for Kaposi sarcoma originating from mesenchymal stem cell through KSHV-induced mesenchymal-to-endothelial transition. *Cancer Res* 78:230–245. <https://doi.org/10.1158/0008-5472.CAN-17-1961>.
116. Liu R, Li X, Tulpule A, Zhou Y, Scehnet JS, Zhang S, Lee JS, Chaudhary PM, Jung J, Gill PS. 2010. KSHV-induced notch components render endothelial and mural cell characteristics and cell survival. *Blood* 115:887–895. <https://doi.org/10.1182/blood-2009-08-236745>.
117. Naipauer J, Rosario S, Gupta S, Premer C, Mendez-Solis O, Schlesinger M, Ponziabbio V, Jain V, Gay L, Renne R, Chan HL, Morey L, Salyakina D, Abba M, Williams S, Hare JM, Goldschmidt-Clermont PJ, Mesri EA. 2019. PDGFRA defines the mesenchymal stem cell Kaposi's sarcoma progenitors by enabling KSHV oncogenesis in an angiogenic environment. *PLoS Pathog* 15:e1008221. <https://doi.org/10.1371/journal.ppat.1008221>.
118. Jabara CB, Jones CD, Roach J, Anderson JA, Swanstrom R. 2011. Accurate sampling and deep sequencing of the HIV-1 protease gene using a primer ID. *Proc Natl Acad Sci U S A* 108:20166–20171. <https://doi.org/10.1073/pnas.1110064108>.
119. Svensson V, Natarajan KN, Ly LH, Miragaia RJ, Labalette C, Macaulay IC, Cvejic A, Teichmann SA. 2017. Power analysis of single-cell RNA-sequencing experiments. *Nat Methods* 14:381–387. <https://doi.org/10.1038/nmeth.4220>.
120. Haghverdi L, Lun ATL, Morgan MD, Marioni JC. 2018. Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nat Biotechnol* 36:421–427. <https://doi.org/10.1038/nbt.4091>.
121. Lin Y, Ghazanfar S, Wang KYX, Gagnon-Bartsch JA, Lo KK, Su X, Han ZG, Ormerod JT, Speed TP, Yang P, Yang JYH. 2019. scMerge leverages factor analysis, stable expression, and pseudoreplication to merge multiple single-cell RNA-seq datasets. *Proc Natl Acad Sci U S A* 116:9775–9784. <https://doi.org/10.1073/pnas.1820006116>.
122. Zhu X, Wolfgruber TK, Tasato A, Arisdakessian C, Garmire DG, Garmire LX. 2017. Granatum: a graphical single-cell RNA-seq analysis pipeline for genomics scientists. *Genome Med* 9:108. <https://doi.org/10.1186/s13073-017-0492-3>.
123. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. 2018. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* 36:411–420. <https://doi.org/10.1038/nbt.4096>.