Widespread Traces of Lytic Kaposi Sarcoma-Associated Herpesvirus in Primary Effusion Lymphoma at Single-Cell Resolution

Nicole C. Rondeau,[a] Michael O. Finlayson,[b] JJ L. Miranda[a]

[a]Department of Biology, Barnard College, Columbia University, New York, New York, USA
[b]JP Sulzberger Columbia Genome Center, Columbia University Irving Medical Center, New York, New York, USA

ABSTRACT Cancer cells of primary effusion lymphoma (PEL) often contain both Kaposi sarcoma-associated herpesvirus (KSHV) and Epstein-Barr virus (EBV). We measured the interplay of human, KSHV, and EBV transcription in a cell culture model of PEL using single-cell RNA sequencing. The data detect widespread trace expression of lytic KSHV genes.

Immortalized cell lines derived from primary effusion lymphoma (PEL) positive for both Kaposi sarcoma-associated herpesvirus (KSHV) and Epstein-Barr virus (EBV) serve as tractable systems for studying transcriptional host-pathogen interactions. We grew BC-1 cells (1) in culture (2) and processed log-phase samples for single-cell RNA deep sequencing (scRNA-seq) with Gel beads in EMulsion microfluidic (GEM) technology (3) using a Chromium Single Cell 3′ Library & Gel Bead Kit v2 (10X Genomics, Pleasanton, CA). The experiment was repeated twice with two independent biological replicates that each yielded ~400 to 700 million paired-end reads of 28 to 101 nucleotides.

We built a bioinformatics pipeline to simultaneously measure human and viral transcription. Reads were demultiplexed and aligned using Cell Ranger (10X Genomics). A custom genome reference was created by appending the viral genomes of EBV (4) (GenBank accession number NC_007605.1) and KSHV (5) (GenBank accession number U75698.1) to the human genome GRCh38. Reads aligning to viral genomes were assigned to features using a heuristic algorithm implemented in Python (https://github.com/jjmirandalab/scrnaseq). Feature definitions were obtained from a published EBV annotation (6) (https://github.com/flemingtonlab/public/blob/master/annotation/chrEBV_B95_8_Raji.ann) and downloaded from the KSHV reference (5) (GenBank accession number U75698.1). First, each virus-mapping read was assigned the union of all RNA genome features its alignment overlapped with. Second, each molecule was assigned the intersection of all feature assignments from reads with its corresponding unique molecular identifier (UMI). If any molecule mapped to multiple genomes, it was excluded. Finally, for each unique set of features assigned to a molecule, counts were recorded for each cell. Cells were filtered using R scripts (https://github.com/jjmirandalab/scrnaseq). Distributions of cell quality metrics were modeled as log-normal for the total UMI count, log-normal for the proportion of expression from mitochondrial genes, and normal for the number of genes detected. Cells with extreme metric values were then identified according to the fit distributions. Data between samples were normalized with sctransform (7). "Sample" was included as the "batch_var" parameter.

Earlier RNA-seq methods only measured expression of exon regions that did not overlap another transcript (8, 9); our approach discards less data. Approximately 20% of the feature sets we count correspond to potentially overlapping transcripts that

TABLE 1 Detection of KSHV and EBV transcripts in BC-1 PEL cells with scRNA-seq

	Positive cells (% of total)[a]		Cells with 1 count (% of positive)	
Transcript	Replicate 1	Replicate 2	Replicate 1	Replicate 2
EBV latent				
A73, BARF0, RPMS1[b]	29.8	30.4	69.7	71.7
EBER-2	4.4	6.6	95.3	97.4
EBNA-1 Q_p	1.1	0.4	97.0	100.0
KSHV latent				
ORF K2 vIL6	98.7	91.0	2.9	14.8
ORF 72 vCyclin	2.1	1.9	98.4	94.6
ORF 73 LANA	99.9	99.7	0.2	0.4
KSHV lytic				
ORF 11	1.1	0.3	80.6	70.0
ORF K3	1.1	0.8	83.9	87.0
ORF K4	16.4	9.9	76.7	85.5
ORF K5	5.2	2.7	83.3	87.2
ORF K6	3.3	0.8	87.2	82.6
ORF 16	45.6	31.2	70.4	80.5
ORF 45	10.4	5.0	87.7	94.5
ORF 47	1.5	0.2	88.1	100.0
ORF 54	11.0	6.8	71.6	75.5
ORF K9	38.2	19.7	63.1	81.1
ORF K10	5.8	2.8	90.4	97.6
ORF K11	1.2	0.3	88.2	100.0
ORF 75	9.7	5.7	90.0	93.5
PAN	16.3	6.8	87.2	92.9
KSHV unclassified				
ORF K2 vIL6, ORF 2[c]	2.3	1.1	95.5	87.5
ORF 2[d]	6.2	2.4	90.4	98.6

[a] Only transcripts present in at least 1% of cells in at least one replicate are shown.
[b] The union of transcripts could be assigned to multiple genes, but all components are latent.
[c] The union of transcripts contains one component of unknown expression timing.
[d] The expression timing of this transcript is not known.

would have previously been ignored. Comprehensive profiling instills more confidence in our interpretations.

Preliminary analysis of our scRNA-seq data reveals an unexpected plethora of lytic KSHV transcripts. Herpesviruses switch between a transcriptionally quiescent latent state and a reactivated lytic state. Genes were classified as latent or lytic based on expression timing (10, 11). For EBV, we detected only latent genes. For KSHV, we detected both latent and lytic genes (Table 1). Transcription of the lytic Open Reading Frames 16, K9, K4, 54, 45, and 75, as well as the lytic polyadenylated nuclear RNA PAN gene, can each be found in ~5 to 50% of cells. While we usually detected more than 1 count in each cell for the latent genes LANA and vIL6, lytic genes appear with only 1 count most of the time. Greater than 50% of cells contain these trace levels of lytic RNA. In contrast, ~98% of BC-1 cells are thought to harbor latent KSHV because lytic protein is not detected (12). Although more detailed analysis is required, our observations prompt a thoughtful redefinition of a "latent" transcriptome.

Data availability. Our data have been deposited in the NCBI Gene Expression Omnibus (13, 14) under accession number GSE154900.

REFERENCES

1. 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.

2. Moquin SA, Thomas S, Whalen S, Warburton A, Fernandez SG, McBride AA, Pollard KS, Miranda JL. 2017. The Epstein-Barr virus episome maneuvers between nuclear chromatin compartments during reactivation. J Virol 92: e01413-17. https://doi.org/10.1128/JVI.01413-17.

3. Zheng GXY, Terry JM, Belgrader P, Ryvkin P, Bent ZW, Wilson R, Ziraldo SB, Wheeler TD, McDermott GP, Zhu J, Gregory MT, Shuga J, Montesclaros L, Underwood JG, Masquelier DA, Nishimura SY, Schnall-Levin M, Wyatt PW, Hindson CM, Bharadwaj R, Wong A, Ness KD, Beppu LW, Deeg HJ, McFarland C, Loeb KR, Valente WJ, Ericson NG, Stevens EA, Radich JP, Mikkelsen TS, Hindson BJ, Bielas JH. 2017. Massively parallel digital transcriptional profiling of single cells. Nat Commun 8:14049. https://doi.org/10.1038/ncomms14049.

4. de Jesus O, Smith PR, Spender LC, Elgueta Karstegl C, Niller HH, Huang D, Farrell PJ. 2003. Updated Epstein-Barr virus (EBV) DNA sequence and analysis of a promoter for the BART (CST, BARF0) RNAs of EBV. J Gen Virol 84:1443–1450. https://doi.org/10.1099/vir.0.19054-0.

5. Russo JJ, Bohenzky RA, Chien MC, Chen J, Yan M, Maddalena D, Parry JP, Peruzzi D, Edelman IS, Chang Y, Moore PS. 1996. Nucleotide sequence of the Kaposi sarcoma-associated herpesvirus (HHV8). Proc Natl Acad Sci U S A 93:14862–14867. https://doi.org/10.1073/pnas.93.25.14862.

6. Concha M, Wang X, Cao S, Baddoo M, Fewell C, Lin Z, Hulme W, Hedges D, McBride J, Flemington EK. 2012. Identification of new viral genes and transcript isoforms during Epstein-Barr virus reactivation using RNA-Seq. J Virol 86:1458–1467. https://doi.org/10.1128/JVI.06537-11.

7. Hafemeister C, Satija R. 2019. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. Genome Biol 20:296. https://doi.org/10.1186/s13059-019-1874-1.

8. 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.

9. Phan AT, Martinez DM, Miranda JL. 2017. RNA-seq detects pharmacological inhibition of Epstein-Barr virus late transcription during spontaneous reactivation. Genom Data 13:5–6. https://doi.org/10.1016/j.gdata.2017.05.012.

10. Yuan J, Cahir-McFarland E, Zhao B, Kieff E. 2006. Virus and cell RNAs expressed during Epstein-Barr virus replication. J Virol 80:2548–2565. https://doi.org/10.1128/JVI.80.5.2548-2565.2006.

11. 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.

12. Miller G, Heston L, Grogan E, Gradoville L, Rigsby M, Sun R, Shedd D, Kushnaryov VM, Grossberg S, Chang Y. 1997. Selective switch between latency and lytic replication of Kaposi's sarcoma herpesvirus and Epstein-Barr virus in dually infected body cavity lymphoma cells. J Virol 71:314–324. https://doi.org/10.1128/JVI.71.1.314-324.1997.

13. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Holko M, Yefanov A, Lee H, Zhang N, Robertson CL, Serova N, Davis S, Soboleva A. 2013. NCBI GEO: archive for functional genomics data sets: update. Nucleic Acids Res 41:D991–D995. https://doi.org/10.1093/nar/gks1193.

14. Edgar R, Domrachev M, Lash AE. 2002. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res 30:207–210. https://doi.org/10.1093/nar/30.1.207.