Supplementary Information for

5-Methylcytosine Modification of an Epstein-Barr Virus Noncoding RNA Decreases its Stability

Belle A. Henry[1], Jack P. Kanarek[1], Annika Kotter[2], Mark Helm[2], Nara Lee[1]*

Supplementary Figure Legends

Supplementary Figure 1. Outline for generating CRISPR-Cas9 knockout cells of NSUN2 and DNMT2. (**A**) Genomic locus of NSUN2 and location of sgRNA target sites (red boxes) for CRISPR knockout are shown. Primer pairs for examining CRISPR-mediated deletion are indicated underneath. (**B**) After nucleofection of plasmids expressing Cas9 and both sgRNAs into BJAB-B1 cells, nucleofected cells were diluted to ~5-10 cells as a starting population of a clonal cell line. Please note that lymphocytes in our hands do not expand when diluted to less than ~5 cells per well. Five of the 96 examined clones (referred to as cl.1 - cl.5) showed undetectable NSUN2 protein levels as determined by Western blot. Anti-Tubulin antibody was used as a loading control. (**C**) The knockout cell lines cl.1 and cl.2 were examined further by genomic PCR for correct gene editing using the primer pairs indicated in A. Primer set #1 produces a 334 bp amplicon if the NSUN2 gene is unedited, while primer set #2 will generate a 398 bp amplicon upon deletion of

the internal region of the NSUN2 gene. Genomic PCR was also conducted with the parental BJAB-B1 cell line. (**D**) Increased EBER1 levels in NSUN2 knockout clones of BJAB-B1 cells is not due to elevated EBV episome copy number inside host cells. The EBV genome persists in latently infected cells as multicopy episomes. The number of episomes can vary between cell lines from 5-800 copies. The relative EBV genome content was measured by qPCR normalized to the *ACTIN* gene locus as previously described (1). Except for clone 4 (cl.4), which exhibits a ~1.8-fold increase in relative EBV copy number compared to the parental cell line, all other clones contain a comparable EBV genome copy number inside host cells. (**E**) Genomic locus of DNMT2 and location of sgRNA target sites (red boxes) for CRISPR knockout are shown. DNMT2 knockout clones were characterized as described for NSUN2 knockout cell lines.

Supplementary Figure 2. ALYREF and YBX1 are not EBER1-specific m^5C-readers. (**A**) Representative autoradiographs following immunoprecipitation of either ALY (top) or YBX1 (bottom) are shown. RNA-protein adducts (red box) that are present only after UV irradiation and represent RNA fragments crosslinked to ALY/YBX1 were excised and converted into a deep-sequencing library. (**B**) HITS-CLIP tracks of ALY and YBX1 across the EBER locus. HITS-CLIP for ALY and YBX1 was carried out as previously described (2,3) using either wildtype or NSUN2 knockout BJAB-B1 cells and either anti-ALY antibody (Abcam cat. no. ab202829) or anti-YBX1 antibody (Abcam cat no. ab76149). The vast majority of the mappable reads aligned to the human genome and only >1% of reads were EBV-specific. Of these, almost all reads mapped to the EBER locus. Deep sequencing data were deposited in the Sequence Read Archive under BioProject ID PRJNA555033 (accession number SRR9696698 for ALY-CLIP with BJAB-B1 wildtype cells, SRR9696699 for ALY-CLIP with BJAB-B1 NSUN2 knockout cells, and SRR10674124 for YBX1-CLIP with BJAB-B1 wildtype cells). YBX1 does not bind EBER1 over the methylation site (marked by dashed line), indicating that it does not act as a m^5C reader of EBER1. Three peaks over EBER1 are detected for ALY, of which one overlaps the methylation

site. However, ALY binding remains unchanged in NSUN2 knockout cells, which completely lack EBER1 m^5C methylation, indicating that ALY is neither a *bona fide* EBER1-specific m^5C reader. (**C**) RNAi against ALY is ineffective in BJAB-B1 lymphocytes. ALY is successfully knocked down in HEK293T cells as shown by Western blot analysis with siGENOME siRNAs from Dharmacon (cat no. M-012078-01-0005). In contrast, even micromolar concentrations of siRNA cannot elicit a knockdown of ALY in BJAB-B1 cells. (**D**) Screen shot taken from the UCSC Genome Browser showing the expression of ALY in various tissues. Markedly elevated expression of ALY is observed in "*Cells - EBV-transformed lymphocytes*" (arrow). (**E**) EBER1 levels are unaffected by knockdown of ALY. While HEK293T cells can be efficiently knocked down by RNAi, EBV-positive lymphocytes were refractory to siRNA-mediated knockdown even at a tenfold higher concentration of siRNAs, which may be due to highly increased ALY mRNA levels in this cell type (D). Thus, we used CRISPR interference to deplete ALY, which employs a catalytically inactive Cas9 endonuclease that retains the ability to be targeted by single guide (sg)RNAs and acts as a road block for RNA polymerase II. ALY was depleted using CRISPRi and dCas9-KRAB using sgRNAs targeting promoter proximal sequences of the ALY locus (sgRNA 1: 5′-TGGCACCGGAAGGCCTCACG-3′; sgRNA 2: 5′-GCCCGCCATGGCCGACAAAA-3′; sgRNA 3: 5′-CGACAAAATGGACATGTCTC-3′). The regions targeted by ALY-specific sgRNAs are indicated. Western blot analysis was carried out to examine ALY depletion. Anti-Nucleolin antibody was used as a loading control. (**F**) EBER1 levels are unaffected by knockdown of YBX1 using siRNAs (Dharmacon cat. no. M-010213-03-0005). Western blot analysis was carried out to measure the efficiency of YBX1 depletion. Anti-Nucleolin antibody was used as a loading control. Northern blots probing for EBER1 and U6 as a loading control are shown.

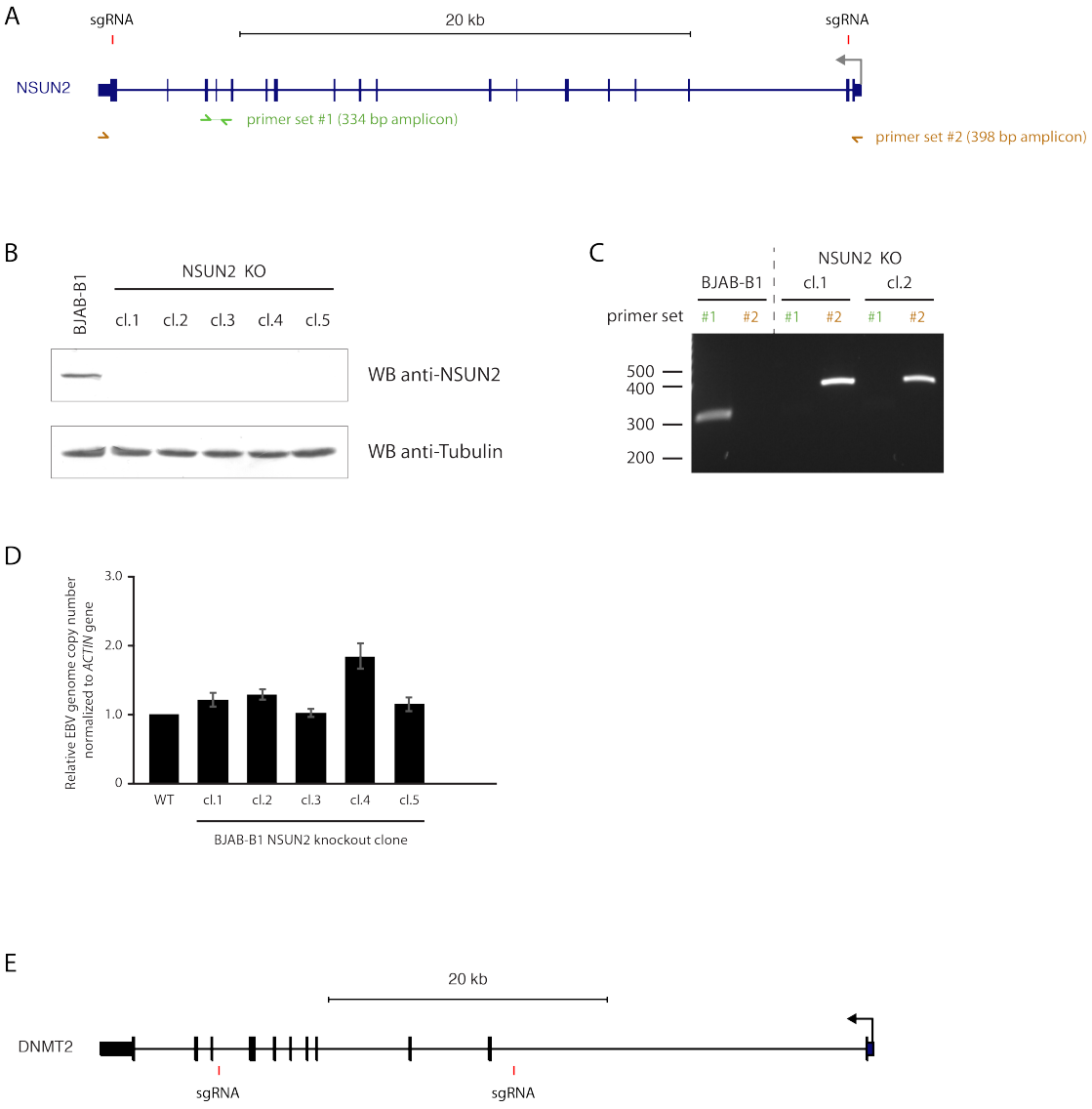
Supplementary Figure 3. m^5C modification of EBER1 is dispensable for EBV lytic replication. (**A**) NSUN2 knockout clones of the lytic-replication permissive HH514-16 cell line were generated. Western blot analysis of NSUN2 for two knockout clones is shown; anti-Tubulin antibody was

used as a loading control. (**B+C**) NSUN2 knockout and thus absence of m^5C modification on EBER1 does not affect viral lytic replication. Lytic cycle was induced by the addition of 3 mM sodium butyrate (NaB) into the culture medium, and EBV genome content inside host cells (B) or the viral titer in the culture supernatant (C) was measured after three days of NaB treatment. Experimental outline for lytic induction has been previously described (1).

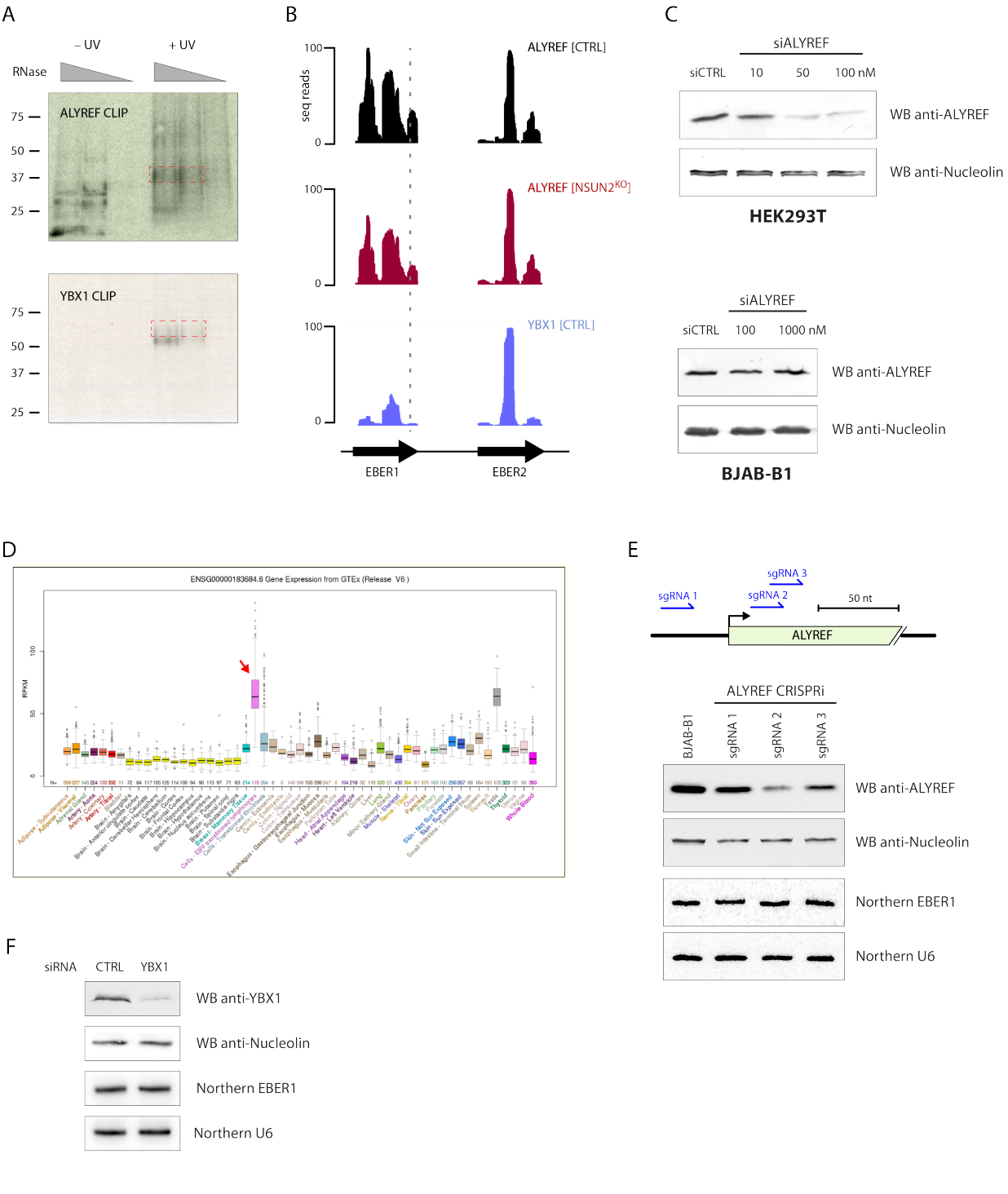
References

1. Lee, N., Moss, W.N., Yario, T.A. and Steitz, J.A. (2015) EBV noncoding RNA binds nascent RNA to drive host PAX5 to viral DNA. *Cell*, **160**, 607-618.
2. Lee, N., Le Sage, V., Nanni, A.V., Snyder, D.J., Cooper, V.S. and Lakdawala, S.S. (2017) Genome-wide analysis of influenza viral RNA and nucleoprotein association. *Nucleic Acids Res*, **45**, 8968-8977.
3. Moore, M.J., Zhang, C., Gantman, E.C., Mele, A., Darnell, J.C. and Darnell, R.B. (2014) Mapping Argonaute and conventional RNA-binding protein interactions with RNA at single-nucleotide resolution using HITS-CLIP and CIMS analysis. *Nat Protoc*, **9**, 263-293.

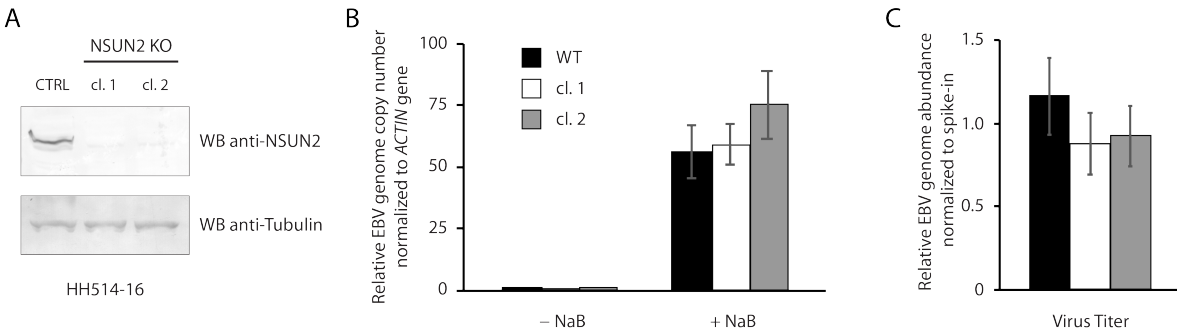
Supplementary Figure 1



Supplementary Figure 2



Supplementary Figure 3



Supplementary Table 1. List of RNA modifications analyzed by LC-MS/MS for presence in EBER1 and EBER2. For EBER1, only m^5C was detected at stoichiometric levels (~1 modification per RNA molecule); m^7G and m^5U, probably stemming from contaminating RNA species, were detected at trace levels only (less than 0.03 modifications per RNA molecule); other modifications were not detected at significant levels.

Name	Abbreviation	Elemental composition	Expected m/z values	Retention time [min]
pseudouridine	Ψ	$C_9O_6N_2H_{12}$	245	3.9
1-methyladenosine	m^1A	$C_{11}O_4N_5H_{15}$	282	4.7
5-hydroxymethylcytidine	hm^5C	$C_{10}O_6N_3H_{15}$	274	5.1
3-methylcytidine	m^3C	$C_{10}O_5N_3H_{15}$	258	6
7-methylguanosine	m^7G	$C_{11}O_5N_5H_{17}$	298	7.4
5-carbamoylmethyluridine	ncm^5U	$C_{11}O_7N_3H_{15}$	302	7.5
5-methylcytidine	m^5C	$C_{10}O_5N_3H_{15}$	258	8.8
2′-O-methylcytidine	Cm	$C_{10}O_5N_3H_{15}$	258	9.5
inosine	I	$C_{10}O_5N_4H_{12}$	269	10
5-methyluridine	m^5U	$C_{10}O_6N_2H_{14}$	259	10.6
2-thiocytidine	s^2C	$C_9O_4N_3H_{13}S_1$	260	10.6
2-thiouridine	s^2U	$C_9O_5N_2H_{12}S_1$	261	11
2′-O-methyluridine	Um	$C_{10}O_6N2H_{14}$	259	11.6
3-methyluridine	m^3U	$C_{10}O_6N_2H_{14}$	259	12.1
1-methylguanosine	m^1G	$C_{11}O_5N_5H_{15}$	298	12.5
5-methoxycarbonylmethyluridine	mcm^5U	$C_{12}O_8N_2H_{16}$	317	12.7
2′-O-methylguanosine	Gm	$C_{11}O_5N_5H_{15}$	298	12.7
N^2-methylguanosine	m^2G	$C_{11}O_5N_5H_{15}$	298	13
4-thiouridine	s^4U	$C_9O_5N_2H_{12}S_1$	261	13.1
2'-O-methylthymidine	m^5Um	$C_{11}O_6N_2H_{16}$	273	14.3
N^2,N^2-dimethylguanosine	$m^{2,2}G$	$C_{12}O_5N_5H_{17}$	312	14.9
5-methyl-2-thiouridine	m^5s^2U	$C_{10}O_5N_2H_{14}S_1$	275	15
2′-O-methyladenosine	Am	$C_{11}O_4N_5H_{15}$	282	16.1
2-methyladenosine	m^2A	$C_{11}O_4N_5H_{15}$	282	16.5
N^6-methyladenosine	m^6A	$C_{11}O_4N_5H_{15}$	282	16.8
N^6,N^6-dimethyladenosine	$m^{6,6}A$	$C_{12}O_4N_5H_{17}$	296	18.6
N^6-isopentenyladenosine	i^6A	$C_{15}O_4N_5H_{21}$	336	22