Supplementary Information

Circular DNA tumor viruses make circular RNAs

Tuna Toptan, Bizunesh Abere, Micheal A. Nalesnik, Steven H. Swerdlow, Sarangarajan Ranganathan, Nara Lee, Kathy H. Shair[a], Patrick S. Moore, Yuan Chang

Patrick S. Moore and Yuan Chang

E-mail: psm9@pitt.edu, yc70@pitt.edu

This PDF file includes:

Fig. S1 to S4

Table S1

Captions for Databases S1 to S4

References for SI references citations

Other supplementary materials for this manuscript include the following:

Datasets S1 to S4

www.pnas.org/cgi/doi/10.1073/pnas.1811728115

Supplementary Figures and Figure Legends

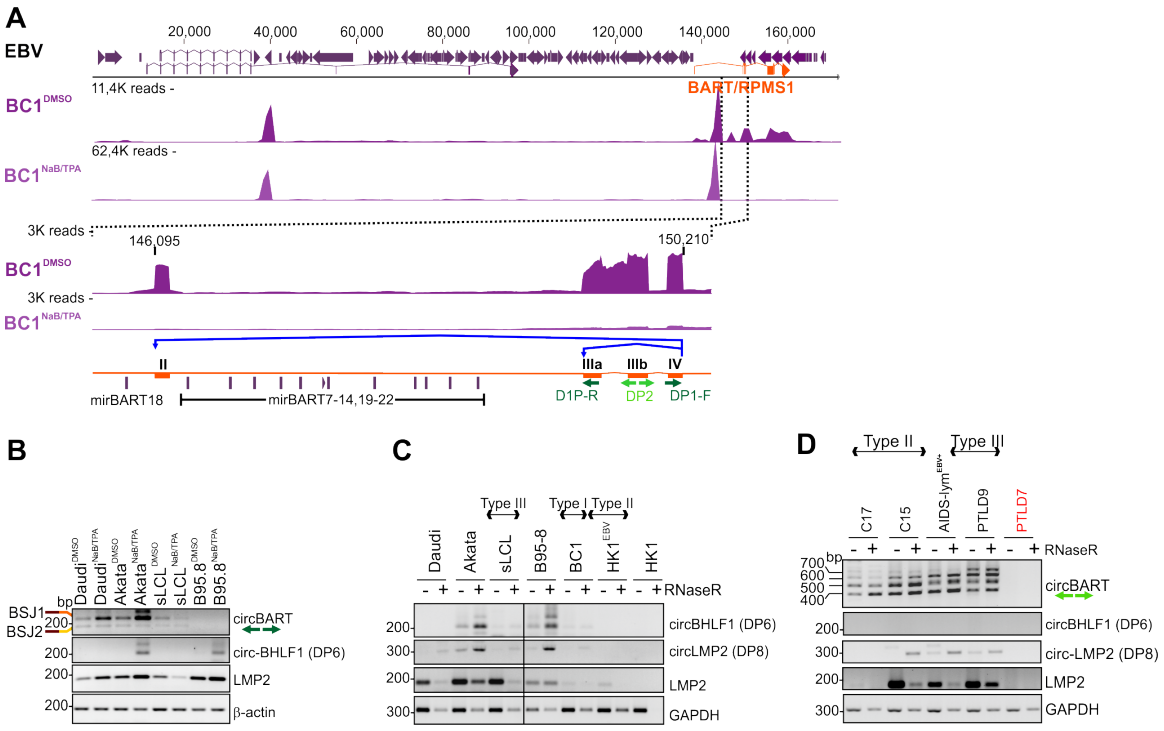

Fig. S1. Identification and validation of EBV circRNAs A. Identification of EBV backsplice junctions in BC1 cells. Ribominus and RNase R-treated RNA sequencing reads from EBV and KSHV co-infected BC1^{DMSO} and BC1$^{NaT/TPA}$ samples were mapped to the EBV reference genome (Mutu strain: KC207813) and read coverage files were generated using CLC Genomics Workbench tool. Top panel illustrates selected mRNAs and ORFs of EBV including RPMS1 in the BART region (orange). The coverage depth is indicated on the Y-axis for each alignment panel. The expanded view between 146-150,2 kb and shows read coverage in latent and reactivated BC1 cells. EBV mRNA and non-coding RNA (ncRNA) between genome position 146-150.2 kb corresponds to BART exons II-IV flanking the intronic region with the miRNAs (mirBART 7-22). CIRI2-predicted BSJ1 and BSJ2 are shown in light and dark blue respectively. BSJ1 is formed by backsplicing of the 3' end of exon IV onto the 5' end of exon II. BSJ2 is formed by 3' end exon IV backsplicing onto end of exon IIIa. Divergent primers designed to verify BSJ1 and BSJ2 are shown in dark green (DP1) and light green (DP2). **B. EBV circRNA expression following lytic induction.** EBV positive Daudi, Akata, sLCL and B95-8 cells were treated with DMSO or NaB/TPA for 48 h. Extracted RNA was analyzed by RT-PCR using junction spanning divergent primers for circBART (DP1, dark green arrows)

and circBHLF1. Viral LMP2 and cellular β-actin linear transcripts were analyzed as internal controls using convergent primers. DP1 RT-PCR amplified circBART-BSJ1 and BSJ2 in all condtions except B95-8 which has a deletion in BART locus. CircBHLF1 BSJ-PCR product (~200bp), was detected in NaB/TPA treated Akata and B95-8. Daudi is a Burkitt's lymphoma cell line which has a deletion in BHLF1 and its promoter region. **C. circBHLF1 and circLMP2 expression in different cell lines.** CIRI2 predicted additional EBV circRNAs in BC1 (**Dataset S1**). RNase R treated (+) or untreated (-) RNAs from cell lines having various forms of EBV latency, were analyzed by RT-PCR using divergent primers spanning BSJs in circBHLF1, circLMP2 and convergent primers (black arrows) for linear LMP2 and GAPDH transcripts. Sequencing analysis of the circBHLF1-BSJ spanning PCR product confirmed the predicted junction site given in Table S1. To confirm the predicted BSJ sites for circLMP2 DP7 was used for RT-PCR (**Table S1**) which produced multiple PCR products ranging between 200-1,200 bp enriched following RNase R treatment with Akata and B95-8 RNA. Following sequencing analysis of the PCR products we found an additional junction between 58nt-1682nt in Mutu strain genome position which was validated by circLMP2 DP8 primers shown here. Basepair (bp), EBV positive (*), nucleotide (nt). **D**. **circBHLF1 and circLMP2 expression in tumor samples.** RNase R treated (+) or untreated (-) RNAs from EBV(+) PTLD9, EBV(-) PTLD7 (shown in red), NPC tumor lines C17, C15 and an EBV (+) AIDS-associated lymphoma, were used for RT-PCR with DP2 (light green arrows) primers to detect circBARTs, circBHLF1 (DP6) and circLMP2 (DP8). Convergent primers (black arrows) for LMP2 (middle panel) and GAPDH (lower panel) linear transcripts were used as internal controls and to assess RNaseR treatment efficiency.

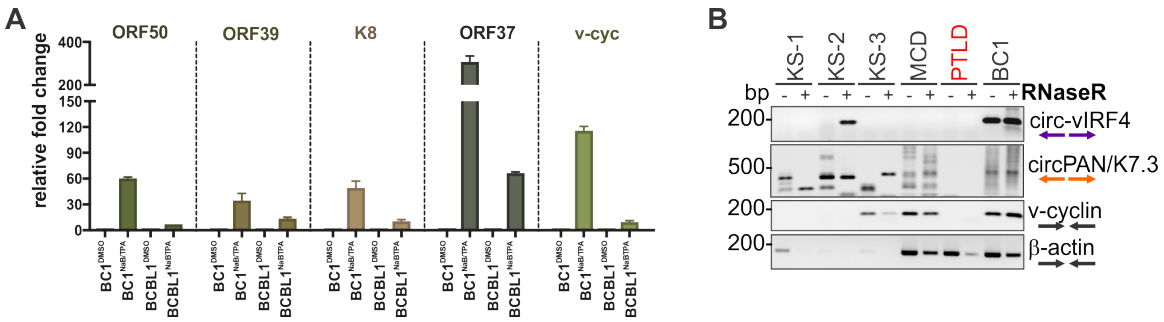

Fig. S2. A. Viral transcript expression analysis of BC1 and BCBL1 cell lines. KSHV (+) primary effusion lymphoma lines were treated with NaB/TPA for 48 h. Extracted RNA was used for Ribominus, RnaseR+ RNA sequencing. cDNA generated from these RNA was analyzed for immediate early (ORF50, ORF39), early (K8, ORF37) and latent (v-cyclin) transcript expression profile to assess the reactivation efficiency. The data are expressed as fold change of the NaB/TPA treated versus untreated (DMSO) cells after normalization to GAPDH. **B. KSHV circRNAs in KS and MCD patient tissues**. RNAs extracted from three KS and one MCD show circPAN/K7.5 BSJ in all KSHV-positive tissues but circvIRF4-BSJ was detected in only one KS sample. The KS specimens, stored in liquid nitrogen from the mid-1990s, showed evidence of RNA degradation with absence or diminished v-cyclin and β-actin mRNA RT-PCR positivity, consistent with the notion that circRNAs are particularly resistant to degradation. PTLD (EBV-negative) and BC-1 RNAs were used as virus negative and positive controls, respectively.

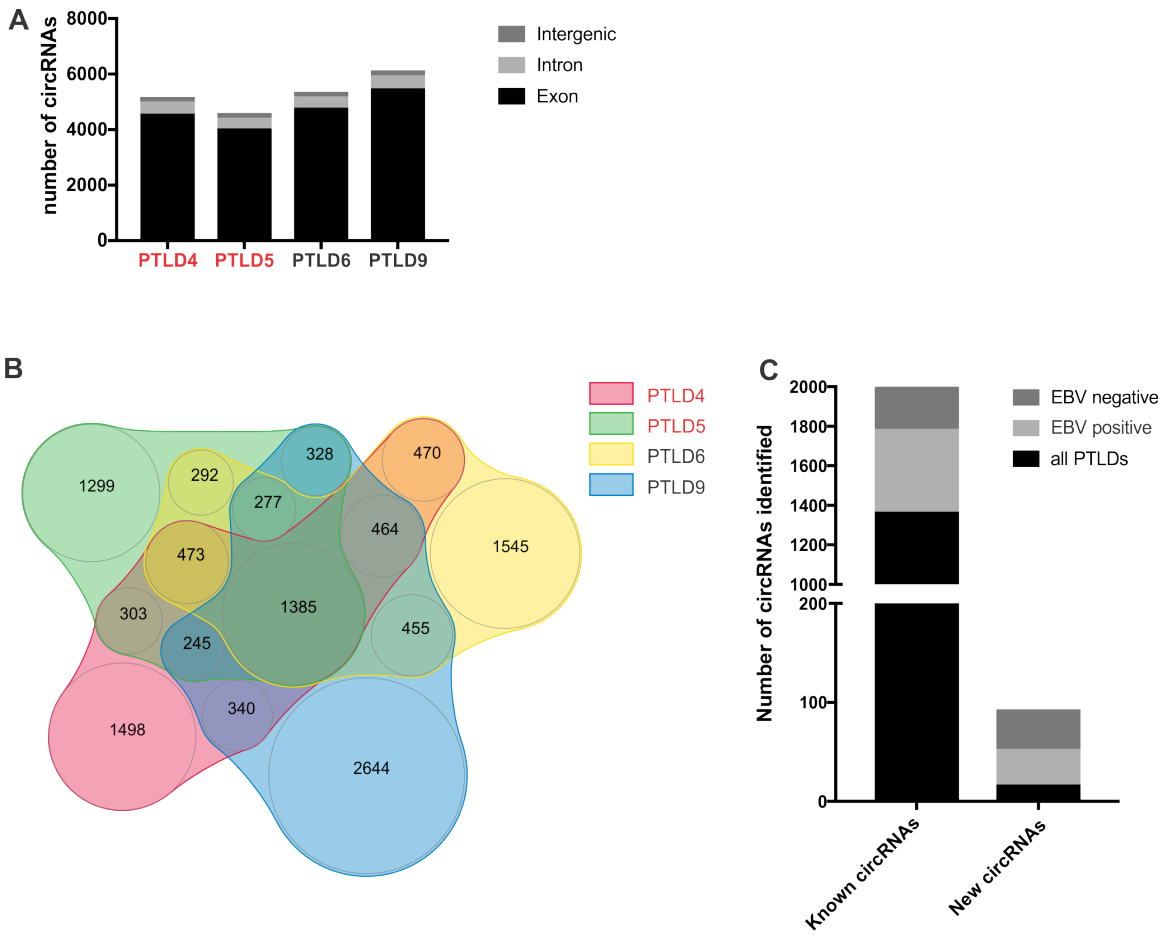

Fig S3. A. Type of predicted human circRNAs in PTLD samples. CIRI2 analysis found a total of 5178, 4602, 5361 and 6138 human circRNAs in PTLD4,5,6 and 9 respectively (**Dataset S3**). ~90% of these are exonic and the rest are generated from intronic and intergenic regions of the human genome. **B-C.** Venn diagram illustrates shared and uniquely expressed human circRNAs predicted by CIRI2 analysis in PTLD samples. EBV (+) PTLD 6 and 9 express 455 human circRNAs in common and 35 of them was identified in our study. EBV (-) PTLD 4 and 5 express 303 human circRNAs in common and we found 40 new circRNAs in this group (**Dataset S3**). ~99% of circRNAs (1385) that are common in all samples were annotated in circBase.

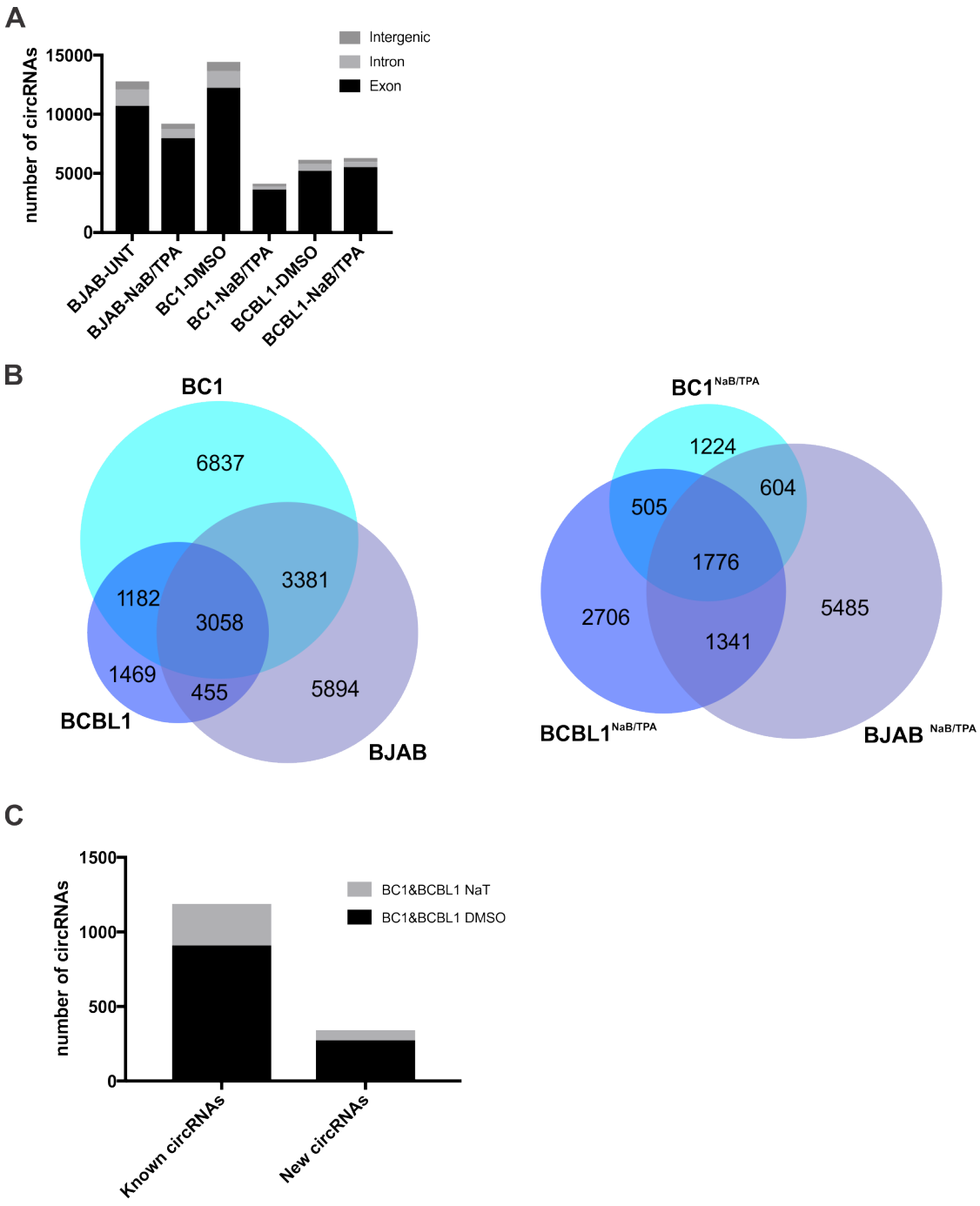

Fig S4. A. Type of predicted human circRNAs in KSHV (+/-) cell lines. CIRI2 analysis found a range of 4,100 to 14,400 cellular circRNAS these cell lines (**Dataset S4**). **B-C.** Venn diagram illustrates shared and uniquely expressed human circRNAs in BJAB and PEL cell lines predicted by CIRI2 analysis. We found 273 new human circRNAs both in latent BC1 and BCBL1 samples. NaB/TPA treated PEL cells have 505

human circRNAs in common, 98 of which have not been previously reported (**Dataset S4**).

Supplementary Table and Datasets

Table S1. Primers and antisense oligos (ASO) used in this study.

Name	Sequence	Figures	PCR product size (bp)
Divergent primers			
DP1 (circBART.BSJ2)	CGCCCGTATTCACACATTCC GACGCTAGTGCTGCATGGG	Fig. 1, Fig, 2, Fig. 4, Fig S1	160-264
DP2 (circBART)	AGCCCTTCTTCGTTATGCAC TGAGGAATACCTCGTTGTCTTCCG	Fig. 1, Fig. 4, Fig S1	400-700
DP3 (circvIRF4)	CAAAGCTACGAGGAGGCAGG CGCCGACACCAACGCATCAAAC	Fig. 3	577
DP4 (circvIRF4)	GGCGATATAACGACTGAACAGA CAAATGCATGGTACACCGAATAC	Fig 3, Fig. 4, Fig. S2	139
DP5 (circPAN/K7.3)	CGCCCACTGGTGTATCAGA AATCGCAGCTTTTGTTCTGC	Fig. 3, Fig. S2	126-668
DP6 (circBHLF1)	CGCTTGCCTGGTCCTGG CAGGCGTACCGGGCCAG	Fig. S1	216
DP7 (circLMP2)	CACCAGCGATTAGCGCG GGTCATTAGATGCTGCCGCTAC	Fig. S1	210-1,178
DP8 (circLMP2)	GCAGCGGCATATGAGCTGG GGTCATTAGATGCTGCCGCTAC	Fig. S1	258
DP9 (circvIRF4)	CATTTGATGAGGAGTGTGATAGAG GAACCGCTATTACAATGTTGGC	entire circvIRF4 amplification and sequencing	632
DP10 (circPAN/K7.3)	TTCTGTGTTTGTCTGATTCTTAG CCGAAACAACGAATGAGCA	Fig. 3	325-744
DP11 (circBART.BSJ1)	GGTCAAGTAGCTGCGTCCAAA GACGCTAGTGCTGCATGGG	Fig. 4	117
Convergent primers used for RT-PCR and qPCR			
GAPDH.F	GTCATCAATGGAAATCCCATCACC	Fig. 1, Fig. 2, Fig. 3	320
GAPDH.R	TGAGTCCTTCCACGATACCAAA		
GAPDH.F	TGCACCACCAACTGCTTAGC	Fig.3	98
GAPDH.R	GGCATGGACTGTGGTCATGAG		
Beta-actin.F	CACACTGTGCCCATCTATGAGG	Fig. 3,Fig. S2	191
Beta-actin.R	TCGAAGTCTAGGGCGACATAGC		
18S.F	CGAACGTCTGCCCTATCAACTT	Fig. 3	115
18S.R	TGTGGTAGCCGTTTCTCAGG		
vIL6.F	TTCAAAACACGCACCGCTTG	Fig. 3	210
vIL6.R	AAACGTGGACGTCATGGAGC		
v-cyc.F	CGCCTGTAGAACGGAAACAT	Fig. 3, Fig.4, Fig. S2	137
v-cyc.R	TTGCCCGCCTCTATTATCAG		
LANA F	TTTAGTGTAGAGGGACCTTGGG	Fig. 3	258
LANA R	TCTCCATCTCCTGCATTGCC		
KSHV.ORF50.F	CAGAGTCTATTCGCCCTGTTAG	Fig. S2	115
KSHV.ORF50.R	CTGGTACAGTCCTTGCAGAATA		
KSHV.K8.F	CCAAGAGGCGACTACATAGAAAG	Fig. S2	111
KSHV.K8.R	GGGTGATGTTCCCTACCTTAAC		
KSHV.ORF37.F	TGGGCGAGTTTATTGGTAGTG	Fig. S2	125
KSHV.ORF37.R	CGCTGATGTGCGTTCATTTG		
KSHV.ORF39.F	CAGGCAGCAGTAGAATCAGATAA	Fig. S2	110
KSHV.ORF39.R	GACGGTCGTGTGGTACATAAA		
LMP2.F	TGCCTGCCTGTAATTGTTGCG	Fig. 1, Fig. 2, Fig. 4	151
LMP2.R	GCAGCGGCATATGAGCTGG		

Antisense oligos (ASO)

circBART_ASO-BSJ2	mG*mA*mC* mA*mC*mG* C*C*G* G*A*C* C*T*T* G*C*C* mC*mG*mU* mC*mG*mA
circBART_ASO-BSJ1	mG*mC*mC*mC*mA*mA*T*G*G*C*A*T*C*T*T*G*C*C*mC*mG*mU*mC*mG*mA
circvIRF4_ASO	mG*mG*mG*mG*mC*mG*C*G*G*G*G*G*C*T*G*A*G*G*mU*mA*mG*mA*mU*mG

107
108

Captions for the Datasets

Dataset S1. CIRI2 EBV circRNA prediction analysis in BC1, PTLD6 and PTLD9 samples.

Dataset S2. CIRI2 EBV circRNA prediction analysis in BC1 and BCBL1 samples.

Dataset S3. CIRI2 human circRNA prediction analysis in EBV negative PTLD 4, PTLD5 and EBV positive PTLD6 and PTLD9 samples.

Dataset S4. CIRI2 human circRNA prediction analysis in KSHV positive PEL (BC1, BCBL1) and KSHV negative BJAB cells.

Column headings indicate chromosome (chr), circRNA strart and end position, strand, junction read counts, SM_MS_SMS, non-junction read counts, junction read ratio, circRNA type, and associated gene IDs for each predicted circRNA. SM_MS_SMS refers to types of paired chiastic clipping (PCC) signals detected by CIRI2 (1). For viral circRNAs calculated reads per mission (RPM) based on total aligned reads are given. In Datasets 3 and 4 circBase IDs for indicated samples are also indicated.

References

1. Gao Y, Wang J, & Zhao F (2015) CIRI: an efficient and unbiased algorithm for de novo circular RNA identification. *Genome Biol* 16:4.