HSV-1 and influenza infection induce linear and circular splicing of the long NEAT1 isoform

Marie-Sophie Friedl[1], Lara Djakovic[2], Michael Kluge[1], Thomas Hennig[2], Adam W. Whisnant[2], Simone Backes[2], Lars Dölken[2], Caroline C. Friedel[1]*

[1] Institute of Informatics, Ludwig-Maximilians-Universität München, Amalienstr. 17, 80333 Munich, Germany

[2] Institute for Virology and Immunobiology, Julius-Maximilians-University Würzburg, Versbacher Straße 7, 97078 Würzburg, Germany

* Corresponding author

Short title: Induction of linear and circular NEAT1 splicing

Abstract

The herpes simplex virus 1 (HSV-1) virion host shut-off (*vhs*) protein cleaves both cellular and viral mRNAs by a translation-initiation-dependent mechanism. Here, we show that *vhs*-mediated degradation of mRNAs leads to an accumulation of circular RNAs (circRNAs) relative to linear mRNAs during HSV-1 infection. Strikingly, we found that circular splicing of the long isoform (NEAT1_2) of the nuclear paraspeckle assembly transcript 1 (NEAT1) was massively induced during HSV-1 infection. In contrast to other circRNAs, induction of the NEAT1_2 circRNA was independent of *vhs* and occurred while NEAT1_2 was still bound to the chromatin. This was associated with induction of linear splicing of NEAT1_2 both within and downstream of the circRNA. NEAT1_2 splicing was absent in uninfected cells but can be induced by ectopic co-expression of the HSV-1 immediate-early proteins ICP22 and ICP27. Interestingly, NEAT1_2 circular and linear splicing was also up-regulated in influenza A virus (IAV) infection but absent in stress conditions, which disrupt transcription termination similar to but not by the same mechanisms as HSV-1 and IAV infection. Large-scale analysis of published RNA-seq data uncovered induction of NEAT1_2 splicing in cancer cells upon inhibition or knockdown of cyclin-dependent kinase 7 (CDK7) or the MED1 subunit of the Mediator complex phosphorylated by CDK7. Finally, CDK7 inhibition also disrupted transcription termination, highlighting a possible link between disruption of transcription termination and NEAT1_2 splicing.

Introduction

Herpes simplex virus 1 (HSV-1) is one of nine herpesviruses known to infect humans [1, 2]. It is most commonly known for causing cold sores but can also result in life-threatening diseases. A key role in HSV-1 lytic infections is played by the virion host shut-off (*vhs*) protein, which is delivered to the infected cell by the incoming virus particles and rapidly starts cleaving both cellular and viral mRNAs [3]. *Vhs* target recognition occurs during translation initiation and requires binding of *vhs* to components of the cap-binding complex eIF4F [4-8]. *Vhs* can also cleave circular RNAs (circRNAs), but only if they contain an internal ribosome entry site (IRES) [9]. CircRNAs form closed RNA-loops and are naturally generated during the splicing process either by “back-splicing” of exons out of their linear order (resulting in non-colinear 3′–5′ junctions) or stabilization of intron lariats with 2′–5′ junctions joining the 5′ and branchpoint nucleotides [10]. Previously considered to occur only rarely [11], recent large-scale RNA sequencing (RNA-seq) studies revealed the existence of thousands of circRNAs in eukaryotic cells [12, 13]. Due to the absence of a 5’ cap or poly(A) tail and their resistance to exonucleases, circRNAs are more stable than linear RNAs [10]. While a function as miRNA sponge has been reported for a small number of circRNAs [14, 15], this is now nevertheless thought to be rare [13]. Other functions reported for individual circRNAs involve regulation of transcription, alternative splicing and translation and some may even serve as templates for protein synthesis [13]. However, for most circRNAs their function remains elusive.

Recently, Shi *et al.* reported on dysregulation of circRNAs during HSV-1 infection, with 188 circRNAs being significantly up-regulated at 48 h post infection (p.i.) [16]. Considering the specific targeting of *vhs* to translated mRNAs, we hypothesized that up-regulation of circRNAs may simply be a consequence of circRNAs escaping *vhs*-

mediated RNA decay. We previously performed 4-thiouridine-(4sU)-tagging followed by sequencing (4sU-seq) to characterize *de novo* transcription in hourly intervals during the first 8 h of lytic wild-type (WT) HSV-1 strain 17 infection of primary human foreskin fibroblasts (HFFs) and combined this with total RNA-seq in two-hourly steps [17, 18] (S1 Fig a, n=2 replicates). More recently, we applied the same experimental set-up to Δ*vhs* infection to study *vhs*-dependent gene regulation [19] (S1 Fig a). In this study, we explore these data to show a general accumulation of circRNAs in a *vhs*-dependent manner during HSV-1 infection.

Fig 1. *vhs*-dependent accumulation of circRNAs in HSV-1 infection.

(a) Genomic coordinates of the NEAT1_1 (polyadenylated) and NEAT1_2 (non-polyadenylated, stabilized by triple helix) transcript (blue) and the circRNA induced in HSV-1 and influenza A virus (IAV) infection (red). **(b)** Number of circRNAs identified for each sample of total RNA for WT and Δ*vhs* infection. Numbers for 4sU-RNA are shown in S2 Fig a. **(c,d)** Scatterplots comparing normalized circRNA counts (normalized to the number of linear junction reads mapped to the host genome) between mock and 8 h post infection (p.i.) WT infection **(c)** and mock and 8 h p.i. Δ*vhs* infection **(d)**. Linear regression analysis across all circRNAs (red line) was used to estimate the accumulation of circRNAs compared to mock infection. The regression estimate for the accumulation is shown on the bottom right. The gray line indicates the diagonal, i.e. equal values on the x- and y-axis. The five most highly expressed circRNAs are marked by name. Corresponding scatterplots for 2, 4 and 6 h p.i. are shown in S3 Fig. **(e)** Percentage of expressed circRNAs (=normalized circRNA count >0 in uninfected cells) at each time-point of WT HSV-1 infection for which at least one exon within the genomic region of the circRNA shows differential

exon usage (determined with DEXSeq, multiple testing adjusted p-value ≤0.005). **(f)** Boxplots showing the distribution of log2 fold-changes for exons located within circRNAs. For each circRNA, only the exon with the maximum absolute log2 fold-change is shown.

Strikingly, this analysis also revealed induction of both circular and novel linear splicing of the long non-coding (linc)RNA nuclear paraspeckle assembly transcript 1 (NEAT1) during HSV-1 infection. In contrast to other circRNAs, this induction was independent of *vhs.* The nuclear lincRNA NEAT1 is one of the most highly expressed lincRNAs and has two isoforms with identical 5' but distinct 3' ends [20] (Fig 1a): A short transcript (NEAT1_1, also denoted as MENε, ~3.7nt), stabilized by a poly(A) tail and expressed in a wide range of cells [21], and a long transcript (NEAT1_2, also denoted as MENβ, ~22.7 nt), stabilized by a triple helical structure [22] but with more limited expression in particular cell types [21]. The HSV-1-induced circRNA is generated from the 3' region unique to the NEAT1_2 transcript (Fig 1a). NEAT1_2 is essential for the structure of paraspeckles [23], nuclear bodies located in the interchromatin space. NEAT1_2 forms a scaffold for paraspeckles by being bound by a number of proteins with functions in transcription and/or RNA processing and splicing (reviewed in [24, 25]). Paraspeckles impact gene expression by nuclear retention of adenosine-to-inosine-edited mRNAs [25]. In viral infection, NEAT1_2 plays a role in interleukin 8 (IL-8) induction by sequestering the IL-8 repressor splicing factor proline/glutamine-rich (SFPQ) to paraspeckles [26]. NEAT1 is up-regulated in HIV-1, dengue, HSV-1, Hantavirus, Hepatitis D and influenza infections [25]. Expression levels of NEAT1_2 have been shown to be positively correlated with presence of paraspeckles [27] and HSV-1 infection induces paraspeckle formation, likely through up-regulation of NEAT1 [28].

The HSV-1 immediate-early proteins ICP27 and ICP22 are both known to interact with the host transcription and RNA processing machinery. ICP27 regulates splicing and polyadenylation by interacting with splicing factors and the mRNA 3' processing factor CPSF [29-32]. We recently demonstrated that ICP27 is sufficient but not necessary for disruption of transcription termination in HSV-1 infection [31]. ICP22 represses RNA Polymerase II (Pol II) transcription elongation by interacting with elongation factors, such as the positive transcription elongation factor b (P-TEFb) and the FACT complex, and inhibits phosphorylation of the Pol II carboxyterminal domain (CTD) at Ser2 residues [33]. Ser2 phosphorylation plays a key role in recruiting splicing and termination factors to Pol II [34]. We thus hypothesized that ICP27 and/or ICP22 may play a role in induction of NEAT1_2 splicing. Indeed, we could show that combined ectopic expression of ICP27 and ICP22 is sufficient to induce both circular and linear splicing of NEAT1_2.

HSV-1 and influenza A (IAV) infection as well as cellular stresses disrupt transcription termination, leading to read-through transcription beyond poly(A) sites for tens-of-thousands of nucleotides for many but not all cellular genes [17, 18, 35-37]. Strikingly, IAV infection, but not salt and heat stress, also induces NEAT1_2 circular and linear splicing. Moreover, selective inhibition and knockdown of cyclin-dependent kinase 7 (CDK7) as well as knockdown of its phosphorylation target the MED1 subunit of the Mediator complex [38] induce both NEAT1_2 splicing and read-through transcription in cancer cell lines. This highlights a possible link between read-through transcription and NEAT1_2 splicing and a potential role of CDK7 in HSV-1 or IAV infection.

Results

Vhs-dependent accumulation of circRNAs in HSV-1 infection

CircRNA detection from RNA-seq reads is based on identifying junction reads connecting a donor splice site of a downstream exon to the acceptor splice site of an upstream exon. Several algorithms for circRNA *de novo* detection are available, but due to high false positive rates combination of at least two algorithms is recommended [39]. We thus employed a pipeline combining circRNA_finder [40] and CIRI2 [41] for circRNA *de novo* detection (see methods and S1 Fig b) and analyzed only circRNAs that were independently discovered by both algorithms. We applied this pipeline on our total RNA-seq and 4sU-seq time-courses for the first 8 h of lytic wild-type (WT) and Δ*vhs* HSV-1 infection (S1 Fig a). This identified a total of 16,463 circRNAs in the WT time-course and 7,306 in the Δ*vhs* infection time-course. Consistent with previous reports [42] and the high stability of circRNAs, only a small number of circRNAs were detected in 4sU-RNA (119 and 194 in at least one sample of WT and Δ*vhs* 4sU-seq, respectively). Most circRNAs (>97%) were only detected in total RNA (Fig 1b, S2 Fig a,b). Almost all of the identified circRNAs (82.6% and 92.7% for the WT and Δ*vhs* infection time-courses, respectively) were already annotated in circBase, a database of circRNAs identified in large-scale RNA-seq studies [43].

Strikingly, in total RNA a substantially larger number of circRNAs was identified for WT infection than for mock and Δ*vhs* infection (Fig 1b). Since the number of reads mapping to the host genome in total RNA was comparable or higher in Δ*vhs* infection than in WT infection (S2 Fig b), the lower number of circRNAs identified in Δ*vhs* infection is not due to lower sequencing depth. To evaluate abundance of individual circRNAs, circRNA read counts were normalized to the total number of reads

mapping to (linear) exon-exon junctions of host protein-coding genes and then averaged between replicates. We used this normalization since the fraction of (linearly) spliced host reads decreased considerably during WT infection (S2 Fig c). This is largely a consequence of HSV-1-induced read-through transcription. Read-through transcripts are rarely spliced downstream of the poly(A) site and are retained in the nucleus [17, 18], thus escaping *vhs*-mediated decay. In *vhs* infection, the fraction of spliced reads did not decrease (S2 Fig d) due to reduced read-through transcription and absence of *vhs*-mediated decay of cytosolic mRNAs. Normalized circRNA counts increased substantially with increasing duration of WT infection (Fig 1c, S3 Fig a-c, red line = linear regression estimate over all circRNAs). By 8 h p.i. WT infection, circRNAs were ~6.5-fold increased in total RNA compared to uninfected cells. In contrast, Δ*vhs* infection exhibited only a very modest ~1.3-fold increase in circRNAs by 8 h p.i. (Fig 1d, S3 Fig d-f), which can likely be attributed to the global loss of host transcriptional activity in HSV-1 infection [44]. Since circRNAs are more stable than linear RNAs, they can consequently accumulate relative to linear RNAs even in absence of *vhs*-mediated decay. Notably, if we normalize by the total numbers of reads mapped to the host, we also observe increased accumulation of circRNAs in WT infection compared to Δ*vhs* infection, albeit to a lower degree (S4 Fig).

One factor that might negatively impact circRNA detection in the Δ*vhs* infection time-course is the shorter read length (76 nt in the Δ*vhs* infection time-course vs. 101 nt in the WT infection time-course, both paired-end sequencing). CIRI2 and circRNA_finder require that at least one read overlaps by at least 19 and 20 nt, respectively, on both sides of identified circular junctions. Thus, longer reads could potentially lead to identification of more circRNAs. To address this issue, we pursued

two alternative approaches. First, we trimmed reads from the WT infection time-course to 76 nt and applied our *de novo* circRNA detection method to these trimmed reads. Second, we implemented an alternative circRNA detection method based on aligning reads against circRNA junction sequences constructed from annotated circRNAs, which allowed to reduce or increase the required overlap on either side of the junction (see methods and S5 Fig). In this case, false positive rates are reduced by the restriction to annotated circRNAs. Both with trimmed reads and the alignment-based pipeline with a required overlap of only 10 nt, we observed approximately the same circRNA enrichment in WT infection as with the *de novo* detection pipeline on full reads (S6 and S7 Fig). This confirms the strong *vhs*-dependent accumulation of circRNAs in WT HSV-1 infection. Unless explicitly noted otherwise, all results reported in the following use the alignment-based pipeline requiring an overlap of at least 10 nt on either side of the junction.

We also analyzed data from a separate RNA-seq experiment in which we obtained total RNA for mock, WT, Δ*vhs* and ΔICP27 infection in the same experiment as well as subcellular fractions (chromatin-associated, nucleoplasmic and cytosolic RNA) (all 8 h p.i., n=2) [18, 19, 45]. This revealed an accumulation of circRNAs in total RNA of WT and ΔICP27 infection, but not Δ*vhs* infection (S8 Fig), Notably, although linear regression analysis over all circRNAs (red line) estimated only a 1.5- to 2-fold accumulation in WT infection, some of the most highly expressed circRNAs accumulated more strongly. We conclude that *vhs*-mediated degradation of mRNAs leads to an accumulation of circRNAs relative to linear mRNAs during HSV-1 infection.

CircRNA reads that do not cross circular junctions can be aligned by standard RNA-seq mapping programs and are thus counted towards circRNA “parent” genes and

exons in standard differential gene expression (DGE) and differential exon usage (DEU) analyses. Consequently, genes encoding for circRNAs tend to have significantly higher fold-changes in DGE analysis of total RNA in WT infection from 6 h p.i. (one-sided Wilcoxon rank sum test, p-value <0.0001, S9 Fig a). This was not observed in Δ*vhs* infection. Similarly, DEU analysis with DEXSeq showed differential exon usage in WT infection for many exons contained in circRNAs (multiple testing adjusted p-value ≤0.005). By 8 h p.i. WT infection, ~17% of circRNAs showed differential exon usage for at least one exon within the genomic range of the circRNA (Fig 1e), with increased exon usage for almost all of these circRNAs (Fig 1f). This effect was most pronounced for highly expressed circRNAs (S9 Fig b). Consistent with the small degree of circRNA accumulation late in Δ*vhs* infection, 3.6% of circRNAs also showed differential exon usage by 8 h p.i. Δ*vhs* infection (S9 Fig c,d). In summary, our results demonstrate that accumulation of circRNAs in WT infection biases both DGE and DEU analyses of total RNA and thus needs to be taken into account in these analyses.

Induction of NEAT1 linear and circular splicing in HSV-1 infection

Strikingly, we identified one circRNA (circBase id: hsa_circ_0003812) that increased massively (up to >500-fold in total and 4sU-RNA) in HSV-1 infection. In contrast to other circRNAs, this could be observed even in absence of *vhs* and in 4sU-, nucleoplasmic and chromatin-associated RNA (Fig 2a, S10 Fig). The latter is particularly noteworthy considering the small number of circRNAs that were recovered in these samples (S2 Fig a,b, S11 Fig). hsa_circ_0003812 is encoded in the unique part of the long NEAT1_2 transcript (Fig 1a) and was essentially absent in mock-infected cells (0-2 reads in 4sU-RNA, total RNA and all subcellular fractions,

≤0.22 normalized circRNA count). During infection, it was induced in all samples of infected cells (with the notable exception of cytosolic RNA), in particular also in ΔICP27 and Δ*vhs* infection, albeit to a lesser degree (Fig 2a, S10 Fig). No other circRNA was induced this strongly, not even in presence of *vhs*. From 6 h p.i., the NEAT1_2 circRNA was among the most highly expressed circRNAs in 4sU-, total, chromatin-associated and nucleoplasmic RNA with an 80- to 1,590-fold enrichment in WT infection compared to mock, 7- to 21-fold enrichment in ΔICP27 infection and 11- to 108-fold enrichment in Δ*vhs* infection (S10 Fig). The latter shows that, in contrast to other circRNAs, the NEAT1_2 circRNA is not induced simply by *vhs*-mediated degradation of the linear transcript. Consistent with the nuclear localization of NEAT1, some enrichment was also observed in cytosolic RNA of WT (but not ΔICP27 and Δ*vhs*) infection, but absolute levels in cytosolic RNA remained low. Other NEAT1_2 circRNAs were also detected in total RNA in WT infection, but only at low levels (Fig 2a). Since nucleoplasmic, chromatin-associated and 4sU-RNA represent very recently transcribed RNA, our data indicate that the hsa_circ_0003812 NEAT1_2 circRNA is synthesized at very high rates during HSV-1 infection. This again contrasts with most other circRNAs, which cannot be recovered in newly transcribed RNA.

Fig 2. Induction of a NEAT1_2 circular and linear splicing in WT, Δ*vhs* and ΔICP27 infection.

(a) Heatmap of normalized circRNA counts (normalized to the number of linear junction reads mapped to the host genome) for all identified NEAT1_2 circRNAs in total and 4sU-RNA time-courses of WT and Δ*vhs* infection and nucleoplasmic, chromatin-associated, cytoplasmic and total RNA for 8 h p.i. WT, Δ*vhs* and ΔICP27

infection. Columns represent individual circRNAs, which are ordered according to their genomic coordinates from the most 5' to the most 3'. The hsa_circ_0003812 NEAT1_2 circRNA is marked by a blue rectangle. **(b,c)** Sashimi plots showing NEAT1 read coverage (overlay of both replicates) and circular (red) and linear (same color as read coverage) splice junctions as arcs connecting acceptor and donor splice site in **(b)** selected samples of the total RNA time-courses of WT and Δ*vhs* infection and **(c)** chromatin-associated RNA in mock, WT, Δ*vhs* and ΔICP27 infection. Number of junction reads are annotated to arcs separately for the two replicates. Junctions are only shown if at least 4 reads align by at least 10 nt on both sides of the junction. Genomic coordinates of NEAT1 transcripts and the HSV-1-induced circRNA are shown at the bottom. Sashimi plots of the full total RNA and 4sU-RNA time-courses and nucleoplasmic and total RNA from the subcellular fractions experiment are shown in S13 and S15 Fig. **(d)** Heatmap of linear splicing read counts normalized to the number of exon-intron reads crossing the corresponding acceptor and donor splice sites for the WT and Δ*vhs* infection time-courses and the subcellular fraction experiment (dark red = high, white = low, see color bar on the right). Columns represent splice junctions, which are ordered according to their genome positions, first by the donor and then by the acceptor splice site. Counts include only junction and exon-intron reads overlapping the junction or exon-intron boundary, respectively, by at least 10 nt on either side. A pseudocount of 0.01 was used to avoid division by zero. Rectangles on top indicate whether the splice junction is within the genomic region or downstream of the hsa_circ_0003812 circRNA. Colors on the left indicate the virus and time-point of infection (light colors early in infection, darkest colors 8 h p.i, cyan = WT, blue = Δ*vhs*, gold = ΔICP27). Numbers in brackets at the bottom refer to the four most

frequent splice junctions outlined in **(e)**. **(e)** Schematic overview on the circular and the four most frequent linear NEAT1_2 splice junctions induced in HSV-1 infection and sequences around splice sites (| indicates exon-intron boundaries in sequences). The height of boxes indicates the relative abundance of each exon (= never or rarely spliced out) and intron (= frequently spliced out) by 8 h p.i. WT infection. Genomic coordinates on chromosome 11 are indicated for the start and end of exons. Numbers in brackets on top of introns indicate the ranking of corresponding splice junctions according to their splicing rate shown in **(d)**. This ranking is also indicated in **(d)** at the bottom.

The hsa_circ_0003812 NEAT1_2 circRNA was previously found to be enriched in RNA-seq data of human fibroblasts after treatment with the 3'->5' RNA exonuclease RNase R [46], confirming that it is a circRNA. Notably, however, NEAT1_2, is partially resistant to RNase R due to the stabilizing triple helical structure at its 3'end, similar to MALAT1 [47]. Thus, the linear NEAT1_2 transcript is not strongly depleted by RNase R treatment even with an improved protocol using A-tailing prior to RNase R treatment and a Li+ buffer (S12 Fig a) [47]. Nevertheless, we identified a study of circRNAs in Akata cells, in which the linear NEAT1_2 transcript was fully depleted by RNase R treatment in total and cytoplasmic RNA, yet partially resistant in nucleoplasmic RNA, while the hsa_circ_0003812 circRNA was always retained (S12 Fig b) [48]. Interestingly, we observed significant enrichment of the circRNA region in our total RNA-seq time-course of WT infection from 6 h p.i. compared to both up- and downstream regions (Fig 2b, S13 Fig a, 1.5- to 1.76-fold increase in the DEXSeq analysis, multiple testing adjusted $p<0.005$). This effect is likely not mediated by *vhs*

since NEAT1_2 is located in the nucleus while *vhs* is active in the cytosol. Furthermore, some enrichment of the circRNA region was also observed in Δ*vhs* infection by 8 h p.i., consistent with an attenuated progression of Δ*vhs* infection (S13 Fig b). Both NEAT1_1 and NEAT1_2 were found to be highly unstable in mouse, with half-lives of ~30 and ~60 min, respectively [49]. Thus, enrichment of the NEAT1_2 circRNA in HSV-1 infection is likely due to its high stability compared to the linear transcript combined with a high synthesis rate.

The following evidence further confirms that hsa_circ_0003812 is a circRNA and not an artefact of circRNA detection. First, it was identified by three distinct algorithms for circRNA detection that employ different approaches: CIRI2, circRNA_finder and our aligment-based approach. Combination of two or more different circRNA detection methods has been reported as highly successful in removing false positive predictions [39]. Second, since false positives in circRNA detection largely originate from repetitive sequences, we performed a BLAST search in the human and HSV-1 genomes for the 20 nt on either side of the circular junction. This confirmed that both sequences are unique. Third, we ran our alignment-based circRNA detection approach with a required overlap of at least 40 nt on either side of circular junctions on the WT 4sU-seq and total RNA-seq time-courses. This still identified up to 170 reads for the NEAT1_2 circRNA in total RNA and up to 44 in 4sU-RNA (17-36% of reads identified with a ≥10 nt overlap) and confirmed the strong induction of this circRNA (S14 Fig a,b). Finally, we implemented a method to determine so-called “confirming read pairs”, i.e. pairs of reads where both reads could be aligned linearly to the genome within a circRNA region but only in crosswise direction not consistent with a linear transcript (S14 Fig c). We selected those confirming read pairs for which their genomic distance would imply a fragment size ≥500 nt (observed on average for

<3.7% of fragments) if they originated from a linear transcript and the fragment size would be smaller if they originated from the circRNA. In this way, we identified up to 45 confirming read pairs per replicate for hsa_circ_0003812 in WT infection (S14 Fig b).

Strikingly, NEAT1_2 circular splicing was accompanied by induction of linear splicing both within the circRNA region as well as downstream of the circRNA (Fig 2b,c, S13 and S15 Fig). None of these linear splicing events were annotated, but some were already observed at low rates in uninfected cells. Normalization to the number of exon-intron reads crossing the corresponding acceptor and donor splice sites showed that linear splicing was indeed induced and not simply observed more frequently due to the previously reported up-regulation of NEAT1 transcription in HSV-1 infection [26, 28] (Fig 2d). Linear splicing was also induced downstream of the identified circRNAs (Fig 2b-d, S13 and S15 Fig) and several of the confirming read pairs for the hsa_circ_0003812 circRNA included these linear splice junctions. Thus, both linear and circular NEAT1_2 transcripts were further spliced in HSV-1 infection. Canonical GT-AG splice signals were used both for the circular and the most frequent linear splice junctions (see Fig 2e for a diagram) and they tended to be conserved at least among primates (S16 Fig). Splicing rates and enrichment compared to mock differed substantially between the different linear splice junctions within the circRNA region (Fig 2d), indicating presence of alternative splicing isoforms of the circRNA. Induction of linear splicing was also observed in 4sU-, chromatin-associated and nucleoplasmic RNA and in ΔICP27 and Δ*vhs* infection (Fig 2c,d, S15 Fig), although it was less pronounced in ΔICP27 infection. In summary, both linear and circular splicing of NEAT1_2 is induced during HSV-1 infection while

the RNA is still associated with the chromatin. Although ICP27 is not required for NEAT1_2 splicing, it may contribute.

Recently, we also performed total RNA-seq of infection with mock, HSV-1 strain F (WT-F) and its ΔICP22 mutant at 8 and 12 h p.i. (n=2) [50]. We furthermore generated telomerase-immortalized human foreskin fibroblasts (T-HFs) that express either ICP22 in isolation (T-HFs-ICP22 cells) or in combination with ICP27 (T-HFs-ICP22/ICP27 cells) upon doxycyclin (Dox) exposure. Total RNA-seq of T-HFs-ICP22 and T-HFs-ICP22/ICP27 cells was performed both with and without Dox exposure (n=2). Strikingly, we observed that ectopic co-expression of ICP22 and ICP27 lead to induction of both circular and linear splicing of NEAT1_2 (Fig 3a). However, ICP22 knockout did not abolish HSV-1-induced NEAT1_2 circular and linear splicing (Fig 3a, S17 Fig), thus ICP22 expression is not required. This also confirmed induction of linear and circular NEAT1_2 splicing in a different HSV-1 strain. Although circRNA reads were recovered after Dox-induced expression of ICP22, this was limited to 2 reads in one replicate (Fig 3b).

Fig 3. Ectopic co-expression of ICP22 and ICP27 is sufficient for induction of NEAT1_2 splicing.

Sashimi plots as in Fig 2 showing read coverage and circular and linear splice junctions in **(a)** T-HFs-ICP22/ICP27 cells without and with Dox-induced co-expression of ICP22 and ICP27 as well as mock, WT-F and ΔICP22 infection at 12 h p.i., **(b)** T-HFs-ICP22 cells without and with Dox-induced ICP22 expression and **(c)** mock and HSV-1 strain KOS infection as well as ICP27 overexpression. A minimum read count of 4 was again required for **(a)**, while in **(b)** and **(c)** all junction reads are shown covering at least 10nt on either side of the junction. For an explanation of

sashimi plots, see caption to Fig 2. Circular splice junctions are marked in red, linear splice junctions in the same color as read coverage.

To investigate whether ICP27 expression alone could induce NEAT1_2 splicing, we analyzed our recently published 4sU-seq data of mock and HSV-1 strain KOS infection and of cells transfected with an ICP27-expressing plasmid [31]. This confirmed induction of both linear and circular NEAT1_2 splicing in HSV-1 KOS infection but recovered only one circular read upon ICP27 overexpression in one replicate (Fig 3c). In summary, our results show that neither ICP27 nor ICP22 are required for induction of linear and circular NEAT1_2 splicing but co-expression of both proteins is sufficient. Considering the small number of circular reads recovered when expressing either one of these proteins alone, we can neither confidently confirm nor exclude that either ICP22 or ICP27 alone may be sufficient to (weakly) induce NEAT1_2 circular splicing.

Influenza A virus infection also induces NEAT1_2 splicing

Since NEAT1_2 is also up-regulated in IAV infection [26], we investigated presence of circRNAs in previously published total RNA-seq time-courses of influenza A/California/04/09 (H1N1), A/Wyoming/03/03 (H3N2), and A/Vietnam/1203/04 (H5N1) HALo virus infection (3, 6, 12 and 18 h p.i. plus time-matched controls) of human tracheobronchial epithelial cells (HTBE, multiplicity of infection (MOI) = 5) and monocyte-derived macrophages (MDM, MOI = 2) [37]. The H5N1 HALo mutant virus is an attenuated virus generated from WT influenza A/Vietnam/1203/04. H1N1 infection of HTBE cells included 24 h p.i. and two H5N1 infection time-courses of MDM cells were performed. This revealed a strong induction of NEAT1_2 circular splicing in H5N1 infection of both cell types and H3N2 infection of MDM cells as well

as a weak induction in H3N2 infection of HTBE cells and H1N1 infection of MDM cells (Fig 4a, S18 Fig). The most highly expressed circRNA was again hsa_circ_0003812 circRNA as for HSV-1 infection (marked by a blue rectangle in Fig 4a) and its abundance increased with the duration of infection. Notably, hsa_circ_0003812 circRNA was also detected in some of the time-matched controls, however at significantly lower levels than in infected cells (with the exception of H1N1 infection of HTBE cells, S18 Fig a). Similar trends were observed for induction of linear splicing events, both within and downstream of the circRNA region (Fig 4b).

Fig 4. Induction of NEAT1_2 circular and linear splicing in IAV infection.

(a) Heatmap of normalized circRNA counts (normalized to number of linear splice junctions mapped to the host genome) for the three most highly expressed NEAT1_2 circRNAs identified in H1N1, H3N2 and H5N1 infection of HTBE and MDM cells or time-matched controls. Columns represent individual circRNAs, which are ordered according to their genomic coordinates from the most 5' to the most 3'. The NEAT1_2 hsa_circ_0003812 also induced in HSV-1 infection is marked by a blue rectangle. IAV strain and time-points of infection are color-coded on the left (white=time-matched control, cyan=H1N1, green=H3N2, purple=H5N1, darker colors indicate later time-points). **(b)** Heatmap of linear splicing read counts normalized to the number of exon-intron reads crossing the corresponding acceptor and donor splice sites in IAV infection and time-matched controls (dark red = high, white = low, see color bar on the right). Columns represent individual splice junctions, which are ordered according to their genome positions. Counts include only junction and exon-intron reads overlapping the junction or exon-intron boundary, respectively, by at least 10 nt on either side. A pseudocount of 0.01 was used to avoid division by zero.

Colors on the left indicate time-points of infection as in **(a)**. Colored rectangles on top indicate whether linear splice junctions were also observed in HSV-1 infection (cyan) or are within (blue) or downstream (gray) of the hsa_circ_0003812 circRNA region. Numbers in brackets at the bottom refer to the introns outlined in the schematic in Fig 2e.

IAV infection also induces the degradation of host RNAs via two mechanisms. Cap-snatching by the viral RNA-dependent RNA polymerase leads to cleavage of host RNAs downstream of the cap and is required for synthesis of viral mRNAs [51]. The PA-X protein produced by all IAV strains [52] selectively degrades RNAs transcribed by Pol II, including long non-coding RNAs, and is localized and acts predominantly in the nucleus [53]. To exclude that induction of the NEAT1_2 circRNA is simply due to degradation of linear transcripts, we also investigated overall accumulation of circRNAs in IAV infections using linear regression analysis on all circRNAs. This indeed showed an accumulation of circRNAs by up to 1.4-fold in H1N1 infection, 6.8-fold in H3N2 infection and 4.6-fold for H5N1 infection (S19 Fig a-g). It also revealed induction beyond this overall accumulation for several other circRNAs, such as a circRNA of the ZC3HAV1 (zinc finger CCCH-type containing, antiviral 1) gene. ZC3HAV1 is induced during influenza A/WSN/33 (WSN/H1N1) infection and ectopic expression of ZC3HAV1 inhibits WSN replication [54]. Induction of the ZC3HAV1 circRNA thus likely results from up-regulation of ZC3HAV1 transcription and associated increased circRNA biosynthesis and is not simply a consequence of host RNA degradation. We compared fold-changes in normalized circRNA counts between infection and time-matched controls for 7077 well-expressed circRNAs

(normalized circRNA count >1 in at least one condition), which showed that the hsa_circ_0003812 NEAT1_2 circRNA was among the 1% most increased circRNAs in H3N2 and H5N1 infection of MDM cells (S19 Fig h). Occasionally, it was even more up-regulated than the ZC3HAV1 circRNA. Moreover, splicing of linear NEAT1_2 transcripts, which should not escape host mRNA degradation in IAV infections, is also induced during infection (Fig 4b). Thus, our results indicate that – at least in H3N2 and H5N1 infection – induction of NEAT1_2 circular splicing is not only a consequence of IAV-induced degradation of host mRNAs.

Since both HSV-1 and IAV infection disrupt transcription termination, we also evaluated read-through for IAV infection to correlate this to the extent of induction of NEAT1_2 splicing. Paralleling NEAT1_2 induction, read-through was higher in MDM cells than in HTBE cells and higher in H5N1 infection than in H3N2 infection (S20 Fig), while H1N1 infection led to the lowest extent of read-through. Disruption of transcription termination in IAV infection is mediated by the viral NS1 protein via inhibition of the CPSF30 subunit of the cleavage and polyadenylation specificity factor (CPSF) [55]. Our results are consistent with previous reports that the NS1 protein of A/California/04/09 H1N1 virus cannot bind and inhibit CPSF30 [56], while NS1 of the H3N2 and H5N1 strains can [57, 58]. However, a recent report indicated that the NS1-CPSF30 interaction is not necessary for disruption of transcription termination in IAV infection [36]. This could explain why read-through is still observed in H1N1 infection. A similar observation can be made in HSV-1 infection. Although HSV-1-induced disruption of transcription termination is mediated by ICP27 via interaction with CPSF subunits [31], reduced but significant read-through is also observed in ΔICP27 infection likely as a stress response [18, 35]. Due to this correlation between induction of NEAT1_2 splicing and read-through, we also

analyzed our previously published 4sU-seq data of 1 and 2 h salt and heat stress regarding NEAT1_2 splicing. However, neither circular nor linear NEAT1_2 splicing was induced in either of the two stress conditions.

CDK7 inhibition induces NEAT1_2 splicing and disrupts transcription termination in cancer cell lines

To shed additional light on the molecular mechanism(s) governing NEAT1_2 splicing, we searched the recount3 database [59] using Snaptron [60] for presence of the novel linear NEAT1_2 splicing events. Recount3 provides read coverage data for >700,000 publicly available human and mouse RNA-seq samples from the Sequence Read Archive (SRA). Snaptron allows rapidly searching samples in recount3 for (linear) splice junctions within specific genomic regions. Our Snaptron query recovered 6,391 human samples in which at least one of the four most frequent NEAT1_2 linear splice junctions (see Fig 2e) induced in HSV-1 infection was covered by ≥10 reads. This included our HSV-1 RNA-seq experiments, several other studies on HSV-1 infection, the IAV infection time-courses, 62 studies focusing on circRNAs and 21 using RNase R (S21 Fig a). The latter provides further support for the association of linear and circular splicing of NEAT1_2. Strikingly, particularly high abundance of NEAT1_2 splice junctions was found in samples of blood cells, such as platelets, erythrocytes, peripheral blood mononuclear cells (PBMCs), erythroleukemia (K562) cells, and more (S21 Fig a). Consistently, one of these studies reported that circRNAs are strongly enriched in platelets and erythrocytes [61], which both lack a cell nucleus and thus *de novo* transcription. In absence of *de novo* transcription, the high stability of circRNAs leads to their accumulation.

Application of our circRNA alignment pipeline to two RNA-seq experiments for platelets [62] and erythrocytes [63] identified multiple NEAT1_2 circRNAs including the hsa_circ_0003812 circRNA induced by HSV-1 infection (S21 Fig b). Interestingly, almost no expression was detected in the ~4 kb between the 3'end of the NEAT1_1 transcript and the 5'end of hsa_circ_0003812, however, expression extended beyond hsa_circ_0003812. This either implied novel linear transcripts beginning near the 5'end of hsa_circ_0003812 or a novel circRNA not included in circBase. Indeed, our *de novo* circRNA detection pipeline identified a novel circular junction connecting the 5' end of hsa_circ_0003812 with the 3'end of the downstream expressed region. This confirms that only circular, but not linear, NEAT1_2 transcripts were present in these cells and that these circRNAs are also linearly spliced. It also shows that NEAT1_2 circRNAs naturally occur in uninfected cells, however at such low levels that they only become detectable upon degradation of linear transcripts. Since NEAT1_2 has been shown to be up-regulated in HSV-1 and IAV infection by RT-qPCR [26], induction of NEAT1_2 circular splicing in HSV-1 and IAV infection cannot be explained by loss of *de novo* transcription followed by degradation of linear transcripts. However, this partly explains the presence of NEAT1_2 circular and linear junctions in uninfected cells from the IAV infection time-courses as MDM cells were obtained from blood.

Interestingly, the Snaptron search also identified induction of NEAT1_2 linear splicing after inhibition of CDK7 by THZ1 in multiple cancer cell lines. CDK7 is an essential part of the TFIIH transcription factor complex, phosphorylates Pol II CTD at Ser5 residues and plays key roles in maintenance of promoter-proximal Poll II pausing and the regulation of transcription elongation [64]. Overexpression of CDK7 has been associated with a number of cancers and correlated to poor prognosis [64]. THZ1 is a

highly selective covalent inhibitor of CDK7 with anti-tumor activity against multiple different types of cancer [65], but also inhibits CDK12 and CDK13 at higher concentrations [66]. Experiments identified with the Snaptron search cover prostate cancer (VCaP, LNCaP and DU145 cells) [38], bladder cancer (HCV-29 cells) [67], esophageal squamous cell carcinoma (TE7 and KYSE510 cells) [68], nasopharyngeal carcinoma (C666-1, HK1 and HNE1 cells) [69], pancreatic ductal adenocarcinoma (BxPC3, MiaPaCa-2 and PANC1 cells) [70] and chordoma (UM-Chor1 cells) [71]. Further literature search identified a second study of chordoma cells (UM-Chor1 and CH22) [72] with up to 24 h THZ1 treatment and a study on THZ1 treatment of B-cell acute lymphocytic leukemia cells (Nalm6) [73]. Realignment of corresponding samples confirmed linear splicing in all cell lines apart from DU145 and showed that induction increased with the duration and dosage of THZ1 treatment (Fig 5a,b, S22 Fig a). One UM-Chor1 experiment also included actinomycin D (Act-D) treatment, which inhibits transcription but not via inhibition of cyclin-dependent kinases (CDKs) [74]. Act-D treatment, while leading to a similar relative increase in the abundance of NEAT1_2 compared to NEAT1_1 as THZ1 treatment, did not induce NEAT1_2 splicing (S22 Fig a). In addition, siRNA-mediated knockdown of CDK7 in VCaP and LNCaP cells also induced NEAT1_2 splicing (Fig 5c), including many more novel splicing events that extended into the NEAT1_1 region. This indicates that indeed inhibition of CDK7, and not of other CDKs, induces NEAT1_2 splicing. Interestingly, siRNA-mediated knockdown of the MED1 component of the Mediator complex in VCaP and LNCaP cells had the same effect (Fig 5c). CDK7 directly phosphorylates MED1 at Threonine 1457 [38], which together with a phosphorylation at Threonine 1032 promotes MED1 association with the Mediator complex [75]. Mediator is responsible for communicating regulatory signals from gene-specific transcription factors to Pol II and in this way impacts transcription at

multiple stages [76]. Our results suggest that the effect of CDK7 inhibition on NEAT1_2 splicing may follow from the loss of MED1 phosphorylation and thus loss of MED1 association with Mediator.

Fig 5. Induction of NEAT1_2 splicing and read-through upon CDK7 inhibition. **(b)** Heatmap of linear splicing read counts normalized to the number of exon-intron reads crossing the corresponding acceptor and donor splice sites upon treatment by DMSO/control, THZ1 or Act-D or knockdown (KD) of CDK7 or MED1 (dark red = high, white = low, see color bar on the right). Columns represent individual splice junctions, which are ordered according to their genome positions. Counts include only junction and exon-intron reads overlapping the junction or exon-intron boundary, respectively, by at least 10 nt on either side. A pseudocount of 0.01 was used to avoid division by zero. Colored rectangles on top indicate whether linear splice junctions were also observed in HSV-1 (cyan) or IAV (purple) infection or are within (blue) or downstream (gray) of the hsa_circ_0003812 circRNA region. Numbers in brackets at the bottom refer to the introns outlined in the schematic in Fig 2e. **(b,c)** Sashimi plots as in Fig 2 showing read coverage and linear splice junctions in **(b)** UM-Chor1 cells after 4, 8, 12 and 24 h THZ1 or control treatment (3 replicates each) and **(c)** CDK7 and MED1 knockdown in VCaP and LNCaP cells (2 replicates each). All junctions are shown that are covered by at least 4 reads aligning to at least 10 nt on either side of the junction. **(d)** Boxplots showing distribution of %read-through (for calculation see methods) after THZ1 treatment in example studies or CDK7 or MED1 knockdown in VCaP and LNCaP cells.

Surprisingly, considering the previously observed strong association between linear and circular NEAT1_2 splicing, no NEAT1_2 circRNAs were recovered for any of these samples neither with our alignment-based nor our *de novo* circRNA discovery pipeline. However, in all cases almost no circRNAs (maximum 105 circRNAs, mean 18) were recovered at all, not even the most ubiquitous circRNAs such as HIPK3, CORO1C or ASXL1. As read length was ≥75 nt for all samples, this cannot be explained by short read length. It rather suggests that circRNAs were generally not or only poorly recovered in all these studies. While a literature search identified no other studies for specific CDK7 inhibition without poly(A) selection, the Snaptron search also recovered induction of linear NEAT1_2 splicing in chromatin-associated RNA of HEK293 cells upon 2 h treatment with DRB [77]. DRB inhibits CDK7 [78] but also the CDK9 kinase component of the positive P-TEFb complex [79] and other kinases [80]. DRB treatment indeed led to induction of the hsa_circ_0003812 NEAT1_2 circRNA induced by HSV-1 and IAV infection, but not any of the other NEAT1_2 circRNAs present in platelets and erythrocytes (S22 Fig b). This provides some evidence that CDK7 inhibition induces NEAT1_2 circular splicing. In any case, identification of linear splicing in absence of circRNAs shows that the linear NEAT1_2 isoform is spliced upon CDK7 inhibition.

Interestingly, CDK7 inhibition also induced read-through transcription albeit to a lesser degree than HSV-1 or H5N1 infection (Fig 5d, example genes in S23 Fig). Read-through transcription after THZ1 treatment was comparable to H1N1 or H3N2 infection of HTBE cells, most pronounced in THZ1-treated Nalm6 cells and increased with the duration and dosage of THZ1 treatment in UM-Chor1 cells. In addition, both CDK7 and MED1 knockdown induced read-through at levels comparable to THZ1-treated Nalm6 cells. Interestingly, THZ1 treatment also resulted in extended

antisense transcription at promoters, explaining the upstream transcription seen in the non-strand-specific RNA-seq for Nalm6 cells after THZ1 treatment (S23 Fig c,d). In summary, our results show that CDK7 inhibition impacts RNA transcription and processing at multiple levels.

Discussion

In this study, we report on novel circular and linear splicing of the long NEAT1_2 isoform induced in both HSV-1 and IAV infection. We focused on NEAT1_2 in this study for three reasons: (i) The NEAT1_2 circRNA was the only circRNA not simply enriched by *vhs*-mediated degradation of linear RNAs. (ii) It is one of the few circRNAs observed at high levels in newly transcribed 4sU-, nucleoplasmic and chromatin-associated RNA during HSV-1 infection, indicating high levels of *de novo* synthesis. In contrast to these other few circRNAs, however, it is not or rarely produced in uninfected cells. (iii) NEAT1_2 has not previously been documented to be spliced either in a circular or linear fashion except for one linear splicing event towards the 3'end of NEAT1_2. This splicing event represented a shorter 793 nt version of the last "intron" in Fig 2e with the same 3' but a different 5' splice site and was recently reported to increase extractability of NEAT1_2 from the protein phase [81]. Although it was also induced in HSV-1 and IAV infection, it was not among the most frequent splicing events. Re-alignment of the RNA-seq data by Chujo *et al.* confirmed this splicing event as well as splicing of the full last "intron" with very few reads (≤3), but no other splicing.

As NEAT1_2 participates in many RNA-protein interactions in paraspeckles, its unspliced form can only be poorly extracted from the protein phase [81]. Similar observations were made for other architectural RNAs (arcRNAs) of nuclear bodies.

As genomic deletion of the 793 nt intron identified by Chujo *et al.* did not significantly affect extractability, splicing itself appeared to be responsible for increased extractability. Chujo *et al.* concluded that NEAT1_2 as well as other arcRNAs must remain unspliced to allow formation of corresponding nuclear bodies. They hypothesized that competition by protein components of paraspeckles for binding of NEAT1_2 normally prevents association of splicing factors with NEAT1_2 and thus NEAT1_2 splicing. This raises the possibility that NEAT1_2 splicing during HSV-1 and IAV infection may simply be a by-product of up-regulation and thus increased availability of NEAT1_2 and consequently reduced competition between splicing factors and paraspeckle protein components. In HSV-1 infection, NEAT1_2 binds to serine and arginine rich splicing factor 2 (SRSF2) and is required for association of SRSF2 with promoters of the HSV-1 genes ICP0 and thymidine kinase (TK), which are upregulated by SRSF2 [82]. However, the SRSF2 binding sites in NEAT1_2 that are enriched upon HSV-1 infection are upstream of the HSV-1-induced circRNA. Thus, binding between SRSF2 and NEAT1_2 is likely not responsible for NEAT1_2 splicing. Moreover, this means that the linear NEAT1_2 transcript is required for SRSF2 association.

We demonstrated that the two immediate-early HSV-1 proteins ICP22 and ICP27 were sufficient to induce NEAT1_2 circular and linear splicing. However, neither of these two proteins were required nor sufficient on their own for (significant) induction of NEAT1_2 splicing. A possible explanation for this observation could be that NEAT1 was upregulated upon ectopic co-expression of ICP22 and ICP27 (2.2-fold, $p<10^{-11}$) and in ΔICP22 (2.1-fold, $p<10^{-4}$) and ΔICP27 (4.9-fold, $p<10^{-12}$) infection, but not upon expression of either ICP22 or ICP27 alone. Thus, up-regulation of NEAT1_2 could be the trigger for NEAT1_2 splicing rather than specific functions of

ICP22 and ICP27. One line of evidence arguing against NEAT1_2 splicing being only a by-product of NEAT1_2 up-regulation is the absence of NEAT1_2 circular or linear splicing in heat stress despite NEAT1_2 also being up-regulated in the heat shock response [83].

Previously, NEAT1_2 knockdown in HeLa cells was shown to reduce HSV-1 glycoprotein density and intensity and inhibit plaque formation, an indicator of mature virus production [28]. The vital role of NEAT1_2 for HSV-1 infection was also confirmed *in vivo* as a thermosensitive gel containing NEAT1_2 siRNA could heal HSV-1-induced skin lesions in mice [28]. This siRNA targeted the first “exon” within the circRNA (see Fig 2e), which is contained in both linear and circular spliced and unspliced NEAT1_2 transcripts. Interestingly, eCLIP data from ENCODE for the two core paraspeckle proteins SFPQ and NONO shows enriched binding of these proteins in NEAT1_2 within the two most frequently spliced “introns” (1 and 2 in Fig 2e), the “exon” between them, upstream of the circRNA 5’ end and towards the 3’end of NEAT1_2, but not in the first two “exons” and “intron” 3. Since NEAT1_2 circular splicing destroys linear transcripts and splicing of the linear NEAT1_2 transcript removes parts important for binding of paraspeckle proteins, up-regulation of NEAT1_2 splicing by unknown mechanisms could thus also be part of the immune response against HSV-1. Finally, SFPQ and NONO have recently been reported to bind around circRNA loci, and SFPQ depletion led to reduced expression of a subset of circRNAs [84]. Thus, these paraspeckle proteins itself could be involved in NEAT1_2 splicing.

Linear NEAT1_2 splicing, but likely also circular splicing, is also induced upon CDK7 inhibition in cancer cell lines and NEAT1_2 circRNAs are abundantly observed in various blood cells in absence of *de novo* transcription. However, the induction of

NEAT1_2 splicing in HSV-1 and IAV infection and upon CDK7 inhibition is not simply due to loss of transcriptional activity as NEAT1 is up-regulated in HSV-1 and IAV infection and transcription inhibition by Act-D does not induce NEAT1_2 splicing. Interestingly, we found that selective CDK7 inhibition by THZ1 also disrupts transcription termination in cancer cell lines, albeit less strongly than HSV-1 and H5N1 infection. This was surprising considering a recent report that 1 h THZ1 treatment suppresses Pol II read-through at gene 3'ends in acute myeloid leukemia [85]. It is, however, consistent with earlier reports of impaired transcription termination and 3′-end processing of an snRNA, a polyadenylated mRNA and a histone RNA upon CDK7 inhibition [86, 87]. Our results thus demonstrate for the first time widespread disruption of transcription termination upon CDK7 inhibition. A possible explanation for the discrepancy to the study by Sampathi *et al.* is that more extensive CDK7 inhibition is necessary to sufficiently disrupt transcription termination as we observed significant read-through only after several hours of THZ1 treatment or CDK7 knockdown. Interestingly, CDK7 inhibition also leads to increased expression of the non-polyadenylated NEAT1_2 isoform relative to the polyadenylated NEAT1_1 isoform (Fig 5b,c, S22 Fig a). This was also observed in HSV-1 and IAV infection, although not consistently, and raises the possibility that read-through transcription of the NEAT1_1 poly(A) site may be involved in induction of NEAT1_2 splicing by increasing abundance of NEAT1_2. Although salt and heat stress did not induce NEAT1_2 splicing, this does not exclude a role of read-through transcription. While there are strong overlaps between the genes affected by read-through in HSV-1 infection and stress conditions, there are also clear context- and condition-specific differences [18]. Furthermore, the mechanisms underlying disruption of transcription termination in stress conditions remain elusive. This contrasts with HSV-1 and IAV infection, where the HSV-1 ICP27 protein and the IAV

NS1 protein have been shown to disrupt transcription termination via interaction with CPSF subunits [31, 37, 55]. It is important to note that our analyses of NEAT1_2 splicing and read-through upon CDK7 inhibition were performed in cancer cell lines, which are particularly susceptible to CDK7 inhibition [65]. Dysregulation of alternative polyadenylation, often associated with wide-spread shortening of 3'UTRs, is common in cancers [88]. NEAT1 also plays different roles in cancer development and both NEAT1 isoforms have been proposed as cancer biomarkers [89], with NEAT1_2 considered a tumor suppressor and NEAT1_1 considered to be oncogenic [90]. Thus, read-through, relative increase of NEAT1_2 expression, or NEAT1_2 splicing upon CDK7 inhibition could be linked to susceptibility of cancer cell lines to CDK7 inhibition.

This study identified several parallels between HSV-1 and IAV infection, on the one hand, and CDK7 inhibition, on the other hand. This not only includes read-through transcription and NEAT1_2 splicing, but also induction of antisense transcription upon CDK7 inhibition. We previously reported on widespread activation of antisense transcription in HSV-1 infection [91]. Moreover, widespread loss of promoter-proximal pausing was recently reported both upon CDK7 inhibition [85, 92] and HSV-1 infection [93]. Previously, LDC4297A, a different selective inhibitor of CDK7, has been shown to have strong antiviral activity against HSV-1 as well as other herpesviruses (but only low efficacy against H1N1) [94] and CDK7 has been found to be a target of CDK inhibitors inhibiting HSV replication [95]. Both studies confirm a role of CDK7 in HSV-1 infection. It is tempting to speculate that recruitment of CDK7-containing complexes, similar to recruitment of elongation factors like FACT by ICP22 [96], to the HSV-1 genome may lead to a depletion of CDK7 on host genes, thus mimicking the effects of CDK7 inhibition. In summary, our findings highlight

potential important roles of NEAT1_2 splicing and CDK7 in HSV-1 and/or IAV infection.

Materials and Methods

RNA-seq data

All RNA-seq data analyzed in this study were downloaded from the SRA. SRA project IDs: HSV-1 WT infection time-course: SRP044766; HSV-1 Δ*vhs* infection time-course: SRP192356; RNA-seq of subcellular fractions: SRP110623, SRP189489, SRP191795 (same experiment but data for mutant viruses submitted separately); T-HFs-ICP22 and T-HFs-ICP22/ICP27 cells, WT strain F and ΔICP22 infection: SRP340110; WT strain KOS, ΔICP27 and ICP27 overexpression: SRP189262; IAV infection time-courses: SRP091886, SRP103821; platelets: ERP003815; erythrocytes: SRP050333; CDK7 inhibition/knockdown: VCaP, LNCaP and DU145 cells: SRP179971; HCV-29 cells: SRP217721; TE7 and KYSE510 cells: SRP068450; C666-1, HK1 and HNE1 cells: SRP101458; BxPC3, MiaPaCa-2 and PANC1 cells: SRP165924; UM-Chor1 and CH22 cells: SRP166943, SRP270819; Nalm6 cells: SRP307127; DRB treatment of HEK293 cells: SRP055770. RNase R treatment: SRP197110, SRP152310.

RNA-seq read mapping

Sequencing reads were downloaded from SRA using the sratoolkit version 2.10.8 and aligned against the human genome (GRCh37/hg19) and human rRNA sequences using ContextMap2 version 2.7.9 [97] (using BWA as short read aligner [98] and allowing a maximum indel size of 3 and at most 5 mismatches). For HSV-1

infection RNA-seq data, alignment also included the HSV-1 genome (Human herpesvirus 1 strain 17, GenBank accession code: JN555585). For the two repeat regions in the HSV-1 genome, only one copy was retained each, excluding nucleotides 1–9,213 and 145,590–152,222 from the alignment.

circRNA de novo detection

For *de novo* detection of circRNAs, we applied CIRI2 [41] and circRNA_finder [40] in parallel to each sample. Both allow *de novo* detection of circRNAs using splice sites not annotated in the human genome. Only circRNAs were further analyzed that were detected independently by both algorithms. Circular reads identified by either algorithm were pooled and subsequently all reads were removed that could be mapped in a linear fashion to the genome using ContextMap2. Only circRNAs with at least two supporting reads in the same sample were further analyzed. The complete pipeline is outlined in S1 Fig b.

Alignment-based circRNA detection

Candidate circular junction sequences were generated from putative spliced circRNAs sequences downloaded from circBase [43]. RNA-seq reads were aligned against circular junction sequences using BWA [98] and reads were retained if they aligned to at least X nt (default X = 10) on either side of the junction and could not be mapped in a linear fashion to the genome using ContextMap2. In case of paired-end sequencing, the second read in the pair also had to align consistently within the circRNA region. circRNAs resulting from repetitive regions were discarded. The complete pipeline is outlined in S5 Fig.

Identification of confirming read pairs

ContextMap2 is also capable of aligning read pairs, where each individual read can be aligned linearly to the genome, but the two reads are aligned in the wrong

orientation relative to each other (=crosswise, see S14 Fig). Such crosswise alignments of read pairs are inconsistent with linear transcripts and were considered a confirming read pair for a circRNA if (i) both reads were aligned within the circRNA region with a genomic distance implying a fragment size ≥500 nt if they originated from a linear transcript and (ii) the fragment size would be smaller if they originated from the circRNA.

Read-through calculation

Number of read fragments per gene and in 5kb windows downstream of genes were determined from read alignments using featureCounts [99] and gene annotations from Ensembl (version 87 for GRCh37/hg19) [100]. For strand-specific RNA-seq, the stranded mode of featureCounts was used. All fragments (read pairs for paired-end sequencing or reads for single-end sequencing) overlapping exonic regions on the corresponding strand by ≥25bp were counted for the corresponding gene. Downstream transcription for a gene was calculated as previously described [18] as the FPKM (Fragments Per Kilobase Million) in the 5kb windows downstream of genes divided by the gene FPKM multiplied by 100. %Read-through transcription was quantified as the difference in downstream transcription between infected/treated and uninfected/control cells, with negative values set to zero.

Other bioinformatics analyses

Sashimi plots were generated with ggsashimi [101]. Heatmaps were generated with the heatmap.2 and pheatmap functions in R. Differential exon usage analysis was performed with the R Bioconductor package DEXSeq [102]. Log2 fold-changes in gene expression for WT and Δ*vhs* infection time-courses were obtained from our

recent study [19]. In this study, we used DESeq2 [103] to compare gene expression between mock and each time-point of infection for 4,162 genes without read-in transcription originating from disrupted transcription termination for an upstream gene. Genes with read-in transcription were excluded from the analysis as read-in transcription can be mistaken for induction of gene expression.

Acknowledgements

This work was supported by the Deutsche Forschungsgemeinschaft (DFG) (FR2938/10-1 to C.C.F. and LD1275/6-1 to L.D.) and by the European Research Council (ERC-2016-CoG 721016 – HERPES to L.D.).

References

1. Roizman B, Knipe DM, R.J. W. Herpes simplex viruses. In Knipe D M, Howley P M (ed), Fields virology, 5th ed Lippincott Williams & Wilkins, Philadelphia, PA. 2007:2501-601.

2. Ablashi D, Agut H, Alvarez-Lafuente R, Clark DA, Dewhurst S, DiLuca D, et al. Classification of HHV-6A and HHV-6B as distinct viruses. Archives of Virology. 2014;159(5):863-70.

3. Kwong AD, Frenkel N. Herpes simplex virus-infected cells contain a function(s) that destabilizes both host and viral mRNAs. Proc Natl Acad Sci U S A. 1987;84(7):1926-30. Epub 1987/04/01.

4. Oroskar AA, Read GS. Control of mRNA stability by the virion host shutoff function of herpes simplex virus. J Virol. 1989;63(5):1897-906. Epub 1989/05/01.

5. Feng P, Everly DN, Jr., Read GS. mRNA decay during herpesvirus infections: interaction between a putative viral nuclease and a cellular translation factor. J Virol. 2001;75(21):10272-80. Epub 2001/10/03.

6. Doepker RC, Hsu WL, Saffran HA, Smiley JR. Herpes simplex virus virion host shutoff protein is stimulated by translation initiation factors eIF4B and eIF4H. J Virol. 2004;78(9):4684-99. Epub 2004/04/14.

7. Sarma N, Agarwal D, Shiflett LA, Read GS. Small interfering RNAs that deplete the cellular translation factor eIF4H impede mRNA degradation by the virion host shutoff protein of herpes simplex virus. J Virol. 2008;82(13):6600-9. Epub 2008/05/02.

8. Page HG, Read GS. The virion host shutoff endonuclease (UL41) of herpes simplex virus interacts with the cellular cap-binding complex eIF4F. J Virol. 2010;84(13):6886-90. Epub 2010/04/30.

9. Shiflett LA, Read GS. mRNA Decay during Herpes Simplex Virus (HSV) Infections: Mutations That Affect Translation of an mRNA Influence the Sites at Which It Is Cleaved by the HSV Virion Host Shutoff (Vhs) Protein. Journal of Virology. 2013;87(1):94-109.

10. Vicens Q, Westhof E. Biogenesis of Circular RNAs. Cell. 2014;159(1):13-4.

11. Nigro JM, Cho KR, Fearon ER, Kern SE, Ruppert JM, Oliner JD, et al. Scrambled exons. Cell. 1991;64(3):607-13.

12. Salzman J, Gawad C, Wang PL, Lacayo N, Brown PO. Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types. PloS one. 2012;7(2):e30733. Epub 2012/02/10.

13. Wilusz JE. A 360° view of circular RNAs: From biogenesis to functions. Wiley Interdisciplinary Reviews: RNA. 2018;9(4):e1478.

14. Hansen TB, Jensen TI, Clausen BH, Bramsen JB, Finsen B, Damgaard CK, et al. Natural RNA circles function as efficient microRNA sponges. Nature. 2013;495:384.

15. Memczak S, Jens M, Elefsinioti A, Torti F, Krueger J, Rybak A, et al. Circular RNAs are a large class of animal RNAs with regulatory potency. Nature. 2013;495(7441):333-8. Epub 2013/03/01.

16. Shi J, Hu N, Mo L, Zeng Z, Sun J, Hu Y. Deep RNA Sequencing Reveals a Repertoire of Human Fibroblast Circular RNAs Associated with Cellular Responses to Herpes Simplex Virus 1 Infection. Cellular Physiology and Biochemistry. 2018;47(5):2031-45.

17. Rutkowski AJ, Erhard F, L'Hernault A, Bonfert T, Schilhabel M, Crump C, et al. Widespread disruption of host transcription termination in HSV-1 infection. Nature communications. 2015;6:7126. Epub 2015/05/21.

18. Hennig T, Michalski M, Rutkowski AJ, Djakovic L, Whisnant AW, Friedl MS, et al. HSV-1-induced disruption of transcription termination resembles a cellular stress response but selectively increases chromatin accessibility downstream of genes. PLoS pathogens. 2018;14(3):e1006954. Epub 2018/03/27.

19. Friedel CC, Whisnant AW, Djakovic L, Rutkowski AJ, Friedl MS, Kluge M, et al. Dissecting Herpes Simplex Virus 1-Induced Host Shutoff at the RNA Level. J Virol. 2021;95(3). Epub 2020/11/06.

20. Sasaki YTF, Ideue T, Sano M, Mituyama T, Hirose T. MENε/β noncoding RNAs are essential for structural integrity of nuclear paraspeckles. Proceedings of the National Academy of Sciences. 2009;106(8):2525-30.

21. Nakagawa S, Naganuma T, Shioi G, Hirose T. Paraspeckles are subpopulation-specific nuclear bodies that are not essential in mice. The Journal of Cell Biology. 2011;193(1):31-9.

22. Wilusz JE, JnBaptiste CK, Lu LY, Kuhn C-D, Joshua-Tor L, Sharp PA. A triple helix stabilizes the 3′ ends of long noncoding RNAs that lack poly(A) tails. Genes & Development. 2012;26(21):2392-407.

23. Clemson CM, Hutchinson JN, Sara SA, Ensminger AW, Fox AH, Chess A, et al. An Architectural Role for a Nuclear Noncoding RNA: NEAT1 RNA Is Essential for the Structure of Paraspeckles. Molecular Cell. 2009;33(6):717-26.

24. Bond CS, Fox AH. Paraspeckles: nuclear bodies built on long noncoding RNA. The Journal of Cell Biology. 2009;186(5):637-44.

25. Prinz F, Kapeller A, Pichler M, Klec C. The Implications of the Long Non-Coding RNA NEAT1 in Non-Cancerous Diseases. International Journal of Molecular Sciences. 2019;20(3):627.

26. Imamura K, Imamachi N, Akizuki G, Kumakura M, Kawaguchi A, Nagata K, et al. Long Noncoding RNA NEAT1-Dependent SFPQ Relocation from Promoter Region to Paraspeckle Mediates IL8 Expression upon Immune Stimuli. Molecular Cell. 2014;53(3):393-406.

27. Yamazaki T, Souquere S, Chujo T, Kobelke S, Chong YS, Fox AH, et al. Functional Domains of NEAT1 Architectural lncRNA Induce Paraspeckle Assembly through Phase Separation. Molecular Cell. 2018;70(6):1038-53.e7.

28. Wang Z, Fan P, Zhao Y, Zhang S, Lu J, Xie W, et al. NEAT1 modulates herpes simplex virus-1 replication by regulating viral gene transcription. Cellular and Molecular Life Sciences. 2017;74(6):1117-31.

29. Tang S, Patel A, Krause PR. Hidden regulation of herpes simplex virus 1 pre-mRNA splicing and polyadenylation by virally encoded immediate early gene ICP27. PLoS pathogens. 2019;15(6):e1007884.

30. Tang S, Patel A, Krause PR. Herpes simplex virus ICP27 regulates alternative pre-mRNA polyadenylation and splicing in a sequence-dependent manner. Proceedings of the National Academy of Sciences. 2016;113(43):12256-61.

31. Wang X, Hennig T, Whisnant AW, Erhard F, Prusty BK, Friedel CC, et al. Herpes simplex virus blocks host transcription termination via the bimodal activities of ICP27. Nature communications. 2020;11(1):293. Epub 2020/01/17.

32. Wang X, Liu L, Whisnant AW, Hennig T, Djakovic L, Haque N, et al. Mechanism and consequences of herpes simplex virus 1-mediated regulation of host mRNA alternative polyadenylation. PLoS Genet. 2021;17(3):e1009263. Epub 2021/03/09.

33. Isa NF, Bensaude O, Aziz NC, Murphy S. HSV-1 ICP22 Is a Selective Viral Repressor of Cellular RNA Polymerase II-Mediated Transcription Elongation. Vaccines. 2021;9(10):1054.

34. Gu B, Eick D, Bensaude O. CTD serine-2 plays a critical role in splicing and termination factor recruitment to RNA polymerase II in vivo. Nucleic Acids Research. 2012;41(3):1591-603.

35. Vilborg A, Passarelli Maria C, Yario Therese A, Tycowski Kazimierz T, Steitz Joan A. Widespread Inducible Transcription Downstream of Human Genes. Molecular Cell. 2015;59(3):449-61.

36. Bauer DLV, Tellier M, Martínez-Alonso M, Nojima T, Proudfoot NJ, Murphy S, et al. Influenza Virus Mounts a Two-Pronged Attack on Host RNA Polymerase II Transcription. Cell Reports. 2018;23(7):2119-29.e3.

37. Heinz S, Texari L, Hayes MGB, Urbanowski M, Chang MW, Givarkes N, et al. Transcription Elongation Can Affect Genome 3D Structure. Cell. 2018;174(6):1522-36.e22.

38. Rasool Ru, Natesan R, Deng Q, Aras S, Lal P, Sander Effron S, et al. CDK7 Inhibition Suppresses Castration-Resistant Prostate Cancer through MED1 Inactivation. Cancer Discovery. 2019;9(11):1538-55.

39. Hansen TB. Improved circRNA Identification by Combining Prediction Algorithms. Frontiers in Cell and Developmental Biology. 2018;6(20).

40. Westholm Jakub O, Miura P, Olson S, Shenker S, Joseph B, Sanfilippo P, et al. Genome-wide Analysis of Drosophila Circular RNAs Reveals Their Structural and Sequence Properties and Age-Dependent Neural Accumulation. Cell Reports. 2014;9(5):1966-80.

41. Gao Y, Zhang J, Zhao F. Circular RNA identification based on multiple seed matching. Briefings in Bioinformatics. 2018;19(5):803-10.

42. Zhang Y, Xue W, Li X, Zhang J, Chen S, Zhang J-L, et al. The Biogenesis of Nascent Circular RNAs. Cell Reports. 2016;15(3):611-24.

43. Glazar P, Papavasileiou P, Rajewsky N. circBase: a database for circular RNAs. RNA. 2014;20(11):1666-70. Epub 2014/09/23.

44. Spencer CA, Dahmus ME, Rice SA. Repression of host RNA polymerase II transcription by herpes simplex virus type 1. J Virol. 1997;71(3):2031-40. Epub 1997/03/01.

45. Whisnant AW, Jurges CS, Hennig T, Wyler E, Prusty B, Rutkowski AJ, et al. Integrative functional genomics decodes herpes simplex virus 1. Nature communications. 2020;11(1):2038. Epub 2020/04/29.

46. Jeck WR, Sorrentino JA, Wang K, Slevin MK, Burd CE, Liu J, et al. Circular RNAs are abundant, conserved, and associated with ALU repeats. RNA. 2013;19(2):141-57.

47. Xiao M-S, Wilusz JE. An improved method for circular RNA purification using RNase R that efficiently removes linear RNAs containing G-quadruplexes or structured 3′ ends. Nucleic Acids Research. 2019;47(16):8755-69.

48. Ungerleider N, Concha M, Lin Z, Roberts C, Wang X, Cao S, et al. The Epstein Barr virus circRNAome. PLoS pathogens. 2018;14(8):e1007206.

49. Clark MB, Johnston RL, Inostroza-Ponta M, Fox AH, Fortini E, Moscato P, et al. Genome-wide analysis of long noncoding RNA stability. Genome Research. 2012;22(5):885-98.

50. Djakovic L, Hennig T, Reinisch K, Milic A, Whisnant A, Wolf K, et al. Nature Portfolio. 2021.

51. Te Velthuis AJ, Fodor E. Influenza virus RNA polymerase: insights into the mechanisms of viral RNA synthesis. Nat Rev Microbiol. 2016;14(8):479-93. Epub 2016/07/12.

52. Shi M, Jagger BW, Wise HM, Digard P, Holmes EC, Taubenberger JK. Evolutionary Conservation of the PA-X Open Reading Frame in Segment 3 of Influenza A Virus. Journal of Virology. 2012;86(22):12411-3.

53. Khaperskyy DA, Schmaling S, Larkins-Ford J, McCormick C, Gaglia MM. Selective Degradation of Host RNA Polymerase II Transcripts by Influenza A Virus PA-X Host Shutoff Protein. PLoS pathogens. 2016;12(2):e1005427.

54. Zhang B, Goraya MU, Chen N, Xu L, Hong Y, Zhu M, et al. Zinc Finger CCCH-Type Antiviral Protein 1 Restricts the Viral Replication by Positively Regulating Type I Interferon Response. Frontiers in Microbiology. 2020;11(1912).

55. Nemeroff ME, Barabino SML, Li Y, Keller W, Krug RM. Influenza Virus NS1 Protein Interacts with the Cellular 30 kDa Subunit of CPSF and Inhibits 3′ End Formation of Cellular Pre-mRNAs. Molecular Cell. 1998;1(7):991-1000.

56. Hale BG, Steel J, Medina RA, Manicassamy B, Ye J, Hickman D, et al. Inefficient Control of Host Gene Expression by the 2009 Pandemic H1N1 Influenza A Virus NS1 Protein. Journal of Virology. 2010;84(14):6909-22.

57. Twu KY, Kuo R-L, Marklund J, Krug RM. The H5N1 Influenza Virus NS Genes Selected after 1998 Enhance Virus Replication in Mammalian Cells. Journal of Virology. 2007;81(15):8112-21.

58. Das K, Ma L-C, Xiao R, Radvansky B, Aramini J, Zhao L, et al. Structural basis for suppression of a host antiviral response by influenza A virus. Proceedings of the National Academy of Sciences. 2008;105(35):13093-8.

59. Wilks C, Zheng SC, Chen FY, Charles R, Solomon B, Ling JP, et al. recount3: summaries and queries for large-scale RNA-seq expression and splicing. bioRxiv. 2021:2021.05.21.445138.

60. Wilks C, Gaddipati P, Nellore A, Langmead B. Snaptron: querying splicing patterns across tens of thousands of RNA-seq samples. Bioinformatics. 2017;34(1):114-6.

61. Alhasan AA, Izuogu OG, Al-Balool HH, Steyn JS, Evans A, Colzani M, et al. Circular RNA enrichment in platelets is a signature of transcriptome degradation. Blood. 2016;127(9):e1-e11.

62. Kissopoulou A, Jonasson J, Lindahl TL, Osman A. Next Generation Sequencing Analysis of Human Platelet PolyA+ mRNAs and rRNA-Depleted Total RNA. PloS one. 2013;8(12):e81809.

63. Doss JF, Corcoran DL, Jima DD, Telen MJ, Dave SS, Chi J-T. A comprehensive joint analysis of the long and short RNA transcriptomes of human erythrocytes. BMC Genomics. 2015;16(1):952.

64. Diab S, Yu M, Wang S. CDK7 Inhibitors in Cancer Therapy: The Sweet Smell of Success? Journal of Medicinal Chemistry. 2020;63(14):7458-74.

65. Li B-B, Wang B, Zhu C-M, Tang D, Pang J, Zhao J, et al. Cyclin-dependent kinase 7 inhibitor THZ1 in cancer therapy. Chronic Diseases and Translational Medicine. 2019;5(3):155-69.

66. Kwiatkowski N, Zhang T, Rahl PB, Abraham BJ, Reddy J, Ficarro SB, et al. Targeting transcription regulation in cancer with a covalent CDK7 inhibitor. Nature. 2014;511(7511):616-20.

67. Zarei M, Du H, Nassar AH, Yan RE, Giannikou K, Johnson SH, et al. Tumors with TSC mutations are sensitive to CDK7 inhibition through NRF2 and glutathione depletion. Journal of Experimental Medicine. 2019;216(11):2635-52.

68. Jiang Y-Y, Lin D-C, Mayakonda A, Hazawa M, Ding L-W, Chien W-W, et al. Targeting super-enhancer-associated oncogenes in oesophageal squamous cell carcinoma. Gut. 2017;66(8):1358-68.

69. Yuan J, Jiang Y-Y, Mayakonda A, Huang M, Ding L-W, Lin H, et al. Super-Enhancers Promote Transcriptional Dysregulation in Nasopharyngeal Carcinoma. Cancer Research. 2017;77(23):6614-26.

70. Lu P, Geng J, Zhang L, Wang Y, Niu N, Fang Y, et al. THZ1 reveals CDK7-dependent transcriptional addictions in pancreatic cancer. Oncogene. 2019;38(20):3932-45.

71. Sharifnia T, Wawer MJ, Chen T, Huang Q-Y, Weir BA, Sizemore A, et al. Small-molecule targeting of brachyury transcription factor addiction in chordoma. Nature Medicine. 2019;25(2):292-300.

72. Sheppard HE, Dall'Agnese A, Park WD, Shamim MH, Dubrulle J, Johnson HL, et al. Targeted brachyury degradation disrupts a highly specific autoregulatory program controlling chordoma cell identity. Cell Reports Medicine. 2021;2(1).

73. Abudureheman T, Xia J, Li M-H, Zhou H, Zheng W-W, Zhou N, et al. CDK7 Inhibitor THZ1 Induces the Cell Apoptosis of B-Cell Acute Lymphocytic Leukemia by Perturbing Cellular Metabolism. Frontiers in Oncology. 2021;11(745).

74. Trask DK, Muller MT. Stabilization of type I topoisomerase-DNA covalent complexes by actinomycin D. Proceedings of the National Academy of Sciences. 1988;85(5):1417-21.

75. Belakavadi M, Pandey PK, Vijayvargia R, Fondell JD. MED1 Phosphorylation Promotes Its Association with Mediator: Implications for Nuclear Receptor Signaling. Molecular and Cellular Biology. 2008;28(12):3932-42.

76. Allen BL, Taatjes DJ. The Mediator complex: a central integrator of transcription. Nature Reviews Molecular Cell Biology. 2015;16(3):155-66.

77. Werner Michael S, Ruthenburg Alexander J. Nuclear Fractionation Reveals Thousands of Chromatin-Tethered Noncoding RNAs Adjacent to Active Genes. Cell Reports. 2015;12(7):1089-98.

78. Yankulov K, Yamashita K, Roy R, Egly J-M, Bentley DL. The Transcriptional Elongation Inhibitor 5,6-Dichloro-1-β-D-ribofuranosylbenzimidazole Inhibits Transcription Factor IIH-associated Protein Kinase (∗). Journal of Biological Chemistry. 1995;270(41):23922-5.

79. Zhu Y, Pe'ery T, Peng J, Ramanathan Y, Marshall N, Marshall T, et al. Transcription elongation factor P-TEFb is required for HIV-1 Tat transactivation in vitro. Genes & Development. 1997;11(20):2622-32.

80. Turinetto V, Porcedda P, Orlando L, De Marchi M, Amoroso A, Giachino C. The cyclin-dependent kinase inhibitor 5, 6-dichloro-1-beta-D-ribofuranosylbenzimidazole induces nongenotoxic, DNA replication-independent apoptosis of normal and leukemic cells, regardless of their p53 status. BMC Cancer. 2009;9(1):281.

81. Chujo T, Yamazaki T, Kawaguchi T, Kurosaka S, Takumi T, Nakagawa S, et al. Unusual semi-extractability as a hallmark of nuclear body-associated architectural noncoding RNAs. The EMBO Journal. 2017;36(10):1447-62.

82. Li K, Wang Z. Speckles and paraspeckles coordinate to regulate HSV-1 genes transcription. Communications Biology. 2021;4(1):1207.

83. Lellahi SM, Rosenlund IA, Hedberg A, Kiær LT, Mikkola I, Knutsen E, et al. The long noncoding RNA NEAT1 and nuclear paraspeckles are up-regulated by the transcription factor HSF1 in the heat shock response. Journal of Biological Chemistry. 2018;293(49):18965-76.

84. Stagsted LVW, O'Leary ET, Ebbesen KK, Hansen TB. The RNA-binding protein SFPQ preserves long-intron splicing and regulates circRNA biogenesis in mammals. eLife. 2021;10:e63088.

85. Sampathi S, Acharya P, Zhao Y, Wang J, Stengel KR, Liu Q, et al. The CDK7 inhibitor THZ1 alters RNA polymerase dynamics at the 5′ and 3′ ends of genes. Nucleic Acids Research. 2019;47(8):3921-36.

86. Glover-Cutter K, Larochelle S, Erickson B, Zhang C, Shokat K, Fisher RP, et al. TFIIH-Associated Cdk7 Kinase Functions in Phosphorylation of C-Terminal Domain Ser7 Residues, Promoter-Proximal Pausing, and Termination by RNA Polymerase II. Molecular and Cellular Biology. 2009;29(20):5455-64.

87. Larochelle S, Amat R, Glover-Cutter K, Sansó M, Zhang C, Allen JJ, et al. Cyclin-dependent kinase control of the initiation-to-elongation switch of RNA polymerase II. Nature Structural & Molecular Biology. 2012;19(11):1108-15.

88. Yuan F, Hankey W, Wagner EJ, Li W, Wang Q. Alternative polyadenylation of mRNA and its role in cancer. Genes & Diseases. 2021;8(1):61-72.

89. Wu Y, Yang L, Zhao J, Li C, Nie J, Liu F, et al. Nuclear-enriched abundant transcript 1 as a diagnostic and prognostic biomarker in colorectal cancer. Molecular Cancer. 2015;14(1):191.

90. Pisani G, Baron B. NEAT1 and Paraspeckles in Cancer Development and Chemoresistance. Non-Coding RNA. 2020;6(4):43.

91. Wyler E, Menegatti J, Franke V, Kocks C, Boltengagen A, Hennig T, et al. Widespread activation of antisense transcription of the host genome during herpes simplex virus 1 infection. Genome Biology. 2017;18(1):209.

92. Nilson Kyle A, Guo J, Turek Michael E, Brogie John E, Delaney E, Luse Donal S, et al. THZ1 Reveals Roles for Cdk7 in Co-transcriptional Capping and Pausing. Molecular Cell. 2015;59(4):576-87.

93. Rivas T, Goodrich JA, Kugel JF. The Herpes Simplex Virus 1 Protein ICP4 Acts as both an Activator and a Repressor of Host Genome Transcription during Infection. Molecular and Cellular Biology. 2021;41(10):e00171-21.

94. Hutterer C, Eickhoff J, Milbradt J, Korn K, Zeitträger I, Bahsi H, et al. A Novel CDK7 Inhibitor of the Pyrazolotriazine Class Exerts Broad-Spectrum Antiviral Activity at Nanomolar Concentrations. Antimicrobial Agents and Chemotherapy. 2015;59(4):2062-71.

95. Schang LM, Bantly A, Knockaert M, Shaheen F, Meijer L, Malim MH, et al. Pharmacological Cyclin-Dependent Kinase Inhibitors Inhibit Replication of Wild-Type and Drug-Resistant Strains of Herpes Simplex Virus and Human Immunodeficiency Virus Type 1 by Targeting Cellular, Not Viral, Proteins. Journal of Virology. 2002;76(15):7874-82.

96. Fox HL, Dembowski JA, DeLuca NA, Imperiale MJ, Weller S, Sandri-Goldin R. A Herpesviral Immediate Early Protein Promotes Transcription Elongation of Viral Transcripts. mBio. 2017;8(3):e00745-17.

97. Bonfert T, Kirner E, Csaba G, Zimmer R, Friedel CC. ContextMap 2: fast and accurate context-based RNA-seq mapping. BMC Bioinformatics. 2015;16:122. Epub 2015/05/01.

98. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25(14):1754-60. Epub 2009/05/20.

99. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014;30(7):923-30.

100. Zerbino DR, Achuthan P, Akanni W, Amode MR, Barrell D, Bhai J, et al. Ensembl 2018. Nucleic Acids Res. 2018;46(D1):D754-D61. Epub 2017/11/21.

101. Garrido-Martín D, Palumbo E, Guigó R, Breschi A. ggsashimi: Sashimi plot revised for browser- and annotation-independent splicing visualization. PLOS Computational Biology. 2018;14(8):e1006360.

102. Anders S, Reyes A, Huber W. Detecting differential usage of exons from RNA-seq data. Genome Research. 2012;22(10):2008-17.

103. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology. 2014;15(12):550.

Supporting information captions

S1 Fig. Experimental setup and de novo circRNA detection.

(a) Experimental setup for the total RNA- and 4sU-seq time-courses in WT and Δ*vhs* infection. Fig taken from [19] (© the authors). Two biological replicates were obtained for each time-point. **(b)** Outline of our pipeline for de novo circRNA detection (see also methods). It combines two different algorithms, circRNA finder [40] and CIRI2 [41], and retains only circRNAs identified by both algorithms. circRNA finder is based on analysis of chimeric reads determined with the RNA-seq mapper STAR. Chimeric reads are reads that align to two distinct regions of the genome in manner not consistent with "normal" linear transcripts. CIRI2 is based on identifying pairs of clipped read alignments (i.e. alignments for substrings of the read, determined with BWA [98]) for the same read where a downstream part of the read aligns upstream of an upstream part of the read (= back-spliced junction (BSJ) reads). Both methods use further filtering steps to remove false positive results.

S2 Fig. Statistics on identified circRNAs in HSV-1 WT and Δ*vhs* infection

(a) Number of circRNAs identified for each sample of 4sU-RNA. **(b)** Scatterplot comparing the number of reads mapped to the human genome against the number of distinct circular splice sites identified by the *de novo* circRNA detection approach outlined in S1 Fig b for the total RNA and 4sU-RNA time-courses in HSV-1 WT and Δ*vhs* infection. Numbers are shown separately for each replicate. **(c,d)** Percentage spliced host reads (= no. reads aligning to known splice junctions of host protein-coding genes / no. of mapped host reads × 100) for both replicates of total RNA-seq for **(c)** WT and **(d)** Δ*vhs* infection.

S3 Fig. Accumulation of circRNAs in WT but not Δ*vhs* infection.

Scatterplots comparing normalized circRNA counts (normalized to the number of linear junction reads mapped to the host genome) between mock and 2, 4 and 6 h p.i. WT infection **(a-c)** and mock and 2, 4 and 6 h p.i. Δ*vhs* infection **(d-f)**. Linear regression analysis across all circRNAs (red line) was used to estimate the accumulation of circRNAs compared to mock infection. The regression estimate for the accumulation is shown on the bottom right. The gray line indicates the diagonal, i.e. equal values on the x- and y-axis. The five most highly expressed circRNAs are marked by name. Corresponding scatterplots for 8 h p.i. are shown in Fig 1c,d.

S4 Fig. Results for circRNA normalization by total number of mapped reads.

Scatterplots comparing normalized circRNA counts normalized to the total number of reads mapped to the host genome between mock and 2, 4, 6 and 8 h p.i. WT infection **(a-d)** and mock and 2, 4, 6 and 8 h p.i. Δ*vhs* infection **(e-h)**. Linear regression analysis across all circRNAs (red line) was used to estimate the accumulation of circRNAs compared to mock infection. The regression estimate for the accumulation is shown on the bottom right. The gray line indicates the diagonal, i.e. equal values on the x- and y-axis. The five most highly expressed circRNAs are marked by name.

S5 Fig. Outline of our alignment-based pipeline for circRNA detection.

For details see Materials and Methods. This approach allows increasing or decreasing the required overlap of a read with either side of a circular junction (=:X).

S6 Fig. Results for trimmed reads.

Scatterplots comparing normalized circRNA counts (normalized to the number of linear junction reads **(a-d)** or to the total number of reads mapped to the host genome **(e-f)**) between mock and 2, 4, 6 and 8 h p.i. WT infection after trimming reads down to 76 nt. Linear regression analysis across all circRNAs (red line) was used to estimate the accumulation of circRNAs compared to mock infection. The regression estimate for the accumulation is shown on the bottom right. The gray line indicates the diagonal, i.e. equal values on the x- and y-axis. The five most highly expressed circRNAs are marked by name.

S7 Fig. Results for the alignment-based circRNA detection pipeline.

Scatterplots comparing normalized circRNA counts (obtained with the alignment-based circRNA detection pipeline outlined in S5 Fig and normalized to the number of linear junction reads mapped to the host genome) between mock and 2, 4, 6 and 8 h p.i. WT infection **(a-d)** and between mock and 2, 4, 6 and 8 h p.i. Δ*vhs* infection **(e-h)**. Linear regression analysis across all circRNAs (red line) was used to estimate the accumulation of circRNAs compared to mock infection. The regression estimate for the accumulation is shown on the bottom right. The gray line indicates the diagonal, i.e. equal values on the x- and y-axis. The five most highly expressed circRNAs are marked by name.

S8 Fig. Results for an independent total RNA-seq experiment including ΔICP27 infection.

Scatterplots comparing normalized circRNA counts (normalized to the number of linear junction reads mapped to the host genome) between mock and **(a)** WT, **(b)** ΔICP27 and **(c)** Δ*vhs* infection. CircRNA read counts were obtained with the alignment-based circRNA detection pipeline outlined in S5 Fig. Linear regression

analysis across all circRNAs (red line) was used to estimate the accumulation of circRNAs compared to mock infection. The regression estimate for the accumulation is shown on the bottom right. The gray line indicates the diagonal, i.e. equal values on the x- and y-axis. The five most highly expressed circRNAs are marked by name.

S9 Fig. Impact of circRNA accumulation on differential gene and exon analyses.

(a) Boxplots showing the distribution of log2 fold-changes in gene expression between mock and 2, 4, 6 and 8 p.i. total RNA in WT and Δv*hs* infection for genes containing at least one circRNA (blue) or no circRNA (gray). log2 fold-changes were taken from our recent study [19] (see also methods). Wilcoxon rank sum tests were used to assess whether log2 fold-changes were significantly increased for circRNA genes compared to other genes at each time-point of infection (p-values shown in red at the bottom). **(b)** Boxplots showing the distribution of normalized circRNA counts for circRNAs with (red) and without (blue) differential exon usage for exons within the circRNA region at each time-point of WT infection. **(c)** Percentage of expressed circRNAs (=normalized circRNA count >0 in uninfected cells) for which at least one exon within the genomic region of the circRNA shows differential exon usage (determined with DEXSeq, multiple testing adjusted p-value ≤ 0.005) for each time-point of Δ*vhs* infection. **(d)** Boxplots showing the distribution of log2 fold-changes for exons located within circRNAs. For each circRNA, only the exon with the maximum absolute log2 fold-change is shown.

S10 Fig. Enrichment of the hsa_circ_0003812 NEAT1_2 circRNA in total and 4sU-RNA and subcellular fractions in HSV-1 and mutant infections.

Enrichment (:=ε) compared to mock infection was calculated as the ratio of normalized circRNA counts between infected and uninfected cells using a pseudocount of 0.1 to avoid division by zero. Numbers on top indicate the number of circRNAs identified in the corresponding condition with a higher normalized circRNA count than hsa_circ_0003812 . NA indicates that hsa_circ_0003812 was not found in the particular condition.

S11 Fig. Scatterplot comparing the number of reads mapped to the human genome against the number of distinct circular splice sites for chromatin-associated, nucleoplasmic, cytosolic and total RNA from the subcellular fractions experiment.

S12 Fig. Partial resistance of NEAT1_2 against RNase R treatment.

Sashimi plots showing NEAT1 read coverage (overlay of replicates) and circular (red) and linear (same color as read coverage) splice junctions as arcs connecting acceptor and donor splice site in RNA-seq data from the studies by **(a)** Xiao and Wilusz [47] and **(b)** Ungerleider et al. [48] (for details see below). Number of junction reads are annotated to arcs separately for replicates. Junctions are only shown if at least 3 and 4 reads, respectively, align by at least 10 nt on both sides of the junction. Genomic coordinates of NEAT1 transcripts and the HSV-1-induced circRNA are shown at the bottom. **(a)** Xiao and Wilusz performed RNA-seq for HeLa cells with and without (=Ctrl) RNase R treatment using two different protocols. The first protocol is the standard approach used for enriching circRNAs and employs RNase R treatment with a KCl-containing buffer for 15 min (2 biological replicates). In the second protocol, RNA was treated by E-PAP followed by digestion with RNase R in a LiCl-containing buffer (3 biological replicates). **(b)** Ungerleider et al. performed

RNase R treatment for total, cytoplasmic and nuclear RNA obtained from Akata cells without (=Ctrl) and with induction of EBV activation. **(a,b)** NEAT1_2 is partially resistant to RNase R treatment due to the stabilizing triple helical structure at its 3'end, thus even after RNase R treatment reads from the linear NEAT1_2 transcript, in particular its 3'end, are obtained for total RNA from HeLa cells and nuclear RNA from Akata cells. For total and cytoplasmic RNA from Akata cells (both induced and noninduced), depletion of the linear NEAT1_2 transcript was more successful.

S13 Fig. Sashimi plots for total RNA time-courses of WT and Δ*vhs* infection. For an explanation of sashimi plots see caption to Fig 2. Circular splice junctions are marked in red, linear splice junctions in the same color as read coverage.

S14 Fig. Confirmation of the hsa_circ_0003812 NEAT circRNA.

(a) Heatmap of normalized circRNA counts (normalized to the number of linear junction reads mapped to the host genome) for all NEAT1_2 circRNAs in total and 4sU-RNA time-courses of WT infection identified with the alignment-based approach with a minimum overlap of at least 40 nt on either side of the circular junction. Columns represent individual circRNAs, which are ordered according to their genomic coordinates from the most 5' to the most 3'. The hsa_circ_0003812 NEAT1_2 circRNA is marked by a blue rectangle. **(b)** Read counts for circular junction reads for hsa_circ_0003812 identified with the alignment-based approach with a minimum overlap of at least 40 nt on either side of the circular junction (i.e. raw read counts used for calculating normalized circRNA counts for **(a)**) and read counts for confirming read pairs for hsa_circ_0003812 identified with the approach outlined in **(c)**.

Numbers are shown separately for both replicates of the total and 4sU RNA time-courses of WT infection. **(c)** Definition and detection of confirming read pairs for circRNAs. For further details see methods.

S15 Fig. Sashimi plots for 4sU-RNA time-courses of WT (a) and Δ*vhs* infection (b) and nucleoplasmic (c) and total RNA (d) from the subcellular fractions experiment for WT, Δ*vhs* and ΔICP27 infection.

For an explanation of sashimi plots see caption to Fig 2. Circular splice junctions are marked in red, linear splice junctions in the same color as read coverage.

S16 Fig. Conservation of splice sites for (a) the hsa_circ_0003812 circular junction and (b-c) the two most frequent linear NEAT1_2 splice junctions (numbered (1) and (2) in Fig 2e).

The left columns show the donor splice site (canonical GT signal at the intron 5' end) and the right column the acceptor splice site (canonical AG at the intron 3'end), with the red vertical lines marking the exon-intron boundaries. Please note that for the circular junction in **(a)**, the donor splice site is downstream of the acceptor splice site on the genome. Genome alignments against the human genome (hg19) were obtained from the UCSC genome browser.

S17 Fig. Sashimi plots for mock, WT-F and ΔICP22 infection at 8 h p.i.

For an explanation of sashimi plots see caption to Fig 2. Circular splice junctions are marked in red, linear splice junctions in the same color as read coverage.

S18 Fig. Normalized circRNA counts for the hsa_circ_0003812 NEAT1_2 circRNA in IAV infection (cyan) and time-matched controls (gray).

Results are shown separately for H1N1, H3N2 and H5N1 infection and HTBE and MDM cells.

S19 Fig. Accumulation of circRNAs in IAV infection.

(a-g) Scatterplots comparing normalized circRNA counts (normalized to the number of linear junction reads mapped to the host genome) for IAV infections at 18 h p.i. against time-matched controls. Linear regression analysis across all circRNAs (red line) was used to estimate the accumulation of circRNAs compared to mock infection. The regression estimate for the accumulation is shown on the bottom right. The gray line indicates the diagonal, i.e. equal values on the x- and y-axis. The most highly expressed circRNAs are marked by name. **(h)** Boxplots showing the distribution of log2 fold-changes between normalized circRNA counts in infection compared to time-matched controls for 7077 well-expressed circRNAs (normalized circRNA count >1 in at least one condition). A pseudocount of 0.1 was used to avoid division by zero. IAV strain and time-points of infection are color-coded (cyan=H1N1, green=H3N2, purple=H5N1, darker colors indicate later time-points). Values for the hsa_circ_0003812 NEAT1_2 circRNA and a strongly induced ZC3HAV1 circRNA are shown as blue squares and red triangles, respectively.

S20 Fig. Boxplots showing distribution of % read-through in H1N1, H3N2 and H5N1 infection of HTBE and MDM cells.

Read-through was calculated as previously described in [18] (see also methods).

S21 Fig. Results of the Snaptron search for NEAT1_2 splice junctions.

(a) Overview of number of projects and samples that were identified with the Snaptron search in recount3 for the four most frequent NEAT1_2 linear splice

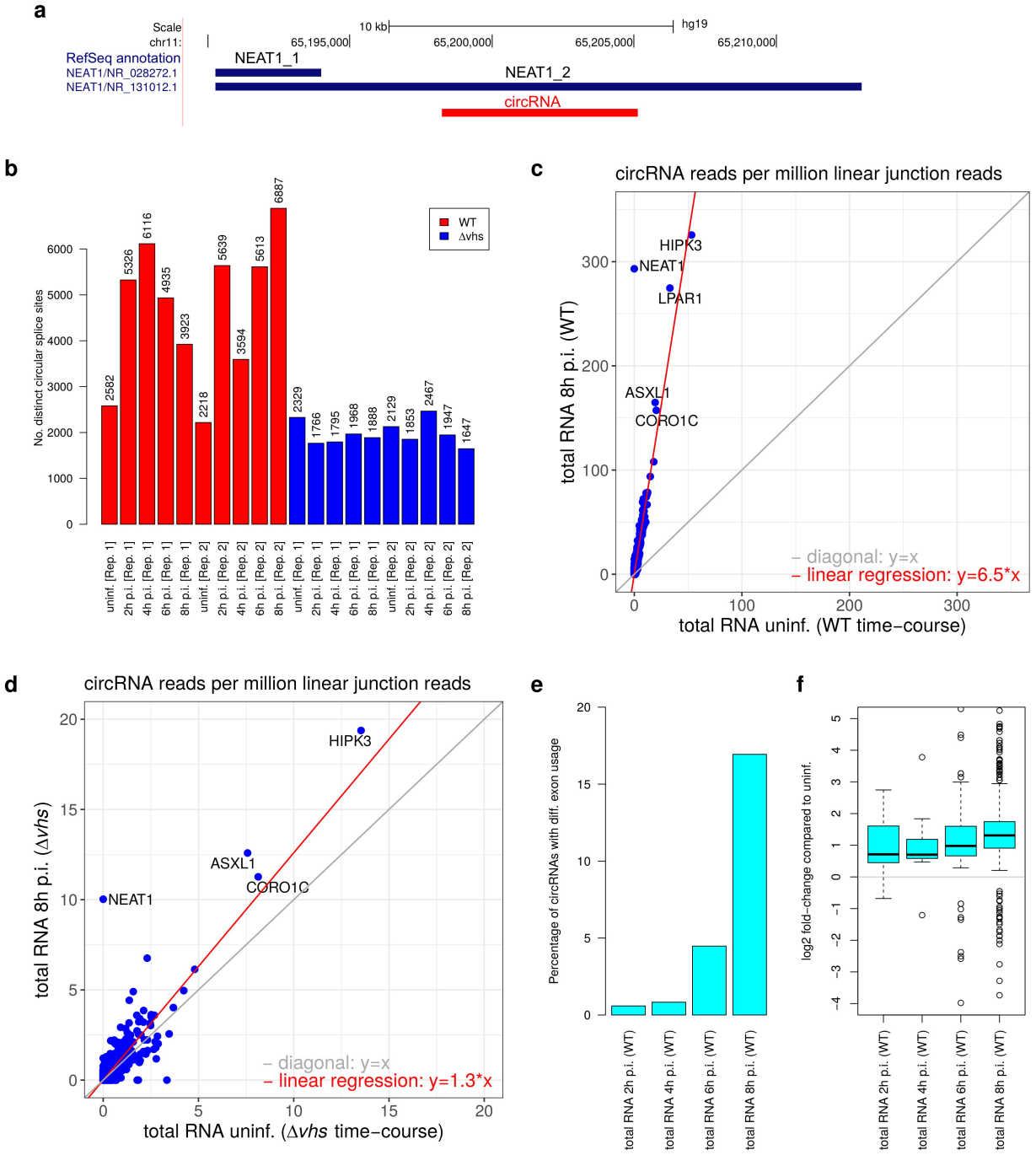
junctions induced in HSV-1 infection. The left-most column indicates regular expressions that were searched in the project and sample descriptions provided by recount3 using grep in R (case insensitive). The table shows only projects and samples for which at least one of the four most frequent NEAT1_2 linear splice junctions is covered by ≥ 10 reads. **(b**) Sashimi plots as in Fig 2 for two RNA-seq experiments (3 replicates each) in erythrocytes [63] and platelets [62]. For an explanation of sashimi plots, see caption to Fig 2. Circular splice junctions are marked in red, linear splice junctions in black.

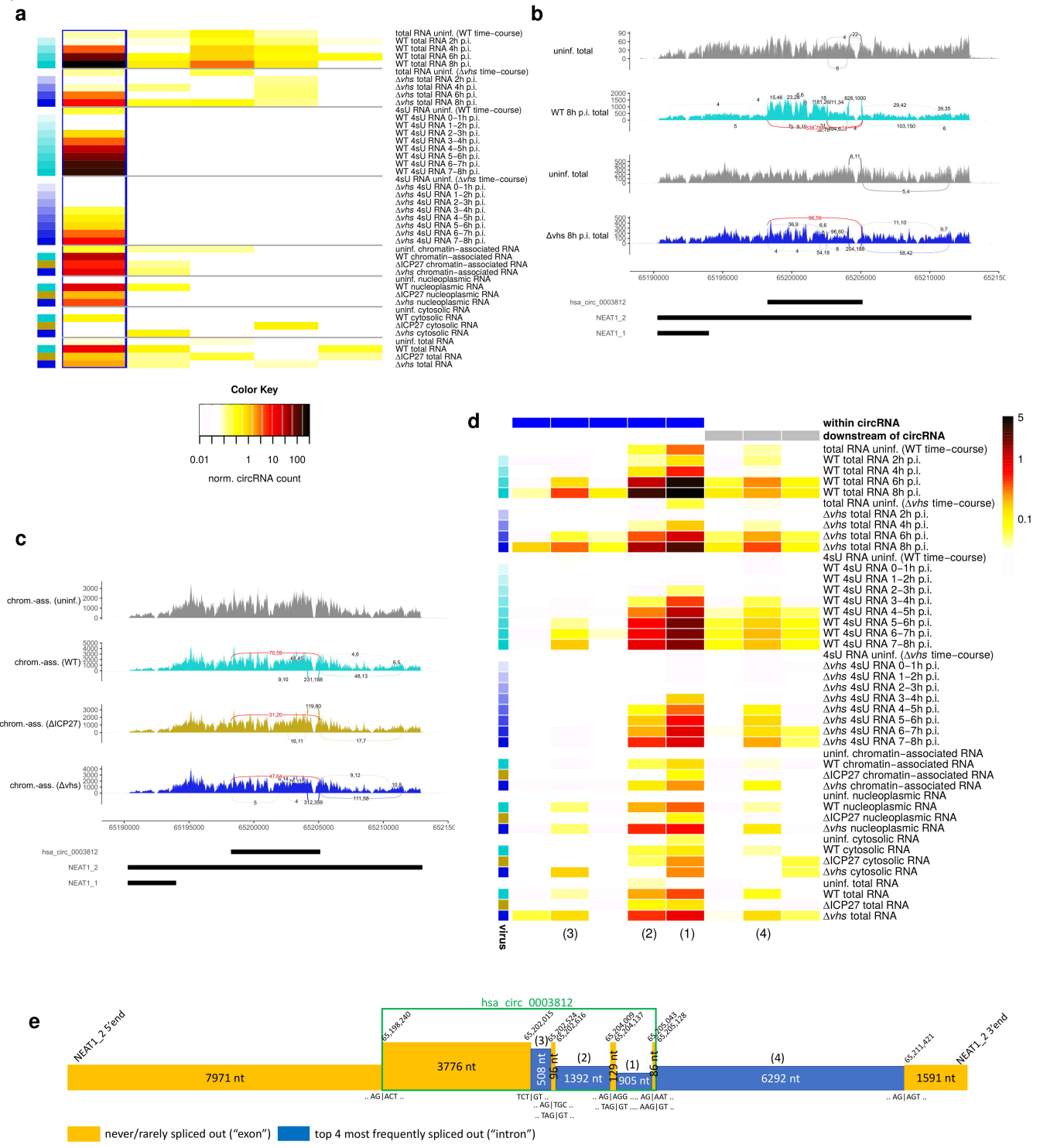
S22 Fig. Sashimi plots showing read coverage and circular and linear splice junctions (a) upon control, THZ1 or Act-D treatment in different concentrations in UM-Chor1 cells (2 replicates each) and (b) upon control or DRB treatment in HEK293 cells (2 replicates each).

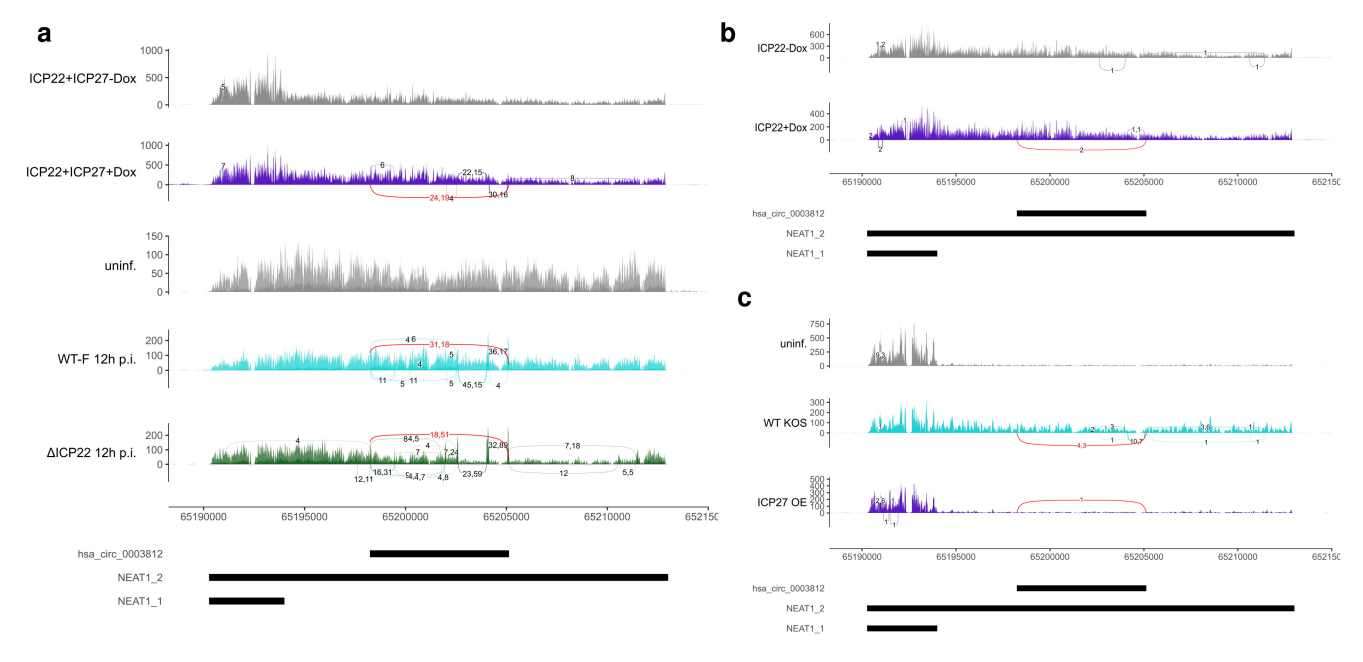
For an explanation of sashimi plots see caption to Fig 2. Circular splice junctions are marked in red, linear splice junctions in the same color as read coverage.

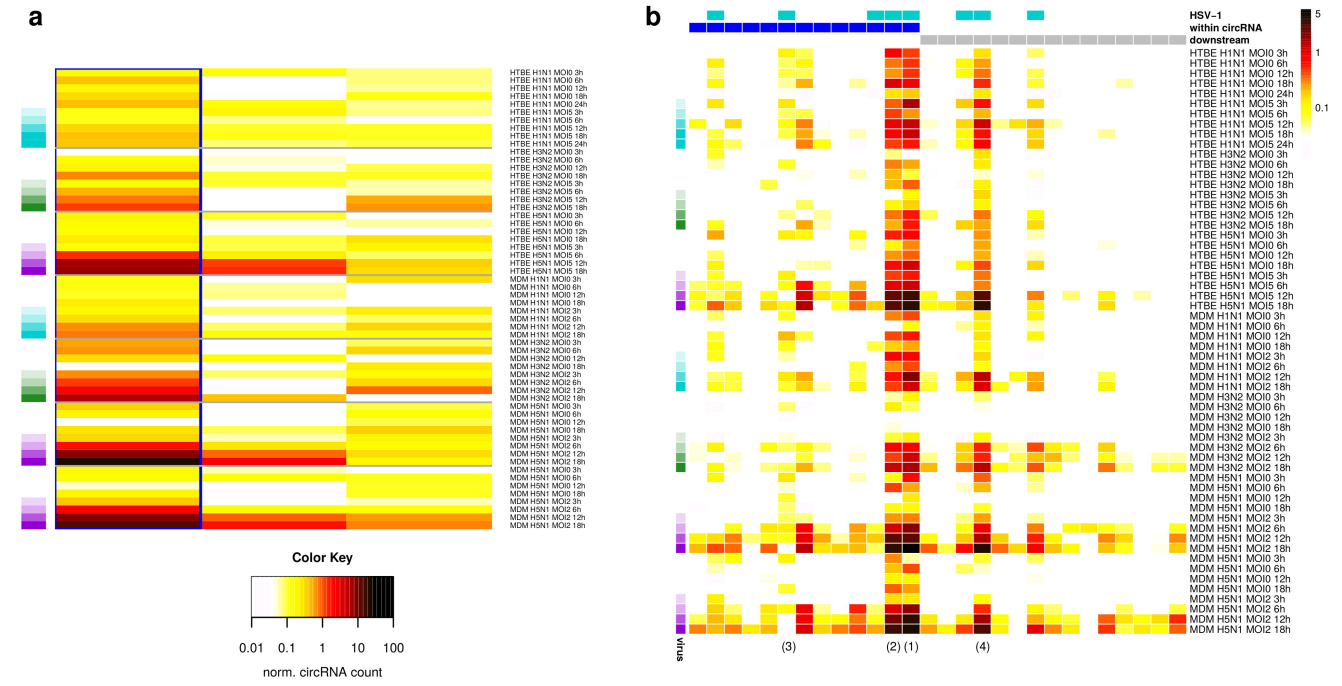
S23 Fig. Sashimi plots illustrating read-through transcription upon THZ1 treatment in UM-Chor1 and Nalm6 cells for example genes.

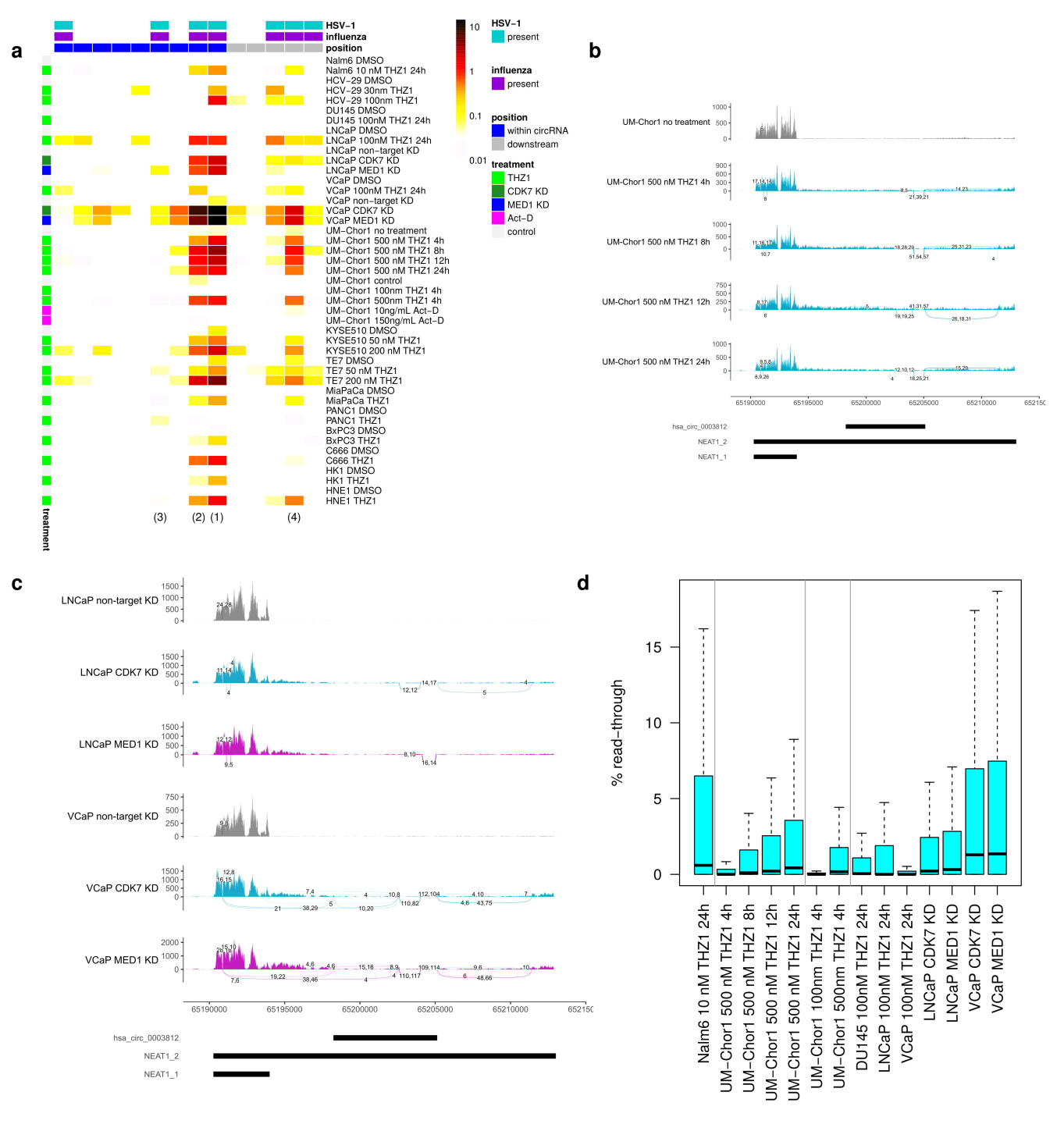
Example genes SRSF3 and SRSF6 also show read-through in HSV-1 and IAV infection and salt and heat stress. Transcript annotation is shown at the bottom, with filled rectangles for exons and lines for introns and arrowheads indicating strand. For UM-Chor1 cells, RNA-seq was performed in a strand-specific manner and only reads mapping to the same strand as the genes are shown. RNA-seq in Nalm6 cells was not strand-specific, thus read coverage upstream of gene 5'ends likely reflects induction of antisense transcription upon THZ1 treatment.


a
Scale
chr11:
10 kb
hg19
65,195,000
65,200,000
65,205,000
65,210,000
RefSeq annotation
NEAT1/NR_028272.1
NEAT1/NR_131012.1
NEAT1_1
NEAT1_2
circRNA
b
No. distinct circular splice sites
WT
Δvhs
2582
5326
6116
4935
3923
2218
5639
3594
5613
6887
2329
1766
1795
1968
1888
2129
1853
2467
1947
1647
uninf. [Rep. 1]
2h p.i. [Rep. 1]
4h p.i. [Rep. 1]
6h p.i. [Rep. 1]
8h p.i. [Rep. 1]
uninf. [Rep. 2]
2h p.i. [Rep. 2]
4h p.i. [Rep. 2]
6h p.i. [Rep. 2]
8h p.i. [Rep. 2]
c
circRNA reads per million linear junction reads
total RNA 8h p.i. (WT)
total RNA uninf. (WT time−course)
HIPK3
NEAT1
LPAR1
ASXL1
CORO1C
diagonal: y=x
linear regression: y=6.5*x
d
circRNA reads per million linear junction reads
total RNA 8h p.i. (Δvhs)
total RNA uninf. (Δvhs time−course)
HIPK3
ASXL1
CORO1C
NEAT1
diagonal: y=x
linear regression: y=1.3*x
e
Percentage of circRNAs with diff. exon usage
total RNA 2h p.i. (WT)
total RNA 4h p.i. (WT)
total RNA 6h p.i. (WT)
total RNA 8h p.i. (WT)
f
log2 fold−change compared to uninf.
total RNA 2h p.i. (WT)
total RNA 4h p.i. (WT)
total RNA 6h p.i. (WT)
total RNA 8h p.i. (WT)


a
total RNA uninf. (WT time−course)
WT total RNA 2h p.i.
WT total RNA 4h p.i.
WT total RNA 6h p.i.
WT total RNA 8h p.i.
total RNA uninf. (Δvhs time−course)
Δvhs total RNA 2h p.i.
Δvhs total RNA 4h p.i.
Δvhs total RNA 6h p.i.
Δvhs total RNA 8h p.i.
4sU RNA uninf. (WT time−course)
WT 4sU RNA 0−1h p.i.
WT 4sU RNA 1−2h p.i.
WT 4sU RNA 2−3h p.i.
WT 4sU RNA 3−4h p.i.
WT 4sU RNA 4−5h p.i.
WT 4sU RNA 5−6h p.i.
WT 4sU RNA 6−7h p.i.
WT 4sU RNA 7−8h p.i.
4sU RNA uninf. (Δvhs time−course)
Δvhs 4sU RNA 0−1h p.i.
Δvhs 4sU RNA 1−2h p.i.
Δvhs 4sU RNA 2−3h p.i.
Δvhs 4sU RNA 3−4h p.i.
Δvhs 4sU RNA 4−5h p.i.
Δvhs 4sU RNA 5−6h p.i.
Δvhs 4sU RNA 6−7h p.i.
Δvhs 4sU RNA 7−8h p.i.
uninf. chromatin−associated RNA
WT chromatin−associated RNA
ΔICP27 chromatin−associated RNA
Δvhs chromatin−associated RNA
uninf. nucleoplasmic RNA
WT nucleoplasmic RNA
ΔICP27 nucleoplasmic RNA
Δvhs nucleoplasmic RNA
uninf. cytosolic RNA
WT cytosolic RNA
ΔICP27 cytosolic RNA
Δvhs cytosolic RNA
uninf. total RNA
WT total RNA
ΔICP27 total RNA
Δvhs total RNA
Color Key
0.01
1
10
100
norm. circRNA count
b
uninf. total
WT 8h p.i. total
uninf. total
Δvhs 8h p.i. total
65190000
65195000
65200000
65205000
65210000
hsa_circ_0003812
NEAT1_2
NEAT1_1
c
chrom.-ass. (uninf.)
chrom.-ass. (WT)
chrom.-ass. (ΔICP27)
chrom.-ass. (Δvhs)
65190000
65195000
65200000
65205000
65210000
hsa_circ_0003812
NEAT1_2
NEAT1_1
d
within circRNA
downstream of circRNA
total RNA uninf. (WT time−course)
WT total RNA 2h p.i.
WT total RNA 4h p.i.
WT total RNA 6h p.i.
WT total RNA 8h p.i.
total RNA uninf. (Δvhs time−course)
Δvhs total RNA 2h p.i.
Δvhs total RNA 4h p.i.
Δvhs total RNA 6h p.i.
Δvhs total RNA 8h p.i.
4sU RNA uninf. (WT time−course)
WT 4sU RNA 0−1h p.i.
WT 4sU RNA 1−2h p.i.
WT 4sU RNA 2−3h p.i.
WT 4sU RNA 3−4h p.i.
WT 4sU RNA 4−5h p.i.
WT 4sU RNA 5−6h p.i.
WT 4sU RNA 6−7h p.i.
WT 4sU RNA 7−8h p.i.
4sU RNA uninf. (Δvhs time−course)
Δvhs 4sU RNA 0−1h p.i.
Δvhs 4sU RNA 1−2h p.i.
Δvhs 4sU RNA 2−3h p.i.
Δvhs 4sU RNA 3−4h p.i.
Δvhs 4sU RNA 4−5h p.i.
Δvhs 4sU RNA 5−6h p.i.
Δvhs 4sU RNA 6−7h p.i.
Δvhs 4sU RNA 7−8h p.i.
uninf. chromatin−associated RNA
WT chromatin−associated RNA
ΔICP27 chromatin−associated RNA
Δvhs chromatin−associated RNA
uninf. nucleoplasmic RNA
WT nucleoplasmic RNA
ΔICP27 nucleoplasmic RNA
Δvhs nucleoplasmic RNA
uninf. cytosolic RNA
WT cytosolic RNA
ΔICP27 cytosolic RNA
Δvhs cytosolic RNA
uninf. total RNA
WT total RNA
ΔICP27 total RNA
Δvhs total RNA
virus
(3)
(2)
(1)
(4)
5
1
0.1
e
hsa_circ_0003812
NEAT1_2 5'end
NEAT1_2 3'end
7971 nt
3776 nt
508 nt
96 nt
1392 nt
129 nt
905 nt
86 nt
6292 nt
1591 nt
(3)
(2)
(1)
(4)
.. AG|ACT ..
TCT|GT ..
.. AG|TGC ..
.. TAG|GT ..
.. AG|AGG .... AG|AAT ..
.. TAG|GT .... AAG|GT ..
.. AG|AGT ..
never/rarely spliced out ("exon")
top 4 most frequently spliced out ("intron")


a
ICP22+ICP27-Dox
ICP22+ICP27+Dox
uninf.
WT-F 12h p.i.
ΔICP22 12h p.i.
65190000
65195000
65200000
65205000
65210000
hsa_circ_0003812
NEAT1_2
NEAT1_1
b
ICP22-Dox
ICP22+Dox
c
uninf.
WT KOS
ICP27 OE


a
HTBE H1N1 MOI0 3h
HTBE H1N1 MOI0 6h
HTBE H1N1 MOI0 12h
HTBE H1N1 MOI0 18h
HTBE H1N1 MOI0 24h
HTBE H1N1 MOI5 3h
HTBE H1N1 MOI5 6h
HTBE H1N1 MOI5 12h
HTBE H1N1 MOI5 18h
HTBE H1N1 MOI5 24h
HTBE H3N2 MOI0 3h
HTBE H3N2 MOI0 6h
HTBE H3N2 MOI0 12h
HTBE H3N2 MOI0 18h
HTBE H3N2 MOI5 3h
HTBE H3N2 MOI5 6h
HTBE H3N2 MOI5 12h
HTBE H3N2 MOI5 18h
HTBE H5N1 MOI0 3h
HTBE H5N1 MOI0 6h
HTBE H5N1 MOI0 12h
HTBE H5N1 MOI0 18h
HTBE H5N1 MOI5 3h
HTBE H5N1 MOI5 6h
HTBE H5N1 MOI5 12h
HTBE H5N1 MOI5 18h
MDM H1N1 MOI0 3h
MDM H1N1 MOI0 6h
MDM H1N1 MOI0 12h
MDM H1N1 MOI0 18h
MDM H1N1 MOI2 3h
MDM H1N1 MOI2 6h
MDM H1N1 MOI2 12h
MDM H1N1 MOI2 18h
MDM H3N2 MOI0 3h
MDM H3N2 MOI0 6h
MDM H3N2 MOI0 12h
MDM H3N2 MOI0 18h
MDM H3N2 MOI2 3h
MDM H3N2 MOI2 6h
MDM H3N2 MOI2 12h
MDM H3N2 MOI2 18h
MDM H5N1 MOI0 3h
MDM H5N1 MOI0 6h
MDM H5N1 MOI0 12h
MDM H5N1 MOI0 18h
MDM H5N1 MOI2 3h
MDM H5N1 MOI2 12h
MDM H5N1 MOI2 18h
MDM H5N1 MOI0 6h
MDM H5N1 MOI0 12h
MDM H5N1 MOI0 18h
MDM H5N1 MOI2 3h
MDM H5N1 MOI2 6h
MDM H5N1 MOI2 12h
MDM H5N1 MOI2 18h
Color Key
0.01
0.1
1
10
100
norm. circRNA count
b
HSV-1
within circRNA
downstream
5
1
0.1
HTBE H1N1 MOI0 3h
HTBE H1N1 MOI0 6h
HTBE H1N1 MOI0 12h
HTBE H1N1 MOI0 18h
HTBE H1N1 MOI0 24h
HTBE H1N1 MOI5 3h
HTBE H1N1 MOI5 6h
HTBE H1N1 MOI5 12h
HTBE H1N1 MOI5 18h
HTBE H1N1 MOI5 24h
HTBE H3N2 MOI0 3h
HTBE H3N2 MOI0 6h
HTBE H3N2 MOI0 12h
HTBE H3N2 MOI0 18h
HTBE H3N2 MOI5 3h
HTBE H3N2 MOI5 6h
HTBE H3N2 MOI5 12h
HTBE H3N2 MOI5 18h
HTBE H5N1 MOI0 3h
HTBE H5N1 MOI0 6h
HTBE H5N1 MOI0 12h
HTBE H5N1 MOI0 18h
HTBE H5N1 MOI5 3h
HTBE H5N1 MOI5 6h
HTBE H5N1 MOI5 12h
HTBE H5N1 MOI5 18h
MDM H1N1 MOI0 3h
MDM H1N1 MOI0 6h
MDM H1N1 MOI0 12h
MDM H1N1 MOI0 18h
MDM H1N1 MOI2 3h
MDM H1N1 MOI2 6h
MDM H1N1 MOI2 12h
MDM H1N1 MOI2 18h
MDM H3N2 MOI0 3h
MDM H3N2 MOI0 6h
MDM H3N2 MOI0 12h
MDM H3N2 MOI0 18h
MDM H3N2 MOI2 3h
MDM H3N2 MOI2 6h
MDM H3N2 MOI2 12h
MDM H3N2 MOI2 18h
MDM H5N1 MOI0 3h
MDM H5N1 MOI0 6h
MDM H5N1 MOI0 12h
MDM H5N1 MOI0 18h
MDM H5N1 MOI2 3h
MDM H5N1 MOI2 6h
MDM H5N1 MOI2 12h
MDM H5N1 MOI2 18h
MDM H5N1 MOI0 3h
MDM H5N1 MOI0 6h
MDM H5N1 MOI0 12h
MDM H5N1 MOI0 18h
MDM H5N1 MOI2 3h
MDM H5N1 MOI2 6h
MDM H5N1 MOI2 12h
MDM H5N1 MOI2 18h
virus
(3)
(2) (1)
(4)


a
HSV−1
influenza
position
treatment
Nalm6 DMSO
Nalm6 10 nM THZ1 24h
HCV−29 DMSO
HCV−29 30nm THZ1
HCV−29 100nm THZ1
DU145 DMSO
DU145 100nM THZ1 24h
LNCaP DMSO
LNCaP 100nM THZ1 24h
LNCaP non−target KD
LNCaP CDK7 KD
LNCaP MED1 KD
VCaP DMSO
VCaP 100nM THZ1 24h
VCaP non−target KD
VCaP CDK7 KD
VCaP MED1 KD
UM−Chor1 no treatment
UM−Chor1 500 nM THZ1 4h
UM−Chor1 500 nM THZ1 8h
UM−Chor1 500 nM THZ1 12h
UM−Chor1 500 nM THZ1 24h
UM−Chor1 control
UM−Chor1 100nm THZ1 4h
UM−Chor1 500nm THZ1 4h
UM−Chor1 10ng/mL Act−D
UM−Chor1 150ng/mL Act−D
KYSE510 DMSO
KYSE510 50 nM THZ1
KYSE510 200 nM THZ1
TE7 DMSO
TE7 50 nM THZ1
TE7 200 nM THZ1
MiaPaCa DMSO
MiaPaCa THZ1
PANC1 DMSO
PANC1 THZ1
BxPC3 DMSO
BxPC3 THZ1
C666 DMSO
C666 THZ1
HK1 DMSO
HK1 THZ1
HNE1 DMSO
HNE1 THZ1
(3) (2) (1) (4)
10
1
0.1
0.01
HSV−1
present
influenza
present
position
within circRNA
downstream
treatment
THZ1
CDK7 KD
MED1 KD
Act−D
control
b
UM−Chor1 no treatment
UM−Chor1 500 nM THZ1 4h
UM−Chor1 500 nM THZ1 8h
UM−Chor1 500 nM THZ1 12h
UM−Chor1 500 nM THZ1 24h
65190000
65195000
65200000
65205000
65210000
hsa_circ_0003812
NEAT1_2
NEAT1_1
c
LNCaP non−target KD
LNCaP CDK7 KD
LNCaP MED1 KD
VCaP non−target KD
VCaP CDK7 KD
VCaP MED1 KD
65190000
65195000
65200000
65205000
65210000
hsa_circ_0003812
NEAT1_2
NEAT1_1
d
% read−through
Nalm6 10 nM THZ1 24h
UM−Chor1 500 nM THZ1 4h
UM−Chor1 500 nM THZ1 8h
UM−Chor1 500 nM THZ1 12h
UM−Chor1 500 nM THZ1 24h
UM−Chor1 100nm THZ1 4h
UM−Chor1 500nm THZ1 4h
DU145 100nM THZ1 24h
LNCaP 100nM THZ1 24h
VCaP 100nM THZ1 24h
LNCaP CDK7 KD
LNCaP MED1 KD
VCaP CDK7 KD
VCaP MED1 KD