

Functional characterization of viral transcriptional regulators reveals extensive rewiring of host pathways

Received: 19 December 2025

Accepted: 3 September 2026

Published online: 17 September 2026

Check for updates

Jaice T. Rottenberg¹, Xing Liu¹, Mohamed Y. ElSadeq², Anna Lagani ^{1,3}, Luis F. Soto-Ugaldi⁴, James E. Corban ⁵, Phillip J. Dexheimer ^{6,7}, Matthew Hass^{6,7}, Clarissa Santoso^{1,3}, Cailing Yin^{6,7}, Berkay Engin¹, Ryan Lane ¹, Sakshi Shah ¹, Kerstin Spirohn-Fitzgerald ^{8,9,10}, Shubham Khetan ¹¹, Cheng-Che Lee¹, George Muñoz-Esquivel ¹², Zhaorong Li², Lucia Martinez-Cuesta^{1,13}, Yunwei Lu¹, Philipp Trollmann ², Tong Hao ^{8,9,10}, S. Stephen Yi^{14,15,16,17}, Nidhi Sahni ¹⁴, Martha L. Bulyk ^{8,11,18}, Michael A. Calderwood ^{8,9,10}, Leah Kottyan ^{6,7}, Matthew T. Weirauch ^{6,7}, Marc Vidal^{8,9,10}, Srivatsan Raman ^{19,20,21,22} & Juan I. Fuxman Bass ^{1,2,3,8,23} ✉

Viral transcriptional regulators (vTRs) reprogram host gene regulatory networks to promote replication, persistence, and immune evasion. Despite the identification of hundreds of vTRs in human viruses, how they rewire host pathways remains unclear. Here, we systematically profiled 95 vTRs from diverse human viruses across multiple functional assays. vTRs perturb immune, cell proliferation/death, and signaling pathways through various mechanisms; some bind DNA directly, others cooperate or antagonize human transcription factors (hTFs), and some remodel chromatin. vTRs can act as activators or repressors and recruit similar but not identical repertoires of proteins as hTFs. These findings reveal vTRs as versatile transcriptional modulators that converge on conserved host “pressure points” while diversifying across pathways to promote viral replication and persistence. Notably, many vTR dysregulate genes within autoimmune, neurological, and cardiovascular risk loci, revealing potential mechanistic links to disease. Together, we provide a large-scale resource for understanding and targeting viral control of human transcription.

Viruses are a leading cause of human disease, through direct effects of acute infection and by increasing the risk of chronic conditions, such as multiple sclerosis, dementia, or cancer^{1–5}. A key mechanism underlying viral pathogenicity is the ability to reprogram host gene expression to promote cell states suitable for viral replication or persistence⁶. In particular, viruses often suppress immune response genes to evade detection and clearance, as well as modulate pathways controlling cell survival and proliferation to prevent elimination of infected cells and enhance viral spread^{7–10}. This viral manipulation of host transcriptional

programs likely plays a central role in disease development and progression.

Viral proteins can modulate host gene expression either directly—by recruitment to cis-regulatory elements (CREs) in the host genome—or indirectly, by altering signaling pathways that control downstream genes⁶. Direct modulation typically involves proteins classified as viral transcription regulators (vTRs), a diverse group of multifunctional viral proteins whose activities affect transcription. Beyond controlling viral gene expression, vTRs can influence host

A full list of affiliations appears at the end of the paper. ✉ e-mail: fuxman@bu.edu

gene expression by binding DNA independently, co-binding with human transcription factors (hTFs), or being recruited to DNA through interactions with chromatin-associated proteins such as hTFs or cofactors (Fig. 1a). For instance, the Epstein-Barr virus (EBV) Zta binds TPA-responsive elements to regulate immune genes^{11,12}, while HBZ from human T-lymphotropic virus 1 (HTLV-1) dimerizes with hTFs JUNB and JUN to modulate their DNA binding activity¹³. In contrast, ChIP-seq studies have shown that vTRs such as EBV EBNA2 are indirectly recruited to human CREs through interactions with hTFs⁶.

More than 400 vTRs have been identified in human viruses, predominantly in double-stranded DNA (dsDNA) viruses⁶. Many vTRs are key mediators of viral pathogenesis, including well-characterized oncoproteins such as E1A from adenovirus, HBx from hepatitis B virus (HBV), and E6 and E7 from human papillomavirus (HPV)^{14–16}. Other vTRs primarily function as immunosuppressive factors, such as vIRF1-3 and Rta from Kaposi's Sarcoma-associated virus (KSHV)¹⁷. Notably, certain vTRs exhibit allele-specific binding, contributing to the pathogenesis of chronic or complex diseases. For example, the EBV EBNA2 is recruited to nearly half of the genetic risk loci associated with autoimmune disorders such as multiple sclerosis and systemic lupus erythematosus, where it shows evidence of allele-dependent recruitment^{18,19}.

Despite their critical role in disease, mechanisms of host transcriptional modulation are poorly understood for most vTRs. For

instance, genomic occupancy has been mapped for only 15 unique vTRs^{6,20,21}, and protein-protein interaction (PPI) data are available for just 74 of the ~400 vTRs²². Although recent transcriptional activity studies have been systematic, they have focused on 80 amino acid regions²³, potentially missing larger, synergistic, or context-dependent activation or repression domains. The lack of comprehensive studies comparing various molecular functions of vTRs across diverse viral families has significantly limited our understanding of their regulatory strategies and contributions to pathogenesis.

Here, we systematically characterize 95 vTRs from 31 human dsDNA viruses and retroviruses using a multimodal framework to define how vTRs reshape the human transcriptome and uncover the principles governing their regulatory functions. This vTR set spans different structural folds, timing of expression, reported functional roles, and likely mechanisms of action. We find that vTRs from phylogenetically distinct viruses often converge on core host pathways, including immunity, signal transduction, cell division, and apoptosis, revealing both shared and divergent regulatory strategies. This transcriptional rewiring is mediated by distinct mechanisms, including direct DNA binding (independently or cooperatively with hTFs), tethering to hTFs, and active disruption of hTF-DNA binding. We show that vTRs can function as transcriptional activators or repressors and engage in PPI networks that exhibit both shared and unique features compared to hTFs. This integrative analysis reveals previously

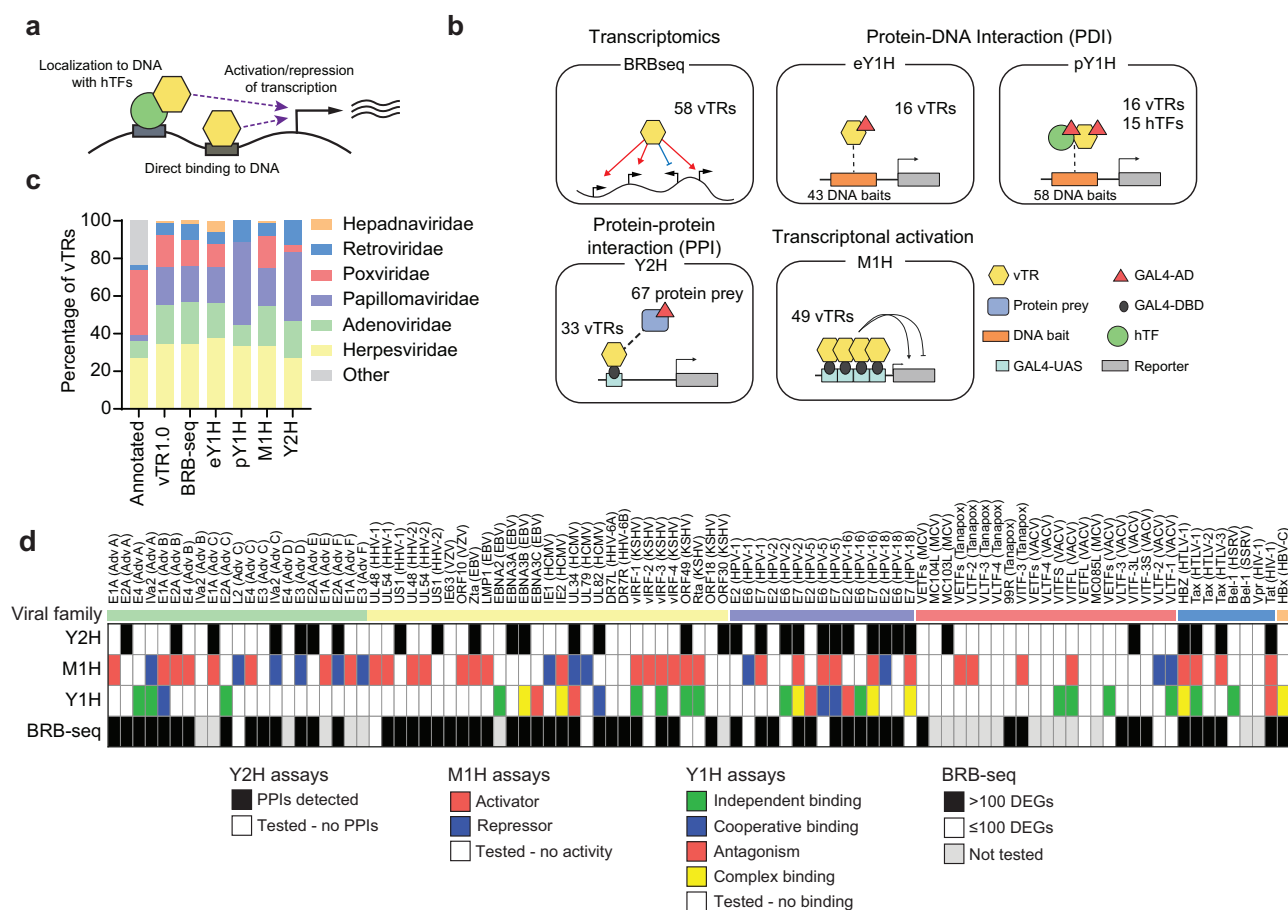


Fig. 1 | vTR1.0: a resource to study the viral transcriptional regulator functional landscape. **a** Schematic overview of vTR modulation of human transcription. **b** Overview of multimodal interrogation of vTR function. eY1H, enhanced yeast one-hybrid; Y2H, yeast two-hybrid; M1H, mammalian one-hybrid. Gal4-AD, Gal4 activation domain; Gal4-DBD, Gal4 DNA-binding domain; Gal4-UAS, Gal4 upstream activation sequence; vTR, viral transcriptional regulator; hTF, human transcription factor. **c** Percentage of vTRs from different

viral families represented across all annotated vTRs, the vTR1.0 resource, and tested in each experimental modality. **d** Data acquired for each vTR across experimental modalities. Columns correspond to vTRs (virus of origin in parentheses). Rows correspond to each experimental modality. Colored, black and white cells indicate data generated for a given vTR in each modality (based on parameters and thresholds for each experiment). Gray indicated vTRs not tested by the assay.

unappreciated redundancies and division of labor among vTRs encoded by the same virus, as well as striking functional divergence among orthologs. Finally, we identify links between vTR-modulated genes and disease risk loci, suggesting that viral regulators may act as environmental sensitizers of genetic susceptibility. Collectively, our work highlights vTRs as a functionally diverse set of regulators with implications for viral pathogenesis and fundamental principles of transcriptional control.

Results

vTRL0: a resource to study the vTR functional landscape

Similar to human transcriptional regulators, vTRs function through direct or indirect recruitment to DNA regions and modulate gene expression through interactions with cofactors and hTFs. To determine the mechanisms by which different vTRs perturb gene regulatory networks, we used a multimodal functional approach integrating transcriptomics, biophysical interactions, and measurements of transcriptional activity (Fig. 1b). We selected a subset of 95 vTRs from the annotated compendium of 419 vTRs for gene synthesis (Supplementary Data 1). These vTRs span all major dsDNA virus and retrovirus families (Fig. 1c) and include several vTRs implicated in viral pathogenesis, such as E6 and E7 from different HPV strains, EBNA2 from EBV, and E1A from adenovirus. Codon-optimized vTR open reading frames (ORFs) were cloned into Gateway-compatible donor vectors, resulting in the vTRL0 entry clone collection, which can be readily transferred to destination vectors for various functional assays (Fig. 1b, c).

To characterize the impact of vTRs on host gene expression, we performed Bulk RNA Barcoding and sequencing (BRB-seq)²⁴ in HEK293T cells transiently expressing 84 different vTRs. Using stringent criteria (>100 differentially expressed genes, $|\log_2\text{FC}| > 0.5$, $p < 0.05$), 58 vTRs induced substantial transcriptional changes, generating a transcriptomic dataset that captures diverse host responses to vTRs.

To evaluate the intrinsic ability of vTRs to bind to DNA regions (i.e., in the absence of other hTFs), we used enhanced yeast one-hybrid (eY1H) assays^{25,26}, a yeast-based reporter assay that detects protein-DNA interactions (PDIs) in a high-throughput manner. We tested the entire vTRL0 collection against a set of 112 DNA-baits derived from the promoter elements (2 kb regions upstream of the transcription start site) of cytokine genes, which are heavily targeted by many viruses^{27,28}. To assess the possible co-binding or sequestration between hTFs and vTRs, we used paired yeast one-hybrid (pY1H) assays²⁸, a variation of eY1H that evaluates pairs of proteins for synergistic or antagonistic binding to DNA. In particular, we tested 113 vTR-hTF pairs—selected based on known PPIs, PPIs identified in this study, and PPIs involving vTR orthologs—against 39 immune gene promoters and 83 promoters of genes involved in cell signaling, tumor suppression, oncogenesis, differentiation, metabolism, replication, and apoptosis. These experiments, together with previous studies, show that vTRs are recruited to DNA elements through multiple binding modalities (direct, cooperative, antagonistic, and indirect recruitment).

To measure the intrinsic transcriptional activity of vTRs, we performed mammalian one-hybrid (MIH) assays²⁹ in HEK293T cells, a method in which proteins are tethered to a firefly luciferase reporter construct to quantify transcriptional activation or repression. Using MIH assays with the entire vTRL0 collection, we identified 36 activator and 13 repressor vTRs. Finally, we used yeast two-hybrid (Y2H) assays to screen vTRL0 against the human ORFeome v9.1³⁰, comprised of 17,472 protein-coding genes, and identified 132 PPIs between 33 vTRs and 67 human proteins.

Although nearly all vTRs were tested across the various functional assays, only a subset exhibited detectable changes in host gene expression, intermolecular interactions, or transcriptional activation/repression (Fig. 1d and Supplementary Fig. 1). Nevertheless, the vTRs with positive results remained broadly representative of the viral

family composition of vTRL0 (Fig. 1c). Across all assays, 91% of tested vTRs exhibited at least one measurable molecular activity, indicating that the vast majority of vTRs engage host regulatory systems through detectable molecular mechanisms and establishing vTRL0 as a resource for systematically comparing how diverse viruses modulate human gene regulation.

vTRs modulate multiple host pathways

To determine the pathways rewired by different vTRs, we used HEK293T cells as a standardized human background to allow systematic comparison across the vTRs profiled. This cell line provides high transfection efficiency and stable, well-characterized transcriptional programs, and low activation by plasmid transfection, enabling robust detection of vTR-driven perturbations. Moreover, HEK293T cells express a broad set of transcriptional cofactors and signaling components, well-suited for other assays within our multimodal profiling. We expressed individual vTRs in HEK293T cells for 72 h, determined the transcriptomic profile using BRB-seq, and identified the set of differentially expressed genes relative to the empty vector. Correlation between replicate experiments was high ($r \approx 0.98$), illustrating the high quality of our transcriptomics dataset (Supplementary Fig. 2). Although we observed some variability in expression across vTRs, we did not detect a correlation between vTR abundance and the number of differentially expressed genes (Supplementary Fig. 3a, b). This argues against vTR expression level being a major determinant of the observed transcriptional effects and suggests that differences among vTRs primarily reflect their underlying regulatory activities.

Gene set enrichment analysis of differentially expressed genes shows that vTRs generally dysregulate pathways associated with immune response, apoptosis, cell cycle, DNA damage repair, and metabolism (Fig. 2a and Supplementary Data 2). These pathways have been broadly associated with viral replication and persistence, and some are known to be affected by vTRs; for example, KSHV vIRFs are known to inhibit p53 activity^{31,32}. Consistent with this, vIRF1/3/4 downregulated several pathways associated with p53 regulation, as well as downstream pathways including those regulating cell cycle checkpoints and DNA double-stranded break repair (Fig. 2a). Similarly, HPV E6 and E7, commonly associated with suppression of apoptosis and innate immunity and the upregulation of cell cycle genes^{33,34}, substantially modulated these pathways.

We next evaluated whether vTRs from the same viral family, vTR homologs from related viruses, and vTRs from the same virus broadly affect similar or different pathways. To accomplish this, we calculated the Pearson correlation coefficient of pathway-normalized enrichment scores for each pair of vTRs. We observed that vTR pairs have an overall low correlation between affected pathways, consistent with the diverse functions and pathways affected by different vTRs, whereas homologous vTRs and vTRs from the same virus have more similar transcriptional effects (Fig. 2b). For example, IE1 and IE2, both expressed from the human cytomegalovirus (HCMV) major immediate early promoter, have highly concordant effects on host pathways, including upregulation of signal transduction and downregulation of metabolism pathways (Fig. 2c). The broad distribution of correlation values between vTRs from the same virus, however, suggests that division of labor, redundancy, cooperation, and opposing effects between vTRs may all be employed by viruses to robustly rewire host cellular pathways. For instance, E1A and IVa2 from adenovirus A exhibit opposite regulation of signal transduction and RNA processing related pathways and mutually exclusive regulation of cell cycle (E1A) and metabolism (IVa2) (Fig. 2d and Supplementary Data 2). The contrasting transcriptional effects of these two vTRs combined with their distinct expression kinetics (E1A - immediate early, IVa2 - late) underscore the specific requirements for each progressive stage of adenovirus infection. Together, this illustrates how vTRs from the same virus may have

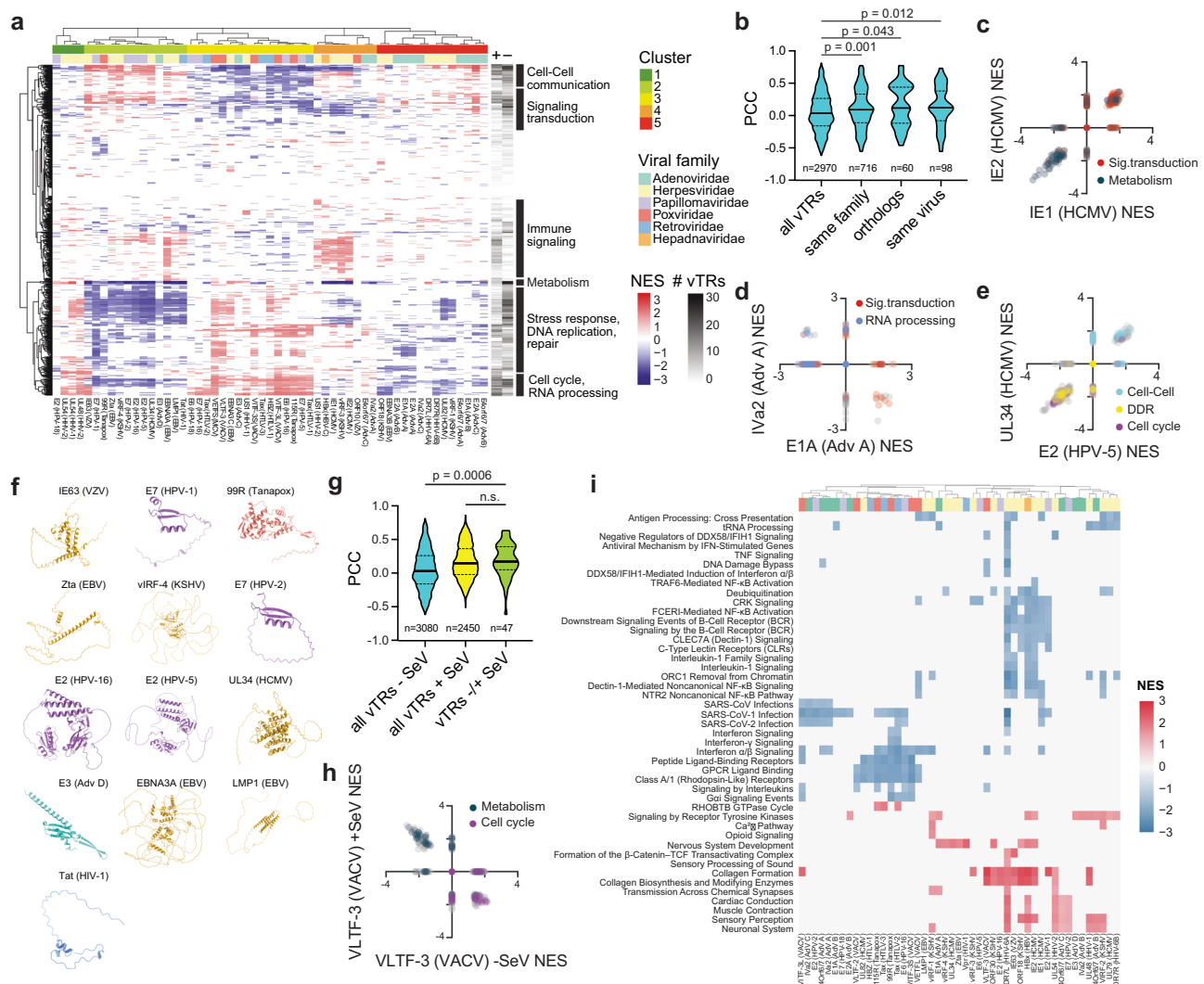


Fig. 2 | Perturbation of human gene expression by vTRs. a Aggregated results of gene set enrichment analysis from cells expressing different vTRs. vTR are clustered based on the CSI-corrected Pearson correlation coefficient. Rows correspond to gene sets clustered based on gene similarity. Normalized enrichment scores (NES) are represented for each vTR as a color gradient. The number of vTRs with a positive or negative NES per pathway is indicated in gray. **b** Distribution of Pearson correlation coefficients (PCC) for all pairs of vTRs, vTRs from the same family, vTR orthologs, and vTRs from the same virus. Solid lines denote the median and dashed lines denote quartiles. Statistical significance determined by two-sided Mann-Whitney's U test. **c–e** Scatter plots comparing the NES of pathways dysregulated by pairs of vTRs. Selected pathways are colored. **c** HCMV IE1 versus IE2. **d** Adenovirus A

EIA versus IVA2. **e** HPV-5 E2 versus HCMV UL34. **f** AlphaFold2 models for vTRs from Cluster 2 in Fig. 2A. Color indicates viral family. **g** Distribution of PCC scores for all vTRs in resting cells, all vTRs tested in Sendai Virus-infected cells, and vTRs compared between untreated and Sendai Virus-infected cells. SeV - Sendai Virus. Solid lines denote the median and dashed lines denote quartiles. Statistical significance determined by two-sided Mann-Whitney's U test. **h** Scatter plot comparing the NES of pathways dysregulated by vaccinia virus (VACV) VLTf-3 in untreated and Sendai Virus-infected cells. **i** Heatmap of NES from GSEA comparing SeV infection versus no infection in vector and vTR conditions. Pathways shown are significantly altered by SeV infection and display opposite responses between vector and vTR.

similar or different transcriptional effects depending on their expression timing during infection.

Interestingly, we identified 77 vTR-pairs that produce highly similar transcriptional profiles ($PCC > 0.5$) from viruses belonging to different viral families. For example, HPV-5 E2 and HCMV UL34 upregulate pathways involved in cell-cell communication and downregulate those associated with cell cycle control and DNA damage repair (Fig. 2e). This marked degree of pathway similarity between HPV-5 E2 and HCMV UL34, as well as other vTRs with similar transcriptional profiles, is not attributable to similarities in structure. For example, vTRs from cluster 2 display very different folds determined by AlphaFold2 (Fig. 2f). Together, these results illustrate how vTRs from distant, unrelated viruses can have similar transcriptional effects, suggesting that viruses have converged to affect overlapping host pathways while retaining specific effects on viral gene regulation.

More generally, we found 5 clusters of vTRs with overall similar transcriptional profiles (Fig. 2a). As expected, homologous vTRs are generally clustered together, illustrating the quality of our clustering approach. Of the clusters identified, Clusters 2 and 3 are notable for exhibiting somewhat opposite effects. vTRs from Cluster 2 downregulate genes involved in cellular stress, such as apoptosis, DNA damage response, hypoxia, and chemical stress, and those related to cell replication, and upregulate immune, cell-cell communication, and intracellular signaling pathways. Instead, vTRs in Cluster 3 include many oncoproteins (e.g., HPV-18 E6, HPV-16/HPV-5 E7, and HTLV-1 Tax^{4,35}) and upregulate pathways related to cell cycle progression, DNA damage, and stress responses, while downregulating those involved in extracellular signaling and antigen presentation. Interestingly, several viruses, such as HPV and EBV, encode vTRs present in both Clusters 2 and 3, suggesting potential balancing mechanisms controlling cell

growth and stress at different stages in the viral replication cycle. In addition, each cluster is composed of vTRs from different viruses and families. Together, our results suggest that unrelated viruses have evolved convergent transcriptional strategies to manipulate common host processes, while retaining distinct programs tailored to their specific requirements for persistence and replication.

vTR regulation of pathways in infected cells

Viral infection greatly affects cell states and signaling pathways, some of which are leveraged by the virus to promote replication or persistence, while others reflect host antiviral responses. Sendai virus (SeV) provides a robust and reproducible activation of antiviral and inflammatory signaling pathways without introducing confounding vTRs, enabling us to assess vTR function in an infection-relevant cellular context. To determine how vTRs function in infected cells, we performed BRB-seq in HEK293T cells infected with SeV following transient transfection of 49 vTRs. Differences in host transcriptional responses were not attributable to variations in SeV infection, as SeV transcript levels were comparable across vTR-expressing and control samples and showed no correlation with vTR abundance (Supplementary Fig. 3c–e). This indicates that the observed effects on gene expression reflect vTR activity rather than differences in viral infection levels.

Overall, transcriptional responses induced by the same vTR in unstimulated and SeV-infected cells were significantly more similar than expected from random pairwise comparisons between different vTRs, suggesting that many vTRs retain characteristic regulatory activities across conditions (Fig. 2g). However, the distribution of correlations was broad, with some vTRs showing weak concordance or even anticorrelated responses, highlighting substantial context dependence in vTR function. For example, VLTf-3 (vaccinia virus) shifts from repressing to activating metabolic/biosynthetic pathways and from activating to repressing cell cycle pathways in SeV-infected cells (PCC = −0.61; Fig. 2h), which may enable the virus to adapt to different cellular states.

Several vTRs, including IE1 (HCMV) and KSHV vIRFs, are known to act as immunomodulatory proteins^{36–39}. To identify additional vTRs that may counteract the antiviral effects elicited by infection, we focused on 45 pathways that are altered under SeV infection relative to resting cells, including interferon response, interleukin signaling, NF- κ B activation, and antigen presentation. We identified 47 vTRs that affect at least one of these pathways in SeV-infected cells (Fig. 2i), of which 39 downregulate immune-associated pathways normally upregulated in the presence of SeV. While several of these associations are known, such as vIRF-1 affecting interferon pathways^{38,39} and IE1/2 downregulating the T-cell receptor and interleukins^{36,37}, our results also identify several additional immunoregulatory functions. For example, Tax proteins from HTLV-2 and HTLV-3, typically associated with NF- κ B regulation⁴⁰, both downregulate type I interferon signaling, whereas DR7L (HHV-6A) and IE63 (Varicella-Zoster virus - VZV) downregulate IL-1-associated pathways (Fig. 2i). In addition to immunomodulatory roles, 31 vTRs upregulate signaling pathways normally suppressed during active infection (Fig. 2i), which may reflect indirect immune evasion strategies that attempt to reestablish cellular processes needed for replication.

To determine whether the immunomodulatory effects observed during SeV infection reflected general antiviral signaling rather than virus-specific responses, we selected 21 vTRs that altered immune-associated pathways in SeV-infected cells and evaluated them in HEK293 cells stimulated with poly(dA:dT), which activates cGAS-STING and RNA polymerase III/RIG-I signaling. HEK293 cells were used because, unlike HEK293T cells, they mount a measurable transcriptional response to poly(dA:dT). Focusing on genes that were induced by both SeV infection and poly(dA:dT) stimulation, including multiple type I and III interferons and inflammatory cytokines, we observed broad concordance in vTR activity across the two conditions

(Supplementary Fig. 4). In particular, several vTRs consistently suppressed expression of IFNL1, IFNL2, IFNL3, and IFNB1 in both stimulation conditions, including Tax (HTLV-3), vIRF-3 and vIRF-4 (KSHV), LMP1 (EBV), UL82 and IE2 (HCMV), and IE63 (VZV). These findings indicate that many vTR-mediated effects on antiviral gene expression are not specific to SeV infection but instead target common host innate immune pathways to reshape antiviral signaling.

vTRs are recruited to host DNA through multiple mechanisms

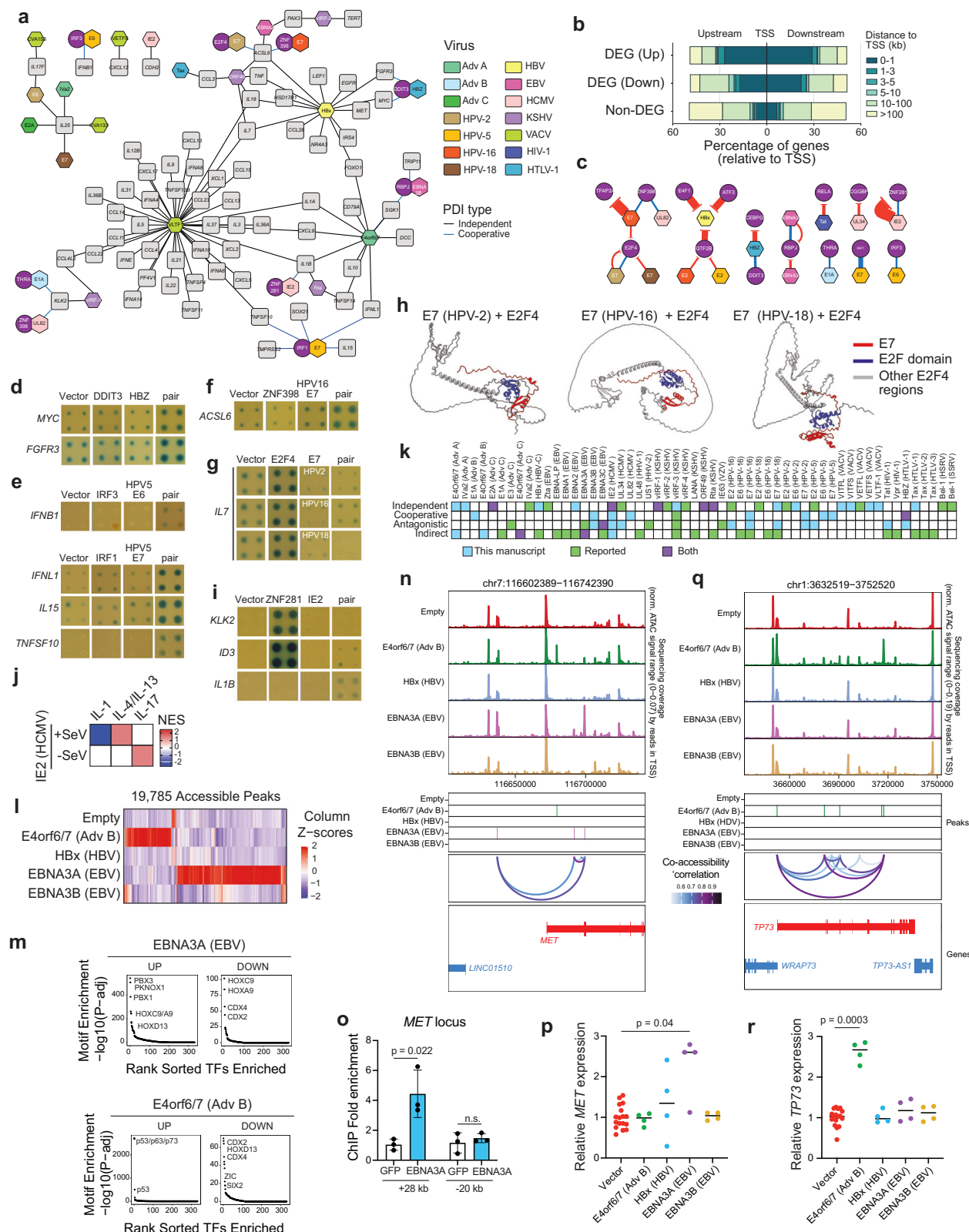
Transcriptional regulators can be recruited to DNA elements using different mechanisms, including direct binding to DNA (either independently or as part of a protein complex) and indirect recruitment mediated by DNA-binding factors. We previously curated 33 ChIP-seq data sets corresponding to 12 vTRs, which showed large variations in target preference and binding distance to transcription start sites⁶. This study showed that ChIP-seq peaks of many vTRs are enriched for hTF motifs, suggesting indirect recruitment of vTRs to genomic sites. Other vTRs have been shown to bind DNA directly, including bZIP Zta (EBV) and E2A (adenovirus)^{12,41}.

To identify which vTRs are recruited to DNA directly and through which mechanisms, we first used PADIT-seq⁴² to assay the DNA-binding activities of ten vTRs to all possible 10-bp DNA sequences. These vTRs were selected based on their likelihood of directly binding to DNA given domain presence (e.g. viral IRFs), YIH data, or previous vTR-DNA interaction data²⁸. To facilitate protein expression, we expressed structured vTR regions and 10–15 flanking amino acids, as previously conducted for hTFs⁴². However, none of these vTR constructs showed significant binding to 10-bp DNA sequences, while our hTF positive controls did. The observed lack of DNA binding activity may be due to misfolding of the vTRs' DNA binding domains in the *in vitro* transcription and translation system, the need for excluded vTR regions for DNA binding and/or PPIs, the interference by the ALFA tag fusion required by the assay, that these vTRs do not bind with sufficient affinity to the 10-mers tested in the PADIT-seq assays, or that the vTRs associate indirectly, rather than directly, with their target genomic sites.

To address these possibilities, we performed eYIH assays, which evaluate interactions between individual proteins and DNA regions of up to 2 kb in the absence of other human or viral proteins. We tested 95 vTRs against 112 immune gene promoters, given that most vTRs dysregulate at least one immune pathway. We also assessed cooperativity with hTFs using pYIH assays, a modified eYIH approach that evaluates DNA binding of protein pairs. We tested 113 vTR-hTF pairs selected based on reported PPIs, homologs of known vTR-hTF interactions, and PPIs identified in this study. Both individual vTRs/hTFs and their pairs were assayed against 39 cytokine promoters and 83 promoters of genes from the Cancer Gene Census, involved in cell proliferation, cell death, and cell signaling, all pathways dysregulated by vTRs (Fig. 3a).

We identified 17 vTRs (18% of vTRs tested) that bind human promoters, with Vaccinia VLTf-1 and hepatitis B X-protein (HBx) having the most PDIs (Fig. 3a and Supplementary Data 3, 4). This proportion is similar to the proportion of hTFs binding these regions (24%)²⁷, suggesting that many vTRs might not rely on hTFs for DNA recruitment, in contrast to previous studies⁶. We validated two of these interactions—EBV EBNA2 binding to the *PAX3* promoter and EBV EBNA3B binding to the *TRIP11* promoter—by ChIP-qPCR (Supplementary Fig. 5a, b). We also confirmed binding of HCMV IE2 to the *CDH2* promoter, a gene previously reported to be upregulated during HCMV infection⁴³, using ChIP-seq (Supplementary Fig. 5c). Together, these results provide independent support for the PDIs identified by YIH assays.

To gain broader insight into how IE2 engages host chromatin beyond the individual promoters identified by YIH assays, we analyzed the genomic distribution of IE2 ChIP-seq peaks. IE2 is a multifunctional vTR known to coordinate viral gene expression and reprogram host pathways involved in cell proliferation, DNA damage responses, and



signaling during HCMV infection^{44–46}. We found that 64% of the 10,217 peaks detected occur within 1 kb of transcription start sites, indicating a strong preference for promoter-proximal regulatory regions, particularly among genes upregulated by IE2 expression (Fig. 3b and Supplementary Fig. 5d). In addition to the expected enrichment for cell-cycle, DNA replication, DNA repair, and RNA-processing pathways, IE2-

bound genes were enriched for developmental and signaling programs, including ROBO/SLIT and Notch signaling, receptor tyrosine kinase and GPCR signaling, extracellular matrix organization, and glycosaminoglycan metabolism (Supplementary Fig. 5e). Several of these pathways were also transcriptionally altered by IE2 expression (Fig. 2a), suggesting that promoter-proximal binding contributes

Fig. 3 | vTRs bind host regulatory elements and alter chromatin structure. **a** Protein-DNA interactions for vTRs (left), vTRs cooperating with hTFs (right), and human promoter sequences. Hexagons represent vTRs and are colored by virus. Purple circles represent hTFs. Squares represent human promoters. **b** Percentage of genes upregulated, downregulated, or not regulated by HCMV IE2 with IE2 ChIP-seq peaks at different positions relative to their transcription start site (TSS). **c** Network of relationships between hTFs and vTRs. Blue edges represent cooperativity, red edges represent antagonism. Edge thickness indicates the number of promoters at which a specific interaction occurs. **d–g, i** Representative pYIH interactions for hTF-vTR pairs. Italicized gene names (left) indicate the tested promoter bait. Yeast strains shown: empty vector control, hTF only, vTR only, hTF-vTR pair. **d** Cooperativity between DDIT3 and HBZ (HTLV-1) at the *MYC* and *FGFR3* promoters. **e** Cooperativity between IRF3 and E6 (HPV-5) at the *IFNB1* promoter. Cooperativity between IRF1 and E7 (HPV-5) at the *IFNL1*, *IL15*, and *TNFSF10* promoters. **f** Cooperativity between ZNF398 and E7 (HPV-16) at the *ACSL6* promoter. **g** Antagonism of E2F4 by multiple E7 homologs at the *IL7* promoter. **i** Antagonism of ZNF281 by IE2 (HCMV) at *KLK2* and *ID3* promoters; cooperativity between ZNF281 and IE2 at the *IL1B* promoter. **h** AlphaFold3 models of interactions between E2F4 and E7 from different HPV strains. **j** Heatmap of NES scores for Interleukin pathways derived from BRB-seq data for IE2 (HCMV) in SeV-infected and uninfected

cells. **k** vTR DNA binding modalities described in this manuscript (light blue), reported in the literature (green), or both (purple). **l** Heatmap of z-score normalized chromatin accessibility at 19,785 peaks that are differentially accessible ($FDR \leq 0.1$ and $\log_2(FC) \geq 1$) in at least one vTR surveyed via PROD-ATAC comparing E4ORF6/7 (Adv B), HBx (HBV), EBNA3A (EBV), and EBNA3B (EBV). Columns of peaks are hierarchically clustered. **m** Representative DNA motif enrichment within increased/decreased differentially accessible peaks for EBV EBNA3A and Adenovirus B E4ORF6/7. Displayed values were calculated using hypergeometric tests and plotted relative to the ranked order of all queried motifs. **n, q** Genome browser tracks displaying sequencing coverage, significant chromatin peaks ($FDR \leq 0.1$ and $\log_2(FC) \geq 1$), and co-accessibility correlation values for vTR peaks in the *MET* (**n**) and *TP73* (**q**) loci. **o** ChIP-qPCR enrichment of EBNA3A relative to GFP control in the *MET* locus. *MET* enhancer (+28 kb): chr7:116,699,142–116,699,226; *MET* negative control region (–20 kb): chr7:116,651,717–116,651,804. Statistical significance determined by two-sided Student's *t* test. Error bars correspond to mean $\pm$ SD of three biological replicates. **p, r** *MET* (**p**) and *TP73* (**r**) expression in HEK293T cells expressing different vTRs relative to vector control. Statistical significance determined by a two-sided Student's *t* test. Biological replicates: vector ($n = 17$), vTRs ($n = 4$).

directly to both canonical cell-cycle programs and broader changes in host signaling, cell adhesion, and developmental regulatory networks. Similarly, analysis of publicly available ChIP-seq datasets revealed significant enrichment of binding by multiple vTRs at the promoters (within 1 kb of the transcription start site) of genes whose expression was altered by the corresponding vTR (Supplementary Fig. 5f), supporting a direct role for promoter occupancy in the transcriptional regulation by many vTRs.

We also identified 9 vTRs that bind DNA cooperatively with hTFs (Fig. 3a, c). For example, we observed co-binding between bZIP HBZ (HTLV-1) and the human bZIP DDIT3 at the *MYC* and *FGFR3* promoters (Fig. 3d), two genes heavily involved in oncogenic pathways. This is consistent with studies showing that HTLV-1 activates *MYC* expression in adult T-cell leukemias⁴⁷. We also found cooperative interactions between HPV-5 E6 and IRF3 at the *IFNB1* promoter, and between HPV-5 E7 and IRF1 at the *IFNL1*, *IL15*, and *TNFSF10* promoters (Fig. 3e), in line with known immunomodulatory roles of E6 and E7^{48–50}. Considering that E6 and E7 are known to inhibit transcriptional activation by IRF3 and IRF1, respectively^{50,51}, our results suggest that E6 and E7 may outcompete IRF dimers to prevent activation of immune response genes. Interestingly, HPV-16 E7 cooperatively binds to DNA with ZNF398 (Fig. 3f), an oncogenic hTF that promotes tumorigenesis through p53 degradation, enhanced cellular metabolism, disruption of ferroptosis, and maintenance of cancer stem cell niches^{52–56}. Our results suggest that E7 may leverage these ZNF398 functions to promote cell proliferation and prevent cell death.

In addition to the detected vTR binding events, we observed 12 vTR-hTF pairs in which a vTR antagonized the binding of an hTF at specific promoters (Fig. 3c). This likely occurs by sequestration, as AlphaFold3 modeling predicts binding of the vTRs to the DNA-binding domain of the interacting hTFs (Supplementary Fig. 6). For example, E2F4 is antagonized by several HPV E7 orthologs, which are predicted to interact with the DNA-contacting interface of the E2F domain (Fig. 3c, g, h). Ingenuity pathway analysis of our BRB-seq data shows that expression of HPV-16 E7 leads to downregulation of E2F targets (z-score = –2.164, $p = 0.0022$). HPV E7 is known to bind the tumor suppressor retinoblastoma protein (RB1) and prevent its association with E2F proteins^{57,58}. Together with our results, this suggests that E7 may disrupt E2F function through two separate mechanisms: inhibition of RB1-E2F association and sequestration of E2F, preventing its association with DNA. We observe a similar pattern for TFAP2A, which was antagonized by HPV-16 E7 at 15 promoters (Fig. 3c) and was previously shown to be downregulated in cervical cancer cell lines expressing E7⁵⁹.

We also observed 36 instances in which HCMV IE2 antagonized the DNA binding of ZNF281 likely by forming an interlocked complex between both proteins that occludes the cyst2His2 zinc fingers of the DNA binding domain (Fig. 3c, i and Supplementary Figs. 6, 7). ZNF281 is an oncogenic hTF associated with several types of cancer and plays a role in inflammation^{60–63}. Interestingly, we detected a cooperative interaction between IE2 and ZNF281 at the *IL1B* promoter (Fig. 3c, i). Although IE2 is primarily an activator, we propose that recruitment of ZNF281 may result in a repressive complex. Indeed, we observed that IE2 repressed IL1-related pathways in our transcriptomics data from SeV-infected cells (Fig. 3j). Our results suggest that IE2 may co-opt ZNF281 to suppress interleukin response during active infection states. This mechanism is distinct and complementary to reported immune suppressor activities of IE2 mediating STING and CD83 degradation, activation of the IL10/STAT3 signaling pathway, and inhibition of antigen presentation^{64–67}.

Together, our results extend previous reports on vTR recruitment to host DNA and show that in addition to indirect and independent DNA binding, vTRs may reprogram host gene expression through cooperative or antagonistic interactions with hTFs (Fig. 3k). Interestingly, we show that 18 vTRs may modulate expression through multiple mechanisms, potentially expanding their regulatory targets (Fig. 3k). This is likely an underestimate given the limited number of DNA sequences tested by YIH assays and the sparsity in currently available vTR ChIP-seq datasets. Altogether, our data uncover an unexpectedly broad repertoire of DNA-engagement mechanisms, showing that vTRs can operate as independent DNA binders, partners that co-opt hTFs, or antagonists that block hTF interactions, often combining multiple modes to rewire host gene expression.

vTRs alter chromatin accessibility

In addition to binding to open regions in the host genome, vTRs can drive changes in chromatin architecture. While these effects have been demonstrated for several vTRs, including Adenovirus E1A, EBV BZLF1, and HIV Tat^{68–70}, the effects of many vTRs on chromatin accessibility remain largely uncharacterized. We performed PROD-ATAC, a single-cell ATAC-seq approach that evaluates the effect of protein overexpression on chromatin accessibility⁷¹, to determine the accessibility landscapes generated by EBV EBNA3A and EBNA3B, Adenovirus B E4ORF6/7, and HBV X-protein expression in HEK293T cells (Supplementary Fig. 8a, b). The data passed strict quality metrics of fragments per cell, reads in peaks, and transcription start site enrichment scores (Supplementary Fig. 8a). Of the vTRs evaluated, EBNA3A produced the largest number of differentially accessible peaks (Fig. 3l), with newly

opened regions enriched for PBX, CDX, and HOX motifs (Fig. 3m). This is consistent with a recent study showing that EBNA3A ChIP-seq peaks and newly opened regions (ATAC-seq) are enriched for PBX motifs²⁰. We also found that regions closed in EBNA3A-expressing cells are enriched for HOX and CDX motifs, though these regions are different from those with HOX/CDX motifs opened in EBNA3A-expressing cells, suggesting that EBNA3A induces the remodeling of chromatin occupied by different homeodomain proteins. To determine whether these accessibility changes are associated with direct regulation of host genes, we focused on the *MET* locus, which contains three regions that become more accessible upon EBNA3A expression (Fig. 3n). ChIP-qPCR confirmed EBNA3A binding to one of these regions located approximately 28 kb downstream of the *MET* transcription start site (Fig. 3o), indicating that EBNA3A directly associates with a newly accessible regulatory element. Increased accessibility at the *MET* locus was accompanied by elevated *MET* expression (Fig. 3p), suggesting that EBNA3A activates *MET* through local chromatin remodeling.

We also detected marked changes in accessibility by adenovirus B E4ORF6/7, which exhibit minimal overlap with those driven by EBNA3A (Fig. 3l). Newly opened chromatin induced by E4ORF6/7 was enriched for p53/p63/p73 motifs (Fig. 3m), whereas closed regions were dominated by HOX/CDX motifs. E4ORF6/7 has been reported to interact with p53; however, unlike its larger splice variant E4ORF6, it does not degrade p53^{72,73}. Notably, E4ORF6/7 induced accessibility at four regions within the *TP73* locus, resulting in a threefold increase in *TP73* expression (Fig. 3q, r). Together with our results, this supports a model in which E4ORF6/7 from Adenovirus B promotes activation of p53-family regulatory networks, occupancy of p53/p63/p73 binding sites, and potentially modulates downstream targets. Indeed, ingenuity pathway analysis shows that Adenovirus B E4ORF6/7 expression leads to the upregulation of p53-responsive genes ($z\text{-score} = 2.092$, $p = 4.2 \times 10^{-5}$). Interestingly, both E4 splice variants are co-expressed early during infection, although E4ORF6/7 is less stable⁷⁴, suggesting that the adenoviral lifecycle may require distinct regulation of p53 and related family members at specific time points during infection.

vTRs have intrinsic transcriptional effector activity

Like hTFs, many vTRs contain effector domains that recruit transcriptional cofactors to activate or repress target gene expression⁷⁵. A previous study using HT-recruit, a high-throughput tiling peptide approach to screen for transcriptional activity, identified hundreds of viral effector domains²³. Because effector domains may function differently in the context of a full-length protein structure, and because some effector domains can be larger than those examined using HT-recruit (80 amino acids)^{76,77}, we interrogated the transcriptional activity of 95 full-length vTRs using MIH assays. MIH uses a luciferase reporter downstream of an enhancer element that contains 4 copies of the Gal4 upstream activating sequence (UAS). A protein (e.g., a full-length vTR) fused to the Gal4 DNA-binding domain is then recruited to the reporter construct and may activate or repress luciferase expression. We used two versions of the MIH assay, using either a minimal or a moderate promoter between the UAS and luciferase, to better capture both activation and repression. Using our minimal promoter system, we observed that 25.2% of vTRs tested are strong transcriptional activators ($\text{Log}_2(\text{fold change}) \geq 5$), compared to only 6.75% for a set of 224 hTFs with broad representation across families and that was evaluated using the same MIH system (Fig. 4a)⁷⁷. This suggests that vTRs may have evolved potent effector functions, enabling viruses to more effectively reprogram viral and host transcriptional networks.

Using our moderate promoter MIH system, we identified 36 activator and 13 repressor vTRs (Fig. 4b and Supplementary Data 5). We detected a significant correlation between MIH activity and the density of activation domains previously determined by HT-recruit (Supplementary Fig. 9a)²³. We also used ADpred to predict activation domains within vTRs and found that activator vTRs—in particular,

strong activators—were significantly enriched for predicted activation domains relative to non-activators or repressor vTRs, providing independent support for the MIH-based functional classifications (Supplementary Fig. 9b, c). We observed different distributions in vTR activity across viral families. Many highly active vTRs belong to herpesviruses or adenoviruses, the latter of which also account for half of repressive vTRs. Comparatively, poxvirus, papillomavirus, and retrovirus vTRs show milder transcriptional effects, which may reflect different roles in direct regulation of transcription.

We next compared the transcriptional activity of vTRs expressed at different stages of the viral lifecycle, and observed striking differences in activity between these groups (Fig. 4c). Immediate-early vTRs, which play integral roles initiating viral expression cascades, have the highest activity among all groups. We also observed that many vTRs expressed late in the lytic cycle act as repressors rather than activators. Notable exceptions are highly active tegument proteins UL48 (HHV-1/2), which although late expressed, initiate transcription of immediate-early genes upon infection or reactivation (Fig. 4c)⁷⁸. Finally, vTRs expressed during latency exhibit low levels of transcriptional activity, suggesting that these vTRs may elicit more subtle or context-specific effects on viral and host expression to favor persistence.

vTRs interact with overlapping but distinct sets of proteins compared to hTFs

hTFs regulate gene expression through interactions with cofactors, chromatin modifying enzymes, other hTFs, and members of the transcription preinitiation complex^{79–83}. To identify potential mechanistic similarities and differences between vTRs and hTFs, we compared their PPI profiles. Many PPIs between viral and host proteins were previously determined^{22,84,85}, mostly using affinity-based methods such as IP-MS; however, these methods tend to favor the detection of abundant proteins and often identify indirect interactors. Because we expected vTRs to interact with hTFs, which are generally lowly abundant, we performed Y2H assays, which do not depend on in vivo protein abundance, as proteins are expressed using a defined promoter. We performed Y2H assays between each vTR in the vTR1.0 collection and the complete human ORFeome (~17,500 proteins)³⁰. Positive hits were retested pairwise in a matrix format to enable PPI comparisons between vTRs. In total, we identified 132 PPIs between 33 vTRs and 67 human proteins, 13 of which had been identified in earlier studies (Fig. 1d and Supplementary Data 6). Although the overlap with previously reported AP-MS/IP-MS datasets was modest, this is consistent with the complementary interaction modalities measured by Y2H and affinity purification approaches, which typically recover largely distinct but overlapping subsets of the interactome^{30,86}.

After integrating our Y2H with available PPI data, we observed that 31.2% of vTR interactions involve hTFs or cofactors, significantly more than for other viral proteins and similar to hTFs (Fig. 4d, e). Strikingly, we detected previously unreported interactions between six vTRs and the cofactor TLE5 via Y2H assays (Fig. 4f). TLE5 belongs to the TLE family of transcriptional co-repressors, which regulate proliferation, survival, cell fate, and immunity^{87,88}. Unlike other TLE proteins, TLE5 lacks repression domains and instead acts as a dominant-negative regulator, interfering with repressive complex formation^{87–89}. We found that US1 (HHV-1), UL82 (HCMV), E7 (HPV-16), and Tax (HTLV-1), with available BRB-seq data, altered pathways typically regulated by TLE proteins, including Notch and Wnt signaling, and RUNX target genes (Supplementary Data 2). We confirmed TLE5 interactions with US1 (HSV-2) and UL82 (HCMV) by co-immunoprecipitation and western blot (Fig. 4g). Structural modeling of these interactions with AlphaFold3 showed that both US1 and UL82 associate with the Q-domain of TLE5 (Fig. 4h), which is required for interaction with other TLE proteins^{88,89}. This suggests that many vTRs from different viral families sequester TLE5 away from other TLE family members, thereby

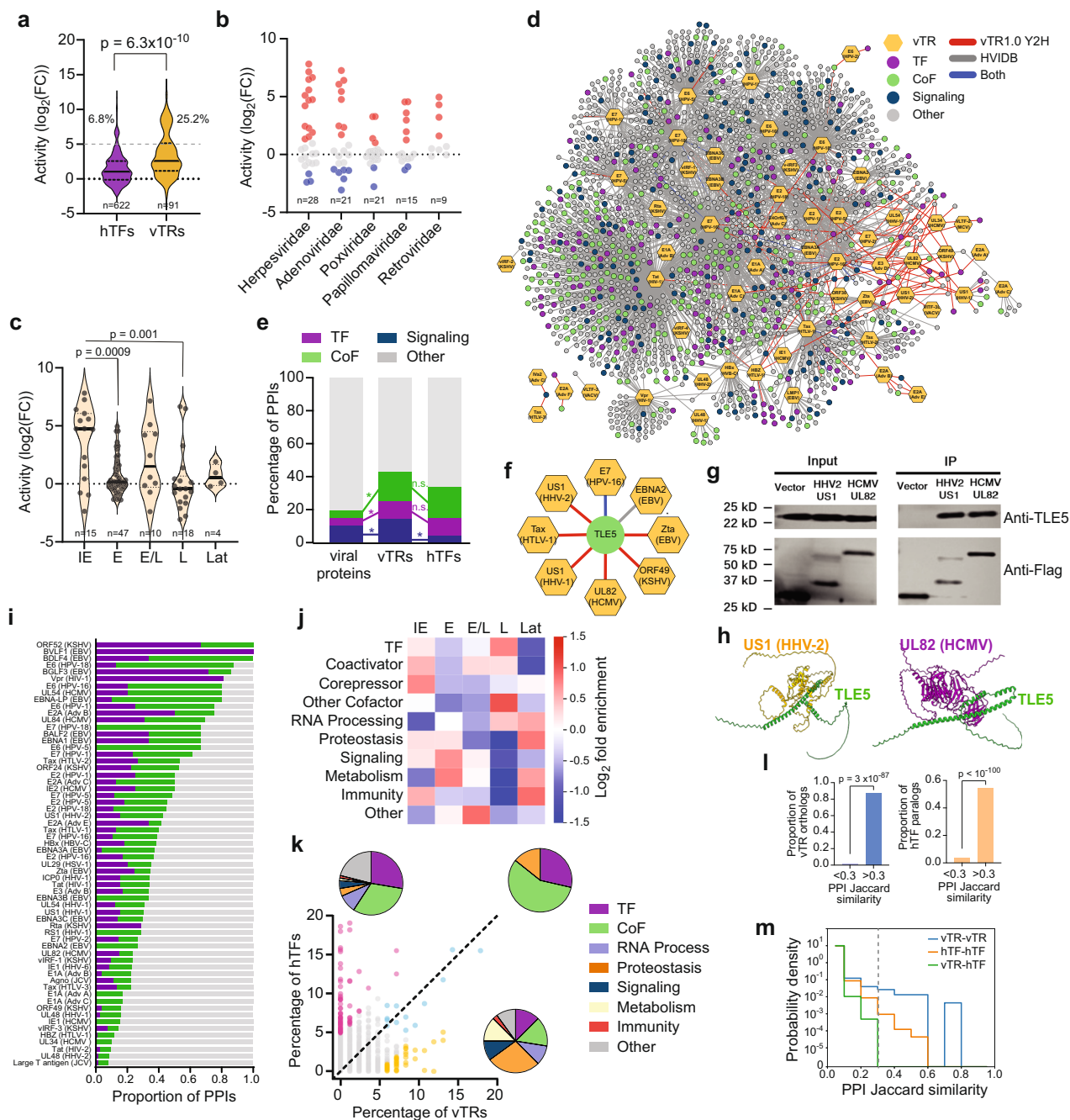


Fig. 4 | vTRs display intrinsic transcription activity and interact with host regulators. **a** Distribution of transcriptional activities elicited by hTFs and vTRs by MIH assays using a minimal promoter. Significance determined by two-sided Mann-Whitney test. **b** Transcriptional activities of vTRs across viral families by MIH assays with a moderate promoter. Red and blue points indicate activating ($\log_2(\text{FC}) \geq 1$) and repressing ($\log_2(\text{FC}) \leq -1$) vTRs, respectively. **c** Distribution of transcriptional activities for vTRs expressed at different timepoints of infection. IE, immediate early; E, early; E/L, early/late; L, late; Lat, latent. Distributions were compared using One-Way ANOVA. Asterisks correspond to multiple comparisons adjusted p -values ($*** P < 0.001$). **d** Protein-protein interaction (PPI) network curated from HVIDB and the Y2H screen. Hexagons indicate vTRs and circles indicate human proteins. Edges indicate the source of the PPI: hTF, human transcription factor; CoF, cofactor. **e** Proportion of PPIs between viral proteins, vTRs, or hTFs and hTFs, cofactors (CoF), signaling proteins, or other proteins. $*p < 0.01$ by two-sided proportion comparison test. **f** PPI Sub-network between 8 vTRs and TLE5. **g** Western blot showing co-immunoprecipitation of TLE5 with FLAG-tagged HHV2 US1 and HCMV

UL82. Representative experiment of three biological replicates. **h** AlphaFold3 model of interactions between TLE5 and HHV2 US1 and HCMV UL82. **i** Proportion of PPIs between vTRs and hTFs, CoFs or other proteins. Only vTRs with at least 3 human PPIs are shown. **j** PPI $\log_2(\text{fold enrichment})$ between vTRs expressed at different timepoints of infection for ten categories of human proteins. **k** Percentage of vTRs or hTFs with PPIs with individual human proteins. Interactions shown are the union of vTR1.0 Y2H, HuRI (<http://www.interactome-atlas.org>), HVIDB, and BioGRID. Magenta dots indicate human proteins preferentially interacting with hTFs, gold dots indicate proteins preferentially interacting with vTRs, and cyan dots indicate proteins that frequently interact with both vTRs and hTFs. Pie charts represent the fraction of different protein classes within each previously described subset of proteins. **l** Proportion of vTR homologs (left) or hTF paralogs (right) represented in comparisons with high Jaccard similarity (> 0.3) or low Jaccard similarity (< 0.3). Statistical significance determined by the proportion comparison test. **m** Probability density distributions of Jaccard PPI profile similarity scores within vTRs (blue), within hTFs (orange), or between vTRs and hTFs (green).

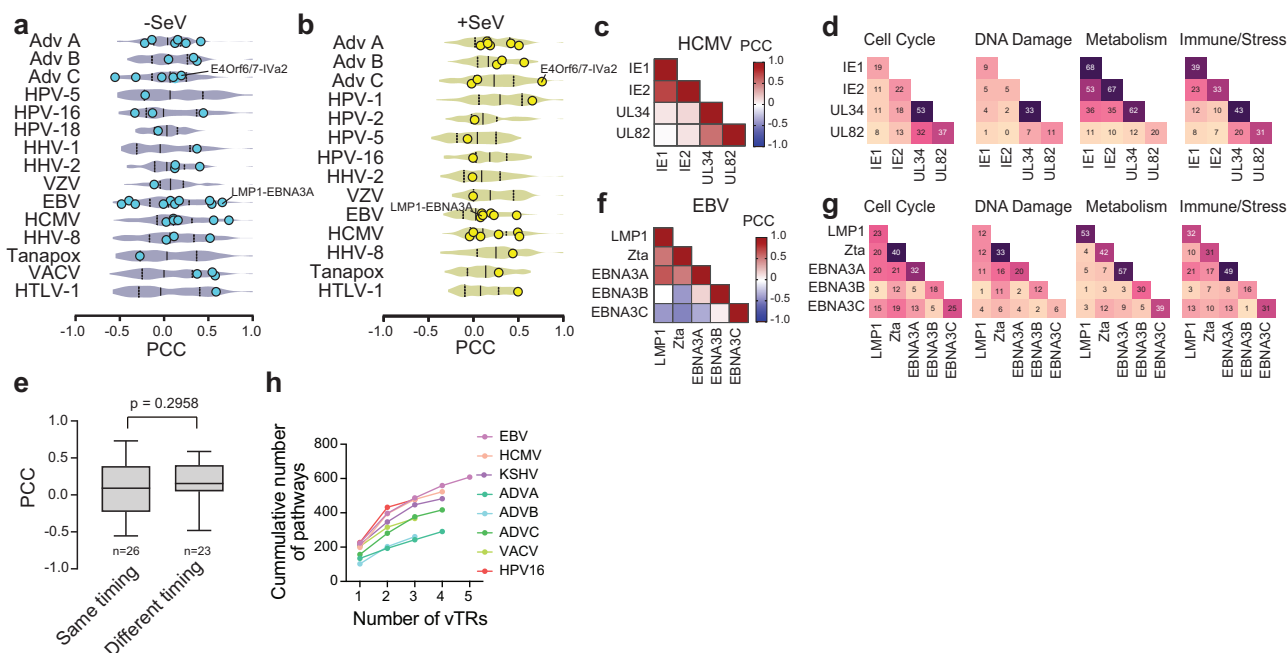


Fig. 5 | Functional diversity and redundancy between vTRs from the same virus.

a, b Pearson correlation coefficient (PCC) for pairs of vTRs encoded by the same virus in resting (**a**) and SeV-infected (**b**) cells. For each virus, points represent within-virus vTR comparisons, shown relative to the distribution of all other possible pairwise comparisons involving these vTRs. **c, f** PCC comparisons for five EBV (**c**) and four HCMV (**f**) vTRs in resting cells. **d, g** Heatmaps showing the number of overlapping pathways regulated by EBV vTRs (**d**) or HCMV vTRs (**g**) across four functional categories. Diagonal cells indicate the total pathways regulated by each

vTR, and off-diagonal cells indicate the number of pathways shared between vTR pairs. **e** PCC distribution for vTR pairs from the same virus with either the same or different expression timing. Box plots show the interquartile range, with the median indicated by the center line and whiskers representing the minimum and maximum values. Statistical significance evaluated by two-sided Mann-Whitney's U-test. **h** Cumulative number of pathways affected by the vTR repertoire of different viruses. Only viruses with at least 3 vTRs in our dataset are shown.

disrupting host transcriptional repression. Beyond TLE5, many other vTRs also modulate Notch and Wnt signaling pathways likely through alternative mechanisms, implicating these pathways as critical for viral replication or immune evasion (Fig. 2a and Supplementary Data 2).

Strikingly, over 60% of vTRs with at least 3 PPIs predominantly interact with non-TF/cofactor proteins (Fig. 4i). This suggests a spectrum of vTR functionality beyond direct transcriptional regulation, consistent with the known multifunctionality of viral proteins, including vTRs⁶. For instance, early vTRs interact with proteins involved in proteostasis, signaling, and metabolism, aligned with their role in manipulating the host cell environment for viral replication (Fig. 4j). Latent vTRs interact with immune, RNA processing, metabolic, and proteostatic proteins, potentially contributing to regulate cellular states, evade antiviral responses, and preserve latency.

Next, we evaluated which individual human proteins preferentially interact with vTRs or hTFs. As expected, many hTFs, cofactors, and RNA processing proteins preferentially interact with hTFs (Fig. 4k). Similarly, a set of 14 proteins, consisting primarily of broadly expressed cofactors (e.g., EP300, CREBBP, and KAT2B) and hTFs (TP53, SPI1, and CEBPB), was found to frequently interact with both hTFs and vTRs. We also found 40 proteins biased towards interactions with vTRs. This includes basal hTFs such as TBP and GTF2B, MEOX1 known to influence development, and NR4A1, RB1 and CCNA2, involved in cell cycle control and cell death^{90–96}. We also found that proteins involved in proteostasis, cell signaling, metabolism, and immunity preferentially interact with vTRs rather than hTFs (Fig. 4k), consistent with the multifunctionality of vTRs⁶.

Proteins with similar interaction partners often perform related molecular functions or participate in common regulatory processes. We therefore examined whether any vTR exhibits a PPI profile similar to that of a hTF, which could suggest that the vTR may have evolved to mimic specific host transcriptional regulators. As expected, vTR

homologs and hTF paralogs were significantly enriched among vTR-vTR and hTF-hTF pairs with similar interaction profiles, indicating that evolutionary relatedness is associated with conservation of interaction partners (Fig. 4l). However, vTR-hTF pairs engaged largely distinct sets of host proteins, as we did not identify an hTF-vTR pair with similar interaction profiles (Jaccard > 0.3, Fig. 4m). This underscores that vTRs have evolved distinct interaction networks rather than mirroring those of hTFs.

vTRs from the same virus exhibit both convergence and specialization in host regulatory programs

dsDNA viruses encode multiple vTRs that regulate different aspects of viral replication and host-cell remodeling⁶. To determine how vTRs from the same virus partition these functions, we compared the host pathways affected by vTRs from viruses represented by two or more vTRs in our BRB-seq dataset. Although vTRs from the same virus tend to affect more similar pathways than unrelated vTR pairs, substantial variability exists, with some pairs exhibiting highly similar and others highly divergent pathway profiles (Figs. 2b and 5a, b). For instance, HCMV vTRs expressed at the same time during infection exhibit the most correlated transcriptional effects (Fig. 5c); immediate early IE1 and IE2 affect similar cell cycle, DNA damage, metabolism, and immune/stress pathways, whereas early UL34 and UL82 regulate cell cycle and immune/stress genes and display no correlation with IE1- and IE2-affected pathways (Fig. 5d). Notably, however, similarity in transcriptional effects was not necessarily associated with the stage of infection at which the vTR is expressed (Fig. 5e). For example, latent EBV vTRs LMP1 and EBNA3A affect similar pathways to the lytic transactivator Zta, but very different pathways than EBNA3B and EBNA3C which are generated by alternative splicing from the same latently expressed transcript as EBNA3A (Fig. 5f, g)⁹⁷. These observations are consistent with a recent study showing that latent EBNA2 and

lytic switch Rta from EBV share many gene targets, suggesting that viruses maintain regulation of key host pathways throughout infection by deploying distinct vTRs that converge on common regulatory programs²⁰.

vTRs from the same virus can also engage in division of labor across cellular conditions. For example, LMP1 and EBNA3A (EBV), which are highly concordant in uninfected cells (PCC = 0.66), diverge under SeV infection (PCC = 0.09), whereas E4ORF6/7 and Iva2 (Adenovirus C) affect different pathways in uninfected conditions, but similar pathways under SeV infection (PCC = 0.19 in uninfected cells; PCC = 0.76 in infected cells) (Fig. 5a, b). Together, this illustrates both redundancy and specialization between vTRs from the same virus, which overall results in an expanded set of pathways affected (Fig. 5h).

vTR homologs display varying levels of functional divergence

Homologous viral genes are often assumed to perform similar functions; however, evolutionary adaptation to distinct hosts, tissues, and transmission strategies can drive functional divergence. To systematically assess the extent of conservation among vTR homologs, we compared 19 homolog groups across our experimental modalities (Fig. 6a). Consistent with an overall conservation of function, homologous vTRs exhibited significantly more similar transcriptional effects than random vTR pairs in our BRB-seq dataset (Fig. 2b). However, the degree of similarity varied widely among homolog groups, indicating that conservation is often partial rather than complete (Fig. 6b). For example, HTLV Tax homologs displayed highly similar transcriptional profiles under both uninfected and infected conditions, whereas ICP22 homologs from alphaherpesviruses (HHV-1 US1, HHV-2 US1, and VZV IE63) exhibited substantial divergence. HHV-1 and HHV-2 US1 down-regulate similar metabolic pathways, while these are upregulated by VZV IE63. Likewise, all three ICP22 homologs altered cell-cell signaling pathways, but the identity and direction of regulation differed substantially (Fig. 6c). Pathways specifically regulated by HHV-2 US1 are primarily involved in cell growth and metabolite trafficking, whereas VZV IE63 preferentially upregulates extracellular matrix and cell-cell interaction pathways. This specialization is consistent with the prominent role of cell-to-cell spread in VZV infection relative to HHV-1 and HHV-2^{98–100}.

Interestingly, transcriptional profile similarity was not correlated with amino acid sequence identity between homologs, suggesting rapid functional divergence across related viruses (Fig. 6d). For example, while E2 from HPV-5 and HPV-16 only share 30% amino acid identity, their PCC transcriptional similarity is 0.77; whereas E2 from HPV-16 and HPV-18 show a negative transcriptional correlation (PCC = −0.33) even though they share a 44% amino acid identity. Furthermore, even homologs that drive similar transcriptional profiles in one condition may differ in others. For instance, while HPV-16 E2 maintains a robust transcriptional effect upon SeV infection, upregulating extracellular signaling and downregulating cell cycle pathways, infection flips the HPV-5 E2 transcriptional effect (Fig. 6e, f). These differences in responses to SeV infection were also observed for other sets of vTR homologs (Fig. 6g). Together, these observations indicate that functional conservation among vTR homologs is often context dependent and cannot be inferred solely from amino acid sequence similarity.

To identify molecular mechanisms underlying this divergence, we compared the transcriptional activity and interaction profiles of vTR homologs. Several homolog groups, including adenovirus E1A and E4ORF6/7 and HPV E7, exhibited varying levels of transcriptional activation, potentially having a different impact on target gene expression (Fig. 6h). Furthermore, HPV E6 and adenovirus E2A contained both activating and repressing homologs (Fig. 6h). These differences were associated with distinct PPI profiles (Fig. 6i, j). For example, the activating adenovirus B E2A interacted with proteins commonly linked to transcriptional activation, including MAPK8 and

MAPK9, whereas the repressive adenovirus F E2A interacted with chromatin regulator PAXIP1 and the hTF ZNF446, which contains a strong repressive KRAB domain (Fig. 6i). Similarly, repressive HPV-1 E6 interacted with corepressors TRIM28 and SUZ12, whereas the strongly activating HPV-5 E6 associated with coactivators CREBBP, EP300, MAML1, and TADA3 (Fig. 6j). These observations suggest that gains and losses of specific coactivator, corepressor, hTF interactions contribute to functional divergence among vTR homologs.

We also observed divergence in DNA targeting and chromatin remodeling activities. Homologs of HPV E6 and E7 displayed different DNA-binding profiles in Y1H assays, which together with the observed variation in transcriptional activity, results in different transcriptional effects within these homolog groups (Figs. 2a, 3a and 6h). In addition, we found marked differences in chromatin accessibility between homologs of adenovirus E4ORF6/7. In particular, adenovirus B E4ORF6/7 induced substantially stronger accessibility changes than other homologs, including increased accessibility and expression at the TP73 locus (Fig. 6l, m and Supplementary Fig. 8b). Consistent with this broader chromatin remodeling activity, adenovirus B E4ORF6/7 altered the expression of approximately five-fold more genes than other E4ORF6/7 homologs (Supplementary Fig. 8c). Despite these quantitative differences, accessible regions induced by E4ORF6/7 homologs were enriched for similar p53/p63/p73 motifs, suggesting that divergence arises primarily from differences in interaction strength, cooperative DNA binding, or pioneering activity rather than entirely distinct targeting mechanisms. (Fig. 6n).

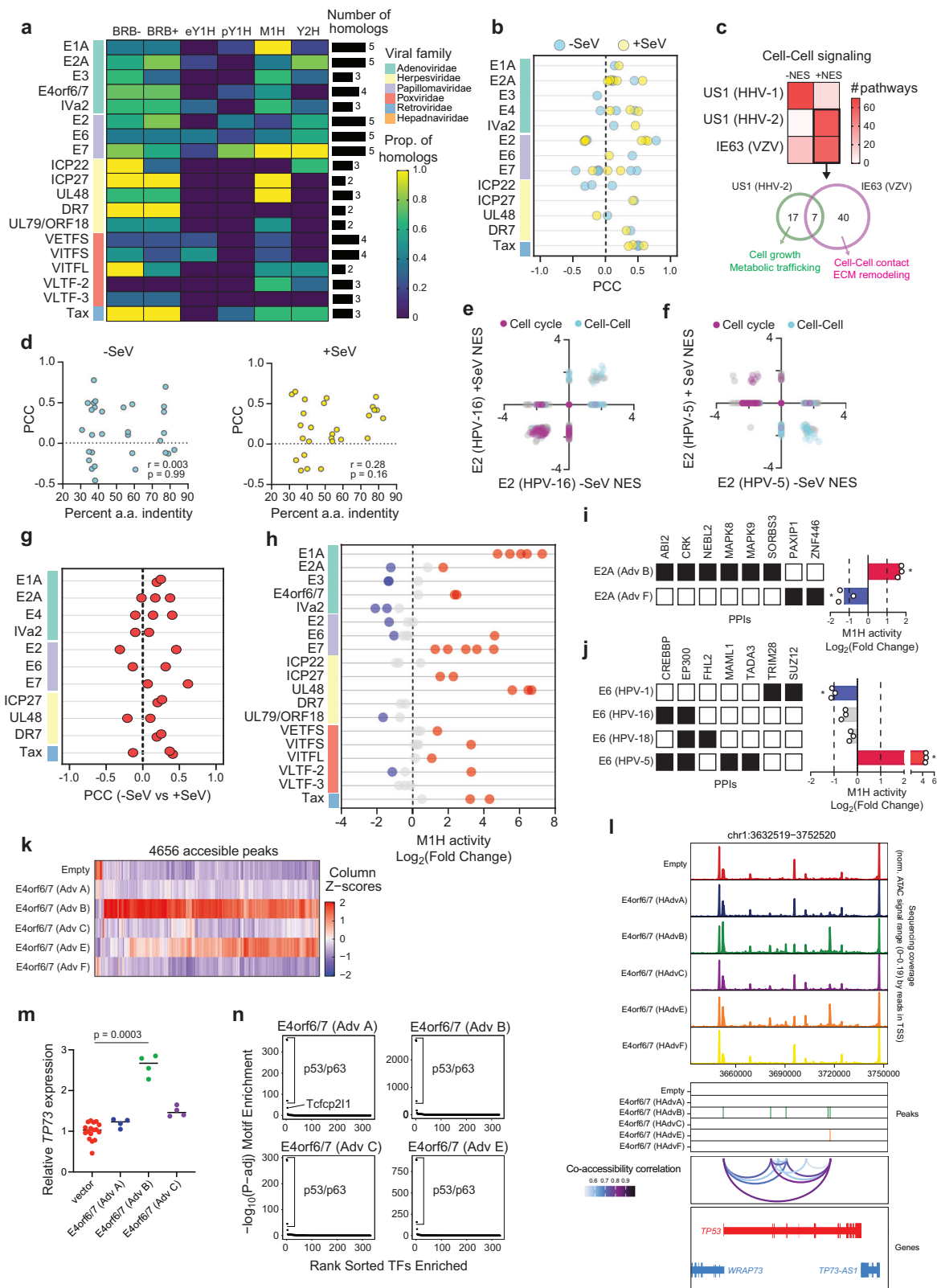
Collectively, these results show that homologous vTRs frequently share a subset of regulatory functions while diverging in others. This divergence is evident across transcriptional responses, transcriptional activity, protein interactions, DNA binding, and chromatin remodeling, resulting in distinct patterns of host target regulation that may contribute to differences in viral biology, pathogenesis, and host adaptation.

vTRs regulate genes associated with disease risk

Viral infections have long been associated with different chronic diseases, including autoimmune and neurological disorders, through molecular mimicry, bystander activation, immune dysregulation, and alteration of immune tolerance^{101,102}. For instance, studies have shown that EBV EBNA2 is recruited to the risk loci of multiple autoimmune diseases through interactions with hTFs such as RBPJ, leading to host gene dysregulation^{18,101–105}. To determine whether this is a unique feature of EBNA2 or whether other vTRs may also alter the expression of genes at disease risk loci, we evaluated the overlap between genes modulated by each vTR and risk loci from genome-wide association studies (GWAS). In total, we identified 40 vTRs with significant associations (FDR < 0.05) with at least one disease, corresponding to autoimmune, cardiovascular, respiratory, and neurological disorders (Fig. 7a).

Of the 109 significant disease-vTR associations identified, 40 are consistent with documented links between the disease and the virus encoding the vTR (Supplementary Data 7). For instance, genes dysregulated by HCMV IE1 overlapped with those associated with multiple neurological disorders, consistent with reports linking HCMV to these diseases and the role of IE1 in perturbing neuronal development and function^{106,107}. We also identified previously unreported vTR-disease associations, including DR7L (HHV-6A) with chronic obstructive pulmonary disease and UL48 (HHV-2) with COVID-19 severity (Fig. 7a). These results suggest that vTR-mediated transcriptional dysregulation frequently impacts genes associated with disease-risk loci, providing a potential molecular link between viral infection and disease-associated gene expression programs.

To determine whether the associations between vTRs and genes associated with disease-risk loci could be mediated by human regulatory factors, we integrated vTR-disease associations with ChIP-seq



binding profiles of interacting hTFs/cofactors at GWAS loci. This analysis identified 11 significant vTR-hTF/cofactor-disease trios (Supplementary Data 8), including several involving hypothyroidism or rheumatoid arthritis, and Parkinson's disease. Many of these trios contained disease-associated SNPs overlapping hTF/cofactor binding sites near genes dysregulated by the corresponding vTRs (Fig. 7b). For example, vIRF-3 (KSHV) dysregulated multiple genes in proximity of

Parkinson's disease risk loci that are also bound by MYC and EP300, two proteins reported to interact with vIRF-3. Notably, MYC binds the promoter of *GBF1*, which contains the Parkinson's disease-associated SNP rs2296887 (Fig. 7c). We further found that vIRF-3 increased *GBF1* expression, consistent with previous reports that vIRF-3 enhances MYC stability and transcriptional activity (Fig. 7d)^{108,109}. Together, these findings support a model in which interactions between vTRs

Fig. 6 | Functional variation across vTR homologs. **a** Survey of data available for sets of homologous vTRs. Cell color gradient indicates the proportion of homologs in each data set. Black bars indicate the total number of vTR homologs per set. BRB-, BRB-seq in resting cells; BRB+, BRB-seq in Sendai Virus-infected cells; eYIH, enhanced yeast one-hybrid; pYIH, paired yeast one-hybrid; MIH, mammalian one-hybrid; Y2H, yeast two-hybrid. **b** Pearson correlation coefficients (PCC) for pairs of vTR homologs. Blue and yellow dots indicate comparisons from uninfected or Sendai virus-infected cells (SeV). **c** Comparison of up- and down-regulated Cell-Cell signaling pathways between HHV ICP22 homologs. Color gradient indicates the number of pathways affected in uninfected conditions. A venn diagram represents the shared and distinct pathways upregulated by both US1 (HHV-2) and IE63 (VZV). NES, normalized enrichment score. **d** Scatter plots representing the relationship between the percent amino acid identity between homologs and PCC. Left plot, uninfected cells; Right plot, infected cells. Pearson r values are indicated in the plot. **e, f** Scatter plots comparing the NES of pathways dysregulated by HPV-16 E2 (**e**) and HPV-5 E2 (**f**) in uninfected versus Sendai Virus-infected conditions. Selected pathways are colored. **g** PCC for vTR homologs in uninfected versus SeV-infected cells.

h Transcriptional activities of vTRs across homolog groups via MIH assays using a moderate promoter. Red and blue points indicate activating ($\log_2(\text{FC}) \geq 1$) and repressing ($\log_2(\text{FC}) \leq -1$) vTRs, respectively. **i, j** PPIs (left) and MIH activity (right) for two Adenovirus E2A (**i**) and four HPV E6 (**j**) and homologs. Statistical significance determined by two-sided Student's t test from three biological replicates. $*p < 0.05$. **k** Heatmap of z-score normalized chromatin accessibility at 4,656 peaks that are differentially accessible ($\text{FDR} \leq 0.1$ and $\log_2(\text{FC}) \geq 1$) via PROD-ATAC comparing E4ORF6/7 homologs from five Adenoviruses. Peak columns are hierarchically clustered. **l** Genome browser tracks displaying sequencing coverage, significant chromatin peaks ($\text{FDR} \leq 0.1$ and $\log_2(\text{FC}) \geq 1$), and co-accessibility correlation values for peaks of adenovirus E4ORF6/7 orthologs in the *TP73* locus. **m** *TP73* expression in HEK293T cells expressing different adenovirus E4ORF6/7 orthologs relative to vector control. Statistical significance determined by two-sided Student's t test. Biological replicates: vector ($n = 17$), vTRs ($n = 4$). **n** Representative DNA motif enrichment within increased differentially accessible peaks for four Adv E4ORF6/7 homologs. Displayed values were calculated using hypergeometric tests and plotted relative to the ranked order of all queried motifs.

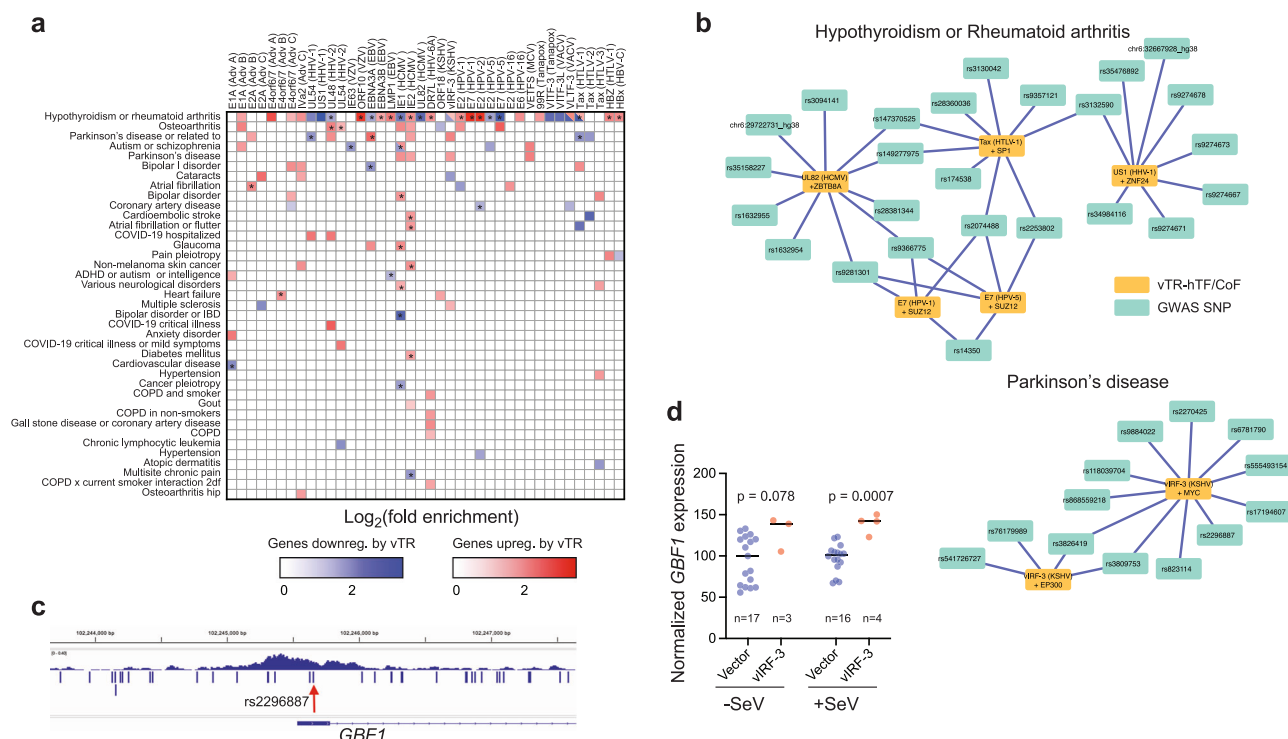


Fig. 7 | Association between vTRs and disease. **a** Fold enrichment of disease risk loci that overlap with genes up- or down-regulated by the expression of different vTRs (gene transcription start site ± 1 kb). Only associations with an $\text{FDR} < 0.05$ are shown. Asterisks indicate cases where the virus expressing a particular vTR has been associated with the corresponding disease in the literature. **b** Network showing significant vTR-hTF/CoF-GWAS locus trios for Hypothyroidism or rheumatoid arthritis, and Parkinson's disease. Edges represent overlap between vTR

DEGs and the disease risk SNP, and between ChIP-seq peaks for the hTF/CoF and the same disease risk SNP, for pairs of vTRs and hTFs/CoFs with reported PPIs. **c** Genome browser tracks displaying publicly available MYC ChIP-seq data at the *GBF1* locus. The Parkinson's disease-associated SNP rs2296887 is shown. **d** *GBF1* expression in HEK293T cells expressing vTRF-3 and subsequently infected with Sendai virus (SeV). Statistical significance determined by a two-sided Student's t test. The number of biological replicates per condition is indicated in the figure.

and human transcriptional regulators contribute to the dysregulation of genes located near disease-associated loci. While these analyses do not demonstrate direct regulation of the causal variants or regulatory elements underlying disease risk, they identify candidate vTR-hTF/cofactor-disease relationships for future mechanistic investigation.

Discussion

We systematically examined 95 vTRs across diverse human viruses, assessing their effects on transcription, DNA binding, transcriptional activity, and PPIs, providing a framework to understand how these proteins rewire host cellular programs and contribute to viral replication, persistence, and pathogenesis. For instance, we observe that

vTRs across viruses consistently perturb host pathways involved in proliferation, metabolism, DNA damage, and immune response, all of which contribute to viral replication and survival. This illustrates how vTRs act as powerful modulators of host cell states, enabling viruses to simultaneously influence multiple host pathways and drive broad shifts in cellular behavior.

vTRs use different mechanisms to target these pathways. Although in some cases this is achieved by direct binding to DNA, in most cases the transcriptional effects are likely driven by hijacking endogenous regulatory proteins. One mechanism involves indirect recruitment by tethering to hTFs, consistent with findings that vTR ChIP-seq peaks are enriched for binding sites of hTFs that physically

interact with the vTR^{6,18,20}. In other cases, vTRs cooperate with hTFs, which could re-direct hTFs to other genomic locations or change their activity. For example, HCMV IE2 prevents the binding of ZNF281 at multiple promoters and enhances ZNF281 binding to the *IL1B* promoter. A less explored strategy involves vTRs preventing the binding of hTFs that regulate cellular programs that interfere with viral replication (e.g., HPV E7 antagonizing E2F4). Together, these strategies enable vTRs to rewire existing transcriptional programs that control specific biological processes. This may reduce the need for direct recruitment of vTRs to large numbers of genomic loci to achieve broad, coordinated regulatory effects. Interestingly, some vTRs function through multiple mechanisms (e.g., directly binding to DNA, cooperating with hTFs, and antagonizing other hTFs), which could be leveraged by viruses with compact genomes to rewire diverse pathways using few vTRs.

Many vTRs are multifunctional proteins that can affect host pathways through a variety of mechanisms, such as targeted degradation or signaling inhibition. For example, in addition to exhibiting potentially direct and cooperative DNA binding, antagonism of hTFs, and intrinsic transcriptional activity, HPV E6 promotes p53 degradation, whereas HPV E7 promotes the degradation of RB1^{10,111}. Thus, some of the transcriptional signatures we observe are likely a combination of direct transcriptional and indirect effects. Further, viruses also encode proteins that disrupt host pathways independently of transcriptional control. These include factors that promote protein degradation (e.g., KSHV K3 and K5), mimic host molecules (e.g., HCMV US28, HIV-1 Gp120), or scaffold host signaling complexes (e.g., HHV-1/2 UL46)^{112–114}. Therefore, vTRs may buffer, enhance, or act orthogonally to these signaling modulators within multi-layered regulatory systems.

We found that vTRs from the same virus can have both overlapping and distinct roles in modulating host pathways. Although viruses have compact genomes with relatively few protein-coding genes, they may preserve functional redundancy for several reasons. Redundant vTRs can provide continuity in host regulation across different stages of infection (e.g., early and late vTRs both suppress immune pathways) or serve as a failsafe to counteract antiviral defenses that disable individual proteins. vTRs can also exhibit a division of labor, either by targeting different pathways that affect the various processes needed for viral replication or persistence, targeting parallel pathways that converge on the same cellular process, or by having opposing effects on the same pathway depending on context. These patterns suggest that redundancy and specialization among vTRs constitute complementary strategies that enable viruses to exert robust and fine-tuned control over their environment.

vTRs from unrelated viral species with no structural homology can produce strikingly similar effects on host pathways, underscoring widespread selective pressure to target conserved ‘pressure points’ in host biology. Conversely, homologous vTRs from closely related viruses often differ in function and mechanism. While broad regulatory agreement is shaped by the fundamental requirements for viral replication and survival, factors such as tropism can drive diversification of otherwise similar vTRs. These functional differences among homologs may also contribute to differences in pathogenicity, as seen in E2/E6/E7 across low- and high-risk HPVs or in the Tax proteins of HTLV-1, -2, and -3^{115–118}.

Our analysis uncovered widespread associations between disease-risk loci and genes modulated by diverse vTRs, suggesting that interactions between viral activity and host genetic variation may be far more pervasive than previously appreciated. Importantly, these associations are based on overlap between vTR-regulated genes and genes assigned to GWAS loci and therefore do not, on their own, establish that vTRs directly regulate the causal variants or regulatory elements underlying disease risk. Although determining causality will require longitudinal cohorts, integrated viral-host genomic and transcriptomic datasets, and functional studies in primary human systems,

our findings support a model in which viral infection acts as a broad environmental trigger that perturbs transcriptional programs linked to host susceptibility loci, potentially influencing disease penetrance, severity, and heterogeneity. Given the high prevalence of many of these viruses, these findings underscore the importance of systematically incorporating viral serostatus into GWAS designs to more accurately assess the combined effects of inherited and viral factors on complex disease risk.

Altogether, our findings highlight vTRs as versatile regulators that leverage diverse molecular mechanisms yet consistently converge on core host pathways essential for replication, persistence, and immune evasion. Beyond defining these principles, this work establishes a broadly applicable resource for the virology community. The multi-modal datasets generated will enable researchers studying a specific vTR or virus to contextualize their findings within a comparative framework spanning diverse viral families. Such comparisons may indicate whether vTRs, vTR orthologs, or viral proteins not included in our study act through strategies represented in our datasets and suggest candidate host proteins, DNA-recruitment mechanisms, and downstream pathways for mechanistic characterization. Likewise, by linking vTRs to host protein interactions, chromatin occupancy, transcriptional effects, and disease-associated genes, the resource can help connect infection phenotypes to specific vTRs and generate experimentally testable hypotheses regarding their molecular functions. Furthermore, by mapping how vTRs perturb host gene regulatory networks, our study identifies host pathways recurrently targeted across viruses, providing a foundation for future studies of virus–host interactions and the development of antiviral strategies for both acute and chronic infections.

Methods

EXPERIMENTAL APPROACHES

Generation of vTR1.0 entry clone collection. 95 vTR ORF sequences (Supplementary Data 1) were synthesized as gene fragments (Twist Bioscience) and subsequently cloned into pDONR221 by Gateway BP reaction (Thermo Fisher Scientific) followed by transformation into *E. coli* DH5 α . Transformed *E. coli* were then plated onto LB agar containing 50 μ g/mL Kanamycin overnight at 37 °C. Colonies for each transformation were confirmed by Sanger sequencing and expanded to generate our vTR entry clone collection.

BRB-seq experiments. To determine the transcriptomic effects of vTRs in human cells, we performed bulk RNA barcoding and sequencing (BRB-seq) in HEK293T cells (ATCC CRL-3216) expressing individual vTRs. Expression vTR clones were generated via Gateway LR cloning into the pEZY3 constitutive expression vector (Addgene plasmid #18672). Following the transformation, plasmids were extracted, and insertion was confirmed via PCR and gel electrophoresis. The full list of vTRs tested by BRB-seq is provided in Supplementary Data 2.

For BRB-seq experiments, HEK293T cells were cultured in DMEM media supplemented with 10% FBS and seeded in two 96-well plates at a density of 18,000 cells/well. The following day, cells were transfected in quadruplicate using Lipofectamine 3000 (Thermo Fisher Scientific) with 80 ng of the pEZY3-vTR plasmids, as well as empty pEZY3 vector controls. After 24 hours, cells were either infected with 4.0×10^6 CEID₅₀/mL Sendai virus (VR-907, ATCC) or left uninfected. RNAs were harvested 2 days post-transfection, a time point that we showed has minimal interferon pathway activation due to transfection.

To determine the effect of vTRs on cytosolic DNA-sensing pathways, BRB-seq was conducted in HEK293 cells (ATCC CRL-1573) treated with poly(dA:dT) (Invivogen). HEK293 were used for these experiments as immune response mediated through cGAS/STING is far more robust than in HEK293T cells^{119–121}. A concentration of 1 μ g/ml poly dA:dT was selected for stimulation, which induces a robust immune response in HEK293 after 24 hours.

Sequencing libraries were prepared using the MERCURIUS DRUG-seq kit (Alithea Genomics) according to the manufacturer's instructions and as described previously²⁴. Briefly, HEK293T cells were permeabilized using a cell lysis buffer on ice 15 min with periodic agitation, followed by centrifugation to clear cellular debris. Individual RNA-containing lysates were then reverse-transcribed using barcoded oligo-dT primers such that the resulting cDNA from each individual sample was uniquely barcode-labeled. cDNA from all samples in a single plate were then pooled and purified using the column-based Zymo Clean & Concentration Kit according to the manufacturer's instructions (Zymo, D4014). The pooled cDNA was then cleared of non-incorporated primers via exonuclease digestion, followed by second strand synthesis and quantification of ds-cDNA via Qubit HS-DNA kit (Thermo Fisher). Tagmentation was then performed on 20 ng cDNA, followed by library indexing and amplification using standard Illumina NGS adapter sequences. Indexed libraries were sequenced at a depth of 5 million raw reads per sample on an Illumina Novaseq platform.

Enhanced yeast one-hybrid (eY1H) assay. To detect PDIs between vTRs and cytokine gene promoters, eY1H assays were performed as previously described^{26,122,123}. vTRs prey clones were generated via Gateway LR into pDEST-AD2 μ (Walhout lab) for all 95 vTRs in vTR1.0, such that each vTR was fused with the yeast Gal4 Activation Domain. vTR-prey strains were generated via transformation into Y α 1867 yeast as previously described^{26,124}. Yeast were inoculated in 1 L liquid YAPD media at a concentration of OD₆₀₀ = 0.15 and incubated at 30 °C, shaking at 200 rpm until they reached OD₆₀₀ = 0.4–0.6, washed with sterile water, and washed again with TE + 0.1 M lithium acetate (TE/LiAc). Yeast were resuspended in TE/LiAc with salmon sperm DNA (ThermoFisher #15632011) at a dilution of 1:10 before adding 250 ng of the vTR clone. Six volumes of TE/LiAc + 40% polyethylene glycol were then added, and samples were mixed gently ten times. Yeast were then incubated at 30 °C without shaking for 30 min followed by 42 °C for 20 min, then resuspended in sterile water. Transformed yeast were plated in selective media lacking tryptophan to select for transformants. vTR prey yeast strains were confirmed by Sanger sequencing.

DNA-bait strains corresponding to 112 human cytokine promoters were previously generated using the Y1HaS2 yeast strain²⁷. Bait-strains carry two integrated copies of the promoter element upstream of *HIS3*, which allows yeast to grow in the absence of histidine and overcome growth inhibition by 3-amino-1, 2, 4-triazole (3AT), and *LacZ*, which causes yeast colonies to turn blue in the presence of 5-bromo-4-chloro-3-indolyl-beta-D-galacto-pyranoside (X-gal).

eY1H was performed as previously described^{26,122,123} using a high-density array ROTOR robotic platform (Singer Instruments). The vTR yeast array and cytokine promoter yeast strains were mated pairwise and spotted on permissive agar media for 1 day at 30 °C followed by selection on agar plates lacking Uracil and Tryptophan for 2 days to select diploid yeast. Selected diploid yeast were then transferred to selective media agar plates lacking uracil, tryptophan, and histidine, with 5 mM 3AT and 320 mg/L X-gal. These readout plates were imaged at 2-, 3-, 4-, and 7-days post-plating. Interactions were tested in quadruplicate and were manually curated by 2 researchers. Only vTR-promoter pairs that resulted in growth and blue color for at least two of the four colonies were considered positive (Supplementary Data 3). At least 90% of interactions were detected with all four colonies.

Paired Yeast one-hybrid (pY1H) assay. To detect cooperative binding and antagonism between vTRs and hTFs, pY1H assays were performed as previously described²⁸. We selected 113 vTR–hTF pairs from three sources: published PPIs, PPIs identified in this study, and inferred interactions involving hTFs that interact with vTR orthologs of previously established vTR–hTF pairs (Supplementary Data 9). Prey strains expressing each vTR–hTF pairs were generated using a

Y α 1867 strain where the *LEU2* gene had been disrupted (Y α 1867 Δ leu2) to accommodate transformation with plasmids containing both *TRP1* and *LEU2* selection markers²⁸. This array vTR–hTF array also included clones with single vTRs or hTFs expressed. DNA-bait strains corresponding to 39 human cytokine promoters and 83 cancer-related gene promoters were previously generated using the Y1HaS2 yeast strain^{27,29}.

pY1H was performed using a high-density array ROTOR robotic platform (Singer Instruments) as previously described²⁸. Briefly, the vTR–hTF yeast array and human promoter yeast strains were mated pairwise (in quadruplicate) on permissive medium agar plates and incubated at 30 °C for 1 d. Mated yeast were then transferred to selective medium agar plates lacking uracil, leucine, and tryptophan to select for successfully mated yeast and incubated at 30 °C for 2 d. These selection plates were imaged and analyzed to identify array locations with failed yeast growth, which were then removed from further analysis. Diploid yeast were finally transferred to selective medium agar plates lacking uracil, leucine, tryptophan, and histidine, with 5 mM 3AT and 320 mg/liter X-gal. These Readout plates were imaged 2, 3, 4, and 7 d after final plating.

Yeast plate images were processed and visualized using DISHA (Detection of Interactions Software for High-throughput Analyses) software²⁸. vTR–hTF pair strains were sorted based on each index (cooperativity, antagonism index 1, and antagonism index 2) separately. Images were then manually analyzed to identify cooperative and antagonistic interactions. To call an interaction, the following criteria were used: 1) vTR–hTF pair, vTR, and hTF yeast strains all showed growth in the mating selection plates before transfer to readout plates; and 2) on readout plates, ≥ 3 out of 4 replicate colonies were uniform for vTR–hTF pair, vTR, and hTF yeast strains. For cooperative interactions, the vTR–hTF pair yeast showed a strong or moderate reporter activity relative to the empty-empty strain, and vTR and hTF yeast showed no or only weak reporter activity. For antagonistic interactions, vTR and/or hTF yeast showed a strong or moderate reporter activity relative to the empty-empty strain and the vTR–hTF pair yeast showed no or only weak reporter activity. Cooperative and antagonistic interactions are provided in Supplementary Data 4.

ChIP-seq and ChIP-qPCR experiments. Flp-In-293 cells were cultured in DMEM supplemented with 10% FBS, 0.1% Anti-anti, and 0.2% normocin. After 48 h, cells were transfected using Lipofectamine 2000 and 0.25 μ g of a plasmid containing the coding sequence of EBV EBNA2, EBNA3A, or EBNA3B or HCMV IE2, each fused to GFP or an unfused GFP control and 2.25 μ g of the pOG44 Flp-Recombinase expression vector (ThermoFisher Scientific #K650001). Three days later, the media was changed to media containing 50 μ g/mL hygromycin, which was gradually increased to 200 μ g/mL to produce a pool of cells stably expressing EBNA2, EBNA3A, EBNA3B, IE2 or the GFP control. Microscopy for GFP and western blots using an anti-GFP antibody (1:1000) were used to confirm vTR expression.

Chromatin crosslinking was performed on 20 million cells resuspended in 20 mL of media by adding 2 mL of 11% formaldehyde, 50 mM HEPES (pH 8.0), 0.1 M NaCl, 1 mM EDTA, and 0.5 mM EGTA for 10 min with rotation. The crosslinking reaction was stopped by adding 1.1 mL of 2.5 M glycine and rotating for 5 min at room temperature. Cells were washed twice with cold PBS, centrifuged at 3000 $\times g$ for 5 min, and then stored in an Ultra-low freezer until further processing. Cross-linked nuclei were prepared by washing the cell pellet with 1 mL of ChIP L1 buffer (10 mM HEPES pH 8.0, 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.25% Triton X-100, and 0.5% NP-40) plus protease inhibitors and rotating at 4 °C for 10 min and centrifuged at 10,000 $\times g$ for 5 min. Cell pellets were then washed with 1 mL of ChIP L2 buffer (10 mM Tris pH 8.0, 1 mM EDTA, 200 mM NaCl, and 0.5 mM EGTA) plus protease inhibitors by rotating at room temperature for 10 min and the nuclei were centrifuged at 10,000 $\times g$ for 5 min. Nuclei were then washed with

0.5 mL of sonication buffer (TE + 0.1% SDS) with protease inhibitors and again centrifuged. The pelleted nuclei were resuspended in 1 mL of sonication buffer containing protease inhibitors, sonicated using a Covaris S220 focused-ultrasonicator at 10% duty cycle, 175 peak power, 200 burst/cycle for 7 min, centrifuged at maximum speed for 10 min, and the supernatant containing soluble chromatin was collected. The salt and detergents were adjusted to final concentrations of 150 mM NaCl, 0.1% NaDOC, 1% Triton X-100, 5% glycerol and protease inhibitors. The chromatin was pre-cleared of Ig by adding 20 μ L of Dynabeads Protein A or G and rotating at 4 °C for 40 min and centrifuged for 10,000 $\times$ g for 10 min.

ChIP was performed using 200 μ L of the pre-cleared chromatin on a SX-8X IP-STAR compact automated system robot from Diagenode, along with the antibody against GFP (Abcam #ab290) of interest and Dynabeads. The ChIP material was eluted using 10 mM TrisHCl, and libraries were prepared by ChIPmentation¹²⁵.

The GFP and HCMV IE2 libraries were sequenced at the CCHMC Genomics Sequencing Facility using either an Illumina NovaSeq 6000 or NovaSeq X Plus (single-end, 100 bp read length). ChIP-seq experiments for HCMV IE2 were performed in triplicate and compared to triplicates from the unfused GFP Flp-In-293 control cells.

ChIP-seq libraries in triplicate from the Flp-In-293 cells stably expressing either GFP, EBNA2-GFP, EBNA3A-GFP, or EBNA3B-GFP were quantified by qubit and then diluted to 0.01 ng/ μ L. We used 5 μ L of diluted ChIP-seq libraries for each as the template for qPCR using primers against EBNA2-GFP, EBNA3A-GFP, or EBNA3B-GFP bound regions or control regions near the bound peaks (Supplementary Data 10) and amplified using iTaq Universal SYBR Green Supermix and quantified on a QuantStudio 6 Flex. The qPCR cycle threshold (Ct) for each region was normalized to the average GFP control Ct for that region, and enrichment by EBNA2-GFP, EBNA3A-GFP, or EBNA3B-GFP determined by dividing the normalized Ct of each sample by the average normalized GFP control Ct for the same region and the significance of enrichment determined by Student's *t* test comparing the normalized values for each vTR to GFP.

Mammalian one-hybrid (M1H) assays. To measure vTR transcriptional activity, M1H assays were performed in HEK293T cells^{29,77}. In this assay, vTRs fused to the Gal4 DNA-binding domain are recruited to the upstream activating sequence (UAS) upstream of the firefly luciferase gene. Changes in expression of firefly luciferase are reflective of the endogenous activating or repressing activity of the vTR.

vTRs were cloned into the DB-pEZY3 vector via Gateway LR such that the vTRs were N-terminally fused to Gal4 DNA-binding domain¹²³. M1H assays were performed using pGL4.23 luciferase reporter constructs. For low-background M1H assays, a 4xUAS + minimal promoter construct was used, as described^{29,77}. We also performed M1H assays using a moderate background activity construct containing 4xUAS + moderate promoter containing the Syn2B10 enhancer sequence previously described¹²⁶ to detect transcriptional repression with higher sensitivity, as well as activation.

HEK293T cells were cultured in DMEM media supplemented with 10% FBS and plated in 96-well white opaque plates (Corning) at a density of 10,000 cells/well and incubated at 37 °C, 5% CO₂ for one day. Cells were then transfected with 80 ng of vTR (DB-pEZY3) plasmid, 20 ng 4xUAS construct, and 10 ng of renilla luciferase plasmid as a transfection normalization control. An empty DB-pEZY3 plasmid and 4xUAS construct plasmid were co-transfected to serve as a negative control. Cells were incubated for 2 days after transfection at 37 °C with 5% CO₂, followed by measurement of both firefly and renilla luciferase activity using the Dual-Glo Luciferase Assay System (Promega). Non-transfected cells were used to subtract background for firefly/renilla luciferase activities. Firefly luciferase activity was then normalized with renilla luciferase activity in each well. vTR activities using both reporter systems are provided in Supplementary Data 5.

PROD-ATAC Pooled library generation. The pooled library of HEK293T cells used to generate PROD-ATAC (Protein-Coding single-cell Assay for Transposase Accessible Chromatin) chromatin accessibility profiles were produced for the described set of vTR sequences as previously described in detail⁷¹. Briefly, 3-wells of a 6-well plate were seeded with 400,000 HEK293T-B4 attP-containing landing pad cells and left to attach overnight in DMEM media supplemented with 10% FBS. At ~50% confluency, each well was transfected (using Lipofectamine 3000 in Opti-MEM) with 1500 ng donor attB library, encoding each of the relevant vTR sequences, and 100 ng of pCAG-NLS-HA-Bxb1 (Addgene, 51271). Liposome-containing media was aspirated 18 hours post-transfection and replaced with fresh complete media containing 0.8 μ g/ml puromycin. After 3 days of puromycin selection, the 3-wells were pooled into a single T-25 flask. Passages were subsequently performed every 2 days (when cells reached ~70% confluency). At 12 days post-selection start, several cell aliquots were frozen (5% DMSO v/v) and stored in the vapor phase of a liquid nitrogen dewar.

PROD-ATAC Screening. In preparation for the PROD-ATAC sequencing run, the HEK293T cell library encoding the described vTR sequences was seeded and allowed to attach overnight in DMEM media supplemented with 10% FBS. Cells were treated with 2 μ g/ml doxycycline to induce library expression for 96 h, with a single split performed at 48 h post-induction. After this point, the PROD-ATAC single cell sequencing library generation procedure was performed as previously described using 10x Genomics scATAC kits with v2 Chemistry, with solely two deviations: 15,000 nuclei (rather than 10,000) were targeted for capture per 10x chip lane, and 8 total lanes were used in the assay⁷¹.

Yeast two-hybrid (Y2H) assays. Y2H assays were performed as previously described^{30,77}. vTRs were cloned into pDESTDB-X vectors via Gateway LR such that each vTR was fused to the yeast Gal4 DNA-binding domain (DB). The competent yeast strain Y8930 (MAT α) was transformed with individual DB-vTR constructs and plated on yeast synthetic media lacking leucine to select for DB-vTR plasmid-containing yeast. Haploid DB-vTR yeast strains were tested for auto-activation of the *GALI::HIS3* reporter gene. Individual DB-vTR yeast strains were spotted on SC-Leu-His+1 mM 3AT media, and any strains showing growth were considered auto-activators (AAs) and removed from the collection of strains to be screened.

A primary Y2H screen was performed in which all non-AA vTR strains were tested against the hORFeome v9.1 collection of 17,472 ORF clones³⁰ tagged with Gal4 activation domain (AD). DB-vTR strains were mated individually against 99 pools of ~188 AD-ORF strains (preys). To perform the mating, fresh overnight cultures of DB-vTR strains were mixed with AD-ORF strain pools and grown overnight at 30 °C in liquid rich media (YEPD). After overnight growth, the mated yeast cells were transferred into liquid SC-Leu-Trp media to select for diploids and again grown overnight at 30 °C. Finally, the yeast cells were spotted onto SC-Leu-Trp-His + 1 mM 3AT solid media to select for activation of the *GALI::HIS3* reporter gene. In parallel, diploid yeast cells were transferred onto SC-Leu-His + 1 mM 3AT solid media supplemented with 1 mg/l cycloheximide to test for spontaneous DB-ORF auto-activators. After 72 h incubation at 30 °C, yeast that grew on SC-Leu-Trp-His + 1 mM 3AT but not SC-Leu-His + 1 mM 3AT + 1 mg/L cycloheximide media were picked into SC-Leu-Trp, grown overnight, then processed to determine the identity of interaction partners by Sanger sequencing. In total, 674 PPIs between vTRs and human proteins were identified, involving 41 vTRs.

Following this first-pass screening, vTRs were tested pairwise for interactions with the candidate human ORFs identified in the first-pass screening in a matrix format (the 41 vTRs against the union of all ORF interactors). Yeast strains corresponding to the identified human interaction partners were picked from archival glycerol stocks,

cultured in liquid medium and mated one-by-one against all 41 vTRs, processed as described above and then scored. Briefly, interactors were inoculated in 200 μ L of selective media and mated overnight at 30 °C in 150 μ L liquid rich media (YEPD). The following day, mated yeast cells were transferred to 150 μ L liquid SC-Leu-Trp media to select for diploids. After overnight incubation at 30 °C, 5 μ L diploid yeast cells were spotted onto SC-Leu-Trp-His+1 mM 3AT solid media to select for activation of the *GAL1::HIS3* reporter gene, as well as on SC-Leu-Trp to control for successful mating. AA tests were included by mating each Y8930:DB-ORF against a Y8800:AD-null (containing no ORF), which was included on each test plate for each individual Y8930:DB-ORF. Spots were scored for growth with scores of 0, 1, 2, 3, 4, and NA for cases where the spotting had failed. Growth scores of 0 and 1 were considered not growing. If a spot corresponding to either the pair or the corresponding AA-test did not grow on SC-Leu-Trp, the pair was scored NA. If the AA-test had a growth score of 4, the corresponding pair was scored NA. Interactions were scored positive if they had higher growth scores on SC-Leu-Trp-His+1 mM 3AT solid media compared to the auto-activator test and had a growth score of at least 2; otherwise they were scored negative. All positive scored colonies were picked, lysed and the identity of the two interacting proteins was confirmed by sequencing. 86 PPIs were validated, involving in 28 vTRs. Next, we curated PPIs between non-AA vTRs and human proteins previously reported in Uniprot. All human proteins that interacted with at least one vTR, with available clones in the ORFeome, were pairwise tested against the non-AA vTRs. This led to 46 interactions, which were combined with the previous 86 PPIs for a total of 132 interactions between 33 vTRs and 67 human proteins (Supplementary Data 6).

Immunoprecipitation and immunoblotting analyses. HEK293T cells stably expressing the corresponding vTRs were lysed with RIPA lysis buffer (20 mM Tris-HCl, 37 mM NaCl, 2 mM EDTA, 1% Triton-X, 10% glycerol, 0.1% SDS and 0.5% sodium deoxycholate) with protease and phosphatase inhibitors (Roche). Protein samples were quantified (Pierce BCA Protein Assay Kit, Thermo Fisher Scientific), and the same amount of cell lysates from vTR overexpressing stable cells and GFP vector control cells were incubated with an anti-Flag M2 Affinity Gel (Sigma) overnight at 4 °C. The beads were washed 3x with TBST buffer and then boiled in the sample buffer. The supernatants were subjected to SDS polyacrylamide gel electrophoresis and transferred to polyvinylidene difluoride membranes. Membranes were incubated with primary antibodies, including anti-AES (1:1000; Proteintech) and anti-Flag (1:2000; Sigma) and subsequent secondary antibody: horseradish peroxidase-conjugated Affinipure Goat Anti-Rabbit IgG(H+L) (1:3000; Proteintech, catalog no. SA00001-2). Uncropped blots are provided in the Source Data file.

QUANTIFICATION AND STATISTICAL ANALYSES

BRB-seq sequence alignment. BRB-seq reads were aligned to a custom human reference based on GRCh38 (primary assembly) with Gencode v46 annotations¹²⁷, augmented by a concatenated FASTA of all VTR sequences and a merged GTF that combined genome and VTR features. Genome indices were built with STAR v2.7.9a¹²⁸ using splice junctions (`--sjdbGTFfile Gencode+VTR`, `--sjdbOverhang 79`). Alignment and counting were performed with STARsolo in CB/UMI simple mode to recover per-barcode gene counts (`--soloType CB_UMI_Simple`; `--soloFeatures Gene`; `--soloStrand Forward`; `--clipAdapterType CellRanger4`; `--outSAMtype BAM SortedByCoordinate`; `--outSAMunmapped Within`; `--outReadsUnmapped Fastx`). Barcodes and UMIs were parsed with a 14-nt cell barcode and 14-nt UMI (`--soloCBstart 1 --soloCBlen 14`; `--soloUMIstart 15 --soloUMIlen 14`), matched to a provided whitelist allowing one mismatch (`--soloCBmatchWtype IMM`), and deduplicated with directional one-mismatch UMI collapsing (`--soloUMIddedup IMM_Directional`). STAR output

included comprehensive SAM attributes (NH HI AS nM CR CY UR UY CB UB GX GN sS sQ sM).

Pre-processing, differential expression and pathways analysis. Raw STARsolo matrices were imported into Seurat¹²⁹ for each plate-condition, and well-level metadata were joined from barcode annotation tables. For each well, we quantified the proportion of reads mapping to the expected vTR and compared it with the most abundant vTR signal within the full vTR panel. Wells with at least 85% of reads assigned to a single vTR were flagged as highly specific and used to assess concordance with the expected assignment. For downstream analyses, we applied the following filters: (i) vTR wells were required to have at least 85% of their vTR-panel reads assigned to the expected construct and a minimum of 500,000 total mapped reads; (ii) control and vector wells were required to show no dominant vTR signal (top VTR < 0.01% of total reads) and at least 500,000 total mapped reads. This corresponded to 84 vTRs. After filtering, counts were merged across wells and joined to Gencode v46 annotations¹²⁷. We retained protein-coding, T-cell receptor, and immunoglobulin genes, while excluding mitochondrial and ribosomal genes as well as lowly expressed genes. Normalization and differential expression were performed using DESeq2¹³⁰, with the vector set as the reference. A pooled matrix design that included batch as a covariate was chosen, as it showed high correlation between replicates, while effectively reducing noise and confounding effects. Results were extracted using Benjamini-Hochberg correction (FDR < 0.1) without independent filtering. Gene Set Enrichment Analysis (GSEA)¹³¹ used msigdb to retrieve MSigDB C2:CP-REACTOME sets¹³² and was run with 10,000 permutations, Benjamini-Hochberg correction, $\text{eps} = 0$, $\text{minGSSize} = 15$, $\text{maxGSSize} = 500$, and p -value cutoff reporting at 1. Pathways with a p -value > 0.05 were considered not significant. Because libraries were aligned to a combined human + SeV reference, SeV burden was quantified per sample as counts per million (CPM): the summed STARsolo counts over all SeV transcripts were divided by the total mapped counts across all (host plus viral) genes and scaled by 10^6 .

Reactome pathway categorization. Individual reactome pathways from BRB-seq experiments were designated under any number of the following categories: cell-cell signaling, signal transduction, immune, metabolism, stress, DNA damage repair, cell cycle, gene regulation/ RNA biogenesis. Pathways were categorized based on reactome pathway descriptions (<https://reactome.org/>) and genes present in the pathway.

Quantification of pathway similarity between vTRs. To quantify sample-sample similarity at the pathway level, we constructed a matrix of normalized enrichment scores (NES) from GSEA Reactome results, setting values to zero when the corresponding pathway had a non-significant enrichment ($p > 0.05$). Pathways represented in fewer than three samples were excluded, and missing values were imputed as zero. Pairwise sample correlations were then computed using Pearson's correlation coefficient across the filtered NES matrix.

To identify clusters of vTRs with similar effects on pathway perturbation, the resulting PCC scores were used as input for Connection Specificity Index (CSI)^{133,134}. For each sample pair (A, B), the CSI was defined as the proportion of third samples (X) for which the maximum of the distances (A-X, B-X) was smaller than the distance between A and B. This procedure was parallelized across all sample pairs, producing a symmetric CSI matrix, which was normalized by the number of possible third samples ($n - 2$) to yield values between 0 (no concordance) and 1 (high concordance). The resulting CSI matrix was then used for downstream hierarchical clustering.

Alphafold3 of vTR structures. Predicted three-dimensional structures of proteins were generated using AlphaFold3¹³⁵. Structural predictions

were run through the publicly accessible AlphaFold3 interface using default parameters. Briefly, amino-acid sequences of each protein of interest were retrieved from UniProt and supplied as input to the AlphaFold3 prediction pipeline. Five structural models were produced per sequence. Predicted Local Distance Difference Test (pLDDT) scores and Predicted Aligned Error (PAE) heatmaps were used to assess overall model quality and to identify well-resolved regions suitable for interpretation. The highest-confidence monomer model was selected for downstream analysis and visualization, which was performed using UCSF ChimeraX¹³⁶.

Protein–protein complex (heterodimer) structures were predicted using AlphaFold3's multimer mode. Human transcription factors, or TLE5, were paired with vTRs and supplied jointly as separate chains in a single multimer prediction run. Five multimer models were generated for each input pair and ranked using pLDDT, inter-chain PAE, and AlphaFold3's interface confidence metrics (iPTM and chain-pair interface scores). The top-ranked multimer model was selected for downstream interpretation, and all structural visualization and interaction inspection were performed using UCSF ChimeraX¹³⁶.

Transcriptomics comparison between stimulated and unstimulated conditions. Differential expression was performed by comparing (i) each SeV-infected vTR against the SeV-infected vector control, and (ii) SeV-infected vector controls against uninfected vector controls. In addition, we compared Poly(dA:dT)-stimulated conditions to water control conditions. Count matrices were constructed from STARsolo outputs and joined with matched metadata to ensure consistent gene and sample annotation. DESeq2 (v1.46)¹³⁰ was used with poscounts normalization to account for sparsity in UMI data, and significance was assessed with Benjamini–Hochberg correction.

Pathway enrichment was carried out with fgsea (v1.32)(<https://doi.org/10.1101/060012>) using curated Reactome sets for both comparisons. Genes were ranked by $\log_2$ (fold change), restricted to those present in Reactome sets, and tested with 10,000 permutations. Non-significant pathways (FDR $\geq$ 0.05) were set to zero before aggregation into a summary NES matrix. To highlight construct-specific effects beyond stimulation alone, discordant pathways were defined as those enriched in opposite directions between stimulated vector controls and specific stimulated vTRs. These pathways were summarized across vTRs, and clustering of normalized enrichment scores was used to visualize the relationship between constructs and pathway-level responses.

Genes induced by poly(dA:dT) were identified by DESeq2 comparison of poly(dA:dT)- versus water-treated samples within each cell group (Wald test, “poscounts” size-factor estimation; significant at adjusted $p < 0.05$ and $|\log_2(\text{fold change})| \geq 1$). The effect of each vTR on innate immune gene expression was assessed by comparing vTR-versus vector-expressing samples with a two-sided Student's t test (Supplementary Fig. 4a, b), and per-gene $\log_2$ (fold changes) of vTR versus vector under each stimulus were computed to display vTRs modulating genes induced by both poly(dA:dT) and Sendai virus (Supplementary Fig. 4c).

HCMV IE2 ChIP-seq analysis. Alignment and peak calling were performed with a version of the ENCODE ChIP-seq pipeline^{137,138} modified to work within a Nextflow environment. A GFP-expressing HEK293 cell line was used as a background control for peak-calling. Alignments were performed using the hg38 genome, and “IDR conservative” peaks (10,217 peaks; narrowPeak format) generated from three IE2 ChIP-seq replicates relative to three monoGFP control replicates were used for downstream analysis. For visualization, per-replicate coverage tracks were generated as CPM-normalized bigWig files, and representative binding was displayed as genome-browser tracks for the three IE2 and three monoGFP replicates together with the IDR peak calls and gene models at the *CDH2* locus (Supplementary Fig. 5c).

To characterize where IE2 binds relative to genes, each peak was reduced to its summit (peak start plus the narrowPeak summit offset) and assigned to the nearest transcription start site (TSS). Strand-aware TSS coordinates were extracted for all genes from the Ensembl GRCh38 (release 110) annotation, and for every IE2 peak, the closest TSS on the same chromosome was identified and the absolute summit-to-TSS distance recorded. The cumulative distribution of these distances was plotted across all 10,217 peaks, and the fraction of peaks falling within 1, 3, and 10 kb of a TSS, together with the median distance, was computed to quantify the promoter-proximal enrichment of IE2 binding (Supplementary Fig. 5d).

To ask whether IE2 binding is concentrated in specific biological pathways, genes were classified by IE2 occupancy: a gene was considered IE2-bound if an IE2 peak summit fell within ± 1 kb of its strand-aware TSS (evaluated across all 10,217 peaks and all GRCh38 release 110 TSS). Reactome gene sets were obtained from MSigDB (collection C2, subcollection CP:REACTOME; *Homo sapiens*). For each such pathway, over-representation of IE2-bound genes was tested with a two-sided Fisher's exact test on a 2×2 table of IE2-bound versus IE2-unbound genes and pathway-member versus non-member genes, using all annotated genes as the universe; p -values were corrected across pathways by the Benjamini–Hochberg method. Results were displayed as a dot plot in which each pathway's odds ratio (capped at 20) is plotted against the pathway, point size reflects the number of IE2-bound genes in the pathway, and color encodes $-\log_{10}(\text{FDR})$ (Supplementary Fig. 5e).

Overlap between vTR ChIP-seq peaks and differentially expressed genes. Enrichment of ChIP-seq peaks near differentially expressed genes was performed with the RELI algorithm¹⁸. Gene regions were defined as the 1 kb up- and down-stream of the transcription start site of each DEG in our BRB-seq dataset, and overlapping regions were merged. ChIP-seq datasets used include the HCMV IE2 generated in this study, EBV vTRs from Dexheimer et al.²⁰, HIV-1 Tat (GSM1603206), and HTLV-1 HBZ (GSM2481679). The null model used for significance calculations contained all ATAC-seq peaks when any EBV vTR was expressed²⁰.

RELI analysis for GWAS loci. To identify which differentially expressed genes (DEGs) resulting from vTR expression were associated with disease risk loci, we utilized the Regulatory Element Locus Intersection (RELI) approach as described by Harley et al.¹⁸ RELI statistically tests for enrichment of GWAS risk loci in user-supplied genomic regions or gene lists, through permutation and comparison to matched controls. We input the list of up- or down-regulated genes resulting from the expression of each vTR into RELI and assessed the enrichment of risk-associated loci from published GWAS datasets proximal to these DEGs (transcription start site ± 1 kb). Only associations with FDR < 0.05 and at least 5 overlapping loci were considered for further analysis. To determine potential evidence in the literature for these associations, we annotated reported cases where the virus encoding for the vTR was associated with the corresponding disease (Supplementary Data 7).

To assess whether hTFs or cofactors (CoFs) may mediate the influence of vTRs on disease-associated loci, we applied RELI to publicly available hTF/CoF ChIP-seq peak sets performed in HEK293 (or related) cells. For each hTF/CoF, RELI was used to quantify the enrichment of GWAS risk loci within ChIP-seq peaks. As in the vTR–DEG analysis, risk loci were expanded to include variants in linkage disequilibrium with published GWAS variants, and enrichment was evaluated by comparing the observed overlap to a null distribution generated through permutation using matched control regions. Significant associations were defined as hTF/CoF–disease pairs with $p < 0.01$. Next, we integrated these results with the vTR–disease associations identified in the DEG-based RELI analysis. For a given disease, we identified hTFs or CoFs that showed significant enrichment at the

same GWAS loci for which a vTR was also significantly associated. To determine whether the vTR could potentially modulate these loci through recruitment by the hTF/CoF, we required evidence of a PPI between the vTR and the corresponding hTF/CoF. We annotated all resulting trios where (i) the vTR DEGs were significantly associated with the disease, (ii) the hTF/CoF ChIP-seq peaks overlapped GWAS loci for that disease, and (iii) a PPI existed between the vTR and the hTF/CoF (Supplementary Data 8).

Comparison between hTF and vTR M1H activity. The transcriptional activities of vTRs were compared to those of hTFs tested by Lambourne et al. using the low-background M1H system^{75,77}. The distribution of $\log_2$ (fold change) for all vTRs tested using the minimal promoter construct described previously was compared to those of reference hTF isoforms previously tested using the same reporter construct. Distributions were compared in Prism (GraphPad 10.60) using the two-sided Mann Whitney U-test.

M1H activity of vTRs with different expression timing. Annotation of the expression timing of vTRs was performed via literature curation (Supplementary Data 11). Taking into account the varied expression kinetics of different viruses, we broadly defined expression timing as follows: Immediate Early (IE) – expressed before replication of the viral genome. Early (E) – expressed following immediate-early and before replication of the viral genome. Late (L) – expressed following replication of the viral genome. Early-Late (E/L) expressed both prior to and following viral genome replication. Latent (Lat) – expressed primarily or exclusively during viral latency. The transcriptional activities of vTRs expressed at different timepoints of infection were compared using Prism (Graphpad 10.60) via Brown-Forsythe and Welch multiple ANOVA.

Prediction of activation domains in vTRs. Putative activation domains (ADs) were predicted in the 95 vTRs from vTR1.0 using ADpred v1.2.8, a convolutional neural network trained to identify transcriptional ADs from amino acid sequence and predicted secondary structure¹³⁹. Because most vTRs are not represented in public structural databases, secondary-structure annotations were generated from predicted protein structures. vTR structures were predicted using ESMFold through the ESM Atlas API. For proteins longer than 400 amino acids, structures were predicted in overlapping 400-amino-acid windows with a stride of 300 amino acids. Secondary-structure states were subsequently assigned using the P-SEA algorithm implemented in Biotite (annotate_sse) and converted to the three-state representation required by ADpred (helix, strand, or coil). ADpred was then applied using the amino acid sequence and corresponding secondary-structure annotations to generate per-residue activation domain probabilities, P(AD). Candidate activation domains were defined as contiguous regions of at least five residues with P(AD) $\geq$ 0.80. Predicted activation domains were summarized for each vTR and compared with experimentally determined transcriptional activation and repression activities.

PROD-ATAC Screen. PROD-ATAC data analysis (of both the single-cell ATAC libraries and genotyping dial-out libraries) was performed on the University of Wisconsin-Madison's Center for High Throughput Computing cluster as previously described, accessed May 2025⁷¹. Briefly, single-cell sequencing datafiles were converted to FASTQ format using cellranger-atac mkfastq (10x Genomics, version 2.1.0) and aligned to the hg38 reference genome via cellranger-atac count. Fragment files generated via cellranger-atac count (along with associated cell-assignment file outputs) were used to build Arrow files in ArchR. ArchR (version 1.0.3) was used for calling vTR marker peaks and performing downstream analysis using the same filter settings and quality cutoffs as described in depth previously⁷¹. Briefly, all nuclei used for

PROD-ATAC data analysis were first filtered for those with TSS enrichment scores ≥ 7 and $\geq 40,000$ fragments. We then generated pseudobulk samples representing each vTR genotype and created in silico replicates to perform statistical tests to identify differentially accessible peaks. The addGroupCoverages command (minCells = 20, maxCells = 500, minReplicates = 2, maxReplicates = 5) merged cells of a known genotype. The resultant coverage file was queried to identify peaks via addReproduciblePeakSet. Quality of the data was evaluated based on fragments per cell, reads in peaks, and TSS enrichment scores (Supplementary Fig. 8a). MACS3 was used with default parameters for peak calling. Patterns of motif enrichment within these differentially accessible peaks were identified by calling peakAnnoEnrichment with default parameters.

Scripts relevant for downstream PROD-ATAC data processing and figure generation used in both this and other PROD-ATAC associated publications are deposited in GitHub (<https://github.com/mfrenkel16/OncofusionPRODATA/>).

Comparing hTF/CoF interactions between viral proteins and vTRs. Protein interaction partner types for vTRs and non-vTR viral proteins were compared by taking the union of our Y2H results and interactions compiled by the Human-Virus Interaction DataBase (HVIDB) (Supplementary Data 12)²². Protein interaction partners of vTRs were categorized into eight categories: hTF, cofactor, RNA processing, proteostasis, signaling, metabolism, immunity, or other (Supplementary Data 12). The TF list was obtained from Lambert et al.⁸⁰, whereas cofactors include proteins that were not classified as hTF but are annotated in the list of human cofactors from Animal TF DB¹⁴⁰. RNA processing, proteostasis, signaling, metabolism, and immunity were non-hTF or cofactor proteins that were defined based on gene ontology (GO) terms generated by UniProt and confirmed via PubMed literature search. Example GO terms for each category are as follows: RNA processing - regulation of RNA splicing (GO:0043484), mRNA processing (GO:0006397), mRNA stabilization (GO:0048255), regulation of transcription by RNA polymerase II (GO:0006357). Proteostasis - protein ubiquitination (GO:0016567), protein folding (GO:0006457), protein stabilization (GO:0050821), intracellular protein transport (GO:0006886). Signaling - signaling (GO:0023052) and related lower terms as described in Lambourne et al.⁷⁷. Metabolism - fatty acid biosynthetic process (GO:0006633), glycolytic process (GO:0006096), cellular respiration (GO:0045333), ATP biosynthetic process (GO:0006754), TOR signaling (GO:0031929). Immune - innate immune response (GO:0045087), adaptive immune response (GO:0002250), inflammatory response (GO:000695), defense response to virus (GO:0051607). Proportions of interactions with hTF, cofactor, or other (all non-hTF or cofactor proteins) were then compared between vTRs and non-vTR viral proteins using a permutation test of a χ^2 statistic; two-sided empirical *P* values were computed.

Enrichment analysis for different PPI partners across vTR timing. To assess how different categories of protein interaction partners vary across vTR timing groups, we first normalized raw interaction counts to per-condition proportions. Proteins were annotated into functional categories (e.g., hTFs, cofactors, RNA processing, proteostasis, signaling, metabolism, immunity, and other). For each group, we aggregated proportions across proteins and compared them to the background average across all conditions. Fold-enrichment values were calculated, $\log_2$ -transformed, and visualized as heatmaps with a fixed scale (± 1.5). This approach highlights partner categories enriched or depleted in specific vTR timing groups relative to the overall interaction distribution.

Comparison between vTR and hTF PPI partners. Similarity of PPI partners between vTRs and hTFs was assessed by comparing the union of hTF-human protein interactions supported by at least one piece of

evidence in BIOGRID¹⁴¹ or reported in HuRI³⁰ with the union of vTR–human protein interactions detected by Y2H (as described) or reported in HVIDB²². For each protein partner represented, we calculated the percentage of hTFs or vTRs with which it interacted. Proteins interacting with more than 5% of hTFs or vTRs were highlighted. Jaccard indices were calculated to quantify the overlap of PPI partners: within vTRs, within hTFs, and between vTRs and hTFs.

Determining structural similarity of homolog vTRs. Structural similarity of homolog vTRs was determined using NCBI blast to compare amino acid identities on a pairwise basis. Amino acid sequences were chosen based on the representative UniProt ID for that given vTR. Amino acid % identity values were rounded to the nearest tenth place.

Statistics & reproducibility. No statistical method was used to pre-determine sample size. Sample sizes were selected based on established practices for the respective experimental approaches and are indicated in the Methods, figure legends, or Source Data, as appropriate. Experiments involving biological replicates were performed using independently prepared samples. The number of biological replicates and the statistical tests used for individual analyses are indicated in the corresponding figure legends and Methods sections. Unless otherwise indicated, statistical tests were two-sided. Multiple-testing correction was applied where appropriate, as described for the individual analyses.

For sequencing-based experiments, samples or features that did not meet predefined quality-control criteria were removed as described in the corresponding Methods sections. In particular, BRB-seq samples were filtered based on sequencing depth and the specificity of reads assigned to the expected vTR construct, and genes were filtered according to the criteria described in the BRB-seq analysis section. No other data were excluded from the analyses. The experiments were not randomized. The Investigators were not blinded to allocation during experiments and outcome assessment.

Materials availability

All unique/stable reagents generated in this study will be made available from the lead contact with a completed material transfer agreement.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The sequencing data generated in this study have been deposited in the Gene Expression Omnibus database under accession codes GSE334309 and GSE336673. The protein interaction data generated in this study have been deposited in the IMEx [<http://www.imexconsortium.org>] consortium through IntAct¹⁴² and assigned the identifier IM-30629. The processed data from all experiments performed in this study are provided in the Supplementary Data, Source Data files, and Zenodo [<https://doi.org/10.5281/zenodo.22233405>]. The structure models are available in ModelArchive [modelarchive.org] with accession codes: ma-zmhd4, ma-wax1w, ma-tds2f, ma-m4hrz, ma-zmn6h, ma-j3052, ma-4nd9m, ma-dq91f, ma-htc12, ma-w7kdu, ma-lmedc, ma-gzwev, ma-1lxau, ma-ok3j4, ma-rpsnt, ma-g892f, ma-sedbj, ma-bh9s0, ma-302gl, ma-rhu4f, ma-qr7qf, ma-r8gl1a, ma-4tkc2, ma-k7y60. Source data are provided as a Source Data file. Source data are provided in this paper.

Code availability

Original code is available in Zenodo [<https://doi.org/10.5281/zenodo.22233405>]

References

- Nikoskelainen, J., Panelius, M. & Salmi, A. E.B. virus and multiple sclerosis. *Br. Med. J.* **4**, 111 <https://doi.org/10.1136/bmj.4.5832.111> (1972).
- Magrath, I. Epidemiology: clues to the pathogenesis of Burkitt lymphoma. *Br. J. Haematol.* **156**, 744–756 <https://doi.org/10.1111/j.1365-2141.2011.09013.x> (2012).
- Marra, F. et al. Effects of herpes zoster infection, antivirals and vaccination on risk of developing dementia: A systematic review and meta-analysis. *Hum. Vaccin Immunother.* **21**, 2546741 <https://doi.org/10.1080/21645515.2025.2546741> (2025).
- Wolf, J. et al. Human papillomavirus infection: Epidemiology, biology, host interactions, cancer development, prevention, and therapeutics. *Rev. Med. Virol.* **34**, e2537 <https://doi.org/10.1002/rmv.2537> (2024).
- Xiao, Q. et al. Viral oncogenesis in cancer: from mechanisms to therapeutics. *Signal Transduct. Target Ther.* **10**, 151 <https://doi.org/10.1038/s41392-025-02197-9> (2025).
- Liu, X. et al. Human Virus Transcriptional Regulators. *Cell* **182**, 24–37 <https://doi.org/10.1016/j.cell.2020.06.023> (2020).
- Garcia-Sastre, A. Ten Strategies of Interferon Evasion by Viruses. *Cell Host Microbe* **22**, 176–184 <https://doi.org/10.1016/j.chom.2017.07.012> (2017).
- Katze, M. G., He, Y. & Gale, M. Viruses and interferon: a fight for supremacy. *Nat. Rev. Immunol.* **2**, 675–687 <https://doi.org/10.1038/nri888> (2002).
- Benedict, C. A., Norris, P. S. & Ware, C. F. To kill or be killed: viral evasion of apoptosis. *Nat. Immunol.* **3**, 1013–1018 <https://doi.org/10.1038/ni102-1013> (2002).
- Mesri, E. A., Feitelson, M. A. & Munger, K. Human viral oncogenesis: a cancer hallmarks analysis. *Cell Host Microbe* **15**, 266–282 <https://doi.org/10.1016/j.chom.2014.02.011> (2014).
- Tillo, D. et al. The Epstein-Barr Virus B-ZIP Protein Zta Recognizes Specific DNA Sequences Containing 5-Methylcytosine and 5-Hydroxymethylcytosine. *Biochemistry* **56**, 6200–6210 <https://doi.org/10.1021/acs.biochem.7b00741> (2017).
- Ramasubramanyam, S. et al. Epstein-Barr virus transcription factor Zta acts through distal regulatory elements to directly control cellular gene expression. *Nucleic Acids Res.* **43**, 3563–3577 <https://doi.org/10.1093/nar/gkv212> (2015).
- Basbous, J. et al. The HBZ factor of human T-cell leukemia virus type I dimerizes with transcription factors JunB and c-Jun and modulates their transcriptional activity. *J. Biol. Chem.* **278**, 43620–43627 <https://doi.org/10.1074/jbc.M307275200> (2003).
- Ferrari, R. et al. Epigenetic reprogramming by adenovirus e1a. *Science* **321**, 1086–1088 <https://doi.org/10.1126/science.1155546> (2008).
- Guerrieri, F. et al. Genome-wide identification of direct HBx genomic targets. *BMC Genom.* **18**, 184 <https://doi.org/10.1186/s12864-017-3561-5> (2017).
- Patrick, D. R., Oliff, A. & Heimbrook, D. C. Identification of a novel retinoblastoma gene product binding site on human papilloma-virus type 16 E7 protein. *J. Biol. Chem.* **269**, 6842–6850 (1994).
- Miyazawa, M. et al. IL-10 promoter transactivation by the viral K-RTA protein involves the host-cell transcription factors, specificity proteins 1 and 3. *J. Biol. Chem.* **293**, 662–676 <https://doi.org/10.1074/jbc.M117.802900> (2018).
- Harley, J. B. et al. Transcription factors operate across disease loci, with EBNA2 implicated in autoimmunity. *Nat. Genet.* **50**, 699–707 <https://doi.org/10.1038/s41588-018-0102-3> (2018).
- Hong, T. et al. Epstein-Barr virus nuclear antigen 2 extensively rewires the human chromatin landscape at autoimmune risk loci. *Genome Res.* **31**, 2185–2198 <https://doi.org/10.1101/gr.264705> (2021).

20. Dexeimer, P. J. et al. Systematic investigation reveals extensive Epstein-Barr virus transcriptional regulation of the human genome. Preprint at <https://doi.org/10.1101/2025.07.18.665561> (2025).
21. Massimino, L. et al. Gut virome-colonising Orthohepadnavirus genus is associated with ulcerative colitis pathogenesis and induces intestinal inflammation in vivo. *Gut* **72**, 1838–1847 <https://doi.org/10.1136/gutjnl-2022-328375> (2023).
22. Yang, X. et al. HVIDB: a comprehensive database for human-virus protein-protein interactions. *Brief. Bioinform.* **22**, 832–844 <https://doi.org/10.1093/bib/bbaa425> (2021).
23. Ludwig, C. H. et al. High-throughput discovery and characterization of viral transcriptional effectors in human cells. *Cell Syst.* **14**, 482–500.e8 <https://doi.org/10.1016/j.cels.2023.05.008> (2023).
24. Alpern, D. et al. BRB-seq: ultra-affordable high-throughput transcriptomics enabled by bulk RNA barcoding and sequencing. *Genome Biol.* **20**, 71 <https://doi.org/10.1186/s13059-019-1671-x> (2019).
25. Fuxman Bass, J. I. et al. Human gene-centered transcription factor networks for enhancers and disease variants. *Cell* **161**, 661–673 <https://doi.org/10.1016/j.cell.2015.03.003> (2015).
26. Reece-Hoyes, J. S. et al. Enhanced yeast one-hybrid assays for high-throughput gene-centered regulatory network mapping. *Nat. Methods* **8**, 1059–1064 <https://doi.org/10.1038/nmeth.1748> (2011).
27. Santoso, C. S. et al. Comprehensive mapping of the human cytokine gene regulatory network. *Nucleic Acids Res.* **48**, 12055–12073 <https://doi.org/10.1093/nar/gkaa1055> (2020).
28. Berenson, A. et al. Paired yeast one-hybrid assays to detect DNA-binding cooperativity and antagonism across transcription factors. *Nat. Commun.* **14**, 6570 <https://doi.org/10.1038/s41467-023-42445-6> (2023).
29. Lu, Y. et al. A large-scale cancer-specific protein-DNA interaction network. *Life Sci. Alliance* **7**, e202402641 <https://doi.org/10.26508/lsa.202402641> (2024).
30. Luck, K. et al. A reference map of the human binary protein interactome. *Nature* **580**, 402–408 <https://doi.org/10.1038/s41586-020-2188-x> (2020).
31. Shin, Y. C. et al. Inhibition of the ATM/p53 signal transduction pathway by Kaposi's sarcoma-associated herpesvirus interferon regulatory factor 1. *J. Virol.* **80**, 2257–2266 <https://doi.org/10.1128/JVI.80.5.2257-2266.2006> (2006).
32. Lee, H.-R., Kim, M. H., Lee, J.-S., Liang, C. & Jung, J. U. Viral interferon regulatory factors. *J. Interferon Cytokine Res.* **29**, 621–627 <https://doi.org/10.1089/jir.2009.0067> (2009).
33. zur Hausen, H. Papillomaviruses and cancer: from basic studies to clinical application. *Nat. Rev. Cancer* **2**, 342–350 <https://doi.org/10.1038/nrc798> (2002).
34. Zhang, Y., Qiu, K., Ren, J., Zhao, Y. & Cheng, P. Roles of human papillomavirus in cancers: oncogenic mechanisms and clinical use. *Signal Transduct. Target Ther.* **10**, 44 <https://doi.org/10.1038/s41392-024-02083-w> (2025).
35. Azran, I., Schavinsky-Khrapunsky, Y. & Aboud, M. Role of Tax protein in human T-cell leukemia virus type-I leukemogenicity. *Retrovirology* **1**, 20 <https://doi.org/10.1186/1742-4690-1-20> (2004).
36. Müller, S. & Dejean, A. Viral immediate-early proteins abrogate the modification by SUMO-1 of PML and Sp100 proteins, correlating with nuclear body disruption. *J. Virol.* **73**, 5137–5143 <https://doi.org/10.1128/JVI.73.6.5137-5143.1999> (1999).
37. Harwardt, T. et al. Human Cytomegalovirus Immediate-Early 1 Protein Rewires Upstream STAT3 to Downstream STAT1 Signaling Switching an IL6-Type to an IFN γ -Like Response. *PLoS Pathog.* **12**, e1005748 <https://doi.org/10.1371/journal.ppat.1005748> (2016).
38. Kirchhoff, S. et al. Viral IFN-regulatory factors inhibit activation-induced cell death via two positive regulatory IFN-regulatory factor 1-dependent domains in the CD95 ligand promoter. *J. Immunol.* **168**, 1226–1234 <https://doi.org/10.4049/jimmunol.168.3.1226> (2002).
39. Lin, R. et al. HHV-8 encoded vIRF-1 represses the interferon antiviral response by blocking IRF-3 recruitment of the CBP/p300 coactivators. *Oncogene* **20**, 800–811 <https://doi.org/10.1038/sj.onc.1204163> (2001).
40. Ross, T. M., Narayan, M., Fang, Z. Y., Minella, A. C. & Green, P. L. Human T-cell leukemia virus type 2 tax mutants that selectively abrogate NF κ B or CREB/ATF activation fail to transform primary human T cells. *J. Virol.* **74**, 2655–2662 <https://doi.org/10.1128/jvi.74.6.2655-2662.2000> (2000).
41. Bertzbach, L. D. et al. The adenovirus DNA-binding protein DBP. *J. Virol.* **98**, e0188523 <https://doi.org/10.1128/jvi.01885-23> (2024).
42. Khetan, S., Carroll, B. S. & Bulky, M. L. Multiple overlapping binding sites determine transcription factor occupancy. *Nature* **646**, 1001–1011 <https://doi.org/10.1038/s41586-025-09472-3> (2025).
43. Ye, Z. et al. HCMV infection disrupts barrier functions and promotes epithelial-mesenchymal transition in a cholangiocyte organoid model. *Nat. Commun.* **17**, 2273 <https://doi.org/10.1038/s41467-026-68962-8> (2026).
44. Wiebusch, L. & Hagemeier, C. Human cytomegalovirus 86-kilodalton IE2 protein blocks cell cycle progression in G(1). *J. Virol.* **73**, 9274–9283 <https://doi.org/10.1128/JVI.73.11.9274-9283.1999> (1999).
45. Song, Y.-J. & Stinski, M. F. Inhibition of cell division by the human cytomegalovirus IE86 protein: role of the p53 pathway or cyclin-dependent kinase 1/cyclin B1. *J. Virol.* **79**, 2597–2603 <https://doi.org/10.1128/JVI.79.4.2597-2603.2005> (2005).
46. Wu, X., Zhou, X., Wang, S. & Mao, G. DNA damage response(DDR): a link between cellular senescence and human cytomegalovirus. *J. Virol.* **20**, 250 <https://doi.org/10.1186/s12985-023-02203-y> (2023).
47. Duyao, M. P., Kessler, D. J., Spicer, D. B. & Sonenshein, G. E. Transactivation of the c-myc gene by HTLV-1 tax is mediated by NF κ B. *Curr. Top. Microbiol. Immunol.* **182**, 421–424 https://doi.org/10.1007/978-3-642-77633-5_53 (1992).
48. Poirson, J. et al. High-risk mucosal human papillomavirus 16 (HPV16) E6 protein and cutaneous HPV5 and HPV8 E6 proteins employ distinct strategies to interfere with interferon regulatory factor 3-mediated beta interferon expression. *J. Virol.* **96**, e0187521 <https://doi.org/10.1128/jvi.01875-21> (2022).
49. Castro-Muñoz, L. J. et al. Alteration of the IFN-Pathway by Human Papillomavirus Proteins: Antiviral Immune Response Evasion Mechanism. *Biomedicines* **10**, 2965 <https://doi.org/10.3390/biomedicines10112965> (2022).
50. Um, S.-J. et al. Abrogation of IRF-1 response by high-risk HPV E7 protein in vivo. *Cancer Lett.* **179**, 205–212 [https://doi.org/10.1016/s0304-3835\(01\)00871-0](https://doi.org/10.1016/s0304-3835(01)00871-0) (2002).
51. Ronco, L. V., Karpova, A. Y., Vidal, M. & Howley, P. M. Human papillomavirus 16 E6 oncoprotein binds to interferon regulatory factor-3 and inhibits its transcriptional activity. *Genes Dev.* **12**, 2061–2072 <https://doi.org/10.1101/gad.12.13.2061> (1998).
52. Li, W.-F. et al. p52-ZER6: a determinant of tumor cell sensitivity to MDM2-p53 binding inhibitors. *Acta Pharm. Sin.* **44**, 647–660 <https://doi.org/10.1038/s41401-022-00973-9> (2023).
53. Huang, C. et al. Zinc-finger protein p52-ZER6 accelerates colorectal cancer cell proliferation and tumour progression through

- promoting p53 ubiquitination. *EBioMedicine* **48**, 248–263 <https://doi.org/10.1016/j.ebiom.2019.08.070> (2019).
54. Qiu, L. et al. p52-ZER6/DAZAP1 axis promotes ferroptosis resistance and colorectal cancer progression via regulating SLC7A11 mRNA stabilization. *Acta Pharm. Sin. B* **15**, 2039–2058 <https://doi.org/10.1016/j.apsb.2025.02.013> (2025).
 55. Tang, Y. et al. The p52-ZER6/G6PD axis alters aerobic glycolysis and promotes tumor progression by activating the pentose phosphate pathway. *Oncogenesis* **12**, 17 <https://doi.org/10.1038/s41389-023-00464-4> (2023).
 56. Li, W. et al. p52-ZER6/IGF1R axis maintains cancer stem cell population to promote cancer progression by enhancing pro-survival mitophagy. *Oncogene* **43**, 2115–2131 <https://doi.org/10.1038/s41388-024-03058-5> (2024).
 57. Chellappan, S. et al. Adenovirus E1A, simian virus 40 tumor antigen, and human papillomavirus E7 protein share the capacity to disrupt the interaction between transcription factor E2F and the retinoblastoma gene product. *Proc. Natl. Acad. Sci. USA* **89**, 4549–4553 <https://doi.org/10.1073/pnas.89.10.4549> (1992).
 58. Wu, E. W., Clemens, K. E., Heck, D. V. & Münger, K. The human papillomavirus E7 oncoprotein and the cellular transcription factor E2F bind to separate sites on the retinoblastoma tumor suppressor protein. *J. Virol.* **67**, 2402–2407 <https://doi.org/10.1128/JVI.67.4.2402-2407.1993> (1993).
 59. Caberg, J.-H. D. et al. Silencing of E7 oncogene restores functional E-cadherin expression in human papillomavirus 16-transformed keratinocytes. *Carcinogenesis* **29**, 1441–1447 <https://doi.org/10.1093/carcin/bgn145> (2008).
 60. Hahn, S. & Hermeking, H. ZNF281/ZBP-99: a new player in epithelial-mesenchymal transition, stemness, and cancer. *J. Mol. Med.* **92**, 571–581 <https://doi.org/10.1007/s00109-014-1160-3> (2014).
 61. Zhao, Q. et al. ZNF281 inhibits mitochondrial biogenesis to facilitate metastasis of hepatocellular carcinoma. *Cell Death Discov.* **9**, 396 <https://doi.org/10.1038/s41420-023-01691-9> (2023).
 62. Lu, C. et al. ZNF281 drives hepatocyte senescence in alcoholic liver disease by reducing HK2-stabilized PINK1/Parkin-mediated mitophagy. *Cell Prolif.* **56**, e13378 <https://doi.org/10.1111/cpr.13378> (2023).
 63. Pierdomenico, M. et al. Transcription Factor ZNF281: A Novel Player in Intestinal Inflammation and Fibrosis. *Front. Immunol.* **9**, 2907 <https://doi.org/10.3389/fimmu.2018.02907> (2018).
 64. Taylor, R. T. & Bresnahan, W. A. Human cytomegalovirus immediate-early 2 gene expression blocks virus-induced beta interferon production. *J. Virol.* **79**, 3873–3877 <https://doi.org/10.1128/JVI.79.6.3873-3877.2005> (2005).
 65. Zhang, X. et al. Human cytomegalovirus-IE2 suppresses antigen presentation of macrophage through the IL10/STAT3 signalling pathway in transgenic mouse. *PLoS ONE* **20**, e0322334 <https://doi.org/10.1371/journal.pone.0322334> (2025).
 66. Heilingloh, C. S. et al. The Major Immediate-Early Protein IE2 of Human Cytomegalovirus Is Sufficient to Induce Proteasomal Degradation of CD83 on Mature Dendritic Cells. *Front. Microbiol.* **8**, 119 <https://doi.org/10.3389/fmicb.2017.00119> (2017).
 67. Kim, J.-E., Kim, Y.-E., Stinski, M. F., Ahn, J.-H. & Song, Y.-J. Human Cytomegalovirus IE2 86 kDa Protein Induces STING Degradation and Inhibits cGAMP-Mediated IFN- β Induction. *Front. Microbiol.* **8**, 1854 <https://doi.org/10.3389/fmicb.2017.01854> (2017).
 68. Horwitz, G. A. et al. Adenovirus small e1a alters global patterns of histone modification. *Science* **321**, 1084–1085 <https://doi.org/10.1126/science.1155544> (2008).
 69. Schaeffner, M. et al. BZLF1 interacts with chromatin remodelers promoting escape from latent infections with EBV. *Life Sci.* *Alliance* **2**, e201800108 <https://doi.org/10.26508/lsa.201800108> (2019).
 70. Pumfery, A. et al. Chromatin remodeling and modification during HIV-1 Tat-activated transcription. *Curr. HIV Res.* **1**, 343–362 <https://doi.org/10.2174/1570162033485186> (2003).
 71. Frenkel, M., Corban, J. E., Hujoel, M. L. A., Morris, Z. & Raman, S. Large-scale discovery of chromatin dysregulation induced by oncofusions and other protein-coding variants. *Nat. Biotechnol.* **43**, 996–1010 <https://doi.org/10.1038/s41587-024-02347-4> (2025).
 72. Nevels, M., Rubenwolf, S., Spruss, T., Wolf, H. & Dobner, T. Two distinct activities contribute to the oncogenic potential of the adenovirus type 5 E4orf6 protein. *J. Virol.* **74**, 5168–5181 <https://doi.org/10.1128/jvi.74.11.5168-5181.2000> (2000).
 73. Yamano, S. et al. Induction of transformation and p53-dependent apoptosis by adenovirus type 5 E4orf6/7 cDNA. *J. Virol.* **73**, 10095–10103 <https://doi.org/10.1128/JVI.73.12.10095-10103.1999> (1999).
 74. Boivin, D., Morrison, M. R., Marcellus, R. C., Querido, E. & Branton, P. E. Analysis of synthesis, stability, phosphorylation, and interacting polypeptides of the 34-kilodalton product of open reading frame 6 of the early region 4 protein of human adenovirus type 5. *J. Virol.* **73**, 1245–1253 <https://doi.org/10.1128/JVI.73.2.1245-1253.1999> (1999).
 75. Soto, L. F. et al. Compendium of human transcription factor effector domains. *Mol. Cell* **82**, 514–526 <https://doi.org/10.1016/j.molcel.2021.11.007> (2022).
 76. Mukund, A. X. et al. High-throughput functional characterization of combinations of transcriptional activators and repressors. *Cell Syst.* **14**, 746–763 <https://doi.org/10.1016/j.cels.2023.07.001> (2023).
 77. Lambourne, L. et al. Widespread variation in molecular interactions and regulatory properties among transcription factor isoforms. *Mol. Cell* **85**, 1445–1466 <https://doi.org/10.1016/j.molcel.2025.03.004> (2025).
 78. Fan, D. et al. The Role of VP16 in the Life Cycle of Alpha-herpesviruses. *Front. Microbiol.* **11**, 1910 <https://doi.org/10.3389/fmicb.2020.01910> (2020).
 79. Chen, H. & Pugh, B. F. What do Transcription Factors Interact With? *J. Mol. Biol.* **433**, 166883 <https://doi.org/10.1016/j.jmb.2021.166883> (2021).
 80. Lambert, S. A. et al. The human transcription factors. *Cell* **172**, 650–665 <https://doi.org/10.1016/j.cell.2018.01.029> (2018).
 81. Yan, C. et al. Transcription preinitiation complex structure and dynamics provide insight into genetic diseases. *Nat. Struct. Mol. Biol.* **26**, 397–406 <https://doi.org/10.1038/s41594-019-0220-3> (2019).
 82. Inge, M. M. et al. Rapid profiling of transcription factor-cofactor interaction networks reveals principles of epigenetic regulation. *Nucleic Acids Res.* **52**, 10276–10296 <https://doi.org/10.1093/nar/gkaf706> (2024).
 83. Panne, D., Maniatis, T. & Harrison, S. C. An atomic model of the interferon-beta enhanceosome. *Cell* **129**, 1111–1123 <https://doi.org/10.1016/j.cell.2007.05.019> (2007).
 84. Guirimand, T., Delmotte, S. & Navratil, V. VirHostNet 2.0: surfing on the web of virus/host molecular interactions data. *Nucleic Acids Res.* **43**, D583–D587 <https://doi.org/10.1093/nar/gku1121> (2015).
 85. Calderone, A., Licata, L. & Cesareni, G. VirusMentha: a new resource for virus-host protein interactions. *Nucleic Acids Res.* **43**, D588–D592 <https://doi.org/10.1093/nar/gku830> (2015).
 86. Yu, H. et al. High-quality binary protein interaction map of the yeast interactome network. *Science* **322**, 104–110 <https://doi.org/10.1126/science.1158684> (2008).

87. Chen, G. & Courey, A. J. Groucho/TLE family proteins and transcriptional repression. *Gene* **249**, 1–16 [https://doi.org/10.1016/S0378-1119\(00\)00161-x](https://doi.org/10.1016/S0378-1119(00)00161-x) (2000).
88. Yu, G. et al. Roles of transducin-like enhancer of split (TLE) family proteins in tumorigenesis and immune regulation. *Front. Cell Dev. Biol.* **10**, 1010639 <https://doi.org/10.3389/fcell.2022.1010639> (2022).
89. Beagle, B. & Johnson, G. V. W. AES/GRG5: more than just a dominant-negative TLE/GRG family member. *Dev. Dyn.* **239**, 2795–2805 <https://doi.org/10.1002/dvdy.22439> (2010).
90. Candia, A. F. et al. Mox-1 and Mox-2 define a novel homeobox gene subfamily and are differentially expressed during early mesodermal patterning in mouse embryos. *Development* **116**, 1123–1136 <https://doi.org/10.1242/dev.116.4.1123> (1992).
91. Candia, A. F. & Wright, C. V. Differential localization of Mox-1 and Mox-2 proteins indicates distinct roles during development. *Int. J. Dev. Biol.* **40**, 1179–1184 (1996).
92. He, L. et al. A Regulation Loop between YAP and NR4A1 Balances Cell Proliferation and Apoptosis. *Cell Rep.* **33**, 108284 <https://doi.org/10.1016/j.celrep.2020.108284> (2020).
93. Fassett, M. S., Jiang, W., D'Alise, A. M., Mathis, D. & Benoist, C. Nuclear receptor Nr4a1 modulates both regulatory T-cell (Treg) differentiation and clonal deletion. *Proc. Natl. Acad. Sci. USA* **109**, 3891–3896 <https://doi.org/10.1073/pnas.1200090109> (2012).
94. Gao, Y. et al. Nr4a1 enhances Wnt4 transcription to promote mesenchymal stem cell osteogenesis and alleviates inflammation-inhibited bone regeneration. *Mol. Ther.* **32**, 1479–1496 <https://doi.org/10.1016/j.ymthe.2024.02.034> (2024).
95. Goodrich, D. W., Wang, N. P., Qian, Y. W., Lee, E. Y. & Lee, W. H. The retinoblastoma gene product regulates progression through the G1 phase of the cell cycle. *Cell* **67**, 293–302 [https://doi.org/10.1016/0092-8674\(91\)90181-w](https://doi.org/10.1016/0092-8674(91)90181-w) (1991).
96. Pagano, M., Pepperkok, R., Verde, F., Ansorge, W. & Draetta, G. Cyclin A is required at two points in the human cell cycle. *EMBO J.* **11**, 961–971 <https://doi.org/10.1002/j.1460-2075.1992.tb05135.x> (1992).
97. Rogers, R. P., Woisetschlaeger, M. & Speck, S. H. Alternative splicing dictates translational start in Epstein-Barr virus transcripts. *EMBO J.* **9**, 2273–2277 <https://doi.org/10.1002/j.1460-2075.1990.tb07398.x> (1990).
98. Cohrs, R. J. & Gilden, D. H. Human herpesvirus latency. *Brain Pathol.* **11**, 465–474 <https://doi.org/10.1111/j.1750-3639.2001.tb00415.x> (2001).
99. Reichelt, M., Brady, J. & Arvin, A. M. The replication cycle of varicella-zoster virus: analysis of the kinetics of viral protein expression, genome synthesis, and virion assembly at the single-cell level. *J. Virol.* **83**, 3904–3918 <https://doi.org/10.1128/JVI.02137-08> (2009).
100. Chen, J. J., Zhu, Z., Gershon, A. A. & Gershon, M. D. Mannose 6-phosphate receptor dependence of varicella zoster virus infection in vitro and in the epidermis during varicella and zoster. *Cell* **119**, 915–926 <https://doi.org/10.1016/j.cell.2004.11.007> (2004).
101. Münz, C., Lünemann, J. D., Getts, M. T. & Miller, S. D. Antiviral immune responses: triggers of or triggered by autoimmunity? *Nat. Rev. Immunol.* **9**, 246–258 <https://doi.org/10.1038/nri2527> (2009).
102. Fujinami, R. S., von Herrath, M. G., Christen, U. & Whitton, J. L. Molecular mimicry, bystander activation, or viral persistence: infections and autoimmune disease. *Clin. Microbiol. Rev.* **19**, 80–94 <https://doi.org/10.1128/CMR.19.1.80-94.2006> (2006).
103. Keane, J. T. et al. The interaction of Epstein-Barr virus encoded transcription factor EBNA2 with multiple sclerosis risk loci is dependent on the risk genotype. *EBioMedicine* **71**, 103572 <https://doi.org/10.1016/j.ebiom.2021.103572> (2021).
104. Viel, K. C. M. F. et al. Shared and distinct interactions of type 1 and type 2 Epstein-Barr Nuclear Antigen 2 with the human genome. *BMC Genomics* **25**, 273 <https://doi.org/10.1186/s12864-024-10183-8> (2024).
105. Ricigliano, V. A. G. et al. EBNA2 binds to genomic intervals associated with multiple sclerosis and overlaps with vitamin D receptor occupancy. *PLoS ONE* **10**, e0119605 <https://doi.org/10.1371/journal.pone.0119605> (2015).
106. Wu, C.-C. et al. Human cytomegalovirus immediate Early 1 protein causes loss of SOX2 from neural progenitor cells by trapping unphosphorylated STAT3 in the nucleus. *J. Virol.* **92**, e00340-18 <https://doi.org/10.1128/JVI.00340-18> (2018).
107. Yu, M. et al. Brain-specific expression of cytomegalovirus immediate early protein 1 disrupts neurodevelopment in mice by inducing neuroinflammation and altering astrocytic metabolism. *Virol. J.* **22**, 258 <https://doi.org/10.1186/s12985-025-02874-9> (2025).
108. Baresova, P., Pitha, P. M. & Lubyova, B. Kaposi sarcoma-associated herpesvirus vIRF-3 protein binds to F-box of Skp2 protein and acts as a regulator of c-Myc protein function and stability. *J. Biol. Chem.* **287**, 16199–16208 <https://doi.org/10.1074/jbc.M111.335216> (2012).
109. Lubyova, B., Kellum, M. J., Frisancho, J. A. & Pitha, P. M. Stimulation of c-Myc transcriptional activity by vIRF-3 of Kaposi sarcoma-associated herpesvirus. *J. Biol. Chem.* **282**, 31944–31953 <https://doi.org/10.1074/jbc.M706430200> (2007).
110. Scheffner, M., Werness, B. A., Huibregtse, J. M., Levine, A. J. & Howley, P. M. The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. *Cell* **63**, 1129–1136 [https://doi.org/10.1016/0092-8674\(90\)90409-8](https://doi.org/10.1016/0092-8674(90)90409-8) (1990).
111. Dyson, N., Howley, P. M., Münger, K. & Harlow, E. The human papilloma virus-16 E7 oncoprotein is able to bind to the retinoblastoma gene product. *Science* **243**, 934–937 <https://doi.org/10.1126/science.2537532> (1989).
112. Alcamí, A. & Koszinowski, U. H. Viral mechanisms of immune evasion. *Trends Microbiol.* **8**, 410–418 [https://doi.org/10.1016/S0966-842X\(00\)01830-8](https://doi.org/10.1016/S0966-842X(00)01830-8) (2000).
113. You, H. et al. Herpes simplex virus 1 tegument protein UL46 inhibits TANK-binding kinase 1-mediated signaling. *mBio* **10**, e00919-19 <https://doi.org/10.1128/mBio.00919-19> (2019).
114. Boomker, J. M., van Luyn, M. J. A., The, T. H., de Leij, L. F. M. H. & Harmsen, M. C. US28 actions in HCMV infection: lessons from a versatile hijacker. *Rev. Med. Virol.* **15**, 269–282 <https://doi.org/10.1002/rmv.468> (2005).
115. Doorbar, J., Egawa, N., Griffin, H., Kranjec, C. & Murakami, I. Human papillomavirus molecular biology and disease association. *Rev. Med. Virol.* **25**, 2–23 <https://doi.org/10.1002/rmv.1822> (2015). Suppl 1.
116. Doorbar, J. Molecular biology of human papillomavirus infection and cervical cancer. *Clin. Sci.* **110**, 525–541 <https://doi.org/10.1042/CS20050369> (2006).
117. Shirinian, M., Kfoury, Y., Dassouki, Z., El-Hajj, H. & Bazarbachi, A. Tax-1 and Tax-2 similarities and differences: focus on post-translational modifications and NF- κ B activation. *Front. Microbiol.* **4**, 231 <https://doi.org/10.3389/fmicb.2013.00231> (2013).
118. Chevalier, S. A. et al. The transcription profile of Tax-3 is more similar to Tax-1 than Tax-2: insights into HTLV-3 potential leukemogenic properties. *PLoS ONE* **7**, e41003 <https://doi.org/10.1371/journal.pone.0041003> (2012).
119. Cheng, W.-Y. et al. The cGAS-Sting signaling pathway is required for the innate immune response against Ectromelia virus. *Front Immunol.* **9**, 1297 <https://doi.org/10.3389/fimmu.2018.01297> (2018).

120. Sun, W. et al. ERIS, an endoplasmic reticulum IFN stimulator, activates innate immune signaling through dimerization. *Proc. Natl. Acad. Sci. USA* **106**, 8653–8658 <https://doi.org/10.1073/pnas.0900850106> (2009).
121. Sun, L., Wu, J., Du, F., Chen, X. & Chen, Z. J. Cyclic GMP-AMP synthase is a cytosolic DNA sensor that activates the type I interferon pathway. *Science* **339**, 786–791 <https://doi.org/10.1126/science.1232458> (2013).
122. Berenson, A. & Fuxman Bass, J. I. Enhanced yeast one-hybrid assays to study protein-DNA interactions. *Methods Mol. Biol.* **2599**, 11–20 https://doi.org/10.1007/978-1-0716-2847-8_2 (2023).
123. Shrestha, S., Liu, X., Santoso, C. S. & Fuxman Bass, J. I. Enhanced yeast one-hybrid screens to identify transcription factor binding to human DNA sequences. *J. Vis. Exp.* <https://doi.org/10.3791/59192> (2019).
124. Fuxman Bass, J. I., Reece-Hoyes, J. S. & Walhout, A. J. M. Performing yeast one-hybrid library screens. *Cold Spring Harb. Protoc.* **2016**, pdb.prot088955 <https://doi.org/10.1101/pdb.prot088955> (2016).
125. Schmidl, C., Rendeiro, A. F., Sheffield, N. C. & Bock, C. ChIPmentation: fast, robust, low-input ChIP-seq for histones and transcription factors. *Nat. Methods* **12**, 963–965 <https://doi.org/10.1038/nmeth.3542> (2015).
126. Schlabach, M. R., Hu, J. K., Li, M. & Elledge, S. J. Synthetic design of strong promoters. *Proc. Natl. Acad. Sci. USA* **107**, 2538–2543 <https://doi.org/10.1073/pnas.0914803107> (2010).
127. Mudge, J. M. et al. GENCODE 2025: reference gene annotation for human and mouse. *Nucleic Acids Res.* **53**, D966–D975 <https://doi.org/10.1093/nar/gkaf1078> (2025).
128. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 <https://doi.org/10.1093/bioinformatics/bts635> (2013).
129. Hao, Y. et al. Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat. Biotechnol.* **42**, 293–304 <https://doi.org/10.1038/s41587-023-01767-y> (2024).
130. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 <https://doi.org/10.1186/s13059-014-0550-8> (2014).
131. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **102**, 15545–15550 <https://doi.org/10.1073/pnas.0506580102> (2005).
132. Liberzon, A. et al. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* **1**, 417–425 <https://doi.org/10.1016/j.cels.2015.12.004> (2015).
133. Fuxman Bass, J. I. et al. Using networks to measure similarity between genes: association index selection. *Nat. Methods* **10**, 1169–1176 <https://doi.org/10.1038/nmeth.2728> (2013).
134. Green, R. A. et al. A high-resolution *C. elegans* essential gene network based on phenotypic profiling of a complex tissue. *Cell* **145**, 470–482 <https://doi.org/10.1016/j.cell.2011.03.037> (2011).
135. Abramson, J. et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500 <https://doi.org/10.1038/s41586-024-07487-w> (2024).
136. Pettersen, E. F. et al. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 <https://doi.org/10.1002/pro.3943> (2021).
137. Hitz, B. C. et al. The ENCODE uniform analysis pipelines. Preprint at <https://doi.org/10.1101/2023.04.04.535623> (2023).
138. Luo, Y. et al. New developments on the Encyclopedia of DNA Elements (ENCODE) data portal. *Nucleic Acids Res.* **48**, D882–D889 <https://doi.org/10.1093/nar/gkz1062> (2020).
139. Erijman, A. et al. A high-throughput screen for transcription activation domains reveals their sequence features and permits prediction by deep learning. *Mol. Cell* **78**, 890–902 <https://doi.org/10.1016/j.molcel.2020.04.020> (2020).
140. Shen, W.-K. et al. AnimalTFDB 4.0: a comprehensive animal transcription factor database updated with variation and expression annotations. *Nucleic Acids Res.* **51**, D39–D45 <https://doi.org/10.1093/nar/gkac907> (2023).
141. Oughtred, R. et al. The BioGRID database: A comprehensive biomedical resource of curated protein, genetic, and chemical interactions. *Protein Sci.* **30**, 187–200 <https://doi.org/10.1002/pro.3978> (2021).
142. Orchard, S. et al. The MintAct project-IntAct as a common curation platform for 11 molecular interaction databases. *Nucleic Acids Res.* **42**, D358–D363 <https://doi.org/10.1093/nar/gkt1115> (2014).

Author contributions

Conceptualization, J.T.R., X.L. and J.I.F.B.; validation, J.T.R., X.L., L.F.S-U. and B.E.; formal analysis, J.T.R., X.L., L.F.S-U., M.Y.E., J.E.C., P.D., B.E., G.M.-E., Z.L., P.T., A.L., S.S., T.H. and J.I.F.B.; investigation, J.T.R., X.L., A. L., C.S., J.E.C., R.L., B.E., Y.L., K.S-F., S.K., S.S., C.-C.L., L.M.-C., M.H., C.Y. and J.I.F.B.; resources, J.I.F.B., M.V., S.R., M.L.B., M.T.W. and L.K.; data curation, J.T.R., X.L., J.I.F.B. and M.T.W.; writing – original draft, J.T.R. and J.I.F.B.; writing – review and editing, J.T.R., J.I.F.B., N.S., S.Y., S.R., M.T.W., A. L., L.F.S-U., M.L.B., S.K., M.V. and P.D.; visualization, J.T.R., M.Y.E., B.E., A. L., J.E.C. and J.I.F.B.; supervision, J.I.F.B., S.R., M.V., M.C., M.L.B., M.T.W. and L.K.; project administration, J.I.F.B.; and funding acquisition, J.I.F.B., M.V., S.R., M.L.B., M.T.W. and L.K.

Funding

The work was supported by NIH grants R35GM128625 (to J.I.F.B.), R35GM133658 (to S.S.Y.), R35GM137836 (to N.S.), U01CA232161 (to J.I.F.B., M.V. and M.L.B.), R01HG010730, U24HG013078, R01NS099068, and R01AI024717 (to M.T.W.), and Alfred P. Sloan Scholar Research Fellowship grant FG-2018-10723 (to N.S. M.Y.E and P.T. were supported by NIH training grant T32GM150533. M.Y.E was supported by NIH training grant T32GM154655. P.D. was supported by NIH training grant T32AR069512. S.K. was supported by NIH grant K99HG013675. S.R. was supported by the Little Warrior Foundation and the Vilas Associate Fellowship.

Competing interests

S.R. is co-founder of Synpha Biosciences and Percepta Biosciences. The remaining authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-026-77780-x>.

Correspondence and requests for materials should be addressed to Juan I. Fuxman Bass.

Peer review information *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2026

¹Department of Biology, Boston University, Boston, MA, USA. ²Bioinformatics Program, Boston University, Boston, MA, USA. ³Molecular Biology, Cellular Biology and Biochemistry Program, Boston University, Boston, MA, USA. ⁴Tri-institutional Program, Computational Biology and Medicine, New York, NY, USA. ⁵Integrated Program in Biochemistry, University of Wisconsin-Madison, Madison, WI, USA. ⁶Center for Autoimmune Genomics and Etiology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA. ⁷Division of Allergy and Immunology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA. ⁸Center for Cancer Systems Biology (CCSB), Dana-Farber Cancer Institute, Boston, MA, USA. ⁹Department of Genetics, Blavatnik Institute, Harvard Medical School, Boston, MA, USA. ¹⁰Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, MA, USA. ¹¹Division of Genetics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA. ¹²Escuela Profesional de Genética y Biotecnología, Facultad de Ciencias Biológicas, Universidad Nacional Mayor de San Marcos, Lima, Peru. ¹³Laboratorio de Virología, Centro de Investigación Veterinaria de Tandil (CIVETAN), UNCPBA-CICPBA-CONICET, Campus Universitario, Tandil, Buenos Aires, Argentina. ¹⁴Department of Neurosurgery, Baylor Research Institute and Baylor College of Medicine, Temple, TX, USA. ¹⁵Dan L Duncan Comprehensive Cancer Center, Baylor College of Medicine, Houston, TX, USA. ¹⁶Dan L Duncan Institute for Clinical and Translational Research, Baylor College of Medicine, Houston, TX, USA. ¹⁷Center for AI and Biomedical Discovery (AIBD), Baylor Research Institute, Temple, TX, USA. ¹⁸Department of Pathology, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA. ¹⁹Department of Biochemistry, University of Wisconsin-Madison, Madison, WI, USA. ²⁰Great Lakes Bioenergy Research Center, University of Wisconsin-Madison, Madison, WI, USA. ²¹Department of Bacteriology, University of Wisconsin-Madison, Madison, WI, USA. ²²Department of Chemical and Biological Engineering, University of Wisconsin-Madison, Madison, WI, USA. ²³Biological Design Center, Boston University, Boston, MA, USA. ✉ e-mail: foxman@bu.edu