Oncogenic HPV promotes the expression of the long noncoding RNA lnc-FANCI-2 through E7 and YY1

Haibin Liu[a,1], Junfen Xu[a,b,1], Yanqin Yang[c], Xiaohong Wang[a], Ethan Wu[a], Vladimir Majerciak[a], Tingting Zhang[a,d], Renske D. M. Steenbergen[e], Hsu-Kun Wang[f,2], Nilam S. Banerjee[f], Yang Li[b], Weiguo Lu[b], Craig Meyers[g], Jun Zhu[c], Xing Xie[b], Louise T. Chow[f,3], and Zhi-Ming Zheng[a,3]

[a]Tumor Virus RNA Biology Section, HIV Dynamics and Replication Program, National Cancer Institute, National Institutes of Health, Frederick, MD 21702; [b]Department of Gynecologic Oncology, Women's Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang 310006, China; [c]Genome Technology Laboratory, Systems Biology Center, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD 20892; [d]Stomatological Hospital of Tianjin Medical University, Tianjin 300070, China; [e]Amsterdam Universitair Medische Centra, Pathology, Cancer Center Amsterdam, Vrije Universiteit Amsterdam, 1081 HV Amsterdam, The Netherlands; [f]Department of Biochemistry and Molecular Genetics, University of Alabama at Birmingham, Birmingham, AL 35294; and [g]Department of Microbiology and Immunology, Penn State University College of Medicine, Hershey, PA 17033

Contributed by Louise T. Chow, November 23, 2020 (sent for review July 7, 2020; reviewed by Lawrence Banks and Haifan Lin)

Long noncoding RNAs (lncRNAs) play diverse roles in biological processes, but their expression profiles and functions in cervical carcinogenesis remain unknown. By RNA-sequencing (RNA-seq) analyses of 18 clinical specimens and selective validation by RT-qPCR analyses of 72 clinical samples, we provide evidence that, relative to normal cervical tissues, 194 lncRNAs are differentially regulated in high-risk (HR)-HPV infection along with cervical lesion progression. One such lncRNA, *lnc-FANCI-2*, is extensively characterized because it is expressed from a genomic locus adjacent to the *FANCI* gene encoding an important DNA repair factor. Both genes are up-regulated in HPV lesions and in in vitro model systems of HR-HPV18 infection. We observe a moderate reciprocal regulation of *lnc-FANCI-2* and *FANCI* in cervical cancer CaSki cells. In these cells, *lnc-FANCI-2* is transcribed from two alternative promoters, alternatively spliced, and polyadenylated at one of two alternative poly(A) sites. About 10 copies of *lnc-FANCI-2* per cell are detected preferentially in the cytoplasm. Mechanistically, HR-HPVs, but not low-risk (LR)-HPV oncogenes induce *lnc-FANCI-2* in primary and immortalized human keratinocytes. The induction is mediated primarily by E7, and to a lesser extent by E6, mostly independent of p53/E6AP and pRb/E2F. We show that YY1 interacts with an E7 CR3 core motif and transactivates the promoter of *lnc-FANCI-2* by binding to two critical YY1-binding motifs. Moreover, HPV18 increases YY1 expression by reducing miR-29a, which targets the 3′ untranslated region of YY1 mRNA. These data have provided insights into the mechanisms of how HR-HPV infections contribute to cervical carcinogenesis.

human papillomavirus | lncRNA | oncoproteins | transcription | YY1

Human papillomaviruses (HPVs) are a large group of small, nonenveloped, double-stranded DNA tumor viruses. Although HPV infections are usually cleared up by a healthy immune surveillance, persistent infections with high-risk (HR)-HPVs (HPV16, HPV18, and HPV31, etc.) can progress to cancer. Cervical cancer is a major cause of mortality in women worldwide, and almost all cases of cervical cancer are caused by HR-HPV infections, with a worldwide estimated 530,000 new cases and 270,000 deaths each year (1).

Gene expression profiling based on microarray and high-throughput technology has been used to analyze the molecular changes during the progression of cervical tumorigenesis. The identified genes with aberrant expression may contribute to cancer diagnosis, classification of cancer subtype, and development of therapy (2–4). In our recent study employing RNA-sequencing (RNA-seq) technology for whole genome transcriptome analysis, we identified an altered genome-wide expression profile of host genes associated with the carcinogenic progression of HPV-infected cervical lesions (5).

To date, only 2.94% of RNA transcripts from the human genome encode proteins, while the majority of the genomic transcripts are noncoding RNAs (6). We have profiled host microRNAs (miRNAs), one class of noncoding RNAs in the size range of 19 to 23 nts (7), in HPV-infected cells and cervical tissues and reported that a subset of miRNAs are regulated by HR-HPV infections (8–13) and may serve as biomarkers for HR-HPV infections and cervical cancer progression (12, 13). Long noncoding RNAs (lncRNAs) are another class of noncoding RNAs that are over 200 nts in length but lack the coding capacity for a protein over 100 amino acid (aa) residues. lncRNAs have been proposed to function in diverse biological processes (14–16) and they are named by their close proximity to annotated genes (17). First, lncRNAs can regulate transcription by recruiting chromatin modifiers to DNA in *cis* or *trans* (18). For instance, the X-inactive-specific transcript (*Xist*) recruits polycomb repressive complex 2 (PRC2) in *cis* to promote H3K27 methylation of *Xist*, leading to X chromosome (chr) inactivation (19), whereas *HOTAIR* recruits PRC2 in *trans* to mediate epigenetic silencing of tumor suppressor genes (20). Second, lncRNAs regulate chromatin/nuclear structure by binding to specific proteins and function as scaffolds. As an example, architectural lncRNA, *NEAT1*, initiates NONO dimer to aggregation, resulting in paraspeckle assembly (21, 22). Third, lncRNAs

Significance

Persistent infections of oncogenic HPVs can lead to cervical, anogenital, and head-and-neck cancers. We report that HPV-infected cervical lesions display aberrant expression of 194 long-noncoding RNAs, including *lnc-FANCI-2*. We show that, in cell and tissue cultures, *lnc-FANCI-2* is induced primarily by the oncogenic HPV E7 protein via interacting with a cellular transcription factor YY1. HPV infection also increases YY1 protein expression by reducing miR-29a, which targets the 3′ UTR of YY1 RNA. In vivo and in vitro, the nearby DNA repair gene *FANCI* is co-upregulated. These findings highlight an aspect on how oncogenic HPV E7 contributes to pathogenesis.

Author contributions: H.L., X.X., L.T.C., and Z.-M.Z. designed research; H.L., J.X., X.W., E.W., V.M., T.Z., R.D.M.S., H.-K.W., N.S.B., Y.L., and C.M. performed research; R.D.M.S., W.L., C.M., J.Z., and L.T.C. contributed new reagents/analytic tools; H.L., J.X., Y.Y., V.M., J.Z., X.X., L.T.C., and Z.-M.Z. analyzed data; and H.L., L.T.C., and Z.-M.Z. wrote the paper.

Reviewers: L.B., International Center for Genetic Engineering and Biotechnology; and H.L., Yale School of Medicine.

The authors declare no competing interest.

[1]H.L. and J.X. contributed equally to this work.

[2]Present address: Life Sciences R&D, MilliporeSigma, Temecula, CA 92590.

[3]To whom correspondence may be addressed. Email: ltchow@uab.edu or zhengt@exchange.nih.gov.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2014195118/-/DCSupplemental.

Published January 13, 2021.

may act as protein decoys to regulate the functions of cytoplasmic or nuclear proteins. The abundant nuclear lncRNA, *MALAT1*, interacts with splicing factors and regulates their recruitment to pre-mRNAs, thereby affecting their alternative splicing (23). Fourth, lncRNA has also been suggested to bind to PTEN-targeting miRNAs to prevent the reduction of *PTEN* mRNA and thus has tumor-suppressive functions (24). However, this function has been challenged by other studies (25).

Until now, only a few lncRNAs have been identified to associate with cervical cancer (26, 27) and some have been linked to HPV oncoproteins, E6 (28, 29) or E7 (30, 31). However, a systematic genome-wide lncRNA profile during cervical lesion progression is lacking and there has been no detailed study of how HPV infections would lead to aberrant expression of lncRNAs that might contribute to cervical tumorigenesis.

In this report, RNA-seq was used to analyze three groups of cervical tissues: normal tissues without HPV infection, HPV-infected cervical intraepithelial neoplasia grades II-III (CIN 2/3) tissues, and HPV-infected cervical cancer tissues. We identified 194 lncRNAs that were differentially expressed along with cervical lesion progression. One lncRNA, *lnc-FANCI-2*, was chosen for detailed study, because its genomic location is near the *FANCI* gene. This latter gene is part of the Fanconi anemia (FA) pathway for DNA repair. Analogous to the association of specific DNA repair genes with certain cancers (32), patients with a genetic defect in the FA pathway have a high incidence of HR-HPV-induced cancers (33, 34) and FA is also important for HR-HPV DNA amplification (33, 35, 36). We found that the expression of *lnc-FANCI-2* and *FANCI* were co-upregulated by HR-HPV infections primarily by the E7 oncoprotein and that the transcription factor YY1 was essential for *lnc-FANCI-2* expression. In association with E7, YY1 up-regulates the *lnc-FANCI-2* promoter by direct binding to two YY1 binding motifs in the promoter. We also found HR-HPV infections reduce the expression of miR-29a which targets YY1 mRNA, thereby increasing YY1 protein levels.

Results

HPV-Positive Cervical Lesions Display Unique lncRNA Expression Profiles Relative to HPV-Negative Normal Cervix. To explore whether HPV infection of the cervix leads to differential expression of lncRNAs, we performed two independent RNA-seq analyses of 18 individual tissue samples. Each dataset comprised three clinical samples from three groups, including HPV-negative normal cervical tissues (N), oncogenic HPV-positive CIN 2/3, and cervical cancer (CA) tissues. We mapped the RNA-seq raw reads (NCBI GSE accession no. 150227) to the reference human genome (hg19) with TopHat software, counted the mapped reads using bedtools, and employed one-way ANOVA with high stringent conditions to identify differentially expressed lncRNA transcripts based on 113,513 lncRNAs from LNCipedia (v.3) (http://www.lncipedia.org/) (37). From the overlaps between the two RNA-seq datasets, we identified the aberrant expression of 194 lncRNAs in HPV-induced cervical lesions (Fig. 1*A* and *SI Appendix*, Table S1), including the increased expression of *lnc-AFMID-1*, *lnc-C9orf57-2*, *lnc-CDHR1-1*, *lnc-CDKN2B-2*, *lnc-COL7A1-1*, *lncDIS3-4*, *lnc-FANCI-2*, *lnc-GNRHR-5*, *lnc-GPX2-1*, *lnc-KRT6A-1*, *lnc-KRT6B-1*, *lnc-KRT9-2*, *lnc-LYNX1-1*, *lnc-MCM2-3*, *lnc-NMU-1*, *lnc-PSCA-5*, *lnc-PUM2-3*, *lnc-TNFRSF13B-8*, *lnc-UBE2C-1*, and *lnc-UCN2*, and the decreased expression of *lnc-AGR3-2*, *lnc-AHR-1*, *lnc-AMY2A-1*, *lnc-CBLN2*, *lnc-CCNA1-1*, *lnc-CXCL5-1*, *lnc-ELMO1-2*, *lnc-EML4-4*, *lnc-ENPP5-2*, *lnc-FGFBP2*, *lnc-FOXA2-2*, *lnc-GLB1L2-1*, *lnc-MAP2K6-4*, *lnc-NAP1L3-4*, *lnc-NDNF-1*, *lnc-PDZD8-2*, *lnc-PGR-3*, *lnc-SCGB1A1-1*, *lnc-TFF2-1*, and *lnc-WFDC2-3*. The up-regulated *lnc-CDKN2B-2* might be responsible for HR-HPV-induced expression of p16 and p14ARF (38, 39). In contrast, the decreased *lnc-CCNA1-1* expression could be associated with HPV infection-induced CCNA1 (cyclin A1) reduction (40).

lnc-FANCI-2 (*LINC00925*) with increased expression was specifically chosen for in depth investigation because the gene is located at the genomic locus chromosome 15q26.1 (chr 15q26.1) near the *FANCI* gene. FA patients have a high incidence of HR-HPV infection-induced squamous cell carcinomas (33, 34), and *FANCI* gene is involved in DNA repair, important for HR-HPV DNA amplification (33, 35, 36). RT-qPCR confirmed an increased expression of *lnc-FANCI-2* in 48 cervical tissues from HPV-induced CIN 2/3 to cervical cancer relative to 24 normal uninfected tissues (Fig. 1*B*). Interestingly, the high level of *lnc-FANCI-2* expression was found to correlate with a better survival rate of the patients with cervical cancer (*SI Appendix*, Fig. S1*A*).

For comparison, we also examined *lnc-GLB1L2-1* from the group with decreased expression. This lncRNA is expressed from the same genomic locus 11q25 as with the lysosomal β-galactosidase (*β-gal*) gene. The endogenous β-gal is highly expressed and accumulates specifically in senescent cells. Thus, the presence of β-gal has been widely used as a biomarker for senescent and aging cells (41, 42). Indeed, RT-qPCR confirmed the decreased expression in the same cervical tissues relative to the normal cervix (Fig. 1 *B* and *C*). These results are completely in agreement with the RNA-seq data and validate our methodology.

To identify the viral genes responsible for *lnc-FANCI-2*, we examined day 10 raft cultures of human foreskin (HFK) and cervical vaginal (HVK) keratinocytes immortalized by HPV16 or HPV18 (43). These cultures exhibited elevated *lnc-FANCI-2* expression relative to control cells (Fig. 1*D*). Enhanced expression of *lnc-FANCI-2* was also observed in independent isolates of primary foreskin keratinocytes (44) separately transduced with E6/E7 of HR-HPV types (i.e., HPV 16, 18, 31, 33, 35, 45, 51, 52, 59, 66, and 70), but not with LR-HPV11 or cutaneous wart-inducing HPV5 and HPV10 (44). HPV16-containing cervical carcinoma cell lines CaSki and SiHa also exhibited elevated *lnc-FANCI-2* (Fig. 1*E*).

***lnc-FANCI-2* and *FANCI* Were Co-Upregulated in Cervical Lesions and in Model Systems In Vitro.** *lnc-FANCI-2* is positioned at the genomic locus chr15q26.1 upstream of the *RHCG* (~73 kb apart) and the *lnc00928* (~106 kb apart) genes, but downstream of the *FANCI* (~62 kb apart) and the *POLG* (~43 kb apart) genes (see Fig. 4*A*). Given that most reported lncRNAs regulate the expression and functions of the neighboring genes (45–47), we next investigated whether these protein-coding genes might also be modulated in HPV-induced cervical lesions. Analysis of RNA-seq reads-coverage map by Integrative Genomics Viewer (IGV) showed that the expression of *RHCG* or *lnc00928* did not change in cervical lesions; but the increased *lnc-FANCI-2* expression was associated with increased *FANCI* and *POLG*, which codes a catalytic subunit γ of the mitochondrial DNA polymerase (48) (Fig. 2 *A* and *B*). *PDCD2*, a chr6-encoded gene responsible for cell death and apoptosis (49), was randomly chosen as a nonneighboring gene control and showed no coexpression with *lnc-FANCI-2* (Fig. 2*A*). We confirmed by RT-qPCR a significant co-upregulation of *lnc-FANCI-2* and *FANCI* concurrently along with cervical lesion progression in the same 72 clinical cervical samples (Fig. 2 *B* and *C*). The enhanced coexpression was confirmed by RT-qPCR in HFK-derived organotypic raft cultures that were productively infected with HPV18 and harvested on day 8, 12, or 16 (50) (Fig. 2*D*), and on day 11 raft cultures of HFKs acutely transduced with a retrovirus expressing HPV18 E6, E7, or E6/E7 (51, 52) (Fig. 2*E*). We confirmed the HPV18 oncogene expression in the transduced raft cultures (*SI Appendix*, Fig. S1*B*) and their selective effects on the expression of specific host targets (*SI Appendix*, Fig. S1*C*).

High-Risk HPV Infection Induces the Up-Regulation of *lnc-FANCI-2* Primarily by Oncoprotein E7. Given that HR-HPV oncogenes are critical in induction of *lnc-FANCI-2* expression (Fig. 1*E*) and viral E7 appears a stronger inducer than E6 (Fig. 2*E*), we then performed comparative knockdown (KD) of HPV18 E6 and E7

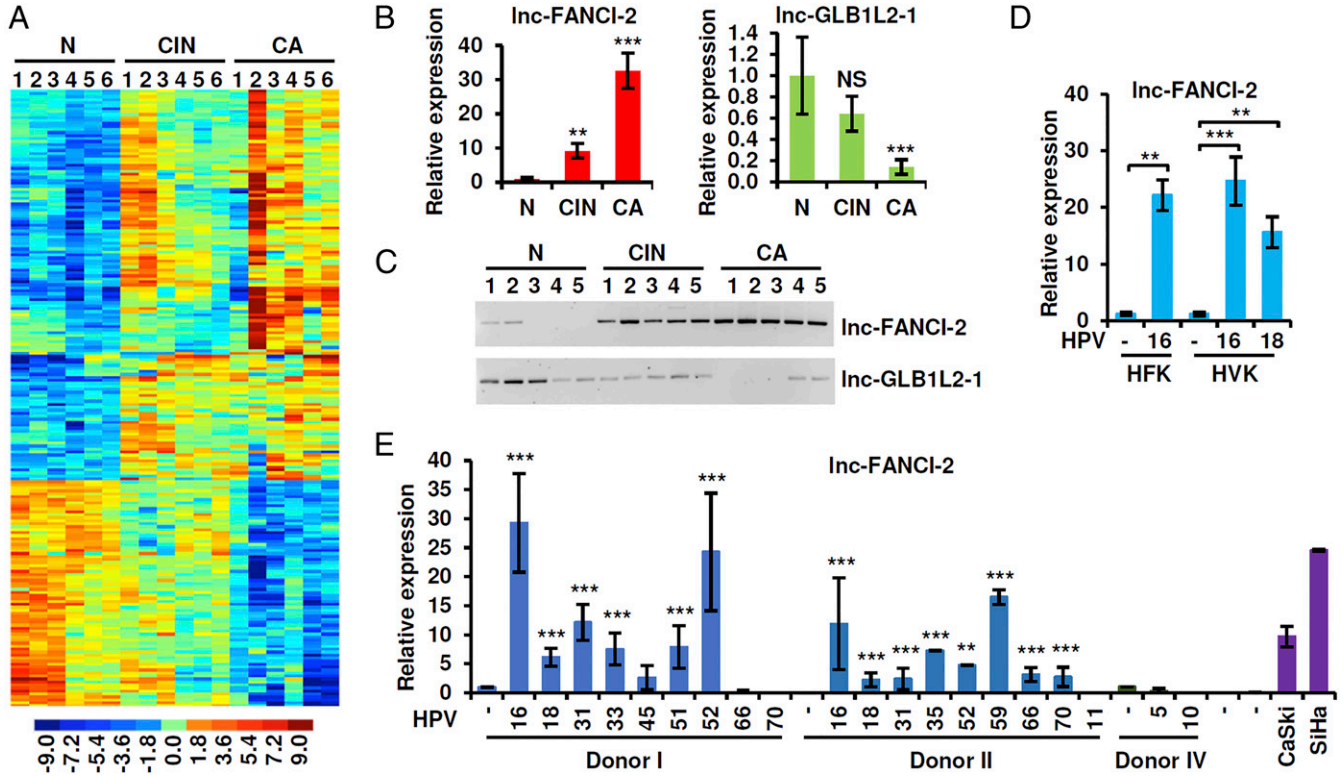

Fig. 1. Oncogenic HPVs promote the expression of *lnc-FANCI-2* during cervical lesion progression and in model systems in vitro. (*A*) Heatmap representation of the 194 lncRNAs expression from HPV-negative normal (N) cervical tissues to oncogenic HPV-positive CIN 2/3 and cervical cancer (CA) tissues. Six individual patient tissues from each group were analyzed by RNA-seq. (*B* and *C*) Increased expression of *lnc-FANCI-2* and decreased expression of *lnc-GLB1L2-1* were verified by RT-qPCR in 24 HPV-negative normal cervix, 25 oncogenic HPV-positive CIN 2/3, and 23 oncogenic HPV-positive CA tissues (*B*). The RT-qPCR products from five clinical samples from each group in *B* were selectively shown using 4% agarose gel electrophoresis (*C*). (*D*) Both HPV16 and HPV18 infections increase *lnc-FANCI-2* expression. Total RNA from HFK or HVK-derived day 10 raft cultures without (–) or with extrachromosomal HPV16 or HPV18 DNA was examined for *lnc-FANCI-2* by RT-qPCR. (*E*) High-risk, but not low-risk, E6/E7-transduced HFK cells exhibit increased expression of *lnc-FANCI-2*. Relative *lnc-FANCI-2* expression in separate donor-derived HFK cells transduced with E6/E7 oncogenes of HR-HPV16, 18, 31, 33, 35, 45, 52, 59, 66, and 70, LR-HPV11, and cutaneous HPV5 and 10 was examined by RT-qPCR, with the *lnc-FANCI-2* expression in HPV16-positive cervical cancer cell lines, CaSki and SiHa cells as controls. $^{**}P < 0.01$; $^{***}P < 0.001$; NS, not significant by two-tailed Student *t* test.

expression with siRNAs (53) (Fig. 3*A*) from HPV18-immortalized HFK monolayer cultures (43) (Fig. 2*F* and *SI Appendix*, Fig. S2*A*) and confirmed the higher activity of E7 than E6 in the induction of *lnc-FANCI-2* expression (Fig. 2*F*, *Upper* bar graph, and *SI Appendix*, Fig. S2*A*). E6-specific siRNA (siE6) reduced E6 without affecting E7. However, E7-specific siRNA (siE7) targeting to both E7 open reading frame (ORF) and E6 RNA 3′-untranslated region (UTR) reduces both E7 and E6 proteins (53, 54). KD efficacy of the siRNAs was verified by the elevated p53 protein (Fig. 2*F*, immunoblots in the *Lower* panel). Collectively, these data are consistent with the observation that HR-HPV E7, more so than HPV E6, is responsible for activation of the FA pathway (55).

To investigate a potential link between the coregulation of *FANCI* and *lnc-FANCI-2*, we subsequently knocked down E6, E7, *FANCI*, or *lnc-FANCI-2* expression in CaSki cells. The results showed that E7 KD had a bigger impact than E6 KD on *lnc-FANCI-2* expression (*SI Appendix*, Fig. S2*B*) and that KD of *FANCI* expression decreased *lnc-FANCI-2* expression by about 25% and vice versa (Fig. 2 *G* and *H*, *Left*). The KD efficacy of the two silnc-FANCI-2 (a and b) was verified by the reduction of *lnc-FANCI-2* (Fig. 2 *H*, *Right*). These observations suggest the existence of a moderate reciprocal regulation of *FANCI* and *lnc-FANCI-2* expression in cervical cancer cells, in addition to being regulated by HR-HPV oncoproteins.

E6 and E7 Regulate *lnc-FANCI-2* Expression Largely Independent of p53/E6AP and pRb/E2f. HR-HPV E6 and E7 oncoproteins destabilize p53 and pRb, respectively, to regulate cell cycle entry and progression (38, 56). To analyze whether the increased *lnc-FANCI-2* expression by viral E6 and E7 might be mediated through p53 and pRb/E2F, KD of E6 or E7 with siRNAs in CaSki cells was conducted as described above. As expected, we found that E6 KD increased the level of p53 (53) (Fig. 3*A*) accompanied by a moderate decrease in the expression of *lnc-FANCI-2*, but this reduction was not altered by simultaneous KD of p53 (Fig. 3*B*). KD of E7 led to a strong reduction of *lnc-FANCI-2*, accompanied by increased p53 level (Fig. 3*C*). E7 KD also reduced the expression of E2F1 and E2F2, but not E2F3 (Fig. 3*D*). This is most likely due to destabilization of the E2F1 protein lacking E7 binding (57) and reduction of E7 transactivating the expression of E2F1 (58) and E2F2 (59). Nevertheless, we assessed whether E2F1 has a role in regulating *lnc-FANCI-2* expression by E2F1 KD. We observed no effect (Fig. 3*C* and *SI Appendix*, Fig. S2*C*). Consistently with the double KD of E6 plus p53 (Fig. 3*B*), neither KD of E6 plus E6AP nor E7 plus E2F1 or pRb led to additional reduction in *lnc-FANCI-2* expression in CaSki cells (*SI Appendix*, Fig. S2 *C* and *D*), although KD of E6 or E7 led to slight reduction in E6AP and pRb, respectively. These results suggest that viral oncoproteins E6 and E7 regulate the expression of *lnc-FANCI-2* expression largely independent of p53/E6AP and pRb/E2F.

Characterization of *lnc-FANCI-2* Gene and Its Expression. The annotated *lnc-FANCI-2* by LNCipedia (v.3) contains seven exons and produces ~35 RNA isoforms (Fig. 4*A* and *SI Appendix*, Fig. S3*A*). During our study, LNCipedia (v.5) and University of California

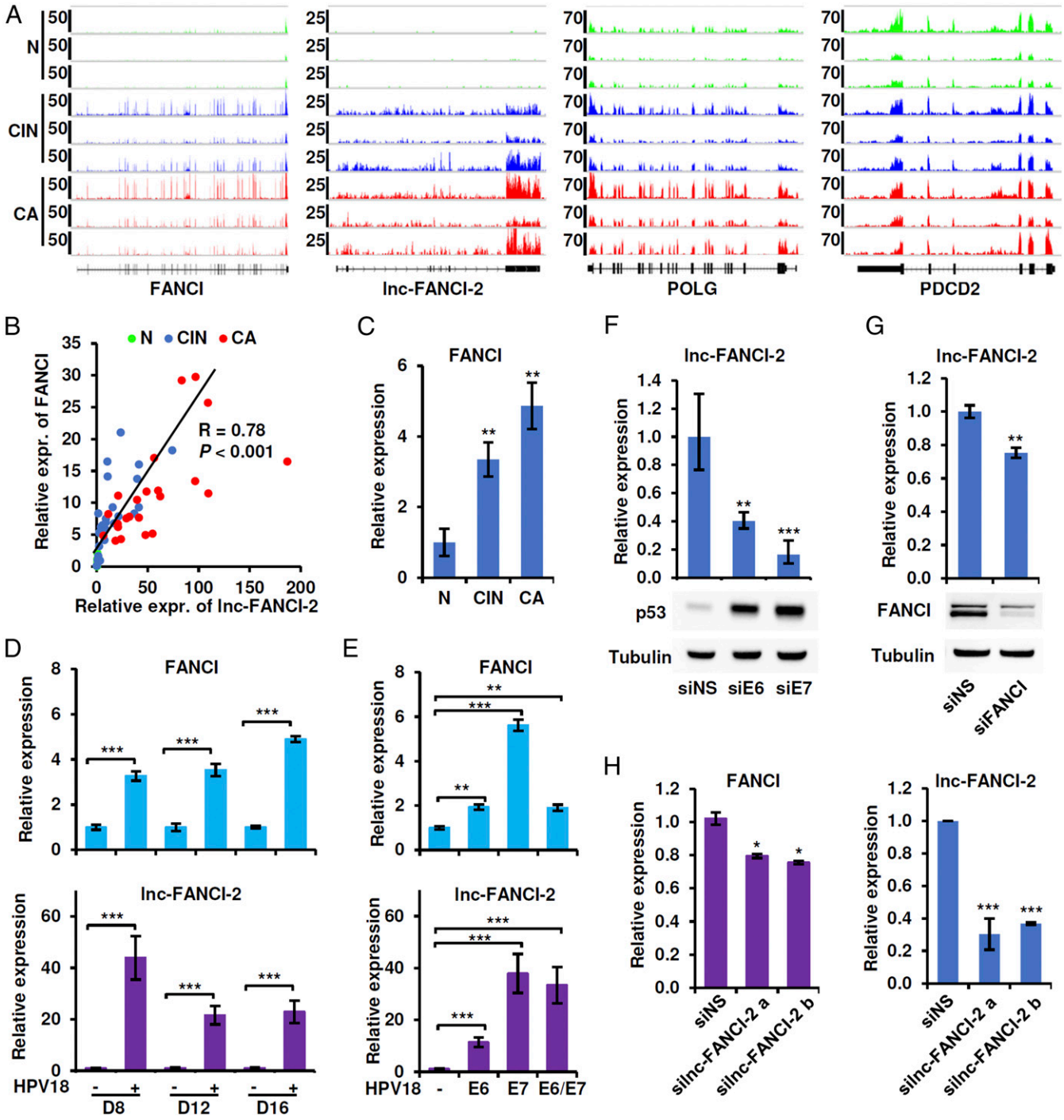

Fig. 2. Oncogenic HPV infection and viral oncoproteins promote the coexpression of *lnc-FANCI-2* and *FANCI*. (*A*) RNA-seq reads-coverage maps by IGV showing the expression levels of *FANCI*, *lnc-FANCI-2*, *POLG*, and *PDCD2* from three groups of cervical tissues, three samples in each group [green, normal (N); blue, CIN 2/3; red, cancer (CA)]. Numbers on the left are reads-coverage depth. (*B* and *C*) Coexpression of *lnc-FANCI-2* and *FANCI* in 24 normal cervix, 25 CIN 2/3, and 23 cervical cancer tissues was quantified by RT-qPCR and displayed by a scatterplot analysis (*B*). *FANCI* expression levels from three separate groups of cervical tissues are shown in a bar graph (*C*). (*D*) Productive HPV18 infection leads to increased expression of *lnc-FANCI-2* and *FANCI*. HFK-derived raft tissues with productive HPV18 infection on day 8 (D8), 12 (D12), and 16 (D16) were examined for the expression of both *FANCI* and *lnc-FANCI-2* RNA by RT-qPCR. (*E*) HPV18 oncoproteins, E6 and E7, are responsible for the increased expression of *FANCI* and *lnc-FANCI-2*. Raft cultures of HFK cells or HFKs acutely transduced with retroviruses expressing HPV18 E6, E7 or both were examined on day 11 for expression of *FANCI* and *lnc-FANCI-2* RNA by RT-qPCR. (*F*) HPV18 E6 and E7 regulates *lnc-FANCI* expression in HPV18-immortalized HFK cells in monolayer cultures. *lnc-FANCI-2* RNA was examined by RT-qPCR after KD of E6 and E7 expression in the HFKs by HPV18 E6- and E7-siRNA (siE6 or siE7) (Fig. 3*A*) (53) or a nonspecific, nontargeting siRNA (siNS) as a control. The increased p53 protein indicated efficient KD of E6 and E7. (*G* and *H*) Reciprocal regulation of *lnc-FANCI-2* and *FANCI* in HPV16-positive CaSki cells. The effects of *FANCI* KD on *lnc-FANCI-2* expression (*G*) or *lnc-FANCI-2* KD on *FANCI* expression were examined in CaSki cells transfected with a pooled *FANCI* siRNA (siFANCI) (*G*) or with two separate siRNAs targeting the *lnc-FANCI-2* exon 6 (silnc-FANCI-2a or 2b) (*H*). siNS was a control. FANCI protein level in *G* was detected by immunoblot to assess the KD efficiency. The expression levels of *FANCI* and *lnc-FANCI-2* RNA were measured by RT-qPCR. **P* < 0.05; ***P* < 0.01; ****P* < 0.001 by two-tailed Student *t* test.

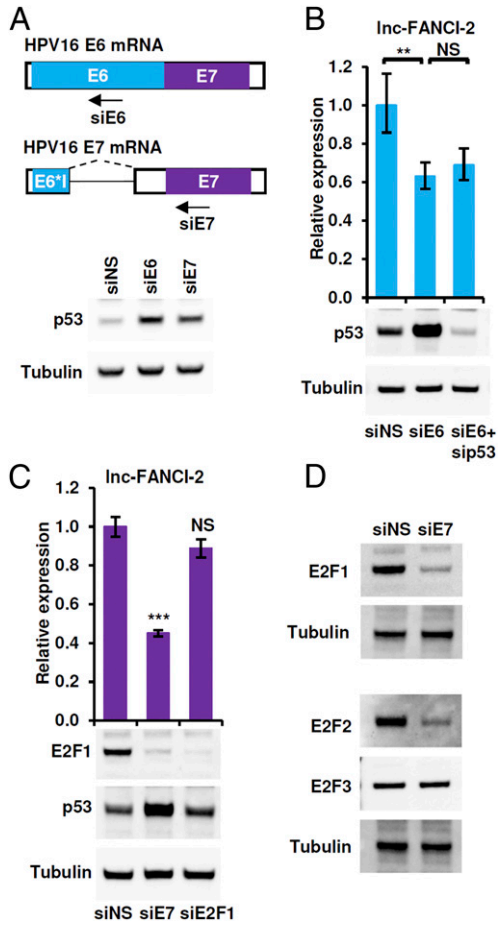

Fig. 3. Viral oncoproteins E6 and E7 in CaSki cells regulate the expression of *lnc-FANCI-2* largely independent of p53 and E2F. (*A*) Diagrams of HPV16 E6 and E7 mRNA transcripts and siRNA strategies for knocking down E6 and E7 expression. As the spliced E6*I, but not the full-length E6 mRNA, serves as an E7 mRNA, the designed siE6 targeting to the intron region of the E6 ORF will not affect the expression of E7. In contrast, siE7 targeting both the E7 ORF and E6 UTR reduced both E6 and E7 proteins. The increased p53 protein levels by immunoblot analysis served as an indication of E6 or E7 KD efficiency in CaSki cells. (*B* and *C*) KD of HPV16 E6 or E7 expression in CaSki cells decreased *lnc-FANCI-2* expression as examined by RT-qPCR upon KD of E6 with siE6, siE6 plus sip53, or a control siNS (*B*), or KD of E7 with siE7, siE2F1, or siNS (*C*). $^{**}P < 0.01$; $^{***}P < 0.001$; NS, not significant by two-tailed Student *t* test. Immunoblots showing increased p53 protein (*B* and *C*) and decreased E2F1 (*C*) are indicative of effective KD of E6 or E7. (*D*) KD of HPV18 E7 expression in HeLa cells as described (53) decreased expression of E2F1 and E2F2, but not E2F3. The samples were examined by immunoblot analysis with indicated antibodies. β-Tubulin served as loading control in each of the panels.

Santa Cruz Genome Browser renamed the *lnc-FANCI-2* as *MIR9-3HG* (NR_015411 or LINC00925) because miR-9-3 coding sequence appears ∼10 kb upstream of the *lnc-FANCI-2* gene (Fig. 4*A*). Since there is no report to validate or characterize the *lnc-FANCI-2* gene or its expression, we experimentally analyzed the gene structure and coding potential of *lnc-FANCI-2* (*SI Appendix*, Fig. S3*B*). Due to the limited patient tissues, we performed these analyses using total RNA extracted from CaSki cells.

As shown in Fig. 4 *A*–*C* and *SI Appendix*, Fig. S4*A*, we identified two major transcription start sites (TSSs) among additional scattered minor TSS sites from this region by 5′ rapid amplification of cDNA ends (RACE), topoisomerase based clonging (TOPO cloning), and sequencing of the RACE products. The TSS2 was more frequently used than the TSS1 (Fig. 4*B* and *SI Appendix*, Fig. S4 *A* and *B* and Table S2). Of the 67 clones sequenced, 16 showed the TSS2 start site at nucleotide 89922290A-92A and 7 mapped to the TSS1 start site at nucleotide 89921335G-37A. The presence of the alternative TSS1 site was confirmed by 5′ RACE using a gene-specific primer Pr5 upstream of the TSS2. Five of 18 clones were mapped to the nucleotide 89921335G-37A positions (Fig. 4*C* and *SI Appendix*, Fig. S4*B* and Table S2). A rare TSS in the intron 4 was also observed by using two *lnc-FANCI-2*-specific primers Pr6 (from exon 5, arrowhead) and Pr7 (from exon 6, arrowhead) in the 5′ RACE assay (*SI Appendix*, Fig. S5*A*).

The 3′ RACE products showed that *lnc-FANCI-2* utilizes one of two alternative polyadenylation (PA) signals for its RNA polyadenylation at two widely separate cleavage sites, a minor unconventional PA1 for cleavage at nucleotide 89939699 and a major conventional PA2 for cleavage at nucleotide 89941718 (Fig. 4*D* and *SI Appendix*, Fig. S4*C*). Northern blot analysis using a ^{32}P-labeled antisense probe Pr8 to exon 6 revealed two major *lnc-FANCI-2* bands in CaSki cells, but not in HPV-negative HEK293T cells (Fig. 4*F*, compare lane 1 with lane 2). The major band of 4-kb size was derived from the PA2 polyadenylation while the minor band of 2-kb size was derived from the PA1 polyadenylation.

Upon additional analyses by using RT-PCR and sequencing, a total of at least 14 *lnc-FANCI-2* isoforms (see *SI Appendix* for individual GenBank accession numbers) were identified and they were derived from the usage of alternative promoters, RNA splicing, and polyadenylation (Fig. 4 *A* and *E* and *SI Appendix*, Figs. S4 *E* and *F* and S5*B*). A full-length *lnc-FANCI-2* has six, not seven exons as described by LNCipedia. A minor set of these isoforms also selected an alternative 3′ splice site b in exon 4 (*SI Appendix*, Fig. S5*C*). These data are consistent with the profile of RNA-seq reads coverage in this genomic region and confirmed that a Refseq transcript *NR-015411* or *LINC00925* (*SI Appendix*, Fig. S4*D*) resembles an isoform of *lnc-FANCI-2* transcribed from the TSS1. We verified this transcript by primer-walking RT-PCR, but did not find exon 1A splicing to exon 1B (from TSS2) as predicted for NR-015411 (Fig. 4*E* and *SI Appendix*, Fig. S4*E*). Moreover, our RNA-seq did not detect any transcripts corresponding to the Refseq transcripts *NR-133001*, *NR-133002*, or *NR-133003* (*SI Appendix*, Fig. S4*D*), nor the *NR-133003* by RT-PCR (*SI Appendix*, Fig. S4*E*, lane 7).

We next examined the subcellular localization of *lnc-FANCI-2* in CaSki cells. RT-qPCR of nuclear and cytoplasmic fractions suggested that *lnc-FANCI-2* appears slightly more enriched in the cytoplasm than the nucleus (Fig. 4*G*). The preferential cytoplasmic localization was confirmed by single molecule RNA in situ hybridization (RNA-ISH) (Fig. 4*H* and Movies S1 and S2). By counting 200 cells, we found each CaSki cell harbors ∼10 copies of *lnc-FANCI-2*. In contrast, ∼25% of protein-coding RNAs and ∼80% of the lncRNAs are present only in 1 or fewer copies per cell (60) and not detectable by Northern blot.

YY1 Binding to *lnc-FANCI-2* Promoter Is Essential for *lnc-FANCI-2* Expression. To understand how HR-HPV E7 regulates *lnc-FANCI-2* transcribed mainly from TSS2, we scanned the genomic ∼500-bp region upstream of TSS2 for potential transcription factor (TF) binding motifs using the JASPAR database (http://jaspar.genereg.net/) and predicted putative YY1, E2F, and TBX binding motifs with a confident score >10 (*SI Appendix*, Table S3) (Fig. 5*A*). We then used biotin-labeled double-stranded DNA oligonucleotides corresponding to the putative TF binding sequences for pulldown-immunoblot assays from either CaSki or HeLa cell nuclear extracts (Fig. 5*B*). Double-stranded SP1 binding oligos served as a positive control. We demonstrated that oligos 4 and 5, that correspond to potential YY1 binding motifs B and A, respectively, pulled down much higher amounts of the YY1 protein than other oligos corresponding to YY1-like motifs. Oligo 1 corresponding to the predicted E2F motif did not show any binding of E2F1 (Fig. 5*B*), E2F2, or E2F3, as beads alone. Oligo 2 matching to the predicted TBX motif did not bind YY1 nor E2F1, but TBX3 was

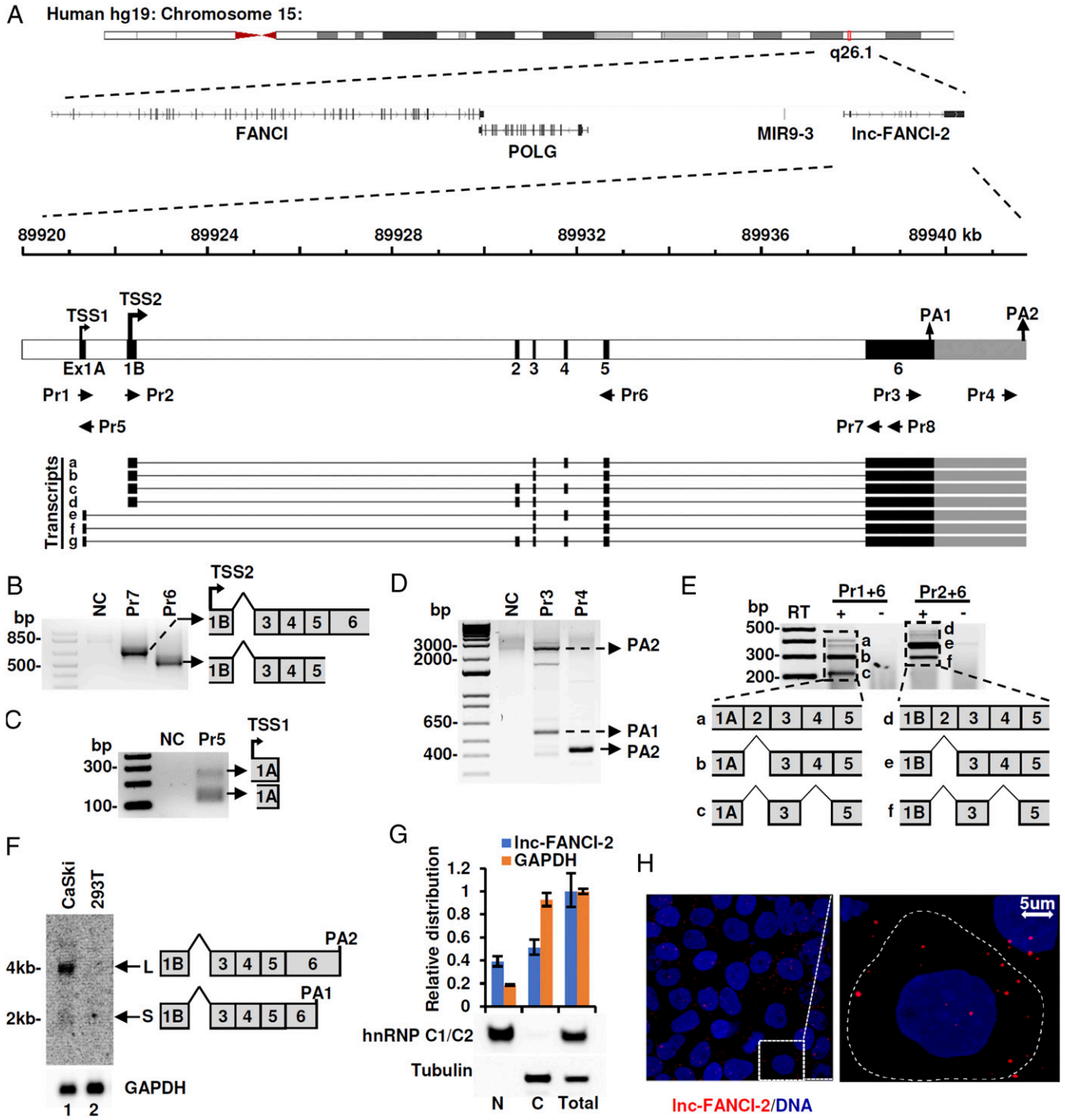

Fig. 4. Characterization of *lnc-FANCI-2* gene structure and expression. (*A*) The genomic locations of *lnc-FANCI-2* and its adjacent upstream genes on chromosome 15. Transcription of all the genes was from left to right. The diagram below represents the genomic position, the two alternative TSSs and two alternative polyadenylation (PA) sites, and six exons (numbered black boxes) in mapped *lnc-FANCI-2* gene using total RNA from CaSki cells. TSS2 or PA2 with a heavier arrow indicates its predominate usage for *lncFANCI-2* expression. The *Lower* panels are verified *lncFANCI-2* RNA isoforms with exons in filled boxes and introns in thin lines. The space between two PA sites is represented by the gray box. Primers used for RACE and RT-PCR are shown under the *lnc-FANCI-2* gene structure. (*B* and *C*) Mapping of *lnc-FANCI-2* TSS by 5′ RACE. 5′ RACE products amplified with different *lnc-FANCI-2* primers were gel purified, cloned, and sequenced. (*D*) Mapping of polyadenylation cleavage sites by 3′ RACE using two separate *lnc-FANCI-2* primers. 3′ RACE products were gel purified and sequenced. The identified two PA cleavage sites are indicated as PA1 and PA2. (*E*) Validation of alternative promoter usage and RNA splicing by primer walking RT-PCR in the presence (+) or absence (–) of reverse transcriptase (RT). RT-PCR were gel purified and sequenced, with the identified RNA isoform of *lnc-FANCI-2* diagrammed below. (*F*) Northern blot analysis of endogenous *lnc-FANCI-2*. Northern blotting was performed with polyA$^+$ RNA selected from CaSki or HEK293T cells and hybridized with a ^{32}P-labeled, lnc-FANCI-2-specific oligo probe. GAPDH polyA$^+$ RNA served as a loading control. (*G*) Subcellular *lnc-FANCI-2* distributions in the cytoplasmic (C) and nuclear (N) fractions of CaSki cells by RT-qPCR with *GAPDH* RNA as a control. Fractionation efficiency of the cytoplasmic (represented by tubulin) and nuclear (represented by hnRNP C1/C2) proteins was verified by immunoblot analysis. (*H*) More abundant cytoplasmic than nuclear *lnc-FANCI-2* in CaSki cells were detected by RNAscope RNA-ISH. See also Movies S1 and S2 for three-dimensional representations.

nonspecifically pulled down by all the oligos and was eliminated for further analyses (Fig. 5*B*). Introduction of point mutations into the putative YY1 binding motif (Fig. 5*C*) confirmed the specific YY1 binding to the wild-type (WT) oligos 4 and 5 (Fig. 5*D*). The specific YY1 binding to the *lnc-FANCI-2* promoter in the CaSki cells was further verified by chromatin immunoprecipitation (ChIP)-PCR assays (Fig. 5*E*) and this binding was reduced approximately threefold by KD of E7 (Fig. 5 *F* and *G*). Furthermore,

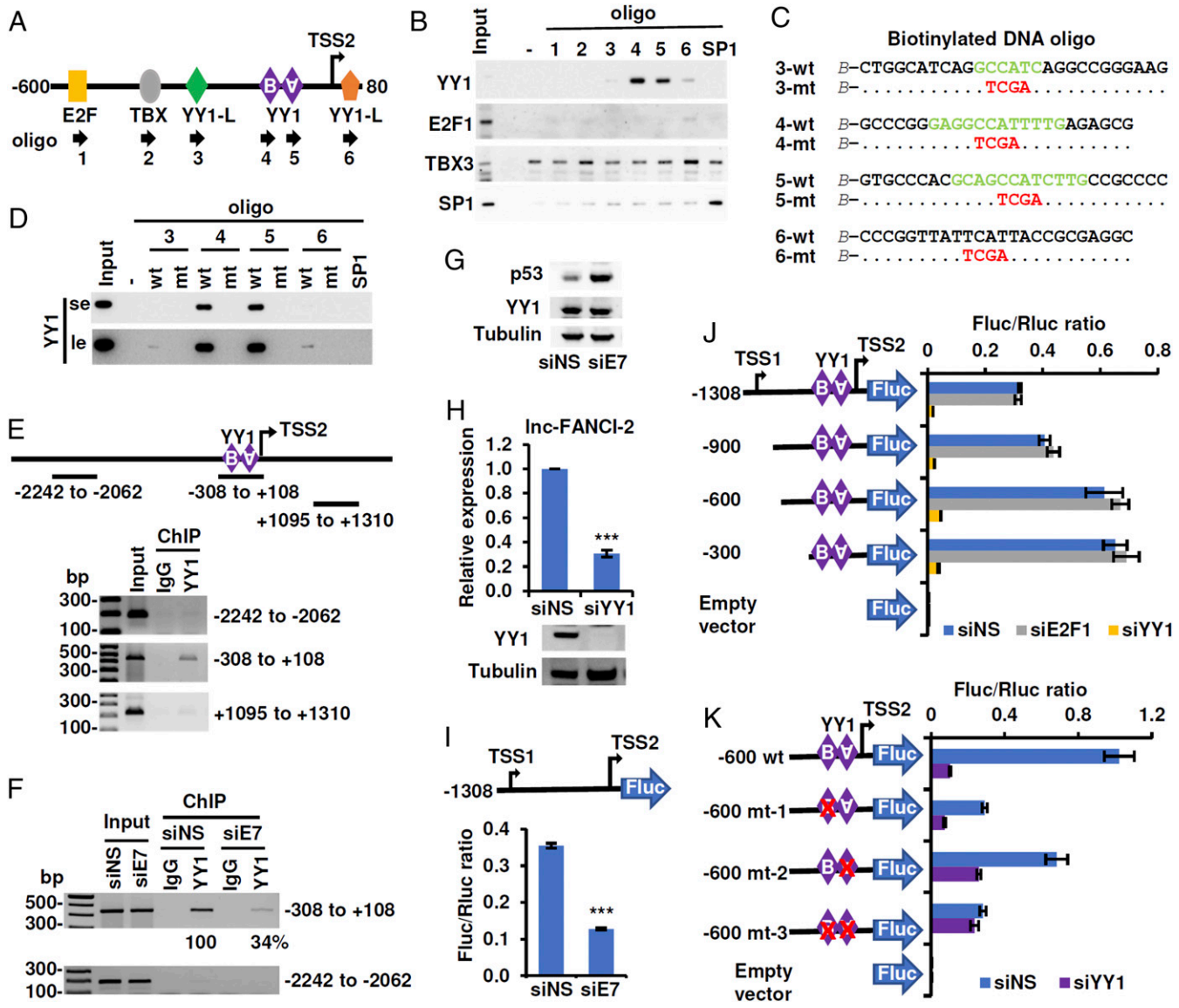

Fig. 5. YY1 is a transcription factor critical for the *lnc-FANCI-2* promoter activity. (*A*) Illustration of putative E2F, TBX, and YY1 binding sites upstream of the transcription start site TSS2 of *lnc-FANCI-2*. YY1-L is a YY1 binding motif-like sequence. Numbered arrows are DNA oligos corresponding to the predicted transcription factor binding sites used for pulldown assays. (*B–D*) YY1 is a major transcription factor. Biotinylated (B) DNA oligos immobilized on avidin beads were mixed with HeLa nuclear extract and the proteins in the pulldowns were immunoblotted as indicated (*B*). Oligos containing a predicted YY1-binding site (in green) without (WT) or with (mt) point mutations (in red) (*C*) were compared in pulldown and immunoblot assays (*D*). Le: long exposure; se, short exposure. (*E*) YY1 binding to the mapped promoter region in CaSki cells. ChIP-PCR was performed with an anti-YY1 antibody on lysates derived from the formaldehyde cross-linked CaSki nuclei by three pairs of primers as diagrammed. The corresponding PCR products from ChIP-enriched and input DNA are shown by agarose gel electrophoresis. (*F–G*) KD of E7 expression reduces the binding of YY1 to the *lnc-FANCI-2* promoter in CaSki cells. ChIP-PCR was performed with an anti-YY1 antibody on chromatin preps from nonspecific siRNA- or E7 siRNA-transfected CaSki cells by two pairs of primers as described in *E*. The PCR band intensity representing the ChIP enrichment was calculated after being normalized to its corresponding input, with YY1 ChIP enrichment in the siNS control as 100%. The levels of p53 for E7 KD efficiency and YY1 protein were examined by immunoblot, with tubulin serving as a protein loading control (*G*). One representative of two ChIP-PCR assays was shown. (*H*) KD of YY1 expression in CaSki cells reduces *lnc-FANCI-2* expression as measured by RT-qPCR. KD efficiency of YY1 was validated by immunoblot below the bar graph, with tubulin as a protein loading control. (*I*) KD of E7 expression affects *lnc-FANCI-2* promoter activity. A Firefly luciferase (FLuc) reporter containing a 1.3-kb DNA fragment (−1308 to +79 nt relative to TSS2) derived from CaSki cells DNA was constructed in a pGL3-basic vector. A dual luciferase assay with Renilla luciferase (Rluc) serving as a promoter control (85) was performed in CaSki cells upon KD by a siNS or HPV16 siE7 (53). ***$P < 0.001$ by two-tailed Student *t* test. (*J*) Delineation of the *lnc-FANCI-2* promoter core region by successive 5′ end deletions of the 1.3-kb fragment in the Fluc reporter vector. The YY1-binding motifs A and B are diagrammed in the promoter region. The promoter activity in CaSki cells treated with siNS, siE2F1, or siYY1 was measured as a ratio of Fluc activity divided by the control Rluc activity. Data were from one representative of three in triplicate. (*K*) Introduction of point mutations into YY1-binding sites in the −600 promoter region identifies the YY1 binding motif B to be crucial for *lnc-FANCI-2* promoter activity in CaSki cells by dual luciferase reporter assays.

we demonstrated that KD of YY1 expression in CaSki cells by YY1-specific siRNA decreased the expression of *lnc-FANCI-2* (Fig. 5*H*) without affecting the short-term growth of the cells (*SI Appendix*, Fig. S6), despite the fact that YY1 has a debatable function in control of viral p97 early promoter activities (61, 62).

We then inserted a CaSki cell-derived 1.3-kb fragment containing TSS1, TSS2, and upstream sequences of the *lnc-FANCI-2* gene in front of the luciferase reporter in a pGL3-basic plasmid and tested its promoter activity in a dual-luciferase assay in CaSki cells (Fig. 5*I*). This promoter activity was regulated by E7 expression, as KD of HPV16 E7 expression significantly reduced its activity (Fig. 5*I*). This result agreed with our observation that E7 KD in CaSki cells reduces *lnc-FANCI-2* expression (Figs. 2*F* and 3*C*). To test whether other transcription factors might also contribute to the regulation of the *lnc-FANCI-2* promoter, a series of successive 5′- to 3′-deletion mutations of the 1.3-kb fragment were examined for their promoter activity. As shown in Fig. 5*J*, all truncated fragments containing the YY1 binding motifs B and A showed strong promoter activities and were equally sensitive to KD of YY1, but not to E2F1. Interestingly, the longer forms displayed lower promoter activities than the shorter forms. Presumably, other factors that bound to the longer (beyond −600 nt) promoter region might negatively regulate the activity (Fig. 5*J*).

We next focused on the −600-nt promoter region by introducing mutations into YY1 binding motifs in CaSki cells with YY1 KD. Point mutations to disrupt individual binding motifs A, B, or both showed that motif B is essential for the promoter activity, whereas motif A positioned in the opposite orientation had partial activity (Fig. 5*K*). Moreover, both motifs' activities were sensitive to YY1 KD. Taking together, we concluded that the expression of *lnc-FANCI-2* is highly dependent on YY1 binding for its promoter activity, in particular the YY1 motif B.

E7 Interacts with YY1 to Promote the Transcription of the *lnc-FANCI-2* Promoter. Because the ability of E7 to regulate the *lnc-FANCI-2* promoter is independent of E2F, we asked whether E7 might interact with YY1. Alternatively, E7 is known to interact with p300, an acetyltransferase, to regulate the activity of YY1 (63–66). GFP-HPV16 E7 (GFP-E7) fusion protein (67) and Flag-YY1 fusion protein were separately expressed in HEK293T cells. By mixing the GFP or GFP-E7 extracts with Flag or Flag-YY1 extracts followed by coimmunoprecipitation (co-IP) using anti-Flag or anti-GFP antibodies and immunoblot assays, we demonstrated anti-Flag antibodies pulling down of Flag-YY1 and associated GFP-E7, but not the GFP (Fig. 6*A*, compare lane 8 to lane 7). Conversely, anti-GFP antibodies pulling down of GFP-E7 and associated Flag-YY1, whereas the expression of GFP alone did not (Fig. 6*B*, compare lane 4 to lane 3). These observations showed that E7 interacts with YY1. This interaction is DNA independent because it was insensitive to DNase treatment of the cell extracts (Fig. 6*C*, compare lane 6 to lane 5). Luciferase reporter assays using the *lnc-FANCI-2* promoter also showed the knockdown of either E7 or YY1 in CaSki cells led to reduction of *lnc-FANCI-2* promoter activities (Fig. 6*D*).

To map the E7-interacting motif with YY1, we constructed a series of HPV16 E7 with deletion or point mutations (Fig. 6*E*) and then fused each of the mutated E7 with GFP for subsequent co-IP analyses. As shown in Fig. 6*F*, deletion of the CR3 domain from E7 (lane 11) or mutations of 68CVQ70AAA in the E7 CR3 core (lane 23) led to >70% reduction of E7 interaction with YY1. In contrast, mutations in CR1 or CR2 had no effect.

Acetyltransferase p300 binds to and acetylates the YY1 transcriptional repression domain and this acetylation is required for the full transcriptional repressor activity of YY1 (65, 66). The literature differs on whether E7 binding to p300 prevents or stimulates p300 acetylation activity (63, 68, 69). We examined the interaction of YY1 and p300 in CaSki cells by co-IP and we detected no interaction between these two proteins (*SI Appendix*, Fig. S7*A*). The lack of YY1-p300 interaction was unrelated to E7 expression since KD of E7 expression in CaSki cells did not promote YY1 interaction with p300 (*SI Appendix*, Fig. S7*B*). YY1 contains a HAT/HDAC interaction domain (HID) which is acetylated by p300 (65) (*SI Appendix*, Fig. S7*C*). We constructed two mutations of the YY1-Flag deleted of the HID domain, Δ170 to 200, or Δ158 to 200 (*SI Appendix*, Fig. S7*C*). By using the dual luciferase assays, we compared the mutations to WT YY1-Flag in rescuing the *lnc-FANCI-2* promoter activity in YY1-KD CaSki cells. This was achieved by transfection of CaSki cells with pooled YY1-specific siRNAs and then with overexpressed ectopic YY1 to titrate the activity of pooled siRNAs. As shown in *SI Appendix*, Fig. S7*D*, a partial rescue of *lnc-FANCI-2* promoter activities was observed in WT YY1-Flag transfected cells as compared with the cells transfected with an empty vector (compare lane 3 to lane 2). Rescue was also seen with the two YY1-Flag deletion mutations Δ170 to 200 and Δ158 to 200 (compare lanes 4 and 5 to lanes 2 and 3). These data support the notion that unacetylated YY1 functions as a transactivator for the *lnc-FANCI-2* promoter.

HPV Infection Promotes YY1 Expression by Inhibition of miR-29a Production. YY1 is a ubiquitous transcription factor in many types of cells and tissues. In HPV18-infected raft cultures, we found that the increased expression of *lnc-FANCI-2* was always accompanied by an increased expression of YY1 (Fig. 7*A*, *Middle* and *Right* bar graphs). Monolayer cultures of HFKs immortalized by episomal HPV18 also exhibited increased expression of YY1 protein (Fig. 7*B*). These data are consistent with the notion that the increased *lnc-FANCI-2* expression (Figs. 1 *A*–*C* and 2*A*) coincided with increased YY1 expression in HPV-positive CIN 2/3 and cervical cancer tissues (Fig. 7*C*).

Increased YY1 expression in cervical cancer cells appears to be associated with miR-29 down-regulation (11). Our study shows that HFK raft cultures with productive HPV18 infection displayed a decreased expression of miR-29a (13) (Fig. 7*A*, *Left* bar graph). Likewise, we found day 11 raft tissues transduced by E6 or E7 retrovirus infection led to a moderate reduction of primary miR-29a transcripts (*SI Appendix*, Fig. S8 *A* and *B*). This was probably mediated through a p53 responsive element (p53RE) in the *MIR29A* gene promoter (*SI Appendix*, Fig. S8 *A* and *C*) as E6 reduces p53 and E7-stabilized p53 is transcriptionally inactive (70, 86). Further study showed that KD of p53 in U2OS cells led to reduced expression of precursor and mature miR-29a (*SI Appendix*, Fig. S8*D*).

To investigate how YY1 protein levels could be linked to miR-29 levels, we used the TargetScan program (http://www.targetscan.org/vert_72/) to scan the *YY1* 3′ UTR against the human miRBase (http://www.mirbase.org/). We found a consensus miR-29a seed match in the *YY1* 3′ UTR, which is broadly conserved among vertebrates. We confirmed the functionality of the miR-29a seed match in CaSki cells. As shown in Fig. 7*D*, blocking miR-29a activity in CaSki cells by transfection of two forms of anti-miR29a, but not a nonspecific anti-miR, increased YY1 expression. Conversely, overexpression of miR-29a by transfection of miR-29 mimics decreased YY1 expression. Next, we substituted the 3′ UTR of the Renilla luciferase gene with a *YY1* 3′ UTR in a psiCHECK plasmid (Fig. 7*E*, *Top* diagram). Luciferase reporters were then constructed to contain the *YY1* 3′ UTR with or without point mutations that disrupt the miR-29a seed match. These reporters were then examined in HEK293T cells where the Firefly luciferase serves as an internal control. Luciferase activities were then measured in response to overexpression of miR-29a mimics. As shown in Fig. 7*E* (bar graphs), miR-29a mimics at the two doses tested showed strong inhibition of the luciferase activities only when the reporter contained the WT 3′ UTR of *YY1*. There was no inhibition when the reporter had a mutated 3′ UTR of *YY1* or the luciferase (parent) 3′ UTR. These data provide strong evidence that miR-29a down-regulates YY1 expression through a

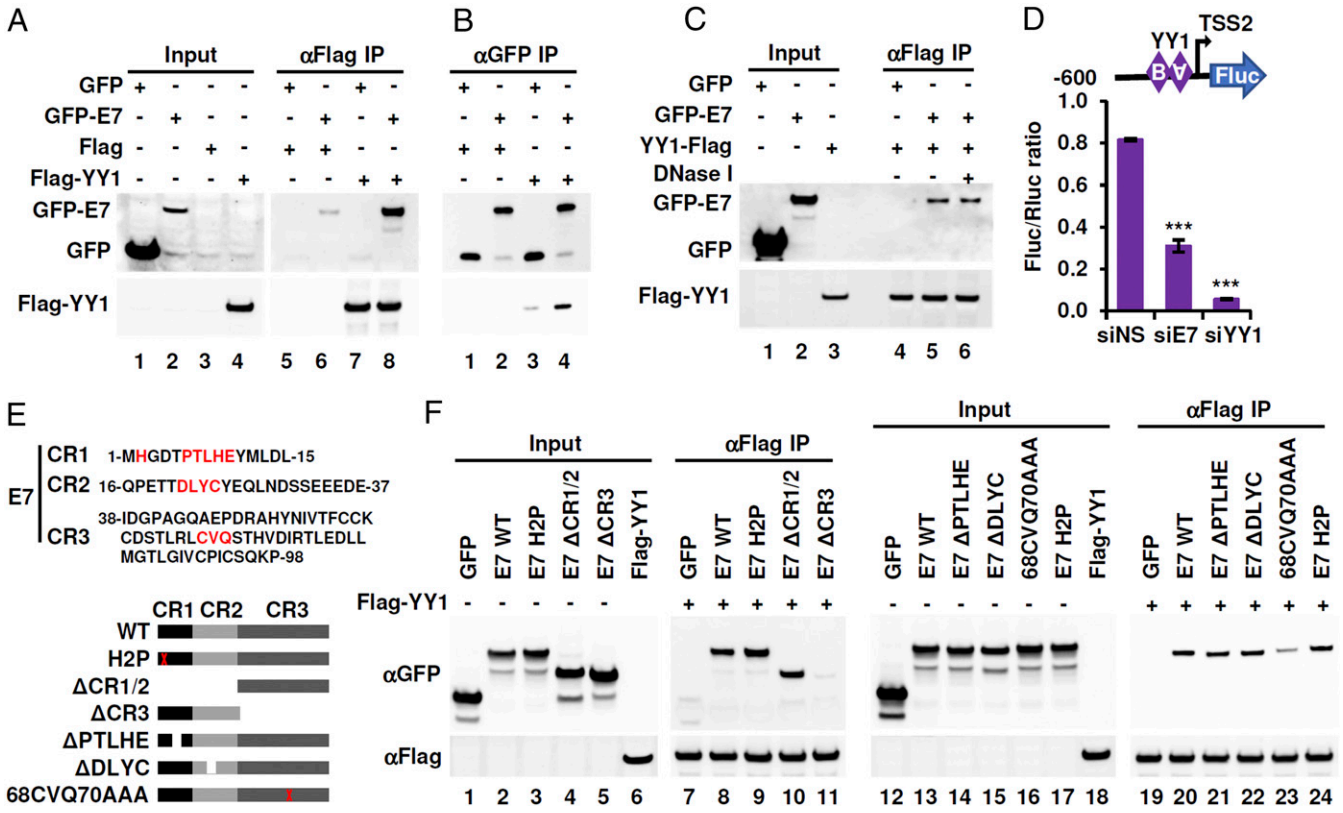

Fig. 6. HPV16 E7 in association with the YY1 protein promotes transcription of *lnc-FANCI-2*. (*A–C*) Ectopic E7 and YY1 interaction by co-IP. HEK293T cells were transfected with YY1-Flag, Flag-control, GFP-E7, or GFP control, respectively. Total cell lysate separately prepared were mixed followed by co-IP with an anti-Flag (*A*) or anti-GFP (*B*) antibody. YY1 and E7 proteins in IPed samples were immunoblotted with an anti-GFP (*A* and *B*) for GFP-E7, and with anti-Flag (*A*) or anti-YY1 (*B*) for Flag-YY1. The cell lysate treated with or without DNase I (*C*) was immunoprecipitated by an anti-Flag antibody and blotted for GFP-E7 and Flag-YY1 as in *A*. (*D*) KD of E7 or YY1 expression affects *lnc-FANCI-2* promoter activity. The activities of a truncated (−600) *lnc-FANCI-2* promoter were evaluated in CaSki cells with or without E7 or YY1 KD by dual luciferase assays as a ratio of Fluc/Rluc activity. ***$P < 0.001$ by two-tailed Student *t* test. (*E*) HPV16 E7 aa residues and conserved regions of CR1, CR2, and CR3 used for deletion mapping. The aa residues in red are mutated or deleted in the diagrammed E7 mutants below. (*F*) The 68CVQ70 of E7 CR3 was mapped to interact with YY1. HEK293T cells were transfected with YY1-Flag, E7-GFP, or mutated E7-GFP. Total cell lysate of YY1-Flag was mixed with cell lysates containing wild-type or mutant form of E7-GFP before IP with an anti-Flag antibody. YY1-Flag and E7-GFP proteins in the immunoprecipitated complexes were immunoblotted with an anti-GFP for E7-GFP or anti-Flag for YY1-Flag.

seed match in the *YY1* 3′ UTR. Thus, HR-HPV infections can elevate YY1 protein levels partially by reducing miR-29a.

Discussion

Oncogenic HPV infections can lead to epithelial cancers in women and men. Mechanistically, HPVs destabilize host tumor suppressor proteins and also dysregulate the expression of a subset of host genes. In this study, we discovered that 194 lncRNAs, including *lnc-FANCI-2*, were differentially expressed in association with cervical lesion progression induced by HR-HPV infection. Importantly, the *lnc-FANCI-2* expression was found to be a biomarker for better survival of the patients with cervical cancer (*SI Appendix*, Fig. S1*A*). We showed that the up-regulation is also observed in vitro in productive infection by HR-HPV in HFKs and HR-HPV immortalized HFKs and HVKs. The induction is mediated primarily by E7, but to a lesser extent by E6, and largely independent of pRB/E2F and p53/E6AP, the well-known targets of HR-HPV oncoproteins.

We have characterized the structure and expression of the *lnc-FANCI-2* gene. We showed that expression of *lnc-FANCI-2* is initiated mainly from the transcription start site TSS2, and the promoter contains two highly conserved YY1 binding motifs upstream of the TSS2. Our studies demonstrate that both YY1 and E7 are essential for the up-regulation of *lnc-FANCI-2* during HR-HPV infection. This YY1 activity depends on E7. It has been reported that YY1 transcriptional repression function requires acetylation by acetyltransferase p300 (65, 66), but it is inactivated by viral E6 and E7 binding (63, 68, 69). However, we did not observe any interaction between YY1 and p300 regardless of E7 (*SI Appendix*, Fig. S7). Thus, p300 appears not to play a significant role in regulating the *lnc-FANCI-2* promoter in our experimental settings.

We discovered that E7 activates *lnc-FANCI-2* transcription through two mechanisms. First, we show that miRNA29a targets the 3′ UTR of YY1 (Fig. 7), thereby decreasing YY1 protein levels. However, HPV infections (Fig. 7), specifically E6 and E7 expression, reduces the production of miR-29a (11, 13), thereby increasing the YY1 expression. Second, we report an association of the E7 CR3 motif with YY1 (Fig. 6) and E7 KD reduces YY1 binding to the *lnc-FANCI-2* promoter in CaSki cells. Thus, we postulate that the complex alters the activity of YY1 bound to the lnc-FANCI-2 promoter, ultimately leading to its transactivation (Fig. 8).

lncRNAs are known to be involved in the *cis* regulation of target genes located at or near the same genomic locus (45–47). We found that *lnc-FANCI-2* and the nearby *FANCI* were co-upregulated in vivo in HR-HPV-infected cervical lesions and in vitro in productive infection of HPV18 in HFK raft cultures and in raft cultures expressing HPV18 oncogenes. Furthermore, in vitro, the expression of *lnc-FANCI-2* affects the expression of *FANCI* or vice versa. Although a possible mechanism for this mutual regulation remains unknown and deletion mutation of individual genes in the context of host genome would facilitate

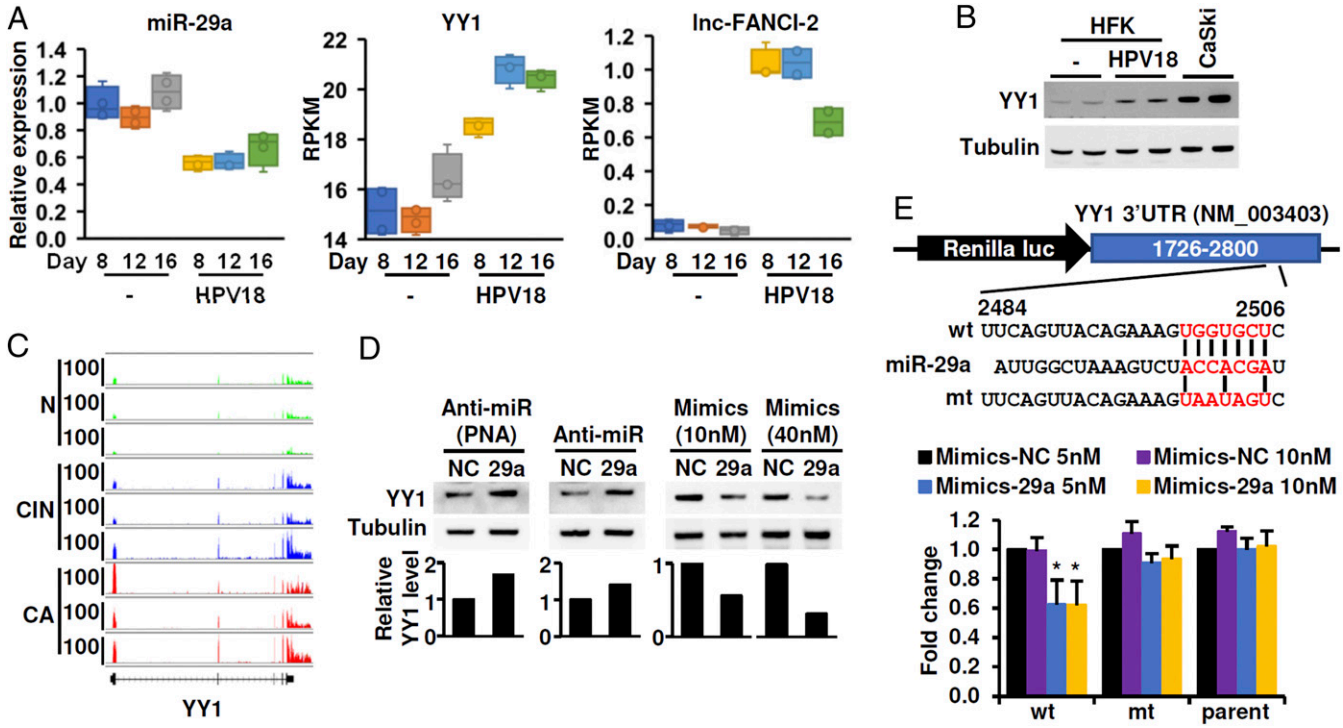

Fig. 7. miR-29a targets the 3′ UTR of YY1 and reduces YY1 protein expression. (*A*) HPV18 infection decreases miR-29a expression, thereby increasing the expression of YY1 and *lnc-FANCI-2*. Total RNA from uninfected or HPV18-infected HFK raft cultures on days 8, 12, and 16 were measured for the expression of miR-29a by RT-qPCR and for *YY1* and *lnc-FANCI-2* by RNA-seq. RPKM, reads per kilobase per million mapped reads. (*B*) HPV18 infection increases the expression of YY1 protein in monolayer HFK cells. HFKs or HFKs immortalized by episomal HPV18 DNA were immunoblotted for YY1 protein. Proteins from CaSki cells served as a loading control. (*C*) RNA-seq reads coverage showing *YY1* expression from normal (N), CIN 2/3, or cancerous (CA) cervical tissues visualized by IGV. Numbers on the *Left* are reads-coverage depth. (*D*) miR-29a in CaSki cells regulates YY1 expression. Lysates of CaSki cells transfected with 40 nM of miR-29a inhibitors (anti-miR-29a) or nonspecific control (NC), one of which was a peptide nucleic acid (PNA)-based form, or with miR-29a mimics or a mimics nonspecific control (mimics-NC) were immunoblotted for YY1 expression with tubulin as a sample loading control. The relative YY1 level in each sample is shown as a bar graph after normalizing to tubulin. (*E*) The 3′ UTR of *YY1* contains a miR-29a seed region sensitive to miR-29-mediated reduction of YY1. The 3′ UTR of the Renilla luciferase reporter in a psiCHECK-2 plasmid was replaced by the *YY1* 3′ UTR with a WT or mt miR-29a seed match (red) and examined for its response to miR-29a mimics in HEK293T cells. Renilla luciferase activities were measured as a ratio to the control Firefly luciferase activity, with the parent psiCHECK-2 plasmid as a control. Data are one representative in triplicate of three independent experiments. $*P < 0.05$ by two-tailed Student *t* test.

this investigation, our observations suggest that *lnc-FANCI-2* might be involved in DNA repair. The FA pathway involves 22 *FANC* genes to carry out repair of damaged DNA. In particular, FANCI and FANCD2 form a heterodimer which clamps DNA at the site of DNA damage in a DNA repair reaction (35, 71–74). HPV infection, specifically the expression of E7 in postmitotic, differentiated keratinocytes promotes S phase reentry and activates both ataxia telangiectasia mutated (ATM)- and ataxia telangiectasia and Rad3-related protein (ATR)-mediated DNA damage response pathways. Both are critical for viral DNA amplification which occurs during a prolonged G2 phase following host DNA replication (75–77). Our observation that the HR-HPV induction of *FANCI* expression along with *lnc-FANCI-2* is in line with these reports.

But the ability of E7 to promote S phase reentry through destabilizing pocket proteins (pRb and p130) in the differentiated keratinocytes is a double-edged sword. E7 also plays a key role in inducing DNA damage, mitotic defects, and host genome instability (55, 78, 79) when it is overexpressed in proliferating cells by promoting excessive host DNA replication, especially when coupled with the E6 protein activity, which inactivates and destabilizes p53, the guardian of the genome (77, 80–82). These unintended detrimental effects of inappropriate overexpression is thought to occur in normally quiescent basal stem or reserve cells during repeated wound and healing in infected epithelium in patients, leading to cell immortalization, transformation, and lesion progression (83). Indeed, FA patients with a genetic defect in the FA pathway are more susceptible to HPV-associated squamous cell carcinomas (34). FA-deficient mice expressing E7, but not E6, also displays high cancer incidence of female lower reproductive tract and head and neck (81, 84).

How does one reconcile this dichotomy of *FANCI* and *lnc-FANCI-2*? On the one hand, FA patients are highly susceptible to developing cervical cancer. But on the other hand, HR-HPV elevated FANCI and *lnc-FANCI-2*. We suggest that because the FA pathway is also important for HR-HPV DNA amplification (33, 35, 36), difficulties in replicating HR-HPV DNA might increase the probability of viral DNA integration into the host chromosome, thereby increasing the likehood of developing cervical cancer. The fact that E7 not only up-regulates the expression of *lnc-FANCI-2* and *FANCI* genes, but also activates ATM and ATR DNA damage response pathways (75–80), suggests that, for both viral productive infection and during viral carcinogenesis, maintaining a steady level of genome stability is crucial for cell viability.

As a final note, our characterization of the *lnc-FANCI-2* gene structure and expression provides strong evidence that *lnc-FANCI-2* expression is independent of miR-9-3 transcription, despite all transcripts from this region containing *lnc-FANCI-2* have been recently annotated as *MIR9-3HG* transcripts in the updated RefSeq database (https://www.ncbi.nlm.nih.gov/refseq/). We suggest restoring the nomenclature to *lnc-FANCI*-2 originally designated in LNCipedia (v.3) for the following reasons: 1) We demonstrate that *lnc-FANCI-2* transcripts are initiated mainly from two alternative promoters, alternatively spliced and terminated at one of two

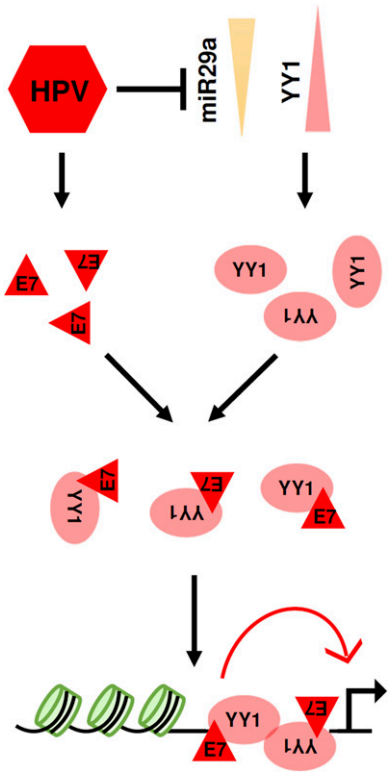

Fig. 8. Proposed model of how HR-HPV E7 enhances *lnc-FANCI-2* transcription during HR-HPV infection through modulation of miR-29a expression and E7 interaction with YY1.

alternative polyadenylation sites, generating at least 14 isoforms (Fig. 4). The major promoter is regulated by HR-HPV E7 and YY1 (Fig. 5). Notably, the minor *lnc-FANCI-2* isoform transcribed from its TSS1 resembles National Center for Biotechnology Information (NCBI) *NR_015411* or *LINC00925*; 2) RNA-seq reads in our study show the predominant coverage in the *lnc-FANCI-2* coding region, with only a few reads in the *MIR9-3* coding region or upstream (*SI Appendix*, Fig. S4*D*); and 3) there are no detectable *miR9-3HG* transcripts extending into the *lnc-FANCI-2* coding region in the CaSki cells by RT-PCR (*SI Appendix*, Fig. S4*E*). Human miR-9 is known to be transcribed from chr1q22 and chr5q14.3 and there is no direct evidence that miR-9 can derive from the chr15q26.1 region. Rather, a mature miR-9 sequence was only computationally predicted from the chr15q26.1 region, near *lnc-FANCI-2*.

In conclusion, we have discovered differential expression of a subset of lncRNAs during HR-HPV-induced cervical carcinogenesis. Characterization of one of these lncRNAs, *lnc-FANCI-2*, led us to demonstrate its coregulation with *FANCI*, a DNA repair gene, and the important roles of E7 and YY1 in regulating *lnc-FANCI-2* transcription. These findings reveal a role of E7 during HPV pathogenesis and might lead to biomarkers for early diagnosis of HR-HPV infection and for development of possible intervention strategies to treat HPV-induced cancers.

Materials and Methods

Detailed descriptions of patient samples, human primary keratinocytes, organotypic (raft) epithelial cultures, cell line cultures, siRNA transfection, RNA extraction, RNA-seq, data processing, RT-PCR and RT-qPCR, 5′ RACE, 3′ RACE, antibodies and immunoblot, Northern blot, RNA-ISH, DNA oligo pulldown assays, ChIP, protein expression vectors, promoter reporter plasmids and dual luciferase assays, IP, cell proliferation assay, and GenBank accession numbers of the identified *lnc-FANCI-2* isoforms are provided in *SI Appendix*.

Data Availability. All study data are included in the article and supporting information. The GenBank has assigned the following accession numbers for the identified *lnc-FANCI-2* RNA isoforms in this report: MT669800 for lnc-FANCI-2a-PA2, MT669801 for lnc-FANCI-2a-PA1, MT669802 for lnc-FANCI-2b-PA2, MT669803 for lnc-FANCI-2b-PA1, MT669804 for lnc-FANCI-2c-PA2, MT669805 for lnc-FANCI-2c-PA1, MT669806 for lnc-FANCI-2d-PA2, MT669807 for lnc-FANCI-2d-PA1, MT669808 for lnc-FANCI-2e-PA2, MT669809 for lnc-FANCI-2e-PA1, MT669810 for lnc-FANCI-2f-PA2, MT669811 for lnc-FANCI-2f-PA1, MT669812 for lnc-FANCI-2g-PA2, and MT669813 for lnc-FANCI-2g-PA1.

ACKNOWLEDGMENTS. This study was supported by the Intramural Research Program of the NIH, the NCI, and the CCR (ZIASC010357 to Z.-M.Z.); the National Key Research and Development Program of China (2016YFC1302900 to W.L.); the National Science Foundation of China (81702552, to J.X. and W.L.); a NIH grant (CA83679 to L.T.C., N.S.B., and H.-K.W.); Anderson Endowment Funds through the University of Alabama at Birmingham (UAB) (to L.T.C.); and pilot project grants from the UAB Comprehensive Cancer Center (5P30CA013148-43 and 316851 to N.S.B.). We thank Karl Munger of Tufts University for his E7 mutants from Addgene and Kai Ge of NIDDK for his assistance on p300.

1. L. A. Torre *et al.*, Global cancer statistics, 2012. *CA Cancer J. Clin.* **65**, 87–108 (2015).
2. Y. Chen *et al.*, Identification of cervical cancer markers by cDNA and tissue microarrays. *Cancer Res.* **63**, 1927–1935 (2003).
3. D. Gius *et al.*, Profiling microdissected epithelium and stroma to model genomic signatures for cervical carcinogenesis accommodating for covariates. *Cancer Res.* **67**, 7113–7123 (2007).
4. T. Rajkumar *et al.*, Identification and validation of genes involved in cervical tumourigenesis. *BMC Cancer* **11**, 80 (2011).
5. J. Xu *et al.*, Genome-wide profiling of cervical RNA-binding proteins identifies human papillomavirus regulation of RNASEH2A expression by viral E7 and E2F1. *MBio* **10**, e02687-18 (2019).
6. ENCODE Project Consortium, An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).
7. V. K. Gangaraju, H. Lin, MicroRNAs: Key regulators of stem cells. *Nat. Rev. Mol. Cell Biol.* **10**, 116–125 (2009).
8. X. Wang *et al.*, Oncogenic HPV infection interrupts the expression of tumor-suppressive miR-34a through viral oncoprotein E6. *RNA* **15**, 637–647 (2009).
9. X. Wang *et al.*, Aberrant expression of oncogenic and tumor-suppressive microRNAs in cervical cancer is required for cancer cell growth. *PLoS One* **3**, e2557 (2008).
10. X. Wang, C. Meyers, M. Guo, Z. M. Zheng, Upregulation of p18Ink4c expression by oncogenic HPV E6 via p53-miR-34a pathway. *Int. J. Cancer* **129**, 1362–1372 (2011).
11. Y. Li *et al.*, Progressive miRNA expression profiles in cervical carcinogenesis and identification of HPV-related target genes for miR-29. *J. Pathol.* **224**, 484–495 (2011).
12. Q. Tian *et al.*, MicroRNA detection in cervical exfoliated cells as a triage for human papillomavirus-positive women. *J. Natl. Cancer Inst.* **106**, dju241 (2014).
13. X. Wang *et al.*, microRNAs are biomarkers of oncogenic human papillomavirus infections. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 4262–4267 (2014).
14. J. E. Wilusz, H. Sunwoo, D. L. Spector, Long noncoding RNAs: Functional surprises from the RNA world. *Genes Dev.* **23**, 1494–1504 (2009).
15. J. L. Rinn, H. Y. Chang, Genome regulation by long noncoding RNAs. *Annu. Rev. Biochem.* **81**, 145–166 (2012).
16. R. W. Yao, Y. Wang, L. L. Chen, Cellular functions of long noncoding RNAs. *Nat. Cell Biol.* **21**, 542–551 (2019).
17. M. W. Wright, A short guide to long non-coding RNA gene nomenclature. *Hum. Genomics* **8**, 7 (2014).
18. N. Brockdorff, Noncoding RNA and polycomb recruitment. *RNA* **19**, 429–442 (2013).
19. J. Zhao, B. K. Sun, J. A. Erwin, J. J. Song, J. T. Lee, Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science* **322**, 750–756 (2008).
20. R. A. Gupta *et al.*, Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* **464**, 1071–1076 (2010).
21. T. Yamazaki *et al.*, Functional domains of NEAT1 architectural lncRNA induce paraspeckle assembly through phase separation. *Mol. Cell* **70**, 1038–1053.e7 (2018).
22. Y. T. Sasaki, T. Ideue, M. Sano, T. Mituyama, T. Hirose, MENepsilon/beta noncoding RNAs are essential for structural integrity of nuclear paraspeckles. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 2525–2530 (2009).
23. V. Tripathi *et al.*, The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol. Cell* **39**, 925–938 (2010).
24. L. Poliseno *et al.*, A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature* **465**, 1033–1038 (2010).
25. R. Denzler, V. Agarwal, J. Stefano, D. P. Bartel, M. Stoffel, Assessing the ceRNA hypothesis with quantitative measurements of miRNA and target abundance. *Mol. Cell* **54**, 766–776 (2014).
26. W. J. Wu *et al.*, Integrated analysis of long non-coding RNA competing interactions revealed potential biomarkers in cervical cancer: Based on a public database. *Mol. Med. Rep.* **17**, 7845–7858 (2018).
27. H. Zhu *et al.*, Long non-coding RNA expression profile in cervical cancer tissues. *Oncol. Lett.* **14**, 1379–1386 (2017).

28. S. Sharma, K. Munger, Expression of the cervical carcinoma expressed PCNA regulatory (CCEPR) long noncoding RNA is driven by the human papillomavirus E6 protein and modulates cell proliferation independent of PCNA. *Virology* **518**, 8–13 (2018).
29. J. A. Barr *et al*., Long non-coding RNA FAM83H-AS1 is regulated by human papillomavirus 16 E6 independently of p53 in cervical cancer cells. *Sci. Rep.* **9**, 3662 (2019).
30. S. Sharma *et al*., Bridging links between long noncoding RNA HOTAIR and HPV oncoprotein E7 in cervical cancer pathogenesis. *Sci. Rep.* **5**, 11724 (2015).
31. H. He *et al*., Human papillomavirus E6/E7 and long noncoding RNA TMPOP2 mutually upregulated gene expression in cervical cancer cells. *J. Virol.* **93**, e01808–e01818 (2019).
32. F. Dietlein, L. Thelen, H. C. Reinhardt, Cancer-specific defects in DNA repair pathways as targets for personalized therapeutic approaches. *Trends Genet.* **30**, 326–339 (2014).
33. E. E. Hoskins *et al*., The fanconi anemia pathway limits human papillomavirus replication. *J. Virol.* **86**, 8131–8138 (2012).
34. D. I. Kutler *et al*., Human papillomavirus DNA and p53 polymorphisms in squamous cell carcinomas from Fanconi anemia patients. *J. Natl. Cancer Inst.* **95**, 1718–1721 (2003).
35. S. Khanal, D. A. Galloway, High-risk human papillomavirus oncogenes disrupt the Fanconi anemia DNA repair pathway by impairing localization and de-ubiquitination of FancD2. *PLoS Pathog.* **15**, e1007442 (2019).
36. C. C. Spriggs, L. A. Laimins, FANCD2 binds human papillomavirus genomes and associates with a distinct set of DNA repair proteins to regulate viral replication. *MBio* **8**, e02340-16 (2017).
37. P. J. Volders *et al*., LNCipedia: A database for annotated human lncRNA transcript sequences and structures. *Nucleic Acids Res.* **41**, D246–D251 (2013).
38. A. Roman, K. Munger, The papillomavirus E7 proteins. *Virology* **445**, 138–168 (2013).
39. H. Kanao *et al*., Correlation between p14(ARF)/p16(INK4A) expression and HPV infection in uterine cervical cancer. *Cancer Lett.* **213**, 31–37 (2004).
40. K. Chalertpet, W. Pakdeechaidan, V. Patel, A. Mutirangura, P. Yanatatsaneejit, Human papillomavirus type 16 E7 oncoprotein mediates CCNA1 promoter methylation. *Cancer Sci.* **106**, 1333–1340 (2015).
41. Y. Tang *et al*., Downregulation of splicing factor SRSF3 induces p53β, an alternatively spliced isoform of p53 that promotes cellular senescence. *Oncogene* **32**, 2792–2798 (2013).
42. G. P. Dimri *et al*., A biomarker that identifies senescent human cells in culture and in aging skin in vivo. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 9363–9367 (1995).
43. M. E. McLaughlin-Drubin, C. Meyers, Propagation of infectious, high-risk HPV in organotypic "raft" culture. *Methods Mol. Med.* **119**, 171–186 (2005).
44. D. M. Schütze *et al*., Differential in vitro immortalization capacity of eleven (probable) [corrected] high-risk human papillomavirus types. *J. Virol.* **88**, 1714–1724 (2014).
45. M. Ebisuya, T. Yamamoto, M. Nakajima, E. Nishida, Ripples from neighbouring transcription. *Nat. Cell Biol.* **10**, 1106–1113 (2008).
46. S. Guil, M. Esteller, Cis-acting noncoding RNAs: Friends and foes. *Nat. Struct. Mol. Biol.* **19**, 1068–1075 (2012).
47. J. M. Engreitz *et al*., Local regulation of gene expression by lncRNA promoters, transcription and splicing. *Nature* **539**, 452–455 (2016).
48. M. A. Graziewicz, M. J. Longley, W. C. Copeland, DNA polymerase gamma in mitochondrial DNA replication and repair. *Chem. Rev.* **106**, 383–405 (2006).
49. B. W. Baron *et al*., Repression of the PDCD2 gene by BCL6 and the implications for the pathogenesis of human B and T cell lymphomas. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 7449–7454 (2007).
50. H. K. Wang, A. A. Duffy, T. R. Broker, L. T. Chow, Robust production and passaging of infectious HPV in squamous epithelium of primary human keratinocytes. *Genes Dev.* **23**, 181–194 (2009).
51. S. Cheng, D. C. Schmidt-Grimminger, T. Murant, T. R. Broker, L. T. Chow, Differentiation-dependent up-regulation of the human papillomavirus E7 gene reactivates cellular DNA replication in suprabasal differentiated keratinocytes. *Genes Dev.* **9**, 2335–2349 (1995).
52. N. J. Genovese, N. S. Banerjee, T. R. Broker, L. T. Chow, Casein kinase II motif-dependent phosphorylation of human papillomavirus E7 protein promotes p130 degradation and S-phase induction in differentiated human keratinocytes. *J. Virol.* **82**, 4862–4873 (2008).
53. S. Tang, M. Tao, J. P. McCoy Jr, Z. M. Zheng, The E7 oncoprotein is translated from spliced E6*I transcripts in high-risk human papillomavirus type 16- or type 18-positive cervical cancer cell lines via translation reinitiation. *J. Virol.* **80**, 4249–4263 (2006).
54. S. Tang, M. Tao, J. P. McCoy Jr, Z. M. Zheng, Short-term induction and long-term suppression of HPV16 oncogene silencing by RNA interference in cervical cancer cells. *Oncogene* **25**, 2094–2104 (2006).
55. N. Spardy *et al*., The human papillomavirus type 16 E7 oncoprotein activates the Fanconi anemia (FA) pathway and causes accelerated chromosomal instability in FA cells. *J. Virol.* **81**, 13265–13270 (2007).
56. S. B. Vande Pol, A. J. Klingelhutz, Papillomavirus E6 oncoproteins. *Virology* **445**, 115–137 (2013).
57. S. G. Hwang, D. Lee, J. Kim, T. Seo, J. Choe, Human papillomavirus type 16 E7 binds to E2F1 and activates E2F1-driven transcription in a retinoblastoma protein-independent manner. *J. Biol. Chem.* **277**, 2923–2930 (2002).
58. L. Qiao, Q. Zhang, W. Zhang, J. J. Chen, The lysine acetyltransferase GCN5 contributes to human papillomavirus oncoprotein E7-induced cell proliferation via up-regulating E2F1. *J. Cell. Mol. Med.* **22**, 5333–5345 (2018).
59. M. S. Longworth, R. Wilson, L. A. Laimins, HPV31 E7 facilitates replication by activating E2F2 transcription through its interaction with HDACs. *EMBO J.* **24**, 1821–1830 (2005).
60. S. Djebali *et al*., Landscape of transcription in human cells. *Nature* **489**, 101–108 (2012).
61. S. H. Tan, C. C. Baker, W. Stünkel, H. U. Bernard, A transcriptional initiator overlaps with a conserved YY1 binding site in the long control region of human papillomavirus type 16. *Virology* **305**, 486–501 (2003).
62. M. May *et al*., The E6/E7 promoter of extrachromosomal HPV16 DNA in cervical cancers escapes from cellular repression by mutation of target sequences for YY1. *EMBO J.* **13**, 1460–1466 (1994).
63. A. Bernat, N. Avvakumov, J. S. Mymryk, L. Banks, Interaction between the HPV E7 oncoprotein and the transcriptional coactivator p300. *Oncogene* **22**, 7871–7881 (2003).
64. A. L. Jansma *et al*., The high-risk HPV16 E7 oncoprotein mediates interaction between the transcriptional coactivator CBP and the retinoblastoma protein pRb. *J. Mol. Biol.* **426**, 4030–4048 (2014).
65. Y. L. Yao, W. M. Yang, E. Seto, Regulation of transcription factor YY1 by acetylation and deacetylation. *Mol. Cell. Biol.* **21**, 5979–5991 (2001).
66. T. Bauknecht, R. H. See, Y. Shi, A novel C/EBP beta-YY1 complex controls the cell-type-specific activity of the human papillomavirus type 18 upstream regulatory region. *J. Virol.* **70**, 7695–7705 (1996).
67. M. Tao, M. Kruhlak, S. Xia, E. Androphy, Z. M. Zheng, Signals that dictate nuclear localization of human papillomavirus type 16 oncoprotein E6 in living cells. *J. Virol.* **77**, 13232–13247 (2003).
68. H. J. Dyson, P. E. Wright, Role of intrinsic protein disorder in the function and interactions of the transcriptional coactivators CREB-binding protein (CBP) and p300. *J. Biol. Chem.* **291**, 6714–6722 (2016).
69. L. Lv *et al*., Mitogenic and oncogenic stimulation of K433 acetylation promotes PKM2 protein kinase activity and nuclear localization. *Mol. Cell* **52**, 340–352 (2013).
70. E. Y. Kho, H. K. Wang, N. S. Banerjee, T. R. Broker, L. T. Chow, HPV-18 E6 mutants reveal p53 modulation of viral DNA amplification in organotypic cultures. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 7542–7549 (2013).
71. R. Ceccaldi, P. Sarangi, A. D. D'Andrea, The fanconi anaemia pathway: New players and new functions. *Nat. Rev. Mol. Cell Biol.* **17**, 337–349 (2016).
72. P. Alcón *et al*., FANCD2-FANCI is a clamp stabilized on DNA by monoubiquitination of FANCD2 during DNA repair. *Nat. Struct. Mol. Biol.* **27**, 240–248 (2020).
73. D. Lopez-Martinez *et al*., Phosphorylation of FANCD2 inhibits the FANCD2/FANCI complex and suppresses the fanconi anemia pathway in the absence of DNA damage. *Cell Rep.* **27**, 2990–3005.e5 (2019).
74. M. Castella *et al*., FANCI regulates recruitment of the FA core complex at sites of DNA damage independently of FANCD2. *PLoS Genet.* **11**, e1005563 (2015).
75. C. A. Moody, Impact of replication stress in human papillomavirus pathogenesis. *J. Virol.* **93**, e01012–e01017 (2019).
76. N. S. Banerjee, D. Moore, C. J. Parker, T. R. Broker, L. T. Chow, Targeting DNA damage response as a strategy to treat HPV infections. *Int. J. Mol. Sci.* **20**, 5455 (2019).
77. N. S. Banerjee, H. K. Wang, T. R. Broker, L. T. Chow, Human papillomavirus (HPV) E7 induces prolonged G2 following S phase reentry in differentiated human keratinocytes. *J. Biol. Chem.* **286**, 15473–15482 (2011).
78. S. Duensing *et al*., The human papillomavirus type 16 E6 and E7 oncoproteins cooperate to induce mitotic defects and genomic instability by uncoupling centrosome duplication from the cell division cycle. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 10002–10007 (2000).
79. S. Duensing, K. Münger, The human papillomavirus type 16 E6 and E7 oncoproteins independently induce numerical and structural chromosome instability. *Cancer Res.* **62**, 7075–7082 (2002).
80. C. A. Moody, L. A. Laimins, Human papillomaviruses activate the ATM DNA damage pathway for viral genome amplification upon differentiation. *PLoS Pathog.* **5**, e1000605 (2009).
81. J. W. Park, M. K. Shin, P. F. Lambert, High incidence of female reproductive tract cancers in FA-deficient HPV16-transgenic mice correlates with E7's induction of DNA damage response, an activity mediated by E7's inactivation of pocket proteins. *Oncogene* **33**, 3383–3391 (2014).
82. B. A. Johnson, H. L. Aloor, C. A. Moody, The Rb binding domain of HPV31 E7 is required to maintain high levels of DNA repair factors in infected cells. *Virology* **500**, 22–34 (2017).
83. L. T. Chow, T. R. Broker, Human papillomavirus infections: Warts or cancer? *Cold Spring Harb. Perspect. Biol.* **5**, a012997 (2013).
84. J. W. Park *et al*., Deficiencies in the Fanconi anemia DNA damage response pathway increase sensitivity to HPV-associated head and neck cancer. *Cancer Res.* **70**, 9959–9968 (2010).
85. S. Tang, K. Yamanegi, Z. M. Zheng, Requirement of a 12-base-pair TATT-containing sequence and viral lytic DNA replication in activation of the Kaposi's sarcoma-associated herpesvirus K8.1 late promoter. *J. Virol.* **78**, 2609–2614 (2004).
86. A. Eichten, M. Westfall, J. A. Pietenpol, K. Munger, Stabilization and functional impairment of the tumor suppressor p53 by the human papillomavirus type 16 E7 oncoprotein. *Virology* **295**, 74–85 (2002).