Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

The TGF-β/ID1 Signaling Pathway Axis Mediates HPV E6/E7 Induced Malignant Transformation of Mammary Epithelial Cells

Xing Xiong
ShanghaiTech University
Shilin Zhang
ShanghaiTech University
Zhi Man
ShanghaiTech University
Meijin Wang
ShanghaiTech University
Wei Xiong
ShanghaiTech University
Zhenhua Li
rmyylizhh@sdsmu.edu.cn

Shandong Second Medical Universityhina
Li Fu
ShanghaiTech University

Research Article

Keywords: Breast cancer, E6/E7, Malignant transformation, TGF-β signaling pathway, ID1

Posted Date: November 16th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11619892/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Breast cancer is the most prevalent malignancy among women worldwide. Increasing evidence implicates human papillomavirus (HPV) in breast cancer, particularly HPV16 and HPV18. However, the underlying mechanisms remain unclear. This study investigated the role of the TGF-β/ID1 signaling axis in HPV16/18 E6/E7 mediated malignant transformation of mammary epithelial cells. Stable MCF-10A cell lines expressing E6/E7 were established. E6/E7 expression promoted malignant transformation of mammary epithelial cells. RNA sequencing revealed extensive transcriptional reprogramming, with enrichment of cell cycle and DNA replication pathways. Gene set enrichment analysis and flow cytometry confirmed accelerated cell cycle progression, with increased S and G2/M phase populations. Mechanistically, E6/E7 upregulated TGFB2 and enhanced SMAD3 phosphorylation, indicating activation of canonical TGF-β signaling. TGFB2 silencing reduced SMAD3 activation and suppressed proliferation. E6/E7 also induced the downstream target ID1, whereas TGFB2 knockdown decreased ID1 expression. ID1 depletion inhibited proliferation and induced cell cycle arrest. In summary, E6/E7 promotes cell cycle progression and tumor initiating capacity via activation of the TGFB2–TGF-β–ID1 axis, providing mechanistic insight into HPV associated breast malignant transformation.

Introduction

Breast cancer is the most prevalent malignant tumor among women worldwide[1,2]. According to the 2022 global cancer statistics released by the World Health Organization, there were 2.3 million new cases of female breast cancer, accounting for 23.8% of all newly diagnosed malignancies in women[3]. The continuous rise in breast cancer incidence underscores the significant public health importance of elucidating its underlying pathogenic mechanisms. Studies have identified multiple risk factors closely associated with breast cancer development, such as hormonal imbalances, mutations in genetic susceptibility genes, radiation exposure, obesity, aging[2,4]. Recent years, viral infection has increasingly attracted attention as a potential etiological contributor[5,6].

Human papillomavirus (HPV) is a circular double stranded DNA virus that exclusively infects humans, with more than 200 genotypes identified to date[7,8]. HPV infection is associated with the development of various epithelial malignancies, particularly HPV16 and HPV18[8–10]. Following HPV infection, the viral oncoproteins E6/E7 promote malignant transformation by degrading or functionally inactivating key tumor suppressors such as p53 and pRb, thereby disrupting cell cycle regulation and driving uncontrolled proliferation[11,12]. E6/E7 proteins can alter the expression of microRNAs (miRNAs) in host cells through direct and indirect effects, such as miR-218[13], miR-23b[14], miR-34a[15] and miR-15a/16–1[16]. In recent years, HPV DNA has also been detected in a subset of breast cancer tissues, with reported prevalence rates of approximately 20%–30%[17–20]. Some studies further suggest that HPV positive breast cancer patients may have an increased risk of disease development and poorer prognosis[21–23]. However, whether E6/E7 proteins contribute to breast cancer initiation by regulating specific signaling pathways has not yet been investigated.

ID1 (Inhibitor of DNA Binding 1) can form nonfunctional heterodimers with bHLH transcription factors, thereby inhibiting their DNA binding activity and regulating cell differentiation and fate[24]. ID1 has been shown to be a key regulator of the tumor initiating cell (TIC) phenotype in multiple solid tumors[25–27]. In breast cancer, ID1 can independently induce a TIC phenotype in mammary epithelial cells and directly promote the acquisition of stem-like cell traits during malignant transformation[28].The transforming growth factor-beta (TGF-β) signaling pathway is a key regulatory axis controlling cell fate determination, proliferation, and differentiation[29,30]. The canonical TGF-β pathway activates SMAD2/3, which then forms a complex with SMAD4 and translocates to the nucleus to regulate the transcription of multiple target genes[31]. This pathway plays diverse roles in the initiation and progression of cancer[32–35]. Under specific molecular contexts, TGF-β can activate downstream targets associated with stemness maintenance and malignant transformation[36]. Studies have confirmed that ID1 is an important regulatory target of the TGF-β signaling pathway[37].

Based on the background, this study focuses on whether E6/E7 promotes stemness maintenance and enhanced tumor initiation capacity by regulating the TGF-β/ID1 pathway, thereby providing a new perspective on the pathogenesis of breast cancer.

Material and Methods

Cell culture

The human mammary epithelial cell line MCF-10A and human embryonic kidney cell line HEK293T were purchased from Zhejiang Bodi Biotechnology Co., Ltd. (Zhejiang, China) in April 2024. Both cell lines were authenticated by the supplier using short tandem repeat (STR) profiling prior to delivery. No additional STR authentication was performed after receipt of the cells. All cell-based experiments were completed by October 2025. MCF-10A cells were cultured in commercially available complete medium (Procell, China) containing DMEM/F12 supplemented with 5% horse serum (HS), 20 ng/mL epidermal growth factor (EGF), 0.5 μg/mL hydrocortisone, 10 μg/mL insulin, 1% non-essential amino acids (NEAA), and 1% penicillin–streptomycin (P/S). HEK293T cells were cultured in DMEM (Gibco) supplemented with 10% fetal bovine serum (FBS) and 1% P/S (Gibco). All cells were maintained at 37°C in a humidified atmosphere containing 5% CO_2.

Real-time PCR (qPCR)

Total RNA was extracted using the AFTMag Quick Tissue/cell RNA extraction Kit for animals (ABclonal) using a fully automated nucleic acid extractor. Subsequently, 1 μg of total RNA was reverse transcribed into cDNA in a 20 μL reaction volume using the HiScript® III All-in-one RT SuperMix Perfect for qPCR (Vazyme, Nanjing, China). Real-time PCR was performed using ChamQ Universal SYBR qPCR Master Mix (Vazyme) with the cDNA synthesized as above. YWHAZ was used as an endogenous control for

normalization. The relative fold changes in expression were analyzed using 2^-△△Ct. All primers used are listed in Table 1.

primerID	sequence(5'-3')
TGFB2-1-F	CCCCGGAGGTGATTTCCATC
TGFB2-1-R	GAAGAAGGGCGGCATGTCTA
TGFB2-2-F	AAAGCCAGAGTGCCTGAACA
TGFB2-2-R	GTGCAGCAGGGACAGTAA
ID1-1-F	GCTCTACGACATGAACGGCT
ID1-1-R	GTCGATGACGTGCTGGAGAA
ID1-2-F	CGTGCTGCTCTACGACATGA
ID1-2-R	TGTAGTCGATGACGTGCTGG
siTGFB2-1	AAGCAGAGUUCAGAGUCUUUCGUUU
siTGFB2-2	CCACAUAUACCAGUGGUGAUCAGAA
siID1-1	CCUACUAGUCACCAGAGACUUTT
siID1-2	CUCGGAAUCCGAAGUUGGAAC

Western blotting

Western blotting

Cells that were treated were collected, rinsed, and broken down on ice with RIPA buffer containing protease inhibitors. The lysates were sonicated, clarified by high speed centrifugation at 4°C, and protein levels were measured using a BCA assay. The protein lysates were mixed with Laemmli sample buffer and heated at 95°C for 5 minutes to denature them. Proteins ranging from 10 to 25 µg were then separated using SDS-PAGE. The wet transfer method was used to transfer proteins onto PVDF membranes, which were then blocked with 5% skimmed milk powder for 60 minutes at room temperature. The primary antibody was used to incubate the membranes overnight at 4°C, with specific dilutions and antibody details outlined in the Table 2. Membranes were treated with the correct HRP-conjugated secondary antibody for one hour at room temperature after being washed with TBST. Membranes underwent further washing and were then exposed to ECL chemiluminescent substrate for 60 seconds at room temperature, shielded from light. A chemiluminescence imaging system was used to capture the images.

antibody	manufactures	product code	dilution rate
HPV16E6	GeneTex	GTX637546	1: 500
HPV16E7	Santa Cruz	sc-6981	1: 100
HPV18E6	Santa Cruz	sc-365089	1: 100
HPV18E7	Abcam	ab100953	1: 1000
ID1	HUABIO	ET1705-49	1: 2000
SMAD2/3	HUABIO	RT1566	1: 1000
pSMAD2	HUABIO	HA722443	1: 1000
pSMAD3	HUABIO	ET1609-41	1: 2000
TGFB2	Proteintech	83167-5-RR	1: 1000
GAPDH	Proteintech	60004-1-IG	1: 20000
YWHAZ	HUABIO	HA500368	1: 1000
Goat Anti-Mouse IgG H&L (HRP)	Abcam	AB205719	1: 10000
Goat Anti-Rabbit IgG	Abcam	AB205718	1: 10000

Cell proliferation assay

During the logarithmic growth phase, 1000 cells were introduced into each well of 96-well plates. After seeding, at the 4h, 24h, 48h, 72h, 96h, and 120h time points, the existing medium was replaced with 100 μL of CCK-8 mixture (90% basal medium + 10% CCK-8 reagent). After one hour incubation period, the plates were shaken for 30 seconds, and absorbance at 450 nm was recorded using a microplate reader. Data points representing cell number were collected on Days 0, 1, 2, 3, 4, and 5 for growth curve construction.

Colony formation assays

For the colony formation assays, 1000 cells suspensions were plated into 6-well plates in triplicate. After incubation for 7 days in a 5% CO_2 incubator set at 37°C, the cells were fixed in formalin and stained with 0.5% crystal violet (dissolved in ddH2O). After fixing the colonies, they were photographed and counted. Only colonies with 50 cells or more were counted.

3D tumor spheroid formation experiment

Prepare a 1% agarose solution in advance and sterilize it. Mix it with an equal volume of MCF-10A special culture medium containing 20% FBS to prepare a 0.5% agarose solution. Add 400 μL to each well

of a 24-well plate. The cells to be tested were digested, centrifuged, resuspended, counted, and 10,000 cells were added to a 0.25% agarose solution. Supplement with 100ul of culture medium every 3 days. On day 0 and day 21, images were taken at 4x and 20x magnification using a microscope, and the results were counted.

Cell cycle analysis

Cells were collected through trypsin digestion, spun down, and the resulting pellets were resuspended in pre-cooled 70% ethanol for fixation for at least two hours at 4°C. Cells that were fixed underwent centrifugation, had ethanol removed, were washed with PBS, and then resuspended in a propidium iodide (PI) staining solution for 30 minutes at room temperature in the dark. Using flow cytometry, researchers analyzed cellular DNA content. The cell cycle phases (G0/G1, S, G2/M) were identified with MODFIT5 software.

Wound healing assay

Cells were seeded into 6-well plates. When confluence was reached, the medium was replaced by a serum-free medium (SFM). Subsequently, two linear, parallel wounds were created per well using a pipette tip positioned vertically against the plate edge. Images were captured at 0 h and 12 h post-wounding at four designated positions per well. Wound closure was analysed using ImageJ software.

RNA sequencing and data analysis

After extracting cellular RNA, perform quality assessment, and then reverse transcribe it into cDNA. Libraries harboring distinct indices were then subjected to multiplexed pooling. Sequencing was conducted on the DNBSEQ T7 sequencing platform using the 2×150 bp paired-end (PE) sequencing mode, in accordance with the manufacturer's instructions. Transcriptomic experiments and preliminary analyses were carried out by Shanghai OE Biotech Co., Ltd.

Cell transfection

pLXSN16E6E7 (#52394, Addgene), pLXSN18E6E7 (#53459, Addgene), psPAX2, and Pmd2.G were packaged into 293T cells using lipo2000 (Thermo fisher, USA). Cells of interest were infected by the virus for 48 h and treated with G418 for selection of the stably transfected cells.

siRNA silencing

Small interference RNAs (siRNAs) were synthesized by RuiMian (Shanghai, China). Cells were transfected with siRNA duplexes (10 nM) using the lipo2000 according to the manufacturer's instructions.

Statistical analysis

All the results are presented as the means ± SDs of triplicate samples from at least three independent experiments. Statistical significance was calculated using a two-tailed Student's t test, or ANOVA where two or more groups were compared. A $p < 0.05$ was considered statistically significant.

Results

E6/E7 expression promotes transformation-associated phenotypes

To investigate the effects of E6/E7 in mammary epithelial cells, we established MCF-10A cells stably expressing these proteins, and western blot confirmed their expression. (Fig. 1A). We evaluated whether E6/E7 expression would lead to transformation-associated phenotypes. E6/E7 expression cells exhibited enhanced proliferation compared with vector (Fig. 1B). In clonogenic assays, E6/E7 expression cells formed markedly increased numbers and larger colonies, indicating enhanced clonogenic potential (Fig. 1C). We also performed 3D tumor spheroid formation assay to assess cell stemness. The number and volume of tumor spheroid produced by cells expressing E6/E7 were higher than vector (Fig. 1D), indicating enhanced self-renewal capacity.

To further validate stemness phenotypes, flow cytometric analysis was conducted to examine the proportion of $CD44^+/CD24^-$ breast stem-like cells. The percentage of $CD44^+/CD24^-$ cells was elevated in E6/E7 expression cells (Fig. 1E). In addition, wound healing assays demonstrated enhanced migratory capacity in HPV18 E6/E7 expression cells, as evidenced by accelerated wound closure (Fig. 1F).

These results indicate that E6/E7 drives malignant transformation of MCF-10A cells by promoting proliferation, stemness, and migration.

E6/E7 induces transcriptomic reprogramming and activate tumor-associated pathways

Based on the phenotypic observations described above, we performed RNA sequencing. Principal component analysis (PCA) showed that the biological replicates in each group were tightly clustered. Meanwhile, E6/E7 expression cells was clearly separated from the vector, suggesting a difference between the two groups at the overall transcriptome level (Fig. 2A). E6/E7 expression triggered widespread transcriptional reprogramming compared to vector (Fig. 2B). Analysis of differentially expressed genes (DEGs) revealed that E6/E7 expression each generated distinct sets of DEGs, while also sharing a subset of common DEGs (Fig. 2C). Consistent with the phenotypic results, these findings suggest that the two HPV exert both shared regulatory effects and distinct molecular characteristics in mammary epithelial cells.

Hierarchical clustering of DEGs further demonstrated clear segregation between E6/E7 expression groups and vector (Fig. 2D). KEGG pathway enrichment analysis showed that the expression of E6/E7 was enriched in multiple pathways related to cell proliferation, cell cycle regulation (top 20) (Fig. 2E),

suggesting that E6/E7 may promote malignant phenotypic changes through the regulation of multiple key molecular pathways.

E6/E7 activates cell cycle programs and promotes TGF-β/SMAD signaling pathways

To further identify the key biological processes associated with E6/E7 mediated transcriptional changes, we performed gene set enrichment analysis (GSEA) on the RNA-seq data. Results showed that gene sets related to the mitotic cell cycle and DNA replication were positively enriched in the E6/E7 expression cells (NES > 2.5, $p < 0.001$, FDR < 0.001) (Fig. 3A), indicating that E6/E7 expression markedly activates proliferation-associated transcriptional programs, particularly those involved in cell cycle progression and DNA replication. We analyzed cell cycle distribution using flow cytometry to validate the transcriptomic findings. Results showed that the proportions of cells in the S phase and G2/M phase were increased in E6/E7 expression cells, while the proportion of cells in the G1 phase was decreased (Fig. 3B), further confirming that E6/E7 promotes cell cycle progression then accelerates cellular proliferation.

To identify potential key regulators, we examined top 10 altered genes within the top 5 enriched pathways. Among the genes associated with cell cycle related pathways, we observed a marked change in the key regulatory factor TGFB2 (Fig. 3C). It may play an important role in the transformation-associated phenotypes induced by E6/E7. We showed that TGFB2 expression was upregulated (Fig. 3D). Given that TGFB2 is an important ligand of the canonical TGF-β signaling pathway, we examined the activation status of this pathway. Western blot analysis revealed that E6/E7 expression increased the phosphorylation level of SMAD3 (pSMAD3) (Fig. 3E), indicating activation of the canonical TGF-β/SMAD signaling pathway.

Collectively, these findings suggest that E6/E7 promotes proliferation of mammary epithelial cells by activating transcriptional programs associated with the cell cycle and DNA replication, accompanied by upregulation of TGFB2 expression and activation of the TGF-β/SMAD signaling pathway.

TGFB2 mediates E6/E7 induced activation of the TGF-β/SMAD–ID1 signaling axis

To further investigate the functional role of TGFB2 in the phenotypic changes mediated by E6/E7, we transfected either negative control siRNA (si-NC) or TGFB2-targeting siRNA (si-TGFB2). qPCR and Western blot showed that after transfection, si-TGFB2 reduced TGFB2 mRNA levels and protein expression (Fig. 4A, B), indicating efficient knockdown. We then examined the effect of TGFB2 silencing on the TGF-β/SMAD signaling pathway. Results showed that si-TGFB2 treatment reduced pSMAD3 compared with si-NC (Fig. 4C), suggesting that TGFB2 acts as a key upstream regulator of E6/E7 induced activation of the TGF-β/SMAD pathway.

TGFB2 knockdown attenuated the growth-promoting effects of E6/E7 expression cells, as evidenced by reduced proliferative capacity and impaired colony formation (Fig. 4D, E). These findings support a critical role for TGFB2 in mediating E6/E7 induced proliferation and clonogenicity. Next, we explored downstream regulators of the TGF-β signaling pathway. We examined the expression changes of downstream effectors of the TGF-β signaling pathway and found that ID1 was upregulated in the E6/E7 expressing group (Fig. 4F). Previous studies have shown that ID1 can promote the tumor initiating capacity of mammary cells in the context of activated TGF-β signaling pathway[28]. Therefore, we hypothesized that E6/E7 may promote malignant transformation-associated phenotypes by activating the TGF-β signaling pathway and subsequently upregulating the downstream effector ID1 (Fig. 4G).

ID1 acts as downstream effector of TGF-β signaling pathway

We examined the expression of ID1, a canonical downstream target of TGF-β signaling. qPCR and Western blot analyses showed that E6/E7 expression increased both the mRNA and protein levels of ID1 (Fig. 5A, B). To determine the role of ID1 in tumor initiating capacity, we examined ID1 expression in cells transfected with siTGFB2. TGFB2 knockdown reduced ID1 protein levels (Fig. 5C), suggesting that ID1 functions downstream of the TGFB2–TGF-β signaling pathway. Subsequently, E6/E7 expression cells were transfected with si-NC or ID1-targeting siRNA (si-ID1). Results showed that si-ID1 reduced ID1 mRNA levels and protein expression (Fig. 5D, E).

Silencing ID1 reduced the proliferative capacity of E6/E7 expression cells, while colony formation remained largely unaffected (Fig. 5F, G). These findings suggest that ID1 primarily contributes to the proliferative effects of E6/E7, whereas clonogenic potential may be preserved by compensatory factors such as ID3. This is in line with previous reports showing that ID1 and ID3 cooperatively maintain tumor initiating capacity and stemness in breast cancer. Furthermore, flow cytometric analysis of cell cycle distribution showed that si-ID1 treatment decreased the proportion of cells in the S phase while increasing the proportions of cells in the G1 and G2/M phases (Fig. 5H). It indicated that ID1 knockdown leads to cell cycle arrest.

Collectively, these results demonstrate that ID1 acts as a key downstream effector of the TGFB2–TGF-β/SMAD signaling pathway and plays an important role in E6/E7 induced cell cycle progression and tumor initiation.

Discussion

In this study, we revealed the mechanism by which HPV oncogenes drive malignant transformation of breast epithelial cells. We demonstrate that E6/E7 expression induces a coordinated transcriptional reprogramming event characterized by activation of cell cycle and DNA replication pathways, enhanced proliferative capacity, increased clonogenicity, and expansion of tumor initiating cell (TIC) like populations. Mechanistically, these phenotypes are mediated through upregulation of TGFB2, activation of canonical TGF-β/SMAD signaling pathway, and subsequent induction of the downstream effector ID1,

establishing a functional TGFB2–TGF-β–ID1 signaling axis that promotes early oncogenic transformation.

These oncogenic activities have classically been attributed to E6-mediated degradation of p53 and E7-driven inactivation of pRb, thereby disrupting cell cycle checkpoints[11,12]. While these mechanisms explain loss of growth control, they do not fully account for the acquisition of stem-like traits and tumor initiating capacity. Our findings demonstrate that HPV oncogenes can actively reshape the host signaling network, thereby regulating cell fate determination and proliferation processes.

TGF-β signaling pathway is widely recognized as context dependent, functioning as a tumor suppressor in normal epithelium, but often acquiring pro-tumorigenic properties during cancer progression[38]. Our data indicate that E6/E7 can redirect TGF-β signaling output during the initial stages of malignant tumors. E6/E7 expression markedly upregulated TGFB2 and enhanced SMAD3 phosphorylation, indicating activation of canonical TGF-β signaling pathway. TGFB2 silencing attenuated SMAD activation and significantly impaired proliferation and clonogenic growth, establishing TGFB2 as a critical upstream mediator of E6/E7-driven phenotypes. These findings support a model in which E6/E7 does not merely activate TGF-β signaling, but functionally repurposes it to support oncogenic transformation.

ID1 emerged as a key downstream effector of this axis[37]. ID1 induction was dependent on TGFB2 signaling, and its depletion resulted in impaired cell cycle progression and reduced proliferative capacity. Notably, ID1 suppression did not completely abolish clonogenicity, suggesting potential functional compensation by other ID family members such as ID3[39]. Nonetheless, our data position ID1 as a central mediator linking TGF-β signaling pathway to E6/E7 induced proliferative reprogramming. Given that ID1 has been implicated in maintenance of stemness independently of EMT, activation of the TGFB2–ID1 axis may represent a mechanism by which E6/E7 enhances tumor initiating capacity during early mammary epithelial transformation.

Several limitations should be acknowledged. The present study is primarily based on in vitro transformation models, and in vivo validation will be necessary to establish tumorigenic capacity in physiological settings. The cooperative contribution of other ID family members warrants further investigation. Given the ongoing debate regarding the prevalence and etiological significance of HPV in breast cancer, validation in clinically annotated HPV-positive breast tumor samples will be essential to determine the translational relevance of this signaling axis.

In summary, our study identifies TGFB2–TGF-β–ID1 signaling as a critical downstream pathway engaged by E6E7 in mammary epithelial cells. By functionally reprogramming TGF-β signaling toward a pro-proliferative and stemness supportive output, E6/E7 promote early malignant transformation. These findings provide mechanistic insight into virus-associated mammary tumorigenesis and suggest that components of the TGFB2–TGF-β–ID1 axis may represent potential therapeutic vulnerabilities in HPV drive breast malignancy.

Declarations

Conflict of Interest

The authors declare no competing interests.

Ethics Statement

Not involved.

Author Contribution

X.X., L.Z., and L.F. designed the study. X.X., W.M., Z.S., M.Z. W.W., L.Z. and L.F. collected the data. X.X. wrote the manuscript., L.Z. and L.F. revised the manuscript. All authors have reviewed the manuscript and approved the final version.

Acknowledgement

This work was supported by the National Natural Science Foundation of China (grants 82173344).

Data and Code Availability

The data underlying this article will be shared on reasonable request to the corresponding author.

References

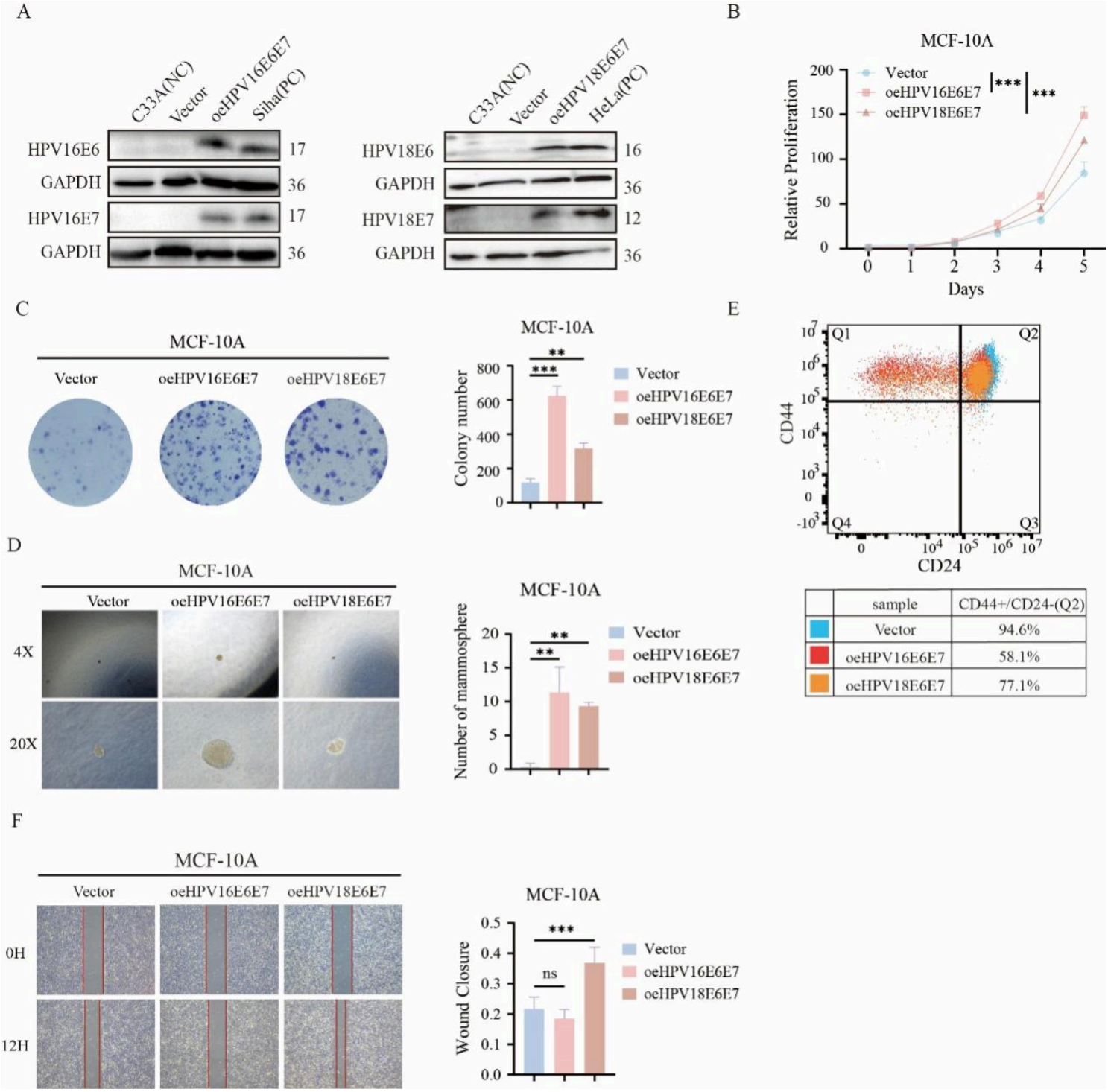
1. Siegel, R. L., Miller, K. D., Wagle, N. S. & Jemal, A. Cancer statistics, 2023. *CA Cancer J Clin* **73**, 17-48, doi:10.3322/caac.21763 (2023).
2. Soerjomataram, I. & Bray, F. Planning for tomorrow: global cancer incidence and the role of prevention 2020-2070. *Nat Rev Clin Oncol* **18**, 663-672, doi:10.1038/s41571-021-00514-z (2021).
3. Bray, F. *et al.* Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Ca-a Cancer Journal for Clinicians* **74**, 229-263, doi:10.3322/caac.21834 (2024).
4. Xiao, Y. *et al.* Comprehensive metabolomics expands precision medicine for triple-negative breast cancer. *Cell Research* **32**, 477-490, doi:10.1038/s41422-022-00614-0 (2022).
5. Wang, Y. *et al.* MMTV-like env gene sequences in human breast cancer. *Arch Virol* **146**, 171-180, doi:10.1007/s007050170201 (2001).

6. Gannon, O. M., Antonsson, A., Bennett, I. C. & Saunders, N. A. Viral infections and breast cancer - A current perspective. *Cancer Lett* **420**, 182-189, doi:10.1016/j.canlet.2018.01.076 (2018).
7. Egawa, N. & Doorbar, J. The low-risk papillomaviruses. *Virus Research* **231**, 119-127, doi:10.1016/j.virusres.2016.12.017 (2017).
8. Mlynarczyk-Bonikowska, B. & Rudnicka, L. HPV Infections-Classification, Pathogenesis, and Potential New Therapies. *Int J Mol Sci* **25**, doi:10.3390/ijms25147616 (2024).
9. Carifi, M., Napolitano, D., Morandi, M. & Dall'Olio, D. Recurrent respiratory papillomatosis: current and future perspectives. *Therapeutics and Clinical Risk Management* **11**, 731-738, doi:10.2147/tcrm.S81825 (2015).
10. Saito, Y. *et al.* Human papillomavirus-related oropharyngeal carcinoma. *Japanese Journal of Clinical Oncology* **52**, 692-698, doi:10.1093/jjco/hyac049 (2022).
11. Münger, K. *et al.* Complex formation of human papillomavirus E7 proteins with the retinoblastoma tumor suppressor gene product. *Embo j* **8**, 4099-4105, doi:10.1002/j.1460-2075.1989.tb08594.x (1989).
12. 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, doi:10.1016/0092-8674(90)90409-8 (1990).
13. Martinez, I. *et al.* Human papillomavirus type 16 reduces the expression of microRNA-218 in cervical carcinoma cells. *Oncogene* **27**, 2575-2582, doi:10.1038/sj.onc.1210919 (2008).
14. Au Yeung, C. L., Tsang, T. Y., Yau, P. L. & Kwok, T. T. Human papillomavirus type 16 E6 induces cervical cancer cell migration through the p53/microRNA-23b/urokinase-type plasminogen activator pathway. *Oncogene* **30**, 2401-2410, doi:10.1038/onc.2010.613 (2011).
15. Wang, X. *et al.* Oncogenic HPV infection interrupts the expression of tumor-suppressive miR-34a through viral oncoprotein E6. *Rna* **15**, 637-647, doi:10.1261/rna.1442309 (2009).
16. Melar-New, M. & Laimins, L. A. Human papillomaviruses modulate expression of microRNA 203 upon epithelial differentiation to control levels of p63 proteins. *J Virol* **84**, 5212-5221, doi:10.1128/jvi.00078-10 (2010).
17. Karachalios, C., Petousis, S., Margioula-Siarkou, C. & Dinas, K. Human papillomaviruses and breast cancer: A systematic review and meta-analysis. *Oncol Lett* **27**, 75, doi:10.3892/ol.2023.14208 (2024).
18. Corbex, M. *et al.* Prevalence of papillomaviruses, polyomaviruses, and herpesviruses in triple-negative and inflammatory breast tumors from algeria compared with other types of breast cancer tumors. *PLoS One* **9**, e114559, doi:10.1371/journal.pone.0114559 (2014).
19. Haghshenas, M. R., Mousavi, T., Moosazadeh, M. & Afshari, M. Human papillomavirus and breast cancer in Iran: a meta- analysis. *Iran J Basic Med Sci* **19**, 231-237 (2016).
20. Awan, U. A. *et al.* An updated systemic review and meta-analysis on human papillomavirus in breast carcinogenesis. *Front Oncol* **13**, 1219161, doi:10.3389/fonc.2023.1219161 (2023).

21. Ren, C. *et al.* Human papillomavirus infection increases the risk of breast carcinoma: a large-scale systemic review and meta-analysis of case-control studies. *Gland Surg* **8**, 486-500, doi:10.21037/gs.2019.09.04 (2019).
22. De Carolis, S. *et al.* HPV DNA Associates With Breast Cancer Malignancy and It Is Transferred to Breast Cancer Stromal Cells by Extracellular Vesicles. *Front Oncol* **9**, 860, doi:10.3389/fonc.2019.00860 (2019).
23. Piana, A. F. *et al.* HPV infection and triple-negative breast cancers: an Italian case-control study. *Virol J* **11**, 190, doi:10.1186/s12985-014-0190-3 (2014).
24. Ruzinova, M. B. & Benezra, R. Id proteins in development, cell cycle and cancer. *Trends Cell Biol* **13**, 410-418, doi:10.1016/s0962-8924(03)00147-8 (2003).
25. Anido, J. *et al.* TGF-β Receptor Inhibitors Target the CD44(high)/Id1(high) Glioma-Initiating Cell Population in Human Glioblastoma. *Cancer Cell* **18**, 655-668, doi:10.1016/j.ccr.2010.10.023 (2010).
26. O'Brien, C. A. *et al.* ID1 and ID3 regulate the self-renewal capacity of human colon cancer-initiating cells through p21. *Cancer Cell* **21**, 777-792, doi:10.1016/j.ccr.2012.04.036 (2012).
27. Zhang, J. *et al.* Targeted inhibition of KDM6 histone demethylases eradicates tumor-initiating cells via enhancer reprogramming in colorectal cancer. *Theranostics* **10**, 10016-10030, doi:10.7150/thno.47081 (2020).
28. Stankic, M. *et al.* TGF-β-Id1 signaling opposes Twist1 and promotes metastatic colonization via a mesenchymal-to-epithelial transition. *Cell Rep* **5**, 1228-1242, doi:10.1016/j.celrep.2013.11.014 (2013).
29. Jiang, W. J. *et al.* Breast Cancer Subtype-Specific Organotropism Is Dictated by FOXF2-Regulated Metastatic Dormancy and Recovery. *Cancer Res* **85**, 644-659, doi:10.1158/0008-5472.Can-24-0479 (2025).
30. Massagué, J. TGFbeta in Cancer. *Cell* **134**, 215-230, doi:10.1016/j.cell.2008.07.001 (2008).
31. David, C. J. & Massagué, J. Contextual determinants of TGFβ action in development, immunity and cancer. *Nat Rev Mol Cell Biol* **19**, 419-435, doi:10.1038/s41580-018-0007-0 (2018).
32. Scheel, C. *et al.* Paracrine and autocrine signals induce and maintain mesenchymal and stem cell states in the breast. *Cell* **145**, 926-940, doi:10.1016/j.cell.2011.04.029 (2011).
33. Pickup, M., Novitskiy, S. & Moses, H. L. The roles of TGFβ in the tumour microenvironment. *Nat Rev Cancer* **13**, 788-799, doi:10.1038/nrc3603 (2013).
34. Wakefield, L. M. & Hill, C. S. Beyond TGFβ: roles of other TGFβ superfamily members in cancer. *Nat Rev Cancer* **13**, 328-341, doi:10.1038/nrc3500 (2013).
35. Tauriello, D. V. F. & Batlle, E. Targeting the Microenvironment in Advanced Colorectal Cancer. *Trends Cancer* **2**, 495-504, doi:10.1016/j.trecan.2016.08.001 (2016).
36. Liu, K. *et al.* Stabilization of TGF-β Receptor 1 by a Receptor-Associated Adaptor Dictates Feedback Activation of the TGF-β Signaling Pathway to Maintain Liver Cancer Stemness and Drug Resistance. *Adv Sci (Weinh)* **11**, e2402327, doi:10.1002/advs.202402327 (2024).

37. Ullmann, P. *et al.* The miR-371∼373 Cluster Represses Colon Cancer Initiation and Metastatic Colonization by Inhibiting the TGFBR2/ID1 Signaling Axis. *Cancer Res* **78**, 3793-3808, doi:10.1158/0008-5472.Can-17-3003 (2018).
38. Colak, S. & Ten Dijke, P. Targeting TGF-β Signaling in Cancer. *Trends Cancer* **3**, 56-71, doi:10.1016/j.trecan.2016.11.008 (2017).
39. Huang, L. *et al.* ID3 Promotes Stem Cell Features and Predicts Chemotherapeutic Response of Intrahepatic Cholangiocarcinoma. *Hepatology* **69**, 1995-2012, doi:10.1002/hep.30404 (2019).

Figures



	sample	CD44+/CD24-(Q2)
	Vector	94.6%
	oeHPV16E6E7	58.1%
	oeHPV18E6E7	77.1%

Figure 1

Expression of E6/E7 promotes transformation-associated phenotypes of MCF-10A cells

(A) Western blot analysis confirming the expression of E6/E7 proteins in MCF-10A cells. (B) Cell proliferation of E6/E7 expression cells was assessed using the CCK-8 assay. (C) Clonogenic assays were performed to evaluate the clonogenic potential of cells expressing E6/E7. (D) 3D tumor spheroid formation assays were conducted to assess the stemness of E6/E7 expression cells. (E) Flow cytometry

was used to detect the levels of breast stem-like cell markers. (F) Wound healing assays were performed to evaluate the migratory capacity of E6/E7 expression cells.

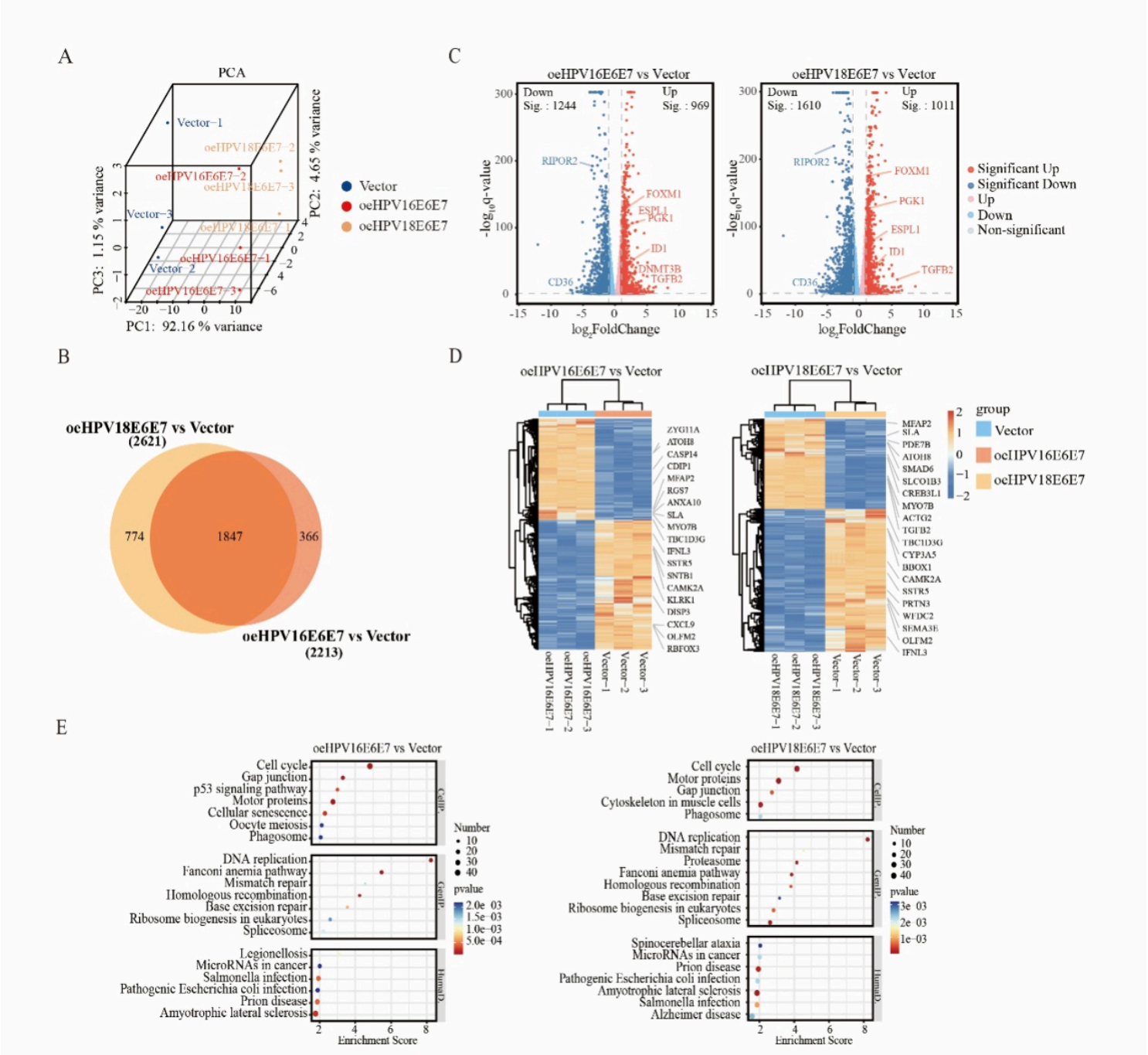


Figure 2

Differentially expressed genes are enriched in cell cycle and DNA replication pathways

(A) Principal component analysis (PCA) of RNA-seq data showing the overall transcriptomic differences between E6/E7 expression cells and vector, as well as the consistency among biological replicates. (B) Volcano plot displaying the statistical significance ($p < 0.05$) and relative expression changes (log_2/fold change/ > 1) of differentially expressed genes. (C) Venn diagram showing the number of differentially

expressed genes and the unique and shared DEGs. (D) Hierarchical clustering heatmap of differentially expressed genes illustrating the clustering patterns. (E) KEGG pathway enrichment analysis showing the top 20 upregulated pathways.

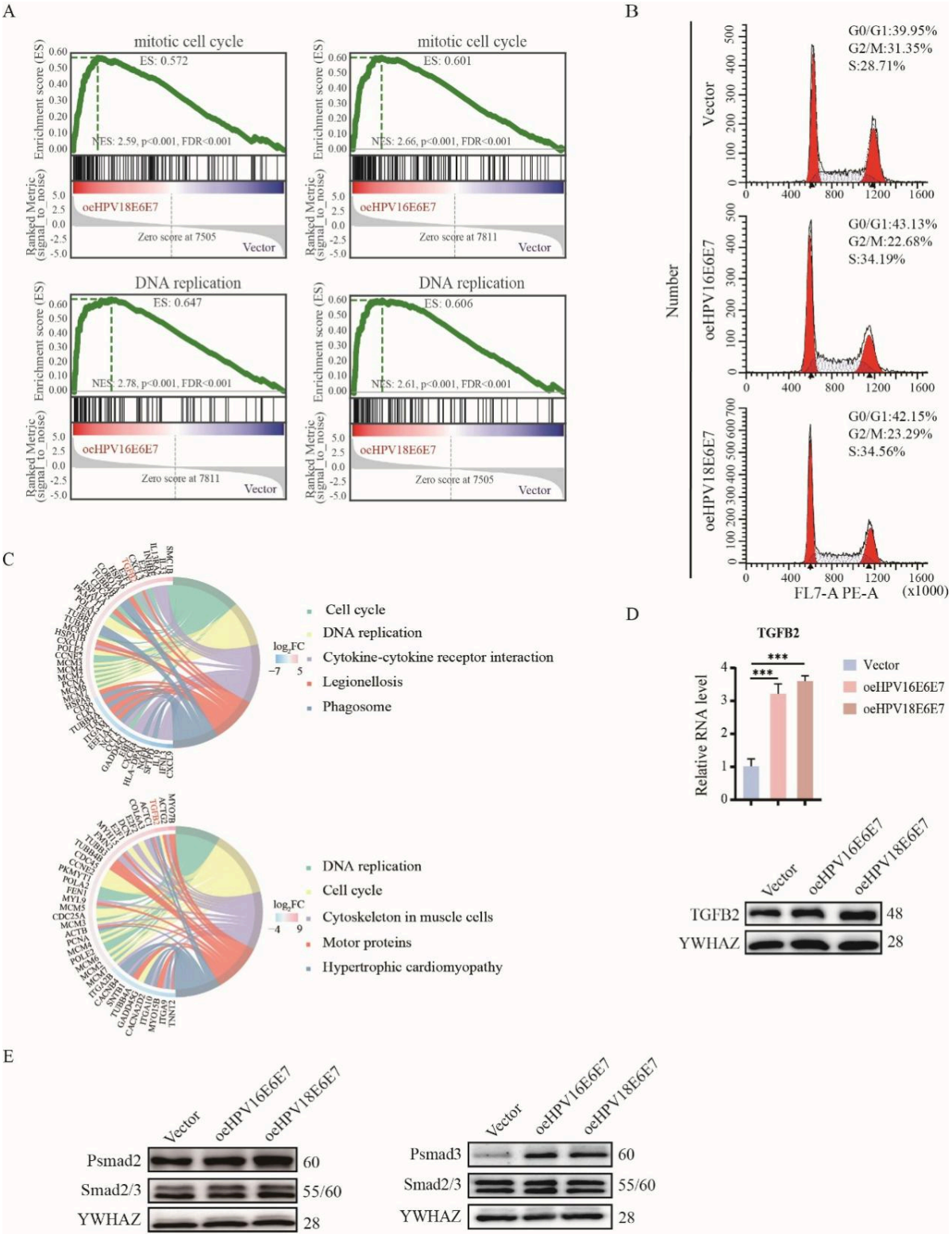


Figure 3

Activation of the TGF-β/SMAD signaling pathway

(A) GSEA enrichment plots showing that gene sets related to the mitotic cell cycle and DNA replication are significantly positively enriched in E6/E7 expression cells (NES > 2.5, $p < 0.001$, FDR < 0.001), indicating activation of proliferation associated transcriptional programs. (B) Cell cycle distribution by flow cytometry. (C) Chord diagram illustrating the top five significantly altered pathways along with the top ten significantly changed genes within each pathway. The TGFB2 gene in the cell cycle related pathway is highlighted. (D) qPCR and Western blot analyses of TGFB2 expression. (E) Western blot analysis of the activation status of the canonical TGF-β/SMAD signaling pathway (pSMAD3).

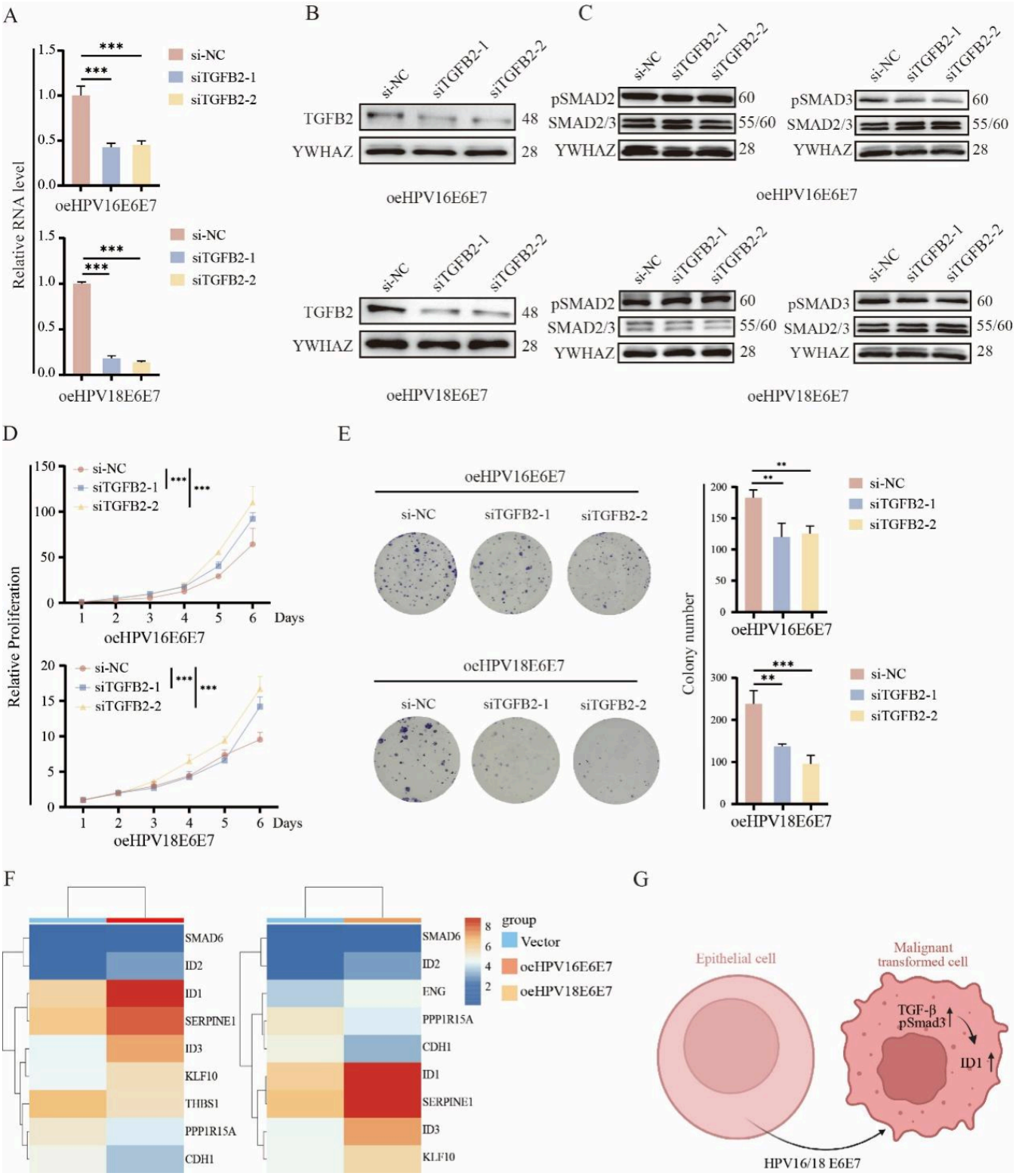


Figure 4

Silencing TGFB2 suppresses the proliferative capacity of transformed cells

(A) E6/E7 expression cells were transfected with negative control siRNA (si-NC) or TGFB2-targeting siRNA (si-TGFB2), and TGFB2 mRNA levels were analyzed by qPCR 48 hours after transfection. (B) Protein expression levels were examined by Western blot analysis 72 hours after transfection. (C) The effect of si-TGFB2 on the TGF-β/SMAD signaling pathway in E6/E7 expression cells was analyzed. (D) Cell proliferation of E6/E7 expression cells transfected with si-NC or si-TGFB2 was assessed using the CCK-8 assay. (E) Clonogenic assays were performed to evaluate colony formation. Si-TGFB2 reduced both the number and size of colonies. (F) Heatmap of TGF-β downstream effector expression. (G) Schematic diagram of epithelial cells expressing E6/E7 proteins activating intracellular TGF-β signaling pathway, thereby increasing the expression of downstream effector ID1 and promoting malignant transformation of cells.

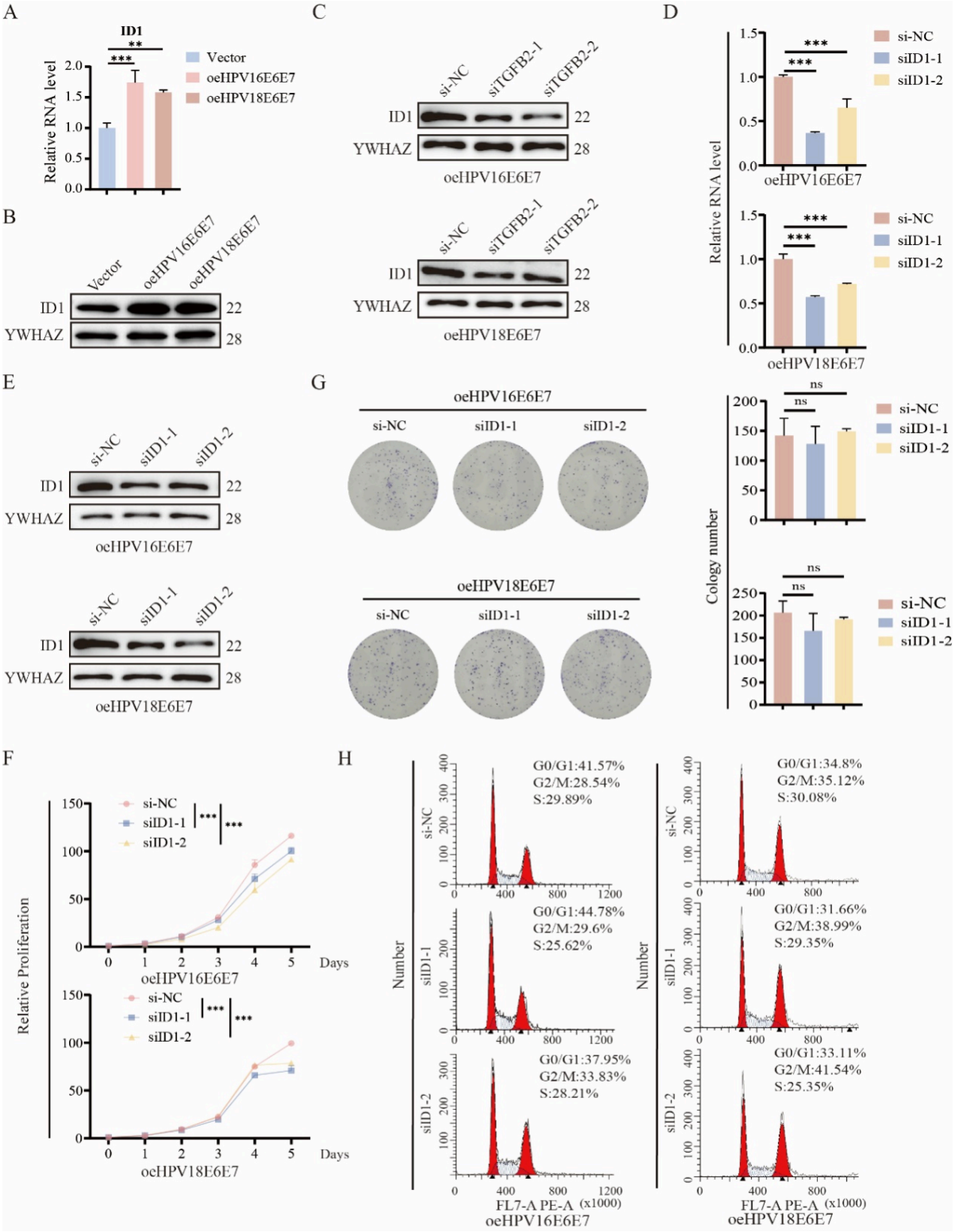


Figure 5

Silencing ID1 induces cell cycle arrest in transformed cells

(A) qPCR and (B)Western blot analyses were performed to examine ID1 expression levels in vector and E6/E7 expression cells. (C) ID1 protein expression was analyzed in which E6/E7 expression cells were transfected with si-NC or si-TGFB2. (D) E6/E7 expression cells were transfected with si-NC or ID1-

targeting siRNA (si-ID1), and ID1 mRNA levels were measured by qPCR 48 hours after transfection. (E) ID1 protein expression was further examined by Western blot analysis 72 hours after transfection. (F) Cell proliferation of E6/E7 expression cells transfected with si-NC or si-ID1 was assessed using the CCK-8 assay. (G) Clonogenic assays were performed to evaluate colony formation transfected with si-NC or si-ID1. (H) Flow cytometry was used to determine the cell cycle distribution of transfected with si-NC or si-ID1.