Journal of Advanced Research 57 (2024) 163–180

Contents lists available at ScienceDirect

Journal of Advanced Research

journal homepage: www.elsevier.com/locate/jare

Original Article

PIBF1 regulates multiple gene expression via impeding long-range chromatin interaction to drive the malignant transformation of HPV16 integration epithelial cells

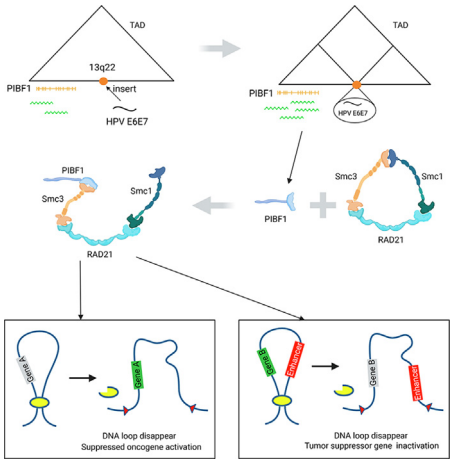
Xiaomin Li [a,b,1], Ci Ren [a,1], Anni Huang [a,1], Yue Zhao [c,1], Liming Wang [a], Hui Shen [a], Chun Gao [a], Bingxin Chen [a], Tong Zhu [a], Jinfeng Xiong [a], Da Zhu [a], Yafei Huang [d], Jianlin Ding [a], Zan Yuan [c,*], Wencheng Ding [a,*], Hui Wang [a,e,*]

[a] National Clinical Research Center for Gynecology and Obstetrics, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, China
[b] Department of Anesthesiology, Hubei Key Laboratory of Geriatric Anesthesia and Perioperative Brain Health, and Wuhan Clinical Research Center for Geriatric Anesthesia, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, China
[c] Annoroad Gene Technology (Beijing) Co., Ltd, Beijing 100176, China
[d] Department of Pathogen Biology, School of Basic Medicine, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China
[e] Zhejiang Provincial Key Laboratory of Precision Diagnosis and Therapy for Major Gynecological Diseases, Women's Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China

HIGHLIGHTS

- A site-specific HPV16 gene knock-in cell model was established.
- PIBF1 was up-regulated as *cis*-perturbation of HPV integration.
- PIBF1 is a chromatin binding protein and could interact with cohesin complex off the chromatin.
- PIBF1 may cause dysregulation of tumorigenesis-related genes through destroying cohesin ring-shaped structure.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 5 December 2022
Revised 31 March 2023
Accepted 19 April 2023
Available online 13 May 2023

Keywords:
HPV16 integration

ABSTRACT

Introduction: Human papillomavirus (HPV) integration can induce gene expression dysregulation by destroying higher-order chromatin structure in cervical cancer.
Objectives: We established a 13q22 site-specific HPV16 gene knock-in cell model to interrogate the changes in chromatin structure at the initial stages of host cell malignant transformation.
Methods: We designed a CRISPR-Cas9 system with sgRNA targeting 13q22 site and constructed the HPV16 gene donor. Cells were cotransfected, screened, and fluorescence sorted. The whole genome sequencing (WGS) was used to confirm the precise HPV16 gene integration site. Western blot and qRT-

Peer review under responsibility of Cairo University.
* Corresponding authors.
E-mail addresses: yfindderam@gmail.com (Z. Yuan), dingwencheng326@163.com (W. Ding), huit71@sohu.com (H. Wang).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.jare.2023.04.015

Malignant transformation
PIBF1
Chromatin structure

PCR were used to measure gene expression. In vitro and in vivo analysis were performed to estimate the tumorigenic potential of the HPV16 knock-in cell model. Combined Hi-C, chromatin immunoprecipitation and RNA sequencing analyses revealed correlations between chromatin structure and gene expression. We performed a coimmunoprecipitation assay with anti-PIBF1 antibody to identify endogenous interacting proteins. In vivo analysis was used to determine the role of PIBF1 in the tumor growth of cervical cancer cells.

Results: We successfully established a 13q22 site-specific HPV16 gene knock-in cell model. We found that HPV integration promoted cell proliferation, invasion and stratified growth in vitro, and monoclonal proliferation in vivo. HPV integration divided the affected topologically associated domain (TAD) into two smaller domains, and the progesterone-induced blocking factor 1 (PIBF1) gene near the integration site was upregulated, although PIBF1 was not enriched at the domain boundary by CUT-Tag signal analysis. Moreover, PIBF1 was found to interact with the cohesin complex off chromatin to reduce contact domain formation by disrupting the cohesin ring-shaped structure, causing dysregulation of tumorigenesis-related genes. Xenograft experiments determined the role of PIBF1 in the proliferation in cervical cancer cells.

Conclusion: We highlight that PIBF1, a potential chromatin structure regulatory protein, is activated by HPV integration, which provides new insights into HPV integration-driven cervical carcinogenesis.

Introduction

Cervical cancer is one of the most common gynecological malignancies, with 604,127 diagnoses and over 341,000 deaths worldwide in 2020[1]. The integration of high-risk HPV (hrHPV) DNA fragments into the host genome is considered a key event in the progression of major cervical lesions [2]. According to previous studies, integration of HPV genes was found in up to 76–94 % of HPV16-positive squamous cervical cancers [3-5]. Additional evidence of nonrandom integration of HPV in the human genome has been found. Many high-frequency integration loci have been reported, such as 3q28, 8q24.21, and 13q22.1, among others [5,6]. Previous studies have focused on hrHPV virus oncogene expression, which can deregulate host gene expression [7]. The integration of HPV can also cause structural variations (SVs) in the host genome due to the insertion of new genetic material into the chromatin. It has been systematically reported that SVs can alter regulatory elements by disrupting higher-order chromatin structures, such as TADs and loops, and could affect gene expression far from the SV breakpoint, which leads to the progression of cancer or other diseases [8].

To date, several studies have characterized the SVs and the effects on higher-order chromatin structures and gene expression driven by HPV integration in cervical cancer [9-11]. These studies have found that integration is often accompanied by instability of the host genome, recurrent focal amplification, rearrangement adjacent to the integrated locus, and even dysregulation of distant gene expression. However, the data reported in these studies were based on clinical cervical cancer specimens or naturally integrated cervical cancer cells, all of which were in a post integration state with massive structural variations in the host genome [9-12]. Therefore, it is difficult to assess the dynamic changes in pathogenesis driven by HPV integration from the premalignancy stage to the carcinoma stage on the basis of these samples.

Although there are several types of HPV integration events that can occur in the cells of polyclonal precancerous lesions, clones with integration at hotspots are often identified as the main clone subtype of cervical cancers [13]. However, the mechanism of the host hotspots site of integration and local genes in determining the progression of cancer remains to be explored. A rigorously designed cell model of HPV integration is urgently needed to facilitate the investigation of the initial stages of phenotypic progression to cervical cancer. Groves et al. chose to study early generations of W12-derived cell lines to minimize any effects of genomic instability caused by deregulated HPV16 oncogene expression, but the W12 cell context itself had some effect on the integration because this line was derived from a clinical low-grade squamous intraepithelial lesion [10]. Therefore, a cell model that maintains the restricted cell differentiation and phenotypic characteristics of the epithelial cell layer is urgently needed.

In this study, we have successfully established a 13q22 site-specific HPV16 knock-in cellular model. We first visually observed the impact of the single integration at this site, without the influence of the other integration sites, on genome structure. We found that PIBF1 may act as a potential chromatin structure regulatory protein. HPV integration could upregulate PIBF1 expression to accelerate the disruption of the cohesin complex and thereby reduce long-range *cis* interactions and disrupt gene expression. The data reported here help us to obtain deep insight into the mechanism of how HPV integration affects the regulation of spatial chromosome structure.

Materials and methods

Plasmid construction and preparation

All sgRNAs were designed based on the hg19 genome using a website tool (https://tools.genome-engineering.org). The gene sequences around the HPV integration site were selected, and we selected the three sgRNAs with the highest score and fewest off-target matches. They were cloned based on the protocol from the Feng Zhang lab. Briefly, oligonucleotides complementary to gRNA templates were ordered separately, annealed, phosphorylated and cloned into the BbsI sites of pSpCas9(BB)-2A-EGFP (Addgene plasmid, Cambridge, MA, USA). The sgRNAs sequences are listed in Fig. S1A.

To construct donor vectors, left and right homology arms (500 bp each) were amplified from HaCaT genomic DNA by PCR and cloned into pDC515, generating pDC515-HAL-HAR (D0). Then, the flag-Loxp + CMV-EGFP-2A-Puro-polyA + loxp cassette was generated and inserted downstream of the left arm of D0, yielding pDC515-HAL-flag-Loxp + CMV-EGFP-2A-Puro-polyA + loxp-HAR (D1), which served as the control donor vector. To obtain HPV16 knock-in donor vectors, the HPV16 URR + E6 + E7 gene was cloned from HPV16 BAC and then reversed and inserted upstream of the EGFP cassette of D1, generating pDC515-HAL-HPV16 E7 + E6 + URR-flag-Loxp + CMV-EGFP-2A-Puro-polyA + loxp-HAR (D2). All plasmids were prepared using endotoxin-free plasmid extraction kit (D6915, Omega, USA,) and stored at −80 °C. All constructs were confirmed by sequencing.

For cell transfection of the donor, linearized D1 (HAL-flag-Loxp + CMV-EGFP-2A-Puro-polyA + loxp-HAR) and D2 (HAL-HPV16 E7 + E6 + URR-flag-Loxp + CMV-EGFP-2A-Puro-polyA + loxp-HAR) fragments were amplified by PCR, using primer 1 (Table S3). All the PCR product was checked by agarose gel and purified by a Cycle Pure Kit (D6492, Omega, USA).

Cell culture and transfection

HaCaT cells (derived from a transformed nontumorigenic human keratinocyte cell line), Human papillomavirus-related cervical cancer cell lines (SiHa) were purchased from ATCC and cultured in fresh Dulbecco's minimum essential medium (DMEM; Invitrogen) supplemented with 10 % fetal bovine serum (FBS, Gibco) and 100 U/mL penicillin streptomycin (Invitrogen). Ect1/E6E7 (immortalized human ectocervical squamous cell line) were purchased from ATCC and cultured in Keratinocyte-Serum Free medium supplemented with human recombinant EGF (0.1 ng/ml), bovine pituitary extract (0.05 mg/ml), and calcium chloride (final concentration 0.4 mM). Cells were grown in a humidified incubator at 37 °C, in a 5 % carbon dioxide atmosphere.

All cells were transfected using linear polyethylenimine (PEI) (Sigma; 764604). PEI powder was diluted with sterilized ddH_2O to 5 μg/μl, and the ratio of PEI to total DNA was 3 μg:1 μg. Each well of a six-well plate received 1 μg sgRNA and 1 μg linearized donor. Before transfection, the cell medium was replaced with 1.8 ml of DMEM, PEI was added to the DNA mixture suspended in 200 μl of Opti-MEM, and the cells were incubated for 15 min at room temperature. Then, the mixture was added to the cultured cells and incubated the cells in a CO2 incubator. After 24 h, the medium was replaced daily with 2 ml of fresh DMEM containing puromycin (50–100 nM) for 3 weeks; the puromycin concentration was decreased gradually with the increase in EGFP positive cells under a fluorescence microscope (Olympus, Tokyo, Japan).

sgRNA targeting analyses

HaCaT cells plated in 6-well plates were transfected with 2 μg of the indicated sgRNAs (separately) for 48 h. Genomic DNA was extracted, and approximately 500 bp of the genomic fragment encompassing the sgRNA target sites was amplified with primer 5 (listed in Table S3). A 200 ng sample of each purified PCR product was denatured and reannealed in 1 × NEB Buffer 2 T7EI (New England Biolabs) for 15 min at 37 °C. The reaction mixtures were then analyzed on a 10 % TBE polyacrylamide gel. The products were also purified using a gel extraction kit to prepare for single cloning into the PMD18-T plasmid vector, and the exact mutant sequences was screened by Sanger sequencing.

Single-cell cloning and analysis

After selection with puromycin for 5 days, cells were digested by trypsin, and EGFP-positive cells were immediately sorted into 96-well plates using BD FACS under sterile conditions and cultured at 37 °C with 5 % CO2. Each well contained 100 μl fresh DMEM supplemented with puromycin (25 mM) and refreshed weekly. After 3 weeks, EGFP-positive cell colonies were digested and transferred into 24-well plates. Then, a portion of the cells was harvested for DNA extraction using the Tissue DNA Kit (QIAGEN, D3396), and every colony was genotyped by PCR (Table S1) and and Sanger sequencing.

Polymerase chain reaction (PCR)

PCR was performed by Taq DNA polymerase (Vazyme; P131) with the following settings: 3 min at 95 °C for pre-incubation, followed by 34 cycles of 15 s at 95 °C, 30 s at 60 °C, and extension for 60 sec/kb depending on the expected product size at 72 °C.

Flow cytometry

To evaluate the percentages of EGFP-positive cells, the cells was suspended in PBS and fixed with 2% paraformaldehyde for 10 min, then subjected to flow cytometry analysis using a FACSAria II cell sorter (BD Biosciences).

Quantitative reverse transcription PCR (RT–PCR)

Total RNA was isolated using the Total RNA Kit I (OMEGA; R6834). RT-PCR was performed using 2 μg of total RNA. Quantitative RT-PCR was performed in a CFX96 Real-Time system (Bio–Rad, USA). The experiments were performed in triplicate, and the primers of all target genes and GAPDH (reference gene) are listed in Table S4.

Colony formation assay

A total of 500 cells were seeded in 12-well plates in triplicate and incubated for 2 weeks. The colonies were stained with 0.04% crystal violet and photographed. Colonies of>50 mm in diameter were counted using an Omnicon 3600 image analysis system.

Cell proliferation assay

Cell proliferation were performed using Cell Counting Kit-8 (Dojindo Molecular Technologies) according to the manufacturer's manual. Briefly, 5 × 10^3 cells were seeded in 96-well plates and cultured for 0 h, 24 h, 48 h, 72 h and 96 h in a 37°C incubator with 5% CO_2. Then CCK-8 dye was incubated at 37°C for 3 h, the absorbance was determined at 450 nm using a microplate reader.

Transwell migration and invasion assay

Cells were diluted with serum-free DMEM to a concentration of 5x10^5 cells/ml, and 100 μl cells were placed on Transwell chambers (Corning Life Sciences). For the cell invasion assay, chambers were precoated with 50 μl Matrigel (BD Biosciences) solution diluted in DMEM (vol/vol 1:8), and cells were added to the upper compartment and incubated for 24 and 48 h at 37°C. DMEM supplemented with 20% FBS was prepared in the lower compartment. Then cells were scored by counting the number of stained cells on the whole filter under a microscope.

Wound healing assay

The cells were seeded in six-well culture plates. When the cell density reached 100%, a wound was made by using a 200 μl pipette tip to scratch the monolayer. The cells was washed with 1x PBS for 3 times to remove cell debris and then incubated in fresh serum-free DMEM. The area of the scratch was photographed at 0 h, 6 h and 12 h, and the distance from the scratch was measured at different time points. Finally, the wound healing effect was analyzed for each group of cells with the following calculation: 100 × (scratch width at 0 h - scratch width at 12 h)/ scratch width at 0 h. Three replicates were performed for each group of experiments.

3D organotypic coculture

We performed the three-dimensional organotypic coculture as described previously [14]. Briefly, a polyethylene terephthalate

hanging cell culture insert (Millipore Merck Germany; MCEP24H48) in a 24-well plate was coated with 50 μl Matrigel. After the Matrigel was polymerized in a 37 °C incubator, 5000 fibroblasts diluted in 100 μl Matrigel mix (50 μl Matrigel, 20 μl FBS, 10 μl 10 × DMEM, 10 μl 1 N NaOH, 10 μl rat tail collagen, type 1) were added to the bottom layer. Then, $1x10^6$ HaCaT cells were added to the top layer for 24 h after fibroblast implantation. When cell was implanted for 10 days, we harvested the organotypic gels and embedded them in paraffin. Hematoxylin and eosin (HE) staining was performed on paraffin sections, and images were taken using an Olympus BX53 microscope (Olympus).

Lentivirus preparation and transduction

The PIBF1 knockdown lentivirus were constructed by Shanghai Genechem Co., Ltd. For the PIBF1 knockdown lentivirus (lv-shPIBF1), the siRNA sequence of PIBF1, ATGAGCTAGTGAATCCATTAA or GCAAGGTGACTACCGTCAAGA, was cloned into the hU6-MCS-CBh-gcGFP-IRES-puromycin vector. TTCTCCGAACGTGTCACGT was cloned into the same lentivirus vector (lv-shCON), and it is served as the negative-control scrambled sequence. All the lentivirus constructs were transduced according to the manufacturer's instructions.

Xenograft experiments

The four-week-old female BALB/c-Nude mice were provided by GemPharmatech (Nanjing, China) and housed at the Experimental Animal Center, Tongji Medical College, Huazhong University of Science and Technology (HUST, Wuhan, China). All mice were divided into groups randomly. 5×10^6 SiHa cells or 1×10^7 HaCaT cells was injected subcutaneously in the right flank of each mouse. The tumor size was measured by a digital caliper and calculated using the following formula: $0.5 \times \text{Length} \times \text{Width}^2$. All the experimental protocols were approved by the Institutional Animal Care and Use Committee of HUST, and we carried out the study in strict accordance with the Guidelines for the Welfare of Animals in Experimental Neoplasia.

Immunohistochemistry staining (IHC)

We euthanized the mice and isolated the xenografts, then fixed them with 4 % paraformaldehyde. IHC staining was performed on the paraffin-embedded sections (5 μm) according to the protocol of antibodies. The slides were incubated with the rabbit anti-Ki67 (1:150, ab16667, Abcam) primary antibody overnight at 4 °C. Antibody detection was performed using DAB. Photographs were taken using cellSens Dimension, and the staining intensity was measured using ImagePro Plus.

Extraction of total cellular DNA

After trypsin digestion, the cultured cells were collected in a 1.5 ml EP tube by centrifugation at 1200 rpm for 5 min. Then, the cells were washed with 1x PBS and centrifuged again to obtain cell precipitates. DNA from cells was extracted using a standard Qiagen DNA extraction kit following the kit's operating instructions. DNA was quantified and quality assayed using a Qubit 2.0 Fluorophotometer (Life Technologies), NanoDrop One/One^C (1.8 < O D260/OD280 < 2.0). DNA was stored at −80 °C.

G-banding chromosome karyotype analysis

The G-banding chromosome karyotype analysis procedure have been described in detail previously [15]. The reagents were prepared as follows: colchicine (Sinochrome), 0.9% saline, trypsin (Thermo Fisher), hypotonic solution (0.4% KCl mixed 1:1 with 0.4% sodium citrate), fixative (3:1 methanol:glacial acetic acid), staining solution (Giemsa stock solution:phosphate buffer, pH 7.4 = 1:6), digestion solution (50 ml of physiological saline placed in 37 °C water for 0.5 h, followed by addition of 1 ml of 0.25% trypsin and rewarming for 30 min.). First, the cultured cells were incubated with colchicine for 2 h, and then the culture medium was transferred into an EP tube. Second, after incubating the cells with trypsin for 2 min, the cell suspension was transferred to the above EP tube and centrifuged at 1500 rpm for 10 min to obtain the cell pellet. Next, the cells were resuspended in hypotonic solution, incubated in 37 °C water for 20 min, prefixed with fixative and centrifuged at 1500 rpm for 10 min to obtain a cell pellet. After repeating the above fixation process 2 times, we resuspended the cell pellet in fixative to obtain the final cell suspension. We then dropped the cell suspension onto a glass slide and allowed it to dry. The slide was placed in digestion solution for 15 s, washed with physiological saline, and then stained with the staining solution. Finally, we counted and recorded 10 split phases under the microscope for analysis.

Western blot assay

The proteins were isolated and resolved by Sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE), transferred to polyvinylidene difluoride membrane (PVDF) and probed with primary antibodies for 12 h at 4 °C. The proteins were visualized using the ECL system (Bio–Rad) and analyzed by Image Lab (6.0.1). The primary antibodies and dilutions used were as follows:

GAPDH (1:2500, rabbit, YM1235, Immunoway), PIBF1 (1:2000, rabbit, NB100-58-834, NOVUS), SMC3 (1:1000, rabbit, ab9263, Abcam), RAD21 (1:1000, rabbit, ab217678, Abcam), H3 (1:500, rabbit, ab32388, Abcam), SP1 (1:1000, rabbit, A14662, ABclonal).

Subcellular protein fractionation

After trypsin digestion, the cultured cells were collected in 1.5 ml EP tubes by centrifugation at 1200 rpm for 5 min. The supernatant was discarded, and the cells were resuspended with 1x PBS and centrifuged again to obtain the cell precipitate. We used the Subcellular Protein Fractionation Kit for Cultured Cells (Thermo Scientific, 78840) to obtain the protein supernatant from the cytoplasm, nuclear pulp, and chromatin-bound components according to the kit instructions. We added protein loading buffer to the protein and boiled the sample at 100 °C for 15 min. Finally, the protein expression in each component was detected by immunoblotting assay.

Immunoprecipitation assay

Cultured cells were lysed for 2 h in ice-cold RIPA lysis buffer (Beyotime, P0013) containing PMSF protease inhibitors (Beyotime, ST505). After centrifugation at 12000 rpm for 5 min at 4 °C, the supernatant containing the total protein extract was transferred to a new EP tube. The preparation for the combination of antibody and beads was performed as follows: first, the primary antibodies against PIBF1 (NOVUS, rabbit, NB100-58834) and anti-IgG (Beyotime, rabbit, A7016) were diluted according to the antibody instructions. Second, the antibody was coincubated with protein A + G agarose beads (Beyotime, A7016) for 2 h with rotation at 4 °C and then centrifuged at 1000*g* for 5 min at 4 °C to obtain the antibody-bound beads. Next, 600 μl of the total protein supernatant obtained above was added to the antibody-conjugated protein A + G agarose beads and rotated overnight at 4 °C to bind the

protein to the antibody. On the second day, the reaction was centrifuged at 1000*g* at 4 °C for 5 min, the supernatant was discarded, and the protein A + G agarose beads were washed 4 times with 1x PBS. In the next step, we added 1.5 × loading buffer to the sample and boiled it at 100 °C for 5 min. Finally, the obtained protein was analyzed by western blot assay.

Immunofluorescence assay

Cells were plated in 6-well plates on glass coverslips. After 24 h, the cells were fixed with 4% paraformaldehyde for 15 min, permeabilized in 0.5% Triton X-100/1X PBS for 20 min and blocked in 3% BSA/1X PBS solution for 30 min. Cells were then incubated with primary antibody overnight at 4 °C, washed with 1x PBST and incubated with HRP secondary antibody for 1 h at 37 °C. Nuclei were stained with 4′6-diamidino-2-phenylindole (DAPI). Images were acquired with fluorescence microscope. The antibodies included PIBF1 (1:100, mouse, sc-166372, Santa Cruz), SMC3 (1:200, rabbit, ab9263, Abcam), RAD21 (1:100, rabbit, ab217678, Abcam), Cy5-goat anti-mouse IgG H&L (ab6563, Abcam), and Cy5-goat anti-rabbit IgG H&L (Ab6564, Abcam).

Liquid chromatography/mass spectroscopy (LC–MS/MS) analysis

The mass spectrometric procedure and detection conditions used have been described in detail previously [16]. Briefly, coimmunoprecipitated proteins were prepared by immunoblotting; the area of the protein in the SDS–PAGE gel was excised and cut into small pieces of approximately 1 mm^3, and then the peptides were prepared according to the experimental procedure for detection and analysis. Mass spectrometry analysis was performed using the Q Exactive Plus LC–MS system (Thermo), and the resulting mass spectra were searched by MaxQuant (V1.6.2.10) using the MaxLFQ database search algorithm. The database searched was the Human Proteome reference database in UniProt.

siRNA

Cells were transfected with small interfering RNA targeting PIBF1 (si-PIBF1-1 and si-PIBF1-2, RiboBio, Guangzhou, China) or a nontargeting control (si-NC, RiboBio, Guangzhou, China, siN0000001-1–5) using Lipo3000 (Thermo, L3000015) according to the manufacturer's instructions. The cells were harvested 72 h after transfection.

Whole-genome sequencing and processing

Genomic DNA extraction of the HaCaT, Vector, and KI cell lines using a QIAamp DNA Mini Kit (QIAGEN). After the sample passes the DNA integrity test, g-Tube (Covaris) was used to shearing 8 μg of genomic DNA and concentrated with AMPure PB magnetic beads. The construction of each SMRTbell library was using the Pacific Biosciences SMRTbell Template Prep Kit 1.0 or Express Template Prep Kit 2.0. We produced the library size-selected by Sage ELF BluePippin™ system 3 for molecules 8–12 kb/14–17 kb, followed by primer annealing and the binding of SMRTbell templates to polymerases with the DNA Polymerase Binding Kit. Sequencing was carried out for 30 h on the Pacific Bioscience Sequel II platform at Annoroad Genomic Company. We used NGLMR [17] to perform the alignments with default parameters.

Hi-C sequencing and processing

(1) Hi-C assay preparation and sequencing.

Adherent cells were crosslinked with 40 ml 2 % formaldehyde solution 15 min under vacuum at room temperature. After that, we used 4.324 ml 2.5 M Gly in order to terminate the crosslinking reaction. Afterward, the supernatant was discarded, and cells retained in the precipitate. Cells were frizzed by liquid nitrogen. We resuspended the cells in 25 ml extraction buffer. Next, we filtered the sample through Miracloth (Calbiochem), and the filtrate was centrifuged at 4000 rpm for 20 min at 4 °C. Thereafter, we resuspended the pellet in 1 ml extraction II buffer (10 mM Tris-HCl pH 8, 1 % Triton X-100, 0.25 M sucrose, 10 mM MgCl2, 0.1 mM PMSF, 13 protease inhibitors and 5 mM mercaptoethanol). After a few moments, the solution was centrifuged 10 min at 14000 rpm and 4 °C. The pellet was then resuspended in 300 μl extraction buffer III (10 mM Tris-HCl pH 8, 2 mM MgCl2, 0.15 % Triton X-100, 1.7 M sucrose, 5 mM β-mercaptoethanol, 1 μl protease inhibitor, 0.1 mM PMSF). Then, add 300 μl clean extraction buffer III, and centrifuged the sample 10 min at 14000 rpm. After the supernatant discarded, the pellet we washed with 500 μl ice-cold 1x CutSmart buffer and the sample was centrifuged for 5 min. The nucleus was remained in the pellet. Thereafter, the pellet was washed twice used restriction enzyme buffer and then transferred the pellet to a safe-lock tube. Then, the chromatin was solubilized with dilute SDS and incubated for 10 min at 65 °C. Triton X-100 overnight was used to quench the SDS. In the next step, a 4-cutter restriction enzyme (400 units MboI) was used to digested the nuclei on a rocking platform at 37 °C. After that, the DNA ends were labeled with biotin-14-dCTP, and the blunt-end were joined between the cross-linking fragments. Thus, the proximal chromatin DNA were ligated. We incubated nuclear complexes with proteinase K at 65 °C and then generated reverse crosslinked nuclear complexes [18]. We purified the reversed crosslinked nuclear complexes by phenol–chloroform extraction after crosslinking. After we cleared the Biotin-C from nonligated fragments by adding T4 DNA polymerase, we sheared the fragments to 200–600 bp by sonication and the fragments were repaired with a mixture of T4 DNA polymerase, DNA Polymerase I Klenow Fragment and T4 polynucleotide kinase. In the next step, enrichment of the biotin-labeled DNA fragments was performed by streptavidin C1 magnetic beads. The Hi-C library was then sequenced on the Illumina HiSeq X Ten platform with paired-end reads of 150 bp. We trimmed the raw read length to 50 bp and filtered with fqtools_plus (https://github.com/annoroad/fqtools_plus) discarding low-quality reads, reads containing adapters (>5 bp adapter nucleotide) and reads with a high N ratio (>5 %).

(2) Hi-C read mapping and normalization.

Corresponding to our experiment, the donor sequence is inserted into the chr13:73814191 of the *Homo sapiens* genome (hg19) as the customized reference of all omics data alignment. We mapped reads to the customized hg19 reference through Bowtie2 (v2.3.4) [19]. Then, HiC-Pro was used to do genome mapping and send the unique mapped paired-end reads to genome MboI simulated digestion fragments [20]. Only the paired-end reads mapped to diffrent MboI fragments uniquely were extracted for processing. We used HiC-Pro software (v2.7.1) build matrix tool to generate interaction matrices at different resolutions (1 Mb, 100 kb, 40 kb, 20 kb) [20]. HiCCUPS software (v2.7.1) was used to constructed 10 kb and 5 kb resolution matrices in hic format. Briefly, ICE (Iterative Correction) method was used to remove potential chromatin interaction bias [20]. And we used a more computational efficiency version (https://github.com/seqyuan/iced) developed by us.

(3) Sample interaction differential and IDE analysis.

Observed/Expected hic matrix was used to analyze HaCaT, Vector, and KI interaction differences (described below in "Identification of A/B compartments") of each sample's intrainteraction matrices. The differences in cis interactions between pairs of samples are displayed in heatmaps. IDE (Interaction Decay Exponents)

describe the interaction frequency or trend in relation to the genome physical distance, the IDE values was calculated by (https://github.com/dekkerlab/cworld-dekker) at 1 Mb resolution. The bin pair distance and interaction frequencies were both log10 transformed, and the interaction and distance exponents were fit a linear model. The slope of the linear model was taken as the IDE value.

A/B compartments identification

First, the expected interaction score of variable distance was calculated by loess smoothed method in each intrachromosomal matrix. Then we calculate the Observed/Expected matrix of each chromosome. After that, PCA (principal component analysis) of the Observed/Expected Pearson correlation matrix was defined as the compartment eigenvalues. After that, we adjusted the PCA eigenvalues by gene density to define the A/B compartments. We also compared the interaction matrices and Pearson's correlation matrices of each chromosome to determine whether they matched each other. In addition, we checked whether the gene expression count was significantly different in the two different compartments. Then, we decided whether eigen1 values or eigen2 values should be used for each chromosome in each sample. In addition, multiplication by -1 was performed if the gene expression level of the chromosome compartment with the positive eigenvalue was significantly lower than that of the other compartment. Finally, we define the compartment with a negative or zero eigenvalues as the B compartment and the other as the A compartment (https://github.com/dekkerlab/cworld-dekker).

Defined contact domain boundaries

CDBs (contact domain boundaries) was identified by HiCDB method [21] at 10 kb resolution matrix by default parameters. The CDBs differences in samples was compared by bedtools [22].

Loop calling

The whole genome chromatin interaction loops were identified by HiCCUPS (Hi-C Computational Unbiased Peak Search), a tool of the Juicer package using a 5-kb resolution matrix [23]. To identified sample-specified loops, all of the sample loop anchors Observed/Expected ratio was calculated in each. And then, we compared these correspond loop OE ratios between samples. If the OE ratio of one sample is twice that of the other, then we define this loop as distinct between the two samples.

RNA-seq analysis

Samples were extracted by the TRIzol method, RNA-seq sequencing libraries were constructed followed by the standard Illumina protocols, and then sequenced on the Illumina HiSeq X-ten system. Each library was performed three biological replicates. The filtered sequencing reads were mapped to the customized hg19 human genome using Hisat2 (v2.1.0) [24] software with the default parameter. The gene expression level was normalized using the FPKM method as described by Traver Hart [25]. DESeq2 package [26] was used to identified differentially expressed genes (DEGs). Significantly differentially expressed genes were considered by parameters: absolute log2-fold change $\geq$ 1 and Benjamini-Hochberg adjusted q-value < 0.05.

ChIP sequencing and processing

We performed chromatin immunoprecipitation libraries according to previously described methods [27]. In brief, 10 million cells were washed twice in a cold PBS buffer, cross-linked with 1 % formaldehyde 10 min at room temperature, and then addition of glycine (125 mmol /L final concentration) were used to quenched. Afterward, the sample is then lysed on ice to produce chromatin. Sonicated was used to obtain soluble sheared chromatin, average DNA length was 200–500 bp. 100 μl of the sheared samples was used for immunoprecipitation with primary antibodies and IgG antibodies (Abcam, ab17870), 20 μl was saved at –20 °C for input DNA. The antibodies used were as follows: H3K27ac (Abcam, ab4729), H3K27me3 (CST9733), CTCF (ABclonal, A13272), PIBF1 (Novus, NB100-58834), RAD21 (Abcam, ab217678), and SMC3 (Abcam, ab9263); At 4 °C, 10 μg antibodies were used overnight for each immunoprecipitate reaction. The next day, 30 μl protein beads were added and incubated for 3 h. The beads were next washed once with 20 mM Tris/HCl (pH 8.1), 1 % Triton X-100, 50 mM NaCl, 0.1 % SDS and 2 mM EDTA; twice with 10 mM Tris/HCl (pH 8.1), 1 mM EDTA, 1 % deoxycholic acid, 1 % NP-40 and 250 mM LiCl; and twice with TE buffer 1× (10 mM Tris-Cl at pH 7.5, 1 mM EDTA). The bound material was then eluted from the beads in 300 μl elution buffer (100 mM NaHCO3, 1 % SDS) and first treated with RNase A (final concentration 8 μg/mL) at 65°C for 6 h. Then proteinase K (final concentration 345 μg/mL) was treated overnight at 45°C. Sequencing libraries were constructed using immunoprecipitated DNA following a protocol provided by Illumina® sequenced INEXTFLEX® ChIP-Seq Library Preparation Kit (NOVA-5143–02, Bioo Scientific). PE 150 sequencing was performed on Illumina HiSeq X Ten platform.

The bowtie2 software [19] aligned ChIP-seq reads to custom hg19 references and uses only unique and non-duplicate mapped reads for downstream analysis. Samtools was used to calculate read coverage and depth [28]. To check the repeatability of the ChIP-seq experiment, deeptools was used to generate correlation graphs for all samples, including the input samples. DeepTools [29] bamCoverage function was used to generated ChIP Signal track files in BigWig format, and normalized to 1 million reads for visualization. The gene body and flanking region heatmap graph according to the normalized signal intensity were also generated by DeepTools. Peak calling was utilized by MACS2 [30], and the following step was annotation of peaks by using bedtools [31].

Cut-Tag sequencing and analysis

We first harvested cells and then centrifuged them for 3 min at 600g at room temperature. Subsequently, the samples were washed twice in wash buffer by gentle pipetting. Then, cell nuclei were prepared and resuspended. The next phase was incubation of the sample and primary antibody overnight at 4 °C on a rotating platform. After centrifugation, the primary antibody was removed carefully. To increase the number of Protein A binding sites for each bound antibody, we incubated the cells for 30 min at RT in secondary antibody diluted in DigWash buffer. Next, the cells were washed using DigWash buffer to remove unbound antibodies. The pA-Tn5 adapter complex dilution was prepared and added to the nucleus with gentle vortexing. Then, the nuclei were incubated at RT for 1 h and washed in wash buffer to remove unbound pA-Tn5 protein. In the subsequent step, we resuspended the nuclei in tagmentation buffer and incubated them for 1 h at 37 °C. To stop the tagmentation reaction, STOP buffer was added to the sample, incubated at 55 °C for 30 min and then heated at 70 °C for 20 min. Library DNA was then purified and amplified for sequencing with the Illumina NovaSeq 6000 [32]. Raw sequence reads were

first processed by FastQC for quality control, and then poor-quality reads and adapter sequences were removed. Quality filtered reads were then mapped to the reference genome using Bowtie [19], and the reads mapped to the genome unique site were kept for next processing. MACS2 [30] was used to call peaks in the next step. The mapped bam files were convert to bigwig format for later visualization analysis.

GO analysis

There are many differentially expressed genes in areas of chromatin alterations, such as A/B compartments switched regions and differential loops or contact domains. We therefore performed GO enrichment analysis to determine whether these genes were significantly enriched in a biological process. For each GO term, we obtained a BH-corrected p value corresponding to a single, independent test for each GO term [33].

KEGG analysis

We also applied KEGG analysis to analyze the genes in areas of chromatin alteration. For the KEGG terms, we obtained the BH-corrected p value corresponding to a single, independent test [34].

Statistical analysis

The relationship between interaction changes were used fisher's test and wilcoxon rank sum test to determine, transcriptome and epigenome signals. Annoroad-OMIC-Viz (https://github.com/Spartanzhao/Annoroad-OMIC-Viz) and trackc (https://github.com/seqyuan/trackC) were used to generate multiomics display figures.

Results

Establishment of a 13q22 site-specific HPV16 gene knock-in cell model

Based on the CRISPR–Cas9 gene editing combined with the homologous end repair mechanism for DNA knock-in at the target site [35], as shown in Fig. 1A, we established a gene-editing design for knock-in of the HPV gene at the 13q22 site in HaCaT cells. Based on the integration site in the human cervical cancer cell line SiHa, we constructed the three highest-scoring sgRNAs targeting the 13q22 site sequence in HaCaT cells (Fig. S1A). We validated the effectiveness of the sgRNAs by T7EI assay and Sanger sequencing. As shown in Fig. S1B-D, CRISPR–Cas9 with sgRNA3 efficiently edited the sequences of the genome target site. Then, the knock-in donor was cloned to efficiently screen the integrated positive cells, containing the Flag-Loxp + CMV-EGFP-2A-Puro-polyA + Loxp cassette and HPV16 E7-E6-URR gene, flanking 500 bp homologous arm sequences (Fig. S1E). We also designed a control donor containing the flag gene only (Fig. S1F).

Previous studies have proven that double-stranded oligodeoxynucleotides (dsODNs) are more efficient for targeted insertion of DNA than circular donors when used as templates in combination with CRISPR–Cas9[36], so we prepared the KI donor (4.5 Kb) and Vector donor (3.2 Kb) in the form of dsODNs (Fig. S1G). Following the procedure, cells were cotransfected, screened, and fluorescence sorted. We finally obtained EGFP-positive monoclonal cell lines (Fig. 1B). Flow cytometry assays showed a high purity of monoclonal cells (Fig. S2A). Both the specific DNA amplification products (Fig. 1C-D) and Sanger sequencing results (Fig. S2B-C) verified the accuracy of the insert sequences. To further precisely characterize the donor DNA knock-in at the 13q22 site, we performed whole-genome long-read sequencing in KI cells and Vector cells. The reads from the KI cells indicated a 3.5 kb insertion at chr13:73814191, and the reads from the Vector cells indicated a 2.2 kb insertion at the same target site (Fig. 1E). The insert sequences of all these reads were aligned to the donor sequence. The results clearly demonstrated that the donor DNA was successfully knocked in at the target site. The difference between the total number of reads aligned to human genomic DNA near the integration site (n = 32) and the number of reads aligned partially to donor DNA (n = 16) indicated heterozygosity in the KI cells. Because a previous study reported that CRISPR–Cas9 caused chromosome breakage during gene editing [37], we confirmed the integrity chromosome 13 in the cell models using G-banding chromosome karyotype analysis (Fig. 1F). Finally, the E6 and E7 mRNA expressions were also detected in KI cells (Fig. 1G). In summary, we have successfully constructed a 13q22 site-specific HPV16 gene knock-in cellular model (KI) and a 13q22 site-specific flag gene knock-in cellular model (Vector). To our knowledge, this represents the first site-specific HPV knock-in cell model.

HPV16 gene knock-in at the 13q22 site alters the cell biological phenotype

After establishing the knock-in cell lines, we performed cell experiments to evaluate the effect of HPV16 gene knock-in on cell biological behavior. As shown in Fig. 2A and 2B, the number of colonies was much higher in KI cells than in HaCaT and Vector cells. In addition, the KI cells grew faster than the control cells, according to the CCK-8 assay (Fig. 2C). The Transwell and wound healing assay results showed that HPV16 gene knock-in promoted cell invasion and migration ability (Fig. 2D-G). To Further explore whether HPV16 gene knock-in affected cells stratified growth, we developed a 3D organotypic coculture model, consisting of ECM Matrigel substrate at the bottom, a layer of fibroblasts in the middle, and a layer of KI cells or control cells at the top, which faithfully represents the histologic and biological microenvironment in vivo. Our results demonstrated that HPV16 gene knock-in enhanced the cell stratified growth (Fig. 2H). To evaluate the tumorigenic potential, three group cells (HaCaT, Vector, and KI) were injected subcutaneously into nude mice. Tumor growth was measured at weekly intervals. As exemplified in tumor growth curves of the three group cells (Fig. 2I), HaCaT and Vector cells formed epidermal cysts and finally growth arrest completely, and KI cells survived and remained vital during the observation period of 7 weeks and the tumor enlarged slowly. Histologically, both H&E and IHC staining showed necrosis in the HaCaT and Vector xenograft tissues, but monoclonal proliferation was detected in the KI xenografts tissues (Fig. 2J). Furthermore, the IHC assays showed that the expression levels of Ki67 increased in the KI xenograft tissues (Fig. 2K). These data suggested that HPV16 gene knock-in at the 13q22 site induces malignant biological behavior in KI cells.

Chromatin architecture undergoes reorganization after HPV16 gene knock-in at the 13q22 site

To explore whether HPV16 gene knock-in at the human genome can alter chromatin architecture and gene expression, we conducted Hi-C analysis on KI and control cells, for which RNA-seq and ChIP-seq datasets had also been produced. Our results showed that approximately 88.5 % of the A/B compartment remained conserved in KI cells compared to HaCaT and Vector cells, and only 6.2 % of the genome underwent A/B or B/A compartment changes (Fig. S3A). Interestingly, a large proportion of A/B compartment transformations occurred on chromosome 13 (Fig. S3B, C) and mainly occurred within 10 Mb near the HPV knock-in site (Fig. 3A). GO and KEGG pathway analyses were performed for ge-

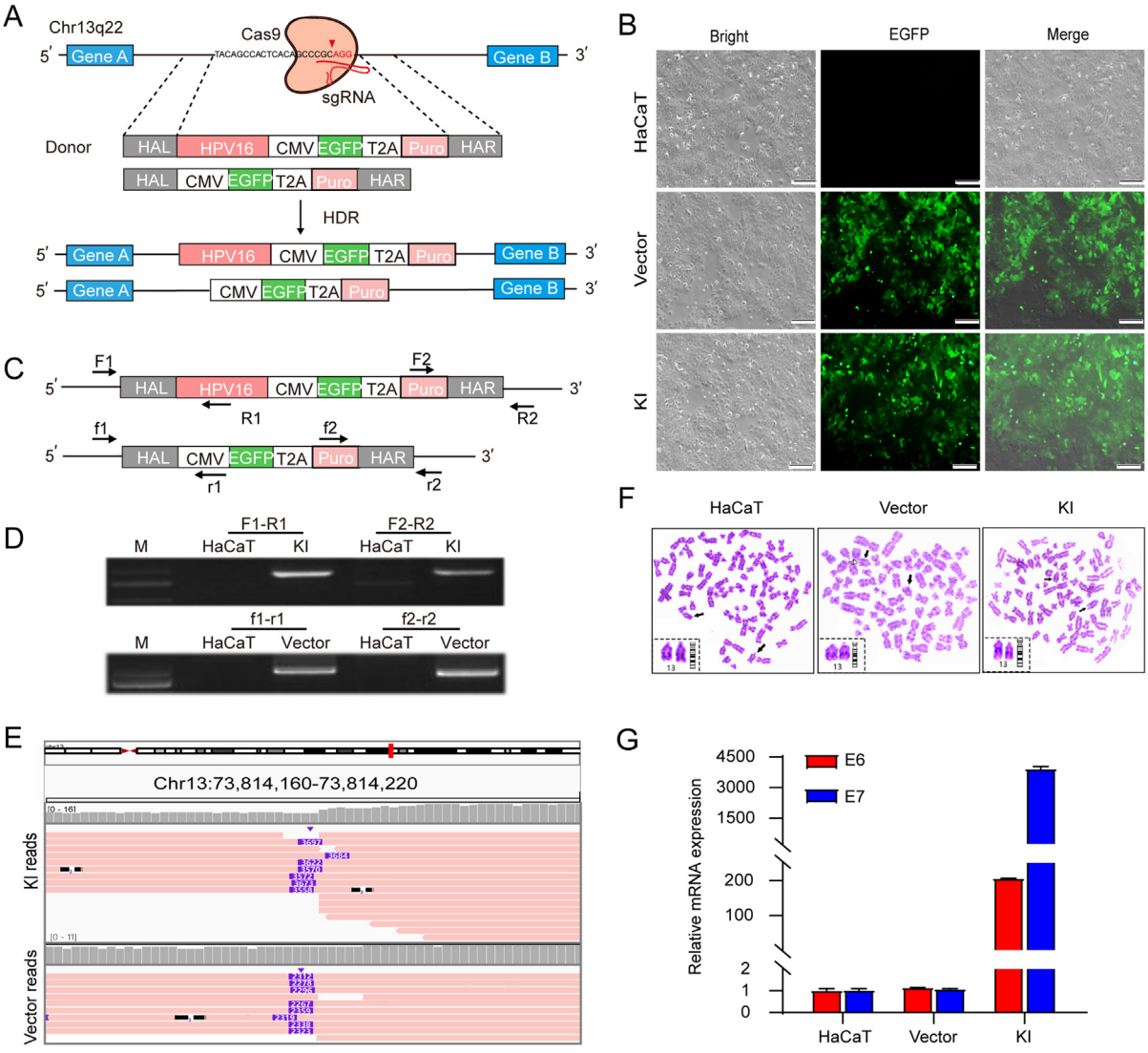


Fig. 1. Establishment of a 13q22 site-specific HPV16 gene knock-in cell model. (A) Schematic overview of the 13q22 site-specific gene knock-in strategy through CRISPR–Cas9 system. The red letters represented the PAM sequence, and the triangle indicates the sgRNA target site. Donor: the upper is for KI, and the lower is for Vector. HAL/HAR, left/right homology arm; HDR: homology-directed repair. (B) Representative EGFP fluorescence images of HaCaT, Vector, and KI cells. HaCaT: parental cells; KI: 13q22 site-specific HPV16 URRE6E7 and flag gene knock-in HaCaT cell model; Vector: 13q22 site-specific flag gene knock-in HaCaT cell model. (C, D) Identification of site-specific gene knock-in by PCR. Schematic diagrams of PCR primers designed for gene knock-in in KI and Vector cells (C) and agarose gel electrophoresis of DNA PCR fragments for the PCR analysis of HaCaT, Vector, and KI cells with primers as indicated in figure C (D). F1-R1 and f1-r1 are the PCR primers at the 5′ end of the knock-in site in KI and Vector cells; F2-R2 and f2-r2 are the PCR primers at the 3′ end of the knock-in site in KI and Vector cells. The direction of the arrows represents the direction of the primers. The primer sequences are listed in the Table S1. (E) An integrative genomics viewer screenshot of long-read whole-genome sequencing shows 16 reads indicating a 3.592 kb insertion at the 13q22 locus in KI cells and 11 reads indicating a 2.312 kb insertion at 13q22 locus in Vector cells (purple rectangles indicate the full-length donor sequence). (F) G-banding chromosome karyotype analysis of HaCaT, Vector, and KI cells. The black arrow indicates chromosome 13. (G) The relative mRNA expression of HPV16 E6 and HPV16 E7 in HaCaT, Vector, and KI cells.

nes associated with A/B compartment changes in KI cells. The results showed that these genes are associated with cancer and metabolism-related pathways, including the regulation of cell adhesion, ubiquitin-mediated proteolysis, and histidine metabolism pathways (Fig. S3D-E). As shown in Fig. 3B, C, the TAD around the target site was highly similar in HaCaT and Vector cells, but in KI cells, the large TAD was split into two smaller TADs after HPV knock-in. The Hi-C map clearly shows that the interaction frequency of genome regions at both ends of the integration site is significantly reduced. A subtraction heatmap (Fig. 3D) and the interaction differences score distribution (Fig. 3E) show this difference more clearly. We also analyzed the genome-wide change in TADs. With the HiCDB method and contact domain boundary (CDB) local relative insulation (LRI) score analysis, we identified 9239 conserved CDBs in the three groups of cells and 2336 specific CDBs in KI cells (Fig. S3F, G). As an indicator of active transcriptional status, the H3K27ac ChIP-seq profile of the CDBs showed enhanced signal in KI cells (Fig. S3H), which may explain the altered gene expression network in KI cells. KI-specific CDB-containing genes were used for pathway enrichment analysis. We identified many enrichment pathways related to cell growth and the cell cycle, such as cell cycle and regulation of cell growth (Fig. S3I, J), which correspond to the biological behavior of rapid cell proliferation described above. In addition, we further focused on the characteristic changes in the chromatin loop after HPV integration. We identified 860 loops with an increase in long-range interaction and 305 loops with a reduction in long-range interaction in KI cells compared with control cells (Fig. 3F, G). Changes in

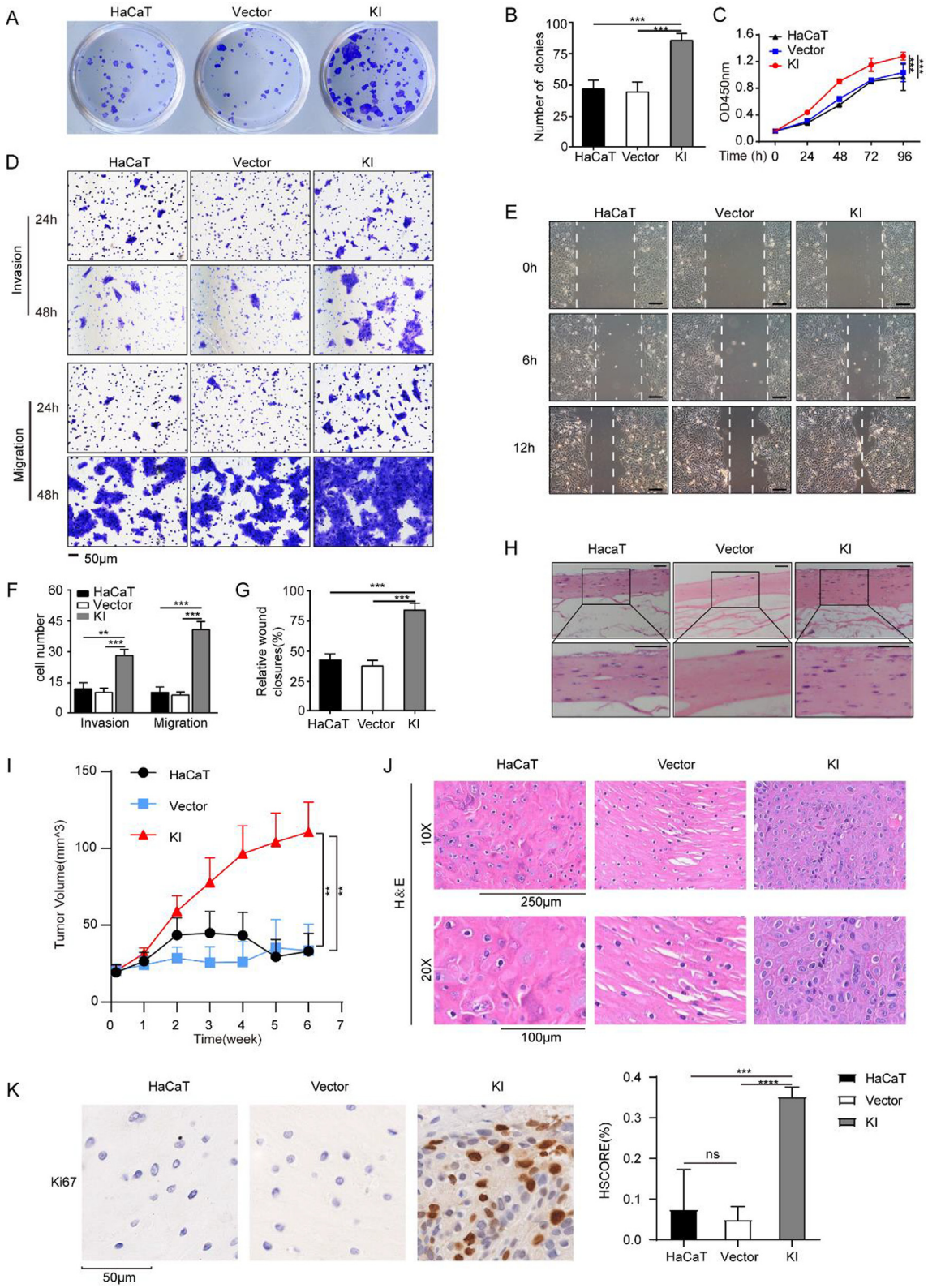


Fig. 2. HPV16 gene integration at the 13q22 site promoted cell proliferation, invasion and migration, and stratified growth. (A, B) Colony-forming ability of cells. Representative images (A) and quantitative analysis results (B) of the colony-forming assay for HaCaT, Vector, and KI cells. The viability of the three groups of cells at 0 h, 24 h, 48 h, and 72 h, as detected by CCK-8 assay. (D) Representative images of the Transwell assay to detect the migration and invasion ability of the three groups of cells at 24 h and 48 h, respectively. Scale bar: 50 μm. (E) Representative images showing the scratch assay results for the three groups of cells at 0 h, 6 h, and 12 h. Scale bar: 200 μm. Computed migration front lines are shown (white). (F) Quantitative analysis of the invasion and migration assay results. (G) The relative scratch healing rates of the three groups of cells. (H) Representative H&E staining results of in vitro three-dimensional cell culture. The lower panel corresponds to the section marked by the black box in the upper panel. Scale bars: 20 μm. (I-J) For the in vivo analyses, 1×10^7 cells (HaCaT, Vector, and KI) were injected subcutaneously into the posterior hip of nude mice. Each group included five nude mice. I)Tumor growth was recorded every week until week 7. J) Represent H&E-stained images in xenograft tumor tissues. (K) Immunohistochemical analysis of Ki67 expression in xenograft tumor tissues and Ki67 protein expression was evaluated using HSCORE. The data shown represented mean ± SD (Student's *t*-test). P value was denoted as *P < 0.05, **P < 0.01, and ***P < 0.001, "ns" represented "not significant".

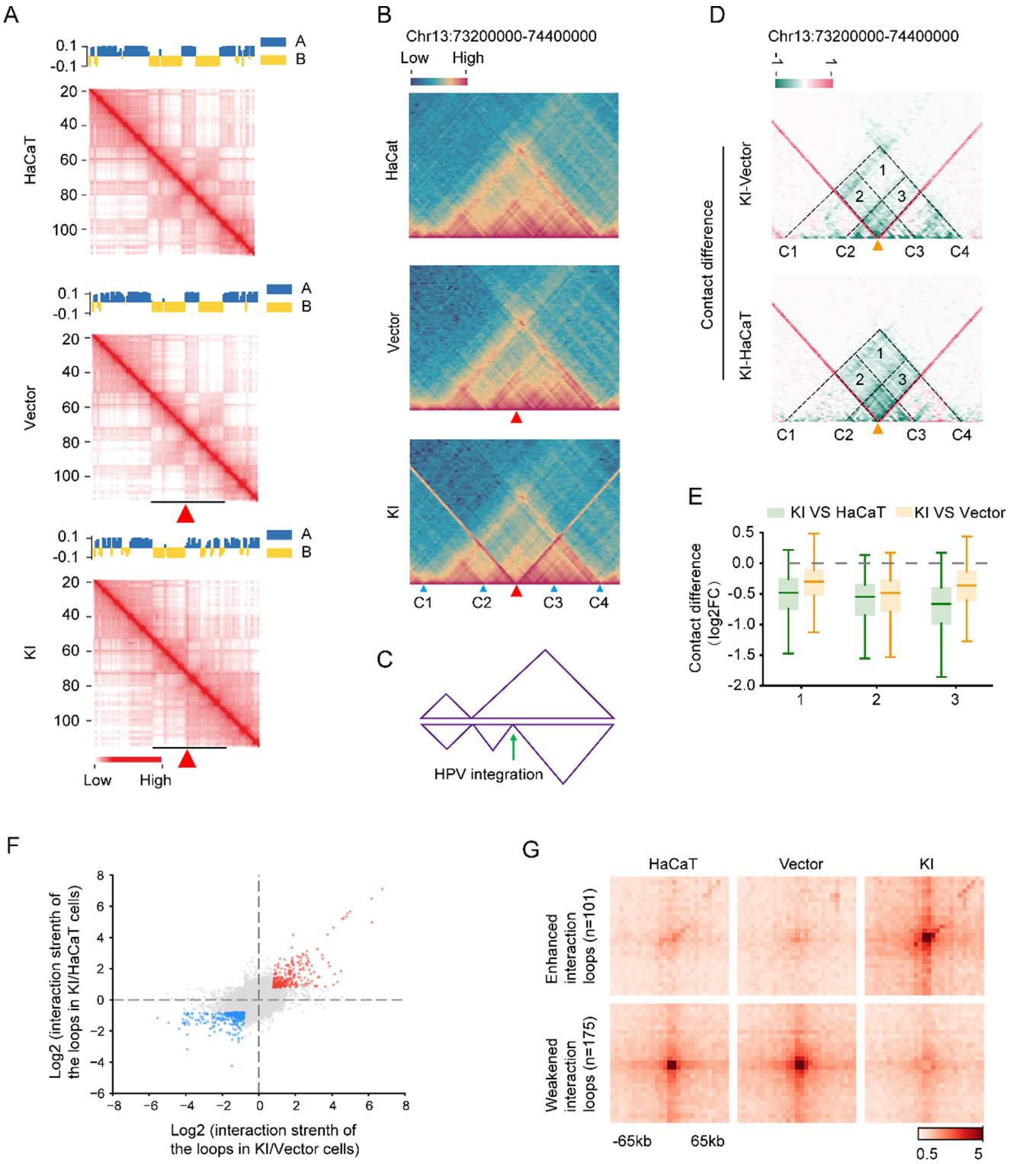


Fig. 3. Chromatin architecture undergoes reorganization after HPV16 gene knock-in at the 13q22 site. (A) Hi-C maps indicating cis contacts of chromosome 13 in three groups of cells (resolution:1 Mb). Corresponding cis compartment tracks are shown at the top. The blackline indicates the region in which the compartment was altered. The red arrow indicates the knock-in site. (B) Hi–C maps of chromatin contacts for the Chr13:73.2 M–74.4 M region of the three groups of cells (resolution: 20 kb) illustrate the local contact domain reorganization near the integration site. The red arrow indicates the integration site. The blue arrows indicate the CDBs (C1, C2, C3, C4). (C) Pattern diagram illustrating the change in chromatin TADs after HPV knock-in. Purple triangle: TAD; green arrow: HPV knock-in. (D) Hi-C difference maps of the Chr13:73.2 M–74.4 M region obtained by subtracting the normalized Hi-C contact frequencies in control cells from those in KI cells. The yellow arrow indicates the integration site. C1-C4 indicate the CDBs. The local contact domain of the TAD where the integration site was located is depicted using the black dotted lines and defined as contact domains 1–3. (E) Distribution of contact difference scores within the contact domain defined in figure D in KI cells compared to control cells. (F) Scatter plot showing the long-range interaction loops identified as enhanced (red dots, n = 860) and weakened (green dots, n = 305) in KI cells compared to control cells. (G) Aggregated peak analysis of KI cell-specific enhanced (top, n = 101) and weakened (bottom, n = 175) long-range interaction loops in comparison to those in control cells.

chromatin loop interactions often correlate with transcriptional changes. Our results showed that the increase in long-range interaction was associated with upregulation of gene expression (Fig. S3K). Moreover, several pathways enriched among these genes were associated with tumorigenesis, such as the AMPK signaling pathway and glycosaminoglycan biosynthetic process (Fig. S3L, M). Altogether, these results show that HPV16 gene knock-in at the genome 13q22 site led to chromatin architecture reorganization.

PIBF1 is a DNA-binding protein and is upregulated in KI cells

Our previous high-throughput sequencing study based on cancer tissues revealed that HPV integration at hotspots affects the expression of key genes near the integration site [5]. In this study, in addition to the TAD changes near the integration site, the H3K27ac ChIP-seq signal was significantly enriched in KI cells compared to control cells (Fig. 4A). The expression of genes near the integration site showed that *PIBF1*, *MZT1*, and *KLF5* were upregulated in KI cells (Fig. 4B). The relationship between *MZT1* or *KLF5* and carcinogenesis has been well recognized [38,39]. While *PIBF1* is known to play a crucial role in maintaining pregnancy through its effects on the immune system, its function in tumorigenesis needs further investigation. Western blot assays detected high expression of PIBF1 in KI cells compared to two control cells (Fig. 4C). Molecular function studies by Beata Polgar et al.[40] suggested that PIBF1 may have transcription factor functions, which was confirmed for the first time in a study of invasion regulation. To further clarify the function of PIBF1 binding to DNA with new methods, we performed a subcellular protein fractionation assay to detect the distribution of PIBF1 protein in the cytoplasm, nucleus, and chromatin

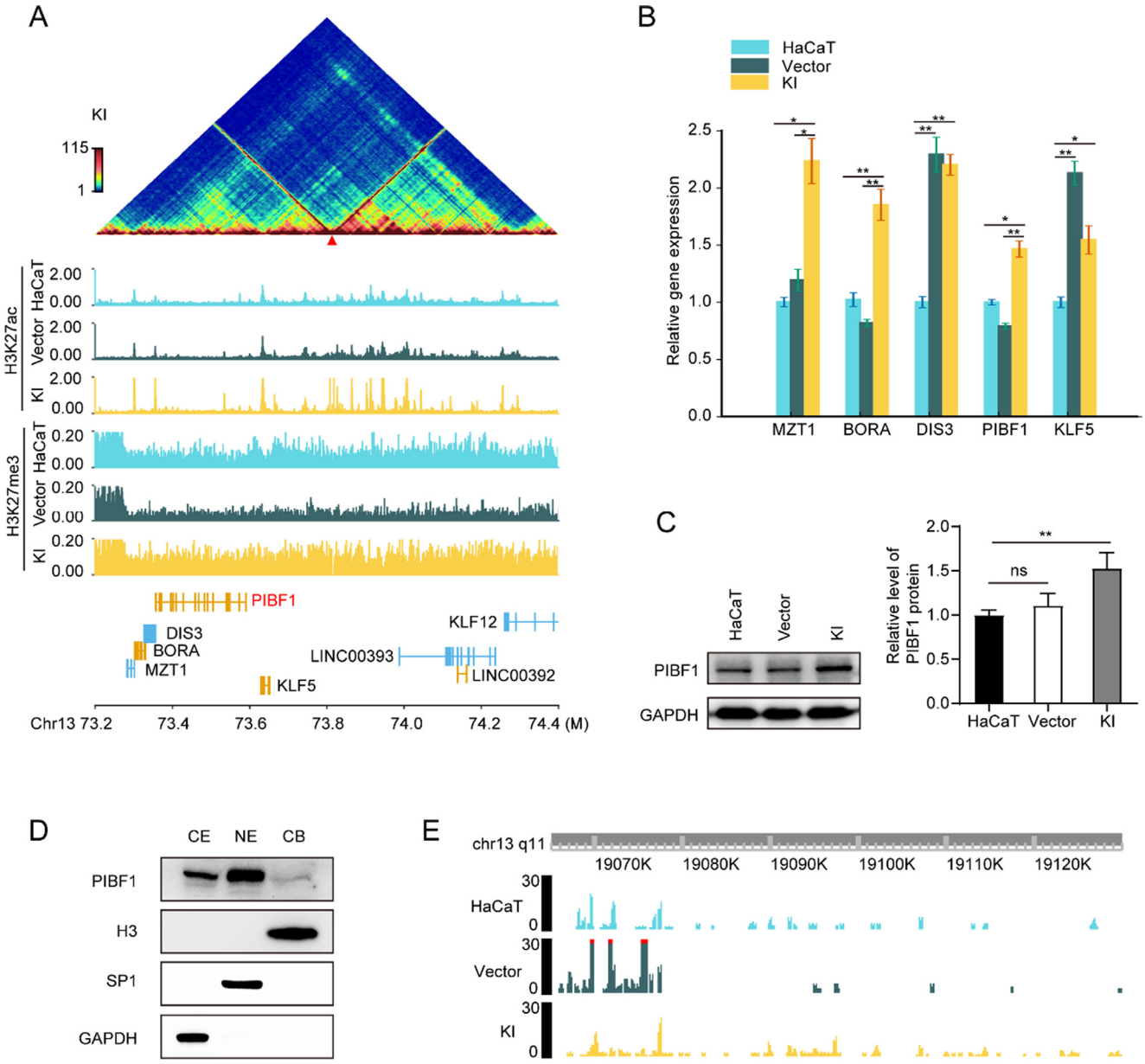


Fig. 4. PIBF1 is a DNA-binding protein and upregulated in KI cells. (A) Hi-C contact heatmap at HPV16 gene knock-in site neighboring regions (Chr13:73.2–74.4 Mb, resolution: 20 kb) as well as matching tracks of the ChIP-seq profile. The gene track is shown in the bottom panel. The red arrow indicates the integration site. (B) The relative mRNA expression of genes in the chromatin TAD where the knock-in site was located. (C) Representative images (left panel) and quantitative analysis results (right panel) of PIBF1 protein expression in the three groups of cells. (D) The distribution of PIBF1 protein expression in cytoplasmic extract (CE), nuclear extract (NE), and chromatin-bound extract (CB). (E) An example region of PIBF1 CUT-Tag profiles in the three groups of cells. P value: two-tailed *t* test, *P < 0.05, **P < 0.01.

(Fig. 4D), which indicated that PIBF1 could bind to chromatin. Subsequently, we identified 7694, 8823, and 13,134 peaks in HaCaT, Vector, and KI cells, respectively, through PBF1 CUT-Tag profiles, which further confirmed the DNA binding function of PIBF1; the peak track of an example region is shown in Fig. 4E. Annotation of these peaks showed that approximately 88 % of the PIBF1 bound to intronic and intergenic regions, while approximately 6 % bound to promoter regions (Fig. S4A). Among all the binding sites observed within the context of all human genes, a weak PIBF1 signal was observed at the promoter region (Fig. S4B). Taken together, our data demonstrate that PIBF1 is a DNA-binding protein and upregulated in KI cells.

PIBF1 interacts with the cohesin complex off chromatin

The PIBF1 protein is predicted to have extensive structural maintenance of chromosome (SMC) structural domains. Two SMC proteins, SMC1 and SMC3, are chromatin structural proteins that interact with RAD21 to form cohesin complexes, which participate in chromatin loop formation [41]. Therefore, we speculated that PIBF1 might interact with SMC3/RAD21. To verify the above hypothesis, we performed a coimmunoprecipitation assay with anti-PIBF1 antibody to identify endogenous interacting proteins. This assay revealed a strong interaction between PIBF1 and SMC3/RAD21 (Fig. 5A). The mass spectrometry spectrum of the PIBF1 immunoprecipitation product showed SMC3 and RAD21 signals (Fig. S5A). Immunofluorescence colocalization assays also showed that PIBF1 colocalized with SMC3/RAD21 in the nucleus (Fig. 5B). Together, the above results indicate that PIBF1 interacts with the cohesin complex. Furthermore, we speculated that PIBF1 might participate in the formation of chromatin loops and CDBs. However, the PIBF1 CUT-Tag profile and SMC3/RAD21 ChIP-seq profiles showed that the PIBF1 peaks did not colocalize with SMC3 and RAD21 peaks genome-wide (Fig. 5C). Moreover, the PIBF1 CUT-Tag signal was not enriched at the CDBs or loop anchors (Fig. 5D, E). Hence, PIBF1 binding sites within the DNA did not correlate with the chromatin loop anchors or CDBs. We conclude that PIBF1 interacts with the cohesin complex off chromatin.

A model for PIBF1 and cohesin complex interaction and the associated chromatin domain disappearance

Based on the above results, we propose that the interaction between PIBF1 and the cohesin complex in KI cells may disrupt its ring-shaped structure such that can no longer participate in forming the chromatin loop anchor and CDB through the extrusion model (Fig. 6A, B). The disappearance of loops/CDBs in certain genome regions may be associated with PIBF1 mediated disruption of the cohesin complex. We summarized the possible effects of chromatin loop disappearance on gene expression. As shown in Fig. 6A, B, PIBF1 interacts with RAD21 or SMC1 and then disrupts the ring-shaped cohesin structure, influencing chromatin loop formation and gene expression. The disappearance of the enhancer-promoter loop would lead to downregulated expression of the cancer suppressor gene after the loss of enhancer regulation. However, the disappearance of the repressive loop would activate specific

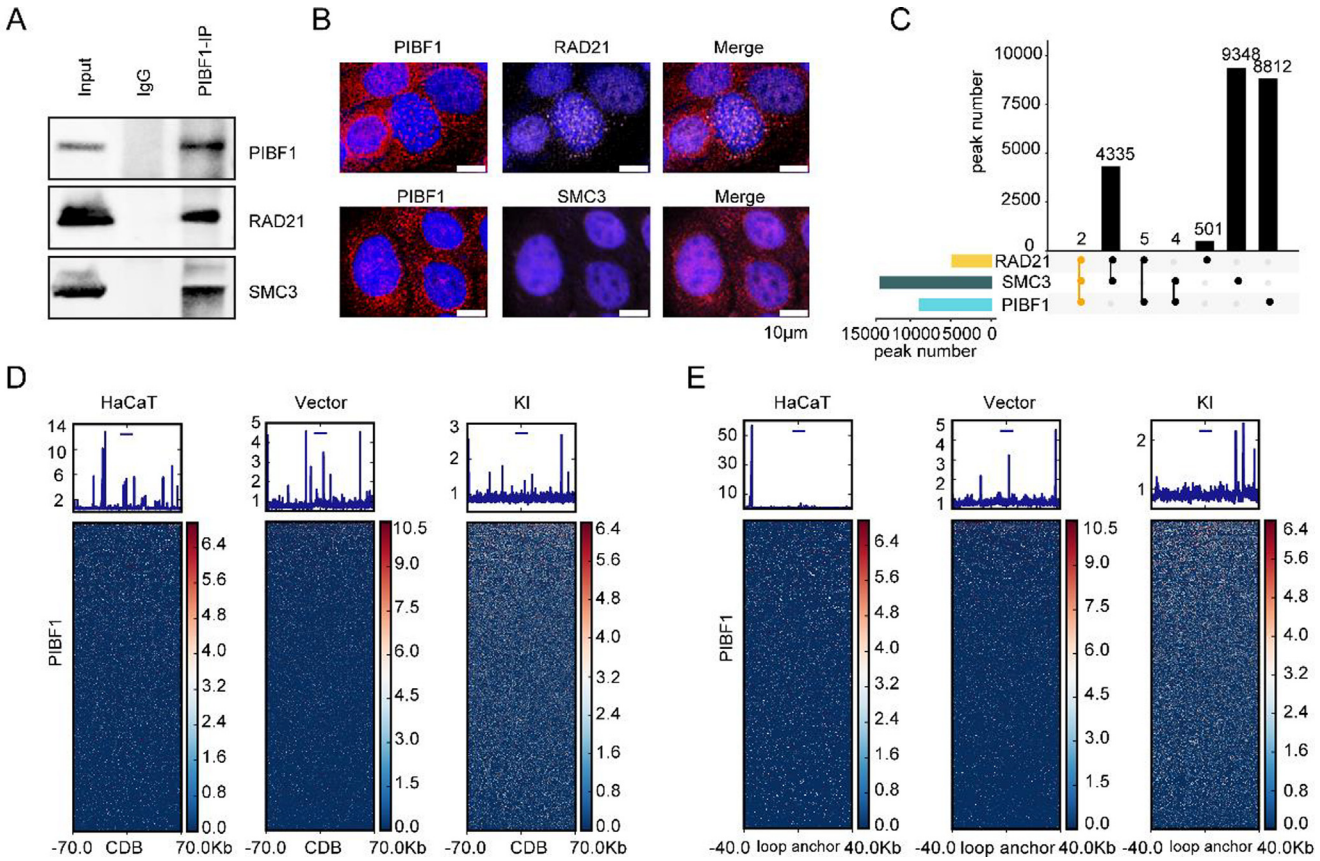
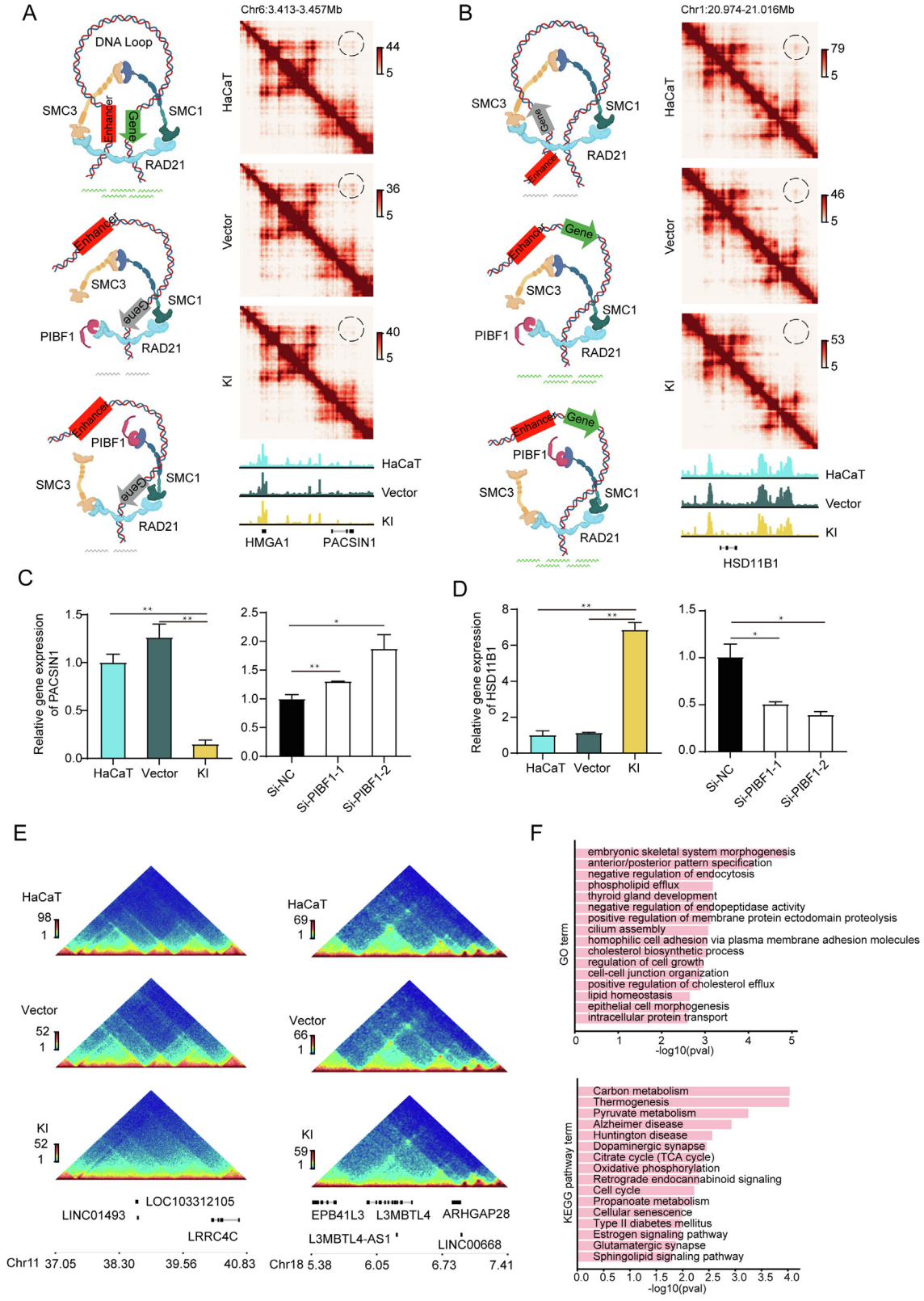


Fig. 5. PIBF1 interacts with the cohesin complex off chromatin. (A) PIBF1, RAD21, and SMC3 were detected by western blot in cell lysates(input) and after PIBF1 immunoprecipitation (IP) experiments, as indicated. (B) Immunofluorescence with PIBF1 (red), RAD21 (pink), SMC3 (pink), and cell nucleus (DAPI) and merged images. Scale bar: 10 μm. (C) The columnar Venn diagram showed the peak numbers of RAD21, SMC3 and PIBF1 identified by ChIP-seq and CUT-Tag profiling. The bars in the bottom left represent all peak numbers of each protein. Dots with connected lines at the bottom represents cobinding peaks of the correspondence proteins. Single dots represent peaks specific to each protein. (D, E) Quantification analysis and heatmap of PIBF1 CUT-Tag enrichment at CDBs (D) or loop anchors (E) in the three groups of cells.

oncogenes with upregulated expression after gaining the nearby enhancer. From the 305 loops with reduced long-range interactions in KI cells, we identified E-P loops and suppressive loops. Examples of E-P loops found at chr6 and chr18 that exhibit a reduction in long-range interactions in KI cells are shown in Fig. 6A and Fig. S6C.

Associated with these loops, the expression of the cancer suppressor genes PACSIN1 and L3MBTL4 was downregulated in KI cells compared to control cells (Fig. 6C and Fig. S6E). In addition, examples of repressive loops found at chr1 and chr11 that exhibit a reduction in long-range interactions in KI cells are shown in Fig. 6B and Fig. S6D; concomitantly, the expression of the oncoge-



nes HSD11B1 and OR5B21 was upregulated in KI cells compared to control cells (Fig. 6D and Fig. S6F). To verify the correlation between PIBF1 overexpression and the dysregulation of gene expression, we transfected KI cells with PIBF1 siRNA. The results showed that the expression of the cancer suppressor genes PACSIN1 and L3MBTL4 was upregulated (Fig. 6C, Fig. S6E), and the expression of the oncogenes HSD11B1 and OR5B21 was downregulated (Fig. 6D, Fig. S6F). These data indirectly suggest a potential effect of PIBF1 on chromatin loop disruption. In addition, we identified weakened CDBs in KI cells compared to control cells (Fig. S7A-B). Examples of Hi-C interaction heatmaps showed the disappearance of CDBs in KI cells (Fig. 6E, Fig. S7C). The functional analysis of genes showed that CDB remodeling-related genes in KI cells were enriched in several cancer-related pathways, such as cell cycle and regulation of cell growth. In summary, the strong interaction of PIBF1 with the cohesin complex may disrupt its ring-shaped structure and impede the formation of chromatin loops and CDBs, resulting in the dysregulation of cancer-related gene expression.

Knockdown of PIBF1 inhibits the proliferation, invasion, and migration of KI cells

To further illustrate the effect of PIBF1 on KI cell proliferation, invasion, and migration, siRNA targeting PIBF1 was transfected into KI cells to investigate whether knockdown of PIBF1 could reverse the malignant biological behavior of KI cells. RT-qPCR and western blotting results showed that PIBF1 was significantly knocked down after siRNA treatment (Fig. 7A, B). Next, CCK8, Transwell, and clone formation assays were performed. As shown in Fig. 7C, the KI cells grew more slowly than the control cells after knockdown of PIBF1. Then, Transwell invasion/migration and wound healing assays were performed to test cell invasion and migration ability. The results showed that knockdown of PIBF1 apparently inhibited KI cell invasion and migration ability (Fig. 7D, E). In addition, knockdown of PIBF1 significantly decreased the number of colonies formed (Fig. 7F). These data indicated the vital role of PIBF1 in the malignant phenotypic changes in KI cells.

Downregulated PIBF1 inhibits the tumor growth of cervical cancer cells

The HPV16 integration event at 13q22 site has been reported in SiHa cell line [5,42]. However, the potential association and implications of HPV integration at PIBF1 was unidentified. In this study, we found that the PIBF1 expression in SiHa was higher than that in Ect1/E6E7 (Fig. 8(A, B)). To further demonstrate the role of PIBF1 in promoting tumorigenesis in cervical cancer after HPV integration, as shown in Fig. 8C, we transfected SiHa cervical cancer cells with lv-shPIBF1 to knockdown PIBF1. Compared with the control cells transfected with lv-shCON, the PIBF1 expression decreased in SiHa-lv-shPIBF1 cells (Fig. 8 (D, E)). The constructed SiHa cells were then injected subcutaneously into the posterior hips of nude mice. We measured the volume of the tumors every five days and euthanized the mice until day 30. Compared to the control tumors transfected with lv-shCON, the tumors transfected with lv-shPIBF1 showed slower growth in PIBF1 downregulated condition (Fig. 8 (F-H)). Furthermore, the IHC assays demonstrated that the expression levels of Ki67 decreased in the tumors transfected with the lv-shPIBF1 lentivirus (Fig. 8 (I, J)). The above results suggested that PIBF1 promoted the tumor growth of cervical cancer.

Discussion

The frequency of HPV integration was covered to increase as CIN progresses to cervical cancer gradually, and it was considered an essential indicator of disease assessment [43,44]. Some researchers believed that HPV integration events give cells a selective growth advantage and promote oncogenic progression [45]. Previous studies of cervical cancer cell lines and clinical samples noted that some regions of the human genome had been repeatedly targeted by HPV integration [46-49]. However, the cause and consequence of integration remains needs further investigation to benefit our understanding. As these authors are known, there is no stable HPV16 genes target integrated cell model up to now to decipher the initial stages of phenotypic progression to cervical cancer. The main reason for this problem may be the difficulty and uncontrollability of HPV gene knock-in. With the help of the CRISPR-Cas9 approach and homologous terminal repair mechanism, our study here established an original cell model with the HPV16 gene integrated at 13q22 site for the first time. The PCR-sequencing and WGS confirmed the precise HPV16 gene integration at 13q22 site, and the sequencing data has not shown integration in off-target sites. The integrity of chromosomes showed by Karyotype analysis has avoided the effects of chromosome structural variations other than HPV16 gene integration. A study has demonstrated that the HPV16 gene integration event has contributed to the malignant progression in W12 cells[50]. The growth advantages of KI cells well-established here, shown by our results in vivo and in vitro, also highlighted the importance of HPV16 gene hotspot integration.

Genomic changes such as duplication, insertion, and deletion have altered chromatin 3D structure, resulting in gene expression dysregulation in disease. For instance, genomic tandem duplication disrupted TAD boundaries in gastric cancer, resulting in oncogene IGF2 overexpression after acquiring enhancer CCNE1[51]. Researchers have also shown that HERV-H Retrotransposons inserted in the human genome can mediate TAD boundaries forming and potentially affect gene transcription regulation[52]. In our study, 3D genome architecture is globally unperturbed after HPV integration, but we do observe local modulations on A/B compartments, TADs, and loops near the integration site. We became aware of a study by Christian Paris et al. in which CTCF binding sites were identified in the HPV gene region[53]. Thus, it could explain the formation of TAD boundary at HPV integration site here, consistent with a previously reported TAD change after HPV integrates at another site[9]. Therefore, the integration of HPV could directly change interaction frequencies on both sides of the integration site and impact gene expression. More interestingly, previous studies have found that HPV integration appears more readily to target

Fig. 6. A model of PIBF1 and cohesin complex interaction and the associated domain disappearance. (A) Scheme illustrating the interaction between PIBF1 and the cohesin complex leading to the disruption of the ring-shaped cohesin complex structure, which in turn leads to the disappearance of enhancer-promoter loop (left panel). Example Hi-C contact heatmap at genome region chr6:34130000–34570000 in three groups of cells followed by the H3K27ac ChIP-seq profile and gene tracks, which indicated a weaker enhancer-PACSIN1 interaction loop in KI cells compared to control cells (right panel). The dashed circles indicate the region with an enhancer-PACSIN1 loop. (B) Scheme illustrating that the interaction between PIBF1 and the cohesin complex leads to the disruption of the ring-shaped cohesin complex structure, which in turn leads to the disappearance of the suppressive loop and results in enhancer acquisition at oncogene (left panel). Example Hi-C contact heatmap at genome region chr1:209740000–210160000 in three groups of cells followed by the H3K27ac ChIP-seq profile and gene tracks, which indicated a weaker chromatin interaction loop of HSD11B1 (right panel). The dashed circles indicate the region with the HSD11B1 loop. (C, D) The relative expression of PACSIN1 and HSD11B1 in different groups of cells. KI cells were transfected with small interfering RNA targeting PIBF1 (si-PIBF1-1 and si-PIBF1-2) or nontargeting control (si-NC). (E) Representative Hi-C contact heatmap at regions chr11:37.05–40.83 Mb and Chr18: 5.83–7.41 Mb displaying the disappearance of CDBs in KI cells. (F) GO and KEGG analysis for genes related to specific CDBs in KI cells. P value: two-tailed *t* test, $^{*}P < 0.05$, $^{**}P < 0.01$.

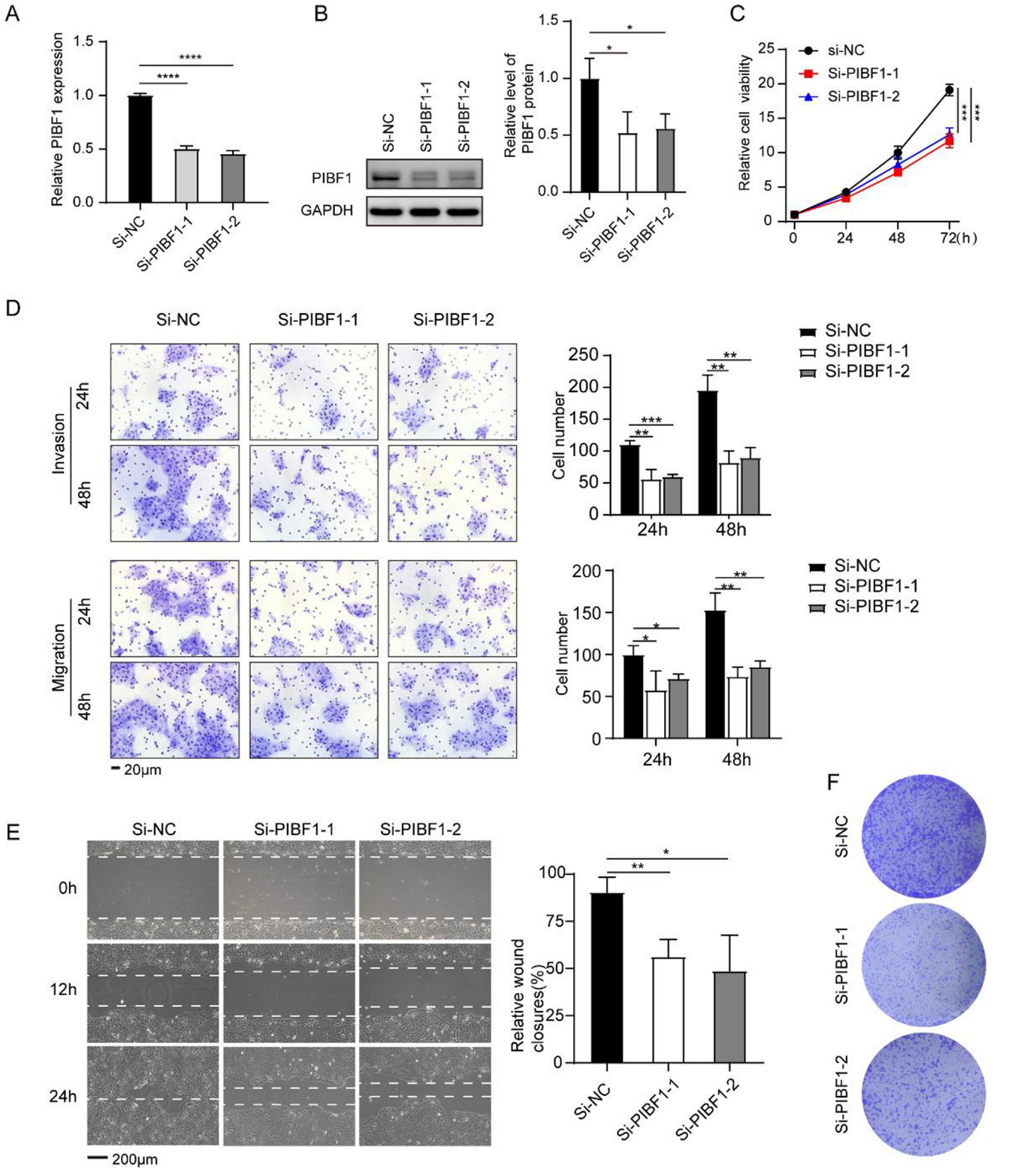


Fig. 7. Knockdown of PIBF1 inhibits the proliferation, invasion, and migration of KI cells. (A) The relative mRNA expression of PIBF1 in KI cells after siRNA treatment. (B) Representative images (left panel) and quantitative analysis results (right panel) of PIBF1 protein expression in KI cells after siRNA treatment. (C) Viability of KI cells at 0 h, 24 h, 48 h, and 72 h after siRNA treatment. (D) Representative images of Transwell assays at 24 h and 48 h in KI cells after siRNA treatment, and the corresponding quantitative analysis results. Scale bar: 20 μm. (E) Representative scratch assay images at 0 h, 12 h, and 24 h in KI cells after siRNA treatment and computed results of the relative scratch healing rates. Computed migration front lines are shown (white). Scale bar: 200 μm. (F) Representative images of the colony-forming assay in KI cells after siRNA treatment. P value: two-tailed *t* test, *P < 0.05, **P < 0.01. ***P < 0.001, ****P < 0.0001.

at transcriptional active regions[46,54-56], which may affect the expression of viral oncogenes and target genes. We have also shown in our KI and control cells that the active histone modification (H3K27ac) was highly enriched at the genome region around the 13q22 integration site. All the above epigenetic variations were contributed to the gene expression changes, such as PIBF1. A new

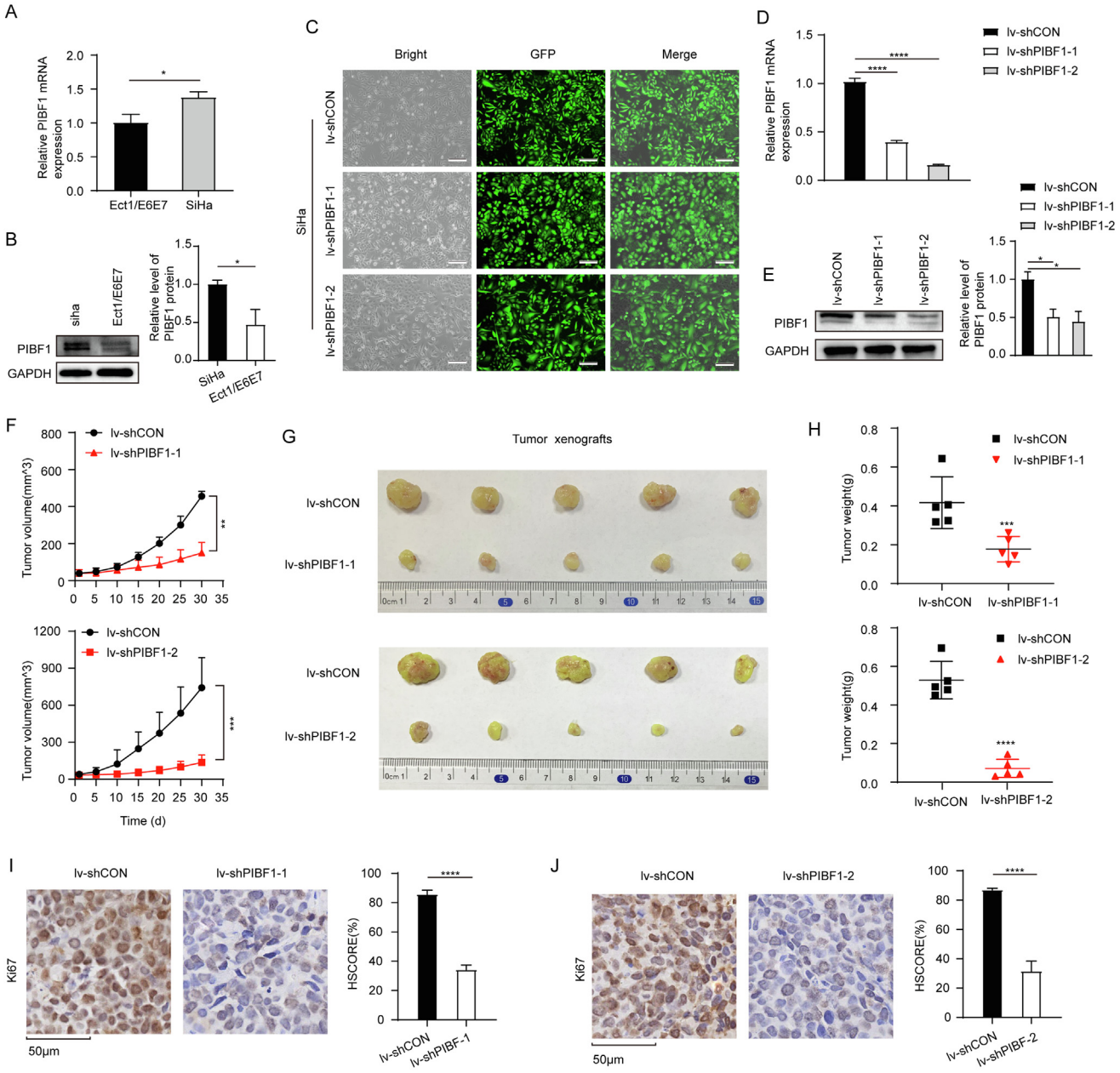


Fig. 8. PIBF1 Knockdown inhibits cervical cancer cells growth. (A, B) The PIBF1 expression was detected in SiHa and Ect1/E6E7 cells. The quantitative analysis results of PIBF1 protein expression were shown (B). (C) Representative GFP fluorescence images of SiHa cells infected with lentivirus. (D, E) PIBF1 expression was detected in SiHa cells infected with lentivirus. The quantitative analysis results of PIBF1 protein expression were shown (E). (F-H) For the in vivo analyses, 5×10^6 SiHa cells infected with lentivirus were injected subcutaneously into the posterior hip of nude mice. F) Tumor growth was recorded every five days until day 30. Error bars represent the SD of the mean (n = 5). J) Image of the tumor specimens after harvested. H) Tumor weight was measured after harvested. Error bars represent the SD of the mean (n = 5). (I, J) Immunohistochemical analysis of Ki67 expression in xenograft tumor tissues, and Ki67 protein expression was evaluated using HSCORE. Each group included five nude mice. The data shown represented mean ± SD (Student's *t*-test). P value was denoted as *P < 0.05, **P < 0.01, and ***P < 0.001, "ns" represents "not significant".

HPV16 knock-in cellular model with promoter nonsense mutation should be constructed in the future to further illustrate whether the proximity of the HPV16 promoter affecting the gene expression changes.

PIBF1, located at 13q22 locus, plays a role in normal pregnancy and tumor progression[57]. It has been noted that PIBF1 protein has ~90% structural maintenance of chromosomes (SMC) domains and interacts with some SMC proteins such as KIAA1328[58]. KIAA1328 was reported to interact with another SMC protein, SMC3[59], one of the three core subunits of the chromatin cohesin complex. Consistent with what researchers suppose, our data has demonstrated through immunofluorescence and immunoprecipitation assay, that PIBF1 interacted with cohesin complex. Chromatin in TAD and loop were important genome 3D structures, and loop extrusion by the cohesin complex contributed to TADs and chromatin loops[60,61]. However, the distribution of PIBF1 binding sites at DNA was inconsistent with SMC3/RAD21, indicating PIBF1 has a competitive combination role rather than a constitutive protein in cohesin construction. The expression of PIBF1 in our model is up-regulated. It may affect the formation of chromatin loops and CDBs by disrupting the ring-like cohesin complex, thereby impacting gene expression, indirectly supported by the differentially expressed genes in the vicinity of loci with altered interaction frequencies of loops and CDBs. For example, chromatin loops inter-

action frequencies change at chr6: 34130000–34570000 and chr1: 209740000–210160000 resulted in cancer suppressor PACSIN1[62] and oncogene HSD11B1[63,64] expression dysregulation, and they were also directly regulated by PIBF1. Phenotypically, PIBF1 overexpression was related to poor prognosis and promoted cells proliferation and colony formation by inhibiting Akt signaling in gastric cancer[65]. Knockdown of PIBF1 could decrease primary lung carcinoma and ovarian cells invasion[66]. It also acted as an oncogene in breast cancer [67]. Hence, the diverse phenotypes reported could result from pleiotropic effects triggered by local perturbations of 3D chromatin architecture when PIBF1 was overexpressed. On the other hand, the application of PIBF1 siRNA succeeded in reversing the malignant biological behavior of integrative cells. Downregulated PIBF1 also inhibits the tumor growth of cervical cancer cells in vivo. Thus, it is convinced that the overexpressed PIBF1 contributed a lot to the malignant transformation of host cells upon HPV16 genes integration.

In conclusion, our study shows that genome 3D structure and epigenome alteration near the 13q22 integration site leads to PIBF1 overexpression, which likely contributes to genome-wide chromatin loops and CDBs change and gene regulation, hence unraveling the relationship between HPV integration and the biological behavior change of host cells. It provided a solid basis for the theory of HPV integrated carcinogenesis. However, further research is needed to verify the HPV integration-induced carcinogenesis, including cell tumorigenesis, the effects of integration byproducts on cells, the dynamic process of genome 3D structure variation mediated by PIBF1, and the role of other elevated expression genes near the HPV integration site. We also need to establish more cell models with HPV integrated at different sites. HPV18 is another important high-risk HPV type. Previous studies have proved that HPV integration is observed in all HPV18-related cervical cancers, and cancers with HPV18 integration are more aggressive which threaten women' life and health seriously. Subsequently, studies have found that HPV18 is more likely to integrate into the fragile site 8q24 region than HPV16. Therefore, an 8q24 site-specific cell model that maintains the restricted cell differentiation and phenotypic characteristics of the epithelial cell layer is also urgently needed. Besides, the KI cells are at a low passage in this study, and we will perform long-term culture to further interrogate the chromosome instability, which will be a good direction in the next research. And more efforts need to be made because many events and processes may contribute to the development of HPV integration, a strong driver of oncogenesis.

Conclusion

We successfully established a 13q22 site-specific HPV16 gene knock-in cell model. HPV16 gene knock-in at the 13q22 site induced malignant transformation of host cell. HPV integration could upregulate PIBF1 expression. Moreover, the interaction of PIBF1 with the cohesin complex may disrupt its ring-shaped structure and impede the formation of chromatin loops and CDBs, resulting in the dysregulation of cancer-related gene expression. The data reported here provides new insights into HPV integration-driven cervical carcinogenesis.

Ethics statement

We carried out the study in strict accordance with the Guidelines for the Welfare of Animals in Experimental Neoplasia. The experimental protocols were approved by the Institutional Animal Care and Use Committee of HUST (Approval no. TJ-IRB20210207).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This project was supported by funds from the National Key R&D Program of China (2021YFC2701204), the National Natural Science Funding of China (81830074, 81772786, 81902667, 82103161). We thank our colleagues from Annoroad Gene Tech. (Beijing) Co., Ltd. for sequencing and analysis. We thank for all members of our work team for helpful discussions.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jare.2023.04.015.

References

[1] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin 2021;71(3):209–49.
[2] Moody CA, Laimins LA. Human papillomavirus oncoproteins: pathways to transformation. Nat Rev Cancer 2010;10(8):550–60.
[3] Liu L, Ying C, Zhao Z, Sui L, Zhang X, Qian C, et al. Identification of reliable biomarkers of human papillomavirus 16 methylation in cervical lesions based on integration status using high-resolution melting analysis. Clin Epigenetics 2018;10:10.
[4] Integrated genomic and molecular characterization of cervical cancer. Nature 2017;543(7645):378–84.
[5] Hu Z, Zhu D, Wang W, Li W, Jia W, Zeng X, et al. Genome-wide profiling of HPV integration in cervical cancer identifies clustered genomic hot spots and a potential microhomology-mediated integration mechanism. Nat Genet 2015;47(2):158–63.
[6] Ojesina AI, Lichtenstein L, Freeman SS, Pedamallu CS, Imaz-Rosshandler I, Pugh TJ, et al. Landscape of genomic alterations in cervical carcinomas. Nature 2014;506(7488):371–5.
[7] Zapatka M, Borozan I, Brewer DS, Iskar M, Grundhoff A, Alawi M, et al. The landscape of viral associations in human cancers. Nat Genet 2020;52(3):320–30.
[8] Spielmann M, Lupianez DG, Mundlos S. Structural variation in the 3D genome. Nat Rev Genet 2018;19(7):453–67.
[9] Cao C, Hong P, Huang X, Lin D, Cao G, Wang L, et al. HPV-CCDC106 integration alters local chromosome architecture and hijacks an enhancer by three-dimensional genome structure remodeling in cervical cancer. J Genet Genomics 2020;47(8):437–50.
[10] Groves IJ, Drane ELA, Michalski M, Monahan JM, Scarpini CG, Smith SP, et al. Short- and long-range cis interactions between integrated HPV genomes and cellular chromatin dysregulate host gene expression in early cervical carcinogenesis. PLoS Pathog 2021;17(8):e1009875.
[11] Adeel MM, Jiang H, Arega Y, Cao K, Lin D, Cao C, et al. Structural Variations of the 3D Genome Architecture in Cervical Cancer Development. Front Cell Dev Biol 2021;9:706375.
[12] Wentzensen N, Vinokurova S, von Knebel DM. Systematic review of genomic integration sites of human papillomavirus genomes in epithelial dysplasia and invasive cancer of the female lower genital tract. Cancer Res 2004;64(11):3878–84.
[13] Groves IJ, Coleman N. Human papillomavirus genome integration in squamous carcinogenesis: what have next-generation sequencing studies taught us? J Pathol 2018;245(1):9–18.
[14] Kenny HA, Lal-Nag M, White EA, Shen M, Chiang CY, Mitra AK, et al. Quantitative high throughput screening using a primary human three-dimensional organotypic culture predicts in vivo efficacy. Nat Commun 2015;6:6220.
[15] Wang J, Wang D, Yin Y, Deng Y, Ye M, Wei P, et al. Assessment of Combined Karyotype Analysis and Chromosome Microarray Analysis in Prenatal Diagnosis: A Cohort Study of 3710 Pregnancies. Genet Res (Camb) 2022;2022:6791439.
[16] Wan W, Wang L, Chen X, Zhu S, Shang W, Xiao G, et al. A Subcellular Quantitative Proteomic Analysis of Herpes Simplex Virus Type 1-Infected HEK 293T Cells. Molecules 2019;24(23).
[17] Sedlazeck FJ, Rescheneder P, Smolka M, Fang H, Nattestad M, von Haeseler A, et al. Accurate detection of complex structural variations using single-molecule sequencing. Nat Methods 2018;15(6):461–8.

[18] Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 2009;326(5950):289–93.
[19] Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat methods 2012;9(4):357–9.
[20] Servant N, Varoquaux N, Lajoie BR, Viara E, Chen CJ, Vert JP, et al. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. Genome Biol 2015;16:259.
[21] Chen F, Li G, Zhang MQ, Chen Y. HiCDB: a sensitive and robust method for detecting contact domain boundaries. Nucleic Acids Res 2018;46(21):11239–50.
[22] Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 2010;26(6):841–2.
[23] Durand NC, Shamim MS, Machol I, Rao SS, Huntley MH, Lander ES, et al. Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. Cell Syst 2016;3(1):95–8.
[24] Sirén J, Välimäki N, Mäkinen V. Indexing Graphs for Path Queries with Applications in Genome Research. IEEE/ACM Trans Comput Biol Bioinform 2014;11(2):375–88.
[25] Hart T, Komori HK, LaMere S, Podshivalova K, Salomon DR. Finding the active genes in deep RNA-seq gene expression studies. BMC Genomics 2013;14:778.
[26] Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15(12):550.
[27] Landt SG, Marinov GK, Kundaje A, Kheradpour P, Pauli F, Batzoglou S, et al. ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. Genome Res 2012;22(9):1813–31.
[28] Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009;25(16):2078–9.
[29] Ramírez F, Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, et al. deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res 2016;44(W1):W160–5.
[30] Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol 2008;9(9):R137.
[31] Guo Q, Song WD, Li HY, Zhou YP, Li M, Chen XK, et al. Scored minor criteria for severe community-acquired pneumonia predicted better. Respir Res 2019;20(1):22.
[32] Kaya-Okur HS, Wu SJ, Codomo CA, Pledger ES, Bryson TD, Henikoff JG, et al. CUT&Tag for efficient epigenomic profiling of small samples and single cells. Nat Commun 2019;10(1):1930.
[33] Harris MA, Clark J, Ireland A, Lomax J, Ashburner M, Foulger R, et al: The Gene Ontology (GO) database and informatics resource. Nucleic Acids Res 2004, 32 (Database issue):D258-261.
[34] Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res 1999;27(1):29–34.
[35] Dickinson DJ, Ward JD, Reiner DJ, Goldstein B. Engineering the Caenorhabditis elegans genome using Cas9-triggered homologous recombination. Nat Methods 2013;10(10):1028–34.
[36] Paquet D, Kwart D, Chen A, Sproul A, Jacob S, Teo S, et al. Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. Nature 2016;533(7601):125–9.
[37] Ledford H. CRISPR gene editing in human embryos wreaks chromosomal mayhem. Nature 2020;583(7814):17–8.
[38] Okabe A, Huang KK, Matsusaka K, Fukuyo M, Xing M, Ong X, et al. Cross-species chromatin interactions drive transcriptional rewiring in Epstein-Barr virus-positive gastric adenocarcinoma. Nat Genet 2020;52(9):919–30.
[39] Chen W, Zhang J, Fu H, Hou X, Su Q, He Y, et al. KLF5 Is Activated by Gene Amplification in Gastric Cancer and Is Essential for Gastric Cell Proliferation. Cells 2021;10(5).
[40] Ott T, Kaufmann L, Granzow M, Hinderhofer K, Bartram CR, Theiß S, et al. The Frog Xenopus as a Model to Study Joubert Syndrome: The Case of a Human Patient With Compound Heterozygous Variants in PIBF1. Front Physiol 2019;10:134.
[41] Zhang L, Xue G, Liu J, Li Q, Wang Y. Revealing transcription factor and histone modification co-localization and dynamics across cell lines by integrating ChIP-seq and RNA-seq data. BMC Genomics 2018;19(Suppl 10):914.
[42] el Awady M, Kaplan J, O'Brien S, Burk RJV. Molecular analysis of integrated human papillomavirus 16 sequences in the cervical cancer cell line SiHa. Virology 1987;159(2):389–98.
[43] Hudelist G, Manavi M, Pischinger K, Watkins-Riedel T, Singer C, Kubista E. Czerwenka KJGo: Physical state and expression of HPV DNA in benign and dysplastic cervical tissue: different levels of viral integration are correlated with lesion grade. Gynecol Oncol 2004;92(3):873–80.
[44] Arias-Pulido H, Peyton C, Joste N, Vargas H. Wheeler CJJocm: Human papillomavirus type 16 integration in cervical carcinoma in situ and in invasive cervical cancer. J Clin Microbiol 2006;44(5):1755–62.
[45] Jeon S, Allen-Hoffmann B. Lambert PJJov: Integration of human papillomavirus type 16 into the human genome correlates with a selective growth advantage of cells. J Virol 1995;69(5):2989–97.
[46] Bodelon C, Untereiner M, Machiela M, Vinokurova S. Wentzensen NJIjoc: Genomic characterization of viral integration sites in HPV-related cancers. Int J Cancer 2016;139(9):2001–11.
[47] Thorland E, Myers S, Persing D, Sarkar G, McGovern R, Gostout B. Smith DJCr: Human papillomavirus type 16 integrations in cervical tumors frequently occur in common fragile sites. Cancer Res 2000;60(21):5916–21.
[48] Dall K, Scarpini C, Roberts I, Winder D, Stanley M, Muralidhar B, et al. Coleman NJCr: Characterization of naturally occurring HPV16 integration sites isolated from cervical keratinocytes under noncompetitive conditions. Cancer Res 2008;68(20):8249–59.
[49] Schmitz M, Driesch C, Beer-Grondke K, Jansen L, Runnebaum I. Dürst MJIjoc: Loss of gene function as a consequence of human papillomavirus DNA integration. Int J Cancer 2012;131(5):E593–602.
[50] Pett M, Alazawi W, Roberts I, Dowen S, Smith D, Stanley M. Coleman NJCr: Acquisition of high-level chromosomal instability is associated with integration of human papillomavirus type 16 in cervical keratinocytes. Cancer Res 2004;64(4):1359–68.
[51] Ooi W, Nargund A, Lim K, Zhang S, Xing M, Mandoli A, et al. CCNE1Integrated paired-end enhancer profiling and whole-genome sequencing reveals recurrent and enhancer hijacking in primary gastric adenocarcinoma. Gut 2020;69(6):1039–52.
[52] Zhang Y, Li T, Preissl S, Amaral M, Grinstein J, Farah E, et al. Transcriptionally active HERV-H retrotransposons demarcate topologically associating domains in human pluripotent stem cells. Nat Genet 2019;51(9):1380–8.
[53] Paris C, Pentland I, Groves I, Roberts D, Powis S, Coleman N, et al. Parish JJJov: CCCTC-binding factor recruitment to the early region of the human papillomavirus 18 genome regulates viral oncogene expression. J Virol 2015;89(9):4770–85.
[54] Christiansen I, Sandve G, Schmitz M, Dürst M. Hovig EJPo: Transcriptionally active regions are the preferred targets for chromosomal HPV integration in cervical carcinogenesis. PLoS One 2015;10(3):e0119566.
[55] Doolittle-Hall J, Cunningham Glasspoole D, Seaman W, Webster-Cyriaque JJC. Meta-Analysis of DNA Tumor-Viral Integration Site Selection Indicates a Role for Repeats. Gene Expression and Epigenetics Cancers (Basel) 2015;7(4):2217–35.
[56] Kelley D, Flam E, Izumchenko E, Danilova L, Wulf H, Guo T, et al. Integrated Analysis of Whole-Genome ChIP-Seq and RNA-Seq Data of Primary Head and Neck Tumor Samples Associates HPV Integration Sites with Open Chromatin Marks. Cancer Res 2017;77(23):6538–50.
[57] Szekeres-Bartho J, Polgar BJAjori: PIBF: the double edged sword. Pregnancy and tumor. Am J Reprod Immunol 2010, 64(2):77-86.
[58] Ott T, Kaufmann L, Granzow M, Hinderhofer K, Bartram C, Theiß S, et al. XenopusThe Frog as a Model to Study Joubert Syndrome: The Case of a Human Patient With Compound Heterozygous Variants in. Front Physiol 2019;10:134.
[59] Gol'dshtein D, Pogorelov A, Chailakhyan T. Smirnov AJBoeb, medicine: Changes in intracellular potassium concentration in a one-cell mouse embryo after enucleation. Bull Exp Biol Med 2004;138(3):243–4.
[60] Bonev B. Cavalli GJNrG: Organization and function of the 3D genome. Nat Rev Genet 2016;17(11):661–78.
[61] Cremer T. Cremer CJNrG: Chromosome territories, nuclear architecture and gene regulation in mammalian cells. Nat Rev Genet 2001;2(4):292–301.
[62] Zimu Z, Jia Z, Xian F, Rui M, Yuting R, Yuan W, et al. Enfang SJFimb: Decreased Expression of PACSIN1 in Brain Glioma Samples Predicts Poor Prognosis. Front Mol Biosci 2021;8:696072.
[63] Hu D, Zhou M. Zhu XJBri: Deciphering Immune-Associated Genes to Predict Survival in Clear Cell Renal Cell Cancer. Biomed Res Int 2019;2019:2506843.
[64] Li C, Liu T, Wang J, Yu S, Chen Y, Fang F, et al. Hydroxysteroid 11-Beta Dehydrogenase 1 Overexpression with Copy-Number Gain and Missense Mutations in Primary Gastrointestinal Stromal Tumors. J Clin Med 2018;7(11).
[65] Chu S, Wang G, Zhang P, Zhang R, Huang Y, Lu Y, et al. Zhu JJJoe et al.: Retraction Note: MicroRNA-203 suppresses gastric cancer growth by targeting PIBF1/Akt signaling. J Exp Clin Cancer Res 2017;36(1):61.
[66] Balassa T, Berta G, Jakab L, Bohonyi N. Szekeres-Bartho JJJori: The effect of the Progesterone-Induced Blocking Factor (PIBF) on E-cadherin expression, cell motility and invasion of primary tumour cell lines. J Reprod Immunol 2018;125:8–15.
[67] Delsite R, Kachhap S, Anbazhagan R, Gabrielson E. Singh KJMc: Nuclear genes involved in mitochondria-to-nucleus communication in breast cancer cells. Mol Cancer 2002;1:6.