



รายงานวิจัยฉบับสมบูรณ์

โครงการ การจับกันของโปรตีนอีสองของไวรัสกับโปรตีนของเซลล์โฮสต์เพื่อยับยั้งโปรโมเตอร์ของยีนก่อมะเร็งของเชื้อไวรัสฮิวแมนแพปิโลมาท่ายป์ 16 ในเซลล์ปากมดลูกที่มีภาวะเป็นจีโนมวงกลมและการแทรกตัวเข้าสู่โครโมโซมของโฮสต์

(Interaction of E2 protein with cellular partner proteins that facilitates transcriptional repressive activity of episomal and integrated human papillomavirus 16 oncogene promoter in cervical cancer cell lines)

โดย ดร. อาคม ไชยวงศ์คต

กันยายน 2562

สัญญาเลขที่ TRG6080008

รายงานวิจัยฉบับสมบูรณ์

โครงการ การจับกันของโปรตีนอีสองของไวรัสกับโปรตีนของเซลล์
โฮสต์เพื่อยับยั้งโปรโมเตอร์ของยีนก่อมะเร็งของเชื้อไวรัสฮิวแมนแพปปี
โลมาทัยป์ 16 ในเซลล์ปากมดลูกที่มีภาวะเป็นจีโนมวงกลมและการ
แทรกตัวเข้าสู่โครโมโซมของโฮสต์

(Interaction of E2 protein with cellular partner proteins
that facilitates transcriptional repressive activity of episomal and
integrated human papillomavirus 16 oncogene promoter in
cervical cancer cell lines)

ดร.อาคม ไชยวงศ์คต

จุฬาลงกรณ์มหาวิทยาลัย

สนับสนุนโดยสำนักงานกองทุนสนับสนุนการวิจัยและต้นสังกัด

(ความเห็นในรายงานนี้เป็นของผู้วิจัย

สกว.และต้นสังกัดไม่จำเป็นต้องเห็นด้วยเสมอไป)

Abstract (บทคัดย่อ)

Project Code : TRG6080008

Project Title : การจับกันของโปรตีนอีสองของไวรัสกับโปรตีนของเซลล์โฮสต์เพื่อยับยั้งโปรโมเตอร์ของยีนก่อมะเร็งของเชื้อไวรัสอีวแมนเปปิโลมาทัยป์ 16 ในเซลล์ปากมดลูกที่มีภาวะเป็นจีโนมวงกลมและการแทรกตัวเข้าสู่โครโมโซมของโฮสต์

(Interaction of E2 protein with cellular partner proteins that facilitates transcriptional repressive activity of episomal and integrated human papillomavirus 16 oncogene promoter in cervical cancer cell lines)

Investigator : ดร. อาคม ไชยวงศ์คต

E-mail Address : arkomchaiwongkot@gmail.com

Project Period : 2 ปี

Abstract

Background: Upregulated expression of high risk human papillomaviruses (HR-HPVs) E6 and E7 oncogenes is associated with cervical carcinogenesis and it has been shown that E2 protein can either activate or repress HPV oncogenes expression.

Objectives: We aimed to detect the effect of E2 protein on host repressor proteins expression that may facilitate the E2 mediated transcriptional repressive activity of human papillomavirus 16 oncogene promoters and the interaction of HPV16 E2 protein and host partner protein. In addition, RNAseq analysis was performed using RNA extracted from five cervical cancer cell lines containing HPV16, 18, 45 and 68b (CaSki, SiHa, HeLa, MS751 and ME180) and HPV negative cervical cancer cell line (C33A) to investigate the transcription patterns of viral and cellular genes.

Methods: Transfection experiment on CaSki and SiHa using plasmid containing HPV16 E2 gene was performed. Viral and host gene expression were detected by real time PCR and western blot.

Results: The major transcripts obtained from four HR-HPV types 16, 18, 45 and 68b were E6*1 with splicing event occur within E6 gene fused with full length E7 ORF, 226^409 (HPV16), 233^416 (HPV18), 230^412 (HPV45) and 129^311(HPV68b) and were associated with full length E7 oncogene expression. CaSki and SiHa cell lines transfected with HPV16E2 showed reduction of full length E7 oncogenes. HDAC1 was found to be upregulated in HPV16 E2 expressing CaSki but not SiHa.

Conclusion: Splicing pattern that occurred within E6 gene resulting in E6*1 transcript facilitates E7 oncogene translation. Thus, HPV major transcript E6*1 and full length E7 oncogenes expressed in HR-HPVs active cervical cancer cell lines could be used as biomarker and targeted therapy. HDAC1 may be involved in HPV16E2 mediated viral oncogene transcripts down expression in CaSki and SiHa cell lines. Further study

should be performed to investigate the viral-host proteins interaction and their effect on viral oncogenes expression..

Keywords: Human papillomavirus16 E2; transcription patterns; host gene expression

บทคัดย่อ

ความรู้เดิม: การแสดงออกของยีนก่อมะเร็งอีหกและอีเจ็ดของไวรัสแปปิลโลมาที่มีความเสี่ยงสูงมีความสัมพันธ์กับการเป็นมะเร็งปากมดลูก และจากการศึกษาพบว่าโปรตีนอีสองสามารถกระตุ้นหรือยับยั้งการแสดงออกของยีนก่อมะเร็งของไวรัสแปปิลโลมาได้

วัตถุประสงค์: เราจะทำการศึกษาผลของโปรตีนอีสองต่อการแสดงออกของโปรตีนของคนที่เกี่ยวข้องกับการยับยั้งการแสดงออกของยีนซึ่งโปรตีนของคนนี้อาจจะมีส่วนร่วมกับโปรตีนอีสองในการยับยั้งการแสดงออกของยีนก่อมะเร็งของไวรัสโดยไปจับตรงโปรโมเตอร์ของยีนก่อมะเร็งของไวรัส และยังศึกษาการจับกันระหว่างโปรตีนอีสองกับโปรตีนของคนด้วย และเราได้ศึกษาเพิ่มเติมโดยทำ RNAseq เพื่อดูรูปแบบการแสดงออกของยีนไวรัสและยีนของคนในเซลล์มะเร็งปากมดลูกที่มีเชื้อไวรัสชนิดต่างๆ ได้แก่ 16, 18, 45 and 68b และไม่มีเชื้อไวรัส

วิธีการศึกษา: นำพลาสมิดที่มียีนอีสองใส่ลงไปในเซลล์เพาะเลี้ยงมะเร็งปากมดลูกที่ติดเชื้อไวรัสแปปิลโลมาชนิดลบหห ได้แก่ CaSki และ SiHa แล้วนำมาสกัดอาร์เอ็นเอเพื่อศึกษาการแสดงออกของยีนของไวรัสและยีนของคน

ผลการศึกษา: เส้นอาร์เอ็นเอของไวรัสแปปิลโลมาที่มีความเสี่ยงสูงชนิด 16, 18, 45 และ 68b ที่มีการแสดงออกเยอะในเซลล์มะเร็งปากมดลูกได้แก่เส้นอาร์เอ็นเอที่เรียกว่า E6*1 ซึ่งเป็นเส้นอาร์เอ็นเอที่มีการขาดหายไปของยีน E6 บางส่วนพร้อมยีน E7 แบบเต็มสายได้แก่ 226⁴⁰⁹ (HPV16), 233⁴¹⁶ (HPV18), 230⁴¹² (HPV45) และ 129³¹¹(HPV68b) และเมื่อมีการกระตุ้นให้มีการแสดงออกของโปรตีน E2 ในเซลล์มะเร็ง CaSki และ SiHa พบว่ามีการแสดงออกของยีน E7 ที่ลดลง และพบว่ายีนของคน HDAC1 มีการแสดงออกที่เพิ่มขึ้นในเซลล์ CaSki อย่างเดียว

สรุปผลการศึกษา: รูปแบบการแสดงออกของเส้นอาร์เอ็นเอที่เรียกว่า E6*1 มีความเกี่ยวข้องกับการแสดงออกของยีน E7 ที่สูงขึ้น ดังนั้นเราสามารถใช้ในการตรวจ E6*1 และยีน E7 ในการเป็นตัวบ่งชี้ทางชีวภาพได้และใช้เป็นยีนเป้าหมายในการรักษา และโปรตีนของคน HDAC1 อาจจะมีส่วนร่วมกับโปรตีนอีสองในการไปกดการแสดงออกของยีนก่อมะเร็งของไวรัส

Keywords: โปรตีนอีสองของไวรัสแปปิลโลมาชนิด 16; รูปแบบการแสดงออกของอาร์เอ; การแสดงออกของยีนของคน

สารบัญ

	หน้า
บทคัดย่อ	1
ความสำคัญ / ความเป็นมา	4
วัตถุประสงค์ของโครงการ	5
ผลการวิจัย	6
สรุปผลการวิจัย	13
เอกสารอ้างอิง	14
ภาคผนวก	18

ความสำคัญ / ความเป็นมา

High-risk human papillomaviruses (hr-HPVs) infection is accepted as the main cause of cervical cancer (Lombard et al. 1998, Walboomers et al. 1999). Among the HR-HPV types associated with anogenital cancers, HPV 16 is the most common type (Clifford et al. 2005, Clifford et al. 2003, de Sanjose et al. 2007, Munoz et al. 2003) accounting for >50% in cervical cancer. Deregulated expression of HR-HPV oncogenes E6 and E7 leading to cervical carcinogenesis by well characterized mechanisms such as binding of E6 protein to p53, the tumor suppressor gene, causes p53 degradation and E7 binds retinoblastoma protein (pRb) causes releasing of E2F protein, E2F is host transcription factor that drives cell into S phase and initiate cell proliferation (Doorbar 2006, Duensing et al. 2000, Duensing and Munger 2003, Munger et al. 2004). In productive HPV infection phase, expression of E6 and E7 oncogenes is controlled by viral E2 protein. The E2 protein can either activate or repress of viral oncogene promoter. Low E2 protein level binds high affinity distal E2BS that activates gene expression and high level of E2 protein binds both distal and proximal E2BSs consequently preventing the binding of Sp1, TFIID and RNA polymerase complex at promoter region as shown by our experiment (**Figure 1**) and others (Demeret et al. 1997, Fujii et al. 2001, Soeda et al. 2006, Stubenrauch and Laimins 1999). Full length E2 protein consists of N-terminal transactivation domain, hinge region and C-terminal DNA binding domain (McBride 2013). The truncated E2 isoform such as E1^ΔE2 and E8^ΔE2C, that lack transactivation domain, are dominant negative transcription factor that can repress both transcription and replication of viral genome (Ammermann et al. 2008, Lace et al. 2008, Stubenrauch et al. 2000, Stubenrauch et al. 2007, Zobel et al. 2003). Deletion of E2 gene due to viral integration is considered to be the main event that causes loss of negative regulatory function of E2 protein on oncogenes expression. E2 gene transcript was not found in squamous cell carcinoma (SCC) (Chen et al. 2014) while integrate-derived oncogene transcripts become more stabilized because of its co-transcribed cellular sequences (Jeon and Lambert 1995) and as a result, increase viral oncogenes expression level. Additionally, multiple copies of integrated HPV genome orientated in a head to tail contain high methylation in HPV16 promoter region whereas no methylation was found in single copy of integrated HPV16 (De-Castro Arce et al. 2012, Kalantari et al. 2008). Oncogene transcripts of high methylated multiple copies of integrated HPV genome were expressed from one active region without DNA methylation (Van Tine et al. 2004). Furthermore, episomal form is also found in cervical cancer cells (Klaes et al. 1999, Matovina et al. 2009, Vinokurova et al. 2008), epigenetic alteration in proximal E2

binding sites (EBS) such as methylation and chromatin remodeling may play significant role to block binding of E2 protein and consequently abrogate its repressive influence on E6 and E7 expression (Chaiwongkot et al. 2013). A study revealed that episome-derived oncogene transcripts that expressed E6 and E7 oncoproteins were sufficient to induce centrosome abnormalities and genomic instability in raft cultures (Duensing et al. 2001). DNA aneuploidy and chromosomal aberrations were also detected in cervical cells containing episomal HPV16 (Hopman et al. 2004, Melsheimer et al. 2004).

Co-immunoprecipitation and mass spectrometric analysis of E2 complexes revealed that full length E2 protein can interact with many cellular proteins involved in chromatin remodeling, histone modifier, mRNA processing, protein transport, DNA replication, transcription and repair, cell cycle, structural maintenance of chromosomes, proteasome ubiquitin system and apoptosis (Jang et al. 2015, McBride 2013, Muller and Demeret 2012). These interactions may facilitate E2 functional activity such as full length E2-Brd4 interaction is required for active viral transcription and replication. These protein complex also stabilizes E2 protein and tether episomal HPV on transcriptionally active host genome (Jang et al. 2009, McBride and Jang 2013). E2 can also interact with Brm-containing SWI/SNF chromatin remodeling complex which enhance transcription from an episomal reporter plasmid (Kumar et al. 2007). On the other hand, E2 protein can also interact with chromatin remodelers and histone modifier proteins such as NuRD, HDAC1/2 and NCOR/SMRT/HDAC3 that may facilitate transcriptional repression (Dreer et al. 2016, Jang et al. 2015). To diminish the virus induced carcinogenesis, decrease in viral oncogene transcriptions and viral genome replication are the key events. Re-expression of HPV E2 protein may help to stop cervical cancer cells growth by repression of E6 and E7 oncogenes expression (Fernandez et al. 2009). We will investigate whether E2 protein can interact with cellular proteins that are involved in chromosome remodeling and transcriptional repressive activity. HDAC class 1 (HDAC1, 2 and 3) will be selected to study their binding activities with HPV16 full length E2 protein and the effect of these interactions on viral oncogene promoter activity will be investigated. Chromatin immunoprecipitation assay revealed that E2 protein can bind E2BS sequences in the human genome such as nuclear receptor corepressor 2 isoform 1 (NCOR2) (Vosa et al. 2012). Functional analysis showed that full length E2-Brd4 complex can induce c-FOS expression as shown by ChIP assay which c-Fos can activate viral oncogene expression (Delcuratolo et al. 2016). E2 can also enhance NF- κ B and STAT3 genes expression that may be involved in carcinogenesis (Prabhavathy et al. 2014). Therefore, we

further hypothesize that at high E2 concentration apart from binding to its E2BS in viral genome, E2 protein may activate expression of HDAC1, 2, 3 and other repressor genes such as NuRD, NCOR, SMRT, CDP and YY-1 that may facilitate E2 mediated transcriptional repressive activity of episomal and integrated human papillomavirus 16 oncogene promoters.

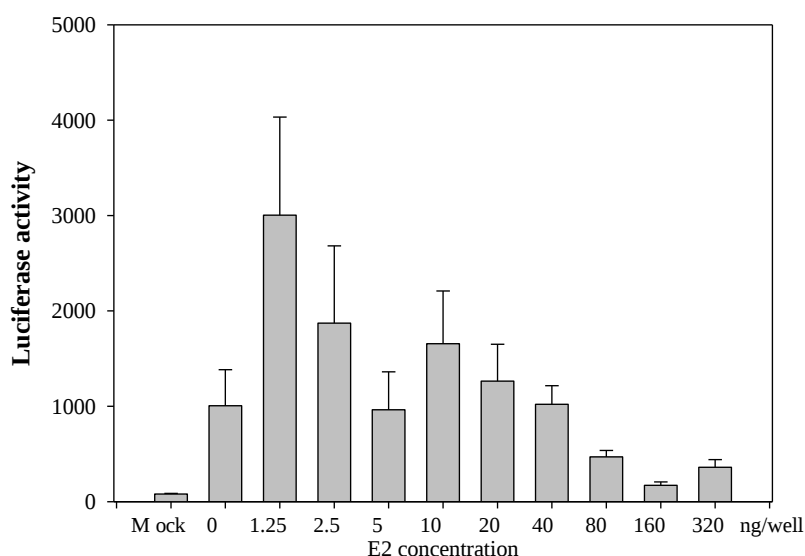


Figure 1. Effect of different E2 concentrations on HPV 16 promoter activity in transient transfection of C33A.

วัตถุประสงค์ของโครงการ

- 1 To study the interaction of HPV16 full length E2 protein with repressor cellular proteins HDAC class 1 (HDAC1, 2 and 3)
- 2 To study the functional activity of full length E2-HDAC complex on HPV16 early promoter activity in episomal and integration form
- 3 To study the binding activity of full length E2 protein on selected cellular gene promoter and their gene expression levels including HDAC1, HDAC2, HDAC3, NuRD, NCoR, SMRT, CDP and YY-1

ผลการวิจัย

1. Effect HPV16 E2 full length on HDAC1 and HDAC3 expression

HPV16 E2 protein expressed from p3662 (Addgene) in cell lines has an effect on increased HDAC1 expression in CaSki cell line but not SiHa cell line detected by real time RT-PCR and western blot as shown in Figure 2A, 2B and 2C. However, there was no upregulation of HDAC3 gene expression in both E2 expressing CaSki and SiHa cell lines. The other cellular gene involved in negatively controlling gene expression included YY1 was not up regulated in E2 expressing cell lines.

2. Effect of HPV16 E2 on HPV16 viral gene expression

It has been shown that HPV16 E6*1 transcript with splicing event within E6 gene fused with full length E7 ORF was highly expressed in both CaSki and SiHa, they were SD226^SA409-SD880^SA3358 (Figure 3A) and SD226^SA409-SD880^SA2709 (Figure 3B), respectively. Most expressed transcripts contain full length E7 ORF. HPV16 E2 could repressed transcripts containing HPV16 full length E7 ORF in CaSki and but has a little effect in SiHa cell lines as shown by real time RT-PCR (Figure 3C). When detecting E6 gene at splicing SA226-SD409, it is shown that E2 has no effect on this site, it seems that transcription pattern of HPV16 transcripts has changed after reintroduction of E2 protein into cell lines that cause less expression of full length E7 and caused accumulation of transcripts with splicing at SA226-SD409 of E6 ORF.

3. Interaction of HPV16 E2 protein with HDAC1 and HDAC3

Co-immunoprecipitation and western blot analysis were done to investigate the E2/HDAC1/HDAC3 complexes. We failed to identify HDAC1 and HDAC3 in co-immunoprecipitated complex using anti-HPV16 E2 antibody as shown in figure 4. However, unknown band (molecular weight between 30-50 kDa) was observed when detecting by anti-HDAC1 antibody, no further experiment was performed to identify this unknown band.

We fail to detect HPV16 proteins included E2 (Figure 2B), E6 and E7 (Figure 5) by western blot analysis either by automated machine or manual detection systems, thus we performed only reverse transcription RT-PCR to detect viral gene expressions.

4. Methylation status of HPV16 L1 gene

As known that HPV16 L1 gene hypermethylation was mostly detected in abnormal cervical cells while, L1 gene hypomethylation was detected in normal cells. We further investigate the effect of E2 protein on methylation status of HPV16 L1 gene whether the repression of E7 protein by E2 protein can lead to demethylation of HPV16 genome as previously reported that E7 protein can activate DNA methyltransferase to induce self-methylation of viral gene. However, there is no change in methylation level between cell lines with E2 and without E2 (Figure 6).

5. Effect of full-length E2 protein on selected cellular gene promoter activity such as HDAC1.

We do not have luminometer machine to detect luciferase expression as reporter gene expressed from E2/promoter complex. I plan to set the experiment using that OD can be detected by microplate reader to investigate the effect of HPV16 E2 protein on cellular promoter activity. This experiment has not finished yet.

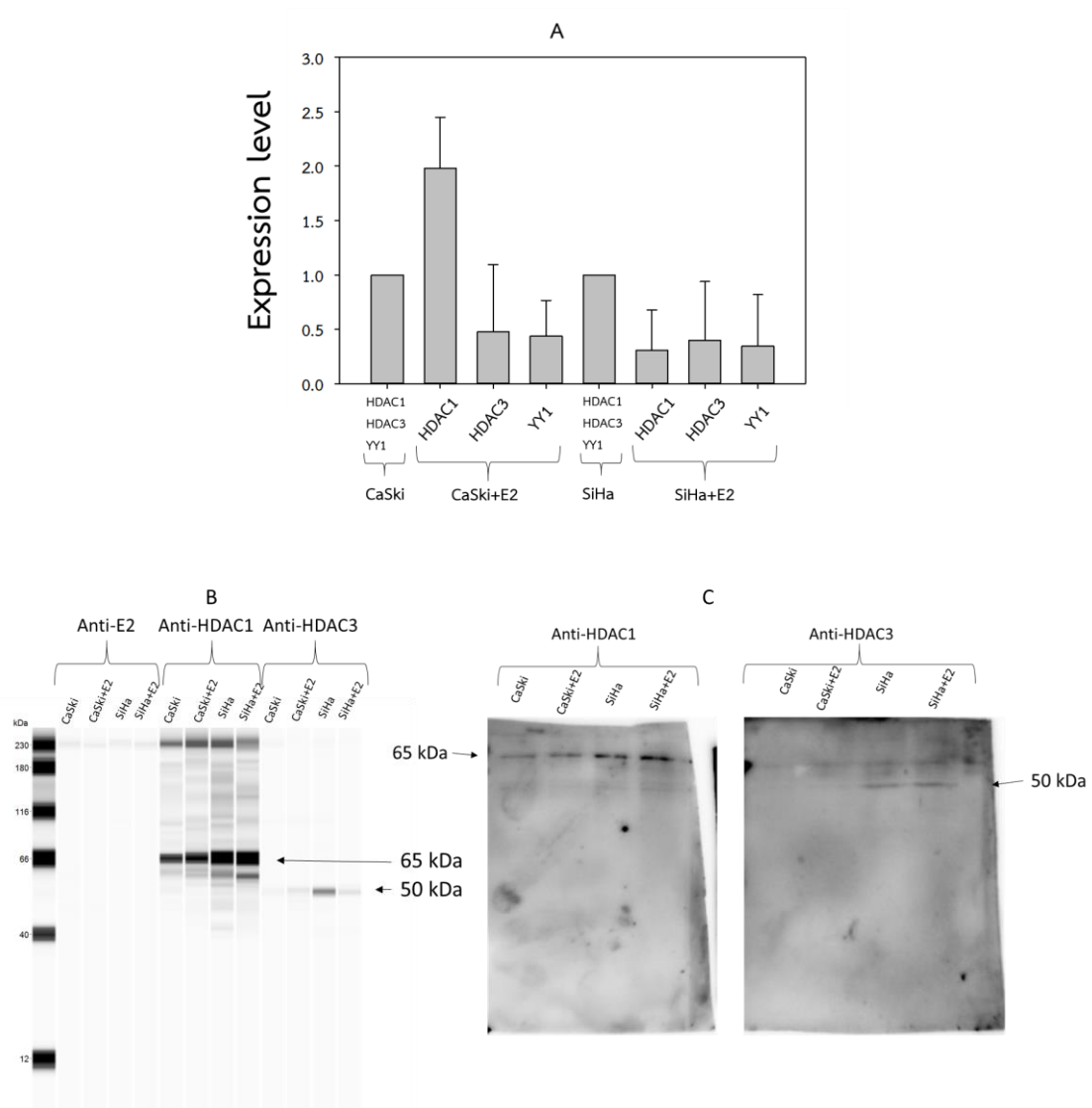


Figure 2. Effect of HPV16 E2 protein on HDAC1 and HDAC3 expression detected by real time PCR (A) and western blot analysis either by automated machine (B) and manual detection system(C).

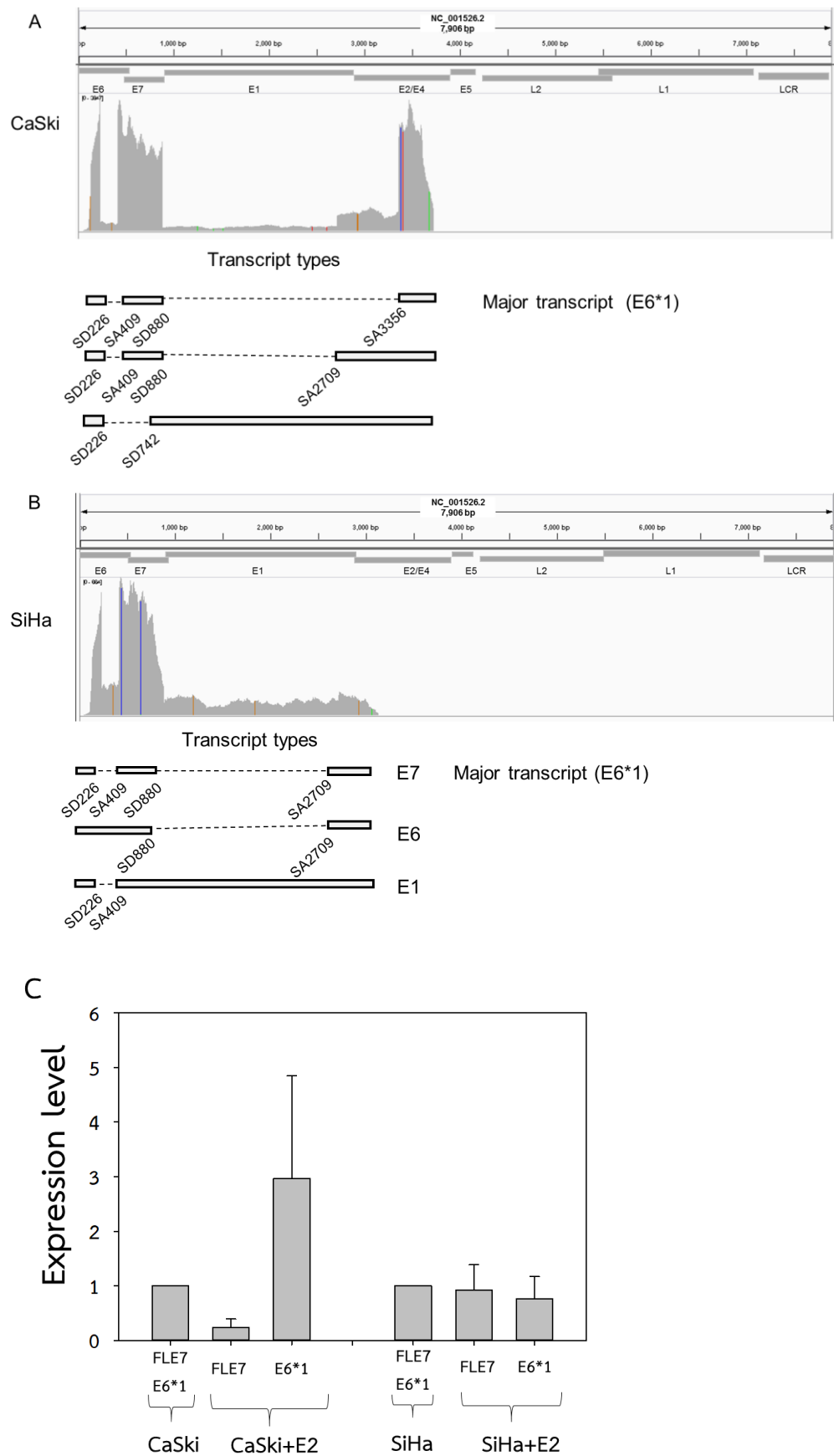


Figure 3. Map of HPV16 major transcripts detected in CaSki (A) and SiHa (B). The level of full length (FL) E7 and E6*1 in E2 expressing cell lines were shown in C.

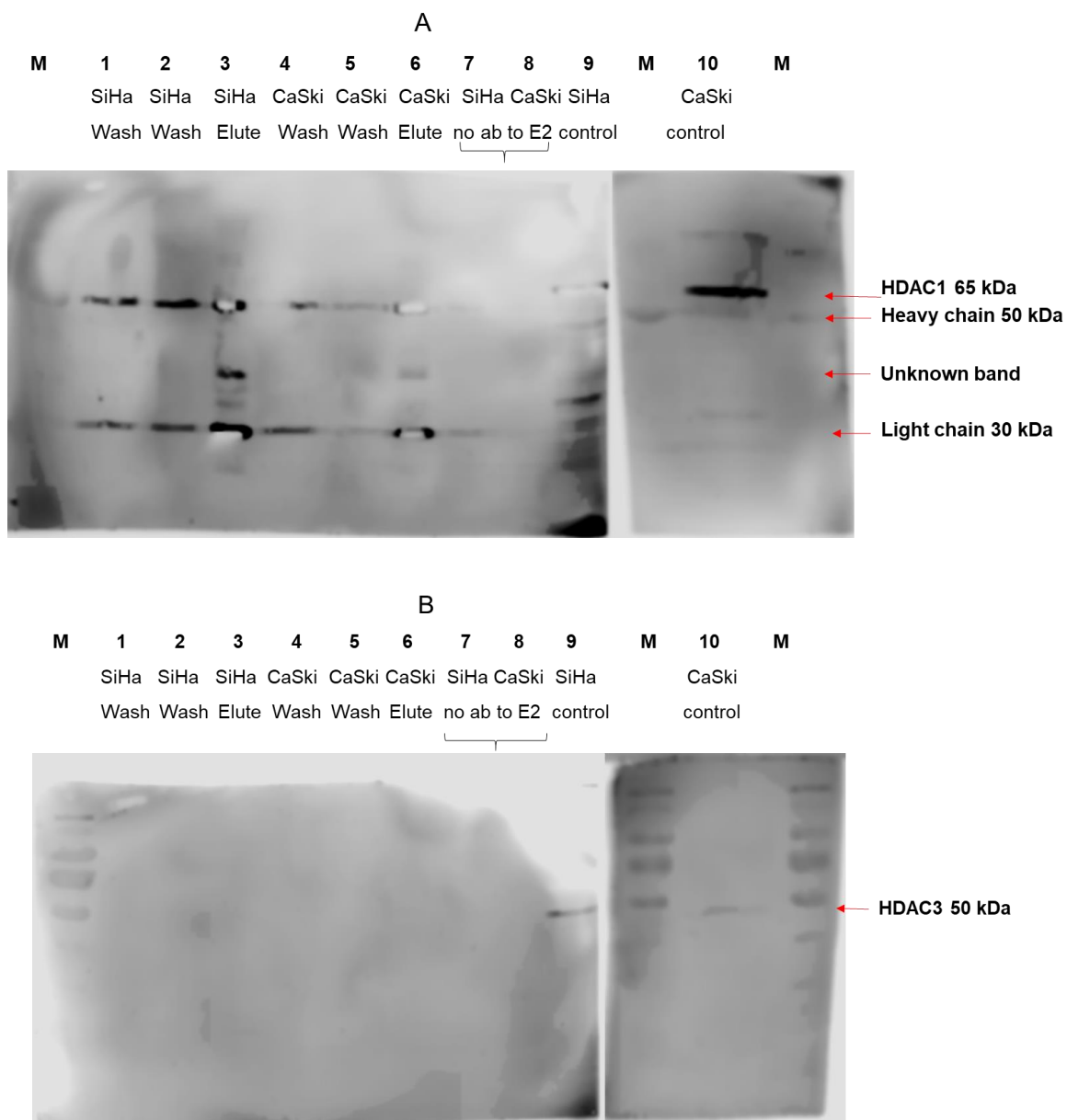


Figure 4. Co-immunoprecipitation using anti HPV16 E2 antibody and western blot analysis using anti- HDAC1 (A) and anti-HDAC3 (B) antibody.

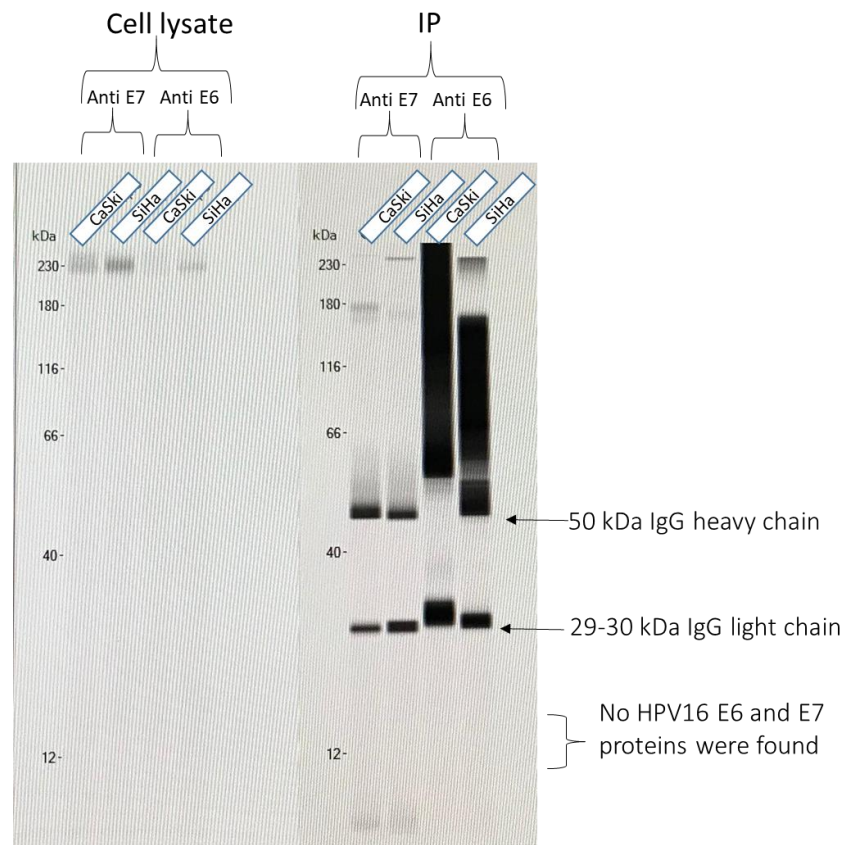


Figure 5. Automated western blot was performed to detect HPV16 E6 and E7 protein in CaSki and SiHa cell lines using cell lysate and immunoprecipitation assay using anti-HPV16 E6 and anti-HPV16 E7 to enrich proteins.

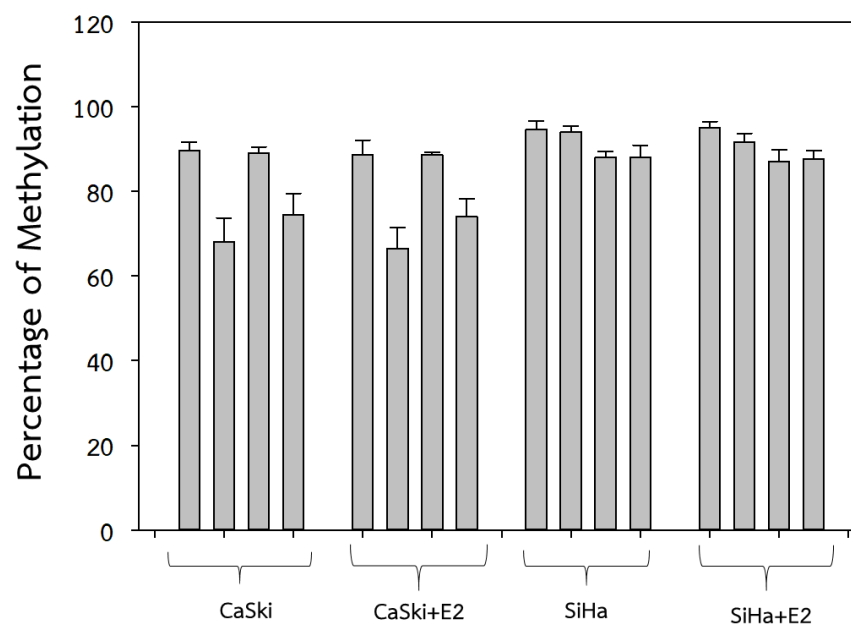


Figure 6. Methylation of HPV16 L1 gene methylation in E2 expressing cell lines.

6. Bioinformatics analysis of cellular gene expression from six cervical cancer cell lines

We have performed RNAseq in high risk HPV positive cervical cancer cell lines and HPV negative cervical cancer cell line. After data normalization, heat map of cellular gene expressions between five HPV positive cervical cancer cell lines (CaSki, SiHa, HeLa, ME180 and MS751) and HPV negative cervical cancer cell line (C33A) was shown in figure 7. Log fold change >4.0 was set as the cut off criteria for the differential gene expression analysis of up regulated genes in HR-HPV infected cell lines compared to HPV negative cell line. There were 1085, 854, 922, 840 and 1057 upregulated genes in CaSki, SiHa, HeLa, MS751 and ME180 respectively. There were 250 overlapped upregulated genes >4.0 fold changes of five cell lines..

Gene-Enrichment and Functional Annotation analysis were analyzed by using g:Profiler and KEGG pathway. 250 overlapped upregulated genes were used for gene ontology (GO) complete analysis, 174 pathways in biological process (BP), 55 cellular components (CC), 7 pathways in molecular function (MF) and 13 (KEGG) pathways were obtained. The most significant biological process included negative regulation of keratinocyte proliferation (GO:0010839), cell differentiation involved in embryonic placenta development (GO:0060706), regulation of keratinocyte proliferation (GO:0010837), establishment of endothelial barrier (GO:0061028) and negative regulation of extrinsic apoptotic signaling pathway via death domain receptors (GO:1902042). For cellular component included; laminin-5 complex (GO:0005610), apicolateral plasma membrane (GO:0016327), recycling endosome membrane (GO:0055038), specific granule (GO:0042581) and cell-substrate junction (GO:0030055), the significantly molecular function pathway included cadherin binding (GO:0045296), cell adhesion molecule binding (GO:0050839), identical protein binding (GO:0042802), enzyme binding (GO:0019899) and protein binding (GO:0005515). The most significant KEGG pathway analysis included; Pathways in cancer (hsa05200), Proteoglycans in cancer (hsa05205), p53 signaling pathway (hsa04115), MicroRNAs in cancer (hsa05206).

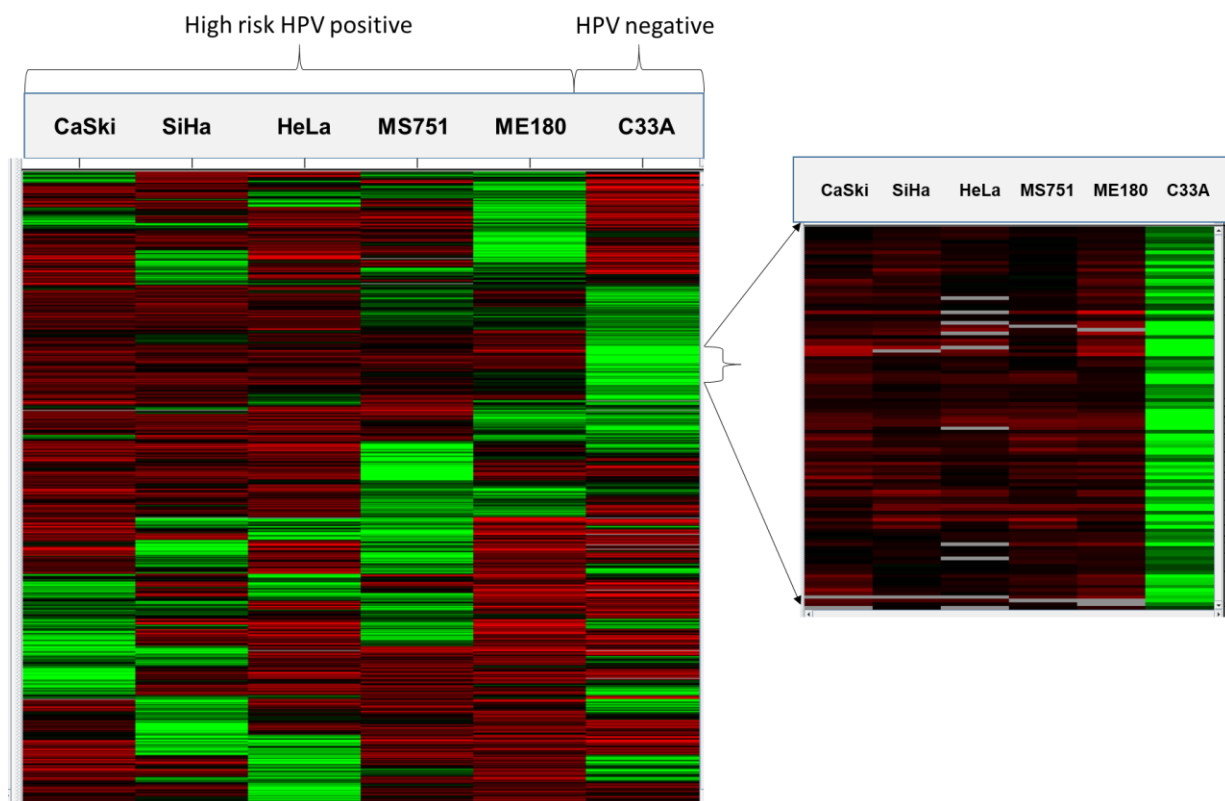


Figure 7. Heat map of host gene expression profile in HPV positive and HPV negative cell lines.

สรุปผลการวิจัย

The present study revealed the association between HPV16E2 protein and HDAC1 expression. HDAC1 is the cellular protein that function as negative regulatory gene expression may be involved in suppression of viral gene expression. However, we failed to detect HDAC1 co-precipitated with HPV16 E2 protein, however, unknown band was observed. We will optimize assay using plasmid containing secreted alkaline phosphatase as reporter gene (SEAP (secreted alkaline phosphatase) for promoter activity detection that would be more suitable in laboratory using microplate ELISA reader.

The transcriptomic analysis of cellular gene expression of high risk HPV compared to HPV negative cervical cancer cell line revealed genes involved in Pathways in cancer that can be used as targeted therapy apart from HPV oncogenes.

For further experiment, we will select cellular genes that are highly expressed in high risk HPV positive cervical cancer cell lines and study the function of that selected genes in term of carcinogenesis involvement.

References

- Ammermann, I., Bruckner, M., Matthes, F., Iftner, T. and Stubenrauch, F. Inhibition of transcription and DNA replication by the papillomavirus E8-E2C protein is mediated by interaction with corepressor molecules *J. Virol.* 2008;82:5127-36
- Chaiwongkot, A., Vinokurova, S., Pientong, C., Ekalaksananan, T., Kongyingyoes, B., Kleebkao, P., et al. Differential methylation of E2 binding sites in episomal and integrated HPV 16 genomes in preinvasive and invasive cervical lesions *Int. J. Cancer* 2013;132:2087-94
- Chen, J., Xue, Y., Poidinger, M., Lim, T., Chew, S. H., Pang, C. L., et al. Mapping of HPV transcripts in four human cervical lesions using RNAseq suggests quantitative rearrangements during carcinogenic progression *Virology* 2014;462-463:14-24
- Clifford, G. M., Rana, R. K., Franceschi, S., Smith, J. S., Gough, G. and Pimenta, J. M. Human papillomavirus genotype distribution in low-grade cervical lesions: comparison by geographic region and with cervical cancer *Cancer Epidemiol. Biomarkers Prev.* 2005;14:1157-64
- Clifford, G. M., Smith, J. S., Plummer, M., Munoz, N. and Franceschi, S. Human papillomavirus types in invasive cervical cancer worldwide: a meta-analysis *Br. J. Cancer* 2003;88:63-73
- De-Castro Arce, J., Gockel-Krzikalla, E. and Rosl, F. Silencing of multi-copy HPV16 by viral self-methylation and chromatin occlusion: a model for epigenetic virus-host interaction *Hum. Mol. Genet.* 2012;21:1693-705
- de Sanjose, S., Diaz, M., Castellsague, X., Clifford, G., Bruni, L., Munoz, N., et al. Worldwide prevalence and genotype distribution of cervical human papillomavirus DNA in women with normal cytology: a meta-analysis *The Lancet Infectious Diseases* 2007;7:453-9
- Delcuratolo, M., Fertey, J., Schneider, M., Schuetz, J., Leiprecht, N., Hudjetz, B., et al. Papillomavirus-Associated Tumor Formation Critically Depends on c-Fos Expression Induced by Viral Protein E2 and Bromodomain Protein Brd4 *PLoS Pathog* 2016;12:e1005366
- Demeret, C., Desaintes, C., Yaniv, M. and Thierry, F. Different mechanisms contribute to the E2-mediated transcriptional repression of human papillomavirus type 18 viral oncogenes *J. Virol.* 1997;71:9343-9
- Doorbar, J. Molecular biology of human papillomavirus infection and cervical cancer *Clin. Sci.* 2006;110:525-41
- Dreer, M., Fertey, J., van de Poel, S., Straub, E., Madlung, J., Macek, B., et al. Interaction of NCOR/SMRT Repressor Complexes with Papillomavirus E8^{E2C} Proteins Inhibits Viral Replication *PLoS Pathog* 2016;12:e1005556
- Duensing, S., Duensing, A., Flores, E. R., Do, A., Lambert, P. F. and Munger, K. Centrosome abnormalities and genomic instability by episomal expression of human papillomavirus type 16 in raft cultures of human keratinocytes *J. Virol.* 2001;75:7712-6

Duensing, S., Lee, L. Y., Duensing, A., Basile, J., Piboonniyom, S., Gonzalez, S., et al. The human papillomavirus type 16 E6 and E7 oncoproteins cooperate to induce mitotic defects and genomic instability by uncoupling centrosome duplication from the cell division cycle *Proc. Natl. Acad. Sci. U. S. A.* 2000;97:10002-7

Duensing, S. and Munger, K. Human papillomavirus type 16 E7 oncoprotein can induce abnormal centrosome duplication through a mechanism independent of inactivation of retinoblastoma protein family members *J. Virol.* 2003;77:12331-5

Fernandez, A. F., Rosales, C., Lopez-Nieva, P., Grana, O., Ballestar, E., Ropero, S., et al. The dynamic DNA methylomes of double-stranded DNA viruses associated with human cancer *Genome Res.* 2009;19:438-51

Fujii, T., Brandsma, J. L., Peng, X., Srimatkandada, S., Li, L., Canaan, A., et al. High and low levels of cottontail rabbit papillomavirus E2 protein generate opposite effects on gene expression *J. Biol. Chem.* 2001;276:867-74

Hopman, A. H., Smedts, F., Dignef, W., Ummelen, M., Sonke, G., Mravunac, M., et al. Transition of high-grade cervical intraepithelial neoplasia to micro-invasive carcinoma is characterized by integration of HPV 16/18 and numerical chromosome abnormalities *J. Pathol.* 2004;202:23-33

Jang, M. K., Anderson, D. E., van Doorslaer, K. and McBride, A. A. A proteomic approach to discover and compare interacting partners of papillomavirus E2 proteins from diverse phylogenetic groups *Proteomics* 2015;15:2038-50

Jang, M. K., Kwon, D. and McBride, A. A. Papillomavirus E2 proteins and the host BRD4 protein associate with transcriptionally active cellular chromatin *J. Virol.* 2009;83:2592-600

Jeon, S. and Lambert, P. F. Integration of human papillomavirus type 16 DNA into the human genome leads to increased stability of E6 and E7 mRNAs: implications for cervical carcinogenesis *Proc. Natl. Acad. Sci. U. S. A.* 1995;92:1654-8

Kalantari, M., Lee, D., Calleja-Macias, I. E., Lambert, P. F. and Bernard, H. U. Effects of cellular differentiation, chromosomal integration and 5-aza-2'-deoxycytidine treatment on human papillomavirus-16 DNA methylation in cultured cell lines *Virology* 2008;374:292-303

Klaes, R., Woerner, S. M., Ridder, R., Wentzensen, N., Duerst, M., Schneider, A., et al. Detection of high-risk cervical intraepithelial neoplasia and cervical cancer by amplification of transcripts derived from integrated papillomavirus oncogenes *Cancer Res.* 1999;59:6132-6

Kumar, R. A., Naidu, S. R., Wang, X., Imbalzano, A. N. and Androphy, E. J. Interaction of papillomavirus E2 protein with the Brm chromatin remodeling complex leads to enhanced transcriptional activation *J. Virol.* 2007;81:2213-20

Lace, M. J., Anson, J. R., Thomas, G. S., Turek, L. P. and Haugen, T. H. The E8--E2 gene product of human papillomavirus type 16 represses early transcription and replication but is dispensable for viral plasmid persistence in keratinocytes J. Virol. 2008;82:10841-53

Lombard, I., Vincent-Salomon, A., Validire, P., Zafrani, B., de la Rochefordiere, A., Clough, K., et al. Human papillomavirus genotype as a major determinant of the course of cervical cancer J. Clin. Oncol. 1998;16:2613-9

Matovina, M., Sabol, I., Grubisic, G., Gasperov, N. M. and Grce, M. Identification of human papillomavirus type 16 integration sites in high-grade precancerous cervical lesions Gynecol. Oncol. 2009;113:120-7

McBride, A. A. The papillomavirus E2 proteins Virology 2013;445:57-79

McBride, A. A. and Jang, M. K. Current understanding of the role of the Brd4 protein in the papillomavirus lifecycle Viruses 2013;5:1374-94

Melsheimer, P., Vinokurova, S., Wentzensen, N., Bastert, G. and von Knebel Doeberitz, M. DNA aneuploidy and integration of human papillomavirus type 16 E6/E7 oncogenes in intraepithelial neoplasia and invasive squamous cell carcinoma of the cervix uteri Clin. Cancer Res. 2004;10:3059-63

Muller, M. and Demeret, C. The HPV E2-Host Protein-Protein Interactions: A Complex Hijacking of the Cellular Network Open Virol J 2012;6:173-89

Munger, K., Baldwin, A., Edwards, K. M., Hayakawa, H., Nguyen, C. L., Owens, M., et al. Mechanisms of human papillomavirus-induced oncogenesis J. Virol. 2004;78:11451-60

Munoz, N., Bosch, F. X., de Sanjose, S., Herrero, R., Castellsague, X., Shah, K. V., et al. Epidemiologic classification of human papillomavirus types associated with cervical cancer N. Engl. J. Med. 2003;348:518-27

Prabhavathy, D., Vijayalakshmi, R., Kanchana, M. P. and Karunagaran, D. HPV16 E2 enhances the expression of NF-kappaB and STAT3 target genes and potentiates NF-kappaB activation by inflammatory mediators Cell. Immunol. 2014;292:70-7

Soeda, E., Ferran, M. C., Baker, C. C. and McBride, A. A. Repression of HPV16 early region transcription by the E2 protein Virology 2006;351:29-41

Stubenrauch, F., Hummel, M., Iftner, T. and Laimins, L. A. The E8E2C protein, a negative regulator of viral transcription and replication, is required for extrachromosomal maintenance of human papillomavirus type 31 in keratinocytes J. Virol. 2000;74:1178-86

Stubenrauch, F. and Laimins, L. A. Human papillomavirus life cycle: active and latent phases Semin. Cancer Biol. 1999;9:379-86

Stubenrauch, F., Straub, E., Fertey, J. and Iftner, T. The E8 repression domain can replace the E2 transactivation domain for growth inhibition of HeLa cells by papillomavirus E2 proteins Int. J. Cancer 2007;121:2284-92

Van Tine, B. A., Kappes, J. C., Banerjee, N. S., Knops, J., Lai, L., Steenbergen, R. D., et al. Clonal selection for transcriptionally active viral oncogenes during progression to cancer J. Virol. 2004;78:11172-86

Vinokurova, S., Wentzensen, N., Kraus, I., Klaes, R., Driesch, C., Melsheimer, P., et al. Type-dependent integration frequency of human papillomavirus genomes in cervical lesions *Cancer Res.* 2008;68:307-13

Vosa, L., Sudakov, A., Remm, M., Ustav, M. and Kurg, R. Identification and analysis of papillomavirus E2 protein binding sites in the human genome *J. Virol.* 2012;86:348-57

Walboomers, J. M., Jacobs, M. V., Manos, M. M., Bosch, F. X., Kummer, J. A., Shah, K. V., et al. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide *J. Pathol.* 1999;189:12-9

Zobel, T., Iftner, T. and Stubenrauch, F. The papillomavirus E8-E2C protein represses DNA replication from extrachromosomal origins *Mol. Cell. Biol.* 2003;23:8352-62

ภาคผนวก

Title: Transcriptomic dataset of high-risk human papillomaviruses 16, 18, 45, 68b transcripts in five cervical cancer cell lines

Author names

Arkom Chaiwongkot^{1,2*}, Masahide Seki³, Yutaka Suzuki³, Teerasit Techawiwattanaboon², Parvapan Bhattarakosol^{1,2}.

Affiliations

¹Applied Medical Virology Research Unit, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand.

²Department of Microbiology, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand.

³Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Chiba, Japan.

***Corresponding author**

Arkom Chaiwongkot, Ph.D.

Applied Medical Virology Research Unit, Department of Microbiology, Faculty of Medicine, Chulalongkorn University 1873 Rama 4 Road, Pathumwan, Bangkok, 10330, Thailand

Fax; +66 02-252-5952 Telephone; +66 02-256-4132 E-mail:

arkomchaiwongkot@gmail.com and Arkom.Cha@chula.ac.th

Funding

This work was supported by Ratchadapiseksompoch Fund, Faculty of Medicine, Chulalongkorn University, grant number RA59/013 and was partly supported by the TRF Grant for new researcher from Thailand Research Fund Grant (Grant number TRG6080008)

ABSTRACT

Upregulated expression of high-risk human papillomaviruses (HR-HPVs) E6 and E7 oncogenes is associated with cervical carcinogenesis. Transcription patterns and major viral transcripts of four HR-HPV types 16, 18, 45 and 68b were analysed by RNAseq using RNA extracted from five cervical cancer cell lines (CaSki, SiHa, HeLa, MS751 and ME180). The major transcripts obtained from four HR-HPV types 16, 18, 45 and 68b were E6*1 with splicing event occur within E6 gene fused with full length E7 ORF, 226^409 (HPV16), 233^416 (HPV18), 230^412 (HPV45) and 129^311(HPV68b) and were associated with full length E7 oncogene expression. Real time PCR revealed that full length HR-HPV E7 oncogene was highly expressed in all five cell lines.

Keywords: Human papillomavirus; transcription patterns; RNAseq; Cervical cancer; Major transcript; E6 and E7 oncogenes

1.Introduction

Long term persistent infection and upregulated expression of high risk human papillomaviruses (HR-HPVs) E6 and E7 oncoproteins leading to chromosomal instability and cervical cancer progression. These E6 and E7 oncoproteins disrupt the normal function of host proteins necessary for normal cell cycle regulation for example, E6 causes degradation of p53 and E7 inactivates retinoblastoma tumor suppressor protein (pRb) leading to cell cycle progression (Munger et al. 2004, Munger et al. 2001). Epidemiological reports showed that HPV16 is mostly found in cervical cancer cases accounting for 50% worldwide, followed by HPV18 (~20%) and other HR-HPV types such as HPV45, 31, 33, 58, 52 depends on regions (Clifford et al. 2003, Munoz et al. 2003).

HPV16, 18 and 31 transcription patterns have been widely studied and revealed that viral oncogene transcripts E6 and E7 are expressed from early promoter as bicistronic or polycistronic transcripts with two or more open reading frames (ORFs) located on the same mRNA with splicing variants found in E6 region in some transcript types (Ajiro et al. 2016, Graham 2010, Graham and Faizo 2017, Milligan et al. 2007, Ozbun 2002, Ozbun and Meyers 1998, Schmitt et al. 2011, Stubenrauch and Laimins 1999, Wang et al. 2011). It has been reported that the transcript encoding the full length of E6 and E7 mRNA translated only the first E6 ORF that located close to promoter (Smotkin et al. 1989, Tang et al. 2006) due to the distance between two ORFs is short, consequently, only the first ORF is translated (Attal et al. 1999). It was reported that monocistronic mRNA containing E7 ORF expressed from promoter located within E6 gene such as p542 for HPV16 is required in order to efficiently translate E7 protein (Braunstein et al. 1999, Glahder et al. 2003). Splicing event within E6 intronic region is the alternative mechanism to help express E7 protein and E6 disruption by splicing event is associated with cervical carcinogenesis (Chen et al. 2014, Lin et al. 2015, Tang et al. 2006). Different splicing patterns within E6 region of HPV16 have been identified with one splicing donor (SD) site at position 226 and three splicing acceptor (SA) sites at position 409, 526 and 742, they are named as E6*I (SD226^SA409), E6*II (SD226^SA526) and E6*X(E6*^E7*)(SD226^SA742) (Graham 2010). Two splicing patterns within E6 gene of HPV18 have been identified including E6*I (SD233^SA416) and E6*^E7* (SD233^SA791) (Ajiro et al. 2016, Toots et al. 2014, Wang et al. 2011). There is no splicing event within E6 region of low risk HPVs (Zheng 2010).

Most of previous published papers focused on HPV transcription patterns, however, there were limited quantitative information of major transcript that are associated with E6 and E7 oncogene expression. Recent study using RNAseq assay reported quantitative results of different HPV16 transcripts (Chen et al. 2014). Other HR-HPVs that are necessary for cervical carcinogenesis should be concerned in order to understand the transcription patterns among HR-HPV types. Thus, we aimed to employ RNAseq analysis to identify major transcripts that are associated with E6 and E7 oncogene expressions of not only HPV16 and HPV18 but also HPV45 and HPV68b in five cervical cancer cell lines. The major transcripts that were associated with E6* and E7 oncogene expressions obtained from four HR-HPV types 16, 18, 45 and 68b with integration form were identified in the present study.

2. Materials and methods

2.1. Cervical cancer cell lines

Five cervical cancer cell lines containing different HR-HPV types were used in the present study including CaSki (containing approximately 600 copies of integrated HPV16 per cell, (CRL-1550 Lot No.3794357)), SiHa (containing approximately 1-2 copies of integrated HPV16 per cell (HTB-35 Lot No. 4031219)), HeLa (containing integrated HPV18), MS751(containing integrated HPV45(HTB-34 Lot No.58078726)) and ME180 (containing integrated HPV68b(HTB-33 Lot 57758662)). All cell lines were maintained in Dullbecco's modified Eagle's medium (DMEM) with 10% fetal bovine serum in CO₂ incubator at 37°C. This study has been approved (COA No. 677/2015) by the Institutional Review Board of the Faculty of Medicine, Chulalongkorn University.

2.2. RNA preparation for RNA sequencing

Total RNA was extracted from cervical cancer cell lines using an RNeasy® Mini Kit (QIAGEN, Valencia, CA) according to the manufacturer's instructions. Approximately 10,000,000 cells were used. The RNA was eluted twice in a total volume of 100 µl of RNase-free H₂O. RNA was quantified by Agilent 2100 Bioanalyzer and samples with RIN value >8.0 were used for RNAseq analysis. Illumina TrueSeq Stranded mRNA Sample

Preparation Kit were used. Samples were sequenced on an Illumina HiSeq2500 using a paired-end run.

2.3. Data analysis

The Hisat2 v2.1.0 aligner was used to map the raw data reads onto the viral reference sequences as follows, HPV16 (NC_001526.2), HPV18 (NC_001357.1), HPV68b (FR751039.1) and HPV45 (X74479.1). Following the alignment, the reads were assembled into transcripts and quantified, using Cufflinks v. 2.2.1. Following the alignment, the reads were assembled into transcripts (Cuffmerge) and quantified (Cuffdiff) as fragments per kilobase per million (FPKM). The results were visualized, using Integrated Genomic Viewer (IGV) to identify the expressed transcripts and splicing junction of each HPV type.

2.4 Validation of HPV transcripts by Real Time Polymerase Chain Reaction (real time-PCR)

Primers were designed to detect the viral transcripts in five cell lines by real time RT-PCR. cDNA was synthesized using Superscript III Reverse Transcriptase (Invitrogen). Real time PCR was performed using SsoAdvanced Universal SYBR Green Supermix (Bio-Rad). Primers sequences were shown in Table 1. The PCR conditions were: initial denaturing at 98 °C for 30 sec, followed by 40 cycles of 95 °C for 15 seconds and 60 °C for 1 minute. GAPDH was used as house keeping gene and $\Delta\Delta Ct$ was calculated to investigate difference in gene expression.

3. Results

3.1. Identification of expressed transcripts

The reads were mapped onto the viral reference sequences to identify the viral transcripts that were expressed in each cervical cancer cell lines (CaSki (HPV16), SiHa (HPV16), HeLa (HPV18), MS751 (HPV45) and ME180 (HPV68b)). Transcripts containing truncated E6 and full length E7 genes were highly detected in all four HR-HPV types 16, 18, 45 and 68b. Full length E6, full length E1 and truncated E1 transcripts were also detected but in less amount. Transcript containing full length E2 and truncated E2 ORF were found in HPV16 positive CaSki cell line (Figure 1).

The splicing junctions of all four HPV types were found within E6 and E1 regions. For E6 region, one splicing donor (SD) at the 5' end and different splicing acceptor (SA) positions at the 3' end were found as follow; three splicing junctions were found in HPV16 positive cervical cancer cell lines, CaSki and SiHa, SD226^SA409(E6*I), SD226^SA526(E6*II) and SD226^SA742(E6*X). Two splicing junctions were found in HPV18 (HeLa), SD233^SA416(E6*I), SD233^SA635, HPV45(MS751); SD230^SA412(E6*I), SD230^SA640 and HPV68b (ME180); SD129^SA311(E6*I), SD129^SA406.

Splicing junctions within E1 region found in CaSki and SiHa were SD880^SA3358, SD880^SA3361, SD880^SA3391, SD880^SA1726, SD880^SA2405, SD880^SA2582(E1C), SD880^SA2709(E2*), SD880^SA3020, SD880^SA3078, SD880^SA3329, SD577^SA6810, SD898^SA1725, SD1302^SA2709, SD1302^SA3358(E2C), SD1760^SA3391, SD1263^SA3391, SD2309^SA3461 and the other forms were SD96^SA1063, SD226^SA2709 (E6*IV), SD226^SA3329, SD226^SA3358(E6*III), SD226^SA3361, SD226^SA3391 and SD579^SA6809. Splicing junctions within E1 region in HPV18(HeLa) were SD929^SA2779, SD977^SA1836, SD1342^SA1436, SD1987^SA2047 and one splicing event within E7 region, SD599^SA619. HPV68b(ME180) were SD839^SA2586, SD683^SA2586, SD839^SA2586 and no E1 splicing junctions were found in HPV45(MS751). Transcripts found in four HR-HPV types 16, 18, 45 and 68b in five cervical cancer cell lines were expressed as bicistronic and polycistronic.

3.2. Major transcripts of HR-HPV types 16, 18, 45 and 68b

According to coverage and FPKM values visualized by IGV software, major transcripts found in CaSki (HPV16), SiHa(HPV16), HeLa(HPV18), MS751 (HPV45) and ME180(HPV68b) were SD226^SA409-SD880^SA3358 (Figure 2A), SD226^SA409-SD880^SA2709 (Figure 2B), SD233^SA416 (Figure 2C), SD230^SA412 (Figure 2D) and SD129^SA331-SD839^SA2586 (Figure 2E), respectively. All major transcripts of four HR-HPVs contain disruption within E6 gene.

3.3 Validation of HPV oncogene expressions by real time PCR

HPV transcripts were validated to confirm the RNAseq results, primers specific for full length and truncated HPV transcript types were designed as shown in table 1. It has been shown that full length E7 transcript of all 4 HR-HPV types was highly expressed. E6*1 transcript was the second most expressed, however full length E6 transcript was less expressed.

4. Discussion

Four HR-HPV types 16, 18, 45 and 68 mRNA obtained from five different cervical cancer cell lines were used for RNAseq analysis to identify major transcripts necessary for expression of viral oncogenes. Expression of viral oncogenes E6 and E7 are detected as bicistronic and polycistronic, however, it has been reported that a bicistronic transcript encoding the full length E6 and E7 ORFs translates only E6 protein (Smotkin et al. 1989, Tang et al. 2006). The mechanism is that the spacing between the E6 ORF and E7 ORF within this full length transcript is only two nucleotides, thus, the scanning ribosome does not have enough time to discard its releasing factors and to efficient reinitiation of E7 protein translation. In order to efficiently translate E7 gene, monocistronic transcript containing the full length E7 ORF transcribed from promoter located within E6 region should be expressed (Braunstein et al. 1999, Glahder et al. 2003). Deep RNA sequencing was performed to identify and quantitate the major transcripts that were expressed in five cervical cancer cell lines containing four different HR-HPV types included HPV16, 18, 45 and 68b. The RNAseq results of four HR-HPV types in the present study revealed less mRNA transcript encoding the full length E6 and monocistronic full length E7 transcripts (Figure 1 and 2). Thus, these two transcript types were not required for increased E6 and E7 proteins expression. The major transcripts that were detected in all four HR-HPV types were transcripts with disruption of E6 intronic region where splicing event occurred that consistent with previous published papers that E7 oncoprotein was efficiently expressed from mRNA with splicing event occurs within E6 gene (Chen et al. 2014, Lin et al. 2015, Tang et al. 2006, Zheng et al. 2004). The explanation is that splicing of E6 intron creates a frameshift, resulting in the spliced mRNA which has enough nucleotide space for the translation termination of truncated E6 protein and re-initiation of full length E7 translation (Tang et al. 2006). However, it has been reported that the distance between the splicing acceptor at E6 region and the start codon of E7 gene must be optimal to efficiently translate E7 protein(Smotkin et al. 1989, Zheng et al. 2004). Quantitative analysis as shown by coverage and FPKM values of major transcripts revealed that transcript types 226^409 (HPV16), 233^416 (HPV18), 230^412 (HPV45) and 129^311(HPV68b) with full length E7 ORF were mostly expressed (Figure 2) and these transcripts were named E6*I (Tang et al. 2006). HPV16, for example, the distance between the SA409 to E7 start codon at nucleotide position 562 is 153 bp that is the optimal space for strongly translate E7 gene when compared to SA526 to E7 start codon 562 (distance 36 bp). The transcript with splicing acceptor within E7 gene and unspliced within E1 gene were proposed to encode for E1 protein(Johansson and Schwartz 2013, Schwartz 2013, Wu et al. 2017). It has been reported that splicing event within E6 gene was found only in HR-HPVs but not Low risk HPVs (Zheng 2010). The level of E6*I transcript found in CaSki and SiHa was not largely different as shown by RNAseq(FPKM 664,626 and 510,559 in CaSki and SiHa,

respectively) and real time PCR (Figure 3), although there is the difference in copy number of integrated HPV16 within each cell line, that may be due to high methylation (>90%) and closed chromatin were found in long control region (LCR) of HPV16 in CaSki cell lines (600 copies per cell), however, no methylation was found in HPV16 LCR in SiHa cell line (1-2 copies per cell) (Chaiwongkot et al. 2013, De-Castro Arce et al. 2012, Stunkel and Bernard 1999).

It has been shown by amplification of papillomavirus oncogene transcripts (APOTs) assay that HPV16 E2 gene was retained in CaSki cell line and splicing acceptor was also found in E2 gene at position 3358 (data not shown). In accordance with RNAseq results, in which major transcript of HPV16 found in CaSki cell line was SD226/SA409-SD880/SA3358. The important of SA3358 for HPV16 E6 and E7 translation was reported (Li et al. 2013). APOTs assay for HPV18 in HeLa revealed splicing donor site at 929 in E1 gene which spliced to cellular sequence (data not shown), however, the major transcript of HPV18 detected in the present study was SD233/SA416 without splicing event within E1 gene (Figure 2C).

When comparing HPV16 transcripts found in cervical cancer cell lines, CaSki and SiHa in the present study with previous report transcripts of W12 cell line (W12 cell lines derived from LSIL cells), it was found that transcript encoding the full length E6 and E7 ORFs with splicing donor at 880 was found only in cervical cancer cell lines (Graham 2010, Milligan et al. 2007). This transcript type was previously reported to be expressed only in CIN3 and squamous cell carcinoma (SCC) (Chen et al. 2014). Transcript type E6*I was mostly detected in cervical cancer (Chen et al. 2014, Schmitt et al. 2011, Zheng 2010). In the present study study, HPV16 oncogene transcripts E6*I containing truncated E6 and full length E7 were present in high levels compared to E6*II and E6*X in both CaSki and SiHa cervical cancer cell lines. E6*I and E7 transcripts were also highly expressed in HR-HPV types 18, 45 and 68b. One study showed that E7 oncoprotein of HPV16 and HPV18 was highly expressed in CaSki and HeLa, respectively, that correlated with the major transcript for E7 gene expression (E6*I) detected in the present study (Seedorf et al. 1987).

The cervical cancer cell line MS751 revealed transcripts of truncated E6, E7 and part of E1 ORFs that consistent with previous report on characterization of HPV45 DNA in MS751 cell line in which E6, E7 and part of E1 ORFs were retained after viral integration into host chromosome (Geisbill et al. 1997). Apart from E6*I, E7 and E1 transcripts, both HPV18 and 45 (Both HPVs are in the same group, Alpha 7 papillomavirus) expressed non coding region, however, their significant is not well understood.

In summary, full length HPV E7 ORF was found in most expressed transcripts, followed by truncated E6. Full length E6 ORF was less detected. E2 transcript was also found in CaSki cell lines. HR-HPV16, 18, 45, 68b expressed common major transcripts with splicing donor and acceptor sites occurred within E6 gene (E6*I) that facilitate E7 oncogene translation. Thus, major transcripts encoded for truncated E6*I isoform and E7 proteins expressed from HR-HPV types may be involved in carcinogenesis (Paget-Bailly et al. 2019) could be used as biomarker and the target for cervical cancer therapy which may help to improve the treatment. could also be used as drug target.

Competing interests

The authors declare that they have no competing interests

Acknowledgements

We thank Vanida Mungmee for her kindly help in cell culture preparation, Megumi Kombu for her help with the RNAseq library construction and Yuuta Kuze for his help with the data analyses. We thank Suzuki lab at the University of Tokyo for providing facilities necessary for RNAseq experiment. This work was supported by Ratchadapiseksompoch Fund, Faculty of Medicine, Chulalongkorn University, grant number RA59/013 and was partly supported by the TRF Grant for new researcher from Thailand Research Fund Grant (Grant number TRG6080008).

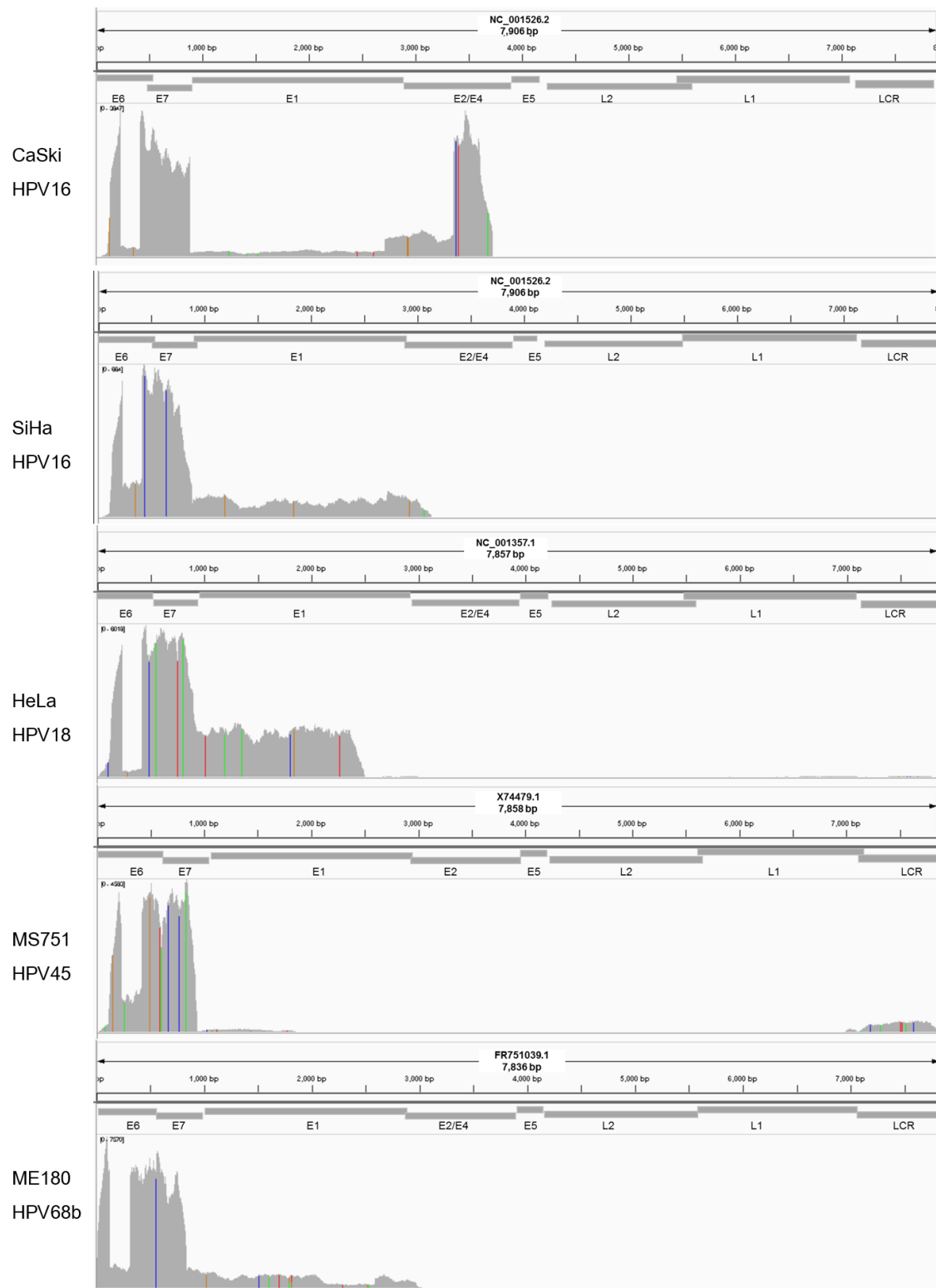
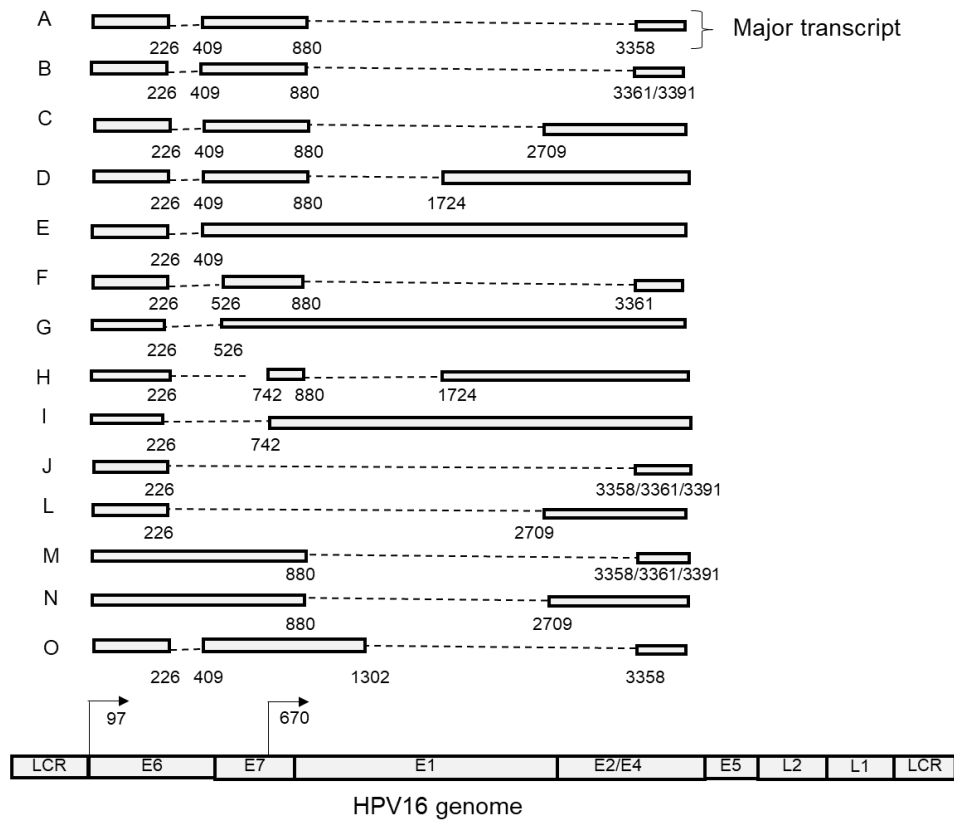
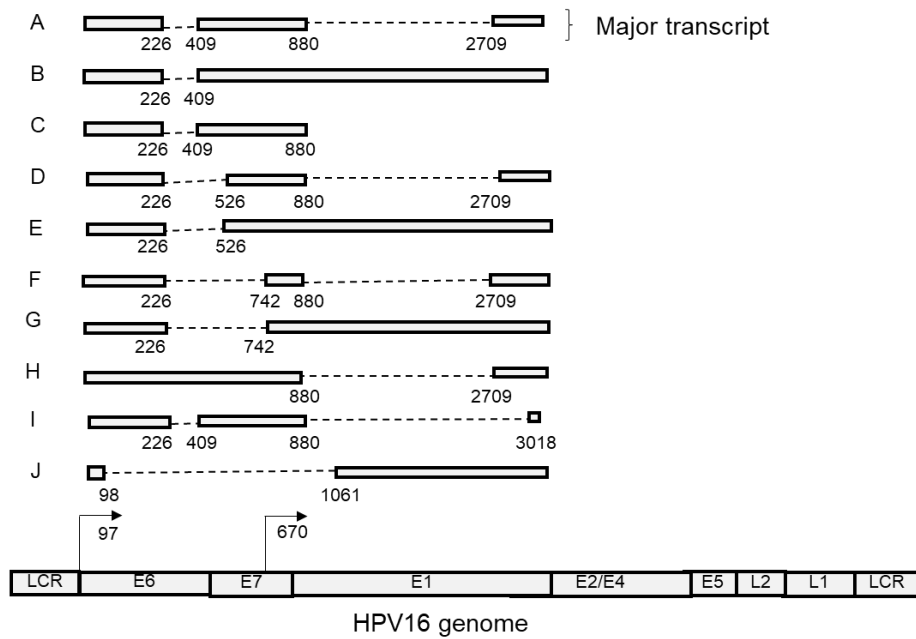


Figure 1. Read depth maps of four HR-HPV type 16, 18, 45 and 68b transcripts obtained from five different cervical cancer cell lines.

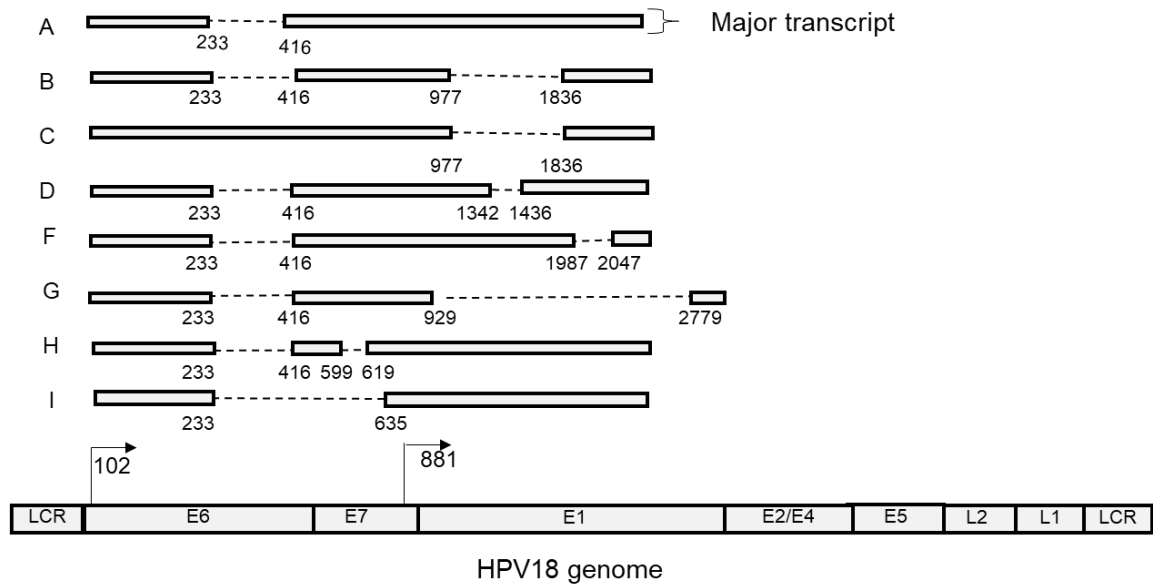
A



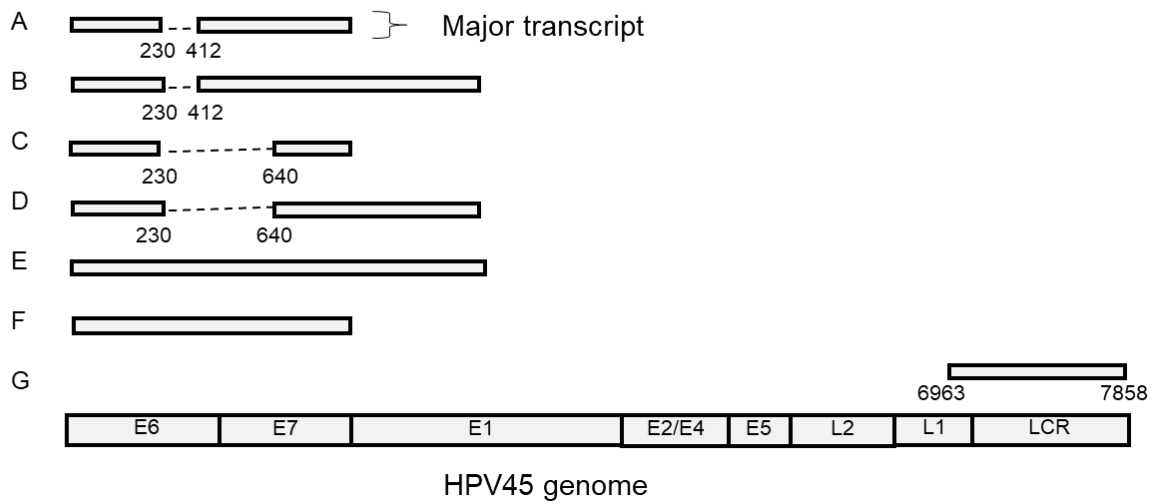
B



C



D



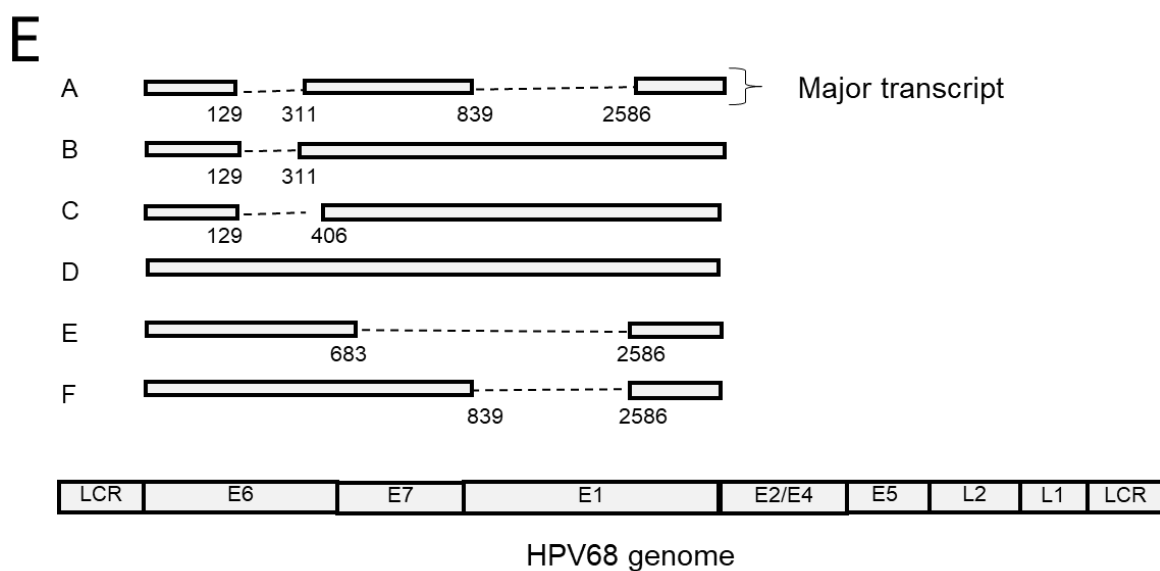


Figure 2. Map of HPV16, 18, 45 and 68b major transcripts detected in five cell lines including A. CaSki (HPV16), B. SiHa (HPV16), C. HeLa (HPV18), D. MS751 (HPV45), E. ME180 (HPV68).

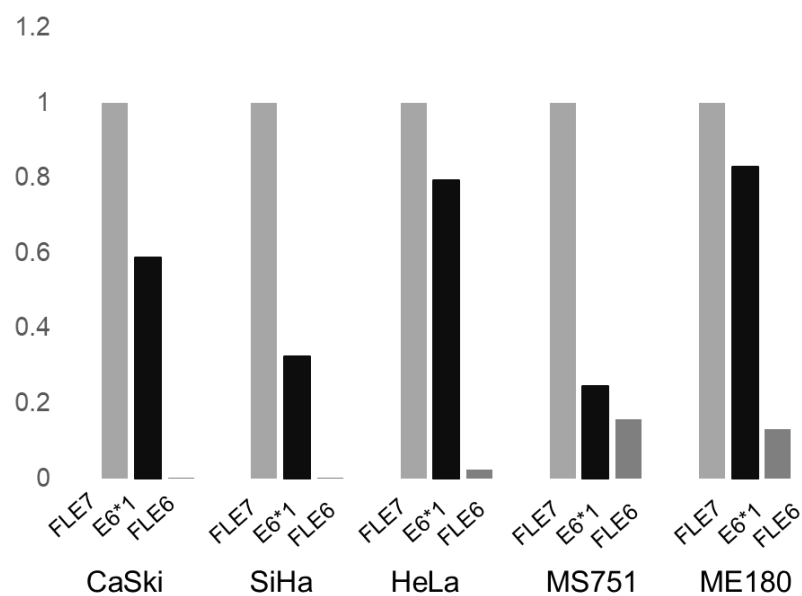


Figure 3. Validation of HPV16, 18, 45 and 68b oncogene expressions (Full length E7(FLE7), E6*1 and FLE6) detected in five cell lines by real time PCR.

Table 1 Primers for validation of HPV transcripts by real time PCR

HPV types	Target gene	Forward (5'-3')	Reverse(5'-3')
HPV16	E7 full length	CAGCTCAGAGGAGGAGGATG	GCCCATTAAACAGGTCTTCCA
	E6 full length	CATTATTGTTATAGTTTGTATGG	TGGAATCTTTGCTTTTGTCC
	E6*I409	ACTGCGACGTGAGGTGTATTAAC	TGGAATCTTTGCTTTTGTCC
	E6*II526	ACTGCGACGTGAGATCATCAAGAACACG	GTCGAGTCTCCTCCTCCTAC
	E6*X742	ACTGCGACGTGAGTGTGACTCTACG	GGTTTCTGAGAACAGATGGGGC
	E6*IV2709	ACTGCGACGTGAGGACGTGGTCCAG	TCTAGGCGCATGTGTTTCCAAT
	E6*III3356	ACTGCGACGTGAGCAGCAACGAAG	GGTCGCTGGATAGTCGTCTG
HPV18	E7 full length	TATGCATGGACCTAAGGC	GTGTGACGTTGTGGTTCGGCTCG
	E6 full length	TCTGTGTATGGAGACACATTGGA	CACTGGCCTCTATAGTGCCC
	E6*416	CTTACAGAGGTGCCTGCGG	CACTGGCCTCTATAGTGCCC
HPV45	E7 full length	CGAGTCAGAGGAGGAAAACG	CACAAAGGACAAGGTGCTCA
	E6 full length	TCTGTATATGGAGAGACACTGG	CTCGGTACTGTCCAGCTATGC
	E6*412	AACGCACAGAGGTGCCTGCGGTGCC	CTCGGTACTGTCCAGCTATGC
	E6*640	AACGCACAGAGAATGAATTAGATC	TTTTGTGACGCTGTGGTTCGGCTCG
HPV68b	E7 full length	ACCCGACCATGCAGTTAATC	TCTAGCTTCCGAGGTTCTC
	E6 full length	GGGACGGGGTACCATTAGCTGC	CATATAACTTTGTATTAGTTATGG
	E6*311	CGGACAGAGGTGCATGTGTTGCCTG	CCTCTCGTTTACTGGTCCAGC
	E6*406	CGGACAGAGGACAGTGTGCCAC	GCATGGACATAACTCTAACAC
GAPDH		GCACCGTCAAGGCTGAGAAC	ATGGTGGTGAAGACGCCAGT

References

- Ajiro, M., Tang, S., Doorbar, J. and Zheng, Z. M. Serine/Arginine-Rich Splicing Factor 3 and Heterogeneous Nuclear Ribonucleoprotein A1 Regulate Alternative RNA Splicing and Gene Expression of Human Papillomavirus 18 through Two Functionally Distinguishable cis Elements J. Virol. 2016;90:9138-52
- Attal, J., Theron, M. C., Puissant, C. and Houdebine, L. M. Effect of intercistronic length on internal ribosome entry site (IRES) efficiency in bicistronic mRNA Gene Expr. 1999;8:299-309
- Braunstein, T. H., Madsen, B. S., Gavnholt, B., Rosenstjerne, M. W., Johnsen, C. K. and Norrild, B. Identification of a new promoter in the early region of the human papillomavirus type 16 genome J. Gen. Virol. 1999;80 (Pt 12):3241-50
- Chaiwongkot, A., Vinokurova, S., Pientong, C., Ekalaksananan, T., Kongyingyoes, B., Kleebkaw, P., et al. Differential methylation of E2 binding sites in episomal and integrated HPV 16 genomes in preinvasive and invasive cervical lesions Int. J. Cancer 2013;132:2087-94
- Chen, J., Xue, Y., Poidinger, M., Lim, T., Chew, S. H., Pang, C. L., et al. Mapping of HPV transcripts in four human cervical lesions using RNAseq suggests quantitative rearrangements during carcinogenic progression Virology 2014;462-463:14-24
- Clifford, G. M., Smith, J. S., Plummer, M., Munoz, N. and Franceschi, S. Human papillomavirus types in invasive cervical cancer worldwide: a meta-analysis Br. J. Cancer 2003;88:63-73
- De-Castro Arce, J., Gockel-Krzkalla, E. and Rosl, F. Silencing of multi-copy HPV16 by viral self-methylation and chromatin occlusion: a model for epigenetic virus-host interaction Hum. Mol. Genet. 2012;21:1693-705
- Geisbill, J., Osmer, U. and Durst, M. Detection and characterization of human papillomavirus type 45 DNA in the cervical carcinoma cell line MS751 J. Gen. Virol. 1997;78 (Pt 3):655-8
- Glahder, J. A., Hansen, C. N., Vinther, J., Madsen, B. S. and Norrild, B. A promoter within the E6 ORF of human papillomavirus type 16 contributes to the expression of the E7 oncoprotein from a monocistronic mRNA J. Gen. Virol. 2003;84:3429-41

- Graham, S. V. Human papillomavirus: gene expression, regulation and prospects for novel diagnostic methods and antiviral therapies *Future microbiology* 2010;5:1493-506
- Graham, S. V. and Faizo, A. A. Control of human papillomavirus gene expression by alternative splicing *Virus Res.* 2017;231:83-95
- Johansson, C. and Schwartz, S. Regulation of human papillomavirus gene expression by splicing and polyadenylation *Nat Rev Microbiol* 2013;11:239-51
- Li, X., Johansson, C., Cardoso Palacios, C., Mossberg, A., Dhanjal, S., Bergvall, M., et al. Eight nucleotide substitutions inhibit splicing to HPV-16 3'-splice site SA3358 and reduce the efficiency by which HPV-16 increases the life span of primary human keratinocytes *PloS one* 2013;8:e72776
- Lin, K., Lu, X., Chen, J., Zou, R., Zhang, L. and Xue, X. E6-associated transcription patterns in human papilloma virus 16-positive cervical tissues *Oncology letters* 2015;9:478-82
- Milligan, S. G., Veerapraditsin, T., Ahamet, B., Mole, S. and Graham, S. V. Analysis of novel human papillomavirus type 16 late mRNAs in differentiated W12 cervical epithelial cells *Virology* 2007;360:172-81
- Munger, K., Baldwin, A., Edwards, K. M., Hayakawa, H., Nguyen, C. L., Owens, M., et al. Mechanisms of human papillomavirus-induced oncogenesis *J. Virol.* 2004;78:11451-60
- Munger, K., Basile, J. R., Duensing, S., Eichten, A., Gonzalez, S. L., Grace, M., et al. Biological activities and molecular targets of the human papillomavirus E7 oncoprotein *Oncogene* 2001;20:7888-98
- Munoz, N., Bosch, F. X., de Sanjose, S., Herrero, R., Castellsague, X., Shah, K. V., et al. Epidemiologic classification of human papillomavirus types associated with cervical cancer *N. Engl. J. Med.* 2003;348:518-27
- Ozbun, M. A. Human papillomavirus type 31b infection of human keratinocytes and the onset of early transcription *J. Virol.* 2002;76:11291-300
- Ozbun, M. A. and Meyers, C. Human papillomavirus type 31b E1 and E2 transcript expression correlates with vegetative viral genome amplification *Virology* 1998;248:218-30
- Paget-Bailly, P., Meznad, K., Bruyere, D., Perrard, J., Herfs, M., Jung, A. C., et al. Comparative RNA sequencing reveals that HPV16 E6 abrogates the effect of E6*I on ROS metabolism *Scientific reports* 2019;9:5938
- Schmitt, M., Dalstein, V., Waterboer, T., Clavel, C., Gissmann, L. and Pawlita, M. The HPV16 transcriptome in cervical lesions of different grades *Mol. Cell. Probes* 2011;25:260-5
- Schwartz, S. Papillomavirus transcripts and posttranscriptional regulation *Virology* 2013;445:187-96
- Seedorf, K., Oltersdorf, T., Krammer, G. and Rowekamp, W. Identification of early proteins of the human papilloma viruses type 16 (HPV 16) and type 18 (HPV 18) in cervical carcinoma cells *EMBO J.* 1987;6:139-44
- Smotkin, D., Prokoph, H. and Wettstein, F. O. Oncogenic and nononcogenic human genital papillomaviruses generate the E7 mRNA by different mechanisms *J. Virol.* 1989;63:1441-7
- Stubenrauch, F. and Laimins, L. A. Human papillomavirus life cycle: active and latent phases *Semin. Cancer Biol.* 1999;9:379-86
- Stunkel, W. and Bernard, H. U. The chromatin structure of the long control region of human papillomavirus type 16 represses viral oncoprotein expression *J. Virol.* 1999;73:1918-30
- Tang, S., Tao, M., McCoy, J. P., Jr. and Zheng, Z. M. The E7 oncoprotein is translated from spliced E6*I transcripts in high-risk human papillomavirus type 16- or type 18-positive cervical cancer cell lines via translation reinitiation *J. Virol.* 2006;80:4249-63
- Toots, M., Mannik, A., Kivi, G., Ustav, M., Jr., Ustav, E. and Ustav, M. The transcription map of human papillomavirus type 18 during genome replication in U2OS cells *PloS one* 2014;9:e116151
- Wang, X., Meyers, C., Wang, H. K., Chow, L. T. and Zheng, Z. M. Construction of a full transcription map of human papillomavirus type 18 during productive viral infection *J. Virol.* 2011;85:8080-92
- Wu, C., Kajitani, N. and Schwartz, S. Splicing and Polyadenylation of Human Papillomavirus Type 16 mRNAs *International journal of molecular sciences* 2017;18:
- Zheng, Z. M. Viral oncogenes, noncoding RNAs, and RNA splicing in human tumor viruses *International journal of biological sciences* 2010;6:730-55
- Zheng, Z. M., Tao, M., Yamanegi, K., Bodaghi, S. and Xiao, W. Splicing of a cap-proximal human Papillomavirus 16 E6E7 intron promotes E7 expression, but can be restrained by distance of the intron from its RNA 5' cap *J. Mol. Biol.* 2004;337:1091-108