



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Maestría y Doctorado en Ciencias Bioquímicas

“SOX2 como un nuevo modulador de la actividad transcripcional de la Región Larga de Control (LCR) del Virus de Papiloma Humano tipo 16 (VPH16)”

TESIS

QUE PARA OPTAR POR EL GRADO DE:

Doctor en Ciencias

PRESENTA:

M. en C. Imelda Martínez Ramírez

TUTOR PRINCIPAL

Dra. Marcela Lizano Soberón

[Instituto de Investigaciones Biomédicas](#)

MIEMBROS DEL COMITÉ TUTOR

Dr. Alejandro García Carrancá

[Instituto de Investigaciones Biomédicas](#)

Dr. Alejandro Zentella Dehesa

[Instituto de Investigaciones Biomédicas](#)

Ciudad de México, noviembre, 2018



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Este trabajo se llevó a cabo en el Laboratorio de Epidemiología y Biología Molecular de Virus Oncogénicos, ubicado en el Instituto Nacional de Cancerología (INCan)/Investigación Biomédica en Cáncer. Con la asesoría de la Dra. Marcela Lizano Soberón perteneciente a la Unidad de Investigación Biomédica en Cáncer, Instituto Nacional de Cancerología (INCan)/Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México (UNAM). Para la realización de este trabajo se recibió apoyo del proyecto INCan 008/058/IBI con número de CONACyT 60722.

Hijo mío,
Si recibieres mis palabras,
Y mis mandamientos guardares dentro de ti,
Haciendo estar atento tu oído a la sabiduría;
Si inclinares tu corazón a la prudencia,
Si clamares a la inteligencia,
Y a la prudencia dieres tu voz;
Si como a la plata la buscares,
Y la escudriñares como a tesoros,
Entonces entenderás en temor de Jehová.

Proverbios 2:1-5

AGRADECIMIENTOS

- * A Dios, por la vida que me ha regalado, por la familia que me ha dado, por todas las personas que ha puesto delante de mí y por cada tropiezo que ha convertido en victoria.

- * A mi amado esposo y a mis amados hijos a quienes he sacrificado tanto y siguen ahí, apoyándome y le han dado sentido a mi vida.

- * A mis amados padres, por la vida que me dieron, por todos y cada uno de sus sacrificios y por cada enseñanza que me han dado.

- * A mis hermanos y sobrinos. Por su apoyo y aliento.

- * A la Dra. Marcela Por la confianza que ha depositado en mí.

- * A Adela e Irma, por su gran apoyo.

- * A los integrantes de mi comité tutorial por su guía en todo el camino del doctorado.

- * A los integrantes de mi comité de evaluación, por su tiempo dedicado en mejorar mi trabajo.

- * A todos mis compañeros de laboratorio por su paciencia y por todas sus valiosas aportaciones.

- * A todos aquellos que han aportado para que mi trabajo se vea finalmente culminado.

AGRADECIMIENTOS

- * Al consejo Nacional de Ciencia y Tecnología (CONACyT), por el sustento otorgado durante este tiempo.
- * Al Programa de Apoyo a los Estudios de Posgrado (PAEP), por el apoyo brindado para compartir, con otros compañeros, el trabajo que hemos realizado.
- * A la Universidad Nacional Autónoma de México (UNAM), por la oportunidad, que me ha dado, de seguir aprendiendo.
- * Al Instituto Nacional de Cancerología (INCan), por todo lo que me ha brindado.

AGRADECIMIENTOS

Unos reglones no me bastan para agradecerle Dra. Marcela, por la confianza que ha depositado en mí. Por creer que en algún momento lograría culminar con esta tarea que me fue asignada.

Gracias por todo lo que he podido aprender de usted.

Es un gran ejemplo a seguir.

Gracias por todos sus excelentes consejos.

Gracias por su excelente dirección.

Gracias también por enseñarme que el conocimiento no está peleado con la calidad humana.

DEDICATORIA

Dedico este trabajo con todo mi corazón:

A mi esposo y a mis hijos,

A mis Padres,

A la Dra. Marcela,

A toda mi gran familia que engloba también a mis amigos.

Y

A mí.

Con el apoyo de Dios y de todos ustedes,

¡Finalmente lo he logrado!

ÍNDICE

RESUMEN.....	1
SUMMARY.....	3
1. INTRODUCCIÓN.....	5
1.1. Características del cáncer.....	5
1.2. El cáncer cérvicouterino.....	6
1.2.1. Factores de riesgo asociados al desarrollo de CaCU.....	8
1.2.2. Tipos de CaCU.....	9
1.3. El virus de Papiloma humano.....	11
1.3.1. Generalidades.....	11
1.3.2. Estructura.....	12
1.3.3. Ingreso del VPH a la célula y al núcleo.....	15
1.3.4. El ciclo viral replicativo del VPH.....	16
1.3.5. Integración del genoma del VPH.....	18
1.4. La región larga de control.....	19
1.5. Regulación transcripcional de <i>E6</i> y <i>E7</i> por <i>E2</i>	21
1.6. Regulación transcripcional del VPH por factores transcripcionales celulares.....	23
2. ANTECEDENTES.....	26
2.1. Análisis <i>in silico</i> con el programa JASPAR CORE.....	26
2.2. El factor transcripcional SOX2.....	27
3. JUSTIFICACIÓN.....	32
4. HIPÓTESIS.....	32
5. OBJETIVO GENERAL.....	33
6. OBJETIVOS ESPECÍFICOS.....	33
7. METODOLOGÍA.....	34
7.1. Cultivo de células.....	34

7.2. Construcción de plásmidos.....	34
7.3. Infección con lentivirus recombinantes.....	35
7.4. Ensayos de transfección y actividad transcripcional.....	35
7.5. Análisis <i>in silico</i>	36
7.6. Ensayo de cambio en el movimiento electroforético (EMSA).....	36
7.7. Ensayo de inmunoprecipitación de la cromatina (ChIP).....	37
7.8. Ensayos de mutagénesis dirigida multi-sitio	38
7.9. Western blots.....	39
7.10. Aislamiento de RNA y síntesis de cDNA.....	39
7.11. PCR cuantitativa.....	40
7.12. Análisis estadístico.....	40
8. RESULTADOS.....	41
8.1. La LCR del VPH16 y del 18 muestran diferente actividad transcripcional en líneas celulares de origen escamoso o de origen glandular.....	41
8.2. SOX2 se une a la LCR del VPH16.....	43
8.2.1. SOX2 se une a sus sitios putativos de unión en la LCR del VPH16.	43
8.2.2. SOX2 se une a la LCR del VPH16 <i>in vivo</i>	46
8.3. SOX2 reprime la actividad transcripcional de la LCR del VPH16.....	47
8.4. Los sitios S1, S2 y S3 de la LCR del VPH16 son necesarios para la represión mediada por SOX2.....	48
8.5. La expresión endógena de <i>E6</i> y <i>E7</i> es regulada por SOX2.....	50
9. DISCUSIÓN.....	52
10. CONCLUSIONES.....	58
11. PERSPECTIVAS.....	59
12. REFERENCIAS.....	60

ÍNDICE DE FIGURAS Y TABLAS

Figura 1. Características de una célula cancerosa.....	5
Figura 2. Mortalidad e incidencia a nivel mundial y en México.....	7
Figura 3. Anatomía del aparato genital femenino y la unión escamo-columnar	8
Figura 4. Organización del genoma del VPH16.....	14
Figura 5. Mecanismo de entrada del VPH a la célula que infecta.....	16
Figura 6. Ciclo replicativo del VPH.....	17
Figura 7. Región larga de control (LCR) del VPH16.....	21
Figura 8. Regulación de la expresión de los genes E6 y E7 por la proteína E2 durante el proceso de integración del genoma viral al genoma celular.....	23
Figura 9. Estructura y organización de los miembros del grupo SOXB1.....	28
Figura 10. Regulación de SOX2.....	29
Figura 11. Las LCRs de los VPH16 y 18 presentan una actividad transcripcional diferente en células con un origen histológico distinto.....	43
Figura 12. SOX2 se une a los elementos de respuesta putativos en la LCR del VPH16.....	45
Figura 13. SOX2 se une a los sitios de unión putativos en la LCR del VPH16 in vivo.....	47
Figura 14. SOX2 reprime la actividad transcripcional de la LCR del VPH16....	48
Figura 15. Electroferogramas donde se identifican los nucleótidos mutados en la región larga de control (LCR) del VPH16 por mutagénesis dirigida multi-sitio	49
Figura 16. Los sitios S1, S2 y S3 de la LCR del VPH16 tienen un papel en la actividad represora de SOX2.....	50
Figura 17. SOX2 reprime la expresión de los oncogenes virales E6 y E7 endógenos.....	51

Tabla 1. Proteínas del virus de Papiloma humano y algunas de sus principales funciones	14
Tabla 2. Secuencias consenso de unión a factores de transcripción en la LCR del VPH16 y VPH18.....	24
Tabla 3. Predicción <i>in silico</i> de sitios de unión a factores de transcripción para la LCR del VPH16 o 18 empleando la base de datos JASPAR CORE.....	27
Tabla 4. Oligonucleótidos utilizados para introducir las mutaciones en los sitios putativos 1, 2 y 3 de unión a SOX2 de la LCR del VPH16.....	38
Tabla 5. Oligonucleótidos empleados para el análisis por qPCR de un fragmento de los genes <i>E6</i> y <i>E7</i> y de un fragmento de la subunidad 18S del ribosoma que es un gen constitutivo.....	40

ABREVIACIONES

VPH-AR	Virus del Papiloma Humano de alto riesgo
CaCU	Cáncer cervicouterino
CCE o SCC	Carcinomas de células escamosas o squamous cell carcinomas
AC	Adenocarcinomas
LCR	Por sus siglas en inglés, Long Control Region (Región larga de control)
FT o TF	Factores de transcripción o transcription factors
EMSA	Por sus siglas en inglés, electrophoretic mobility shift assay (Ensayo de cambio en el movimiento electroforético)
ChIP	Por sus siglas en inglés, Chromatin Immunoprecipitation (Inmunoprecipitación de la cromatina)
OMS	Organización Mundial de la Salud
DES	Dietilestilbestrol
NETs	Por sus siglas en inglés, neuroendocrine tumors (tumores neuroendócrinos)
NICs	Neoplasias Intraepiteliales Cervicales
EGFR	Por sus siglas en inglés, epidermal growth factor receptor (Receptor del factor de crecimiento epidérmico)
TGN	Por sus siglas en inglés, Trans-Golgi network (Red Trans Golgi)
PBD	Por sus siglas en inglés, PDZ Binding Domain (Dominio de unión a PDZ)
Dlgl	Disk large homolog 1
GRE	Por sus siglas en inglés, glucocorticoids response elements (elementos de respuesta a glucocorticoides)
SOX	Sex determining region Y-box2
AEG	Anoftalmia esofágica genital
SDS	Por sus siglas en inglés, sodium dodecyl sulfate (dodecilsulfato de sodio)
EE	Error Estandar
SQ	Por sus siglas en inglés, squamous (Escamoso)
NS	Por sus siglas en inglés, non specific (No específica)

CS	Por sus siglas en inglés, consensus sequence (Secuencia Consenso)
³² P	Por sus siglas en inglés, ³² Phosphorus (³² Fosforo)
SS	Por sus siglas en inglés, super shift (super retardo)
GAPDH	Por sus siglas en inglés, Glyceraldehyde-3-Phosphate Dehydrogenase (Gliceraldehido-3-fosfato deshidrogenasa)
GFP	Por sus siglas en inglés, Green fluorescent protein (proteína verde fluorescente)
M1	Mutación en sitio 1
M2	Mutación en sitio 2
M3	Mutación en sitio 3
M1,2,3	Mutación en sitio 1,2,3
TFBS	Por sus siglas en inglés, Transcription Factor Binding Sites (Sitios de unión a factores de transcripción)

RESUMEN

Las infecciones persistentes con el Virus del Papiloma Humano de alto riesgo (VPH-AR) constituyen el principal factor de riesgo para el desarrollo del cáncer cérvicouterino (CaCU). El VPH16 es el tipo más frecuente asociado a carcinomas de células escamosas (CCE), seguido del VPH18 el cual se asocia con más frecuencia a los Adenocarcinomas (AC), segundo tipo más común de CaCu.

La región larga de control (LCR, por sus siglas en inglés Long Control Region) en el genoma del VPH contiene el origen de replicación y secuencias reconocidas por factores de transcripción (FT) celulares y virales que controlan la transcripción del virus. La alteración de la expresión de los oncogenes virales *E6* y *E7*, modulada por la LCR, causa modificaciones en vías celulares, como la de proliferación, conduciendo a una transformación maligna. Mediante análisis bioinformático se analizó la LCR del VPH16 y del 18 utilizando la base de datos JASPAR CORE e identificamos tres elementos de respuesta para el factor transcripcional SOX2 exclusivos en la LCR del VPH16. Dado el papel de SOX2 como regulador de varios procesos celulares importante como por ejemplo pluripotencia y diferenciación celular escamosa, así como por su participación en carcinogénesis, nos interesó determinar su efecto en la regulación transcripcional del VPH16.

Conocer la regulación transcripcional de los VPH-AR en distintos contextos celulares, ayudará a entender su participación en la carcinogénesis cervical.

OBJETIVO: El objetivo de este trabajo fue determinar el papel de SOX2 en la regulación transcripcional de la LCR del VPH16 en distintos contextos celulares.

METODOLOGÍA: Mediante ensayos con vectores reporteros en líneas celulares con origen escamoso (CaSki, HaCaT) y adenocarcinoma (HeLa), se determinó la actividad transcripcional basal de la LCR del VPH16 comparada con la del 18 así como el papel de la sobreexpresión de SOX2 sobre la actividad transcripcional de las mismas LCRs. Mediante ensayo de cambio en el movimiento electroforético (EMSA, por sus siglas en inglés

electrophoretic mobility shift assay) y ensayos de inmunoprecipitación de la cromatina (ChIP, por sus siglas en inglés Chromatin Immunoprecipitation), se determinó la unión de SOX2 a los tres elementos de respuesta identificados en la LCR del VPH16. Por medio de mutagénesis dirigida multi-sitio, se determinó el papel y contribución de cada uno de los sitios de unión a SOX2. Finalmente, mediante ensayos de RT-qPCR se corroboró el papel de SOX2 sobre sus blancos directos *E6* y *E7*.

RESULTADOS: Demostramos que SOX2 se une a la LCR del VPH16 mediante ensayos *in vitro* (EMSA) e *in vivo* (ChIP). La sobreexpresión de este FT reprimió la actividad transcripcional de la LCR del VPH16, como se mostró mediante ensayos con genes reporteros de luciferasa y mediante la regulación negativa de oncogenes endógenos del VPH (*E6* y *E7*). Los ensayos de mutagénesis dirigida revelaron que los tres sitios putativos de unión a SOX2 están involucrados en la regulación de la actividad de la LCR del VPH16.

CONCLUSIONES: El factor transcripcional SOX2 reconoce y se une a sus tres elementos de respuesta ubicados en la LCR del VPH16 y actúa como un represor transcripcional disminuyendo la expresión de los oncogenes *E6* y *E7* en un contexto de CCE.

Los tres elementos de respuesta a SOX2 sobre la LCR del VPH16 conforman una entidad represora.

SUMMARY

Persistent infections with high-risk human papillomavirus (HR-HPV) constitute the main risk factor for the development of cervicouterine cancer (CaCU). HPV16 is the most frequent type associated with squamous cell carcinomas (SCC), followed by HPV18 which is most frequently associated with Adenocarcinomas (AC), the second most common type of CaCU.

The long control region (LCR) in the HPV genome contains the origin of replication and sequences recognized by cellular and viral transcription factors (TFs) controlling viral transcription. Altered expression of *E6* and *E7* viral oncogenes, modulated by the LCR, causes modifications in cellular pathways, such as proliferation, leading to malignant transformation. Bioinformatics analysis of HPV16-LCR and the HPV18-LCR, using the JASPAR CORE database, led to the identification of three putative response elements for the transcriptional factor SOX2, found exclusively in the HPV16-LCR. Due to the SOX2 participation as a regulator of several cellular processes, for example, pluripotency and squamous cellular differentiation as well as its participation in carcinogenesis, we became interested in elucidating the role of SOX2 in the HPV16-LCR transcriptional regulation.

The study of HR-HPV transcriptional regulation in distinct cellular contexts could help to understand their role in cervical carcinogenesis

OBJECTIVE: The objective of this work was to determine the role of SOX2 in the transcriptional regulation of the LCR of HPV16 according to the cell context.

METHODOLOGY: Through reporter vectors assays, in cell lines with squamous origin (CaSki and HaCaT) and adenocarcinoma (HeLa), the basal transcriptional activity of the HPV 16 LCR was compared with that of 18, as well as the role of the overexpression of SOX2 on the LCR16 and 18 activities. Using Electrophoretic mobility shift assays (EMSAs) and Chromatin immunoprecipitation (ChIP), we determined the binding of SOX2 to the three response elements located in the LCR of the HPV16. By means of directed mutagenesis multi-site, the role and contribution of each of the SOX2 binding sites was determined. Finally, using RT-qPCR assays, we corroborated the role of SOX2 on its direct targets *E6* and *E7*.

RESULTS: We demonstrated the binding of SOX2 to the HPV16-LCR through *in vitro* (EMSA) and *in vivo* (ChIP) assays. Overexpression of SOX2 repressed the transcriptional activity of the HPV16-LCR, as shown through reporter plasmid assays and by down-regulation of endogenous HPV oncogenes (*E6* and *E7*). Multi-site directed mutagenesis assays revealed that all three putative SOX2 binding sites are involved in the repression of the HPV16-LCR activity.

CONCLUSIONS: The transcriptional factor SOX2 recognizes and binds to the three response elements located in the LCR of the HPV16 and acts as a transcriptional repressor decreasing the expression of oncogenes *E6* and *E7* in a SCC context.

The three SOX2 response elements on the HPV16-LCR constitute a repressor entity.

1. INTRODUCCIÓN

1.1. Características del cáncer

El cáncer es un conjunto de enfermedades relacionadas entre sí que se origina por el resultado de daños genéticos y epigenéticos acumulados en el genoma ¹. La tumorigénesis en humanos es un proceso complejo que dirige la transformación progresiva de una célula normal humana a derivados malignos, cuyo proceso es análogo a la evolución Darwiniana, en la que cambios genéticos sucesivos conducen a una conversión gradual de la célula normal humana a célula cancerosa. Muy interesantemente, se ha determinado que el microambiente tumoral, formado por células del estroma asociadas al tumor tienen una participación activa en la tumorigénesis, contribuyendo al desarrollo y a la expresión de ciertas capacidades distintivas de la célula tumoral ².

Hanahan D y cols. en 2011 sugieren que una célula cancerosa se caracteriza por presentar ocho alteraciones esenciales en su fisiología, que son: Autosuficiencia en señales de crecimiento o señalización proliferativa sostenida, insensibilidad a señales inhibitorias del crecimiento, evasión a la muerte celular programada (apoptosis), potencial replicativo ilimitado, angiogénesis sostenida, invasión de tejidos y metástasis, reprogramación del metabolismo energético y evasión de la destrucción inmunológica, las cuales se ilustran en la **Figura 1** ².

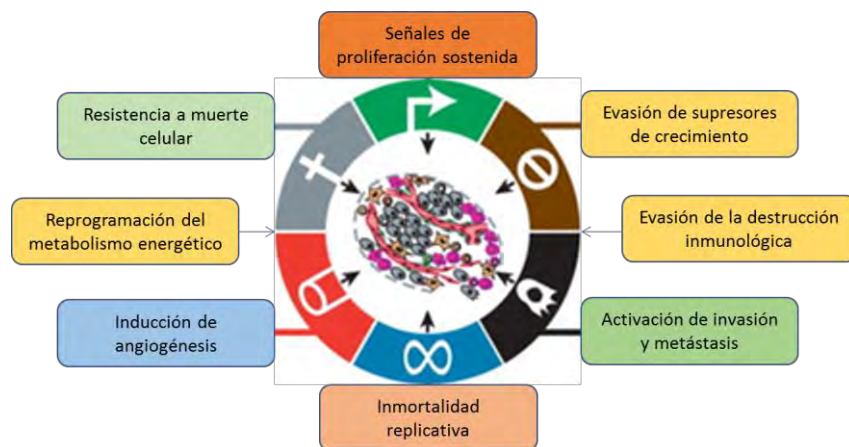


Figura 1. Características de una célula cancerosa. Las células cancerosas van adquiriendo características que se mantienen en prácticamente todas o en la mayoría de estas células como son: autosuficiencia en señales de crecimiento o señalización proliferativa sostenida, insensibilidad a señales inhibitorias del crecimiento, evasión a la muerte celular programada (apoptosis), potencial replicativo ilimitado, angiogénesis sostenida, invasión de tejidos y metástasis, reprogramación del metabolismo energético y evasión de la destrucción inmunológica (Modificada de Hanahan D and Weinberg RA., 2000) ³.

El cáncer es la segunda causa de muerte en el mundo ⁴. Por cada seis muertes una es ocasionada por cáncer y la mayoría ocurre en países de ingresos medios y bajos. Según la Organización Mundial de la Salud (OMS), el índice de masa corporal elevado, la ingesta reducida de frutas y verduras, la falta de actividad física, el consumo de tabaco y alcohol son los principales factores de riesgo relacionados con la conducta y dieta que ocasionan un tercio de muerte por cáncer. Según la OMS, el tabaquismo es el principal factor de riesgo y es el responsable del 22% de las muertes por cáncer, mientras que el 25% de los casos de cáncer es ocasionado por las infecciones oncogénicas, por ejemplo, aquellas ocasionadas por el virus de la hepatitis o por Papilomavirus humanos ⁴.

1.2. El cáncer cérvicouterino

El cáncer cérvicouterino (CaCU) es el cuarto cáncer más común en mujeres en el mundo y ocupa la segunda causa de muerte por cáncer en mujeres en México ⁵ (**Figura 2**). En el año 2018, a nivel mundial, se presentan alrededor de 569,847 nuevos casos mientras que aproximadamente el número de muertes por este cáncer alcanzaron los 311,365 casos ⁵. Las gráficas de la **figura 2** muestran la incidencia y la mortalidad del CaCU a nivel mundial y en México, en relación con los cánceres de mayor frecuencia.

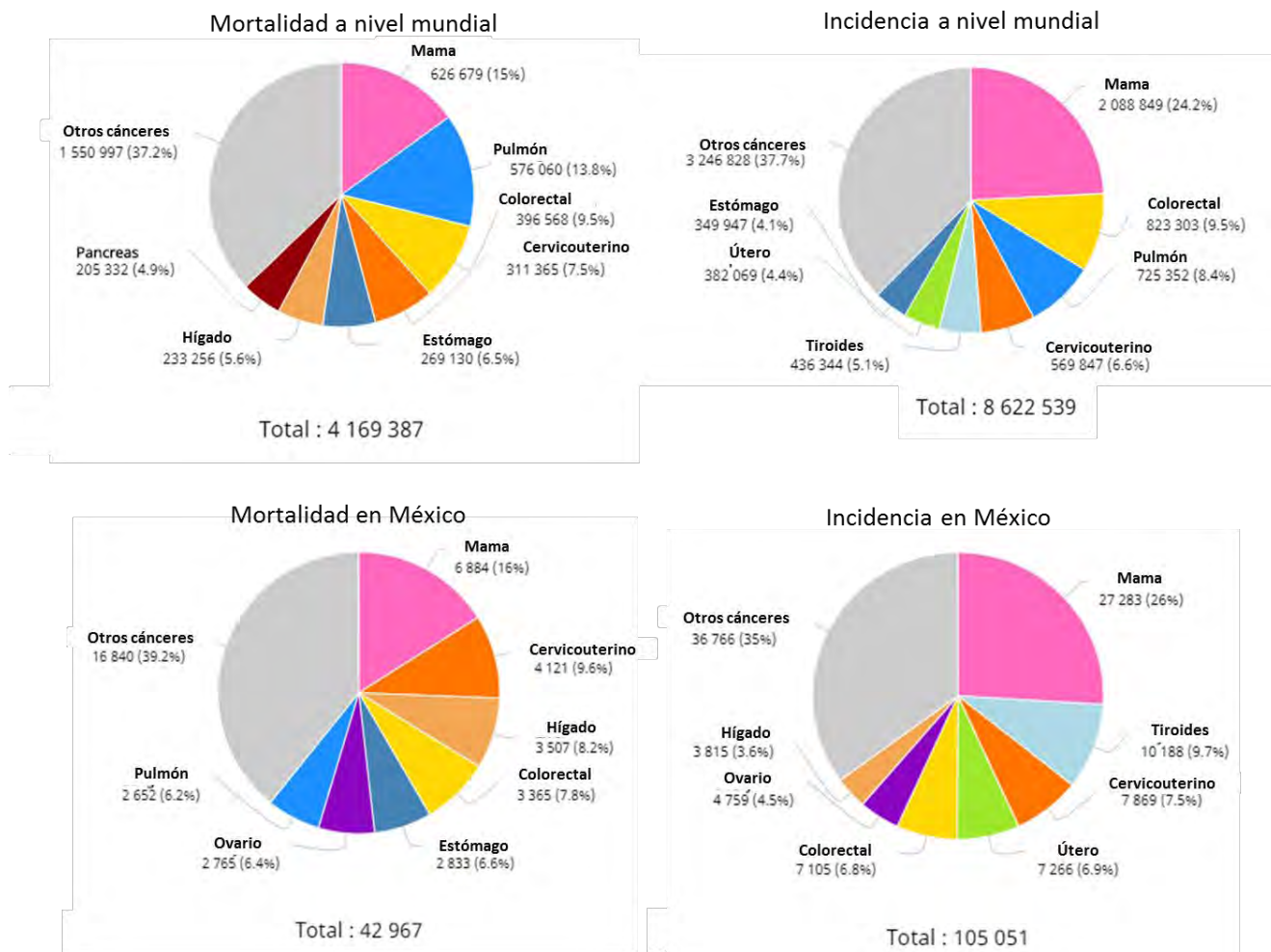


Figura 2. Mortalidad e incidencia a nivel mundial y en México por cáncer en mujeres. En la parte superior se ilustra la mortalidad y la incidencia por cáncer en mujeres a nivel mundial. En la parte inferior se representan los niveles de mortalidad e incidencia en México (Tomado de Globocan 2018) ⁵.

El CaCU se origina en la unión escamocolumnar del cuello uterino o cérvix. La unión escamocolumnar es el sitio de unión del epitelio glandular que recubre el endocérvix (canal endocervical) con el epitelio estratificado del exocérvix (porción que protruye dentro de la vagina). La posición de esta unión varía dependiendo del estado hormonal, la anatomía cervical y la edad. En la mayoría de las mujeres, la acidificación del pH vaginal que se da en la adolescencia induce el reemplazamiento de una porción del epitelio columnar endocervical por un epitelio escamoso metaplásico conocido como zona de transformación ⁶(Figura 3).

Herfs Michael y cols. (2012), describieron una pequeña población de células en la unión escamocolumnar del cérvix con un perfil de expresión que podría ser responsable para la mayoría de carcinomas cervicales asociados al VPH ⁶. A diferencia de los tumores benignos o no cancerosos como los fibromas y adenomas que no crecen de manera ilimitada y agresiva, las células tumorales malignas invaden y destruyen tejido cercano saludable o viajan a otras partes del cuerpo a través de vasos linfáticos y sanguíneos. En el caso del cáncer de cérvix este se puede difundir más frecuentemente al intestino, vaso, pulmón e hígado ⁷.

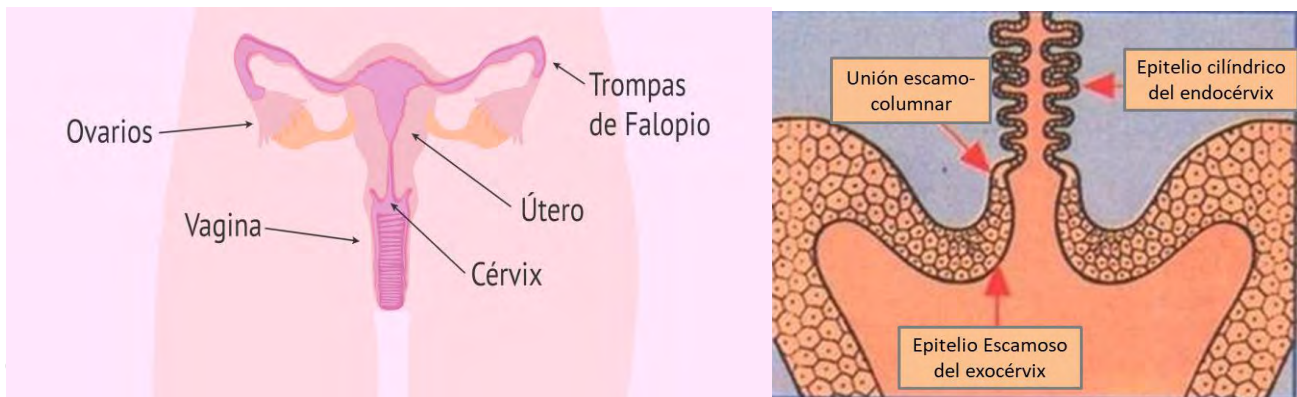


Figura 3. Anatomía del aparato genital femenino y la unión escamo-columnar. A la izquierda se muestra la anatomía del aparato reproductor femenino formado por el útero, las trompas de Falopio y ovarios, así como la vagina y el cérvix, señalado con flechas. A la derecha se muestra el cérvix, constituido por endocérvix, exocérvix y la unión escamo-columnar. El endocérvix está formado por células de tipo glandular o cilíndricas, el exocérvix por células de tipo escamosos y la unión escamo-columnar que corresponde a la unión de los dos epitelios. (Tomadas y modificadas de <https://www.reproduccionasistida.org/fertilidad-en-la-mujer/fertilidad-mujer/> ⁸ y <https://www.eurocytology.eu/es/course/465> ⁹, respectivamente).

1.2.1. Factores de riesgo asociados al desarrollo de CaCU

Existen varios factores de riesgo para el desarrollo de CaCU, como:

- 1.- Número elevado de partos, asociados a infección con el VPH ¹⁰.
- 2.- Tabaquismo e infección por el VPH ¹¹.
- 3.- Consumo por largo tiempo de anticonceptivos orales e infección con el VPH ^{12,13}.
- 4.- Inmunodepresión ^{14,15}.
- 5.- Inicio de relaciones sexuales a edad temprana ¹⁶.

6.- Tener varias parejas sexuales ¹⁶.

7.- Haber estado expuesta al compuesto dietilestilbestrol (DES) en el vientre de la madre ¹⁷.

Pero la infección persistente con el virus de Papiloma humano de alto riesgo (VPH-AR) es el principal factor de riesgo ^{18,19,20,21}.

La mayoría de las infecciones por estos virus son eliminadas por el sistema inmunológico del hospedero en aproximadamente 18 meses ²²; sin embargo, del 10-15 % de las mujeres no eliminan la infección y por lo tanto se mantiene como persistente ²¹. Finalmente, un pequeño porcentaje de estas mujeres posee una mayor probabilidad de desarrollar cáncer, favorecido por la persistencia viral, probablemente la carga viral, la integración del DNA viral al genoma celular, el papel de variantes virales (cambios en la secuencia de un mismo tipo viral, menores al 2%), etc. La progresión de lesiones epiteliales infectadas con VPH de alto riesgo a cáncer invasor es un proceso largo que involucra la acumulación de alteraciones en el DNA de genes celulares en el hospedero así como cambios tanto genéticos como epigenéticos en oncogenes y genes supresores de tumores que afectan la expresión global ²³.

1.2.2. Tipos de CaCU

Los cánceres son nombrados dependiendo de las áreas en la que se desarrollan. En el caso del cérvix su superficie contiene dos tipos diferentes de células: las células epiteliales escamosas y las células epiteliales columnares. Las células epiteliales escamosas que recubren el exocervix son delgadas y planas y se amontonan una sobre otra mientras que las células columnares tienen una morfología alargada. Una teoría dice que si el cáncer cervical se desarrolla a partir de células epiteliales escamosas, se origina un carcinoma de células escamosas, y que si se desarrolla a partir de células epiteliales columnares se origina un adenocarcinoma ⁷. En contraste con la teoría previamente descrita, Herfs Michael y cols. (2012) quienes proponen que estos tipos de cáncer se originan de un solo grupo de células ⁶.

El CaCU frecuentemente se presenta como carcinoma de células escamosas (CCE), que corresponde al 70-80% de todos los casos de cáncer de cérvix, seguido por los adenocarcinomas (AD), con un origen glandular (más del 10% de los casos aproximadamente) ^{24,25}. Otros tipos de CaCU mucho menos frecuentes son los tumores adeno-escamosos que representan del 5-6% de los cánceres cervicales y están compuestos por una mezcla de elementos glandulares y escamosos ^{25,26}, y los carcinomas de células pequeñas que crecen rápidamente y representan aproximadamente un 3% de los cánceres cervicales. Los tumores neuroendócrinos (NETs, por sus siglas en inglés Neuroendocrine tumors), dentro de los tumores de células pequeñas, representan del 0.9-1.5 % de tumores de cérvix uterino y se han considerado una entidad única en cáncer ginecológico por tener peor pronóstico y alta recurrencia, comparado con otros tipos de tumores de cérvix. Los NETs presentan una diseminación general en un estado temprano de su desarrollo por lo que este carcinoma de células pequeñas de cérvix es uno de los tumores más agresivos del tracto genital femenino por su rápida recurrencia y su asociación con alta mortalidad ^{25,27,28}.

Carcinomas de células escamosas y adenocarcinomas de cérvix

Los carcinomas de células escamosas, CCE, son el tipo histológico más abundante de cáncer cérvico-uterino seguido del adenocarcinoma, AD. Los CCE se desarrollan a partir de las Neoplasias Intraepiteliales Cervicales (NICs) las cuales, dependiendo de la proporción del tejido afectado, se han clasificado de leve (NIC1) a severa (NIC3) ²³. NIC2 y NIC3 se han considerado como lesiones de alto grado. Poco se sabe acerca de las lesiones precursoras de los AD además del AD *in situ*²³. En cuanto a análisis epidemiológicos, se ha determinado que el tabaquismo se asocia más con CCE que con AD; mientras que éstos últimos se asocian más con otros factores comunes como multiparidad ²⁹ y obesidad ³⁰. Histológicamente los AD tienen una distribución muy variada de tipos, el más común es el endocervical, y hay otras variedades como el enteroide, villoglandular, endometroide, de células claras, el seroso, etc.

Los CCE en su mayoría son los no-queratinizante ³¹, hay otros tipos como los queratinizantes de células grandes, los no queratinizantes de células grandes y los no queratinizantes de células pequeñas. También se pueden encontrar otras variedades menos comunes como los basaloides, verrucosos, condilomatosos, papilar, tipo linfoepitelioma y escamoso-transicional Según los lineamientos de la Red Nacional Integral de Cáncer (NCCN), no hay diferencias en las estrategias de tratamiento entre CCE y AD aunque varios estudios sugieran que los AD tienen un peor pronóstico ³¹.

1.3. El Virus de Papiloma Humano

1.3.1. Generalidades

Los VPH son pequeños virus no envueltos de 50 nm, de doble cadena de DNA circular, de aproximadamente 8 pares de kb que se encuentra dentro de una cápside de proteínas, pertenecen a la familia *Papillomaviridae* y específicamente infectan el epitelio. Los Papilomavirus humanos se pueden clasificar en 5 géneros conocidos: Alfa, Beta, Gama, Mu y Nu de los cuales el género beta se caracteriza por infectar principalmente piel y el género alfa, mucosas ³².

A la fecha se han descrito más de 200 tipos del VPH con preferencia por epitelio cutáneo o mucoso ³³. Aproximadamente 40 tipos de VPH infectan el tracto genital y se han clasificado en tipos de bajo y de alto riesgo, en base a su potencial oncogénico ^{20,34}.

Cabe señalar que los tipos de VPH16 y 18 son los responsables de aproximadamente el 80 % de los casos de cáncer cervical en el mundo. En México también son los tipos virales más frecuentes. Otros tipos de VPH menos frecuentes, pero que también se han clasificado como oncogénicos son 31, 33, 35, 39, 45, 51, 52, 56, 58 y 59 ³⁴.

VPH16 and 18 son los tipos virales más frecuentemente identificados en CCE de cérvix (aproximadamente 60% y 16%, respectivamente) ³⁵. Sin embargo, en AD de cérvix la proporción de casos positivos al VPH18 es comparable a la del 16 o incluso está incrementada ^{35,36,37}. Esta clara asociación de la frecuencia de VPH16 y 18 en adenocarcinoma y carcinomas escamosos de cérvix sugiere que los contextos específicos para cada estirpe histológica pueden ser determinantes en el establecimiento de la infección, incidiendo en la replicación, transcripción viral o en el establecimiento de infecciones persistentes. Es decir, es posible que cierto perfil celular, establecido por la proporción o presencia exclusiva de determinados factores transcripcionales, influya de manera favorable o negativa en la actividad transcripcional de promotores de los diferentes tipos de VPH. Además del cáncer cérvicouterino, se ha reportado un papel etiológico del VPH con otros tipos de cáncer como anal, vulvar y cáncer de cabeza y cuello ^{38,39,40}.

En la actualidad existen vacunas disponibles en el mercado dirigidas contra algunos de los tipos virales de VPH de alto riesgo más frecuentes, como el 16, 18, 31, 33, 45, 52, y 58 e incluso contra aquellos que están asociados al desarrollo de verrugas genitales como son los tipos de bajo riesgo 6 y 11. Estas vacunas se basan en partículas parecidas al virión que están constituidas por la proteína L1 de la cápside viral, pero carecen de genoma viral. Estas vacunas van dirigidas a niñas y mujeres jóvenes que aún no han estado expuestas al virus y en las cuales, se ha determinado una disminución significativa de las enfermedades asociadas con el VPH ⁴¹. Es importante resaltar que los títulos de anticuerpos que se producen de una infección natural no proporcionan protección contra la infección con algún tipo viral ⁴².

1.3.2. Estructura

El genoma viral para su estudio se ha dividido en tres regiones: La región larga de control (LCR, por sus siglas en inglés Long Control Region) una región de expresión temprana (E, por su sigla en inglés, Early) y una región de expresión tardía (L, por su sigla en inglés, Late).

La región de expresión temprana, E, codifica seis marcos de lectura, para las proteínas E1, E2, E4, E5, E6, E7 durante la infección temprana además, en esta región también se ha descrito a la proteína de fusión E8-E2C la cual al parecer tiene una función reguladora sobre la transcripción y replicación del VPH. La región tardía, L, contiene dos genes tardíos que son L1, que codifica la proteína más abundante de la cápside viral y el gen L2 que codifica la proteína menos abundante. Estas proteínas están presentes en las capas superiores y más diferenciadas del epitelio. Finalmente, la LCR se encuentra situada entre los marcos de lectura que codifican L1 y E6 como se observa en la **Figura 4**. Esta secuencia contiene elementos de respuesta que son reconocidos por proteínas reguladoras como por ejemplo, activadores o represores transcripcionales o terminadores e iniciadores de replicación ^{43,44,45,46}. En la LCR se encuentra el promotor temprano en el extremo 3'(P97 para VPH16 y P105 para VPH18) que regula principalmente la transcripción de la región viral E ⁴⁷.

Las proteínas virales E1 y E2 son necesarias para la replicación del genoma vírico; E4 tiene un papel en la liberación de viriones; E5 activa el receptor del factor de crecimiento epidérmico (EGFR) ⁴⁸; y E6 y E7 son las oncoproteínas virales más importantes ⁴⁹. La expresión alterada de las oncoproteínas E6 y E7 puede causar modificaciones en las vías celulares críticas implicadas en la proliferación celular y la apoptosis, dando lugar a la inmortalización celular y posteriormente la transformación maligna ^{50,51,52}. En la **tabla 1** se muestran algunas funciones descritas para las proteínas del VPH.

Tabla 1. Proteínas del virus de Papiloma humano y algunas de sus principales funciones.

Proteínas del VPH	Funciones de las proteínas del VPH
E1	Es una proteína muy importante en la replicación del DNA, y puede actuar como ATPasa y DNA helicasa.
E2	La proteína E2 regula la transcripción, así como la replicación del DNA. Puede actuar como un represor de la transcripción de los oncogenes E6 y E7.
E4	Es una proteína que interrumpe la red de citoqueratina y participa en el arresto del crecimiento celular.
E5	La proteína E5 modifica la actividad del receptor del factor de crecimiento epidérmico, induce la vía de señalización de células inflamatorias, participa en transformación maligna junto con E6 y E7, aumenta proliferación, aumenta angiogénesis.
E6	Esta proteína tiene la capacidad de unirse a varias proteínas celulares dentro de las que se encuentra p53 y lleva a cabo transformación y tumorigénesis.
E7	La proteína E7 se ha demostrado que actúa con inhibidores de quinasas dependientes de ciclina, deacetilasas de histona y ciclinas, alterando el control de ciclo celular.
L1	Es la proteína más abundante de la cápside viral.
L2	Es la proteína menos abundante de la cápside viral y se encarga del ensamblaje de viriones.

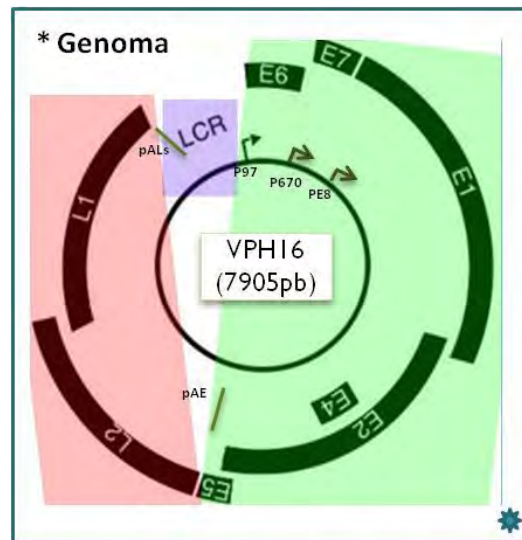


Figura 4. Organización del genoma del VPH16. Se muestra la ubicación de los genes codificados por el genoma viral. Con flechas se muestra la ubicación de los principales promotores del virus y con una línea pequeña perpendicular al genoma viral, se señalan los sitios de poliadenilación, tanto el temprano (pAE) como el tardío (pALs) (Modificado de Taylor & Francis Group. 2004) ⁵³.

1.3.3. Ingreso del VPH a la célula y al núcleo

Se han sugerido varias vías de entrada de los VPH a las células que infecta, dependiendo del tipo viral. En general los VPH infectan las células basales no diferenciadas del epitelio, a través de micro lesiones. Además, la accesibilidad a células cercanas a la unión escamo-columnar que forman una sola capa, aumenta la posibilidad de infección por el VPH⁴⁹. A la fecha se conoce poco sobre los receptores de entrada a las células de los VPH así como su mecanismo exacto, algunos autores han propuesto algunos modelos, de los cuales el más aceptado para el VPH16 sugiere que la proteína L1 de la cápside viral se une a proteoglicanos de heparán sulfato funcionando como receptores iniciales de entrada, estos proteoglicanos se encuentran en la matriz extracelular sobre la superficie celular ^{54,55} con la participación de $\alpha 6$ -integrina y laminin 5 como co-receptores y se lleva a cabo una infección viral eficiente. El ingreso del virus a la célula involucra cambios conformacionales en la cápside y receptores, por lo cual dicha entrada es muy lenta y puede llevarse más de 12 h ⁵⁶. Después de la unión del virus al receptor, el virus entra a la célula vía endocitosis dependiente de actina mediada por clatrina o caveolina ⁵⁷. Cuando las partículas virales han ingresado a la célula, estas viajan a través de la ruta endosomal en donde la acidificación ayuda a que las partículas virales se vayan desensamblando y la cápside pueda ser degradada ⁵⁸. Primeramente, el DNA viral se desensamblen de la proteína de la cápside viral L1 con ayuda de ciclofilinas de la célula, permitiendo el transporte a la red trans-Golgi (TGN por sus siglas en inglés: Trans-Golgi network) y otros compartimentos distales vía la ruta del tráfico retrógrada ⁵⁹. Después de salir de la TGN, el complejo L2-DNA viral, viaja al núcleo donde se asocia con los cuerpos PML nucleares e inicia la transcripción y replicación viral temprana (**Figura 5**) ⁶⁰. La translocación del DNA viral al núcleo requiere de la localización de L2 y DNA viral en la red trans-Golgi y que la célula entre a mitosis ⁶¹. En la **figura 5** se muestran los pasos que sigue el VPH para entrar a la célula.

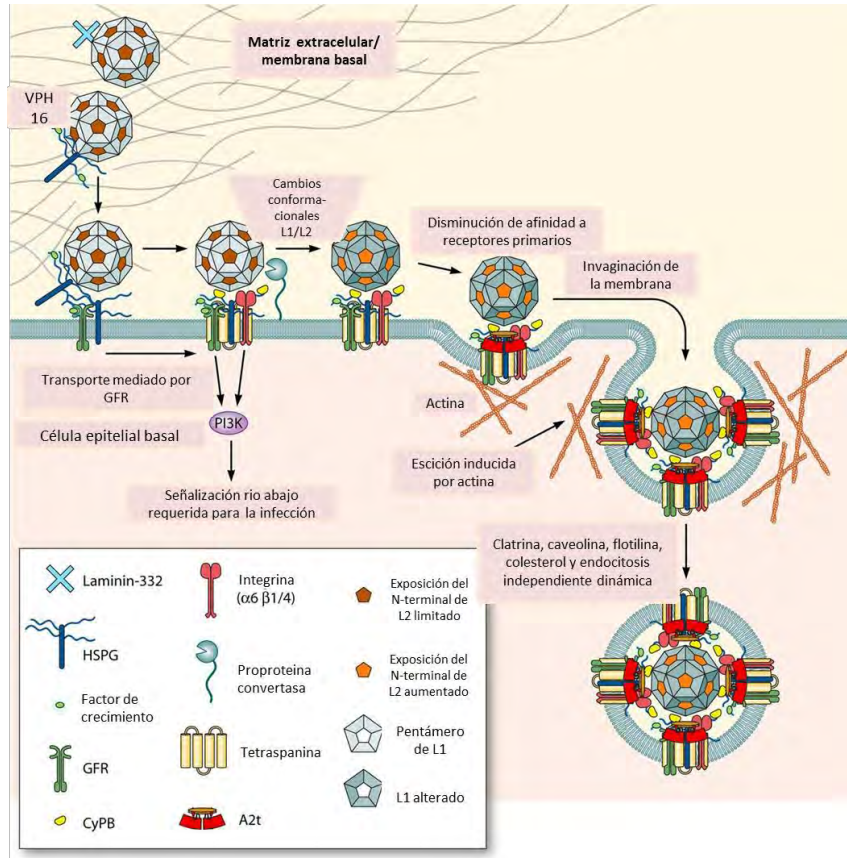


Figura 5. Mecanismo de entrada del VPH a la célula que infecta. Con flechas gruesas se señala el recorrido de las partículas virales, así como sus interacciones con diversas proteínas de la célula hasta llegar al citoplasma. Se anexa un cuadro con el significado de la simbología empleada (Tomada de Raff AB y cols., 2013) ⁶².

1.3.4. El ciclo viral replicativo del VPH

Una vez que el virus logra ingresar a la célula y posteriormente al núcleo, durante la etapa de amplificación inicial y establecimiento de la infección, el VPH replica en un bajo número de copias (10-200 copias por célula) ⁶³. Conforme las células infectadas por el virus se diferencian ocurre un incremento en la replicación del genoma viral, por lo tanto, los niveles de las proteínas E1 y E2, implicadas en este proceso también aumentan ⁶⁴. Posteriormente en la última capa del epitelio, que ha sufrido una diferenciación terminal, se expresan las proteínas L1 y L2 de la cápside del virus culminando el ciclo replicativo con el ensamblaje de las nuevas

partículas virales que se desprenderán junto con las células muertas que descaman del epitelio y estarán listas para reinfectar otro sitio o a un nuevo hospedero ⁵⁰. Dependiendo de la proporción del epitelio infectado por el virus se pueden detectar lesiones intraepiteliales escamosas de bajo grado, alto grado o cáncer invasor que se caracteriza por presentar todas las células del epitelio infectadas por el virus e infiltración de la capa basal del epitelio ⁶⁵. Como se muestra en la **Figura 6**.

Aunque la integración del genoma viral al genoma celular no forma parte del ciclo replicativo del VPH, en algunas ocasiones y por causas aún no definidas, el DNA de los VPHs de alto riesgo se integra al genoma de la célula hospedera, lo cual se ha observado en muestras de tejido de carcinomas de células escamosas de cérvix ⁶⁶. Se ha propuesto que este proceso de integración del genoma viral al genoma celular pueda ser un evento temprano de progresión de lesiones intraepiteliales escamosas de bajo grado a alto grado ⁶⁷.

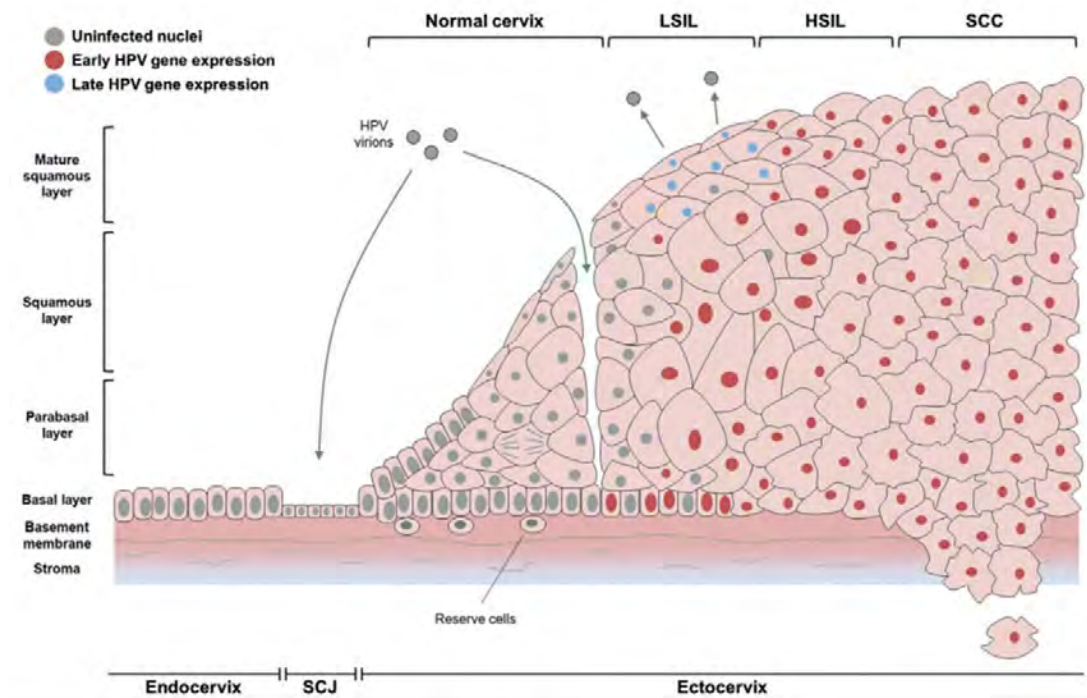


Figura 6. Ciclo replicativo del VPH. El virus ingresa a la capa basal del epitelio por medio de microlesiones o por la accesibilidad que tiene en la unión escamo-columnar. Ahí se replica en un bajo número de copias. Conforme el epitelio se va diferenciando se van expresando las proteínas virales. En las capas superiores se

expresan las proteínas de la cápside viral para finalmente formar nuevas partículas virales que son liberadas y potencialmente pueden reinfectar el epitelio u otros epitelios. Dependiendo de la proporción de células infectadas con el virus se tendrán diferentes tipos de lesiones que van de lesiones intraepiteliales escamosas de bajo y alto grado, hasta cáncer invasor, que se origina como consecuencia de la integración del genoma viral al genoma celular (Tomada de Woodman CB. y cols., 2007) ⁶⁵.

1.3.5. Integración del genoma del VPH

Las infecciones con los virus de Papiloma humano, por lo general, son transitorias y originan lesiones benignas. En algunas ocasiones la infección con virus de alto riesgo u oncogénicos como VPH16, 18 y 31 promueven el desarrollo de lesiones malignas. En el proceso de infección productiva no se lleva a cabo transformación celular puesto que las células que expresan los oncogenes virales se pierden por la renovación constante del epitelio escamoso cervical. En las células menos diferenciadas, como son las células basales del epitelio que expresan los oncogenes, estos se encargan de inhibir diferenciación e inducir inestabilidad cromosómica, lo que les da a esas células la posibilidad de sufrir eventos de transformación ⁶⁸ .⁶⁹.

La expresión de E6 y E7 induce que las células entren nuevamente a fase S y por lo tanto activen otra vez la maquinaria de replicación, promoviendo la síntesis de viriones. E6 y E7 son las principales oncoproteínas virales y tienen diversas funciones como por ejemplo, regular el ciclo celular mediante su interacción con diversas proteínas celulares dentro de las que se encuentran p53 regulada por E6 y pRB regulada por E7 ^{70,71}. Estas interacciones abaten la función normal de estos genes supresores, lo que lleva a la progresión del ciclo celular, inhibición de apoptosis, daño al DNA y finalmente evasión de la respuesta inmune del hospedero ⁷². E6 también se puede unir a proteínas con dominios PDZ mediante el sitio PBD (por sus siglas en inglés, PDZ Binding Domain) en su extremo carboxi-terminal. Por ejemplo, cuando E6 se une a las proteínas con dominio PDZ disk large homolog 1 (Dlg1), Scribble y a guanilato cinasa asociadas a membrana (MAGI-1,2,3), estas proteínas se encaminan a degradación por la vía del proteosoma. Las proteínas E6 de VPH de alto riesgo pueden interactuar con proteínas con dominios PDZ dando como resultado cambios en la morfología

celular o reorganización de microfilamentos que finalmente están asociados con la inmortalización de queratinocitos ^{72,73}. E7, que es la principal proteína transformante de VPH de alto riesgo, interactúa con proteínas como, pRB, c-Jun, varios elementos epigenéticos como deacetilasa de histonas 1 (HDAC1), DNA metiltransferasas (DNMT1) y CBP/p300 ⁷⁴.

La integración del genoma viral al genoma celular es un proceso accidental que frecuentemente interrumpe al gen E2 ^{75,76}. Cuando el genoma viral se encuentra en estado episomal se expresa en un bajo número de copias y la expresión de los genes virales es regulado principalmente por la proteína E2 ⁷⁷, por lo tanto, cuando se da el proceso de integración del genoma viral al genoma celular esta regulación se pierde, provocando una alta expresión de E6 y E7 ^{78,79}.

Dado que en la integración se rompe el gen E2 con alta frecuencia, se pierde el efecto represor de la proteína E2 sobre los oncogenes E6 y E7 ⁸⁰. Sin embargo, existen otros mecanismos por los cuales E6 y E7 pueden desregular. Por ejemplo: se ha determinado que en ciertos casos existen copias de DNA de VPH que se encuentran integradas en tándem y aunque en este caso el gen de E2 no se encuentra roto, los sitios de unión a E2 en la LCR se encuentran metilados, por lo que E2 no puede llevar a cabo su efecto represor ⁸¹. Por otra parte, se ha reportado que la integración en tandem del genoma del VPH también puede provocar que se formen elementos parecidos a un super-enhancer que fuertemente activan la expresión de E6/E7 ⁸².

1.4. La región larga de control

La LCR es la región más variable entre los VPHs y en la mayoría de los papilomavirus abarca aproximadamente el 10% de su genoma. Se ubica entre el final del marco de lectura del gen L1 y el inicio del marco de lectura del gen E6 (**Figura 4**). En el VPH16, la LCR abarca 850 pares de bases (pb) de las 7906 pb de su genoma. Para su estudio se puede dividir en tres regiones: la región 5', la región central y la región 3'. Dentro de la región 3', se localiza el

principal sitio de inicio de la transcripción del promotor de E6, el cual tiene un arreglo muy particular en todos los alfa papilomavirus, la RNA Pol II lo transcribe, contiene una caja TATA 31 – 26 pb río arriba del sitio de inicio de la transcripción y es donde se unirá el complejo de pre-iniciación que contiene a TFIID con la subunidad de unión a la caja TATA (TBP). Río arriba de caja TATA se encuentra un sitio de unión al factor transcripcional activador en células epiteliales, SP1, un hexámero rico en G (GGGCGT), el VPH16 tiene un solo sitio SP1 y es necesario para la actividad del promotor de E6 ⁷⁷. La LCR no tiene cajas CCAAT conservadas, que son elementos regulatorios relevantes de muchos genes celulares. El factor transcripcional SP3 se une al mismo sitio que SP1 y en fibroblastos funciona como un represor del promotor E6. Dentro de esta región 3' también se encuentra el origen de replicación viral, en el cual hay un sitio de unión a E1 rico en AT de alrededor de 20 pb y sitios de unión a E2. La región central se caracteriza por contener una región enhancer, que abarca varios elementos reguladores que hiper-activan la transcripción viral ⁸³. Esta región se puede ubicar entre dos de los sitios de unión a la proteína viral E2, altamente conservados entre los alfa papilomavirus y cuya actividad transcripcional depende de los factores transcripcionales celulares como AP-1, NF-1, TEF-1 y Oct-1 que combinados pueden conferirle especificidad de tejido ⁸⁴. Finalmente, dentro de la región 5' se encuentran sitios que codifican la señal de terminación de la transcripción de los transcritos tardíos (**Figura 7**). Del mismo modo, se ha informado que algunos FT como EPOC-1 / Skn-1a, C / EBP- α , - β , c-Myb, NFATx, Pax5 y WT1 están implicados en la expresión, dependiente de la diferenciación, de genes virales tempranos y tardíos ⁸⁵. Además, las LCR de diferentes tipos de VPH contienen sitios de unión FT potenciales distintivos, que podrían estar asociados a patrones de expresión viral asociados a cada tipo viral, en respuesta a la diferenciación de células huésped ⁸⁶. Por ejemplo, se ha demostrado que GATA3 y TFAP2B afectan específicamente la transcripción del VPH18, mientras que dichos sitios de unión no están presentes en la LCR del VPH16 ⁸⁷. Dentro del genoma del papilomavirus se pueden activar promotores tardíos en la capa granular o epitelial mediante el proceso de diferenciación. Algunos de estos promotores se localizan en la secuencia que codifica E7 como son P670 y P142 para el VPH16

y VPH31, respectivamente en la LCR ⁸⁸, en la región L1 y dentro de la región que codifica E6, pero al parecer tienen actividad muy débil ^{89,90}.

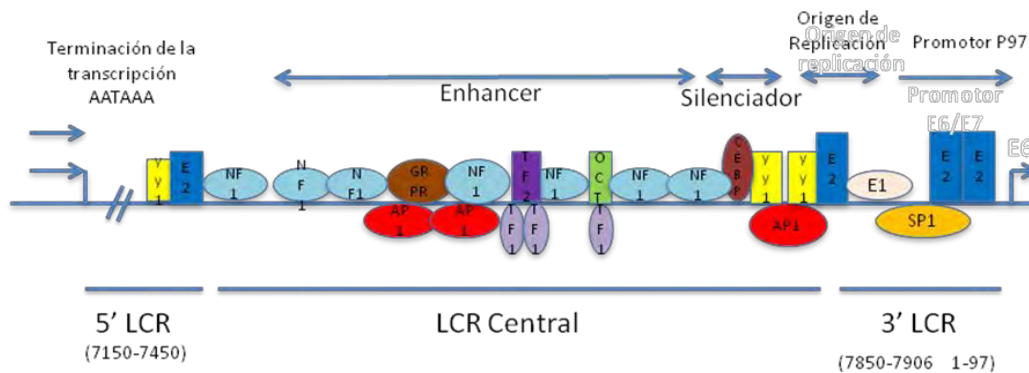


Figura 7. Región larga de control (LCR) del VPH16. Se muestran diferentes sitios de unión a factores de transcripción. Las flechas señalan a la región enhancer, el origen de replicación y el promotor, así como el término de la transcripción. Con líneas se señalan la región 5'LCR, la región Central y la región 3'LCR (Modificada de O'Connor MJ y cols., 1996) ⁹¹.

Los VPHs tienen un patrón de transcripción policistrónico, por lo que producen varios RNAs mensajeros que son procesados por splicing alternativo ⁹². En VPHs de bajo riesgo, los genes *E6* y *E7* son transcritos de dos promotores independientes mientras que en VPHs de alto riesgo estos genes se transcriben como un solo pre-RNAm policistrónico a partir del promotor temprano. Otra característica importante de los VPHs de alto riesgo es que el RNAm policistrónico para *E6* y *E7*, contiene un sitio de splicing donador y uno aceptor que puede llevar a cabo un proceso de splicing alternativo induciendo la expresión de una variedad de transcritos de splicing de *E6* los cuales han sido llamados *E6**; mientras que los de bajo riesgo no sufren splicing en esta región ^{93,94,95}.

1.5. Regulación transcripcional de *E6* y *E7* por *E2*

La proteína *E2* del VPH es una proteína que se une al DNA viral en secuencias específicas y se encarga de regular el ciclo de replicación viral mediante su asociación con la proteína *E1*,

también conocida como factor de replicación viral ⁹⁶. Tiene 360 aminoácidos y se divide en tres dominios: un dominio de transactivación amino terminal, un dominio bisagra y un dominio de unión al DNA (DBD, por sus siglas en inglés DNA Binding Domine) en la región carboxi-terminal. La proteína E2 se une como dímero a sus sitios dentro de la LCR ⁹⁷. E2 modula al ciclo viral y tiene un papel importante en la regulación transcripcional y en la replicación del DNA viral ⁹⁸.

Dependiendo de la ubicación de los sitios de unión a E2 dentro del promotor de cada tipo viral, E2 puede actuar como un activador o un represor de la expresión de los genes virales. Para el caso de los promotores de los VPHs tipo 16 y 18 la expresión de los genes *E6* y *E7* es regulada negativamente por la unión de E2 a sus sitios de reconocimiento dentro del promotor, lo que se acompaña con el desplazamiento de los factores transcripcionales celulares que se unen al promotor viral ^{99,100}. La pérdida de la expresión de E2 se ha asociado fuertemente con el desarrollo de carcinomas inducidos por el VPH. Como E2 tiene la capacidad de suprimir la expresión de *E6/E7*, su interrupción, como en el caso de la integración del genoma viral al genoma celular, provoca una expresión desregulada de *E6* y *E7* y por lo tanto contribuye a la progresión carcinogénica (**Figura 8**) ⁸⁰.

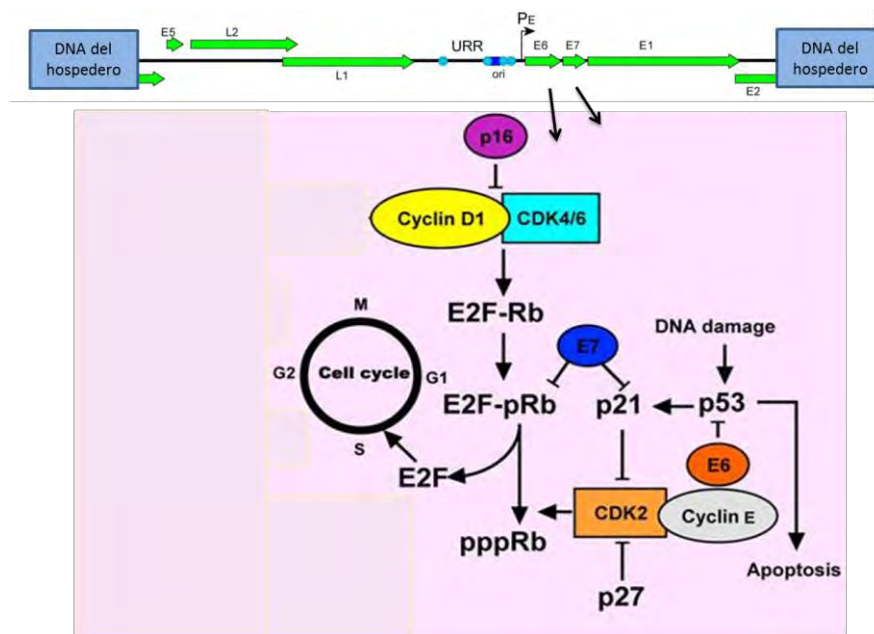


Figura 8. Regulación de la expresión de los genes *E6* y *E7* por la ruptura del gen *E2* durante el proceso de integración del genoma viral al genoma celular. Se muestra la ruptura del genoma viral que se da frecuentemente en la región que codifica *E2*, por lo tanto, se lineariza y se pierde la expresión de *E2*, así como la represión transcripcional que ejerce sobre la expresión de los genes *E6* y *E7*, aumentando entonces los niveles de estas oncoproteínas. Con flechas se señalan los blancos sobre los cuales van a actuar estas oncoproteínas para desregular el ciclo celular, ejemplificado con un círculo (Modificada de Koontongkaew S. 2013) ¹⁰¹.

1.6. Regulación transcripcional del VPH por factores transcripcionales celulares.

La transcripción temprana del genoma del VPH está regulada por secuencias de nucleótidos específicas ubicadas en la LCR conocidas como elementos de respuesta que actúan en cis que además de unir factores transcripcionales virales también unen una variedad de factores de transcripción celulares **Tabla 2**. Es importante hacer notar que la composición de una región de control transcripcional, así como el espacio entre diferentes elementos reguladores, tienen influencia importante en la actividad transcripcional ^{102,103}. Debido a su alta prevalencia, los tipos de VPH16 y VPH18 han sido los principales modelos en el estudio de la regulación transcripcional del VPH. Factores de transcripción celulares que modulan la actividad de la LCR de los VPH de alto riesgo, como SP-1, AP-1, NF-1, Oct-1, YY-1 y el receptor de glucocorticoides GRE (Figura 2) ¹⁰⁴, también son importantes en el crecimiento celular normal, desarrollo y diferenciación.

Tabla 2. Secuencias consenso de unión a factores de transcripción en la LCR del VPH16 y VPH18.

Virus type	Binding site (Transcription factor)	Sequences 5'→3'	Nucleotide Position
HPV16	AP-1	TTAATCA	7631-7637
		TGAGTCA	7648-7654
		TTGTCA	7810-7816
	NF-1	TTGGC	7473-7477
		TTGGC	7558-7554
		TTGGC	7591-7587
		TTGGC	7679-7675
		TTGGC	7711-7715
		TTGGC	7742-7746
		TTGGC	7770-7766
	GRE	TGTACATTGTGTCAT	7641-7655
	Oct-1	AATTGCAT	7732-7739
	Sp-1	GGGCGT	28-33
KRF	ND	ND	
HPV18	AP-1	TTAGTCA	7608-7614
		TTAGTCA	7798-7792
	NF-1	CTGGCN5TGCAAA	7513-7528
		TTGGCN8TTGGC	7569-7586
		TTGGC	7731-7735
	Oct-1	GCCTTGCAT	7644-765
		AATTGCAT	7721-772
	KRF-1	TGCATAACTATATCC	7647-7669
	GRE	ACTCCCTA	7839-7853
	Sp-1	AGCACATACTATACT	35-40
		GGGAGT	

Se indica la secuencia y la posición en la que se encuentra el sitio de unión para cada factor transcripcional en la LCR del VPH16 y del VPH18.

AP-1

El factor transcripcional AP-1 es un factor importante que regula la transcripción del genoma viral. Está constituido por heterodímeros de proteínas de la familia Jun y Fos. Dentro de las funciones de AP-1 se encuentran el control de la expresión de genes implicados en la proliferación celular, diferenciación y transformación, a través de diversos estímulos como

citocinas, factores de crecimiento, estrés e infecciones virales o bacterianas ¹⁰⁵. JunB y Fra2 son componentes de AP-1, que se encuentran expresados en alta cantidad en queratinocitos y modulan su diferenciación a través de la regulación de varios genes como involucrina, loricrina, citoqueratinas o integrinas ¹⁰⁶. La LCR de VPH18 contiene dos secuencias consenso de unión a AP-1, uno ubicado dentro de la región enhancer y otra dentro del promotor proximal y se ha demostrado que juegan un papel clave en la activación del promotor del VPH18 ¹⁰⁴.

Sp-1

Sp-1 es un factor de transcripción cuya estructura contiene dedos de zinc que se unen a motivos ricos en GC de muchos promotores. Sp-1 juega un papel muy importante en diversos procesos celulares como diferenciación y crecimiento celular, apoptosis, respuesta inmune, respuesta al daño del DNA y remodelación de la cromatina. Se han descrito tres isoformas de Sp-1. Junto con Ap-1 juegan un papel muy importante en la regulación transcripcional de VPH de alto y bajo riesgo. Sitios de unión a Sp-1 en la LCR del VPH16 y VPH18 generalmente se encuentran cerca del sitio de inicio de la transcripción y en ensayos de actividad transcripcional con estos sitios mutados se ha observado una fuerte reducción de la actividad transcripcional, reflejando su importancia en la activación transcripcional de *E6 / E7* ¹⁰⁴.

Oct-1

El factor de transcripción Oct-1 forma parte de la familia de factores de transcripción POU. Los miembros de esta familia se caracterizan por contener el Dominio POU, que es una región de 160 aminoácidos necesaria para la unión del DNA a la secuencia octamérica ATGCAAAT ⁴⁶. Se ha reportado que para VPH11 y VPH18, Oct-1 funciona como un activador transcripcional, ya que cuando se mutan sus sitios de unión consenso, se reduce significativamente la actividad transcripcional ¹⁰⁷.

YY-1

YY-1 es un factor transcripcional con dedos de Zinc distribuido ubicuamente que regula procesos celulares importantes como crecimiento, diferenciación y desarrollo y juega un papel importante en la tumorigénesis ¹⁰⁸ e infección por VPH ¹⁰⁹ YY-1 puede reprimir ¹¹⁰ y activar ¹¹¹ diversos promotores tanto celulares como virales, pero además tiene la capacidad de dirigir deacetilasas de histonas y acetiltransferasas a un promotor para activarlo o reprimirlo. En VPH18 ¹¹² y VPH11 ¹¹³ YY-1 actúa como un represor. En VPH8, YY-1 tiene un papel dual activador y represor ¹¹⁴; mientras que en VPH16, YY-1 reprime la transcripción *E6* y *E7* ¹¹⁵.

Elementos de respuesta a glucocorticoides (GRE)

Las hormonas esteroideas como los glucocorticoides pueden actuar como cofactores para el VPH en el desarrollo de cáncer de cuello uterino. Los glucocorticoides, como la dexametasona, que es sintética, mejoran la tasa de transformación de células transfectadas con VPH de alto riesgo ^{116,117}. VPH11, VPH16 y VPH18 tienen respuesta activa a elementos de respuesta a glucocorticoides (GRE). Dentro de la LCR del VPH18 se encontró una secuencia AGCACATACTATACT, que confiere respuesta a dexametasona y progesterona. La LCR de VPH31 contiene al menos un GRE funcional y tres potenciales ¹¹⁸.

2. ANTECEDENTES

2.1. Análisis *in silico* con el programa JASPAR CORE

En un estudio bioinformático preliminar se analizó la LCR completa del VPH16 en comparación con la del 18 empleando la base de datos JASPAR CORE ¹¹⁹ e identificamos varios sitios de unión a factores de transcripción (FT) o elementos de respuesta. De estos, el factor transcripcional SOX2 nos pareció muy interesante para iniciar nuestro estudio ya que encontramos tres elementos de respuesta para SOX2 específicamente en la LCR del VPH16, tipo viral reportado más frecuentemente asociado con carcinomas de células escamosas,

mientras que estos sitios no estaban presentes en la LCR del VPH18 (**Tabla. 3**). Se ha reportado que SOX2 puede inducir diferenciación escamosa, que es un factor importante de células troncales, y que tiene un papel importante en cáncer. De este modo, en el presente estudio nos interesó determinar cuál es el papel de SOX2 en la regulación transcripcional de la LCR del VPH16 y si esto puede ser afectado según la estirpe celular.

Tabla 3. Predicción *in silico* de sitios de unión a factores de transcripción para la LCR del VPH16 o 18 empleando la base de datos JASPAR CORE ¹¹⁹.

	LCR-HPV16			LCR-HPV18	
Arnt	TGTTGTTTGTG	(nt 7180)	Ddit3::Cebpa	TGATTG	(nt 7205)
	TGTTGTTTGTG	(nt 7180)	Gata1	GTACTGTA	(nt 7504)
Foxa2	CACGTG	(nt 7268)	MZF1-1-4	CAATTGCT	(nt 7639)
	CGTGTGTA	(nt 7270)	NFE2L2	ATAGTTA	(nt 7651)
	GTGTATGTGTTT	(nt 7273)	NKX3-2	TAATTGCA	(nt 7720)
FoxD1	TGACACAAT	(nt 7303)		AATTA	(nt 7720)
	TGACACA	(nt 7305)	TAL1::TCF3	AGTATATA	(nt 22)
FoxF2	ATAAGTT	(nt 7325)			
HNF1B	TTTCCT	(nt 7569)			
	TCAGGAA	(nt 7570)			
MAFb	TTCCATTGTT	(nt 7597)			
MAX	TCCATTGTT	(nt 7598)			
Mycn	GTAAAAAA	(nt 7605)			
SOX2	CATTGTT	(nt 7600)			
	CATTGTT	(nt 7822)			
	CACATGG	(nt 7843)			

2.2. El factor transcripcional SOX2

El gen SOX2 (sex determining region Y-box2) pertenece a la familia SOX (abreviación para Sry-related HMG box) que consiste de un grupo de factores de transcripción que participan

en el desarrollo y en la biología de las células troncales. Comprende a un grupo de proteínas que contienen un dominio distintivo de un grupo de alta movilidad HMG (por sus siglas en inglés High Mobility Group) cuya similitud en sus aminoácidos es del 50% o más al dominio HMG de Sry. En 1990 descubrieron al gen Sry el cual se nombró así por su ubicación en la región que determina el sexo en el cromosoma Y y codifica al factor determinante del testículo en mamíferos. El gen *SOX2* fue descubierto y caracterizado en humanos en 1994, se localiza en el cromosoma 3q26.3-q27 y pertenece al grupo SOXB1 de la familia SOX. Codifica una proteína de 317 aminoácidos ¹²⁰. Alteraciones del gen *SOX2* se han asociado con enfermedades del desarrollo como el síndrome de anoftalmia esofágica genital (AEG) y diversas investigaciones que se han acumulado asocian a la proteína *SOX2* con cáncer. La proteína se puede dividir para su estudio en tres dominios: el amino terminal, el dominio HMG y el dominio de transactivación (**Figura 9**).

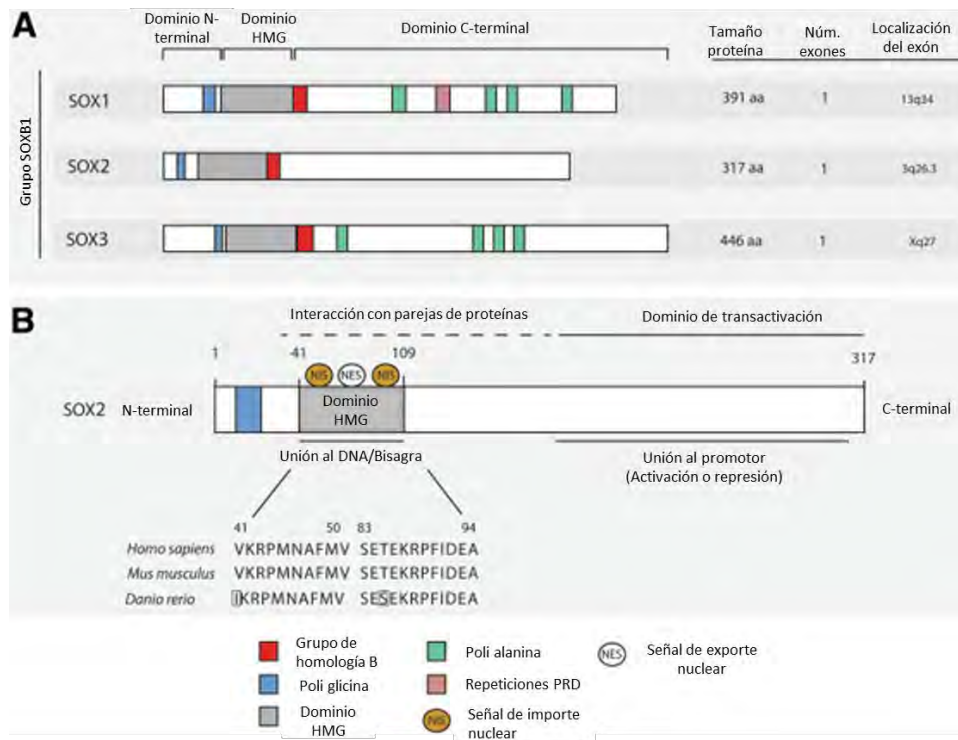


Figura 9. Estructura y organización de los miembros del grupo SOXB1. *SOX2* pertenece al grupo SOXB1. A) Dominios que tienen los miembros de la familia SOXB1. Se señala el tamaño de la proteína, el número de exones y la localización cromosómica. B) Se detalla la secuencia de aminoácidos del dominio HMG de *SOX2* de *Homo sapiens* y se compara con otras especies. (Tomada de Weina K and Utikal J., 2014) ¹²¹.

La actividad de la proteína SOX2 se puede modificar de diferentes formas, por ejemplo los microRNAs pueden regular a SOX2. Tal es el caso del micro-RNA-145 que podría inhibir la auto-renovación de células troncales embrionarias y la expresión de genes de pluripotencia mediante la represión directa de algunos de sus blancos como son OCT4, SOX2 y KLF4 ¹²². SOX2 puede regularse post-traduccionalmente sufriendo diversos tipos de modificaciones como: fosforilación, acetilación, ubiquitinación, metilación y SUMOilación y por lo tanto su actividad se verá afectada en relación al tipo de modificación que presente (Figura 10). Por ejemplo la SUMOilación en la lisina 247, inhibe la actividad transcripcional de SOX2 ya que altera su capacidad de unión al DNA ¹²³. La acetilación mediada por p300/CBP en el residuo de la lisina 47, en el dominio de unión al DNA, promueve exporte nuclear de SOX2 y por lo tanto los niveles aumentados de SOX2 acetilado provocan ubiquitinación y degradación proteosomal de SOX2 ¹²⁴. La fosforilación de SOX2 en la Treonina 118, mediante Akt, mejora su estabilidad y por lo tanto aumenta su actividad transcripcional ¹²⁵. La metilación sobre Arginina 113 de SOX2, mediada por CARM1, facilita la transactivación mediada por SOX2 ¹²⁶. La SUMOilación de SOX2 inhibe la expresión de NANOG ¹²⁷.

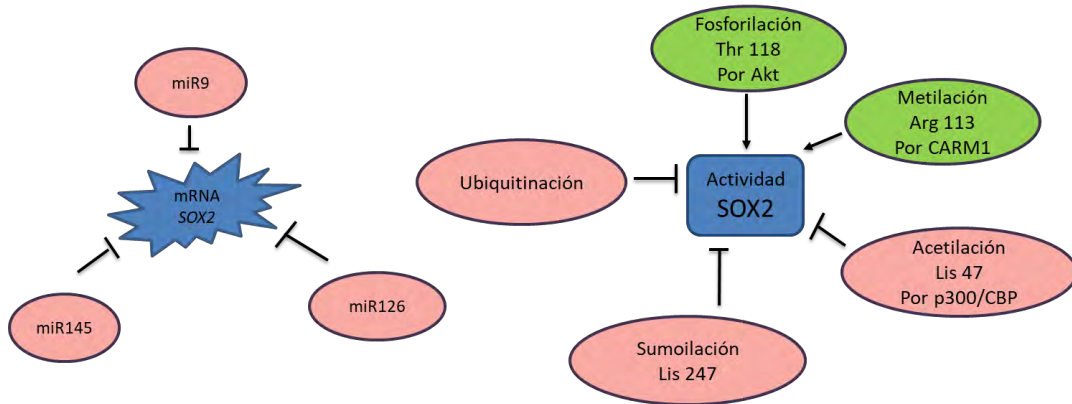


Figura 10. Regulación de SOX2. SOX2 puede regularse traduccional o post-traduccionalmente. SOX2 puede sufrir fosforilación, ubiquitinación, acetilación, sumoilación y metilación, que repercuten directamente en su funcionamiento activando o reprimiendo su función. microRNAs regulan negativamente a SOX2 ¹²⁸.

El factor transcripcional SOX2 puede formar complejos con otros factores transcripcionales como NANOG y OCT4 y tiene la capacidad de estimular la reprogramación de células adultas a células troncales pluripotentes inducidas y mantener propiedades de células troncales cancerosas ¹²⁹. Se ha reportado que SOX2 está involucrado en varios procesos importantes de la célula como regulación del ciclo celular, reparación del DNA y auto-renovación de células troncales ¹³⁰ y también se ha visto que en células cancerosas está asociado con quimioresistencia, mantenimiento de propiedades de células troncales y tumorigénesis ¹³¹. La sobre-expresión de SOX2 está relacionada con migración e invasión en líneas celulares de cáncer cervical ¹³² y de manera controversial hay algunos estudios que han demostrado que la expresión de SOX2 se puede asociar con sobrevida prolongada en pacientes con cáncer cervical ¹³³ o que la amplificación de SOX2 y la expresión de la proteína están asociadas con una sobrevida favorable en cáncer de células escamosas de pulmón ¹³⁴. Incluso hay reportes que hablan de un comportamiento diferencial según el tipo de células, por ejemplo, la sobreexpresión de SOX2 en carcinomas de células escamosas de pulmón se asoció con un pronóstico favorable, mientras que en adenocarcinoma del tracto gastrointestinal se asoció con una menor sobrevida ¹³⁵. Aunque se ha reportado que SOX2 y OCT4 funcionan cooperativamente mediante el complejo SOX2/OCT4, también se ha reportado que los tres factores transcripcionales SOX2, OCT4 y Nanog, pueden formar complejos individuales por ejemplo con nucleofosmina ¹³⁶. Ji J, Zheng PS (2010) reportó la expresión del factor transcripcional SOX2 específico de células troncales embrionarias, en tejido de cáncer cervical primario y en esferas tumorales formadas por células de carcinoma cervical primario, demostrándose que puede funcionar como un oncogén en carcinogénesis cervical ya que tiene la capacidad de promover crecimiento celular y tumorigenicidad ¹³⁷. Los resultados de Liu XF y cols., (2014) sugirieron que SOX2 puede ser un marcador potencial para células troncales cancerosas de cérvix ya que líneas celulares de cáncer cervical (SiHa y C33A) transfectadas con un plásmido que expresa SOX2 mostraron características de este tipo de células como auto-renovación, con la propiedad de iniciar el desarrollo tumoral. Las células positivas a SOX2 también expresaron niveles incrementados de otros factores

transcripcionales como OCT4 y Bmil, así como el marcador de células troncales cancerosas ALDH1 y más interesante mantuvieron características mesenquimales intrínsecas como la expresión de antígenos comúnmente asociados con células troncales mesenquimales como vimentina, snail y β -catenina ¹³⁸.

Muy interesantemente, se ha reportado que la sobreexpresión del factor transcripcional SOX2 en células de cáncer pancreático indujo la expresión de ciclina D3, específica de G1/S y por lo tanto M Herreros-Villanueva y cols (2013) demostraron que p21^{Cip1}, p27^{Kip1} y ciclina D3 son blancos auténticos de SOX2 ¹³⁹.

3. JUSTIFICACIÓN

El contexto celular incluye una serie de factores transcripcionales que pueden o no estar presentes en algunas estirpes celulares o estar presentes, en diferentes concentraciones, lo que puede influir en la actividad transcripcional de promotores virales específicos. Esto puede explicar, en parte, que el virus del papiloma humano tipo 16 sea el de mayor prevalencia en los carcinomas de células escamosas de cérvix aun cuando otros tipos virales como el VPH18 también tienen un alto potencial oncogénico. Sin embargo, este aspecto aún no está esclarecido. De este modo, resulta de gran relevancia poder conocer si factores transcripcionales específicos, como es el caso de SOX2, pueden tener un impacto determinante en la actividad transcripcional de los tipos del VPH de alto riesgo, lo que eventualmente puede repercutir en el establecimiento de infecciones persistentes y en el desarrollo del cáncer. El factor transcripcional celular SOX2 tiene un papel importante en cancer y se ha caracterizado como un factor transcripcional que puede inducir diferenciación escamosa. Dado que en un análisis *in silico* utilizando la base de datos JASPAR CORE se encontró que la LCR del VPH16 contiene sitios de unión putativos a SOX2, que no contiene la LCR del VPH18, resulta de gran importancia determinar el efecto de SOX2 sobre la regulación de su actividad transcripcional y si este efecto es dependiente del contexto celular.

4. HIPÓTESIS

SOX2 regula la transcripción de la LCR del Virus de Papiloma Humano tipo 16 mediante la unión a sus elementos de respuesta en la LCR, y es dependiente del contexto celular.

5. OBJETIVO GENERAL

Determinar el papel de SOX2 en la regulación transcripcional de la LCR del VPH16 según el contexto celular.

6. OBJETIVOS ESPECÍFICOS

- 1.- Determinar los niveles de expresión de SOX2 y la actividad transcripcional de la LCR del VPH16 en líneas celulares con diferente origen histológico (glandular y escamoso).
- 2.- Determinar la unión de SOX2 a sus elementos de respuesta putativos en la LCR del VPH16.
- 3.- Evaluar el efecto de SOX2 sobre la regulación transcripcional de la LCR del VPH16.
- 4.- Evaluar el papel de los elementos de respuesta a SOX2 en la LCR del VPH16 en su regulación transcripcional.
- 5.- Evaluar el efecto de SOX2 en la expresión de los oncogenes *E6* y *E7* del VPH16.

7. METODOLOGÍA

7.1. Cultivo de células

Se cultivaron las siguientes líneas celulares: CaSki, derivada de carcinoma de células escamosas, positiva al VPH16, American Type Culture Collection, (ATCC # CRL-1550) en medio RPMI-1640 (Life Technologies, Gibco, Grand Island, Nueva York, EE. UU.), con suero bovino fetal al 10% (SBF) (Life Technologies, Gibco, Grand Island, Nueva York, EE. UU.); la línea celular HeLa, adenocarcinoma, positiva al VPH18 (ATCC # CCL-2) y la línea celular HaCaT, queratinocito humano no tumorigénica (amablemente donada por el Dr. P. Gariglio, Centro de Investigación y Estudios Avanzados del Instituto Politécnico Nacional, CINVESTAV, Ciudad de México), se cultivaron en medio Dulbecco's Modified Eagle Medium: Nutrient Mixture F-12 (DMEM-F12) Medium (Life Technologies, Gibco, Grand Island, Nueva York, EE. UU.), con 10% de SFB, a 37 °C en una atmósfera húmeda al 5% de CO₂. Las líneas celulares CaSki, HeLa y HaCaT fueron autenticadas mediante perfiles cortos de repetición en tándem en agosto de 2016, por el Instituto Nacional de Medicina Genómica (INMEGEN), México. No mostraron evidencia de contaminación cruzada con líneas celulares conocidas de ATCC.

7.2. Construcción de plásmidos

La LCR del VPH16 se amplificó por PCR a partir de biopsias de cáncer cervical positivas para el VPH16. Los oligonucleótidos usados para la amplificación de la LCR del VPH16 fueron los siguientes: Sentido SmaI 5'-GGGCCCCGGGCAGACCTAGATCAGTTTCCT-3' y antisentido HindIII 5'-GGG AAGCTTGCAGTTCTCTTTTGGTGCAT-3', estos amplifican la región que comprende los nucleótidos del 7007-7905 y del 1-102 del genoma del VPH16. Los productos de PCR se insertaron en el vector pGL2-Basic, que contiene al gen de luciferasa como reportero. Tanto los productos amplificados por PCR como el vector pGL2-Basic fueron digeridos empleando las enzimas de restricción SmaI / HindIII (Promega corporation, Madison, Wisconsin, EE. UU.), generando el vector pGL2-LCR16. El plásmido pGL2-LCR18

se construyó previamente, a partir de productos de la LCR amplificados por PCR, obtenidos a partir de biopsias de CaCU positivo para el VPH18¹⁴⁰, este plásmido contiene los nucleótidos del 7201-7857 y del 1-124 del genoma del VPH18. Las construcciones de plásmidos se verificaron por secuenciación automática por el método de Sanger. Los plásmidos pcDNA-SOX2 (Addgene Cambridge, MA) y pcDNA3.1 Zeo (+) (Life Technologies, Invitrogen, Carlsbad, CA) se usaron para la sobreexpresión de SOX2 y como un control negativo, respectivamente.

Los vectores lentivirales Pax2 y PMD2.G (Addgene) se usaron para producir lentivirus recombinantes que contienen al gen SOX2 y / o al gen que codifica la proteína verde fluorescente (GFP, por sus siglas en inglés Green Fluorescent Protein). La línea celular de empaquetamiento 293T se cotransfectó con los plásmidos de expresión lentiviral SOX2 y / o GFP, Pax2 y PMD2.G, usando lipofectamina plus (Invitrogen). Dos días después de la transfección, se colectaron los sobrenadantes celulares que contenían los lentivirus recombinantes y se almacenaron a -80 °C hasta su uso.

7.3. Infección con lentivirus recombinantes

Las células CaSki se sembraron en placas de 24 pozos a una densidad celular de 3×10^4 células / pozo. Después de 24 h, en base a las curvas estándar, las células se transdujeron con 50 microlitros de sobrenadante lentiviral añadido a cada pozo y la incubación se continuó durante 72 h. El vector lentiviral GFP se usó como control de transducción.

7.4. Ensayos de transfección y actividad transcripcional

Para los ensayos de transfección transitoria, se sembraron 30,000 células CaSki, HeLa o HaCaT en placas de 24 pozos con medio completo; después de 24 horas, las células se transfectaron usando lipofectamina 2000 (Invitrogen) con 200 ng de vector reportero de luciferasa y 1 ng de plásmido de Renilla (pRL-CMV) para la normalización. Como control positivo, se usó pGL3-Control, que contiene el promotor del Virus Simian 40 (SV40). Los extractos celulares se prepararon 48 h después de la transfección y se ensayaron para

actividades de luciferasa y renilla usando el kit Dual Luciferase (empresa Promega, Madison, Wisconsin, EE. UU.). Las unidades relativas de luciferasa se midieron en un luminómetro Monolight 3010 (PharMingen, BD Biosciences, San José CA, EE. UU.). Los experimentos se realizaron por triplicado.

7.5. Análisis *in silico*

Se analizaron los sitios de unión putativos a factores de transcripción (FTs) dentro de la LCR del VPH16 y 18 utilizando la base de datos JASPAR CORE ¹¹⁹ y alineamiento con las secuencias consenso. La secuencia completa de la LCR del VPH16 y 18 fue capturada y comparada con los datos existentes.

7.6. Ensayo de cambio en el movimiento electroforético (EMSA)

El ensayo de cambio en el movimiento electroforético (EMSA, por sus siglas en inglés electrophoretic mobility shift assay), se realizó en un volumen final de 20 μ L que contenía extracto de proteína nuclear de células HeLa (15 μ g) en el buffer de incubación (Tris 30 mM, pH 7.5, KCl 300 mM, MgCl₂ 15 mM y 30 % de glicerol), adicionados con 1M de ditioneitol (DTT), 10 μ g / μ L de BSA, 1 μ g / μ l de poli (dI-dC) (Amersham Biosciences, Buckinghamshire, RU) y 1 μ L de Oligonucleótidos de ADN marcados con γ -ATP (³²P). Los oligonucleótidos utilizados para identificar los sitios de unión a SOX2 fueron: para la secuencia de unión consenso (CS, por sus siglas en inglés Consensus sequences) a SOX2 (CS): 5' 'TAGTCATAACAATGTTTCATTT 3'; sitio 1 putativo de unión a SOX2 (S1): 5'-AACCATTCCATTGTTTTTTA-3'; sitio 2 putativo de unión a SOX2 (S2): 5'-TAGTCATACATTGTTTCATTT-3'; sitio 3 putativo de unión a SOX2 (S3): 5'-AACTGCACATGGGTGTGTGCAAACCGT-3'; y oligonucleótidos no específicos (NS, por sus siglas en inglés Non-specific): 5' 'AACTGCACATGGGTGTGTGCAAACCGT 3'. El competidor de doble cadena no marcado también se incluyó antes de la adición de oligonucleótidos marcados, cuando se especifica. Para el análisis de super-retardo (SS, por sus siglas en inglés

super shift), se usaron anticuerpos anti-SOX2 de conejo (Santa Cruz Biotechnology, Dallas, Texas, EE. UU., Sc-20088) y factor de unión anti-CCCTC de conejo (CTCF) (sc-5916X). Las reacciones de unión se incubaron durante 20 minutos a temperatura ambiente. Los complejos proteína-ADN se separaron del oligonucleótido libre en un gel de 6,0% de poliacrilamida / Tris borato-ácido etilen-diamino-tetra-acético no desnaturizante y se resolvieron a 130 V durante 2 h. Los geles se visualizaron en un Typhoon 9400 (GE Healthcare, Uppsala, Suecia) y se analizaron usando el software Image Quant (Biociencias AB GE Healthcare Life Sciences, Uppsala, Suecia).

7.7. Ensayo de inmunoprecipitación de la cromatina (ChIP)

El ensayo de inmunoprecipitación de la cromatina (ChIP, por sus siglas en inglés Chromatin Immunoprecipitation) se llevó a cabo con las células CaSki ya que es la línea celular que contiene copias endógenas del VPH16 y se utilizó el kit One Day ChIP (Diagenode, Denville, NJ, EE. UU.), siguiendo las instrucciones del fabricante. 1×10^6 células se entrecruzaron con formaldehído al 1% y la reacción se inactivó con 0.125 M de glicina. Las células se lisaron con un buffer que contenía dodecilsulfato de sodio al 1% (SDS) y se sometieron a sonicación con el disruptor sónico Thermo Fisher Scientific (Thermo Fisher Scientific, Waltham, MA, EE. UU., Modelo 505). La cromatina fragmentada (principalmente de 300 a 500 pb) se inmunoprecipitó utilizando anticuerpos: anti-SOX2 (conejo) (Abcam, Cambridge, MA, EE. UU., Ab59776); anti-p65 (conejo) (Santa Cruz Biotechnology, sc-109), como control de unión a la LCR¹⁴¹; y anti-CTCF (conejo) (Millipore, Billerica, MA, EE.UU., 07-729), como control de no unión¹⁴². Como control de ruido de fondo, se incubó una fracción de cromatina con anticuerpo de IgG de conejo (Santa Cruz Biotechnology, sc-2027).

La PCR cuantitativa (qPCR) se realizó por triplicado en una placa de reacción de 96 pozos (Applied Biosystems, Carlsbad, CA). Las reacciones se realizaron utilizando oligonucleótidos que amplifican una región de la LCR: Sentido 5'-GTAAACTACACATGGGTGTGTG-3' y Antisentido 5'-GTCTGCTTTTATACTACACCGG-3'; con Thermo Maxima SYBR Green / ROX1 PCR Master Mix (Thermo Scientific, Applied Biosystems, Woolston, Warrington, RU,

K0222), en el sistema StepOnePlus Real-Time PCR (Applied Biosystems, 4376600). Como control positivo de la unión de SOX2 a un promotor, se amplificó por qPCR un fragmento del promotor del factor transcripcional Lefty con los siguientes oligonucleótidos: sentido 5'-GTCAGGGTAGGGATCACTC-3'; y antisentido 5'-TACAATGGTCTGAAGTGTTGC-3'.

7.8. Ensayos de Mutagénesis Dirigida Multi-Sitio

El plásmido pGL2-LCR16 se mutó en los tres sitios putativos de unión a SOX2 mediante mutagénesis dirigida multi-sitio (QuikChange Multi-Site-Directed Mutagenesis kit, Agilent Technologies, USA). Los oligonucleótidos utilizados en este proceso se muestran en la **tabla 4**. Se obtuvieron cuatro mutantes diferentes: mutante LCR-Sitio 1 (M1), mutante LCR-Sitio 2 (M2), mutante LCR-Sitio 3 (M3) y la LCR-triple mutante (M1,2,3). Los plásmidos así obtenidos cuentan con las mutaciones respectivas en los sitios putativos de unión a SOX2. Los plásmidos mutantes de la LCR se confirmaron mediante secuenciación automática Sanger. El sitio 1 se mutó en dos bases, mientras que los sitios 2 y 3 se mutaron en tres bases cada uno.

Tabla 4. Oligonucleótidos utilizados para introducir mutaciones en los sitios putativos 1, 2 y 3 de unión a SOX2 de la LCR del VPH16.

Primer	Sitio	Secuencia
Sentido	1	GCACTGCTTGCCAACCATTC <u>CC</u> TTTTTTTTTACTACTGCACTATGTG (A) (G)
Antisentido	1	CACATAGTGCAGTGTAAAAAAAAGGGAATGGTTGGCAAGCAGTG
Sentido	2	CCTAGTTCAT <u>AC</u> ATGAACTGTGTAAGGTTAGTCAT <u>CC</u> TTTTTCATTTGTAAACTG CAC (A)(A) (G)
Antisentido	2	GTGCAGTTTTACAAATGAAAAAGGGATGACTAACCTTTACACAGTTCATGTATGAAC TAGG
Sentido	3	CATTGTTCAATTTGTAAACTG <u>CCCGTA</u> GGTGTGTGCAAACCGTTTTGGG (A) (A)(G)
Antisentido	3	CCCAAACGGTTTGCACACACCTACGGGCAGTTTTACAAATGAACAATG

Las nucleótidos subrayados, indican la región de los sitio de unión putativos respectivos, a SOX2. Los nucleótidos que están entre paréntesis fueron los que se sustituyeron por los nucleótidos que están en negritas, cambiando la secuencia original.

7.9. Western Blots

Los lisados totales de las células se obtuvieron empleando el buffer RIPA (cloruro de sodio 150 mM, Triton-X100 al 0.1%, dodecilsulfato sódico al 0.5% y Tris 50 mM pH8 en ddH₂O) y se prepararon en presencia de inhibidores de proteasas (mini EDTA-free de Roche, Mannheim, Alemania). Las concentraciones de proteína se determinaron en base a la cuantificación realizada mediante la solución de ácido bicinonínico y la solución de sulfato de cobre (II) (Sigma-Aldrich, St. Louis, EE. UU.). Se corrieron 30 µg de proteína en la electroforesis en geles de poliacrilamida con dodecil sulfato sódico al 12% (SDS-PAGE) y se transfirieron a una membrana de difluoruro de polivinilideno Immobilon-P (PVDF) (Amersham, Reino Unido) para transferencia. Las membranas se bloquearon con leche descremada al 10% en TBS-T (Tris 10 mM, pH 7,5, NaCl 100 mM y Tween-20 al 0,1%) y se incubaron con los siguientes anticuerpos primarios (en leche descremada al 5%) durante toda la noche a 4°C. Se analizaron los niveles de proteína SOX2 y Lefty con anti-SOX2 de conejo (Abcam, Cambridge, MA, EE.UU. ab59776) (dilución 1: 2000) y anticuerpos anti-Lefty de conejo (Santa Cruz Biotechnology, sc-48836) (dilución 1: 1000). Se usaron anti- α -actinina de ratón, anti- β -actina o anti-GAPDH (dilución 1: 1000) (Santa Cruz Biotechnology, sc-17829, sc-47778, sc-32233, respectivamente) como controles de carga de proteína. Se usaron anticuerpos secundarios anti-conejo o anti-ratón (dilución 1: 10,000) (Santa Cruz Biotechnology, sc-2030; sc-2005, respectivamente), conjugados con HRP (peroxidasa de rábano). Las proteínas se detectaron utilizando el kit de sustrato de HRP quimioluminiscente de Immobilon Western (Merck Millipore, Billerica, MA, EE.UU., WBKLSO500).

7.10. Aislamiento de RNA y síntesis de cDNA

Las células CaSki se resuspendieron en 800 µl de Trizol Reagent (Life Technologies, Invitrogen, Carlsbad, CA). El RNA se aisló de acuerdo con el protocolo del fabricante. Los sedimentos de RNA se disolvieron en 20 µL de H₂O libre de DNasa. El RNA se cuantificó y se trató con DNasa I (Invitrogen). El cDNA se sintetizó usando la SuperScript II RNaseH-Reverse Transcriptase (Life Technologies, Invitrogen, Carlsbad, CA).

7.11. PCR cuantitativa

El cDNA generado a partir de células transfectadas se envió a qPCR usando la mezcla de PCR SYBR green (Applied Biosystems, 4309155), de acuerdo con el protocolo del fabricante. Los oligonucleótidos utilizados se señalan en la tabla 5.

Tabla 5. Oligonucleótidos empleados para el análisis por qPCR de un fragmento de los genes *E6*, *E7* y de la subunidad 18S del ribosoma.

	Oligonucleótidos	Secuencia
E6-VPH16	Sentido	GGGGGATCCATGTTTCAGGACCCACAGGAGCGA
E6-VPH16	Antisentido	GGGGAATTCTTACAGCTGGGTTTCTCTACGTGT
E7-VPH16	Sentido	GGGGGATCCATGCATGGAGATACACCTACATTGC
E7-VPH16	Antisentido	GGGAAGCTTTTATGGTTTCTGAGAACAGATGGGGCA
18S	Sentido	AACCCTTGAACCCATT
18S	Antisentido	CCATCCAATCGGTAGTAGCG

7.12. Análisis estadístico

La P se calculó mediante la prueba t de "Student" a partir de los datos presentados como media \pm Error Estándar (EE). Se aceptaron diferencias significativas a $P \leq 0.05$, como se indica.

8. RESULTADOS

8.1. La LCR del VPH16 y del 18 muestran diferente actividad transcripcional en líneas celulares de origen escamoso o de origen glandular

La regulación de la actividad transcripcional del VPH durante su ciclo replicativo puede ser influida por variaciones en las cantidades de FT o cofactores a lo largo del epitelio estratificado, así como en diferentes contextos celulares. Para determinar el efecto del contexto celular (escamoso o glandular) sobre la actividad transcripcional de la LCR del VPH16 en comparación con la del VPH18, se realizaron ensayos con vectores que contienen el gen de luciferasa como reportero bajo el control de la LCR16 o LCR18 en líneas celulares CaSki, HaCaT y HeLa. Se observó una actividad significativamente mayor del vector pGL2-LCR16 en comparación con el vector control pGL2-B (sin promotor) en las tres líneas celulares (**Figura 11A-C**). Sin embargo, la actividad fue considerablemente mayor en células con origen escamoso (SQ, por sus siglas en inglés squamous) (**Figura 11A, B**). Además, la actividad de la LCR16 en células HeLa (origen glandular), aunque significativo, no fue tan alta como en las células de origen escamoso (**Figura 11C**).

Por el contrario, la actividad de la LCR18 fue significativamente menor que la de la LCR16 en células CaSki (**Figura 11A**). Como se esperaba para un control positivo, el promotor SV40 presentó una alta actividad en líneas celulares CaSki y HeLa (**Figura 11A, C**), mientras que en células HaCaT fue 10 veces superior al vector pGL2-B, control negativo (**Figura 11B**); probablemente debido a la menor eficacia de transfección que se ve comúnmente en la línea celular HaCaT en comparación con las otras ¹⁴³. Sin embargo, está claro que la actividad de la LCR16 en células HaCaT es similar a la del control positivo de SV40. Estos resultados muestran un comportamiento diferente para la LCR del VPH16 con respecto a la del 18 dependiendo del contexto celular del origen tumoral (SQ o glandular). Para la LCR16, su actividad transcripcional se ve favorecida en el epitelio escamoso, posiblemente debido a la variación de la proporción de factores activadores con respecto a factores represores. Por otro

lado, las diferencias en la actividad transcripcional entre tipos de VPHs podría deberse a variaciones en el número, posición o afinidad de elementos de respuesta compartidos o a elementos de respuesta que están presentes en una LCR y ausentes en la otra.

Como previamente se había detectado que existían elementos de respuesta de SOX2 en la LCR del VPH16 y no en la LCR del VPH18, quisimos determinar los niveles de proteína del factor transcripcional SOX2 en las diferentes líneas celulares empleadas (HeLa, CaSki y HaCaT) y posteriormente revisar la posible correlación con la actividad transcripcional de la LCR del VPH16. El análisis mostró que la línea celular HeLa presentó la mayor cantidad de proteína, seguida por CaSki y finalmente HaCaT (**Figura 11D**). Al hacer una correlación de los niveles de la proteína SOX2 con la actividad transcripcional de la LCR del VPH16 se observó una correlación inversa con la actividad transcripcional, es decir, a mayor cantidad de SOX2 (HeLa) menor actividad transcripcional de la LCR del VPH16 (2.5 veces) y a menor cantidad de SOX2 (CaSki y HaCaT) mayor actividad transcripcional (CaSki, 18 veces y HaCaT, aunque 8 veces, prácticamente la actividad transcripcional se observó tan alta como la actividad transcripcional del control positivo en esta línea celular).

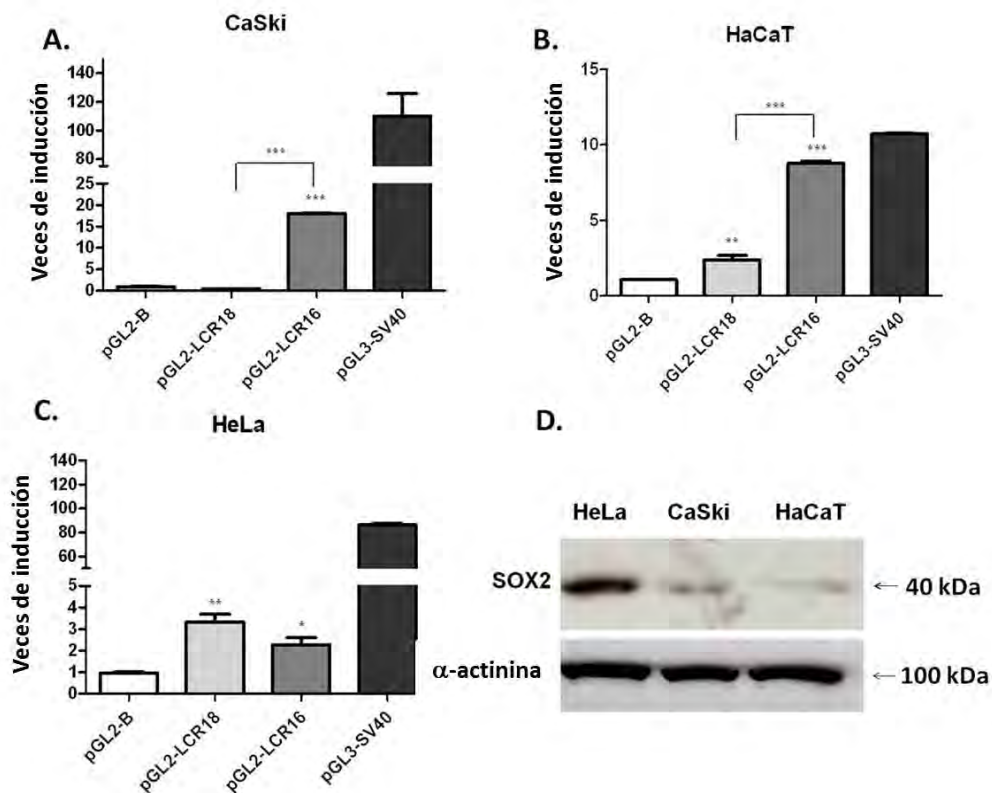


Figura 11. Las LCRs de los VPH16 y 18 presentan una actividad transcripcional diferente en células con un origen histológico distinto. Los plásmidos reporteros se transfectaron en líneas celulares CaSki (A), HaCaT (B) y HeLa (C). La actividad transcripcional se muestra como veces de inducción, en comparación con el vector pGL2-B. Se muestra el error estándar del promedio; * p <0.05, ** p <0.01, *** p <0.001. Los corchetes muestran diferencias entre las actividades transcripcionales de las LCRs. El plásmido pGL3-SV40 que alberga el promotor fuerte de Simian virus 40 (SV40), se usó como control positivo. Los gráficos muestran experimentos representativos llevados a cabo por triplicado y repetidos tres veces; (D) Western blot representativo de los niveles de proteína basal SOX2 en líneas celulares HeLa, CaSki y HaCaT. Para el control de carga, se usó α -actinina.

Analizamos los niveles de proteína SOX2 endógena en las diferentes líneas celulares mediante western blot. Como se muestra en la Figura 1D, las células HeLa expresaron el nivel más alto de SOX2, seguido de CaSki; mientras que HaCaT presentó los niveles más bajos (Figura 1D). Aunque muchos factores podrían estar involucrados en las diferencias de transcripción de las LCRs observadas entre nuestras líneas celulares, es notable que los niveles más bajos de proteína SOX2 coinciden con una actividad transcripcional relativamente más alta de la LCR16.

8.2. SOX2 se une a la LCR del VPH16

8.2.1. SOX2 se une a sus sitios putativos de unión en la LCR del VPH16

Para determinar si el factor transcripcional SOX2 reconoce y se une a cada uno de los tres elementos putativos de respuesta a SOX2 encontrados en la LCR del VPH16, se realizó EMSA. Se emplearon extractos de proteínas nucleares de la línea celular HeLa, ya que esta línea celular presentó la mayor cantidad del factor transcripcional SOX2, según el western blot de la **Figura 11D**. Se diseñaron oligonucleótidos cortos de doble cadena de DNA, con la secuencia consenso (CS, por sus siglas en inglés consensus sequence, como control positivo), la secuencia de cada uno de los sitios putativos de unión a SOX2 y una secuencia no específica (NS, por sus siglas en inglés non specific, como control negativo), en la parte

central de cada oligonucleótido. La formación de un complejo principal de DNA-proteína, señalado con un asterisco (*), se observó cuando los extractos nucleares se incubaron con la SC de unión a SOX2, marcada con γ -ATP ³² fosforo (por sus siglas en inglés, ³²Phosphorus) y no se observó cuando los extractos nucleares se incubaron con los oligonucleótidos que contienen la secuencia NS (**Figura 12A**). Con el fin de evaluar si los sitios S1, S2 y S3 son capaces de competir con el complejo DNA-proteína generado por el CS, se usaron diferentes cantidades de sondas frías (no marcadas con ³²P) de S1, S2 o S3. La formación del complejo con la sonda CS se compitió fuertemente con la adición de un exceso de 60 (**a**) y principalmente con 260 (**b**) veces la sonda fría o no marcada radioactivamente; mientras que el oligonucleótido frío no específico (NS), no compitió (**Figura 12A**). Para las sondas S1, S2 o S3, se observaron dos complejos principales de DNA-proteína (*) (**Figura 12B-D**). Se observó un comportamiento similar cuando los tres sitios SOX2 (S1, S2 y S3) competían inversamente contra un exceso de 50 (**a**) y 250 (**b**) veces de sonda fría CS, donde se observó una clara disminución en los complejos DNA-proteína, que no se observó con la sonda fría NS (**Figura 12B-D**). Estos datos sugieren que SOX2 puede unirse a los tres sitios de unión propuestos que se encuentran en la LCR del VHP 16.

Para demostrar que la proteína SOX2 está presente en el complejo DNA-proteína, se realizó un ensayo de super-retardo. La presencia de SOX2 en los complejos de DNA-proteína se confirmó después de la adición de 50 ng (**c**) o 2500 ng (**d**) de anticuerpo anti-SOX2, dando como resultado una banda super-retardada (SS, por sus siglas en inglés super shift) en el complejo CS cuando se usan 2500 ng de anticuerpo (**d**) (**Figura 12E**). Se observó una reducción significativa de los complejos ADN-proteína de las sondas S1, S2 y S3 con 2500 ng (**d**) de anticuerpo SOX2 (**Figura 12F-H**). Se confirmó la especificidad del anticuerpo SOX2, ya que no se produjo ningún cambio con 50 o 2500 ng (**c** y **d**) de anticuerpo anti-CTCF que reconoce a un factor transcripcional que se sabe que no se une a la LCR del VPH16 (**Figura 12E-H**). Por lo tanto, nuestros resultados *in vitro* apoyan la interacción de SOX2 con los tres sitios putativos de unión presentes en la LCR del VPH16.

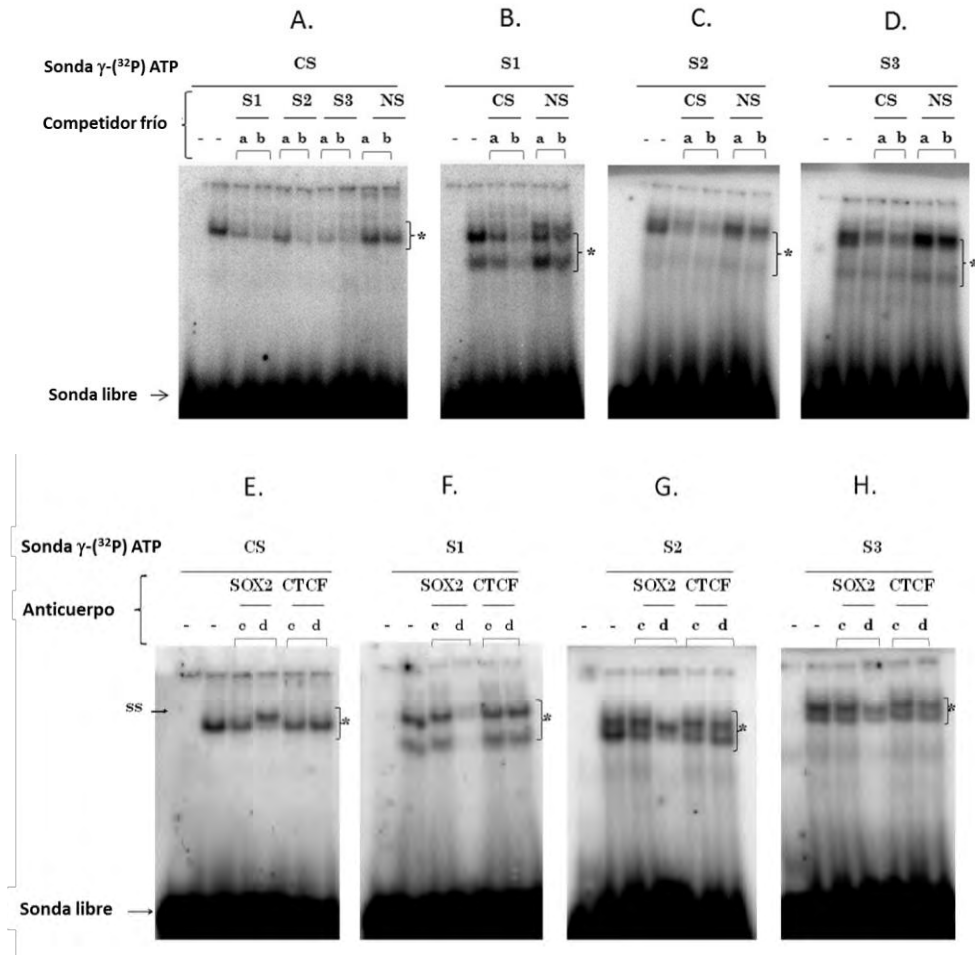


Figura 12. SOX2 se une a los elementos de respuesta putativos en la LCR del VPH16. Los extractos de proteína nuclear HeLa se usaron en el EMSA para analizar las interacciones entre SOX2 y oligonucleótidos de DNA de doble cadena que contienen los sitios putativos de unión a SOX2: S1, S2 o S3 y el sitio de unión consenso (CS). Los primeros carriles de todos los EMSA no contienen extracto nuclear; los carriles restantes contienen 15 μ g de extracto nuclear. Las competencias con 60 (a) o 260 (b) veces de exceso de S1, S2, S3, NS (no específicas) y CS sondas no marcadas o frías se muestran (A-D). Los complejos DNA-proteína se indican mediante (*). La sonda libre marcada se muestra en la parte inferior de los EMSAs. Los EMSAs E-H muestran los complejos DNA-proteína con los respectivos anticuerpos añadidos (anti-SOX2 o anti-CCCTC (CTCF), como se indica, 50 ng (c) o 2500 ng (d). SS es el super-retardo formado con la sonda CS y 2500 ng (d) de anticuerpo anti-SOX2 (E). En F-H, se observó una disminución clara en los complejos DNA-proteína de cada sonda (S1, S2 y S3) con 2500 ng (d) de anti-SOX2.

8.2.2. SOX2 se une a la LCR del VPH16 *in vivo*

Con el fin de confirmar que SOX2 se une a la LCR del VPH16 en las células, se realizaron ensayos de Inmunoprecipitación de la cromatina (ChIP) con extractos de cromatina obtenidos de células CaSki, que contienen al VPH16. Los resultados de ChIP demuestran que SOX2 se une a la LCR del VPH16 en las células (*in vivo*). Como control positivo empleamos anticuerpos anti-p65, ya que se sabe que p65 se une a VPH16-LCR¹⁴¹[31]. Previamente, se informó que CTCF no se recluta por la LCR del VPH16, por lo tanto, utilizamos anticuerpos anti-CTCF como control de no unión¹⁴².

Los resultados de ChIP muestran que el fragmento de DNA que contiene los sitios de unión a SOX2 en la LCR del VPH16 se amplificó por qPCR a partir de la muestra inmunoprecipitada con anti-SOX2 (**Figura 13**) lo que significa que hubo unión. Tal enriquecimiento fue mayor que el 8% observado cuando se inmunoprecipita con anticuerpo anti-p65 que va dirigido contra un factor transcripcional que se sabe se une a la LCR (control positivo). Cuando se usó anti-CTCF como control negativo de unión, no hubo una amplificación importante del fragmento de la LCR. Es interesante mencionar que la proporción de unión de SOX2 a la LCR es incluso mayor que la de su blanco transcripcional Lefty (5%), utilizado como control positivo de la unión de SOX2. Estos resultados demuestran que SOX2 se une al VPH16-LCR en las células CaSki *in vivo*.

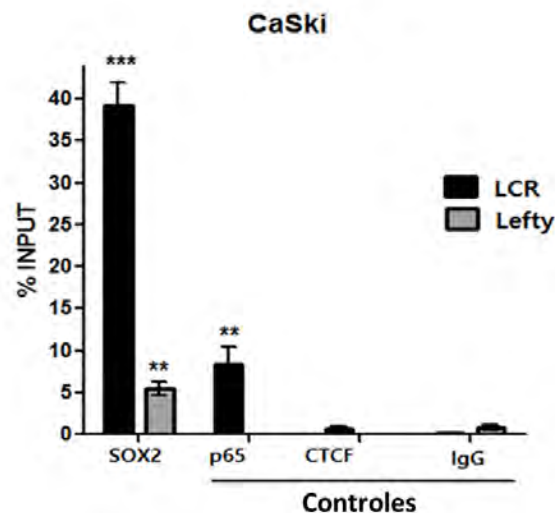


Figura 13. SOX2 se une a los sitios de unión putativos en la LCR del VPH16 *in vivo*. El ensayo de inmunoprecipitación de cromatina (ChIP) de SOX2 se realizó evaluando la región de la LCR del VPH16 de las células CaSki. Los anticuerpos anti-p65 y -CTCF se emplearon como controles positivo y negativo, respectivamente. El anticuerpo IgG no específico se incluyó como control negativo. El análisis de PCR cuantitativa (qPCR) se realizó con el DNA obtenido del ensayo ChIP, que evalúa la región de la LCR del VPH16 (barras negras) y la región promotora del gen Lefty (barras grises). Como un control positivo de la unión de SOX2 (n = 2). El gráfico representa el porcentaje de INPUT recuperado después de la inmunoprecipitación con anticuerpos anti-SOX2, -p65 o -CTCF. Se realizaron tres experimentos biológicos independientes. Se muestra el error estándar de la media de las mediciones de qPCR por triplicado. ** p <0.01, *** p <0.001.

8.3. SOX2 reprime la actividad transcripcional de la LCR del VPH16

Con el fin de dilucidar el efecto de SOX2 en la regulación transcripcional de la LCR del VPH16, las células CaSki se transdujeron con vectores lentivirales que portaban SOX2 y / o GFP como control. Después de 24 h, las células se transfectaron con plásmidos reporteros pGL2-LCR16 o pGL2-LCR18. Las células se procesaron 48 h después. Para evaluar la presencia de un SOX2 funcional sobreexpresado, los niveles de proteína de Lefty, un blanco conocido del gen SOX2, también se midieron en días diferentes después de la transducción. Como se esperaba, los niveles de la proteína Lefty disminuyeron a medida que aumentaba SOX2 (**Figura 14A**). A las 48 h después de la transfección, se observó una represión significativa de la LCR16 debido a la sobreexpresión de SOX2, permaneciendo de 3.5 veces por debajo de las células transducidas con GFP (**Figura 14B**). La actividad de la LCR18 no fué afectada por la sobreexpresión de SOX2. Estos resultados indican que SOX2 está actuando como un represor de LCR16 en células CaSki.

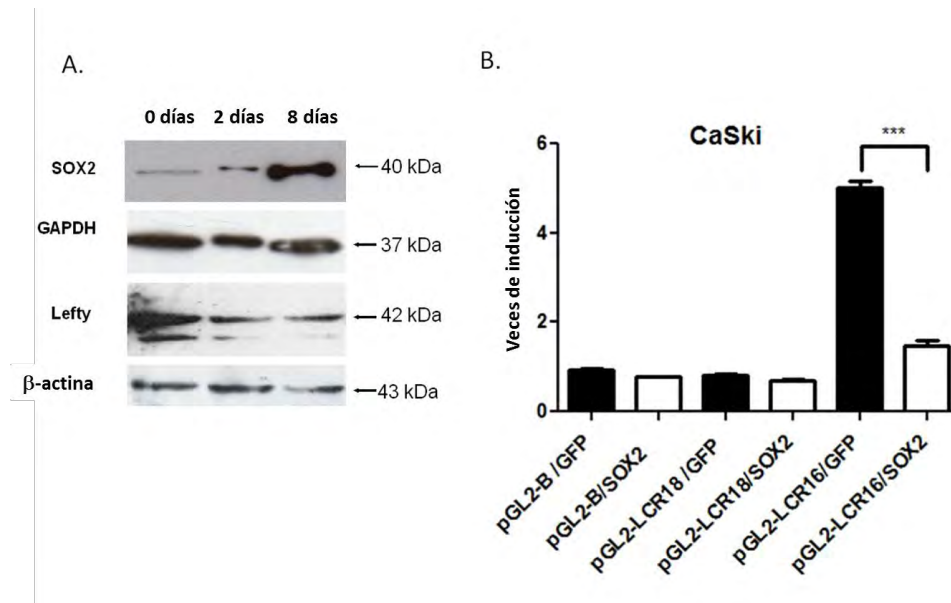


Figura 14. SOX2 reprime la actividad transcripcional de la LCR del VPH16. (A) SOX2 se sobreexpresó en células CaSki usando un sistema lentiviral. Se muestran los niveles representativos de expresión de SOX2 y su blanco Lefty, en diferentes días después de la transducción (0, 2 y 8 días). Glyceraldehyde-3-Phosphate Dehydrogenase (GAPDH) y anticuerpos β-actina se utilizaron como controles de carga de proteína; (B) Se transfectaron células CaSki transducidas con vectores lentivirales que expresan SOX2 y / o proteína fluorescente verde (GFP) con plásmidos reporteros de luciferasa que contienen LCR16 o LCR18. La actividad transcripcional después de 48 h, se muestra como veces de inducción, en comparación con el vector pGL2-B / GFP. El gráfico muestra un experimento representativo llevado a cabo por triplicado y repetido tres veces. Se muestra el error estándar de la media. *** $p < 0.001$.

8.4. Los sitios S1, S2 y S3 de la LCR del VPH16 son necesarios para la represión mediada por SOX2

Para probar si la unión de SOX2 a sitios S1, S2 y S3 está implicada en la represión transcripcional de la LCR16 observada, se generaron mutaciones en el vector pGL2-LCR16 (wt, tipo salvaje): 2 pares de bases (pb) mutados en el sitio S1 (M1), 3-pb mutado en cada uno de los sitios S2 y S3 (M2 y M3, respectivamente), y las mutaciones de 8-pb se introducen en la triple mutante (M1,2,3) como se muestra en los electroferogramas de la **Figura 15**. Como se muestra a través de la actividad de luciferasa, se observó claramente una recuperación de la actividad transcripcional de la LCR16, en comparación con la LCR16wt, con M1, M2 y M1,2,3 (barras blancas), sugiriendo que las

mutaciones respectivas escapan de los efectos represivos de SOX2 endógeno. Para determinar si SOX2 está involucrado en este efecto, la proteína se sobre-expresó ectópicamente (barras negras). La actividad de la LCR16wt se reprimió hasta 66.2% y M1, M3 y M1,2,3 no fueron reprimidos por sobreexpresión de SOX2; mientras que M2 fue levemente reprimido en una proporción de 27.5%. Estos datos confirmaron que los tres sitios de unión a SOX2: S1, S2 y S3 contribuyen a la actividad represora de SOX2 (**Figura 16**) y que existe una interacción directa de SOX2 con los tres sitios.

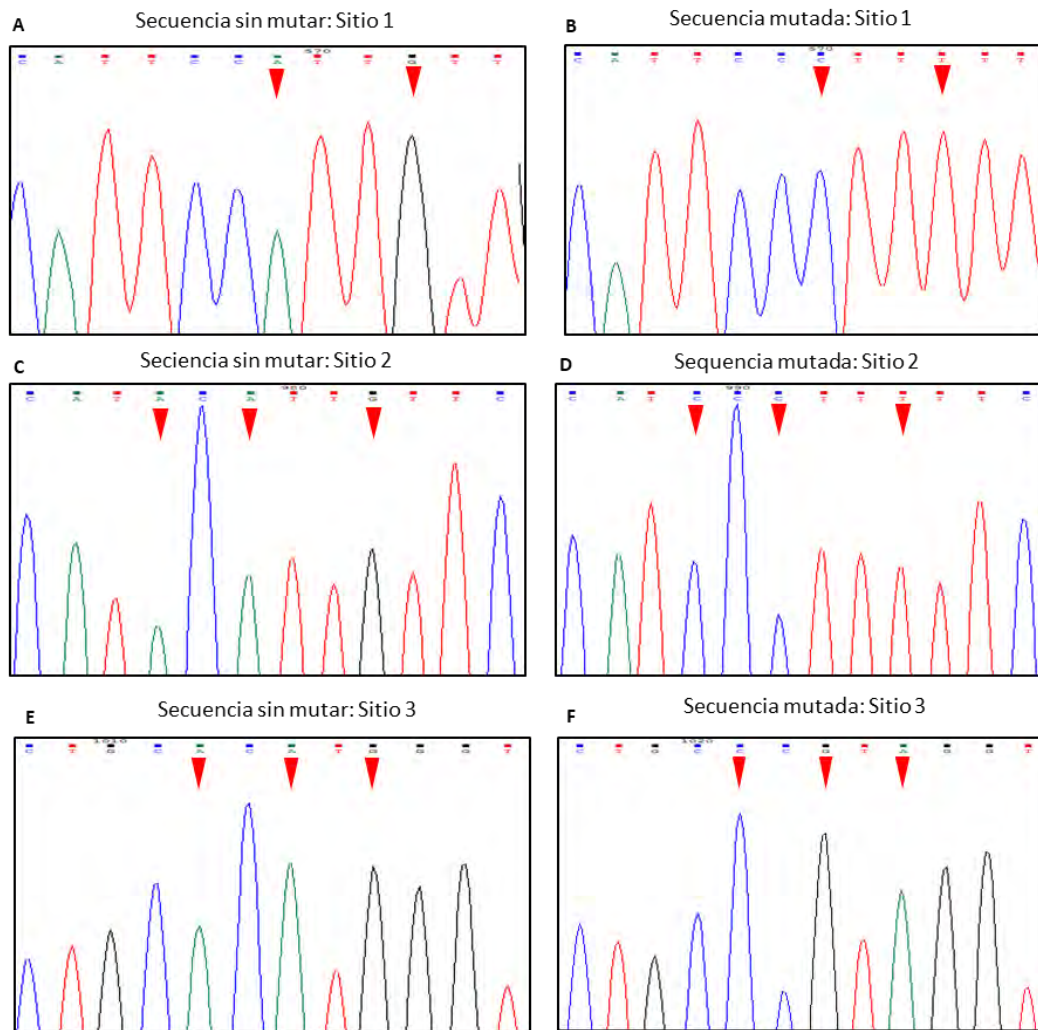


Figura 15. Electroferogramas donde se identifican los nucleótidos mutados en la región larga de control (LCR) del VPH16 por mutagénesis multi-sitio dirigida. Se muestra la secuencia sin mutaciones de los sitios de unión a SOX2 1, 2 y 3 (A, C y E respectivamente). Las imágenes B, D y F muestran las secuencias mutadas de los sitios de unión a SOX2. Los nucleótidos señalados con una flecha son aquellos que fueron cambiados por mutagénesis.

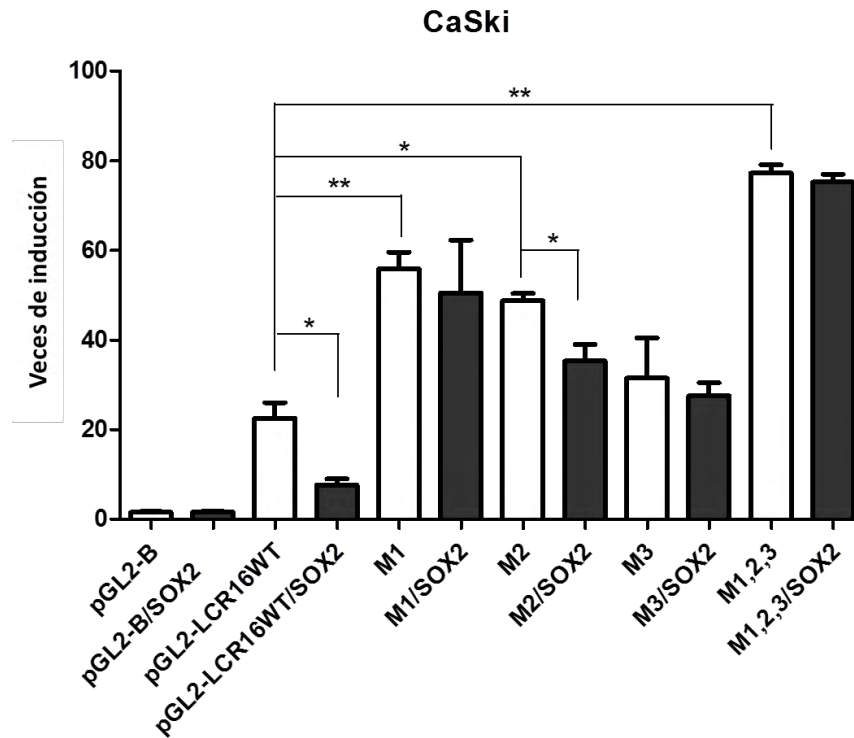


Figura 16. Los sitios S1, S2 y S3 de la LCR del VPH16 participan de manera diferente en la actividad represora de SOX2. La mutagénesis multi sitio-dirigida se realizó en el plásmido pGL2-LCR16 (wt, tipo salvaje), generando M1 (mutante LCR-Sitio 1), M2 (mutante LCR-Sitio 2), M3 (mutante LCR-Sitio 3) y M1, 2, 3 mutantes (triple mutante). La actividad transcripcional de los plásmidos reporteros se midió en células CaSki en presencia (barras negras) o ausencia (barras blancas) de SOX2 expresada ectópicamente. La inducción de la actividad de la LCR16 se calculó en relación con el vector control pGL2-B. Las actividades transcripcionales de las mutantes de LCR16 se compararon con actividad del plásmido pGL2-LCR16wt (barras blancas). La actividad transcripcional de cada LCR16 (M1, M2, M3 o M1,2,3) se comparó con su condición respectiva con y sin SOX2 (barras blancas contra negras). Se muestra el error estándar de la media. Prueba t * p < 0.05, ** p < 0.01.

8.5. La expresión endógena de E6 y E7 es regulada por SOX2

Dado que los oncogenes virales E6 y E7 son regulados directamente por la LCR, analizamos si SOX2 estaba actuando como un represor transcripcional de tales oncogenes *in vivo*. Para analizar el efecto de SOX2 sobre los niveles endógenos de los transcritos E6 y E7 dos cantidades de plásmido que expresa el factor de transcripción SOX2 (150 y 300 ng) fueron transfectadas en células CaSki (Figura 17A). Como se muestra por western blot, la proteína SOX2 se sobreexpresó en células CaSki (Figura 17A). El análisis de PCR cuantitativa (qPCR) muestra que los niveles de transcripción

E6 (Figura 17B) y *E7* (Figura 17C) disminuyeron significativamente en presencia de SOX2 (aproximadamente 40 a 33%, respectivamente). No se observó represión dosis dependiente.

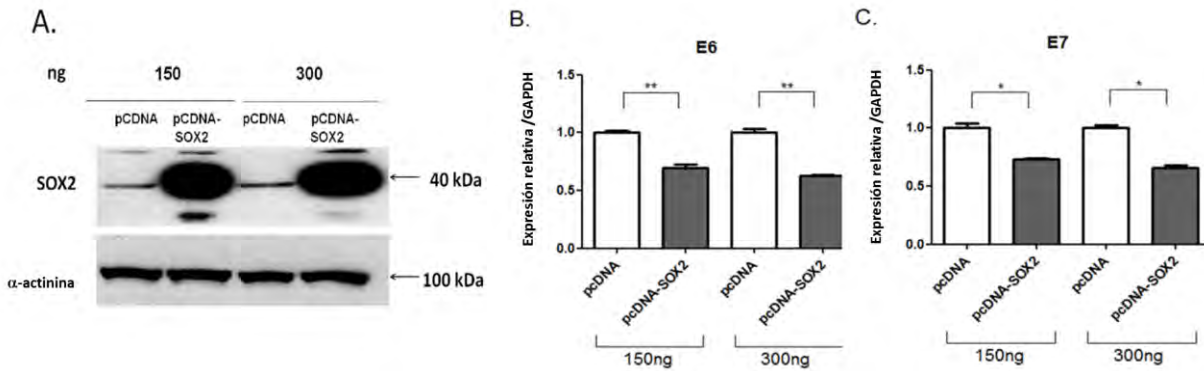


Figura 17. SOX2 reprime la expresión de los oncogenes virales *E6* y *E7* endógenos. Las células CaSki se transfectaron con 150 o 300 ng de plásmido que expresa SOX2. (A) Se muestran los niveles de proteína SOX2 después de la transfección. Para el control de carga de proteína, se analizaron los niveles de α -actinina. Se realizó una reacción en cadena de la polimerasa de transcripción reversa cuantitativa (qRT-PCR) de los oncogenes virales *E6* (B) y *E7* (C) endógenos en presencia o ausencia de SOX2 sobreexpresado. El vector de pCDNA vacío se usó como control. Los niveles de expresión *E6* y *E7* disminuyeron significativamente en presencia de SOX2. Los datos se normalizaron con respecto al gen constitutivo *GAPDH*. Se muestra el error estándar de la media. * $p < 0.05$; ** $p < 0.01$.

9. DISCUSIÓN

A nivel mundial y en diversas poblaciones analizadas, el VPH16 es el tipo de VPH más prevalente en los cánceres de cérvix ³⁵. Esto puede ser reflejo del grado de oncogenicidad del VPH16. Sin embargo, diversos factores, aun no esclarecidos, del hospedero, de las células infectadas o propias del tipo viral, podrían determinar la alta prevalencia del VPH16 observada particularmente en el carcinoma de células escamosas del cuello uterino, en relación con otros VPHs oncogénicos. Entre ellos, es posible que un conjunto de factores de transcripción específicos pueda guiar diferencialmente la actividad transcripcional viral en las diferentes líneas celulares. Nuestro interés inicial fue determinar el efecto que el contexto celular puede ejercer en la actividad transcripcional de las LCRs de los VPH16 y VPH18. Esto fue analizado mediante ensayos con plásmidos reporteros. Nuestros resultados indicaron que en las líneas celulares con un contexto escamoso, CaSki y HaCaT, la actividad de la LCR16 fue significativamente más alta que la de la LCR18 como lo esperabamos. Esto nos llevó a pensar que posiblemente un conjunto particular de factores transcripcionales que estén presentes en células escamosas favorece la actividad del VPH16. Por el contrario, en la línea celular HeLa, que es un adenocarcinoma, la actividad transcripcional de las LCRs, son más bajas para ambos tipos virales y similares entre ellos. Esto sugiere que el contexto celular escamoso puede favorecer la actividad de la LCR16, mientras que un contexto de factores transcripcionales presentes en las células con origen glandular, puede reprimirla. Por otra parte, hay que considerar que la actividad diferencial entre la LCR16 y LCR18 en las tres líneas celulares podría deberse a las diferencias en número, posición o afinidad de sitios de unión a factores transcripcionales (TFBS) compartidos en las secuencias virales, o a la presencia de TFBS específicos para cada LCR.

Para identificar nuevos FT que pudieran modular la actividad transcripcional diferencial de la LCR16 y LCR18, en un programa *in silico* (JASPAR CORE) buscamos sitios de unión putativos para FT específicos. Como era de esperar, encontramos algunos TFBS compartidos entre ambos tipos de virus; sin embargo, había otros

específicos para cada tipo viral. De manera interesante se encontraron sitios de unión consenso para el factor transcripcional SOX2, exclusivamente en la secuencia de la LCR del VPH16. SOX2 es miembro de la familia de factores transcripcionales SOX (región determinante del sexo Y (SRY) y es un regulador clave de la embriogénesis y el destino celular durante el desarrollo ¹⁴⁴, participando además en el mantenimiento de la pluripotencia ¹⁴⁵. Aunque el papel predominante de SOX2 se ha descrito como un activador transcripcional, recientemente se ha demostrado que desempeña un doble papel, dado que también actúa como un represor de la expresión génica ^{146,147}. Este factor está involucrado en el desarrollo normal, así como en los procesos malignos ¹²⁸.

SOX2 se ha relacionado con las características distintivas del cáncer, promoviendo la proliferación celular en cáncer de mama, páncreas, próstata y cervical ^{137,148,139,149,121}, así como evadiendo la apoptosis en cáncer de próstata, gástrico y pulmón ^{149,150,151}. Además, la expresión de SOX2 se ha asociado con eventos tempranos de carcinogénesis cervical ¹³³, ya que está altamente expresada en lesiones premalignas, así como en cáncer invasor, mientras que su expresión es baja en el epitelio cervical normal. Se ha reportado que la sobreexpresión de SOX2 en células cancerosas del cuello uterino aumenta la proliferación, la clonogenicidad y la tumorigenicidad ¹³⁷. Estos resultados sugieren que SOX2 tiene un papel esencial en la carcinogénesis cervical. Sin embargo, el efecto de SOX2 en la regulación transcripcional de los VPHs de alto riesgo no se ha estudiado previamente.

En este estudio, demostramos que SOX2 se une a tres sitios de unión identificados en la LCR16 a través de interacciones directas, como se muestra en los ensayos de EMSA y CHIP. También confirmamos que SOX2 es un represor transcripcional de la LCR16, ya que la actividad disminuyó 3.5 veces cuando SOX2 se sobreexpresó en células CaSki (**Figura 14**). Como se informó anteriormente, la LCR del VPH16 y VPH18 comparten un repertorio de TFBS que regulan su actividad ⁸⁷. Sin embargo, consideramos que los TFBS que son específicos para cada tipo de LCR pueden explicar, en parte, los datos

epidemiológicos relacionados con la diferente distribución de los tipos de VPH oncogénicos, dependiendo de las características histológicas celulares.

En el presente trabajo demostramos la interacción directa de SOX2 con sus sitios putativos de unión en la LCR16 a través del análisis de la actividad transcripcional de los sitios de unión mutados. Se observó una recuperación significativa de la actividad de luciferasa con las mutantes M1, M2 y M1,2,3 en el estado basal, lo que sugirió que las mutantes de la LCR16 escaparon de la represión endógena de SOX2. Además, cuando SOX2 fue sobreexpresada ectópicamente, no se observó represión de la LCR16 M1 y M3, mientras que la mutante M2 mostró una represión suave (27.5%), en contraste con lo observado para la LCR16wt, que mostró una fuerte represión (66.2%). Estos datos demuestran que los tres sitios de unión a SOX2 contribuyen a la actividad represora de SOX2 en la LCR del VPH16.

La actividad represora de SOX2 sobre la LCR16 se confirmó en células CaSki que contienen secuencias endógenas del VPH16, donde los oncogenes *E6* y *E7* son naturalmente expresados. El análisis de RT-PCR cuantitativo mostró que los niveles de transcripción de *E6* y *E7* disminuyeron significativamente en las células a las que se les sobreexpresó SOX2, aunque no se observó un efecto dosis dependiente, al utilizar las distintas dosis de plásmido de expresión de SOX2, probablemente debido a una saturación de los sitios putativos de unión a este factor transcripcional.

Con estos datos, demostramos un efecto directo de SOX2 sobre la LCR del VPH16, aunque no podemos descartar la contribución de otros factores que posiblemente sean afectados por SOX2 y que a su vez modulen a la LCR16. Por ejemplo, Bass et al. (2009) demostraron que un aumento en la expresión de varios FT de células troncales se correlacionaba con altos niveles de expresión de SOX2 en SCC de pulmón ¹⁵²; por lo tanto, no podemos descartar la posibilidad de que otros factores de transcripción, regulados por SOX2, también puedan contribuir a la modulación de la actividad de la LCR.

Es interesante que el análisis de los niveles de la proteína SOX2 en las diferentes líneas celulares mostró que las células HeLa tuvieron los niveles de proteína más altos, en comparación con las líneas SCC (CaSki y HaCaT). Estos datos coinciden con las actividades de transcripción más altas observadas para LCR16 en el contexto escamoso, en el que la transcripción del VPH16 podría verse favorecida debido a un contexto menos represivo (menos factores transcripcionales represores). Por lo tanto, bajo estas condiciones, SOX2 no actúa como un activador de la LCR16 en las células CCE, sino como un represor en las células de adenocarcinomas como HeLa, lo que sugeriría que en este contexto, las LCRs de virus tipo 16 estarían más reprimidas. De esta manera, altos niveles de SOX2 en células HeLa podrían tener una participación en la baja actividad observada para la LCR16 en este tipo de células. Sin embargo, es importante considerar que otros factores transcripcionales que no analizamos en este trabajo seguramente influyen en la actividad de las LCRs de los VPH16 y VPH18.

Es importante señalar que, aunque en nuestro trabajo no profundizamos en las distintas actividades transcripcionales de la LCR del VPH18 observadas en los contextos celulares analizados, es un tema que merece estudios adicionales. Los ensayos con vectores reporteros mostraron una clara disminución de la actividad transcripcional de la LCR18 en un contexto de células escamosas. Además, el análisis *in silico* también mostró supuestos sitios putativos de unión a factores transcripcionales que son específicos para la LCR del VPH18. Por lo tanto, sería relevante estudiar qué factores novedosos podrían modular la actividad de esta LCR.

La regulación de la LCR del VPH16 por SOX2 podría ser biológicamente relevante en las infecciones tempranas por el VPH. El VPH puede infectar células basales y células troncales^{153,154} donde se podría esperar una sobreexpresión de SOX2. De hecho, se han demostrado altos niveles de SOX2 en el cáncer de cuello de útero¹⁵⁵ y en particular en la subpoblación de células tumorales con propiedades similares a las troncales¹⁵⁶. En tal escenario, una expresión controlada de los oncogenes E6 y E7 por SOX2 podría retardar el daño celular y por lo tanto, podría favorecer la persistencia viral. Profundizar en el

estudio de este tema podría ayudar a explicar, en parte, la alta prevalencia del VPH16 en el carcinoma de células escamosas.

Es importante mencionar que se ha reportado una alta expresión de *SOX2* en carcinoma de cérvix de células escamosas; sin embargo, también hay trabajos que muestran la pérdida de *SOX2* en un subgrupo de casos de cáncer cervical donde se ha asociado con un peor pronóstico ¹³³. Esto nos lleva a pensar que en este último caso, la pérdida de *SOX2* podría favorecer la sobreexpresión de E6 y E7 de VPH16, lo cual podría explicar el desenlace poco favorable para estas pacientes. Apoyando esta idea, cabe mencionar que se ha demostrado que el aumento en la expresión del RNAm de oncogenes del VPH se asocia con un peor pronóstico en pacientes con cáncer de cuello uterino ¹⁵⁷. De esta manera, si *SOX2* actuara como un represor transcripcional de la LCR del VPH16 *in vivo*, su disminución en las células cancerosas podría eventualmente promover un aumento en la actividad de la LCR, que se reflejaría en la sobreexpresión de los oncogenes *E6* y *E7*. Para estudiar más a fondo esta posibilidad, sería necesario determinar si en las biopsias de cáncer cervical positivo para VPH16, la cantidad de *SOX2* está inversamente relacionada con los niveles de expresión de los oncogenes *E6* y *E7*.

Un aspecto que resulta importante considerar es que *SOX2* se asocia con diversas proteínas celulares, incluyendo coactivadores y co-represores transcripcionales y que tales asociaciones pueden modular diferencialmente a los genes blanco ¹⁵⁸. En células madre neuronales se ha demostrado que cuando *SOX2* actúa como un represor, requiere la interacción con miembros de la familia de represores Groucho ¹⁵⁹. Por lo tanto, el análisis de estas posibles interacciones en células de cáncer de cuello uterino podría proporcionar información relevante para comprender los mecanismos de regulación transcripcional del VPH por *SOX2*.

La regulación negativa de la LCR16 por *SOX2* sobre podría explicar en parte la menor prevalencia de VPH16 en adenocarcinomas de cérvix comparada con la de carcinoma de células escamosas, dado que es probable que en células con origen de

adenocarcinoma, como HeLa, existan mayores niveles de SOX2. Asimismo, la alta prevalencia de VPH16 en carcinomas de células escamosas podría deberse al ambiente celular menos represivo y por lo tanto más favorable para la actividad transcripcional de VPH16.

En este estudio, identificamos a SOX2 como un nuevo protagonista en el control transcripcional del VPH16. Una mayor comprensión de la regulación diferencial de tipos de VPH de acuerdo al contexto celular ayudará a explicar las diferencias en su comportamiento biológico, considerando además del tipo celular, si está diferenciada o no y el grado de avance de la lesión.

10. CONCLUSIONES

- El factor de transcripción SOX2 se une a la región larga de control del VPH16 y reprime su actividad transcripcional.
- La sobreexpresión de SOX2 reprime los niveles de *E6* y *E7* en la línea celular CaSki.
- SOX2 no regula la actividad transcripcional de la LCR de VPH18.
- Existe una relación inversa entre los niveles de SOX2 y la actividad transcripcional de VPH16 en diferentes contextos celulares.

11. PERSPECTIVAS

Resulta relevante analizar los niveles de SOX2, así como de los transcritos de *E6* y *E7* en muestras de tejido sano, lesiones premalignas y cáncer invasor de cérvix para determinar si hay alguna correlación inversa entre los niveles de estas proteínas, que asemeje a lo observado en este trabajo y que explique en parte la regulación de VPH16 durante la evolución hacia cáncer.

Asimismo, dado que otros factores transcripcionales pueden tener relevancia en la regulación transcripcional diferencial de las LCRs de los tipos virales 16 y 18, es importante establecer cuál es el efecto de otros factores transcripcionales sobre la regulación de promotores virales específicos.

12. REFERENCIAS

1. Genética del cáncer - National Cancer Institute.
<https://www.cancer.gov/espanol/cancer/causas-prevencion/genetica#ui-id-2>. Accessed October 11, 2018.
2. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011;144(5):646-674. doi:10.1016/j.cell.2011.02.013
3. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell*. 2000;100(1):57-70.
<http://www.ncbi.nlm.nih.gov/pubmed/10647931>. Accessed May 18, 2018.
4. OMS. <http://www.who.int/es/news-room/fact-sheets/detail/cancer>. Accessed November 5, 2018.
5. Ferlay J, Colombet M, Soerjomataram I, et al. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. *Int J Cancer*. October 2018. doi:10.1002/ijc.31937
6. Herfs M, Yamamoto Y, Laury A, et al. A discrete population of squamocolumnar junction cells implicated in the pathogenesis of cervical cancer. *Proc Natl Acad Sci*. 2012;109(26):10516-10521. doi:10.1073/pnas.1202684109
7. Hasan H. *Cervical Cancer : Current and Emerging Trends in Detection and Treatment*. Rosen Pub; 2009.
8. Reproducción asistida. <https://www.reproduccionasistida.org/fertilidad-en-la-mujer/fertilidad-mujer/>.
9. Eurocytology. <https://www.eurocytology.eu/es/course/465>.
10. Muñoz N, Franceschi S, Bosetti C, et al. Role of parity and human papillomavirus in cervical cancer: the IARC multicentric case-control study. *Lancet (London, England)*. 2002;359(9312):1093-1101. doi:10.1016/S0140-6736(02)08151-5
11. Plummer M, Herrero R, Franceschi S, et al. Smoking and cervical cancer: pooled analysis of the IARC multi-centric case--control study. *Cancer Causes Control*. 2003;14(9):805-814. <http://www.ncbi.nlm.nih.gov/pubmed/14682438>. Accessed July 6,

2018.

12. Moreno V, Bosch FX, Muñoz N, et al. Effect of oral contraceptives on risk of cervical cancer in women with human papillomavirus infection: the IARC multicentric case-control study. *Lancet (London, England)*. 2002;359(9312):1085-1092. doi:10.1016/S0140-6736(02)08150-3
13. International Collaboration of Epidemiological Studies of Cervical Cancer, Appleby P, Beral V, et al. Cervical cancer and hormonal contraceptives: collaborative reanalysis of individual data for 16,573 women with cervical cancer and 35,509 women without cervical cancer from 24 epidemiological studies. *Lancet (London, England)*. 2007;370(9599):1609-1621. doi:10.1016/S0140-6736(07)61684-5
14. Abraham AG, D'Souza G, Jing Y, et al. Invasive cervical cancer risk among HIV-infected women: a North American multicohort collaboration prospective study. *J Acquir Immune Defic Syndr*. 2013;62(4):405-413. doi:10.1097/QAI.0b013e31828177d7
15. Grulich AE, van Leeuwen MT, Falster MO, Vajdic CM. Incidence of cancers in people with HIV/AIDS compared with immunosuppressed transplant recipients: a meta-analysis. *Lancet (London, England)*. 2007;370(9581):59-67. doi:10.1016/S0140-6736(07)61050-2
16. International Collaboration of Epidemiological Studies of Cervical Cancer. Cervical carcinoma and reproductive factors: collaborative reanalysis of individual data on 16,563 women with cervical carcinoma and 33,542 women without cervical carcinoma from 25 epidemiological studies. *Int J cancer*. 2006;119(5):1108-1124. doi:10.1002/ijc.21953
17. Hoover RN, Hyer M, Pfeiffer RM, et al. Adverse Health Outcomes in Women Exposed In Utero to Diethylstilbestrol. *N Engl J Med*. 2011;365(14):1304-1314. doi:10.1056/NEJMoa1013961
18. Globocan 2012 - Home. <http://globocan.iarc.fr/Default.aspx>. Accessed April 3, 2018.
19. Schiffman M, Castle PE, Jeronimo J, Rodriguez AC, Wacholder S. Human papillomavirus and cervical cancer. *Lancet (London, England)*. 2007;370(9590):890-907. doi:10.1016/S0140-6736(07)61416-0

20. Trottier H, Franco EL. The epidemiology of genital human papillomavirus infection. *Vaccine*. 2006;24 Suppl 1:S1-15. doi:10.1016/j.vaccine.2005.09.054
21. Bodily J, Laimins LA. Persistence of human papillomavirus infection: keys to malignant progression. *Trends Microbiol*. 2011;19(1):33-39. doi:10.1016/j.tim.2010.10.002
22. Stanley M. Immunobiology of HPV and HPV vaccines. *Gynecol Oncol*. 2008;109(2 Suppl):S15-21. doi:10.1016/j.ygyno.2008.02.003
23. Wilting SM, Steenbergen RDM. Molecular events leading to HPV-induced high grade neoplasia. *Papillomavirus Res (Amsterdam, Netherlands)*. 2016;2:85-88. doi:10.1016/j.pvr.2016.04.003
24. Green J, Berrington de Gonzalez A, Sweetland S, et al. Risk factors for adenocarcinoma and squamous cell carcinoma of the cervix in women aged 20-44 years: the UK National Case-Control Study of Cervical Cancer. *Br J Cancer*. 2003;89(11):2078-2086. doi:10.1038/sj.bjc.6601296
25. Types and grades | Cervical cancer | Cancer Research UK. <https://www.cancerresearchuk.org/about-cancer/cervical-cancer/stages-types-grades/types-and-grades>. Accessed October 12, 2018.
26. Amălinei C, Balan R, Stolnicu S, Rădulescu D, Boeru C, Cotuțiu C. Adenosquamous cervical carcinoma morphological characteristics. *Rev Med Chir Soc Med Nat Iasi*. 2019;109(2):343-346. <http://www.ncbi.nlm.nih.gov/pubmed/16607797>. Accessed August 8, 2018.
27. Abeler VM, Holm R, Nesland JM, Kjørstad KE. Small cell carcinoma of the cervix. A clinicopathologic study of 26 patients. *Cancer*. 1994;73(3):672-677. <http://www.ncbi.nlm.nih.gov/pubmed/8299089>. Accessed August 8, 2018.
28. Ganesan R, Hirschowitz L, Dawson P, et al. Neuroendocrine Carcinoma of the Cervix: Review of a Series of Cases and Correlation With Outcome. *Int J Surg Pathol*. 2016;24(6):490-496. doi:10.1177/1066896916643385
29. Wang SS, Sherman ME, Hildesheim A, Lacey J V, Devesa S. Cervical adenocarcinoma and squamous cell carcinoma incidence trends among white women and black women

- in the United States for 1976-2000. *Cancer*. 2004;100(5):1035-1044. doi:10.1002/cncr.20064
30. Castellsagué X, Díaz M, de Sanjosé S, et al. Worldwide human papillomavirus etiology of cervical adenocarcinoma and its cofactors: implications for screening and prevention. *J Natl Cancer Inst*. 2006;98(5):303-315. doi:10.1093/jnci/djj067
 31. Fujiwara K, Monk B, Devouassoux-Shisheboran M. Adenocarcinoma of the uterine cervix: why is it different? *Curr Oncol Rep*. 2014;16(12):416. doi:10.1007/s11912-014-0416-y
 32. de Villiers E-M, Fauquet C, Broker TR, Bernard H-U, zur Hausen H. Classification of papillomaviruses. *Virology*. 2004;324(1):17-27. doi:10.1016/j.virol.2004.03.033
 33. Papillomavirus Episteme. <https://pave.niaid.nih.gov/>.
 34. Bouvard V, Baan R, Straif K, et al. A review of human carcinogens--Part B: biological agents. *Lancet Oncol*. 2009;10(4):321-322.
<http://www.ncbi.nlm.nih.gov/pubmed/19350698>. Accessed August 8, 2018.
 35. Li N, Franceschi S, Howell-Jones R, Snijders PJF, Clifford GM. Human papillomavirus type distribution in 30,848 invasive cervical cancers worldwide: Variation by geographical region, histological type and year of publication. *Int J cancer*. 2011;128(4):927-935. doi:10.1002/ijc.25396
 36. Guan P, Clifford GM, Franceschi S. Human papillomavirus types in glandular lesions of the cervix: a meta-analysis of published studies. *Int J cancer*. 2013;132(1):248-250. doi:10.1002/ijc.27663
 37. Siriaunkgul S, Utaipat U, Suthipintawong C, Tungsinmunkong K, Triratanachat S, Khunamornpong S. HPV genotyping in adenocarcinoma of the uterine cervix in Thailand. *Int J Gynecol Obstet*. 2013;123(3):226-230. doi:10.1016/j.ijgo.2013.06.034
 38. Benevolo M, Donà MG, Ravenda PS, Chiocca S. Anal human papillomavirus infection: prevalence, diagnosis and treatment of related lesions. *Expert Rev Anti Infect Ther*. 2016;14(5):465-477. doi:10.1586/14787210.2016.1174065
 39. Lee LJ, Howitt B, Catalano P, et al. Prognostic importance of human papillomavirus (HPV) and p16 positivity in squamous cell carcinoma of the vulva treated with

- radiotherapy. *Gynecol Oncol.* 2016;142(2):293-298. doi:10.1016/j.ygyno.2016.05.019
40. Fakhry C, Gillison ML. Clinical Implications of Human Papillomavirus in Head and Neck Cancers. *J Clin Oncol.* 2006;24(17):2606-2611. doi:10.1200/JCO.2006.06.1291
 41. Muñoz N, Kjaer SK, Sigurdsson K, et al. Impact of Human Papillomavirus (HPV)-6/11/16/18 Vaccine on All HPV-Associated Genital Diseases in Young Women. *JNCI J Natl Cancer Inst.* 2010;102(5):325-339. doi:10.1093/jnci/djp534
 42. Olsson S-E, Kjaer SK, Sigurdsson K, et al. Evaluation of quadrivalent HPV 6/11/16/18 vaccine efficacy against cervical and anogenital disease in subjects with serological evidence of prior vaccine type HPV infection. *Hum Vaccin.* 2009;5(10):696-704. <http://www.ncbi.nlm.nih.gov/pubmed/19855170>. Accessed August 8, 2018.
 43. Bouallaga I, Massicard S, Yaniv M, Thierry F. An enhanceosome containing the Jun B/Fra-2 heterodimer and the HMG-I(Y) architectural protein controls HPV 18 transcription. *EMBO Rep.* 2000;1(5):422-427. doi:10.1093/embo-reports/kvd091
 44. Apt D, Watts RM, Suske G, Bernard HU. High Sp1/Sp3 ratios in epithelial cells during epithelial differentiation and cellular transformation correlate with the activation of the HPV-16 promoter. *Virology.* 1996;224(1):281-291. doi:10.1006/viro.1996.0530
 45. Chan WK, Chong T, Bernard HU, Klock G. Transcription of the transforming genes of the oncogenic human papillomavirus-16 is stimulated by tumor promoters through AP1 binding sites. *Nucleic Acids Res.* 1990;18(4):763-769. <http://www.ncbi.nlm.nih.gov/pubmed/2156229>. Accessed August 8, 2018.
 46. Chong T, Apt D, Gloss B, Isa M, Bernard HU. The enhancer of human papillomavirus type 16: binding sites for the ubiquitous transcription factors oct-1, NFA, TEF-2, NF1, and AP-1 participate in epithelial cell-specific transcription. *J Virol.* 1991;65(11):5933-5943. <http://www.ncbi.nlm.nih.gov/pubmed/1656080>. Accessed August 8, 2018.
 47. Thierry F, Heard JM, Dartmann K, Yaniv M. Characterization of a transcriptional promoter of human papillomavirus 18 and modulation of its expression by simian virus 40 and adenovirus early antigens. *J Virol.* 1987;61(1):134-142. <http://www.ncbi.nlm.nih.gov/pubmed/3023691>. Accessed August 8, 2018.

48. DiMaio D, Petti LM. The E5 proteins. *Virology*. 2013;445(1-2):99-114.
doi:10.1016/j.virol.2013.05.006
49. Doorbar J, Quint W, Banks L, et al. The biology and life-cycle of human papillomaviruses. *Vaccine*. 2012;30 Suppl 5:F55-70. doi:10.1016/j.vaccine.2012.06.083
50. Doorbar J. The papillomavirus life cycle. *J Clin Virol*. 2005;32 Suppl 1:S7-15.
doi:10.1016/j.jcv.2004.12.006
51. Longworth MS, Laimins LA. Pathogenesis of human papillomaviruses in differentiating epithelia. *Microbiol Mol Biol Rev*. 2004;68(2):362-372.
doi:10.1128/MMBR.68.2.362-372.2004
52. Wallace NA, Galloway DA. Novel Functions of the Human Papillomavirus E6 Oncoproteins. *Annu Rev Virol*. 2015;2(1):403-423. doi:10.1146/annurev-virology-100114-055021
53. Prendiville W, Davies P. *THE HEALTH PROFESSIONAL'S Editors-in-Chief*.
www.tandf.co.uk. Accessed August 8, 2018.
54. Mudhakir D, Harashima H. Learning from the viral journey: how to enter cells and how to overcome intracellular barriers to reach the nucleus. *AAPS J*. 2009;11(1):65-77.
doi:10.1208/s12248-009-9080-9
55. Johnson KM, Kines RC, Roberts JN, Lowy DR, Schiller JT, Day PM. Role of Heparan Sulfate in Attachment to and Infection of the Murine Female Genital Tract by Human Papillomavirus. *J Virol*. 2009;83(5):2067-2074. doi:10.1128/JVI.02190-08
56. Sapp M, Bienkowska-Haba M. Viral entry mechanisms: human papillomavirus and a long journey from extracellular matrix to the nucleus. *FEBS J*. 2009;276(24):7206-7216.
doi:10.1111/j.1742-4658.2009.07400.x
57. Day PM, Lowy DR, Schiller JT. Papillomaviruses infect cells via a clathrin-dependent pathway. *Virology*. 2003;307(1):1-11. <http://www.ncbi.nlm.nih.gov/pubmed/12667809>.
Accessed August 8, 2018.
58. Bienkowska-Haba M, Williams C, Kim SM, Garcea RL, Sapp M. Cyclophilins facilitate dissociation of the human papillomavirus type 16 capsid protein L1 from the L2/DNA

- complex following virus entry. *J Virol.* 2012;86(18):9875-9887. doi:10.1128/JVI.00980-12
59. Day PM, Thompson CD, Schowalter RM, Lowy DR, Schiller JT. Identification of a Role for the trans-Golgi Network in Human Papillomavirus 16 Pseudovirus Infection. *J Virol.* 2013;87(7):3862-3870. doi:10.1128/JVI.03222-12
 60. Day PM, Baker CC, Lowy DR, Schiller JT. Establishment of papillomavirus infection is enhanced by promyelocytic leukemia protein (PML) expression. *Proc Natl Acad Sci U S A.* 2004;101(39):14252-14257. doi:10.1073/pnas.0404229101
 61. Calton CM, Bronnimann MP, Manson AR, et al. Translocation of the papillomavirus L2/vDNA complex across the limiting membrane requires the onset of mitosis. Meyers C, ed. *PLoS Pathog.* 2017;13(5):e1006200. doi:10.1371/journal.ppat.1006200
 62. Raff AB, Woodham AW, Raff LM, et al. The evolving field of human papillomavirus receptor research: a review of binding and entry. *J Virol.* 2013;87(11):6062-6072. doi:10.1128/JVI.00330-13
 63. Doorbar J. Papillomavirus life cycle organization and biomarker selection. *Dis Markers.* 2007;23(4):297-313. <http://www.ncbi.nlm.nih.gov/pubmed/17627064>. Accessed August 9, 2018.
 64. McKinney CC, Hussmann KL, McBride AA. The Role of the DNA Damage Response throughout the Papillomavirus Life Cycle. *Viruses.* 2015;7(5):2450-2469. doi:10.3390/v7052450
 65. Woodman CBJ, Collins SI, Young LS. The natural history of cervical HPV infection: unresolved issues. *Nat Rev Cancer.* 2007;7(1):11-22. doi:10.1038/nrc2050
 66. Kalantari M, Blennow E, Hagmar B, Johansson B. Physical state of HPV16 and chromosomal mapping of the integrated form in cervical carcinomas. *Diagn Mol Pathol.* 2001;10(1):46-54. <http://www.ncbi.nlm.nih.gov/pubmed/11277395>. Accessed August 9, 2018.
 67. Kulmala S-MA, Syrjänen SM, Gyllensten UB, et al. Early integration of high copy HPV16 detectable in women with normal and low grade cervical cytology and histology. *J Clin Pathol.* 2006;59(5):513-517. doi:10.1136/jcp.2004.024570

68. Duensing S, Münger K. The human papillomavirus type 16 E6 and E7 oncoproteins independently induce numerical and structural chromosome instability. *Cancer Res.* 2002;62(23):7075-7082. <http://www.ncbi.nlm.nih.gov/pubmed/12460929>. Accessed August 9, 2018.
69. Munger K, Baldwin A, Edwards KM, et al. Mechanisms of Human Papillomavirus-Induced Oncogenesis. *J Virol.* 2004;78(21):11451-11460. doi:10.1128/JVI.78.21.11451-11460.2004
70. Vande Pol SB, Klingelhutz AJ. Papillomavirus E6 oncoproteins. *Virology.* 2013;445(1-2):115-137. doi:10.1016/j.virol.2013.04.026
71. Roman A, Munger K. The papillomavirus E7 proteins. *Virology.* 2013;445(1-2):138-168. doi:10.1016/j.virol.2013.04.013
72. Yeo-Teh NSL, Ito Y, Jha S. High-Risk Human Papillomaviral Oncogenes E6 and E7 Target Key Cellular Pathways to Achieve Oncogenesis. *Int J Mol Sci.* 2018;19(6):1706. doi:10.3390/ijms19061706
73. Massimi P, Shai A, Lambert P, Banks L. HPV E6 degradation of p53 and PDZ containing substrates in an E6AP null background. *Oncogene.* 2008;27(12):1800-1804. doi:10.1038/sj.onc.1210810
74. Jansma AL, Martinez-Yamout MA, Liao R, Sun P, Dyson HJ, Wright PE. The high-risk HPV16 E7 oncoprotein mediates interaction between the transcriptional coactivator CBP and the retinoblastoma protein pRb. *J Mol Biol.* 2014;426(24):4030-4048. doi:10.1016/j.jmb.2014.10.021
75. Cullen AP, Reid R, Campion M, Lörincz AT. Analysis of the physical state of different human papillomavirus DNAs in intraepithelial and invasive cervical neoplasm. *J Virol.* 1991;65(2):606-612. <http://www.ncbi.nlm.nih.gov/pubmed/1846186>. Accessed August 9, 2018.
76. Daniel B, Mukherjee G, Seshadri L, Vallikad E, Krishna S. Changes in the physical state and expression of human papillomavirus type 16 in the progression of cervical intraepithelial neoplasia lesions analysed by PCR. *J Gen Virol.* 1995;76 (Pt 10)(10):2589-

2593. doi:10.1099/0022-1317-76-10-2589

77. Gloss B, Bernard HU. The E6/E7 promoter of human papillomavirus type 16 is activated in the absence of E2 proteins by a sequence-aberrant Sp1 distal element. *J Virol.* 1990;64(11):5577-5584. <http://www.ncbi.nlm.nih.gov/pubmed/2170687>. Accessed August 9, 2018.
78. Dürst M, Glitz D, Schneider A, zur Hausen H. Human papillomavirus type 16 (HPV 16) gene expression and DNA replication in cervical neoplasia: analysis by in situ hybridization. *Virology.* 1992;189(1):132-140. <http://www.ncbi.nlm.nih.gov/pubmed/1318602>. Accessed August 9, 2018.
79. Cheng S, Schmidt-Grimminger DC, Murrant T, Broker TR, Chow LT. Differentiation-dependent up-regulation of the human papillomavirus E7 gene reactivates cellular DNA replication in suprabasal differentiated keratinocytes. *Genes Dev.* 1995;9(19):2335-2349. <http://www.ncbi.nlm.nih.gov/pubmed/7557386>. Accessed August 9, 2018.
80. McBride AA, Warburton A. The role of integration in oncogenic progression of HPV-associated cancers. Spindler KR, ed. *PLoS Pathog.* 2017;13(4):e1006211. doi:10.1371/journal.ppat.1006211
81. Chaiwongkot A, Vinokurova S, Pientong C, 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(9):2087-2094. doi:10.1002/ijc.27906
82. Dooley KE, Warburton A, McBride AA. Tandemly Integrated HPV16 Can Form a Brd4-Dependent Super-Enhancer-Like Element That Drives Transcription of Viral Oncogenes. *MBio.* 2016;7(5):e01446-16. doi:10.1128/mBio.01446-16
83. Swift F V, Bhat K, Youngusband HB, Hamada H. Characterization of a cell type-specific enhancer found in the human papilloma virus type 18 genome. *EMBO J.* 1987;6(5):1339-1344. <http://www.ncbi.nlm.nih.gov/pubmed/3038518>. Accessed August 9, 2018.
84. Chiang CM, Ustav M, Stenlund A, Ho TF, Broker TR, Chow LT. Viral E1 and E2 proteins support replication of homologous and heterologous papillomaviral origins.

Proc Natl Acad Sci U S A. 1992;89(13):5799-5803.

<http://www.ncbi.nlm.nih.gov/pubmed/1321423>. Accessed August 9, 2018.

85. Kajitani N, Satsuka A, Kawate A, Sakai H. Productive Lifecycle of Human Papillomaviruses that Depends Upon Squamous Epithelial Differentiation. *Front Microbiol.* 2012;3:152. doi:10.3389/fmicb.2012.00152
86. García-Vallvé S, Iglesias-Rozas JR, Alonso A, Bravo IG. Different papillomaviruses have different repertoires of transcription factor binding sites: convergence and divergence in the upstream regulatory region. *BMC Evol Biol.* 2006;6(1):20. doi:10.1186/1471-2148-6-20
87. Sichero L, Sobrinho JS, Villa LL. Identification of novel cellular transcription factors that regulate early promoters of human papillomavirus types 18 and 16. *J Infect Dis.* 2012;206(6):867-874. doi:10.1093/infdis/jis430
88. Grassmann K, Rapp B, Maschek H, Petry KU, Iftner T. Identification of a differentiation-inducible promoter in the E7 open reading frame of human papillomavirus type 16 (HPV-16) in raft cultures of a new cell line containing high copy numbers of episomal HPV-16 DNA. *J Virol.* 1996;70(4):2339-2349.
<http://www.ncbi.nlm.nih.gov/pubmed/8642661>. Accessed August 9, 2018.
89. Glahder JA, Hansen CN, Vinther J, Madsen BS, 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(Pt 12):3429-3441. doi:10.1099/vir.0.19250-0
90. Rosenstjerne MW, Vinther J, Hansen CN, Prydsøe M, Norrild B. Identification and characterization of a cluster of transcription start sites located in the E6 ORF of human papillomavirus type 16. *J Gen Virol.* 2003;84(Pt 11):2909-2920. doi:10.1099/vir.0.19332-0
91. O'Connor MJ, Tan SH, Tan CH, Bernard HU. YY1 represses human papillomavirus type 16 transcription by quenching AP-1 activity. *J Virol.* 1996;70(10):6529-6539.
<http://www.ncbi.nlm.nih.gov/pubmed/8794287>. Accessed August 9, 2018.
92. Milligan SG, Veerapraditsin T, Ahamet B, Mole S, Graham S V. Analysis of novel

- human papillomavirus type 16 late mRNAs in differentiated W12 cervical epithelial cells. *Virology*. 2007;360(1):172-181. doi:10.1016/j.virol.2006.10.012
93. Zheng Z-M, Baker CC. Papillomavirus genome structure, expression, and post-transcriptional regulation. *Front Biosci*. 2006;11:2286-2302.
<http://www.ncbi.nlm.nih.gov/pubmed/16720315>. Accessed August 9, 2018.
94. Mesplède T, Gagnon D, Bergeron-Labrecque F, et al. p53 degradation activity, expression, and subcellular localization of E6 proteins from 29 human papillomavirus genotypes. *J Virol*. 2012;86(1):94-107. doi:10.1128/JVI.00751-11
95. Ajiro M, Zheng Z-M. Oncogenes and RNA splicing of human tumor viruses. *Emerg Microbes Infect*. 2014;3(9):e63-e63. doi:10.1038/emi.2014.62
96. McBride AA, Romanczuk H, Howley PM. The papillomavirus E2 regulatory proteins. *J Biol Chem*. 1991;266(28):18411-18414. <http://www.ncbi.nlm.nih.gov/pubmed/1655748>. Accessed August 9, 2018.
97. de Prat-Gay G, Gaston K, Cicero DO. The papillomavirus E2 DNA binding domain. *Front Biosci*. 2008;13:6006-6021. <http://www.ncbi.nlm.nih.gov/pubmed/18508638>. Accessed August 9, 2018.
98. Thierry F. Transcriptional regulation of the papillomavirus oncogenes by cellular and viral transcription factors in cervical carcinoma. *Virology*. 2009;384(2):375-379. doi:10.1016/j.virol.2008.11.014
99. Steger G, Corbach S. Dose-dependent regulation of the early promoter of human papillomavirus type 18 by the viral E2 protein. *J Virol*. 1997;71(1):50-58. <http://www.ncbi.nlm.nih.gov/pubmed/8985322>. Accessed August 9, 2018.
100. Rank NM, Lambert PF. Bovine papillomavirus type 1 E2 transcriptional regulators directly bind two cellular transcription factors, TFIID and TFIIB. *J Virol*. 1995;69(10):6323-6334. <http://www.ncbi.nlm.nih.gov/pubmed/7666533>. Accessed August 9, 2018.
101. Koontongkaew S. The tumor microenvironment contribution to development, growth, invasion and metastasis of head and neck squamous cell carcinomas. *J Cancer*.

2013;4(1):66-83. doi:10.7150/jca.5112

102. Fromental C, Kanno M, Nomiya H, Chambon P. Cooperativity and hierarchical levels of functional organization in the SV40 enhancer. *Cell*. 1988;54(7):943-953. <http://www.ncbi.nlm.nih.gov/pubmed/2843294>. Accessed August 9, 2018.
103. Ondek B, Gloss L, Herr W. The SV40 enhancer contains two distinct levels of organization. *Nature*. 1988;333(6168):40-45. doi:10.1038/333040a0
104. Butz K, Hoppe-Seyler F. Transcriptional control of human papillomavirus (HPV) oncogene expression: composition of the HPV type 18 upstream regulatory region. *J Virol*. 1993;67(11):6476-6486. <http://www.ncbi.nlm.nih.gov/pubmed/8411351>. Accessed August 9, 2018.
105. Angel P, Karin M. The role of Jun, Fos and the AP-1 complex in cell-proliferation and transformation. *Biochim Biophys Acta*. 1991;1072(2-3):129-157. <http://www.ncbi.nlm.nih.gov/pubmed/1751545>. Accessed August 9, 2018.
106. Mehic D, Bakiri L, Ghannadan M, Wagner EF, Tschachler E. Fos and jun proteins are specifically expressed during differentiation of human keratinocytes. *J Invest Dermatol*. 2005;124(1):212-220. doi:10.1111/j.0022-202X.2004.23558.x
107. Zhao W, Chow LT, Broker TR. A distal element in the HPV-11 upstream regulatory region contributes to promoter repression in basal keratinocytes in squamous epithelium. *Virology*. 1999;253(2):219-229. doi:10.1006/viro.1998.9478
108. Gordon S, Akopyan G, Garban H, Bonavida B. Transcription factor YY1: structure, function, and therapeutic implications in cancer biology. *Oncogene*. 2006;25(8):1125-1142. doi:10.1038/sj.onc.1209080
109. Baritaki S, Sifakis S, Huerta-Yepez S, et al. Overexpression of VEGF and TGF-beta1 mRNA in Pap smears correlates with progression of cervical intraepithelial neoplasia to cancer: implication of YY1 in cervical tumorigenesis and HPV infection. *Int J Oncol*. 2007;31(1):69-79. <http://www.ncbi.nlm.nih.gov/pubmed/17549406>. Accessed August 9, 2018.
110. Montalvo EA, Cottam M, Hill S, Wang YJ. YY1 binds to and regulates cis-acting

- negative elements in the Epstein-Barr virus BZLF1 promoter. *J Virol.* 1995;69(7):4158-4165. <http://www.ncbi.nlm.nih.gov/pubmed/7769675>. Accessed August 9, 2018.
111. Mills LK, Shi Y, Millette RL. YY1 is the cellular factor shown previously to bind to regulatory regions of several leaky-late (beta gamma, gamma 1) genes of herpes simplex virus type 1. *J Virol.* 1994;68(2):1234-1238. <http://www.ncbi.nlm.nih.gov/pubmed/8289358>. Accessed August 9, 2018.
112. Bauknecht T, Angel P, Royer HD, zur Hausen H. Identification of a negative regulatory domain in the human papillomavirus type 18 promoter: interaction with the transcriptional repressor YY1. *EMBO J.* 1992;11(12):4607-4617. <http://www.ncbi.nlm.nih.gov/pubmed/1330541>. Accessed August 9, 2018.
113. Ralph WM, Liu K, Auborn KJ. CCAAT/enhancer-binding protein beta represses human papillomavirus 11 upstream regulatory region expression through a promoter-proximal YY1-binding site. *J Gen Virol.* 2006;87(Pt 1):51-59. doi:10.1099/vir.0.81207-0
114. Pajunk HS, May C, Pfister H, Fuchs PG. Regulatory interactions of transcription factor YY1 with control sequences of the E6 promoter of human papillomavirus type 8. *J Gen Virol.* 1997;78 (Pt 12)(12):3287-3295. doi:10.1099/0022-1317-78-12-3287
115. Lace MJ, Yamakawa Y, Ushikai M, Anson JR, Haugen TH, Turek LP. Cellular factor YY1 downregulates the human papillomavirus 16 E6/E7 promoter, P97, in vivo and in vitro from a negative element overlapping the transcription-initiation site. *J Gen Virol.* 2009;90(10):2402-2412. doi:10.1099/vir.0.012708-0
116. Dürst M, Gallahan D, Jay G, Rhim JS. Glucocorticoid-enhanced neoplastic transformation of human keratinocytes by human papillomavirus type 16 and an activated ras oncogene. *Virology.* 1989;173(2):767-771. <http://www.ncbi.nlm.nih.gov/pubmed/2556855>. Accessed August 10, 2018.
117. Mittal R, Pater A, Pater MM. Multiple human papillomavirus type 16 glucocorticoid response elements functional for transformation, transient expression, and DNA-protein interactions. *J Virol.* 1993;67(9):5656-5659. <http://www.ncbi.nlm.nih.gov/pubmed/8394465>. Accessed August 10, 2018.

118. Bromberg-White JL, Meyers C. The upstream regulatory region of human papillomavirus type 31 is insensitive to glucocorticoid induction. *J Virol.* 2002;76(19):9702-9715. <http://www.ncbi.nlm.nih.gov/pubmed/12208949>. Accessed August 10, 2018.
119. Khan A, Fornes O, Stigliani A, et al. JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. *Nucleic Acids Res.* 2018;46(D1):D260-D266. doi:10.1093/nar/gkx1126
120. Stevanovic M, Zuffardi O, Collignon J, Lovell-Badge R, Goodfellow P. The cDNA sequence and chromosomal location of the human SOX2 gene. *Mamm Genome.* 1994;5(10):640-642. <http://www.ncbi.nlm.nih.gov/pubmed/7849401>. Accessed August 10, 2018.
121. Weina K, Utikal J. SOX2 and cancer: current research and its implications in the clinic. *Clin Transl Med.* 2014;3(1):19. doi:10.1186/2001-1326-3-19
122. Xu N, Papagiannakopoulos T, Pan G, Thomson JA, Kosik KS. MicroRNA-145 Regulates OCT4, SOX2, and KLF4 and Represses Pluripotency in Human Embryonic Stem Cells. *Cell.* 2009;137(4):647-658. doi:10.1016/j.cell.2009.02.038
123. Tsuruzoe S, Ishihara K, Uchimura Y, et al. Inhibition of DNA binding of Sox2 by the SUMO conjugation. *Biochem Biophys Res Commun.* 2006;351(4):920-926. doi:10.1016/j.bbrc.2006.10.130
124. Baltus GA, Kowalski MP, Zhai H, et al. Acetylation of Sox2 Induces its Nuclear Export in Embryonic Stem Cells. *Stem Cells.* 2009;27(9):2175-2184. doi:10.1002/stem.168
125. Jeong C-H, Cho Y-Y, Kim M-O, et al. Phosphorylation of Sox2 cooperates in reprogramming to pluripotent stem cells. *Stem Cells.* 2010;28(12):2141-2150. doi:10.1002/stem.540
126. Zhao H, Zhang Y, Dai H, Zhang Y, Shen Y. CARM1 mediates modulation of Sox2. Jin D-Y, ed. *PLoS One.* 2011;6(10):e27026. doi:10.1371/journal.pone.0027026
127. Wu Y, Guo Z, Wu H, et al. SUMOylation Represses Nanog Expression via Modulating Transcription Factors Oct4 and Sox2. Akagi T, ed. *PLoS One.* 2012;7(6):e39606.

doi:10.1371/journal.pone.0039606

128. Liu K, Lin B, Zhao M, et al. The multiple roles for Sox2 in stem cell maintenance and tumorigenesis. *Cell Signal*. 2013;25(5):1264-1271. doi:10.1016/j.cellsig.2013.02.013
129. Rodda DJ, Chew J-L, Lim L-H, et al. Transcriptional regulation of nanog by OCT4 and SOX2. *J Biol Chem*. 2005;280(26):24731-24737. doi:10.1074/jbc.M502573200
130. Peng C, Li N, Ng Y-K, et al. A Unilateral Negative Feedback Loop Between miR-200 microRNAs and Sox2/E2F3 Controls Neural Progenitor Cell-Cycle Exit and Differentiation. *J Neurosci*. 2012;32(38):13292-13308. doi:10.1523/JNEUROSCI.2124-12.2012
131. Tian T, Zhang Y, Wang S, Zhou J, Xu S. Sox2 enhances the tumorigenicity and chemoresistance of cancer stem-like cells derived from gastric cancer. *J Biomed Res*. 2012;26(5):336-345. doi:10.7555/JBR.26.20120045
132. Chang X, Zhang J, Huang C, et al. Sex-determining region Y-related high mobility group box (SOX)-2 is overexpressed in cervical squamous cell carcinoma and contributes cervical cancer cell migration and invasion in vitro. *Tumour Biol*. 2015;36(10):7725-7733. doi:10.1007/s13277-015-3450-y
133. Kim BW, Cho H, Choi CH, et al. Clinical significance of OCT4 and SOX2 protein expression in cervical cancer. *BMC Cancer*. 2015;15(1):1015. doi:10.1186/s12885-015-2015-1
134. Wilbertz T, Wagner P, Petersen K, et al. SOX2 gene amplification and protein overexpression are associated with better outcome in squamous cell lung cancer. *Mod Pathol*. 2011;24(7):944-953. doi:10.1038/modpathol.2011.49
135. Wang Q, He W, Lu C, et al. Oct3/4 and Sox2 are significantly associated with an unfavorable clinical outcome in human esophageal squamous cell carcinoma. *Anticancer Res*. 2009;29(4):1233-1241. <http://www.ncbi.nlm.nih.gov/pubmed/19414369>. Accessed August 10, 2018.
136. Johansson H, Simonsson S. Core transcription factors, Oct4, Sox2 and Nanog, individually form complexes with nucleophosmin (Npm1) to control embryonic stem


- (ES) cell fate determination. *Aging (Albany NY)*. 2010;2(11):815-822.
doi:10.18632/aging.100222
137. Ji J, Zheng P-S. Expression of Sox2 in human cervical carcinogenesis. *Hum Pathol*. 2010;41(10):1438-1447. doi:10.1016/j.humpath.2009.11.021
138. Liu X-F, Yang W-T, Xu R, Liu J-T, Zheng P-S. Cervical cancer cells with positive Sox2 expression exhibit the properties of cancer stem cells. Asselin E, ed. *PLoS One*. 2014;9(1):e87092. doi:10.1371/journal.pone.0087092
139. Herreros-Villanueva M, Zhang J-S, Koenig A, et al. SOX2 promotes dedifferentiation and imparts stem cell-like features to pancreatic cancer cells. *Oncogenesis*. 2013;2(8):e61. doi:10.1038/oncsis.2013.23
140. López-Saavedra A, González-Maya L, Ponce-de-León S, García-Carrancá A, Mohar A, Lizano M. Functional implication of sequence variation in the long control region and E2 gene among human papillomavirus type 18 variants. *Arch Virol*. 2009;154(5):747-754. doi:10.1007/s00705-009-0362-4
141. Fontaine V, van der Meijden E, de Graaf J, ter Schegget J, Struyk L. A functional NF-kappaB binding site in the human papillomavirus type 16 long control region. *Virology*. 2000;272(1):40-49. doi:10.1006/viro.2000.0363
142. Paris C, Pentland I, Groves I, et al. CCCTC-binding factor recruitment to the early region of the human papillomavirus 18 genome regulates viral oncogene expression. Imperiale MJ, ed. *J Virol*. 2015;89(9):4770-4785. doi:10.1128/JVI.00097-15
143. Yong W, Peng D, Wang L, Dong Z, He B. Screening of HaCaT clones for CCL20 gene knockout and preliminary exploration of gene-targeting vector transfection approaches in this cell line. *Med Sci Monit Basic Res*. 2015;21:21-28. doi:10.12659/MSMBR.893143
144. Sarkar A, Hochedlinger K. The Sox Family of Transcription Factors: Versatile Regulators of Stem and Progenitor Cell Fate. *Cell Stem Cell*. 2013;12(1):15-30. doi:10.1016/j.stem.2012.12.007
145. Chew J-L, Loh Y-H, Zhang W, et al. Reciprocal transcriptional regulation of Pou5f1 and Sox2 via the Oct4/Sox2 complex in embryonic stem cells. *Mol Cell Biol*. 2005;25(14):6031-

6046. doi:10.1128/MCB.25.14.6031-6046.2005
146. Chaturvedi AK, Engels EA, Pfeiffer RM, et al. Human Papillomavirus and Rising Oropharyngeal Cancer Incidence in the United States. *J Clin Oncol*. 2011;29(32):4294-4301. doi:10.1200/JCO.2011.36.4596
 147. Zhou C, Yang X, Sun Y, Yu H, Zhang Y, Jin Y. Comprehensive profiling reveals mechanisms of SOX2-mediated cell fate specification in human ESCs and NPCs. *Cell Res*. 2016;26(2):171-189. doi:10.1038/cr.2016.15
 148. Stolzenburg S, Rots MG, Beltran AS, et al. Targeted silencing of the oncogenic transcription factor SOX2 in breast cancer. *Nucleic Acids Res*. 2012;40(14):6725-6740. doi:10.1093/nar/gks360
 149. Jia X, Li X, Xu Y, et al. SOX2 promotes tumorigenesis and increases the anti-apoptotic property of human prostate cancer cell. *J Mol Cell Biol*. 2011;3(4):230-238. doi:10.1093/jmcb/mjr002
 150. Hütz K, Mejías-Luque R, Farsakova K, et al. The stem cell factor SOX2 regulates the tumorigenic potential in human gastric cancer cells. *Carcinogenesis*. 2014;35(4):942-950. doi:10.1093/carcin/bgt410
 151. Chen S, Li X, Lu D, et al. SOX2 regulates apoptosis through MAP4K4-survivin signaling pathway in human lung cancer cells. *Carcinogenesis*. 2014;35(3):613-623. doi:10.1093/carcin/bgt371
 152. Bass AJ, Watanabe H, Mermel CH, et al. SOX2 is an amplified lineage-survival oncogene in lung and esophageal squamous cell carcinomas. *Nat Genet*. 2009;41(11):1238-1242. doi:10.1038/ng.465
 153. Ortiz-Sánchez E, Santiago-López L, Cruz-Domínguez VB, et al. Characterization of cervical cancer stem cell-like cells: phenotyping, stemness, and human papilloma virus co-receptor expression. *Oncotarget*. 2016;7(22):31943-31954. doi:10.18632/oncotarget.8218
 154. Lopez J, Ruíz G, Organista-Nava J, Gariglio P, García-Carrancá A. Human Papillomavirus Infections and Cancer Stem Cells of Tumors from the Uterine Cervix. *Open Virol J*. 2012;6(1):232-240. doi:10.2174/1874357901206010232

155. Boumahdi S, Driessens G, Lapouge G, et al. SOX2 controls tumour initiation and cancer stem-cell functions in squamous-cell carcinoma. *Nature*. 2014;511(7508):246-250. doi:10.1038/nature13305
156. Feng D, Peng C, Li C, et al. Identification and characterization of cancer stem-like cells from primary carcinoma of the cervix uteri. *Oncol Rep*. 2009;22(5):1129-1134. <http://www.ncbi.nlm.nih.gov/pubmed/19787230>. Accessed August 15, 2018.
157. de Boer MA, Jordanova ES, Kenter GG, et al. High Human Papillomavirus Oncogene mRNA Expression and Not Viral DNA Load Is Associated with Poor Prognosis in Cervical Cancer Patients. *Clin Cancer Res*. 2007;13(1):132-138. doi:10.1158/1078-0432.CCR-06-1568
158. Cox JL, Mallanna SK, Luo X, Rizzino A. Sox2 uses multiple domains to associate with proteins present in Sox2-protein complexes. Qi RZ, ed. *PLoS One*. 2010;5(11):e15486. doi:10.1371/journal.pone.0015486
159. Liu Y-R, Laghari ZA, Novoa CA, et al. Sox2 acts as a transcriptional repressor in neural stem cells. *BMC Neurosci*. 2014;15(1):95. doi:10.1186/1471-2202-15-95

Article

SOX2 as a New Regulator of HPV16 Transcription

Imelda Martínez-Ramírez ¹, Víctor del-Castillo-Falconi ¹, Irma B. Mitre-Aguilar ²,
Alfredo Amador-Molina ¹, Adela Carrillo-García ¹, Elizabeth Langley ¹,
Alejandro Zentella-Dehesa ^{2,3}, Ernesto Soto-Reyes ¹ , Alejandro García-Carrancá ¹,
Luis A. Herrera ^{1,3} and Marcela Lizano ^{1,3,*}

- ¹ Unidad de Investigación Biomédica en Cáncer, Instituto Nacional de Cancerología (INCan)/Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México (UNAM), Ciudad de México 14080, Mexico; immara02@yahoo.com.mx (I.M.-R.); vdcfalconi@gmail.com (V.d.-C.-F.); aamadorm@incan.edu.mx (A.A.-M.); adcarrillo2004@yahoo.com.mx (A.C.-G.); langleyemx@gmail.com (E.L.); epigenetics.cancer@gmail.com (E.S.-R.); carranca@biomedicas.unam.mx (A.G.-C.); herreram@biomedicas.unam.mx (L.A.H.)
- ² Unidad de Bioquímica, Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán (INCMNSZ)/Unidad Periférica del Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México (UNAM), Ciudad de México 14080, Mexico; irma.mitre@gmail.com (I.B.M.-A.); azentell@biomedicas.unam.mx (A.Z.-D.)
- ³ Departamento de Medicina Genómica y Toxicología Ambiental, Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México (UNAM), Ciudad de México 04510, Mexico
- * Correspondence: lizanosoberon@gmail.com; Tel.: +52-555-573-4662

Academic Editor: Joanna Parish

Received: 15 May 2017; Accepted: 27 June 2017; Published: 5 July 2017

Abstract: Persistent infections with high-risk human papillomavirus (HPV) constitute the main risk factor for cervical cancer development. HPV16 is the most frequent type associated to squamous cell carcinomas (SCC), followed by HPV18. The long control region (LCR) in the HPV genome contains the replication origin and sequences recognized by cellular transcription factors (TFs) controlling viral transcription. Altered expression of *E6* and *E7* viral oncogenes, modulated by the LCR, causes modifications in cellular pathways such as proliferation, leading to malignant transformation. The aim of this study was to identify specific TFs that could contribute to the modulation of high-risk HPV transcriptional activity, related to the cellular histological origin. We identified sex determining region Y (SRY)-box 2 (SOX2) response elements present in HPV16-LCR. SOX2 binding to the LCR was demonstrated by *in vivo* and *in vitro* assays. The overexpression of this TF repressed HPV16-LCR transcriptional activity, as shown through reporter plasmid assays and by the down-regulation of endogenous HPV oncogenes. Site-directed mutagenesis revealed that three putative SOX2 binding sites are involved in the repression of the LCR activity. We propose that SOX2 acts as a transcriptional repressor of HPV16-LCR, decreasing the expression of *E6* and *E7* oncogenes in a SCC context.

Keywords: human papillomavirus; HPV; LCR; SOX2; transcriptional regulation

1. Introduction

Cervical cancer (CC) is the third most common cause of cancer death in women from Latin America [1]. This type of cancer commonly presents as squamous cell carcinoma (SCC), corresponding to 80% of all cervical cancers, followed by adenocarcinoma (AD), with a glandular origin, which accounts for approximately 15% of the cases [2]. Persistent infection with high-risk Human papillomavirus (HPV) is the main risk factor for development of cervical cancer [3,4]. An etiological role for HPV in other types of cancer has also been reported, such as non-cervical anogenital cancer [5,6], head and neck squamous carcinoma [7] and oropharyngeal cancers [8]. More than 170 viral types

have been described, with 10% of genomic differences among them [9]. About 40 HPV types infect the anogenital mucosa, and are characterized as high-risk, associated with malignant transformation, or low-risk, which cause benign warts and premalignant lesions [10,11]. HPV16 and 18 are the most frequent viral types associated with SCC (nearly 60% and 16%, respectively) [12]. However, in cervical adenocarcinoma, the proportion of HPV18 positive cases is comparable to that of HPV16, or even increases to double [12–15].

The HPV double-stranded DNA genome contains a long control region (LCR) or upstream regulatory region (URR) located between the last codon of the *L1* gene and the beginning of *E6*. This sequence contains response elements that are recognized by regulatory proteins such as transcriptional activators, repressors, terminators and initiators of replication [16–19], as well as an early promoter in the 3' terminus (P97 for HPV16 and P105 for HPV18) [20] which mainly regulates transcription of the early (E) viral region, encoding six open reading frames. Viral proteins E1 and E2 are necessary for viral genome replication; E4 has a role in virion release; E5 activates the epidermal growth factor receptor (EGFR) [21]; and E6 and E7 are the most important viral oncoproteins [22]. Altered expression of E6 and E7 oncoproteins may cause modifications in critical cellular pathways involved in cell proliferation and apoptosis, leading to malignant transformation [23].

The LCR contains an “epithelial-specific” enhancer, encompassing several regulatory elements that hyper-activate viral transcription [24]. Transcription factors (TFs) that activate the LCR have been thoroughly characterized, including AP-1, NF1, Oct1, SP1 and YY1, which can activate and/or repress viral transcription [25]. Likewise, some TFs such as EPOC-1/Skn-1a, C/EBP- α , - β , c-Myb, NFATx, Pax5, and WT1, have been reported to be involved in differentiation dependent expression of early and late viral genes, [26]. Moreover, LCRs of different HPV types contain distinctive potential TF binding sites, which could be associated to viral expression patterns associated to each viral type, in response to host-cell differentiation [27]. For example, it has been shown that GATA3 and TFAP2B specifically affect HPV18 transcription, while no such binding sites are present in the HPV16-LCR [28].

The identification of additional TFs regulating the LCR could widen the knowledge of HPV type-specific biological behavior during the development of cervical cancer. The aim of this study was to identify specific TFs that could contribute to the modulation of high-risk HPV transcriptional activities, related to the cellular histological origin. Through *in silico* analysis of the HPV16- and HPV18-LCRs we identified a variety of response elements exclusive for each LCR. Interestingly, three putative sex determining region Y (SRY)-box 2 (SOX2) response elements were present in HPV16-LCR. SOX2 binding to these sites was demonstrated by *in vivo* and *in vitro* assays. The overexpression of this TF repressed HPV16-LCR transcriptional activity, as shown through reporter gene plasmid assays and by the down-regulation of the expression of endogenous HPV oncogenes. Site-directed mutagenesis revealed that the three putative SOX2 binding sites are involved in the repression of the LCR activity. We propose that SOX2 acts as a transcriptional repressor of HPV16-LCR, modulating *E6* and *E7* oncogene expression in a squamous cell context.

2. Materials and Methods

2.1. Cell Culture

CaSki HPV16 positive SCC cell line (American Type Culture Collection, ATCC # CRL-1550) was grown in RPMI-1640 medium (Life Technologies, Gibco, Grand Island, NY, USA), with 10% fetal bovine serum (FBS) (Life Technologies, Gibco). HeLa HPV18 positive adenocarcinoma cell line (ATCC # CCL-2) and the non-tumorigenic human keratinocyte HaCaT cell line (kindly donated by Dr. P. Gariglio, Centro de Investigación y de estudios Avanzados del Instituto Politécnico Nacional, CINVESTAV, Mexico), were grown in Dulbecco's Modified Eagle Medium: Nutrient Mixture F-12 (DMEM-F12) medium (Life Technologies, Gibco), with 10% FBS, at 37 °C in a moist atmosphere of 95% air and 5% CO₂. CaSki, HeLa and HaCaT cell lines were authenticated by short tandem repeat

profiling in August 2016, by the Instituto Nacional de Medicina Genómica (INMEGEN), México and exhibited no evidence of cross contamination with known ATCC cell lines.

2.2. Plasmid Constructs.
The HPV16-LCR was PCR-amplified from HPV16 positive cervical cancer biopsies. Primers used for HPV16-LCR amplification were as follows: Forward (Fwd) SmaI 5'-GGGCCC GGCAGACCTAGATCAGTTTCCT-3', and Reverse (Rev) HindIII 5'-GGG AAGCTTGCAGTTCTCTTTTGGTGCAT-3', these amplify the region comprising nucleotides 7007–7905 to 1–102 of the HPV16 genome. PCR products were inserted into a SmaI/HindIII-treated pGL2-Basic vector (Promega corporation, Madison, Wisconsin, USA), which harbors luciferase gene as reporter, generating the pGL2-LCR16 vector. The pGL2-LCR18 was previously built from PCR-amplified LCR products obtained from HPV18 positive cervical cancer biopsies [29]; this plasmid contains nucleotides 7201–7857 to 1–124 of HPV genome. Plasmid constructs were verified by DNA sequencing. Plasmids pcDNA-SOX2 (Addgene, Cambridge, MA, USA) and pcDNA3.1 Zeo(+) (Life Technologies, Invitrogen, Carlsbad, CA, USA) were used for SOX2 overexpression and as a negative control, respectively.

The lentiviral vectors Pax2 and PMD2.G (Addgene) were used to produce recombinant lentiviruses containing SOX2 and/or green fluorescent protein (GFP) genes. The packaging cell line 293T was co-transfected with lentiviral expression plasmids SOX2 and/or GFP, Pax2, and PMD2.G, using lipofectamine plus (Invitrogen). Two days after transfection, cell supernatants containing the recombinant lentiviruses, were collected and stored at -80°C until use.

2.2. Infection with Recombinant Lentiviruses

CaSki cells were seeded in 24-well plates at a cell density of 3×10^4 cell/well. After 24 h, based on standard curves, cells were transduced with 50 microliters of lentiviral supernatant added to each well and incubation was continued for 72 h. Lentiviral vector GFP was used as transduction control.

2.3. Transfection Assays and Transcriptional Activity

For transient transfection assays, 30,000 CaSki, HeLa or HaCaT cells were seeded in 24-well plates with complete medium; after 24 h the cells were transfected using lipofectamine 2000 (Invitrogen) with 200 ng of luciferase reporter vector and 1 ng of Renilla plasmid (pRL-CMV) for normalization. As a positive control pGL3-Control, containing simian virus 40 (SV40) promoter was used. Cell extracts were prepared 48 h post-transfection and assayed for luciferase and renilla activities using the Dual Luciferase kit (Promega corporation, Madison, WI, USA). Relative luciferase units were measured in a Monolight™ 3010 luminometer (Pharmingen, BD Biosciences, San José CA, USA). Experiments were performed in triplicate.

2.4. In Silico Analysis

Putative TFs binding sites within the LCR sequences of HPV16 and HPV18 were analyzed using the JASPAR CORE database [30] and alignment with consensus sequences. The complete LCR sequences of HPV16 and 18 were captured and compared with the existing data.

2.5. Electrophoretic Mobility Shift Assay

The Electrophoretic Mobility Shift Assay (EMSA) was performed in a final volume of 20 μL containing nuclear protein extract from HeLa cells (15 μg), in the incubation buffer (30 mM Tris pH 7.5, 300 mM KCl, 15 mM MgCl₂, and 30% glycerol), 1M dithiothreitol (DTT), 10 $\mu\text{g}/\mu\text{L}$ BSA, 1 $\mu\text{g}/\mu\text{L}$ poly (dI-dC) (Amersham Biosciences, Buckinghamshire, UK) and 1 μL of double-stranded γ -ATP (³²P) end-labelled DNA oligonucleotides. Primers used to identify SOX2 binding sites were: SOX2 consensus binding sequence (CS): 5' TAGTCATACAATGTTTCATT 3'; putative SOX2 binding site 1 (S1): 5'-AACCATTCCATTGTTTTTAA-3'; putative SOX2 binding site 2 (S2): 5'-TAGTCATACATTGTTTCATT-3'; putative SOX2 binding site 3 (S3): 5'-AACTGCACATGGGTGTGTGCAAACCGT-3'; and non-specific oligonucleotide (NS): 5' AACTGCACATGGGTGTGTGCAAACCGT 3'. The unlabeled double-stranded

competitor was also included prior to the addition of labeled oligonucleotides, when specified. For the super-shift analysis, rabbit anti-SOX2 antibody (Santa Cruz Biotechnology, Dallas, TX, USA, sc-20088) and rabbit anti-CCCTC-binding factor (CTCF) (sc-5916X) were used. Binding reactions were incubated for 20 min at room temperature. Protein-DNA complexes were separated from free oligonucleotide on a 6% non-denaturing polyacrylamide/Tris borate-ethylenediaminetetraacetic acid (EDTA) gel, and were resolved at 130 V for 2 h. The gels were visualized in a Typhoon 9400 (GE Healthcare, Uppsala, Sweden) and analyzed using the Image Quant software (Bio-sciences AB GE Healthcare Life Sciences; Uppsala, Sweden).

2.6. Chromatin Immunoprecipitation Assay

The chromatin immunoprecipitation (ChIP) assay with CaSki cells was carried out using the OneDay ChIP kit (Diagenode, Denville, NJ, USA), following the manufacturer's instructions. 1×10^6 cells were cross-linked with 1% formaldehyde and quenched with 0.125 M of glycine. Cells were lysed with a buffer containing 1% sodium dodecyl sulfate (SDS) and sonicated with Thermo Fisher Scientific sonic disruptor (Thermo Fisher scientific, Waltham, MA, USA, Model 505). The fragmented chromatin (mainly 300 to 500 pb) was immunoprecipitated with: rabbit anti-SOX2 (Abcam, Cambridge, MA, USA, ab59776); rabbit anti-p65 (Santa Cruz Biotechnology, sc-109), as LCR-binding control [31]; and rabbit anti-CTCF (Millipore, Billerica, MA, USA, 07-729), as non-binding control [32]. As a background control, a chromatin fraction was incubated with normal rabbit IgG antibody (Santa Cruz Biotechnology, sc-2027).

Quantitative PCR (qPCR) was performed in triplicate in a fast optical 96-well qPCR reaction plate (Applied Biosystems, Carlsbad, CA, USA). Reactions were performed with primers amplifying a LCR region using the following primers: Fwd 5'-GTAAACTACACATGGGTGTGTG-3' and Rev 5'-GTCTGCTTTTATACTACACCGG-3'; with Thermo Maxima SYBR Green/ROX1 PCR Master Mix (Thermo Scientific, Applied Biosystems, Woolston, Warrington, UK, K0222), in StepOnePlus Real-Time PCR System (Applied Biosystems, 4376600). As a positive control of SOX2 interaction, a fragment of the Lefty promoter was PCR amplified with the following oligonucleotides: Fwd 5'-GTCAGGGTAGGGATCACTC-3'; and Rev 5'-TACAATGGTCTGAAGTGTTC-3'.

2.7. Multi-Site-Directed Mutagenesis Assays

Plasmid pGL2-LCR16 was mutated in the three putative binding SOX2-sites through site-directed mutagenesis (QuikChange Multi-Site-Directed Mutagenesis kit, Agilent Technologies, USA), with primers shown in Table S1. Four different mutants were obtained: LCR-Site 1 mutant (M1), LCR-Site 2 mutant (M2), LCR-Site 3 mutant (M3) and the LCR triple mutant (M1,2,3) which is mutated in all SOX2 putative sites. LCR mutants were confirmed by direct sequencing. Site 1 was mutated in two bases, while sites 2 and 3 were mutated in three bases each (Figure S1).

2.8. Western Blots

Whole cell lysates were collected using RIPA buffer (150 mM sodium chloride, 0.1% Triton-X 100, 0.5% sodium dodecyl sulphate and 50 mM Tris pH8 in ddH₂O) and prepared in the presence of complete mini EDTA-free protease inhibitor tablets (Roche, Mannheim, Germany). Protein concentrations were determined following the instructions of bicinchoninic acid solution and Copper(II) sulfate solution (Sigma-Aldrich, St. Louis, MO, USA). 30 µg of protein loaded for each lane were resolved in 12% sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to an Immobilon-P polyvinylidene difluoride (PVDF) membrane (Amersham, UK) for western blot. The membranes were blocked with 10% skim milk in TBS-T (10 mM Tris pH 7.5, NaCl 100 mM and 0.1% Tween-20) and incubated with primary antibodies (in 5% skim milk) overnight at 4 °C. Protein levels of SOX2 and Lefty were analyzed with rabbit anti-SOX2 (Abcam, ab59776) (1:2000 dilution) and rabbit anti-Lefty antibodies (Santa Cruz Biotechnology, sc-48836) (1:1000 dilution). Mouse anti-α-actinin, anti-β-actin and anti-GAPDH (1:1000 dilution) (Santa Cruz Biotechnology,

sc-17829, sc-47778, sc-32233, respectively), were used as convenient for protein load control. Secondary antibodies anti-rabbit or anti-mouse (1:10,000 dilution) (Santa Cruz Biotechnology, sc-2030; sc-2005, respectively), conjugated to HRP (horseradish peroxidase-conjugated), were used. Proteins were detected using Immobilon Western chemiluminescent HRP Substrate kit (Merck Millipore, Billerica, MA, USA, WBKLSO500).

2.9. RNA Isolation and cDNA Synthesis

Cells were resuspended in 800 μ L of Trizol Reagent (Life Technologies, Invitrogen). RNA was isolated according to the manufacturer's protocol. RNA pellets were dissolved in 20 μ L DNase-free H₂O. RNA was quantified and treated with amplification grade DNase I (Invitrogen). cDNA was synthesized using the SuperScript II RNaseH-Reverse Transcriptase (Life Technologies, Invitrogen).

2.10. Quantitative PCR

cDNA generated from transfected cells was submitted to qPCR using SYBR green PCR master mix (Applied Biosystems, 4309155), according to the manufacturer's protocol. The oligonucleotides used were for HPV16 E6, Fwd 5'-GGGGGATCCATGTTTCAGGACCCACAGGAGCGA-3', and Rev 5'-GGGGAATTCTTACAGCTGGGTTTCTCTACGTGT-3'; for HPV16 E7, Fwd 5'-GGGGGATCCATGCATGGAGATACCTACATTGC-3', and Rev 5'-GGGAAGCTTTTATGGTTTC TGAGAACAGATGGGGCA-3'; and for 18S, Fwd 5'-AACCTTGAACCCATT-3', and Rev 5'-CCATCCAATCGGTAGTAGCG-3'.

2.11. Statistical Analysis

P was calculated by Student's *t*-test from data presented as mean \pm standard deviation (SD). Significant differences were accepted at $P \leq 0.05$, as indicated.

3. Results

3.1. The LCR of HPV16 Contains Putative Binding Sites for SOX2

In order to identify transcription factor binding sites (TFBS) specific for HPV16- or HPV18-LCRs, we performed an *in silico* prediction analysis of probable TFBS using the JASPAR CORE database and alignment with consensus sequences. The results showed that although HPV16- and HPV18-LCRs shared TF binding elements, a group of sites that were restricted to each LCR, was identified. From this group, we found binding elements related to TFs involved in cellular processes such as proliferation and differentiation: FOX family (forkhead box); xenobiotic metabolism: Arnt (aryl hydrocarbon receptor nuclear translocator), FOXA2 (forkhead box); apoptosis: Ddit3 (DNA damage-inducible transcript 3); embryogenesis and tissue specificity: FOXD1, FOXF2, SOX2, MAFb (musculoaponeurotic fibrosarcoma homolog B) (Table 1).

The presence of SOX2 TFBS, found exclusively in the enhancer region of HPV16-LCR, was particularly interesting since SOX2 participates in stem cell pluripotency [33], squamous differentiation [34], and in cervical carcinogenesis [35,36]. In our analysis, 3 putative binding SOX2 response elements were found in the HPV16-LCR: Site 1, Site 2 and Site 3 (Figure 1A). These sites were similar to the reported SOX2 consensus sequence (Figure 1B) [37]. We became interested in analysing the role of SOX2 in HPV regulation since there were no previous reports indicating its function in this context.

Table 1. In silico prediction of transcription factor binding sites exclusive for HPV16- or HPV18-LCR.

	LCR-HPV16			LCR-HPV18		
Arnt	TGTTGTTTGTG	(nt 7180)	Ddit3::Cebpa	TGATTG	(nt 7205)	
	TGTTGTTTGTG	(nt 7180)		Gata1	GTACTGTGA	(nt 7504)
Foxa2	CACGTG	(nt 7268)	MZF1-1-4	CAATTGCT	(nt 7639)	
	CGTGTGTA	(nt 7270)		NFE2L2	ATAGTTA	(nt 7651)
FoxD1	GTGTATGTGTTT	(nt 7273)	NKX3-2	TAATTGCA	(nt 7720)	
	TGACACAAT	(nt 7303)		AATTA	(nt 7720)	
FoxF2	TGACACA	(nt 7305)	TAL1::TCF3	AGTATATA	(nt 22)	
	ATAAGTT	(nt 7325)				
HNF1B	TTTCCT	(nt 7569)				
	TCAGGAA	(nt 7570)				
MAFb	TTCCATTGTT	(nt 7597)				
MAX	TCCATTGTT	(nt 7598)				
Mycn	GTAAAAAA	(nt 7605)				
SOX2	CATTGTT	(nt 7600)				
	CATTGTT	(nt 7822)				
	CACATGG	(nt 7843)				

Transcription factor binding sites (TFBS) exclusive for each long control region (LCR). The sequence and nucleotide (nt) positions of putative TFBS are identified in human papillomavirus (HPV) 16 and HPV18 LCR sequences. Arnt (aryl hydrocarbon receptor nuclear translocator); Fox A2, Fox D1, Fox F2 (forkhead box); HNF1B (hepatocyte nuclear factor 1 homeobox B); MAFb (musculoaponeurotic fibrosarcoma homolog B); MAX (Myc-associated factor X); Mycn (avian myelocytomatosis viral oncogene neuroblastoma); SOX2 (sex determining region Y (SRY)-box 2); DDIT3 (DNA damage-inducible transcript 3); Cebpa (CCAAT/enhancer-binding protein α); GATA1 (globin transcription factor 1); MZF1-1-4 (myeloid zinc finger 1); NFE2L2 (nuclear factor erythroid 2 like 2); NKX3-2 (homeobox protein NK-3 homolog B); TAL1 (T-cell acute lymphocytic leukemia protein 1); TCF3 (transcription factor 3).

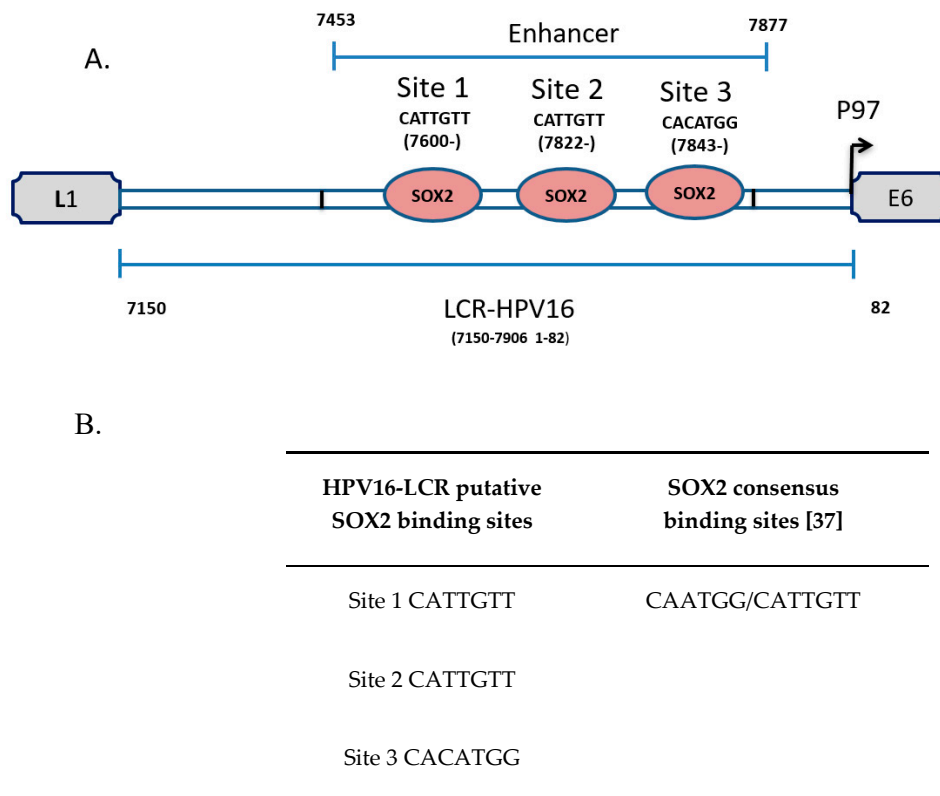


Figure 1. The LCR enhancer region of HPV16 contains sex determining region Y (SRY)-box 2 (SOX2) binding sites. **(A)** The nucleotide positions of the identified putative SOX2 binding sites (Site 1, Site 2 and Site 3) on the LCR enhancer region are shown; **(B)** Comparison of putative SOX2 binding sites to the reported SOX2 consensus sequences [37].

3.2. HPV16- and HPV18-LCRs Show Different Transcriptional Activities in Squamous Carcinoma and Adenocarcinoma Cell Lines

The regulation of HPV transcriptional activity during the viral life cycle may be influenced by variations in the amounts of TFs or cofactors along the stratified epithelium, as well as under different cellular contexts. In order to determine the effect of the cellular context (squamous or glandular) on the transcriptional activity of HPV16- and HPV18-LCRs, promoter-reporter assays were performed in CaSki, HaCaT and HeLa cell lines with reporter vectors containing the luciferase gene under the control of LCR16 or LCR18. A significant activity of pGL2-LCR16 compared to the control pGL2-B vector was observed in the three cell lines (Figure 2A–C). Nevertheless, the activity was considerably higher in cells with a squamous cell origin (SQ), with a mean 19 and 8.3 times higher than the activity of the pGL2-B control vector in CaSki and HaCaT cells, respectively (Figure 2A,B). Additionally, LCR16 activity in HeLa cells, although significant, was only 2.4 times above the control (Figure 2C).

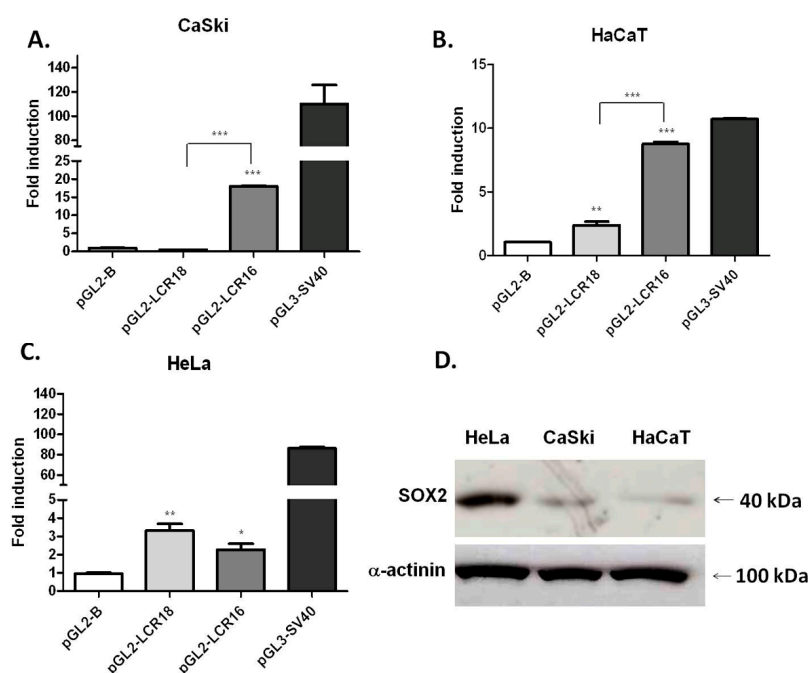


Figure 2. HPV16- and HPV18-LCRs present a different transcriptional activity in cells with distinct histological origin. Reporter plasmids were transfected in CaSki (A), HaCaT (B) and HeLa (C) cell lines. The transcriptional activity is shown as the fold induction, compared to the pGL2-B vector. The standard error of the mean is shown; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The square brackets show differences between transcriptional activities of LCRs from HPV16 and HPV18. pGL3-simian virus 40 (SV40) plasmid harboring SV40 promoter was used as a positive control. The graphs show representative experiments carried out in triplicate and repeated three times; (D) Representative western blot of SOX2 basal protein levels in HeLa, CaSki and HaCaT cell lines. For loading control α -actinin was used.

In contrast, LCR18 activity was significantly lower than that of LCR16 in CaSki cells (Figure 2A). In HaCaT, LCR18 activity was nearly double that of the control, but 3.7 times lower than LCR16 (Figure 2B). Furthermore, in HeLa cells LCR18 activity was 3.5 times above the control, and although not significant, it was slightly higher than that of LCR16 (Figure 2C). As expected for a positive control, the SV40 promoter presented a high activity in CaSki and HeLa cell lines (116 and 90 times above the control, respectively) (Figure 2A,C); while in HaCaT cells it was 10 times above pGL2-B control vector (Figure 2B). This is probably due to lower transfection efficiencies commonly seen in the HaCaT cell line compared to the others [38]. However, it is clear that the activity of LCR16 in HaCaT cells

is similar to that of the SV40 positive control. These results show a different behavior for HPV16- or HPV18-LCR depending on the cellular context of the tumor of origin (SQ or glandular). For LCR16, its transcriptional activity is favored in squamous epithelium, possibly from variations in the ratio of activating versus repressive factors. Conversely, the differences in transcriptional activity between HPV types could be due to variations in number, position or affinity of shared TFBS; or to TFBS which are present in one LCR, and absent in the other.

We analyzed endogenous SOX2 protein levels in the different cell lines by western blot. As shown in Figure 2D, HeLa cells expressed the highest level of SOX2, followed by CaSki; whereas HaCaT presented the lowest SOX2 protein levels (Figure 2D). Although many factors could be involved in the observed LCR transcriptional differences among our cell lines, it is noteworthy that the lower levels of SOX2 protein coincide with a relatively higher transcriptional activity of LCR16.

3.3. SOX2 Binds to the HPV16-LCR

3.3.1. SOX2 Binds to Its Putative Binding Sites in LCR16 In Vitro

To determine whether SOX2 recognizes the three putative SOX2 response elements found in the HPV16-LCR, EMSA was performed using nuclear proteins from HeLa cells. The formation of one major DNA-protein complex (*) was observed when nuclear extracts were incubated with the SOX2 binding consensus probe (CS), labelled with γ -ATP (32 P) (Figure 3A). In order to evaluate if the S1, S2 and S3 sites are able to compete with the DNA-Protein complex generated by the CS, different amounts of cold probes from S1, S2 or S3 were used. The complex formation with CS probe was strongly competed with the addition of 60 (a) and 260 (b) fold excess of the three cold probes; while cold non-specific oligonucleotide (NS) did not compete (Figure 3A). For S1, S2 or S3 probes we observed two major DNA-protein complexes (*) (Figure 3B–D). A similar behavior was observed when SOX2 sites (S1, S2 and S3) were inversely competed against a 50 (a) and 250 (b) fold excess of CS cold probe, where a clear drop in the DNA-protein complexes was observed, which was not observed with the NS cold probe (Figure 3B–D). These data suggest that SOX2 can bind to the three proposed binding sites found in the HPV16-LCR.

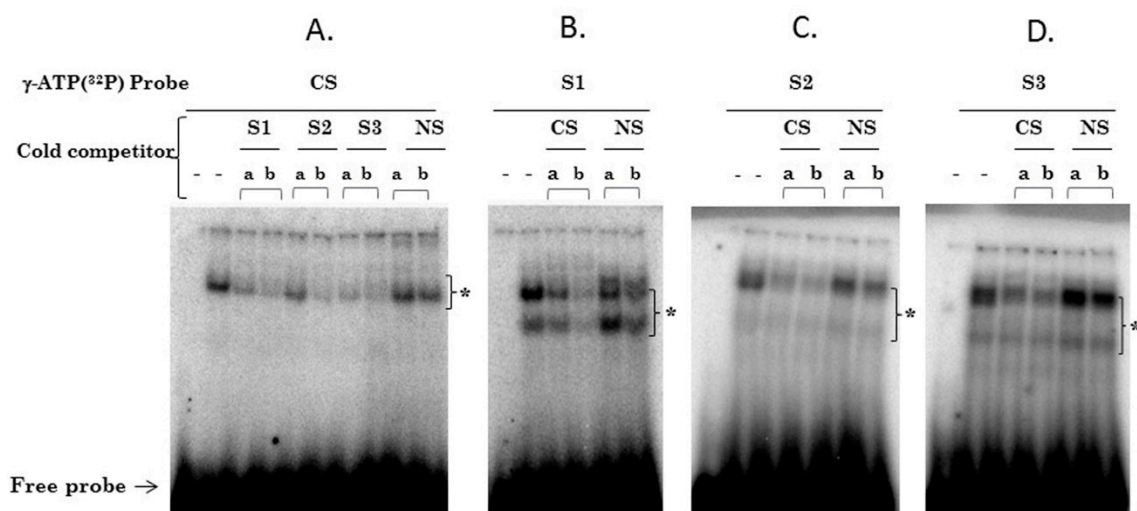


Figure 3. Cont.

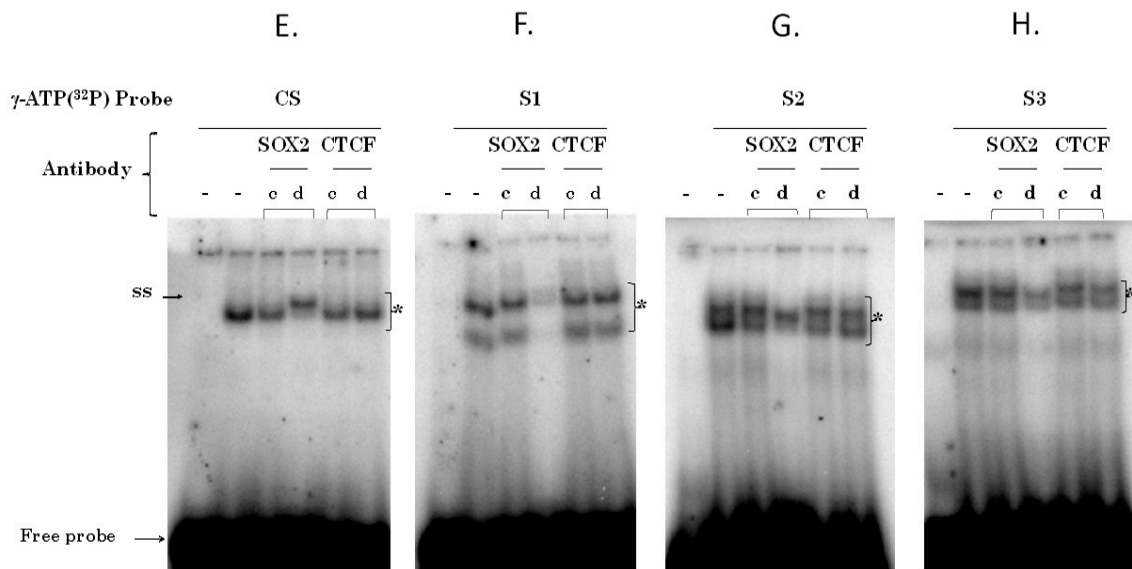


Figure 3. SOX2 binds to the putative SOX2 elements in HPV16-LCR. HeLa nuclear protein extracts were used in electrophoretic mobility shift assay (EMSA), to analyze the interactions between SOX2 and double-stranded DNA probes containing the three putative SOX2 binding sites (S1, S2 and S3) and the consensus binding site (CS). The first lanes of all the EMSAs are without nuclear extract; all the other lanes contain 15 μ g of nuclear extract. Competitions with 60 (a) or 260 (b) fold excess of S1, S2, S3, NS (non-specific) and CS cold probes are shown (A–D). DNA-protein complexes are denoted by (*). Free-labeled probes are shown at the bottom of the EMSAs. EMSAs E–H show the DNA-protein complexes with the respective antibodies added (anti-SOX2 or anti-CCCTC-binding factor (CTCF), as indicated) at 50 ng (c) or 2500 ng (d). SS is the supershift formed with the CS probe and 2500 ng (d) of anti-SOX2 antibody (E). In F–H, a clear decrease in the DNA-protein complexes from each probe (S1, S2 and S3) was observed with 2500 ng (d) of anti-SOX2.

In order to demonstrate that SOX2 protein is present in the DNA-protein complex, a supershift assay was performed. The presence of SOX2 in the DNA-protein complexes was confirmed after the addition of 50 ng (c) or 2500 ng (d) of anti-SOX2 antibody, resulting in a supershift band (SS) in the CS complex when using 2500 ng of antibody (d) (Figure 3E). A significant reduction of the DNA-protein complexes of S1, S2 and S3 probes was observed with 2500 ng (d) of SOX2 antibody (Figure 3F–H). The specificity of SOX2 antibody was confirmed, since no change occurred with 50 or 2500 ng (c and d) of anti-CTCF antibody (Figure 3E–H). Thus, our *in vitro* results support the novel interaction of SOX2 with the three putative binding sites present in the HPV16-LCR.

3.3.2. SOX2 Binds to the HPV16-LCR *In Vivo*

In order to confirm that SOX2 binds to the HPV16-LCR *in vivo*, ChIP assays were performed with chromatin extracts obtained from CaSki cells. ChIP results demonstrate that SOX2 binds *in vivo* to the HPV16-LCR. As a positive control we employed anti-p65 antibody, since it is known that p65 binds to HPV16-LCR [31]. Previously, it has been reported that CTCF is not recruited to HPV16-LCR, therefore we used anti-CTCF antibody as a non-binding control [32].

ChIP results show that the DNA fragment containing the HPV16-LCR SOX2 binding sites was qPCR-amplified from the anti-SOX2 immunoprecipitated sample, with near 40% above IgG background (Figure 4). Such enrichment was higher than the 8% observed when immunoprecipitating with anti-p65 antibody. When using anti-CTCF as a negative control, there was no important LCR amplification. The proportion of SOX2 binding to the LCR is even higher than that to its target Lefty (5%), used as positive control of SOX2 binding. These results demonstrate that SOX2 binds to the HPV16-LCR *in vivo*.

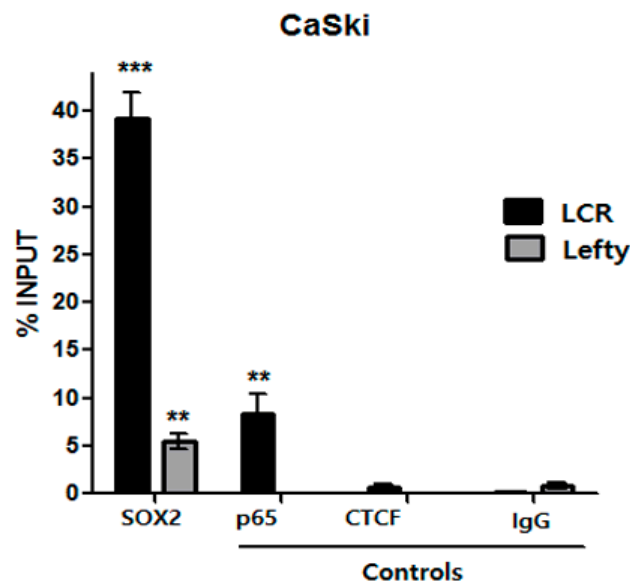


Figure 4. SOX2 binds to HPV16-LCR putative binding sites in vivo. Chromatin immunoprecipitation assay (ChIP) of SOX2 were performed evaluating the HPV16-LCR region (LCR) from CaSki cells. Anti-p65 and -CTCF antibodies were employed as positive and negative controls, respectively. Non-specific IgG antibody was included as a negative control. Quantitative PCR (qPCR) analysis was performed on the DNA obtained from the ChIP assay evaluating the HPV16-LCR region (black bars), and the Lefty (gray bars) gene promoter region, as a positive control of SOX2 binding ($n = 2$). The graphic represents the percentage of input recovered after immunoprecipitation with anti-SOX2, -p65 or -CTCF antibodies. Three independent biological experiments were performed. The standard error of the mean of triplicate qPCR measurements is shown. ** $p < 0.01$, *** $p < 0.001$.

3.4. SOX2 Represses HPV16-LCR Transcriptional Activity

In order to elucidate the effect of SOX2 on the transcriptional regulation of HPV16-LCR, CaSki cells were transduced with lentiviral vectors carrying SOX2, and/or GFP as control. After 24 h, cells were transfected with reporter plasmids pGL2-LCR16 or pGL2-LCR18. Cells were processed 48 h later. To assess the presence of a functional overexpressed SOX2, protein levels of Lefty, a known SOX2 gene target, were also measured at different days post transduction. As expected, Lefty protein levels decreased as SOX2 increased (Figure 5A). At 48 h post-transfection, significant repression of the LCR16 was observed due to SOX2 overexpression, remaining 3.5 times beneath the GFP transduced cells (Figure 5B). LCR18 activity was not affected by SOX2 overexpression. These results indicate that SOX2 is acting as a repressor of LCR16 in CaSki cells.

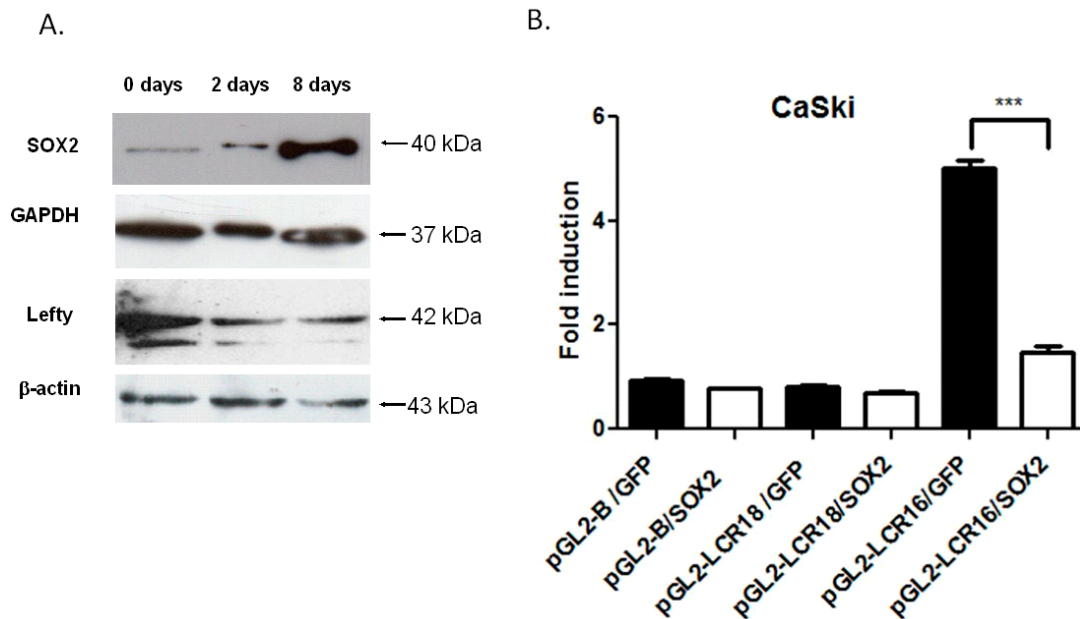


Figure 5. SOX2 represses the transcriptional activity of HPV16-LCR. (A) SOX2 was overexpressed in CaSki cells using a lentiviral system. Shown is a representative western blot of the expression levels of SOX2 and its target Lefty, at different days post-transduction (0, 2 and 8 days). Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and β -actin antibodies were used as protein loading controls; (B) CaSki cells transduced with lentiviral vectors expressing SOX2 and/or green fluorescent protein (GFP) were transfected with luciferase reporter plasmids containing LCR16 or LCR18. The transcriptional activity is shown as fold induction, compared to the pGL2-B/GFP vector. The graph shows a representative experiment carried out in triplicate and repeated three times. The standard error of the mean is shown. *** $p < 0.001$.

3.5. HPV16-LCR Sites S1, S2 and S3 Are Required for SOX2 Mediated Repression

To test whether SOX2 binding to S1, S2 and S3 sites is involved in the observed LCR16 transcriptional repression, mutations were generated in the pGL2-LCR16 vector (wt, wild type): 2-base pairs (bp) mutated in S1 site (M1), 3-bp mutated in each of the S2 and S3 sites (M2 and M3, respectively), and 8-bp mutations are introduced in the triple mutant (M1,2,3) (Figure S1). As shown through luciferase activity, a recovery of LCR16 transcriptional activity, compared to LCR16wt, is clearly observed with M1, M2 and M1,2,3 (white bars), suggesting that the respective mutations escape from the repressive effects of endogenous SOX2. To determine if SOX2 is involved in this effect, the protein was ectopically overexpressed (black bars); LCR16wt activity was repressed up to 66.2% and M1, M3 and M1,2,3 were not repressed by SOX2 overexpression; while M2 was mildly repressed in a proportion of 27.5%. These data confirmed that the three SOX2 putative binding sites S1, S2, and S3, contribute to SOX2 repressor activity (Figure 6) and that there is a direct interaction of SOX2 with the three sites.

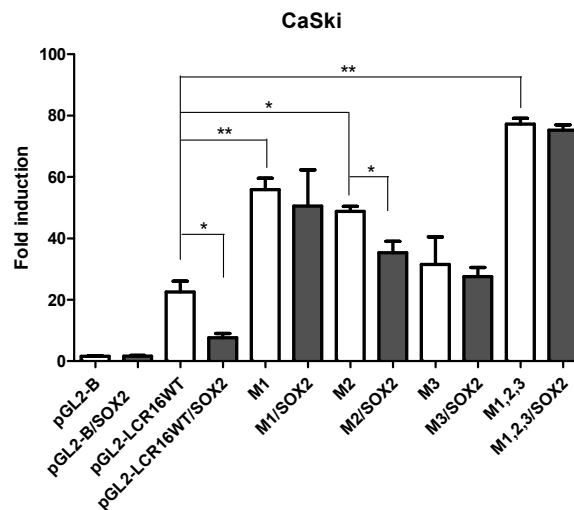


Figure 6. HPV16-LCR sites S1, S2 and S3 have a role in SOX2 repressor activity. Site directed mutagenesis was performed on the HPV16-LCR (wt, wild type) plasmid, generating M1 (LCR-Site 1 mutant), M2 (LCR-Site 2 mutant), M3 (LCR-Site 3 mutant) and M1,2,3 (LCR triple mutant) mutants. Transcriptional activity of luciferase reporter plasmids was measured in CaSki cells in the presence (black bars) or absence (white bars) of ectopically expressed SOX2. Fold induction of LCR16 activity was calculated in relation to the control vector pGL2-B. LCR16 mutant transcriptional activities were compared with pGL2-LCR16wt activity (white bars). Transcriptional activity of each LCR16 (wt, M1, M2, M3, or M1,2,3) was compared to its respective condition with and without SOX2 (white vs. black bars). Standard error of the mean is shown. * $p < 0.05$, ** $p < 0.01$.

3.6. E6 and E7 Endogenous Expression Is Down Regulated by SOX2

Since E6 and E7 viral oncogenes are directly regulated by the LCR, we next analyzed whether SOX2 was acting as a transcriptional repressor of such oncogenes in vivo. Two amounts of SOX2 transcription factor expressing plasmid (150 and 300 ng) were transfected into CaSki cells to analyze its effect on the endogenous levels of E6 and E7 transcripts (Figure 7). As shown in the western blot, SOX2 protein was overexpressed in SOX2 transfected cells (7A). Quantitative PCR analysis shows that E6 (7B) and E7 (7C) transcript levels were significantly decreased in the presence of SOX2 (about 40 to 33%, respectively). No SOX2 dose-dependent repression of E6 and E7 was observed.

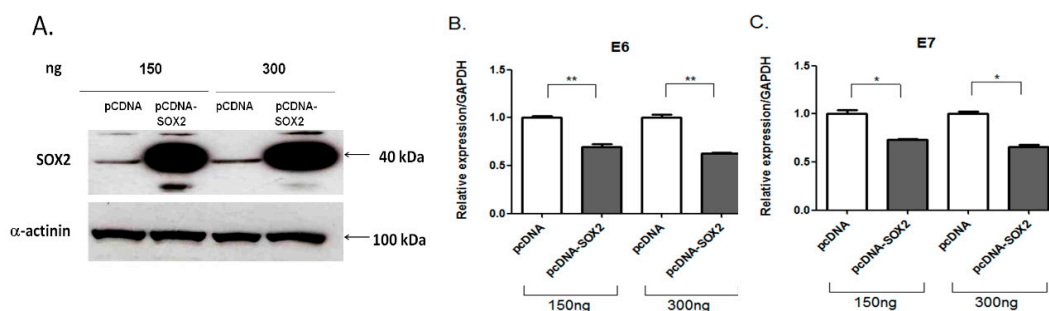


Figure 7. SOX2 represses endogenous E6 and E7 viral oncogenes expression. CaSki cells were transfected with 150 or 300 ng of SOX2 expressing plasmid. (A) SOX2 protein levels after transfection are shown. For protein loading control, α -actinin levels were analyzed. Quantitative Reverse transcription-polymerase chain reaction (qRT-PCR) of endogenous E6 (B) and E7 (C) viral oncogenes in the presence or absence of overexpressed SOX2, was performed. The empty vector pCDNA was used as a control. E6 and E7 expression levels were significantly diminished in the presence of SOX2. Data were normalized to the housekeeping gene GAPDH. Standard error of the mean is shown. * $p < 0.05$; ** $p < 0.01$.

4. Discussion

Several factors, including host, cellular or viral features, could determine the high prevalence of HPV16 observed in squamous cell carcinoma of the cervix, in relation to other oncogenic HPVs. Among them, it is possible that a set of specific transcription factors could differentially guide viral transcriptional activity in the different cell lines. To investigate the effect of the cellular context in the transcriptional activity of HPV16- and 18-LCRs, we performed assays with reporter plasmids. We found that LCR16 activity was significantly higher than that of LCR18 in CaSki and HaCaT cell lines, which have a SQ context, possibly due to a particular set of TFs shared in squamous cells. In contrast, in the adenocarcinoma HeLa cell line, LCR transcriptional activities are lower but similar for both HPV16- and 18-LCRs. This suggests that SQ context may favor LCR16 activity or that a glandular context (TFs found in cells with glandular origin) may repress it. A differential activity between LCR16 and LCR18 in the three cell lines was evident, which could be due in part to differences in number, position or affinity of shared TFBS, or to the presence of different TFBS for each LCR.

To identify novel TFs that could modulate the transcriptional activity of LCR16 and LCR18, we searched for putative binding sites for specific TFs. As expected, we found some TFBS shared between both viral types; however, there were others specific for each of the LCRs. We became interested in SOX2 transcriptional factor; due to the fact that consensus-binding sites for this factor were exclusively identified in the HPV16-LCR sequence and that it is involved in normal development as well as in malignant processes [39]. This TF is a member of the SOX (sex determining region Y (SRY)-box) transcriptional factor family and is a key controller of embryogenesis and cell fate during development [40], additionally participating in the maintenance of pluripotency [33]. Although the predominant role of SOX2 has been described as a transcriptional activator, it has recently been shown to play a dual role, also acting as a repressor of gene expression [41,42].

SOX2 has been linked to cancer hallmarks, promoting cell proliferation in breast, pancreatic, prostate and cervical cancers [36,43–46], as well as evading apoptosis in prostate, gastric and lung cancers [45,47,48]. Furthermore, SOX2 expression has been associated with early events of cervical carcinogenesis [49], since it is highly expressed in premalignant lesions, as well as in cervical cancer, while its expression is low in normal cervical epithelium. In addition, SOX2 overexpression in cervical cancer cell lines causes an increase in proliferation, clonogenicity and tumorigenicity [36]. These results suggest that SOX2 has an essential role in cervical carcinogenesis. Nevertheless, the effect of SOX2 in the transcriptional regulation of high-risk HPVs has not been previously studied.

In this study, we demonstrated that SOX2 binds to three putative binding sites identified in the enhancer sequence of LCR16 through direct interactions, as shown in EMSA and ChIP assays. We also confirmed that SOX2 is a transcriptional repressor of LCR16, since its activity decreased 3.5 times when SOX2 was over-expressed in CaSki cells (Figure 5). As has been previously reported, HPV16- and HPV18-LCRs share a repertoire of TFBS regulating their activity [28]; however, TFBS that are specific for each LCR type may partly explain the epidemiological data related to the different distribution of oncogenic HPV types, depending on cellular histological characteristics.

The direct interaction of SOX2 on the LCR16 putative binding sites was demonstrated through the analysis transcriptional activities of the mutated binding sites. A significant recovery of luciferase activity was observed with M1, M2 and M1,2,3 mutants in the basal state, which suggested that LCR16 mutants escaped from endogenous SOX2 repression. Moreover, when SOX2 was ectopically overexpressed, no repression of LCR16 M1 and M3 was observed and M2 mutant showed a mild repression (27.5%), in contrast to the observed for LCR16wt (66.2%). These data demonstrate that the three SOX2 putative binding sites contribute to the repressor activity of SOX2 in the HPV16-LCR.

SOX2 repressor activity on LCR16 was confirmed *in vivo* in CaSki cells which contain HPV16 endogenous sequences, where E6 and E7 oncogenes are naturally expressed. Quantitative PCR analysis showed that E6 and E7 transcript levels significantly decreased in cells overexpressing SOX2, although no dose dependent effect of SOX2 was observed, probably due to a saturation of the putative sites at the tested SOX2 doses. With these data, we demonstrate a direct effect of SOX2 on HPV16-LCR,

although we cannot discard the contribution of indirect factors. Bass et al. (2009), showed that an increase in the expression of various stem cell TFs was correlated with high levels of SOX2 expression in SCC lung tumors [34]; therefore, we cannot rule out the possibility that other transcription factors, regulated by SOX2, could also contribute to the modulation of LCR activity.

It is interesting that the analysis of SOX2 protein levels in the different cell lines showed that HeLa cells had the higher SOX2 protein levels, compared to the SCC lines (CaSki and HaCaT). These data match the higher transcriptional activities observed for LCR16 in the SQ context, in which HPV16 transcription could be favored due to a less repressive context. Therefore, under this background, it is possible that SOX2 is not acting as an activator of the LCR16 in SCC cells, but instead, as a repressor in adeno cells. Hence, high levels of SOX2 in HeLa cells could have a participation in the low activity observed for LCR16 in this cell type. However, it is important to consider that TFs that we did not analyze in this work surely influence the activity of the HPV LCRs.

It is important to note that even though in our study we did not delve further into the distinct transcriptional activities of HPV18-LCR observed in the analyzed cell contexts, it is a subject that deserves additional studies. The assays with reporter vectors showed a clear down-regulation of the LCR18 under a SQ cell context. Furthermore, the *in silico* analysis also showed putative TFs binding sites that were specific for HPV18-LCR. Thus, it would be relevant to study which novel transcriptional factors could be modulating LCR18 activity.

SOX2 regulation of HPV16-LCR could be biologically relevant in early HPV infections. HPV may infect basal as well as stem cells [50,51], where an overexpression of SOX2 in the latter would be expected. In fact, high levels of SOX2 have been demonstrated in cervical cancer [52] and particularly, in a subpopulation of tumor initiating cells with stem-like properties [53]. In such scenario, a controlled expression of E6 and E7 oncogenes, not only through viral but also through cellular proteins such as SOX2, may retard cellular damage and hence, viral persistence may be favored. The study of this landscape could help to explain, in part, the high prevalence of HPV16 in squamous cell carcinoma.

It is worth mentioning that in cervical SCCs a high expression of SOX2 has been reported; however, the loss of SOX2 in a subgroup of cervical cancer cases, has been associated with poor prognosis [49]. Furthermore, it has been shown that the increase in HPV oncogene mRNA expression is associated with poor prognosis in cervical cancer patients [54]. Consequently, if SOX2 acts as a transcriptional repressor of HPV16-LCR, its down-regulation in cancer cells could eventually promote an increase in the LCR activity, reflected as an overexpression of the E6 and E7 oncogenes. To further study this possibility, it would be necessary to determine whether in HPV16 positive cervical cancer biopsies the amount of SOX2 is inversely associated with the expression levels of E6 and E7 oncogenes. In this situation, it should also be considered that SOX2 associates with many partner proteins, including co-activators and co-repressors, and such diverse associations may differentially modulate target genes [55]. In neural stem cells it has been shown that when SOX2 acts as a repressor, it requires the interaction with members of the Groucho family of repressors [41]. Therefore, the analysis of such possible interactions in cervical cancer cells could provide relevant information for understanding the mechanisms of transcriptional regulation of HPV by SOX2.

In this study, we identified SOX2 as a new player in the transcriptional control of HPV16. A greater comprehension of the differential regulation of HPV types according to the cellular context will help to explain differences in their biological behavior.

Supplementary Materials: The following are available online at www.mdpi.com/1999-4915/9/7/175/s1, Figure S1: Identification of the nucleotides mutated in the Long Control Region (LCR) from HPV-16 by multi site-directed mutagenesis. Table S1: Primers used to introduce mutations in the putative SOX2 binding sites 1, 2 and 3 of the HPV16-LCR.

Acknowledgments: This work was part of the Ph.D. thesis of Imelda Martínez-Ramírez from the Programa de Maestría y Doctorado en Ciencias Bioquímicas, Universidad Nacional Autónoma de México (UNAM). The work was partially supported by Institutional funds (INCan 008/058/IBI) and by CONACyT grant 60722, and scholarship 45562. We thank María Alexandra Rodríguez-Sastre for technical assistance.

Author Contributions: M.L. and I.M.-R. conceived and designed the experiments; I.M.-R., A.A.-M., A.C.-G., V.d.C.-F. and I.B.M.-A. performed the experiments; I.M.-R., M.L. and E.L. analyzed the data; A.G.-C., E.S.-R., L.A.H. and A.Z.-D. contributed with analysis tools; M.L. and I.M.-R. wrote the paper.

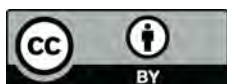
Conflicts of Interest: The authors declare that they have no conflict of interest.

References

1. GLOBOCAN 2012: Estimated Cancer Incidence, Mortality and Prevalence Worldwide in 2012. Available online: <http://globocan.iarc.fr/Default.aspx> (accessed on 4 October 2016).
2. Green, J.; de Gonzalez, A.B.; Sweetland, S.; Beral, V.; Chilvers, C.; Crossley, B.; Deacon, J.; Hermon, C.; Jha, P.; Mant, D.; et al. Risk factors for adenocarcinoma and squamous cell carcinoma of the cervix in women aged 20–44 years: The UK National Case-Control Study of Cervical Cancer. *Br. J. Cancer* **2003**, *89*, 2078–2086. [[CrossRef](#)] [[PubMed](#)]
3. Schlecht, N.F.; Kulaga, S.; Robitaille, J.; Ferreira, S.; Santos, M.; Miyamura, R.A.; Duarte-Franco, E.; Rohan, T.E.; Ferenczy, A.; Villa, L.L.; et al. Persistent human papillomavirus infection as a predictor of cervical intraepithelial neoplasia. *JAMA* **2001**, *286*, 3106–3114. [[CrossRef](#)] [[PubMed](#)]
4. Munoz, N.; Bosch, F.X.; de Sanjose, S.; Herrero, R.; Castellsague, X.; Shah, K.V.; Snijders, P.J.; Meijer, C.J.; International Agency for Research on Cancer Multicenter Cervical Cancer Study Group. Epidemiologic classification of human papillomavirus types associated with cervical cancer. *N. Engl. J. Med.* **2003**, *348*, 518–527. [[CrossRef](#)] [[PubMed](#)]
5. Benevolo, M.; Dona, M.G.; Ravenda, P.S.; Chiocca, S. Anal human papillomavirus infection: Prevalence, diagnosis and treatment of related lesions. *Expert Rev. Anti-Infect. Ther.* **2016**, *14*, 465–477. [[CrossRef](#)] [[PubMed](#)]
6. Lee, L.J.; Howitt, B.; Catalano, P.; Tanaka, C.; Murphy, R.; Cimbak, N.; DeMaria, R.; Bu, P.; Crum, C.; Horowitz, N.; et al. Prognostic importance of human papillomavirus (HPV) and p16 positivity in squamous cell carcinoma of the vulva treated with radiotherapy. *Gynecol. Oncol.* **2016**, *142*, 293–298. [[CrossRef](#)] [[PubMed](#)]
7. Fakhry, C.; Gillison, M.L. Clinical implications of human papillomavirus in head and neck cancers. *J. Clin. Oncol.* **2006**, *24*, 2606–2611. [[CrossRef](#)] [[PubMed](#)]
8. Chaturvedi, A.K.; Engels, E.A.; Pfeiffer, R.M.; Hernandez, B.Y.; Xiao, W.; Kim, E.; Jiang, B.; Goodman, M.T.; Sibug-Saber, M.; Cozen, W.; et al. Human papillomavirus and rising oropharyngeal cancer incidence in the United States. *J. Clin. Oncol.* **2011**, *29*, 4294–4301. [[CrossRef](#)] [[PubMed](#)]
9. De Villiers, E.M. Cross-roads in the classification of papillomaviruses. *Virology* **2013**, *445*, 2–10. [[CrossRef](#)] [[PubMed](#)]
10. Smith, J.S.; Lindsay, L.; Hoots, B.; Keys, J.; Franceschi, S.; Winer, R.; Clifford, G.M. Human papillomavirus type distribution in invasive cervical cancer and high-grade cervical lesions: A meta-analysis update. *Int. J. Cancer* **2007**, *121*, 621–632. [[CrossRef](#)] [[PubMed](#)]
11. Lacey, C.J.; Lowndes, C.M.; Shah, K.V. Chapter 4: Burden and management of non-cancerous HPV-related conditions: HPV-6/11 disease. *Vaccine* **2006**, *24* (Suppl. 3), S35–S41. [[CrossRef](#)] [[PubMed](#)]
12. Li, N.; Franceschi, S.; Howell-Jones, R.; Snijders, P.J.; Clifford, G.M. Human papillomavirus type distribution in 30,848 invasive cervical cancers worldwide: Variation by geographical region, histological type and year of publication. *Int. J. Cancer* **2011**, *128*, 927–935. [[CrossRef](#)] [[PubMed](#)]
13. Guan, P.; Clifford, G.M.; Franceschi, S. Human papillomavirus types in glandular lesions of the cervix: A meta-analysis of published studies. *Int. J. Cancer* **2013**, *132*, 248–250. [[CrossRef](#)] [[PubMed](#)]
14. Tenti, P.; Romagnoli, S.; Silini, E.; Zappatore, R.; Spinillo, A.; Giunta, P.; Cappellini, A.; Vesentini, N.; Zara, C.; Carnevali, L. Human papillomavirus types 16 and 18 infection in infiltrating adenocarcinoma of the cervix: PCR analysis of 138 cases and correlation with histologic type and grade. *Am. J. Clin. Pathol.* **1996**, *106*, 52–56. [[CrossRef](#)] [[PubMed](#)]
15. Siriaunkgul, S.; Utaipat, U.; Suthipintawong, C.; Tungsinmunkong, K.; Triratanachat, S.; Khunamornpong, S. HPV genotyping in adenocarcinoma of the uterine cervix in Thailand. *Int. J. Gynaecol. Obstet.* **2013**, *123*, 226–230. [[CrossRef](#)] [[PubMed](#)]

16. Bouallaga, I.; Massicard, S.; Yaniv, M.; Thierry, F. An enhanceosome containing the Jun B/Fra-2 heterodimer and the HMG-I(Y) architectural protein controls HPV 18 transcription. *EMBO Rep.* **2000**, *1*, 422–427. [[CrossRef](#)] [[PubMed](#)]
17. Apt, D.; Watts, R.M.; Suske, G.; Bernard, H.U. High Sp1/Sp3 ratios in epithelial cells during epithelial differentiation and cellular transformation correlate with the activation of the HPV-16 promoter. *Virology* **1996**, *224*, 281–291. [[CrossRef](#)] [[PubMed](#)]
18. Chan, W.K.; Chong, T.; Bernard, H.U.; Klock, G. Transcription of the transforming genes of the oncogenic human papillomavirus-16 is stimulated by tumor promoters through AP1 binding sites. *Nucleic Acids Res.* **1990**, *18*, 763–769. [[CrossRef](#)] [[PubMed](#)]
19. Chong, T.; Apt, D.; Gloss, B.; Isa, M.; Bernard, H.U. The enhancer of human papillomavirus type 16: Binding sites for the ubiquitous transcription factors oct-1, NFA, TEF-2, NF1, and AP-1 participate in epithelial cell-specific transcription. *J. Virol.* **1991**, *65*, 5933–5943. [[PubMed](#)]
20. Thierry, F.; Heard, J.M.; Dartmann, K.; Yaniv, M. Characterization of a transcriptional promoter of human papillomavirus 18 and modulation of its expression by simian virus 40 and adenovirus early antigens. *J. Virol.* **1987**, *61*, 134–142. [[PubMed](#)]
21. DiMaio, D.; Petti, L.M. The E5 proteins. *Virology* **2013**, *445*, 99–114. [[CrossRef](#)] [[PubMed](#)]
22. Doorbar, J.; Quint, W.; Banks, L.; Bravo, I.G.; Stoler, M.; Broker, T.R.; Stanley, M.A. The biology and life-cycle of human papillomaviruses. *Vaccine* **2012**, *30* (Suppl. 5), F55–F70. [[CrossRef](#)] [[PubMed](#)]
23. Wallace, N.A.; Galloway, D.A. Novel functions of the human papillomavirus E6 oncoproteins. *Ann. Rev. Virol.* **2015**, *2*, 403–423. [[CrossRef](#)] [[PubMed](#)]
24. Swift, F.V.; Bhat, K.; Younghusband, H.B.; Hamada, H. Characterization of a cell type-specific enhancer found in the human papilloma virus type 18 genome. *EMBO J.* **1987**, *6*, 1339–1344. [[PubMed](#)]
25. Carson, A.; Khan, S.A. Characterization of transcription factor binding to human papillomavirus type 16 DNA during cellular differentiation. *J. Virol.* **2006**, *80*, 4356–4362. [[CrossRef](#)] [[PubMed](#)]
26. Kajitani, N.; Satsuka, A.; Kawate, A.; Sakai, H. Productive lifecycle of human papillomaviruses that depends upon squamous epithelial differentiation. *Front. Microbiol.* **2012**, *3*, 152. [[CrossRef](#)] [[PubMed](#)]
27. Garcia-Vallve, S.; Iglesias-Rozas, J.R.; Alonso, A.; Bravo, I.G. Different papillomaviruses have different repertoires of transcription factor binding sites: Convergence and divergence in the upstream regulatory region. *BMC Evol. Biol.* **2006**, *6*, 20. [[CrossRef](#)] [[PubMed](#)]
28. Sichero, L.; Sobrinho, J.S.; Villa, L.L. Identification of novel cellular transcription factors that regulate early promoters of human papillomavirus types 18 and 16. *J. Infect. Dis.* **2012**, *206*, 867–874. [[CrossRef](#)] [[PubMed](#)]
29. Lopez-Saavedra, A.; Gonzalez-Maya, L.; Ponce-de-Leon, S.; Garcia-Carranca, A.; Mohar, A.; Lizano, M. Functional implication of sequence variation in the long control region and E2 gene among human papillomavirus type 18 variants. *Arch. Virol.* **2009**, *154*, 747–754. [[CrossRef](#)] [[PubMed](#)]
30. The JASPAR Database. Available online: <http://jaspar2014.genereg.net/> (accessed on 13 February 2015).
31. Fontaine, V.; van der Meijden, E.; de Graaf, J.; ter Schegget, J.; Struyk, L. A functional NF- κ B binding site in the human papillomavirus type 16 long control region. *Virology* **2000**, *272*, 40–49. [[CrossRef](#)] [[PubMed](#)]
32. Paris, C.; Pentland, I.; Groves, I.; Roberts, D.C.; Powis, S.J.; Coleman, N.; Roberts, S.; Parish, J.L. CCCTC-binding factor recruitment to the early region of the human papillomavirus 18 genome regulates viral oncogene expression. *J. Virol.* **2015**, *89*, 4770–4785. [[CrossRef](#)] [[PubMed](#)]
33. Chew, J.L.; Loh, Y.H.; Zhang, W.; Chen, X.; Tam, W.L.; Yeap, L.S.; Li, P.; Ang, Y.S.; Lim, B.; Robson, P.; et al. Reciprocal transcriptional regulation of *Pou5f1* and *Sox2* via the Oct4/Sox2 complex in embryonic stem cells. *Mol. Cell. Biol.* **2005**, *25*, 6031–6046. [[CrossRef](#)] [[PubMed](#)]
34. Bass, A.J.; Watanabe, H.; Mermel, C.H.; Yu, S.; Perner, S.; Verhaak, R.G.; Kim, S.Y.; Wardwell, L.; Tamayo, P.; Gat-Viks, I.; et al. SOX2 is an amplified lineage-survival oncogene in lung and esophageal squamous cell carcinomas. *Nat. Genet.* **2009**, *41*, 1238–1242. [[CrossRef](#)] [[PubMed](#)]
35. Yang, Z.; Pan, X.; Gao, A.; Zhu, W. Expression of Sox2 in cervical squamous cell carcinoma. *J. BUON* **2014**, *19*, 203–206. [[PubMed](#)]
36. Ji, J.; Zheng, P.S. Expression of Sox2 in human cervical carcinogenesis. *Hum. Pathol.* **2010**, *41*, 1438–1447. [[CrossRef](#)] [[PubMed](#)]
37. Rodda, D.J.; Chew, J.L.; Lim, L.H.; Loh, Y.H.; Wang, B.; Ng, H.H.; Robson, P. Transcriptional regulation of nanog by OCT4 and SOX2. *J. Biol. Chem.* **2005**, *280*, 24731–24737. [[CrossRef](#)] [[PubMed](#)]

38. Yong, W.; Peng, D.; Wang, L.; Dong, Z.; He, B. Screening of HaCaT clones for CCL20 gene knockout and preliminary exploration of gene-targeting vector transfection approaches in this cell line. *Med. Sci. Monit. Basic Res.* **2015**, *21*, 21–28. [[PubMed](#)]
39. Liu, K.; Lin, B.; Zhao, M.; Yang, X.; Chen, M.; Gao, A.; Liu, F.; Que, J.; Lan, X. The multiple roles for Sox2 in stem cell maintenance and tumorigenesis. *Cell. Signal.* **2013**, *25*, 1264–1271. [[CrossRef](#)] [[PubMed](#)]
40. Sarkar, A.; Hochedlinger, K. The Sox family of transcription factors: Versatile regulators of stem and progenitor cell fate. *Cell Stem Cell* **2013**, *12*, 15–30. [[CrossRef](#)] [[PubMed](#)]
41. Liu, Y.R.; Laghari, Z.A.; Novoa, C.A.; Hughes, J.; Webster, J.R.; Goodwin, P.E.; Wheatley, S.P.; Scotting, P.J. Sox2 acts as a transcriptional repressor in neural stem cells. *BMC Neurosci.* **2014**, *15*, 95. [[CrossRef](#)] [[PubMed](#)]
42. Zhou, C.; Yang, X.; Sun, Y.; Yu, H.; Zhang, Y.; Jin, Y. Comprehensive profiling reveals mechanisms of SOX2-mediated cell fate specification in human ESCs and NPCs. *Cell Res.* **2016**, *26*, 171–189. [[CrossRef](#)] [[PubMed](#)]
43. Stolzenburg, S.; Rots, M.G.; Beltran, A.S.; Rivenbark, A.G.; Yuan, X.; Qian, H.; Strahl, B.D.; Blancafort, P. Targeted silencing of the oncogenic transcription factor SOX2 in breast cancer. *Nucleic Acids Res.* **2012**, *40*, 6725–6740. [[CrossRef](#)] [[PubMed](#)]
44. Herreros-Villanueva, M.; Zhang, J.S.; Koenig, A.; Abel, E.V.; Smyrk, T.C.; Bamlet, W.R.; de Narvajias, A.A.; Gomez, T.S.; Simeone, D.M.; Bujanda, L.; et al. SOX2 promotes dedifferentiation and imparts stem cell-like features to pancreatic cancer cells. *Oncogenesis* **2013**, *2*, e61. [[CrossRef](#)] [[PubMed](#)]
45. Jia, X.; Li, X.; Xu, Y.; Zhang, S.; Mou, W.; Liu, Y.; Liu, Y.; Lv, D.; Liu, C.H.; Tan, X.; et al. SOX2 promotes tumorigenesis and increases the anti-apoptotic property of human prostate cancer cell. *J. Mol. Cell Biol.* **2011**, *3*, 230–238. [[CrossRef](#)] [[PubMed](#)]
46. Weina, K.; Utikal, J. SOX2 and cancer: Current research and its implications in the clinic. *Clin. Transl. Med.* **2014**, *3*, 19. [[CrossRef](#)] [[PubMed](#)]
47. Hutz, K.; Mejias-Luque, R.; Farsakova, K.; Ogris, M.; Krebs, S.; Anton, M.; Vieth, M.; Schuller, U.; Schneider, M.R.; Blum, H.; et al. The stem cell factor SOX2 regulates the tumorigenic potential in human gastric cancer cells. *Carcinogenesis* **2014**, *35*, 942–950. [[CrossRef](#)] [[PubMed](#)]
48. Chen, S.; Li, X.; Lu, D.; Xu, Y.; Mou, W.; Wang, L.; Chen, Y.; Liu, Y.; Li, X.; Li, L.Y.; et al. SOX2 regulates apoptosis through MAP4K4-survivin signaling pathway in human lung cancer cells. *Carcinogenesis* **2014**, *35*, 613–623. [[CrossRef](#)] [[PubMed](#)]
49. Kim, B.W.; Cho, H.; Choi, C.H.; Ylaya, K.; Chung, J.Y.; Kim, J.H.; Hewitt, S.M. Clinical significance of OCT4 and SOX2 protein expression in cervical cancer. *BMC Cancer* **2015**, *15*, 1015. [[CrossRef](#)] [[PubMed](#)]
50. Ortiz-Sanchez, E.; Santiago-Lopez, L.; Cruz-Dominguez, V.B.; Toledo-Guzman, M.E.; Hernandez-Cueto, D.; Muniz-Hernandez, S.; Garrido, E.; Cantu-De Leon, D.; Garcia-Carranca, A. Characterization of cervical cancer stem cell-like cells: Phenotyping, stemness, and Human Papilloma Virus co-receptor expression. *Oncotarget* **2016**, *7*, 31943–31954. [[CrossRef](#)] [[PubMed](#)]
51. Lopez, J.; Ruiz, G.; Organista-Nava, J.; Gariglio, P.; Garcia-Carranca, A. Human papillomavirus infections and cancer stem cells of tumors from the uterine cervix. *Open Virol. J.* **2012**, *6*, 232–240. [[CrossRef](#)] [[PubMed](#)]
52. Boumahdi, S.; Driessens, G.; Lapouge, G.; Rorive, S.; Nassar, D.; Le Mercier, M.; Delatte, B.; Caauwe, A.; Lenglez, S.; Nkusi, E.; et al. SOX2 controls tumour initiation and cancer stem-cell functions in squamous-cell carcinoma. *Nature* **2014**, *511*, 246–250. [[CrossRef](#)] [[PubMed](#)]
53. Feng, D.; Peng, C.; Li, C.; Zhou, Y.; Li, M.; Ling, B.; Wei, H.; Tian, Z. Identification and characterization of cancer stem-like cells from primary carcinoma of the cervix uteri. *Oncol. Rep.* **2009**, *22*, 1129–1134. [[PubMed](#)]
54. De Boer, M.A.; Jordanova, E.S.; Kenter, G.G.; Peters, A.A.; Corver, W.E.; Trimbo, J.B.; Fleuren, G.J. High human papillomavirus oncogene mRNA expression and not viral DNA load is associated with poor prognosis in cervical cancer patients. *Clin. Cancer Res.* **2007**, *13*, 132–138. [[CrossRef](#)] [[PubMed](#)]
55. Cox, J.L.; Mallanna, S.K.; Luo, X.; Rizzino, A. Sox2 uses multiple domains to associate with proteins present in Sox2-protein complexes. *PLoS ONE* **2010**, *5*, e15486. [[CrossRef](#)] [[PubMed](#)]





Research Signpost
37/661 (2), Fort P.O.
Trivandrum-695 023
Kerala, India

Cervical Cancer: From Public Health to Molecular Pathogenesis, 2015: 25-45
ISBN: 978-81-308-0555-9 Editors: Luis del Pozo-Yauner and Jorge Meléndez Zajjla

2. Transcriptional regulation of Human Papillomavirus genes

Marcela Lizano-Soberón¹, Imelda Martínez-Ramírez¹ and
Adriana Contreras-Paredes¹

¹*Unidad de Investigación Biomédica en Cáncer. Instituto Nacional de Cancerología, México*
Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México

Abstract. Human papillomaviruses (HPVs) comprise a large family of pathogens of the cutaneous or mucosal squamous epithelia. Despite the knowledge of the genomic sequences of many HPV types, the regulation of viral gene expression is still uncertain. Viral gene expression in infected cells depends on cell differentiation and is tightly regulated at the transcriptional and post-transcriptional levels. All papillomaviruses contain a 400-850 bp long noncoding region (referred to as long control region LCR), which is positioned between the L1 and E6 genes and contains most of the elements that regulate transcription of the viral genome. Only one strand of viral DNA is transcribed and more than one promoter is involved in the transcription initiation process, resulting multiple mRNAs with several open reading frames. The early promoters of HPV-16 (P97) and HPV-18 (P105) have been well characterized, as well as the late differentiation-regulated promoter of HPV-16 (P670). The LCR exhibits recognition sites for positive and negative cellular factors, such as AP-1, Sp-1, YY1 and Oct-1; and binding sites for the viral E2 proteins. The LCR also contains glucocorticoid response

Correspondence/Reprint request: Dr. Marcela Lizano-Soberón, Instituto Nacional de Cancerología, Av. San Fernando No. 22, Col., Sección XVI Delegación Tlalpan, C.P. 14080, D.F, México
E-mail: lizanosoberon@gmail.com

elements (GRE) and an enhancer component that specifically activates transcription in epithelial cells through the synergism between a diversity of transcription factors. Gene regulation occurs at transcriptional and post-transcriptional levels. Papillomavirus are transcribed as bicistronic or polycistronic forms that frequently contain two or more ORFs, and are polyadenylated at the early and late poly(A) sites. This Chapter will focus on the mechanisms of transcriptional regulation of HPV genes.

Introduction

Human Papillomaviruses (HPVs) show a great diversity; however their genomic organization is well conserved. These viruses are non-enveloped and contain a double-stranded circular DNA genome of approximately 8 kb in size divided according to its function into three main regions: Early (E), Late (L) and the Long Control Region (LCR or upstream regulatory region URR). Two polyadenylation sites separate these regions. Papillomaviruses have a polycistronic transcription pattern, which yields a series of mRNAs processed by alternative splicing and polyadenylation (1). The early region encode six open reading frames that are expressed early during infection, generating the E1, E2, E4, E5, E6 and E7 proteins, and more recently described, the fusion protein E8-E2C also appear to have a regulatory function in viral transcription and replication. Two L genes are expressed lately producing L1 major and L2 minor capsid proteins. The genetic regulation resides mainly in the non-coding region known as the LCR.

The Early proteins are involved in critical functions such as HPV genome replication and transcription, control of the cell cycle, apoptosis and immune modulation. Some important functions of HPV proteins are described in Table 1.

A differential expression of viral early and late genes occur in oncogenic HPV infections, due to the activation of the viral early promoter in undifferentiated keratinocytes and the viral late promoter in highly differentiated keratinocytes (2, 3, 4). E1, E2, E5, E6 and E7 proteins are expressed along the viral cycle throughout the differentiating epithelia, with reduced expression in the more differentiated layers. Expression of E1, E4, L1 and L2 proteins occur in the upper spinous and more differentiated granular layers of the epithelium (5, 6).

The LCR is the most variable region within Papillomaviruses, consisting of 850 to 1000 base pairs. It carries the origin of replication with binding sites for E1 and E2 viral proteins, and multiple recognition sites for cellular transcription factors, important for transcription initiation through the RNA polymerase II (16). Promoters of HPV-18, HPV-31 and HPV-16 have been well characterized (17, 18).

Table 1. HPV proteins and functions.

Protein	Function
E1	Viral DNA replication, ATPase, and DNA helicase (7,8).
E2	Regulation of viral gene transcription and DNA replication. During early stages of viral infection, represses the transcription of the oncogenes E6 and E7 (9).
E4	Disruption of cytokeratin networks/cell growth arrest (10).
E5	Modifies the activity of the epidermal growth factor receptor, induces inflammatory cell signaling pathway, contributes to malignant transformation complementing the roles of E6 and E7, enhances angiogenesis and proliferation and may contribute to cancer progression (10,11).
E6	Transformation and tumorigenesis (binds to p53 amongst other proteins). In addition, a large number of interactions between the E6 and cellular proteins have been reported (12,13.)
E7	Transformation (binds to pRb amongst other proteins), alter cell cycle control through interactions with histone deacetylases, cyclins and cyclin-dependent kinase inhibitors (14).
L1	The only Major capsid protein (15).
L2	Minor capsid protein (15).

HPV genome integration

Papillomavirus infections are usually transient or result in benign lesions. Nevertheless, sometimes infections with “high risk” or oncogenic HPVs such as HPV-16, HPV-18 and HPV-31 persist, promoting the development of malignant lesions.

A common consequence of cancer associated to persistent infections with high-risk HPVs is the integration of the viral genomes into the cell genome, frequently disrupting the viral E2 regulator gene (19,20). In the normal viral life cycle, HPV genomes exist in a circular or episomal state and are retained in the basal cells of the squamous epithelium at approximately 50 to 100 copies per cell (21). In this state, viral gene expression is mainly regulated by E2 protein (22). High-level expression of E6 and E7 oncogenes is supposed to occur only in suprabasal post-mitotic cells (23,24), where the viral oncogenes induce unprogrammed re-entry into S-phase, activating the host replication machinery. This cell environment is needed for viral genome amplification prior to virion synthesis. In

productive infections, viral oncogenes fail to transform since they are expressed in cells destined to be lost due to the constant renewal of the cervical squamous epithelium. Meanwhile, oncogene expression in basal epithelial cell inhibits differentiation and induces chromosomal instability, representing the potential cells for transformation events (25,26).

In the majority of cervical carcinomas an over-expression of E6 and E7 oncogenes is observed in cells in which truncated viral genomes are integrated into the host DNA. HPV genomic integration correlates with increased viral gene expression and cellular growth advantage, promoting the development of epithelial displasia and eventually cancer (27). Viral integration often occurs in the E1 and E2 region and results in the loss of negative feedback control of oncogene expression. In this situation, interactions of E6 and E7 with cellular suppressor proteins such as p53 and pRb increase, and consequently suppressor functions are inhibited (28,29). Many different interactions of oncogenes E6 and E7 with cellular proteins have been described and the majority of those interactions seem to affect cell cycle regulation (30,31). This is the case for hDlg, which is a member of the membrane-associated guanylate kinase (MAGUK) family of proteins, involved in the inhibition of cell cycle (32). *In vitro* assays as well as results obtained through cervical cancer cell line analysis, have demonstrated that E6 proteins from high risk HPV, bind to hDlg promoting its degradation via the ubiquitin proteosome pathway (33,34). It has also been found that E6 intratype variants of HPV-18 differentially modulate hDlg degradation, rebounding in levels of the suppressor protein PTEN and activated PKB, which are important proteins in survival signal transduction pathways (35). Through these interactions HPV-18 E6 variants are also able to strongly upregulate other elements implicated in cell proliferation.

Genomes of “Low-risk” or non-oncogenic HPVs, such as HPV-6 and HPV-11, are not found integrated. These viral types can induce some type of dysplasia or benign condyloma acuminata, but it is established that they do not promote cancer (26).

The mechanism of HPV integration is not well understood, since it is not a normal event in viral cell cycle. Therefore integration must represent a chance occurrence; probably through events that increase the frequency of double-strand breaks in host and viral DNA (36).

HPV transcriptional regulatory region

The integration of the viral genome into the cellular genome is an essential feature for carcinogenesis. The integrated viral segment contains the regulatory region of the virus, known as “long control region (LCR)” or

“upstream regulatory region (URR)” and the sequence encoding E6 and E7 oncogenes. These oncogenes are transcribed from the promoter located in the regulatory region.

The LCR is responsible for carrying out the genetic regulation of the papillomaviruses. Its sequence from 800-1000 bp is extremely variable depending on the viral type; although common regulatory elements, such as AP-1 and SP-1, are conserved (37). The LCR contains binding sites for the E2 protein on the viral DNA replication origin, where E1 and E2 proteins play a very important role. HPV early promoters, such as P105 for HPV-18 and P97 for HPV-16 (38); and P90, P270 and P680 for HPV6 and HPV11, are active throughout the process of viral genome replication (39,40). Those promoters contain the cellular enhancer region whose transcriptional activity depends on cellular transcription factors like, AP-1, NF-1, TEF-1 and Oct-1, whose combination may confer some tissue specificity, (41). Within the papillomavirus genome, late promoters are activated in the upper epithelial or granular layer, through the cell differentiation process. Some of those promoters are located within the E7 gene, such as P670 and P142, for HPV-16 and HPV-31, respectively (42). Additionally, HPV-16 promoter activities have been also detected within the 5' region of the LCR, also in the L1 region, and within E6 ORF (43,44), whose importance are not yet known although they appear as weak promoters (45).

Within the first portion of the LCR is also the TATA box where transcriptional factors bind in conjunction with polymerase II, carrying out transcription of the viral genome.

LCR transcriptional regulation

Transcriptional regulation by E2 viral protein

The E2 viral protein has approximately 360 amino acids and contains three functional domains: an N-terminal transactivation domain, an internal “hinge” domain, and a DNA binding domain (DBD) located at the C-terminus (Fig. 1). Both the C-terminal and N-terminal domains are relatively well conserved within the papillomaviruses. Based on the results of X-ray crystallography, structural and molecular models of both the N- and C-terminal domains have been constructed (46,47). E2 binds as a dimer at DNA-binding sites through its C-terminal DBD. Some evidence suggests that the transactivation domain mediates linking activity between E2 molecules bound at distant E2-binding sites, and as a consequence DNA loops are formed (48,47).

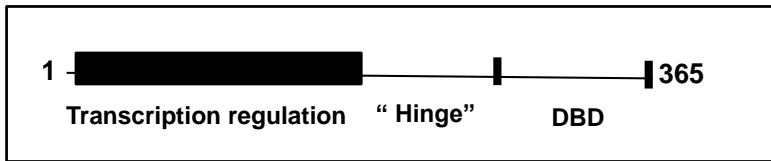


Figure 1. E2 Protein domains. A region of approximately 200 amino acids at the N-terminus (thick line) is relatively well conserved among the papillomaviruses and encodes the transactivation domain. The 85 C-terminal amino acids comprise the DNA-binding and dimerization domain (DBD).

The papillomavirus E2 protein has demonstrated many functions in the viral cycle, including roles in transcriptional regulation and in viral DNA replication (49). Besides these historical roles, which rely on their aptitude to bind to specific viral DNA sequences, E2 also modulates the host cells through direct protein interactions to several cellular proteins, such as Brd4, cNAP1, Gps2 and p300 (50-54). E2 interacts with topoisomerase II binding protein 1 (TopBP1), which enhances its ability to activate transcription and replication (55). The binding of p53 to E2 alters its transcriptional activity and also inhibits HPV DNA replication (56,57). Finally, E2 binds to APC activators Cdh1 and Cdc20 through which the high-risk E2 proteins might induce genomic instability (58).

On the other hand, E2 regulates replication of the viral genome, due to its interaction with E1 viral protein. It also regulates transcription of viral early genes; E2 can repress E6/E7 promoter from promoter-proximal binding sites (49). E2 also plays important roles in the segregation of HPV genomes between daughter cells after cell division (59,60).

In several HPV types, a spliced RNA product, that expresses a fusion protein consisting of the small E8 domain fused to E2C (61), is produced. The functions of the E8^ΔE2C protein have been mainly investigated in HPV-31. E8^ΔE2C knockout HPV-31 genomes display a strong over-replication of viral genomes in short-term analyses (61). Yet, in stable cell lines, these knockout genomes are not maintained as episomes but are found integrated into the host chromosomes, suggesting that E8^ΔE2C is required for the long-term extrachromosomal maintenance of viral genomes (61). Genetic and biochemical analyses of the 31E8^ΔE2C protein have demonstrated that the 31E8 domain is required for transcriptional repression. This is due to the recruitment of cellular co-repressors, such as the histone deacetylase 3 (HDAC3)/N-CoR complex, by the 31E8 domain (62). E8^ΔE2C activities may vary among different papillomaviruses, since analysis in the context of

HPV-16 show that 16E8^{E2C} is not required for stable maintenance of HPV-16 episomes (63).

The understanding of the roles of E2 proteins has been evolving over the years. Binding of E2 to DNA palindromic sequences has been associated with dimerization of the protein. E2 recognizes the consensus sequence, 5'-ACCGNNNNcGGT-3', with nucleotide positions 4 and 9 allowing some variability (64,65). The sequence of the 4-nucleotide spacer varies by HPV type, and is thought to be critical for determining E2 binding affinity and gene regulation (66). The E2 homodimer binds the DNA by the alpha helices of each monomer by contact with two successive major grooves of the target site (67,47). Four E2 binding sites are conserved in the LCR of most papillomaviruses, assigned with numbers according to distance from the early promoter (68). Each site is differentially regulated by variable binding affinity for the E2 protein, resulting in varying replication and transcriptional effects during the viral life cycle (51).

Regarding LCR activity regulation by E2 during a viral infection, mutation of the promoter-proximal E2 binding site in HPV-31 genome induces transcriptional repression (69). In an HPV-Epstein-Barr virus (EBV) model system the HPV-16 E2 protein represses the transcription of the HPV-16 LCR (70). By other side, it has been suggested that transcriptional activity of the HPV-31 E2 protein is not required during the viral life cycle, in this context being critical the replication function of E2 (71).

HPV E2 is among the first early proteins expressed from the HPV genome involved in viral transcription, replication and maintenance of viral genomes (72,73). E2 is known to regulate the viral DNA replication. During this process the papillomavirus E1 helicase is recruited by E2 to the viral origin of replication, where it assembles into a double hexamer that begins replication of the viral genome (74,75). On the other hand, the role of E2 in the control of HPV gene expression is not clearly known. There are several ways in which the E2 proteins can either activate or repress transcription, depending on the position of the E2-binding sites and the nature of the E2 polypeptides. In one case, binding of the full-length E2 transactivator to a number of sites located upstream from the promoter activates transcription (74). In these cases, the E2 sites behave like conditional enhancer elements. On the other hand, E2 can also repress transcription, when the E2-binding motif are situated close to the transcriptional start site, overlapping with binding sites for essential cellular transcription factors. This mechanism suggests that the role of E2 transcription repressor or activator depends of binding sites within the LCR. The promoter-proximal E2 binding sites are flanked by a TATA box sequence and a Sp-1-binding site. E2 sites are located just upstream from the TATA box and binding of E2 to these sites may sterically hinder the binding of the cellular factor TFIID, and/or

interfere with the assembly of the transcriptional initiation complex (75). In addition to competing with Sp-1 and the TATA box binding protein (TBP) for binding to DNA, E2 can also interact with both of these proteins (76,77). It has been proposed that E2 levels might determine its affinity to promoter E2 binding sites, since at early stages of infection there are low levels of E2 that might increase the binding of this protein to a distal promoter, resulting in transcription activation; whereas at later times of infection higher E2 protein levels might bind to the proximal promoter E2 sites inducing to the repression of transcription. However, other HPV types E2 proteins binds equally well to all of its sites within the genome (78,79).

HPV transcriptional regulation through cellular factors

Papillomavirus early gene transcription is regulated by sequences located in the LCR known as cis-acting response elements. Some response elements located in HPV-16 and -18 LCRs are described in table 2. These elements bind a variety of cellular transcription factors. Is worth to mention that both, the composition of a transcriptional control region, as well as the space between different regulatory elements, pressure the transcriptional activity (80,81). Due to its high prevalence, HPV-16 and HPV-18 types have been the main models in the study of HPV transcriptional regulation.

Transcription factors which modulate LCR activity of high-risk HPVs, as SP-1, AP-1, NF-1, Oct-1, YY-1 and the glucocorticoid receptor GRE (Figure 2) (82), are also important in normal cell growth, development and differentiation.

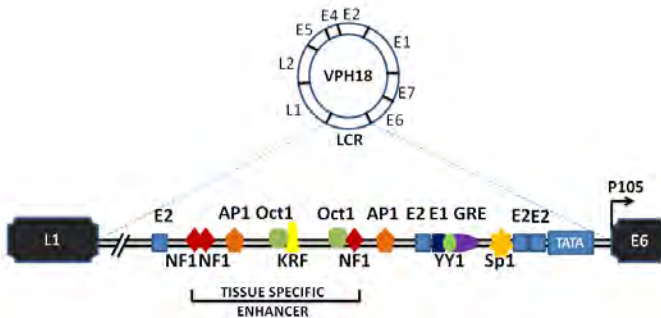


Figure 2. Schematic representation of the HPV-18 LCR, showing binding sites for cellular and viral regulatory proteins. Four binding sites for the E2 viral protein are shown. The origin of replication is defined by E1 together with the adjacent E2 binding sites. These sites are conserved among mucosal HPV types, as well as the Sp-1 and AP-1 response elements.

Table 2. HPV-16 and -18 LCR response elements.

Virus type	Binding site (Transcription factor)	Sequences 5'→3'	Nucleotide Position
HPV16	AP-1	TTAATCA	7631-7637
		TGAGTCA	7648-7654
		TTGTCA	7810-7816
	NF-1	TTGGC	7473-7477
		TTGGC	7558-7554
		TTGGC	7591-7587
		TTGGC	7679-7675
		TTGGC	7711-7715
		TTGGC	7742-7746
		TTGGC	7770-7766
	GRE	TGTACATTGTGTCAT	7641-7655
	Oct-1	AATTGCAT	7732-7739
	Sp-1	GGGCGT	28-33
KRF	ND	ND	
HPV18	AP-1	TTAGTCA	7608-7614
		TTAGTCA	7798-7792
	NF-1	CTGGCN5TGCAAA	7513-7528
		TTGGCN8TTGGC	7569-7586
		TTGGC	7731-7735
	Oct-1	GCCTTGCAT	7644-765
		AATTGCAT	7721-772
	KRF-1	TGCATAACTATATCC	7647-7669
	GRE	ACTCCCTA	7839-7853
	Sp-1	AGCACATACTATACT	35-40
		GGGAGT	

Positions of consensus binding sequences for each transcription factor on HPV-16 and -18 LCR.

Sp-1

Sp-1 is a zinc finger transcription factor that binds to GC-rich motifs of many promoters. This factor has three isoforms and is involved in many

cellular processes such as cell differentiation, cell growth, apoptosis, immune responses, response to DNA damage, and chromatin remodeling. Post-translational modifications significantly affect Sp-1 activity. In conjunction with AP-1, Sp-1 plays a critical role in transcriptional regulation of mucosal high- and low-risk HPVs. Sp-1 binding sites in HPV-16 and HPV-18 are generally found close to the transcriptional initiation site. Mutation of this site in HPV-18 leads to a strong reduction of the transcriptional activity in the URR, reinforcing its importance in transcriptional activation of E6/E7 expression (82). Sp-1 performs its transcriptional activation in cooperation with other cis elements found within the viral enhancer.

AP-1

AP-1 is a transcription factor constituted by heterodimers of Jun and Fos family proteins by hydrophobic interactions of “Leucine-Zippers” (83). AP-1 controls the expression of genes involved in cellular proliferation, differentiation and transformation (84). JunB and Fra2, components of AP-1, are highly expressed in keratinocyte and modulate its differentiation through regulation of several genes such as involucrin, loricrin, cytokeratins or integrins (85).

HPV-18 URR contains two AP-1 elements, one located within the constitutive enhancer region and other within the promoter-proximal region, playing a key role in activation of the HPV-18 E6/E7 promoter (82).

Glucocorticoid Response Element (GRE)

Steroid hormones may act as cofactors for HPV in cervical cancer development. Glucocorticoids in particular, such as the synthetic dexamethasone, have demonstrated to enhance the transformation rate of high-risk HPV-transfected cells (86,87).

HPV-11, HPV-16 and HPV-18 have active glucocorticoid response elements (GREs), since glucocorticoids activate transcription from their early promoters. In the proximal promoter region of HPV-18 LCR it is found AGCACATACTATACT sequence, which confers dexamethasone and progesterone response. The LCR HPV-31 contains at least one functional GRE, although sequence analysis identified three potential GREs (88).

Oct-1 and NF-1

The Oct-1 transcription factor was among the first identified members of the POU transcription factor family. Members of this family contain the POU domain, a 160-amino acid region necessary for DNA binding to the octameric sequence ATGCAAAT (89). An Oct-1-like site located near a NF-1 element at the enhancer of HPV-16 has been identified, that even being aberrant it can activate the promoter of E6/E7. Functional differences among high-risk HPV genomic variants have been described, suggesting repercussions on their oncogenic potential. The LCR of HPV-18 variants have distinct transcriptional activities in different Cervical Cancer cell lines (90). It has been proposed that these differences may be due to variations in transcription factor binding sites KRF-1/OCT-1 (positions 7651, 7658) and OCT-1/TEF-1 (positions 7726, 7730).

Neurofibromin 1 (NF-1) is a Ras-GAP protein, which negatively regulates Ras signaling. Spliced transcript variants encoding different isoforms have been described for this gene (91,92). However, the precise cellular function of neurofibromin has yet to be clarified. Using proteomic strategies, a set of neurofibromin-associating cellular proteins was identified (92). The consensus NF-1 transcription factor-binding site is (5'- TTGGC-3'). In HeLa cells, NF-1 synergizes with Oct-1 bound to a juxtaposed site. NF-1 has an activating function in HPV its role in HPV-18 is less clear (82).

YY-1

The transcription factor YY-1 is a ubiquitously distributed zinc finger protein that regulates important cellular processes such as growth, differentiation and development and plays an important role in tumorigenesis (93) and HPV infection (94). It is involved in repressing (95) and activating (96) a diverse number of cellular and viral promoters. YY-1 may direct histone deacetylases and histone acetyltransferases to a promoter in order to activate or repress. In HPV-18 (97) and HPV-11 (98) YY-1 represses major early promoters when bound to a distal upstream site. In HPV-8 YY-1 appears both to repress and to activate E6 promoter (99); while in HPV-16, YY-1 repress the E6 and E7 transcription when bound upstream of the P97 start site (100).

Post-transcriptional regulation

HPV gene expression depends on cell differentiation and is regulated at the transcriptional and post-transcriptional levels. Papillomavirus transcripts are

transcribed as a bicistronic or polycistronic from genes frequently containing two or more ORFs, polyadenylated at an early and a late poly (A) sites.

Post-transcriptional regulation of gene expression can involve regulation at various steps in mRNA biogenesis, as splicing (101), polyadenylation (102), nuclear export (103), RNA stability (104), and translation (105). For example, addition of the 7 methyl guanosine cap stabilizes the RNA, and this is essential for export to the cytoplasm and for translation on polysomes; splicing removes introns from the pre-mRNA, joining exonic coding regions. Alternative splicing in viral RNA is a means of production of multiple proteins. Splicing is also important for nuclear export of mRNA to the cytoplasm and for appropriate translation (106).

How keratinocyte differentiation regulates HPV gene expression remains to be determined, but it is believed that infection of basal cells leads to the expression of the early region of the viral genome, producing the six non-structural viral proteins (E1, E2, E4, E5, E6 and E7). In more differentiated keratinocytes the two structural viral capsid proteins are expressed (L1 and L2) from the late region of the viral genome.

The study of the complexity of papillomavirus gene expression has been a hard task and limited knowledge exists concerning the proteins translated from each viral transcript. Much of the information has been obtained from HPV-16 transcript mapping.

HPV RNA splicing

RNA splicing regulates viral gene expression at the post-transcriptional level, giving rise to a bunch of possibilities in HPV coding potential of its pre-mRNAs. Extensive RNA mapping in HPV infected cells and raft tissues have led to the establishment of transcription maps from various HPV genomes (107,108), which provide important information for further HPV studies (Fig. 3). One splice donor and two alternative splice acceptor sites within the E6 transcript are unique to high-risk HPV types (109); accordingly, low-risk HPV types do not undergo splicing, being essentially monocistronic. Transcripts in HPV-16 could be spliced through three alternative 3' splice sites, positioned in the first intron at nt 409, 526 or 742 (110); and three other alternative 3' splice acceptor sites at nt 2582, 2709, 3358 in the second intron (111,112,113). All combinations can potentially produce more than 14 species of mRNA transcripts with various coding potentials. All the early region transcripts contain a downstream 3'exon that includes E2C/E4/E5a/E5b ORFs, while all the late region transcripts contain a 3'exon that encodes the L1 major capsid protein. Consequently, most mRNAs contain multiple ORFs, being some of them polycistronic.

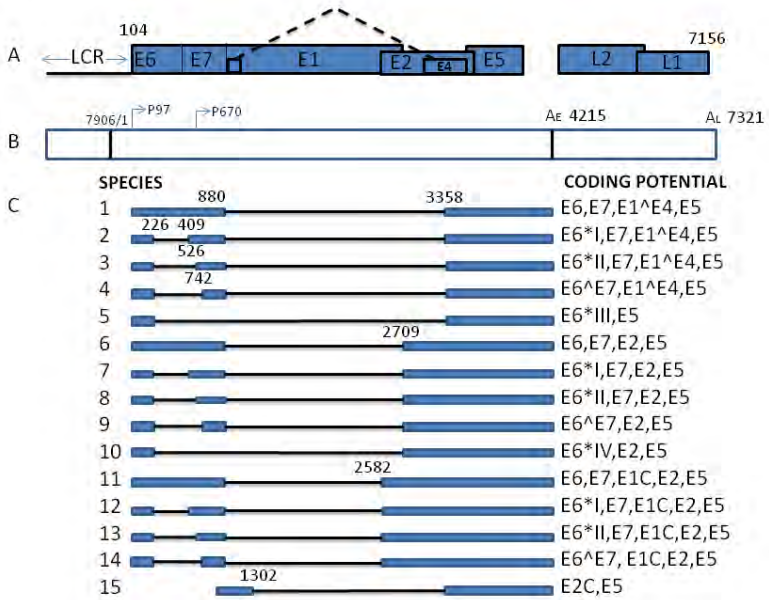


Figure 3. Genome structure and transcription map of Early genes of HPV-16. A) Blue boxes represent open reading frames of early and late genes. Formation of E4 requires RNA splicing (upper dashed lines). LCR is the long control region. B) Linear form of the virus genome with promoters (arrows) and early (A_E) and late (A_L) polyadenylation sites. C) Some early RNA species derived from alternative promoter usage and alternative splicing. Exons (blue thick lines) and introns (thin lines) are exemplified on each species of the RNA, with the indicated splice site nucleotide positions.

Five small Nuclear Ribonucleoproteins (snRNP) named U, participate in viral RNA splicing removing an intron and synchronizing the binding of individual exons and splicing factors. Classical intronic splicing signals have three cis-elements consisting in a 5' splice site (5'ss), a branch site, and a 3' splice site (3'ss) (110). The set up of the exon-intron boundary is the first step in the correct recognition of an intron 5'ss and an intron 3'ss, which involves the interaction of the 5'ss with U1 snRNP of the branch site, with U2 snRNP, and the 3'ss with other heterodimers of U2 auxiliary factors. This process requires many cellular splicing factors, such as SR proteins (serine/arginine-rich), and exonic splicing enhancers (ESEs). Moreover, regulation of HPV-16 splicing late pre-mRNA also appears to utilize splicing silencers (ESS). The presence of ESS immediately upstream of the

nt 3632 5'ss (114) may be necessary to counteract U1 recognition of the nt 3632 5'ss and enhance recognition of the early polyadenylation site (pA) at nt 4215 by the cellular polyadenylation machinery (115). Meanwhile, a 79 nt RNA regulatory element, the late regulatory element (LRE), involved in the regulation of L1 and L2 expression in response to epithelial differentiation is sited at the 3' end of the L1 gene and extends in the late 3' untranslated region (3' UTR). This element represses late gene expression in differentiating epithelia and may be active in differentiated cells (116).

Alternative splicing of HPV RNA regulates the final concentration of E6 and E7 proteins of the high-risk papillomaviruses. Splicing in the E6 ORF can give rise to the smaller species E6*I, E6*II, and E6*III, affecting the amount of full length E6 protein (20). Some authors propose that this kind of splicing also regulates translation of the E7 ORF by enhancing translation initiation at the AUG in the E7 ORF (117,110,118,119)). Splicing of E6/E7 RNA has been shown to provide more E7 RNA templates and to promote E7 translation, whereas a lack of RNA splicing was found to lower the levels of E7 proteins. Nevertheless, the excess of an efficient splicing of E6/E7 transcripts could prevent the expression of enough amounts of full-length E6. Even there is no evidence of how does keratinocyte differentiation regulates viral RNA splicing, it is possible that the splicing process could be precisely controlled by cellular factors differentially expressed according to the cell differentiation status. As a result, high risk HPVs appear to limit the activities of the full-length E6 protein, under certain conditions, although it is not known whether this occurs in stage of infection dependant manner.

DNA methylation in papillomavirus gene expression

As a consequence of HPV integration into the host DNA, the viral genome becomes a potential target for cellular control mechanisms acting on the integration site. Thus, DNA methylation is suggested to be another mechanism involved in HPV transcriptional regulation. The introduction of methyl groups to cytosine residues at CpG dinucleotides often leads to a transcriptional silencing; therefore DNA methylation regulates cellular and viral gene expression. When the methylation is held in a promoter region, it prevents transcription factors from binding to their recognition sites. Alternatively, methylated CpG dinucleotides interact with methyl CpG binding proteins, which recruit histone deacetylase in association with chromatin. Histone deacetylation then leads to changes in the chromatin configuration inhibiting transcription.

Significance of the methylation in papillomavirus gene expression and carcinogenesis is not fully studied; even some authors showed that CpG

methylation within the binding sites of HPV16 E2 protein, disturb DNA-binding activity of E2 (120,121), affecting viral gene expression. By other side, in the host genome context CpG methylation is found to be more common in cervical cancer than in premalignant lesions (122), being a frequent event throughout the host genome. It has been reported that in HPV-16 and HPV-18 genomes, methylation occurs more often in LCR regions and part of L1 ORF region, which is not transcribed when the viral genome is integrated (123,124). Some authors propose that CpG methylation in integrated HPV DNA may prevent most viral DNA from transcription activation, leaving only few transcription units per cell (125).

Papillomavirus RNA polyadenylation

Viral gene expression, as well as eukaryotic gene expression, is highly regulated by RNA polyadenylation. Cleavage of the nascent transcript in the pre-mRNA and the addition of a poly(A) tail with 150-200 adenylate residues are coupled events. RNA polyadenylation is an important feature of post-transcriptional regulation of gene expression, since it confers stability to the nascent RNA and favours ribosome recruitment improving translation. At least five cellular factors are involved in RNA polyadenylation, including poly(A) polymerase (PAP), cleavage and polyadenylation specificity factor CPSF, cleavage stimulation factor (CstF), and cleavage factors I and II (126). The late transcription unit overlaps at part of the early transcription unit and contains two polyadenylation sites (early and late) (Fig. 3) which potentially are in competition with each other and are mutually exclusive events. Early to late switch in papillomavirus gene expression could be due to changes in either splicing or polyadenylation or both. In addition, negative cis-elements (NRE), spanning from the 3' end of the L1 ORF into the late 3' UTR regulate HPV-16 late transcripts (127). The NRE binds multiple RNA processing factors, including U2AF65, CstF-64, HuR, and SF2/ASF (128-131). Even the exact function of the HPV-16 NRE is not fully understood; it is known that two NRE-binding proteins (U1A and U2F65) may inhibit polyadenylation. The NRE may also compete with downstream CstF binding polyadenylation element for CstF, leading to a less stable polyadenylation complex and reduced polyadenylation efficiency (132).

Polyadenylation and splicing are both co-transcriptional processes that can modulate each other. A 65-nt AC rich splicing enhancer located approximately 100-nts downstream of the splice site 3358 3' of HPV-16 stimulates splicing and up-regulates polyadenylation of the early transcripts at the early poly(A) signal (114), mainly leading to the expression of E1⁺E4 early transcript. By other side, L1 mRNA splicing of the nt 3632 5' splice

site to the 5639 3' site is supposed to compete with early polyadenylation since these two events are mutually exclusive.

Conclusion

Papillomavirus transcription is restricted to epithelial cells of human origin, more specifically to keratinocytes. These viruses encode multiple proteins in a very compact genome, taking advantage of alternative splicing to differentially express the proteins in a temporal manner, linked with the cell differentiation program. Viral transcription is controlled by response elements present in the long control region. Tissue specificity is modulated by epithelial cellular factors acting in the LCR transcriptional enhancer. HPV gene expression relies on a complex relationship of differential promoter use, selection of the appropriate polyadenylation site, generation of alternative splice isoforms of polycistronic RNAs transcribed from the genome, and translation efficiency of the mRNAs generated. All these processes are controlled by cellular DNA or RNA-binding factors, some of which are known to be expressed in a differentiation stage-specific manner. Significant progress has been made in the comprehension of the mechanisms of transcriptional and post-transcriptional regulation of viral gene expression. However, this knowledge must be strengthened to profoundly understand papillomavirus life cycle, host interactions and viral oncogenesis.

References

1. Milligan, S., Veerapradistin, T., Ahamat, B., Mole, S., and Graham, S. 2007, *Virology*, 360, 172.
2. Jia, R., Liu, X., Tao, M., Kruhlak, M., Guo, M., Meyers, C., Baker, C.C., and Zheng, Z.M. 2009, *J. Virol.*, 83, 167.
3. Middleton, K., Peh, W., Southern, S., Griffin, H., Sotlar, K., Nakahara, T., El-Sherif A., Morris, L., Seth, R., Hibma, M., Jenkins, D., Lambert, P., Coleman, N., and Doorbar, J. 2003, *J. Virol.*, 77, 10186.
4. Wang, H.K., Duffy, A.A., Broker, T., and Chow, L.T. 2009, *Genes Dev.*, 23, 181.
5. Doorbar, J. 2005, *J. Clin. Virol.*, 32, S7.
6. Longworth, M.S., and Laimins, L.A. 2004, *Microbiol. Mol. Biol. Rev.*, 68, 362.
7. Lehoux, M., Fradet-Turcotte, A., Lussier-Price, M., Omichinski, J.G., and Archambault, J. 2012, *J. Virol.*, 86, 3486.
8. Moody, C.A., and Laimins, L.A. 2010, *Nat. Rev. Cancer*, 10, 550.
9. Hegde, R.S. 2002, *Annu. Rev. Biophys. Biomol. Struct.*, 31, 343.
10. Hamid, N.A., Brown, C., and Gaston, K. 2009, *Cell. Mol. life Sci.*, 66, 1700.

11. Kim, M.K., Kim, H.S., OH, J.M., Han J.Y., Lim, J.M. Juhnn, Y.S., and Song, Y.S. 2010, *Biochem Pharmacol*, 80, 1930.
12. Ghittoni, R., Accardi, R., Hasan, U., Gheit, T., Sylla, B., and Tommasin, M. 2010, *Virus Genes*, 40, 1.
13. Tungteakkhun, S.S., and Duerksen-Hughes, P. 2008, *Arch. Virol.*, 153, 397.
14. Prendiville, W., and Davies, P. 2004, *The Health Professional's HPV Handbook: Human papillomavirus and Cervical Cancer*, Oxford (Ed), Taylor and Francis, 11.
15. McMurray, H.R., Nguyen, D., Westbrook, T.F., and McAnce, D.J. 2001, *Int. J. Exp. Pathol.*, 82, 15.
16. Bernard, H.U. 2002, *Antivir Ther.*, 7, 219.
17. Bednarek, P.H., Lee, B.J., Gandhi, S., Lee, E., and Phillips, B. 1998, *J. Virol.*, 72, 708.
18. Sen, E., Alam, S., and Meyers, C. 2004, *Virol.*, 78, 612.
19. Cullen, A.P., Reid, R., Champion, M., and Lorincz, A.T. 1991, *J. Virol.*, 65, 606.
20. Daniel, B., Mukherjee, G., Seshadri, L., Vallikad, E., and Krishna, S. 1995, *J. Gen. Virol.*, 76, 2589.
21. Bedell, M., Hudson, J.B., Golub, T.R., Turyk, M.E., Hosken, M., Wilbanks, G., and Laimins, L.A. 1991, *J. Virol.*, 65, 2254.
22. Gloss, B., and Bernard, H.U. 1990, *J. Virol.*, 64, 5577.
23. Durst, M., Glitz, D., Schneider, A., and zur Hausen, H. 1992, *Virology*, 189, 132.
24. Cheng, S., Schmidt-Grimminger, D.C., Murant, T., Broker, T.R., and Chow, L.T. 1995, *Genes Dev.*, 9, 2335.
25. Duensing, S., and Munger, K. 2002, *Cancer Res.*, 62, 7075.
26. Munger, K., Baldwin, A., Edwards, K.M., Hayakawa, H., Nguyen, C.L., Owens, S.M., Grace, M., and Huh, K. 2004, *J. Virol.*, 78, 11451.
27. Jeon, S., Allen-Hoffmann, B.L., and Lambert, P.F. 1995, *J. Virol.*, 69, 2989.
28. Adams, P.D. 2001, *Biochim. Biophys. Acta*, 1471, M123.
29. Scheffner, M., Huibregtse, J.M., Vierstra, R.D., and Howley, P.M. 1993, *Cell*, 75, 495.
30. Malanchi, I., Accardi, R., Diehl, F., Smet, A., Androphy, E., Hoheisel J., and Tommasino, M. 2004, *J. Virol.*, 78, 13769.
31. Motoyama, S., Ladines-Llave, C.A., and Villanueva L.I., Maruo T. 2004, *Kobe J. Med. Sci.*, 50, 9.
32. Ishidate, T., Matsumine, A. Toyoshima, A.K., and Akiyama T. 2000, *Oncogene*, 19, 365.
33. Massimi, P., Gammoh, N., Thomas, M., and Banks, L. 2004, *Oncogene*, 23, 8033.
34. Thomas, M., Glaunsinger, B., Pim, D., Javier, R., and Banks, L. 2001, *Oncogene*, 20, 5431.
35. Contreras-Paredes, A., De la Cruz-Hernández, E., Martínez-Ramírez, I., Duenas-González, A., and Lizano, M. 2009, *Virology*, 383, 78.
36. Pett, M., and Coleman, N. 2007, *J. Pathol.*, 212, 356.
37. Gius, D., Grossman, S., Bedell, M.A. and Laimins, L.A. 1988, *J. Virol.*, 62, 665.

38. Thierry, F., and Yaniv, M. 1987, *EMBO J.*, 6, 3391.
39. Thierry, F., and Howley, P.M. 1991, *New Biol.*, 3-90.
40. Thierry, F., Heard, J.M., Dartmann, K., and Yaniv, M. 1987, *J. Virol.*, 61, 134.
41. Chiang, C.M., Ustav, M., Stenlund, A., Ho, T.F., Broker, and T., Chow, L.T. 1992, *Proc. Natl. Acad. Sci. USA*, 89, 5799.
42. Grassmann, K., Rapp, B., Maschek, H., Petry, K.U., and Iftner, T. 1996, *J. Virol.*, 70, 2339.
43. Glahder, J.A., Hansen, C.N., Vinther, J., Madsen, B.S., and Norrild, B. 2003, *J. Gen. Virol.*, 84, 3429.
44. Tan, S., Baker, C.C., Stunkel, W., and Bernard, H.U. 2003, *Virology*, 305, 486.
45. Rosenstierne, M.W., Vinther, J., Hansen, C.N., Prydsoe, M., and Norrild, B. 2003, *J. Gen. Virol.*, 84, 2909.
46. Hegde, R.S., Grossman, S.R., Laimins, L.A., and Sigler, P.B. 1992, *Nature*, 359, 505.
47. de Prat-Gay, G., Gaston, K., and Cicero, D.O. 2008, *Front. Biosci.*, 13, 6006.
48. Giri, I., and Yaniv, M. 1988, *EMBO J.*, 7, 2823.
49. Thierry, F. 2009, *Virology*, 384, 375.
50. Breiding, D.E., Sverdrup, F., Gossel, M.J., Moscufo, N., Broonchai, W., and Androphy, E.J. 1997, *Mol. Cell. Biol.*, 17, 7208.
51. Donaldson, M.M., Boner, W., and Morgan, I.M. 2007, *J. Virol.*, 81, 4338.
52. Ilves, I., Maemets, K.T., Silla, T., Janikson, K., and Ustav, M. 2006, *J. Virol.*, 80, 3660.
53. Muller, A., Ritzkowsky, A., and Steger, G. 2002, *J. Virol.*, 76, 11042.
54. Rehtanz, M., Schmidt, H.M., Warthorst, U., and Steger, G. 2004, *Mol. Cell. Biol.*, 24, 2153.
55. Boner, W., Taylor, E.R., Tsirimonaki, E., Yamane, K., Campo, M.S., and Morgan, I.M. 2002, *J. Biol. Chem.*, 277, 22297.
56. Massimi, P., Pim, D., Bertoli, C., Bouvard, V., and Banks, L. 1999, *Oncogene*, 18, 7748.
57. Brown, C., Kowalczyk, A.M., Taylor, E.R., Morgan, I.M., and Gaston, K. 2008, *Virol. J.*, 5, 5.
58. Bellanger, S., Blachon, S., Mechali, F., Bonne-Andrea, C., and Thierry, F. 2005, *Cell Cycle*, 4, 1608.
59. Zheng, P. S., Brokaw, J., and McBride, A.A. 2005, *J. Virol.*, 79, 1500.
60. You, J., Croyle, J.L., Nishimura, A., Ozato, K. and Howley, P.M. 2004, *Cell*, 117, 349.
61. Stubenrauch, F., Hummel, M., Iftner, T., and Laimins, L.A. 2000, *J. Virol.*, 74, 1178.
62. Fertey, J., Ammermann, I., Winkler, M., Stoger, R., Iftner, T. and Stubenrauch, F. 2010, *J. Virol.*, 84, 9505.
63. Lace, M.J., Anson, J.R., Thomas, G.S., Turek, L.P., and Haugen, T.H. 2008, *J. Virol.*, 82, 10724.
64. Li, R., Knight, J., Bream, G., Stenlund, A., and Botchan, M. 1989, *Genes Dev.*, 3, 510.
65. Thain, A., Webster, K., Emery, D., Clarke, A.R., and Gaston, K. 1997, *J. Biol. Chem.*, 272, 8236.

66. Sanchez, I.E., Dellarole, M., Gaston, K., and de Prat-Gay, G. 2008, *Nucleic Acids Res.*, 36, 756.
67. Androphy, E.J., Lowy, D.R., and Schiller, J.T. 1987, *Nature*, 325, 70.
68. Cripe, T.P., Haugen, T.H., Turk, J.P., Tabatabai, F., Schmid, P.G., Durst, M., Gissmann, L., Roman, A., and Turek, L. 1987, *EMBO J.*, 6, 3745.
69. Stubenrauch, F., Lim, H.B., and Laimins, L.A. 1998, *J. Virol.*, 72, 1071.
70. Soeda, E., Ferran, M.C., Baker, C.C., and McBride, A. A. 2006, *Virology*, 351, 29.
71. Stubenrauch, F., Colbert, A.M., and Laimins, L.A. 1998, *J. Virol.* 72, 8115.
72. Ushikai, M., Lacey, M.J., Yamakawa, Y., Kono, M., Anson, J., Ishiji, T., Parkkinen, S., Wicker, N., Valentine, M.E., and Davidson, I. 1994, *J. Virol.*, 68, 6655.
73. Steger, G., Ham, J. and Yaniv, M. 1996, *Methods Enzymol.*, 274, 173.
74. Bouvard, V., Storey, A., Pim, D., and Banks, L. 1994, *EMBO J.*, 13, 5451.
75. Tan, S.H., Leong, L.E., Walker, P.A., and Bernard, H.U. 1994, *J. Virol.*, 68, 6411.
76. Demeret, C., Yaniv, M., and Thierry, F. 1994, *J. Virol.*, 68, 7075.
77. Steger, G., Ham, J., Lefebvre, O., and Yaniv, M. 1995, *EMBO J.*, 14, 329.
78. Dell, G., Wilkinson, K.W., Tranter, R., Parish, J., Leo, B.R., and Gaston, K. 2003, *J. Mol. Biol.*, 334, 979.
79. Alexander, K.A. and Phelps, W.C. 1996, *Biochemistry*, 35, 9864.
80. Fromental, C., Kanno, M., Nomiyama, H., and Chambon, P. 1988, *Cell*, 54, 943.
81. Ondek, B., Gloss, L., and Herr, W. 1988, *Nature (London)*, 333, 40.
82. Butz, K., and Hoppe-Seyler, F. 1993, *J. Virol.*, 67, 6476.
83. Angel, P., Hattori, K., Smeal, T., and Karin, M. 1988, *Cell*, 55, 875.
84. Angel, P., and Karin, M. 1991, *Biochem Biophys Acta*, 1072, 129.
85. Mehic, D., Bakiri, L., Ghannadan, M., Wagner, E.F., and Tschachler, E. 2005, *J. Invest. Dermatol.*, 124, 212.
86. Durst, M., Gallahan, D., Jay, G., and Rhim, J.S. 1989, *Virology*, 173, 767.
87. Mittal, R., Pater, A., and Pater, M.M. 1993, *J. Virol.*, 67, 5656.
88. Bromberg-White, J.L., and Meyers, C. 2002, *J. Virol.*, 76, 9702.
89. Chong, T., Apt, D., Gloss, B., Isa, M., and Bernard, H.U. 1991, *J. Virol.*, 65, 5933.
90. Lopez-Saavedra, A., González-Maya, L., Ponce-de-Leon, S., Garcia-Carranca, A., Mohar, A., and Lizano, M. 2009, *Arch. Virol.*, 154, 747.
91. Kossler, N., Stricker, S., Rödelsperger, C., Robinson, P.N., Kim, J., Dietrich, C., Osswald, M., Kühnisch, J., Stevenson, D.A., Braun, T., Mundlos, S., and Kolanczyk, M. 2011, *Hum. Mol. Genet.*, 20, 2697.
92. Patrakitkomjorn, S., Kobayashi, D., Morikawa, T., Wilson, M.M., Tsubota, N., Irie, A., Ozawa, T., Aoki, M., Arimura, N., Kaibuchi, K., Saya, H., and Araki, N. 2008, *J. Biol. Chem.*, 283, 9399.
93. Gordon, S., Akopyan, H., Garban, H., and Bonavida, B. 2006, *Oncogene*, 25, 1125.
94. Baritaki, S., Sifakis, S., Huerta-Yepez, S., Neonakis, I.K., Soufla, G., Bonavida, B., and Spandidos, D.A. 2007, *Int. J. Oncol.*, 31, 69.

95. Montalvo, E.A., Cottam, M., Hill, S., and Wang, Y.J. 1995, *J. Virol.*, 69, 4158.
96. Mills, L.K., Shi, Y., and Millette, R.L. 1994, *J. Virol.*, 68, 1234.
97. Bauknecht, T., Angel, P., Royer, H.D., and Zur Hausen, H. 1992, *EMBO J.*, 11, 4607.
98. Ralph, W.M., Jr Liu, K., and Auborn, K.J. 2006, *J. Gen. Virol.*, 87, 51.
99. Pajunk, H.S., May, C., Pfister, H., and Fuchs, P.G. 1997, *J. Gen. Virol.*, 78, 3287.
100. Lace, M.J., Yamakawa, Y., Ushikai, M., Anson, J.R., Haugen, T.H., and Turek, L.P. 2009, *J. Gen. Virol.*, 90, 2402.
101. Wang, J., and Manley, J. 1997, *Curr. Opin. Gen. Dev.*, 7, 205.
102. Shatkin, A.J., and Manley, J.L. 2000, *Nat. Struct. Biol.*, 7, 838.
103. Cullen, B.R. 2003, *J. Cell Sci.*, 116, 587.
104. Mitchell, P., and Tollervey, D. 2000, *Curr. Opin. Gen. Dev.*, 10, 193.
105. Proud, C.G. 2001, *Essays Biochem.*, 37, 97.
106. Graham, S.V. 2008, *Frontiers in Bioscience*, 13, 5646.
107. Baker, C., and Calef, C. 1996, *Maps of papillomavirus mRNA transcripts. The Human Papillomavirus Compendium. Los Alamos National Laboratory (ed.)*, iii-3.
108. Wang, X., Meyers, C., Wang, H.K., Chow, L.T., and Zheng, Z.M. 2011, *J. Virol.*, 85, 8080.
109. Schwarz, E., Freese, U.K., Gissmann, L., Mayer, W., Roggenbuck, B., Stremlau, A., and zur Hausen, H. 1985, *Nature*, 314, 111.
110. Zheng, Z.M., Tao, M., Yamanegi, K., Bodaghi, S., and Xiao, W. 2004, *J. Mol. Biol.*, 337, 1091.
111. Zheng, Z.M., and Baker, C.C. 2006, *Front. Biosci.*, 11, 2286.
112. Doorbar, J., Parton, A., Hartley, K., Banks, L., Crook, T., Stanley, M., and Crawford, L. 1990, *Virology*, 178, 254.
113. Sherman, L., and Alloul, N. 1992, *Virology*, 191, 953.
114. Rush, M., Zhao, X., and Schwartz, S. 2005, *J. Virol.*, 79, 12002.
115. Zhao, X., Rush, M., and Schwartz, S. 2004, *J. Virol.*, 78, 10888.
116. Cumming, S.A., Chuen-Im, T., Zhang, J., and Graham, S.V. 2009, *Virology*, 383, 142.
117. Sedman, S.A., Barbosa, M.S., Vass, W.C., Hubbert, N.L., Haas, J.A., Lowy, D.R., and Schiller, J.T. 1991, *J. Virol.*, 65, 4860.
118. Zhao, X., Oberg, D., Rush, M., Fay, J., Lambkin, H., and Schwartz, S. 2005, *J. Virol.*, 79, 4270.
119. Sherman, L., Alloul, N., Golan, I., Durst, M., and Baram, A. 1992, *Int. J. Cancer*, 50, 356.
120. Rosl, F., Arab, A., Klevenz, B., and zur, H. 1993, *J. Gen. Virol.*, 74, 791.
121. Thain, A., Jenkins, O., Clarke, A.R., and Gaston, K. 1996, *J. Virol.*, 70, 7233.
122. Kalantari, M., Calleja-Macias, I.E., Tewari, D., Hagmar, B., Lie, K., Barrera-Saldana, H.A., Wiley, D.J., and Bernard, H.U. 2004, *J. Virol.*, 78, 12762.
123. Badal, V., Chuang, L.S., Tan, E.H., Badal, S., Villa, L.L., Wheeler, C.M., Li, B.F., and Bernard, H.U. 2003, *J. Virol.*, 77, 6227.

124. Badal, S., Badal, V., Calleja-Macias, I.E., Kalantari, M., Chuang, L.S., Li, B.F., and Bernard, H.U. 2004, *Virology*, 324, 483.
125. Van Tine, B.A., Kappes, J.C., Banerjee, N.S., Knops, J., Lai, L., Steenbergen, R.D., Meijer, C.L., Snijders, P.J., Chatis, P., Broker, T.R., Moen, P.T., and Chow, L.T. 2004, *J.Virol.*, 78, 1172.
126. Ryan, K., Calvo, O., and Manley, J.L. 2004, *RNA*, 10, 565.
127. Kennedy, I.M., Haddow, J.K., and Clements, J.B. 1990, *J.Virol.*, 64, 1825.
128. Cumming, S.A., McPhillips, M.G., Veerapraditsin, T., Milligan, S.G., and Graham, S.V. 2003, *J.Virol.*, 77, 5167.
129. Dietrich-Goetz, W., Kennedy, I.M., Levins, B., Stanley, M.A., and Clements, J. B. 1997, *Proc. Natl. Acad. Sci.*, 94, 163.
130. Koffa, M.D., Graham, S.V., Takagaki, Y., Manley, J.L., and Clements, J.B. 2000, *Proc. Natl. Acad. Sci.*, 97, 4677.
131. McPhillips, M.G., Veerapraditsin, T., Cumming, S.A., Karali, D., Milligan, S.G., Boner, W., Morgan, I.M., and Graham, S.V. 2004, *J.Virol.*, 78, 10598.
132. Ko, B., and Gunderson, S.I. 2002, *J. Mol. Biol.*, 318, 1189.