

**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**

---

**POSGRADO EN CIENCIAS  
BIOLÓGICAS**

**FACULTAD DE CIENCIAS**

**SECUENCIA COMPLETA DEL GEN E2 DE DOS VARIANTES DE  
HPV16 PRESENTE EN LA POBLACIÓN MEXICANA Y ANÁLISIS  
FUNCIONAL DE SU EFECTO SOBRE LA TRANSCRIPCIÓN**

**T E S I S**

**QUE PARA OBTENER EL GRADO ACADEMICO DE  
MAESTRA EN CIENCIAS (BIOLOGIA CELULAR)**

**P R E S E N T A**

**MIRIAM CECILIA GUIDO JIMÉNEZ**

**DIRECTOR DE TESIS: DOCTOR M. ALEJANDRO GARCÍA CARRANCÁ**



Universidad Nacional  
Autónoma de México



**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.

**Esta tesis se realizó bajo la tutoría del Dr. Alejandro García Carrancá, en el Departamento de Biología Molecular del Instituto de Investigaciones Biomédicas UNAM y en colaboración con el Dr. Jaime Berumen en el Loyola Medical Center en Chicago Illinois USA.**

## DEDICATORIAS

A Vania Cecilia, mi aliento y mi razón de ser.

A mi abuelita... Te extraño todavía.  
A mi tia CECI... eres mi angel...

A mi Mamá: Amor, inteligencia,  
por enseñarme con el ejemplo,  
que el conocimiento es una manera de disfrutar la vida  
y que nunca se deja de aprender, te amo infinito....

A mi Papá: Inteligencia, sabiduría y perfección,  
por enseñarme a demostrar el amor con acciones,  
te amo profundamente.

Marguin: Hermana mayor,  
por ese gran impulso con amor, gracias ....

Tío José Luis: (Pepino) Hermano mayor,  
Por ese apoyo con amor incondicional.

Tia Conchis; amor de mi otra madre.

A Marilú, por su tezón, sabiduría  
Eres uno de mis mayores pilares.

A Migue Vero, Fran, Bolis, Vale, Jessi, y Alexa  
por ser tan hermosa familia.

A mis hermanos menores, Erik, Marisol, Mariel,  
Jorge Arturo y Jose Luis con orgullo...

A Morris, amigo por siempre..., que más hermano...

Alfredo, no tengo palabras..., por tu amor incondicional y tu apoyo... por lo que tanto trabajo nos costó ser..., por lo que somos ahora... y por lo mucho que seremos... TE AMO.

También dedico esta tesis a todos mis amigos y compañeros que han estado en el laboratorio, saben todo lo que los quiero, han sido mi familia durante todo este tiempo y he aprendido mucho de ustedes:

Para los pioneros: Cesar, Lupita, Mayra, Marce, Greco, Rosenda, Enrique, Marco, Ronit, Humberto, Lupita Bilvao, Amaranta, Gaby, Aileen, Verónica, Olga, Néstor, Manuel Ramos, Laura, Edith, Elsa, Carla, Benito, Paloma, Salomón, Jorge, Víctor, Emilio, Lety Rocha, Fabio

Por los que siguieron: Hilda, Ana Laura, Paty, Cecilia, Elisa, Rogelio, Mariana, Citlali, Jorge Aurelio, Paulina, Paula, Verónica, Ileana, Karina, Moisés, Ascari, Jorge

Por los nuevos: Cristian, Lety, Alexandra, Silvia, Luis, Brenda, Valerio, Alejandro, Martha, Dalia, Claudia Cirigo, Angeles, Esther, Oskar, Hugo, Adriana, Samantha y Rocio.

Para los actuales: Michel, Gaby Anaya, Gaby Toledo, Idalia, Salvador, Jose Luis, Jorge, Jaqueline y Manuel.

A Alex, gracias por enseñarme ciencia, darme la oportunidad de lograr mis metas y por estar en tu laboratorio disfrutando de mi trabajo. Eres un gran jefe, una gran persona y un gran amigo.

## AGRADECIMIENTOS

Agradezco a mis sinodales

Dra. Carmen Gómez. Dr. Alejandro García Carrancá, Dr Baltazar Becerril, Dr. Luis Felipe Jiménez y Dr. Vicente Madrid

Agradezco al Dr. Jaime Berumen por su asesoría durante el tiempo en que se desarrolló parte de este trabajo en el Loyola Medical Center en Chicago USA

Agradezco a la MC Patricia Padilla, MC Silvia Galvan Huerta y MC Rocio Susana Martínez por sus contribuciones a esta tesis.

A Guadalupe Peñaloza por su invaluable ayuda con la bibliografía de esta tesis y a Raquel López Paniagua por la ayuda durante la escritura de esta tesis en el Incan.

Agradezco a Eva Méndez, Beatriz Vera y la Sra. Julia el apoyo durante el desarrollo de esta tesis en IIB.

*Y no me cansare de agradecer enormemente al Dr. Efraín Garrido y al Biol. Cesar Bravo, gracias a ellos conocí la ciencia en toda la extensión de la palabra.*

Y finalmente agradezco al tiempo haberme dado la oportunidad de lograr esta meta...

## I. INDICE GENERAL

I.	Indice general	
II.	Resumen	
1.	<b>INTRODUCCION</b> .....	1
1.1.	Cáncer Cérvicouterino.....	1
1.2.	Biología del Papilomavirus Humano (VPH).....	4
1.3.	Variantes del VPH16.....	8
1.4.	Organización del genoma .....	12
1.4.1.	La región larga de control (LCR) y elementos que controlan la replicación y transcripción del virus.....	15
1.4.2.	La proteína E2 de VPH.....	16
1.4.2.1.	La proteína E2 y la replicación del virus.....	17
1.4.2.2.	Dominios y estructura de la proteína E2.....	19
1.4.2.3.	La proteína E2 y su secuencia blanco en el LCR .....	20
1.4.2.4.	Unión al DNA y dimerización.....	20
1.4.2.5.	Transactivación y represión por E2.....	22
1.4.2.6.	Dominio de activación de E2.....	23
1.4.2.7.	Interacción con proteínas celulares.....	24
2.	<b>OBJETIVOS</b> .....	25
3.	<b>HIPÓTESIS</b> .....	25
4.	<b>JUSTIFICACIÓN</b> .....	25
5.	<b>MATERIAL Y METODOS</b> .....	26
5.1.	Amplificación de E2 por PCR de tumores positivos para VPH16.....	27
5.2.	Secuenciación del gen E2.....	28
5.2.1.	Electroforesis del producto secuenciado.....	28
5.3.	Construcciones de DNA.....	28
5.4.	Ensayo de retardamiento ADN/proteína.....	30
5.4.1.	Transcripción-Traducción in vitro de la proteína E2.....	30
5.4.2.	Marcaje radioactivo del oligo.....	30
5.4.3.	Preparación de gel de poliacrilamida nativo.....	31
5.4.4.	Interacción ADN/proteína.....	31
5.5.	Transfección en células C33.....	31
5.5.1.	Preparación de las células.....	31
5.5.2.	Precipitación del ADN/fosfato de calcio.....	31
5.5.3.	Ensayo del gen reportero CAT.....	31
6.	<b>RESULTADOS</b> .....	33
6.1.	Secuenciación de los genes E2 de VPH16 AA-a, y AA-c.....	33
6.2.	Expresión de distintas proteínas E2 e interacción con el ADN.....	35
6.3.	Actividad basal de los LCRs.....	41
6.4.	Efecto de E2 sobre la actividad de los LCRs.....	44
7.	<b>DISCUSION</b> .....	48
8.	<b>CONCLUSIONES</b> .....	51
9.	<b>LISTA DE TABLAS Y FIGURAS</b> .....	52
10.	<b>REFERENCIAS</b> .....	54
11.	<b>ANEXO.</b> Artículo publicado.....	60

## RESUMEN

En el desarrollo de cáncer Cérvicouterino (CaCU) el principal factor implicado es la presencia del virus de papiloma Humano (VPH). Hasta ahora se han descrito cerca de 100 tipos de Papilomavirus algunos de ellos de alto riesgo (Ej. Tipos, 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68, 73 y 82). En México el más prevalente es el VPH16 seguido por el VPH18. Estos virus de ADN contienen genes tempranos (E), genes tardíos (L) y una región reguladora de la transcripción (LCR). Dos de los genes tempranos codifican para las proteínas oncogénicas del virus (E6 y E7) y para la proteína E2 que funciona como factor de transcripción regulando la expresión de E6 y E7. Los genes tardíos L1 y L2 codifican para proteínas de la cápside. Los Papilomavirus son clasificados como variantes moleculares cuando presentan un 2% de diferencia en la secuencia del gen L1 con respecto al papilomavirus prototipo. Las variantes Asiático-Americanas del papilomavirus tipo 16 están ligadas con una alta incidencia de CaCU en México, y se sugiere que son más oncogénicas que las variantes Europeas (E).

Durante el desarrollo de este trabajo se realizó y analizó la secuencia del gen E2 (1200pb) de dos variantes de VPH16 (AA-a y AA-c) presentes en biopsias de mujeres mexicanas, que desarrollaron CaCU, observándose que tienen 25 cambios en nucleótidos con respecto al VPH16 Europeo. Se analizó la capacidad de ésta proteína para unirse a su secuencia blanco por medio de geles de retardo, demostrándose que la proteína E2 europea se une más fuertemente a su secuencia blanco que la proteína E2 de las variantes AA-a y AA-c. También se analizó la actividad reguladora de la LCR de estas dos variantes (usando como gen reportero a la enzima CAT: Cloranfenicol acetil-transferasa) observándose que la actividad basal de la LCR de la variante AA-c es 1.8 veces más alta que la de la LCR VPH16 Europeo, y la actividad de la variante AA-a es 1.6 más baja que la LCR de VPH16 E. De la misma manera se observó que las proteínas E2 de las variantes no están reprimiendo sus propios promotores a la misma magnitud que la proteína E2 de VPH16 E. Estos resultados nos sugieren que tanto los LCRs como las proteínas E2 de las variantes no reprimen la expresión de los oncogenes E6 y E7 como en VPH16 Europeo, lo que nos puede ayudar a concluir que tal vez es la causa por las que las lesiones con estas variantes son más agresivas.

## 1. INTRODUCCION

### 1.1 Cáncer Cérvicouterino.

El cáncer Cervicouterino (CaCU) es el segundo cáncer más común en mujeres en el mundo, especialmente en países en desarrollo donde los métodos de diagnóstico rutinarios como el Papanicolaou no están realmente disponibles a toda la población y la calidad del servicio es aun deficiente (Lazcano-Ponce et al., 2003). Se estima que cada año se diagnostican 500,000 nuevos casos de cáncer del cérvix en el mundo, de los cuales mueren el 45%. En México en el año 2002, la tasa de defunción por CaCU fue de 4323 defunciones de mujeres por cada 100,000 (Información estadística Secretaria de Salud). En los países desarrollados, debido a la aplicación masiva de campañas de detección oportuna del CaCU, la realización del Papanicolaou y a los avances terapéuticos en esta área, ha disminuido notablemente la tasa de mortalidad de este padecimiento en las últimas dos décadas (Nelson et al., 1984) En México representa el 20% de todos los cánceres y el 31% de los cánceres que afectan a la población femenina, ocupando el primer lugar en incidencia y mortalidad. Entre otros países de Latinoamérica, México ocupa el tercer lugar en la tasa de mortalidad por CaCU. Se presentan 3,500 casos nuevos por año, la tasa de mortalidad por cáncer del cérvix es de 11 por 100,000 mujeres y se calcula que la tasa de incidencia en mujeres mayores de 25 años es de 115 por 100,000 (Dirección General de Epidemiología, Secretaría de Salud, 1991; Instituto Nacional de Salud Publica, 1993).

En mujeres con cáncer cervical, las células malignas se desarrollan de tejido citológicamente anormal, que puede ser desde displasia media (neoplasia intraepitelial cervical grado I; NIC1), a displasia severa moderada (NIC2 – 3). Las células o koilocitos comunes en tejidos infectados con Papilomavirus Humanos (VPH) caen dentro del espectro de NIC1 y ocasionalmente en NIC2. Las mujeres con anomalías indicativas de infección por Papilomavirus Humanos (VPH) tienen más alta probabilidad de contraer cáncer cervical que las que tienen citología normal (Koutsky, 1997). Consistente con esto, el PCR y la captura de híbridos, métodos muy sensibles para detección de ADN de VPH, identifican un 93% de ADN de VPH en cánceres cervicales. Estos y otros datos nos proveen de un argumento convincente para la asociación etiológica entre infección de VPH y cáncer cervical.

Hasta la fecha se han identificado cerca de 100 tipos diferentes de VPH de los cuales 80 se han secuenciado. De todos ellos, aproximadamente 20 tipos se han encontrado asociados a tumores de la región genital. Los tipos virales más comunes en las lesiones genitales benignas son el VPH 6 y 11 que se encuentran en el 90-100% de las lesiones benignas de la región genital (condiloma acuminado y condiloma plano) y en las lesiones neoplásicas precursoras y el cáncer invasor los tipos 16 y 18, presentes en el 10-20% de las lesiones precancerosas tempranas del cuello uterino (Condiloma atípico y NIC I) y en lesiones avanzadas (NIC 2/3 y cáncer invasor). El 99.7% de carcinomas cervicales humanos contienen ADN de VPH de alto riesgo (Kreider et al., 1985; Walboomers et al., 1999)

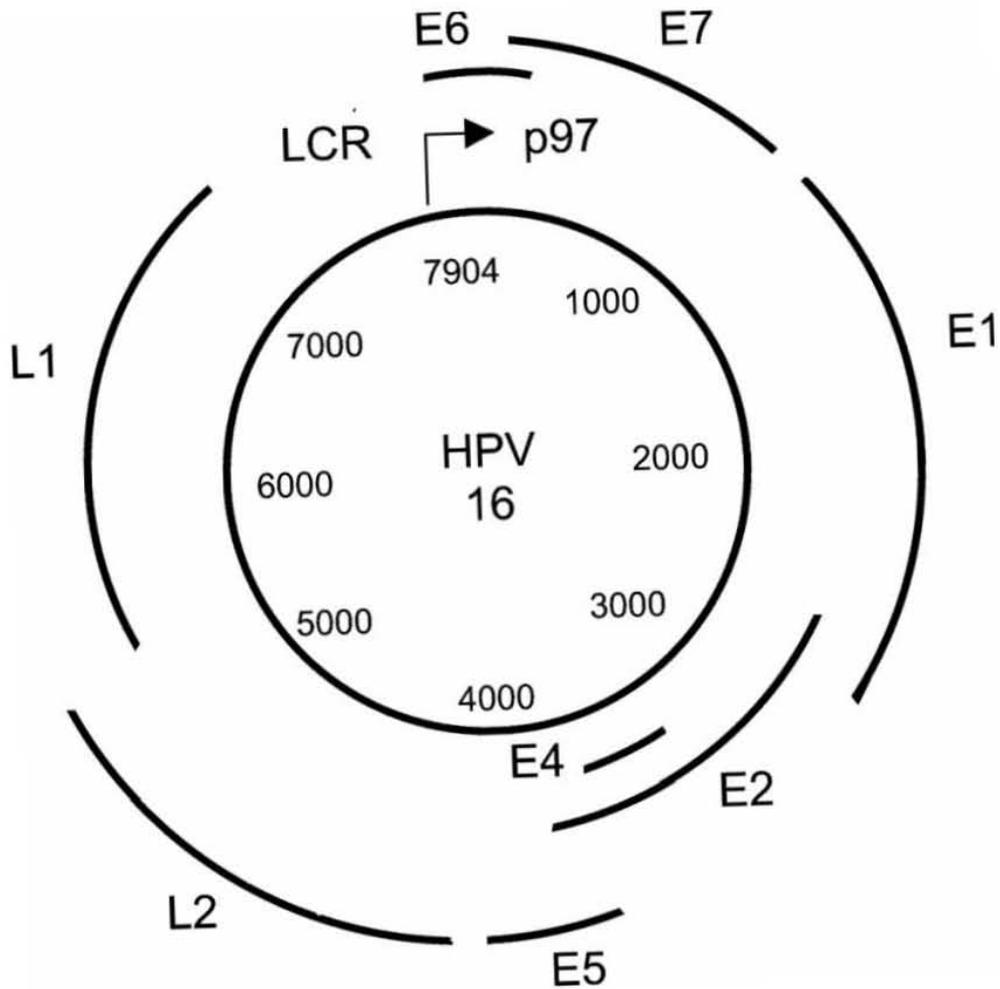
La detección y purificación de este virus de las células del CaCU, cumple con los postulados de Koch para asignar etiologicamente a un agente infeccioso con una enfermedad, se han obtenido

resultados parciales exponiendo epitelios de cérvix sanos a VPH 11 injertados debajo de la cápsula renal de ratones atímicos, observándose que los epitelios se transformaron en displásicos (Kreider et al., 1985; Kryszke et al., 1987). Además, mediante experimentos de transfección con el VPH16, se ha logrado la transformación maligna *in vitro* de fibroblastos y queratinocitos humanos (Pirisi et al., 1987; Yasumoto et al., 1986). Por otro lado se ha demostrado que el BPV-1 (Papilomavirus bovino tipo 1), clonado en el vector pBR322, es capaz de transformar células de ratón susceptibles, y a diferencia de otros virus que inducen tumores, BPV-1 no requiere la integración del genoma viral (Heilman et al., 1982).

Cabe mencionar que la investigación con VPH's se ha visto beneficiada principalmente por las tecnologías de ADN recombinante en virología, principalmente porque los VPH's no podían ser crecidos en cultivo o transmitidos a otras especies animales, por lo que con solo esta herramienta se debía de buscar información acerca de los mecanismos involucrados en la carcinogénesis. Existen ahora algunos métodos para propagarlo en cultivo de células en el laboratorio, que involucran la exposición de piezas de tejido susceptibles (como epitelio cervical o piel) a virus y después transplantando el tejido en la cápsula renal de ratón atímico desnudo. Después de un periodo de 3 – 6 meses, se empiezan a formar las verrugas y se puede detectar el virus infeccioso, de esta manera se ha logrado obtener el papilomavirus humano tipo 11 (VPH 11). Se ha demostrado que transfectando queratinocitos humanos normales con la secuencia completa de VPH18 y VPH31, se obtienen líneas celulares que mantienen los genomas virales como elementos extracromosomales (Frattini et al., 1996), en el caso de VPH18 se observo que todas las líneas celulares contenían aproximadamente de 100 a 200 copias de VPH18 por célula, mediante crecimiento organotípico (raft) de estas líneas celulares, se observa la expresión de los genes virales dependientes de la diferenciación celular, la amplificación de los genomas virales y la biosíntesis de los viriones (Frattini et al., 1997).

En VPH's que infectan genitales se estima que el periodo de incubación es de aproximadamente tres meses con un rango de infectividad de aproximadamente el 65%. (Park et al., 1995). Aunque puede existir un periodo de latencia de entre 10 y 25 años en lesiones preinvasoras para que evolucione a CaCU ((Nelson et al., 1984)

Existen otros factores asociados al cáncer del cérvix, se ha observado que la presencia de VPH por si sola no ha podido explicar algunos de los hechos relacionados con la iniciación y progresión del tumor y parece no ser suficiente para producir el CaCU. Esto se puede demostrar porque solo un porcentaje de las neoplasias cervicales preinvasoras (cerca del 10%) y que presentan VPH evolucionan a CaCU (Pirisi et al., 1987; Reid et al., 1984). Por otro lado existen estudios epidemiológicos que relacionan otros factores etiológicos con la génesis del CaCU como son las hormonas sexuales, la dieta, infecciones crónicas del cérvix, el virus del Hérpes y el estado del sistema inmunológico (Halpert et al., 1986) los cuales pudieran estar participando en alguna de las etapas del desarrollo del tumor.



**Figura 1. Mapa genómico del Papilomavirus Humano tipo 16 (VPH16).** El genoma de DNA es de doble cadena, circular. Posee 6 regiones de expresión temprana (E) y dos regiones de expresión tardía (L). Entre el final de la región tardía y el principio de la región temprana se encuentra la LCR (Región larga de control) que contiene los sitios de reconocimiento a factores celulares y virales que regulan la transcripción y el origen de replicación (Crosbie & Kitchener, 2006).

En hombres la infección de VPH está asociada con la falta de uso del condón, en mujeres aun no es claro aunque hay estudios que apoyan la idea de que el condón también tiene un efecto protector. En cuanto al uso de anticonceptivos orales, existen estudios que sugieren que este tipo de hormonas incrementan el riesgo de desarrollar cáncer cervical. En un estudio, se demostró que el uso de hormonas no tuvo efecto en la incidencia de lesiones intraepiteliales escamosas de bajo grado, aunque incrementó el riesgo de lesiones intraepiteliales de alto grado. Este hecho sugirió que las hormonas anticonceptivas actúan alterando la progresión de la enfermedad en individuos infectados con VPH, más que incrementar la infección por VPH (Koutsky, 1997).

## 1.2 Biología del Papilomavirus Humano (VPH)

Los papilomavirus infectan las células basales del epitelio estratificado y establecen sus genomas como episomas con bajo número de copias, en esas células los episomas se replican en sincronía con los cromosomas celulares.

La síntesis vegetativa del ADN viral empieza en las células suprabasales de capa espinosa y es seguido por la expresión tardía de los genes, es decir, de los polipéptidos estructurales del virus, y el ensamble de los viriones ocurre casi exclusivamente dentro del núcleo de la capa de células granulares. La expresión fenotípica tanto de papilomavirus cutáneos como genitales es similar. La infección por VPH se manifiesta por proliferación de las células espinosas (acantosis), vacuolización degenerativa citoplásmica (koilocitosis), alteraciones nucleares (arrugamiento y piknosis) de las células de la capa granular y la producción excesiva de queratina. (Park et al., 1995).

Los papilomavirus pertenecen a la Familia: Papillomaviridae, Genero: papillomavirus, Especie; *Papillomavirus humanos*, poseen una cadena doble de ADN circular covalente de aproximadamente 8000 pb (Figura 1), la cual está empaquetada en una cápside icosaédrica de 55 nm de diámetro, con 72 capsómeros, su transcripción es en un solo sentido (Munger, 2002). Los papilomavirus están distribuidos en la naturaleza, infectan humanos y animales (bovinos, caninos, conejos, delfines, y venados). Aunque los papilomavirus humanos y animales poseen una organización genómica similar, ellos son altamente especie específicos y no infectan otras especies (Park et al., 1995). Los tipos de epitelio susceptible y las lesiones que provocan entre los diferentes papilomavirus difieren dando lugar a distintos síntomas clínicos dependiendo del virus (Tabla 1). Las proteínas de la cápside son antigénicamente muy similares entre todos los papilomavirus, estudios previos indican que la mayoría de los genotipos genitales representan serotipos separados, de la neutralización cruzada observada entre los VPHs de alto riesgo, se puede sugerir que la clasificación de serotipos es similar a la de los genotipos (Combita et al., 2002). Se define un tipo de VPH nuevo cuando la secuencia del gen L1 es 10% diferente de cualquier otro tipo de VPH (de Villiers et al., 2004). Un subtipo se define cuando la secuencia de L1 es entre 2% -10% diferente del tipo al que pertenece lo que la coloca entre los tipos y variantes, y una variante se define cuando la secuencia de L1 es menor al 2% diferente del tipo al que pertenece. Existen numerosas variantes de cada tipo, mientras que los subtipos son extraordinariamente raros, la razón de esto aun no

**Tabla 1. Tipos de Papilomavirus Humanos, y lesiones de las cuales fueron aislados.**

<b>CUTANEOS NO GENITALES</b>	<b>TIPO DE VPH</b>
Verrugas comunes (verrucae vulgaris)	1, 2, 4, 26, 27, 29, 41, 57, 65
Verrugas plantares (myrmecias)	1, 2, 4, 63
Verrugas planas (verrucae plana)	3, 10, 27, 28, 38, 41, 49
Verrugas de Butcher (verrugas comunes de gente que maneja carne cruda y pescado).	1, 2, 3, 4, 7, 10, 28
Verrugas de mosaico	2, 27, 57
Carcinoma de células escamosas Ungual	<b>16</b>
Epidermodysplasia verruciformis (benigno)	2, 3, 10, 12, 15, 19, 36, 46, 47, 50
Epidermodysplasia verruciformis (maligno o benigno)	5, 8, 9, 10, 14, 17, 20, 21, 22, 23, 24, 25, 37, 38
Lesiones de la piel no verrugosas	37, 38

**ENFERMEDADES MUCOSAS NO GENITALES**

Papilomatosis respiratoria	6 11
Carcinoma de células escamosas de pulmón	6, 11, <b>16, 18</b>
Papiloma laringeo	6, 11, 30
Carcinoma laringeo	<b>16, 18</b>
Papiloma del seno maxilar	57
Carcinoma de células escamosas de las sinusoides	<b>16, 18</b>
Papilomas de la conjuntiva	6, 11
Carcinoma de la conjuntiva	<b>16</b>
Hiperplasia Epitelial focal oral (enfermedad Heck)	13, 32
Carcinoma oral	<b>16, 18</b>
Leucoplakia Oral	<b>16, 18</b>
Carcinoma celular escamoso del esófago	<b>16, 18</b>

**ENFERMEDADES ANOGENITALES**

Condylomata acuminata	6, 11, 30, 42, 43, 44, 45, 51, 52, 54
Papulosis Bowenoides	<b>16, 18</b> , 34, 39, 42, 45
Enfermedad Bowenoides	<b>16, 18</b> , 31, 34
Condiloma gigante (Tumores Buschke-Löwenstein)	6, 11
Neoplasia intraepitelial no especificada	30, 34, 39, 40, 53, 57, 59, 61, 62, 64, 66, 67, 68, 69
Neoplasia intraepitelial de bajo grado	6, 11, 43
Neoplasia intraepitelial intermedia	31, 33, 35, 42, 44, 45, 51, 52
Neoplasia intraepitelial de alto grado	<b>16, 18</b> , 56, 58
Carcinoma de vulva	6, 11, <b>16, 18</b>
Carcinoma de vagina	<b>16</b>
Carcinoma de cervix	<b>16, 18</b> , 31
Carcinoma de ano	16, 31, 32, 33
Carcinoma <i>in situ</i> de pene (eritroplasia de Queyrat)	16
Carcinoma de pene	16, 18

es entendida (Calleja-Macias et al., 2005).

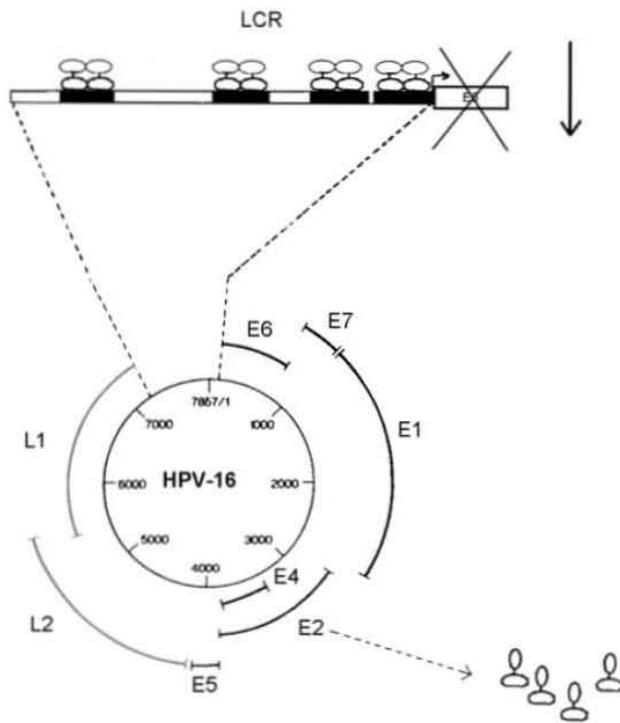
Las proteínas virales estructurales constituyen cerca del 80% de la masa del virión. Su ADN está asociado a cuatro polipéptidos que son similares a histonas celulares. Los estudios en ADN de VPH son dependientes de la extracción de ADN no integrado superhelicoidal de viriones purificados o directamente de lesiones asociadas a VPH, también depende del tipo de VPH (Park et al., 1995).

Todos los tipos de VPH conocidos son epiteliotrópicos, infectan tipos específicos de epitelios y frecuentemente producen proliferaciones epiteliales focales en los sitios de infección. Los VPH cutáneos tipos 1, 2, 4, 26, 27, 41 y 57 son los más comunes identificados en verrugas de manos o en manejadores de carne. Se han encontrado 23 tipos diferentes que infectan epitelios del tracto anogenital femenino y masculino entre los que se encuentran los tipos 6, 11, 16, 18, 30, 31, 33, 35, 39, 40, 42-45, 51-58 y 61, los cuales causan desde Condyloma acuminata hasta carcinoma celular escamoso (Ver Tabla 1). Los VPHs tipos 6, 11, 16 y 35 son los más comúnmente encontrados en mujeres sin enfermedades vulvares o cervicales. El VPH6 es el más comúnmente encontrado en mujeres u hombres con condilomas exofíticos del tracto anogenital, éstos tipos de lesiones no están casi nunca asociados al VPH18 o con VPH1 o VPH2 (Park et al., 1995).

Con el uso de métodos moleculares sensibles como es la reacción en cadena de la polimerasa (PCR) o captura de híbridos, el VPH ha sido detectado en más del 80% de las mujeres que tienen lesiones cervicales que histológicamente son de bajo riesgo. Mientras que el VPH16 y 18 se han encontrado frecuentemente asociados con lesiones de cáncer invasivo cervical. Basados en su asociación con lesiones de alto grado de la vulva, vagina o cérvix o con lesiones epiteliales benignas, los virus se han clasificado en tres categorías diferentes dependiendo de su riesgo oncogénico. El primer grupo es llamado de "bajo riesgo oncogénico" a él pertenecen los VPHs tipos 6, 11, 42, 43 y 44, estos virus están asociados con *C. acuminata* y raramente se encuentran asociados a cáncer cervical. El segundo grupo es llamado de "alto riesgo oncogénico", a él pertenecen los VPHs tipos 16, 18, 31 y se asocian con lesiones de cáncer invasivo del cérvix, vulva, pene y ano. El tercer grupo de VPHs es llamado de "riesgo intermedio oncogénico", y están más comúnmente asociados con lesiones de alto grado oncogénico pero no están asociados con cáncer invasivo anogenital, y son los VPHs tipos 33, 35, 39, 51 y 52 (Park et al., 1995).

En lesiones benignas cervicales, el ADN viral es mantenido de forma extracromosomal circular llamada "episoma". Sin embargo en muchos cánceres cervicales, en líneas celulares derivadas de ellos y en queratinocitos transformados *in vitro* con VPH, el ADN se encuentra integrado en el genoma celular. Berumen y cols. (Berumen et al., 1994) encontraron que el número de copias del virus por célula, fue significativamente más bajo en tumores positivos a VPH18, que en tumores positivos a VPH16, y la región E1/E2 estuvo ausente en todos los tumores positivos para VPH18 y presente en más de la mitad de los tumores positivos a VPH16, lo que sugiere que VPH18 se encontró integrado mientras que VPH16 se encontró en estado episomal. En un estudio llevado a cabo por Cullen y cols. (Cullen et al., 1991) se demostró que la integración ocurre más frecuentemente en carcinomas asociados con VPH18 que en

**Figura 2. Mapa genómico del VPH 16 y esquema de regulación de la transcripción por E2.**



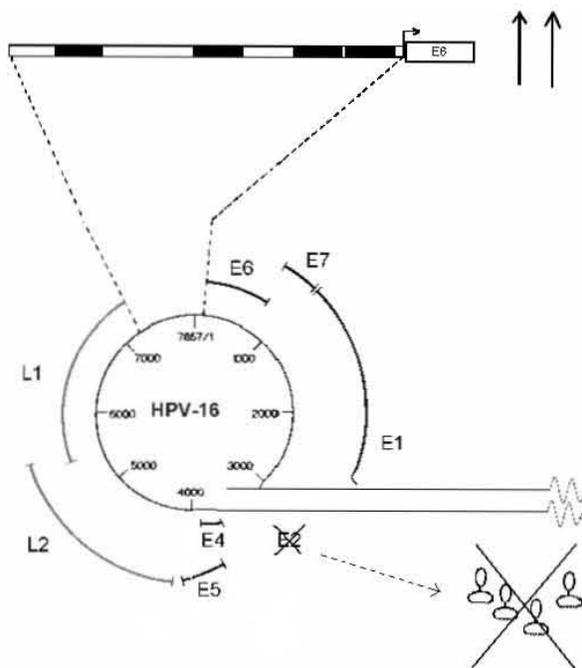
a) La proteína E2 reprime la transcripción de los oncogenes virales.



Dímero de la proteína E2



Sitio de unión a E2: ACCG N4 CGGT



b) La integración del virus al DNA hospedero comúnmente rompe la región E1/E2 con la consecuente pérdida de estos genes, permitiendo la transcripción continua de los oncogenes E6 y E7.



Proteína E2

- Sitio de unión a E2: ACCG N4 CGGT

aquellos asociados con VPH16, en algunos casos encontraron tanto VPH16 episomal como VPH18 integrado en la misma muestra de tejido. Los datos de la frecuencia de integración de VPH en ADN genómico hospedero son poco consistentes (Park et al., 1995).

Se ha demostrado que la integración tiene un efecto importante en la expresión de los genes virales y en el comportamiento biológico de las células infectadas con VPH. El acto físico de la integración comúnmente interrumpe la región E1/E2 del VPH, manteniendo intacta la región que codifica para los oncogenes E6/E7 (Figura 2a y b). Se sabe que la proteína E2 actúa como un supresor de la transcripción de los oncogenes E6 y E7, en virus de alto riesgo (Guido et al., 1992) la disrupción del gen E2 y la consecuente pérdida de la proteína que codifica, da lugar a la transcripción incontrolada de E6 y E7 (Park et al., 1995).

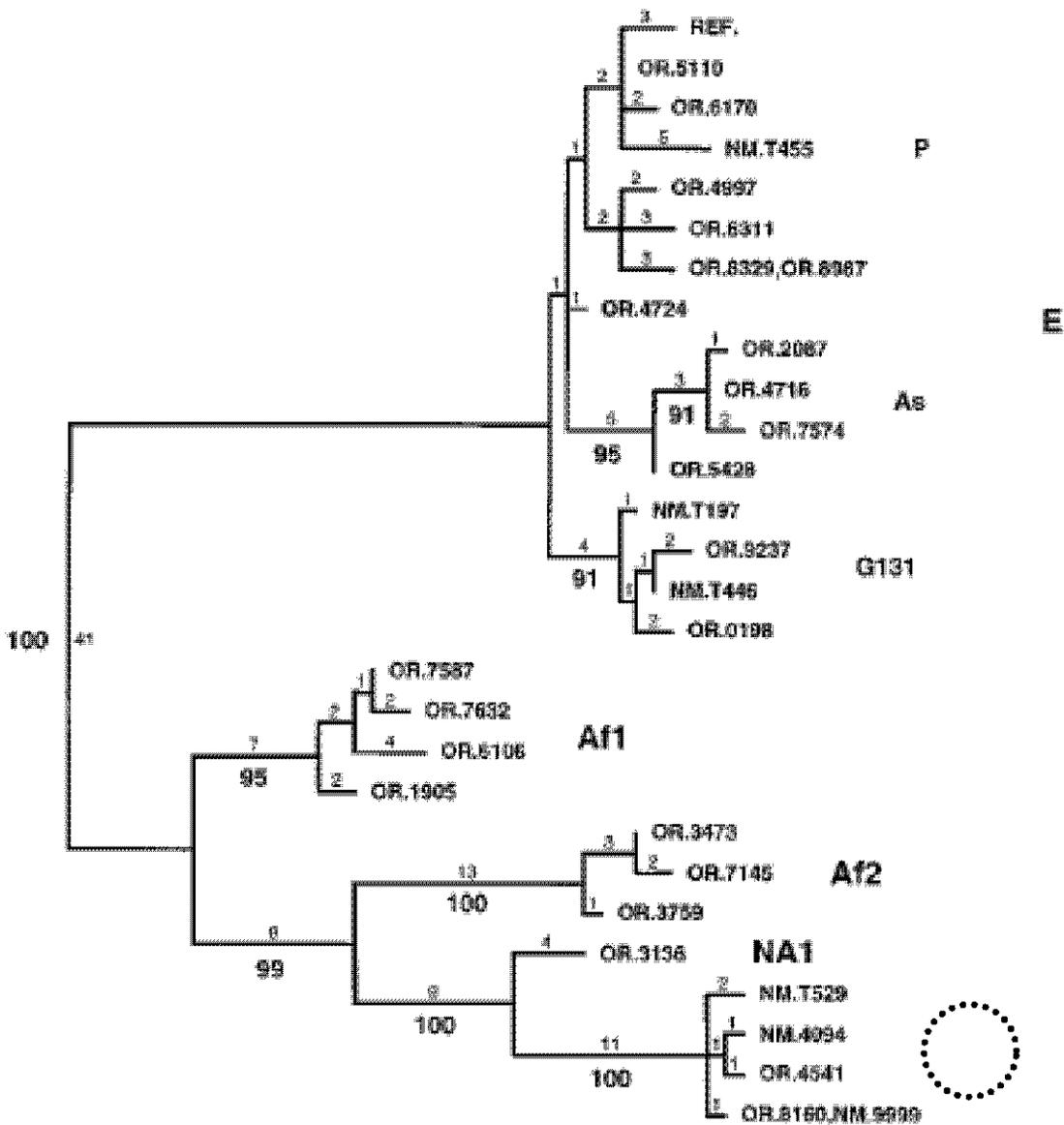
El sitio en el cual se integra el VPH en el ADN cromosomal hospedero es muy variable, y no se ha identificado un locus preferencial de integración. Sin embargo en algunas líneas celulares de tumores que contienen VPH y en queratinocitos de cultivos primarios de humanos, el ADN viral se ha mapeado en o dentro de sitios frágiles, que son regiones cromosomales particularmente susceptibles a la integración de ADN extraño. También se ha detectado integración de ADN de VPH cerca de oncogenes celulares, como son *c-myc* y *N-myc*. La interrupción y rearreglo de estas regiones puede resultar en inestabilidad genómica o activación de oncogenes celulares y promover la progresión celular. Sin embargo la integración del ADN viral no parece ser un requisito importante en el desarrollo de muchos carcinomas cervicales, ya que existen algunos tumores que contienen solamente ADN episomal (Park et al., 1995).

### **1.3 Variantes del VPH16**

Todos los mamíferos y probablemente muchos otros vertebrados contienen papilomavirus especie-específicos, y aun dentro de hospederos individuales como humanos o ciertos ungulados, existe un gran número de tipos de virus diferentes y poco relacionados que dan lugar a diferentes patologías. Consecuentemente la evolución puede ser estudiada en correlación con diferentes variables biológicas como es la especie del hospedero o por el tropismo preferencial de muchos tipos de VPHs por un tipo de epitelio o de célula. La relación que existe entre los diferentes tipos de papilomavirus puede ser evaluada filogenéticamente ya que sus genomas codifican para proteínas con varios dominios altamente conservados que pueden ser alineados sin ambigüedad (Ho et al., 1993).

La filogenia inferida por Ho y cols. (1993) se basó en el análisis de la secuencia de nucleótidos de la región larga de control (LCR) de VPH16, un segmento no codificante que regula la transcripción y contiene el origen de replicación del virus.

En un nivel de evolución más reciente, los aislados de papilomavirus tipo 16 y probablemente todos los demás tipos de papilomavirus humanos pueden ser distinguidos a través de mutaciones



**FIGURA 3: Arbol filogenético de las variantes de VP16 basado en análisis de parsimonia.** Las secuencias combinadas de E6, L2, L1 y LCR de 30 aislados (incluyendo el VP16 de referencia), dieron lugar a alineamiento de 129 posiciones variables. Las porciones del árbol correspondientes a las clases variantes E, Af1, Af2, NA1 y AA están indicadas por marcas largas. Y los linajes E de subclases P, As y G131 están indicados con marcas pequeñas (Yamada et al., 1995).

puntuales. La diversidad de secuencia entre aislados en la región larga de control (LCR) puede ser de 5% mientras que esta no exceda el 2% dentro de los demás genes. Los eventos de mutación en la LCR son raros, ya que los patrones de mutación de aislados de VPH16 de Brasil, Alemania, Singapur y Tanzania, forman grupos geográficos, y todos los aislados, aun de lugares remotos uno de otro, están relacionados en la forma de un árbol evolutivo. Estos aislados o variantes contienen un detallado récord molecular de la evolución del VPH16 (Ho et al., 1993).

Estudios de heterogeneidad genómica de VPH16 han demostrado la presencia de múltiples formas variantes en todas las poblaciones examinadas hasta ahora. Es concebible que las variantes naturales de VPH16 en una población humana, pudieran tener diferentes comportamientos biológicos (Xi et al., 1997).

Los aislados de VPH históricamente se clasificaron en tipos, comparando su ADN con un set de genomas de VPHs de referencia mediante hibridación del ADN. Los aislados que pertenecen al mismo tipo de VPH están muy cercanamente relacionados con respecto a la secuencia de bases de sus nucleótidos y representan un discreto grupo filogenético. Aunque todos los aislados de VPH16 están cercanamente relacionados, investigaciones de Ho et al, en 1993, comparando las variantes de VPH16 de diferentes grupos étnicos y localizaciones geográficas, identificó cinco grandes brazos filogenéticos. Los brazos se designaron como E (Europeo), As (Asiático), AA (Asiático-Americano), Af1 (Africano-1) y Af2 (Africano-2). El genoma de VPH16 de referencia, denominado G11, es un aislado Alemán y es miembro del linaje E. La Figura 3 muestra el árbol filogenético de las variantes de VPH16 inferidas por Yamada y cols. (1995).

Mediante el análisis comparativo de la LCR, los brazos E y As son distintos pero muy cercanamente relacionados, pareciera ser que As es una subclase del linaje E. AA, Af1 y Af2 están muy separados entre sí y de los brazos E y As. La mayoría de ellos aunque no todos, reflejan la etnicidad del grupo en el cual vive el paciente infectado, existe una desviación de esta correlación que se le atribuye quizá a las migraciones de grupos humanos infectados.

Ho y cols (1993), proponen que los genomas de VPH16 actuales representan un grado de diversidad que evolucionó en un periodo largo de tiempo, que probablemente excede los 200,000 años, derivado de genomas precursores que pudieron haberse originado en África. Ellos encontraron que para VPH16 y VPH18 existe una topología del árbol filogenético similar a la del árbol evolutivo de las tres principales razas humanas, Africanos, Caucásicos y Asiáticos del Este, esto nos dice que la coevolución de los humanos y papilomavirus es más probable que la diseminación de VPH16 puramente por geografía, sin embargo esto no excluye la idea de la transmisión reciente, entre miembros de grupos étnicos diferentes, tal es el caso de variantes del brazo Europeo en algunos pacientes negros en los Estados Unidos y ocasionalmente en África (Ho et al., 1993).

En un estudio de mujeres jóvenes, Xi y cols. (1997), encontraron 16 diferentes variantes de VPH16. Una de las variantes se encontró persistentemente durante un lapso largo de tiempo, mientras que otras variantes se detectaron transitoriamente.

Se han encontrado variaciones en la secuencia de ADN entre aislados de VPH16 y VPH18 en varias localidades geográficas y grupos étnicos. En la población Mexicana se encontraron dos variantes de VPH16 antes descritas, Bb-2 e IND-8, las cuales pertenecen a un grupo de variantes Brasileñas e Indias respectivamente y son consideradas representantes de una línea evolutiva del VPH16 esta variante se ha agrupado junto con otras dentro del brazo Asiático-Americano. Se ha sugerido que la diversidad viral correlaciona con las características étnicas de la población más que con la geográfica. En México los grupos étnicos mezclados están enriquecidos mayoritariamente por Amerindios más que por Caucásicos o Africanos. Es de particular interés, que además de la clona de referencia, la única variante de VPH16 detectada en la población Mexicana se haya encontrado previamente en Brasil (IND-8), en una población de origen Amerindio. Esta variante se encontró presente en tumores de origen escamoso y en mujeres más jóvenes (40 años o menos) con respecto a la clona de referencia de VPH16 en el que las pacientes tienen más de 40 años de edad. Esto sugiere que la variante encontrada en México en mujeres con características Amerindias, puede ser más agresiva y de alto riesgo. Esta posibilidad podría explicar en parte la alta incidencia de carcinoma cervical observada en las mujeres Latino-Americanas (Lizano et al., 1997). El adenocarcinoma, el carcinoma adenoescamoso y el carcinoma de células pequeñas, son reportados como de baja incidencia, pero clínicamente importantes, exhiben un comportamiento biológico más agresivo y tienen un pronóstico más pobre comparados con carcinoma escamoso. El VPH 18 se ha asociado predominantemente con adenocarcinomas y carcinomas adenoescamosos, se ha observado que existe una asociación exclusiva entre VPH18 var-2 con carcinoma escamoso, y de VPH18 var-1 al igual que VPH8 referencia con adenocarcinoma (Lizano et al., 1997).

El VPH tipo 16 se asocia principalmente con los carcinomas de células escamosas (Bosch et al., 1995), actualmente se tiene evidencia que sugiere que sus variantes están asociadas a distintos tipos histológicos de cáncer y que tienen un distinto comportamiento biológico. El riesgo de desarrollar una lesión no es el mismo para todas las variantes del VPH16. En un estudio en el que se detectaron variantes a través de la región del LCR se llegó a la conclusión que el riesgo de desarrollar una NIC II-III no es el mismo para todas las variantes, sino que aquellas no parecidas al prototipo (As, AA y Af) tienen un riesgo más elevado (Xi et al., 1997). De manera similar en tumores de cuello uterino era más predominante la presencia de las variantes diferentes al prototipo (Zehbe et al., 1998). Además se ha reportado que una variante del VPH16 encontrada en los tumores de mujeres mexicanas y que pertenece a la rama AA pudiera presentar un comportamiento más agresivo, ya que se encontró en un grupo de mujeres más jóvenes que la prototipo (Lizano et al., 1997). Casi un cuarto de todos los CaCU en México contienen las variantes Asiática-Americana (AA) de VPH16 y principalmente las subclases AA-a y AA-c, y se han asociado con mujeres 8 años más jóvenes y con tumores más agresivos comparada con la variante Europea (Berumen et al., 2001).

En un estudio que se llevó a cabo con mujeres Mexicanas se reportó la secuencia de los genes L1 y E6 de 76 aislados de ADN de cáncer Cérvico-uterino positivos para VPH16, de las cuales 33 (43.4%) muestras presentaron la variante AA, 42 (55.3%) la variante E y solo una de ellas la variante NA (Norte-Americana). La subclase AA-c es la más frecuente. La edad promedio de las pacientes positivas para VPH16 es de 51 años, para las AA de 47 años y para la E de 53 años. Es de interés recalcar que para las subclases del linaje AA, se encontró que la edad promedio más baja fue para AA-c de 42 años. Además de que se encontró en adenocarcinomas, lo que nos pudiera estar hablando de un comportamiento más agresivo (Casas et al., 1999).

Se sugiere que la alta oncogenicidad de las variantes de VPH16 puede deberse a tres razones; incrementada actividad de los oncogenes E6/E7, replicación viral mas eficiente o mejor estimulación de la expresión de los oncogenes virales. La proteína E6 de la variante AA-a (aislado 512) produce más suero y colonias resistentes a diferenciación por calcio en cultivos primarios de queratinocitos de piel y estimulan la degradación de p53 mejor que las proteína E6 de VPH16 Europeo (Stoppler et al., 1996). Por otro lado el número de partículas virales de ADN de las variantes AA es significativamente más alta que la de las variantes E en carcinomas cervicales, lo que sugiere que las variantes AA se replican mejor (Casas et al., 1999).

#### 1.4 Organización del genoma viral

Aunque existe muy poca similitud entre algunos genomas de papilomavirus de animales y humanos entre ellos tienen muchas propiedades en común lo cual valida la utilización de BPV1 como modelo. Todos los papilomavirus presentan una organización genómica que consiste al menos de 10 secuencias que codifican para proteínas o marcos de lectura abierta (MLA) en una sola de las cadenas de ADN. El genoma de papilomavirus se divide en dos regiones codificantes separadas por una región no codificante de entre 400 y 1000 pb. La región temprana (E) contiene los genes de E1 a E8, incluyendo la información necesaria para la replicación y la transformación celular y una región tardía (L) que contiene los genes L1 y L2 los cuales son expresados únicamente en queratinocitos diferenciados y codifican para las proteínas de la cápside (Sousa et al., 1990) En la Figura 1 se esquematiza la organización genómica del papilomavirus humano tipo 16 (VPH16). A continuación se describe cada uno de los genes de los Papilomavirus Humanos.

**Gen E1:** Codifica para una proteína de entre 68-76 kD esencial para la replicación del ADN. El producto E1 completo es una proteína fosforilada nuclear que se une al origen de replicación en el LCR (Región Larga de Control). Se ha demostrado que une ATP y se une *in vitro* a la proteína E2 la cual es transactivadora de la transcripción. La unión a E2 aumenta la afinidad de E1 para unirse al origen de replicación del ADN. En VPH16 E1 tiene efectos indirectos en la inmortalización.(Conger et al., 1999). La proteína E1 de BPV tiene actividad de helicasa y la de VPH11 tiene actividad de ATPasa. Junto con la proteína E2 es requerida para la replicación extracromosomal (Park et al., 1995).

**Gen E2:** Este gen codifica para la proteína E2 que tienen efectos regulatorios en la transcripción y replicación del ADN viral. Pesa de 40-58 kD, activa la transcripción de BPV-1 por unión a su elemento enhancer (aumentador de la transcripción) responsivo a E2. Se ha observado que dependiendo de la distancia que exista entre su secuencia blanco y el inicio de la transcripción, esta proteína es capaz de reprimir (entre 3 a 67 pb) o activar (más de 100 pb) la transcripción en promotores heterólogos (Guido et al., 1992). En VPH16 y VPH18 la proteína E2 completa reprime la transcripción de E6 y E7 que son proteínas transformantes. Cuando VPH-16 se integra en el cromosoma de la célula humana hospedera, generalmente los genes que se rompen son E1/E2, permitiendo así que la represión de E6 y E7 se libere con la consecuente sobreexpresión de estas proteínas transformantes. Más adelante se detallará ampliamente.

**Gen E4:** Este gen se encuentra completamente dentro del gen E2. Frecuentemente pierde su codón de inicio y es expresado de transcritos que sufren corte y empalme. El producto de E4 se encuentra principalmente en queratinocitos superficiales donde es abundante. Este gen fue considerado como temprano, actualmente se sabe que participa en encapsidación por lo que se le considera también como un gen tardío. Se desconocen mucho acerca de su función. Existe una proteína que resulta del “splicing” (corte y empalme) del RNA mensajero de la región N-terminal de E1/E4 en VPH16, ésta proteína no parece poseer propiedades transformantes, se encuentra asociada a la membrana celular y se acumula en el citoplasma celular. En queratinocitos humanos esta proteína induce el colapso de la red de citoqueratinas citoplásmicas sugiriendo que estos cambios facilitan la liberación de partículas virales, este fenómeno es el que podría estar causando también el aclaramiento nuclear o halo nuclear conocido como koilocitosis observado histológicamente en células productivamente infectadas con VPH (Park et al., 1995).

**Gen E5:** El gen E5 codifica una proteína transformante celular. Es una de las menos conservadas entre los PV's. Su producto es la proteína transformante más pequeña identificada aproximadamente 7 kD. Causa proliferación en fibroblastos dérmicos en fibropapilomas. Contiene un dominio hidrofóbico en su región amino-terminal, el cual es el responsable de su unión a la membrana plasmática en la célula infectada. La proteína E5 es capaz de incrementar la actividad de los receptores para el factor de crecimiento epidérmico (EGF) y el factor estimulante de colonias (CSF1) en células NIH 3T3 a través de la inhibición de la regulación interna de los receptores (Martin et al., 1989). Recientemente se ha encontrado que E5 interactúa con una proteína celular de 16 kD la cual es un componente de la ATPasa vacuolar, esta interacción juega un papel en el bombeo de protones interfiriendo con la internalización y degradación de los receptores (Goldstein et al., 1991).

**Gen E6:** Codifica para una proteína de 16-19 kD transformante. El producto del gen E6 contiene cuatro motivos Cys-X-X-Cys, lo que indica que tiene un sitio de unión a zinc; por lo que puede actuar como una proteína de unión a ADN. En virus de alto riesgo como VPH-16 las proteínas E6 y E7 son necesarias y suficientes para inmortalizar a sus células epiteliales escamosas hospederas. Se ha demostrado que la proteína E6 de VPH's de alto riesgo forma un complejo con p53 la cual es una proteína endógena celular,

supresora de tumores y promueve su degradación. La caracterización de esta proteína indica que E6 pudiera promover la degradación de p53 por medio de la vía proteolítica dependiente de ubiquitina (Scheffner et al., 1990). La ubiquitinación y degradación de p53 es mediada por una proteína celular llamada E6-AP (proteína asociada a E6), comparada con la proteína E6 de VPHs de bajo riesgo oncogénico (VPH 6 y 11), la proteína E6 de VPHs de alto riesgo parecen tener una alta afinidad por p53 y/o acelera su degradación (Park et al., 1995).

**Gen E7:** La proteína E7 es una fosfoproteína de 98 aminoácidos (14kD) localizada en el núcleo y el citoplasma celular (Smotkin & Wettstein, 1987). Los primeros 37 aminoácidos del extremo amino terminal de los VPH16 y VPH18 son capaces de unirse a la proteína p105-RB, que es la proteína supresora de retinoblastoma, que está involucrada en el control de la proliferación de las células normales (Dyson et al., 1989; Munger et al., 1989). En células normales no infectadas con VPHs de alto riesgo, las formas hipofosforiladas de Rb, p107 y p130, forman complejos con factores de transcripción de la familia E2F, esta unión es regulada precisamente a través de la acción de cinasas dependientes de ciclinas en estados específicos del ciclo celular. Durante las fases G<sub>0</sub> y G<sub>1</sub> del ciclo celular las proteínas p107 y p130 forman complejos con las proteínas hipofosforiladas de Rb, estos complejos regulan negativamente el ciclo celular reprimiendo la transcripción de genes dependientes de E2F. Conforme las células progresan hacia la fase G<sub>1</sub>/S del ciclo celular, los complejos E2F-Rb se disocian y la E2F libre estimula la transcripción de los genes dependientes de esta proteína y permite la replicación celular. La proteína E7 de VPH16 altera esta regulación tan precisa, ya que se une a Rb y disocia el complejo E2F-Rb, así la E2F liberada puede estimular la transcripción de genes necesarios para la replicación del ADN, permitiendo que se replique el ADN y la célula incontroladamente. La proteína E7 de VPH6 que es un virus de bajo riesgo, tiene baja afinidad por Rb, mientras que la E7 de VPH16 y 18 tiene alta afinidad por Rb. Esto es debido a que la E7 de VPH6 tiene una sustitución de un aminoácido (Park et al., 1995).

Los oncogenes E6 y E7 por si solos no han sido capaces de transformar células normales en cultivo a menos que sean co-tranfectadas con un oncogén como c-Ha-Ras (Matlashewski et al., 1987) y cuando se fusionan células tumorales de CaCU con células normales, se inhibe la expresión de los oncogenes virales y se revierte el fenotipo maligno, esto indica que durante el proceso maligno se requiere la pérdida de alguna función celular que regula la expresión viral (zur Hausen, 1991).

**Gen L1 y L2:** Estos genes son tardíos y codifican para las proteínas de cápside; L1 es la mayor de las dos pesa 56-60 kD, la proteína L2 tiene un peso molecular de 49-60 kD. L2 esta relativamente bien conservada entre los papilomavirus. Contiene dos sitios de localización nuclear en el dominio carboxilo-terminal.

**LCR:** región larga de control, algunas veces nombrada como región regulatoria 5' hacia arriba (URR) o como región no codificante. Es de entre 800 y 1000 pb, está bien conservada entre todos los VPH's. Está localizada entre el final de L1 y el principio de E6. No posee ningún marco de lectura abierta, es decir que no codifica para ninguna proteína y posee varias secuencias de reconocimiento a factores celulares y

virales, (García-Carranca et al., 1988) en ella se encuentran el origen de replicación y el sitio de inicio de transcripción de los oncogenes E6 y E7 (Ver Figura 4).

Algunas de las proteínas que reconoce son

GRE: Es el elemento de respuesta a glucocorticoides, une al receptor de glucocorticoides, es un palíndromo parcial que contiene las bases 5'-TGTACAN<sub>3</sub>TGTCAT-3'. En VPH16 los glucocorticoides activan la transcripción del promotor P<sub>97</sub>. Esta secuencia está presente en muchos de los papilomavirus que infectan mucosas.

Sitio de unión a YY1: Esta proteína represora se une a una gran diversidad de secuencias. En papilomavirus los sitios de unión más comunes son similares al sitio consenso 5'-AANATGGNS-3' o a la secuencia 5'-CTCCATTTT-3'. YY1 en VPH16 y 18 se une a su secuencia silenciadora y ejerce su acción probablemente interfiriendo con el inicio de transcripción.

NF-1: Reconoce la secuencia 5'-TTGGC-3' y actúa aumentando la transcripción.

SP-1: Reconoce la secuencia 5'-NGGNGN-3' y está involucrado en regulación de la transcripción.

AP-1: Reconoce la secuencia 5'-TGACTAA-3' involucrado en regulación de la transcripción.

Sitios de unión a E2: Estos sitios presentan la secuencia consenso 5'-ACC N<sub>6</sub>G (GT)T-3', el cual es un palíndromo y funciona como un elemento regulatorio a través de la cual interactúan varias proteínas E2. La secuencia central N<sub>6</sub> es determinante en la afinidad con la que se unen los productos de E2. Estos sitios están bien conservados en el LCR de los papilomavirus.

En el siguiente párrafo se detalla más acerca de la actividad de esta región.

#### **1.4.1 La región larga de control (LCR) y elementos que controlan la replicación y transcripción del virus.**

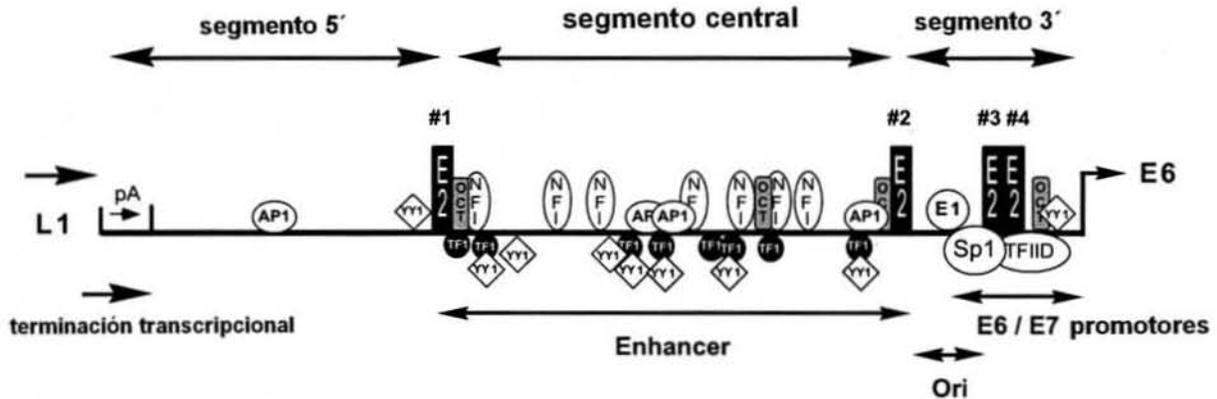
Durante la infección natural del VPH, las células basales del epitelio, son las primariamente infectadas. Su actividad transcripcional dentro del epitelio está fuertemente ligada al estado de diferenciación de los queratinocitos. La transcripción temprana se observa en células basales y suprabasales. La transcripción de los genes tardíos con la síntesis de las proteínas de cápside se lleva a cabo en los queratinocitos terminalmente diferenciados cercanos a la superficie epitelial. Sin embargo los transcritos tardíos son también detectables en capas inferiores, pero permanecen en el núcleo debido al bloqueo del procesamiento del RNA (Geisen & Kahn, 1996).

Existen varios estudios que demuestran que la región larga de control o LCR de los Papilomavirus contienen elementos regulatorios muy importantes como son el origen de replicación, varios promotores, una secuencia de mantenimiento plasmídico o episomal y un potenciador (o enhancer) de la transcripción (García-Carranca et al., 1988).

Mediante análisis de mutaciones se identificó un potenciador de la transcripción de 224 nucleótidos (nt) keratinocito-dependiente en la LCR de VPH16 (Gloss et al., 1987). Experimentos similares con VPH18 mostraron que existe un potenciador de 230 nt (Gius et al., 1988). Experimentos de "Footprinting" (huella impresa) con ADNsa, revelaron la presencia de varias secuencias dentro del

potenciador de estos virus, dentro de los que se encuentran las secuencias de reconocimiento a la proteína E2 (Garcia-Carranca et al., 1988). Otra región potenciadora diferente de la inducible por E2 encontrada en VPH16, es también inducible por glucocorticoides (Gloss et al., 1987).

**Figura 4: LCR de VPH16 y factores que reconoce.**



#### 1.4.2 La proteína E2 de VPH

La modulación de la expresión de los oncogenes E6 y E7 por E2 tiene implicaciones muy importantes para la progresión maligna. En lesiones benignas y pre-malignas que contienen VPH el ADN viral se encuentra preferencialmente como un episoma estable. En contraste en aproximadamente el 5% de las lesiones avanzadas del cérvix (CIN 2 y 3), al igual que en el 80% de las biopsias de tumores de carcinoma cervical, líneas celulares derivadas de neoplasias cervicales y queratinocitos humanos inmortalizados, el ADN viral se encuentra integrado en el ADN hospedero. Interesantemente la integración del ADN parece estar asociado con el establecimiento de un fenotipo inmortalizado en células en cultivo.

Se ha sugerido que la integración del VPH a los cromosomas humanos no es un evento al azar, este sucede preferencialmente en los MLA de E1-E2 (Figura 2), dejando intactos a los genes que codifican para los oncogenes E6 y E7. La disrupción del gen E2 da lugar a la pérdida de regulación de los oncogenes E6 y E7 produciendo una sobreexpresión de estas oncoproteínas. Por esta razón la integración del ADN viral se considera como un paso importante en la progresión de las lesiones de VPH a cáncer. Sin embargo es importante recalcar que el 13% del 90% de los tumores de CaCU que contienen integrado el ADN viral, contienen también ADN episomal. Se ha encontrado la coexistencia de formas episomales e integradas de ADN viral dentro de una sola célula. Este hecho es difícil de reconciliar con la hipótesis de que la disrupción del gen E2 debida a la integración, es un paso esencial hacia la

transformación maligna, ya que la presencia de copias episomales con el potencial de expresar la proteína E2, puede reemplazar las funciones regulatorias del gen E2 disrumpido. La aclaración de este punto necesita más estudios tanto de la expresión de las proteínas como de la regulación de la transcripción en el VPH (Rocha-Zavaleta et al., 1997). Los papilomavirus han revolucionado un sistema sofisticado de control mediado por interacciones proteína-ADN y proteína-proteína que involucra tanto a proteínas celulares como a proteínas virales. La fosfoproteína E2 nuclear de masa molecular relativa de 45000 (Mr 45K) tiene dos papeles claves en este control. Funciona como el principal factor codificado por el virus y en asociación con la proteína E1 viral crea un complejo molecular para llevar a cabo la replicación del ADN viral (Antson et al., 2000).

#### **1.4.2.1 La proteína E2 y la replicación del virus**

El ciclo de replicación completo de VPH resultante en la producción de viriones está fuertemente ligado al estado de diferenciación de las células infectadas. La persistencia viral se lleva a cabo a través del mantenimiento de genomas extracromosomales virales de 50-200 copias por célula infectada. Durante la diferenciación epitelial, los genomas de VPH se amplifican a niveles de 1000-10,000 copias por célula. En el caso de VPH31b los niveles de expresión de RNAs de E1 y E2 correlaciona con la replicación vegetativa del ADN viral, los niveles de E1 y E2 son disímiles a través del ciclo viral, el RNA de E2 permanece relativamente constante mientras que los niveles de RNA de E1 son más elevados durante la máxima amplificación de los genomas virales y la biosíntesis de los viriones (Ozbun & Meyers, 1998).

La mayoría de los datos sobre replicación genómica de los papilomavirus es derivada de ensayos transitorios usando papilomavirus Bovino tipo 1 (BPV1), y recientemente se ha mostrado que los VPHs muestran muchas de las actividades de replicación que presenta BPV1. El origen de replicación del ADN de VPH está localizado dentro de la región reguladora del virus (LCR), hacia el extremo 5' del marco de lectura abierta del oncogen E6. Los productos de los genes virales, las proteínas E1 y E2, son necesarios para la replicación del virus. La proteína E1 de VPHs es una fosfoproteína que une ATP y tiene actividad de ADN helicasa, esta proteína se une a una secuencia conservada rica en A/T en el sitio de origen de replicación del ADN viral. La proteína E2 que es también una proteína nuclear y que puede activar la transcripción, (como se mencionará mas adelante), utiliza un sitio de unión al ADN justo hacia el extremo 3' a la secuencia rica en A/Ts a la cual se une la proteína E1 del origen de replicación. La proteína E2 forma un complejo con E1 y estimula la replicación mediada por E1. La afinidad de E1 por el origen de replicación es incrementada por E2 y requiere un sitio adyacente de unión a E2. Se sabe que la alteración en la relación E1: E2 afecta el patrón de replicación de los factores que se unen al origen viral in vitro. Sin embargo el mecanismo que permite la regulación de la expresión de E1 y E2 y la replicación vegetativa del ADN viral durante el ciclo viral es poco conocido (Frattini & Laimins, 1994).

E2 parece reclutar y aumentar la unión de la proteína E1 al origen de replicación donde E1 puede actuar como ATPasa dependiente de ADN, helicasa dependiente de ADN y tener actividad de desdoblamiento. E2 puede también interactuar con el factor celular de replicación A (RPA) el cual es

requerido para la iniciación de la replicación y puede ayudar a disrumpir la estructura de la cromatina cerca del origen de replicación, permitiendo a los factores de replicación virales y celulares tengan acceso al origen de replicación(Li & Botchan, 1994).

Un gran número de estudios se han realizado para determinar cuales regiones de E2 son requeridas para la actividad de replicación y de interacción con E1, la conclusión de estos estudios indicó que las mutaciones en el dominio N-terminal de E2 bajan la actividad de replicación. Sin embargo la interpretación de estos estudios son complicados por la limitada información estructural acerca de la proteína E2. Esencialmente todas las deleciones en el dominio de activación de E2 inactivan tanto la replicación como la transcripción y es probablemente el resultado de una proteína no plegada, es decir con la estructura tridimensional alterada. Recientemente también se ha mostrado que mutaciones puntuales en el dominio de activación tienen efectos en la estructura de la proteína (Berg & Stenlund, 1997). Se ha demostrado previamente que las proteínas E1 y E2 pueden funcionar mezcladas y en combinaciones apareadas. Por ejemplo, E1 de BPV1 puede funcionar con E2 de VPH11 para replicar un plásmido con el origen (Ori) de VPH, lo que indica que el sitio Ori es muy conservado, sin embargo existen algunas excepciones, E1 de BPV1 y E2 de VPH11 no funcionan para replicación en el Ori de BPV1.

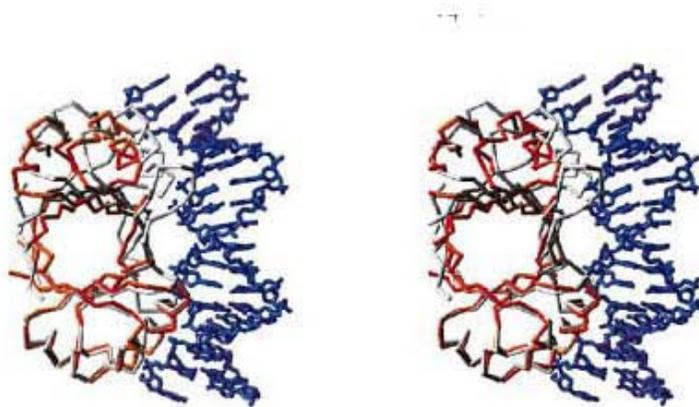
Berg y Stenlund en 1997, generando proteínas E2 quiméricas BPV-VPH para mapear los dominios de E2 que son requeridos para interacción con E1 en el Ori, demostraron que son requeridas dos tipos de interacciones, una de ellas ocurre entre E1 y el dominio de activación de E2, esta interacción es conservada y se presenta bien entre E2 de BPV1 y de VPH11, y es suficiente para la actividad de replicación. El segundo tipo de interacción necesario es entre E1 y el dominio de unión al ADN de E2 de BPV. Esta interacción no es conservada entre E2 de BPV1 y E2 de VPH11. Sin embargo se observó que cambiando la distancia entre Ori y el sitio de unión a E2 también se inhibe la interacción E1- E2, lo que indica que esta interacción particular depende del contexto estérico. La región de E2 que específicamente es requerida para su actividad en la replicación es la región N-terminal activadora, en contraste la región de bisagra y la de unión a ADN no son específicamente requeridas, sin embargo el dominio de unión es necesario para la unión de E1 y E2 a Ori (Berg & Stenlund, 1997). Como para la proteína E2 de BPV1, la región N-terminal de la proteína E2 de VPH16 parece mediar la interacción con la proteína E1, ya que anticuerpos monoclonales que reconocen los aminoácidos 18 a 41, interrumpen la interacción con la proteína E1 de VPH16 (Hibma et al., 1995). Mediante ensayos de sustitución de aminoácidos de la región N-terminal de la proteína E2 de VPH11, VPH16 y BPV1 por alaninas, mostraron que la sustitución en el aminoácido E39 inhibe la replicación, pero no la transactivación (Cooper et al., 1998).

#### 1.4.2.2 Dominios y estructura de la proteína E2

Análisis de la secuencia de aminoácidos de todas las proteínas E2, muestra que poseen dos dominios muy conservados. Uno N-terminal de aproximadamente 200 aminoácidos y uno C-terminal de aproximadamente 100 aminoácidos, separados por una región no conservada de longitud variable o “bisagra”, rica en prolina y flexible, esta región sobrelapa el marco de lectura abierta del gen E4 el cual es muy divergente entre los papilomavirus, mientras que los dominios conservados de la proteína E2 son similares entre si en aproximadamente el 35%.

La proteína E2, forma dímeros a través de su región carboxilo terminal que a su vez reconocen un dominio conservado de unión al ADN (ACCG N<sub>4</sub> CGGT) este dominio dimeriza para formar un barril  $\beta$  plegado con hélices de reconocimientos que los flanquean posicionados en el surco mayor del ADN como se muestra en la Figura 5. (Hegde & Androphy, 1998). El dominio N-terminal de transactivación es de aproximadamente 200 aminoácidos y es diferente de muchos dominios de transactivación ya caracterizados para otras proteínas. En E2 este dominio parece poseer una estructura muy constreñida que es muy fácilmente interrumpida por delección de ciertas mutaciones puntuales no conservadas. La secuencia de aminoácidos de casi todas las proteínas E2, en esta región, demuestra que forma dos  $\alpha$ -hélices. Sin embargo hasta ahora no existe evidencia experimental de que estas estructuras secundarias existan en dominios de transactivación.

La región de bisagra de las proteínas E2 varía tanto en longitud como en composición de aminoácidos. Se ha postulado que esta región forma una unión flexible entre los otros dos dominios, un estudio de la proteína E2 de VPH16 confirmó que la bisagra es una región no estructurada (McBride et al., 1989).



**Figura 5.** Vista estereoscópica de la súper imposición de las estructuras de E2 VPH16, con el ADN en la conformación encontrada en el complejo de E2 BPV-1 al ADN (Hegde & Androphy, 1998).

### 1.4.2.3 La proteína E2 y su secuencia blanco en el LCR

La región larga de control (LCR) de todos los papilomavirus humanos tienen cuatro sitios de reconocimiento a dímeros de la proteína E2, (Figura 4) las posiciones de estos sitios son altamente conservadas, lo que nos indica que existen mecanismos similares entre todos los papilomavirus para que E2 regule la expresión del promotor de E6 y E7. Se ha demostrado que en el caso del VPH18 se requiere únicamente del sitio de unión a E2 más proximal a TATA (sitio BS-4), para reprimir la transcripción, mientras que en VPH16 se requieren dos sitios de unión a E2 (BS-3 y BS4), para que la proteína E2 reprima al promotor p97. Estos dos sitios de unión a E2 están espaciados por 1 o 3pb de secuencias que reconocen factores celulares como SP1 y la proteína de unión a TATA (TBP). Se ha demostrado en ensayos *in vitro* que en VPH11, 16 y 18 la unión de E2 a los sitios BS3 y BS4, compite con el ensamblaje del complejo de preinicio de la transcripción y la unión a SP1. El sitio BS-2 que se encuentra flanqueando el sitio 3' del enhancer (potenciador de la transcripción) se ha encontrado involucrado en la replicación del ADN viral. La función del sitio BS-1 que está más distal del promotor, aun no se ha determinado, sin embargo la conservación notable de este sitio entre los VPHs genitales, sugiere que este sitio puede ser importante para la regulación de la expresión de los genes virales. (Steger & Corbach, 1997).

### 1.4.2.4 Unión al ADN y dimerización

El dominio C-terminal de E2 se une específicamente al ADN como un dímero. La secuencia consenso de reconocimiento a E2 en el ADN es ACCG N<sub>4</sub> CGGT, los nucleótidos internos y las secuencias flanqueantes, pueden influenciar la afinidad de unión de E2 en varios órdenes de magnitud (McBride et al., 1989; Hines et al., 1998). La unión de la proteína E2 a su secuencia blanco induce un doblez en el ADN y este doblez puede ser inhibido por metilación en los nucleótidos CpG en su secuencia blanco.

La secuencia de aminoácidos del dominio de unión al ADN de las proteínas E2 de BPV1 y VPH16 son 33% idénticos y 51% similares. E2 de VPH16 se une con más afinidad a su secuencia blanco en el ADN cuando el espaciador es rico en Adeninas y Timidinas. E2 de VPH16 es más sensible a variaciones de secuencias dependientes de ADN que la proteína E2 de BPV (Hines et al., 1998).

La estructura cristalográfica de rayos-X de los 85 aminoácidos del dominio Carboxilo-terminal de E2 del BPV1, unido a su secuencia de ADN blanco fue el primer ejemplo de una estructura de unión a ADN en forma de barril- $\beta$  anti-paralelo. Como se muestra en la Figura 5 un dímero del dominio de E2 de unión al ADN forma un barril- $\beta$  anti-paralelo hecho de cuatro cadenas de cada subunidad. Un par de hélices- $\alpha$  simétricamente posicionada en la parte externa del barril contienen los residuos que son requeridos para la interacción específica con el ADN. Las dos hélices - $\alpha$  del dímero de E2 se insertan dentro de los surcos mayores del ADN y se acomodan de tal manera que provocan una suave y simétrica deformación del ADN (Hegde & Androphy, 1998).

El complejo E2-ADN se forma de las interacciones directas entre cuatro aminoácidos conservados en cada hélice de reconocimiento; Asn 336, Lys 339, Cys 340 y Phe 343, tres de éstos aminoácidos forman un total de siete uniones de hidrógeno directos unidos a las pares de bases ACC, que son parte del sitio consenso de unión a la proteína E2, el cuarto aminoácido forma un contacto de van der Waals con el grupo 5' metil de Timidina. Además existen dos interacciones específicas con las bases, mediadas por agua; diez contactos directos de fosfato y catorce interacciones de fosfato mediadas por agua, todas estas uniones son por cada monómero de E2 (Hegde et al., 1992).

La proteína E2 forma un dímero muy fuerte que puede ser interrumpido solamente bajo condiciones desnaturizantes como son Urea 5M. La interacción que permite esta unión dimérica entre dos proteínas es una red que embona internamente entre las cadenas  $\beta$  del barril con un total de 10 uniones de hidrógeno. Existe un corazón hidrofóbico en el centro del barril que contiene un aminoácido Trp conservado en la posición 360 que cuando es reemplazado por un residuo polar en experimentos de análisis mutacionales, se interrumpe el dímero. Aunque no perfectamente, los anillos indol del Trp 360 de las subunidades opuestas están en contacto van der Waals (Hegde et al., 1992).

El plegamiento de la proteína es muy similar a la estructura cristalográfica del dominio de BPV-1, pero la hélice de reconocimiento al ADN parece ser mas flexible como se ha observado en otras proteínas de unión al ADN. El dominio de unión al ADN de la proteína EBNA1 del virus de Epstein Barr tiene una estructura de unión a ADN muy similar a la de E2 aunque no tienen secuencias similares (Bochkarev et al., 1995).

La hélice de reconocimiento contiene un residuo de cisteína altamente conservado en la posición 340 que es muy sensible a oxidación (Li & Botchan, 1994). Este residuo hace contacto directo con el ADN, proteínas E2 con substituciones en este residuo (glicina, serina, alanina) son capaces de unirse al ADN. Sin embargo estas proteínas no son capaces de activar la transcripción eficientemente en células de mamífero (Dowhanick et al., 1995).

Hegde y Androphy, (Hegde & Androphy, 1998) obtuvieron información de la estructura tridimensional de la proteína E2, el ADN de reconocimiento a la proteína libre y del complejo ADN-proteína. Sus conclusiones fueron que la unión de la proteína E2 a su ADN causa alteraciones estructurales cuaternarias a la proteína E2 de BPV1 y que esta interacción con el ADN es simétrica y está acompañada por una significativa distorsión del ADN.

Las propiedades de unión a ADN y de dimerización de este dominio no pueden ser separadas por análisis de delección; todas las delecciones que se han probado, eliminan las dos propiedades del dominio C-terminal. Esto es porque la dimerización involucra una interface extensiva de la subunidad que consiste de uniones intra-bisagras de hidrógeno entre las hebras- $\beta$ , la interacción entre las cadenas laterales de la hebra- $\beta$  en cada subunidad y un corazón altamente hidrofóbico. Este corazón hidrofóbico contiene un residuo de tryptófano altamente conservado en la posición 360 el cual ha sido designado como puente de tryptófano (McBride et al., 1989). Proteínas E2 mutadas en esta posición (W360) son funcionales pero la sustitución de W360 por residuos polares interrumpe la dimerización. El dímero de E2

es la unidad funcional al ADN, mutaciones que interrumpen al dímero inhiben la unión de E2 al ADN (Prakash et al., 1992). Mutaciones no conservadas en otras partes del dominio C-terminal pueden también eliminar tanto la unión al ADN como dimerización interrumpiendo la estructura de la proteína (Breiding et al., 1997).

#### 1.4.2.5 Transactivación y represión por E2

En contraste con la proteína E2 trans activadora del Papilomavirus Bovino (BPV-1), las funciones de los productos del gen E2 completo de VPH-16 o de otros papilomas genitales humanos han sido difíciles de establecer. Aunque el gen E2 de VPH y de BPV-1 muestra homologías en su secuencia en los dominios de unión al ADN y de transactivación, estudios previos han sugerido que los productos de E2 de VPHs genitales pueden ejercer diferentes funciones. Los vectores que expresan la proteína E2 completa de BPV-1 puede trans-activar genes reporteros con fragmentos de ADN que tienen secuencias blanco para E2. En contraste, se ha observado que el promotor de E6 de VPH16 y otros Papilomavirus genitales pueden ser reprimidos por los dos sitios proximales (Guido et al., 1992), parece ser que la proteína E2 puede funcionar activando o reprimiendo la actividad del promotor en diferentes líneas celulares y dependiendo del contexto en el que se encuentre el sitio de unión al E2 en el ADN (Guido et al., 1992).

Otro tipo diferente de represión transcripcional puede ser mediado por el transactivador completo E2-TA y este depende de la posición de los sitios de unión a E2 con respecto a los elementos en los promotores proximales. En muchos papilomavirus humanos se encuentran posicionados dos sitios de E2 entre un sitio SP1 conservado y la caja TATA. Se ha demostrado que la unión de E2 en esos sitios inhibe la unión de estos factores celulares resultando en la represión de la actividad promotora basal (Guido et al., 1992).

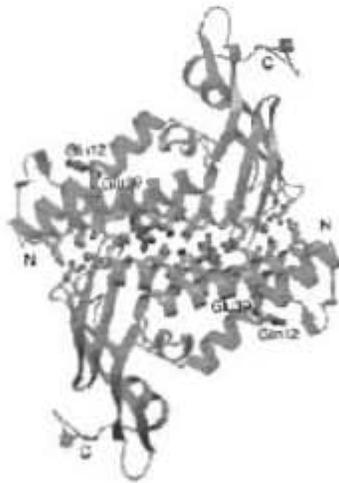
En los VPHs el número y la distribución de sitios de reconocimiento a E2, son diferentes y E2 es solamente un activador moderado. Por ejemplo, el promotor tardío de VPH8, el cual infecta la piel es activado a niveles muy bajos por E2, y la represión de la actividad basal del promotor ocurre sólo a altos niveles de proteína E2. La activación es debido a la unión de E2 a sitios de alta afinidad, mientras que la represión es mediada por unión de E2 a sitios de baja afinidad., los cuales sobrelapan con sitios de reconocimiento para factores celulares de transcripción necesarios para la actividad promotora basal (Steger & Corbach, 1997).

Estudios de microscopía electrónica han mostrado que la proteína E2 completa permite la formación de loops en moldes de ADN que se encuentran a mas de 560pb, mientras que la proteína E2 que solo contiene el dominio C-terminal no permitió la formación de loops (Knight et al., 1991). Antson y cols, (2000) observaron por cristalografía la formación de dímeros de la región amino-terminal de la proteína E2, con dos monómeros formados unidos por las hélices  $\alpha_2$  y  $\alpha_3$  del dominio N1 y los residuos 142-144 del dominio N2. Estos resultados sugieren que la dimerización de los módulos de E2NT pueden

ligar dímeros de E2 y permitir la formación de loops en el ADN, sirviendo de mecanismo por el cual acercaría factores distales al inicio de la transcripción (Antson et al., 2000).

#### 1.4.2.6 Dominio de activación de E2

Existen reportes de la estructura cristalográfica del dominio amino Terminal (dominio de activación) de las proteínas E2 de VPH16 (Antson et al., 2000) y 18 (Harris & Botchan, 1999). En el caso de VPH16 esta región consiste de dos dominios N1 y N2, el dominio N1 (residuos 1-92) está compuesto de tres hélices- $\alpha$  largas en posición antiparalela una a otra formando un plano enrollado. Existe un asa ("loop") muy fuerte (y corto) entre la hélice  $\alpha_1$  y  $\alpha_2$  y uno más extendido entre las hélices  $\alpha_2$  y  $\alpha_3$ . El dominio N2 (residuos 110-201) está conformado casi enteramente por una estructura de hojas  $\beta$  antiparalelas. El dominio N1 de activación de VPH16 es dimerico en la estructura cristalográfica con residuos sensibles a mutación conservados, que ocupan la interface del dímero. Se ha sugerido que el dominio de activación mediante oligomerización con otros dominios de activación de E2, pueden jugar un importante papel en la interacción entre moléculas de E2 unidas a sitios de unión de ADN distantes, propiciando la formación de loop en el ADN y otras estructuras de alto orden (Antson et al., 2000).



**Figura 6 . Vista estereoscópica del dímero amino-terminal de E2 VPH16.** Se observan las hélices alfa y beta de cada monómero y la extensión de la interfase entre las dos subunidades (Antson et al., 2000).

#### **1.4.2.7 Interacción con proteínas celulares**

Se ha demostrado que E2-TA de BPV-1 interactúa con la proteína celular de replicación RPA y el factor celular SP1 (Lambert et al., 1989). El dominio de 127 aminoácidos carboxilo terminal de proteínas E2 de fusión o las que contienen 160 aminoácidos de este mismo dominio son incapaces de interactuar con SP1, por lo que se deduce que es requerido el dominio N-terminal para esta interacción (Lambert et al., 1989).

La proteína E2 de BPV-1 interactúa y cooperativamente se une al ADN con los factores celulares de transcripción TFIID y TFIIB. Los residuos de 310-410 son suficientes para interacción proteína-proteína con esos factores, sin embargo se requiere la región de bisagra además del dominio de unión a ADN (residuos 204-410) para la unión cooperativa al ADN con TFIID o TFIIB (McBride et al., 1989; Muller et al., 1997), interactúa también con proteínas de unión proximales al promotor como Sp1 y otros factores celulares como AMF-1 el cual positivamente afecta la activación transcripcional de E2. No se sabe cuales de estas interacciones son suficientes o necesarias para llevar a cabo activación transcripcional, estas proteínas funcionan como activadores transcripcionales usando contactos proteína-proteína específicos para reclutar los componentes de la maquinaria general de transcripción a un promotor con la intención de reclutar a la RNA polimerasa.

## 2. OBJETIVOS

- Secuenciar el gen E2 de dos variantes del VPH tipo 16 presentes en tumores de mujeres Mexicanas.
- Caracterizar la capacidad de la proteína E2 para unirse al ADN, mediante geles de retardo.
- Hacer análisis funcional de la proteína E2 mediante cotransfección con el LCR clonado frente a un gen reportero en líneas celulares derivadas de CaCU para determinar su actividad *in vitro*.

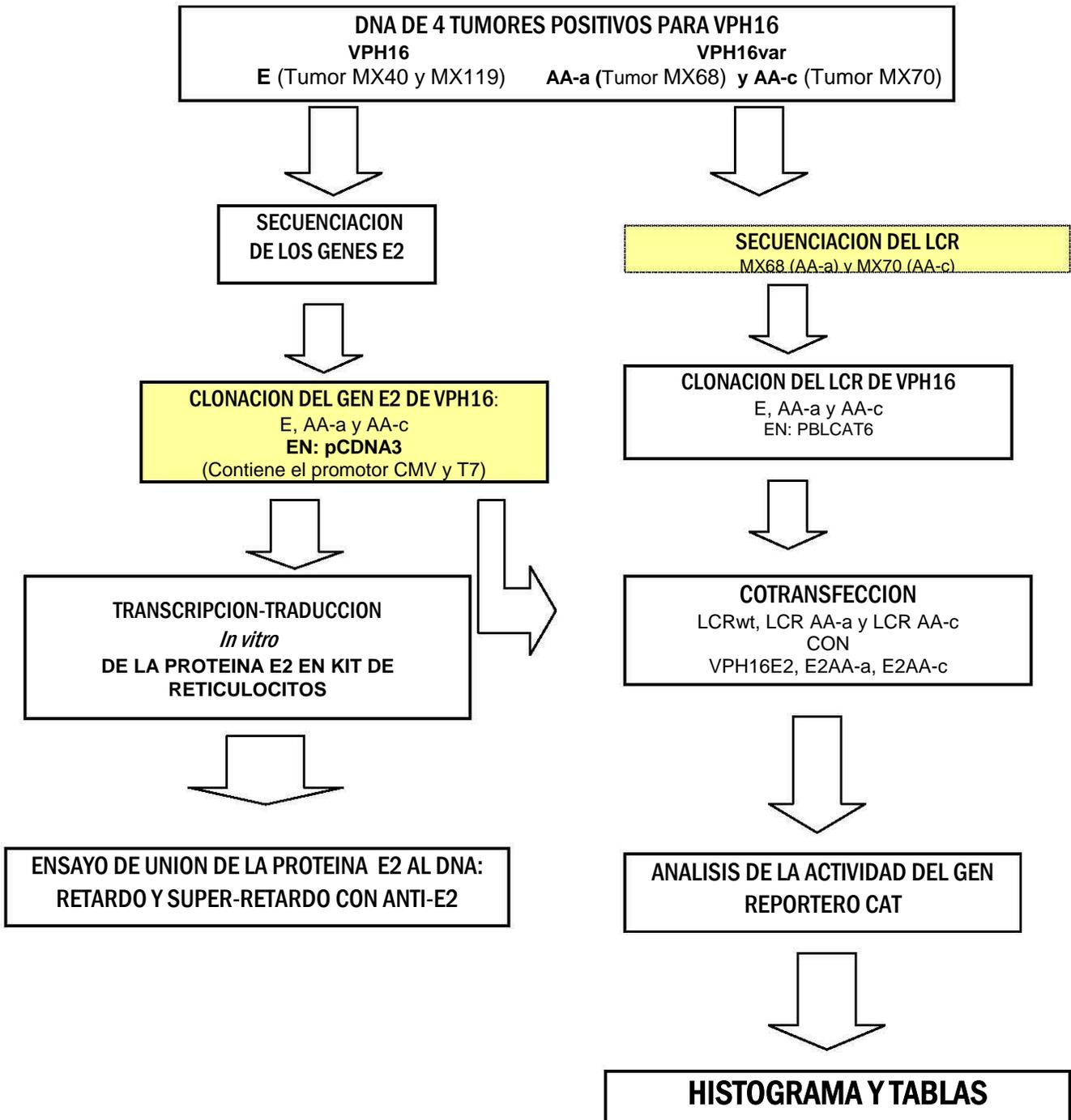
## 3. HIPOTESIS

Se sabe que la LCR y se sugiere que el gen que codifica para la proteína E2 de VPH16 AA-a y AA-c tienen cambios nucleotídicos en el ADN con respecto a VPH16E, estas diferencias pudieran estar confiriendo a los VPHs diferencias en la expresión de las proteínas oncogénicas del VPH.

## 4. JUSTIFICACION

Las lesiones con VPH16 variante, son mas agresivas comparadas con lesiones que tienen el VPH16 Europeo y se ha demostrado que el LCR de VPH16 AA-c es más activo que el Europeo lo que puede aumentar la oncogenicidad del virus. Aun no se saben los mecanismos que marcan estas diferencias, por lo que es necesario secuenciar los genes E2 y hacer análisis de la actividad transcripcional tanto de las proteínas E2 como de la LCR de estas variantes (AA-a y AA-c) y analizar de qué manera están regulando la transcripción.

Figura 7. MATERIAL Y METODOS



## 5. MATERIAL Y METODOS

### 5.1 Amplificación de E2 por PCR de tumores positivos para VPH16

Los “primers” utilizados para amplificar por PCR (Polimerase Chain Reaction; Reacción en cadena de la Polimerasa) un fragmento de 1227 pb (pares de bases) que contiene el gen E2 fueron:

NOMBRE	LOCALIZACION	SECUENCIA
16-E2.1	2670 – 2689	5´ TGA GCT TAA TGA TAA GAA CT 3´
16-E2.2	3896 - 3879	5´ CAA AAA GCA CGC CAG TAA 3´

Para amplificar el gen E2 por PCR se utilizaron de 1-3 µg de ADN de cada uno de los tumores (MX68, MX70: Variantes AA y MX40, MX119: Europeos), 1µM /µl de los “primers” antes mencionados, 200 µM/reacción de cada uno de los cuatro deoxyribonucleótidos (dATP, dCTP, dGTP, dTTP), 2mM de MgCl<sub>2</sub>, buffer de PCR y 1 unidad de Taq-Polimerasa (Perkin Elmer) en un volumen total de 50 µl. La reacción fue sometida a 40 ciclos que comprenden las siguientes tres temperaturas: Desnaturalización, 1 min. a 95°C; alineamiento específico de los primers al ADN, 1 min. a 50°C, y extensión del ADN a amplificar, 2 min. a 72°C, el tiempo de extensión se prolongó 6 seg. conforme aumentó el número de ciclos para asegurar la extensión de todas las hebras de ADN. Estos productos de PCR se purificaron en agarosa 1%, y se limpiaron con el Kit de QIAEX (Quiagen, Chatsworth, CA, USA) (Barnard et al, 1994) eluyendo el ADN en 20 µl de buffer TE (10mM Tris-HCl, pH 8.0, 1mM EDTA, pH 8.0). El producto de PCR se cuantificó para su posterior secuenciación.

### 5.2 Secuenciación del gen E2

La secuenciación del gen E2 se llevó a cabo de acuerdo al Kit de Sequenase versión 2 (Marca USB Corp.), con las modificaciones de Casanova y colaboradores (Casanova et al., 1990) ellos variaron tres parámetros; la temperatura de alineamiento (-70°C), el tiempo de marcaje (30 seg. a temperatura ambiente) y el radio oligonucleótido:molde (1:25), estas modificaciones permiten un intercambio rápido de calor dando lugar a resultados mas consistentes en el alineamiento específico de los primers (Barnard et al., 1994).

Los “primers” que se usaron para secuenciar el producto de PCR derivado del ADN de los tumores fueron los utilizados para la amplificación del gen E2 por PCR, además de los siguientes:

NOMBRE	LOCALIZACION	SECUENCIA
16-E2 S1	3192 -3211	5´ CTG TGG TAG AGG GTC AAG TT 3´
16-E2 S2	3384 – 3367	5´ ATT TCA GGA GAG GAT ACT 3´

Se utilizaron de 0.2 - 0.5 pmoles (aprox. 150 - 400 ng) de ADN molde (producto de PCR mencionado anteriormente), se mezcló con 10 pmol de los "primers" 16-E2 S1 o 16-E2 S2 (radio primer:molde 1/25) y 2  $\mu$ l de buffer de reacción 5X en un volumen final de 10  $\mu$ l. Para desnaturalizar el ADN esta mezcla se colocó en un baño de agua a 100°C por 2 min., inmediatamente después se colocó en hielo seco:etanol por 30 seg. para permitir el alineamiento de los "primers" al ADN molde. Posteriormente se le agregó a esta mezcla 3.75  $\mu$ l de una solución pre-mezclada de marcaje que contenía 1  $\mu$ l de dithiothreitol 100mM (DTT), 1.6  $\mu$ l de buffer de marcaje diluido (1:5), y 0.75  $\mu$ l de ( $\alpha$ -<sup>32</sup>S)dATP (5 $\mu$ Ci) (NEN Du Pont, Boston, MA, USA), inmediatamente después de que se descongeló la mezcla se le agregó 2  $\mu$ l de Secuensa diluida (1:8) (versión modificada de la ADN polimerasa del bacteriófago T7), la solución se mezcló suavemente y se incubó por 30 seg. a temperatura ambiente. Para dar lugar a la terminación de la reacción 3.5  $\mu$ l de esta mezcla se agregó a cada uno de cuatro tubos que contenían 2.5  $\mu$ l de ddNTPs (ddATP, ddCTP, ddGTP y ddTTP previamente incubados 1 min. a 40-45°C) y se incubaron a 40-45°C por 2 min. La Secuensa sintetiza la cadena complementaria de la matriz en presencia de los cuatro dNTPs (presentes en el buffer de marcaje) en el que se encuentra uno marcado con  $\alpha$ -<sup>32</sup>S dATP. Agregando ddNTPs se interrumpe al azar la elongación. Para parar la reacción se le agregó a cada tubo 4 $\mu$ l de una solución que contenía; 95% de formamida, 2mM EDTA y los colorantes de migración; azul de bromofenol y xilene-cyanol. En este paso se guardaron a -20°C hasta que se les hizo electroforesis.

### 5.2.1 Electroforesis del producto secuenciado

Se utilizó un gel desnaturalizante de poliacrilamida al 6%, urea 7.6 M , 1X TBE (10.8 gr de Tris Base; 11 gr de Ácido Bórico; 4 ml de EDTA 0.5M pH 8), para polimerizarlo se le agregaron; 180 $\mu$ l de APS (Persulfato de amonio) 10% y 54  $\mu$ l de TEMED. El gel se precorrió 30 min., a 1800 Volts (aprox. 60 Wats). La reacción de secuenciación se calentó por 2 min., a 75-80°C, y se colocó inmediatamente una alícuota de 3 $\mu$ l de cada muestra en los pozos del gel. La corrida se realizó a 60 Wats, durante 2 hrs., repitiendo el procedimiento con la segunda y tercer alícuota. Después el gel se colocó en una solución de Ácido acético 10%, Metanol 10% durante 15 min, se lavó con agua de la llave y se secó durante 1hr en el horno a 80°C. Se expuso en una placa X-Omat marca Kodak a diferentes tiempos y se reveló para analizar el ADN secuenciado.

### 5.3 Construcciones de ADN:

Las construcciones básicas utilizadas en este estudio, se esquematizan en la Figura 8, contienen al gen E2 de VPH16 Europeo (pE2-16), y E2 de VPH16 variantes, AA-a y AA-c (pE2-G68 y pE2-G70 respectivamente), clonados en el sitio de restricción *Bam*HI y *Eco*RI del polilinker del vector pCADN3 (Marca Invitrogen), la expresión de las proteínas clonadas en este vector esta dirigida por dos tipos de promotores, uno de ellos es el promotor T7 de bacteriófago, el cual nos permitió expresar la proteína un sistema de transcripción-traducción *in vitro*, para hacer ensayos de unión de la proteína E2 a su secuencia blanco, mediante retardos en geles de acrilamida. El otro es el promotor fuerte del virus



Citomegalovirus (CMV), este promotor nos permitió expresar la proteína E2 mediante transfección en células eucariontes.

Para hacer los ensayos de transfección en células eucariontes y analizar la actividad de la proteína E2 sobre el promotor (LCR) europeo y variante, mediante la expresión de un gen reportero (CAT), se utilizaron construcciones que contienen la región reguladora de VPH16 europeo (p16LCR-E) y variantes (p16LCR-AA-a y p16LCR-AA-c, Figura 8), clonados entre los sitios de restricción *HindIII* y *BamHI* del polilinker del vector pBLCAT6 (Boshart et al., 1992), el cual contiene el gen reportero CAT (Cloranfenicol acetiltransferasa). PBLCAT6 nos permitió analizar la actividad represora de la proteína E2, sobre la región reguladora. En la Tabla 3 se muestran los cambios en el LCR de los VPH 16 AA-a y AA-c

## 5.4 Ensayo de retardamiento DNA/proteína

### 5.4.1 Transcripción-Traducción *in vitro* de la proteína E2

Para realizar esta técnica se utilizó el kit acoplado de lisado de reticulocitos de conejo llamado TNT T7 (Cat. No. L4610) de marca Promega, la ventaja que ofrece es que la proteína a estudiar se puede transcribir y traducir *in vitro* en tiempos cortos y se obtiene buen rendimiento de la proteína en cuestión. Este sistema está disponible para transcribir y traducir genes clonados "hacia arriba" del promotor T7 RNA polimerasa. Para usarlo se agregan 0.2 - 2.0µg de ADN plasmídico circular, 25µl de lisado de reticulocito, buffer de la reacción, 1µl de T7 RNA polimerasa, 1µl de una mezcla de aminoácidos 1mM, 1µl de Inhibidor de ribonucleasas (RNAsin) 40u/µl, y agua libre de nucleasas, en un volumen final de 50 µl. La reacción se incubó a 30°C por 60-120 min. Inmediatamente se congela en alícuotas a -70°C y se guarda hasta que se lleve a cabo la electroforesis.

### 5.4.2 Marcaje radioactivo del oligo

Para llevar a cabo este ensayo se utilizó un oligonucleótido de 38 pb de ADN, que reconoce a la proteína E2 y tiene el sitio SP1 mutado (en negritas se muestran las bases cambiadas) de tal manera que no haya interferencia de este factor con la proteína E2, llamado E2-38/32:



Como es de extremos cohesivos, se marcó radiactivamente utilizando la ADN polimerasa de *E. coli* (Fragmento Klenow), esta enzima es capaz de sintetizar la cadena complementaria de los extremos 5' del oligonucleótido en presencia de desoxyribonucleótidos de entre los que se incluye uno marcado con fósforo 32 en posición α.

A un volumen final de 20µl se le agrega; 40ng de oligonucleótido, 2µl del buffer de reacción, 10 mCi de dATP α <sup>32</sup>P, 1mM de dCTP, dGTP y dATP y 5 unidades de la enzima Klenow. Se incubó 30 min.

a temperatura ambiente, se coloca a 80°C para desnaturalizar a la polimerasa y se diluye con agua a un volumen final de 80µl, se utiliza 1µl de esta reacción.

#### **5.4.3 Preparación del gel de poliacrilamida nativo**

A fin de conocer si la proteína E2 se está uniendo a su secuencia blanco, se llevó a cabo la técnica de electroforesis en gel de poliacrilamida nativo, que nos permite separar los fragmentos de ADN libres, de los complejos ADN-proteína que retardan su migración.

A un gel con 5% de poliacrilamida, en TBE 0.5X, se le agregaron; 180µl de APS (Persulfato de amonio) 10% y 50 µl de TEMED para polimerizarlo. El gel se recorrió 60 min. a 150 Volts. Ya depositada la muestra se corrió 2 hrs. a 150 Volts. Para fijar el ADN, el gel se colocó en una solución de Ácido acético 10%, Metanol 10% durante 15min. El gel se transfiere a un papel Wathman 3MM, secado con vacío y expuesto en una placa X-Omat marca Kodak por aproximadamente 24 hr

#### **5.4.4 Interacción ADN/Proteína**

2 µl del lisado de reticulocitos, el cual tiene a la proteína E2, se preincuba con 10µl de buffer de unión 2X (BB; 24mM Hépes, 1mM EDTA, 8mM MgCl<sub>2</sub>, 120mM KCl, 0.2% NP40, 35%glicerol), en presencia de ADNs competidores: cantidades variadas de (1-3 µg) de Poly dl-dC y 1µg de ADN de esperma de Salmón sonicado; que serán reconocidos por proteínas inespecíficas. 4mM de MgCl<sub>2</sub>, 5µg de BSA, 0.25 M de Dithiothreitol (DTT), y 5ng de la sonda u oligo marcado radiactivamente en un volumen final de 20µl. Se deja incubar por 10min. Opcionalmente se le agregó un anticuerpo contra la región carboxilo-terminal de la proteína E2 (Rocha-Zavaleta et al., 1997) se incubó 30 min., y se puso a correr en el gel nativo de poliacrilamida.

### **5.5 Transfección en células C33A**

Las células C33 fueron transfectadas por el método estándar de coprecipitación de fosfato de calcio, descrito por Wigler (Wigler et al., 1977) en cajas de cultivo petri de 60mm de diámetro.

#### **5.5.1 Preparación de las células**

La línea celular C33A utilizada es derivada de un carcinoma de cérvix indiferenciado, su número de catálogo en el ATCC es HTB31, crece en monocapa adherida al fondo de la caja y se cultiva en medio DMEM, con 7% de Suero Fetal de ternera.

El modo en que se preparan para transfectar es: Se siembran  $7 \times 10^5$  células por caja p60. Se dejan creciendo toda la noche. Al menos 4 horas antes de la transfección, se le cambia el medio de crecimiento agregándole 2 ml de medio fresco a cada caja.

### **5.5.2 Precipitación del ADN/fosfato de calcio**

La cantidad de ADN debe ser determinada de acuerdo al promotor que se utiliza, generalmente va de 1 - 10  $\mu\text{g}$ . La preparación del precipitado debe realizarse a temperatura ambiente. El ADN circular debe estar completamente puro, preparado por columna o con gradiente de Cloruro de Cesio.

El ADN se afora a 250  $\mu\text{l}$  con agua estéril y se le agregan 250  $\mu\text{l}$  de buffer HEBS 2X (NaCl 137 mM, KCl 5mM, Na<sub>2</sub>HPO<sub>4</sub> 2H<sub>2</sub>O 0.7mM, Glucosa 6 mM y HEPES 20 mM, pH a 7.05), se mezcla con vórtex y se deja 10 min. a temperatura ambiente. Se le agregan 25  $\mu\text{l}$  de CaCl<sub>2</sub> 2.5 M, durante este paso se debe estar mezclando el tubo con vórtex constantemente y se deja incubando 10 min. a temperatura ambiente para permitir que se forme el precipitado. Inmediatamente después se agrega el contenido a las células cuidando de no expulsarlo sobre la monocapa, para lo cual se ladea un poco la caja y sobre el medio se coloca la solución.

Después de una noche de incubación (máximo 16 hrs), se lava la monocapa dos veces con TD (Para 4 lt: 32 gr de NaCl, 1.5 gr de KCl, 0.4gr deNa<sub>2</sub>HPO<sub>4</sub>, 12 gr de Tris base, pH 7.45) y se le agregan 5 ml de medio fresco. Después de 40 hrs, se obtiene el extracto celular.

### **5.5.3 Ensayo del gen reportero (CAT)**

El extracto celular se obtiene raspando las células y resuspendiéndolas en 200 $\mu\text{l}$  de buffer TGD (Tris 250mM, ditioneitol 5mM, 15% glicerol, pH8), se realizan seis ciclos de congelamiento-descongelamiento en una mezcla de etanol-hielo seco, se centrifuga 10min a 4°C, y se recupera el sobrenadante. Posteriormente se lleva a cabo la determinación de proteínas presentes en el lisado celular por la metodología de Bradford (Bradford, 1976), se utiliza la misma cantidad de proteínas totales para estandarizar el experimento, por lo que el lisado celular se afora a 300 $\mu\text{l}$  con buffer TGD, se le agregan 8 $\mu\text{l}$  de Acetil CoA 20mM y 50 $\mu\text{Ci}$  de cloranfenicol marcado con <sup>14</sup>C. Se incuba de 1-4 hrs. a 37°C, se le agregan 250 $\mu\text{l}$  de acetato de etilo, se mezcla vigorosamente, se centrifuga y se obtiene la fase orgánica superior, se deja evaporar, se resuspende en 10 $\mu\text{l}$  de acetato de etilo y se coloca en puntos finos en una placa de cromatografía de silica-gel durante 45 min., en un solvente que contiene cloroformo-metanol (19:1), se deja secar y se pone a exponer durante aproximadamente 48 hrs. en una placa X-Omat de Kodak. Después de revelar la placa, la cromatografía de silica-gel, es cortada basándose en la placa colocando cada muestra en un vial con líquido de centelleo (PPO 5gr/lt, POPOP 0.3gr/lt).

## 6. RESULTADOS

### 6.1. Secuenciación de los genes E2 de VPH16 AA-a, y AA-c.

Se secuenciaron los genes E2 de 4 tumores positivos para VPH16, dos de ellos el G68 (AA-a) y el G70 (AA-c) presentaron un patrón alterado en Southern blot con la enzima de restricción *Pst*I, compatible con la existencia de una mutación puntual que cambia un aminoácido en el gen E2 (Casas et al., 1999) y otros dos G40 (E) y G111 (E) los cuales no presentaban este patrón alterado, los resultados de esta secuencia se encuentran publicados (Casas et al., 1999).

Se encontraron un total de 23 cambios nucleotídicos en G68 y 24 en G70, (Ver Tabla 2). Comparados con el gen E2 de VPH16 de referencia o europeo (1098 pares de bases), además de analizar parte del gen L1 y E6 (Casas et al., 1999), se llegó a la conclusión de que G68 muestra cambios en un 2% y G70 en un 1.9%, desde este momento se les consideró como variantes relacionadas pero diferentes. En cuanto a cambios en aminoácidos (aa), G68 presentó cambios en 16 aa y G70 en 17aa (dos mutaciones contiguas cambian el mismo codon, 254). G68 presenta cinco cambios de aminoácidos en la región transactivadora, siete en la región de bisagra y 3 en la región de unión al ADN, para G70, se encontraron cinco cambios en la región transactivadora, ocho cambios en la región de bisagra y tres en la región de unión al ADN. En el dominio de transactivación la sustitución H35Q cae dentro de un sitio identificado de interacción con la proteína E1, residuos 18-41; (Piccini et al., 1995) y cuatro cambios caen en la región 3' del dominio de transactivación. En el dominio de unión al ADN, la sustitución T310K se encuentra flanqueando la hélice (292-309) y las sustituciones W313C y D344E se encuentran en sitios putativos de dimerización. Los cambios T135K, T254N y F271V se localizan en tres de los sitios más reactivos de epítopes de Células B (residuos 121-140; 241-260 y 271-290) y los cambios W341C y D344E se encuentran dentro de epítopes mayores de IgG e IgA (residuos 328-345; (Dillner et al., 1996).

Nueve de las 24 mutaciones también afectan al gen E4, el cual está localizado dentro del gen E2, pero en diferente marco de lectura. Cuatro de las nueve mutaciones que afectan al gen E4 resultan en tres substituciones de aminoácidos. Estas tres substituciones, L62T, Q69P y H78Q, se localizan en la región E4 similarmente a la proteína actina, residuos 48 a 114 en VPH1 (Roberts et al., 1994). Dos mutaciones G142N y A143T que se encuentran en la región amino-terminal, afectan la interface de unión al dímero (Antson et al., 2000).

Tabla 2. Variaciones en la secuencia de ADN y aminoácidos de E2 en las variantes VPH16.

Aminoácido	Nucleótido	E2 E	E2 AAa	E2 AAc	Cambio de aa: AA-a	Cambio de aa: AAc
<b>DOMINIO DE TRANSACTIVACION</b>						
35	2860	C	A	A	His H / Gln Q	His H / Gln Q
57	2926	A	G	G	Gln Q / =	Gln Q / =
135	3135	C	A	A	Thr T / Lys K	Thr T / Lys K
136	3161	C	--	T	--	His H / Tyr Y
142	3181	A	C	--	Glu E / Asp E	--
143	3184	G	A	A	Ala A / Thr T	Ala A / Thr T
157	3224	T	A	A	Leu L / Ile I	Leu L / Ile I
<b>DOMINIO DE BISAGRA</b>						
165	3249	G	A	A	Arg R / Gln Q	Arg R / Gln Q
203	3362	A	G	G	Asn N / Asp D	Asn N / Asp D
208	3377	C	G	G	Pro P / Ala A	Pro P / Ala A
211	3387	T	C	C	Ile I / Thr T	Ile I / Thr T
219	3410	*C	T	T	Pro P / Ser S	Pro P / Ser S
221	3416	G	--	A	--	Ala A / Thr T
232	3449	G	A	A	Glu E / Lys S	Glu E / Lys S
254	3516	C	A	A	Thr T / Asn N	Thr T / Asn N
254	3517	T	C	C	--	--
261	3538	A	C	C	Ser S / =	Ser S / =
271	3566	T	G	G	Phe F / Val V	Phe F / Val V
<b>DOMINIO CARBOXILO-TERMINAL</b>						
303	3664	T	C	C	Tyr Y / =	Tyr Y / =
310	3684	C	A	A	Thr T / Lys K	Thr T / Lys K
313	3694	T	A	A	Thr T / =	Thr T / =
317	3706	T	C	C	Ser S / =	Ser S / =
341	3778	G	T	T	Trp W / Cys C	Trp W / Cys C
344	3787	C	A	A	Asp D / Glu E	Asp D / Glu E
350	3805	T	G	G	Val V / =	Val V / =
<b>TOTAL</b>		<b>1</b>	<b>23</b>	<b>24</b>	<b>16</b>	<b>17</b>

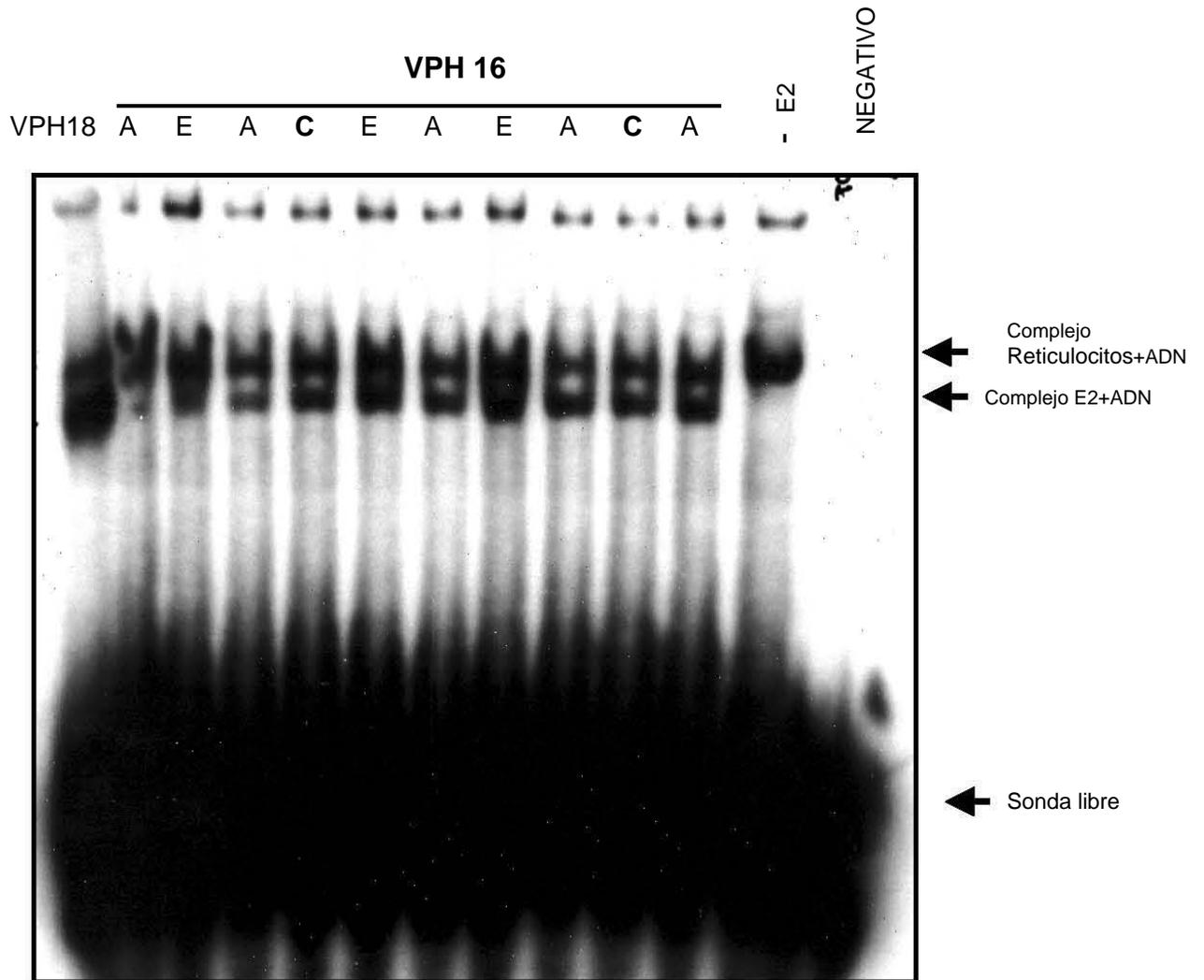
\* Cambio también encontrado en el ADN con HPV16 de referencia.

## 6.2 Expresión de distintas proteínas E2 e interacción con el ADN

La proteína E2 de VPH16 es una fosfoproteína de 43 kDa que se une como homodímero a una secuencia palindrómica de 12 pares de bases en la región que controla la transcripción del genoma del VPH. Después de unirse, E2 puede activar o reprimir la transcripción de los promotores adyacentes dependiendo del tipo celular, de la cantidad de proteína E2 presente y de su distancia al sitio de inicio de la transcripción, esta regulación controla la expresión de los oncogenes E6 y E7 del virus.

Usamos el ensayo de retardamiento para corroborar que la proteína E2 de las variantes (AA-a y AA-c) se estuviera uniendo a su secuencia blanco y tratar de observar diferencias en la unión de las proteínas E2 variantes con respecto a E2 Europea, utilizando un oligonucleótido que contiene un sitio de reconocimiento al homodímero de la proteína E2 (E2-38/32), marcado radiactivamente con  $^{32}\text{P}$ .

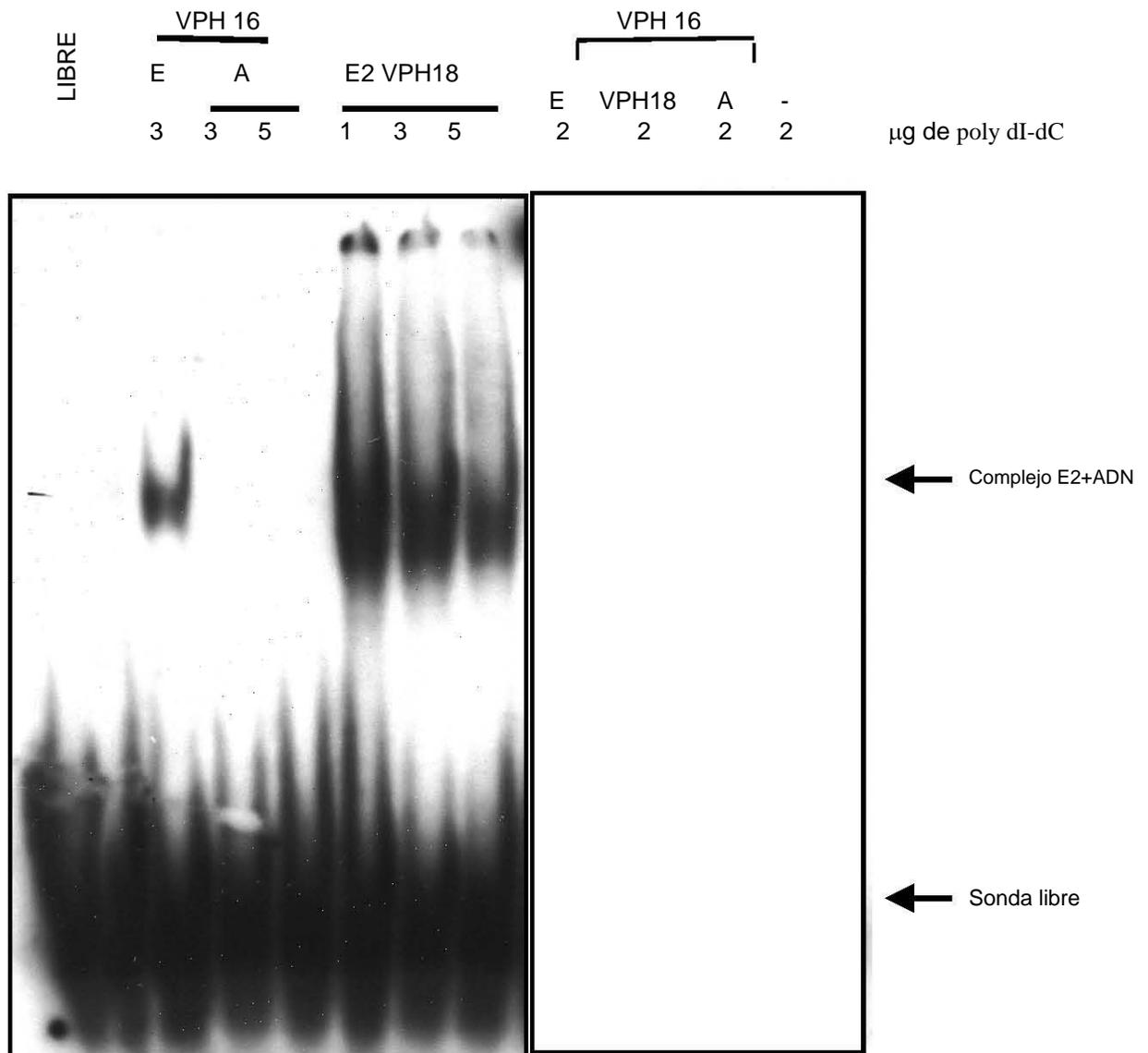
Cuando se incubó el oligonucleótido E2-38/32 con las diferentes proteínas E2 obtenidas mediante transcripción-traducción en reticulocitos de conejo, con 1  $\mu\text{g}$  de Poly dl-dC, se observa la formación de dos complejos (Figura. 9), el complejo ADN+reticulocitos inespecífico y un complejo ADN+E2. Como control positivo utilizamos E2 de VPH18. Como se puede observar en la Figura 9, la banda retardada de la proteína E2 Europea es más intensa, lo que sugiere que esta intensidad se debe a unión más fuerte de ésta proteína a su secuencia blanco, que la proteína E2 perteneciente a las variantes AA-a o AA-c, en las que se observa una banda menos intensa. La banda del complejo ADN+reticulocitos esta formada por alguna proteína inespecífica presente en el lisado de reticulocitos como se observa en el penúltimo carril de la Figura 9 en el que solo se agregó lisado a la sonda marcada radioactivamente.



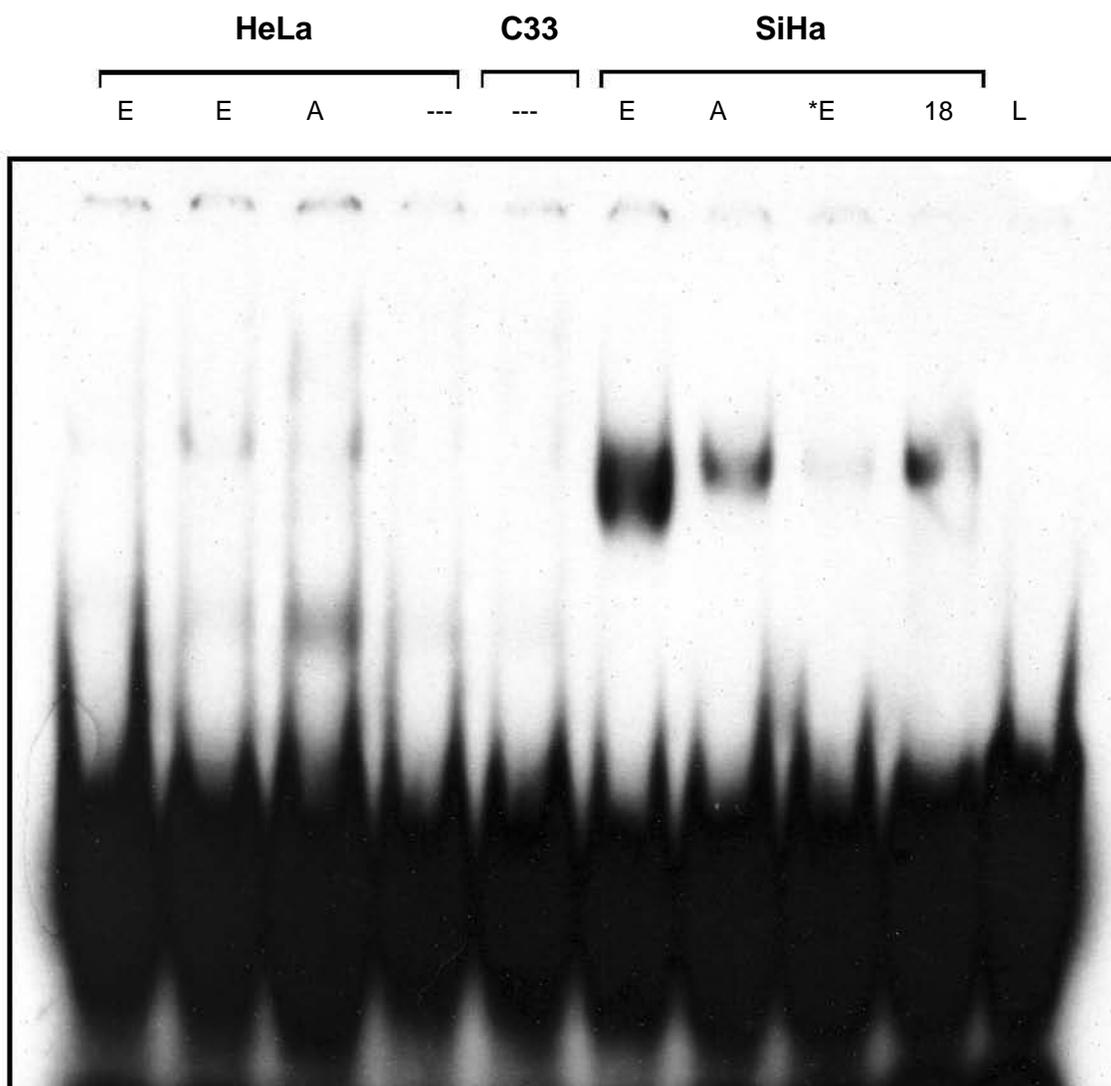
**Figura 9. La proteína E2 Europea, se une más fuertemente a su secuencia blanco, que la proteína E2 Asiático Americana AA-a y AA-c.** Ensayo de retardamiento DNA/proteína; la sonda de reconocimiento a E2 (38/32) marcada con  $^{32}\text{P}$  se incubó con la proteína E2 traducida "in vitro" en reticulocitos. El primer carril contiene la proteína E2 de VPH18. (E) proteína E2 Europea, (A) proteína E2 AA-a y (C) proteína E2 AA-c. Se utilizaron 2  $\mu\text{l}$  de E2 transcrito en reticulocito y 1  $\mu\text{g}$  de Poly dl-dC en todos los carriles. Se observa el complejo de una banda retardada por los reticulocitos que es inespecífica y la banda retardada por la proteína E2 de VPH.

Para tratar de eliminar la banda inespecífica proveniente del lisado de reticulocitos de conejo, se realizó un ensayo de retardamiento (Figura 10), utilizando 2, 3 y 5 $\mu$ g de Poly dl-dC, este compuesto es un copolímero de doble cadena compuesto de dideoxy-inosinas y dideoxy-citidinas, se usa para bloquear la unión no específica de proteínas al DNA de interés, facilitando la detección del complejo específico, por lo que esperábamos que se observara una banda retardada más clara y específica de ADN-E2, conforme se agregara más Poly dl-dC. Observamos que la banda inespecífica de reticulocitos desaparece, pero también se ve una disminución en la intensidad de la banda retardada que contiene la proteína E2 AA-a, lo que sugiere que la unión de ésta proteína es débil con respecto a la proteína E2 E, que mantiene una banda a pesar de usar 3  $\mu$ g de Poly dl-dC. Para tratar de analizar la expresión y unión de las diferentes proteínas E2 a su secuencia blanco en un sistema diferente y tratar de discernir si factores inespecíficos del lisado de reticulocitos estuviera compitiendo por la unión al ADN con E2, se transfectaron los vectores que expresan E2 (plásmidos pE2 E, AAa y AA-c: Figura 8a) en dos líneas celulares derivadas de cáncer Cérvico-Uterino, HeLa y SiHa. El extracto celular, se incubó con la sonda marcada llevando a cabo el ensayo de retardamiento y se corrió en un gel de acrilamida, como se puede observar (Figura 11) la proteína E2 variante AA-a, se une más débilmente a su secuencia blanco que la Europea, tal como habíamos observado en sistema de reticulocitos. En células SiHa la expresión de la proteína E2 es más fuerte que en células HeLa, esto puede ser debido a eficiencia de transfección, o a expresión diferencial de los vectores, sin embargo se observa una banda más intensa en el carril en el que se utilizó la proteína E2 Europea, esto siguió sugiriendo que se une más fuertemente independientemente de la cantidad de Poly dl-dC. Como control del experimento se utilizó extracto de la línea celular C33 derivada CaCU que no tiene VPH.

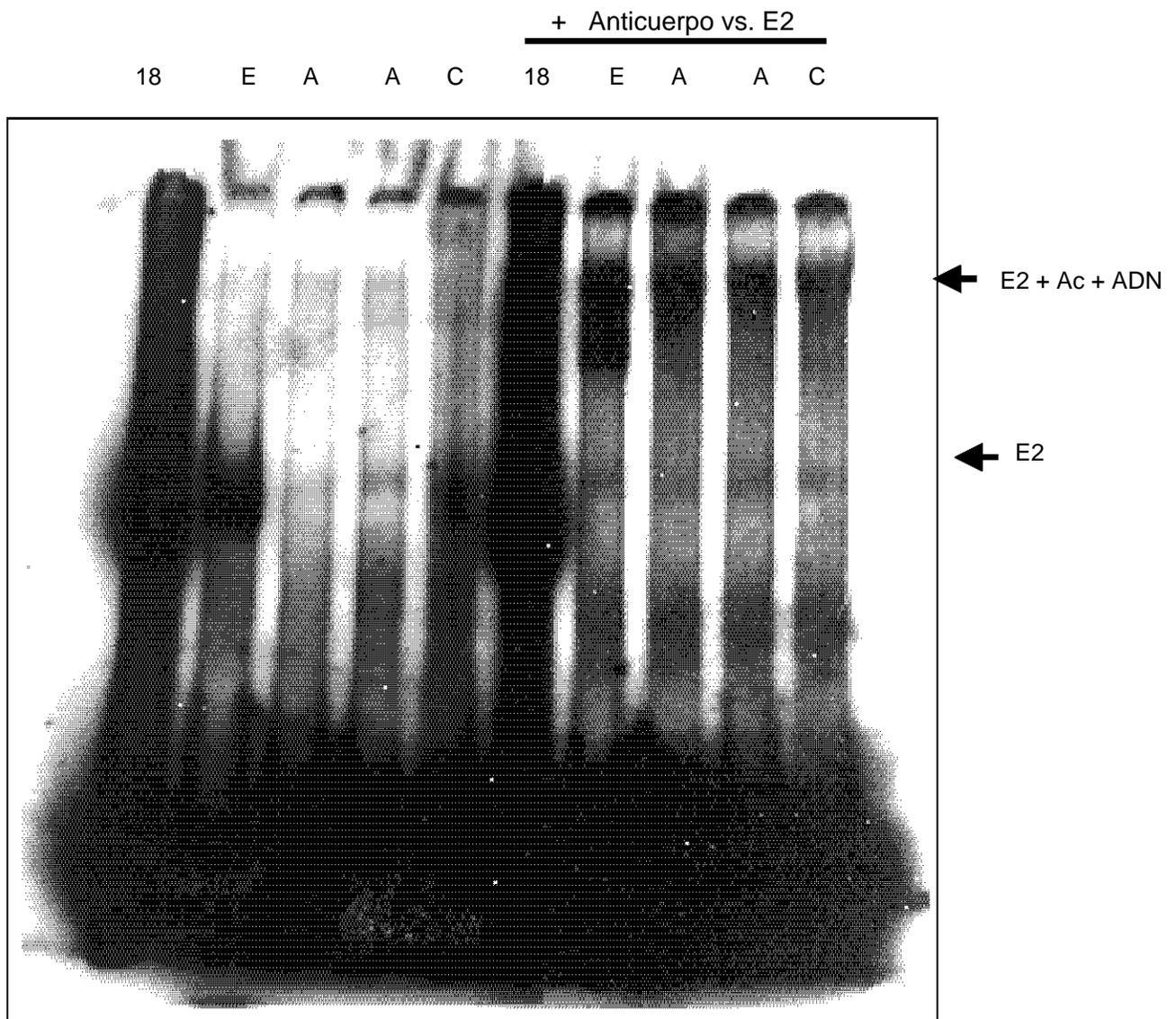
Para corroborar que las diferentes proteínas E2 se estaban uniendo a su secuencia blanco, y descartar que se tratara de algún otro factor inespecífico, se hizo un ensayo de súper-retardamiento utilizando un anticuerpo específico que reconoce el dominio C-terminal de la proteína E2 de VPH16 donado por la Dra. Leticia Rocha, este anticuerpo se obtuvo inyectando lisados de la proteína E2 (expresada en *Escherichia coli*), en conejos (Rocha-Zavaleta et al., 1997). Utilizamos lisado nuclear de células C33 transfectadas con plásmidos que expresan las proteínas E2 (pE216-E, AAa y AAC, Figura 8a) y el extracto se incubó con la sonda marcada por 10 min, agregando después 1ul de anticuerpo e incubando 30 min adicionales. Se puede observar claramente en la Figura 12, la banda super-retardada perteneciente al complejo ADN-E2-Anticuerpo, este resultado nos aseguró que la proteína está en conformación que nos permitió la dimerización y la unión a su sitio específico.



**Figura 10. La Unión de la proteína E2 HPV16 AA-a unida a su secuencia blanco, desaparece al agregar 2, 3 y 5 μg de Poly dI-dC, al igual que la banda inespecífica de reticulocitos.** Ensayo de retardamiento ADN/proteína; la proteína E2 traducida "in vitro" (reticulocitos) se incubó con la sonda de reconocimiento a E2 (38/32) marcada con <sup>32</sup>P. Se utilizaron 2 μl de proteína E2 traducida "in vitro", en todos los carriles (E) Proteína E2 Europea, (A) Proteína E2 AA-a, (18) proteína E2 de PVH18. Se utilizaron 1, 2, 3 y 5 μg de Poly dI-dC como se muestra en la Figura. 2 μg son suficientes para evitar la unión de E2 AA-a al DNA, mientras que 3 μg no inhiben la unión de E2 E.



**Figura 11. La banda de la proteína E2 AA-a es menos intensa que la proteína E2 E en líneas celulares, lo que sugiere que la unión al DNA es más débil.** La proteína E2 se obtuvo de extractos celulares transfectados con un vector plasmídico que la expresa. Se utilizaron 3  $\mu\text{g}$  de extracto celular y 3  $\mu\text{g}$  de Poly dl-dC en todos los carriles. (E) Proteína E2 Europea, (\*E) se utilizó un vector plasmídico diferente del que se muestra en los carriles mostrados como E durante la trasfección. (A) Proteína E2 AA-a, (18) proteína E2 de PVH18. Se observa la banda retardada por la proteína E2 de VPH.



**Figura 12. Unión específica de la proteína E2+DNA+Ac en la proteína E2 E, E2 AA-a y E2 AA-c. Super-retardo de la proteína E2 con su secuencia blanco.** La proteína E2 se obtuvo de extractos de células C33 transfectadas con un vector plasmídico que la expresa. Se utilizaron 2 µg de extracto celular y 1 µg de pdl-dC en todos los carriles. En los carriles señalados se agregó 1 ul de anticuerpo anti-E2 de VPH16 anti conejo. (E) Proteína E2 E. (A) Proteína E2 AA-a (C) Proteína E2 de AA-c, (18) proteína E2 de PVH18 . Se observa la banda retardada por la proteína E2 de VPH y la banda super-retardada con el anticuerpo.

### 6.3 Actividad basal de los LCRs.

En la Tabla No. 2 se muestra una revisión de los cambios nucleotídicos presentes en la región reguladora de las variantes AA, (Kammer et al., 2000; Veress et al., 1999; Yamada et al., 1997). La región larga de control (LCR) de las variantes AA tienen incrementada su actividad transcripcional 1.7 veces comparadas con aislados Europeos. Se sugiere que esta actividad incrementada puede deberse a los cambios nucleotídicos encontrados en el final 3' del LCR (nt 3660-7890).

El cambio en el nucleótido 7689 se ubica en el sitio de reconocimiento a TEF-1, el cual puede cambiar la habilidad de este factor para unirse a este sitio, el cambio en el nucleótido 7743 se ubica en el sitio de reconocimiento del regulador Oct-1 NF1 una base se cambia, entre dos bases que actúan como espaciadoras, el cambio en el nucleótido 7786 se observa entre dos sitios de unión para YY1, sin embargo haciendo mutaciones puntuales de cada uno de estos sitios en un vector que contiene el LCR de VPH16 E, no existió un cambio estadísticamente significativo en la actividad transcripcional (Veress et al., 1999), otros sitios de reconocimiento que no se han investigado son el 7339 y 7394, que se encuentran en el sitio de reconocimiento a GRE 1, 7485 y 7489 en el sitio GRE 2, el sitio 7521 en YY1, 7689 en Tef-1 y 7764 en el sitio NF1. Por medio de mutagénesis dirigida se determinó que la actividad aumentada de AA-c en el LCR se debió al cambio en el nucleótido 7729 (presente también en variante NA1) (Kammer et al., 2000). Cabe mencionar que no se han analizado las regiones exclusivas de cada subclase en las que puede residir la diferencia de expresión entre la LCR E, AA-a, AA-b y AA-c.

	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	
	0	1	2	3	3	3	4	4	5	5	6	6	7	7	7	7	7	8	8	8
	6	9	3	3	9	9	8	8	0	2	6	8	2	4	5	6	8	2	8	9
	0	3	3	9	4	5	5	9	7	1	9	9	9	3	2	4	6	7	6	4
<b>E</b>	<b>G</b>	<b>G</b>	<b>A</b>	<b>A</b>	<b>C</b>	<b>C</b>	<b>A</b>	<b>G</b>	<b>A</b>	<b>G</b>	<b>C</b>	<b>C</b>	<b>A</b>	<b>T</b>	<b>A</b>	<b>C</b>	<b>C</b>	<b>T</b>	<b>C</b>	<b>A</b>
AA-a	T	T	C	T	T	T	C	A	-	A	T	A	C	G	C	T	T	.	G	.
AA-c	T	T	C	T	T	-	C	A	G	A	T	A	C	-	-	T	T	C	G	C
AA-b						T?	C	A	-	A	T	A	C	G	-	T	T			
Na1	T	T	C	-	T	-	C	A	-	A	T	A	C	-	-	T	T	-	-	-
Factore																				
S		Tef1			Gre 1		Gre 2		YY1		Tef1		OCT		NF1		NF1			

**Tabla 3. Resumen de cambios nucleotídicos en la LCR de las variantes AA-a, AA-b, AA-c y NA1.** Se muestra el nucleótido correspondiente al cambio con respecto al LCR del VPH 16 Europeo (Kammer et al., 2000; Veress et al., 1999; Yamada et al., 1997) Algunos de los cambios nucleotídicos se ubican en sitios de reconocimiento a factores celulares. Se marcaron en colores los cambios específicos de cada variante. Verde claro; variante a y c. Azul; variante a y b. Naranja: variante c. Amarillo; variante a.

Debido a que existen cambios nucleotídicos en el LCR de las variantes (Kammer et al., 2000; Veress et al., 1999; Yamada et al., 1997) con respecto al LCR del VPH E, y existen diferencias (tres cambios diferentes entre AA-a y AA-c, nucleótidos 7507, 7725 y 7827) exclusivas de cada variante, lo primero que se analizó fué la actividad basal de los vectores que contienen la región promotora y de regulación (LCR) de la transcripción del papilomavirus VPH16 y las variantes AA-a y AA-c (p16LCRwt, p16LCRAA-a y p16LCRAA-c) estos cambios se encuentran cerca o dentro de sitios de unión para proteínas celulares. Se clonó la LCR de VPH16, AA-a y AA-c en el vector pBLCAT6 (Figura 8b) que contiene al gen reportero CAT. Se transfectaron células C33 con estos vectores y se midió la actividad de la enzima CAT.

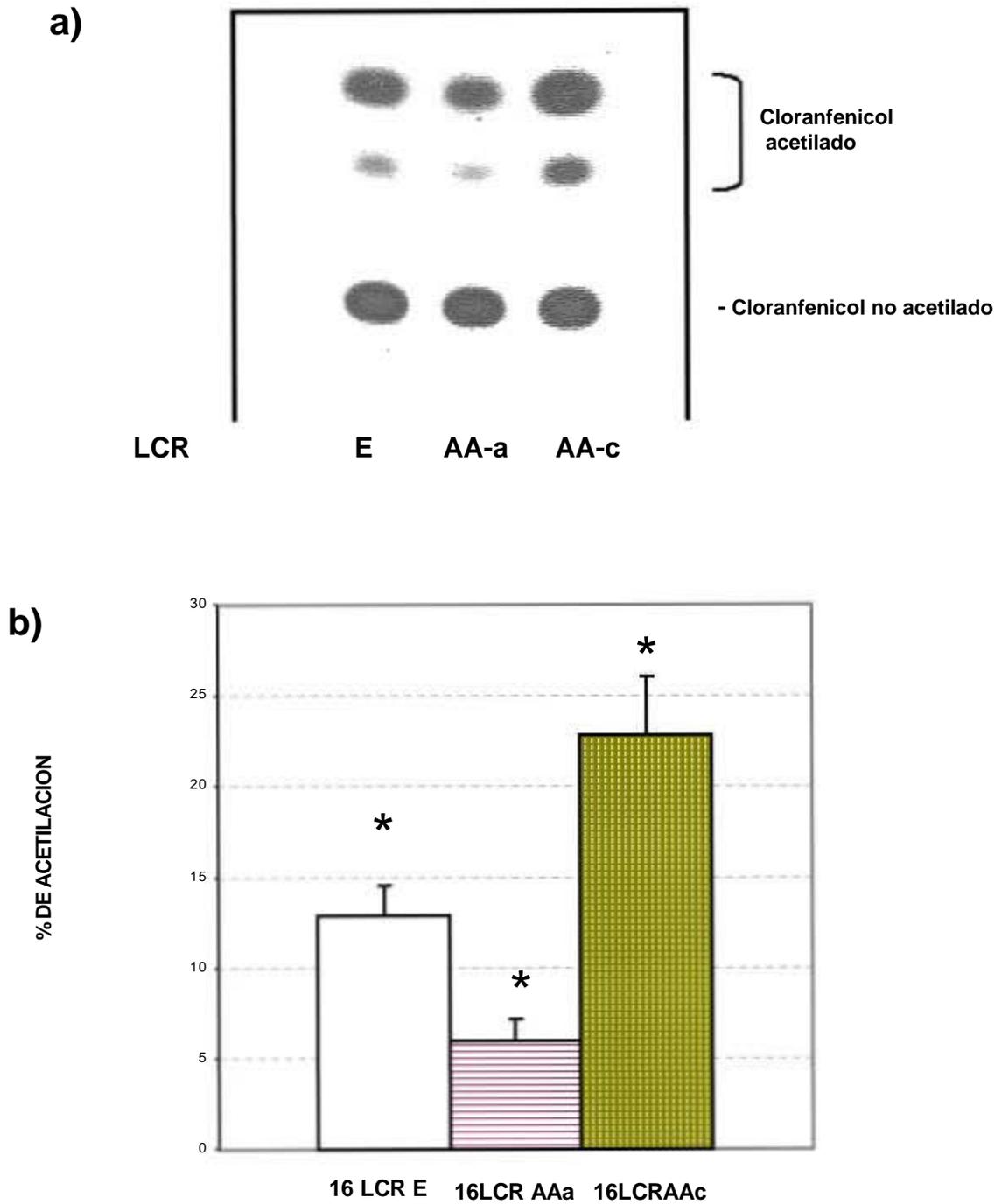
CAT inactiva el cloranfenicol por la formación de derivados mono y di-acetilados, en este trabajo se realizó el ensayo en el cual la acetilación del cloranfenicol marcado con  $^{14}\text{C}$  es medida por cromatografía de capa fina en silica gel. Este ensayo es muy sensible y altamente específico, se separan fácilmente las formas mono y di-acetiladas que se cortan y se cuentan por separado en líquido de centelleo. Estas cuentas por minuto son expresadas en números que nos permiten identificar la actividad transcripcional de cada promotor obteniendo el porcentaje de cuentas del cloranfenicol acetilado. Se hicieron tres experimentos de cada promotor, se obtuvo la media y la desviación estandar (Tabla 4) que se graficaron en un histograma (Figura 13).

Con la prueba de Turkey se compararon las actividades transcripcionales entre grupos pareados de variantes virales y se consideraron valores de  $P < 0.050$  como diferencias significativamente diferentes. Los análisis se llevaron a cabo con el software SigmaStat, versión 3.1.

Como se observa en la Tabla 4 y Figura 13, la actividad basal del gen reportero CAT en el vector p16LCR AA-c es la mas alta de las tres construcciones, aproximadamente 3.8 veces mayor a la actividad basal en el vector p16LCRAA-a (LCR de VPH16 Variante AA-a), y 1.7 veces mas alta que el vector p16LCR E y la actividad basal del vector p16LCR E es 2.1 veces más alta que la del vector p16LCRAA-a. Las diferencias entre las tres actividades basales resultaron ser estadísticamente significativas ( $P < 0.050$ ).

	MEDIA PORCENTAJE DE ACETILACION	DESV. EST.
16LCR E	12.76	1.74
16LCR AA-a	5.9	1.3
16LCRAA-c	22.8	3.24

**Tabla 4. Porcentaje de acetilacion de la enzima CAT, que reflejan la actividad transcripcional basal de cada promotor.** Se obtuvo la media y desviación estandar de tres experimentos independientes. La Figura, 13 muestra un ensayo autoradiografico representativo y un histograma con los resultados de la actividad basal.



**Figura 13.** La actividad basal de la LCR AA-c es mas alta que la actividad basal de la LCR E y AA-a, mientras que la actividad basal de la LCR de AA-a es menor que para las otras dos. **(a)** Autoradiografía de ensayo CAT. Expresión basal de la región larga de control Europeo (16LCR-E), variante AA-a (16LCR AA-a) y variante AA-c (16LCR AA-c). **(b)** Histograma que muestra, el porcentaje de acetilación del gen reportero CAT en los plásmidos que contienen la región reguladora de VPH16 E, AA-a y AA-c. El resultado es el promedio de tres experimentos independientes ( $P < 0.050$  con la prueba de Turkey de comparaciones múltiples).

#### 6.4 Efecto de E2 sobre la actividad de los LCR

Para poder determinar la actividad transcripcional de las proteínas E2 de VPH16 E, AA-a y AA-c en sus propios LCR's, utilizamos los vectores de pCDNA3 con los diferentes genes E2 clonados frente al promotor de Citomegalovirus, (Figura. 7a) y los vectores que contienen los diferentes LCRs (Figura. 7b) cotransfectados en células C33.

Se ha reportado que la proteína E2 de VPH16 E, reprime la transcripción de su propio promotor en varias órdenes de magnitud, para determinar si las proteínas E2 de las variantes también presentan la misma actividad llevamos a cabo cotransfecciones cruzadas entre los diferentes vectores que expresan las proteína E2 y los LCRs, tanto Europeas como variantes.

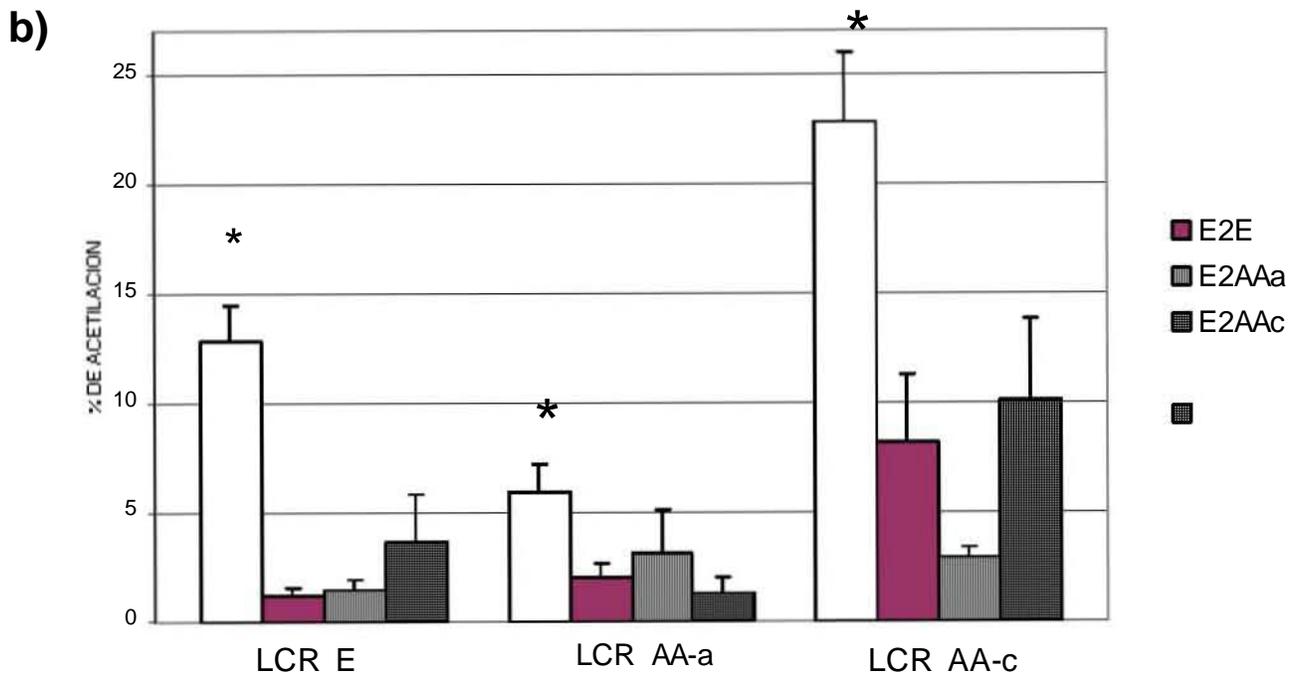
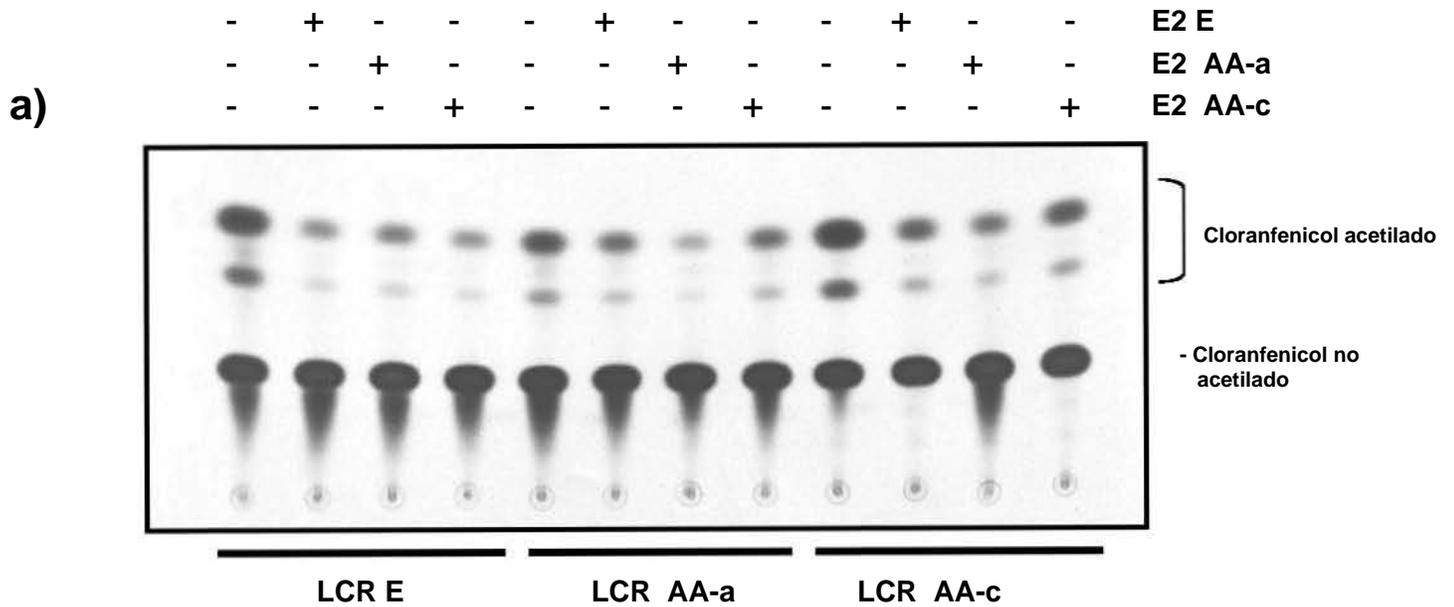
De acuerdo con los resultados reportados en la Tabla No 4, Figura 14, se puede deducir que las proteínas E2 E, AA-a y AA-c están reprimiendo la transcripción tanto en la región reguladora propia como en la cruzada (LCR E, AA-a y AA-c) a diferentes magnitudes.

Se compararon las actividades transcripcionales de cada promotor por separado con respecto a cada proteína E2 con la que fue transfectado por la prueba de Turkey y se consideraron valores de  $P < 0.050$  como diferencias estadísticamente significativas. Los análisis se llevaron a cabo con el software SigmaStat, versión 3.1 y se analizaron los resultados de tres experimentos independientes.

Como se puede observar en la Figura 14 la actividad basal de la LCR E es reprimida en un 90% por la proteína E2 E, 88.3% por la proteína E2 AA-a y 71.7% por la proteína E2 AA-c, estas diferencias son significativas de acuerdo a la prueba de Turkey ( $P < 0.050$ ), sin embargo no existen diferencias significativas si se comparan las actividades represoras de las diferentes construcciones con las proteínas E2 variantes y Europea. La actividad basal de la LCR A es reprimida en un 66.2% por la proteína E2 E, y un 78% por la proteína E2 AA-c diferencias que son estadísticamente significativas (Prueba de Turkey  $P < 0.050$ ), mientras que es reprimida solo un 47.5% por la proteína E2 AA-a en la que no se encontró diferencia significativa. Y por último la actividad basal de la LCR AA-c es reprimida en un 64.1% por la proteína E2E, hasta 87.3% por la proteína E2 AA-a y 55.7% por la proteína E2 AA-c, los valores de la represión de la proteína E2 en este promotor también representan una diferencia significativa con respecto a la actividad basal del propio promotor (Prueba de Turkey de comparaciones múltiples de  $P < 0.050$ ).

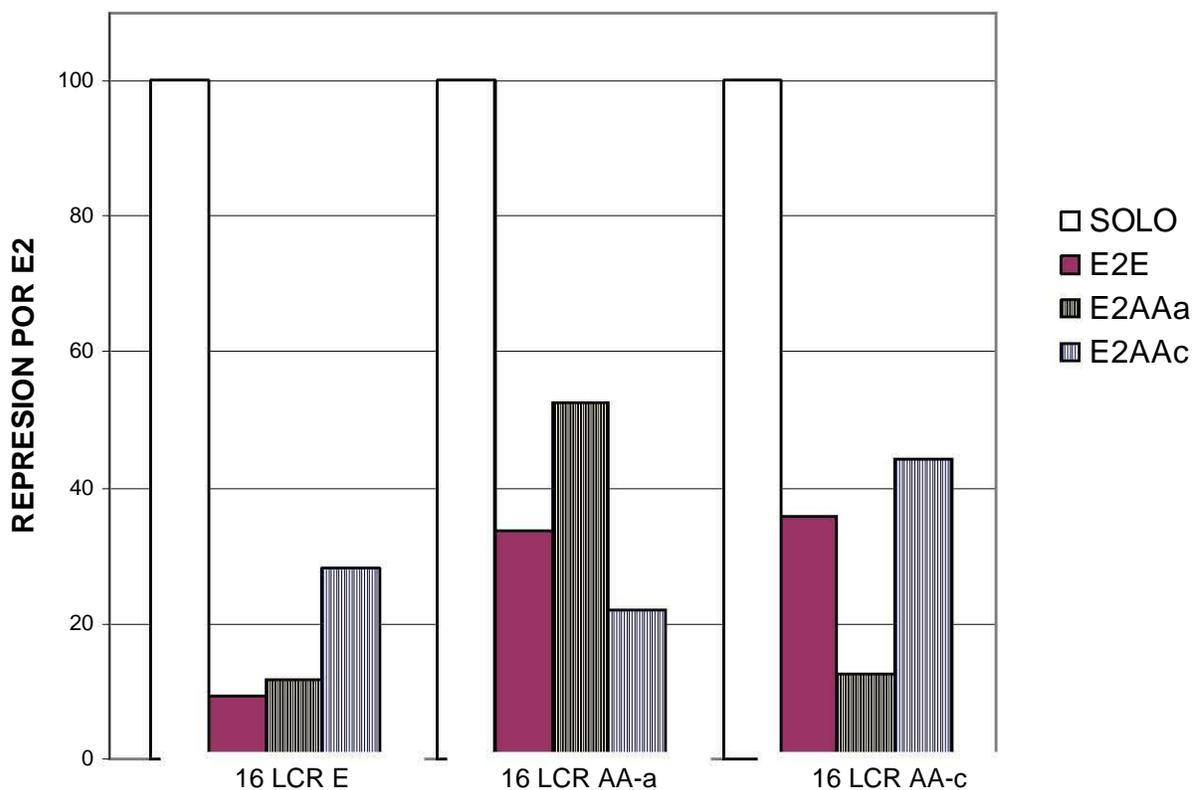
	PROMEDIO DEL PORCENTAJE DE ACETILACION			
	BASAL	E2E	E2AA-a	E2AA-c
16LCR E	12.8	<b>1.2</b>	1.5	3.6
16LCR AA-a	5.9	2	<b>3.1</b>	1.3
16LCRAA-c	22.8	8.2	2.9	<b>10.1</b>

**Tabla 5. Porcentaje de acetilacion de la enzima CAT, que reflejan la actividad transcripcional basal de cada promotor y su actividad al cotransfectarlo con los diferentes vectores que expresan la proteína E2.** La Figura, 13 muestra un ensayo autoradiografico representativo y un histograma con los resultados de la actividad basal y la Figura 14 muestra la actividad cotransfectando con las proteínas E2 de las variantes.



**Figura 14. La actividad de cada LCR es reprimida cuando el vector reportero es cotransfectado con los vectores de expresión para E2. (a)** Autoradiografía de ensayo CAT representativa. Se describe el VPH16 Europeo (16LCR-E), VPH16 variante AA-a (16LCR AA-a) y VPH 16 variante AA-c (16LCR AA-c). **(b)** Histograma que muestra, el porcentaje de acetilación del gen reportero CAT en los plásmidos que contienen la región reguladora de VPH16 E, AA-a y AA-c. Se observan diferencias significativas de represión entre cada LCR cuando se cotransfecta con E2, ( $P < 0.050$ ) a excepción del LCR AA-a que no muestra una represión significativa cuando se cotransfecta con su propia proteína E2 AA-a.

Debido a que cada LCR tiene una actividad basal diferente normalizamos la actividad basal de cada promotor a 100% para conocer la actividad real de cada proteína E2 (Figura. 15), se puede observar u que la proteína E2 E, reprime mejor su propio promotor que el promotor de AA-a y AA-c. La proteína E2 AA-a, reprime menos su promotor y más LCR AA-c y LCR E, y la proteína E2 AA-c reprime mejor al LCR AA-a y seguido por la LCR E y muy poco la LCR AA-c. Con respecto a sus propios LCRs, la proteína E2 de LCR E, reprime más su propio LCR que las otras dos proteínas. La proteína AA-a reprime mejor a los LCRs E y AA-c que a su propio promotor, y la proteína E2 AA-c reprime mejor a la LCR AA-a y E que su propio promotor. En conclusión, se sugiere que las proteínas E2 de las variantes no están reprimiendo sus propios promotores a la misma magnitud que la proteína E2 de VPH16 E. Sin embargo haciendo análisis estadísticos se observa que no existen diferencias significativas entre las actividades de las E2 variantes y Europea, por lo que pudieramos concluir que los cambios en las proteínas no repercuten en su actividad represora o que son necesarios más experimentos para poder llegar a una conclusión más acertada.



**Figura. 15. Normalizando a 100% el porcentaje de acetilación del gen reportero CAT, de cada uno de los LCRs (VPH16 E, AA-a y AA-c) se observa la actividad real de cada una de las proteínas E2. El resultado es el promedio de tres experimentos independientes, la transfección se llevó a cabo con 3µg de ADN plasmídico en células C33.**

## 7. DISCUSION

Nosotros quisimos analizar en detalle al gen E2 porque los productos de este gene tienen papeles significativos en la regulación de la transcripción de las oncoproteínas virales E6 y E7 en la LCR. Además la proteína E2 está involucrada en replicación del ADN episomal del virus. Pensamos que ciertos cambios genéticos (deleciones o mutaciones puntuales) en el gen E2 pudieran alterar la regulación de la transcripción de los promotores tempranos del virus. En este trabajo presentamos el análisis de variación de la secuencia del gen E2 de 4 muestras positivas a HPV16 dos de ellas pertenecientes a HPV16 Europeo (MX 119 y 40) y dos a variantes Asiático Americanas subclase AA-a (MX68) y AA-c (MX70), estas variantes de E2 tienen cambios de bases distribuidos en el gen, incluyendo los dominios de transactivación y de unión al ADN. En un estudio analizando más muestras y las secuencias de L1 (Casas et al., 1999) se determinó que eran pertenecientes a la clase Asiático-Americana sub-clase AA-a y AA-c (Yamada et al., 1997).

Estos cambios revelan la distancia genética entre las variantes AA-a y AA-c, esta cantidad de cambios no se observan en las regiones L1, MY y LCR, en el gen E6 la secuencia de las dos subclases es idéntica, mientras que en el L1, MY y LCR difieren en solo dos bases, en E2 existen 5 diferentes posiciones entre las sub-clases AA y una mutación más en AA-a con respecto a AA-c.

En estudios anteriores se demostró que a las variantes Asiático-Americanas se les asocia por presentar alto número de copias, integradas al ADN hospedero en tandem o como episomas multiméricos (Berumen et al., 1994).

Las substituciones encontradas en los aminoácidos de E2 pudieran estar alterando su actividad molecular. Tres de los cambios (T310K, W3641C y D344E) presentes en AA-a y AA-c, se encuentran en el dominio de unión al DNA el cual es esencial para la función regulatoria de E2. De especial importancia es el cambio de una treonina a lysina en la posición 310, justo a un lado de la hélice de unión al ADN (aa 292-309), pudiera ser que se esté alterando la conformación de esta hélice y la capacidad de unirse al DNA se vea afectada.

En la región N2 amino-terminal de la proteína E2 de las variantes se encuentran 7 cambios de aa, justo en la región en la que E2 se une a TFIIB y comprende los aminoácidos 134 a 216 (Breiding et al., 1997) y los cambios en los aminoácidos 142 y 143 se encuentran en sitios de dimerización. Estos cambios pudieran estar afectando la actividad de unión a TFIIB y de dimerización, alterando su actividad transcripcional.

Se sabe que la proteína E2 es blanco de respuesta inmune humoral y celular (Dillner et al., 1996; Konya et al., 1997). Existen varios epítomos humorales en la proteína E2, uno en el dominio de transactivación (121-140), tres en la región de bisagra (181-200, 241-260 y 271-290) y uno en el dominio de unión al ADN (328-346). Todas excepto una de estas regiones (181-200) contienen cambios en las variantes AA, lo que sugiere que la respuesta humoral a estos epítomos pudiera ser diferente entre E2 Europea y las variantes AA. Terry y colaboradores (Terry et al., 1997), describieron las variantes Europeas de E2 asociadas con cambios en epítomos inmunodominantes en la región de bisagra, 11 de

estos cambios en codones se encuentran en las variantes AA. La infección con variantes AA de PVH16 pudiera explicar la pobre correlación entre la presencia de anticuerpos contra E2 de HPV16 E y la detección de ADN de HPV16 (De Sanjose et al., 1996).

En cuanto a la expresión de las proteínas E2, y su unión al ADN mediante geles de retardo, se sugiere que la proteína E2 Europea se une más fuertemente a su secuencia blanco en el DNA que la proteína E2 AA, esta observación correlaciona con el cambio T310L, observada en la secuencia de la proteína E2. No se han publicado ensayos de retardamiento con las proteínas variantes y tampoco se han hecho análisis de unión al ADN, se hacen necesarios más ensayos de este tipo. De la misma manera hicieron falta ensayos de competencia con sonda fría.

Veres y cols (1999) reportaron que el aislado 33, que pareciera ser perteneciente a la variante AA-c, tiene incrementada la actividad basal de su LCR 1.7 veces comparada con VPH16 E, estos resultados son similares a los obtenidos en este estudio en el que se observa que la actividad basal del promotor AA-c es 1.6 más alta que la actividad basal del LCR E, sin embargo ellos no realizaron experimentos para determinar la actividad transcripcional de la variante AA-a, y no hicieron ensayos para demostrar la actividad de cada LCR con su proteína E2 correspondiente.

Kammer y colaboradores (2000) hicieron mutagénesis sitio-dirigida al nucleótido 7729 en el LCR de las variantes AA-c y NA1 revirtiéndola de A/C y demostraron que la actividad basal se redujo comparada con las clonas originales AA-c y NA1, con niveles de expresión comparables a la del LCR E, sin embargo este cambio se encuentra en todas las variantes analizadas hasta la fecha (Tabla 2) incluyendo la variante AA-a que tiene actividad mas baja, nosotros sugerimos que este cambio no está afectando a la variante AA-a, Veres y colaboradores (1999) demostraron que la actividad enhancer de la variante AA que analizaron se localizaba hacia la región 3' del LCR, después del nucleótido 7722, estos resultados correlacionan con lo encontrado por Kammer (Kammer et al., 2000).

Nosotros observamos que la actividad basal de la LCR AA-a es 3.8 veces menor que la LCR AA-c y 2.1 veces menor que la LCR E, se necesitan ensayos de mutagénesis dirigida en la LCR de la variante AA-a para determinar cual cambio de nucleótido está contribuyendo a que la actividad basal sea mas baja comparada con la LCR E y la LCR AA-c. Yamada y colaboradores (1997) reportan un cambio en el nucleótido 7752 que se encuentra únicamente en el LCR AA-a, este cambio ocurre cerca del sitio reconocimiento para la proteína Oct1 y NF1.

Nosotros demostramos que la construcción con LCR E es mejor reprimida por su propio E2 y menos por E2 AA-a y AA-c, condición que concuerda con Ordoñez (Ordonez et al., 2004) sin embargo ellos no analizaron la actividad de cada proteína E2 en el contexto de su propio promotor, la principal diferencia entre la subclase AA-a y AA-c a nivel de la proteína E2 es que en AA-c se cambia una histidina por una tirosina en la posición 136 en el extremo 3' del dominio de transactivación, tal vez en este sitio erradique la diferencia de actividad entre las dos subclases aunado a que la LCR de la subclase AA-a también tiene una actividad basal baja.

Mientras que para VPH16 E la disrupción de E1/E2 ocurre más frecuentemente que en las variantes AA (Casas et al., 1999) lo que demuestra que el genoma viral de E permanece en la capa basal de células indiferenciada. En contraste en células bien diferenciadas, el genoma viral de E que contiene E1/E2 intactos pudieran inactivar a E2 por “splicing”, y en este caso se observan transcritos del gen E1^E4. En la proteína E2 de las variantes AA no existiría esta presión selectiva que induzca la disrupción o “splicing” de E2. Por lo que la aumentada oncogenicidad de las variantes AA pudiera estar asociada con el transcrito E6^E7 encontrado por Ordoñez (Ordonez et al., 2004) y con la baja represión de E2 al promotor p97 de las variantes.

Ordoñez y colaboradores (2004) demostraron que 20 muestras que tienen E1/E2 intacto tienen mayor cantidad de proteínas E6/E7. Esto se puede deber a que la proteína E2 de AA-a y AA-c se une menos fuertemente a su sitio blanco que E2 E, y reprimen menos sus propios promotores que la proteína E2 E.

## 8. CONCLUSIONES

- Por análisis de secuencia del DNA se encontraron 23 cambios en el gen E2 de la variante AA-a y 24 cambios en la variante AA-c estos cambios se encuentran en sitios importantes para la dimerización y unión al DNA, para la unión de factores de transcripción y en epítomos de la proteína,
- Por los resultados obtenidos con geles de retardo se sugiere que la proteína E2 de las variantes se une más débilmente a su secuencia blanco que la proteína E2 Europea.
- Usando transfecciones con un gen reportero en células C33A, se observó aumento en la transcripción de la LCR en el caso de AAc, que ya estaba reportado y una disminución en la transcripción en la LCR de AA-a, se hace necesario hacer un análisis más detallado de estas diferencias particularmente del cambio en el nucleótido 7752 (cambio de A por C) en la LCR de la variante AA-a.
- La variación en la secuencia del gen E2 no parece ser un mecanismo que altere la transcripción del promotor p97, a excepción de la proteína E2 de AA-a que reprime la transcripción de la LCR AA-a solo un 47.5%, es decir a muy baja magnitud comparada con la proteína E2 E y AA-c.

## 9. LISTA DE FIGURAS Y TABLAS

**Figura 1.** Mapa genómico del Papilomavirus Humano tipo 16 (VPH16)

**Figura 2.** Mapa genómico del VPH 16 y esquema de regulación de la transcripción por E2.

**Figura 3.** Arbol filogenético de las variantes de VPH16 basado en analisis de parsimonia.

**Figura 4.** LCR de VPH16 y factores que reconoce.

**Figura 5.** Vista estereoscópica de la súper imposición de las estructuras de E2 VPH16, con el ADN en la conformación encontrada en el complejo de E2 BPV-1 al ADN.

**Figura 6.** Vista estereoscópica del dímero amino-terminal de E2 VPH16)

**Figura 7.** Material y Métodos.

**Figura 8.** Plásmidos que se utilizaron para la realización de este trabajo.

**Figura 9.** La proteína E2 Europea, se une más fuertemente a su secuencia blanco, que la proteína E2 Asiático Americana AA-a y AA-c.

**Figura 10.** La Unión de la proteína E2 HPV16 AA-a unida a su secuencia blanco, desaparece al agregar 2, 3 y 5 ug de Poly dl-dC, al igual que la banda inespecífica de reticulocitos.

**Figura 11.** La banda de la proteína E2 AA-a es menos intensa que la proteína E2 E en líneas celulares, lo que sugiere que la unión al DNA es más débil.

**Figura 12.** Unión específica de la proteína E2+DNA+Ac en la proteína E2 E, E2 AA-a y E2 AA-c. Super-retardo de la proteína E2 con su secuencia blanco.

**Figura 13.** La actividad basal de la LCR VPH16 AA-c es más alta que la actividad basal de la LCR VPH16 E y AA-a, mientras que la actividad basal de la LCR de VPH16 AA-a es menor que para las otras dos.

**Figura 14.** La actividad de cada LCR es reprimida cuando el vector reportero es cotransfectado con los vectores de expresión para E2.

**Figura 15.** Normalizando a 100% el porcentaje de acetilación del gen reportero CAT, de cada uno de los LCRs (VPH16 E, AA-a y AA-c) cotransfectados con los plásmidos que expresan las diferentes proteínas E2 se observa la actividad real de cada una de las proteínas.

**TABLAS:**

**Tabla 1.** Tipos de Papilomavirus Humanos, y lesiones de las cuales fueron aislados.

**Tabla 2.** Variaciones en la secuencia de ADN y aminoácidos de E2 en las variantes VPH16.

**Tabla 3.** Resumen de cambios nucleotídicos publicados en la LCR de las variantes AA-a, AA-b, AA-c y NA1.

**Tabla 4.** Porcentaje de acetilacion de la enzima CAT, que reflejan la actividad transcripcional basal de cada promotor.

**Tabla 5.** Porcentaje de acetilacion de la enzima CAT, que reflejan la actividad transcripcional de cada promotor cotransfectado con los diferentes vectores que expresan la proteína E2.

## 10. REFERENCIAS

Lazcano-Ponce E, Alonso P, Ruiz-Moreno JA, Hernandez-Avila M (2003) Recommendations for cervical cancer screening programs in developing countries. The need for equity and technological development. *Salud Publica Mex* **45 Suppl 3**: S449-462.

Nelson JH, Jr., Averette HE, Richart RM (1984) Dysplasia, carcinoma in situ, and early invasive cervical carcinoma. *CA Cancer J Clin* **34**: 306-327.

Koutsky L (1997) Epidemiology of genital human papillomavirus infection. *Am J Med* **102**: 3-8.

Kreider JW, Howett MK, Wolfe SA, Bartlett GL, Zaino RJ, Sedlacek T, Mortel R (1985) Morphological transformation in vivo of human uterine cervix with papillomavirus from condylomata acuminata. *Nature* **317**: 639-641.

Walboomers JM, Jacobs MV, Manos MM, Bosch FX, Kummer JA, Shah KV, Snijders PJ, Peto J, Meijer CJ, Munoz N (1999) Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J Pathol* **189**: 12-19.

Kryszke MH, Piette J, Yaniv M (1987) Induction of a factor that binds to the polyoma virus A enhancer on differentiation of embryonal carcinoma cells. *Nature* **328**: 254-256.

Pirisi L, Yasumoto S, Feller M, Doniger J, DiPaolo JA (1987) Transformation of human fibroblasts and keratinocytes with human papillomavirus type 16 DNA. *J Virol* **61**: 1061-1066.

Yasumoto S, Burkhardt AL, Doniger J, DiPaolo JA (1986) Human papillomavirus type 16 DNA-induced malignant transformation of NIH 3T3 cells. *J Virol* **57**: 572-577.

Heilman CA, Engel L, Lowy DR, Howley PM (1982) Virus-specific transcription in bovine papillomavirus-transformed mouse cells. *Virology* **119**: 22-34.

Frattini MG, Lim HB, Laimins LA (1996) In vitro synthesis of oncogenic human papillomaviruses requires episomal genomes for differentiation-dependent late expression. *Proc Natl Acad Sci U S A* **93**: 3062-3067.

Frattini MG, Lim HB, Doorbar J, Laimins LA (1997) Induction of human papillomavirus type 18 late gene expression and genomic amplification in organotypic cultures from transfected DNA templates. *J Virol* **71**: 7068-7072.

Park TW, Fujiwara H, Wright TC (1995) Molecular biology of cervical cancer and its precursors. *Cancer* **76**: 1902-1913.

Reid R, Crum CP, Herschman BR, Fu YS, Braun L, Shah KV, Agronow SJ, Stanhope CR (1984) Genital warts and cervical cancer. III. Subclinical papillomaviral infection and cervical neoplasia are linked by a spectrum of continuous morphologic and biologic change. *Cancer* **53**: 943-953.

Halpert R, Fruchter RG, Sedlis A, Butt K, Boyce JG, Sillman FH (1986) Human papillomavirus and lower genital neoplasia in renal transplant patients. *Obstet Gynecol* **68**: 251-258.

Crosbie EJ, Kitchener HC (2006) Human papillomavirus in cervical screening and vaccination. *Clin Sci (Lond)* **110**: 543-552.

Munger K (2002) The role of human papillomaviruses in human cancers. *Front Biosci* **7**: d641-649.

Combita AL, Touze A, Bousarghin L, Christensen ND, Coursaget P (2002) Identification of two cross-neutralizing linear epitopes within the L1 major capsid protein of human papillomaviruses. *J Virol* **76**: 6480-6486.

de Villiers EM, Fauquet C, Broker TR, Bernard HU, zur Hausen H (2004) Classification of papillomaviruses. *Virology* **324**: 17-27.

Calleja-Macias IE, Kalantari M, Allan B, Williamson AL, Chung LP, Collins RJ, Zuna RE, Dunn ST, Ortiz-Lopez R, Barrera-Saldana HA, Cubie HA, Cuschieri K, Villa LL, Bernard HU (2005) Papillomavirus subtypes are natural and old taxa: phylogeny of human papillomavirus types 44 and 55 and 68a and -b. *J Virol* **79**: 6565-6569.

Berumen J, Casas L, Segura E, Amezcua JL, Garcia-Carranca A (1994) Genome amplification of human papillomavirus types 16 and 18 in cervical carcinomas is related to the retention of E1/E2 genes. *International journal of cancer* **56**: 640-645.

Cullen AP, Reid R, Campion M, Lorincz AT (1991) Analysis of the physical state of different human papillomavirus DNAs in intraepithelial and invasive cervical neoplasm. *J Virol* **65**: 606-612.

Guido MC, Zamorano R, Garrido-Guerrero E, Gariglio P, Garcia-Carranca A (1992) Early promoters of genital and cutaneous human papillomaviruses are differentially regulated by the bovine papillomavirus type 1 E2 gene product. *The Journal of general virology* **73 ( Pt 6)**: 1395-1400.

Ho L, Chan SY, Burk RD, Das BC, Fujinaga K, Icenogle JP, Kahn T, Kiviat N, Lancaster W, Mavromara-Nazos P, et al. (1993) The genetic drift of human papillomavirus type 16 is a means of reconstructing prehistoric viral spread and the movement of ancient human populations. *J Virol* **67**: 6413-6423.

Yamada T, Wheeler CM, Halpern AL, Stewart AC, Hildesheim A, Jenison SA (1995) Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, L2, and L1 coding segments. *J Virol* **69**: 7743-7753.

Xi LF, Koutsky LA, Galloway DA, Kuypers J, Hughes JP, Wheeler CM, Holmes KK, Kiviat NB (1997) Genomic variation of human papillomavirus type 16 and risk for high grade cervical intraepithelial neoplasia. *J Natl Cancer Inst* **89**: 796-802.

Lizano M, Berumen J, Guido MC, Casas L, Garcia-Carranca A (1997) Association between human papillomavirus type 18 variants and histopathology of cervical cancer. *J Natl Cancer Inst* **89**: 1227-1231.

Bosch FX, Manos MM, Munoz N, Sherman M, Jansen AM, Peto J, Schiffman MH, Moreno V, Kurman R, Shah KV (1995) Prevalence of human papillomavirus in cervical cancer: a worldwide perspective. International biological study on cervical cancer (IBSCC) Study Group. *J Natl Cancer Inst* **87**: 796-802.

Zehbe I, Wilander E, Delius H, Tommasino M (1998) Human papillomavirus 16 E6 variants are more prevalent in invasive cervical carcinoma than the prototype. *Cancer Res* **58**: 829-833.

Berumen J, Ordonez RM, Lazcano E, Salmeron J, Galvan SC, Estrada RA, Yunes E, Garcia-Carranca A, Gonzalez-Lira G, Madrigal-de la Campa A (2001) Asian-American variants of human papillomavirus 16 and risk for cervical cancer: a case-control study. *J Natl Cancer Inst* **93**: 1325-1330.

Casas L, Galvan SC, Ordonez RM, Lopez N, Guido M, Berumen J (1999) Asian-american variants of human papillomavirus type 16 have extensive mutations in the E2 gene and are highly amplified in cervical carcinomas. *International journal of cancer* **83**: 449-455.

Stoppler MC, Ching K, Stoppler H, Clancy K, Schlegel R, Icenogle J (1996) Natural variants of the human papillomavirus type 16 E6 protein differ in their abilities to alter keratinocyte differentiation and to induce p53 degradation. *J Virol* **70**: 6987-6993.

Sousa R, Chung YJ, McAllister WT, Wang BC, Lafer EM (1990) Single crystals of a chimeric T7/T3 RNA polymerase with T3 promoter specificity and a nonprocessive T7 RNAP mutant. *J Biol Chem* **265**: 21430-21432.

Conger KL, Liu JS, Kuo SR, Chow LT, Wang TS (1999) Human papillomavirus DNA replication. Interactions between the viral E1 protein and two subunits of human dna polymerase alpha/primase. *J Biol Chem* **274**: 2696-2705.

Martin P, Vass WC, Schiller JT, Lowy DR, Velu TJ (1989) The bovine papillomavirus E5 transforming protein can stimulate the transforming activity of EGF and CSF-1 receptors. *Cell* **59**: 21-32.

Goldstein DJ, Finbow ME, Andresson T, McLean P, Smith K, Bubb V, Schlegel R (1991) Bovine papillomavirus E5 oncoprotein binds to the 16K component of vacuolar H(+)-ATPases. *Nature* **352**: 347-349.

Scheffner M, Werness BA, Huibregtse JM, Levine AJ, Howley PM (1990) The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. *Cell* **63**: 1129-1136.

Smotkin D, Wettstein FO (1987) The major human papillomavirus protein in cervical cancers is a cytoplasmic phosphoprotein. *J Virol* **61**: 1686-1689.

Dyson N, Howley PM, Munger K, Harlow E (1989) The human papilloma virus-16 E7 oncoprotein is able to bind to the retinoblastoma gene product. *Science* **243**: 934-937.

Munger K, Phelps WC, Bubb V, Howley PM, Schlegel R (1989) The E6 and E7 genes of the human papillomavirus type 16 together are necessary and sufficient for transformation of primary human keratinocytes. *J Virol* **63**: 4417-4421.

Matlashewski G, Schneider J, Banks L, Jones N, Murray A, Crawford L (1987) Human papillomavirus type 16 DNA cooperates with activated ras in transforming primary cells. *Embo J* **6**: 1741-1746.

zur Hausen H (1991) Viruses in human cancers. *Science* **254**: 1167-1173.

Garcia-Carranca A, Thierry F, Yaniv M (1988) Interplay of viral and cellular proteins along the long control region of human papillomavirus type 18. *J Virol* **62**: 4321-4330.

Geisen C, Kahn T (1996) Promoter activity of sequences located upstream of the human papillomavirus types of 16 and 18 late regions. *The Journal of general virology* **77 ( Pt 9)**: 2193-2200.

Gloss B, Bernard HU, Seedorf K, Klock G (1987) The upstream regulatory region of the human papilloma virus-16 contains an E2 protein-independent enhancer which is specific for cervical carcinoma cells and regulated by glucocorticoid hormones. *Embo J* **6**: 3735-3743.

Gius D, Grossman S, Bedell MA, Laimins LA (1988) Inducible and constitutive enhancer domains in the noncoding region of human papillomavirus type 18. *J Virol* **62**: 665-672.

Rocha-Zavaleta L, Jordan D, Pepper S, Corbitt G, Clarke F, Maitland NJ, Sanders CM, Arrand JR, Stern PL, Stacey SN (1997) Differences in serological IgA responses to recombinant baculovirus-derived human papillomavirus E2 protein in the natural history of cervical neoplasia. *Br J Cancer* **75**: 1144-1150.

Antson AA, Burns JE, Moroz OV, Scott DJ, Sanders CM, Bronstein IB, Dodson GG, Wilson KS, Maitland NJ (2000) Structure of the intact transactivation domain of the human papillomavirus E2 protein. *Nature* **403**: 805-809.

Ozbun MA, Meyers C (1998) Human papillomavirus type 31b E1 and E2 transcript expression correlates with vegetative viral genome amplification. *Virology* **248**: 218-230.

Frattini MG, Laimins LA (1994) Binding of the human papillomavirus E1 origin-recognition protein is regulated through complex formation with the E2 enhancer-binding protein. *Proc Natl Acad Sci U S A* **91**: 12398-12402.

Li R, Botchan MR (1994) Acidic transcription factors alleviate nucleosome-mediated repression of DNA replication of bovine papillomavirus type 1. *Proc Natl Acad Sci U S A* **91**: 7051-7055.

Berg M, Stenlund A (1997) Functional interactions between papillomavirus E1 and E2 proteins. *J Virol* **71**: 3853-3863.

Hibma MH, Raj K, Ely SJ, Stanley M, Crawford L (1995) The interaction between human papillomavirus type 16 E1 and E2 proteins is blocked by an antibody to the N-terminal region of E2. *Eur J Biochem* **229**: 517-525.

Cooper CS, Upmeyer SN, Winokur PL (1998) Identification of single amino acids in the human papillomavirus 11 E2 protein critical for the transactivation or replication functions. *Virology* **241**: 312-322.

Hegde RS, Androphy EJ (1998) Crystal structure of the E2 DNA-binding domain from human papillomavirus type 16: implications for its DNA binding-site selection mechanism. *J Mol Biol* **284**: 1479-1489.

McBride AA, Bolen JB, Howley PM (1989) Phosphorylation sites of the E2 transcriptional regulatory proteins of bovine papillomavirus type 1. *J Virol* **63**: 5076-5085.

Steger G, Corbach S (1997) Dose-dependent regulation of the early promoter of human papillomavirus type 18 by the viral E2 protein. *J Virol* **71**: 50-58.

Hines CS, Meghoo C, Shetty S, Biburger M, Brenowitz M, Hegde RS (1998) DNA structure and flexibility in the sequence-specific binding of papillomavirus E2 proteins. *J Mol Biol* **276**: 809-818.

Hegde RS, Grossman SR, Laimins LA, Sigler PB (1992) Crystal structure at 1.7 Å of the bovine papillomavirus-1 E2 DNA-binding domain bound to its DNA target. *Nature* **359**: 505-512.

Bochkarev A, Barwell JA, Pfuetzner RA, Furey W, Jr., Edwards AM, Frappier L (1995) Crystal structure of the DNA-binding domain of the Epstein-Barr virus origin-binding protein EBNA 1. *Cell* **83**: 39-46.

Dowhanick JJ, McBride AA, Howley PM (1995) Suppression of cellular proliferation by the papillomavirus E2 protein. *J Virol* **69**: 7791-7799.

Prakash SS, Grossman SR, Pepinsky RB, Laimins LA, Androphy EJ (1992) Amino acids necessary for DNA contact and dimerization imply novel motifs in the papillomavirus E2 trans-activator. *Genes Dev* **6**: 105-116.

Breiding DE, Sverdrup F, Grossel MJ, Moscufo N, Boonchai W, Androphy EJ (1997) Functional interaction of a novel cellular protein with the papillomavirus E2 transactivation domain. *Mol Cell Biol* **17**: 7208-7219.

Knight JD, Li R, Botchan M (1991) The activation domain of the bovine papillomavirus E2 protein mediates association of DNA-bound dimers to form DNA loops. *Proc Natl Acad Sci U S A* **88**: 3204-3208.

Harris SF, Botchan MR (1999) Crystal structure of the human papillomavirus type 18 E2 activation domain. *Science* **284**: 1673-1677.

Lambert PF, Hubbert NL, Howley PM, Schiller JT (1989) Genetic assignment of multiple E2 gene products in bovine papillomavirus-transformed cells. *J Virol* **63**: 3151-3154.

Muller F, Giroglou T, Sapp M (1997) Characterization of the DNA-binding activity of the E1 and E2 proteins and the E1/E2 complex of human papillomavirus type 33. *The Journal of general virology* **78 ( Pt 4)**: 911-915.

Casanova JL, Pannetier C, Jaulin C, Kourilsky P (1990) Optimal conditions for directly sequencing double-stranded PCR products with sequenase. *Nucleic Acids Res* **18**: 4028.

Barnard GF, Puder M, Begum NA, Chen LB (1994) PCR product sequencing with [ $\alpha$ -<sup>32</sup>P] and [ $\alpha$ -<sup>32</sup>P]dATP. *Biotechniques* **16**: 572-573.

Wigler M, Silverstein S, Lee LS, Pellicer A, Cheng Y, Axel R (1977) Transfer of purified herpes virus thymidine kinase gene to cultured mouse cells. *Cell* **11**: 223-232.

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**: 248-254.

Piccini A, Storey A, Massimi P, Banks L (1995) Mutations in the human papillomavirus type 16 E2 protein identify multiple regions of the protein involved in binding to E1. *The Journal of general virology* **76 ( Pt 11)**: 2909-2913.

Dillner J, Kallings I, Brihmer C, Sikstrom B, Koskela P, Lehtinen M, Schiller JT, Sapp M, Mardh PA (1996) Seropositivities to human papillomavirus types 16, 18, or 33 capsids and to Chlamydia trachomatis are markers of sexual behavior. *The Journal of infectious diseases* **173**: 1394-1398.

Roberts S, Ashmole I, Gibson LJ, Rookes SM, Barton GJ, Gallimore PH (1994) Mutational analysis of human papillomavirus E4 proteins: identification of structural features important in the formation of cytoplasmic E4/cytokeratin networks in epithelial cells. *J Virol* **68**: 6432-6445.

Kammer C, Warthorst U, Torrez-Martinez N, Wheeler CM, Pfister H (2000) Sequence analysis of the long control region of human papillomavirus type 16 variants and functional consequences for P97 promoter activity. *The Journal of general virology* **81**: 1975-1981.

Veress G, Szarka K, Dong XP, Gergely L, Pfister H (1999) Functional significance of sequence variation in the E2 gene and the long control region of human papillomavirus type 16. *The Journal of general virology* **80 ( Pt 4)**: 1035-1043.

Yamada T, Manos MM, Peto J, Greer CE, Munoz N, Bosch FX, Wheeler CM (1997) Human papillomavirus type 16 sequence variation in cervical cancers: a worldwide perspective. *J Virol* **71**: 2463-2472.

Konya J, Eklund C, af Geijersstam V, Yuan F, Stuber G, Dillner J (1997) Identification of a cytotoxic T-lymphocyte epitope in the human papillomavirus type 16 E2 protein. *The Journal of general virology* **78 ( Pt 10)**: 2615-2620.

Terry G, Ho L, Cuzick J (1997) Analysis of E2 amino acid variants of human papillomavirus types 16 and 18 and their associations with lesion grade and HLA DR/DQ type. *International journal of cancer* **73**: 651-655.

De Sanjose S, Hamsikova E, Munoz N, Bosch FX, Hofmannova V, Gili M, Izarzugaza I, Viladiu P, Tormo MJ, Moreo P, Munoz MT, Ascunce N, Tafur L, Shah KV, Vonka V (1996) Serological response to HPV16 in CIN-III and cervical cancer patients. Case-control studies in Spain and Colombia. *International journal of cancer* **66**: 70-74.

Ordonez RM, Espinosa AM, Sanchez-Gonzalez DJ, Armendariz-Borunda J, Berumen J (2004) Enhanced oncogenicity of Asian-American human papillomavirus 16 is associated with impaired E2 repression of E6/E7 oncogene transcription. *The Journal of general virology* **85**: 1433-1444.

**11. ANEXO. Artículo publicado.**

## ASIAN-AMERICAN VARIANTS OF HUMAN PAPILLOMAVIRUS TYPE 16 HAVE EXTENSIVE MUTATIONS IN THE *E2* GENE AND ARE HIGHLY AMPLIFIED IN CERVICAL CARCINOMAS

Leonora CASAS<sup>1</sup>, Silvia C. GALVAN<sup>1,2</sup>, Rosa M. ORDOÑEZ<sup>1</sup>, Nora LOPEZ<sup>1,3</sup>, Miriam GUIDO<sup>2</sup> and Jaime BERUMEN<sup>1,4\*</sup>

<sup>1</sup>Multidisciplinary Research Laboratory, Military School of Medical Graduates and Army School of Medicine, University of the Army and Air Force, Mexico City, Mexico

<sup>2</sup>Biomedical Research Institute, UNAM, University City, Mexico City, Mexico

<sup>3</sup>Department of Cellular Biology, CINVESTAV, IPN, Mexico City, Mexico

<sup>4</sup>Applied Science and Advanced Technology Research Center, IPN, Mexico City, Mexico

**Human-papillomavirus (HPV)-*E2* protein is involved in gene-expression regulation and replication of HPV genome. Disruption of the *E2* gene during viral integration has been proposed as a mechanism of tumoral progression, since the expression of *E6/E7* viral oncogenes is allowed. However, retention of *E1/E2* genes and high viral amplification are frequently found in HPV16-positive carcinomas of some populations. In this study, we investigated whether retention of *E1/E2* and viral amplification are associated with particular HPV16 *E2* variants in cervical carcinomas. HPV16 detection, *E1/E2* integrity and viral amplification were explored by Southern blot in 123 cervical carcinomas. HPV16 variants were identified by Southern blot and by sequencing *E6*, *L1/MY* and *E2* regions. Of 46 HPV16-positive tumors, 34 were positive for *E1/E2* and 14 of them showed a variant restriction pattern by mutations in *E2*. All 14 were Asian-American (AA) variants and, of 11 sub-classified, 6 were AA-a and 5 AA-c. Two *E1/E2*-negative tumors also contained the AA-c variant, while the remaining HPV16-positive tumors contained only European variants. The *E2* gene of AA variants showed 24 mutations, 19 identical in both sub-classes. The 24 mutations were distributed throughout the entire gene and 19 result in 18 amino-acid changes. The AA variants were associated with *E1/E2*-positive carcinomas with more than 50 viral copies/cell ( $p = 0.035$ ). The association of Asian-American *E2* variants with retention of *E1/E2* suggests that *E2* variation may be an alternative mechanism de-regulating the expression of viral oncogenes. Int. J. Cancer 83:449–455, 1999.**

© 1999 Wiley-Liss, Inc.

The *E2* protein of human papillomavirus (HPV) is involved in the regulation of early gene expression and, with the *E1* protein, in the replication of viral genome. Disruption of the *E2* gene during integration of the viral DNA into the cellular genome has been proposed as an activation mechanism for progression from advanced pre-invasive lesions to invasive cervical cancer (Schneider-Maunoury *et al.*, 1987). It has been repeatedly shown that viral DNA integration destroys the *E2* gene (Wilczynsky *et al.*, 1988), thereby relieving *E2*-dependent negative control of the *E6/E7* promoter, increasing the expression of the *E6* and *E7* viral oncogenes. *In vitro* studies have shown that cells with integrated HPV16 DNA have a selective growth advantage over those with episomal DNA; and, in clinical studies, disruption of *E2* has been associated with poor disease-free survival from cervical cancer (Vernon *et al.*, 1997).

Integration of the viral genome and disruption of *E1/E2* genes have been found almost in all cell lines derived from cervical carcinomas (Schneider-Maunoury *et al.*, 1987). However, HPV16 DNA is not always integrated in carcinomas, and may also be found as an episome or as a combination of both integrated and episomal viral forms (Vernon *et al.*, 1997; Cullen *et al.*, 1991; Cooper *et al.*, 1991; Matsukura *et al.*, 1989). In addition, HPV16 integration may not always result in complete loss of the *E1/E2* genes. While single-copy HPV-integration events nearly always interrupt *E1/E2*, multi-copy head-to-tail tandem integration preserves intact internal copies, and only the viral copy flanking the cellular DNA is interrupted in the *E1* or *E2* region (Wagatsuma *et al.*, 1990). These

data indicate that the disruption of *E2* is not essential for the development of HPV16-associated carcinomas, and raise questions about alternative mechanisms for increasing transcription of viral oncogenes.

There is wide variation in the reported frequency of *E2* detection in HPV16-positive cervical carcinomas; and in some populations, *e.g.*, Mexican, American black and Japanese, the presence of *E1/E2* genes is observed at fairly high frequency (Table I). Differences in detection methods, populations or distribution of HPV16 variants may contribute to this variation. Nucleic-acid sequencing data show that many natural HPV16 variants exist (Ho *et al.*, 1993; Yamada *et al.*, 1997), distributed differently among the 5 continents. The Asian-American (AA) variants are found principally in Central and South America and Spain, African variants (Af) in Africa, Asian variants (As) in Southeast Asia and European variants (E) in all regions other than Africa (Yamada *et al.*, 1997). It is clear that some variations in the HPV genome may determine biological behaviors, *e.g.*, the sequence variation of L1 protein of HPV16 affects virus assembly (Yamada *et al.*, 1995), while some *E6* variants of HPV16 show differences in immortalizing activity and p53 degradation (Yamada *et al.*, 1997). *E2* variation has been poorly studied. Terry *et al.* (1997) reported HPV16 European *E2* variants associated with changes in immunodominant epitopes of the hinge region of *E2* protein.

In this study, we found that AA variants are highly frequent in Mexican patients with cervical cancer, have extensive nucleotide changes in the *E2* gene and are associated with a high viral-copy number and retention of *E1/E2* genes.

### MATERIAL AND METHODS

#### Samples and experimental design

A total of 123 patients with invasive cervical cancer diagnosed between January and June 1992 at the Hospital de Oncología, Centro Médico Nacional, IMSS of Mexico City, were recruited for the study. Tissue samples were obtained by direct open biopsy; one fragment was used for routine pathological diagnosis and the other was snap-frozen in liquid nitrogen and stored at  $-70^{\circ}\text{C}$  for molecular analysis. Squamous-cell carcinomas comprised 87%, adeno- or adenosquamous carcinomas 6.5% and undifferentiated carcinomas 6.5%. The detection of HPV16, the presence of *E1/E2* and the level of viral amplification were analyzed by Southern blot. *E2*, *E6* and *L1/MY* regions of the HPV16 DNA were amplified by polymerase chain reaction (PCR) and directly sequenced by the

Grant sponsor: Secretariat of National Defense and CONACYT, Mexico; Grant number: L0036-M9607.

\*Correspondence to: Laboratorio Multidisciplinario de Investigación, Escuela Militar de Graduados de Sanidad, Apartado Postal 35-556, 11649 México D.F., México. Fax: 52-5202-0303. E-mail: lamili@mail.internet.com.mx

Received 4 January 1999; revised 29 March 1999

**TABLE I** – FREQUENCY OF EPISOMAL VIRAL FORMS OR E2-GENE DETECTION IN HPV16-POSITIVE INVASIVE CERVICAL CARCINOMAS IN DIFFERENT COUNTRIES<sup>1</sup>

Country	n	Episomal forms (%) <sup>2</sup>	Method	Reference
USA	14	21	SB	Wilczynski <i>et al.</i> , 1988
South Africa/ England	41	15	ISH	Cooper <i>et al.</i> , 1991
India	42	31	SB	Das <i>et al.</i> , 1992
USA	45 <sup>3</sup>	38	ISH	Vernon <i>et al.</i> , 1997
USA/Canada	40	47	SB	Cullen <i>et al.</i> , 1991
England	31	77	SB	Meanwell <i>et al.</i> , 1987
Japan	34	76	SB	Matsukura <i>et al.</i> , 1989
<i>E2</i> -positive (%)				
USA	14	36	SB	Wilczynski <i>et al.</i> , 1988
USA	45 <sup>3</sup>	64	PCR	Vernon <i>et al.</i> , 1997
Mexico	25 <sup>4</sup>	64	PCR	Berumen <i>et al.</i> , 1994

<sup>1</sup>HPV16, human papillomavirus type 16; SB, Southern blot; ISH, *in situ* hybridization; PCR, polymerase chain reaction. Only reports on more than 10 HPV16-positive cervical carcinomas and from different laboratories were included. <sup>2</sup>Episomal forms include carcinomas with episome only and those with both episome and integrated HPV16 genomes. <sup>3</sup>Essentially black population. <sup>4</sup>Mestizo population.

fluorescent cycle-sequencing method and analyzed in an Automatic Genetic Analyzer.

#### Detection of HPV16 by Southern blot

DNA was purified by phenol extraction. For Southern blots, 10 µg of sample DNA was digested with restriction enzymes BamHI and PstI (GIBCO BRL, Life Technologies, Gaithersburg, MD), in conditions recommended by the manufacturer. In each experiment, the first lane contained the molecular-weight standard (Phage lambda/HindIII) and the second lane contained 50 copies/cell of HPV16 linear DNA (725 pg) and 1450 pg of HPV16 DNA plus 10 µg of human placental genome (HPG; Sigma, St. Louis, MO), as carrier, digested with BamHI/PstI. HPV16 DNA was used as probe in stringent conditions, and hybridization was conducted as described (Berumen *et al.*, 1994). The intensity of *E6/E7* bands on autoradiograms was analyzed by the Kodak Digital Science Electrophoresis Documentation and Analysis System (Eastman Kodak, Rochester, NY), and the copy number per cell was calculated as described (Berumen *et al.*, 1994).

#### Sequencing of E2, E6 and L1/MY of the HPV16 genome from different tumor DNAs

The *E2*, *E6* and *L1/MY* regions were amplified using primer pairs specific for each region. The sequence and location of the primers are given in Table II. The PCR mix was as follows: 500 ng of tumor DNA was added to achieve a final volume of 50 µl containing 30 mM Tris-HCl (pH 8.8), 2 mM MgCl<sub>2</sub>, 50 mM KCl, 1 mM DTT, 200 µM of each dATP, dTTP, dGTP and dCTP, 1 µM of each primer, 100 ng/µl BSA and 1.0 U of Taq DNA polymerase (GIBCO BRL-Life Technologies). After de-naturing for 3 min at 94°C, the reaction was carried out over 40 cycles in a DNA GeneAmp PCR System 9600 (Perkin-Elmer, Branchburg, NJ). For *L1/MY* the station 1 was set at 94°C for 30 sec, station 2 at 55°C for 1.5 min and station 3 at 72°C for 1.5 min; thereafter, the reaction was kept at 72°C for another 7 min. The same conditions were used to amplify the *E6* and *E2* regions, but the annealing temperature was set at 54°C. The amplified fragments were purified using the Quiaex II kit (Quiagen, Valencia, CA), as specified by the manufacturer. The products were then sequenced in both directions using the fluorescent cycle-sequencing method (BigDye Terminator Ready Reaction Kit, Perkin-Elmer). The PCR mix and cycling conditions were used as recommended by the manufacturer and the annealing temperature was set at 50°C. In addition to the primers used in the amplification reactions, internal primers were used for the sequencing reactions (Table II). Sequence analysis was performed in the ABI PRISM 310 Genetic Analyzer System (Perkin-

**TABLE II** – LOCATION AND SEQUENCE OF PRIMERS USED IN PCR AMPLIFICATION AND SEQUENCING REACTIONS

Region	Name	Location	Sequence (5'-3')	
<i>E2</i>	F2670	2670–2689	TGAGCTTAATGATAAGAAGT	
	F2975	2975–2994	GAACTGCAACTAACGTTAGA	
	F3192	3192–3211	CTGTGGTAGAGGGTCAAGTT	
	F3500	3500–3519	AACCCCTGCCACACCACTAA	
	R4086	4086–4067	CGGTGCATGTGTATGTATTA	
	R3624	3624–3605	TGTACTATGGGTGTAGTGTT	
	R3384	3384–3367	ATTTCAGGAGAGGATACT	
<i>L1/MY</i> <sup>1</sup>	R3124	3124–3105	TCCATCAAACCTGCACCTTCCA	
	MY11	6584–6603	GCACAGGGCCACAATAATGG	
	F6775	6771–6790	TGTGCAAAATAACCTTAACT	
	MY09	7035–7014	CGTCCTAAGGAAAACCTGATCTA	
	R6859	6859–6840	TTGTAGACCAAATTCAGT	
	<i>E6</i>	F048	48–67	GAACCGAAACCGGTTAGTAT
		F221	221–240	CGTGAGGTATATGACTTTGC
R374		374–355	GCTGTTCTAATGTTGTTCCA	
R622		622–603	CAGTTGTCTCTGGTTGCAAA	

<sup>1</sup>Primers MY09 and MY11 are specific only for HPV16.

Elmer). Data analysis was accomplished with the DNASIS software (Hitachi Software Engineering, Yokohama, Japan) and BLAST via Internet. HPV16 sequences and base positions are numbered according to the 1997 sequence database (Los Alamos National Laboratory, NM).

#### Nucleotide-sequence accession numbers

Nucleotide sequences have been submitted to GenBank and will be released upon publication. Accession numbers have been assigned as follows: AF115831, AF115832.

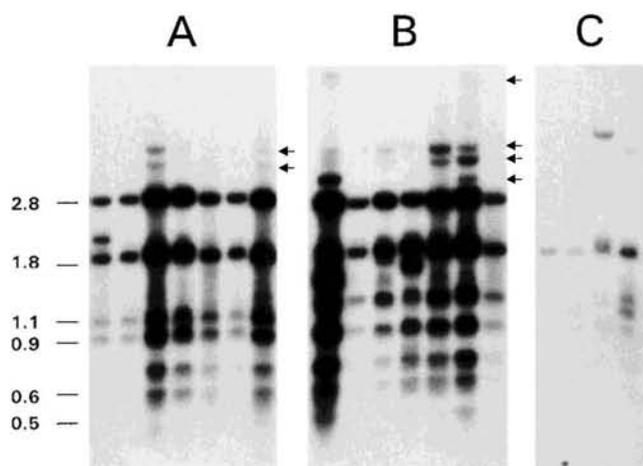
#### Statistical analysis

For copy number, the results are expressed as the mean ± standard error of the mean (SEM). The 2-tailed Mann-Whitney U-test was used to assess the significance of differences on the copy number between the groups. The Chi-squared test was used to assess the significance of differences in the frequency of HPV16 variants among the groups classified by the copy number or *E1/E2* status. The association between rows (r) and columns (c) of copy number and *E1/E2* groups in 3 × 3 contingency tables, for all HPV16- or HPV16 AA-positive samples, was tested by Chi squared. A difference was considered significant at a level of  $p < 0.05$ .

## RESULTS

#### Finding of a new highly frequent E2 variant by analysis of Southern-blot patterns

From the 123 samples explored by Southern blot, 46 (37.4%) were positive for HPV16; 34 (73.9%) of them had the *E1/E2* genes intact (2.8-kb band in Fig. 1a,b) and 12 (26.1%) lost this region (Fig. 1c). All tumors with an intact *E1/E2* region showed the authentic 7-cut PstI/BamHI cleavage pattern of HPV16; however, 18 of them also showed off-sized bands above the *E1/E2* band (lanes 3, 4 and 7 in Fig. 1a and lanes 1, 3, 5 and 6 in Fig. 1b). The authentic PstI/BamHI banding pattern was not identical in all *E1/E2*-positive tumors. The first (2.8-kb) and third (1.1-kb) bands, corresponding to regions A and C respectively on the HPV16 map (Fig. 2b), migrated slightly different from reference HPV16 (Seedorf *et al.*, 1985) in 14 tumors (see Fig. 2a, lanes b, c and h). Band A (*E1/E2*) migrated faster and its molecular weight decreased by approximately 116 bp (2702 vs. 2818 bp of reference type) and band C (part of *E2*, *E5* and part of *L2*) migrated slower and its molecular weight increased by approximately 138 bp (1202 vs. 1064 bp of prototype). These DNA fragments are contiguous in the HPV16 genome (see map in Fig. 2b) and the size variations could result from loss of the PstI site at position 3693 and the acquisition of a new one 116 to 138 bp upstream from this site (Fig. 2b). In fact, a putative PstI-like site is found in the position 3561 (CTGCAT),



**FIGURE 1** – Signal intensity and restriction pattern of European and Asian-American HPV16 variants in the Southern blot. Panels show signal intensity and restriction pattern of DNA samples derived from cervical carcinomas digested with PstI and BamHI restriction enzymes and hybridized with the HPV16 DNA probe in the Southern-blot experiments. The tumoral DNAs shown in (a) and (b) contain the *E1/E2* band (2.8 kb) and are positive for European and Asian-American variants respectively. Samples in (c) are negative for *E1/E2*. Lanes 3, 4 and 7 in (a) and lanes 1, 3, 5 and 6 in (b) show off-sized extra bands (arrows). The restriction pattern of the reference HPV16 genome digested with PstI/BamHI restriction enzymes and size of bands in kb are indicated at left.

132 bases from the 3693 position (see Fig. 2b) and it could become a true PstI recognition site by mutation of only one base, T per G in the position 3566. These predicted base changes were confirmed by PstI digestion of PCR amplified fragments containing the *E1/E2* region of these cases (data not shown) and by sequencing this region (see below and Fig. 3).

#### Variant E2 gene has extensive nucleotide changes

To determine the nature and precise location of *E2* sequence variation, the *E2* gene from 8 cases with variants identified by Southern blot was amplified, sequenced and compared with HPV16 reference: a total of 24 nucleotide changes were identified; while 19 changes were consistent in all 8 cases, the remaining 5 changes fell into 2 groups. Five cases had the pattern exemplified by MX68 isolate (3 changes) and 3 cases had the pattern exemplified by MX70 isolate (2 changes; see Fig. 3). Compared with the *E2* reference sequence (1098 bp), MX68 showed changes in 2.0% and MX70 in 1.9% of bases. Therefore, the isolates MX68 and MX70 should be considered as 2 different but related variants.

Of the 19 consistent nucleotide substitutions, 15 result in 14 amino-acid changes (2 contiguous mutations change the same codon, 254). Five amino-acid (a.a.) substitutions were identified in the transactivation domain, 6 in the hinge region and 3 in the DNA binding domain (Fig. 4). In the transactivation domain, the substitution H35Q fell into an identified site of interaction with the E1 protein (residues 18–41; Piccini *et al.*, 1995) and 4 changes fell into the 3' region of the transactivation domain (Fig. 4). In the DNA binding domain, the substitution T310K fell just beside a DNA-binding helix (292–309) and together with substitutions W341C and D344E in a putative dimerization domain. Of the 3 additional base changes in the MX68 isolate, 2 (A3181C and T3387C) result in amino-acid substitutions, E142D located in the transactivation domain and I211T in the hinge region. Both additional changes (C3161T and G3416A) in the MX70 isolate result in amino-acid substitutions, H136Y in the transactivation domain and A221T in the hinge region. Changes T135K, T254N and F271V were located in 3 of the most reactive B-cell epitopes (residues 121–140;

241–260 and 271–290) and changes W341C and D344E into a major IgG and IgA epitope (residues 328–345; Dillner *et al.*, 1990).

Of the 24 mutations, 9 also affect the *E4* gene, which is located in part of the *E2* DNA region but in a different reading frame. Among the 9 mutations affecting *E4*, 4 result in 3 amino-acid substitutions, L62T, Q69P and H78Q, located in the *E4* region with similarity to actin protein (residues 48 to 114 in HPV1; Roberts *et al.*, 1994).

#### *E2* variants correspond to Asian-American E6 and L1/MY variants

In order to investigate whether these *E2* variants correspond to HPV16 variants previously described on the basis of *L1* and *E6* sequences, we amplified and sequenced the MY fragment of the *L1* gene (6584–7035) and the full *E6* gene (from nucleotides 83 to 559) from all 8 sequenced *E2* variants and from 3 additional Southern-blot-identified *E2* variant tumor DNAs. All 11 samples were Asian-American (AA) variants, 6 of the sub-class AA-a (equal to isolates OR.8160 and OR.4541) and 5 of the sub-class AA-c (equal to isolate IS.53), according to the classification of Yamada *et al.* (1995, 1997). Co-variation between *E6* and *L1* was fully observed, all 6 of the AA-a class-sub-class *E6* variants showed *L1/MY* sequences consistent with AA-a class-sub-class variant and all 5 of the AA-c class-sub-class *E6* variant contained *L1* nucleotide sequences consistent with the AA-c class-sub-class variant. However, although all *E2* variants were of the AA class, the sub-class co-variation between *L1/E6* and *E2* was not fully maintained. All 3 isolates with the MX70 *E2* sequence and one (isolate MX69) of the 5 MX68 *E2* sequence were of the AA-c type in *L1* and *E6*, while the other 4 isolates with MX68 *E2* sequence were of type AA-a sub-class in the *E6* and *L1* genes (Fig. 3).

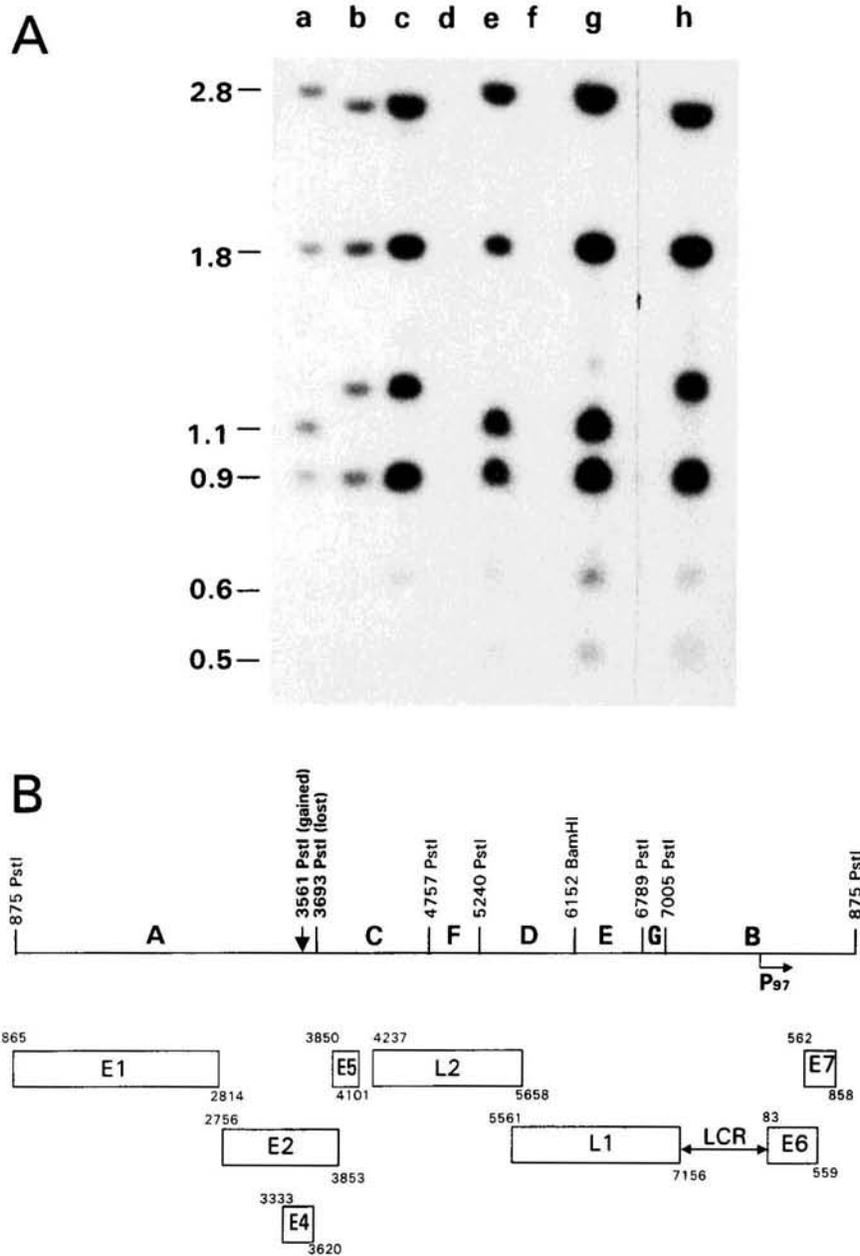
To determine the HPV16-variant types in cancers not showing the variant pattern on Southern blot, the *L1/MY* region and the *E6* gene were amplified and sequenced in all 12 *E1/E2*-negative cases as well as in the remaining *E1/E2*-positive samples ( $n = 20$ ). The variant AA-c was found only in 2 out of 12 (16.7%) tumors negative for *E1/E2* and in none of the *E1/E2*-positive samples without the shift band in Southern blot. All AA-negative samples were European (E) variants. Since intergene co-variation between *L1/E6* and *E2* genes was found in most isolates, it can be reasonably assumed that AA-c variants found in the *E1/E2*-negative samples would have the corresponding base changes in *E2*. The *E2* gene was amplifiable in one of these 2 isolates (MX97), and did have the AA-c sub-class *E2* sequence (Fig. 3), supporting the intergene co-variation. In this study, the frequency of AA variants was 34.8% of all HPV16-positive samples (16 out of 46), with almost equal distribution between sub-classes AA-c (53.8%) and AA-a (46.2%).

The frequency of base changes in the *E6*, *L1/MY* and *E2* regions of HPV16 was very similar. However, the frequency of non-synonymous amino-acid substitutions in *E2* was higher than in *E6* and *L1/MY* (see Table III). In fact, the non-synonymous-to-synonymous ratio was 2.8-fold and 7-fold higher in *E2* than in *E6* and *L1/MY*, respectively.

#### Correlation between viral copy number, Southern-blot pattern and HPV16 variants

The viral-copy number per cell varied according to the presence of *E1/E2* and the off-sized bands in the Southern blots. In agreement with previous observations, the detection of *E1/E2* correlates with high viral-copy number (mean =  $134.1 \pm 21.1$ , range 5–522), the disruption of *E1/E2* with low viral-copy number (mean = 15.2, range 1–60; 2-tailed Mann-Whitney U-test,  $p = 0.0001$ ; Fig. 1). However, *E1/E2*-positive samples with off-sized bands had higher viral-copy numbers (mean =  $207.8 \pm 29.2$ , range 44–522) than those positive for *E1/E2* without off-sized bands ( $51.3 \pm 11.7$ ; 2-tailed Mann-Whitney U-test,  $p = 0.0001$ ).

Copy number per cell was also significantly different for AA and E variants, the mean copy number of AA ( $170 \pm 38.9$ ) was higher than for E ( $67.6 \pm 13.5$ ) variants (2-tailed Mann-Whitney U-test;



**FIGURE 2** – Identification of HPV16 *E2* variants by changes in Southern-blot pattern. (a) Southern-blot pattern of 6 HPV16-positive DNAs derived from cervical carcinomas digested with PstI and BamHI. Samples in lanes a, e and g show the reference pattern, while lanes b, c and h show the variant pattern. (b) PstI/BamHI restriction and genome map of reference HPV16. The PstI sites gained (position 3561) and lost (position 3693) in the *E2* gene of the variant HPV16 are indicated. Positions of restriction sites are located at first nucleotide of the sites. Fragments A, B, C, D, E, F and G in the map correspond to the bands of 2.8, 1.8, 1.1, 0.9, 0.6 and 0.5 kb, respectively, in the Southern blot of (a).

$p = 0.0065$ ; Fig. 1). If tumors are distributed by copy number per cell into 3 groups, 1–10, 11–50 and >50 (Table IV), the AA variants were concentrated (12 out of 16) in the high-copy-number category (more than 50 copies/cell), whereas the E variants were distributed evenly among the 3 groups (Chi-squared test;  $p = 0.024$ ; see Table IV). However, this difference disappears when the tumors were classified according to the presence or absence of *E1/E2*. Although, the distribution of AA and E variants was not statistically different among the *E1/E2* groups (Chi-squared test;  $p = 0.054$ ), the AA variants predominate in the *E1/E2*-positive samples with off-sized bands. In fact, when the copy numbers and the *E1/E2* groups were considered together in the analysis, the difference was

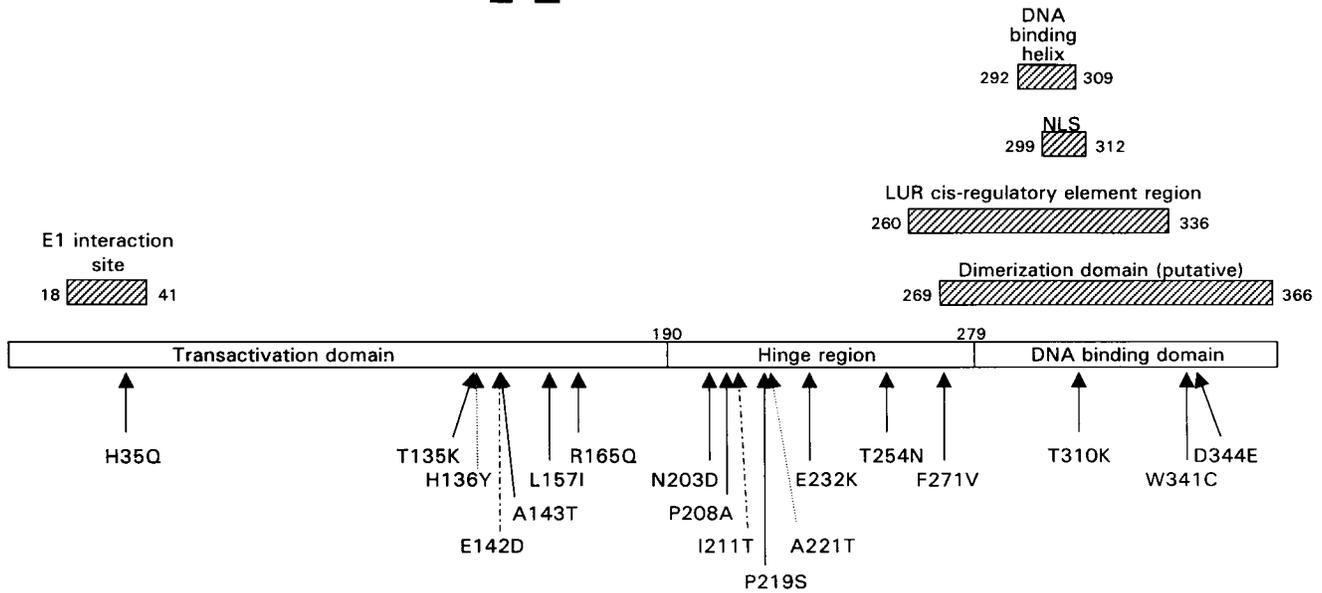
evident; the AA-positive samples were concentrated (62.5%, 10 of 16) in the group positive for *E1/E2* with off-sized bands and more than 50 copies per cell (Chi squared test;  $3 \times 3$  contingency table;  $p = 0.035$ ), whereas the E-positive samples were distributed similarly among the groups (Table IV).

#### DISCUSSION

In this report, we present the analysis of the *E2*-gene sequence of 2 Asian-American variants of HPV16 and its association with the presence of *E1/E2* genes and a high level of viral amplification in cervical carcinomas. These *E2* variants contain extensive base



## E 2



**FIGURE 4** – Localization of 18 amino-acid changes throughout the Asian-American E2 variant proteins. Figure shows the distribution of 18 amino-acid changes along a schematic representation of the E2 protein, including the transactivation, hinge region and DNA-binding domains. Continuous arrows represent substitutions shared by both variants; dotted arrows, exclusive substitutions for AA-c E2 isolates; dashed arrows, exclusive substitutions for AA-a E2 isolates. Functional domains are shown in hatched boxes above the E2 protein scheme. NLS, nuclear localization signal; LUR, late upstream region.

**TABLE III** – COMPARISON OF NUCLEOTIDE- AND AMINO-ACID-SEQUENCE VARIABILITY OF E6, L1/MY AND E2 REGIONS BETWEEN ASIAN-AMERICAN CLASS/SUB-CLASS AA-a AND E REFERENCE SEQUENCE

Region	Number of bases sequenced	Mutations <sup>1</sup>		Amino acids	Synonymous or silent substitutions <sup>2</sup>		Non-synonymous substitutions <sup>3</sup>		Non-synonymous/synonymous substitutions
		Number	%		Number	%	Number	%	
E6	456	6	1.3	152	3	2.0	3	2.0	1.0
L1/MY	411	7	1.7	137	5	3.6	2	1.5	0.4
E2	1098	22	2.0	366	6	1.6	16	4.4	2.8
Aggregate <sup>4</sup>	1965	35	1.8	655	14	2.1	21	3.2	1.5

<sup>1</sup>Number and percentage of mutations of AA-a isolates compared with reference E sequence. –<sup>2</sup>Number and percentage divergence for synonymous changes between AA-a and E reference sequence. –<sup>3</sup>Number and percentage divergence for non-synonymous changes between AA-a and E reference sequence. –<sup>4</sup>Aggregate based on combined data from E6, L1/MY and E2 sequences.

than those with episomal pattern. The strong correlation of AA variants with high copy number and with the presence of E1/E2 plus off-sized bands in the Southern blots suggest that these variants are tandem-integrated or as multimeric episomes.

It is conceivable that the extensive amino-acid substitutions in the E2 protein, a viral protein with major regulatory functions, may alter its molecular activity as well as the outcome of HPV infections. Of these changes, three (T310K, W341C and D344E), both in AA-a and in AA-c variants, fell into the DNA-binding domain, which is essential for the regulatory function of E2. Especially important is the change from a threonine to a lysine in position 310, just beside the DNA-binding helix (a.a. 292–309) of the E2 protein; the conformation structure of this helix may be disturbed if the binding capacity of the E2 protein is impaired. Although the genome-replication capability of E2 variants was not measured *in vitro*, the strong signals in the Southern-blot experiments suggest that E2-variant proteins have conserved the capability of viral replication. Only one base change was found, in the 5' region of the transactivation domain (H35Q), suggesting that this region of the HPV16 E2 is subjected to strong negative selection pressure. The main difference between AA-a and AA-c at the E2-protein level is that AA-c changed a histidine for a tyrosine at position 136 in the 3' extreme of the transactivation domain.

**TABLE IV** – DISTRIBUTION OF HPV16-POSITIVE CERVICAL CARCINOMAS AMONG COPY NUMBER/CELL, SOUTHERN-BLOT PATTERN AND HPV16 VARIANTS

Copy number/cell	E1/E2-negative (n = 12)	E1/E2-positive		All <sup>3</sup>
		Without off-sized bands <sup>1</sup> (n = 16)	With off-sized bands <sup>2</sup> (n = 18)	
<b>Asian-American variants<sup>4</sup></b>				
1–10	0	1	0	1
11–50	1	2	0	3
>50	1	1	10	12
Total	2	4	10	16
<b>European variants</b>				
1–10	7	1	0	8
11–50	3	8	1	12
>50	0	3	7	10
Total	10	12	8	30

<sup>1</sup>Samples with the authentic PstI/BamHI cleavage pattern without off-sized bands in the Southern blot. –<sup>2</sup>Samples with the authentic PstI/BamHI cleavage pattern plus off-sized bands in the Southern blot. –<sup>3</sup>Distribution of E and AA positive samples among copy number groups was significantly different (Chi-squared;  $p = 0.024$ ). –<sup>4</sup>Distribution of AA variants was significantly different among copy number and E1/E2 groups (Chi squared;  $p = 0.035$ ).

The HPV16 E2 protein appears to be the target both of humoral (Dillner, 1990; De SanJose *et al.*, 1996) and of cellular (Kónya *et al.*, 1997) immune response. Several humoral epitopes are known in the E2 protein, one in the transactivation domain (121–140), 3 in the hinge region (181–200, 241–260 and 271–290) and one in the DNA-binding domain (328–346; Dillner, 1990). All but one (181–200) of these epitopes changed in the E2 protein of AA variants, suggesting that the humoral response to these epitopes may be different for the E reference and for the AA variants. Terry *et al.* (1997) reported HPV16 European E2 variants associated essentially with changes in immunodominant epitopes of the hinge region of E2 protein: they described base changes in 11 positions, but only 5 changed the codified amino acid, 2 in the transactivation domain and 3 in the hinge region. Surprisingly, all 11 codon positions were also changed in the AA variants, but, unlike the European variants, 10 changed the amino-acid codification. The presence of HPV16 variants may explain the poor correlation between the presence of antibodies to HPV16 E2 and the detection of HPV16 DNA (De SanJose *et al.*, 1996).

The high frequency of E1/E2-positive samples in this and other reported tumor samples (Berumen *et al.*, 1994) may be related in part to the presence of AA variants in a high percentage of patients with cervical cancer. However, other factors not related to AA E2 variation may be associated with the high frequency of E1/E2-positive tumor samples, since only 41.2% of E1/E2-positive tumors had AA variants. The low frequency of disruption of E1/E2 indicates that disruption of the E2 gene is not necessary to de-regulate the expression of E6 and E7 viral oncogenes. The association of Asian-American E2 variants with retention of E1/E2 suggests that E2 variation may be an alternative mechanism for de-regulating the expression of viral oncogenes.

#### ACKNOWLEDGEMENTS

We are grateful to Dr. E.R. Unger for stimulating discussions and helpful suggestions regarding this report and to Dr. L. Benítez, who provided the tumor samples. We thank Drs. M. Castañeda, R. Lisker and C. Gorodezky for reading the manuscript.

#### REFERENCES

- BERUMEN, J., CASAS, L., SEGURA, E., AMEZCUA, J.L. and GARCÍA-CARRANCA, A., Genome amplification of human papillomavirus types 16 and 18 in cervical carcinomas is related to retention of E1/E2 genes. *Int. J. Cancer*, **56**, 640–645 (1994).
- BONATTO, S.L. and SALZANO, F.M., Diversity and age of the four major mtDNA haplogroups, and their implications for the peopling of the New World. *Amer. J. hum. Genet.*, **61**, 1413–1423 (1997).
- COOPER, K., HERRINGTON, C.S., GRAHAM, A.K., EVANS, M.F. and MCGEE, J.O'D., *In situ* evidence for HPV 16, 18, 33 integration in cervical squamous-cell cancer in Britain and South Africa. *J. clin. Pathol.*, **44**, 406–409 (1991).
- CULLEN, A.P., REID, R., CAMPION, M. and LÖRINCZ, A.T., Analysis of the physical state of different human papillomavirus DNAs in intraepithelial and invasive cervical neoplasm. *J. Virol.*, **65**, 606–612 (1991).
- DAS, B.C., SHARMA, J.K., GOPALAKRISHNA, V. and LUTHRA, U.K., Analysis by polymerase chain reaction of the physical state of human papillomavirus type 16 DNA in cervical preneoplastic and neoplastic lesions. *J. gen. Virol.*, **73**, 2327–2336 (1992).
- DE SANJOSÉ, S. and 14 OTHERS, Serological response to HPV16 in CIN-III and cervical-cancer patients. Case-control studies in Spain and Colombia. *Int. J. Cancer*, **66**, 70–74 (1996).
- DILLNER, J., Mapping of linear epitopes of human papillomavirus type 16: the E1, E2, E4, E5, E6 and E7 open reading frames. *Int. J. Cancer*, **46**, 703–711 (1990).
- HO, L. and 21 OTHERS, The genetic drift of the human papillomavirus type 16 is a means of reconstructing prehistoric viral spread and the movement of ancient human populations. *J. Virol.*, **67**, 6413–6423 (1993).
- KÓNYA, J., EKLUND, C., GEJERSSTAM, V., YUAN, F., STUBER, G. and DILLNER, J., Identification of a cytotoxic T-lymphocyte epitope in the human papillomavirus type 16 E2 protein. *J. gen. Virol.*, **78**, 2615–2620 (1997).
- MATSUKURA, T., KOI, S. and SUGASE, M., Both episomal and integrated forms of human papillomavirus type 16 are involved in invasive cervical cancers. *Virology*, **172**, 63–72 (1989).
- MEANWELL, C.A., COX, M.F., BLACKLEDGE, G. and MAITLAND, N.J., HPV16 DNA in normal and malignant cervical epithelium: implications for the aetiology and behaviour of cervical neoplasia. *Lancet*, **I**, 703–707 (1987).
- PICCINI, A., STOREY, A., MASSIMI, P. and BANKS, L., Mutations in the human papillomavirus type 16 E2 protein identify multiple regions of the protein involved in binding to E1. *J. gen. Virol.*, **76**, 2909–2913 (1995).
- ROBERTS, S., ASHMOLE, I., GIBSON, L.J., ROOKES, S.M., BARTON, G.J. and GALLIMORE, P.H., Mutational analysis of human-papillomavirus E4 proteins: identification of structural features important in the formation of cytoplasmic E4/cytokeratin networks in epithelial cells. *J. Virol.*, **68**, 6432–6445 (1994).
- SCHNEIDER-MAUNOURY, S., CROISSANT, O. and ORTH, G., Integration of human papillomavirus type 16 DNA sequences: a possible early event in the progression of genital tumors. *J. Virol.*, **61**, 3295–3298 (1987).
- SEEDORF, K., KRÄMMER, G., DÜRST, M., SUHAI, S. and RÖWEKAMP, W.G., Human papillomavirus type 16 DNA sequence. *Virology*, **145**, 181–185 (1985).
- TERRY, G., HO, L. and CUZICK, J., Analysis of E2 amino acid variants of human papillomavirus types 16 and 18 and their associations with lesion grade and HLA DR/DQ type. *Int. J. Cancer*, **73**, 651–655 (1997).
- VERNON, S.D., UNGER, E.R., MILLER, D.L., LEE, D.R. and REEVES, W.C., Association of human papillomavirus type 16 integration in the E2 gene with poor disease-free survival from cervical cancer. *Int. J. Cancer*, **74**, 50–56 (1997).
- WAGATSUMA, M., HASHIMOTO, K. and MATSUKURA, T., Analysis of integrated human papillomavirus type 16 DNA in cervical cancers: amplification of viral sequences together with cellular flanking sequences. *J. Virol.*, **64**, 813–821 (1990).
- WILCZYNSKI, S.P., PEARLMAN, L. and WALKER, J., Identification of HPV 16 early genes retained in cervical carcinomas. *Virology*, **166**, 624–627 (1988).
- YAMADA, T., MANOS, M.M., PETO, J., GREER, C.E., MUÑOZ, N., BOSCH, F.X. and WHEELER, C.M., Human papillomavirus type 16 sequence variation in cervical cancers: a worldwide perspective. *J. Virol.*, **71**, 2463–2472 (1997).
- YAMADA, T., WHEELER, C.M., HALPERN, A.L., STEWART, A.C.M., HILDESHEIM, A. and JENISON, S.A., Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, L2 and L1 coding segments. *J. Virol.*, **69**, 7743–7753 (1995).