



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
POSGRADO EN CIENCIAS BIOLÓGICAS

FACULTAD DE MEDICINA
BIOMEDICINA

PATRONES DE EXPRESIÓN GENÉTICA ASOCIADOS A LOS
VIRUS DEL PAPILOMA HUMANO EN CÁNCER DE CÉRVIX

TESIS

QUE PARA OPTAR POR EL GRADO DE:

DOCTORA EN CIENCIAS

PRESENTA:

Ana María Espinosa García

DIRECTOR DE TESIS: DR. JAIME BERUMEN CAMPOS
FACULTAD DE MEDICINA

COMITÉ TUTOR

DR. RAFEL CAMACHO CARRANZA
INSTITUTO DE INVESTIGACIONES BIOMEDICAS

DR. MAURICIO SALCEDO VARGAS
POSGRADO EN CIENCIAS BIOLÓGICAS

MÉXICO, D.F., AGOSTO, 2013.



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
POSGRADO EN CIENCIAS BIOLÓGICAS

FACULTAD DE MEDICINA

BIOMEDICIA

PATRONES DE EXPRESIÓN GENÉTICA ASOCIADOS A LOS
VIRUS DEL PAPILOMA HUMANO EN CÁNCER DE CÉRVIX

TESIS

QUE PARA OPTAR POR EL GRADO DE:

DOCTORA EN CIENCIAS

PRESENTA:

Ana María Espinosa García

DIRECTOR DE TESIS: DR. JAIME BERUMEN CAMPOS
FACULTAD DE MEDICINA

COMITÉ TUTOR

DR. RAFEL CAMACHO CARRANZA
INSTITUTO DE INVESTIGACIONES BIOMEDICAS

DR. MAURICIO SALCEDO VARGAS
INSTITUTO MEXICANO DEL SEGURO SOCIAL

MÉXICO, D.F., AGOSTO, 2013.



Dr. Isidro Ávila Martínez
Director General de Administración Escolar, UNAM
Presente

Me permito informar a usted que en la reunión ordinaria del Subcomité de Biología Experimental y Biomedicina del Posgrado en Ciencias Biológicas, celebrada el día 11 de marzo de 2013, se aprobó el siguiente jurado para el examen de grado de **DOCTORA EN CIENCIAS** de la alumna **ESPINOSA GARCÍA ANA MARÍA** con número de cuenta **95808354** con la tesis titulada **PATRONES DE EXPRESIÓN GENÉTICA ASOCIADOS A LOS VIRUS DEL PAPILOMA HUMANO EN CÁNCER DE CERVIX** realizada bajo la dirección del **DR. JAIME BERUMEN CAMPOS**:

Presidente: DRA. MARCELA LIZANO SOBERÓN
Vocal: DRA. VILMA MALDONADO LAGUNA
Secretario: DR. RAFAEL CAMACHO CARRANZA
Suplente: DRA. LETICIA ROCHA ZAVALA
Suplente: DR. MAURICIO SALCEDO VARGAS

Sin otro particular, me es grato enviarle un cordial saludo.

ATENTAMENTE
"POR MI RAZA HABLARA EL ESPIRITU"
Cd. Universitaria, D.F., a 6 de junio de 2013

DRA. MARIA DEL CORO ARIZMENDI ARRIAGA
COORDINADORA DEL PROGRAMA

AGRADECIMIENTOS

Al Posgrado en Ciencias Biológicas, UNAM por permitirme realizar estos estudios.

Al CONACYT, por los apoyos recibidos durante la realización de esta tesis, inicialmente como becaria de CONACYT No. 194872 y este proyecto además contó con el financiamiento CONACYT-Salud-2002-C01-8135/A-1.

Al Dr. Jaime Berumen Campos, Director de la tesis. Gracias por permitirme realizar este proyecto y trabajar en su laboratorio.

Al Dr. Rafael Camacho Carranza y al Dr. Mauricio Salcedo Vargas, miembros de comité tutor, y además miembros de jurado. Gracias por sus comentarios y aportaciones durante las evaluaciones así como también en la revisión final de la tesis.

Agradecimientos a título personal

A la Dra. Marcela Lizano Soberón, Dra. Vilma Maldonado Lagunas y Dra. Leticia Rocha Zabaleta, miembros de jurado para el examen de grado. Gracias por sus valiosos comentarios y aportaciones, de gran ayuda cada uno de ellos en la revisión final de la tesis.

Al Dr. Edgar Román Basaure, Jefe del Servicio de Oncología del HGM; Dr. Sergio Muñoz Cortez y Dr. Manuel Bagues Ibáñez del Servicio de Ginecología del HGM; Dra. Patricia Alonso del Servicio de Patología del HGM, a todos ustedes gracias por su ayuda en la realización de este proyecto.

A mis compañeros y amigos de la unidad de Medicina Genómica América, Maythe, Cynthia, Anita, Maura, Elenita, Lilia, Lupita, Luis y Jonatan, con los que he compartido momentos gratos y difíciles durante el largo camino realizado para esta tesis.

A la Dra. Ícela Lara Palma, a quien tuve la oportunidad de conocer durante la realización de la IH y de quien he tenido una gran enseñanza, gracias.

A Marco Gudiño, por la ayuda y asesoría en la micrografía, diseño y edición final de imágenes, muchísimas gracias.

A mis compañeros estudiantes, que en algún momento compartimos la estancia en el laboratorio, Sonia, Zoraida, Ángeles, Edna, Ivonne, Valeria, Ingrid, Ulises, Oscar y David.

A todas las personas que de alguna u otra forma colaboraron con la realización de este proyecto.

DEDICATORIA:

♥ A Yariela y Manuel, quienes representan todo en mi vida.

♥ Hijita porque tú sigues y seguirás siendo la luz en mi camino, TQM.

♥ Mi cucho, mi esposo, gracias por tu apoyo.

ÍNDICE GENERAL

	Página
Agradecimientos.....	I
Dedicatoria.....	II
Índice General.....	III
Índice de figuras.....	IV
Abreviaturas.....	V
Resumen.....	VII
I. INTRODUCCIÓN1	
1 Cáncer cervical.....	1
2 Histología.....	2
3 Principales factores de riesgo asociados al desarrollo del Ca-Cu	4
4 Tipos virales y variantes.....	6
5 Biología del VPH.....	8
6 El ciclo de vida de los VPH en los epitelios en diferenciación.....	12
7 Genes celulares alterados o inducidos por los Oncogenes virales del VPH.....	13
8 Vacunas.....	16
9 Marcadores tumorales.....	16
10 Pruebas de diagnóstico, pronóstico y de seguimiento.....	18
II. HIPOTESIS.....	21
III. OBJETIVOS.....	22
IV. MATERIAL Y MÉTODOS.....	23
1 Selección de pacientes.....	23
2. Diseño experimental.....	26
3. Extracción de DNA y RNA.....	27
4. Detección y tipificación del VPH.....	27
5. Análisis del perfil de expresión génica y análisis de datos.....	29
6. Validación de la expresión génica global por un segundo ensayo de microarreglos de alto rendimiento (Microarray HG-ST1.0).....	30
7. Validación de la expresión génica global por RT-PCR en	

tiempo real cuantitativa (qRT-PCR).....	31
8. Inmunohistoquímica.....	33
9. Análisis de supervivencia de pacientes con cáncer.....	35
10. Análisis de clasificación de Ontología de genes.....	37
11. Anotación de los genes y el análisis de datos.....	37
V. RESULTADOS.....	39
1. Análisis de expresión de 8,638 genes en cáncer cervical.....	39
2. Validación de genes expresados diferencialmente con el microarreglo HG-ST1.0.....	42
3. Validación por PCR cuantitativa a tiempo real (qRT-PCR).....	45
4. Análisis de la expresión de <i>CCNB2</i> , <i>CDC20</i> , <i>PRC1</i> , <i>SYCP2</i> , <i>NUSAP1</i> y <i>CDKN3</i> en neoplasias pre-invasivas.....	50
5. Verificación de la expresión de genes candidatos a marcadores tumoraes a nivel de proteínas por inmunohistoquímica.....	54
6. Blancos moleculares en cáncer cervical asociados con la baja sobrevivida.....	60
7. Clasificación de los genes con expresión diferencial ente las muestras de cáncer y las muestras control.....	62
VII. DISCUSIÓN.....	68
VIII. CONCLUSIONES.....	78
IX. REFERENCIAS.....	79

ÍNDICE DE FIGURAS Y TABLAS.

- Figura 1. Progresión de cáncer cervical mediada por HPV.
- Figura 2. Zona de transformación, en epitelio cervical.
- Figura 3. Principales etapas de desarrollo en cáncer cervical.
- Figura 4. Distribución de VPH 16 /18 en Ca-Cu en América central y del sur
- Figura 5. Distribución global de las clases y subclases de VPH por continente.
- Figura 6. Esquema de la organización del genoma de VPH.
- Figura 7. Secuencia de expresión de genes virales a través de la diferenciación de queratinocitos en el epitelio cervical estratificado.
- Figura 8. Mecanismo de transformación mediado por las proteínas virales E6 y E7 en cáncer cervical.
- Figura 9. Electroforesis en gel de agarosa al 1% de RNA proveniente de biopsias. A) Controles, B) tumores, ejemplos representativos.
- Figura 10. Distribución de genes desregulados según el cambio en número de veces (FC) y los valores de Δ -score.
- Figura 11. Agrupamiento de las muestras del tumor y control según la expresión de genes desregulados.
- Figura 12. Correlación global de la intensidad de la expresión de 997 genes examinados por microarreglos HG-focus y HG-ST1.0.
- Figura 13. Correlación de la expresión de la intensidad de 23 genes entre 2 microarreglos, examinados por microarreglos HG-focus y HG-ST1.0
- Figura 14. Validación de la expresión génica de 9 marcadores genéticos por qRT-PCR.
- Figura 15 Validación de la expresión de marcadores que están disminuidos por qRT-PCR.
- Figura 16. Microarreglo de tejidos (MAT).
- Figura 17. Análisis histológico de los genes marcadores.
- Figura 18. Porcentaje de las células tumorales con señales positivas para los marcadores probados por inmunohistoquímica.
- Figura 19. Análisis de sobrevida de mujeres con cáncer cervical según la clasificación de FIGO y expresión del gen *CDKN3*.
- Figura 20. Vías canónicas en donde están implicados los genes desregulados.

TABLAS.

- Tabla I. Tamaño y función de las proteínas del papiloma virus
- Tabla II. Revisión de artículos que evalúan sensibilidad y especificidad en muestras de cáncer cervical.
- Tabla III. Resumen de los datos clínicos, tipo viral, edad, arreglo empleado para la expresión.
- Tabla IV. Sondas TaqMan para ensayos de expresión de Applied Biosystems.
- Tabla V. Seguimiento de pacientes para evaluación de la sobrevida por al menos 42 meses.

Tabla VI. Genes explorados por qRT-PCR.

Tabla VII. Análisis ROC y cálculos de sensibilidad, especificidad y valores predictivos

Tabla VIII. Análisis ROC de 4 genes marcadores seleccionados para las lesiones de alto grado y CC.

Tabla IX. Análisis de la 10 proteínas en 26 muestras de CC y 10 controles de epitelio cervical normal por IH y el cálculo de sensibilidad, especificidad y valores predictivos

Tabla X. Análisis funcional con la herramienta DAVID con alto rigor para los 997 genes alterados en Cáncer cervical.

Tabla XI. Análisis funcional con la herramienta DAVID con alto rigor para los 100 genes más alterados en Cáncer cervical en comparación con el epitelio cervical normal

Tabla XII. Comparación de los 23 genes identificados y validados en este trabajo, con reportes de análisis con microarreglos documentados

Abreviaturas

Ca-Cu	Cáncer cervical
NIC	Neoplasia intraepitelial cervical
VPH	Virus del Papiloma Humano
DNA	Acido desoxirribonucleico
VIH	Virus inmunodeficiencia humana
EST	Enfermedades de transmisión sexual
PAP	Papanicolaou
VPH-BR	Virus del Papiloma Humano de bajo riesgo
VPH-AR	Virus del Papiloma Humano de alto riesgo
AA	Asiático-Americanas
Af-1	Africana-1
AF-2	Africana-2
As	Asiáticas
Eu	Europeas
E	Temprana
L	Tardía
LCR	Región larga de control
RNA	Acido ribonucleico
RNA _m	RNA mensajero
HGM	Hospital General de México
FIGO	Federación Internacional de ginecología y obstetricia
CCE,	Carcinoma de células escamosas
ACC	Adenocarcinoma
ASCC	Adenoescamoso
ECN	Epitelio cervical normal
NIC,	Neoplasia intraepitelial cervical
MA,	Microarreglos
qRT-PCR	Transcripción reversa en tiempo real PCR
IH	Inmunohistoquímica
ND	No determinado
cRNA	RNA complementario
cDNA	DNA complementario
ds-cDNA	cDNA de doble cadena
PE	Ficoeritrina
GAPDH	Gliceraldehido-3 fosfato deshidrogenasa
RMA	Robust Multichip Average
SAM	Significance Analysis of Microarrays
FDR	False discovery rate
dNTP	Deoxinucleotido trifosfato
RT	Transcriptasa reversa
FC	Fold-change
MW	Mann–Whitney

MAT	Microarreglos de tejidos
H&E	Hematoxilina y Eosina
PBS	Amortiguador de fosfatos salino
AUC	Área bajo la curva (curvas ROC)
DAVID	The Database for Annotation, Visualization, and Integrated Discovery
IPA	Ingenuity Pathway Analysis IPA
VPP	Valor predictivo positivo
VPN	Valor predictivo negativo
PCA	Análisis de componente principal
pRB	Proteína de retinoblastoma
FDA	Food and Drug Administration
HR	Hazard ratio
HC2	Captura de híbridos
CBL	Citología de Base líquida
ELISA	Enzyme-Linked ImmunoSorbent Assay
DBS	Doble cadena rota

Resumen

El efecto de la vacunación preventiva contra el virus del papiloma humano (VPH) en la reducción del problema del cáncer cervical (Ca-Cu) no se conocerá hasta dentro de unos 30 años. Por lo tanto, es necesario mejorar los procedimientos de la detección y el tratamiento del Ca-Cu. El objetivo de este estudio fue identificar y caracterizar los blancos celulares que podrían considerarse posibles marcadores para la detección o blancos terapéuticos, utilizando una estrategia piramidal. Inicialmente la expresión de 8.638 genes se comparó entre 43 Ca-Cu positivo para el VPH16 y 12 epitelios cervicales sanos utilizando microarreglos. Encontrando un total de 997 genes desregulados y de estos, los 21 genes que mostraron la mayor desregulación adicionalmente fueron validados mediante qRT-PCR. Los 6 genes que se sobre expresaron en mayor proporción (*CCNB2*, *CDC20*, *PRC1*, *SYCP2*, *NUSAP1*, *CDKN3*) pertenecen a la vía de la mitosis. Estos genes se exploraron también en neoplasias intraepitelial cervical de alto y bajo grado, 29 de bajo grado (NIC1) y 21 NIC de alto grado (NIC2/3), para investigar si podía distinguir Ca-Cu y NIC2/3 (NIC2+) de NIC1 y controles. Los genes *CCNB2*, *PRC1* y *SYCP2* se asociaron mayormente con Ca-Cu y los genes *CDC20*, *NUSAP1* y *CDKN3* también se asociaron con NIC2/3. La sensibilidad y especificidad de *CDKN3* y *NUSAP1* para detectar NIC2+ fue aproximadamente el 90%. Las proteínas codificadas por los 6 genes mostraron también expresarse en Ca-Cu por inmunohistoquímica. La Asociación de estos marcadores con sobrevida, fue investigada en 42 pacientes seguidas por al menos 42 meses. Sólo el gen *CDKN3* se asoció con baja sobrevida y esto fue de forma independiente de la etapa clínica (HR = 5.9, IC del 95% = 1.4–23.8, p = 0.01). Los genes *CDKN3* y *NUSAP1* pueden ser blancos potenciales para el desarrollo de métodos de detección temprana. Sin embargo, se requieren estudios adicionales futuros con un mayor número de muestras para definir el nivel óptimo de sensibilidad y especificidad. Así mismo la inhibición de la mitosis es una estrategia bien conocida para combatir el cáncer. Por lo tanto, *CDKN3* puede ser no sólo un marcador de detección y sobrevida sino

un potencial blanco terapéutico en Ca-Cu. Sin embargo, aun hace falta demostrar si es indispensable para el crecimiento del tumor.

Abstract

The effect of preventive human papillomavirus (HPV) vaccination on the reduction of the cervical cancer (CC) burden will not be known for 30 years. Therefore, it's still necessary to improve the procedures for CC screening and treatment. The objective of this study was to identify and characterize cellular targets that could be considered potential markers for screening or therapeutic targets. A pyramidal strategy was used. Initially the expression of 8,638 genes was compared between 43 HPV16-positive CCs and 12 healthy cervical epitheliums using microarrays. A total of 997 genes were deregulated, and 21 genes that showed the greatest deregulation were validated using qRT-PCR. The 6 most upregulated genes (CCNB2, CDC20, PRC1, SYCP2, NUSAP1, CDKN3) belong to the mitosis pathway. They were further explored in 29 low grade cervical intraepithelial neoplasias (CIN1) and 21 high-grade CIN (CIN2/3) to investigate whether they could differentiate CC and CIN2/3 (CIN2+) from CIN1 and controls. CCNB2, PRC1, and SYCP2 were mostly associated with CC and CDC20, NUSAP1, and CDKN3 were also associated with CIN2/3. The sensitivity and specificity of CDKN3 and NUSAP1 to detect CIN2+ was approximately 90%. The proteins encoded by all 6 genes were shown upregulated in CC by immunohistochemistry. The association of these markers with survival was investigated in 42 CC patients followed up for at least 42 months. Only CDKN3 was associated with poor survival and it was independent from clinical stage (HR = 5.9, 95%CI = 1.4–23.8, p = 0.01). CDKN3 and NUSAP1 may be potential targets for the development of screening methods. Nevertheless, further studies with larger samples are needed to define the optimal sensitivity and specificity. Inhibition of mitosis is a well-known strategy to combat cancers. Therefore, CDKN3 may be not only a screening and survival marker but a potential therapeutic target in CC. However, whether it's indispensable for tumor growth remains to be demonstrated..

INTRODUCCIÓN.

1. Cáncer Cervical.

El Cáncer cérvico-uterino (Ca-Cu) es el segundo cáncer más común en las mujeres alrededor del mundo y en los países en desarrollo es la primera causa de muerte, el 80% de los 500,000 casos nuevos que ocurren anualmente en el mundo se presentan en países en desarrollo [1,2]. Su tasa de incidencia varía de 6 casos por 100,000 en muchas naciones industrializadas hasta más de 50 por 100,000 en algunos países en desarrollo [3], reportándose las mayores tasas de incidencia en América Latina, África, sur de Asia y suroeste de Asia. En México es el segundo cáncer más común en las mujeres y tiene una incidencia de 44 casos por 100,000 mujeres [4].

El Ca-Cu es una enfermedad que se desarrolla por etapas. La clasificación de estas etapas está basada en los cambios morfológicos del epitelio que define el grado de la lesión. La clasificación más utilizada es la neoplasia intraepitelial cervical (NIC) que se divide en grados 1, 2 y 3 (Figura 1). Los criterios para el diagnóstico de neoplasia intraepitelial varían según el criterio del médico-patólogo; aunque los aspectos más importantes a considerar son la desorganización celular, la atipia nuclear y el aumento de la actividad mitótica.

La infección con el Virus del Papiloma Humano (VPH) es el agente etiológico para el desarrollo del cáncer cervical. El DNA del VPH se encuentra en el 99.7% de los casos de Ca-Cu, de los cuales el 50% de los carcinomas son positivos para el Virus del Papiloma Humano 16 (VPH16) [5-8]. El VPH se transmite por vía sexual y representa la infección de transmisión sexual más frecuente en todo el mundo. La prevalencia de la infección en las mujeres sexualmente activas varía de 2 a 44%, los porcentajes mayores predominan en las mujeres menores de 25 años principalmente y disminuye conforme se incrementa la edad [9,10].

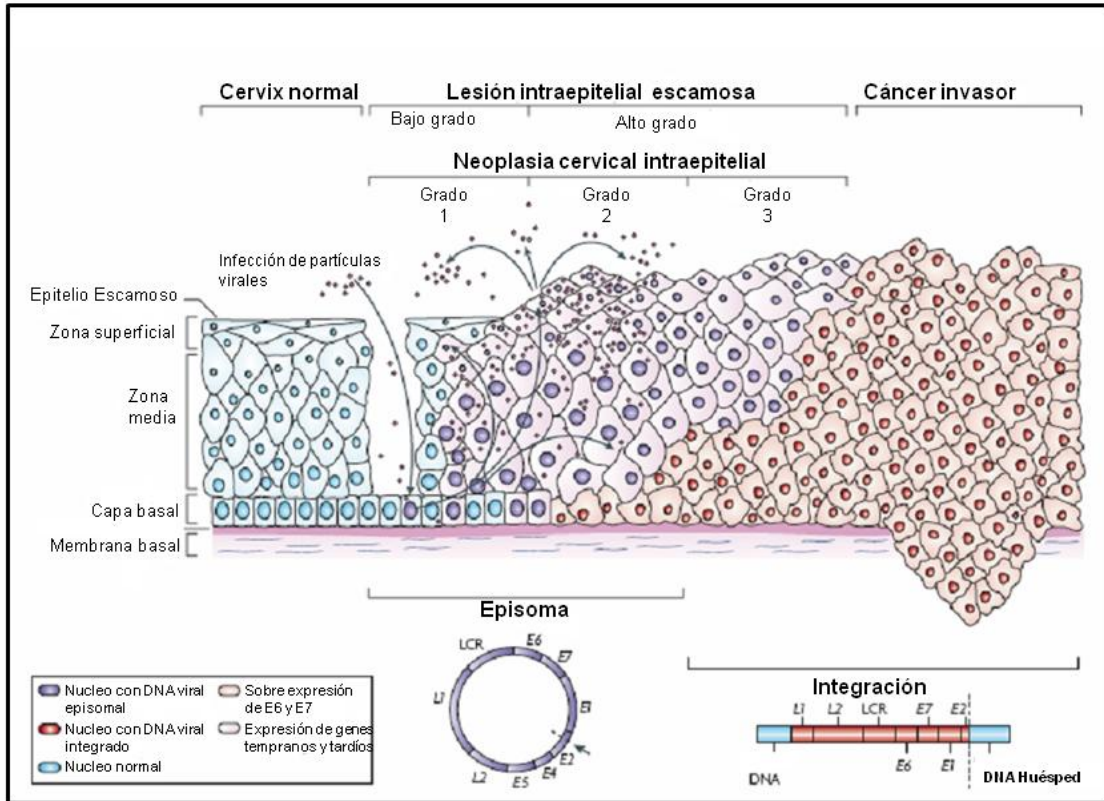


Figura 1. Progresión de cáncer cervical mediada por VPH. Las células basales se encuentran en la parte inferior en el epitelio cervical, por encima de la membrana basal, la cual es soportada a su vez por la dermis. Se ha propuesto que el mecanismo de los VPH para acceder a las células basales es través micro-abrasiones en el epitelio cervical. Después de la infección, se expresan los genes tempranos del VPH, E1, E2, E4, E5, E6 y E7 y el DNA viral se replica de forma episomal (núcleos púrpuras). En las capas superiores del epitelio (zona media y zona superficial) el genoma viral se replica aún más y se expresan los genes tardíos L1 y L2 y el gen E4. Los genes L1 y L2 para encapsular el genoma viral en la forma de viriones en el núcleo. Los nuevos virus vertidos pueden iniciar una nueva infección. Las lesiones intraepiteliales de bajo grado soportan la replicación viral, un número desconocido de infecciones con VPH de alto riesgo progresan a NIC-AG. La progresión de lesiones no tratadas hacia cáncer microinvasor e invasor que se asocia con la integración del genoma del VPH en los cromosomas del huésped (núcleos rojos), con pérdida o ruptura de E2 y la posterior sobreexpresión de los oncogenes de E6 y E7. LCR, región larga de control. (Tomado y modificado de Woodman y cols. [11]).

Sin embargo, la mayoría de las infecciones son eliminadas por función efectiva del sistema inmune y sólo un porcentaje escapa, hacia el desarrollo de lesiones preinvasoras [9].

2. Histología.

El cuello uterino está compuesto por dos tipos de epitelio, el exocervical, y el endocervical. El sitio anatómico en donde se localiza la unión de ambos epitelios

es muy variable, depende de la acción de las hormonas, lo que a su vez se relaciona con la edad. Este sitio de unión de estos dos epitelios se denomina zona de transición escamo-columnar o zona de transformación y es una región dinámica sujeta a cambios constantes (Figura 2). El epitelio cilíndrico mucosecretor, que reviste al conducto endocervical y el epitelio escamoso que recubre al exocérvix. La zona de unión entre estos epitelios relacionados con la acción de hormonas ováricas como el estrógeno, fenómenos proliferativos con el cambio de epitelio cilíndrico a escamoso (metaplasia) y donde se originan con mayor frecuencia las lesiones intraepiteliales.

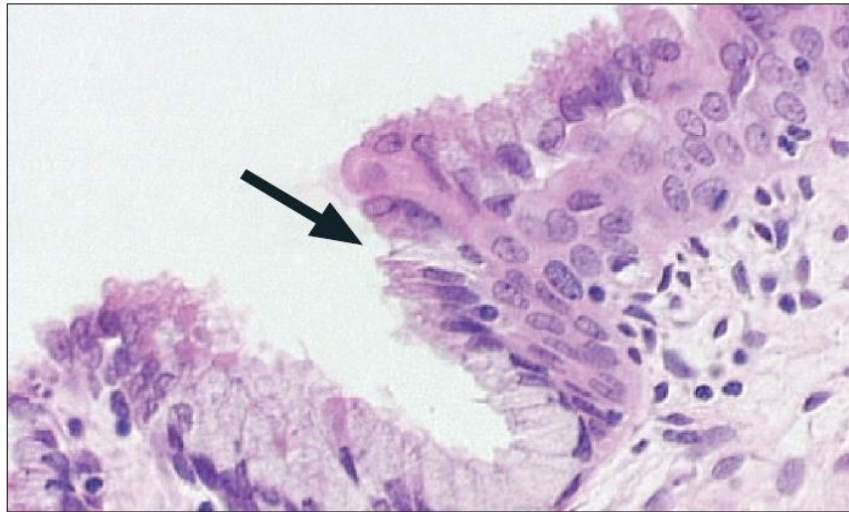


Figura 2. Zona de transformación, en el epitelio cervical. La zona de transformación cervical (indicada con la flecha), es un anillo activo metaplasia escamosa en donde el epitelio escamoso estratificado de ectocervix se debilita progresivamente y reemplaza el epitelio glandular del endocervix. Por razones no completamente claras, el tejido metaplásico es especialmente susceptible al potencial carcinógeno de las infecciones persistentes con VPH (Tomado de Schiffman y cols. [12]).

El conocimiento del origen y comportamiento del carcinoma invasor del cérvix se ha incrementado en los últimos años. Se conoce ahora que cáncer cervical se manifiesta en distintos tipos histológicos, cuya nomenclatura se basa en su comportamiento y no en características morfológicas universales que definen al cáncer. Las variedades más comunes del carcinoma de células escamosas son el de células grandes queratinizantes y no queratinizantes. Estas tienen en común

una diferenciación celular escamosa tanto en la formación intracelular de puentes como en la queratinización [12].

El tipo epidermoide es el más frecuente (90-95%), se origina en el ectocérvix (exocervix), tiene patrón laminar o en nidos de células neoplásicas fusiformes a cúbicas con queratinización individual y presencia de desmosomas. El tipo adenocarcinoma es menos frecuente, aunque recientes reportes indican el aumento de este tipo cáncer cervical, se origina del endocervix y se caracteriza por un patrón glandular recubierto por células neoplásicas cilíndricas, productoras de moco, son en general células grandes [13].

3. Principales factores de riesgo asociados al desarrollo del Ca-Cu.

El cáncer de cuello uterino se presenta mediante una serie de cuatro pasos: la transmisión del VPH, la persistencia viral, la progresión de alguna clona de células pre-cancerosas infectadas y la invasión (Figura 3). La infección por el VPH se ha demostrado que es la causa necesaria, pero no suficiente para el desarrollo del cáncer cervical. La gran mayoría de las mujeres infectadas por algún tipo de VPH de alto riesgo oncogénico no presenta Ca-Cu. Esto sugiere que otros factores adicionales actúan conjuntamente con el VPH y que influyen en el riesgo de aparición de la enfermedad. Co-factores como la paridad, el uso de anticonceptivos orales, el tabaquismo, la inmunosupresión particularmente la relacionada con el virus de la inmunodeficiencia humana (VIH), la infección por otras enfermedades de transmisión sexual (EST) y la desnutrición se han asociado. Sin embargo, la causa específica en la aparición del Ca-Cu no es completamente clara aún, otros factores también referidos como de riesgo son: la edad del comienzo de las relaciones sexuales, el número de compañeros sexuales que se han tenido y otras características de la actividad sexual se vinculan con la probabilidad de contraer el VPH. La enfermedad si se detecta en etapas tempranas teóricamente es prevenible, sin embargo en países en vías de desarrollo no es así, debido a que los programas de detección oportuna, basados

en la citología (Papanicolaou, PAP), han sido poco eficientes, o de escasa cobertura aun, por lo que esto también explicaría la alta incidencia de cáncer cervico-uterino en los países en vías en desarrollo [13-15].

El virus del papiloma humano se transmite por contacto, durante las relaciones sexuales, el uso del preservativo disminuye el contagio sexual, pero no lo elimina. Algunos tipos de virus del papiloma humano tienen alta capacidad oncogénica los denominados de alto riesgo y provocan mitosis sin control, displasia, carcinoma *in situ*, y cáncer de cuello de útero. Las mujeres se contagian al comienzo de sus relaciones sexuales, pero en la mayoría de los casos la infección se elimina por medios naturales, espontáneamente, un funcionamiento en forma adecuada del sistema inmune, en promedio solo 2 de cada 1000 mujeres desarrollara un Ca-Cu, si la infección se mantiene por más de una década [13].

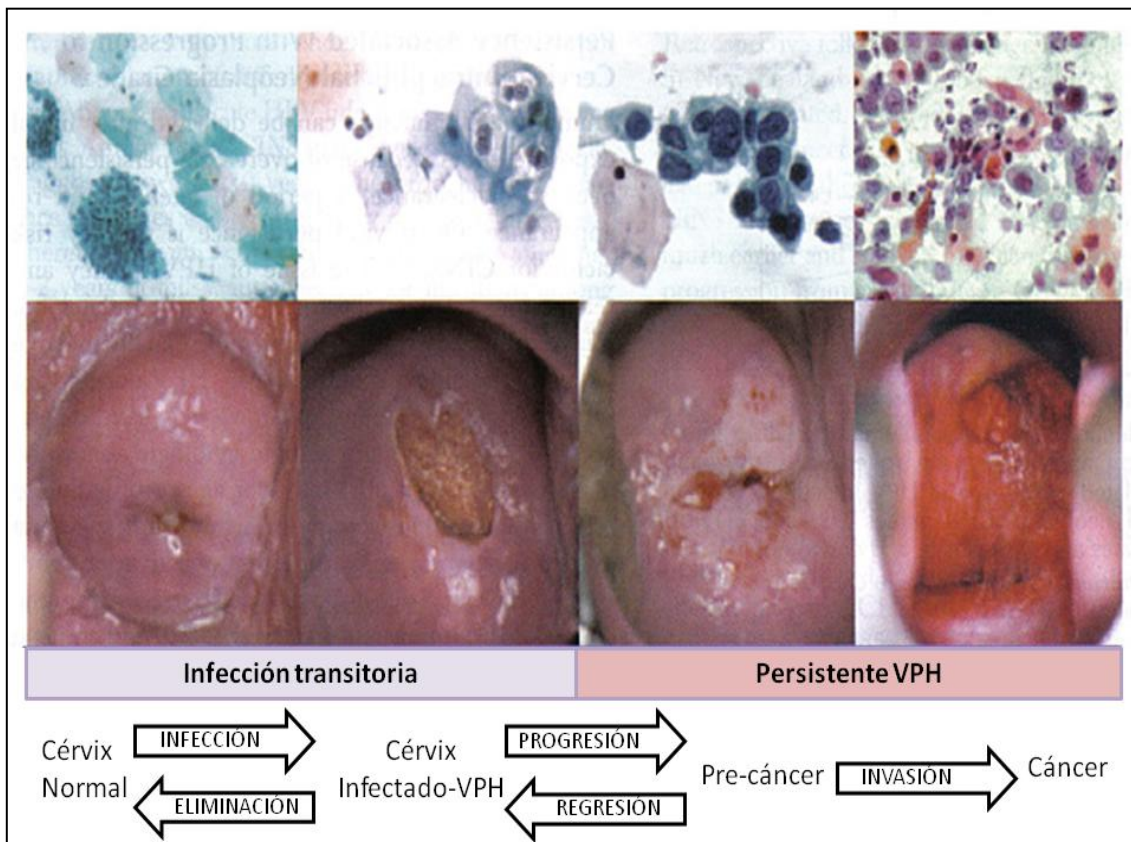


Figura 3. Principales etapas en el desarrollo de cáncer cervical. La figura muestra la comparación entre la citología parte superior y la colposcopia parte inferior. Las etapas

importantes en el desarrollo del cáncer de cuello uterino se pueden entender mejor en relación a la edad de inicio de la primera relación sexual como una aproximación a la edad de la primera infección con VPH, la cual es similar a otras infecciones transmitidas sexualmente; puede existir un gran pico en la infección inicial. El promedio de edad de la infección puede variar en las distintas culturas. La infección con el VPH, puede ser medida y monitoreada con pruebas moleculares, sin embargo la infección con VPH no necesariamente indica una lesión (tomado y modificado de Shiffman y cols. [12]).

La infección permanente es de evolución lenta, puede llevar hasta diez años o más para producir lesiones precancerosas, y otros diez años para producir carcinoma invasor. Asimismo se ha documentado que el 70% de las mujeres sexualmente activas han tenido infección por algún tipo de papiloma virus a lo largo de su vida, y que el 1% de ellas desarrollará cáncer *in situ*. Sin embargo, países como México constituyen aun es una de las principales causas de muerte por cáncer en la mujer desde hace ya varios años (datos de la secretaría de salud, www.salud.gob.mx), solamente rebasado por cáncer de mama [2].

4. Tipos virales y variantes.

Se han identificado alrededor de 200 tipos diferentes de VPH [16], pero solo 100 han sido caracterizados, de los cuales alrededor de 40 son conocidos como VPH genitales y se dividen en dos grupos: los de bajo riesgo (VPH-BR: 6,11,40,42,43,44,54,61,70,72,81) y los de alto riesgo oncogénicos (VPH-AR: 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68, 73, 82). Los tipos 26, 53 y 66 se consideran VPH-AR [17]. Estudios epidemiológicos y moleculares demostraron que la infección con VPH-AR es necesaria, aunque no suficiente, para el desarrollo de cáncer cervical [18]. Los tipos virales más predominantes en la región de América latina son el tipo 16 y el tipo 18 (Figura 4) [19].

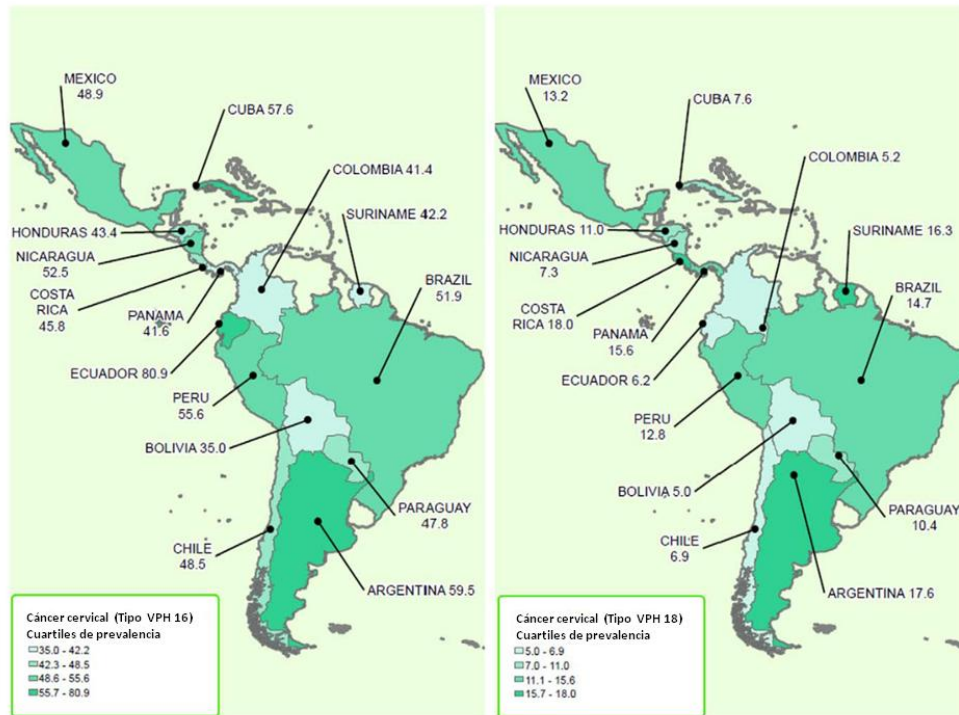


Figura 4. Distribución de VPH 16/18 en Ca-Cu en América central y del sur (tomado de Ciapponi y cols. [19]).

Los tipos virales se consideran diferentes cuando la secuencia de nucleótidos de su genoma difiere en más del 10%. Además los tipos de VPH tienen variantes genéticas, que difieren hasta en un 2% del genoma viral, y algunas de ellas se han asociado a lesiones más avanzadas o tipos histológicos más agresivos [20-22]. Para el VPH16 se han definido cinco linajes filogenéticos, los cuales se distribuyen de manera distinta entre los cinco continentes (Figura 5). Estas se clasifican en: variantes Asiático-Americanas (AA), se encuentran principalmente en Centro y Sudamérica y en España, las variantes Africana-1 (Af-1) y Africana-2 (Af-2) en África, las variantes Asiáticas (As) en el Sudeste de Asia y las variantes Europeas (Eu) en todas las regiones excepto en África [23,24].

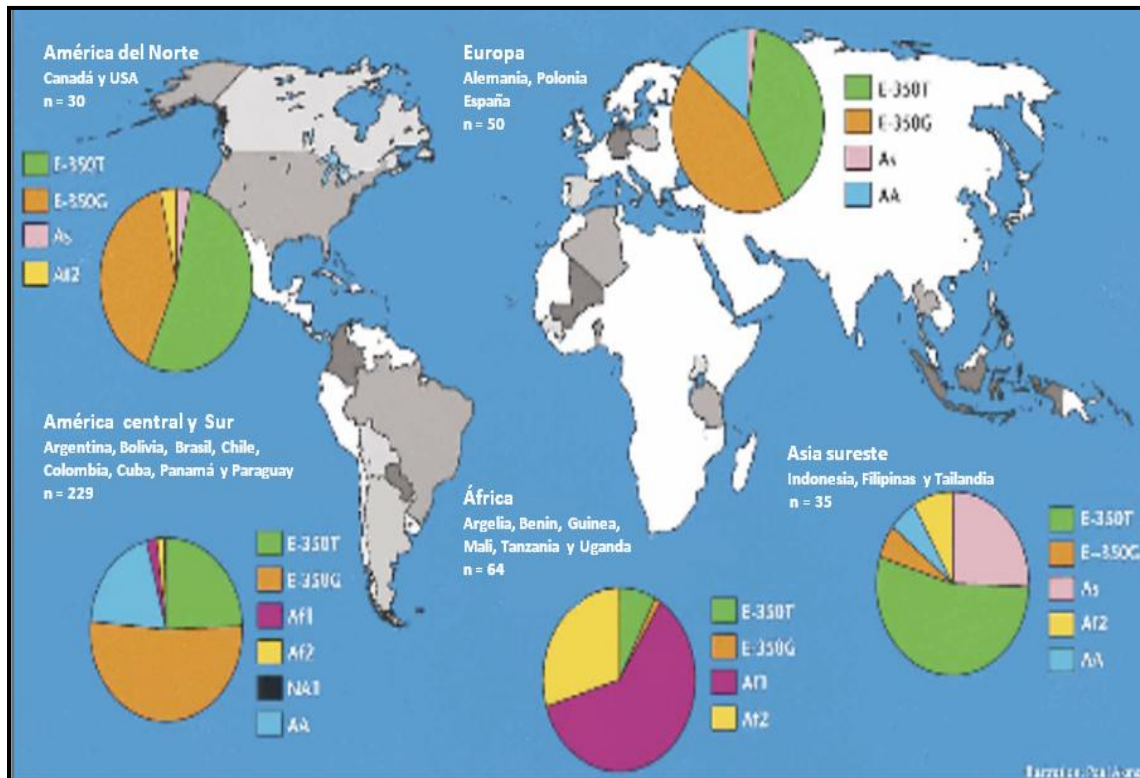


Figura 5. Distribución global de las clases y subclases de VPH16 por continente (tomado de Yamada y cols. [24]).

5. Biología del VPH.

Los VPH infectan epitelios estratificados queratinizados con una alta especificidad y están asociados con la aparición y persistencia de neoplasias benignas y malignas. Todos los VPH presentan un genoma de DNA circular de doble cadena de aproximadamente 8000 pares de bases, y de una cápside icosaédrica compuesta de 72 capsómeros, constituidos de 2 proteínas estructurales, y no poseen envoltura. El genoma viral de los VPH es bastante conservado, todos presentan seis genes de expresión temprana (E) y dos genes de expresión tardía (L), así como una región reguladora no codificadora (Figura 6) [12].

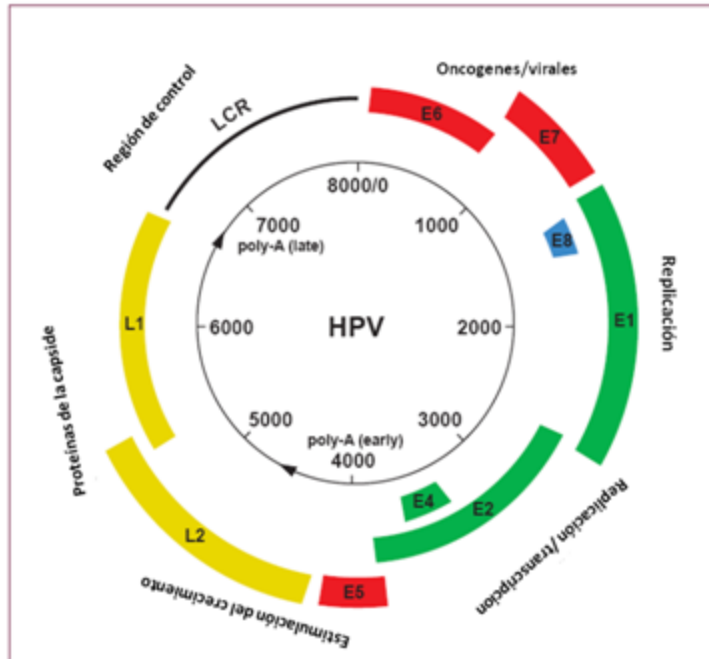


Figura 6. Esquema de organización del genoma de VPH (tomado y modificado de Prendiville y col. [25]).

Todos sus genes están codificados en una sola cadena, aunque en diferentes marcos de lectura y usan un procesamiento diferencial de corte y empalme para la expresión individual de cada uno de sus genes (ver Tabla I). La capacidad oncogénica del VPH reside en tres productos virales: las proteínas tempranas E6, E7 y E5 cuya expresión depende de un gran número de factores celulares, y de la presencia de la proteína viral reguladora E2. Asociada a E2, la proteína E1 cumple funciones importantes en la replicación del DNA viral, mientras que los genes tardíos L1 y L2, codifican para las proteínas de la cápside viral [25].

La proteína temprana E4 se expresa solamente en células diferenciadas, su marco de lectura está localizado dentro de la región temprana; sin embargo al encontrarse bajo la acción de un promotor específico de la diferenciación se expresa como un gen tardío, formando el transcrito llamado E1^{E4} [25].

Tabla I. Tamaño y función de las proteínas del papiloma virus.

Proteína viral/ elemento gen	Peso molecular/ tamaño	Función
Elementos no codificantes		
Región larga de control	500-1000 pb	Origen de replicación y regulación de la expresión de genes de VPH
Proteínas tempranas		
E1	68-85 kD	Función de helicasa, esencial para la replicación viral y el control de la transcripción de genes, similar en algunos tipos.
E2	48 kD	Factor de transcripción viral; esencial para la replicación viral y el control de la transcripción de genes; segregación del genoma y encapsidación.
E3	No conocida	Función no conocida, solo presente en pocos tipos de VPH
E1 [^] E4	10-44 kD	Se une a proteínas del citoesqueleto
E5	14 kD	Interacción con los receptores de EGF/PDGF
E6	16-18 kD	Interacción con diversas proteínas del ciclo celular; degradación de p53 y activación de telomerasa
E7	~10 kD	Interacción con diversas proteínas celulares; interacción con pRB y transactivación de promotores dependientes de E2F
E8-E2C	20 kD	Larga distancia de transcripción y replicación de proteínas represoras
Proteínas tardías.		
L1	57 kD	Proteína principal de la cápside
L2	43-53 kD	Proteína menor de la cápside

La regulación genética de los VPH reside en una porción no codificadora del genoma viral conocida como región larga de control LCR. La LCR es un segmento cuyo tamaño oscila entre los 800 y 1000 pb dependiendo del tipo de VPH. Una característica interesante de la LCR es que su secuencia nucleotídica es extremadamente variable entre los diferentes tipos virales, por lo que se ha utilizado para distinguir entre los diferentes tipos de VPH. La proteína viral E2 juega un papel central en la regulación genética de los VPH, reprimiendo o activando la transcripción de los genes [25].

La actividad represora de E2 tiene interesantes implicaciones clínicas. En lesiones benignas o pre-malignas el DNA genómico del VPH se encuentra de forma circular (episomal, no integrada) la expresión de E6 y E7 está regulada por E2. En cambio en lesiones tumorales se ha demostrado la presencia del DNA viral integrado al genoma celular, usualmente con el gen E2 interrumpido. Por lo tanto, los genes E6 y E7 de VPH no son regulados negativamente por E2 [12].

El gen E6 es uno de los genes virales tempranos expresados durante la infección por VPH y juega un papel importante en el ciclo de vida viral así también en la inmortalización y transformación celular. La participación de E6 en la transformación, esta mediada por la interacción con una gran variedad de proteínas celulares. La más estudiada es la proteína supresora de tumor p53, que es marcada por E6 para su degradación a través de E6-AP una proteína ligasa de la vía celular de ubiquitina.

La oncoproteína E7 interactúa con la proteína supresora de tumores p105 Rb e impide que se asocie con el factor transcripcional celular E2F, responsable de la activación de múltiples promotores de oncogenes celulares. De esta forma, la función de las oncoproteínas E6 y E7 en el ciclo de vida viral está relacionada con la inhibición de algunos mecanismos de la regulación celular [26].

6. El ciclo de vida de los VPH en los epitelios en diferenciación.

Los virus del papiloma son especie-específicos e inducen proliferaciones epiteliales o fibroepiteliales benignas de la piel y mucosas en humanos y varias especies animales. Estos virus tienen un tropismo específico y absoluto por las células epiteliales escamosas y su ciclo reproductivo completo sólo ocurre en estas células. Se cree que la infección por el VPH empieza en las células basales, que son mitóticamente activas. Después de la infección, el virus puede permanecer latente, replicarse y producir partículas infecciosas o integrarse al genoma celular.

La infección productiva se divide en varias etapas que dependen del estado de diferenciación de las células epiteliales en cuestión. El ciclo completo, que incluye la síntesis de DNA viral, la producción de las proteínas de la cápside viral y el ensamblaje de los viriones, ocurre selectivamente en queratinocitos terminalmente diferenciados. Las células de la capa basal que proliferan, a pesar de que contiene DNA del VPH, parecen ser poco activas en la expresión de algunas proteínas virales (Figura 7). Aparentemente existen factores celulares que regulan negativamente la transcripción viral de estas células.

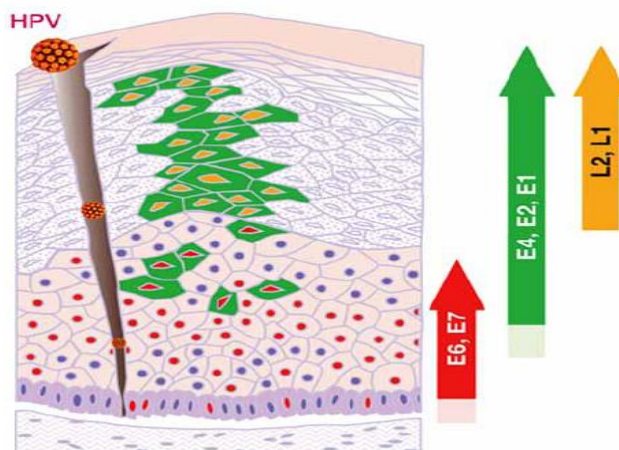


Figura 7. Secuencia de expresión de genes virales a través de la diferenciación de queratinocitos en el epitelio cervical estratificado. Inicialmente los virus de alto riesgo infectan los queratinocitos de la capa basal (tomado de Schiffman [12]).

Esta regulación se libera, cuando las células infectadas migran hacia arriba del epitelio en la capa granulosa, donde sufren un programa de diferenciación hasta que ya no pueden dividirse. En estas células empieza la transcripción activa de secuencias virales tempranas y tardías, se sintetizan proteínas virales y las partículas virales se ensamblan en algunas de las células superficiales [25,13].

7. Genes celulares alterados o inducidos por los oncogenes virales del VPH.

Los oncogenes virales E6 y E7 se expresan en bajos niveles en las células basales, y se sobre expresan conforme las células entran en la etapa de diferenciación terminal o progresan a NIC-AG o cáncer invasor. La expresión sostenida de los genes E6 y E7 es necesaria para mantener el fenotipo maligno. Las proteínas E6 y E7 alteran la diferenciación celular, reactivan la síntesis de DNA, y estimulan la progresión del ciclo celular, E6 transactiva o reprime la expresión de múltiples genes y se une a numerosas proteínas celulares incluyendo E6AP, una proteína ligasa de la vía ubiquitina de la proteólisis. El complejo E6/E6AP se une a la proteína supresora de tumores p53 marcándola para ser degradada por los proteosomas. La proteína p53 induce la apoptosis y frena el ciclo celular en respuesta al estrés genotóxico y daño al DNA. El control del ciclo celular lo ejerce en parte a través de inducir la expresión de p21, un inhibidor del complejo cinasa dependiente de ciclinas G1/S-Cdk, necesario para la progresión del ciclo celular. Por otra parte, es un activador de la transcripción de múltiples genes celulares asociados a la respuesta al estrés celular y al daño del DNA (Figura 8). La pérdida de p53 conduce a la inestabilidad genética y una rápida progresión a malignidad.

La proteína E7 del VPH16 se une a la proteína de retinoblastoma (pRb) y a otros miembros de esa familia [13]. La unión de E7 y pRB, ocurre con la forma hipofosforilada de pRB provocando la liberación del factor de transcripción E2F, el cual estimula la expresión de múltiples genes involucrados en la progresión del ciclo celular y la síntesis de DNA. La proteína E7 también se une y altera la función de factores de transcripción AP-1, los cuales regulan la expresión genética

de muchos genes. E6 y E7 ejercen un efecto traslapado en el control del ciclo celular y en combinación, inmortalizan eficientemente queratinocitos humano [12].

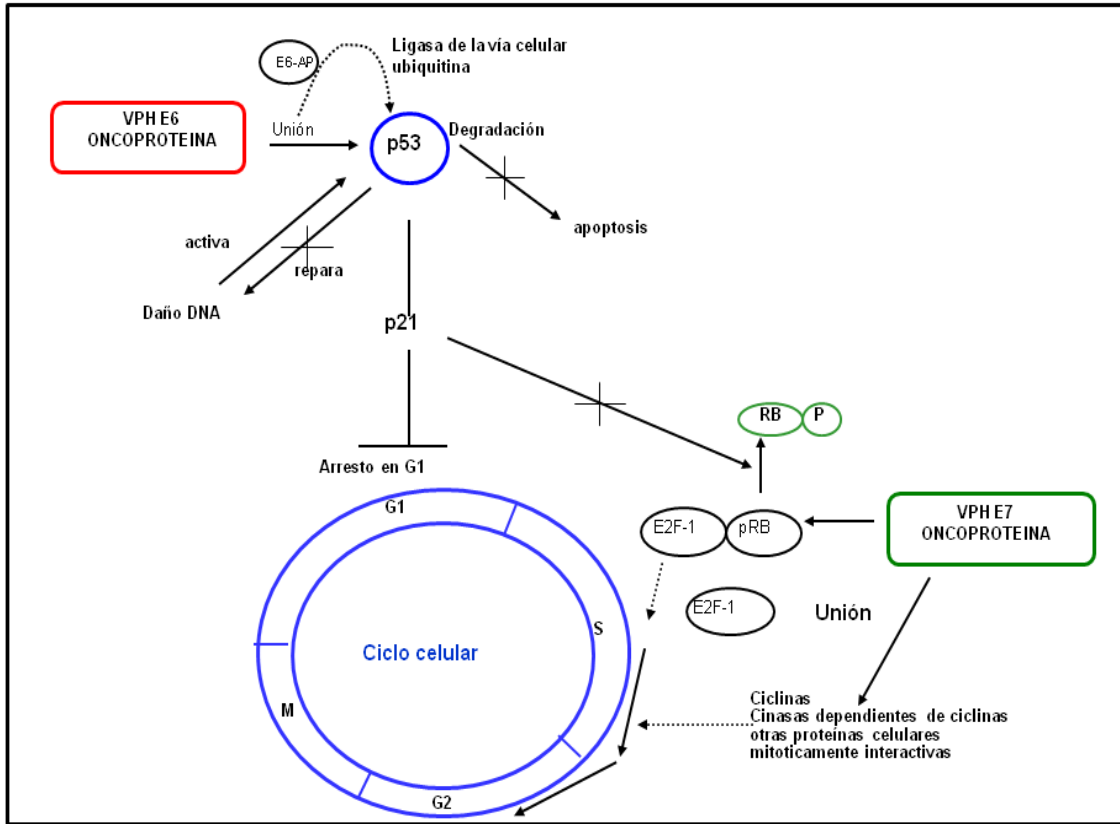


Figura 8. Mecanismo de transformación mediado por las proteínas virales E6 y E7 en cáncer cervical, (modificado Burd [13]).

Así, E6 y E7 interactúan con numerosos reguladores de la transcripción celular, tienen el potencial para modificar, directamente o a través de su efecto sobre p53 y Rb, la expresión de redes de genes en los queratinocitos. La búsqueda de patrones de expresión característicos en el cáncer cervical ha llevado al análisis de proteínas involucradas específicamente en las vías afectadas por los oncogenes virales. Por ejemplo, se ha demostrado la sobre expresión de algunos genes como p16, survivina, FasL, Fascina, Ciclina E, metaloproteinasas entre otros, y la regulación negativa de algunas proteínas de citoesqueleto y adhesión [27,28].

Recientemente, el surgimiento de los microarreglos de DNA ha permitido investigar patrones globales de expresión de genes en cánceres humanos, específicamente en cáncer cérvico-uterino [29-37]. Se han utilizado distintos abordajes, en algunos casos se han analizado queratinocitos transfectados con los oncogenes virales E6/E7 con el objetivo de investigar sus efectos directos sobre la expresión génica. También se han utilizado líneas celulares positivas para VPH y recientemente se han reportado un mayor número de estudios que analizan muestras tumorales [37].

En su conjunto, los resultados obtenidos en distintos trabajos con microarreglos de DNA han mostrado la alteración de la expresión de genes involucrados en la regulación del ciclo celular y apoptosis (regulación negativa de las vías dependientes de p53 y Rb), proteínas de citoesqueleto, proteínas de respuesta al daño al DNA y proteínas inmunomoduladoras. Se ha observado, por ejemplo, que se inhibe la expresión de TGF- β_2 y de múltiples genes regulados por esta interleucina, así como también los genes de respuesta a interferón como: interferón- α , interferón- β ; y por otra parte se induce la expresión de: NF- $\kappa\beta$ y AP-1 y de múltiples genes inducidos por que participan en la señalización de TNF- α o NF- $\kappa\beta$ y AP-1 [26].

Muchos de estos genes están asociados a la resistencia a la infección y la función inmune del huésped. Por ejemplo, los interferones estimulan la expresión de las moléculas HLA I y TAP1 y TAP2, proteínas involucradas en el procesamiento y presentación antigénica. De hecho, la expresión de estas proteínas se encuentra disminuida o abolida en las NIC-AG y cánceres invasores del cérvix, lo que permite que el tumor pueda evadir la respuesta inmune celular [38]. Sin embargo aun cuando existe algo más de un centenar de artículos de microarreglos en Ca-Cu, no se ha reportado que se enfoquen a la búsqueda de marcadores de diagnóstico temprano.

8. Vacunas.

En la actualidad, existen vacunas en el mercado que previenen la infección por VPH oncogénicos tipo 16 y 18, que se asocian con 65–70% de Ca-Cu en todo el mundo [39]. Estas vacunas tienen muy alta eficacia para la prevención de la infección y el desarrollo de neoplasias intraepiteliales cervicales de alto grado (NIC2/NIC3) [40,41]. Sin embargo, las mujeres aún cuando estén vacunadas, deben asistir a los programas para la detección temprana de Ca-Cu, ya que estas vacunas sólo protegen contra ciertos tipos de virus y todavía no se sabe cuánto tiempo durará la respuesta inmune de protección contra los restos del VPH-AR restantes [42,43]. En muchos países las vacunas preventivas para HPV 16 y 18 se han incorporado en el programa nacional de vacunación, para niñas de 9 a 12 años de edad [44,45]. Sin embargo, debido a que la incidencia de pico de Ca-Cu se produce en las mujeres 45–50 años de edad, el efecto de estos programas de vacunación preventiva en la reducción de la prevalencia de Ca-Cu no se conocerá aun con claridad durante los siguientes 30 años. Por lo tanto, es necesario mejorar los procedimientos de detección oportuna del Ca-Cu y el tratamiento.

9. Marcadores tumorales.

De acuerdo con los Institutos Nacionales de Salud de Estados Unidos (NIH); un biomarcador tiene la característica de ser medido objetivamente como un indicador de un proceso biológico normal, proceso patológico, o respuesta farmacológica a una intervención terapéutica. “Una molécula biológica que se encuentra en la sangre, en otros fluidos corporales o tejidos, es una señal de un proceso normal o anormal, de una condición o de enfermedad. Un biomarcador puede servir para ver cómo el cuerpo responde a un tratamiento en una enfermedad o condición. Los biomarcadores también se llaman marcador molecular y firma molecular” [111].

El análisis de los perfiles de expresión, está usándose fuertemente para caracterizar e identificar nuevos y mejores marcadores diagnóstico marcadores

tumorales (marcadores específicos de tumor o asociados a tumor). Estos perfiles son generados por el análisis de RNA mensajero (mRNA); patrones de metilación del DNA, con el uso de sondas específicas que son utilizadas para detectar su complementario y perfiles de expresión de proteínas en tumores.

Existen claros ejemplos de marcadores con aplicación clínica: ya sea de predicción, detección, diagnóstico o pronóstico. Sin embargo, pocos biomarcadores son altamente sensibles y específicos al mismo tiempo, para la detección en cáncer [111]. La sobreexpresión del gen *Her2/Neu*, para cáncer de mama es un ejemplo de un biomarcador predictivo; el Antígeno Prostático Específico (PSA): para Cáncer de próstata; CA 125, en Cáncer de ovario; Gonadotropina Coriónica Humana y Alfafetoproteína, para el Cáncer de testículo; Tiroglobulina y Calcitonina, en Cáncer de Tiroides; *EGFR*, en cáncer de pulmón por mencionar algunos ejemplos [46,111].

En el cáncer cervical se han descrito diversos biomarcadores, que pueden ser útiles como método de detección, sin embargo, estos no han sido completamente validados con estudios con la alta calidad que se requiere, evaluando la especificidad y sensibilidad de las pruebas. El conocimiento acerca como el VPH es el agente causal del cáncer cervical, que ha sido estudiado durante décadas y esto ha abierto la posibilidad de improvisar estrategias de prevención y prácticas de detección, que mejoren las existentes. Algunos de los genes propuestos como biomarcadores funcionales incluyen marcadores del ciclo celular tal como CDKN2A, Ki-67, PCNA, p53, proteína de retinoblastoma (pRB), p21, p27, MCM5, CDC6, ciclina A, E y D, marcadores de diferenciación escamosas citoqueratina (CK) tal como CK14 y CK13 y otras moléculas como involucrina, telomerasa (TERT), survivina, VEGF, FHIT y otras más [1].

10. Pruebas de diagnóstico, pronóstico y de seguimiento.

Durante muchos años, la prueba de Papanicolaou (Pap) ha sido el procedimiento de pesquisa más importante para la detección temprana de Ca-Cu, y su aplicación

masiva en los países desarrollados ha disminuido la incidencia de la Ca-Cu por más de 50% en los últimos 40 años [47]. Las mujeres con pruebas Pap anormales son remitidas a la colposcopia, para que se confirme, descarte o aclare el diagnóstico junto con el estudio histopatológico. Sin embargo, la sensibilidad promedio de citología para la detección de lesiones NIC es 50 a 60%; Aunque la especificidad es muy elevada, aproximadamente un 90% [48].

Como es ampliamente conocido que el virus VPH es indispensable para el desarrollo de Ca-Cu, varios procedimientos para detectar el genoma del VPH se ha incorporado la detección del Ca-Cu. En comparación con la citología convencional, pruebas de DNA de VPH tiene mayor sensibilidad pero baja especificidad para la detección de lesiones CIN2 o superior (CIN2+). La alta sensibilidad y alto valor predictivo negativo (VPN) de pruebas de DNA de VPH para la detección de lesiones CIN2+ sugiere que podría ser utilizado para ampliar los intervalos de seguimiento [49]. En la Tabla II se resumen algunos de los trabajos publicados en donde se compara la sensibilidad y especificidad para la prueba de DNA de VPH, en y se muestra incluso el valor predictivo de esta prueba.

Tabla II. Revisión de artículos que evalúan sensibilidad y especificidad en muestras de cáncer cervical

Referencia	Año	Número total de muestras	casos/controles	VPH-AR	sensibilidad	especificidad	VPP	VPN
Kulasingam y cols. ^[50]	2002	4358	1074/3318	1047	68.1	90		98.4
Coste y cols. ^[51]	2003	2585	389/1405		96	82		
Petry y cols. ^[52]	2003	7908	521/8083	521	97.8	95.3	10.9	100
Brigas y cols. ^[53]	2005	13842	497/1036	437	97	92.4		
Mayrand y cols. ^[54]	2007	9977	5059/5095	399	97.4	94.3	7	100
Cardénas-Turanzas y cols. ^[55]	2008	1850	443/910	240	69	93	17	99
Del Mistro y cols. ^[56]	2010	749	147/582	123	93.1	78.6	14.9	
Lazcano-Ponce y cols. ^[57]	2010	50,159	1106/49053	4,418	93.3	89.2		

VPH-AR= virus de papiloma humano alto riesgo. VPP=Valor predictivo positivo, VPN=Valor predictivo

Sin embargo, la baja especificidad de las pruebas de DNA de VPH aumentaría el número de exámenes de seguimiento y derivaciones de la colposcopia, lo que así mismo acrecentaría el costo de la investigación [58]. Además en países como México el acceso a este diagnóstico aun es difícil por los pocos equipos disponibles y personal calificado para el desarrollo de esta técnica. Por lo tanto, la necesidad de desarrollar nuevos métodos para la detección temprana de Ca-Cu con alta sensibilidad y especificidad es evidente.

Múltiples marcadores de tumor asociados con CIN2 han sido identificados, especialmente CDKN2A, TOP2A y MCM2. Sin embargo, estos marcadores no se han propuesto para la detección, están más dirigidos para diagnóstico, pronóstico o tratamiento [1,27]. El tratamiento actual para el cáncer cervical invasivo es la cirugía, la quimioterapia, la radioterapia o una combinación de estas terapias, dependiendo de la fase clínica de la enfermedad.

El éxito de las terapias convencionales y la supervivencia de las pacientes disminuyen conforme la enfermedad progresa a etapas más avanzadas [59]. De hecho, el porcentaje de mujeres que sobreviven 5 años disminuye de 93% hasta un 15%, entre el estadio IA al estadio IVB (www.cancer.org). A diferencia de otros tipos de cáncer, en los cuales ya varios medicamentos moleculares específicos se han desarrollado [60]; hasta el momento para el cáncer cervical no hay terapias moleculares específicas dirigidas. La mayoría de los medicamentos contra blancos específicos en cáncer se dirige hacia las proteínas mutantes, especialmente proteínas cinasas [61]. Sin embargo, algunos fármacos que utilizan como blanco normales que son sobre-expresadas, como HER2/neu en el cáncer de mama [62,63]. El primer paso para el desarrollo de un medicamento molecular es la identificación de blancos moleculares universales que están presentes en los pacientes con Ca-Cu y ausente en mujeres sanas.

HIPÓTESIS.

Existe un patrón de genes asociados al Ca-Cu, causado por la infección con VPH, estos genes que aun no han sido descritos, pueden participar en procesos como progresión o invasión en Ca-Cu. Estos genes estarán alterados en la mayoría de los tumores y podrían servir como blancos para detección temprana o terapia en cáncer.

PROBLEMA DE INVESTIGACIÓN.

El Ca-Cu es un problema de salud pública por ser uno de los tumores más frecuentes en mujeres en México. Existe una asociación entre el desarrollo de este tipo de cáncer y la infección con VPH en las mujeres, por lo cual esperamos que exista un grupo de genes expresados solo en muestras tumorales positivas a VPH y no en tejido sano. Estos genes pueden ser prometedores como blancos para detección temprana por nuevos métodos sensibles y específicos, así como para terapia. En la actualidad, existe la necesidad de desarrollar nuevos métodos de detección temprana para el Ca-Cu con alta sensibilidad y especificidad al mismo tiempo.

Existen preguntas que requieren ser contestadas:

¿Cuáles son los genes que participan en la invasión en el Ca-Cu?

¿Existen alteraciones genéticas comunes en todo el Ca-Cu que puedan utilizarse como blancos para diagnóstico y terapia?

¿Existen alteraciones genéticas comunes en las lesiones pre-invasoras de Ca-Cu que puedan utilizarse como blancos para detección temprana y terapia?

OBJETIVO GENERAL:

Identificar y caracterizar blancos celulares presentes en la mayoría de muestras Ca-Cu positivas para VPH16 y que no se expresen en el epitelio normal del cuello uterino, con una diferencia significativa entre los dos grupos para ser considerados como posibles marcadores, con una sensibilidad y especificidad cercana al 100%, para la detección temprana en Ca-Cu.

Objetivos secundarios:

- Identificar genes asociados al proceso de invasión en el cáncer de cuello uterino, en muestras positivas para VPH16, mediante el uso de microarreglos de expresión.
- Identificar genes universales alterados en el cáncer de cuello uterino que sean útiles como blancos diagnósticos o terapéuticos.
- Calcular la sensibilidad y especificidad de los marcadores para cáncer de cuello uterino.
- Validar marcadores útiles para pesquisa en las lesiones de bajo y alto grado.

MATERIAL Y MÉTODOS.

1. Muestras.

En este trabajo se incluyeron un total de 144 muestras de pacientes, las cuales pertenecen a cuatro grupos de estudio: **I)** 69 biopsias de pacientes con Ca-Cu invasor, diagnosticadas en el Servicio de Oncología en el HGM; **II)** 29 muestras de pacientes con lesiones de bajo grado (NIC1); **III)** 21 muestras de pacientes de alto grado NIC (NIC2 ó NIC3) y **IV)** 25 epitelios de mujeres con cuello uterino normal, obtenidos en el Servicio de Ginecología y Obstetricia en el HGM.

Las muestras de Ca-Cu invasor pertenecen a una colección total de 462 biopsias de pacientes, que forman parte de un banco de tejidos de la unidad de medicina Genómica y se colectaron de forma secuencial durante un período desde noviembre de 2003 hasta julio de 2007, y que además representan aproximadamente el 80% de las pacientes de recién diagnóstico para Ca-Cu. Para cada paciente participante en el estudio, una carta de consentimiento informado fue obtenida y el estudio fue aprobado por el comité de ética del HGM.

Así mismo, los criterios de selección de las 69 muestras de Ca-Cu se basaron en la disponibilidad de una biopsia de tumor fresco para la extracción de RNA, con más del 70% de células tumorales en el análisis morfológico (ver abajo), en su mayoría FIGO etapa II y el tipo viral. Este grupo de Ca-Cu, incluyó 47 muestras positivas para el VPH16 y 22 muestras positivas para otro tipo de virus incluyendo VPH 18, 31, 33, 45, 51, 58 y 59. Del total de las 69 muestras de Ca-Cu, 54 fueron carcinomas de células escamosas, 14 adenocarcinomas y 1 muestra de un carcinoma adenoescamoso. La edad promedio de las pacientes con cáncer fue de 48 años (rango 23-78, Tabla III). Todas las pacientes recibieron evaluaciones clínicas completas, la clasificación de los tumores de las pacientes se realizó según los últimos protocolos internacionales aprobados para el cáncer ginecológico [64].

Tabla III. Resumen de los datos clínicos, tipo viral, edad, arreglo utilizado para la expresión génica.

Muestra ^a	Histología	Etapas tumor	VPH tipo	Edad (años)	Técnica
R015	CCE	IIB	16	42	MA1/qRT-PCR
R035	CCE	IB2	16	48	MA1-MA2/qRT-PCR
R052	ADC	IIB	16	54	MA1/qRT-PCR/IH
R057	ADC	IB1	16	32	MA1/qRT-PCR/IH
R070	CCE	IIB	16	74	qRT-PCR
R072	CCE	IB1	16	61	MA1-MA2/qRT-PCR
R081	ADC	IB1	16	41	qRT-PCR
R093	ADC	IB1	16	57	MA1/qRT-PCR/IH
R094	CCE	IB1	16	45	MA1/qRT-PCR
R124	CCE	IB1	16	42	MA1-MA2
R170	CCE	IIB	16	67	MA1/qRT-PCR
R183	CCE	IB1	16	64	MA1-MA2/qRT-PCR
R189	ADC	IB1	16	47	MA1-MA2
R221	CCE	IB2	16	41	MA1-MA2/qRT-PCR
R230	CCE	IB1	16	62	MA1-MA2
R232	CCE	IB2	16	45	MA1/qRT-PCR/IH
R240	CCE	IIIB	16	31	MA1-MA2/qRT-PCR/IH
R245	CCE	IIA	16	64	MA1/qRT-PCR/IH
R251	CCE	IB1	18	61	qRT-PCR
R254	CCE	IB1	33	23	qRT-PCR/IH
R255	CCE	IIA	16	45	MA1/qRT-PCR/IH
R256	CCE	IIA	45	78	qRT-PCR
R258	CCE	IB1	16	36	MA1/qRT-PCR/IH
R261	CCE	IB1	18	43	qRT-PCR
R265	CCE	IB1	16	46	MA1-MA2/qRT-PCR/IH
R266	CCE	IIA	18	51	qRT-PCR/IH
R268	CCE	IIB	16	34	qRT-PCR
R269	CCE	IB1	16	66	qRT-PCR
R278	CCE	IB1	45	54	qRT-PCR/IH
R284	ADC	IB2	16	33	MA1/qRT-PCR
R296	CCE	IB2	31	48	qRT-PCR/IH
R308	ADC	IB1	16	45	MA1/qRT-PCR/IH
R312	ADC	IB2	16	34	MA1/qRT-PCR
R315	CCE	IIIB	16	41	MA1-MA2/qRT-PCR
R316	CCE	IB1	51	67	qRT-PCR
R324	CCE	IB2	16	28	MA1-MA2/qRT-PCR/IH
R330	CCE	IB1	16	72	MA1-MA2/qRT-PCR/IH
R333	CCE	IIB	16	56	MA1/qRT-PCR
R335	ADC	IB1	16	37	MA1/qRT-PCR
R336	CCE	IB2	16	36	MA1-MA2/qRT-PCR
R339	CCE	IB2	16	31	MA1-MA2/qRT-PCR
R343	CCE	IB1	31	49	qRT-PCR/IH
R359	ADC	IB2	16	34	MA1/qRT-PCR
R367	CCE	IB2	18	45	qRT-PCR
R369	CCE	IB1	16	50	MA1-MA2/qRT-PCR
R378	CCE	IB2	16	42	MA1-MA2/qRT-PCR
R379	CCE	IB2	18	59	qRT-PCR/IH
R381	CCE	IB2	31	46	qRT-PCR
R385	CCE	IB2	45	47	qRT-PCR
R393	CCE	IB1	31	42	qRT-PCR
R396	ADC	IB2	16	53	MA1/qRT-PCR
R400	CCE	IB2	58	32	qRT-PCR
R403	CCE	IIB	16	34	MA1/qRT-PCR
R405	CCE	IB1	33	51	qRT-PCR/IH
R409	CCE	IB2	16	68	MA1-MA2/qRT-PCR
R411	CCE	IB1	16	34	MA1-MA2/qRT-PCR
R412	CCE	IB2	16	33	MA1-MA2/qRT-PCR
R413	CCE	IB2	31	54	qRT-PCR
R415	ADE	IIB	16	55	MA1/qRT-PCR
R425	CCE	IB1	59	64	qRT-PCR
R434	ADC	IB1	16	34	MA1/qRT-PCR
R441	ADC	IIB	16	24	MA1/qRT-PCR/IH
R443	CCE	IB1	16	34	MA1/qRT-PCR
R446	ADC	IB1	16	43	MA1/qRT-PCR
R451	CCE	IB2	45	36	qRT-PCR/IH
R469	CCE	IVB	51	60	qRT-PCR
R482	CCE	IB2	16	61	MA1/qRT-PCR/IH
R487	CCE	IB1	45	73	qRT-PCR/IH
R488	CCE	IIA	58	55	qRT-PCR
C01X	ECN		Negativo	33	MA1/qRT-PCR
C04X	ECN		Negativo	45	qRT-PCR
C05X	ECN		Negativo	36	MA1-MA2/qRT-PCR
C07X	ECN		Negativo	37	MA1-MA2/qRT-PCR
C10X	ECN		Negativo	45	MA1-MA2
C11X	ECN		Negativo	39	qRT-PCR
C12X	ECN		Negativo	30	qRT-PCR
C13X	ECN		Negativo	29	MA1-MA2/qRT-PCR

C16X	ECN	Negativo	33	qRT-PCR
C19X	ECN	Negativo	30	MA1-MA2/qRT-PCR
C21X	ECN	Negativo	34	qRT-PCR
C29X	ECN	Negativo	30	qRT-PCR
C30X	ECN	Negativo	29	qRT-PCR
C01E	ECN	Negativo	33	MA1/qRT-PCR
C03E	ECN	Negativo	52	MA1/qRT-PCR
C06E	ECN	Negativo	42	MA1/qRT-PCR
C10E	ECN	Negativo	45	MA1/qRT-PCR
C11E	ECN	Negativo	39	qRT-PCR
C12E	ECN	Negativo	30	qRT-PCR
C13E	ECN	Negativo	29	MA1/qRT-PCR
C16E	ECN	Negativo	33	qRT-PCR
C17E	ECN	Negativo	34	qRT-PCR
C19E	ECN	Negativo	30	MA1/qRT-PCR
C20E	ECN	Negativo	37	qRT-PCR
C28E	ECN	Negativo	31	qRT-PCR
C32E	ECN	Negativo	31	qRT-PCR
N008	NIC1	ND	24	qRT-PCR
N013	NIC1	ND	19	qRT-PCR
N014	NIC1	ND	41	qRT-PCR
N017	NIC1	ND	45	qRT-PCR
N025	NIC1	ND	26	qRT-PCR
N027	NIC1	ND	40	qRT-PCR
N029	NIC1	ND	46	qRT-PCR
N030	NIC1	ND	36	qRT-PCR
N031	NIC1	ND	49	qRT-PCR
N038	NIC1	ND	22	qRT-PCR
N040	NIC1	ND	56	qRT-PCR
N044	NIC1	ND	38	qRT-PCR
N046	NIC1	ND	44	qRT-PCR
N051	NIC1	ND	19	qRT-PCR
N053	NIC1	ND	49	qRT-PCR
N066	NIC1	ND	56	qRT-PCR
N067	NIC1	ND	32	qRT-PCR
N071	NIC1	ND	54	qRT-PCR
N073	NIC1	ND	20	qRT-PCR
N077	NIC1	ND	19	qRT-PCR
N080	NIC1	ND	34	qRT-PCR
N081	NIC1	ND	27	qRT-PCR
N082	NIC1	ND	25	qRT-PCR
N097	NIC1	ND	27	qRT-PCR
N098	NIC1	ND	31	qRT-PCR
N101	NIC1	ND	38	qRT-PCR
N102	NIC1	ND	45	qRT-PCR
N103	NIC1	ND	42	qRT-PCR
N104	NIC1	ND	39	qRT-PCR
N002	NIC2	ND	42	qRT-PCR
N010	NIC2	ND	42	qRT-PCR
N018	NIC2	ND	38	qRT-PCR
N085	NIC2	ND	20	qRT-PCR
N089	NIC2	ND	56	qRT-PCR
N096	NIC2	ND	51	qRT-PCR
N006	NIC3	ND	48	qRT-PCR
N007	NIC3	ND	79	qRT-PCR
N011	NIC3	ND	37	qRT-PCR
N016	NIC3	ND	31	qRT-PCR
N023	NIC3	ND	42	qRT-PCR
N065	NIC3	ND	40	qRT-PCR
N084	NIC3	ND	25	qRT-PCR
N087	NIC3	ND	37	qRT-PCR
N088	NIC3	ND	20	qRT-PCR
N090	NIC3	ND	21	qRT-PCR
N091	NIC3	ND	32	qRT-PCR
N105	NIC3	ND	50	qRT-PCR
N106	NIC3	ND	33	qRT-PCR
N107	NIC3	ND	46	qRT-PCR
N108	NIC3	ND	50	qRT-PCR

CCE, carcinoma de células escamosas; ACC, adenocarcinoma; ASCC, Adenoescamoso; ECN, epitelio cervical normal; NIC, Neoplasia intraepitelial cervical; MA, microarreglos; MA1, HG-Focus; MA2, HG-ST1.0; qRT-PCR, PCR cuantitativa en tiempo real; IH, Inmunohistoquímica; ND, no determinado.

^a. El código de las muestras control C terminado con X es exocervix y el terminado E es endocervix.

Se tomaron una o dos biopsias de tumores seleccionadas por evaluación colposcópica, y de estas biopsias una de ellas se dividió en dos partes iguales,

una parte se fijó en formol tamponado para análisis morfológico y la otra parte, junto con la segunda biopsia, se congelaron completamente en hielo seco y se almacenaron a -80°C hasta su análisis. Todas las pacientes se refirieron para recibir tratamiento por cirugía, radio o quimioterapia, o la combinación de estos mismos de acuerdo con las guías de la Sociedad Americana de Cáncer. Para la obtención de las muestras de cérvix control, las pacientes fueron diagnosticadas en el Servicio de Ginecología del Hospital General de México, de estas pacientes que se sometieron a histerectomía debido a miomatosis, se tomaron las biopsias de epitelio normal, previo a la operación para cada una de ellas, se evaluó y se registró el cérvix normal aparentemente, por citología y colposcopia. En el quirófano, inmediatamente después de recibir el fragmento de cérvix, el epitelio endocervical o exocervical, este se disecó bajo un microscopio estereoscópico para evitar las células estromales. Los tejidos se congelaron en nitrógeno líquido y se almacenaron a -80°C hasta su uso.

2. Diseño experimental.

Para la detección y tipificación de los VPH, una muestra de exudado cervical de endocervix y exocervix se colectó con un citobrush tanto de pacientes como de controles (previo a la toma de biopsia), y las células fueron suspendidas en un vial con buffer de extracción y estas se almacenaron a -20°C hasta su análisis. El análisis de la expresión génica global (8,623 genes), se realizó con el RNA aislado de 43 biopsias de tumores frescos y de 12 muestras de epitelio cervical normal, empleando el microarreglo de HG-Focus. Adicionalmente la expresión global de los genes se validó en 24 muestras, incluyendo 19 Ca-Cu, y 5 controles de epitelio cervical normal. Con un segundo microarreglo de alto rendimiento HG-ST1.0.

Los 23 genes que mostraron los mayores valores de desregulación se validaron por PCR en tiempo real (qRT-PCR) en 44 muestras de Ca-Cu positivas para VPH16 y 25 muestras controles. Los 7 de estos genes se evaluaron además en 22 muestras positivas para otros tipos de VPH. Los 6 genes con la mayor expresión diferencial (*CCNB2*, *CDC20*, *PRC1*, *SYCP2*, *NUSAP1* y *CDKN3*) se exploraron posteriormente en lesiones pre-invasoras, 29

de bajo grado (NIC1) y 21 de alto grado (NIC2/3), para investigar si estos genes podían distinguir entre Ca-Cu y NIC2/3 (NIC2+) de los NIC1 y controles. La validación por inmunohistoquímica (IH) se realizó para 10 de las proteínas seleccionadas, en 26 muestras de Ca-Cu y 10 muestras control. Finalmente la asociación de 6 marcadores con la sobrevida que se analizó en 42 pacientes de Ca-Cu positivas para VPH16, que fueron seguidos por al menos 42 meses.

3. Extracción de DNA y RNA.

El DNA se purificó de los exudados cervicales y de algunas biopsias mediante el kit de DNA de PureLink (Invitrogen, Grand Island NY) y almacenado a -20°C hasta su análisis. El RNA total fue obtenido de la mitad de una biopsia con el uso del reactivo Trizol (Invitrogen, Carlsbad, CA), siguiendo las recomendaciones de manufactura del proveedor. La calidad del RNA fue confirmada por el gel de la agarosa verificando que la presencia del RNA ribosomal estuviera intacto con las bandas 28s y 18s, además 28s el doble de intensa que 18s (Figura 9). La pureza de las muestras del RNA también fue evaluada por espectrofotometría donde la proporción OD 260/280 fue mayor de 1.8. La concentración fue expresada en ng/ μ L y fue ajustada a 1 μ g/ μ L en todas las muestras.

4. Detección y tipificación del VPH.

La detección de VPH se realizó mediante PCR con primers universales para el gen L1 del VPH MY09/MY11, GP5+/6+ y L1C1 como se ha descrito anteriormente [65-67]. El gen de beta globina (*HBB*), fue utilizado como control interno para evaluar la calidad del DNA. Los tipos de VPH fueron identificados mediante la secuenciación de las bandas amplificadas en muestras positivas utilizando un método de secuenciación del ciclo fluorescente (BigDye Terminator Applied Biosystems, Foster city, CA). El análisis de la secuencia se realizó mediante un analizador genético 3130xl ABI PRISM (Applied Biosystems). Cada secuencia de las muestras positivas de VPH se analizó con la herramienta de similaridad FASTA [68]. La identidad del porcentaje promedio de estas secuencias para tipos de VPH fue 98.7% (rango, 91–100%).

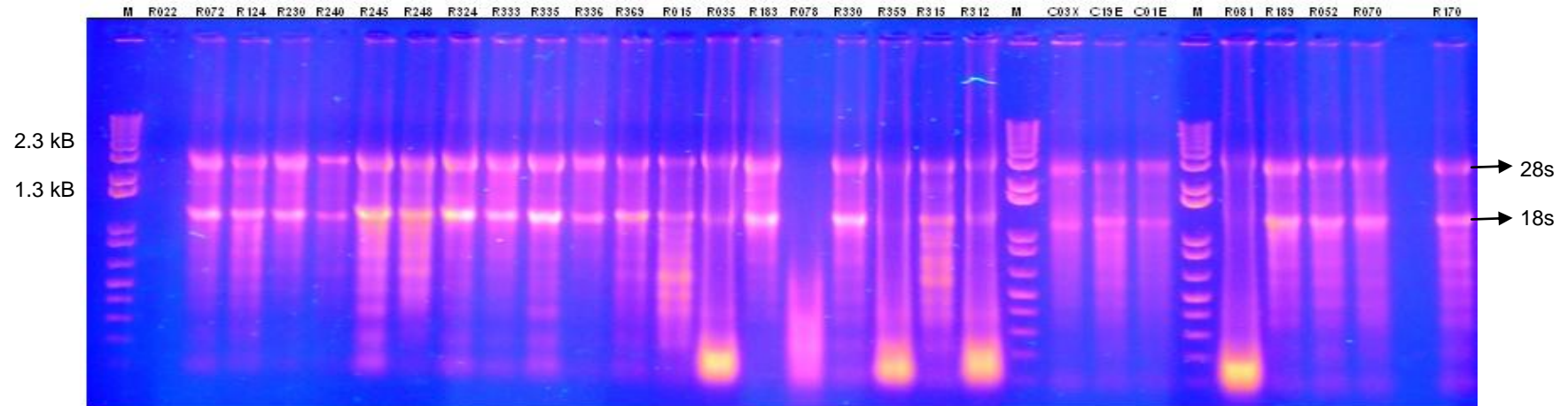


Figura 9. Electroforesis en gel de agarosa al 1% de RNA total proveniente de biopsias muestras de tumores y muestras controles, M indica marcador de peso molecular. La figura muestra como indicadores la banda 28s y 18s ribosomal. 28s debe de ser el doble de intenso que 18S. El mRNA aparece 0.5 a 9 kB. La figura muestra algunos ejemplos representativos de las muestras usadas en este estudio.

5.- Análisis del perfil de expresión génica y análisis de datos

Los perfiles de expresión génica de 43 muestras de Ca-Cu positivos para VPH16 y 12 controles sanos fueron examinados usando el microarreglo de oligonucleótidos (MA) de expresión Human Gene Focus (HG Focus, Affymetrix, Santa Clara, CA). Ese arreglo contiene ~8794 sondas que corresponden con 8638 genes humanos caracterizados en la base de datos Gene Reference. Para la síntesis de cRNA, se utilizó 10 µg de RNA total, siguiendo los protocolos del proveedor (Affymetrix, GeneChip Expression Assay manual). Brevemente: para la síntesis de cDNA de doble cadena (ds-cDNA) se utilizó 10 µg de RNA, la enzima Transcriptasa Reversa SuperScrib II (Invitrogen) y el primer oligo(dT) conteniendo un promotor de RNA para síntesis de la primera cadena. El marcaje del cDNA con biotina fue obtenido por transcripción *in vitro* después de la adición de la enzima polimerasa T7 y nucleótidos biotinilados (Enzo, Biochem). El cDNA marcado fue fragmentado e hibridado en los GeneChips HG-Focus, a 42° C por 16 h.

Posteriormente los MA se tiñeron con el complejo ficoeritrina (PE) conjugada con estreptavidina y se lavaron con la ayuda de la estación de fluidos 400. Finalmente los MA se escanearon y se analizaron evaluando los parámetros de control de calidad de la hibridación y de todo el experimento en su conjunto. Para evaluar la calidad de los experimentos inicialmente se utilizaron las recomendaciones del proveedor, de los siguientes parámetros: expresión aumentada de controles de poly-A exógenos (Lys, Phe, Thr, Dap), la presencia de oligo B2 usada para los alineamientos en rejilla, con un margen de ruido de fondo aceptable de 20 a 100, a través de todas las muestras, el porcentaje de los llamados presentes mayor al 50%, una relación 3'/5' de un gen constitutivo (GAPDH o β-actina) menor de 3, y la expresión incrementada de los controles de hibridación (BioB<BioC<BioD<cre). Finalmente sólo se analizaron los ensayos que cumplieron con todos los parámetros de calidad óptima. Durante la realización de estos experimentos, algunas muestras se analizaron por duplicado para evaluar la reproducibilidad de los experimentos, y como resultado fue superior al 99%.

Los valores de la intensidad de los ensayos de MA, fueron normalizados con el algoritmo RMA (Robust Multichip Average), utilizando el programa FlexArray [69]. Los valores de intensidad normalizada fueron referidos como unidades de intensidad (UI). Los genes expresados de forma diferente entre los tumores y controles fueron identificados mediante el algoritmo SAM (Significance Analysis of Microarrays, SAM versión 3.0, <http://www.stat.stanford.edu/~tibs/SAM>), utilizando los valores de corte: un cambio en número de veces (fold change (FC) de ≥ 1.5 , un valor de FDR (false discovery rate) de 1% y un FDR local de 10% [70]. El análisis de agrupamiento jerárquico y el análisis de componente principal fueron realizados con el programa dChip (versión 1.6, www.dCHIP.org) y el lenguaje R de la plataforma Java's respectivamente.

6. Validación de la expresión génica Global por con un segundo ensayo de micro arreglos de alto rendimiento (Microarray HG-ST1.0).

El perfil de expresión génica de 24 muestras exploradas con el HG-Focus microarray, incluyendo 19 Ca-Cu y 5 epitelios cervicales control, también fue examinado con el Gene Human ST1.0, microarreglo de oligonucleotidos (Affymetrix, Santa Clara, CA). Este arreglo contiene 33,297 sondas que corresponden a aproximadamente 20,741 genes de la base de datos de referencia de genes humanos según la UCSC Genome Browser Assembly Mar. 2006 NCBI 36/hg18, disponible en <http://genome.ucsc.edu/>. Para los ensayos un total de 300 ng de RNA total fue requerido, durante la síntesis del cDNA se realizó el marcaje en la síntesis de DNA, la hibridación, escaneo y análisis de imagen, se realizaron siguiendo los protocolos del proveedor (Affymetrix, GeneChip Expression Assay manual). Para evaluar la calidad de los experimentos, los siguientes parámetros se evalúan: expresión de los controles exógenos de poli-A, la presencia de oligo B2 solo para hacer los alineamientos las alineaciones de rejilla y los valores del área bajo la curva (AUC) tienen que estar por encima de 0,8., y al igual que con el otro microarreglo, sólo se analizaron los controles de calidad óptima. Los microarreglos se normalizaron utilizando el algoritmo RMA en la consola de expresión de Affymetrix. Los valores de intensidad normalizada fueron contemplados como unidades de intensidad (UI). Las intensidades normalizadas (\log_2 valores) de la 8,370 genes

que fueron examinados en ambos microarreglos (HG ST1 y HG focus) se compararon, y el nivel de correlación se evaluó con el coeficiente de correlación de Pearson.

7. Validación de la expresión génica global por RT-PCR en tiempo real cuantitativa (qRT-PCR).

La transcripción reversa de RNA total fue realizada usando el kit “High Capacity cDNA Archive Kit” (Applied Biosystem, CA) en un volumen final de 20 μ L. La mezcla de reacción incluía RNA total 2 μ g, buffer RT 10X 2 μ L, dNTP 100 mM 0.8 μ L, random primer RT 10X 2 μ L, Transcriptasa Reversa MultiScribe™ (5 U/ μ L), 1 μ L y inhibidor de RNasa (2 U/ μ L), 1 μ L. Las reacciones fueron incubadas por 10 min., a 25°C, por 120 min., a 37° C y 5 seg., 98°C y posteriormente el cDNA fue almacenado a -20° C hasta su uso.

La expresión se validó para los 23 genes con mayor significancia estadística en 44 Ca-Cu VPH16-positivo y 25 muestras de control sano epitelio cervical con qRT-PCR usando sondas TaqMan (Applied Biosystem, CA). Los genes incluidos son *CCNB2*, *CDC2*, *CDC20*, *CDKN2A*, *CDKN3*, *CKS2*, *MCM2*, *MKI67*, *NUSAP1*, *PCNA*, *PRC1*, *RFC4*, *RRM2*, *SMC4*, *SYCP2*, *TOP2A*, *TYMS*, *ZWINT*, *CFD*, *EDN3*, *NDN*, *SLC18A2* y *WISP2*. El gen *GAPDH* fue utilizado como control interno (Tabla IV), la utilización de este gen se ha descrito previamente [71]. Adicionalmente 7 genes fueron también explorados en muestras de 22 Ca-Cu positivas para otros tipos de VPH (*CCNB2*, *CDC20*, *CDKN3*, *PRC1*, *SYCP2*, *NUSAP1*, *TYMS*). Además en las lesiones precursoras 29 NIC bajo y 21 NIC de alto grado, también se exploró la expresión de los 6 primeros de ellos, junto con los genes *CDKN2A*, *PCNA* y *MKI67*.

Todos los experimentos fueron corridos por duplicado en un volumen final de reacción de 20 μ L, incluyendo cDNA molde 200 ng, mezcla de Universal PCR Master mix (Applied Biosystems) 10 μ L y sonda Taqman 20X, 1 μ L y agua libre de RNasa 7 μ L. Las condiciones de reacción en el termociclador fueron las siguientes: etapa inicial de activación a 50° C por 2 min., seguido de 10 min., a

95° C y 40 ciclos cada uno con una etapa de fusión por 15 seg., a 95° C y alineación/extensión por 1 min., a 60° C. El termociclador utilizado fue el Rotor-Gene de Corbett Research (Sydney, Australia). Los valores de la mediana de las desviaciones estándar de los C_t en duplicados oscilaron entre 0,09 y 0.24 (media = 0,16) entre los 23 genes, lo que sugiere que las variaciones entre los duplicados fueron muy pequeñas [72].

Tabla IV. Sondas TaqMan para ensayos de expresión de Applied Biosystems.

Símbolo del gen	No. de Assay ID	Tamaño pb
<i>CCNB2</i>	Hs00270424_m1	73
<i>CDC2</i>	Hs00364293_m1	92
<i>CDC20</i>	Hs00415851_g1	108
<i>CDKN2A</i>	Hs00233365_m1	117
<i>CDKN3</i>	Hs00193192_m1	83
<i>CFD</i>	Hs00157263_m1	72
<i>CKS2</i>	Hs00854958_g1	88
<i>EDN3</i>	Hs00171177_m1	78
<i>GAPDH</i>	Hs02758991_g1	93
<i>MCM2</i>	Hs00170472_m1	82
<i>MKI67</i>	Hs00606991_m1	137
<i>NDN</i>	Hs00267349_s1	109
<i>NUSAP1</i>	Hs00251213_m1	80
<i>PCNA</i>	Hs00427214_g1	138
<i>PRC1</i>	Hs01597831_m1	112
<i>RFC4</i>	Hs00427469_m1	73
<i>RRM2</i>	Hs00357247_g1	79
<i>SLC18A2</i>	Hs00996835_m1	145
<i>SMC4</i>	Hs00374522_m1	91
<i>SYCP2</i>	Hs00204558_m1	123
<i>TOP2A</i>	Hs01032127_g1	76
<i>TYMS</i>	Hs00426591_m1	87
<i>WISP2</i>	Hs00180242_m1	81
<i>ZWINT</i>	Hs00199952_m1	73

La medición de la expresión de los genes fue basada en curvas estándar relativas, de 5 órdenes de magnitud distinta, que van desde 500 a 0.05 ng construida a partir de una mezcla de cDNAs de Ca-Cu o epitelio normal. La primera curva se utilizó para calcular los valores de los genes sobre expresados y la segunda curva los valores de los genes que disminuyeron su expresión. Se probaron las curvas para cada gen en tres diferentes experimentos y los valores de los coeficientes de correlación (r) fueron superiores a 0.98. El nivel expresión de genes blanco se normalizó para cada tumor y ejemplo de control de la intensidad de la referencia interna.

En la normalización se midieron los valores de intensidad en ng/mL. Una prueba de normalidad (Shapiro-Wilk) se llevó a cabo para probar una distribución normal de datos de expresión genética. El cambio en la expresión en “Fold-change” FC se obtuvo dividiendo la mediana intensidad normalizada de las muestras tumor por la mediana de intensidad normalizada de las muestras control. La significancia estadística entre las medianas de tumores y controles se calculó con la prueba no paramétrica de Mann–Whitney (MW). Las correlaciones entre los resultados de MA y los qRT-PCR fue realizada usando los valores log₂ y medido usando el coeficiente de correlación de Pearson.

8. Inmunohistoquímica.

La expresión de 10 genes a nivel de proteínas fue determinada en 26 muestras de Ca-Cu y 10 muestras control por inmunohistoquímica (IH). La construcción de dos microarreglos de tejidos (MAT), fue realizada en el laboratorio, uno de ellos contiene: 14 muestras de Ca-Cu positivas para VPH16 y 5 controles y en el otro contenía 12 muestras de Ca-Cu positivas para otros tipos virales de VPH y 5 muestras controles. La proteína NUSAP1 fue explorada solamente en el primer MAT. Las muestras cilíndricas de regiones seleccionadas en el bloque de tejido incluido en parafina, se les realizó una escisión retirando del bloque una porción de 2mm de diámetro (previamente seleccionados en las laminillas teñidas con hematoxilina y eosina (H&E) y se transfirieron en un recipiente para definir la posición de las muestras y ser nuevamente re-incluidos en parafina. Todos los bloques de tejidos de las pacientes se obtuvieron del Servicio de Patología del HGM.

Para realizar la inmunohistoquímica, secciones seriadas de 4 µm del TMA fueron cortados y el decimo corte fue teñido con H&E para confirmar el diagnóstico histopatológico. Para retirar la parafina, los cortes fueron inmersos en xileno y posteriormente se rehidrataron con alcohol en diferentes grados (70%, 95%, 90%, 80% y 100% v/v en agua). La recuperación del antígenos se realizó colocando las laminillas de los MAT, en una solución amortiguadora “Target Retrieval Solution”, pH 6.0 (Dako, Carpinteria, CA) de recuperación antigénica, por 5 min., a 121° C en una olla a presión. La actividad de peroxidasa

endógena fue bloqueada por incubación de las laminillas de MAT con peróxido de hidrógeno al 1% diluido en PBS por 10 min., así como también un bloqueo de proteínas inespecíficas, por la adición de el reactivo Universal Blocking Reagent 10X, (Biogenex, Fremont CA) 100 μ L y se incubó durante 10 min. Los anticuerpos primarios dirigidos contra **PCNA** (sc-53407); p16 para **CDKN2A** (sc-71804); **PRC1** (sc-56345); Ciclina B2 para **CCNB2** (sc-81241); **CDC2** para p34 (sc-70822); SCP-2 para SYCP2 (sc-20048); y CDKN3 (sc-475); se obtuvieron de Santa Cruz Biotechnologica (Santa Cruz, CA). Los anticuerpos contra Ki-67 para **MKI67** (cat. M7187); CDC20 (cat. 34–1900), y NUSAP1 (cat. H00051203-B01) se obtuvieron de Invitrogen, Dako (Glostrup, Dinamarca) y Nova biológico (Littleton, CO), respectivamente. Los anticuerpos remarcados en negrillas hacen referencia a anticuerpos monoclonales, el resto son policlonales.

La dilución utilizada para todos los anticuerpos fue 1:100, excepto CDC2, (1:50) y NUSAP1 (1:250) y el diluyente de anticuerpo utilizado fue de Dako. En un volumen final de 300 μ L, adicionado a cada sección y las laminillas se incubaron durante la noche a 4°C en cámara húmeda. El complejo antígeno-anticuerpo se detecto por el método avidina-biotina peroxidasa utilizando como un sustrato cromogénico 3,3'-diaminobenzidina-tetrahydrochloride (Cat. KO679 LSAB+ Sys/HRP; Dako Cytomation Carpinteria, CA), y los cortes finalmente se contratiñeron con hematoxilina. Todos los ensayos se realizaron por triplicado.

En cada experimento se incluyó un control positivo para los anticuerpos SYCP2, PRC1, CCNB2, CDKN3, CDC2 y CDC20. El anticuerpo para SYCP2 se probó en tejido de testículo de recién nacido; Los anticuerpos para PRC1, CDC2 y CCNB2 se probaron en tejido de cáncer de colon; y CDKN3 se probó en tejido de cáncer de pulmón. Todos los bloques de tejidos se obtuvieron de material de archivos del servicio de patología del HGM. El porcentaje de células teñidas se calculó en un análisis de 10 campos sucesivos de células neoplásicas. Se identificó la localización celular de la reacción inmune, y la intensidad de la señal clasificó de 0 a 4+, en donde 0 no indica ninguna coloración alguna. Para todos los anticuerpos, raramente las señales de la reacción inmune se encontraron en el estroma y no se registró para el análisis.

Las laminillas se analizaron por 2 patólogos. Los pocos casos que llegaron a ser discordantes en sus mediciones se reevaluaron y se registraron nuevamente para tener una opinión consenso.

9. Análisis de supervivencia de pacientes con cáncer.

Las pacientes con cáncer cervical recibieron un tratamiento individualizado de acuerdo a la clasificación FIGO y siguiendo las recomendaciones de tratamiento para cáncer cervical de la American Cancer Society (ver Tabla V). Después del término de tratamiento, cada paciente se evaluó clínicamente cada 3 ó 6 meses por un oncólogo experimentado. Posteriormente, para el estudio de seguimiento los datos clínicos complementarios se obtuvieron de los registros médicos de pacientes. Además, una trabajadora social realizó llamadas telefónicas y visitas al hogar las pacientes cada 6 meses durante el estudio, 42 meses después tratamiento. Se obtuvieron los registros completos de pacientes vivas, los casos desconocidos y de pacientes fallecidas, se siguieron según el número de meses indicado en la Tabla V.

En este análisis los casos señalados como desconocidos, se refieren a aquellas pacientes que se perdieron durante el periodo del estudio de seguimiento o que fallecieron por otras causas distintas al cáncer cervical. Las pacientes se consideraron pérdidas cuando no asistían a las citas médicas, para el control de la enfermedad, cuando no se encontraron en las visitas a domicilio o no respondían a las llamadas telefónicas. Así mismo, en esta cohorte, las pacientes que se registraron como fallecidas, refiere a las mujeres que murieron como causa principal por el tumor primario de cáncer de cuello uterino. La causa de la muerte de todas, se confirmó por la historia clínica y el certificado de defunción, excepto de una paciente fallecida durante el seguimiento. Solo 42 de 44 pacientes con Ca-Cu positivas para VPH16 que se exploraron con para medir expresión por qRT-PCR y se incluyeron en el estudio de seguimiento. Tres casos fueron registrados como desconocidos y 9 registrados como fallecidos.

Tabla V. Seguimiento de pacientes para evaluación de la sobrevida por al menos 42 meses.

Muestra	Histología ^a	Etapas del tumor	Edad (años)	Tratamiento ^b	Seguimiento (meses)	Estado ^c
R093	ADC	IB1	57	HR	53	Vive
R446	ADC	IB1	43	HR	59	Vive
R081	ADC	IB1	41	HR	62	Vive
R094	CCE	IB1	45	HR	62	Vive
R369	CCE	IB1	50	HR	65	Vive
R057	ADC	IB1	32	HR	93	Vive
R072	CCE	IB1	61	HR+TELE+BRAQUI	86	Vive
R411	CCE	IB1	34	HR+TELE+BRAQUI	60	Vive
R434	ADC	IB1	34	HR+TELE+BRAQUI	61	Vive
R443	CCE	IB1	34	HR+TELE+BRAQUI	61	Vive
R258	CCE	IB1	36	HR+TELE+BRAQUI	68	Vive
R335	ADC	IB1	37	HR+TELE+BRAQUI+QUIMIO	65	Vive
R183	CCE	IB1	64	TELE+BRAQUI	61	Vive
R308	ADC	IB1	45	TELE+BRAQUI	61	Vive
R265	CCE	IB1	46	TELE+BRAQUI	67	Vive
R330	CCE	IB1	72	TELE+BRAQUI	54	Vive
R035	CCE	IB2	48	TELE+QUIMIO+HR	73	Vive
R232	CCE	IB2	45	TELE+BRAQUI+QUIMIO	33	Muerta
R409	CCE	IB2	68	TELE+BRAQUI+QUIMIO	42	Vive
R339	CCE	IB2	31	TELE+BRAQUI+QUIMIO	13	Muerta
R324	CCE	IB2	28	TELE+BRAQUI+QUIMIO	14	Muerta
R359	ADC	IB2	34	TELE+BRAQUI+QUIMIO	17	Muerta
R378	CCE	IB2	42	TELE+BRAQUI+QUIMIO	56	Vive
R312	ADC	IB2	34	TELE+BRAQUI+QUIMIO	58	Vive
R482	CCE	IB2	61	TELE+BRAQUI+QUIMIO	60	Vive
R284	ADC	IB2	33	TELE+BRAQUI+QUIMIO	63	Vive
R412	CCE	IB2	33	TELE+BRAQUI+QUIMIO	63	Vive
R336	CCE	IB2	36	TELE+BRAQUI+QUIMIO	64	Vive
R255	CCE	IIB	45	TELE+BRAQUI+QUIMIO	42	Muerta
R052	ADC	IIB	54	TELE+BRAQUI+HR	19	Muerta
R170	CCE	IIB	67	TELE+QUIMIO+HR	82	Vive
R403	CCE	IIB	34	TELE+BRAQUI+QUIMIO	64	Vive
R015	CCE	IIB	42	TELE+HR	66	Vive
R268	CCE	IIB	34	TELE+BRAQUI+QUIMIO	58	Vive
R415	CAE	IIB	55	TELE+BRAQUI+QUIMIO	59	Vive
R333	CCE	IIB	56	TELE+BRAQUI+QUIMIO	66	Vive
R315	CCE	IIIB	41	TELE+BRAQUI+QUIMIO	8	Muerta
R240	CCE	IIIB	31	TELE+BRAQUI+QUIMIO	11	Muerta

a. ADC, Adenocarcinoma. CCE, carcinoma de células escamosas. CAE, carcinoma adenoescamoso.

b. HR, Histerectomía radical. Tele, teleterapia. Braqui, braquiterapia. Quimio, quimio terapia con cisplatino.

c. Estado, vive en el último registro del seguimiento y la muerte fue causada por el tumor primario del cáncer de cuello uterino.

El tiempo promedio de seguimiento de las 42 pacientes fue de 50.5 meses. La asociación del estadio FIGO y la expresión de genes con la sobrevida (*PRC1*, *CCNB2*, *CDC20*, *CDKN3*, *NUSAP1*, *SYCP-2*, *CDKN2A*, *PCNA*, *MKI67*), se investigó por un análisis de sobrevida. En conjunto toda la muestra, un sistema de cambios aleatorios en 21 muestras fue creado para cada gen explorado. Para categorizar los datos de expresión génica cuantificados por qRT-PCR, un

análisis del área bajo la curva (curvas ROC), fue realizado a los diferentes grupos. Este análisis se utilizó para establecer un nivel de cohorte para los valores de expresión génica, que representa esos valores con sensibilidad y especificidad más alta, para diferenciar entre pacientes vivas y muertas.

Posteriormente, el conjunto de muestra se analizó con los promedios del valor de cohorte, calculado a partir de los valores de formación de conjuntos 500. Las muestras con los valores expresión génica por encima del nivel de cohorte se establecieron en 1 y los valores por debajo del nivel de cohorte se establecieron en 0. El tiempo total acumulado de la supervivencia se calculó mediante el Método de Kaplan-Meier y se analizaron por la prueba de Log-Rank. La etapa FIGO y la expresión génica se incluyeron como covariables en un modelo de riesgos proporcional de Cox.

10. Análisis de clasificación de Ontología de genes.

Para este análisis se utilizó, la base de datos y herramienta para anotación, visualización e integración de las anotaciones funcionales, DAVID (The Database for Annotation, Visualization, and Integrated Discovery, <http://david.ABCC.ncifcrf.gov>), [73,74]; herramienta de anotación funcional y el sistema de análisis de las vías Ingenuity Pathway Analysis (IPA; Sistemas de Ingenuity, www.ingenuity.com) utilizado para clasificar los genes expresados diferencialmente. Los genes se clasificaron usando anotación funcional agrupamiento considerando la ontología génica de los procesos biológicos. La clasificación fue astringente fijando los niveles en alto y medio.

11. Anotación de los genes y el análisis de datos.

La posición física de los genes fue asignada de acuerdo al UCSC Genome Browser Assembly marzo de 2006 NCBI 36/hg18, disponible en <http://genome.ucsc.edu/>. El análisis de datos se realizó empleando el programa Access 2010 (Microsoft Inc.). Los datos en crudo de los MA se organizaron depositaron según los requerimientos del MIAME y se depositaron en la base de datos GEO (<http://www.ncbi.nih.gov/geo/>), con el numero de acceso

GSE39001. El análisis del área bajo la curva (AUC) y el índice de Youden fue calculado [75], para seleccionar los mejores puntos de corte y así distinguir las muestras control de los tumores y los NIC2+ de los NIC1, usando los valores de expresión obtenidos de la qRT-PCR de los genes. Para cada marcador, la sensibilidad, la especificidad, el valor predictivo positivo (VPP) y el valor predictivo negativo (VPN) se calcularon según fórmulas previamente descritos [76]. Todas las pruebas de dos caras y los valores de p menores de 0.005 para ser considerados estadísticamente significativos. Para realizar este tipo de análisis de datos se utilizaron los programas Sigma Stat y SPSS ver. 17.

RESULTADOS.

1. Análisis de expresión de 8,638 genes en cáncer cervical.

La comparación entre 43 muestras de Ca-Cu positivas para VPH16 y 12 muestras de cérvix normal, con el uso del microarreglo HG-Focus, el cual contiene 8,638 transcritos, dio un total de 997 genes que se expresaron diferencialmente entre el grupo de cáncer y el grupo control; para 600 de ellos se encontró que aumentó la expresión y en 397 que su expresión disminuyó (Tabla suplementaria 1 (TS1)). Encontrando tanto en los genes que aumentan como en los que disminuyen un cambio en el FC ("fold change") en el rango de 1.5–2.0 y el número de genes en ambos grupos disminuye linealmente ($r=-0.8$, $p=0.002$), así como el valor de FC se incrementa (Figura 10).

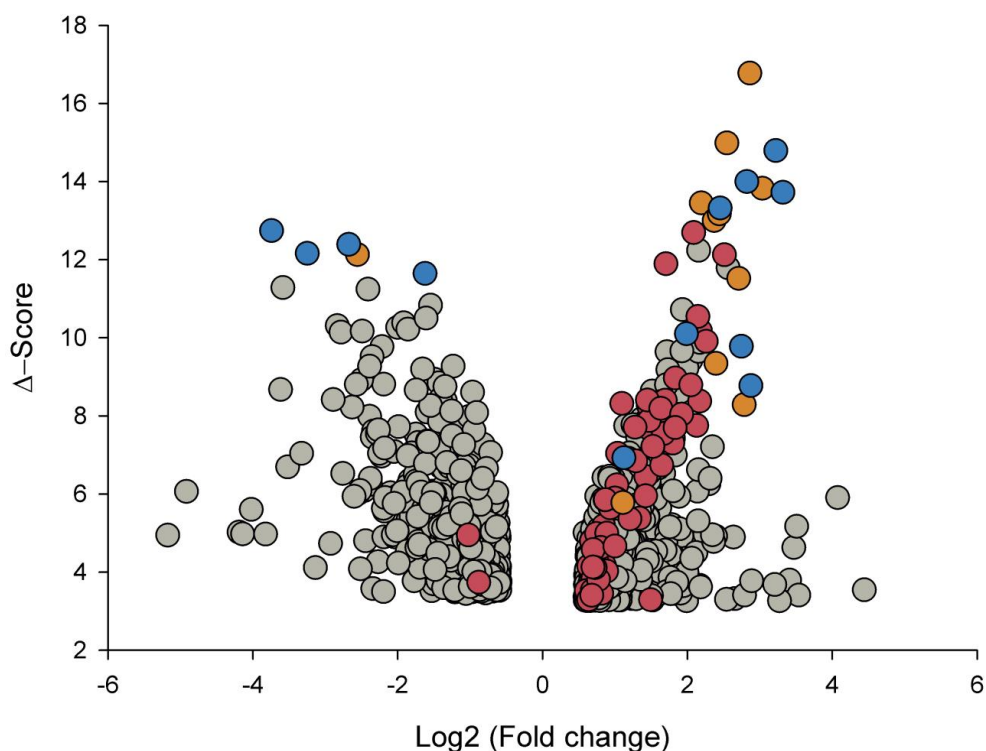


Figura 10. Distribución de genes desregulados según el cambio en número de veces (FC) y los valores de Δ -score. Todos los 997 genes (círculos) que fueron desregulados en tumores de cáncer de cuello uterino (Ca-Cu) en comparación con las muestras control por el método de SAM. La Figura muestra una gráfica de volcán, el eje de las X representa el cambio en la expresión génica expresado en número de veces (FC). (Ejemplo de muestra de cáncer/control) expresado en Log_2 y el eje de las Y representa el valor de Δ -score absoluto, calculado con el método de SAM (prueba de t modificada), los valores mayores de Δ -score, mayor será la significancia estadística. Los valores de Log_2 (FC) son positivos para los genes sobre expresados y los negativos para los genes que disminuyen. Círculos marcados en rojo y naranja representan los genes implicados en la fase M del ciclo celular y los colores en azul y naranja son los genes que fueron validados por qRT-PCR.

El análisis de componente principal (PCA) y el análisis de agrupamiento jerárquico supervisado (panel A en la Figura 11), se realizaron con los valores de expresión de los 997 genes y claramente se observa la separación entre las muestras de cáncer y las del grupo control. Sin embargo la expresión de la mayoría de los genes no fue completamente uniforme en todas las muestras de cáncer, especialmente en el grupo de genes que aumentan su expresión (señales que se muestran en rojo Figura 11a). Muchos de esos genes se sobre-expresaron en algunos tumores y disminuyen en otros. Sin embargo esto contrasta con la uniformidad en las señales de expresión en el grupo de las muestras control.

Los genes que pueden ser probados como marcadores para pesquisa o como blancos con potencial terapéutico, se seleccionaron de acuerdo al valor Δ -score (una prueba de T de student's modificada, usada en SAM), el FC o si han sido descritos como marcadores para cáncer cervical. Para los 997 asociados con las muestras de cáncer, 163 se han reportado previamente como marcadores en diferentes tipos de cáncer (IPA, Ingenuity systems) incluyendo MCM2, TOP2A, y CDKN2A, los cuales se han usado como marcadores diagnóstico en cáncer cervical [77]. Los 997 genes están en orden decreciente por Δ -score (TS1).

Posteriormente, un total de 23 genes (18 sobre expresados y 5 que disminuyen) se seleccionaron para la validación por qRT-PCR (indicado en negrita en la TS1 y Tabla VI; círculos de color azul y naranja en la Figura 10). Los genes que disminuyen (*CFD*, *NDN*, *WISP2*, *END3* y *SLC18A2*) y 10 de los 18 genes de que aumentan su expresión (*PRC1*, *CKS2*, *TYMS*, *RFC4*, *RRM2*, *NUSAP1*, *MCM2*, *CCNB2*, *SMC4* y *CDC2*) fueron seleccionados de acuerdo con el Δ -Score.

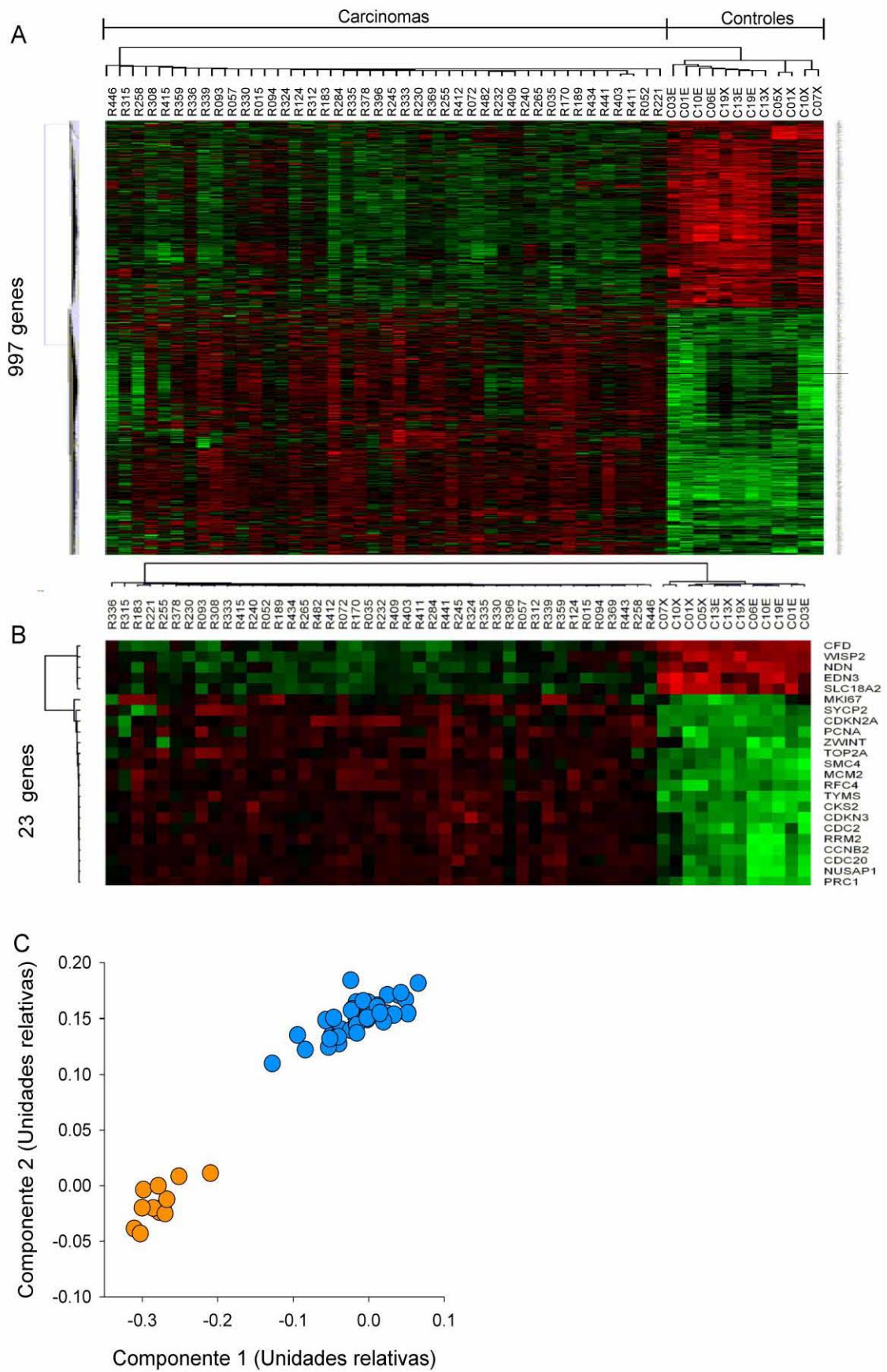


Figura 11. Agrupamiento de las muestras del tumor y control según la expresión de genes desregulados. Análisis de agrupamiento jerárquico no supervisado de 43 muestras de Ca-

Cu y 12 muestras de epitelios cervicales sanos utilizando los valores de la expresión obtenidos con el microarreglo de expresión HG Focus, un total de 997 genes desregulados estadísticamente significativos. Panel A) o en la parte media los 23 genes mejor clasificados y seleccionados para la validación por qRT-PCR, Panel B). Cada renglón representa un gen y cada columna representa una muestra. La longitud y la subdivisión de las ramas representan las relaciones entre las muestras basadas en la intensidad de la expresión génica. El agrupamiento es indicado en color, el código de color rojo es para los genes que se sobre expresan, y el código de color verde para los que disminuyen su expresión y negro para los genes en los que la expresión permanecen sin cambios. Panel C) Análisis de los componentes principales (PCA) usando los valores del panel B; los círculos azules representan las muestras de Ca-Cu (n = 43) y los círculos amarillos representan los controles (n = 12). Ambos grupos de genes separan claramente las muestras en los dos grupos principales con ambos tipos de análisis.

Los otros siete genes restantes de los sobre expresados se encuentran en la lista de los 50 genes mejores clasificados en este estudio; de ellos dos genes ya han propuesto como marcadores en Ca-Cu (*CDKN2A* y *TOP2A*), también los otros cuatro genes (*CDC20*, *CDKN3*, *ZWINT* y *SYCP2*) se seleccionaron con base en el valor FC y finalmente el gen *PCNA*, junto con el gen de *MKI67*, que ubicó en lugar 139, se incluyeron porque estos genes son marcadores ya comúnmente utilizados para medir la proliferación celular (TS1). Sin embargo, la expresión de la mayoría de los genes no fue completamente uniforme en todas las muestras de cáncer, especialmente en el grupo de genes.

Como resultado de ambos análisis, el PCA y el agrupamiento jerárquico, se demostró que los 23 genes seleccionados, también permiten completamente la separación de las muestras en los dos grupos diferentes, controles y tumores. Para ambos genes tanto los que aumentan como los que disminuyen, la diferencia en la intensidad de señal fue bastante uniforme entre las muestras de los 2 grupos (Figura 11, los paneles B y C).

2. Validación de genes expresados diferencialmente con el microarreglo HG-ST1.0.

El número total de genes 8.370 genes que incluye el microarreglo de HG-Focus, se validaron de manera adicional con otro microarreglo de alto rendimiento HG-ST1.0. El análisis con este otro microarreglo (HG-ST1.0) forma parte de otros proyectos realizados en el laboratorio. La comparación incluyó la verificación de 24 muestras (exploradas también con HG-Focus), de las cuales

19 muestras de Ca-Cu y 5 epitelios cervicales sanos. Encontrando una correlación positiva entre los valores de los dos Microarreglos HG-ST1.0 y HG-Focus y esta fue altamente significativa ($p < 1 \times 10^{-15}$, correlación de Pearson). La correlación global entre los 2 arreglos fue de 0.68 y los coeficientes de correlación entre los tumores de forma individual estuvo en el rango de 0.57 a 0.72 (promedio, 0.68), (Figura 12).

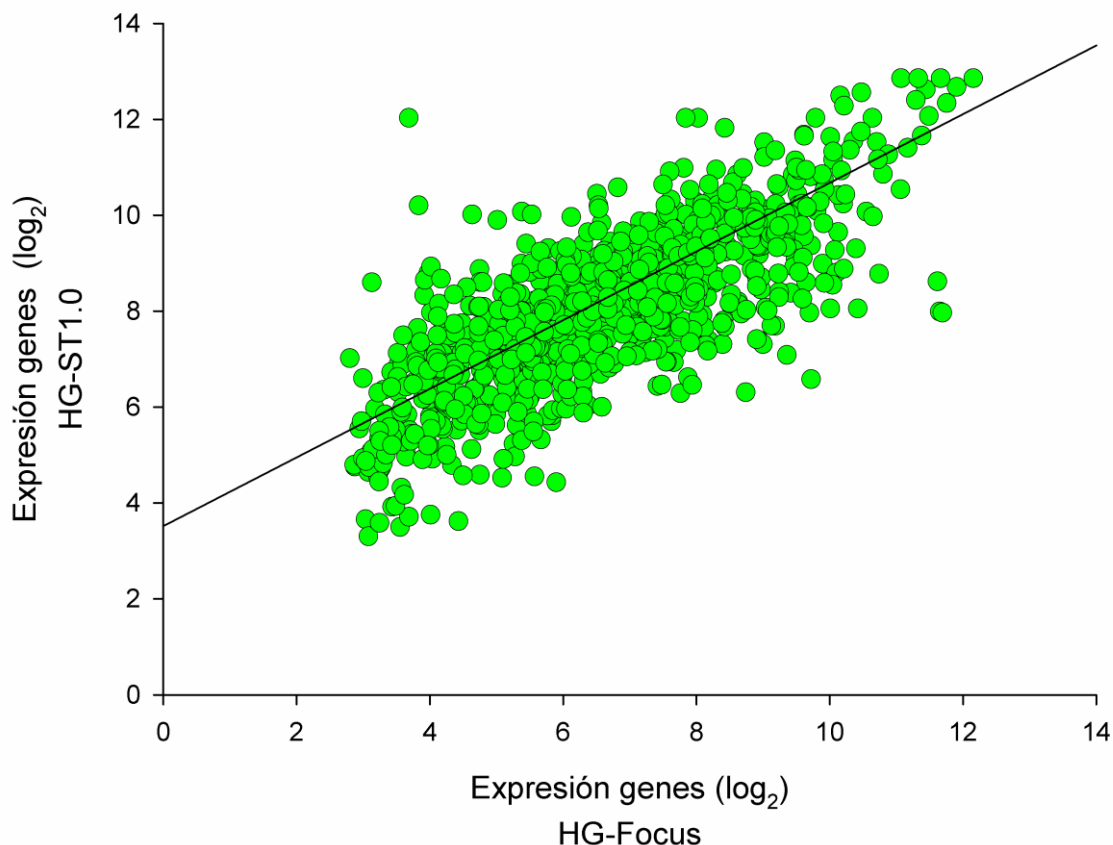


Figura 12. Correlación global de la intensidad de la expresión de 997 genes examinados por microarreglos HG-focus y HG-ST1.0. Se graficaron los valores Log₂ de la señal de la intensidad estandarizada (valores RMA) de los 997 genes de un tumor típico (R230) examinada los 2 microarreglos. La tendencia lineal (línea negra) está incluida, que se calculó con la prueba de correlación de Pearson: $r = 0,78$; $p < 1 \times 10^{-15}$.

Los valores de expresión génica para 826 genes de los 997 (82.8%), expresados diferencialmente entre cáncer y controles, incluyen los 23 genes seleccionados para validación (Figura 13), mostrando una correlación positiva ($p < 0.05$, correlación de Pearson) entre los 2 microarreglos. Los coeficientes de correlación entre los genes individuales oscilan entre 0,34 y 0.95 y el promedio fue de 0,63.

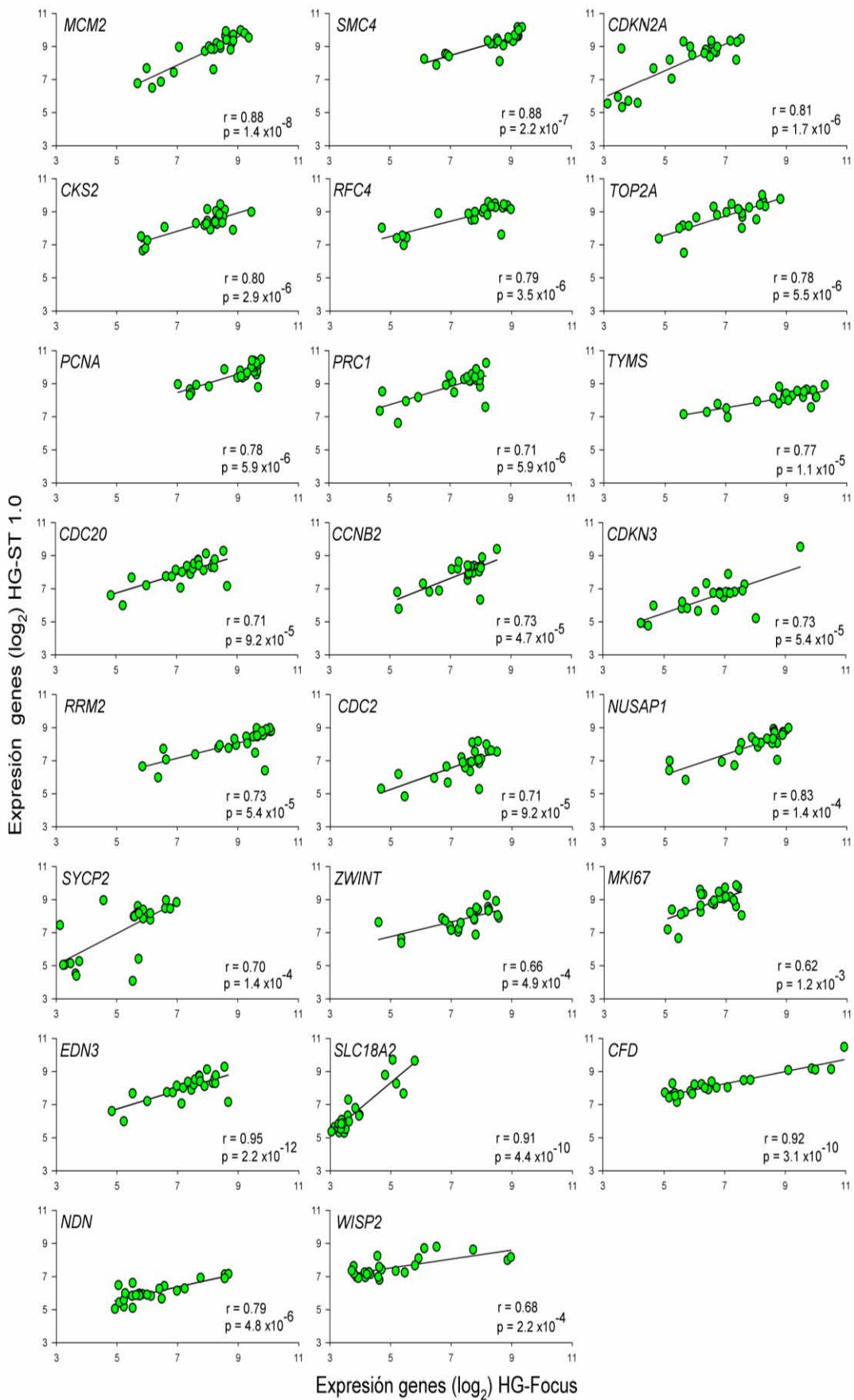


Figura 13. Correlación de la expresión de la intensidad de 23 genes entre los dos microarreglos, examinados (microarreglos HG-focus y HG-ST1.0). Se graficaron los valores

de la intensidad de señal estandarizados en Log₂ (valores RMA) de 23 genes examinados en los 2 micro arreglos de 19 muestras de Ca-Cu y 5 muestras de epitelios cervical normal. La línea de tendencia (línea negra) es incluida, y se calculó con la prueba de correlación de la Person, r = coeficiente de correlación, p = valor de p .

3. Validación por PCR cuantitativa a tiempo real (qRT-PCR).

Por otra parte, la expresión de los genes de 23 seleccionados para validación se midió con la qRT-PCR en un total de 44 muestras de cáncer positivas para el VPH16 y 25 epitelios cervicales sanos, que incluyeron casi todas las muestras previamente analizadas con los MA (Tabla III). Una correlación positiva altamente significativa ($p < 0.0001$, correlación de Pearson) se encontró entre los ensayos de qRT-PCR y MA para 21 de los 23 genes medidos, los valores expresados en log₂. Los coeficientes de correlación variaron de 0.31 a 0.85 y la mediana fue de 0,73. Sin embargo en 2 genes no tuvieron correlaciones no significativas (*NDN* y *SLC18A2*), estos genes se excluyeron del resto de los análisis. Estos datos indicaron que los valores de la expresión que se calcularon a partir de los MA fueron bastante confiables porque el 91% de los genes validados tuvieron una correlación significativa. Debido a que los valores de expresión de la qRT-PCR no seguían una distribución normal, para realizar este análisis se uso la mediana para los cálculos del FC. Los genes son listados en orden decreciente en la Tabla VI, y el gen con el mayor valor de FC de esa lista fue *MKI67*, que es seguido orden decreciente por *CDKN2A*, *SYCP2*, *PCNA*, *NUSAP1* y *CDC2*.

En este trabajo es notorio el incremento en el FC de *MKI67* (1,651) y *CDKN2A* (387) en por lo menos 5 veces más que la FC del gen que sigue en la lista (*SYCP2*; FC = 73,8). De los 10 genes sobre expresados, 2 de los genes mejor clasificados no han sido asociados con cáncer previamente (*NUSAP1* y *CDKN3*), mientras que los otros 8 genes, de 4 de estos, *SYCP2*, *PRC1*, *CCNB2* y *CDC20*, existen pocos reportes de su asociación con Ca-Cu, mientras que *MKI67*, *CDKN2A*, *CDC2*, y *PCNA*, se han asociado ampliamente con Ca-Cu. Los genes *MCM2* y *TOP2A*, clasificados en este estudio el lugar 15 y 18 respectivamente, también están ampliamente reportados asociados con Ca-Cu.

Tabla VI. Genes explorados por qRT-PCR.

Gen ^a	Cambio en # de veces ^b	
	VPH16 positivos	Otros VPHs ^c
Aumentan		
<i>MKI67*</i>	1651	-
<i>CDKN2A*</i>	387	-
<i>SYCP2</i>	74	14
<i>PCNA*</i>	65	-
<i>NUSAP1</i>	26	15
<i>CDC2</i>	23	-
<i>CDC20</i>	17	13
<i>CCNB2</i>	14	6
<i>TYMS</i>	12	2
<i>PRC1</i>	9	4
<i>SMC4</i>	8	-
<i>CDKN3</i>	7	5
<i>RRM2</i>	6	-
<i>CKS2</i>	5	-
<i>MCM2</i>	4	-
<i>ZWINT</i>	4	-
<i>RFC4</i>	4	-
<i>TOP2A</i>	3	-
Disminuyen		
<i>EDN3</i>	1426	-
<i>WISP2</i>	168	-
<i>CFD</i>	25	-
<i>NDN</i>	1	-
<i>SLC18A2</i>	0.3	-

a. Genes marcados en negro se seleccionaron para explorarse en muestras pre-invasoras. Los marcados con el asterisco se han reportado previamente estar asociados con Ca-Cu.

b. El análisis se realizó con 44 Ca-Cu positivas para VPH16, 22 muestras de Ca-Cu positivas para otros tipos virales VPH y 25 muestras controles de cérvix normal. El incremento en número de veces (FC) se calculó con los valores de la mediana como se describe en materiales y métodos: tumor/control para los genes que aumentan su expresión y control/tumor para los genes que disminuyen (ver Material y Métodos). Una prueba de Mann-Whitney Rank Sum se realizó para ver la diferencia entre los grupos que fue estadísticamente significativos ($p < 1 \times 10^{-15}$) para todos, excepto 2 genes (*NDN*, *SLC18A2*). El valor de $p > 0.05$ para *NDN* y *SLC18A2*.

c. Incluye carcinomas positivos para VPH-18 (5), -31 (4), -33 (2), -45 (5), -51 (2), -58 (2) y -59 (2).

Los tres genes que disminuyen su expresión tuvieron una correlación significativa y también un alto valor de FC (controles contra cánceres) especialmente en *END3* (FC = 1, 425.7) y *WISP2* (FC = 167.7; Tabla VI). Las gráficas de cajas y bigotes (Figuras 14 y 15) muestran claramente la diferencia en la expresión génica entre los grupos control y cáncer ($p < 1 \times 10^{-15}$, para todos los genes, prueba U de Mann-Whitney).

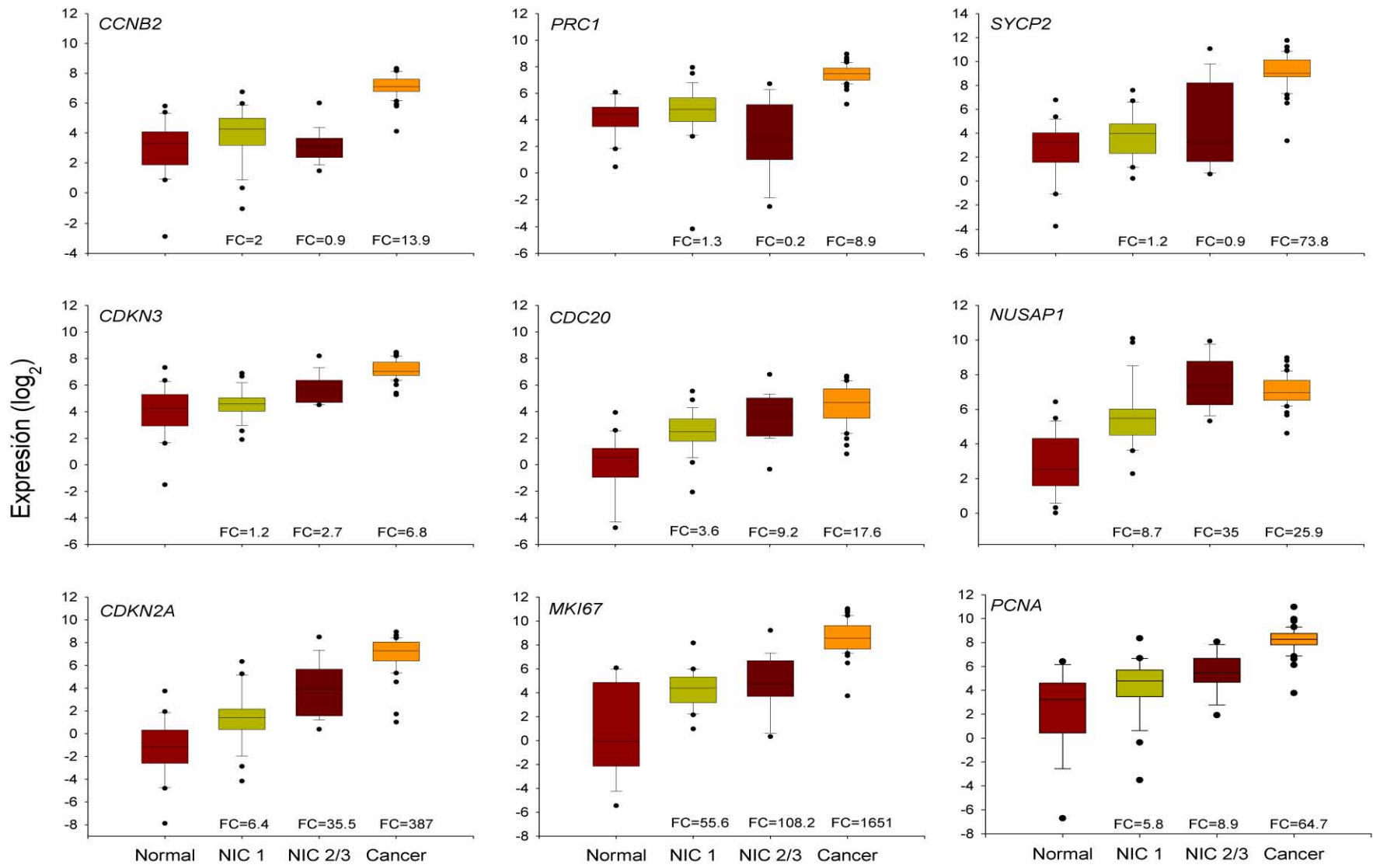


Figura 14. Validación de la expresión génica de 9 marcadores genéticos por qRT-PCR. La intensidad de la expresión génica, es expresada en valores en Log₂, mostrada en los diagramas de caja. La expresión de los 6 genes validados en este estudio (*CCNB2*, *PRC1*, *SYCP2*, *CDKN3*, *CDC20* y *NUSAP1*) y los 3 genes ya conocidos que se asocian a Ca-Cu (*CDKN2A*, *MKI67* y *PCNA*) se comparan entre los 4 grupos, incluyendo epitelio cervical sano (Normal, n = 25), lesiones de bajo grado NIC, (NIC1, n = 29), NIC de alto grado (NIC2/3, n = 21) y Ca-Cu invasor (cáncer, n = 44). Los límites superiores e inferior de las cajas representan los percentiles al 75 y 25, respectivamente. La línea negra dentro de la caja representa el valor de la mediana, y los bigotes representan el mínimo y máximo valores que se encuentran en 1.5 x del rango intercuartíl: desde el extremo de la caja. Valores fuera de este rango están representados por círculos negros. El cambio en número de veces (FC), se calculó dividiendo la mediana de cada grupo patológico por la mediana del grupo control.

Para establecer una línea de separación entre los dos grupos y el valor potencial de estos genes como marcadores de cáncer de cuello uterino, se determinó un valor de cohorte establecido mediante el análisis de las curvas ROC. En general, las curvas ROC con un valor de área bajo la curva (ABC) de 0.97 tiene un valor clínico alto, mientras que un $ABC \leq 0.75$ no son útiles clínicamente [78]. El valor de ABC para los 11 genes que aumentaron su expresión (*CDKN2A*, *MKI67*, *PRC1*, *CDC2*, *CCNB2*, *SYCP2*, *PCNA*, *NUSAP1*, *TYMS*, *CDC20* y *CDKN3*) y un gen que disminuye (*CFD*) fue de >0.97 , ver Tabla VII.

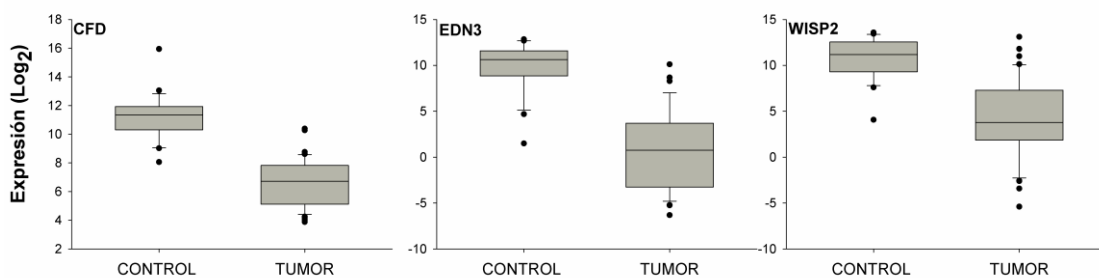


Figura 15. Validación de la expresión de marcadores que están disminuidos por qRT-PCR. La intensidad de la expresión génica, es expresada en valores en Log₂, mostrada en los diagramas de caja. La expresión de los 3 genes validados en este estudio (*CFD*, *EDN3* y *WISP2*) se comparan entre epitelio cervical sano (Normal, n = 25) y Ca-Cu invasor (cáncer, n = 44).

Tabla VII. Análisis ROC y cálculos de sensibilidad, especificidad y valores predictivos.

Genes	ABC	Valor de cohorte ^a	Controles		Tumores		p-valor ^b	Sensibilidad	Especificidad	VPP	VPN	Índice J
			FFP	FVN	FVN	FFN						
<i>MKI67</i>	0.99	79	0	23	40	1	8.2E-14	0.98	1	100	95.8	0.98
<i>PRC1</i>	0.99	73	1	22	40	1	8.2E-14	0.98	1	97.6	95.7	0.98
<i>CDKN2A</i>	0.99	18	0	23	39	2	5.3E-13	0.95	1	100	92.0	0.95
<i>CCNB2</i>	0.99	58	0	23	39	2	5.3E-13	0.95	1	100	92.0	0.95
<i>CDC2</i>	0.99	85	0	23	39	2	5.3E-13	0.95	1	100	92.0	0.95
<i>SYCP2</i>	0.99	66	1	22	40	1	8.2E-14	0.96	0.98	97.6	95.7	0.94
<i>PCNA</i>	0.98	92	1	22	39	2	5.3E-13	0.95	1	97.5	91.7	0.95
<i>NUSAP1</i>	0.98	48	2	21	40	1	6.7E-13	0.98	0.96	95.2	95.5	0.94
<i>TYMS</i>	0.98	40	1	22	39	2	4.3E-12	0.95	0.96	97.5	91.7	0.91
<i>CDC20</i>	0.97	6	1	22	36	5	4.9E-10	0.88	0.96	97.3	81.5	0.84
<i>CDKN3</i>	0.97	90	1	22	36	5	4.9E-10	0.88	0.96	97.3	81.5	0.84
<i>SMC4</i>	0.95	235	4	19	39	2	1.2E-09	0.95	0.87	90.7	90.5	0.82
<i>RFC4</i>	0.93	221	3	20	38	3	1.1E-09	0.93	0.87	92.7	87.0	0.80
<i>RRM2</i>	0.92	103	3	20	39	2	2.0E-10	0.93	0.87	92.9	90.9	0.80
<i>MCM2</i>	0.90	124	3	20	37	4	7.8E-10	0.90	0.91	92.5	83.3	0.81
<i>TOP2A</i>	0.88	128	4	19	40	1	2.0E-10	0.98	0.83	90.9	95.0	0.81
<i>ZWINT</i>	0.83	59	5	18	36	5	5.3E-07	0.85	0.78	87.8	78.3	0.63
<i>CKS2</i>	0.79	239	5	18	33	8	1.5E-05	0.81	0.78	86.8	69.2	0.59
			TPF	FNF	VPF	TNF						
<i>CFD</i>	0.98	478	22	1	2	39	4.30E-12	0.96	0.95	91.7	97.5	0.91
<i>EDN3</i>	0.96	21	22	1	7	34	6.70E-09	0.96	0.83	75.9	97.1	0.79
<i>WISP2</i>	0.95	186	22	1	9	32	6.70E-08	0.96	0.78	71.0	97.0	0.74

ABC=área bajo la curva, FFP=fracción falsos positivos, FVN=fracción verdaderos negativos, FVN=fracción verdaderos negativos,

FFP=Fracción falsos negativos, VPP=Valor predictivo positivo, VPN=Valor predictivo negativo.

a. Valor óptimo de cohorte (ng/ml) fue seleccionado de acuerdo al análisis ROC.

b. prueba de χ^2 .

J=sensibilidad+especificidad-1.

De hecho, la mayoría de estos genes tienen una sensibilidad y especificidad superior al 95%, sugiriendo que podrían ser buenos marcadores para pesquisa entre muestras sanas y cánceres invasores.

Interesantemente, este subconjunto de genes que incluye los 2 (*NUSAP1* y *CDKN3*) que no se han asociado con Ca-Cu y los 5 se han asociado poco con Ca-Cu (*PRC1*, *SYCP2*, *CCNB2*, *TYMS* y *CDC20*); estos 7 genes fueron evaluados con qRT-PCR en otras 22 muestras de Ca-Cu positivas para otros tipos virales, incluyendo los tipos VPH18, 31, 33, 45, 51, 58 y 59. Todos los genes se sobre expresaron en las 22 muestras de tumores; sin embargo, los valores de FC fueron menores a los obtenidos en los tumores positivos para VPH16 (Tabla VI). Estos datos sugieren que estos 7 genes pueden estar sobre expresados en todos los cánceres invasores independientemente del tipo viral y además pueden ser considerados como marcadores potenciales para la detección de Ca-Cu.

4. Análisis de la expresión de *CCNB2*, *CDC20*, *PRC1*, *SYCP2*, *NUSAP1*, y *CDKN3* en neoplasias pre-invasivas.

Para las pruebas de detección, es importante identificar no sólo Ca-Cu invasor, sino también las lesiones de alto grado (NIC2/3) y distinguirlas de las lesiones de bajo grado NIC (NIC1) y los controles sanos. Por consiguiente, para investigar si estos genes pueden distinguir una lesión de NIC2+ de una NIC1-, para lo cual, se analizó la expresión en 2 grupos adicionales de muestras: 29 NICs bajo grado y 21 NICs de alto grado. La expresión fue comparada con 3 marcadores conocidos asociados con Ca-Cu (*PCNA*, *MKI67* y *CDKN2A*), y estos a su vez se clasificaron dentro de los 10 primeros en el análisis previo de qRT-PCR (véase arriba). La gráfica de los datos experimentales es mostrada en la Figura 14 y la significancia estadística de las diferencias se calculó mediante la prueba de MW.

De acuerdo a la mediana y la distribución de los datos, en la grafica de caja y bigotes, los 9 marcadores pueden ser clasificados dentro de 3 grupos; el primer grupo incluye los marcadores asociados exclusivamente a la invasión (*CCNB2*,

PRC1) o mayormente (*SYCP2*), el cual claramente distinguió los tumores invasores de las lesiones pre-invasoras CIN de alto y bajo grado y las muestras de cérvix normal. La expresión de estos marcadores en el grupo de control y las lesiones pre-invasoras de bajo y alto grado fue similar ($p > 0.05$, MW). En contraste, entre las muestras Ca-Cu y las muestras control, la diferencia fue bastante amplia, como se estableció en el análisis anterior (Tabla VI).

Similarmente, la comparación del FC para las lesiones NIC-AG, fue muy alto, especialmente para *SYCP2* (FC = 84.8; $p < 1 \times 10^{-15}$ MW), seguido por *PRC1* (FC = 39.4; $p < 1 \times 10^{-15}$, MW) y *CCNB2* (FC = 15.9; $p < 1 \times 10^{-15}$ MW). Por otra parte, la especificidad para la detección de Ca-Cu y no de otras lesiones, osciló entre 0,85 (*SYCP2*) y 0.98 (*CCNB2*); el límite del valor óptimo fue un cambio en el FC de 4.5 veces más. La baja especificidad mostrada por *SYCP2* fue porque 7 lesiones pre-invasoras (5 NIC2/3 y 2 NIC1) tenían un valor de FC mayor que el óptimo valor de corte de para ese gen (7.9).

El segundo grupo incluye 4 genes (*CDC20*, *NUSAP1*, *CDKN2A* y *CDKN3*) la expresión de que tendió a aumentar desde el grupo de control al grupo Ca-Cu (*CDC20*, *CDKN2A* y *CDKN3*) o el grupo NIC de alto grado (*NUSAP1*). Para *NUSAP1*, la expresión fue similar entre CIN2/3 y Ca-Cu (Figura 14). Estos 4 genes podrían distinguir lesiones NIC2+ de las lesiones NIC1/2 ($p < 1 \times 10^{-15}$, MW; Figura 14).

El tercer grupo incluye a los genes *MKI67* y *PCNA*, para estos la expresión de que aumentó desde el grupo de control al grupo NIC de bajo grado ($p < 0.05$, MW), siendo similar en los grupos de lesiones NIC de alto y bajo ($p > 0.05$, MW) y luego aumenta para el grupo Ca-Cu ($p < 1 \times 10^{-15}$, MW; Figura 14).

Es claro que los genes del primer y tercer grupo no serían buenos marcadores por que no pueden distinguir de las lesiones NIC de alto grado y Ca-Cu de las lesiones NIC de bajo grado y las muestras control. Sin embargo, para explorar el potencial de los genes del segundo grupo (*CDC20*, *CDKN2A*, *CDKN3* y *NUSAP1*) como marcadores para la detección nuevamente se realizó un análisis ROC, empelando los datos de estos genes en conjunto; no obstante, el

resultado mostró que en ninguno de ellos tuvieron valores ABC iguales o superiores a 0,97; el valor más alto de las ABC se obtuvo con *CDKN2A* (0.92), seguido por *NUSAP1* (0.917), *CDKN3* (0.91) y *CDC20* (0.86) (Tabla VIII).

Los marcadores (*NUSAP1* y *CDKN3*) mostraron una sensibilidad ligeramente mayor que el gen de *CDKN2A*, mientras que lo contrario fue verdadero para la especificidad (Tabla VIII). Interesantemente, tanto la sensibilidad como la especificidad se incrementan cuando los datos individuales de *CDKN3*, *NUSAP1* y *CDKN2A* se combinaron (Tabla VII). Esas combinaciones mostraron un índice de Jouden alto. De éstos, sólo el gen *CDKN3*, mostro que puede también discriminar además de Ca-Cu de los NIC2/3 (FC valor de cohorte = 4.4) con una alta sensibilidad (0.9) y especificidad (0.84).

Tabla VIII. Análisis ROC de 4 genes marcadores seleccionados para las lesiones de alto grado y Ca-Cu.

	ABC	VC ^a	<NIC1 (n=50)		>NIC2/3 (n=60)		Valor p^b	Sensibilidad	Especificidad	VPP	VPN	Índice J
			FFP	FVN	FVN	FFN						
<i>CDKN2A</i>	0.92	14	4	46	49	11	6.00E-14	0.81	0.92	92.5	80.7	0.73
<i>CDKN3</i>	0.91	30	13	37	54	6	2.86E-11	0.90	0.74	80.6	80	0.64
<i>NUSAP1</i>	0.90	71	6	44	53	7	6.00E-15	0.88	0.88	89.8	86.3	0.76
<i>CDC20</i>	0.87	5	15	35	51	9	1.45E-08	0.85	0.70	77.3	79.5	0.55
<i>CDKN3, CDKN2A, CDC20</i>			9	41	55	5	2.80E-14	0.92	0.82	85.9	89.1	0.74
<i>CDKN3, CDKN2A, NUSAP1</i>			6	44	55	5	1.00E-15	0.92	0.88	90.2	89.8	0.80
<i>CDKN3, CDC20, NUSAP1</i>			8	42	53	7	5.50E-14	0.88	0.84	86.9	85.7	0.72
<i>CDKN2A, CDC20, NUSAP1</i>			5	45	53	7	1.00E-15	0.88	0.90	91.4	86.5	0.78

ABC=área bajo la curva, FFP=fracción falsos positivos, FVN=fracción verdaderos negativos, FVN=fracción verdaderos negativos, FFP=Fracción falsos negativos, VPP=Valor predictivo positivo, VPN=Valor predictivo

a. Valor óptimo de cohorte (ng/ml) fue seleccionado de acuerdo al análisis ROC.

b. prueba de X^2

c. J=sensibilidad+especificidad-1.

5. Verificación de la expresión de genes candidatos a marcadores tumorales a nivel de proteínas por inmunohistoquímica.

Para investigar si los genes validados *PRC1*, *CDKN3*, *CCNB2*, *SYCP2*, *NUSAP1* y *CDC20*, también se sobre expresan en el nivel de proteína, se realizó un análisis de la expresión a nivel de proteínas por IH. En estos experimentos se incluyeron también las proteínas PCNA, CDKN2A, MKI67 y CDC2, que han sido ampliamente estudiadas y reportadas en cáncer cervical. Todas proteínas fueron evaluadas por IH en los MAT que contenían, 36 muestras, 10 controles y 26 Ca-Cu, de estas 14 muestras positivas para VPH16 y 12 muestras positivas para otros tipos de VPH, (Figura 16), a excepción de la proteína NUSAP1, que solamente fue explorada en el MAT que contenía Ca-Cu positivos para VPH16 y 5 muestras controles.

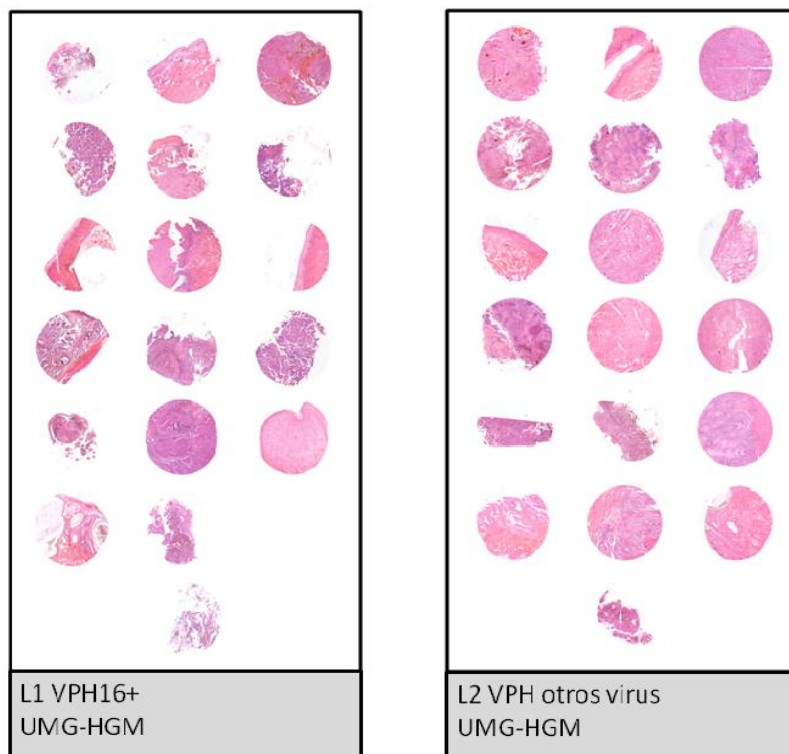


Figura 16. Microarreglos de tejidos (MAT). Tinción con H&E de una laminilla de cada uno de los MA, el MAT L1 contiene 14 muestras de Ca-Cu positivas para VPH16 y 5 controles de cervix sanos. El MAT L2 contiene 12 muestras de Ca-Cu positivas para otros tipos virales y 5 controles de cervix sanos.

A diferencia de los controles casi todas las muestras fueron positivas para los 10 antígenos (Figura 17). El porcentaje más alto de tumores positivos y las señales más intensas se observaron en el antígeno PCNA con un 96.2%; seguido de

CDKN2A y CDKN3, con el 84.6%; CCNB2 y CDC2, un 80.8%; NUSAP1, con 79%; MKI67, SYCP2 y PRC1 con el 76,9%; finalmente CDC20, con un 73.1%. Sin embargo, para tres antígenos inesperadamente un número considerable de controles dieron señal positiva para CDC20 en el 60%; NUSAP1 en un 40% y también SYCP2 con un 50%. Las señales encontradas en las muestras control en estos antígenos tuvieron un patrón de distribución distinto, para CDC20 las señales sólo se observaron en los núcleos de las células en la capa basal, para NUSAP1 las señales se observaron tanto en el núcleo como en el citoplasma de las células en las capas basal y parabasal y para el antígeno SYCP2, la señal se encontró en el polo basal en las células epiteliales de las capas superficiales e intermedias principalmente.

Asimismo, para el resto de antígenos, las diferencias de positividad entre los dos grupos (tumores y controles), coinciden con los datos obtenidos con los ensayos de qRT-PCR (Tabla IX). Las señales para los antígenos de CDKN3, SYCP2, PRC1, CDC2, NUSAP1, y CDKN2A se observaron tanto en el núcleo como en el citoplasma, mientras para el antígeno de CCNB2 sólo se observaron señales de en citoplasma, y las señales para los antígenos CDC20, PCNA y MKI67, sólo se observaron en el núcleo (Figura 18).

Es importante mencionar que como era esperado, las señales de IH señales no fueron completamente uniformes en todas las células de todos los tejidos, debido a que la distribución fue heterogénea, indicando que no todas las células están en la misma etapa del ciclo celular. Las señales para el antígeno PCNA mostraron la distribución más uniforme, y se encontró en promedio en el 70% de los núcleos fueron positivos, lo que sugiere que aproximadamente el 70% de las células en los tejidos se encontraban en la fase S del ciclo celular. Para el resto de las proteínas, las señales nucleares se observaron del 10 al 50% de las

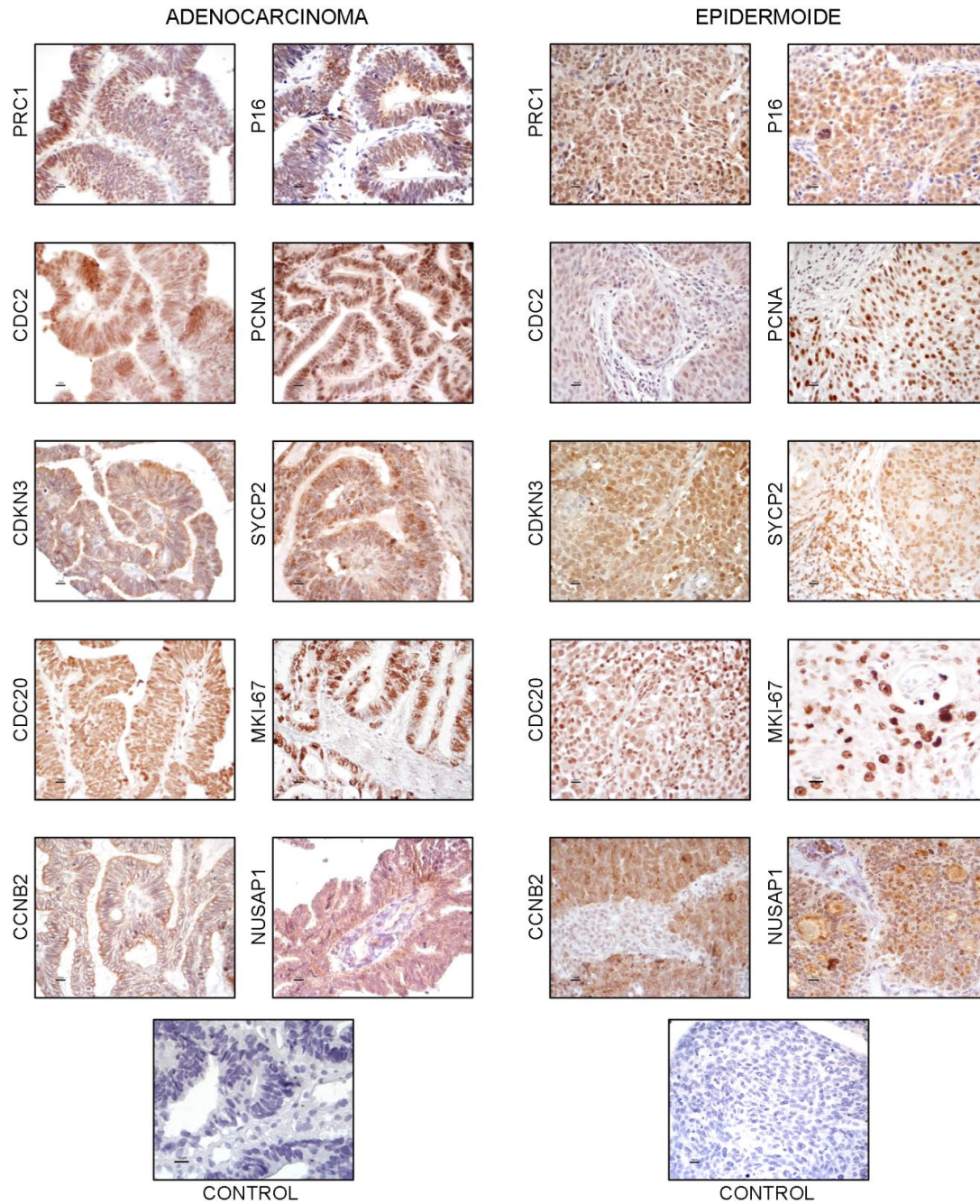
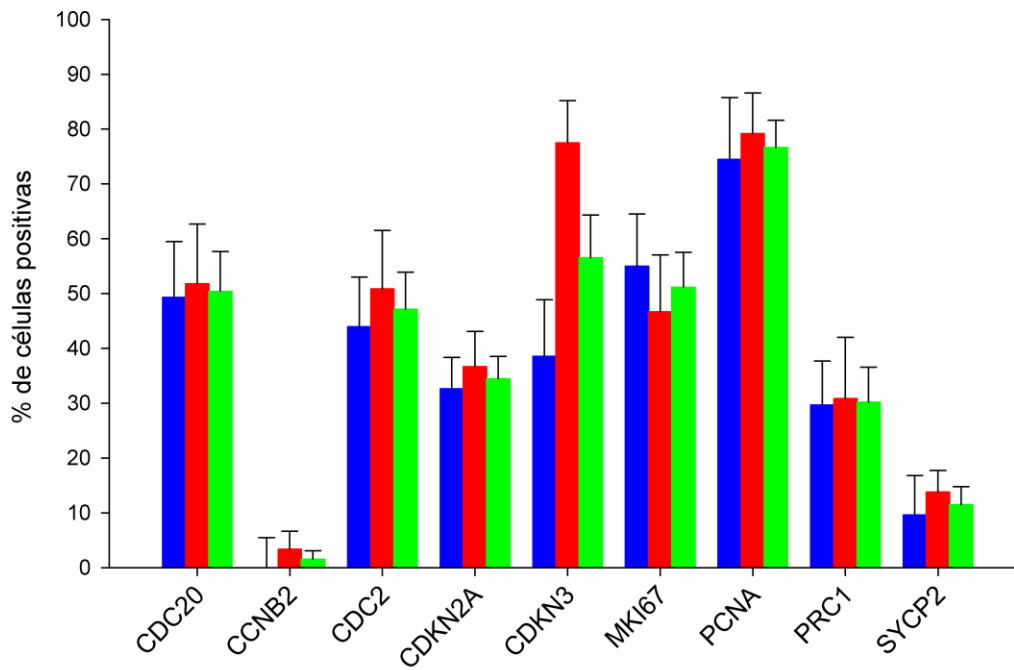


Figura 17. Análisis histológico de los genes marcadores. La expresión a nivel de proteína se determinó por inmunohistoquímica, utilizando secciones de tejido fijadas en formalina del tejido embebido en parafina. Las proteínas exploradas son PRC1, p16, CDC2, PCNA, CDKN3A, SYCP2, CDC20, MKI-67, CCNB2, y NUSAP1. La Figura muestra algún ejemplo representativo de los experimentos en tumores de adenocarcinomas y epidermoides. Las señales específicas aparecen como tinción en color marrón (contratinción con hematoxilina; ampliación original, X400; barras, 10 μ m).

células (Figura 18A). Asimismo, las señales de las proteínas que se localizan en el citoplasma de las células se observaron en promedio en el 40-50% (Figura 18B).

A



B

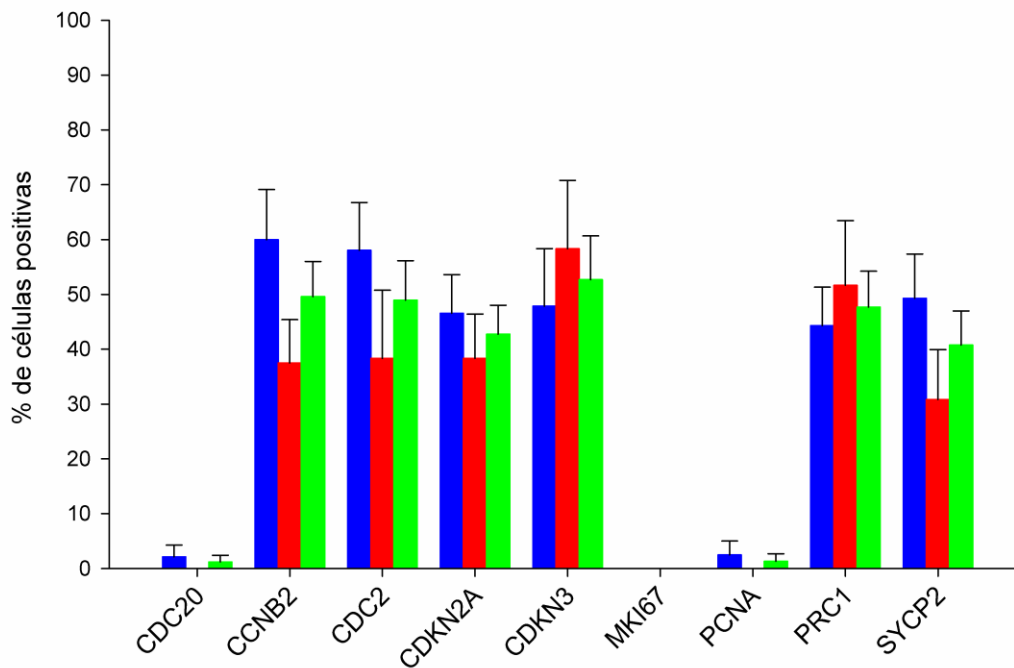


Figura 18. Porcentaje de células tumorales con señales positivas para los marcadores evaluados por inmunohistoquímica. En la gráfica, se muestra la media de los porcentajes de las células tumorales con señales positivas en el núcleo (panel A) o en el citoplasma (panel B) para CDC20, CCNB2, CDC2, CDKN2A, CDKN3, MKI67, PCNA, PRC1 y SYCP2. Los tumores positivos para VPH16 (barras azules) se comparan con los tumores positivos para otros tipos de VPH (barras rojas). El porcentaje total para ambos grupos de tumores también es mostrado (barras verdes). Las líneas por encima de las barras indican el error estándar.

Debido a que todas estas proteínas participan en la Fase M del ciclo celular, los datos sugieren que el 40% de las células están en alguna etapa de esa fase. Casualmente, el porcentaje de células positivas para las proteínas CCNB2, CDC2 y SYCP2 fue alto en los tumores positivos para los VPH16, que en los tumores positivos para otros tipos de VPH y lo opuesto se observó para la proteína CDKN3 (Figura 18). La proteína NUSAP1 sólo fue explorada en Ca-Cu positivos para VPH16, el resultado en porcentaje promedio de células tumorales positivas fue en el núcleo 32,5% y el citoplasma 68.4%.

La capacidad predictiva de los resultados obtenidos por IH también fue evaluada. Sin embargo, comparando con los resultados de los ensayos de qRT-PCR, los datos muestran que la sensibilidad fue baja para todas las proteínas, pero la especificidad fue alta en todas las proteínas, excepto para SYCP2, NUSAP1 y CDC20 (Tabla IX).

Tabla IX. Análisis de la 10 proteínas en 26 muestras de Ca-Cu y 10 controles de epitelio cervical normal por IH y el cálculo de sensibilidad, especificidad y valores predictivos.

Proteínas*	CONTROLES		Ca-Cu		Sensibilidad	Especificidad	PPV	NPV	índice (J)
	FFN	FVN	FVN	FFN					
Señales Citoplasmicas									
PCNA	0	10	1	25	0.04	1.00	100	28.6	0.04
CDKN2A	0	10	21	5	0.81	1.00	100	66.7	0.81
CDKN3	0	10	17	9	0.65	1.00	100	52.6	0.65
CDC2	0	10	18	8	0.69	1.00	100	55.6	0.69
CCNB2	0	10	21	5	0.81	1.00	100	66.7	0.81
SYCP2	5	5	19	7	0.73	0.50	79.2	41.7	0.23
MKI67	0	10	0	26	0.00	1.00	0	16.1	0.00
PRC1	0	10	20	6	0.77	1.00	100	62.5	0.77
CDC20	0	10	1	25	0.04	1.00	100	28.6	0.04
NUSAP1	2	3	11	3	0.79	0.60	78.6	50.0	0.3
Señales Nucleares									
PCNA	0	10	25	1	0.96	1.00	100	90.9	0.96
CDKN2A	0	10	22	4	0.85	1.00	100	71.4	0.85
CDKN3	0	10	18	8	0.69	1.00	100	55.6	0.69
CDC2	0	10	19	7	0.73	1.00	100	58.8	0.73
CCNB2	0	10	1	25	0.04	1.00	100	28.6	0.04
SYCP2	5	5	10	16	0.38	0.50	66.7	23.8	-0.12
MKI67	0	10	20	6	0.77	1.00	100	62.5	0.77
PRC1	0	10	15	11	0.58	1.00	100	47.6	0.58
CDC20	6	4	19	7	0.73	0.40	76	36.4	0.13
NUSAP1	2	3	7	7	0.50	0.60	70.0	30.0	-0.2
Señales citoplasmicas y/o nucleares									
PCNA	0	10	25	1	0.96	1.00	100	90.9	0.96
CDKN2A	0	10	22	4	0.85	1.00	100	71.4	0.85
CDKN3	0	10	22	4	0.85	1.00	100	71.4	0.85
CDC2	0	10	21	5	0.81	1.00	100	66.7	0.81
CCNB2	0	10	21	5	0.81	1.00	100	66.7	0.81
SYCP2	5	5	20	6	0.77	0.50	80	45.5	0.27
MKI67	0	10	20	6	0.77	1.00	100	62.5	0.77
PRC1	0	10	20	6	0.77	1.00	100	62.5	0.77
CDC20	6	4	19	7	0.73	0.40	76	36.4	0.13
NUSAP1	2	3	11	3	0.79	0.60	78.6	50.0	0.3

FFP=fracción falsos positivos, FVN=fracción verdaderos negativos, FVN=fracción verdaderos negativos, FFP=fracción falsos negativos, VPP=valor predictivo positivo, VPN=valor predictivo negativo.

J=sensibilidad +especificidad-1

* NUSAP1 solamente fue explorada en 14 muestras de Ca-Cu positivas para VPH16 y 5 controles.

6. Blancos moleculares en cáncer cervical asociados con la baja sobrevida.

Una manera de investigar si estos blancos moleculares están asociados con el cáncer cervical es durante la progresión en un estudio de sobrevida. Por lo tanto, se realizó un análisis de sobrevida con los datos de los ensayos de expresión obtenidos por qRT-PCR de los genes *PRC1*, *CDC20*, *CCNB2*, *CDKN3*, *NUSAP1*, *SYCP2*, *CDKN2A*, *PCNA* y *MKI67*, y además, de incluir también el estadio FIGO, en 42 pacientes con Ca-Cu positivas para el VPH16, a las cuales se les hizo un seguimiento durante 3.5 años después del diagnóstico inicial y tratamiento, para evaluar la progresión (Tabla V).

En este grupo de pacientes se incluyeron los estadios FIGO IB1 (n = 16), IB2 (n = 14), IIA (n = 1), IIB (n = 9), el IIIB (n = 2). La tasa de supervivencia en general fue alta en todas las muestras del 79.6% y para las distintas etapas del estadio FIGO IB1, IB2, IIA, IIB y IIIB, fue de 100%, 69,2%, 0%, 85.7% y 0%, respectivamente. Las diferencias fueron estadísticamente significativas ($p < 0.001$, prueba de Log-Rank test; Figura 19).

Por otra parte, cuando se realizó el análisis para los 9 genes analizados mediante las curvas de Kaplan-Meier, sólo el gen *CDKN3* se asoció con la baja sobrevida ($p = 0,004$, prueba de Log-Rank test; Figura 18B). La tasa de supervivencia en las pacientes con altos niveles de expresión del gen *CDKN3* (FC >15) fue 42,9% y el tiempo medio de la sobrevida fue de 33 meses. En contraste, los niveles más bajos de expresión del gen *CDKN3*, tuvieron una tasa de sobrevida en general del 87.5%.

Tanto el estadio FIGO y la expresión del gen *CDKN3* se analizaron individualmente en un modelo de riesgo relativo de Cox. Debido a las diferencias en el tamaño de muestras de cada una de las etapas FIGO, las muestras de las pacientes se reasignaron en 2 grupos, el primero grupo que incluyo los estadios FIGO IB1 y IB2 (n = 30) y el otro grupo FIGO IIA, IIB y IIIB (n = 12).

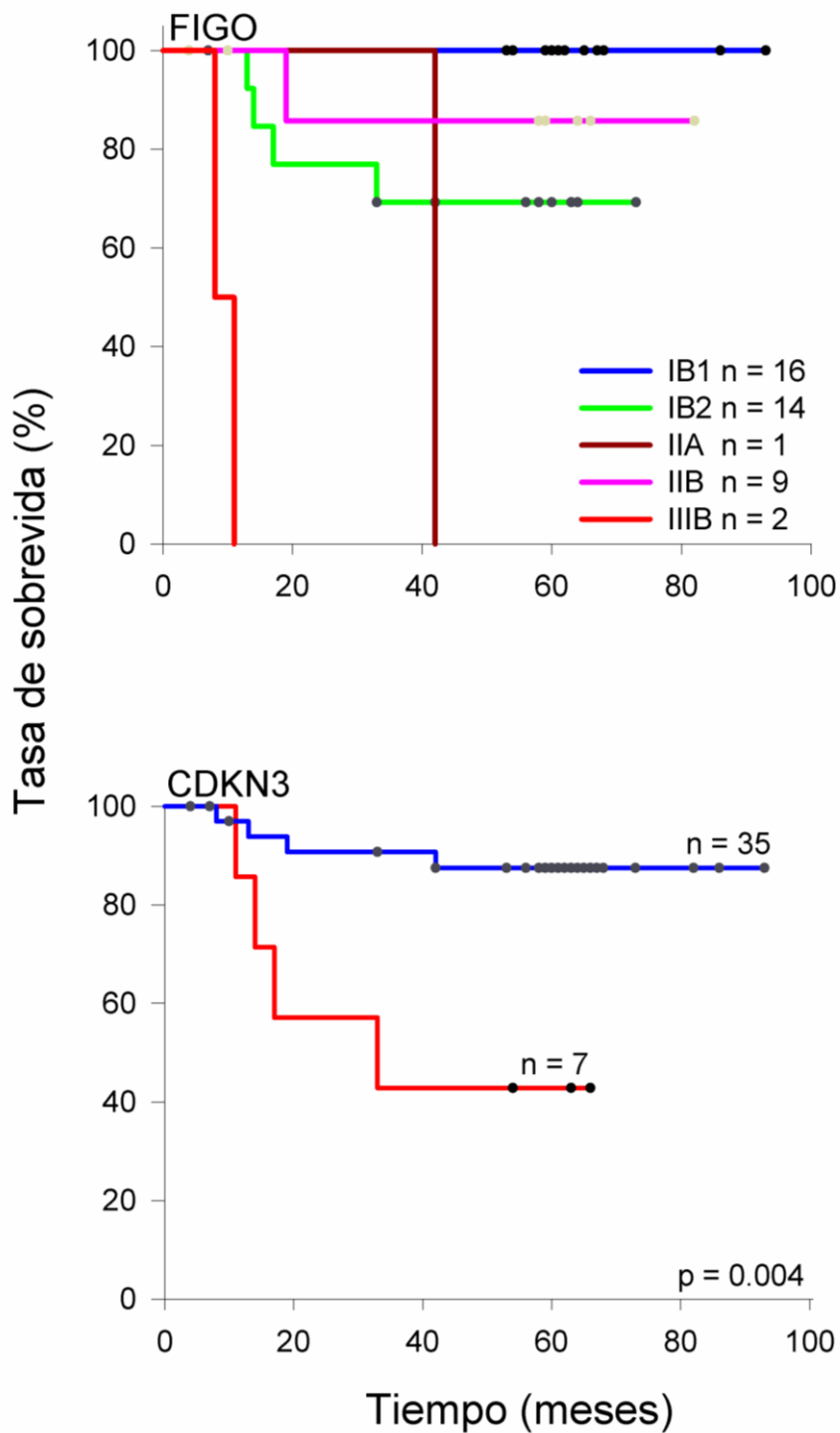


Figura 19. Análisis de sobrevida de mujeres con cáncer cervical según la clasificación del FIGO y la expresión de gen *CDKN3*. El análisis de las curvas de Kaplan-Meier muestran la clasificación de FIGO y del gen *CDKN3*. Las pacientes fueron seguidas por al menos 42 meses. Para la expresión génica, se compararon las pacientes con cáncer de mayor expresión (línea roja) y menor (línea azul) que doblan valores (ver materiales y métodos). Se calculó el valor de p comparando las curvas con el Log-rank test. Las pacientes que fallecieron se etiquetaron con puntos negros, pero cuatro de ellas murieron antes del período mínimo de seguimiento (42 meses).

Individualmente, el probabilidad de riesgo (hazard ratio, HR) del gen *CDKN3* fue 5.9 (95% IC 1.4–24.1, $p = 0.01$) y para la clasificación del grupo FIGO, 3.3 (95% IC 0.83–13.3, $p = 0.08$). Sin embargo, la falta de un significado estadístico en HR, para las clasificación FIGO podría explicarse por las diferencias en el tamaño de muestra y la relación invertida de sobrevida en las muestras de la etapa FIGO IB2 y IIB. Por otra parte, cuando estas 2 covariables fueron incluidas en el mismo modelo de probabilidad de riesgo, el gen *CDKN3* permaneció invariablemente significativo con un HR de 5.9 (IC 95%:1.4–23.8, $p = 0.01$). Estos resultados sugieren que el gen *CDKN3* puede ser un factor pronóstico de sobrevida que es independiente de la etapa FIGO en que se encuentre. Sin embargo, para confirmar y asegurar este resultado es necesario incluir un tamaño de muestra mayor.

7. Clasificación de los genes con expresión diferencial ente las muestras de cáncer y las muestras control.

Con la ayuda de la herramienta disponible en línea, para la anotación funcional de genes DAVID (<http://david.abcc.ncifcrf.gov>), se realizaron distintos análisis comparativos para poder identificar los procesos biológicos; esta herramienta de anotación funcional de DAVID fue utilizada bajo los parámetros de clasificación de mayor y mediano rigor, para catalogar los 997 genes expresados diferencialmente y conocer tanto en un valor medio y como en el de mayor astringencia los procesos biológicos en los que están involucrados estos genes. La base de datos realiza sus comparaciones directamente con la base de datos del genoma humano, los tres principales procesos con los valores de p más significativos, estos son: procesos asociados al ciclo celular, procesos del metabolismo del DNA y procesos relacionados con la regulación de la actividad de ligasa de la proteína ubiquitina (Tabla X).

Interesantemente, cuando el análisis se realizó con el mayor rigor permitido, se encontraron genes fuertemente asociados en cada uno de los grupos. Los principales grupos incluyen mitosis y fase M de mitosis, procesos del ciclo celular ocupando los lugares 1^{ro}, 2^{do} y 5^{to} respectivamente, (ver Tabla S2). Cabe destacar

que, entre los procesos del ciclo celular, ninguno excepto la fase M de mitosis que se enriqueció significativamente (Tabla X).

Tabla X. Grupos del análisis de anotación funcional DAVID en alta astringencia de los 997 genes expresados diferencialmente en cáncer cervical*.

Grupo	Valor de enriquecimiento	Proceso biológico	Número de Genes	Valor de p	FC
1	16.24	Mitosis	54	3.00E-17	3.8
		División nuclear	54	3.00E-17	3.8
		Fase M del ciclo celular mitótico	54	6.80E-17	3.7
		Fusión de organelos	54	1.90E-16	3.6
2	14.02	Regulación positiva de la actividad de proteína ligasa de ubiquitina durante el ciclo celular mitótico	28	3.60E-15	6.3
		Regulación positiva de la actividad ligasa	28	2.90E-14	5.9
3	5.71	Empalme de RNA, vía reacciones de transesterificación	28	1.90E-06	2.8
		Empalme nuclear de RNAm, vía espliceosoma	28	1.90E-06	2.8
4	3.66	Regulación positiva de la apoptosis	49	1.90E-04	1.7
5	3.03	Ciclo celular ensamble de huso mitótico	6	3.90E-04	8.3
		Regulación de la transición mitótica metafase/anafase	6	1.20E-02	4.2
6	2.81	Diferenciación de células mesenquimales	11	1.50E-03	3.3
7	2.34	Recombinación del ADN de células somáticas	7	3.70E-03	4.5
		Receptores inmunes de diversificación somática	7	6.90E-03	4.0
8	2.01	Proceso metabólico del ácido carboxílico	52	8.70E-02	1.4
		Proceso metabólico de las cetonas celulares	52	1.20E-02	1.4
9	1.96	Establecimiento de la localización del RNA	14	1.00E-02	2.2
		Transporte de RNA	14	1.00E-02	2.2
10	1.82	Glicolisis	10	3.00E-03	3.3
		Procesos catabólicos de la hexosa	10	3.50E-02	2.2
11	1.73	Formación del tubo epitelial embrionario	8	2.00E-02	2.8
12	1.73	Recombinación somática de los segmentos de genes de inmunoglobulinas	6	4.90E-03	5.1
		Producción de moléculas mediadoras de la respuesta inmune	6	4.90E-02	3.0
13	1.56	Integración de provirus	4	1.20E-02	7.6
		Integración del DNA	4	1.40E-01	3.1
14	1.54	Citotoxicidad mediada por células NK	4	1.20E-02	7.6
		Citotoxicidad mediada por leucocitos	4	2.40E-02	6.1
15	1.39	Regulación de la actividad de caspasa	11	3.20E-02	2.1
		Regulación de la actividad de endopeptidasa	11	5.30E-02	2.0
16	1.38	Recombinación somática de genes de inmunoglobulinas durante la respuesta inmune	4	3.90E-02	5.1
		Producción de inmunoglobulinas durante la repuesta inmune	4	4.90E-02	4.7

*El valor de enriquecimiento esta en Log_{10} del promedio del valor de p en los términos del agrupamiento. FC es el cambio en número de veces, es la relación entre la lista de genes seleccionados contra la base de datos de referencia del Genoma Humano "Human Gene Reference database".

Interesantemente, cuando el análisis se realizó con los 100 genes mejor clasificados, los 50 genes sobre expresados y los 50 genes que disminuyen su expresión, el grupo de mitosis fue la vía principal y la mas enriquecida 3.3 veces en comparación con los otros grupos alterados por todos los genes (n = 997) expresados diferencialmente (Tabla XI).

Tabla XI. Grupos del análisis de anotación funcional DAVID en alta astringencia de los 100 genes principales desregulados en cáncer cervical comparados con epitelio normal*.

Grupo	Valor de enriquecimiento	Proceso biológico	Número de Genes	Valor de p	FC
1	13.9	Mitosis	19	1.0E-14	12
		División nuclear	19	1.0E-14	12
		Fase M del ciclo celular mitótico	19	1.4E-14	12
		Fusión de organelos	19	2.0E-14	12
2	3.49	Regulación positiva de la actividad ligasa	6	2.10E-04	11
3	2.87	Regulación negativa de la actividad de la ligasa de ubiquitina	5	1.20E-03	11
4	2.83	Regulación positiva de la mitosis	4	6.70E-04	23
		Regulación positiva del ciclo celular	4	7.20E-03	10
5	2.75	Establecimiento de la localización de huso mitótico	3	1.30E-03	53
6	2.58	Proceso catabólico de proteínas dependiente del complejo proteasoma dependiente de ubiquitina promotor de anafase	5	1.10E-03	11
		Regulación de la actividad de la ligasa de la proteína ubiquitina durante el ciclo celular mitótico	5	1.50E-03	10
7	2.31	Regulación negativa de los procesos de modificación de proteínas	6	1.40E-03	7.2
		Regulación negativa de procesos de modificación de proteínas	6	9.70E-03	4.6
8	1.58	Regulación de procesos de modificación de proteínas	8	4.40E-03	3.9
		Regulación de procesos metabólicos de proteínas celulares	8	4.70E-02	2.4
9	1.53	Ensamble del nucleosoma	4	2.10E-02	6.8
		Ensamble de la cromatina	4	2.30E-02	6.6
10	1.42	Regulación de la actividad de proteína cinasa	7	3.30E-02	2.9
11	1.31	Regulación de la fosforilación	8	4.30E-02	2

*El valor de enriquecimiento esta en \log_{10} del promedio del valor de p en los términos del agrupamiento. FC es el cambio en número de veces, es la relación entre la lista de genes seleccionados contra la base de datos de referencia del Genoma Humano "Human Gene Reference database".

De esta forma el análisis de estos datos indicó, que los genes implicados en la mitosis no solo fueron los más enriquecidos, también son los más diferentes en términos de FC y Δ -score, comparando las muestras de Ca-Cu con las muestras control (círculos en rojo y naranja; Figura 10). De hecho, 11 de los 21 genes asociados con el Ca-Cu y que se validaron en este trabajo (*CCNB2*, *CDC20*,

PRC1, SYCP2, NUSAP1, CDKN3, CDC2, CKS2, MKI67, SMC4 y ZWINT), participan en la fase M del ciclo celular.

Una serie de análisis adicionales, se realizaron a los datos de genes expresados diferencialmente, con la ayuda del programa IPA, los resultados fueron similares a los obtenidos con la base de datos anterior DAVID, especialmente cuando el análisis de DAVID se realiza con rigor medio (TS2).

En coincidencia con el análisis realizado con DAVID, la vía de ubiquitinación de proteína fue la segunda vía canónica superior en el todo el conjunto de genes desregulados (Figura 20A) y cuando se analizaron los 100 genes mejor clasificados genes con las mayores diferencias, el resultado de este análisis da entonces como el principal proceso la participación de la mitosis a través de la cinasa semejante a polo (polo-like kinase) (Figura 20B).

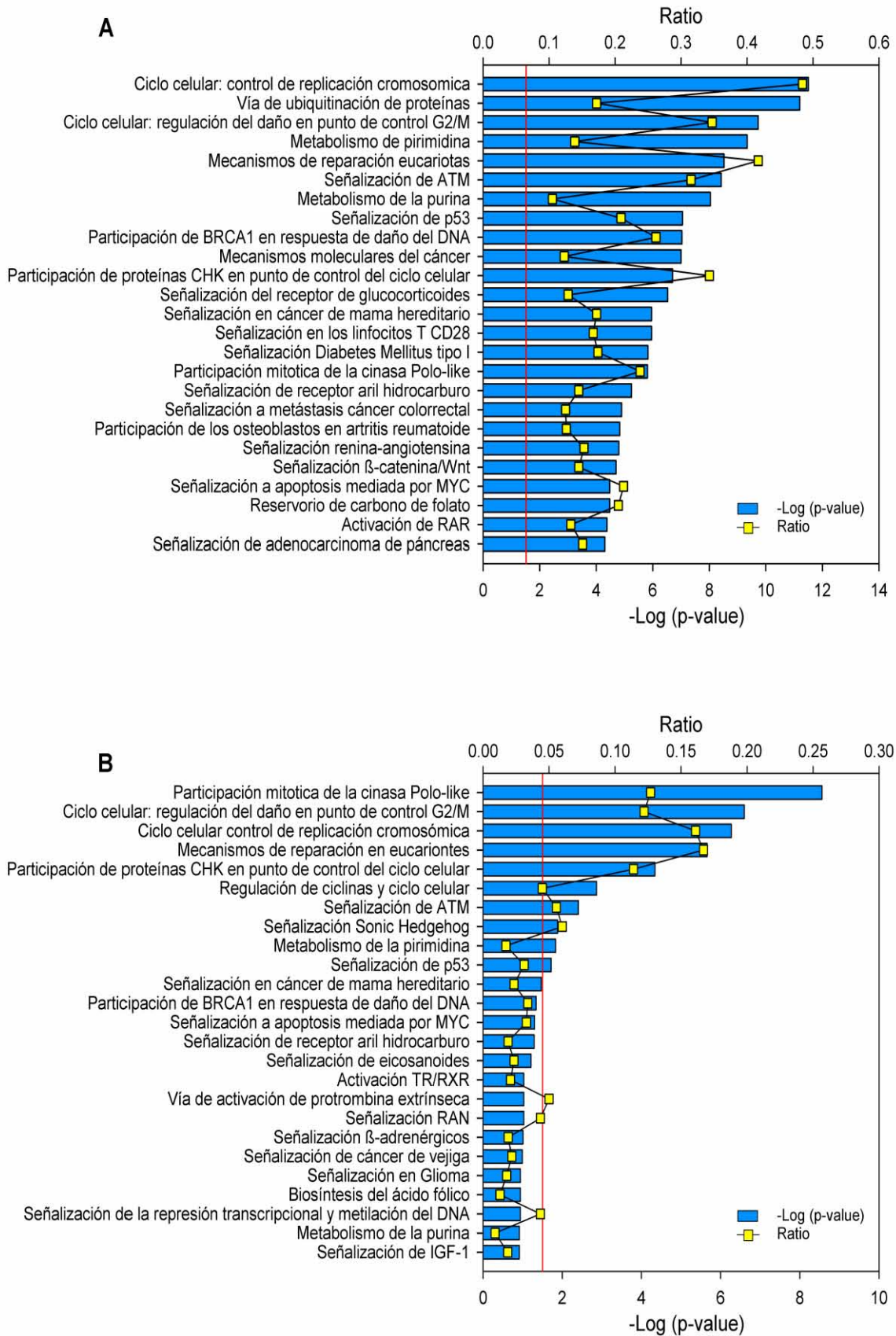


Figura 20. Vías canónicas en donde están implicados los genes desregulados. En A) se muestra las 25 principales vías canónicas identificadas para los 997 genes alterados entre la comparación de tumores y controles. En B) las 25 vías canónicas de los 100 genes mayormente desregulados, los 50 genes que más aumentaron la expresión y los 50 genes

más disminuyen. Las vías canónicas se identificaron mediante el análisis de vías con el programa Ingenuity System (IPA). Los valores en Log-2 (p -valor) en barras azules y la proporción que es indicada con puntos amarillos, esto fue calculado comparando el número de genes presentes en las vías en las bases de datos en comparación con las base de datos humana. El valor p se calculó con la prueba de χ^2 o prueba exacta de Fisher según corresponda y los valores de $-\text{Log}(p\text{-valor}) > 1.3$ (línea roja) que corresponden a valor de $p < 0.05$.

DISCUSIÓN.

En este trabajo se identificaron 6 genes (*PRC1*, *CCNB2*, *SYCP2*, *CDKN3*, *NUSAP1* y *CDC20*) asociados con el cáncer invasor y que pueden ser utilizados como marcadores o blancos terapéuticos. Mientras que los genes *PRC1*, *CCNB2* y *SYCP2* se asociaron en su mayoría con Ca-Cu, los genes *CDKN3*, *NUSAP1* y *CDC20*, se encontraron asociados también con las lesiones de alto grado NIC. En resultados recientes obtenidos en el laboratorio muestran de igual forma la expresión en líneas celulares derivadas del cáncer de cuello uterino (SiHa, CaSki, HeLa y Calo) para estos 4 genes, utilizando el microarreglo de HG-ST1.0., aunque la sobreexpresión no fue uniforme en las líneas de 4 celulares, en el análisis global para todos los genes se observó tener un aumento en la expresión, a excepción del gen *SYCP2*, este gen solo aumentó su expresión en las células CaSki y SiHa [79]. Estos hallazgos indican que estos marcadores son correctos para la predicción de cáncer de cuello uterino.

Los genes *CDKN3*, *NUSAP1* y *CDC20*, y en algo el gen *CDKN2A*, puede distinguir claramente las lesiones de alto grado NIC2/3 y los Ca-Cu, de las lesiones de bajo grado NIC1 y del cuello uterino normal; por lo tanto, podrían utilizarse como marcadores para las pruebas de detección temprana. Además, los resultados mostraron que la sobreexpresión del gen *CDKN3* se asoció con la baja supervivencia de pacientes con cáncer; por lo tanto, podría utilizarse también como un marcador de supervivencia. Los datos indicaron además que la sensibilidad y especificidad de los genes *CDKN3*, *NUSAP1* y *CDC20*, para identificar las lesiones de alto grado NIC2/3 y para los tumores Ca-Cu, fue del 91% y 93%, respectivamente.

Varios estudios han utilizado microarreglos para identificar genes asociados al cáncer de cuello uterino [29–37]. Sin embargo, la mayoría de estos estudios han incluido un pequeño número de tumores y controles, muestras heterogéneas positivas para diferentes tipos de VPH o no es determinado el tipo viral, y esencialmente el diseño de esos estudios no fue enfocado en identificar marcadores para la detección o marcadores de detección. Por lo tanto, los 23 marcadores validados en este trabajo, inicialmente se analizaron con MA en

muestras Ca-Cu positivas para un solo tipo viral VPH16 Ca-Cu, estos genes *CDKN3*, *NUSAP1*, *SMC4* y *WISP2*, no se han identificado previamente o los genes *CCNB2*, *CDC20*, *CKS2* y *RRM2* se han identificado raramente en otros estudios (Tabla XII). Por el contrario, más de la mitad o 40% [31-34], de estos genes se han identificados en otros estudios individuales y muchos de ellos han sido identificados en 3 o más estudios (*CDKN2A*, *MCM2*, *PCNA*, *RFC4*, *SYCP2* y *TYMS*) (Tabla XII) [37].

Tabla XII. Comparación de los 23 genes identificados y validados en este trabajo, con reportes de análisis con microarreglos documentados

Gen	Referencia										
	29	80	31 ^a	30	32 ^a	33	34	81	35	36	37
<i>CCNB2</i>		X ^b	X								
<i>CDC2</i>			X				X				
<i>CDC20</i>		X ^b	X				X ^c	X			
<i>CDKN2A</i>			X		X	X	X				X
<i>CDKN3</i>											
<i>CKS2</i>			X					X ^b			
<i>MCM2</i>			X			X	X				X
<i>MKI67</i>		X ^b	X								
<i>NUSAP1</i>								X ^b			
<i>PCNA</i>			X				X			X	X
<i>PRC1</i>			X			X		X ^b			
<i>RFC4</i>				X		X					X
<i>RRM2</i>		X ^b	X								
<i>SMC4</i>											
<i>SYCP2</i>							X			X	X
<i>TOP2A</i>	X	X ^b	X								
<i>TYMS</i>			X				X		X		X
<i>ZWINT</i>		X ^b	X								
<i>CFD</i>							X				
<i>EDN3</i>							X				
<i>NDN</i>							X				
<i>SLC18A2</i>											
<i>WISP2</i>											

a. El cambio en la expresión de los genes es indicado con una X que fue validado por RT-PCR.

b. Detectados en células HeLa.

c. Detectados en la línea celular NIKS-16, queratinocitos humanos inmortalizados.

El hecho es que los genes *CDKN3*, *CDC20* y *NUSAP1* mostraron una sobreexpresión en las muestras de lesiones de alto grado NIC2+ positivas para los virus VPH16 y para otros tipos virales, esto se demostró por qRT-PCR, sugiriendo que estos genes podrían considerarse como posibles marcadores para la detección de Ca-Cu sin importar el tipo viral que tengan. Hasta el momento, no se han reportado marcadores para la detección que tengan alta sensibilidad y especificidad. Además de la prueba de citología Papanicolaou y la citología de Base líquida (CBL), así como diferentes pruebas convencionales para detección de VPH, se utilizan para la pesquisa. La tecnología de captura de híbridos (HC2) es la herramienta más frecuentemente utilizada para pesquisa, particularmente para la medición de los virus de alto riesgo. Este método, fue aprobado por la FDA en Estados Unidos, tiene una sensibilidad promedio del 95% (rango, 62–98) para la detección de lesiones de alto grado y cáncer invasor. Sin embargo, esta metodología tiene baja especificidad, especialmente en mujeres jóvenes, ya que la mayoría de las infecciones no están asociada a lesiones neoplásicas [82-85]. En las mujeres con más de 30 años, la especificidad es mucho mayor; sin embargo, los datos son bastante variables entre los estudios y depende en parte de la prevalencia de VPH, en la población de estudio [86]. Además, en la mayoría de los estudios, el VPP es muy bajo, menos del 30%, lo que indica que sólo un porcentaje de mujeres infectadas tienen lesiones de alto grado.

Es importante hacer énfasis en cual es el principal valor de identificar biomarcadores en cáncer cervical y porque se tiene como objetivo el desarrollar procedimientos para pesquisa, es para mejorar la especificidad más que la sensibilidad en relación a la prueba del VPH [1]. Las pruebas de pesquisa primario con DNA del VPH, junto con la citología incrementan la especificidad de forma similar a lo que hace la citología convencional [43,82].

Sin embargo, el uso de estos procedimientos en países en vías de desarrollo como el nuestro, tiene aun varios problemas logísticos, ya sea porque el alto porcentaje de mujeres con pruebas de resultado positivo para el VPH, que no regresan a la prueba de la citología o debido a la manipulación de las muestras cuando esta toma de muestras para citología se realiza en pacientes de primera

vez. Adicionalmente además, ya que evita la automatización, y esto parece poco práctico.

El uso simultáneo de método ya establecido como el de captura de híbridos para los virus de alto riesgo y con un método molecular que distinga las lesiones de alto grado NIC2+ de las de bajo grado NIC1- puede incrementar la especificidad y el VPP, con las ventajas de ser más rápido y tener la posibilidad de automatizarse en comparación de la citología convencional.

De los marcadores asociados con Ca-Cu, hasta ahora el más estudiado es la proteína supresora de tumor p16 [1]. Esta proteína se acumula en el núcleo y el citoplasma de las células transformadas por los VPH de alto riesgo y generalmente se detecta por IH. Se relaciona con la cantidad de p16 con la severidad de la neoplasia cervical y se considera un marcador de CIN2+. De forma reciente p16 se ha implementado con éxito para la clasificación de enfermedades relacionadas con VPH.

En biopsias de tejido tomada con pinza o de cono, evaluadas para p16 por IH, se ha reportado reducir las diferencias intraobservadores cuando se compara con el diagnóstico obtenido con la tinción de H&E. Recientemente también p16 ha surgido como un complemento en el diagnóstico sensible y específico para detectar las lesiones NIC2+ en muestras de citología cervical [1]. Distintos estudios de esta proteína la han propuesto por tener una alta sensibilidad (80-95) para la detección de NIC2+; sin embargo la especificidad es baja para la citología (~50%) [86,87]. Esto es porque aproximadamente el 38% de las lesiones de bajo grado NIC, que están infectadas con los VPH de alto riesgo, expresan este marcador [86]. Sin embargo la baja especificidad de este marcador y la necesidad de un patólogo para la interpretación de los ensayos de IH, son las principales razones por las cuales p16 no se ha adoptado completamente para pesquisa inicial.

Recientemente el grupo de Wentzensen N., desarrollaron un método para detectar la proteína p16 en lisados celulares de muestras de exudados cervicales empleando la técnica de ELISA (Enzyme-Linked ImmunoSorbent Assay). La

sensibilidad de este método de ELISA para la identificación de lesiones de alto riesgo fue del 84%, y la especificidad fue del 87% [88]. De acuerdo con los datos obtenidos en nuestro estudio, la detección de mRNA de *CDKN2A* que codifica p16, tuvo una especificidad para pesquisa de los NIC2+ que fue superior 93%, (Tabla VII).

Otros dos marcadores nuevos identificados en este trabajo, los genes *CDKN3* y *NUSAP1* que en conjunto con *CDKN2A*, mostraron una alta especificidad (93%) y VPP (93,4%); por lo tanto, estos pueden ser buenos candidatos para su uso con HC2 como una estrategia de primera línea en un programa de pesquisa. El alcance de este estudio fue el realizar la evaluación para determinar si los niveles de mRNA de los nuevos genes identificados en muestras de Ca-Cu, permitirían la detección de lesiones de alto grado NIC o lesiones invasoras con una alta sensibilidad y especificidad. Sin embargo, el potencial de la sensibilidad reportada en el análisis probablemente se ha sobreestimado en comparación con lo encontrado en práctica clínica, como aquellas de las lesiones NIC2+ que tienen una mayor proporción de desarrollar cáncer cervical (el cual es fácil de identificar) con alguno de los métodos de detección.

En contraste, la especificidad observada puede ser subestimada, dada la gran proporción que había de NIC1- y NIC1. Por lo tanto, estos resultados no se reportan como datos concluyentes sobre la sensibilidad, la especificidad y los valores predictivos de los ensayos. Posteriormente es necesario realizar un mayor número de estudios para determinar los niveles de mRNA, de los genes *CDKN3*, *NUSAP1* y *CDC20* o incluso evaluarse a nivel de proteína en muestras cervicales de una población mayor para detección temprana y obtener más información sobre los valores predictivos y así poder definir un factor de ajuste óptimo entre la sensibilidad y especificidad para la detección de las lesiones CIN2+.

Los genes *PRC1*, *CCNB2* y *SYCP2* son marcadores asociados exclusivamente con el cáncer cervical invasor. Junto con los genes *NUSAP1*, *CDKN3*, y *CDC20*, estos genes representan blancos potenciales específicos para el tratamiento de Ca-Cu avanzado, particularmente el gen *CDKN3*, se encontró estar asociado con la baja sobrevida. Estos genes codifican para proteínas implicadas en el ciclo

celular, específicamente en la fase M (mitosis y citocinesis). De acuerdo con los datos obtenidos en los ensayos de IH, aproximadamente el 30% de las células tumorales en Ca-Cu pueden estar en la fase M.

La participación de estos genes se ha visto en el control de anafase, segregación cromosómica, en el control de la entrada/salida de mitosis. En tiempo cuando la activación de cinasas dependientes de ciclinas (Cdks) dirige a la célula a mitosis, la salida mitótica depende en una inhibición de la actividad de Cdks, principalmente a través de la degradación de ciclinas mitóticas por el complejo promotor de la anafase (APC/C) y la acumulación de proteínas inhibitorias de Cdk, la desfosforilación de proteínas fosforiladas por Cdks. Cuatro de las seis proteínas validadas (CCNB2, CDC20, CDKN3, PRC1) en este trabajo, son esenciales para ese proceso.

La ciclina B2 (CCNB2), la ciclina semejante a B1 (CCNB1), se unen a CDK1 (CDC2) para formar el complejo M-CDK, el cual es esencial para el control del ciclo celular y la transición de G2/M. Sin embargo cuando el complejo de ciclina B1-CDK1 causan condensación cromosómica, reorganización de microtubulos y desensamble de la lámina nuclear y el aparato de Golgi durante la mitosis [89,90]. En adición con esos datos, la ciclina B2 fue localizada exclusivamente en el citoplasma de los Ca-Cu examinados en este artículo (Figura 16).

Interesantemente, la expresión de la ciclina B1 en esos tumores no difiere que en las muestras control (TS1). Esa ciclina es degradada por el complejo APC/C, como una vía regulatoria de la transición de metafase-anafase [91], para seguir la progresión de mitosis de metafase a anafase [91]. El gen *CCNB2* se ha asociado poco con el cáncer cervical [31]; sin embargo se ha reportado estar asociado en otros tipos de cáncer, por ejemplo este gen es sobre expresado en cáncer de colon [92], pulmón y tracto digestivo [93].

Un incremento en la cantidad de CDC20, que es una proteína clave regulatoria del complejo APC/C durante la anafase, y esto puede explicar la ausencia de ciclina B1. Por otra parte CDC20, junto con UBE2C (también conocida como UBCH10), la cual también se incrementa en Ca-Cu (visto en este estudio TS1), es

requerida para la actividad completa de la ligasa de ubiquitina del complejo APC/C y puede conferir la actividad de sustrato sobre el complejo. La proteína CDC20 es regulada negativamente por MAD2L1 y BUB1B (también conocida como BUBR1). En la metafase, el complejo terciario de MAD2L1-CDC20-APC/C es inactivo, mientras que en la etapa de anafase el complejo binario de CDC20-APC/C está activo en la degradación de sustratos. Interesantemente, los transcritos de MAD2L1 y BUB1B, también se incrementan en Ca-Cu (TS1), sugiriendo que corresponden a proteínas que pueden incrementarse y prevenir la activación de APC/C. Sin embargo parte de la proteína de CDC20 puede permanecer libre para asociarse y activar el complejo APC/C, esto se ha demostrado en ensayos de células transfectadas que expresan las proteínas E6/E7 [91].

El gen *CDC20* se ha reportado estar sobre expresado en cáncer de pulmón, páncreas y gástrico [94], así como también en Ca-Cu [31,81]. *CDKN3* es una proteína con actividad dual fosfatasa del grupo de fosfatasas de Cdc14 que interactúa con CDK1 (CDC2) y así inhibe su actividad [95,96]. *CDKN3* y otras fosfatasas como Cdc14, no han sido bien estudiadas aun; sin embargo, estas se han visto ser esenciales por la actividad antagonista de Cdk en la mitosis tardía, llevando las células a salir de la mitosis en la telofase. La regulación de la citocinesis puede ser una función conservada de la fosfatasa Cdc14. Siguiendo la sobre expresión de *CDKN3* se ha asociado con la inhibición de la proliferación celular en líneas derivadas de cáncer de colon [97], también se ha reportado estar sobre expresada en cáncer de mama, próstata y pulmón [98-100]. De acuerdo con los datos obtenidos en este estudio, el gen *CDKN3*, junto con otros genes, se han asociados con la baja supervivencia de pacientes con adenocarcinoma de pulmón [98]. Este es el primer reporte en el cual el gen *CDKN3* es asociado con el cáncer cervical.

El gen *PRC1* está involucrado en la citocinesis y es esencial para contrarlar la formación espacio-temporal de la zona media y la completa citocinesis [101,102]. Este gen además es requerido para la localización en el uso central y cuerpo medio por los miembros de la familia de cinecinas 14 (KIF14) [103] y la cinasa semejante a polo 1 (PLK1) [104]. La supresión de *PRC1* bloquea la división

celular. La transcripción de PRC1 es reprimido por p53 y es una de las rutas por las cuales p53 frena el ciclo celular y el arresto en el punto de control del ciclo G2/M [105]. La oncoproteína E6 del virus VPH16 induce la degradación de p53 en los proteosomas, esto probablemente es porque en carcinomas cervicales *PRC1* este también sobre expresado vía este mecanismo. Este gen se ha reportado estar asociado con cáncer de hígado [106] y Ca-Cu [31,33].

La proteína NUSAP1 es una proteína nucleolar asociada al nucleolar-huso-asociado que tiene una participación en la organización del huso en los microtubulos. Este gen no se ha descrito su asociación con Ca-Cu, pero se ha encontrado estar sobre expresado en cáncer de mama y melanoma [107]. Por otra parte, el gen *SYCP2* es un componente principal del complejo sinaptonémico. Este complejo promueve que se rompa la doble cadena (DSB) y son reparados por recombinación homóloga vía meiosis [108]. Los altos niveles de expresión de gen *SYCP2* en Ca-Cu, evaluados en este trabajo sugieren que el rompimiento de la doble cadena (DSB) es muy común en algunas muestras de Ca-Cu y que el gen *SYCP2* puede estar involucrado en la reparación de DSB, por estimulación de la vía de recombinación homóloga. Interesantemente, este gen se ha encontrado por otros trabajos estar sobre expresado en Ca-Cu [36,37] y en carcinomas orofaríngeos de células escamosas positivos para VPH16, pero no en carcinomas negativos a VPH [109].

Es este estudio el ciclo celular es el proceso más alterado en Ca-Cu y es el principal clasificado en todos los artículos de Ca-Cu en los cuales se han analizado los procesos biológicos [37]. De igual forma en este trabajo, cuando los genes fueron de analizados con la ayuda la base de datos en línea DAVID, en un nivel medio, el proceso de ciclo celular se mostro estar más enriquecido y además ser el numero 1 de la lista (TS2). Sin embargo, cuando con los datos obtenidos en el laboratorio se realizo el análisis pero ahora con el mayor rigor posible, el proceso de fase-M fue el más enriquecido, sugiriendo que la fase M es la principal fase del ciclo celular alterada en Ca-Cu. Estos hallazgos son consistentes con las alteraciones en el ciclo celular y mitosis causadas por el VPH que se ha demostrado *in vivo* por distintos estudios. [80, 81, 110] y esto ha sido poco estudiado en Ca-Cu [81].

Las oncoproteínas de E6 y E7 de los virus de alto riesgo inducen numerosos defectos mitóticos, incluyendo mitosis multipolar, segregación cromosómica, ruptura de la anafase y aneuploidia. Algunas células con mitosis anormal son normalmente dirigidas para la muerte celular, las proteínas E6 y E7 actúan cooperativamente llevando esas células con centrosomas anormales para acumularlos y así, relajando la respuesta del punto de control G2/M y la inhibición de la señalización para apoptosis [110].

En adición a esos datos, el análisis de vías canónicas dio como resultado: regulación del punto de control DNA, regulación de proteínas de CHK en el ciclo células encontradas en la segunda posición y en la quinta de las vías canónicas alteradas en Ca-Cu. En otros trabajos, las oncoproteínas E6 y E7 inducen mecanismos para evitar el punto de control de la mitosis. Para los genes E6/E7 se ha demostrado que inducen la sobre expresión de CDC20 y UBCH10, las cuales activan el complejo de la ligasa de ubiquitina APC/C [91]. El enriquecimiento de actividad por la regulación positiva de proteína ligasa de ubiquitina durante el ciclo celular mitótico se encontró en Ca-Cu (Tabla X), es complementario con esos resultados encontrados *in vivo*.

Asimismo, la inhibición de la mitosis es una estrategia bien conocida para combatir el cáncer. Distintos fármacos que perturban el proceso de división celular se han propuesto por ser efectivos en las terapias anti-cáncer. Ejemplos bien conocidos y documentados de esos fármacos que alteran la formación del uso mitótico, tal como los taxanos y los alcaloides de la vinca. Sin embargo, ellos tienen índices terapéuticos remarcablemente bajos y pocas ventanas terapéuticas. Su eficacia es restringida porque ellos también alteran la red de microtubulos de las células que no están en división. Causando efectos neurotóxicos y afectando la función endotelial de la célula. Para resolver ese problema, para una nueva generación de agentes antimitóticos se ha desarrollado con blancos como las cinecinas y las cinasas con participaciones específicas en la mitosis, tal como *KIF11*, *PLK1* y aurora cinasa A (*AURKA*) [104].

Interesantemente, en este estudio los transcritos para esos 3 genes, también se encontraron en los genes sobre expresados en Ca-Cu (TS1). Los genes *AURKA*,

KIF1 y *PLK1*, en este trabajo se ubicaron en los lugares 19, 72 y 263, respectivamente. Además, esos fármacos podrían también ser probados en el tratamiento de cáncer cervical. En los resultados obtenidos en este trabajo, donde se identificaron y se validaron los genes expresados diferencialmente, en especial para los genes *CDKN3*, *CDC20* y *SYCP2*, comparados con el control el FC fue alto, indicando el potencial que tienen estos genes como blancos para la terapia en cáncer cervical. Sin embargo es necesario y de continuidad para el trabajo, demostrar que estos genes pueden ser indispensables para el crecimiento del tumor.

CONCLUSIONES.

- En este trabajo se identificaron 6 genes (*PRC1*, *CCNB2*, *SYCP2*, *CDKN3*, *NUSAP1* y *CDC20*) asociados con el cáncer invasor
- Los genes identificados se encuentran desregulados tanto en muestras positivas para el VPH16 como para otros tipos virales.
- Es el primer reporte de la identificación de los genes *CDKN3*, *NUSAP1*, *SMC4* y *WISP2*, en cáncer cervical.
- Los genes *CCNB2*, *CDC20*, *CKS2* y *RRM2* han sido escasamente reportados en Ca-Cu por otros estudios.
- Estos genes se evaluaron y se demostró su expresión en lesiones pre-invasoras y no en tejido normal, por lo que pueden ser considerados como candidatos para la detección temprana.
- Estos genes tienen alto valor de sensibilidad y especificidad para detectar lesiones de alto grado.
- Este estudio identifica a estos genes y permite ver el potencial que tienen como posibles marcadores blanco para la terapia en cáncer cervical.
- El estudio en conjunto o de forma individual, de estos genes *PRC1*, *CCNB2*, *SYCP2*, *CDKN3*, *NUSAP1* y *CDC20*, presentan una oportunidad para desarrollar alguna herramienta que favorezca la detección temprana.

REFERENCIAS.

- 1) Hwang, S. J., & Shroyer, K. R. (2011). Biomarkers of cervical dysplasia and carcinoma. *Journal of Oncology*, 2012.
- 2) Ferlay, J., Shin, H. R., Bray, F., Forman, D., Mathers, C., & Parkin, D. M. (2011). GLOBOCAN 2008 v1. 2, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 10 [Internet]. International Agency for Research on Cancer, Lyon, France.
- 3) Parkin, D. M., Bray, F., Ferlay, J., & Pisani, P. (2005). Global cancer statistics, 2002. *CA: a cancer journal for clinicians*, 55(2), 74-108.
- 4) Lazcano-Ponce, E., Herrero, R., Munoz, N., Cruz, A., Shah, K. V., Alonso, P., ... & Hernandez, M. (2001). Epidemiology of HPV infection among Mexican women with normal cervical cytology. *International Journal of Cancer*, 91(3), 412-420.
- 5) Bosch, F. X., Lorincz, A., Munoz, N., Meijer, C. J. L. M., & Shah, K. V. (2002). The causal relation between human papillomavirus and cervical cancer. *Journal of clinical pathology*, 55(4), 244-265.
- 6) Schiffman, M., Wentzensen, N., Wacholder, S., Kinney, W., Gage, J. C., & Castle, P. E. (2011). Human papillomavirus testing in the prevention of cervical cancer. *Journal of the National Cancer Institute*, 103(5), 368-383.
- 7) Londesborough, P., Ho, L., Terry, G., Cuzick, J., Wheeler, C., & Singer, A. (1996). Human papillomavirus genotype as a predictor of persistence and development of high-grade lesions in women with minor cervical abnormalities. *International Journal of Cancer*, 69(5), 364-368.
- 8) Berumen, J., Casas, L., Segura, E., Amezcua, J. L., & 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(5), 640-645.
- 9) Clifford, G. M., Gallus, S., Herrero, R., Munoz, N., Snijders, P. J. F., Vaccarella, S., ... & Franceschi, S. (2005). Worldwide distribution of human papillomavirus types in cytologically normal women in the International Agency for Research on Cancer HPV prevalence surveys: a pooled analysis. *The Lancet*, 366(9490), 991-998.
- 10) Baseman, J. G., & Koutsky, L. A. (2005). The epidemiology of human papillomavirus infections. *Journal of Clinical Virology*, 32, 16-24.
- 11) Woodman, C. B., Collins, S. I., & Young, L. S. (2007). The natural history of cervical HPV infection: unresolved issues. *Nature Reviews Cancer*, 7(1), 11-22.
- 12) Schiffman, M., Castle, P. E., Jeronimo, J., Rodriguez, A. C., & Wacholder, S. (2007). Human papillomavirus and cervical cancer. *The Lancet*, 370(9590), 890-907.
- 13) Burd, E. M. (2003). Human papillomavirus and cervical cancer. *Clinical microbiology reviews*, 16(1), 1-17.
- 14) Hernández Avila, M., Lazcano Ponce, E. C., Berumen Campos, J., Cruz Valdéz, A., Alonso de Ruíz, P. P., & González Lira, G. (1997). Human papilloma virus 16-18 infection and cervical cancer in Mexico: A case-control study. *Archives of Medical Research*, 28(2), 265-71.
- 15) Östör, A. G. (1993). Natural history of cervical intraepithelial neoplasia: a critical review. *International Journal of Gynecologic Pathology*, 12(2), 186.
- 16) Walboomers, J. M., Jacobs, M. V., Manos, M. M., Bosch, F. X., Kummer, J. A., Shah, K. V., ... & Munoz, N. (1999). Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *The Journal of pathology*, 189(1), 12-19.

- 17) Munoz, N., Bosch, F. X., de Sanjose, S., Herrero, R., Castellsagué, X., Shah, K. V., ... & Meijer, C. J. (2003). Epidemiologic classification of human papillomavirus types associated with cervical cancer. *New England Journal of Medicine*, 348(6), 518-527.
- 18) Bosch, F. X., & Muñoz, N. (2002). The viral etiology of cervical cancer. *Virus research*, 89(2), 183.
- 19) Ciapponi, A., Bardach, A., Glujovsky, D., Gibbons, L., & Picconi, M. A. (2011). Type-specific HPV prevalence in cervical cancer and high-grade lesions in Latin America and the Caribbean: systematic review and meta-analysis. *PloS one*, 6(10), e25493
- 20) 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 Research*, 58(4), 829-833.
- 21) Xi, L. F., Koutsky, L. A., Galloway, D. A., Kiviat, N. B., Kuypers, J., Hughes, J. P., ... & Holmes, K. K. (1997). Genomic variation of human papillomavirus type 16 and risk for high grade cervical intraepithelial neoplasia. *Journal of the National Cancer Institute*, 89(11), 796-802.
- 22) Lizano, M., Guido, M. C., García-Carrancá, A., Berumen, J., & Casas, L. (1997). Association between human papillomavirus type 18 variants and histopathology of cervical cancer. *Journal of the National Cancer Institute*, 89(16), 1227-1231.
- 23) Yamada, T., Manos, M. M., Peto, J., Greer, C. E., Munoz, N., Bosch, F. X., & Wheeler, C. M. (1997). Human papillomavirus type 16 sequence variation in cervical cancers: a worldwide perspective. *Journal of virology*, 71(3), 2463-2472.
- 24) Ho, L., Chan, S. Y., Burk, R. D., Das, B. C., Fujinaga, K., Icenogle, J. P., ... & Mavromara-Nazos, P. (1993). The genetic drift of human papillomavirus type 16 is a means of reconstructing prehistoric viral spread and the movement of ancient human populations. *Journal of Virology*, 67(11), 6413-6423.
- 25) Prendiville W, Davies P (2004). The health professional's VPH handbook. Abingdon, Taylor and Francis Group Ed, New York, NY. Paginas 9-73.
- 26) Nees, M., Geoghegan, J. M., Hyman, T., Frank, S., Miller, L., & Woodworth, C. D. (2001). Papillomavirus type 16 oncogenes downregulate expression of interferon-responsive genes and upregulate proliferation-associated and NF-κB-responsive genes in cervical keratinocytes. *Journal of virology*, 75(9), 4283-4296.
- 27) Boulet, G. A., Horvath, C. A., Depuydt, C. E., & Bogers, J. J. (2010). Biomarkers in cervical screening: quantitative reverse transcriptase PCR analysis of P16INK4a expression. *European Journal of Cancer Prevention*, 19(1), 35-41.
- 28) Gupta, N., Srinivasan, R., & Rajwanshi, A. (2010). Functional biomarkers in cervical precancer: an overview. *Diagnostic cytopathology*, 38(8), 618-623.
- 29) Chen, Y., Miller, C., Mosher, R., Zhao, X., Deeds, J., Morrissey, M., ... & Schlegel, R. (2003). Identification of cervical cancer markers by cDNA and tissue microarrays. *Cancer research*, 63(8), 1927-1935.
- 30) Narayan, G., Bourdon, V., Chaganti, S., Arias-Pulido, H., Nandula, S. V., Rao, P. H., ... & Murty, V. V. (2007). Gene dosage alterations revealed by cDNA microarray analysis in cervical cancer: identification of candidate amplified and overexpressed genes. *Genes, Chromosomes and Cancer*, 46(4), 373-384.
- 31) Rosty, C., Sheffer, M., Tsafirir, D., Stransky, N., Tsafirir, I., Peter, M., ... & Sastre-Garau, X. (2005). Identification of a proliferation gene cluster associated with HPV E6/E7 expression level and viral DNA load in invasive cervical carcinoma. *Oncogene*, 24(47), 7094-7104.

- 32) Gius, D., Funk, M. C., Chuang, E. Y., Feng, S., Huettner, P. C., Nguyen, L., ... & Rader, J. S. (2007). Profiling microdissected epithelium and stroma to model genomic signatures for cervical carcinogenesis accommodating for covariates. *Cancer research*, 67(15), 7113-7123.
- 33) Zhai, Y., Kuick, R., Nan, B., Ota, I., Weiss, S. J., Trimble, C. L., ... & Cho, K. R. (2007). Gene expression analysis of preinvasive and invasive cervical squamous cell carcinomas identifies HOXC10 as a key mediator of invasion. *Cancer research*, 67(21), 10163-10172.
- 34) Pyeon, D., Newton, M. A., Lambert, P. F., Den Boon, J. A., Sengupta, S., Marsit, C. J., ... & Ahlquist, P. (2007). Fundamental Differences in Cell Cycle Deregulation in Human Papillomavirus–Positive and Human Papillomavirus–Negative Head/Neck and Cervical Cancers. *Cancer research*, 67(10), 4605-4619.
- 35) Biewenga, P., Buist, M. R., Moerland, P. D., van Themaat, E. V. L., van Kampen, A. H., ten Kate, F. J., & Baas, F. (2008). Gene expression in early stage cervical cancer. *Gynecologic oncology*, 108(3), 520-526.
- 36) Scotto, L., Narayan, G., Nandula, S. V., Arias-Pulido, H., Subramaniam, S., Schneider, A., ... & Murty, V. V. (2008). Identification of copy number gain and overexpressed genes on chromosome arm 20q by an integrative genomic approach in cervical cancer: potential role in progression. *Genes, Chromosomes and Cancer*, 47(9), 755-765.
- 37) Buitrago-Pérez, Á., Garaulet, G., Vázquez-Carballo, A., Paramio, J. M., & García-Escudero, R. (2009). Molecular signature of HPV-induced carcinogenesis: pRb, p53 and gene expression profiling. *Current genomics*, 10(1), 26.
- 38) Mehta, A. M., Jordanova, E. S., Kenter, G. G., Ferrone, S., & Fleuren, G. J. (2008). Association of antigen processing machinery and HLA class I defects with clinicopathological outcome in cervical carcinoma. *Cancer Immunology, Immunotherapy*, 57(2), 197-206.
- 39) de Sanjose, S., Quint, W. G., Alemany, L., Geraets, D. T., Klaustermeier, J. E., Lloveras, B., ... & Puras, A. (2010). Human papillomavirus genotype attribution in invasive cervical cancer: a retrospective cross-sectional worldwide study. *The lancet oncology*, 11(11), 1048-1056.
- 40) Kjaer, S. K., Sigurdsson, K., Iversen, O. E., Hernandez-Avila, M., Wheeler, C. M., Perez, G., ... & Haupt, R. M. (2009). A pooled analysis of continued prophylactic efficacy of quadrivalent human papillomavirus (Types 6/11/16/18) vaccine against high-grade cervical and external genital lesions. *Cancer Prevention Research*, 2(10), 868-878.
- 41) Lehtinen, M., Paavonen, J., Wheeler, C. M., Jaisamrarn, U., Garland, S. M., Castellsagué, X., ... & Dubin, G. (2012). Overall efficacy of HPV-16/18 AS04-adjuvanted vaccine against grade 3 or greater cervical intraepithelial neoplasia: 4-year end-of-study analysis of the randomised, double-blind PATRICIA trial. *The lancet oncology*, 13(1), 89-99.
- 42) Romanowski, B. (2011). Long term protection against cervical infection with the human papillomavirus: review of currently available vaccines. *Human Vaccines*, 7(2), 161-169.
- 43) Cuzick, J. (2010). Long-term cervical cancer prevention strategies across the globe. *Gynecologic oncology*, 117(2), S11-S14.
- 44) Markowitz, L. E., Dunne, E. F., Saraiya, M., Lawson, H. W., Chesson, H., & Unger, E. R. (2007). Quadrivalent human papillomavirus vaccine. *Morbidity and Mortality weekly report*, 56(RR02), 1-24.
- 45) Natunen, K., Lehtinen, J., Namujju, P., Sellors, J., & Lehtinen, M. (2011). Aspects of prophylactic vaccination against cervical cancer and other human papillomavirus-related cancers in developing countries. *Infectious diseases in obstetrics and gynecology*, 2011.

- 46) Morantz, C. (2004). ACS guidelines for early detection of cancer. *American family physician*, 69(8), 2013.
- 47) McCrory, D. C., Matchar, D. B., Bastian, L., Datta, S., Hasselblad, V., Hickey, J., ... & Nanda, K. (1999). Evaluation of cervical cytology. *Evidence report/technology assessment (Summary)*, (5), 1.
- 48) Wright Jr, T. C. (2007). Cervical cancer screening in the 21st century: is it time to retire the PAP smear?. *Clinical obstetrics and gynecology*, 50(2), 313.
- 49) Belinson, S. E., & Belinson, J. L. (2010). Human Papillomavirus DNA Testing for Cervical Cancer Screening. *Molecular Diagnosis & Therapy*, 14(4), 215-222.
- 50) Kulasingam, S. L., Hughes, J. P., Kiviat, N. B., Mao, C., Weiss, N. S., Kuypers, J. M., & Koutsky, L. A. (2002). Evaluation of human papillomavirus testing in primary screening for cervical abnormalities. *JAMA: the journal of the American Medical Association*, 288(14), 1749-1757.
- 51) Coste, J., Cochand-Priollet, B., de Cremoux, P., Galès, C. L., Isabelle, C., Vincent, M., ... & Vielh, P. (2003). Cross sectional study of conventional cervical smear, monolayer cytology, and human papillomavirus DNA testing for cervical cancer screening. *Bmj*, 326(7392), 733.
- 52) Petry, K. U., Menton, S., Menton, M., van Loenen-Frosch, F., de Carvalho Gomes, H., Holz, B., ... & Iftner, T. (2003). Inclusion of HPV testing in routine cervical cancer screening for women above 29 years in Germany: results for 8466 patients. *British journal of cancer*, 88(10), 1570-1577.
- 53) Bigras, G., & De Marval, F. (2005). The probability for a Pap test to be abnormal is directly proportional to HPV viral load: results from a Swiss study comparing HPV testing and liquid-based cytology to detect cervical cancer precursors in 13 842 women. *British journal of cancer*, 93(5), 575-581.
- 54) Mayrand, M. H., Duarte-Franco, E., Rodrigues, I., Walter, S. D., Hanley, J., Ferenczy, A., ... & Franco, E. L. (2007). Human papillomavirus DNA versus Papanicolaou screening tests for cervical cancer. *New England Journal of Medicine*, 357(16), 1579-1588.
- 55) Cárdenas-Turanzas, M., Nogueras-Gonzalez, G. M., Scheurer, M. E., Adler-Storthz, K., Benedet, J. L., Beck, J. R., ... & Cantor, S. B. (2008). The performance of human papillomavirus high-risk DNA testing in the screening and diagnostic settings. *Cancer Epidemiology Biomarkers & Prevention*, 17(10), 2865-2871.
- 56) Del Mistro, A., Frayle-Salamanca, H., Trevisan, R., Matteucci, M., Pinarello, A., Zambenedetti, P., ... & Minucci, D. (2010). Triage of women with atypical squamous cells of undetermined significance (ASC-US): results of an Italian multicentric study. *Gynecologic oncology*, 117(1), 77-81.
- 57) Lazcano-Ponce, E., Lörintz, A. T., Salmerón, J., Fernández, I., Cruz, A., Hernández, P., ... & Hernández-Ávila, M. (2010). A pilot study of HPV DNA and cytology testing in 50,159 women in the routine Mexican Social
- 58) Leinonen, M., Nieminen, P., Kotaniemi-Talonen, L., Malila, N., Tarkkanen, J., Laurila, P., & Anttila, A. (2009). Age-specific evaluation of primary human papillomavirus screening vs conventional cytology in a randomized setting. *Journal of the National Cancer Institute*, 101(23), 1612-1623.
- 59) Andrae, B., Andersson, T. M., Lambert, P. C., Kemetli, L., Silfverdal, L., Strander, B., ... & Sparén, P. (2012). Screening and cervical cancer cure: population based cohort study. *BMJ: British Medical Journal*, 344.

- 60) Torti, D., & Trusolino, L. (2011). Oncogene addiction as a foundational rationale for targeted anti-cancer therapy: promises and perils. *EMBO molecular medicine*, 3(11), 623-636.
- 61) Knight, Z. A., Lin, H., & Shokat, K. M. (2010). Targeting the cancer kinome through polypharmacology. *Nature Reviews Cancer*, 10(2), 130-137.
- 62) Saxena, R., & Dwivedi, A. (2012). ErbB family receptor inhibitors as therapeutic agents in breast cancer: current status and future clinical perspective. *Medicinal research reviews*, 32(1), 166-215.
- 63) Yarden, Y., & Pines, G. (2012). The ERBB network: at last, cancer therapy meets systems biology. *Nature Reviews Cancer*, 12(8), 553-563.
- 64) Pecorelli, S. (2010). Corrigendum to revised FIGO staging for carcinoma of the vulva, cervix, and endometrium. *International Journal of Gynaecology and Obstetrics*, 108(2).
- 65) Manos, M. M., Ting, Y., Wright, D. K., Lewis, A. J., Broker, T. R., & Wolinsky, S. M. (1989). Use of polymerase chain reaction amplification for the detection of genital human papillomaviruses. *Cancer cells*, 7(17), 209-14.
- 66) Van den Brule, A. J., Meijer, C. J., Bakels, V., Kenemans, P., & Walboomers, J. M. (1990). Rapid detection of human papillomavirus in cervical scrapes by combined general primer-mediated and type-specific polymerase chain reaction. *Journal of clinical microbiology*, 28(12), 2739-2743.
- 67) Yoshikawa, H., Kawana, T., Kitagawa, K., Mizuno, M., Yoshikura, H., & Iwamoto, A. (1991). Detection and typing of multiple genital human papillomaviruses by DNA amplification with consensus primers. *Cancer Science*, 82(5), 524-531.
- 68) Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J., & Lopez, R. (2010). A new bioinformatics analysis tools framework at EMBL–EBI. *Nucleic acids research*, 38(suppl 2), W695-W699.
- 69) Blazejczyk, M., Miron, M., & Nadon, R. (2007). FlexArray: a statistical data analysis software for gene expression microarrays. *Genome Quebec, Montreal, Canada, URL <http://genomequebec.mcgill.ca/FlexArray>*.
- 70) Tusher, V. G., Tibshirani, R., & Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. *Proceedings of the National Academy of Sciences*, 98(9), 5116-5121.
- 71) Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT–PCR. *Nucleic acids research*, 29(9), e45-e45.
- 72) D'haene, B., Vandesompele, J., & Hellemans, J. (2010). Accurate and objective copy number profiling using real-time quantitative PCR. *Methods*, 50(4), 262-270.
- 73) Da Wei Huang, B. T. S., & Lempicki, R. A. (2008). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols*, 4(1), 44-57.
- 74) Sherman, B. T., & Lempicki, R. A. (2009). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic acids research*, 37(1), 1-13.
- 75) Bewick, V., Cheek, L., & Ball, J. (2004). Statistics review 13: receiver operating characteristic curves. *Critical care*, 8(6), 508.
- 76) Lalkhen, A. G., & McCluskey, A. (2008). Clinical tests: sensitivity and specificity. *Continuing Education in Anaesthesia, Critical Care & Pain*, 8(6), 221-223.

- 77) Dehn, D., Torkko, K. C., & Shroyer, K. R. (2007). Human papillomavirus testing and molecular markers of cervical dysplasia and carcinoma. *Cancer Cytopathology*, 111(1), 1-14.
- 78) Worster, A., Fan, J., & Upadhye, S. (2006). Understanding receiver operating characteristic (ROC) curves. *CJEM*, 8(1), 19-20.
- 79) Vazquez-Mena, O., Medina-Martinez, I., Juárez-Torres, E., Barrón, V., Espinosa, A., Villegas-Sepulveda, N., ... & Berumen, J. (2012). Amplified genes may be overexpressed, unchanged, or downregulated in cervical cancer cell lines. *PloS one*, 7(3), e32667.
- 80) Thierry, F., Benotmane, M. A., Demeret, C., Mori, M., Teissier, S., & Desaintes, C. (2004). A genomic approach reveals a novel mitotic pathway in papillomavirus carcinogenesis. *Cancer research*, 64(3), 895-903.
- 81) Teissier, S., Khalifa, Y. B., Mori, M., Pautier, P., Desaintes, C., & Thierry, F. (2007). A new E6/P63 pathway, together with a strong E7/E2F mitotic pathway, modulates the transcriptome in cervical cancer cells. *Journal of virology*, 81(17), 9368-9376.
- 82) Leinonen, M., Nieminen, P., Kotaniemi-Talonen, L., Malila, N., Tarkkanen, J., Laurila, P., & Anttila, A. (2009). Age-specific evaluation of primary human papillomavirus screening vs conventional cytology in a randomized setting. *Journal of the National Cancer Institute*, 101(23), 1612-1623.
- 83) Baleriola, C., Millar, D., Melki, J., Coulston, N., Altman, P., Rismanto, N., & Rawlinson, W. (2008). Comparison of a novel HPV test with the Hybrid Capture II (hcII) and a reference PCR method shows high specificity and positive predictive value for 13 high-risk human papillomavirus infections. *Journal of Clinical Virology*, 42(1), 22-26.
- 84) Whitlock, E. P., Vesco, K. K., Eder, M., Lin, J. S., Senger, C. A., & Burda, B. U. (2011). Liquid-based cytology and human papillomavirus testing to screen for cervical cancer: a systematic review for the US Preventive Services Task Force. *Annals of internal medicine*, 155(10), 687-697.
- 85) Giorgi-Rossi, P., Franceschi, S., & Ronco, G. (2012). HPV prevalence and accuracy of HPV testing to detect high-grade cervical intraepithelial neoplasia. *International Journal of Cancer*, 130(6), 1387-1394.
- 86) Tsoumpou, I., Arbyn, M., Kyrgiou, M., Wentzensen, N., Koliopoulos, G., Martin-Hirsch, P., ... & Paraskeva, E. (2009). p16^{INK4a} immunostaining in cytological and histological specimens from the uterine cervix: A systematic review and meta-analysis. *Cancer treatment reviews*, 35(3), 210-220.
- 87) Galgano, M. T., Castle, P. E., Atkins, K. A., Brix, W. K., Nassau, S. R., & Stoler, M. H. (2010). Using biomarkers as objective standards in the diagnosis of cervical biopsies. *The American journal of surgical pathology*, 34(8), 1077.
- 88) Wentzensen, N., Hampl, M., Herkert, M., Reichert, A., Trunk, M. J., Poremba, C., ... & von Knebel Doeberitz, M. (2006). Identification of high-grade cervical dysplasia by the detection of p16INK4a in cell lysates obtained from cervical samples. *Cancer*, 107(9), 2307-2313.
- 89) Draviam, V. M., Orrechia, S., Lowe, M., Pardi, R., & Pines, J. (2001). The localization of human cyclins B1 and B2 determines CDK1 substrate specificity and neither enzyme requires MEK to disassemble the Golgi apparatus. *The Journal of cell biology*, 152(5), 945-958.
- 90) Gong, D., & Ferrell, J. E. (2010). The roles of cyclin A2, B1, and B2 in early and late mitotic events. *Molecular biology of the cell*, 21(18), 3149-3161.
- 91) Patel, D., & McCance, D. J. (2010). Compromised spindle assembly checkpoint due to altered expression of Ubch10 and Cdc20 in human papillomavirus type 16 E6-and E7-expressing keratinocytes. *Journal of virology*, 84(21), 10956-10964.

- 92) Park, S. H., Yu, G. R., Kim, W. H., Moon, W. S., Kim, J. H., & Kim, D. G. (2007). NF- κ B-Dependent Cyclin B2 Expression in Colorectal Adenocarcinoma. *Clinical cancer research*, 13(3), 858-867.
- 93) Mo, M. L., Chen, Z., JIE, L., Li, H. L., Sheng, Q., Ma, H. Y., ... & Zhou, H. M. (2010). Use of serum circulating CCNB2 in cancer surveillance. *The International journal of biological markers*, 25(4), 236-242.
- 94) Nakayama, K. I., & Nakayama, K. (2006). Ubiquitin ligases: cell-cycle control and cancer. *Nature Reviews Cancer*, 6(5), 369-381.
- 95) Demetrick, D. J., Matsumoto, S., Hannon, G. J., Okamoto, K., Xiong, Y., Zhang, H., & Beach, D. H. (2008). Chromosomal mapping of the genes for the human cell cycle proteins cyclin C (CCNC), cyclin E (CCNE), p21 (CDKN1) and KAP (CDKN3). *Cytogenetic and Genome Research*, 69(3-4), 190-192.
- 96) Gyuris, J., Golemis, E., Chertkov, H., & Brent, R. (1993). Cdi1, a human G1 and S phase protein phosphatase that associates with Cdk2. *Cell*, 75(4), 791-803.
- 97) Galamb, O., Spisak, S., Sipos, F., Toth, K., Solymosi, N., Wichmann, B., ... & Molnar, B. (2010). Reversal of gene expression changes in the colorectal normal-adenoma pathway by NS398 selective COX2 inhibitor. *British journal of cancer*, 102(4), 765-773.
- 98) MacDermid, D. M., Khodarev, N. N., Pitroda, S. P., Edwards, D. C., Pelizzari, C. A., Huang, L., ... & Weichselbaum, R. R. (2010). MUC1-associated proliferation signature predicts outcomes in lung adenocarcinoma patients. *BMC medical genomics*, 3(1), 16.
- 99) Julien, S. G., Dubé, N., Hardy, S., & Tremblay, M. L. (2010). Inside the human cancer tyrosine phosphatome. *Nature Reviews Cancer*, 11(1), 35-49.
- 100) Taylor KJ, Sims AH, Liang L, Faratian D, Muir M, et al. (2010) Dynamic changes in gene expression in vivo predict prognosis of tamoxifen-treated patients with breast cancer. *Breast Cancer Res* 12:R39.
- 101) Mollinari, C., Kleman, J. P., Jiang, W., Schoehn, G., Hunter, T., & Margolis, R. L. (2002). PRC1 is a microtubule binding and bundling protein essential to maintain the mitotic spindle midzone. *The Journal of cell biology*, 157(7), 1175-1186.
- 102) Subramanian, R., Wilson-Kubalek, E. M., Arthur, C. P., Bick, M. J., Campbell, E. A., Darst, S. A., ... & Kapoor, T. M. (2010). Insights into antiparallel microtubule crosslinking by PRC1, a conserved nonmotor microtubule binding protein. *Cell*, 142(3), 433-443.
- 103) Gruneberg, U., Neef, R., Li, X., Chan, E. H., Chalamalasetty, R. B., Nigg, E. A., & Barr, F. A. (2006). KIF14 and citron kinase act together to promote efficient cytokinesis. *The Journal of cell biology*, 172(3), 363-372.
- 104) Lens, S. M., Voest, E. E., & Medema, R. H. (2010). Shared and separate functions of polo-like kinases and aurora kinases in cancer. *Nature Reviews Cancer*, 10(12), 825-841.
- 105) Li, C., Lin, M., & Liu, J. (2004). Identification of PRC1 as the p53 target gene uncovers a novel function of p53 in the regulation of cytokinesis. *Oncogene*, 23(58), 9336-9347.
- 106) Wang, S. M., Ooi, L. L. P., & Hui, K. M. (2011). Upregulation of Rac GTPase-activating protein 1 is significantly associated with the early recurrence of human hepatocellular carcinoma. *Clinical Cancer Research*, 17(18), 6040-6051.
- 107) Kretschmer, C., Sterner-Kock, A., Siedentopf, F., Schoenegg, W., Schlag, P. M., & Kemmner, W. (2011). Identification of early molecular markers for breast cancer. *Mol Cancer*, 10(1), 15.

- 108) Li, X. C., Bolcun-Filas, E., & Schimenti, J. C. (2011). Genetic evidence that synaptonemal complex axial elements govern recombination pathway choice in mice. *Genetics*, 189(1), 71-82.
- 109) Martinez, I., Wang, J., Hobson, K. F., Ferris, R. L., & Khan, S. A. (2007). Identification of differentially expressed genes in HPV-positive and HPV-negative oropharyngeal squamous cell carcinomas. *European journal of cancer (Oxford, England: 1990)*, 43(2), 415.
- 110) Moody, C. A., & Laimins, L. A. (2010). Human papillomavirus oncoproteins: pathways to transformation. *Nature Reviews Cancer*, 10(8), 550-560.
- 111) Mishra, A., & Verma, M. (2010). Cancer biomarkers: are we ready for the prime time?. *Cancers*, 2(1), 190-208.

ANEXOS

TABLA SUPLEMENTARIA 1 (TS1). Lista de 997 genes expresados diferencialmente en cáncer cervical comparado con tejido de epitelio normal control.

ID-Affymetrix	SIMBOLO ^a	NOMBRE GEN	CROMOSOMA	INCREMENTO ^b	Δ Score	MARCADOR
Sobreexpresados						
218009_s_at	PRC1	protein regulator of cytokinesis 1	15q26.1	7.30	16.76	
204170_s_at	CKS2	CDC28 protein kinase regulatory subunit 2	9q22	5.86	14.98	
202589_at	TYMS	thymidylate synthetase	18p11.32	9.36	14.78	D, DP, E, P, RT
204023_at	RFC4	replication factor C (activator 1) 4, 37kDa	3q27	7.09	13.99	
218039_at	NUSAP1	nucleolar and spindle associated protein 1	15q15.1	8.23	13.82	P
209773_s_at	RRM2	ribonucleotide reductase M2 polypeptide	2p25-p24	10.01	13.71	D
202705_at	CCNB2	cyclin B2	15q22.2	4.58	13.44	
202107_s_at	MCM2	minichromosome maintenance complex component 2	3q21	5.50	13.31	D, DP, E, P, UP
201664_at	SMC4	structural maintenance of chromosomes 4	3q26.1	5.45	13.16	
203213_at	CDC2	cell division cycle 2, G1 to S and G2 to M	10q21.1	5.19	12.99	
203755_at	BUB1B	BUB1 budding uninhibited by benzimidazoles 1 homolog beta (yeast)	15q15	4.27	12.68	
204962_s_at	CENPA	centromere protein A	2p24-p21	4.45	12.24	D
202954_at	UBE2C	ubiquitin-conjugating enzyme E2C	20q13.12	5.72	12.11	
203046_s_at	TIMELESS	timeless homolog (Drosophila)	12q12-q13	3.27	11.89	
216237_s_at	MCM5	minichromosome maintenance complex component 5	22q13.1	5.91	11.79	
202870_s_at	CDC20	cell division cycle 20 homolog (S. cerevisiae)	1p34.1	6.56	11.51	
222077_s_at	RACGAP1	Rac GTPase activating protein 1	12q13.13	3.81	10.72	
203554_x_at	PTTG1	pituitary tumor-transforming 1	5q35.1	4.44	10.53	P
208079_s_at	AURKA	aurora kinase A	20q13.2-q13.3	4.55	10.17	D, E, P
201202_at	PCNA	proliferating cell nuclear antigen	20pter-p12	3.98	10.09	D, DP, P, RT
209714_s_at	CDKN3	cyclin-dependent kinase inhibitor 3	14q22	6.15	9.97	
210052_s_at	TPX2	TPX2, microtubule-associated, homolog (Xenopus laevis)	20q11.2	4.81	9.90	
218755_at	KIF20A	kinesin family member 20A	5q31	4.51	9.88	
201292_at	TOP2A	topoisomerase (DNA) II alpha 170kDa	17q21-q22	6.73	9.77	D, E, RT
202633_at	TOPBP1	topoisomerase (DNA) II binding protein 1	3q22.1	3.80	9.66	
218350_s_at	GMNN	geminin, DNA replication inhibitor	6p22.2	4.45	9.65	
200853_at	H2AFZ	H2A histone family, member Z	4q24	3.30	9.64	
201897_s_at	CKS1B	CDC28 protein kinase regulatory subunit 1B	1q21.2	3.90	9.55	
204510_at	CDC7	cell division cycle 7 homolog (S. cerevisiae)	1p22	4.14	9.42	
204026_s_at	ZWINT	ZW10 interactor	10q21-q22	5.28	9.33	
204092_s_at	AURKA	aurora kinase A	20q13.2-q13.3	3.11	9.27	
221521_s_at	GINS2	GINS complex subunit 2 (Psf2 homolog)	16q24.1	3.99	9.22	
203209_at	RFC5	replication factor C (activator 1) 5, 36.5kDa	12q24.2-q24.3	3.32	9.18	
204162_at	NDC80	NDC80 homolog, kinetochore complex component (S. cerevisiae)	18p11.32	3.58	8.96	

201930_at	<i>MCM6</i>	minichromosome maintenance complex component 6	2q21	3.45	8.95	P
202338_at	<i>TK1</i>	thymidine kinase 1, soluble	17q23.2-q25.3	3.27	8.79	D, P
219148_at	<i>PBK</i>	PDZ binding kinase	8p21.2	4.15	8.78	
207039_at	<i>CDKN2A</i>	cyclin-dependent kinase inhibitor 2A (melanoma, p16, inhibits CDK4)	9p21	7.38	8.76	D, DP, E, P, RT,S
203358_s_at	<i>EZH2</i>	enhancer of zeste homolog 2 (Drosophila)	7q35-q36	3.82	8.76	D, DP, E, P
204767_s_at	<i>FEN1</i>	flap structure-specific endonuclease 1	11q12	3.34	8.69	
201555_at	<i>MCM3</i>	minichromosome maintenance complex component 3	6p12	2.86	8.63	
218355_at	<i>KIF4A</i>	kinesin family member 4A /// kinesin family member 4B	5q33.1 /// Xq13.1	3.75	8.62	
201589_at	<i>SMC1A</i>	structural maintenance of chromosomes 1A	Xp11.22-p11.21	2.74	8.39	
205046_at	<i>CENPE</i>	centromere protein E, 312kDa	4q24-q25	3.25	8.39	
203362_s_at	<i>MAD2L1</i>	MAD2 mitotic arrest deficient-like 1 (yeast)	4q27	4.52	8.37	
203418_at	<i>CCNA2</i>	cyclin A2	4q25-q31	3.11	8.35	E, P
204709_s_at	<i>KIF23</i>	kinesin family member 23	15q23	2.14	8.30	P
201697_s_at	<i>DNMT1</i>	DNA (cytosine-5-)-methyltransferase 1	19p13.2	2.75	8.28	D
206546_at	<i>SYCP2</i>	synaptonemal complex protein 2	20q13.33	6.91	8.27	
<i>Disminuyen su expresión</i>						
205382_s_at	<i>CFD</i>	complement factor D (adipsin)	19p13.3	0.08	-12.73	
209550_at	<i>NDN</i>	necdin homolog (mouse)	15q11.2-q12	0.16	-12.38	
205792_at	<i>WISP2</i>	WNT1 inducible signaling pathway protein 2	20q12-q13.1	0.11	-12.15	D
208399_s_at	<i>EDN3</i>	endothelin 3	20q13.2-q13.3	0.17	-12.12	
205857_at	<i>SLC18A2</i>	solute carrier family 18 (vesicular monoamine), member 2	10q25	0.33	-11.63	
201525_at	<i>APOD</i>	apolipoprotein D	3q26.2-qter	0.08	-11.28	
202920_at	<i>ANK2</i>	ankyrin 2, neuronal	4q25-q27	0.19	-11.24	
205752_s_at	<i>GSTM5</i>	glutathione S-transferase M5	1p13.3	0.34	-10.83	D, P
205794_s_at	<i>NOVA1</i>	neuro-oncological ventral antigen 1	14q	0.33	-10.50	
219213_at	<i>JAM2</i>	junctional adhesion molecule 2	21q21.2	0.26	-10.38	
205433_at	<i>BCHE</i>	butyrylcholinesterase	3q26.1-q26.2	0.14	-10.31	
205384_at	<i>FXYP1</i>	FXYP domain containing ion transport regulator 1 (phospholemman)	19q13.1	0.25	-10.26	
212843_at	<i>NCAM1</i>	neural cell adhesion molecule 1	11q23.1	0.28	-10.22	E, P
219440_at	<i>RAI2</i>	retinoic acid induced 2	Xp22	0.18	-10.17	
204359_at	<i>FLRT2</i>	fibronectin leucine rich transmembrane protein 2	14q24-q32	0.15	-10.14	
204570_at	<i>COX7A1</i>	cytochrome c oxidase subunit VIIa polypeptide 1 (muscle)	19q13.1	0.22	-9.77	
204749_at	<i>NAP1L3</i>	nucleosome assembly protein 1-like 3	Xq21.3-q22	0.20	-9.51	
205620_at	<i>F10</i>	coagulation factor X	13q34	0.42	-9.27	
204955_at	<i>SRPX</i>	sushi-repeat-containing protein, X-linked	Xp21.1	0.19	-9.27	
205110_s_at	<i>FGF13</i>	fibroblast growth factor 13	Xq26.3	0.32	-9.20	
205392_s_at	<i>CCL14</i>	chemokine (C-C motif) ligand 14	17q11.2	0.36	-8.93	
206481_s_at	<i>LDB2</i>	LIM domain binding 2	4p16	0.18	-8.92	
207154_at	<i>DIO3</i>	deiodinase, iodothyronine, type III	14q32	0.37	-8.84	

205200_at	<i>CLEC3B</i>	C-type lectin domain family 3, member B	3p22-p21.3	0.22	-8.81	
213568_at	<i>OSR2</i>	odd-skipped related 2 (Drosophila)	8q22.2	0.17	-8.80	
213249_at	<i>FBXL7</i>	F-box and leucine-rich repeat protein 7	5p15.1	0.39	-8.75	
209815_at	<i>PTCH1</i>	patched homolog 1 (Drosophila)	9q22.3	0.08	-8.67	
57588_at	<i>SLC24A3</i>	solute carrier family 24 (sodium/potassium/calcium exchanger), member 3	20p13	0.30	-8.66	
204134_at	<i>PDE2A</i>	phosphodiesterase 2A, cGMP-stimulated	11q13.4	0.51	-8.61	
206580_s_at	<i>EFEMP2</i>	EGF-containing fibulin-like extracellular matrix protein 2	11q13	0.38	-8.43	
212187_x_at	<i>PTGDS</i>	prostaglandin D2 synthase 21kDa (brain)	9q34.2-q34.3	0.13	-8.41	
205399_at	<i>DCLK1</i>	doublecortin-like kinase 1	13q13	0.39	-8.39	
205529_s_at	<i>RUNX1T1</i>	runt-related transcription factor 1; translocated to, 1 (cyclin D-related)	8q22	0.34	-8.26	
201540_at	<i>FHL1</i>	four and a half LIM domains 1	Xq26	0.16	-8.22	
205883_at	<i>ZBTB16</i>	zinc finger and BTB domain containing 16	11q23.1	0.42	-8.12	
209735_at	<i>ABCG2</i>	ATP-binding cassette, sub-family G (WHITE), member 2	4q22	0.33	-8.09	D, E, P
204287_at	<i>SYNGR1</i>	synaptogyrin 1	22q13.1	0.53	-8.08	
204719_at	<i>ABCA8</i>	ATP-binding cassette, sub-family A (ABC1), member 8	17q24	0.19	-7.98	
209737_at	<i>MAGI2</i>	membrane associated guanylate kinase, WW and PDZ domain containing 2	7q21	0.50	-7.98	
209710_at	<i>GATA2</i>	GATA binding protein 2	3q21.3	0.37	-7.86	
206167_s_at	<i>ARHGAP6</i>	Rho GTPase activating protein 6	Xp22.3	0.44	-7.78	
204115_at	<i>GNG11</i>	guanine nucleotide binding protein (G protein), gamma 11	7q21	0.25	-7.72	
204731_at	<i>TGFBR3</i>	transforming growth factor, beta receptor III	1p33-p32	0.21	-7.64	
210751_s_at	<i>RGN</i>	regucalcin (senescence marker protein-30)	Xp11.3	0.52	-7.63	
206227_at	<i>CILP</i>	cartilage intermediate layer protein, nucleotide pyrophosphohydrolase	15q22	0.20	-7.61	
209541_at	<i>IGF1</i>	insulin-like growth factor 1 (somatomedin C)	12q22-q23	0.21	-7.55	D, DP, E, P, RT,S
219935_at	<i>ADAMTSS5</i>	ADAM metalloproteinase with thrombospondin type 1 motif, 5 (aggrecanase-2)	21q21.3	0.30	-7.53	
206893_at	<i>SALL1</i>	sal-like 1 (Drosophila)	16q12.1	0.42	-7.49	
209869_at	<i>ADRA2A</i>	adrenergic, alpha-2A-, receptor	10q24-q26	0.20	-7.46	

a: Gene seleccionados con mayor valor de D score y los genes marcados en negro son seleccionados para el análisis con RT-PCR en tiempo real.

b: Incremento es calculado dividiendo el valor de media de la intensidad de señal entre tumores/controles de cada gen

*Aplicación de Biomarcadores: D: diagnóstico; DP, progresión de la enfermedad; P: pronóstico; E: eficacia; RT, respuesta a terapia; UP, aplicaciones no conocidas; S, seguridad.

Mitosis Is a Source of Potential Markers for Screening and Survival and Therapeutic Targets in Cervical Cancer

Ana Maria Espinosa^{1,2}, Ana Alfaro^{1,2}, Edgar Roman-Basaure³, Mariano Guardado-Estrada², Ícela Palma⁴, Cynthia Serralde², Ingrid Medina^{1,2}, Eligia Juárez², Miriam Bermúdez², Edna Márquez⁵, Manuel Borges-Ibáñez⁶, Sergio Muñoz-Cortez⁶, Avissai Alcántara-Vázquez⁷, Patricia Alonso⁷, José Curiel-Valdez⁷, Susana Kofman^{1,8}, Nicolas Villegas⁹, Jaime Berumen^{1,2*}

1 Departamento de Medicina Experimental, Facultad de Medicina, Universidad Nacional Autónoma de México, México City, México, **2** Unidad de Medicina Genómica, Hospital General de México, México City, México, **3** Servicio de Oncología, Hospital General de México, México City, México, **4** Escuela Superior de Medicina, Instituto Politécnico Nacional, México City, México, **5** Instituto de Investigación en Matemáticas Aplicadas, Universidad Nacional Autónoma de México, México City, México, **6** Servicio de Ginecoobstetricia, Hospital General de México, México City, México, **7** Servicio de Patología, Hospital General de México, México City, México, **8** Servicio de Genética, Hospital General de México, México City, México, **9** Departamento de Biomedicina Molecular, Centro de Investigación y Estudios Avanzados del Instituto Politécnico Nacional, México City, México

Abstract

The effect of preventive human papillomavirus (HPV) vaccination on the reduction of the cervical cancer (CC) burden will not be known for 30 years. Therefore, it's still necessary to improve the procedures for CC screening and treatment. The objective of this study was to identify and characterize cellular targets that could be considered potential markers for screening or therapeutic targets. A pyramidal strategy was used. Initially the expression of 8,636 genes was compared between 43 HPV16-positive CCs and 12 healthy cervical epitheliums using microarrays. A total of 997 genes were deregulated, and 21 genes that showed the greatest deregulation were validated using qRT-PCR. The 6 most upregulated genes (*CCNB2*, *CDC20*, *PAC1*, *SYCP2*, *NUSAP1*, *CDKN3*) belong to the mitosis pathway. They were further explored in 29 low-grade cervical intraepithelial neoplasias (CIN1) and 21 high-grade CIN (CIN2/3) to investigate whether they could differentiate CC and CIN2/3 (CIN2+) from CIN1 and controls. *CCNB2*, *PAC1*, and *SYCP2* were mostly associated with CC and *CDC20*, *NUSAP1*, and *CDKN3* were also associated with CIN2/3. The sensitivity and specificity of *CDKN3* and *NUSAP1* to detect CIN2+ was approximately 90%. The proteins encoded by all 6 genes were shown upregulated in CC by immunohistochemistry. The association of these markers with survival was investigated in 42 CC patients followed up for at least 42 months. Only *CDKN3* was associated with poor survival and it was independent from clinical stage (HR = 5.0, 95%CI = 1.4–23.8, p = 0.01). *CDKN3* and *NUSAP1* may be potential targets for the development of screening methods. Nevertheless, further studies with larger samples are needed to define the optimal sensitivity and specificity. Inhibition of mitosis is a well known strategy to combat cancers. Therefore, *CDKN3* may be not only a screening and survival marker but a potential therapeutic target in CC. However, whether it's indispensable for tumor growth remains to be demonstrated.

Citation: Espinosa AM, Alfaro A, Roman-Basaure E, Guardado-Estrada M, Palma I, et al. (2013) Mitosis Is a Source of Potential Markers for Screening and Survival and Therapeutic Targets in Cervical Cancer. PLOS ONE 8(2): e55975. doi:10.1371/journal.pone.0055975

Editor: Michael Schreiber, Baylor College of Medicine, United States of America

Received: August 15, 2012 **Accepted:** January 4, 2013 **Published:** February 6, 2013

Copyright: © 2013 Espinosa et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by the National Council of Science and Technology (CONACYT, www.conacyt.mx), grant numbers 8133A1, 24341 (to JB) and 80680 (to SA), and the National University of Mexico (www.unam.mx), grant number S2/PTD/05.2 (to JB, AME, AA and IMW) were recipients of a scholarship from CONACYT. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: jaimeberumen@hotmail.com

Introduction

The human papilloma virus (HPV) is the main causal factor for the development of invasive cervical cancer (CC), and HPV is found in nearly 100% of these tumors [1,2]. CC results from the progression of preinvasive cervical intraepithelial neoplasia (CIN), which is histologically graded into mild (CIN 1), moderate (CIN 2), or severe (CIN 3) dysplasia. CC occurs mainly from CIN3 and CIN2, but rarely from CIN1; the estimated progression rates of these lesions to CC are 12%, 5% and 1%, respectively [3]. Currently, there are vaccines on the market that prevent infection by oncogenic HPV types 16 and 18, which are associated with 65–70% of CCs worldwide [4]. These vaccines have very high efficiency for the prevention of infection and the development of

high-grade cervical intraepithelial neoplasias (CIN2+/CIN3) [5,6]. However, vaccinated women must still attend programs for early detection of CC since these vaccines only protect against certain virus types, and it is not yet known how long the immune protection against the target virus remains [7,8]. In many countries preventive vaccines for HPV 16 and 18 have been incorporated into the national vaccination program, for girls from 9 to 12 years of age [9,10]. However, because the peak incidence of CC occurs in women 45–59 years old, the effect of these preventive vaccination programs on reducing the prevalence of CC will not be known for 30 years. Therefore, it is necessary to improve the procedures for CC screening and treatment. Because each year 530,000 new cases of CC and 275,000 CC deaths are

reported worldwide, the incidence to mortality ratio is approximately 50% [11,12].

For many years, the Papanicolaou (Pap) test has been the most important screening procedure for early detection of CC, and its massive application in developed countries has decreased the incidence of CC by more than 50% in the last 40 years [13]. Women with abnormal Paps are referred for colposcopy to confirm, discard, or clarify the diagnosis with a histopathological study. However, the average sensitivity of cytology for detection of CIN lesions is 50–60%; although the specificity is very high, approximately 90% [14]. Since HPV is indispensable for the development of CC, several procedures to detect the HPV genome have been incorporated into CC screening. Compared with conventional cytology, HPV DNA testing has higher sensitivity but lower specificity for the detection of CIN2 lesions or higher (CIN2+). The high sensitivity and high negative predictive value (NPV) of HPV DNA tests for the detection of CIN2+ lesions suggest that it could be used to extend screening intervals. However, the low specificity of HPV DNA tests would increase the number of follow-up tests and colposcopy referrals, which would increase the cost of screening [15]. Therefore, the need to develop new methods for early detection of CC with high sensitivity and specificity is clear. Multiple tumor markers associated with CIN2+ have been identified, especially *CDKN2A*, *TOT2A*, and *MCM2*. However, these markers have been proposed not for screening, but for diagnosis, prognosis, or clinical management [11,16].

Invasive cervical cancer is currently treated with surgery, chemotherapy, radiotherapy, or a combination of these therapies, depending on the clinical stage of the disease. The success of these conventional therapies and patient survival diminishes as the disease progresses to more advanced stages [17]. In fact, the percentage of women who survive 5 years decreases from 53% for stage IA to 15% for stage IVB (www.cancer.org). In contrast to other types of cancer, for which several specific molecular drugs have been developed [18], there are no specific molecular-targeted therapies for CC. The majority of drugs against specific targets in cancer are directed toward mutated proteins, especially protein kinases [19]; however, some drugs target normal proteins that are overexpressed, such as *HER2/neu* in breast cancer [20,21]. The first step in developing a specific molecular drug is identifying universal molecular targets that are present in patients with CC and absent in healthy women.

The objective of this study was to identify and characterize cellular targets present in most CCs and absent from normal cervical tissue that differ enough between the 2 groups to be considered either as potential markers for screening, with a sensitivity and specificity close to 100%, or as potential therapeutic targets.

Methods

Ethics Statement

The study protocol was approved by the Scientific and Ethics Committees of the Hospital General de México (approval number DIC/03/311/04/051) and was performed in accordance with the ethical principles described in the 1964 Declaration of Helsinki. Informed written consent was obtained from all participants prior to their inclusion in the study.

Subjects, Samples, and Experimental Design

The study subjects included 69 patients with invasive cervical cancer (CC) diagnosed in the Department of Oncology, 29 patients with low-grade CIN (CIN1), 21 patients with high-grade CIN (CIN2 and CIN3), and 25 women with normal cervical

epithelium evaluated in the Department of Obstetrics and Gynecology at the Hospital General de México in Mexico City. The CC samples were a subset selected from a total of 462 patients with CC who were recruited sequentially from November 2003 through July 2007, which represented approximately 80% of the patients newly diagnosed with CC during this period due to the restrictive inclusion criteria (no previous treatment, incident case, born in Mexico with Mexican ancestry for 2 generations). The selection criteria for the CC subset were based on the availability of a fresh tumor biopsy for RNA extraction with more than 70% tumor cells in the morphological analysis (see below), mostly FIGO stages I/II, and viral type. This subset included 47 samples positive for HPV16 and 22 samples positive for other virus types, including HPV18, 31, 33, 45, 51, 58, and 59. Among them, 54 samples were of squamous cell carcinomas, 14 samples were of adenocarcinomas, and 1 sample was of an adenosquamous carcinoma. The average age of patients with cancer was 48 years (range, 23–78 years; Table S1). All patients received complete clinical evaluations. The tumors of CC patients were staged according to the last international revised protocol for gynecologic cancer [22]. One or two biopsies, conducted under colposcopy examination, were taken from tumors. One biopsy was divided in 2 equal parts, 1 part was fixed in buffered formal for morphological analysis and the other part, together with the second biopsy, was snap-frozen on dry ice and stored at -80°C until analysis. All CC patients were referred for surgery, radiation, chemotherapy, or a combination of these treatments according to the guidelines of the American Cancer Society (see below). Control cervical specimens were obtained from patients undergoing hysterectomy due to myomatosis at the Gynecology Service of the Hospital General de México. They were previously diagnosed with a normal cervix by cytology and colposcopy. Immediately after receiving a cervix fragment from the operating room, the exocervical and endocervical epitheliums were dissected under a stereoscopic microscope to avoid the stromal cells. The tissues were then snap frozen in liquid nitrogen and stored at -80°C until use. For HPV detection and typing, a scrape from the endocervix and exocervix was collected with a cytobrush from the patients and controls, the cells were suspended in a vial with extraction buffer, and then stored at -20°C until analysis. Analysis of global gene expression (8,633 genes) was performed in RNAs extracted from 43 fresh tumor biopsies positive for HPV16 and from 12 samples of normal cervical epithelium using the HG-Focus microarray. Global gene expression was validated in 24 samples, including 19 CCs and 5 cervical epithelium controls, by a second high throughput microarray (HG-ST1.0). The 23 genes that showed the greatest deregulation were validated by real time PCR (qRT-PCR) in 44 HPV16-positive CC and 25 control samples. The 6 most differentially expressed genes (*CCNB2*, *CDG20*, *PRC1*, *SYCF2*, *NUSAP1*, and *CDKN7*) were further explored in 29 low-grade cervical intraepithelial neoplasias (CIN1) and 21 high-grade CIN (CIN2/3) to investigate whether they could differentiate CC and CIN2/3 (CIN2+) from CIN1 and controls. Immunohistochemistry (IHC) was performed for 10 selected proteins in 26 CC samples and 10 control samples. The association of 9 markers with survival was investigated by survival analysis of 49 patients with HPV16-positive CC who were followed up for at least 42 months.

DNA and RNA Isolation

DNA was purified from cervical scrapes and some biopsy specimens using the PureLink Genomic DNA Kit (Invitrogen, Grand Island NY) and maintained at -20°C until analysis. Total RNA was isolated from one half of the divided biopsy using TRIzol reagent (Invitrogen), according to the manufacturer's

protocol. The quality of the RNA was confirmed by agarose gel electrophoresis, as demonstrated by the presence of intact ribosomal RNA, with the 28s band twice as intense as the 18s band.

Detection and HPV Typing

HPV detection was performed by PCR using universal primers located in the HPV *L1* gene *MT09/MT11*, *GP5+/6+*, and *L1C1* as described previously [23–25]. The *hBB* gene was used as an internal control to assess the quality of DNA. The HPV types were identified by sequencing the amplified bands in positive samples using a fluorescent cycle-sequencing method (BigDye Terminator Ready Reaction Kit, Applied Biosystems, Foster city, CA). Sequence analysis was performed using an ABI PRISM 3130xl genetic analyzer (Applied Biosystems). Each sequence from the HPV positive samples was analyzed with the FASTA sequence similarity tool [26]. The average percentage identity of these sequences to HPV types was 98.7% (range, 91–100%).

Gene Expression Profiling and Data Analysis

The gene expression profile of 43 CCs positive for HPV16 and 12 healthy control cervical epitheliums was examined using the Human Gene Focus (HG Focus) oligonucleotide Microarray (MA) (Affymetrix, Santa Clara, CA). This array contains 8,744 probe sets corresponding to 8,638 characterized human genes in the Gene Reference database. Total RNA preparation (10 µg), labeled cRNA synthesis, hybridization, scanning, and image analysis were performed according to the manufacturer's protocols (Affymetrix GeneChip Expression Assay manual). To assess the quality of the experiments, the following parameters were used: increased expression of exogenous poly-A controls (Lys<Phe<Thr<Dap), the presence of oligo B2 used to make grid alignments, background with an acceptable range of 20 to 100, equal noise across all samples, percentage of present calls greater than 50%, a 3'/5' ratio of a constitutive gene (GAPDH or β-actin) of less than 3, and increased expression of the hybridization controls (BioB<BioC<BioD<cre). Only those MAs with optimal quality controls were analyzed. Furthermore, some samples were performed in duplicate to evaluate the reproducibility of the experiment, which was higher than 99%.

MA intensity values were normalized using the Robust Multi-chip Average (RMA) algorithm using FlexArray software [27]. The normalized intensity values were referred to as units of intensity (UI). Genes expressed differently between the tumors and controls were identified using the algorithm Significance Analysis of Microarrays (SAM version 3.0, <http://www.stat.stanford.edu/~tibsh/SAM>) using the cut-off values of a fold change (FC) of ≥ 1.5 , a general false discovery rate (FDR) of 1%, and a local FDR of <10% [28]. Unsupervised hierarchical clustering and principal component analysis (PCA) were performed using dChip software (version 1.6, www.dchip.org/) and R language in Java's platform, respectively.

Validation of Global Gene Expression by a Second High throughput Microarray (HG-ST1.0)

The gene expression profile of 24 samples explored with the HG-Focus microarray, including 19 CCs and 5 cervical epithelium controls, was also examined using the Human Gene 1.0 ST oligonucleotide microarray (Affymetrix, Santa Clara, CA). This array contains 33,297 probe sets that correspond to approximately 20,741 genes of the human gene reference database according to the USC Genome Browser Assembly Mar. 2006 NCBI 35/hg18, available at <http://genome.ucsc.edu/>. Total RNA preparation

(500 ng), labeled DNA synthesis, hybridization, scanning, and image analysis were performed according to the manufacturer's protocols (Affymetrix GeneChip Expression Assay manual). To assess the quality of the experiments, the following parameters were used: expression of the exogenous poly-A controls, the presence of oligo B2 used to make grid alignments, and area under the curve (AUC) values above 0.8. Only those microarrays with optimal quality controls were analyzed. Microarrays were normalized using the RMA algorithm in the Affymetrix expression console. The normalized intensity values were referred to as units of intensity (UI). The normalized intensities (log₂ values) of the 8,376 genes that were examined on both microarrays (HG ST1 and HG Focus) were compared, and the level of correlation was assessed with Pearson's correlation coefficient.

Validation of Global Gene Expression by Real-time Quantitative Retrotranscription PCR (qRT-PCR)

Reverse transcription of total RNA was performed using the High-Capacity cDNA Archive kit (Applied Biosystems) in a total volume of 20 µL. The mix included 2 µg of RNA, 2 µL of 10× RT buffer, 0.8 µL of 100 mM dNTPs, 2 µL of 10× RT Random Primers, 1 µL of MultiScribe™ reverse transcriptase (5 U/µL), and 1 µL of RNase inhibitor (2 U/µL). Reactions were incubated at 37°C for 120 min, and then stored at -20°C. A set of 23 genes was used to validate gene expression in 14 HPV16-positive CC and 25 healthy cervical epithelium control samples with qRT-PCRs using TaqMan probes. The genes included are *CCNB2*, *CDC2*, *CDC20*, *CDKN2A*, *CDKN3*, *CKS2*, *MCM2*, *MKI67*, *NUSAP1*, *PCNA*, *PRC1*, *RFU4*, *RRM2*, *SMC4*, *SYCP2*, *TOP2A*, *TYMS*, *ZWINT*, *CFD*, *EDN3*, *NDN*, *SLC18A2*, and *WISP2*. *GAPDH* was used as internal control. TaqMan gene expression assays were used (Table S2, Applied Biosystems). Seven genes were also explored in 22 CC positive for other HPVs (*CCNB2*, *CDC20*, *CDKN3*, *PRC1*, *SYCP2*, *NUSAP1*, *TYMS*), and the first 6 of them, along with *CDKN2A*, *PCNA*, *MKI67* genes, were further explored in 29 low-grade CINs and 21 high-grade CINs. The experiments were run in duplicate in a final volume of 20 µL, including 200 ng of cDNA template, 10 µL of 2× TaqMan Universal PCR Master Mix (Applied Biosystems), 1 µL of 20× TaqMan Gene Expression Assay, and 7 µL of RNase-free water. The cycling program was run in a Rotor-Gene (Corbett Research, Sydney, Australia), which was set as follows: an initial PCR activation step at 50°C for 2 min followed by 95°C for 10 min, then 40 cycles of melting at 95°C for 15 s and annealing/extension at 60°C for 1 min. The median of the Ct standard deviations in duplicates ranged from 0.09 to 0.24 (mean = 0.16) among the 23 genes, suggesting that the variations between the duplicates were very small [29]. Measurement of gene expression was based on relative standard curves constructed from a 10-fold serially diluted pool of CC or normal cervical epithelium cDNAs ranging from 500 to 0.05 ng. The first curve was used to calculate the values of upregulated genes and the second curve the values of downregulated genes. Curves for each gene were tested in three different experiments run in duplicate and the averages of the correlation coefficients (r) were higher than 0.98. The expression of target genes was normalized in each tumor and control sample to the intensity of the internal reference (*GAPDH*) using a previously described method [30]. The normalized intensity values were measured in ng/µL. A normality test (Shapiro-Wilk) was carried out to test for a normal distribution of gene expression data. The fold-change expression was calculated by dividing the median normalized intensity of tumor samples by the median normalized intensity of the control samples. The statistical significance between the medians of tumors and controls was calculated with the Mann-Whitney

(MW) non-parametric test. The correlations between the MA results and the qRT-PCR data were performed using \log_2 values and measured using Pearson's correlation coefficient.

Immunohistochemistry

The protein expression of 10 genes was determined in 26 CC and 10 control samples with IH. Two homemade tissue microarrays (TMA) were built, one containing 14 HPV16-positive CCs and 5 controls and the other 12 CC positive for other HPVs and 5 controls. NUSAP1 was explored only in samples of the first TMA. Cylindrical samples from representative regions of the paraffin embedded tissue blocks, previously selected by H&E stained slides, were taken with a punch-biopsy needle (2 mm diameter), transferred to recipient paraffin blocks in defined array positions and newly embedded in paraffin. All the tissues blocks of matched patients were obtained from the Pathology Department of the hospital. Serial sections (4 μ m thick) of the TMA were cut and the 10th slide was stained with H&E to confirm the histopathological diagnosis. Sections were immersed in xylene to remove paraffin and then rehydrated with graded alcohol (100%, 95%, 90%, 80%, and 70% v/v in water). Epitope retrieval was performed by heating the slides, and introducing them into Target Retrieval Solution, pH 6.0 (Dako, Carpinteria, CA) at 121°C for 5 min in a pressure cooker. Endogenous peroxidase activity was blocked by incubating the slides with 1% hydrogen peroxide in PBS for 10 min. Then, a non-specific background blocker was added and incubated for 10 min. Primary antibodies against PCNA (sc-53407); p16 for CDKN2A (sc-71804); SCP-2 for SYCP2 (sc-20048); PRG1 (sc-56345); cyclin B2 for CCNB2 (sc-81241); CDKN3 (sc-475); and CDC2 for p34 (sc-70322), were obtained from Santa Cruz Biotechnology (Santa Cruz, CA). The antibodies against CDC20 (cat. 34-1900), Ki-67 for MK167 (cat. M7187) and NUSAP1 (cat. H10051703-H11) were obtained from Invitrogen, Dako (Glostrup, Denmark), and Nova Biological (Littleton, CO), respectively. The dilution used for all antibodies was 1:100, except for CDC2, (1:50) and NUSAP1 (1:250), and the antibody diluent used was from Dako. A total volume of 300 μ L was added to each section, and the slides were incubated overnight at 4°C in a moist chamber. Antigen-antibody complexes were detected by the avidin-biotin peroxidase method, using 3,3'-diaminobenzidine-tetrahydrochloride as a chromogenic substrate (Cat. KO679 LSAB1 Sys/HRP; Dako-Cytomation Carpinteria, CA), and the sections were counterstained with hematoxylin. Assays were performed in triplicate. The antibodies for SYCP2, PRG1, CCNB2, CDKN3, CDC2, and CDC20 were tested in tissues known to express those antigens. SYCP2 was tested in neonate testis; PRG1, CDC2, and CCNB2 were tested in colon cancer; and CDKN3 was tested in lung cancer biopsies. All tissues were obtained from the archives of the Pathology Department. The percentage of stained cells was calculated from an analysis of 10 successive high-power fields of neoplastic cells. The cellular localization of the immunoreaction was identified, and the intensity of the immunoreaction was scored from 0 to 4, where 0 indicated no staining. Immune reaction signals were found rarely in the stroma with all antibodies and were not scored for the analysis. Immunostained slides were analyzed and scored by 2 pathologists, who were blinded to the outcomes. Rare cases with discordant scores were reevaluated and scored based on consensus opinion.

Survival Analysis of Cancer Patients

According to FIGO staging patients with cervical cancer received individualized treatment based on the treatment

guidelines for cervical cancer of the American Cancer Society (See Table 1). After the treatment was completed, each patient was clinically evaluated every 3 or 5 months by an experienced oncologist. Clinical data of the follow-up study was obtained from the patients medical record. Also, a social worker performed phone calls and home visits to the patients every 5 months during the study. Patients recorded as alive in the study were successfully followed up for at least 42 months after treatment. Censored and deceased patients were followed up for the number of months indicated in Table 1. The cases designated as censored referred to those patients who were lost to the study in the follow-up period or deceased from causes other than cervical cancer. Patients were considered lost when did not attend to medical appointments for disease control, were not found at home visits or did not answer phone calls. In this cohort, patients recorded as deceased were only those women who died by cervical cancer primary tumor as a main cause. The cause of death of all but one patient who died during the follow up was confirmed by the medical record and the death certificate. Only 42 of 44 patients with HPV16-positive CC explored with qRT-PCR were included in the followed up study. Four cases were considered right censored and eight deaths were registered. The mean following time of the 42 patients was 50.5 months. The association of FIGO and gene expression (*PRG1*, *CCNB2*, *CDC20*, *CDKN3*, *NUSAP1*, *SYCP2*, *CDKN2A*, *PCNA*, *MKI67*) with survival was investigated by survival analysis. With the whole sample set, 500 training sets of 21 samples were randomly created for each gene explored. To categorize the gene expression data quantified by qRT-PCR, ROC analysis was performed in each training set. This analysis was done to set a cut-off for gene expression that represented those values with the highest sensitivity and specificity to differentiate between dead and surviving patients. The whole sample set was then analyzed with the average cut-off, calculated from the values of the 500 training sets. Samples with gene expression values above the cut-off were set to 1 and those with values below the cut-off were set to 0. The cumulative overall survival time was calculated by the Kaplan-Meier method and analyzed by the log-rank test. FIGO staging and the gene expression were included as covariates in a Cox proportional hazard model.

Gene Ontology Classification Analysis

The Database for Annotation, Visualization, and Integrated Discovery (DAVID) functional annotation tool (<http://david.abcc.ncifcrf.gov>) [31,32] and the Ingenuity Pathway Analysis (IPA; Ingenuity® Systems, www.ingenuity.com) were used to classify the deregulated genes. Genes were classified using functional annotation clustering considering the gene ontology biological processes. Classification stringency was set at medium and maximum level.

Gene Annotation and Data Analysis

The physical position of genes was mapped according to the UCSC Genome Browser Assembly Mar. 2006 NCBI 36/hg18, available at <http://genome.ucsc.edu/>. Data analysis was performed using Access 2010 (Microsoft Inc.). The raw MA data is MIAME compliant and has been deposited in a MIAME compliant database (GEO, <http://www.ncbi.nlm.nih.gov/geo/>) under the accession number GSE35001. Receiver operator characteristic (ROC) curve analysis was performed and Youden index was used [33] to select the best cut-off points to distinguish tumors from controls and CIN2+ from CIN1 using the expression values of selected genes obtained by qRT-

Table 1. Patients followed up for at least 42 months for survival evaluation.

Sample	Histology ^a	Tumor Stage	Age (years)	Treatment ^b	Follow up (months)	Status ^c
R053	ACC	IB1	57	HT	53	Alive
R446	ACC	IB1	43	HT	59	Alive
R081	ACC	IR1	41	HT	67	Alive
R054	SCC	IB1	45	HT	62	Alive
R369	SCC	IB1	50	HT	65	Alive
R057	ACC	IB1	32	HT	53	Alive
R077	SCC	IR1	61	HT+TFI F+BRACHY	66	Alive
R411	SCC	IB1	31	HT+TELE+BRACHY	60	Alive
R434	ACC	IB1	34	HT+TELE+BRACHY	61	Alive
R445	SCC	IB1	34	HT+TELE+BRACHY	61	Alive
R758	SCC	IR1	36	HT+TFI F+BRACHY	68	Alive
R335	ACC	IB1	37	HT+TELE+BRACHY+CHEMO	65	Alive
R183	SCC	IB1	64	TELE+BRACHY	61	Alive
R308	ACC	IB1	46	TELE+BRACHY	61	Alive
R765	SCC	IR1	46	TFI F+BRACHY	67	Alive
R330	SCC	IB1	72	TELE+BRACHY	54	Alive
R035	SCC	IB2	48	TELE+CHEMO+HT	73	Alive
R221	SCC	IB2	41	TELE+BRACHY+CHEMO	35	Death*
R777	SCC	IB2	45	TFI F+BRACHY+CHEMO	33	Death
R409	SCC	IB2	63	TELE+BRACHY+CHEMO	42	Alive
R339	SCC	IB2	31	TELE+BRACHY+CHEMO	13	Death
R324	SCC	IB2	23	TELE+BRACHY+CHEMO	14	Death
R359	ACC	IR2	34	TFI F+BRACHY+CHEMO	17	Death
R378	SCC	IB2	42	TELE+BRACHY+CHEMO	56	Alive
R356	ACC	IB2	53	TELE+BRACHY+CHEMO	7	Unknown
R312	ACC	IB2	34	TELE+BRACHY+CHEMO	58	Alive
R467	SCC	IR2	61	TFI F+BRACHY+CHEMO	60	Alive
R284	ACC	IB2	33	TELE+BRACHY+CHEMO	63	Alive
R412	SCC	IB2	33	TELE+BRACHY+CHEMO	63	Alive
R336	SCC	IB2	36	TELE+BRACHY+CHEMO	64	Alive
R755	SCC	IIA	45	TFI F+BRACHY+CHEMO	47	Death
R052	ACC	IIB	51	TELE+BRACHY+HT	19	Death
R070	SCC	IIB	74	TELE+BRACHY+CHEMO	4	Unknown
R170	SCC	IIB	67	TELE+CHEMO+HT	82	Alive
R473	SCC	IIR	34	TFI F+BRACHY+CHEMO	64	Alive
R015	SCC	IIB	42	TELE+HT	66	Alive
R268	SCC	IIB	34	TELE+BRACHY+CHEMO	58	Alive
R416	ASCC	IIB	56	TELE+BRACHY+CHEMO	59	Alive
R441	ACC	IIR	74	TFI F+BRACHY+CHEMO	10	Unknown
R333	SCC	IIB	56	TELE+BRACHY+CHEMO	66	Alive
R315	SCC	IIIB	41	TELE+BRACHY+CHEMO	8	Death
R240	SCC	IIIB	51	TELE+BRACHY+CHEMO	11	Death

^aACC, Adenocarcinoma. SCC, Squamous Cell Carcinoma. ASCC, Adenosquamous Cell Carcinoma.

^bHT, Radical Hysterectomy. Tele, teletherapy. Brachy, brachytherapy. Chemo, chemotherapy with Cisplatin.

^cStatus alive was registered at the last follow up. death was caused by primary tumor of cervical cancer, except the case labeled with an asterisk, and unknown cases were lost during the follow up study. The cause of death of case labeled with an asterisk was unknown.

doi:10.1371/journal.pone.0055975.t001

²CR. For each marker, the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were calculated according to previously described formulas [34].

All tests were 2 sided, and p-values less than 0.05 were considered statistically significant. Data analysis was performed using Sigma Stat and SPSS ver. 17 software.

Results

Expression Analysis of 8,638 Genes in Cervical Cancer

The amount of mRNA transcribed from 8,638 genes was compared between 43 CC samples positive for HPV16 and 12 normal cervical epithelial samples using the HG-Focus microarray. A total of 997 genes were differentially expressed between the cancer and control groups, 600 were upregulated and 397 were downregulated (Table S3). Almost one-half of the upregulated and downregulated genes had FCs in the range of 1.5–2.0, and the number of genes in both groups decreased linearly ($r = -0.8$, $p = 0.002$) as the FC value increased (Figure 1). The principal component analysis (PCA; data not shown) and the non-supervised hierarchical clustering (panel A in Figure 2) performed with all 997 gene expression values clearly separated the cancer samples from the control group. However, the expression of many genes was not completely uniform among the cancer samples, especially in the group of upregulated genes (signals shown in red in Figure 2A). Many of those genes were upregulated in some tumors and downregulated in other tumors. This was in contrast to the uniformity of the expression signals in the control group samples. Genes to be tested as markers for screening or as potential therapeutic targets were selected according to Δ -score rank (a modified t-test, used in SAM), FC or whether they were previously used as markers for cervical cancer. From the 997 genes associated with the cancer samples, 163 have been previously reported as markers for different types of cancer (IPA, Ingenuity Systems), including *MCM2*, *TOP2A*, and *CDKN2A*, which have been used as markers for diagnosis in cervical cancer [35]. The 997 genes were listed in decreasing order by Δ -score (Table S3). A total of 23 genes (18 upregulated and 5 downregulated) were selected for validation by qRT-PCR (marked in bold in Table S3 and Table 2; circles colored in blue and orange in Figure 1). All downregulated genes (*GDF*, *NDN*, *WISP2*, *END3*, and *SLC18A2*) and 10 of the 18 upregulated genes (*PRC1*, *CKS2*, *TYMS*, *RFCA*, *RRM2*, *NUSAP1*, *MCM2*, *CCNB2*, *SMC4*, and *CDC2*) were selected according to Δ -score rank. Seven of the remaining upregulated genes are on the list of the 50 best ranked genes, 2 of them are genes that have been previously proposed as markers in CC (*CDKN2A* and *TOP2A*), 4 (*CDC20*, *CDKN3*, *ZWINT*, and *SYCP2*) were selected based on the FC value, and *PCNA* (Table S3), together with *MKI67*, which ranked in 139th place, were included because these markers are commonly used to measure cell proliferation. The PCA analysis and hierarchical clustering showed that the 23 selected genes also allowed for segregation of the samples into the 2 different groups. For both the upregulated and downregulated genes, the difference in signal intensities was quite uniform among the samples from the 2 groups (Figure 2, panels B and C).

Validation of Genes with HG-ST1.0 Microarray and Quantitative Real-time PCR (qRT-PCR)

A total of 8,970 genes were validated with microarray HG-ST1.0 in 24 samples explored with the HG-Focus microarray, including 19 CC samples and 5 healthy cervical epitheliums. Highly significant positive correlations ($p < 1 \times 10^{-15}$, Pearson's correlation) were found between the HG-ST1.0 and HG-Focus microarray values. The global correlation between the 2 arrays was 0.68 and the correlation coefficients between the individual tumors ranged from 0.57 to 0.72 (average, 0.68). Gene expression values of 826 out of 997 (82.8%) genes expressed differently between the cancer and control samples, including the 23 genes selected for validation (Figure 3), showed significant positive correlations ($p < 0.05$, Pearson's correlation) between the 2 microarrays. The correlation coefficients between the individual

genes ranged from 0.34 to 0.95 and the average was 0.63 (Figure S1).

On the other hand, the expression of the 23 genes selected for validation was measured with qRT-PCR in a total of 44 HPV16-positive cancer samples and 25 healthy cervical epitheliums, including almost all samples previously determined with MA (Table S1). A highly significant positive correlation ($p < 0.0001$, Pearson's correlation) was found between the qRT-PCR and MA \log_2 values in 21 of the 23 measured genes. The correlation coefficients ranged from 0.31 to 0.85 and the median was 0.73. The 2 genes that had non-significant correlations (*NDN* and *SLC18A2*) were excluded from the rest of the analysis. These data indicated that the expression values calculated from the microarrays were fairly reliable because 91% of validated genes had a significant correlation. Since the qRT-PCR expression values of 87% of the measured genes did not follow a normal distribution, the median rather than the mean was used for the calculations of FCs. Genes were listed in decreasing order by the FC (Table 2) and at the top of this list is *MKI67*, which is followed in decreasing order by *CDKN2A*, *SYCP2*, *PCNA*, *NUSAP1*, and *CDC2*. It is worth noting that the FCs of *MKI67* (1,551), and *CDKN2A* (387) are at least 5 times higher than the FC of the gene that follows on the list (*SYCP2*; FC = 73.8). Of the top 10 ranked upregulated genes, 2 have not been previously reported as associated with cervical cancer (*NUSAP1* and *CDKN3*), while the other 8 have been associated with cervical cancer either scantily (*SYCP2*, *PRC1*, *CCNB2* and *CDC20*) or widely (*MKI67*, *CDKN2A*, *CDC2*, and *PCNA*). *MCM2* and *TOP2A*, which have been widely reported as associated with cervical cancer, ranked 15th and 18th on the list, respectively. The 3 downregulated genes that had a significant Pearson's correlation also had a high FC (controls vs. cancers), especially *FMO3* (FC = 1,425.7) and *WISP2* (FC = 167.7; Table 2). The box plots (Figure 4 and Figure S2) clearly show the difference in gene expression between the cancer and control groups ($p < 1 \times 10^{-5}$ for all genes, Mann-Whitney *U* test).

To establish a separation line between the 2 groups and the potential value of these genes as markers of cervical cancer, cut-off values were established by analyzing ROC curves. In general, ROC curves with an area under the curve (AUC) ≤ 0.75 are not clinically useful, while an AUC of 0.97 has a very high clinical value [35]. The AUC of 11 upregulated genes (*CDKN2A*, *MKI67*, *PRC1*, *CDC2*, *CCNB2*, *SYCP2*, *PCNA*, *NUSAP1*, *TYMS*, *CDC20*, and *CDKN3*) and 1 downregulated gene (*WISP2*) was ≥ 0.97 (Table 3). In fact, most of these genes had a sensitivity and specificity greater than 95%, suggesting that they could be good markers for screening between healthy samples and invasive cancers. Interestingly, this subset included 2 genes that were not reported to be associated with CC (*NUSAP1* and *CDKN3*) and 5 genes that were scantily reported to be associated with CC (*PRC1*, *SYCP2*, *CCNB2*, *TYMS*, and *CDC20*). These 7 genes were examined with qRT-PCR in the 22 CC samples positive for other viral types, including HPV18, 31, 33, 45, 51, 53, and 59. All of them were upregulated in these 22 tumors; however, the FCs were lower than those obtained in the HPV16-positive tumors (Table 2). These data suggest that these 7 genes might be upregulated in all invasive CCs regardless of viral type; therefore, they could be considered potential markers for CC screening.

Analysis of *CCNB2*, *CDC20*, *PRC1*, *SYCP2*, *NUSAP1*, and *CDKN3* Expression in Pre-invasive Neoplasias

For screening tests, it is important to detect not only CC, but also high-grade lesions (CIN2/3) and to distinguish them from low-grade CIN lesions (CIN1) and healthy controls. Therefore, to investigate whether these genes can differentiate CIN2+ from

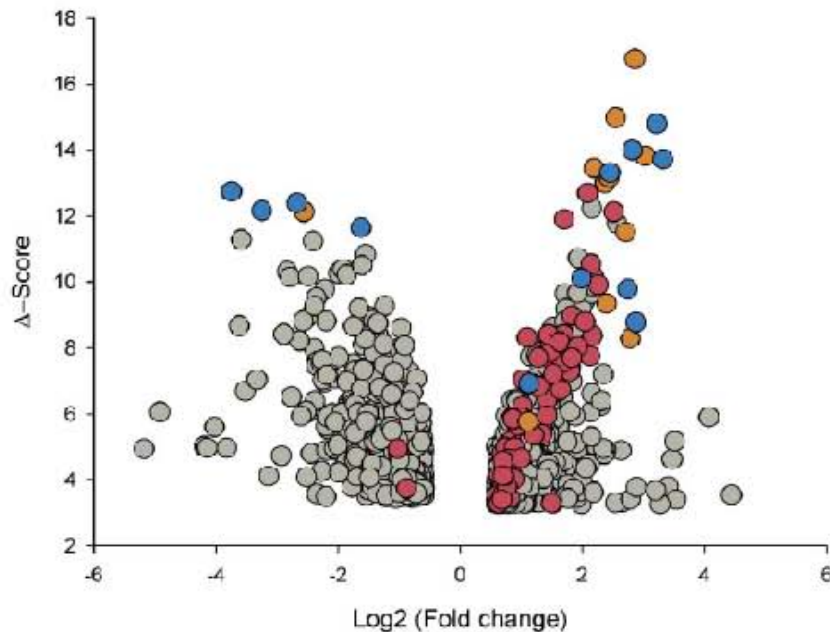


Figure 1. Distribution of deregulated genes according to the fold change (FC) and Δ -score values. All 997 genes (circles) that were deregulated in cervical cancer (CC) tumors compared to the control samples by the SAM method are graphed in a volcano plot. The x-axis represents the FC in gene expression (cancer sample/control sample) expressed in Log₂ and the y-axis display the absolute Δ -score, a modified t-test calculated with the SAM method, the higher the Δ -score values, the higher the statistical significance. The Log₂ (FC) values are positive for upregulated genes and negative for downregulated genes. Circles colored in red and orange represent the genes involved in M-phase of the cell cycle and those colored in blue and orange are the genes that were validated by qRT-PCR. doi:10.1371/journal.pone.0055975.g001

CIN1+, expression was analyzed in 2 additional groups of samples: 29 low-grade CINs and 21 high-grade CINs. Their expression was compared with that of 3 known markers associated with CC (*PCNA*, *MKI67*, and *CDKN2A*), which were ranked in the top 10 in the previous qRT-PCR analysis (see above). Experimental data were box plotted (Figure 4) and the statistical significance of differences was calculated using the MW test. According to the median and distribution of the data in the box plots, the 9 markers can be classified into 3 groups; the first group included markers linked exclusively (*CCNB2*, *PRC1*) or mostly (*SYCP2*) to invasion, which clearly differentiated invading tumors from high-grade CIN, low-grade CIN, and normal cervixes. The expression of these markers in the control group, and in low-grade and high-grade lesions was similar ($p > 0.05$, MW). In contrast, the difference between the CC and control samples was quite large, as was established in the previous analysis (Table 2). Similarly, the FC compared to high-grade CIN was also very high, especially for *SYCP2* (FC = 84.8; $p < 1 \times 10^{-15}$), followed by *PRC1* (FC = 39.4; $p < 1 \times 10^{-15}$, MW) and *CCNB2* (FC = 15.9; $p < 1 \times 10^{-15}$, MW). Moreover, the specificity for detecting just CC, and not other lesions, ranged from 0.05 (*SYCP2*) to 0.98 (*CCNB2*); the optimal cut-off values were at a change well over 4.5 fold. The lower specificity shown by *SYCP2* was because 7 preinvasive lesions (5 CIN2/3 and 2 CIN1) had a FC greater than the optimal cut-off value for this gene (7.9). The second group included 4 genes (*CDC20*, *NUSAP1*, *CDKN2A*, and *CDKN3*) the expression of which tended to increase from the control group to the CC group (*CDC20*, *CDKN2A*, and *CDKN3*) or the high-grade CIN group (*NUSAP1*). For *NUSAP1*, the expression in CIN2/3 and CC was similar (Figure 4). These 4 genes could distinguish CIN2+ lesions

from CIN1- lesions ($p < 1 \times 10^{-15}$, MW; Figure 4). The third group included *MKI67* and *PCNA*, the expression of which increased from the control group to the low-grade CIN group ($p < 0.05$, MW), was similar in the low-grade and high-grade CIN groups ($p > 0.05$, MW), and then increased in the CC group ($p < 1 \times 10^{-15}$, MW; Figure 1). It is clear that genes in the first and third groups would not be good markers for screening since they cannot distinguish high-grade CIN and CC lesions from low-grade CIN lesions and control samples. ROC analysis was performed to explore the potential of the genes in the second group (*CDC20*, *CDKN2A*, *CDKN3*, and *NUSAP1*) as markers for screening. None of them had AUC values equal to or greater than 0.97; the highest AUC value was obtained with *CDKN2A* (0.92), followed by *NUSAP1* (0.917), *CDKN3* (0.91) and *CDC20* (0.86) (Table 4). However, the new markers (*NUSAP1* and *CDKN3*) showed a slightly greater sensitivity than *CDKN2A*, while the opposite was true for the specificity (Table 4). Interestingly, the sensitivity and specificity increased when individual data for *CDKN3*, *NUSAP1*, and *CDKN2A* were combined (Table 4). This combination showed the highest Youden index. From these, only *CDKN3* can also discriminate CC from CIN2/3 (FC cut-off = 4.4) with high sensitivity (0.9) and specificity (0.94).

Verification of the Protein Expression of Selected Tumor Marker Candidates by Immunohistochemistry

To investigate whether the validated genes (*PRC1*, *CDKN3*, *CCNB2*, *SYCP2*, *NUSAP1* and *CDC20*) were also overexpressed at the protein level, the coding proteins were assessed by IH. The expression of *PCNA*, *CDKN2A*, *MKI67*, and *CDC2* was also examined. All but one (*NUSAP1*) proteins were explored in 36

Figure 2. Segregation of tumor and control samples according to the expression of deregulated genes. Unsupervised hierarchical cluster analysis of 43 CCs and 12 healthy cervical epitheliums using the expression values obtained with the HG-Focus microarray of all 997 deregulated genes (panel A) or the 23 top ranked genes selected for validation (panel B). Each row represents a gene and each column represents a sample. The length and the subdivision of the branches represent the relationships among the samples based on the intensity of gene expression. The cluster is color-coded using red for upregulation, green for downregulation, and black for unchanged expression. Panel C shows the principal components analysis (PCA) using the values in panel B; blue circles represent the CCs (n=43) and yellow circles represent the controls (n= 12). Both sets of genes clearly separated the samples into the 2 main groups using both types of analysis. doi:10.1371/journal.pone.0055975.g002

samples (10 controls and 26 CCs, 14 positive for HPV16 and 12 positive for other HPV). NUSAP1 was explored only in HPV16-positive CCs and 5 controls. Unlike the controls, almost all CCs were positive for the 10 antigens (Figure 5, Figure S3). A higher percentage of positive tumors and more intense signals were observed for PCNA (56.2%), followed by CDKN2A and CDKN3 (84.6%), CCNB2 and CDC2 (80.8%), NUSAP1 (79%), MKI67, SYCP2 and PRCl (76.9%), and CDC20 (73.1%). Unexpectedly, a considerable number of controls were positive

for CDC20 (60%), NUSAP1 (40%) and SYCP2 (50%); however, for CDC20 the signals were only observed in the nuclei of cells in the basal layer, for NUSAP1 the signals were observed in the nuclei and cytoplasm of cells in the basal and parabasal layers and for SYCP2 in the basal pole of epithelial cells of superficial and intermediate layers. For the rest of antigens, the differences in positivity between the 2 groups agreed with the data obtained with qRT-PCR (Table S4). Signals for CDKN3, SYCP2, PRCl, CDC2, NUSAP1, and CDKN2A were observed in both the cytoplasm and the nucleus, while signals for CCNB2 were only observed in the cytoplasm, and signals for CDC20, PCNA, and MKI67 were only observed in the nucleus (Figure 5, Figure S3). As expected, the IH signals were not uniform in all cells of all tissues, but rather the distribution was heterogeneous, indicating that not all cells are at the same stage of the cell cycle. The PCNA signals showed the most uniform distribution, and on average 70% of the nuclei were positive, suggesting that approximately 70% of the cells in the tissues were in S phase of the cell cycle. For the rest of the proteins, nuclear signals were observed in 10–50% of cells (Figure 6A). Signals for the proteins localized in the cytoplasm were observed in 40–50% of cells on average (Figure 6B). Given that all these proteins are involved in the M phase of the cell cycle (see below and discussion), the data suggest that 30–40% of the cells are in some stage of this phase. Interestingly, the percentage of cells positive for CCNB2, CDC2, and SYCP2 was higher in tumors positive for HPV16 than in tumors positive for other HPVs, and the opposite was observed for CDKN3 (Figure 5). The predictive capability of IH was also evaluated. Compared to the RT-PCR results, the sensitivity was lower for all proteins, but the specificity was higher for all proteins, except for SYCP2, NUSAP1 and CDC20 (Table S4).

Table 2. Genes explored by qRT-PCR.

Gene ^a	Fold change ^b	
	HPV16 positive	Other HPVc
Upregulated		
<i>MKI67</i>	1651	–
<i>CDKN2A</i>	387	–
<i>SYCP2</i>	74	14
<i>PCNA</i>	65	–
<i>NUSAP1</i>	26	15
<i>CDC2</i>	23	–
<i>CDC20</i>	17	13
<i>CCNB2</i>	14	6
<i>TYMS</i>	12	2
<i>PRCl</i>	9	4
<i>SMC4</i>	8	–
<i>CDKN3</i>	7	5
<i>RRM2</i>	6	–
<i>QSOX2</i>	5	–
<i>MCM2</i>	4	–
<i>ZWINT</i>	4	–
<i>RFC4</i>	4	–
<i>TOP2A</i>	3	–
Downregulated		
<i>EDN3</i>	1426	–
<i>HMSP2</i>	168	–
<i>CFD</i>	25	–
<i>NDN</i>	1	–
<i>SLC18A2</i>	0.3	–

^aGenes in bold were selected to be explored in pre-invasive samples.

^bThe analysis was performed with 44 HPV16-positive CC, 22 CC positive for other HPVs and 15 cervical controls. Fold change (FC) was calculated with the median values as follows: tumor/control for upregulated genes and control/tumor for downregulated genes (see Materials and Methods). The difference between the groups was statistically significant ($p < 1 \times 10^{-15}$; Mann-Whitney Rank Sum Test) for all but 2 genes (NDN, SLC18A2). NDN and SLC18A2 had a $p > 0.05$.

^cIncluded carcinomas positives for HPV-10 (5), -31 (5), -33 (2), -45 (5), -51 (2), -59 (7) and -59 (1).

doi:10.1371/journal.pone.0055975.t002

Molecular Targets in Cervical Cancer Associated with Poor Survival

One way to investigate whether or not these molecular targets are associated with cervical cancer progression is a survival study. Therefore, a survival analysis using the qRT-PCR expression values of *PRCl*, *CCNB2*, *CDC20*, *CDKN3*, *NUSAP1*, *SYCP2*, *CDKN2A*, *PCNA*, and *MKI67* and FIGO staging was conducted on 42 patients with HPV16-positive CC whose progress was followed-up for at least 3.5 years after their diagnosis and initial treatment (Table 1). This subset included FIGO stages IB1 (n = 16), IB2 (n = 14), IIA (n = 1), IIB (n = 9), and IIIB (n = 2). The overall survival rate for the whole sample was 79.6% and for FIGO stages IB1, IB2, IIA, IIB, and IIIB were 100%, 69.2%, 0%, 85.7%, and 0%, respectively. These differences were statistically significant ($p < 0.001$, log-rank test; Figure 7A). Of the 9 genes analyzed using Kaplan-Meier curves, only *CDKN3* was associated with poor survival ($p = 0.004$, log-rank test; Figure 7B). The overall survival rate of patients with the higher levels of *CDKN3* (FC > 15) was 42.9%, and the median survival time was 33 months. In contrast, those with lower levels of *CDKN3* had an overall survival rate of 87.5%.

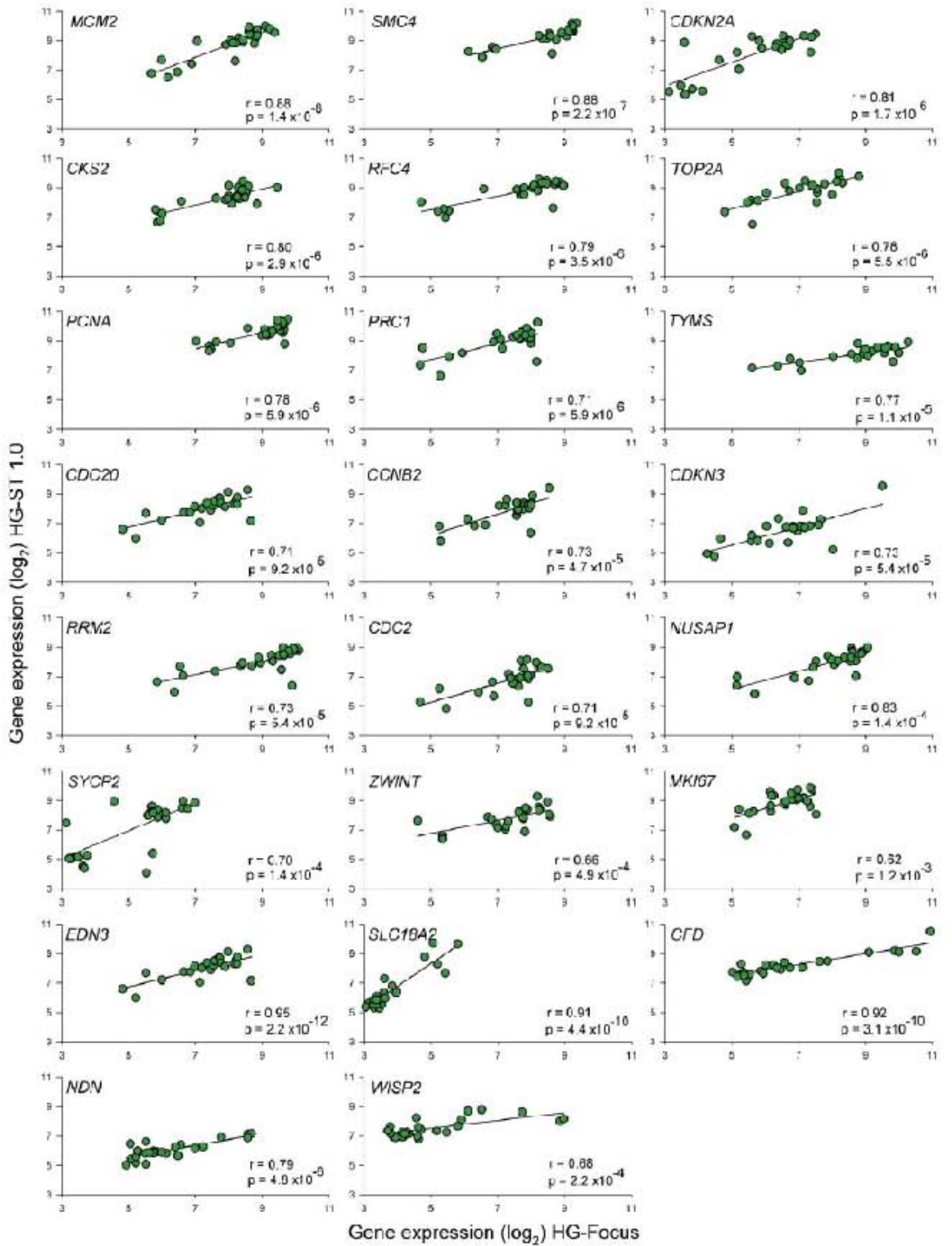


Figure 3. Correlation of expression intensity of 23 genes examined by HG-Focus and HG-ST1.0 microarrays. The Log_2 values of the standardized intensity signals (RMA values) of 23 genes examined by the 2 microarrays in 19 CC and 5 normal cervical epithelium were plotted. The linear trend (black line) is included, which was calculated with Pearson's correlation test. r = correlation coefficient, p = p -value. doi:10.1371/journal.pone.0055975.g003

FIGO staging and *CDKN3* expression were analyzed individually and together in Cox proportional hazard models. Because of the differences in the sample size among the FIGO stages analyzed, patients were reassigned to 2 groups, one including FIGO IB1 and IB2 ($n=30$) and the other FIGO IIA, IIB, and IIB ($n=12$). Individually, the hazard ratio (HR) of *CDKN3* was 5.9 (95% CI 1.4–24.1, $p=0.01$) and of the grouped FIGO, 3.5 (95% CI 0.83–13.3, $p=0.08$). The lack of significance in the HR of grouped FIGO could be explained by differences in the sample size and the inverted survival rates of the individual FIGO stages IB1 and IB. When these 2 covariates were included in the same proportional hazard model, *CDKN3* remained invariably significant with an HR of 5.9 (95% CI 1.4–23.8, $p=0.01$). These results suggest that *CDKN3* could be a prognostic factor for survival that is independent of FIGO

staging. However, a larger sample size is needed to confirm these results.

Classification of Genes with Differential Expression between Cancer and Control Samples

The DAVID functional annotation tool (<http://david.abcc.ncifcrf.gov>) was used at medium and highest stringency to identify the biological processes where the 997 differentially expressed genes are involved. Compared with the human genome database, the 3 most enriched clusters, and with the lowest p values at medium stringency, were cell cycle-associated processes, DNA metabolic processes, and processes associated with the regulation of ubiquitin-protein ligase activity (Table S5). Interestingly, at the highest stringency, where more lightly associated genes in each group are expected, the clusters including mitosis and M-phase of

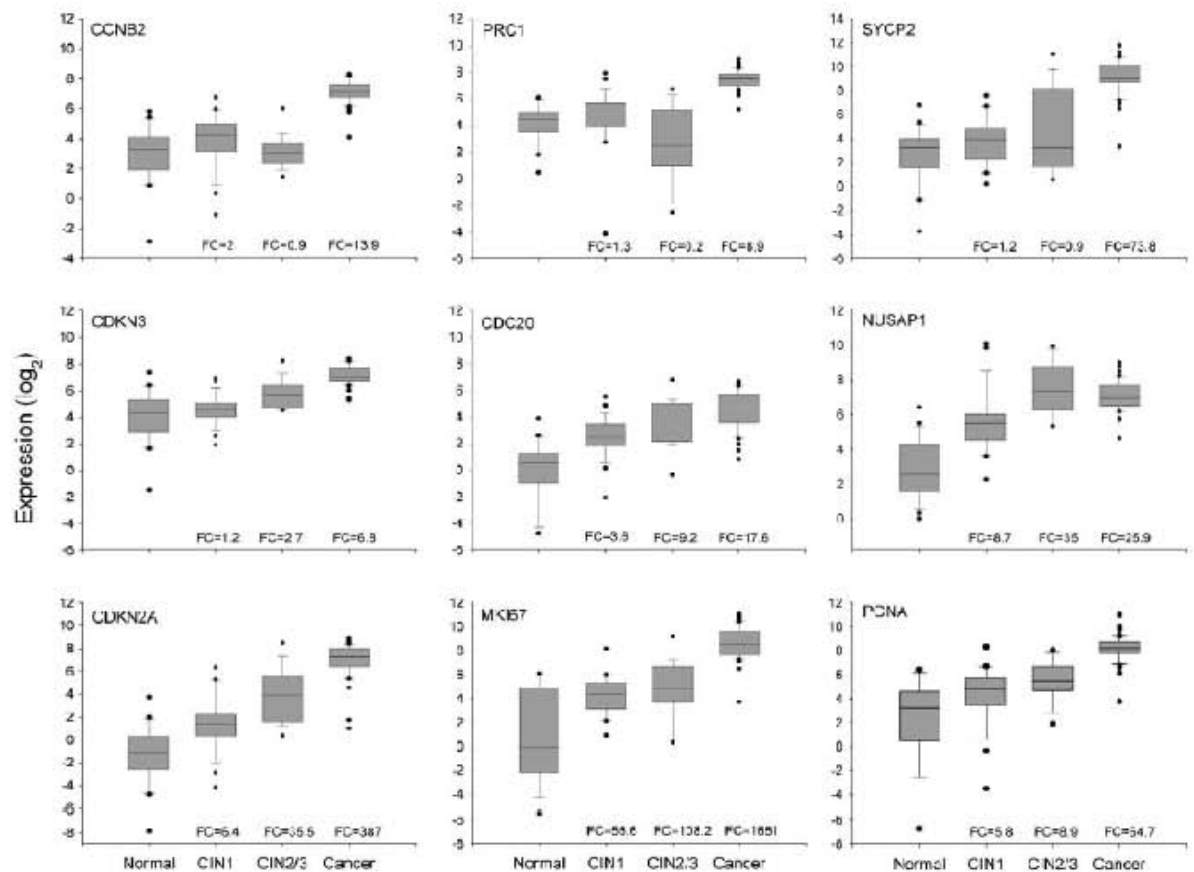


Figure 4. Validation of gene expression of 9 genetic markers by qRT-PCR. The intensity of gene expression, expressed in Log_2 values, is shown in box plots. Expression of the 6 genes validated in this study (*CCNB2*, *PRC1*, *SYCP2*, *CDKN3*, *CDC20*, and *NUSAP1*) and the 3 well-known genes (*CDKN2A*, *MKI67*, and *PCNA*) associated with CC are compared among the 4 groups, including healthy cervical epithelium (Normal, $n=25$), low-grade CIN (CIN1, $n=29$), high-grade CIN (CIN2/3, $n=21$), and invasive CC (cancer, $n=44$). The upper and lower boundaries of the boxes represent the 75th and 25th percentiles, respectively. The black line within the box represents the median value, and the whiskers represent the minimum and maximum values that lie within 1.5 \times the interquartile range from the end of box. Values outside this range are represented by black circles. The fold change (FC) was calculated by dividing the median of each pathological group by the median of the control group. doi:10.1371/journal.pone.0055975.g004

Table 3. ROC analysis and calculus of sensitivity, specificity and predictive values.

Genes	AUC	Cut-off value ^a	Controls (n = 25)				Cervical Cancer (n = 44)		Sensitivity	Specificity	PPV	NPV	Youden index ^d
			FPF	TNF	TPF	FNF	TPF	FNF					
CDKN2A	0.996	18	0	25	42	2	<1 × 10 ⁻¹⁰	0.95	1	100	92.6	0.95	
CCNB2	0.995	58	0	25	43	1	<1 × 10 ⁻¹⁰	0.98	1	100	96.2	0.98	
MKI67	0.995	75	0	25	43	1	<1 × 10 ⁻¹⁰	0.98	1	100	96.2	0.98	
PRC1	0.995	90	0	25	43	1	<1 × 10 ⁻¹⁰	0.98	1	100	96.2	0.98	
CDC2	0.995	85	0	25	42	2	<1 × 10 ⁻¹⁰	0.95	1	100	92.6	0.95	
SYCP2	0.992	115	0	25	42	2	<1 × 10 ⁻¹⁰	0.95	1	100	92.6	0.95	
NUSAP1	0.990	48	1	24	43	1	<1 × 10 ⁻¹⁰	0.98	0.96	97.7	96.0	0.94	
PCNA	0.990	100	0	25	42	2	<1 × 10 ⁻¹⁰	0.95	1	100	92.6	0.95	
TYMS	0.985	46	0	25	41	3	<1 × 10 ⁻¹⁰	0.93	1	100	89.3	0.93	
CDC20	0.971	3	3	22	42	2	<1 × 10 ⁻¹⁰	0.95	0.98	93.3	91.7	0.83	
CDKN3	0.970	83	1	24	41	3	<1 × 10 ⁻¹⁰	0.93	0.96	97.6	88.9	0.89	
SMC4	0.960	431	1	24	40	4	<1 × 10 ⁻¹⁰	0.91	0.96	97.6	85.7	0.87	
AFC4	0.905	221	4	21	42	2	<1 × 10 ⁻¹⁰	0.95	0.84	91.3	91.3	0.79	
RRM2	0.905	103	5	20	41	3	3 × 10 ⁻⁹	0.93	0.8	89.1	87.0	0.73	
TOP2A	0.866	128	5	20	43	1	<1 × 10 ⁻¹⁰	0.98	0.8	89.6	95.2	0.78	
MCM2	0.846	121	4	21	40	4	2.5 × 10 ⁻⁹	0.91	0.84	90.9	84.0	0.75	
ZWINT	0.827	39	7	18	39	5	1.1 × 10 ⁻⁶	0.89	0.72	84.8	78.3	0.61	
CKS2	0.815	239	5	20	35	9	5 × 10 ⁻⁶	0.80	0.8	87.5	69.0	0.60	
			TPF	FNF	FPF	TNF							
CFD	0.982	478	24	1	2	42	<1 × 10 ⁻¹⁰	0.96	0.95	97.7	92.3	0.91	
EDNS	0.968	42	25	2	4	40	<1 × 10 ⁻¹⁰	0.92	0.91	95.2	85.2	0.83	
WSP2	0.926	151	24	1	10	34	2.1 × 10 ⁻⁸	0.96	0.77	97.1	70.6	0.73	

AUC: area under the curve, FPF: false positive fraction, TNF: true negative fraction, TPF: true positive fraction, FNF: false negative fraction, PPV: Positive predictive value, NPV: Negative predictive value
^aOptimal cut-off values (ng/ml) were selected according to the ROC analysis.
^bChi square test.
^cJ = sensitivity + specificity - 1.
^ddoi:10.1371/journal.pone.0055975.t003

Table 4. ROC analysis of 4 gene markers selected for detection of CIN2/3 and CC.

Marker	AUC	Cut-off value	≥CIN1 (n = 54) ^a				≥CIN2/3 (n = 65) ^a		Sensitivity	Specificity	PPV	NPV	Youden Index
			FPF	TNF	TPF	FNF	TPF	FNF					
CDKN2A	0.920	14	4	50	52	13	0.80	0.93	92.9	75.4	0.73		
NUSAP1	0.917	71	6	48	59	6	0.91	0.89	90.8	86.9	0.80		
CDKN3	0.909	50	8	46	55	10	0.85	0.85	87.3	82.1	0.70		
CDC20	0.854	11	7	47	46	19	0.71	0.87	86.8	71.2	0.58		
CDKN3, CDKN2A, CDC20			4	50	55	10	0.85	0.93	93.2	83.3	0.77		
CDKN3, CDKN2A, NUSAP1			4	50	57	8	0.88	0.93	93.4	86.2	0.80		
CDKN3, CDC20, NUSAP1			6	48	57	8	0.88	0.89	90.5	85.7	0.77		
CDKN2A, CDC20, NUSAP1			4	50	53	12	0.82	0.93	93.0	80.6	0.71		

See legends of Table 3.
 The last 4 rows included the combined analysis of CDKN3, NUSAP1, CDC20 and CDKN2A as indicated. Samples were considered positive when at least 2 of the 3 markers were positive.
^aAll comparisons gave a p-value <1 × 10⁻⁹, chi square.
 doi:10.1371/journal.pone.0055975.t004

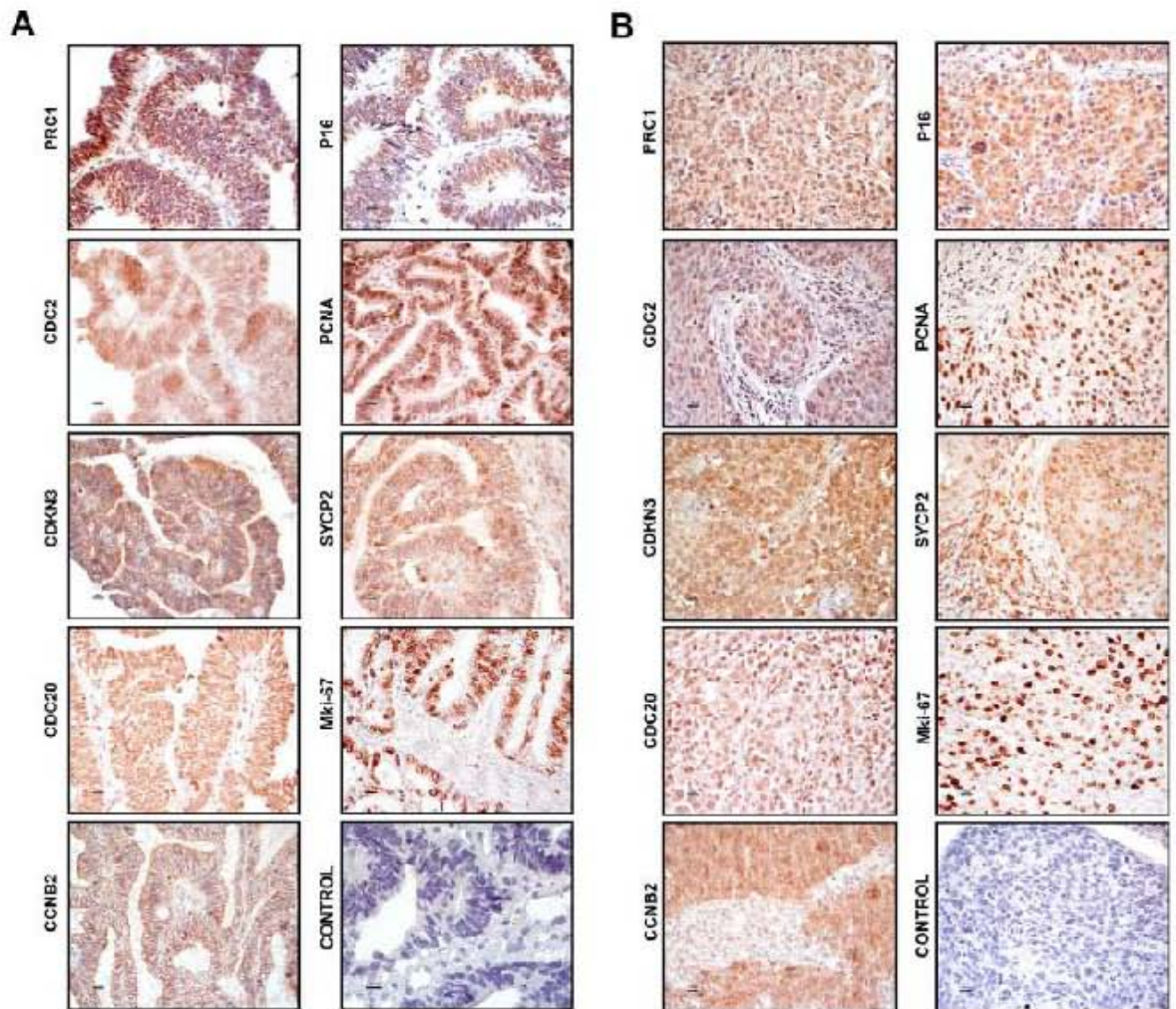


Figure 5. Histological analysis of marker genes. Protein expression was determined by immunohistochemistry using sections from formalin-fixed, paraffin-embedded tissue. Proteins explored were CDKN3A, SYCP2, PRC1, CDC20, CCNE2, PCNA, CDKN2A, MKI67, and CDC2. Representative experiments in adenocarcinomas (panel A) and squamous cell carcinomas (panel B) are shown. The specific signals are shown as brown staining (counterstained with hematoxylin; original magnification, $\times 400$; bars, 10 μm). doi:10.1371/journal.pone.0055975.g005

mitotic cell-cycle processes were ranked in the 1st, 2nd, and 5th places (Table 5). It is worth noting that, among the cell cycle processes, none except the M-phase was enriched significantly (Table 5). Remarkably, in the 100 top ranked genes subset (50 upregulated and 50 downregulated), the mitosis cluster was also the most enriched pathway and compared with the whole set ($n=997$) it was enriched over 3.3 fold (Table 6). This data indicated that the genes involved in mitosis were not only the most enriched, but also the most different in terms of FC and A-score, when compared with the control samples (circles in red and orange; Figure 1). In fact, 11 of the 21 genes associated with CC and validated in this work (*CCNE2*, *CDC20*, *PRC1*, *SYCP2*, *NUSAP1*, *CDKN3*, *GDC2*, *CKS2*, *MAI67*, *SMC1* and *ZWINT*) are involved in the M phase of the cell cycle. The data were also analyzed with the IPA Ingenuity system and the findings were similar to those obtained with DAVID, especially when the

DAVID analysis was run with medium stringency (Table S5). In agreement with the DAVID analysis, the protein ubiquitination pathway was the second top canonical pathway in the entire set of deregulated genes (Figure 8A) and the mitotic roles of polo-like kinase was the top in the subset of the 100 top ranked genes (Figure 8B).

Discussion

In this work we identified 6 genes (*PRC1*, *CCNE2*, *SYCP2*, *CDKN3*, *NUSAP1*, and *CDC20*) associated with invasive cervical cancer that could be used either as markers for diagnosis or as therapeutic targets. While *PRC1*, *CCNE2*, and *SYCP2* were associated mostly with CC, *CDKN3*, *NUSAP1*, and *CDC20* were found to also be associated with high-grade CIN. We recently examined the expression of these genes in 4 cell lines (SiHa, CaSki, HeLa, and Cal5) by using the HG-ST1.0 microarray. Although

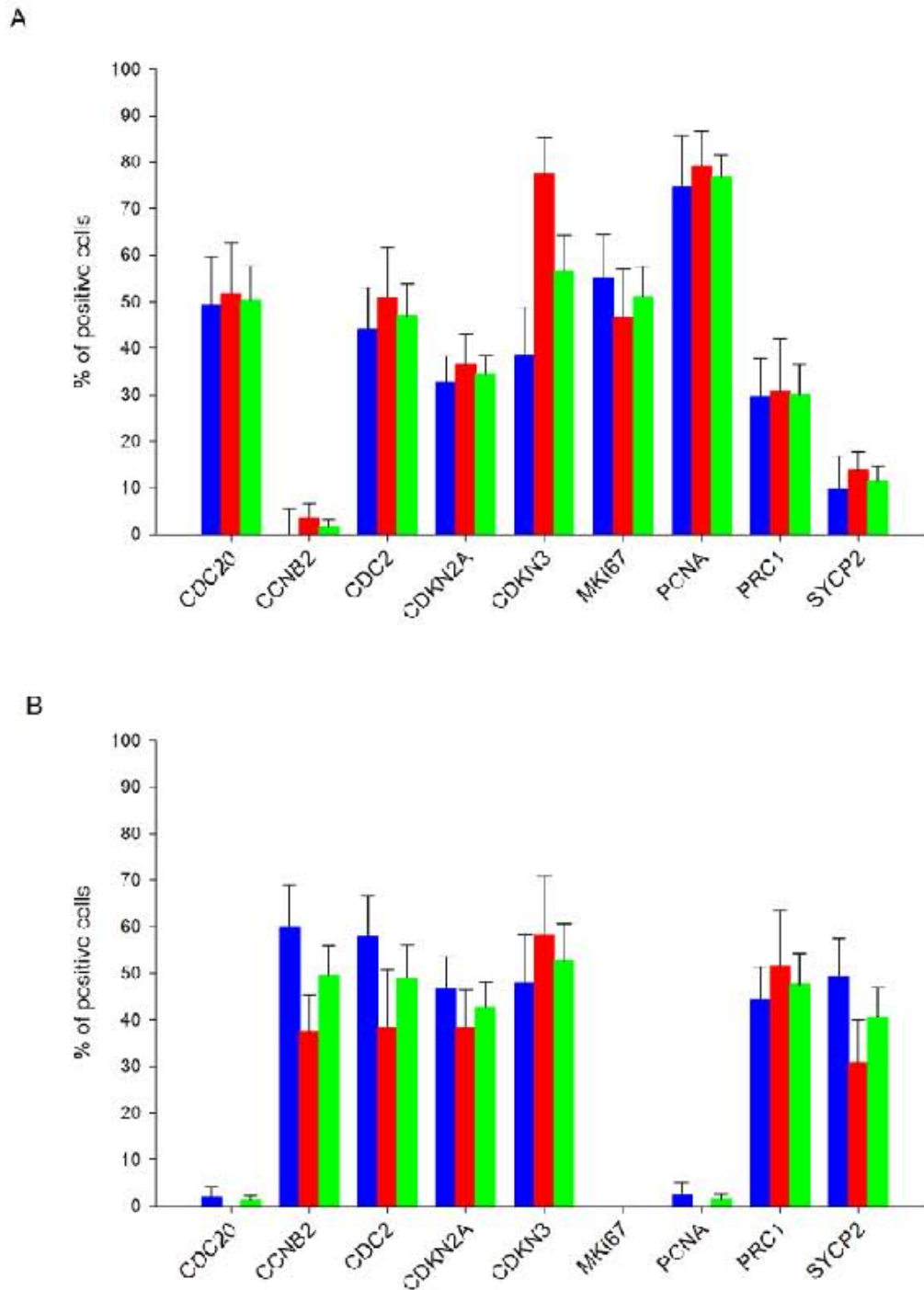


Figure 5. Percentage of tumor cells with positive signals for the markers tested by immunohistochemistry. The average percentages of tumor cells with positive signals in the nucleus (panel A) or cytoplasm (panel B) for CDC20, CCNB2, CDC2, CDKN2A, CDKN3, MKI67, PCNA, PRC1, and SYCP2 are plotted. Tumors positive for HPV16 (blue bars) are compared with tumors positive for other HPV's (red bars). As NUSAP1 was only explored in HPV16-positive CCs, the average percentages of tumor cells with positive signal in nuclei (32.5%) and cytoplasm (68.1%) were not included in this figure. The numbers for the entire set of tumors is also included (green bars). The standard error is shown.
doi:10.1371/journal.pone.0055975.g006

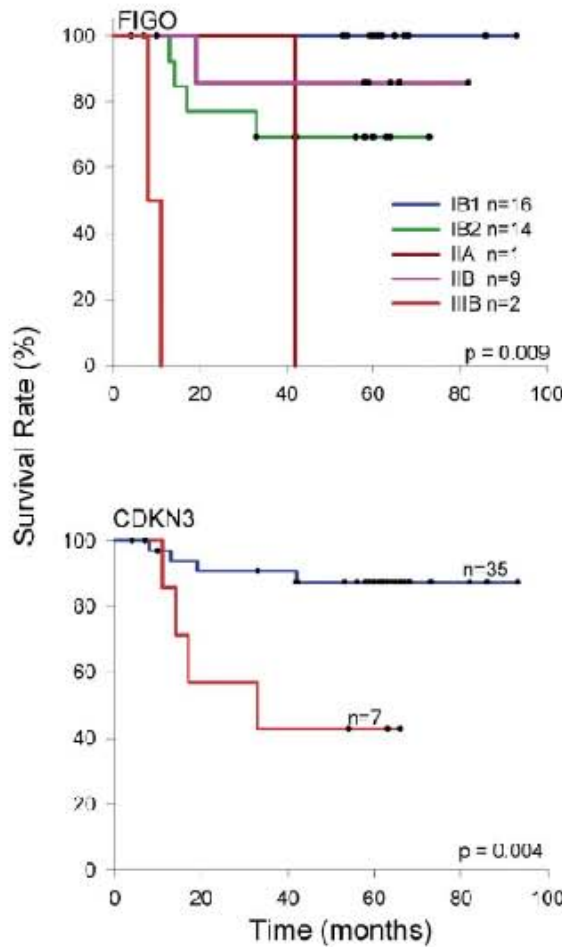


Figure 7. Survival analysis of women with cervical cancer according to FIGO staging and *CDKN3* expression. The Kaplan-Meier curves for FIGO staging and *CDKN3* are shown. Patients were followed-up for at least 42 months. For gene expression, cancer patients with higher (red line) and lower (blue line) fold change values were compared (see material and methods). The p value was calculated by comparing the curves with the log-rank test. Censored patients are labeled with black dots, but only four of them were censored before the minimal period of follow-up (42 months). doi:10.1371/journal.pone.0055975.g007

the overexpression was not uniform in the 4 cell lines, in the global analysis all genes except 1 (*SYCP2*) were observed to be upregulated; *SYCP2* was upregulated only in CaSki and SiHa cells [37]. This finding indicates that these markers are correctly predictive of cervical cancer. *CDKN3*, *NUSAP1*, and *CDC20*, along with *CDKN24*, can differentiate CC and high-grade CIN from low-grade CIN and normal cervix; therefore, they could be used as markers for screening tests. Furthermore, high expression of *CDKN3* was associated with poor survival of cancer patients; therefore, it also could be used as a survival marker. The sensitivity and specificity of *CDKN3*, *NUSAP1*, and *CDC20* to identify high-grade CIN and CC were as high as 91% and 93%, respectively. Several studies have used microarrays to identify genes associated with cervical cancer [38–46]. However, most of them included heterogeneous samples positive for different or undetermined

HPV types, a small number of tumors and controls, and essentially their design was not sufficient to identify markers for screening. Therefore, several of the 23 markers validated in this study, in which only HPV16-positive CCs were analyzed with MA, either have not been identified previously (*CDKN3*, *NUSAP1*, *SAC4*, *WISP2*) or have rarely been (*CCNE2*, *CDC20*, *CKS2*, *RRM2*) identified in other studies (Table S6). In contrast, over half (Table S6) [40] or 40% (Table S6) [43] of them have been found in individual studies, and many of them have been identified in 3 or more studies (*CDKN24*, *MCM2*, *PCNA*, *RFC4*, *SYCP2*, and *TYMS*) [46]. The fact that *CDKN3*, *CDC20* and *NUSAP1* were demonstrated upregulated in CIN2+ positive for other or undetermined HPV types by RT-PCR, suggests these genes could be considered potential markers for CC screening regardless of viral type.

Markers for screening that have high sensitivity and specificity have not been identified or reported. In addition to the conventional Pap and Liquid Base Cytology (LBC), different tests for HPV detection are used for screening. Hybrid Capture 2 Technology (HC2) is the methodology most frequently used for screening, particularly for measuring high-risk virus. This method, approved by the FDA in United States, has an average sensitivity of 95% (range, 62–98) for detecting high-grade lesions and invasive cancer. However, this methodology has low specificity, especially in young women, since the majority of infections are not associated with neoplastic lesions [15,47,48]. In women over the age of 30 years, the specificity is much higher; however, it is quite variable among studies, and depends in part on the prevalence of HPV in the study population [49]. Furthermore, in most studies, the PPV is very low, less than 30%, which indicates that only that percentage of infected women have high-grade lesions.

It is important to emphasize that the primary value of cervical cancer biomarkers and the goal of developing procedures for cervical screening is to improve the specificity rather than sensitivity relative to HPV testing [11]. Primary HPV DNA screening with cytology triage increases the specificity similar to that of conventional cytology [9,15]. However, use of this procedure in developing countries creates logistical problems, either because a high percentage of women who test positive for HPV do not return for cytology or due to the handling of samples when taking a sample for cytology from all patients at the first visit. In addition, because it prevents automatization, it seems impracticable. The simultaneous use of HC2 for high-risk viruses with a molecular method that distinguishes CIN2+ from CIN1– would increase the specificity and the PPV, with the advantages of being faster and having the potential to be automated compared to triaged cytology.

Of the markers associated with CC, p16, a tumor suppressor protein, is the most studied [11]. This protein accumulates in the nucleus and cytoplasm of cells transformed by high-risk HPV and is usually detected by IH. The amount of p16 is related to the severity of cervical neoplasia and is considered a marker of CIN2+. P16 has been successfully deployed for the classification of HPV-related disease. For cervical tissue punch and cone biopsies, IH for p16 has been reported to reduce interobserver disagreement when compared with diagnosis of H&E stained sections. P16 has also recently emerged as a sensitive and specific diagnostic adjunct for underlying CIN2+ lesions in cervical cytology specimens [11]. It consistently exhibits high sensitivity (80–95%) for detection of CIN2+; however, the specificity is lower than that for cytology (~50%) [50,51]. This is because approximately 33% of low-grade CIN lesions, those infected with high-risk HPV types, express this marker [50]. The relatively low specificity of this marker and the need for a pathologist to interpret the IH are the main reasons why this marker has not been adopted for primary screening. Recently,

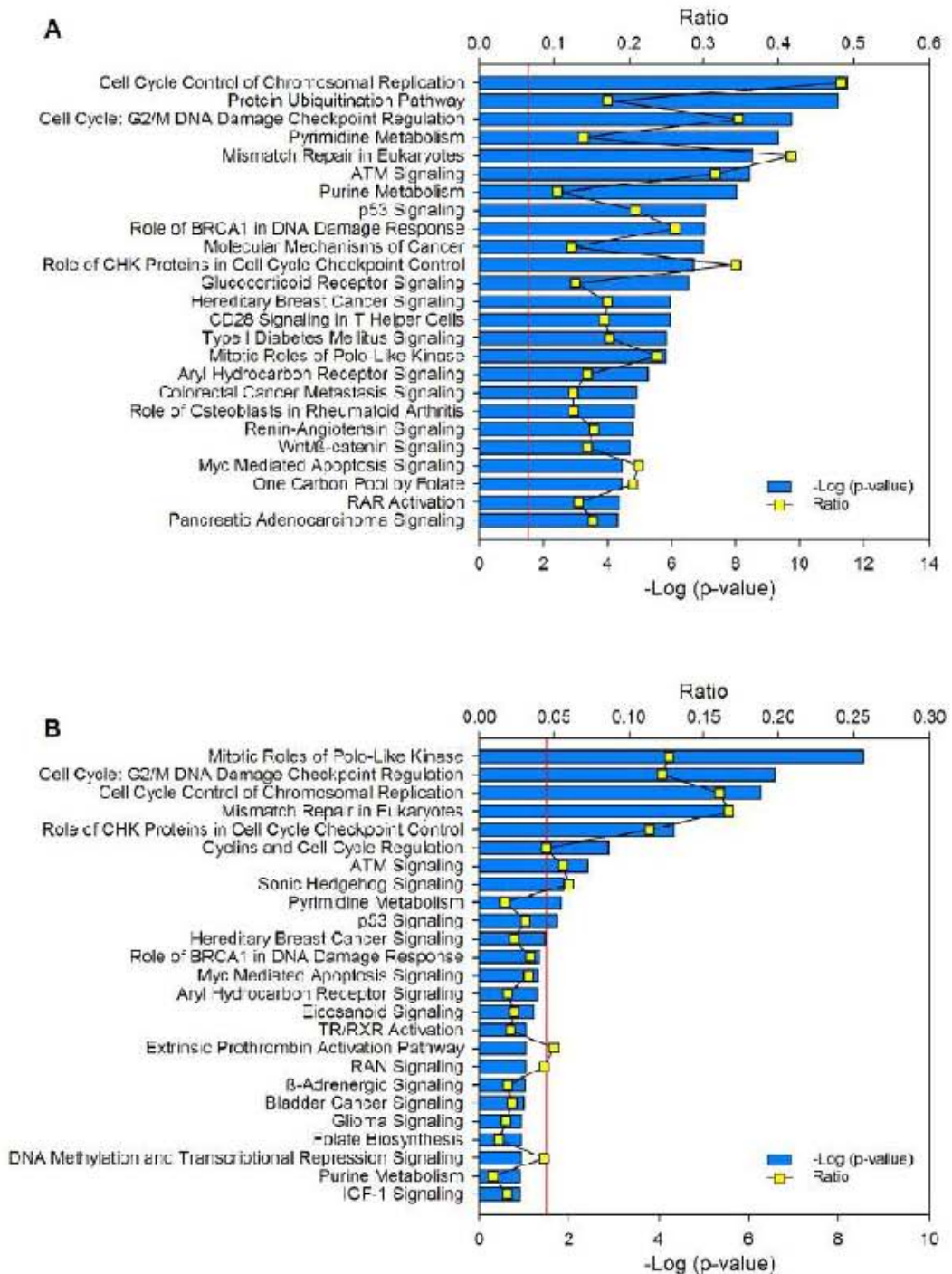


Figure 8. Canonical pathways where deregulated genes are involved. Top 25 canonical pathways identified in the set of 997 deregulated genes between the tumors and controls (A) and in the subset of the 100 best-ranked genes (50 up-regulated and 50 down-regulated; B). The canonical pathways were identified using the Ingenuity Pathway Analysis (IPA) system. The $-\log(p\text{-value})$ (blue bars) and the ratio (yellow dots) were calculated by comparing the number of genes in the pathways present in the datasets versus the human database. The $p\text{-value}$ was calculated using the chi square test or Fisher's exact test as appropriate, and the $-\log(p\text{-value})$ values >1.3 (red line) correspond to $p < 0.05$. doi:10.1371/journal.pone.0055975.g008

Table 5. DAVID functional annotation cluster analysis at the highest stringency of 997 genes deregulated in cervical cancer^a.

Cluster	Enrichment Score	Biological Process	No. Genes	p-value	Fold Change
1	16.24	Mitosis	54	3.00E-17	3.8
		nuclear division	54	3.00E-17	3.8
		M phase of mitotic cell cycle	54	6.80E-17	3.7
		organelle fission	54	1.90E-16	3.6
2	14.02	positive regulation of ubiquitin-protein ligase activity during mitotic cell cycle	28	3.60E-15	6.3
		positive regulation of ligase activity	28	2.90E-14	5.9
3	5.71	RNA splicing, via transesterification reactions	28	1.90E-06	2.8
		nuclear mRNA splicing, via spliceosome	28	1.90E-06	2.8
4	3.66	positive regulation of apoptosis	40	1.90E-04	1.7
5	3.03	mitotic cell cycle spindle assembly checkpoint	6	3.90E-04	8.3
		regulation of mitotic metaphase/anaphase transition	6	1.70E-02	4.7
6	2.81	mesenchymal cell differentiation	11	1.50E-03	3.3
7	2.54	somatic cell DNA recombination	7	3.70E-05	4.5
		somatic diversification of immune receptors	7	6.90E-03	4.0
8	2.01	carboxylic acid metabolic process	52	8.70E-03	1.4
		cellular ketone metabolic process	52	1.70E-02	1.4
9	1.96	establishment of RNA localization	14	1.00E-02	2.2
		RNA transport	14	1.00E-02	2.2
10	1.82	Glycolysis	10	3.00E-03	3.3
		hexose catabolic process	10	3.50E-02	2.2
11	1.73	embryonic epithelial tube formation	8	2.00E-02	2.8
12	1.73	somatic recombination of immunoglobulin gene segments	6	4.90E-03	5.1
		production of molecular mediator of immune response	6	4.90E-02	3.0
13	1.56	provirus integration	4	1.20E-02	7.6
		DNA migration	4	1.40E-01	3.1
14	1.54	natural killer cell mediated cytotoxicity	4	1.20E-02	7.6
		leukocyte mediated cytotoxicity	4	2.40E-02	6.1
15	1.39	regulation of caspase activity	11	3.20E-02	2.1
		regulation of peptidase activity	11	5.30E-02	2.0
16	1.38	somatic recombination of immunoglobulin genes during immune response	4	3.90E-02	5.1
		immunoglobulin production during immune response	4	4.90E-02	4.7

^aEnrichment Score is the $-\log_{10}$ of the average p-value of the terms in the cluster. Fold change is the ratio of the proportion of genes in the tested list versus the Human Gene Reference database.

doi:10.1371/journal.pone.0055975.t005

Wentzensen *et al.*, developed methods to detect p16 protein in cell lysates of cervix exudates using ELISA. The sensitivity of this ELISA method for the identification of high-risk lesions was 84%, and the specificity was 87% [52]. In agreement with these data, the specificity of *CDKN2A* mRNA detection, which encodes p16, in screening for CIN2+ was very close (93%, Table 4).

Two new markers identified in this work (*CDKN3* and *NUSAP1*), along with *CDKN2A*, showed a high specificity (93%) and PPV (93.4%); therefore, they might be good candidates to use with HC2 as a first-line strategy in a screening program. The scope of this study was to perform a feasibility evaluation to ascertain whether determining the mRNA levels of novel genes in cervical samples would allow for the identification of high-grade CIN or invasive lesions with high sensitivity and specificity. However, the potential sensitivities reported in this analysis are most likely overestimated compared to those likely to be found in clinical practice, as those with CIN2+ have a higher proportion of cervical

cancer (which is easy to identify) than that expected in any screening setting. In contrast, the specificity seems to be underestimated, given that a large proportion of CIN1+ had CIN1. Therefore, we did not expect to obtain conclusive data on the sensitivity, specificity or predictive values of the assays. Further studies are needed to determine the levels of *CDKN3*, *NUSAP1*, and *CDC20* mRNA or protein in cervical samples from a screening population to obtain information about the predictive values and to define the optimal trade-off between sensitivity and specificity for the detection of CIN2+.

PRC1, *CCNB2*, and *SYCP2* are markers exclusively associated with invasive cervical cancer. Together with *NUSAP1*, *CDKN3*, and *CDC20*, these genes represent potential specific targets for the treatment of advanced CC, particularly *CDKN3*, which was found to be associated with poor survival. These genes encode proteins involved in the cell cycle, specifically in the M phase (mitosis and cytokinesis). According to the IHC data, approximately 30% of

Table 6. DAVID functional annotation cluster analysis at the highest stringency of the 100 genes most deregulated in cervical cancer compared with normal cervical epithelium*.

Cluster	Enrichment Score	Biological Process	No. Genes	p-value	Fold Change
1	13.89	Mitosis	19	1.00E-14	12.0
		nuclear division	19	1.00E-14	12.0
		M phase of mitotic cell cycle	19	1.40E-14	12.0
		organelle fission	19	2.00E-14	12.0
2	3.49	regulation of ubiquitin-protein ligase activity	6	2.10E-04	11.0
3	2.07	negative regulation of ubiquitin-protein ligase activity	5	1.20E-03	11.0
4	2.83	positive regulation of mitosis	4	6.70E-04	23.0
		positive regulation of cell cycle	4	7.20E-03	10.0
5	2.75	establishment of mitotic spindle localization	3	1.30E-03	53.0
6	2.50	anaphase-promoting complex-dependent proteasomal ubiquitin-dependent protein catabolic process	5	1.10E-03	11.0
		regulation of ubiquitin-protein ligase activity during mitotic cell cycle	5	1.50E-03	10.0
7	2.31	negative regulation of protein modification process	6	1.40E-03	7.1
		negative regulation of protein metabolic process	6	9.70E-03	4.6
8	1.58	regulation of protein modification process	8	4.40E-03	3.5
		regulation of cellular protein metabolic process	8	4.70E-02	2.4
9	1.53	nucleosome assembly	4	2.10E-02	6.8
		chromatin assembly	4	2.50E-02	6.6
10	1.42	regulation of protein kinase activity	7	3.30E-02	2.5
11	1.31	regulation of phosphorylation	8	4.30E-02	2.4

*Enrichment Score is the $-\log_{10}$ of the average p-value of the terms in the cluster. Fold change is the ratio of the proportion of genes in the tested list versus the Human Gene Reference database.

doi:10.1371/journal.pone.0055975.t006

tumor cells in CC could be in the M phase. These genes participate in anaphase control, chromosome segregation, and mitotic entrance/exit. While activation of cyclin dependent kinases (Cdks) drives cells into mitosis, mitotic exit depends on inhibition of Cdks activity, mainly through degradation of mitotic cyclins by the anaphase-promoting complex (APC/C) and accumulation of Cdk inhibitor proteins, and dephosphorylation of proteins phosphorylated by CDKs. Four (CCNB2, CDC20, CDKN3, PRC1) of the six proteins validated in this paper seem essential in this process. Cyclin B2 (CCNB2), like cyclin B1 (CCNB1), binds to CDK1 (CDK2) to form the complex M-CDK, which is essential for control of the cell cycle at the G2/M transition. However, while cyclin B1-CDK1 causes chromosome condensation, reorganizes microtubules, and disassembles the nuclear lamina and the Golgi apparatus, cyclin B2-CDK1 is restricted to the cytoplasm and disassembles the Golgi apparatus during mitosis [53,54]. In agreement with these data, cyclin B2 was localized exclusively in the cytoplasm of the CCs examined in this paper (Figure 5). Interestingly, the expression of cyclin B1 in these tumors did not differ from that in the control samples (Table S3). This cyclin is degraded by the APC/C, a key regulator of the metaphase-to-anaphase transition, to allow progression of mitosis from metaphase to anaphase [55]. *CCNB2* has been scantily associated with cervical cancer [40]; however, it has been reported to be associated with other types of cancer. For instance, it is upregulated in cancers of the colon [56], lung, and digestive tract [57]. The increased amount of CDC20, a key regulatory protein of APC/C complex during anaphase, could explain the absence of cyclin B1. CDC20, together with UBE2C (also known as UBCH10), which was also increased in CC (Table S3), is required

for full ubiquitin ligase activity of the APC/C complex and may confer substrate specificity upon the complex. CDC20 is negatively regulated by MAD2L1 and BUB1B (also known as BUBR1). In metaphase the MAD2L1-CDC20-APC/C ternary complex is inactive, while in anaphase the CDC20-APC/C binary complex is active in degrading substrates. Interestingly, the MAD2L1 and BUB1B transcripts were also increased in CC (Table S3) suggesting that the corresponding proteins could be increased and prevent activation of APC/C. However, part of the CDC20 protein could remain free to bind and activate APC/C, as has been shown in transfected cells expressing the E6/E7 proteins [55]. CDC20 has been found to be upregulated in lung, pancreatic, and gastric cancers [53], as well as in CC [40,55]. CDKN3 is a dual-specificity protein phosphatase of the Cdc14 phosphatase group that interacts with CDK1 (CDC2) and inhibits their activity [60,61]. CDKN3 and other Cdc14 phosphatases have not been well studied; however, they seem to be essential for antagonizing Cdk activity in late mitosis, allowing cells to exit mitosis in telophase. Regulation of cytokinesis may be the 1 conserved function of the Cdc14 phosphatases. Although overexpression of CDKN3 has been associated with inhibition of cell proliferation in colon cancer cell lines [52], it has also been found to be overexpressed in breast, prostate, and lung cancers [63–65]. In agreement with our data, *CDKN3*, along with other genes, has been found to be associated with lower survival of patients with lung adenocarcinomas [63]. This is the first report in which *CDKN3* was associated with cervical cancer (Table S6). PRC1 is involved in cytokinesis and is essential for controlling the spatiotemporal formation of the midzone and successful cytokinesis [66,67]. It is required for kinesin-family member 14 (KIF14)

[68] and polo-like kinase 1 (PLK1) [69] localization to the central spindle and midbody. The suppression of FRC1 blocks cell division. The transcription of *FRC1* is repressed by p53 and is one of the routes by which p53 stops the cell cycle at the G2/M checkpoint [70]. Since the E6 oncoprotein of HPV16 induces degradation of p53 in proteasomes, it is likely that in cervical carcinomas FRC1 is being overexpressed via this mechanism. It has been reported to be associated with liver cancer [71] and CC [40,42]. NUSAP1 is a nuclear-spindle-associated protein that plays a role in spindle microtubule organization. This gene has not been described as associated with CC, but has been found to be upregulated in breast and melanoma cancers [72]. SYCP2 is a major component of the synaptonemal complex. This complex promotes that double strand breaks (DSB) are repaired by the homologous recombination pathway in meiosis [73]. The high levels of *SYCP2* expression in the CCs examined in this work suggests that DSB are very common in some CC samples and that SYCP2 could be involved in DSB repair by the stimulation of homologous recombination pathway. Interestingly, this gene has been found to be upregulated in CC [45,46] and oropharyngeal squamous cell carcinomas positive for HPV16, but not in HPV-negative carcinomas [74].

Cell cycle is the main process altered in CC and is top ranked in all CC papers where biological processes have been analyzed [46]. Similarly, in the present paper, when the gene dataset was analyzed using the DAVID tool at medium stringency, the cell cycle process was shown to be the most enriched and it ranked at the top of the list (Table S5). However, the fact that M-phase processes were the most enriched in our dataset when the analysis was done at high stringency, suggests that the M-phase is the main altered cell-cycle phase in CC. These findings are consistent with the alterations in the cell cycle and mitosis caused by HPV *in vitro* [59,75,76] and correlated in few CC studied [59]. The E6 and E7 oncoproteins of high-risk HPVs induce numerous mitotic defects, including multipolar mitoses, chromosomal missegregation, anaphase bridges, and aneuploidy. Although cells with abnormal mitoses are normally targeted for cell death, E6 and E7 act cooperatively to allow cells with abnormal centrosomes to accumulate by relaxing the G2/M checkpoint response and inhibition of apoptotic signaling [76]. In agreement with these data, the canonical pathways of G2/M DNA Damage Checkpoint Regulation and the Role of CHK Proteins in Cell Cycle Checkpoint Control ranked at the second and fifth positions of the altered canonical pathways in CC. On the other hand, E6 and E7 induce mechanisms to avoid mitosis checkpoint. The E6/E7 genes have been shown to induce the overexpression of GDC20 and UBC10, which activate the APC/C ubiquitin ligase complex [55]. The enrichment of positive regulation of ubiquitin-protein ligase activity during mitotic cell cycle found in CCs (Table 5) completely agree with these *in vitro* results.

Inhibition of mitosis is a well-known strategy to combat cancers. Drugs that perturb the process of cell division have proved to be effective anticancer therapies. Well-known examples of these drugs are those that perturb the formation of the mitotic spindle, such as taxanes and vinca alkaloids. However, they have remarkably low therapeutic indices and narrow therapeutic windows. Their efficacy is restricted because they also perturb the microtubule network of non-dividing cells, causing neurotoxic effects and affecting endothelial cell function. To resolve this issue, a new generation of antimitotic agents has been developed that target kinases and kinases with unique roles in mitosis, such as KIF11, PLK1, and aurora kinase A (AURKA) [69]. Interestingly, the transcripts of these 3 genes were also upregulated in the CCs (Table S3), *AURKA* ranked in 19th place, *KIF11* ranked in 72nd

place, and *PLK1* ranked in 263rd place. Therefore, these drugs could be tested for the treatment of cervical cancer. On the other hand, the high FC of the novel genes calculated in this work, especially *CDKN3*, *CTC3*, and *STUP2*, compared with the control samples, makes these genes potential targets for CC therapy. However, it is still necessary to demonstrate whether they are indispensable for tumor growth.

Supporting Information

Figure S1 Correlation of expression intensity of 997 genes examined by HG-Focus and HG-ST1.0 microarrays. The \log_2 values of the standardized intensity signals (RMA values) of 997 genes in a typical tumor (R230) examined by the 2 microarrays were plotted. The linear trend (black line) is included, which was calculated with Person's correlation test: $r=0.78$, $p<1\times 10^{-13}$. (TIFF)

Figure S2 Validation of gene expression of 3 genetic markers by qRT-PCR. The intensity of gene expression, expressed in \log_2 values, is shown in box plots. Expression of the 3 novel downregulated genes revealed in this study (*CFD*, *EDN3*, *WISP2*) associated with CC are compared between healthy cervical epithelium (Control, $n=25$) and invasive CC (Tumor, $n=44$). See legend of Figure 4. (TIFF)

Figure S3 Histological analysis of NUSAP1. Protein expression was determined by immunohistochemistry using sections from formalin-fixed, paraffin-embedded tissue. Representative experiments in adeno cell carcinomas (left panel) and squamous cell carcinomas (right panel) are shown. The specific signals are shown as brown staining (counterstained with hematoxylin; original magnification, $\times 400$; bars, 10 μm). (TIFF)

Table S1 Summary of clinical data, HPV type and methodology to explore gene expression. SCC, squamous cell carcinoma; ACC, adenocarcinoma; ASCC, Adenosquamous; NCE, normal cervical epithelium; CIN, cervical intraepithelial neoplasia; MA, microarray; MA1, HG-Focus; MA2, HG-ST1.0; qRT-PCR, real-time RT-PCR; IH, Immunohistochemistry; ND, not done. a. Control samples codes ending with X are exocervix and ending with E are endocervix. (XLSX)

Table S2 TaqMan gene expression assays from Applied Biosystems. (XLSX)

Table S3 List of the 997 genes differentially expressed in cervical carcinomas compared with normal cervical epithelium. a. Genes were ranked by Δ -score and those marked in bold were selected for analysis with real time RT-PCR. b. The tumor/control fold change (FC) of each gene was calculated using the mean values of signal intensity obtained with the SAM method. c. Genes which have been used as biomarkers in different diseases, according with the IPA Ingenuity system: D, Diagnosis; DP, Disease Progression; P, Prognosis; E, Efficacy; RT, Response to Therapy; UP, Unspecified Application; S, Safety. (XLSX)

Table S4 Analysis of 10 proteins in 26 CC and 10 normal cervical epithelium samples by IH and calculus of sensitivity, specificity and predictive values. (XLSX)

Table S5 DAVID functional annotation cluster analysis at medium stringency of 997 genes deregulated in cervical cancer². * Enrichment Score is the $-\log_{10}$ of the average p-value of the terms in the cluster. Fold change is the ratio of the proportion of genes in the tested list versus the Human Gene Reference database.
(XLSX)

Table S6 Comparison of the 23 genes identified and validated in this study with previously reported microarray analysis. a. Ccl1 line immortalized human keratinocyte lines NIKS-16. b. Detected in HeLa cells.
(XLSX)

References

1. Bosch FX, Lorincz A, Muñoz N, Meijer CJ, Shah KV (2002) The causal relation between human papillomavirus and cervical cancer. *J Clin Pathol* 55: 244–252.
2. Schiffman M, Wertzman N, Wacholder S, Kinney W, Gage JC, et al. (2011) Human papillomavirus testing in the prevention of cervical cancer. *J Natl Cancer Inst* 103: 356–35.
3. Oster AG (1995) Natural history of cervical intraepithelial neoplasia: a critical review. *Int J Gynecol Pathol* 12(2): 186–92.
4. de Sanjose S, Quinó W, Alemany L, Gemen DT, Klumpp-Schadler JE, et al. (2010) Human papillomavirus genotype attribution in invasive cervical cancer: a retrospective cross-sectional worldwide study. *Lancet Oncol* 11: 1048–56.
5. Kjaer SK, Sigurdson K, Iversen OE, Hernandez-Avila M, Wheeler CM, et al. (2009) A pooled analysis of continued prophylactic efficacy of quadrivalent human papillomavirus (Types 6/11/16/18) vaccine against high-grade cervical and external genital lesions. *Cancer Prev Res (Phila)* 2: 869–78.
6. Lehtinen M, Paavonen J, Wheeler CM, Jaisankaran U, Garland SM, et al. (2012) Overall efficacy of HPV-16/18 AS04-adjuvanted vaccine against grade 3 or greater cervical intraepithelial neoplasia: 4-year end-of-study analysis of the randomised, double-blind PATRICIA trial. *Lancet Oncol* 13: 893–903.
7. Romanowski B (2011) Long-term protection against cervical infection with the human papillomavirus: review of currently available vaccines. *Hum Vaccin* 7: 19–29.
8. Curick J (2010) Long-term cervical cancer prevention strategies across the globe. *Gynecol Oncol* 117: S1–4.
9. Markowitz LE, Dunne EF, Sarnaik M, Lawson HW, Chenou H, et al. (2007) Quadrivalent Human Papillomavirus Vaccine: Recommendations of the Advisory Committee on Immunization Practices (ACIP), MMWR Recomm Rep 56: 1–24.
10. Nahmen K, Lehtinen J, Nannija P, Sellers J, Lehtinen M (2011) Aspects of prophylactic vaccination against cervical cancer and other human papillomavirus-related cancers in developing countries. *Infect Dis Obstet Gynecol* 2011: 673858.
11. Hwang SJ, Stroyer KR (2012) Biomarkers of cervical dysplasia and carcinoma. *J Oncol* 2012: 301280.
12. Ferlay J, Shin HR, Bray F, Forman D, Maier C, et al. (2008) GLOBOCAN 2008. *Cancer Incidence and Mortality Worldwide: IARC Cancer Base No. 10* [Internet]. Lyon, France: International Agency for Research on Cancer; 2010. Available: <http://globocan.iarc.fr>.
13. McCrory DC, Maczhar DE, Bastar L, Datta S, Hamedelad V, et al. (2009) Evaluation of cervical cytology. *Evid Rep Technol Assess (Surrey)* 1995: 1–6.
14. Wright TC (2007) Cervical cancer screening in the 21st century: is it time to retire the PAP smear? *Clin Obstet Gynecol* 50: 313–23.
15. Leinonen M, Nieminen P, Kotaniemi-Talonen L, Malin N, Tarkkanen J, et al. (2009) Age-specific evaluation of primary human papillomavirus screening vs conventional cytology in a randomized setting. *J Natl Cancer Inst* 101: 1612–23.
16. Gupta N, Srinivasan R, Rajawansi A (2010) Functional Biomarkers in cervical precancer: an overview. *Diagn Cytopathol* 38: 619–25.
17. Andrae B, Andersson TM, Lambert PC, Kessel L, Silfverdal L, et al. (2012) Screening and cervical cancer cure: population based cohort study. *BMJ* 344: e900.
18. Tort D, Trussardi L (2011) Oncogene addiction as a foundational rationale for targeted anti-cancer therapy: promises and perils. *EMBO Mol Med* 3: 622–626.
19. Krighi Z, Liu H, Shklar KM (2010) Targeting the cancer genome through polypharmacology. *Nature Cancer Rev* 12: 130–137.
20. Saxena R, Dwevedi A (2012) Ecd5 family receptor inhibitors as therapeutic agents in breast cancer: current status and future clinical perspective. *Med Res Rev* 32: 166–215.
21. Yarden Y, Peles C (2012) The ERBB network: at last, cancer therapy meets systems biology. *Nat Cancer Rev* 12: 553–563.
22. Pictorini S (2009) Revised FIGO staging for carcinoma of the vulva, cervix, and endometrium. *Int J Gynecol Obstet* 105: 103–4.

Acknowledgments

This study was performed as part of the requirements needed to obtain the Ph.D. degree of AME at the Graduate Program in Sciences, Universidad Nacional Autónoma de México (UNAM). The authors thank Marco Gudiño for his help in histological imaging.

Author Contributions

Evaluated clinically healthy women and patients with CIN lesions and provided the cervical specimens from these groups: SMC, MBI. Evaluated clinically CIN patients and provided the cervical specimens from these patients: ERB. Performed the social work: MB. Conceived and designed the experiments: JB, AME, MGE. Performed the experiments: AME, IMM, EJT, AA, IP, CS. Analyzed the data: JB, AME, AA, MCE, EM, AAV, JCV, PA, NVS, MB. Contributed reagents/materials/analysis tools: MGE, EM, SK. Wrote the paper: JB, MGE, AME.

23. Manos MM, Ting Y, Wright DK, Lewis AJ, Beeber TR, et al. (1989) The Use of Polymerase Chain Reaction Amplification for the Detection of Genital Human Papillomavirus. *J Clin Pathol* 42: 209–214. Cold Spring Harbor, New York.
24. van den Brule AJ, Meijer CJ, Bakula V, Koeman P, Walboomers JM (1996) Rapid detection of human papillomavirus in cervical scrapes by combined general p16^{ink4}-mediated and type-specific polymerase chain reaction. *J Clin Microbiol* 34: 3730–37.
25. Yoshikawa H, Kawana T, Kitagawa K, Mizuno M, Yoshikura H, et al. (1991) Detection and typing of multiple genital human papillomaviruses by DNA amplification with consensus primers. *Cancer Sci* 82: 524–531.
26. Cejudo M, McWilliam H, Li W, Valencia F, Squizzato S, et al. (2010) A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res* 38: W695–W699.
27. Bliznjec M, Mirar M, Nadro R (2007) FlexArray: A statistical data analysis software for gene expression microarrays. *Genome Quebec*, Montreal, Canada. URL: <http://genomeq.usherbrooke.ca/FlexArray>.
28. Tusher VG, Tibshirani R, Chu G (2001) Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci USA* 98: 5116–5121.
29. D'Amico B, Vasdolez J, Hellmann J (2010) Accurate and objective copy number profiling using real-time quantitative PCR. *Methods* 50(4): 262–70.
30. Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 29: 2001–2007.
31. Huang DW, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4: 44–57.
32. Huang DW, Sherman BT, Lempicki RA (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* 37: 1–13.
33. Roselli V, Cheek I, Bai J (2004) Statistics review 13: receiver operating characteristic curves. *Crit Care* 8: 508–12.
34. Leisen AG, McCluskey A (2008) Statistics VI: Clinical tests: sensitivity and specificity. *Contin Educ Anesth Crit Care* 3: 221–30.
35. Deha D, Tonko KC, Stroyer KR (2007) Human papillomavirus testing and molecular markers of cervical dysplasia and carcinoma. *Cancer* 111: 1–14.
36. Faa J, Upadhye S, Weston A (2006) Understanding receiver operating characteristic (ROC) curves. *QJEM* 8: 19–20.
37. Vazquez-Mena O, Medina-Marin L, Juárez-Torres E, Valeria Barrón E, Espinosa A, et al. (2012) Amplified genes can be over expressed, unchanged or down regulated in cervical cancer cell lines. *PLoS ONE* 2012: 7(3), e32667.
38. Chen Y, Miller C, Mosher R, Zhao X, Doss J, et al. (2005) Identification of cervical cancer markers by cDNA and tissue microarrays. *Cancer Res* 65: 1927–35.
39. Narzary G, Eorand V, Chaganti S, Arias-Pañizo H, Nandini SV, et al. (2007) Gene dosage alterations revealed by cDNA microarray analysis in cervical cancer: identification of candidate amplified and overexpressed genes. *Genes Chromosomes Cancer* 46: 373–84.
40. Eddy C, Shaffer M, Trafer D, Swendsy N, Tracis J, et al. (2005) Identification of a proliferation gene cluster associated with HPV E6/E7 expression level and viral DNA load in invasive cervical carcinoma. *Oncogene* 24: 3094–3104.
41. Grais D, Funk MC, Chuang EY, Feng S, Huettner PC, et al. (2007) Profiling microdissected epithelium and stroma to model genomic signatures for cervical carcinogenesis: accommodating for covariates. *Cancer Res* 67: 7113–23.
42. Zhai Y, Kueck R, Nan B, Osa I, Weiss SJ, et al. (2007) Gene expression analysis of premalignant and invasive cervical squamous cell carcinomas identifies HOXC10 as a key indicator of invasion. *Cancer Res* 67: 11613–22.
43. Pyeon D, Newton M, Lambert P, den Boon J, Sengupta S, et al. (2007) Fundamental Differences in Cell Cycle Deregulation in Human Papillomavirus Positive and Human Papillomavirus Negative Head/Neck and Cervical Cancers. *Cancer Res* 67(10): 10193–7.

44. Biewenga F, Buik M, Moerland P, Ver Lorenz van Themaat E, van Kampen AIL, et al. (2008) Gene expression in early stage cervical cancer. *Gynecologic Oncology* 100(2): 520–526.
45. Scott L, Narayan G, Nandula S, Arias-Pulido H, Subramaniam S, et al. (2008) Identification of Copy Number Gain and Overexpressed Genes on Chromosome Arm 20q by an Integrative Genomic Approach in Cervical Cancer: Potential Role in Progression. *Genes, Chromosomes, Cancer* 47(9): 755–765.
46. Buitrago Pérez A, Cuzick L, Vázquez-Carballe A, Paranteo JM, García Escudero R (2009) Molecular Signature of HPV-Induced Carcinogenesis: p16, p53 and Gene Expression Profiling. *Current Genomics* 10(1): 26–34.
47. Baleriola C, Millar D, Melki J, Coulston N, Altman P, et al. (2008) Comparison of a novel HPV test with the Hybrid Capture II (hcII) and a reference PCR method shows high specificity and positive predictive value for 13 high-risk human papillomavirus infections. *J Clin Virol* 42 22–6.
48. Whitford EP, Vesce KK, Eker M, Liu JS, Sengen CA, et al. (2011) Liquid-based cytology and human papillomavirus testing to screen for cervical cancer: a systematic review for the U.S. Preventive Services Task Force. *Ann Intern Med* 155: 687–97.
49. Giorgi-Rossi P, Franceschi S, Ronco G (2012) HPV prevalence and accuracy of HPV testing to detect high-grade cervical intraepithelial neoplasia. *Int J Cancer* 130: 1387–94.
50. Tsoumpas I, Arzyu M, Kyggen M, Weitzenzer N, Kolopoulou G, et al. (2009) p16INK4a immunostaining in cytological and histological specimens from the uterine cervix: a systematic review and meta-analysis. *Cancer Treat Rev* 35: 210–20.
51. Galgani MT, Castle PE, Athlin KA, Bris WK, Nassau SR, et al. (2010) Using biomarkers as adjunctive standards in the diagnosis of cervical biopsies. *Am J Surg Pathol* 34: 1077–87.
52. Weitzenzer N, Hampf M, Herkert M, Reikien A, Frank MJ, et al. (2006) Identification of high-grade cervical dysplasia by the detection of p16INK4a in cell lysates obtained from cervical scrapes. *Cancer* 107: 2307–13.
53. Davarian VM, Orrechoa S, Lowe M, Pardi R, Pines J (2001) The localization of human cyclins B1 and B2 determines CDK1 substrate specificity and neither enzyme requires MEK to disassemble the Golgi apparatus. *J Cell Biol* 152: 945–58.
54. Gong D, Ferrell JE (2010) The roles of cyclin A2, B1, and B2 in early and late mitotic events. *Mol Biol Cell* 21: 3145–51.
55. Patel D, McCance DJ (2010) Compromised spindle assembly checkpoint due to altered expression of Ubr1 and Cdk20 in human papillomavirus type 16 E6- and E7-expressing keratinocytes. *J Virol* 84: 10936–46.
56. Park SH, Yu GR, Kim WH, Moon WS, Kim JH, et al. (2007) NF- κ B-dependent cyclin B2 expression in colorectal adenocarcinoma. *Clin Cancer Res* 13: 650–67.
57. Mo ML, Chen Z, Li J, Li HL, Sheng Q, et al. (2010) Use of serum circulating G2NB2 in cancer surveillance. *Int J Biol Markers* 25: 236–42.
58. Nakayama KI, Nakayama K (2006) Ubiquitin ligases: cell-cycle control and cancer. *Nat Rev Cancer* 6: 269–81.
59. Teisler S, Ben Khalfa J, Morf M, Pantke F, Desaintes C, et al. (2007) New E6/E6E2 Pathway, Together with a Strong E7/E2F Mitotic Pathway, Modulates the Tumorigenesis in Cervical Cancer Cells. *Journal of virology* 81(17): 9363–9376.
60. Demerick DJ, Matsuura S, Hayashi CJ, Okazaki K, Xiang Y, et al. (1995) Chromosomal mapping of the genes for the human cell cycle proteins cyclin C (CCNG), cyclin E (CCNE), p21 (CDKN1) and KAP (CDKN3). *Cytogenet Cell Genet* 69: 190–2.
61. Gyuris J, Coleman K, Chertkov H, Brent R (1993) Cdk1, a human G1 and S phase protein phosphatase that associates with Cdk2. *Cell* 73: 391–403.
62. Galarró O, Spisák S, Sjöqvist F, Tóth K, Solymosi N, et al. (2010) Reversal of gene expression changes in the colorectal normal-adenoma pathway by NS390 selective COX2 inhibitor. *Br J Cancer* 102: 765–73.
63. MacDonnell DM, Kuznetsov NN, Pirotski SP, Edwards DC, Polzella CA, et al. (2010) MUC1-associated proliferation signature predicts outcomes in lung adenocarcinoma patients. *BMC Med Genomics* 3: 16.
64. Jilka SC, Dahl N, Hardy S, Tenenblay ML (2011) Inside the human cancer lysosome phosphatase. *Nat Rev Cancer* 11: 35–40.
65. Taylor KJ, New AH, Liang L, Farahan D, Muir M, et al. (2010) Dynamic changes in gene expression in vivo predict prognosis of tamoxifen-treated patients with breast cancer. *Breast Cancer Res* 12: R50.
66. Melloni C, Klemm JP, Jiang W, Schreiber C, Hunter T, et al. (2007) PRC1 is a microtubule-binding and -bundling protein essential to maintain the mitotic spindle midzone. *J Cell Biol* 157: 1175–89.
67. Subramaniam R, Wilson-Kubacki EM, Arthur CP, Bick MJ, Campbell EA, et al. (2010) Insights into antiparallel microtubule crosslinking by PRC1, a conserved microtubule-associated binding protein. *Cell* 142: 433–42.
68. Grunberg U, Neef K, Li X, Chan EH, Chahumian RE, et al. (2005) KIF14 and chro- kinase act together to promote efficient cytokinesis. *J Cell Biol* 172: 363–72.
69. Lee SM, Vesce FF, Madern RH (2010) Shared and separate functions of polo-like kinases and aurora kinases in cancer. *Nat Rev Cancer* 10: 325–41.
70. Li C, Lin M, Liu J (2004) Identification of PRC1 as the p53 target gene uncovers a novel function of p53 in the regulation of cytokinesis. *Oncogene* 23: 9336–47.
71. Wang SM, Ooi LL, Hui KM (2011) Upregulation of Rac GTPase-activating protein 1 is significantly associated with the early recurrence of human hepatocellular carcinoma. *Clin Cancer Res* 17: 6040–51.
72. Kretschmer C, Strasser-Kock A, Hiedinger T, Schoenberg W, Schleg PM, et al. (2011) Identification of early molecular markers for breast cancer. *Mol Cancer* 10: 15.
73. Li XC, Bolam-Fluas E, Schmeffel JC (2011) Genetic evidence that synaptonemal complex axial elements govern recombination pathway choice in mice. *Genetics* 185: 71–82.
74. Martínez I, Wang J, Helzer KF, Ferris RL, Khan SA (2005) Identification of differentially expressed genes in HPV-positive and HPV-negative oropharyngeal squamous cell carcinomas. *Eur J Cancer* 43: 415–32.
75. Tsiang T, Beato M, Dawson C, Mori M, Tsiang S, et al. (2004) A Genomic Approach Reveals a Novel Mitotic Pathway in Papillomavirus Carcinogenesis. *Cancer Res* 64(3): 359–365.
76. Moody CA, Laikins LA (2010) Human papillomavirus oncoprotein pathways to transformation. *Nat Rev Cancer* 10: 550–60.