

11281



UNIVERSIDAD NACIONAL AUTONOMA
DE MEXICO

PROGRAMA DE DOCTORADO EN CIENCIAS BIOMEDICAS
FACULTAD DE MEDICINA

"PATRON DE EXPRESION DE CELULAS SIN NEOPLASIA Y
CELULAS INVASORAS DEL CERVIX UTERINO HUMANO"

TESIS QUE PARA OBTENER EL GRADO DE:

DOCTOR EN CIENCIAS BIOMEDICAS

PRESENTA LA INGENIERO BIOTECNOLOGO AMBIENTAL:

GUELAGUETZA VAZQUEZ ORTIZ

TUTOR: DR. MAURICIO SALCEDO VARGAS

LABORATORIO DE ONCOLOGIA GENOMICA, UNIDAD DE INVESTIGACION
MEDICA EN ENFERMEDADES ONCOLOGICAS, CENTRO MEDICO NACIONAL
SIGLO XXI - IMSS.



IMSS

MEXICO, D.F.

2005.

m343158



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.

El presente trabajo fue realizado en el Laboratorio de Oncología Genómica de la Unidad de Investigación Médica en Enfermedades Oncológicas del Centro Médico Nacional Siglo XXI-IMSS. Durante la realización de este trabajo, se contó con el apoyo económico del Consejo Nacional de Ciencia y Tecnología a través de los proyectos No. 34686-M y 7114; durante este trabajo la estudiante fue becaria de CONACyT No.153069. Además se contó con las becas complementarias de la Dirección General de Estudios de Postgrado de la Universidad Nacional Autónoma de México. Parte del análisis de los datos fueron obtenidos en colaboración con el Laboratorio de microarreglos del Instituto de Regulación Genómica y con el Laboratorio de Biología Molecular y Farmacia de la Universidad de Barcelona en España.

AGRADECIMIENTOS

A mis papás por haberme apoyado cuando lo he necesitado, por haberme alentado a seguir y por todo el cariño que me han dado durante toda mi vida.

A mis compañeros del laboratorio Pato, Karla, Brenda, Rich, Gordo, Patito, Carlos y Luz por hacerme la ciencia y la vida, más fácil y divertida.

A los miembros de mi comité académico, Dr. Fabio Salamanca Gómez y Dr. Héctor Mayani Viveros, especialmente a mi tutor el Dr. Mauricio Salcedo por su confianza, apoyo y amistad.

Al Instituto Mexicano del Seguro Social, a la Dirección General de Estudios de Postgrado de la Universidad Nacional Autónoma de México y al Consejo Nacional de Ciencia y Tecnología y al Dr. Carlos Ciudad de la Universidad de Barcelona por darme la oportunidad de continuar con mi superación académica.



Doctorado en Ciencias Biomédicas

Voto Aprobatorio

27 Febrero, 2005

Ing. Leopoldo Silva Gutiérrez
Director General de Administración Escolar,
UNAM
P R E S E N T E

Estimado Ing. Silva,

Después de haber evaluado la Tesis titulada "PATRON DE EXPRESIÓN EN CELULAS SIN NEOPLASIA Y CELULAS INVASORAS DEL CERVIX UTERINO HUMANO" de la alumna, GUELAGUETZA VAZQUEZ ORTIZ con número de cuenta 50006716-7, inscrita en el Doctorado en Ciencias Biomédicas y aspirante al grado de **Doctora en Ciencias**, considero que la Tesis reúne los requisitos suficientes y la acepto para ser presentada en el examen de grado. Por tanto emito mi VOTO APROBATORIO al respecto.

Agradezco de antemano la atención que se sirva prestar a la presente.

Atentamente,

DR. SERGIO MANUEL ENCARNACION GUEVARA.

CENTRO DE CIENCIAS GENOMICAS



Instituto de Fisiología Celular
Universidad Nacional Autónoma de México
Apartado postal
04510 México, D.F.

Tel.

Voto Aprobatorio

27, Enero, 2005

Ing. Leopoldo Silva Gutiérrez
Director General de Administración Escolar,
UNAM
PRESENTE

Estimado Ing. Silva,

Después de haber evaluado la Tesis titulada "**PATRON DE EXPRESIÓN EN CELULAS SIN NEOPLASIA Y CELULAS INVASORAS DEL CERVIX UTERINO HUMANO**" de la alumna, GUELAGUETZA VAZQUEZ ORTIZ con número de cuenta **50006716-7**, inscrita en el Doctorado en Ciencias Biomédicas y aspirante al grado de **Doctora en Ciencias**, considero que la Tesis reúne los requisitos suficientes y la acepto para ser presentada en el examen de grado. Por tanto emito mi VOTO APROBATORIO al respecto.

Agradezco de antemano la atención que se sirva prestar a la presente.

Atentamente,

DR. JORGE RAMÍREZ SALCEDO.

INSTITUTO DE FISIOLÓGÍA CELULAR



Doctorado en Ciencias Biomédicas

Voto Aprobatorio

8 Febrero, 2005

Ing. Leopoldo Silva Gutiérrez
Director General de Administración Escolar,
UNAM
P R E S E N T E

Estimado Ing. Silva,

Después de haber evaluado la Tesis titulada "PATRON DE EXPRESIÓN EN CELULAS SIN NEOPLASIA Y CELULAS INVASORAS DEL CERVIX UTERINO HUMANO" de la alumna, GUELAGUETZA VAZQUEZ ORTIZ con número de cuenta 50006716-7, inscrita en el Doctorado en Ciencias Biomédicas y aspirante al grado de Doctora en Ciencias, considero que la Tesis reúne los requisitos suficientes y la acepto para ser presentada en el examen de grado. Por tanto emito mi VOTO APROBATORIO al respecto.

Agradezco de antemano la atención que se sirva prestar a la presente.

Atentamente,

DRA. MARCELA LIZANO SOBERON.

FACULTAD DE MEDICINA



UNIVERSIDAD NACIONAL
AUTÓNOMA DE
MÉXICO

DR. ALFREDO VARELA ECHAVARRÍA
INSTITUTO DE NEUROBIOLOGIA

CAMPUS UNAM-UAQ JURIQUILLA
APDO. POSTAL 1-1141
QUERÉTARO, QRO. 76001
TEL. 52-442-238-1032 FAX. 52-442-238-1038
varela@inb.unam.mx

Voto Aprobatorio

9 de Marzo, 2005

Ing. Leopoldo Silva Gutiérrez
Director General de Administración Escolar,
UNAM
PRESENTE

Estimado Ing. Silva,

Después de haber evaluado la Tesis titulada "**PATRON DE EXPRESIÓN EN CELULAS SIN NEOPLASIA Y CELULAS INVASORAS DEL CERVIX UTERINO HUMANO**" de la alumna, GUELAGUETZA VAZQUEZ ORTIZ con número de cuenta **50006716-7**, inscrita en el Doctorado en Ciencias Biomédicas y aspirante al grado de **Doctor en Ciencias**, considero que la Tesis reúne los requisitos suficientes y la acepto para ser presentada en el examen de grado. Por tanto emito mi VOTO APROBATORIO al respecto.

Agradezco de antemano la atención que se sirva prestar a la presente.

Atentamente,

DR. ALFREDO VARELA ECHAVARRIA



Instituto de Fisiología Celular
Universidad Nacional Autónoma de México
Apartado postal
04510 México, D.F.

Tel.

Voto Aprobatorio

27, Enero, 2005

Ing. Leopoldo Silva Gutiérrez
Director General de Administración Escolar,
UNAM
PRESENTE

Estimado Ing. Silva,

Después de haber evaluado la Tesis titulada "**PATRON DE EXPRESIÓN EN CELULAS SIN NEOPLASIA Y CELULAS INVASORAS DEL CERVIX UTERINO HUMANO**" de la alumna, GUELAGUETZA VAZQUEZ ORTIZ con número de cuenta **50006716-7**, inscrita en el Doctorado en Ciencias Biomédicas y aspirante al grado de **Doctora en Ciencias**, considero que la Tesis reúne los requisitos suficientes y la acepto para ser presentada en el examen de grado. Por tanto emito mi VOTO APROBATORIO al respecto.

Agradezco de antemano la atención que se sirva prestar a la presente.

Atentamente,

DR. JORGE RAMÍREZ SALCEDO.

INSTITUTO DE FISIOLÓGÍA CELULAR



Doctorado en Ciencias Biomédicas

Voto Aprobatorio

17 del 02, 2005

Ing. Leopoldo Silva Gutiérrez
Director General de Administración Escolar,
UNAM
PRESENTE

Estimado Ing. Silva,

Después de haber evaluado la Tesis titulada "PATRON DE EXPRESIÓN EN CELULAS SIN NEOPLASIA Y CELULAS INVASORAS DEL CERVIX UTERINO HUMANO" de la alumna, GUELAGUETZA VAZQUEZ ORTIZ con número de cuenta 50006716-7, inscrita en el Doctorado en Ciencias Biomédicas y aspirante al grado de **Doctora en Ciencias**, considero que la Tesis reúne los requisitos suficientes y la acepto para ser presentada en el examen de grado. Por tanto emito mi VOTO APROBATORIO al respecto.

Agradezco de antemano la atención que se sirva prestar a la presente.

Atentamente,

DR. MARCO ANTONIO CERBON CERVANTES.

FACULTAD DE MEDICINA



Doctorado en Ciencias Biomédicas

Voto Aprobatorio

10 Febrero, 2005

Ing. Leopoldo Silva Gutiérrez
Director General de Administración Escolar,
UNAM
PRESENTE

Estimado Ing. Silva,

Después de haber evaluado la Tesis titulada "PATRON DE EXPRESIÓN EN CELULAS SIN NEOPLASIA Y CELULAS INVASORAS DEL CERVIX UTERINO HUMANO" de la alumna, GUELAGUETZA VAZQUEZ ORTIZ con número de cuenta 50006716-7, inscrita en el Doctorado en Ciencias Biomédicas y aspirante al grado de Doctora en Ciencias, considero que la Tesis reúne los requisitos suficientes y la acepto para ser presentada en el examen de grado. Por tanto emito mi VOTO APROBATORIO al respecto.

Agradezco de antemano la atención que se sirva prestar a la presente.

Atentamente,

DR. ALFONSO DUEÑAS GONZALEZ.

FACULTAD DE MEDICINA

PATRÓN DE EXPRESIÓN DE CÉLULAS SIN NEOPLASIA Y CÉLULAS INVASORAS DEL CÉRVIX UTERINO HUMANO

INDICE

Resumen	1
Abstract	2
INTRODUCCIÓN	
I. Cáncer Cérvico Uterino.....	3
a) Epidemiología.....	3
b) Epidemiología del CaCU en México.....	3
II) CaCU y Virus de Papiloma Humano.....	4
a) Prevalencia e Incidencia.....	4
b) Historia natural de la infección por VPH.....	5
III) Genómica.....	8
a) Genómica Estructural y Funcional.....	8
b) Micro y macroarreglos.....	10
c) Captura de la imagen y generación de bases de datos.....	14
d) Bioinformática.....	15
e) Normalización y “ruido” de los datos crudos.....	17
f) Métodos de análisis de datos.....	18
IV. Arreglos de Tejido.....	19
V. Microarrays en el estudio de Cáncer Cérvico Uterino.....	19
JUSTIFICACIÓN	21
HIPÓTESIS	22
OBJETIVO GENERAL Y OBJETIVOS PARTICULARES	23
MATERIALES Y MÉTODOS	24
I. Tejidos frescos y líneas celulares.....	24
II. Tejidos embebidos en parafina para construir arreglos de tejido.....	25
III. Obtención de Acidos nucleicos.....	26
IV. Cuantificación y verificación de la calidad del ADN.....	27
V. Detección y Tipificación de VPH.....	27
VI. Extracción de ARN mensajero.....	29
VII. Síntesis y marcaje de las sondas.....	30

VIII. Comprobación de la eficiencia de marcado.....	31
IX. Características de los arreglos.....	31
X. Hibridación.....	32
XI. Lavados Post-hibridación.....	32
XII. Captura de la Imagen.....	32
a) Plantilla.....	33
b) Eliminación de "Detritos".....	33
c) Fondo.....	33
d)) Normaliza con diferentes opciones.....	33
e) Base de datos.....	33
XIII. Análisis de los datos.....	34
XIV. RT-PCR semicuantitativa.....	35
XV. Hibridación In Situ (HIS).....	36
XVI. Inmunohistoquímica (IHC).....	38
RESULTADOS.....	41
I. Extracción de ADN y tipificación.....	41
II. Extracción de ARN y obtención de ARN mensajero.....	42
III. Resultados de la hibridación.....	46
IV. Normalización.....	46
V. Expresión global del Cáncer Cérvico Uterino.....	51
VI. Procesos biológicos alterados.....	54
VII. Búsqueda de posibles marcadores moleculares de la enfermedad.....	63
VIII. Búsqueda de posibles marcadores de invasión.....	63
DISCUSIÓN.....	67
CONCLUSIONES.....	70
REFERENCIAS.....	72
APÉNDICE 1	
APÉNDICE 2	
APÉNDICE 3	
APÉNDICE 4	
APÉNDICE 5	
APÉNDICE 6	
APÉNDICE 7	

RESUMEN

El carcinoma cérvico uterino (CaCU), representa un importante problema de salud mundial, constituyendo la segunda causa de mortalidad femenina en el mundo. En nuestro país, este tumor ocupa el segundo lugar de incidencia en la población femenina. Se han llevado a cabo diversos experimentos de genómica funcional principalmente de microarreglos, con el fin de definir el transcriptoma parcial de esta neoplasia. Estos trabajos han llevado a determinar algunos patrones de expresión como respuesta a radioterapia. Sin embargo, existen muy pocos reportes que identifiquen a posibles marcadores moleculares de la enfermedad derivados de un patrón de expresión diferencial entre una célula cervical normal y una célula tumoral infectada con virus de papiloma humano (VPH). En este trabajo utilizamos metodologías de análisis genómico, tales como microarreglos de cADN y microarreglos de tejido, así como diversas herramientas de análisis bioinformático y de biología molecular para determinar genes específicos que son blanco de alteraciones en su nivel de expresión en esta neoplasia. Además se llevaron a cabo análisis *in silico* de las diferentes rutas metabólicas y procesos celulares afectados por estas alteraciones en la expresión génica. A través de esta aproximación fue posible determinar la expresión diferencial de la carcinogénesis cervical, caracterizada por la sobreexpresión en tejidos invasores de algunos genes tales como IL6, VEGF, Claudina 1, NDUFB6, Lamc2, Receptor de glucocorticoides alfa entre otros; así como la supresión de los genes Grancalcina, Receptor de interleucina 18, Inhibidor de la trombina placentaria entre otros. Por otro lado se identifico a los genes MMP11, MMP12 y Catepsina F como posibles marcadores potenciales de la progresión del cáncer, ya que en las diferentes lesiones premalignas estudiadas su expresión fue restringida a las células displásicas. Por otro lado, el análisis bioinformático nos permitió identificar diferentes rutas metabólicas o procesos celulares que presentan alteraciones en la expresión génica en presencia del cáncer. Todas los procesos celulares estudiados presentan tanto sobreexpresión o supresión en los genes que los componen como en el caso de los genes homeobox, interleucinas, ciclo celular etc. Sin embargo es importante recalcar que la única ruta metabólica que se encontró totalmente sobreexpresada fue la ruta de señalización de Wnt. En conclusión, este trabajo permitió identificar el transcriptoma parcial y el patrón de expresión diferencial entre la célula cervical normal y la célula tumoral en presencia de VPH; así como la detección de genes que presentan diferencias en su expresión génica que potencialmente podrían ser utilizados como marcadores biológicos útiles para el diagnóstico o pronóstico de esta enfermedad.

ABSTRACT

Uterine cervix carcinoma (UCC) represents an important health problem, representing the second cause of death among the female population worldwide. In our country, this disease occupies the second place in incidence among the female population. Several functional genomic analysis methods, mainly microarrays, have been used to define the partial transcriptome of this neoplasm. These works have led to determine some expression patterns as an answer for radiotherapy. Nevertheless, there are only a few reports which identify possible molecular markers for the disease, based on the differential expression pattern of a normal cervical cell and a HPV infected tumor cell. In the present work we used genomic analysis methods, such as cDNA microarrays, tissue microarrays, as well as several bioinformatic analysis and molecular biology tools in order to determine specific expression patterns in this neoplasm. Furthermore, we performed an *in silico* analysis of the metabolic pathways and cellular processes that might be affected by these alterations. Thanks to this approach, it was possible to determine the differential expression of the cervical carcinogenesis, characterized by the overexpression of IL6, VEGF, Claudin 1, NDUFB6, Lamc2, Glucocorticoid alpha receptor among others and the underregulation of Grancalcin, IL18R, Thrombin Inhibitor placental among others in tumor samples. On the other hand, MMP11, MMP12 and Cathepsin F have been identified as potential molecular markers for the cancer progression, based on its restricted expression to dysplastic cells in all low grade lesions studied. The bioinformatic analysis allowed us to identify several metabolic pathways and cellular processes which present expression alterations in cancer presence. All cellular processes showed over and under expression in all genes just like in the homeobox, interleukins, cell cycle... etc. Nevertheless, it is worth important to note that Wnt signalling pathway presented only overexpression in all genes. In conclusion, this work allowed us to identify the partial transcriptome and the differential expression pattern between normal and HPV tumor cervical cells, as well as the detection of genes which present differences in their expression that might be used as potential biological markers for the diagnosis or prognosis of this disease.

INTRODUCCIÓN

I. Cáncer Cérvico Uterino

a) Epidemiología.

El carcinoma cérvico uterino (CaCU) es la segunda causa de mortalidad femenina, con 288,000 muertes por año a nivel mundial. Su frecuencia, estimada por la Organización Mundial de la Salud es de aproximadamente 510,000 casos por año; de los cuales el 80% ocurre en países en vías de desarrollo. En ausencia de programas de detección temprana, como el uso rutinario de la prueba de Papanicolau, el CaCU se detecta de manera tardía y provoca la muerte en la mayoría de los pacientes. Aún cuando las pacientes con tumores invasores reciban tratamiento adecuado, el 50-55% de ellas morirán debido a la enfermedad (1).

b) Epidemiología del Cáncer Cérvico Uterino en México

Hasta el año 2000, el CaCU representaba la primera causa de muerte entre la población femenina de México y actualmente nuestro país ocupa uno de los primeros lugares en morbilidad de esta neoplasia a nivel mundial (Carcinoma invasor: 41.1, Ca. *In situ*: 49.3 por cada 100,000 mujeres de 25 años o más). De acuerdo a datos provenientes de la Secretaría de Salud, este tumor representa más del 24% de todas las neoplasias malignas reportadas en México; y cada día 12 mujeres mexicanas mueren de esta enfermedad. Desde 1996 ocurrió un incremento en la incidencia de esta patología, haciendo de ella uno de los problemas de salud más graves en nuestro país (2).

II) CaCU y Virus de Papiloma Humano

a) Prevalencia e Incidencia.

Existen diversos factores de riesgo asociados al desarrollo del CaCU, incluyendo el número de parejas sexuales, enfermedades sexualmente transmisibles, inicio de vida sexual activa a edad temprana, número de paras, el uso de anticonceptivos orales y el tabaquismo. Sin embargo, el factor etiológico más importante para el desarrollo del CaCU es la infección con virus de papiloma humano de alto riesgo (VPHAr) (3-6). Un tipo viral se define en base a la diferencia de al menos del 10% en la secuencia nucleotídica. Actualmente, han sido identificados más de 85 VPHs, y aproximadamente 120 posibles nuevos tipos han sido parcialmente caracterizados (7). De todos los tipos virales existentes, aproximadamente 40 pueden infectar el tracto anogenital humano y han sido encontrados en tumores anogenitales. Por otro lado, es importante mencionar que el 20% de los tumores de la orofaringe se han encontrado secuencias que corresponden a estos mismos tipos virales (8).

La prevalencia de infección por VPH en el mundo es de alrededor de 630 millones de personas infectadas. Sin embargo, solamente cerca de 190 millones presentan manifestaciones clínicas. La prevalencia de infección por VPH entre mujeres sexualmente activas va del 18 al 25%, especialmente en ciertos grupos de adolescentes sexualmente activas. La única patología clínica dependiente de la infección por VPH que se sabe que se transmite de madre a hijo en el momento del nacimiento por vía vaginal es la papilomatosis respiratoria (9).

Debido a su incidencia en lesiones con diferentes grados de malignidad, los VPHs anogenitales han sido divididos en: alto riesgo, que son aquellos que predominan en tumores invasores; los de mediano riesgo y de bajo riesgo se encuentran presentes en lesiones benignas, como los condilomatosis (10). Estudios recientes a nivel mundial han determinado que los tipos de VPH que deben ser considerados de alto riesgo son los

tipos: 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68, 73 y 82, mientras que los tipos 23, 53 y 66 sólo deben ser considerados como potencialmente carcinogénicos (11).

Diferentes reportes indican que la prevalencia de infección con VPHAr en CaCU es superior al 99.7% de los casos y que la probabilidad de desarrollar una lesión intraepitelial cervical de alto grado, se incrementa al menos 100 veces en mujeres infectadas con VPHAr, comparadas con mujeres no infectadas (12).

En 1995 la Agencia Internacional de Investigación del Cáncer (IARC), clasificó los tipos virales 16 y 18 como agentes carcinogénicos para el ser humano y los tipos 31 y 33 como probables carcinógenos (13).

Un análisis reciente, en más de 10,000 casos, con el fin de determinar la prevalencia de los diferentes tipos de VPH a nivel mundial, indicó que los tipos de VPHAr más comunes en los tumores invasores son, en orden decreciente: VPH16, 18, 45, 31, 33, 58, 52, 35, 59, 56, 6, 51, 68, 39, 82, 73, 66 y 70. En los carcinomas epidermoides, el VPH 16 fue el más prevalente (46-63%), seguido de los tipos 18 (10-14%), 45 (2-8%), 31 (2-7%) y 33 (3-5%) en todas las regiones geográficas, excepto Asia, donde los tipos 58 (6%) y 52 (4%), fueron identificados con mayor frecuencia. En los adenocarcinomas, la prevalencia de VPH fue significativamente menor (76.4%) que en los tumores escamosos (87.3%), siendo el VPH18 el tipo predominante en todas las regiones geográficas (37.4-41%), seguido de los tipos 16 (26-36%) y 45 (5-7%) (14).

b) Historia natural de la infección por VPH

Las lesiones relacionadas con VPH se desarrollan más comúnmente en el tracto anogenital de pacientes del sexo femenino. Sin embargo, el virus también juega un papel en el desarrollo de carcinomas de ano y de pene. El varón es considerado como uno de los mayores reservorios para la transmisión de VPH (15-16). No obstante, las diferencias entre la incidencia de cáncer de pene contra la de CaCU, sugiere que el epitelio cervical

debe ser mas susceptible que el epitelio del pene a la transformación por VPH. Para establecer una infección productiva, el VPH debe tener acceso a la célula madre epitelial, localizada en el estrato basal del epitelio cervical. La descamación natural de las células de la superficie del epitelio infectado contiene VPH y sirve como vector de transmisión para el epitelio normal. El alojamiento de estas células infectadas en sitios con microlesiones, como las que se producen después del acto sexual, permiten que el virus "cruce" el epitelio y llegue a las células basales. Una vez que el VPH infecta a la célula madre epitelial, el virus toma el control sobre su maquinaria de replicación de DNA, con la finalidad de replicar su propio material genético. En esta etapa, el genoma viral se encuentra en forma circular libre o episoma. Estos episomas se mantienen dentro del núcleo de la célula madre infectada y se distribuyen a cada célula hija durante la división celular (17). Es importante mencionar que en las células basales el DNA del VPH esta presente en bajo número de copias.

Por otro lado, la reparación de las microlesiones puede estimular la división de las células basales permitiendo también la replicación del VPH (18-19). Algunas de estas células hijas, localizadas en el estrato suprabasal, empiezan a expresar las proteínas virales E1 y E2, así como las oncoproteínas virales E6 y E7, y son capaces de replicar y amplificar el genoma viral hasta cientos de copias por célula. Finalmente, las proteínas codificadas por los genes virales E4, E5 y las proteínas tardías L1 y L2 se expresan en el estrato espinoso, donde los nuevos viriones son ensamblados y liberados mediante la exfoliación de las células escamosas en la superficie del epitelio que contienen las cargas virales más altas de todo el epitelio (20-22) (ver Figura 1). El cáncer es el resultado de la pérdida de control sobre el crecimiento celular normal. Normalmente los genes supresores de tumor p53 y pRB en la célula hospedera inspeccionan el proceso de división celular buscando errores o mutaciones en la replicación, de manera que cualquier desviación de la estructura genética es corregida o bien la célula es destruida. De forma importante, la

función normal de estos genes es interrumpida por la producción de las proteínas virales E6 y E7 que no permiten el arresto ni la muerte celular permitiendo la acumulación de lesiones genéticas (23-25). Los VPH de bajo riesgo tienen una capacidad mucho menor de inducir la transformación maligna con respecto a los de Alto riesgo debido principalmente a que las proteínas virales E6 y E7 de tipos virales de bajo riesgo tienen menor afinidad por las proteínas celulares p53 y pRB (26-27).

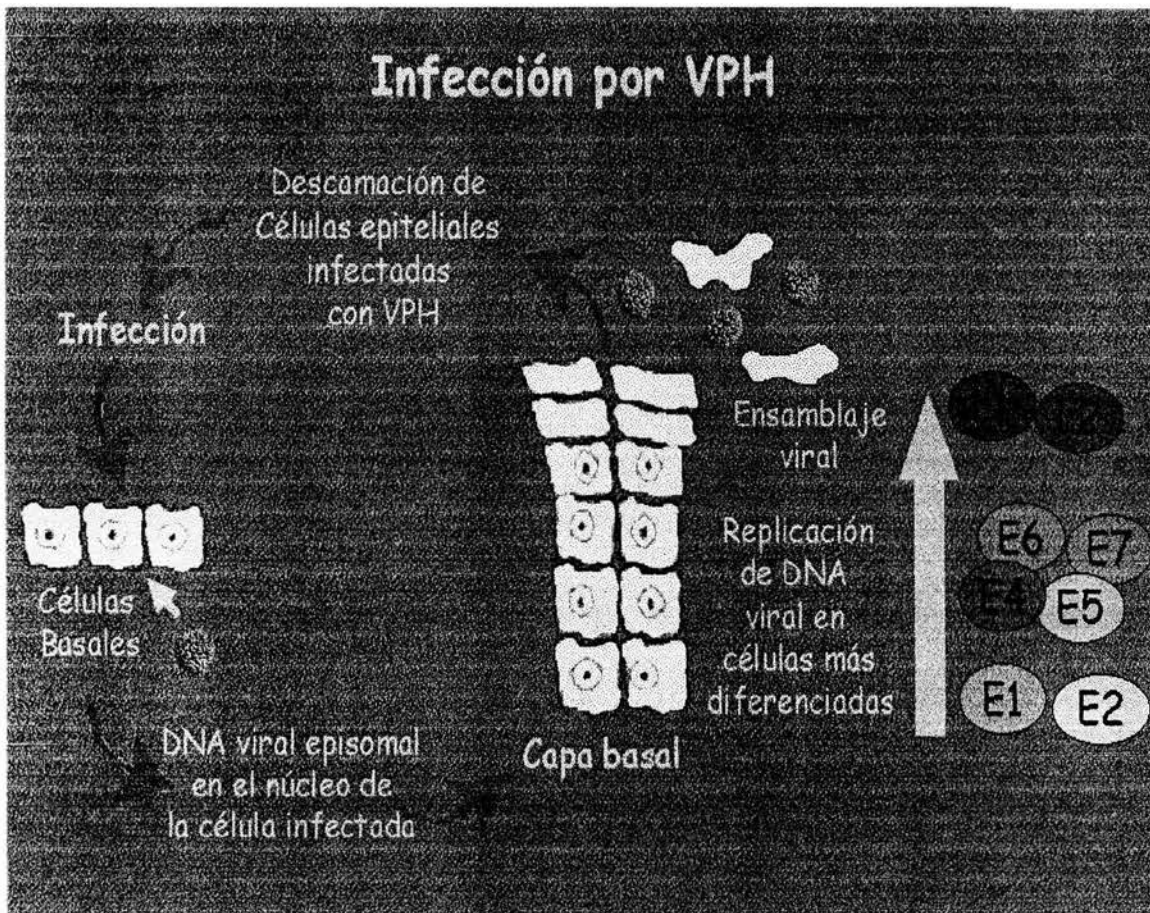


Figura 1.

Historial Natural de la Infección por Virus de Papiloma Humano

En la mayoría de las lesiones pre-malignas el DNA viral permanece en forma episomal, mientras que en lesiones de alto grado y en tumores invasores se encuentra un alto

porcentaje de genomas de VPH integrados al genoma celular. De ahí que la integración del VPH en la célula hospedera parece ser un paso final en la transformación de ésta hacia una célula inmortalizada con capacidad de invadir. La integración viral ocurre cuando se presenta la pérdida del control de la célula hospedera con persistencia del VPH de alto riesgo (28-29).

El hecho de que no todas las pacientes infectadas con VPHAr desarrollen cáncer indica que la presencia de VPHAr es un factor de riesgo necesario pero no suficiente para el desarrollo de una neoplasia cervical, y que es probable que existan otros mecanismos moleculares y celulares que, en conjunto con el virus, provoquen el desarrollo de las neoplasias cervicales. La infección por VPH altera una gran cantidad de mecanismos moleculares en la célula infectada, cada uno de los cuales conferirá a la célula nuevas capacidades y ventajas sobre las células normales. De ahí que, una de las metas a seguir en el estudio del cáncer cervical sea el poder identificar y caracterizar productos de genes celulares que participen directamente en el desarrollo de esta patología. Una herramienta muy poderosa para identificar estos genes es la implementación de la genómica funcional mediante el uso de microarreglos (también conocidos como biochips, microconjuntos, micromatrices, etc.).

III) Genómica

a) Genómica Estructural y Funcional

La obtención del borrador del Proyecto del Genoma Humano (PGH), en donde se ha secuenciado la mayoría de los genes humanos; nos sitúa en lo que se conoce como era postgenómica. Para obtener algo en claro de la inmensa cantidad de secuencias obtenidas hay que analizarlas con el fin de extraer información. La ciencia genómica en conjunto con algunas otras como la bioinformática, representan herramientas para entender las preguntas básicas de la genética (la relación genotipo - fenotipo) pero, esta

vez, a un nivel integrativo de todos los componentes macromoleculares de una célula. De ahí que actualmente muchas ciencias estén empleando las secuencias y los datos que se están generando por el PGH, para dar explicación a dos cuestiones de vital importancia: la causa de una enfermedad y su posible tratamiento.

La genómica es la ciencia que estudia el genoma de un organismo y el papel que juegan en él los genes que lo componen, desde un punto de vista estructural y funcional. En síntesis, la Genómica Estructural identifica, aísla, sitúa y caracteriza el conjunto de genes al que llamamos Genoma; mientras que la Genómica Funcional se ocupa de describir la función biológica de los genes mediante el conocimiento de su actividad en:

- Los rasgos que determinan
- La regulación a la que se ven sometidos
- La interacción con otros genes
- La identificación de un patrón de comportamiento de un gen, dependiendo de las condiciones que lo rodean
- La actividad que desarrolla un gen cuando está alterado en relación a su actividad normal

La Genómica Funcional asocia rasgos o patologías con el gen o los genes que los producen. Los avances tecnológicos están permitiendo determinar qué genes están implicados y la naturaleza de la interacción en la asociación gen-rasgo o bien gen-patología. Mientras no ocurran mutaciones, cada célula del organismo tiene el mismo material genético durante toda su vida. Sin embargo, la expresión del gen (su actividad) varía de unas células a otras, de unos estadios de desarrollo a otros, en procesos normales o patológicos y en función de las condiciones ambientales. El tratar de entender la expresión génica nos permitirá entender los procesos biológicos moleculares en los que intervienen los genes. Es importante recalcar que un rasgo o una patología puede venir determinado por un gen; o bien que varios genes determinen a una sola patología; como

es el caso del cáncer cérvico uterino. De ahí que una buena aproximación para el estudio de la expresión de este cáncer sea por medio de análisis global es decir; que se estudien los niveles de expresión de varios genes al mismo tiempo (30).

b) Micro y Macroarreglos

Los microarreglos, genechips o biochips ofrecen una plataforma sólida para el estudio del análisis global de la expresión. Los arreglos utilizan una de las propiedades del DNA duplex; la complementariedad en la secuencia de sus dos cadenas. Es notable que una molécula de tanta complejidad estructural pueda rehibridarse con una fidelidad tan perfecta después de haber estado separada. El inicio de los arreglos se remonta al descubrimiento de que una sola cadena de ADN se une fuertemente a membranas de nitrocelulosa de manera tal, que previene a las cadenas reasociarse unas con otras, pero permite la hibridación con ARN complementario (31-33). A esto siguió la aplicación directa de esto, y se denominaron "Dot-blot" (Revisado en Anexo 2). Sin embargo, los arreglos modifican ligeramente esta estrategia. Todo el proceso se basa en una técnica que utiliza moléculas de ácidos nucleicos marcados con fluorescencia.

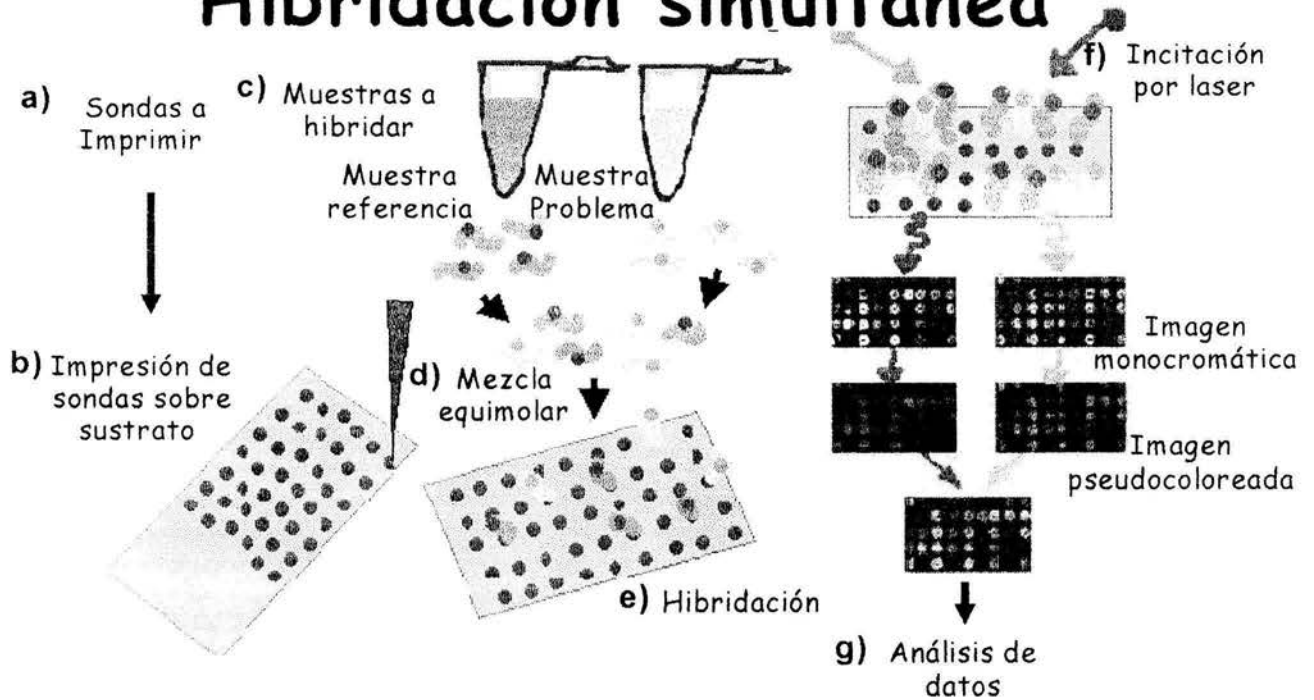
Un arreglo consiste de una superficie sólida que puede ser de nylon, vidrio, plástico o sílice sobre la cual son depositados o sintetizados transcritos de genes conocidos. Estos transcritos se colocan con un orden espacial predeterminado tipo matricial, permitiendo a estos polinucleótidos fungir como sonda blanco en un análisis de hibridación. Por otro lado, se utilizan moléculas de ácidos nucleicos marcadas con fluorescencia que se utilizan como "sondas móviles" para identificar a las moléculas complementarias o sondas blanco inmovilizadas. Tomando como ejemplo el análisis de dos tipos de células: células tipo 1 o células "sanas", y células tipo 2 o células que presentan cierta patología; y partiendo de que ambos tipos de células contienen un conjunto de cuatro genes, a, b, c y d de los cuales se necesita determinar su patrón de expresión.

Se procede de la siguiente manera: inicialmente se aísla el ARN mensajero de cada una de las células y se utiliza como templado para generar ADN complementario marcado con fluorescencia. Los transcritos de cada una de las muestras se marcan con fluorescencia ya sea roja o con verde, de tal manera que las muestras puedan ser diferenciadas en los pasos subsecuentes. Las dos muestras se mezclan en cantidades equimolares y se incuban sobre el microarreglo que contiene a las sondas blanco inmovilizadas a, b, c y d. Las moléculas marcadas se hibridan con su contraparte unida al arreglo, el cual representa los transcritos expresados por la célula. Posteriormente se determina la abundancia relativa de los transcritos presentes en este conjunto de muestras mediante una comparación directa entre las dos sondas (Figura 2). A este tipo de hibridación se le denomina hibridación simultánea. El esquema de hibridación es muy similar cuando se utilizan sondas marcadas con radiactividad, sólo que no se lleva a cabo una hibridación simultánea, sino una hibridación en paralelo entre las dos muestras (Ver Figura 2) (34).

Existen dos tipos de arreglos: los macro y los microarreglos. Los macroarreglos son generalmente membranas de nylon con productos de PCR impresos, cada uno de los cuales se denomina "*spot*", que tienen un diámetro de 0.5 a 1 mm y el espacio entre los puntos (*pitch*) es de aproximadamente 2 mm. Las densidades máximas que pueden llegar a tener los macroarreglos son del orden de 25 impresiones en punto de oligonucleótidos o spots por cm². Las membranas pueden ser de diferentes tamaños, llegando a ser incluso del tamaño de media hoja carta. Normalmente en los macroarreglos se utilizan sondas marcadas con radiactividad en volúmenes muy grandes de hibridación (varios ml), utilizando así una hibridación en paralelo y la adquisición de los resultados se hace por medio de un sistema de pantallas fósfoluminiscentes, cuyas imágenes posteriormente son capturadas con programas computacionales específicos que digitalizan las imágenes.

Los microarreglos tienen dimensiones mucho más pequeñas, su diámetro de spot es de 200 a 500 nm, sus densidades máximas pueden alcanzar hasta varios miles de spots por

Hibridación simultánea



Hibridación en paralelo

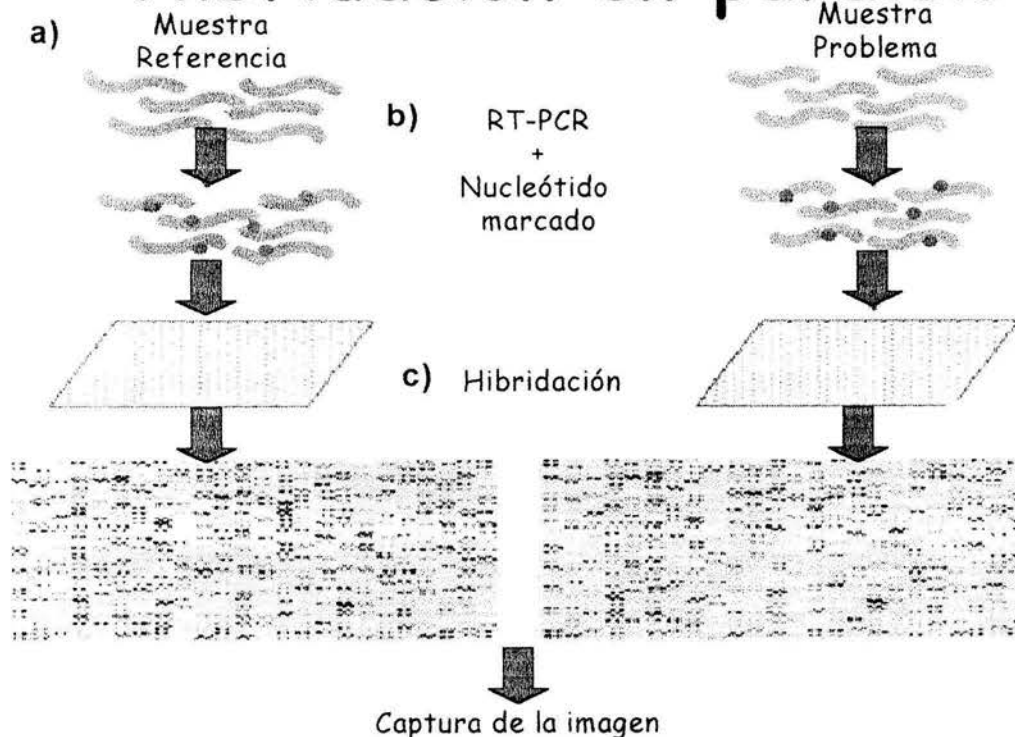


Figura 2.

Esquema de microarrays. Se obtienen los templados para los genes de interés y se amplifican por PCR. Posterior a la purificación y al proceso de control de calidad las alícuotas son impresas sobre la superficie sólida utilizando robots dirigidos por computadoras. El RNA total tanto de la muestra referencia como de la problema se marcan ya sea con fluorescencia (Cy3 y Cy5 respectivamente) o bien con radiactividad mediante RT-PCR. En el caso de la hibridación simultanea, las sondas fluorescentes se mezclan en cantidades equimolares y esta mezcla se hibrida sobre las clonas inmovilizadas en la superficie sólida en condiciones de astringencia. La excitación del láser sobre los marcajes hibridados genera una emisión con características de espectro, que se mide utilizando un escáner que utiliza los principios de la microscopia confocal. En el caso de la hibridación en paralelo cada una de las muestras marcadas se hibrida sobre diferentes arrays. La imagen se captura mediante un Phosphorimager y estas imágenes igualmente son importadas a programas computacionales mediante los cuales se obtendrá el resultado final. Estos programas computacionales agregan la información sobre las clonas, el nombre identificador de la clona, los valores de la intensidad, las razones de la intensidad, la constante de normalización. etc.

cm². Habitualmente, los microarreglos se elaboran en vidrio y pueden llegar a contener hasta 10,000 spots en una superficie de 2 a 8 cm². La hibridación se realiza con una sonda marcada con fluorescencia (hibridación en paralelo) y la imagen se adquiere con un lector que utiliza los principios de la microscopia confocal (35-37).

c) Captura de la imagen y generación de bases de datos.

Después de la hibridación, los microarreglos se colocan en un lector que consiste de varias fuentes de laser, un microscopio y una cámara. El laser excita los transcritos fluorescentes, y el microscopio y la cámara, en conjunto, generan una imagen digitalizada del arreglo. Los datos son almacenados en una computadora, la cual mediante un programa especializado analiza la imagen digital del arreglo calculando la razón de las fluorescencias (rojo/verde) o, bien, sustrayendo el fondo de cada uno de los *spots* que forman el arreglo. Al calcular la razón, el programa crea una base de datos que contiene la razón de la intensidad de la fluorescencia (rojo/verde) para cada uno de los *spots*. De ahí que de esta comparación se obtengan valores de intensidad relativa para cada uno de los *spots*, o bien para cada uno de los genes contenidos en el arreglo. Estos programas computacionales agregan además de la asignación de valores para cada gen, información sobre cada uno de los genes, como el nombre, Unigene ID, GeneBank accesión ID, etc..., generándose así para cada una de las muestras analizadas una base de datos de mediciones de intensidad como la que se muestra a continuación en la Tabla 1. Siguiendo con el ejemplo que se ha mencionado anteriormente, el análisis computacional podría arrojar que ambos tipos celulares presentan el mismo nivel de expresión en el gen "a", que la célula tipo 1 expresa mas el gen "b", que la célula tipo 2 expresa mas el gen "c" y que ninguna de las dos células expresa el gen "d". Sin embargo, cabe mencionar que esto sólo es un ejemplo ya que los arreglos pueden constar hasta de 10,000 o más genes.

d) **Bioinformática**

Es importante enfatizar que, como se vio previamente, por cada experimento de microarreglos se genera una base de datos tan grande como la cantidad de genes que en un determinado arreglo se están analizando. Por lo tanto, es imprescindible encontrar la forma de almacenarlos y analizarlos cuando se requiera. La informática actualmente juega un papel decisivo para poder transformar los datos en información y ésta en conocimiento. Una definición aceptada de la bioinformática es "disciplina científica que se interesa por todos los aspectos relacionados con la adquisición, almacenamiento, procesamiento, distribución, análisis e interpretación de información biológica, mediante la aplicación de técnicas y herramientas de las matemáticas, de la biología y de la informática, con el propósito de comprender el significado biológico de una gran variedad de datos" (Comentario Personal Dr. Lauro Sumoy Van Dyck, Laboratorio de Microarreglos, Instituto de Regulación Genómica, Barcelona, España). Como se mencionaba, el objetivo final de esta disciplina es obtener mediante esta combinación tecnológica nuevas perspectivas biológicas, así como poder crear una perspectiva global a partir de la cual puedan inferirse diferentes principios biológicos básicos. En un principio, la bioinformática se dedicó a coleccionar y organizar en bases de datos resultados tales como secuencias de nucleótidos o proteínas. Además de esta labor de colección y almacenaje, también se diseñaron nuevas "bibliotecas virtuales" o acervos de información disponible donde se reportaban datos obtenidos experimentalmente por diferentes grupos de investigación. Estos acervos deberán ahora permitir acceder a cualquier investigador, tanto para analizar los datos almacenados, como para contribuir con datos nuevos al acervo. Esta ciencia es, además, una herramienta integradora de los diversos datos contenidos en las diferentes bases de datos biológicas existentes.

A_Intensity	B_Intensity	Ratio	Int_Diff	Clone_ID	GenBank_ID	Unigene_ID	Gene_Name
250232.5	100937.43	2.4791	149295.07	2223118	AI026721	Hs.46798	Homo sapiens mRNA, cDNA DKFZp434K152 (from clone DKFZp434K15
171681.5	56412.44	3.0433	115269.06	2515988	AW966798	Hs.19555	prostate tumor over expressed gene 1
278855.5	212716.49	1.3109	66139.01	2319421	AI360052	Hs.173051	ESTs
282087	216147.09	1.3051	65939.91	630051	AA040945	Hs.20107	ESTs
246235	181973.99	1.3531	64261.01	2887170	AA706272	Hs.163666	ESTs
232892	176803.64	1.3172	56088.36	99727	AI620057	Hs.190356	ESTs, Moderately similar to ALU2_HUMAN ALU SUBFAMILY SB SEQUE
152373	104922.02	1.4522	47450.98	2922560	W51999	Hs.24359	Homo sapiens cDNA FLJ11174 fis, clone PLACE1007367
116107	72949.55	1.5916	43157.45	2252822	AI080050	Hs.4932	hypothetical protein FLJ20605
251413.5	220508.55	1.1402	30904.95	2108454	BE007322	Hs.38489	ESTs
53907.3	23590.33	2.2851	30316.97	2512456	AA884787	Hs.173946	hypothetical protein FLJ10486
44135.85	17740.42	2.4879	26395.43	816159	AW157581	Hs.182982	golgin-67
37536.65	11230.63	3.3423	26306.02	2868138	AA554735	Hs.111334	ferritin, light polypeptide
29896.9	4443.1	6.7288	25453.8	641127	XM_005780	Hs.25040	zinc finger protein 239
39274.55	14428.97	2.7219	24845.58	2232065	AA506687	Hs.109929	hypothetical protein MPMGp800B12492Q3
27435.65	3262.63	8.4091	24173.02	604856	NM_006169	Hs.76669	nicotinamide N-methyltransferase
32791.7	9507.86	3.4489	23283.84	882522	AI912413	Hs.27865	Homo sapiens cDNA: FLJ21333 fis, clone COL02535
24764.2	1873.21	13.2202	22890.99	2132815	AW969881	Hs.179779	ribosomal protein L37
31165.75	9124.57	3.4156	22041.18	1943408	AA464096	Hs.193482	Homo sapiens cDNA FLJ11903 fis, clone HEMBB1000030
23220.65	1182.45	19.6378	22038.2	212912	BE881018	Hs.8297	ribonuclease 6 precursor
23515.6	2131.97	11.03	21383.63	2853091	AF029403	Hs.144877	cytochrome P450, subfamily VIIB (oxysterol 7 alpha-hydroxylase), polypept
30429.35	9744.44	3.1227	20684.91	3230705	BG499021	Hs.184297	KIAA0808 gene product

Tabla 1.

Ejemplo de un fragmento de base de datos generada por el programa computacional Atlas Image para una muestra analizada en este trabajo.

Al tiempo que surgen las técnicas de análisis global, la bioinformática coevoluciona con ellas con el fin de combinar la información biológica experimental, la informática y la matemática con el fin de ofrecer una visión más global sobre el funcionamiento celular en condiciones normales, anormales o patológicas. De ahí que esta disciplina sea la herramienta por excelencia para el análisis de datos obtenidos por microarreglos.

e) Normalización y “ruido” de los datos crudos.

Con el fin de analizar los datos obtenidos a partir diferentes arreglos que forman un solo experimento, primero hay que unificarlos en una sola base datos. Para llevar a cabo esto, las diferentes mediciones (niveles de intensidad) de cada uno de los arreglos deben ser normalizadas con el fin de poderlas hacer comparables entre sí. Por ejemplo; en experimentos donde se utilizan poblaciones homogéneas de células similares y una sola tecnología de microarrays y además la medición de la expresión se lleva a cabo en un solo momento; la normalización debería de ser solamente: el ajuste del exceso de brillo de los pixeles de cada microarray digitalmente capturado; siempre y cuando se asuma que la cantidad de RNA hibridada es la misma (38). Por otro lado, otros métodos de normalización que actualmente se emplean incluyen: utilizar los niveles de expresión de los genes constitutivos (39), o bien, dar por hecho que la mayoría de los genes no cambian entre los diferentes experimentos (40), o la utilización de técnicas estadísticas con parámetros no lineales (41-42). Sin embargo, los experimentos que se llevan a cabo en el área de la genómica funcional en la actualidad requieren métodos de normalización compuestos que incluyen dos o más de las metodologías de normalización mencionadas. Por otro lado, se han realizado diferentes estudios con el fin de analizar la reproducibilidad de los datos obtenidos por microarreglos mediante réplicas de cada microarreglo por experimento (43), o bien comparando las mediciones de los datos obtenidos por microarreglos con las obtenidas por northern blot (44-45) y SAGE (46). Además, se han

evaluado algunas estrategias de transcripción *in vitro* (47). Como resultado de dichos análisis, ahora se cuenta con métodos de normalización que permiten hacer comparables los niveles de expresión entre los diferentes arreglos de un solo experimento, con la certeza de que los valores de expresión obtenidos son reflejo de la expresión real del transcrito celular. Por otro lado, se ha comprobado que hay diferentes causas de “ruido” en los datos obtenidos por microarreglos y que las variaciones intra- e inter arreglos pueden sesgar la interpretación de los datos de expresión; de ahí la existencia de programas computacionales específicos para solucionar estos y algunos otros problemas.

f) Métodos de análisis de datos

Los métodos actuales en el área de la bioinformática para el análisis de datos obtenidos por arreglos pueden dividirse en dos grandes categorías: métodos supervisados, que sirven para determinar genes que encajen en un patrón predeterminado; y los métodos no supervisados, que permiten caracterizar los componentes de una base de datos sin tener un conocimiento *a priori* de lo que se espera encontrar. Los métodos supervisados se utilizan, generalmente, para los dos siguientes propósitos: a) encontrar genes con niveles de expresión significativa y diferente entre muestras y b) encontrar genes que pronostiquen una característica de la muestra. Por otro lado; los métodos no supervisados tratan de encontrar una estructura interna o una relación entre los datos obtenidos. Estos métodos se utilizan para: a) determinar genes con características interesantes sin buscar un patrón específico, b) determinar grupos de genes o muestras con patrones similares de expresión y c) elaborar gráficas que representen interacciones gen-gen o gen-fenotipo. (Ver Anexo 2).

Debido a que los arreglos muchas veces cubren gran parte del transcriptoma medible, el usuario pareciera no necesitar tener en mente a un grupo de genes candidatos específicos que quiera estudiar. Sin embargo; esto no significa que los experimentos de

microarreglos se conciban y trabajen sin una hipótesis. Al contrario, el reto actual es que para contestar la hipótesis que se formuló se pueda encontrar la técnica bioinformática más apropiada.

IV. Arreglos de Tejido.

Debido al alto costo de los microarreglos; una pequeña cantidad de muestras son analizadas por esta metodología. Sin embargo, los posibles marcadores moleculares que se obtienen a partir de este experimento inicial deben ser validados en un gran número de muestras clínicas. Actualmente los estudios de genómica funcional de alto rendimiento como los microarreglos aplicados al estudio de las neoplasias, han generado una gran cantidad de biomarcadores que potencialmente podrían ser utilizados en el ámbito clínico como marcadores de diagnóstico o pronóstico. Este punto representa actualmente uno de los principales problemas que se presentan en los estudios de la investigación genómica, deteniendo la aplicación clínica directa de las ventajas obtenidas a través del análisis genómico. Una solución a este problema lo constituye el uso de arreglos de tejidos. Esta metodología permite realizar la validación de marcadores tumorales en forma masiva, ya que se pueden analizar de manera paralela un gran número de muestras clínicas bajo las mismas condiciones experimentales, lo cual representa un ahorro considerable de tiempo y recursos (49-50).

V. Microarreglos en el estudio de Cáncer Cérvico Uterino.

El primer trabajo que fue publicado utilizando la técnica de microarreglos fue un estudio de *Arabidopsis* en el año de 1995 llevado a cabo por Patrick Brown (51). Sin embargo, no es hasta el año 1996 que se realiza el primer trabajo en cáncer utilizando líneas celulares derivadas de melanoma (52), y desde este año; hasta 1998 se genera una gran cantidad de trabajos, principalmente revisiones, en donde básicamente la técnica de microarreglos

es explicada a fondo y donde se mencionan posibles aplicaciones de la misma en diferentes modelos de estudio. Cabe mencionar que en este periodo existen muy pocos experimentos, la mayoría de ellos llevados a cabo en levaduras (53). En 1998 se realizan los primeros estudios experimentales con tejidos humanos (54), y es aquí donde surge el auge de los estudios de expresión por medio de arreglos en distintos modelos de estudio. Es importante mencionar, que el cáncer cérvico uterino fue uno de los primeros tumores sólidos en estudiarse mediante esta tecnología. El primer estudio realizado sobre CaCU mediante microarrays reveló la sobreexpresión de los genes de proteínas de unión a guaninas, Factor nuclear NF45 y beta catenina; siendo la aportación más importante de este trabajo: el haber utilizado muestras procedentes de pacientes y haber determinado a p16 como posible marcador del cáncer (55).

Posteriormente, otro grupo reportó las diferencias en expresión génica en queratinocitos infectados con E6 y E7 de VPH16. Por ejemplo, se describió que la oncoproteína viral E6 altera la expresión de genes que afectan la resistencia de la célula hospedera a la infección y la función inmune de la misma (56). Por otro lado, con la misma tecnología, el se pudo relacionar directamente a los genes de la NADH deshidrogenasa 4 y el de la proteína ribosómica S12 con la progresión y presencia del cáncer cervical (57). En otro trabajo se mostró la expresión diferencial del gen IGFBP-3 causada por la transducción de las proteínas E6/E7 de papilomas de alto y bajo riesgo, que es positiva en lesiones de alto grado y en cáncer invasor; mas no en cérvix normal (58).

Los cambios en la expresión debidos a la integración viral en el genoma del hospedero también se han estudiado, demostrando que la integración de VPH 16 en queratinocitos aumenta la expresión de genes de respuesta a IFN (59) y recientemente los perfiles de expresión en microarreglos se han utilizado para la clasificación molecular de diferentes enfermedades con base en su respuesta a radioterapia, lo que permite predecir el comportamiento tumoral post-radiación en un paciente (60).

JUSTIFICACIÓN

El cáncer cérvico uterino es la primera causa de muerte en la población femenina de México, por lo que es un problema prioritario de salud de nuestro país.

La detección temprana de este tipo de neoplasia permitiría su curación en la mayoría de los casos. Hasta el momento la única forma de determinar la presencia de una lesión es mediante la prueba de Papanicolau, que se basa principalmente en el estudio de figuras morfológicas sugerentes de alteraciones neoplásicas en las células obtenidas durante el análisis citológico. La evidencia clínica indica que además de la infección por VPH de alto riesgo, que es el factor de riesgo más importante asociado a esta enfermedad, existen factores moleculares y/o celulares adicionales necesarios para el desarrollo del cáncer cérvico uterino. La aplicación de metodologías genómicas de alto rendimiento como los microarreglos ofrece una mayor posibilidad de determinar los genes blanco que pudieran funcionar como posibles genes marcadores o pronósticos de la enfermedad, además de su posible papel patogénico. Aunque algunos reportes existentes utilizan esta tecnología en CaCU, estos sólo describen la relación entre la infección por tipos específicos de VPH y los cambios de expresión génica en las células. Sin embargo, en la actualidad no hay estudios sobre los posibles marcadores de expresión de la enfermedad como una entidad patológica en la población mexicana, que además de ser étnicamente distinta, presenta tipos virales diferentes.

HIPÓTESIS

Tomando en cuenta que las células tumorales del CaCU son células francamente transformadas. Encontraremos diferencias en los patrones de expresión de estas células con respecto a su contraparte normal.

OBJETIVO GENERAL

Analizar el patrón de expresión de células normales y células invasoras del Cáncer Cérvico Uterino humano.

OBJETIVOS PARTICULARES

1. Establecer y estandarizar la técnica de microarreglos de cDNA y normalización de los mismos con el fin de llevar a cabo un análisis genómico de alto rendimiento en muestras tumorales invasoras y líneas celulares derivadas de CaCU.
2. Crear una base de datos que contenga los resultados obtenidos de las muestras estudiadas.
3. Determinar el patrón de expresión del CaCU.
4. Mediante el uso de diferentes programas computacionales, identificar genes nuevos que presenten expresión diferencial entre las células tumorales y las normales.
5. Estudiar la expresión y regulación de los genes mediante RT-PCR semicuantitativa.
6. Estudiar la expresión de los genes identificados en microarreglos de tejidos en el laboratorio con el fin de validar los resultados obtenidos mediante hibridación *In situ* y/o inmunohistoquímica.

MATERIALES Y MÉTODOS

I. Tejidos frescos y líneas celulares

En el periodo comprendido entre 1999 y 2004 se capturaron 84 muestras con CaCU en el Departamento de Ginecología Oncológica del Hospital de Oncología del Centro Médico Nacional Siglo XXI-IMSS, el Departamento de Patología del Hospital de la Mujer, y del Centro Nacional de Displasias del Hospital General de México, SS (Tabla 2). En cuanto a las líneas celulares, se analizaron las líneas CaSki (VPH16), Hela (VPH18), SiHa (VPH16) y C33 (VPH-), establecidas a partir de una metástasis al intestino de un tumor epidermoide del cérvix, un adenocarcinoma del cérvix, y dos tumores epidermoides primarios, respectivamente. Por otro lado, las líneas CALO (VPH18) e INBL (VPH18) se establecieron a partir de tumores epidermoides de pacientes mexicanas en el laboratorio de Diferenciación celular y cáncer de la Escuela Superior de Estudios Zaragoza, UNAM (61-62).

Se colectaron treinta tejidos de cérvix normal provenientes de mujeres premenopáusicas obtenidos en un tiempo no mayor a dos horas post-mortem en el Departamento de Patología del Hospital General de México, de mujeres que no hubiesen muerto de causas relacionadas a cáncer o a enfermedades ginecológicas y que además no presentaran miomatosis uterina. Todos los procedimientos fueron evaluados y aprobados por el comité de ética del Instituto Mexicano del Seguro Social.

Los tejidos obtenidos se dividieron en tres secciones, la parte central se congeló en nitrógeno líquido inmediatamente después de la toma de la muestra y los extremos fueron fijados en etanol 70% durante toda la noche y embebidos en parafina. De estos bloques se cortaron secciones de 5 μm y se realizó la tinción de Hematoxilina/Eosina con el fin de que con la ayuda del Departamento de Patología del CMN Siglo XXI, se clasificara a los tejidos de acuerdo a la FIGO (Federación Internacional de Ginecología y Obstetricia) y

además que se pudieran seleccionaran las muestras de Carcinoma invasor que contuvieran al menos un 80% de células tumorales. También se verificó que el total de la muestra de los tejidos normales consistiera en tejido microscópicamente normal. Las muestras colectadas fueron clasificadas como: 10 adenocarcinomas y 74 carcinomas epidermoides. Los adenocarcinomas fueron almacenados para un uso futuro.

II. Tejidos embebidos en parafina para construir arreglos de tejido

Treinta y cinco tejidos cervicales embebidos en parafina y diagnosticados de la siguiente manera: 10 lesiones intraepiteliales de bajo grado, 10 lesiones de alto grado y 15 tejidos invasores fueron colectados de los archivos del Centro Nacional de Displasias de México. Por otro lado, quince de los tejidos cervicales normales frescos se fijaron en formamida, después de lo cual se embebieron en parafina para preparar los arreglos de tejido. De cada bloque de parafina que contenía tejidos diferentes se hicieron cortes de 5 μm para tinción de hematoxilina-eosina con el fin de identificar la región de interés. Es decir, una región que contuviera exactamente la lesión que se pretendía analizar. Posteriormente, se procedió a la construcción del arreglo utilizando una aguja para extracción de médula ósea calibre 16G mediante la cual se extrajeron cilindros de parafina de 3 mm de altura por 2 mm de diámetro de un bloque de parafina de 2.5 X 2.5 cm "en blanco", delimitándose una matriz de 5 x 5 elementos. Una vez elaborada la matriz, se utilizó una aguja de calibre 14G para obtener cilindros de tejido de cada una de las muestras de tejido embebidas en parafina, se tomo como guía el corte teñido con hematoxilina y eosina. Cada uno de estos cilindros que contenían al tejido, fueron inyectados en el bloque en blanco que ya había sido previamente perforado y una vez completada la matriz de 25 tejidos, se vertió una pequeña cantidad de parafina líquida sobre el arreglo para igualar la altura de los cilindros. Posteriormente, el arreglo se calentó en un horno a 60° C por 15 minutos para homogeneizar la parafina de los cilindros inyectados con la del

bloque en blanco y que no quedaran huecos entre los cilindros de las muestras y el bloque, los cuales pudieran causar problemas al momento de cortarlos. Finalmente el arreglo se enfrió durante 5 minutos en hielo y se obtuvieron cortes de 5 μm con un microtomo rotatorio.

III. Obtención de Ácidos nucleicos

El ADN y el ARN fueron obtenidos de los tejidos frescos utilizando homogenizadores de vidrio y 1 ml TRIZOL (Gibco BRL, Life Technologies, Grand Island NY USA). El homogenizado se incubó 15 min en Trizol a temperatura ambiente en agitación constante. Se agregó 1 ml de cloroformo, se agitó la mezcla hasta homogenizar y posteriormente se centrifugó 15 min a 12,000 g. Para la extracción del ARN se separó la fase superior de la tricapa obtenida del centrifugado y pasándose a otro tubo, haciéndose lo mismo con la extracción de ADN.

Para obtener el ARN, a la solución obtenida mediante el Trizol se agregaron 0.5 ml de Isopropanol, se mezcló y se dejó incubar por 15 min. Posteriormente se centrifugo de la misma manera que la vez anterior. Se eliminó el sobrenadante y la pastilla se lavó con etanol al 75%. Finalmente, el RNA se suspendió en 20-50 μl de agua MilliQ. El posible DNA contaminante de las muestras se eliminó mediante el tratamiento con DNAsal libre de RNAsa (Ambion Inc. Austin Texas USA).

Para obtener el ADN se procedió de la siguiente manera: la fase obtenida del Trizol se resuspendió en 1 ml de agua MilliQ, y se agregaron 200 $\mu\text{g/ml}$ de Proteinasa K. Esto se incubó toda la noche a 55° C y al día siguiente se extrajo el ADN con el kit Wizard Genomic DNA Purification kit (Promega), siguiendo las instrucciones del proveedor. Brevemente, al tubo con Proteinasa K se le agregaron 3 μl de solución de RNAsa, se mezcló y se incubó a 37°C por 20 min, se dejó enfriar a temperatura ambiente y

posteriormente se agregaron 200 μ l de solución de precipitación de proteínas. El tubo se centrifugó a 12,000g por 5 min y el sobrenadante se pasó a otro tubo, al cual se le agregaron 600 μ l de isopropanol, se mezcló suavemente por inversión y se volvió a centrifugar bajo las mismas condiciones, se eliminó el sobrenadante y se dejó secar la pastilla. Esta se resuspendió en 100 μ l de agua MilliQ.

IV. Cuantificación y Verificación de la calidad del ADN

Tanto el ADN como el ARN fueron cuantificados por espectrofotometría y con el fin de verificar su calidad, las muestras se corrieron en un gel de agarosa al 1% con bromuro de etidio, a 100 Volts durante 90 min. El gel se visualizó con el analizador de imágenes Eagle-eye (Stratagene).

V. Detección y Tipificación de VPH

La detección de VPH se realizó con la técnica de PCR con los iniciadores consenso MY09 y MY11 para la región L1 del genoma viral, seguida de secuenciación automatizada siguiendo el protocolo de Sequenase™ (Applied Biosystems). Tabla 2.

Muestra	Diagnostico	E6	MY09/MY11	Tipo Viral
1	CaCU	-	+	53
2	CaCU	+		16
3	CaCU	+		16
4	CaCU	-	+	X
5	CaCU	-	+	X
6	CaCU	-	+	53
7	CaCU	+		16
8	CaCU	-	+	X
9	CaCU	+		16
10	CaCU	-	+	69
11	CaCU	-	+	X
12	CaCU	-	+	X
13	CaCU	-	+	X
14	C. In situ	-	+	X
15	CaCU	+		16
16	CaCU	+		16

17	CaCU	+		16
18	CaCU	+		16
19	CaCU	+		16
20	CaCU	+		16
21	CaCU	+		16
22	CaCU	+		16
23	CaCU	+		16
24	CaCU	+		16
25	CaCU	+		16
26	CaCU	+		16
27	CaCU	+		16
28	CaCU	+		16
29	CaCU	+		16
30	CaCU	+		16
31	CaCU	+		16
32	CaCU	+		16
33	CaCU	+		16
34	CaCU	+		16
35	CaCU	+		16
36	CaCU	+		16
37	CaCU	+		16
38	CaCU	+		16
39	CaCU	+		16
40	CaCU	+		16
41	CaCU	+		16
42	CaCU	+		16
43	CaCU	+		16
44	CaCU	+		16
45	CaCU	+		16
46	CaCU	+		16
47	CaCU	+		16
48	CaCU	+		16
49	CaCU	+		16
50	CaCU	+		16
51	CaCU	+		16
52	CaCU	+		16
53	CaCU	+		16
54	CaCU	+		16
55	CaCU	+		16
56	CaCU	+		16
57	Adenocarcinoma	+		16
58	CaCU	+		16
59	CaCU	+		16
60	CaCU	-	-	x
61	CaCU	-	-	x
62	Adenocarcinoma	+		16
63	CaCU	-	-	x

64	CaCU	-	-	x
65	CaCU	-	-	x
66	CaCU	-	-	x
67	Adenocarcinoma	+		16
68	Adenocarcinoma	-	-	x
69	CaCU	-	-	x
70	CaCU	+		16
71	Adenocarcinoma	-	-	x
72	Adenocarcinoma	+		16
73	Adenocarcinoma	+		16
74	Adenocarcinoma	+		16
75	CaCU	+		16
76	CaCU	-	+	18
77	CaCU	+		16
78	Adenocarcinoma	-	-	x
79	Adenocarcinoma	-	-	x
80	CaCU	+		16
81	CaCU	+		16
82	CaCU	+		16
83	CaCU	+		16
84	CaCU	+		16
85	CaCU	+		16

Tabla 2.

Muestras de Tejido de Cáncer Cérvico Uterino colectadas en el período de tiempo comprendido entre 1999 y 2004.

VII. Extracción de ARN mensajero

Una vez seleccionadas las muestras a hibridar, las que dieron la calidad y la cantidad de ARN adecuadas se sometieron al kit PolyAtract mRNA Isolation System (Promega) para obtener los 500 ng de mRNA necesario para hibridar los arreglos. En breve, se agregó oligo-dT biotinilado, el cual hibrida con el poliA de los ARNs mensajeros. Posteriormente se agregaron al tubo miniesferas para-magnéticas cubiertas con estreptavidina, a las cuales se adhiere la biotina del oligo-dT con su contraparte (poliA del mensajero) unida. A continuación, se capturan los mensajeros en columnas magnetizadas y el resto se eluye. Los mensajeros obtenidos se resuspendieron en 20 µl de agua MilliQ.

VII. Síntesis y marcaje de las sondas.

Para llevar a cabo la síntesis y el marcaje de la sonda se utilizó la técnica de PCR radiactivo. Brevemente, se utilizaron los mARNs de HeLa, SiHA, CaSKi, INBL, CaLO, C33, tres muestras de tejido normal VPH- y 8 muestras de tejido invasor VPH16 como sondas. Como control interno de la reacción, se utilizaron 6 regiones de *S. cerevisiae*. Cada una de las muestras se mezcló con este control que funcionó como control interno de la transcripción y como sonda control de la hibridación.

Se mezclaron los siguientes reactivos y se calentaron a 42°C durante 2 horas.

2 ul	10x OligodT anclado
500 ng	mRNA de la muestra a analizar
1 ul	RNA control de <i>S. Cerevisiae</i>
2 ul	Buffer de reacción
2 ul	10x Strip dNTPs
3 ul	(α - ³³ P) dATP
2 ul	MMLV RT (200 U/ul)

La mezcla se llevó a un volumen final de 20 μ l en agua MilliQ.

Se eliminaron los nucleótidos no incorporados agregando 600 μ l de agua y posteriormente se colocó la mezcla en una columna de purificación (NucAway Spin Column, Ambion). Se centrifugó durante 1 min a 735g. La columna se pasó a otro tubo, y se eluyó con 50 μ l de agua. Para su cuantificación, se colocó una alícuota del eluido final en un vial con líquido de centelleo. Se utilizaron sólo aquellas muestras que contenían 10-15 x 10⁶ cpm del cADN purificado, finalmente estas se diluyeron con 0.5 M EDTA para llevar a una concentración de 10 mM.

VIII. Comprobación de la eficiencia de marcado

La eficiencia del marcado se comprobó en una reacción idéntica a la descrita, en donde se sustituyó el nucleótido marcado con el (α - ^{33}P) dATP por un nucleótido marcado con biotina. Una vez marcada, esta sonda se sometió a electroforesis en gel de agarosa al 0.8%, después de correr a 60 V por 2 h, el gel se lavó dos veces durante 15 minutos en 0.25 M HCl, seguido de dos lavados por 15 min en 0.5 M NaOH /1.5 M NaCl y un lavado de 15 min en 20X SSC (3.0 M NaCl, 0.3 M NaAc). Finalmente, la sonda fue transferida del gel a una membrana de nylon mediante vacío durante 90 min en 10x SSC. Una vez transferida la sonda, se secó la membrana y se entrecruzó con luz ultravioleta durante 3 min. La marca se detectó mediante la unión de un complejo avidina-fosfatasa alcalina a la biotina, lo cual se reveló con azul de tetrazolio, obteniéndose un precipitado azul en las áreas de reacción.

IX. Características de los arreglos.

Los arreglos utilizados fueron ULTRArray Advantage System, Ambion Inc., que son membranas de nylon impresas con diferentes productos de PCR con 16,800 spots por membrana, y 8,400 clonas individuales, aproximadamente, ya que cada clona está impresa por duplicado, además de que el arreglo contiene 10 controles internos positivos, 3 negativos y 8 controles de alineación para la lectura. El tamaño de la membrana es de aproximadamente media hoja carta. Los controles positivos son genes de expresión constitutiva como: la cadena ligera de la clatrina A, la proteína ribosómica L17, la Fosfolipasa A2 del grupo IV, las proteínas ribosómicas L37, S24, y L29, el translocador del nucleótido adenina, actina, rubulina, y la creatina cinasa. Los controles negativos son secuencias de PCR que no hibridan con secuencias derivadas de mamíferos, estas secuencias fueron generadas por el proveedor así como las secuencias de alineación. El

tamaño de las secuencias inmovilizadas en la membrana es de aproximadamente 500 bases.

X. Hibridación

Inicialmente se calentó el buffer de hibridación a 68° C por 30 min y los arreglos (ULTRArray Advantage System, Ambion Inc) se enrollaron y se colocaron en botellas de hibridación, teniendo cuidado de que no se traslaparan. Se agregó el buffer y los arreglos se pre-hibridaron a 60°C durante 1 hora con agitación constante.

Por otro lado, se calentaron 50 µl de ADN de esperma de salmón sonicado durante 5 min a 95°C, que así se agregaron al buffer de hibridación con el fin de evitar hibridación cruzada entre secuencias repetitivas. Inmediatamente después, se agregó la sonda previamente marcada, calentada a 95° C por 5 min y la botella se mezcló por inversión. Los arreglos y las sondas se dejaron hibridando en rotación continua a 60°C por 16 horas.

XI. Lavados Post-hibridación.

Las soluciones de lavado se calentaron a 60° C y posteriormente se realizaron dos lavados, con 50 ml, por 30 min, a 60° C, con buffer de baja astringencia (2x SSC, 0.5% SDS) y el procedimiento se repitió pero con el buffer de alta astringencia (0.5x SSC, 0.5% SDS).

XII. Captura de la Imagen

Para llevar acabo el análisis del CaCU mediante microarreglos se capturaron las imágenes de la siguiente manera, primero, cada arreglo fue humedecido con agua y cuidadosamente envuelto en plástico. Estos se alinearon y se colocaron sobre pantallas intensificadoras del PhosphorImager, con la cara hacia la pantalla. El tiempo de exposición fue de 3 horas, ya que se noto que a mayor tiempo de exposición la señal de

fondo aumentaba. Posteriormente, la imagen se capturó en el Phosphorimager generando imágenes con terminación .gel. Las condiciones fueron las siguientes: Bits/Píxel 16, Min Value= 0, Max Value = 99999.886, Pixel size 200 micrones, Resolución 50 dots/cm. El análisis de la imagen fue llevado a cabo con el Software ArrayVision Versión 8.0 Rev 4.0TM, que comprende los siguientes pasos:

a) Plantilla. Este software contiene una plantilla para el análisis de los arreglos utilizados. La plantilla es una cuadrícula que se sobrepone digitalmente en la imagen digitalizada del microarreglo, esta tiene un diseño matricial, como el arreglo, que permite identificar a cada spot como una clona determinada.

b) Eliminación de "Detritos". Un detrito es definido como un pequeño componente cuyo valor del píxel varía con respecto a los que lo rodean dentro del spot. Este programa permite excluir estos artefactos que pueden hacer variar la intensidad final del spot (promedio de todas las intensidades de los píxeles que conforman el spot).

c) Fondo. El fondo para nuestro análisis se definió por nosotros por puntos discretos en las cuatro esquinas de los spot estudiados, formando un cuadrado. El área formada entre este cuadrado y el diámetro del spot se consideró como área de "fondo". Se reportó el fondo para cada spot, y se sustrajo del valor de la intensidad del spot.

d) Normalización con diferentes opciones.

e) Base de datos. Los datos se almacenaron en forma tabular (Tabla 1). Cada columna contiene datos característicos como: intensidad de A, Intensidad de B, Nombre de las clonas etc. Cada muestra se almacenó en una base de datos independiente.

XIII. Análisis de los datos

Se creó una base de datos nueva en donde se obtuvo la relación T/N para cada transcrito presente en los arreglos. Cada gen está impreso por duplicado en cada arreglo. Posteriormente se obtuvo el logaritmo base 2 de cada una de las razones y este valor se normalizó para llevar a cabo el análisis de agrupamiento jerárquico. Este análisis se utilizó para determinar, con base en similitudes de expresión, si las muestras presentan patrones parecidos o compartidos. Esto se realizó como un primer abordaje del análisis de nuestros datos. El análisis de agrupamiento jerárquico se llevó a cabo utilizando el programa "Cluster". Para visualizar el dendrograma, se usó el programa TreeView (<http://rana.lbi.gov/EisenSoftware.htm>).

Con el fin de actualizar los identificadores de los genes presentes en los arreglos, se obtuvieron las tablas "UniGene_2", que contienen 115,693 registros de genes agrupados en la base de datos UniGene, distribuidos a lo largo de todo el genoma humano, así como la tabla "KnownGene", que contiene los datos referentes a 43,232 genes conocidos a lo largo del genoma, a través de la página del navegador del genoma humano de la Universidad de California en Santa Cruz (<http://genome.ucsc.edu/cgi-bin/hgText>) y el programa Genome Browser (<http://genome.ucsc.edu>). Estos identificadores deben ser frecuentemente actualizados debido a que la base de datos de Genbank cambia sus identificadores aproximadamente cada cuatro meses, interrumpiendo así la continuidad de los análisis de datos.

Posteriormente, con el fin de determinar cuantitativamente las diferencias de la expresión génica entre células tumorales y normales en diferentes rutas metabólicas, se utilizó la base de datos del Gene Ontology (GO) Consortium (<http://www.geneontology.org>), Fatigo (<http://fatigo.bioinfo.cnio.es>) y el programa GenMapp (<http://www.genmapp.org/>). El GO describe los productos con base en: los procesos biológicos a los que pertenecen, los componentes celulares con que interactúan y funciones moleculares que realizan o a

las que están sujetos. Este análisis muestra un grupo de términos y relaciones a las cuales se ve sometido un gen, por lo cual se puede interpretar la función de un gen en particular o bien un producto de un gen o bien un grupo de productos de genes involucrados en una patología.

Con la colaboración activa del grupo de Biología Teórica de la Universidad LaSalle y utilizando el paquete limma: Linear Models for Microarray Data y "R" un programa computacional estadístico, pudimos obtener algunos genes que fueran específicos de las líneas celulares y no de las muestras, y viceversa, así como genes que sólo estuvieran presentes en las muestras infectadas con tipos virales 16.

XIV. RT-PCR semicuantitativa

Con el fin de validar los resultados obtenidos por los arreglos, se valoró la sobreexpresión de algunos transcritos mediante RT-PCR convencional en muestras tumorales, líneas celulares y muestras normales. Doscientos ng de ARN total de las muestras a estudiar fueron reversamente transcritos con el kit para RT-PCR de un solo paso Access System (Promega, Madison, WI). Los iniciadores para cada uno de los genes de interés (Tabla 3) se diseñaron con el programa disponible en la red (<http://biotools.idtdna.com/gateway>). El criterio para seleccionar cada uno de estos genes dependió de la pregunta que se quería responder. Ver Anexos 3 al 7. Las condiciones de la RT-PCR para sintetizar la primera cadena del cDNA fueron: 48° C por 45 min; 94° C por 2 min para desnaturalizar el templado y para sintetizar la segunda cadena y la amplificación del DNA: 94° C por 30 seg, temperatura de TM específica para cada uno de los iniciadores por 1 min y 69° C por 2 min en un total de 24 ciclos, seguidos por un solo paso a 68° C por 7 min. Los productos se visualizaron en geles de agarosa al 1.5 % teñidos con Bromuro de etidio, y la señal de las bandas se capturó con el Eagle eye system (Stratagene, La Jolla, Ca) y se analizaron con el MetaView analysing system (version 4.5 Universal Imaging Corp., USA). Las

concentraciones relativas de cADN se expresaron como la razón de las lecturas densitométricas de las bandas de los genes y su relación con el cADN de β -actina.

Nombre del Gen	Secuencia	TM	Tamaño pb
Matriz Metaloproteasa 10	5'-CATTTCAGTCTCTCTACGGACCT-3' 5'-CAGCTATTAGTCTAGGGAAGCC-3'	60	509
Interleucina 6	5'-AGTTGCCTTCTCCCTGG-3' 5'-TGAGGGCTCTTGGGCAAAT-3'	55	621
NDUFB6	5'-GTACACTCCGGATGAGAACTG-3' 5'-CTCCAGAATTGTATCACCAGGG-3'	60	328
Importin 7	5'-TTGGCCTGATCGAGAAACAGCA-3' 5'-CTTGGAACATTGCCAGGGCTT-3'	60	649
Laminin gama 2	5'-TGTTTAGCTCAGCCCAACGACT-3' 5'-AGCAGCTGAACCCGTTATGACA-3'	56	696
Claudina 1	5'-GTGGAGGATTTACTCCTATGCC-3' 5'-CTTGGTGTGGGTAAGAGGT-3'	55	504
Catepsina F	5'-GTGCTGATCAGAGTGGCTGCTGC-3' 5'-AGTTTCCYGGACATGGAGAGGGAC-3'	67	396
Matriz Metaloproteasa 12	5'-TCACGAGATTGGCCATTCTT-3' 5'-TCTGGCTTCAATTCATAAGC-3'	55	370
Matriz Metaloproteasa 11	5'-CCATGGCAGTTGGTGCAGGAGCAG-3' 5'-TGCAGTCATCTGGGCTGAGACTCA-3'	66	399
FZD2	5'-CTGAGCTCAAGTTCTTCTGTG-3' 5'-CAGGAAGTGGTAGTTGAGGTAG-3'	60	386
GSK3B	5'-GTACTATAGGGCACCAGAGTTG-3' 5'-GTAGCCAGAGGTGGATTACTTG-3'	60	504
C-Myc	5'-AGCTTGTACCTGCAGGATCTGA-3' 5'-TGACACTGTCCAACCTTGACCCT-3'	60	469
PPARD	5'-GGTTACCCTTCTCAAGTATGGC-3' 5'-CTCTGAAGACAGGAGACAGAAG-3'	60	86
β -actin	5'-TGAAGTCTGACGTGGACATC-3' 5'-GTTTCGTTCTCATACTGCTCA-3'	55	243

Tabla 3.

Iniciadores diseñados para utilizarse en RT-PCR semicuantitativa y utilizados para validar los resultados obtenidos por los microarreglos

XV. Hibridación *In Situ* (HIS)

Se obtuvieron cortes de 5 μ M de los arreglos de tejido y de algunas otras muestras incluidas de manera individual en parafina. Estos cortes fueron posteriormente montados en laminillas cargadas positivamente. Las laminillas con los cortes se desparafinaron con

xileno y se rehidrataron en baños de etanol (100, 90, 70, y 30%) para finalmente transferir a una solución de PBS (137 mM NaCl, 2.7 mM KCl, 4.3 mM Na₂HPO₄, 1.4 mM KH₂PO₄) por 10 min. Los tejidos se trataron con solución de DNAsa (1 µg/ml) por 10 min a 37°C y se lavaron tres veces con PBS. La inactivación de la peroxidasa endógena se llevó a cabo incubando las muestras en peróxido de hidrógeno al 3% (H₂O₂) en metanol. Las sondas sentido y antisentido se generaron por PCR de cadena sencilla utilizando cDNA obtenido del RNA células SiHa como templado. Estas fueron marcadas con biotina-16-dUTP (Roche). Cada tejido se cubrió con 50 µl de la mezcla de hibridación y un cubreobjetos. Esta mezcla consiste en formamida al 50%, dextran sulfato al 10%, 2x SSC (20x SSC: 3 M NaCl, 300 mM Na₃C₆H₅O₇), PBS, SDS al 2%, 100 µg/ml de DNA de esperma sonificado de salmón y 50 ng de la sonda (sentido o antisentido) marcada. Tanto el RNA del tejido como de la muestra se desnaturalizaron a 65° C por 10 min. Después de la hibridación, los cubreobjetos fueron removidos con una solución de tris y tween 20 1x (TBST 10x: 500 nmol/L TrisHCL, pH 7.6, 3 mol/L NaCl, 1% Tween 20), y los tejidos se incubaron a 55°C por 20 min en una solución de alta astringencia. Inmediatamente se aplicó un anticuerpo marcado con peroxidasa (GenPoint System, DAKO) sobre los tejidos, que se incubaron en cámara húmeda por 15 min. Los cortes se lavaron en TBST 1x. Se aplicó botini tiamida directamente sobre los cortes para amplificar la señal y e incubó a temperatura ambiente por 15 min, posteriormente se lavó con TBST 1x. Se realizó una segunda vuelta con el anticuerpo marcado con peroxidasa y la reacción colorida se reveló con Diaminobenzidina al 0.06% en 3% H₂O₂. Finalmente, las laminillas se lavaron, se tiñeron con HyE, se deshidrataron y se montaron. Los controles negativos de la hibridación *in Situ* se llevaron a cabo con la sonda sentido o bien tratando con RNAsa (100 mg/ml durante 30 min a 37° C) antes de la hibridación.

XVI. Inmunohistoquímica (IHC)

Los diferentes cortes tanto de microarreglos de tejido, como de tejidos únicos embebidos en parafina se desparafinaron y rehidrataron en alcohol. La peroxidasa se inactivó de la misma manera que en la HIS. Las laminillas se preincubaron con el suero DAKO Protein Block Serum free Cat X0909 (Carpinteria, Ca) durante 30 min a 37°C para prevenir reacciones inespecíficas. El exceso se decantó y se incubó el anticuerpo primario a la concentración requerida durante toda la noche a 4°C. Posteriormente, las laminillas se lavaron en PBS y se revelaron con el anticuerpo secundario unido a un polímero sintético (DAKO Envision System Peroxidase from DAKO, Carpinteria, Ca), incubándose a temperatura ambiente por 30 min. Se lavaron y el color de la reacción se desarrolló con diaminobenzidina al 0.06% en 3% H₂O₂. Finalmente, las laminillas se lavaron, se contratiñeron y se montaron. Los datos del arreglo de tejidos, así como los resultados de la inmunohistoquímica se depositaron y manejaron con el programa Tissue MicroArray Workbook de Genetech.

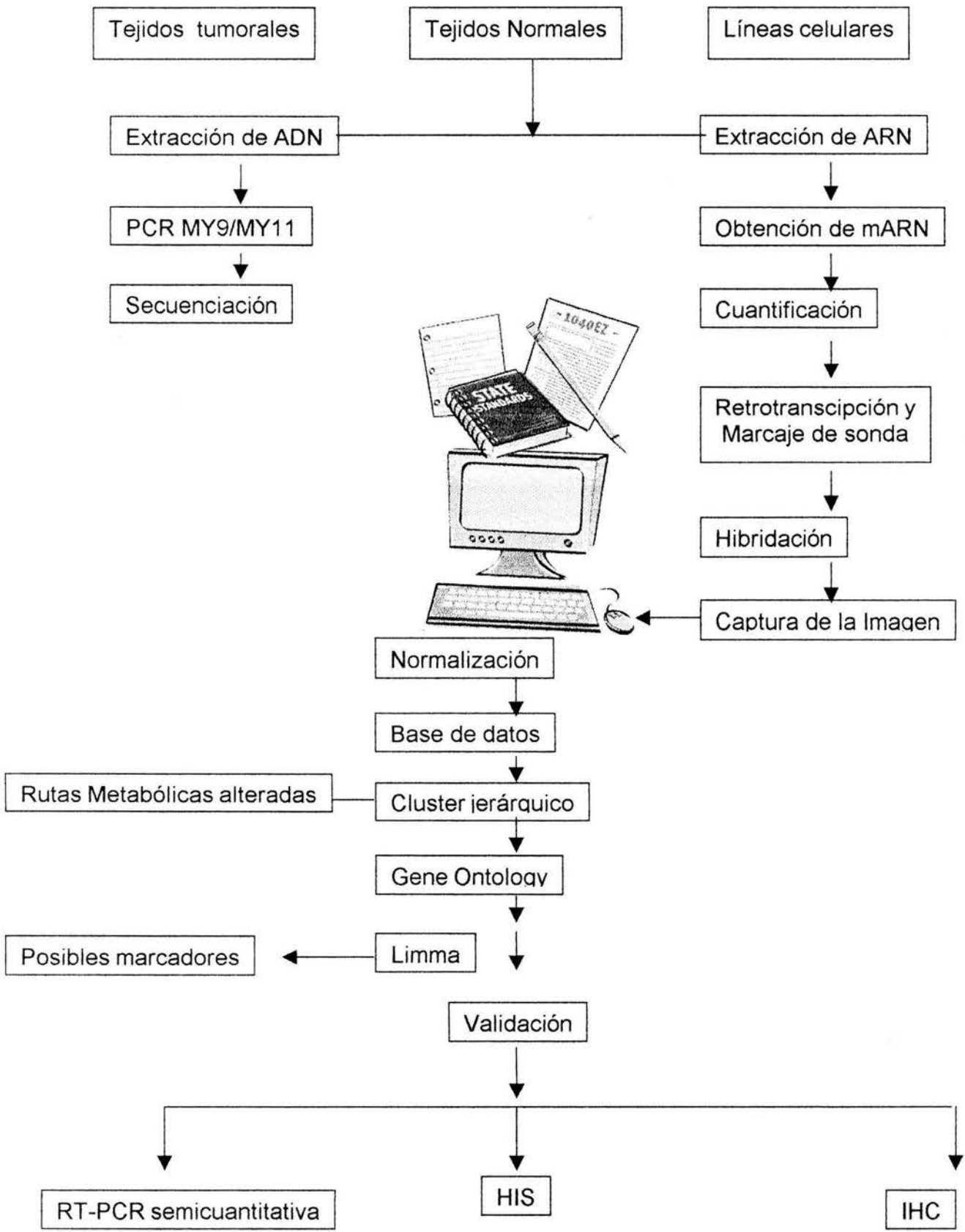


Figura 3.

Diagrama de flujo que muestra todos los pasos que se llevaron a cabo para la elaboración de este experimento.

RESULTADOS

Con la finalidad de determinar posibles marcadores moleculares del carcinoma cérvico uterino, llevamos a cabo experimentos de microarreglos de cADN para determinar el patrón de expresión específico de este tipo de neoplasia. Una vez determinado este patrón de expresión, se creó una base de datos que contiene la información de todas las muestras estudiadas tanto de los datos crudos, como de los normalizados. Esta base de datos se analizó utilizando diversos métodos bioinformáticos con el fin de determinar los posibles marcadores moleculares de la enfermedad y para encontrar las variaciones de expresión de las rutas metabólicas celulares en presencia del cáncer en el epitelio cervical humano. Finalmente, la información obtenida por los microarreglos se validó mediante RT-PCR semicuantitativa, Hibridación *In Situ* e Inmunohistoquímica; en estas dos últimas metodologías se utilizaron arreglos de tejido con el fin de evaluar a los potenciales marcadores moleculares.

1. Extracción de ADN y tipificación

Se realizó la colecta de 84 muestras de CaCU durante el periodo de tiempo comprendido entre 1999 y 2004 en el Departamento de Ginecología Oncológica del CMN Siglo XXI, IMSS, Departamentos de Patología del Hospital de la Mujer y del Centro Nacional de Displasia del Hospital General de México, SS.; de todas las muestras 73 fueron carcinomas epidermoides y 10 fueron determinados como adenocarcinomas. Estos tejidos tumorales así como los tejidos normales y las líneas celulares HeLa, CasKi, SiHA, INBL, CaLO, C33 fueron tratados con TRIZOL con el fin de obtener sus ácidos nucleicos. El DNA genómico obtenido fue analizado por PCR con los iniciadores específicos para el gen que codifica para la proteína viral E6 que detectan específicamente al tipo HPV16, esta reacción genera un fragmento de 126 pb. Por otro lado los DNAs que no resultaron

positivos para HPV16 fueron analizados por PCR con los iniciadores MY09/MY11 que detectan más de 30 tipos virales, y el fragmento que se genera de esta reacción es de 450 pb. Estos fragmentos fueron posteriormente secuenciados para obtener los tipos virales específicos. La figura 4 muestra ejemplos representativos de la calidad del ADN purificado a partir de los tejidos frescos, comparados con DNA genómico obtenido a partir de un cultivo celular. La tabla 2 de la sección de material y métodos presenta los datos de presencia de VPH en los tejidos analizados. Las figuras 5 y 6 muestran una región de la secuencia de los genes E6 y L1, respectivamente, que se utilizaron para identificar el tipo viral presente en la muestra.

Como se esperaba, las líneas Caski y SiHa fueron positivas para VPH16 y HeLa, INBL y CaLO para el tipo VPH18. El tipo viral 16 se detectó en 63 tumores invasores, el tipo viral 18 se encontró en 1 muestra, el tipo 53 se detectó en dos muestras y el tipo 69 en otra. En once casos no fue posible detectar ADN viral con los iniciadores utilizados y en siete no se llevó a cabo la secuenciación, pero se sabe que son muestras positivas para HPV, ya que el fragmento del gen viral se amplificó mediante PCR.

2. Extracción de ARN Total y obtención de ARN mensajero

Por otro lado, utilizando también Trizol se extrajo el ARN total de las muestras antes mencionadas, y solamente a aquellas muestras que cubrían los requisitos en cantidad y calidad se les extrajo el ARN mensajero. (Ver Figura 7)



Figura 4.

Corrimiento electroforetico de diferentes DNAs purificados. Medio microgramo de DNA de cada muestra fue sujeta a electroforesis y fue tenido con bromuro de etidio. Carril 1) T07, 2) T14, 3) T61, 4) T63, 5) T11, 6) T15, 7) T17, 8) T71, 9) T83, 10) T73, 11) HeLa, 12) CaLO, 13) INBL, 14) CaSKi, 15) SiHa, 16) C33, 17) N03.

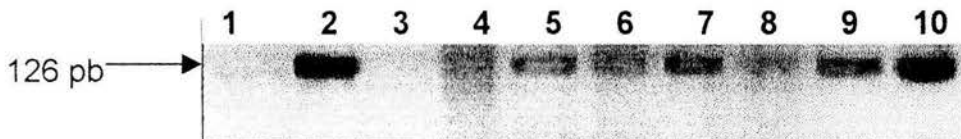


Figura 5.

Detección de amplicones de PCR con iniciadores que detectan un fragmento de 126 pb del gen E del VPH. Carril 1) MPM 100 pb; 2) control positivo; 3) control negativo sin DNA, 4) T45, 5) T56, 6) T02, 7) T23, 8) SiHa, 9) CaSKi, 10) HeLa.

a)



b)

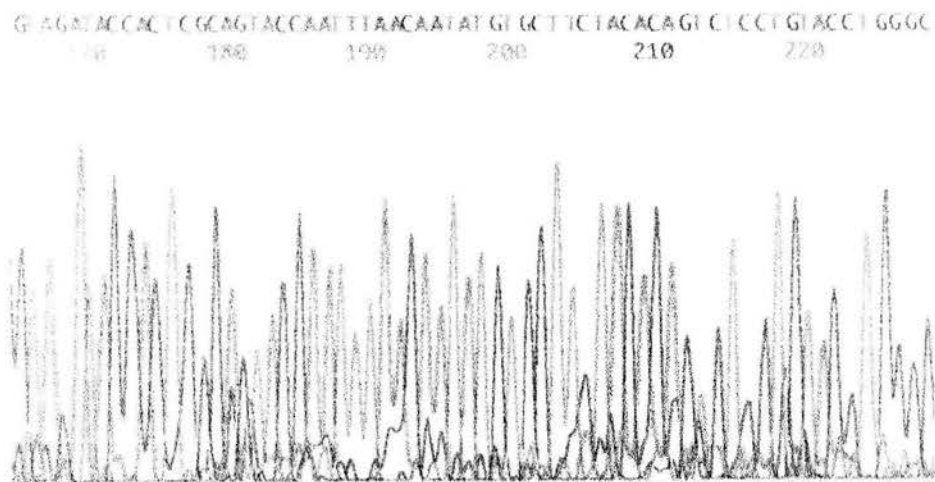


Figura 6.

Panel a) Detección de la amplicones de PCR con iniciadores que detectan un fragmento de 450 pb de la proteína L1. Carril 1) MPM 100 pb; 2) control positivo; 3) control negativo sin DNA, 4) INBL, 5) CaLO, 6) HeLa, 7) T01, 8) muestra negativa, 9) T06, 10) T10. Panel b, ejemplo de secuencia parcial obtenida a partir del uso de los iniciadores MY09/MY11.

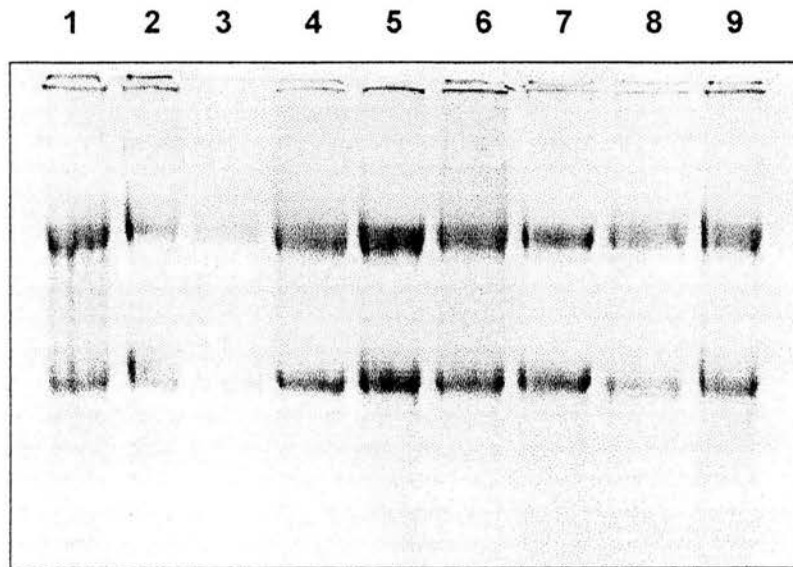


Figura 7.

Calidad del ARN mediante electroforesis. Ejemplos de ARN total purificado a partir de tejidos frescos y líneas celulares a) Carril 1) T07, 2) T11, 3) T61, 4) HeLa, 5) CaLO, 6) INBL, 7) SiHa, 8) CaSKi, 9) C33.

3. Resultados de la hibridación

Con el propósito de buscar e identificar algunos marcadores moleculares del CaCU, se utilizó una plataforma del análisis genómico de alto rendimiento: arreglos de cADN. Durante los experimentos fue necesario marcar las sondas con radiactividad y después hibridar sobre los arreglos que contenían 8,400 sondas blanco que codifican un gen único. Posteriormente, se capturó la imagen con un Phosphorimager. En la figura 8 se pueden apreciar las imágenes capturadas.

4. Normalización

Se realizó una gráfica de box plot con el fin de observar la distribución de los datos (Figura 9). Esta figura muestra los diferentes box plots para cada una de las muestras. La mediana es la línea que cruza la “caja” a la mitad. De la mediana a la parte inferior de la caja se encuentre 25 % de la distribución de los datos, así como de la mediana hacia la parte superior. Es importante hacer notar que las medianas no se encuentran alineadas, por lo que fue necesario realizar una normalización intra-arreglo, mediante LOESS (63), el cual es un procedimiento estadístico que implementa un método no paramétrico para estimar la superficie de la regresión. Este método es muy flexible ya que no asume ni necesita ningún tipo de forma paramétrica para la superficie de la regresión. LOESS es muy utilizado en situaciones en las que se desconoce cual es la mejor opción para determinar la forma paramétrica. Posterior a la utilización de LOESS se repitió la gráfica de box plot (Figura 10) y así se pudieron observar los datos normalizados, en los que la mediana se encuentra alineada entre todas las muestras.

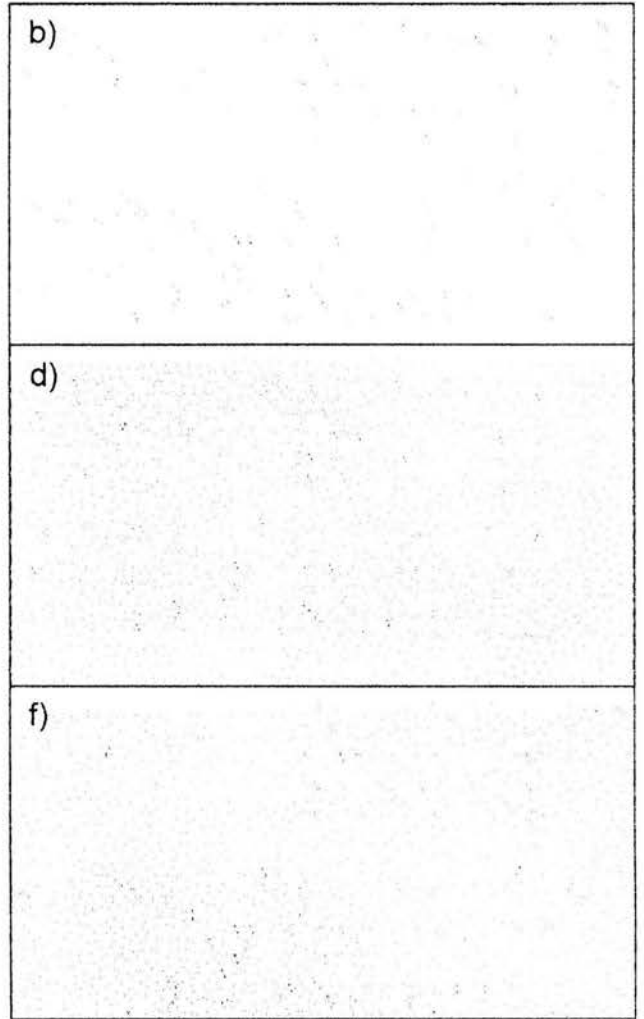
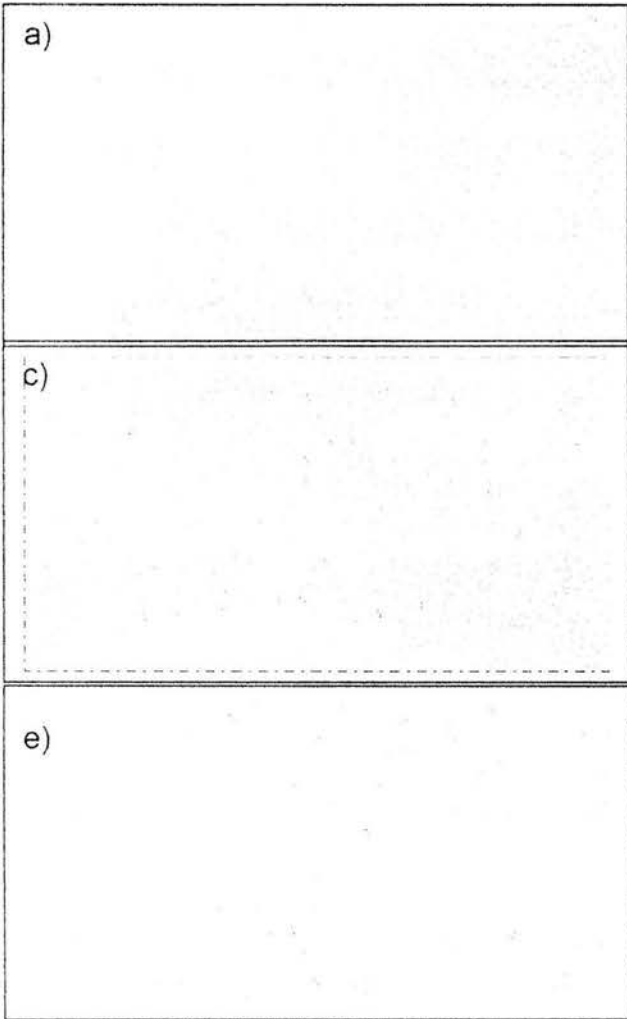


Figura 8.

Ejemplo de Imágenes capturadas con el PhosphorImager. a) Hibridación con sonda proveniente de ARN de HeLa marcada con $P^{33}\alpha$ - dATP, b) hibridación con sonda proveniente de RNA de CaLO marcada bajo las mismas condiciones que a, c) hibridación con sonda de SiHA, d) hibridación con sonda de CaSKi, e) hibridación con sonda de una muestra tumoral e f) hibridación con sonda de una muestra normal.

Original data

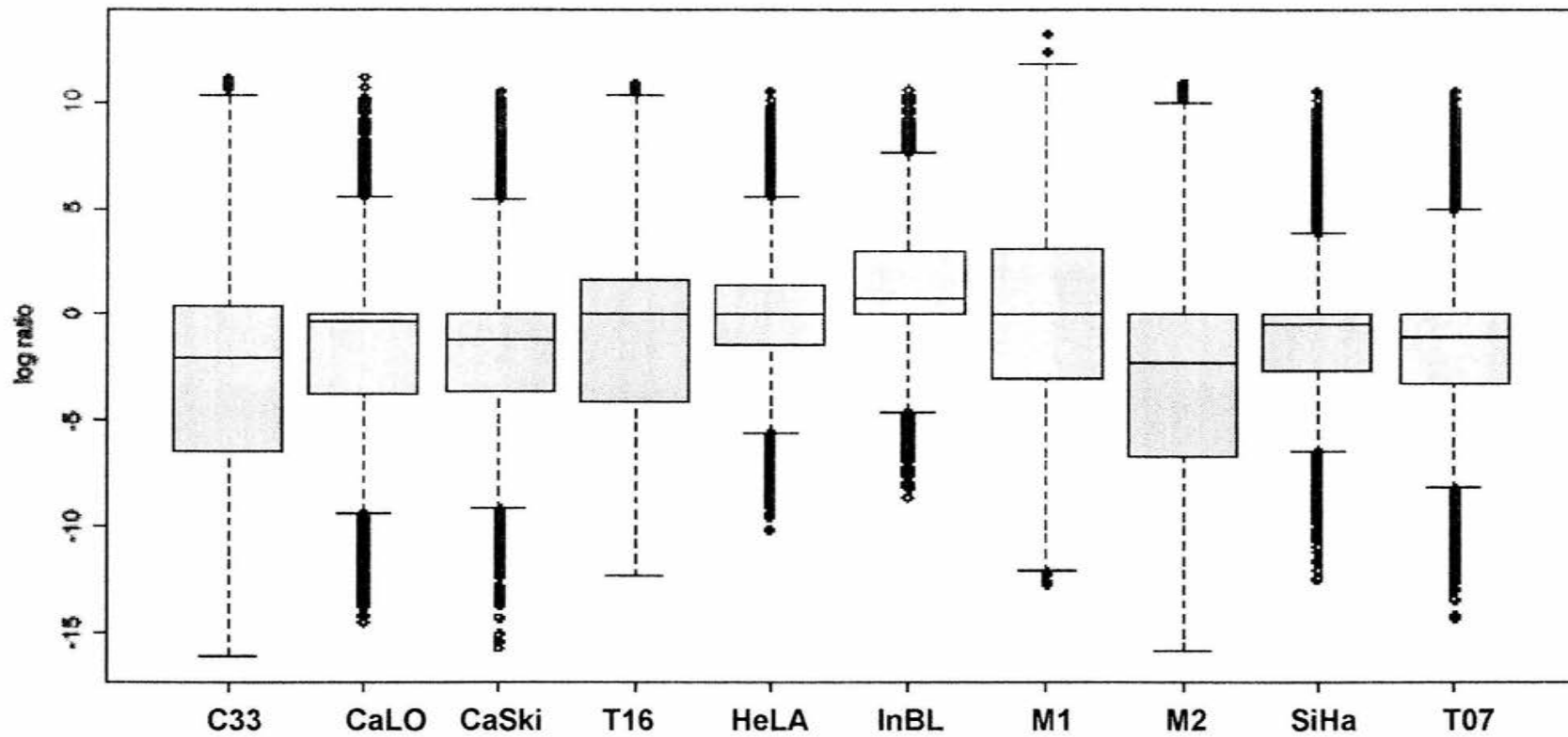


Figura 9.

Grafica de Cajas (Box Plot) de datos crudos. En el eje "Y" se grafica los datos de intensidad de expresión en logaritmo base 2 y en el eje de "X" se grafica a cada una de las muestras

Normalized data

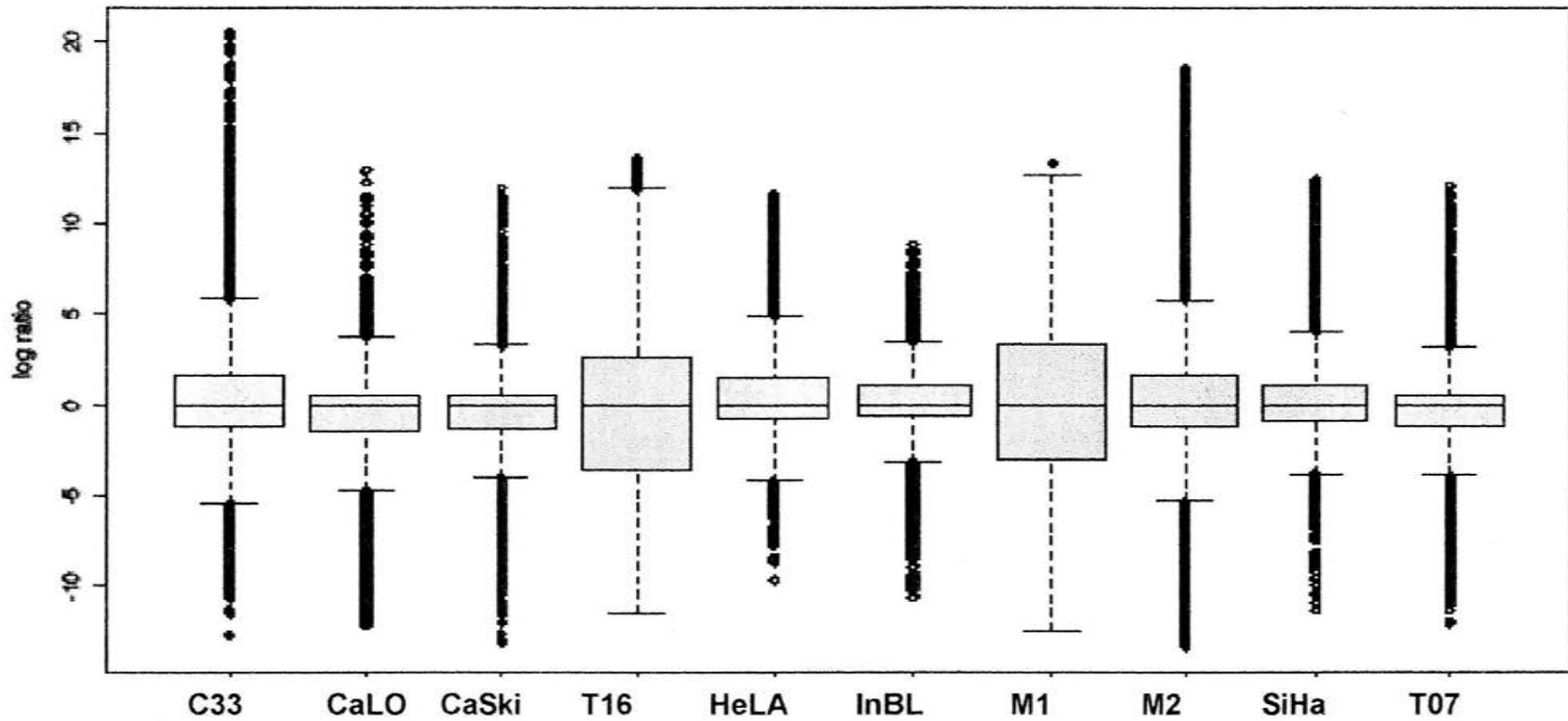


Figura 10.

Gráfica de cajas (Box Plot) de datos normalizados. En el eje "Y" se grafica los datos de intensidad de expresión en logaritmo base 2 y en el eje de "X" se grafica a cada una de las muestras

Con el fin de determinar la distribución de los datos se procedió a elaborar y analizar por medio de histogramas cada una de las muestras. Así, las muestras estudiadas no tienen una distribución normal (Figura 11); por lo tanto, el método de normalización utilizado fue el correcto. Asimismo para verificar la distribución también se realizó un QQ plot (Quantile-Quantile). Si la distribución fuera normal, ésta debería ser casi lineal y los genes que salen de esta línea se encuentran diferencialmente expresados. Con base en este análisis, que sirve como primer abordaje, podemos concluir parcialmente que si hay genes diferencialmente expresados.

5. Expresión global del Cáncer Cérvico Uterino.

Con el propósito de determinar el patrón de expresión global de CaCU, se realizó un análisis de cluster jerárquico. Esto se llevó a cabo mediante la determinación de la razón de cada una de las muestras tumorales con respecto a la mediana de los tres tejidos normales para cada uno de los genes. Estos resultados se analizaron utilizando el programa de Cluster jerárquico, el cual básicamente agrupa por niveles de expresión los genes con mayores similitudes entre sí, así como las muestras con más similitudes. Es importante hacer notar que este análisis generó un dendograma en donde se observa que las muestras que contienen el mismo tipo de VPH forman dos grupos: VPH16 y VPH18 (Ver anexo 3). Sugiriendo que probablemente, por niveles de expresión, los tipos virales tienen características que los distinguen entre si.

Por otro lado, siguiendo el sondeo del comportamiento en expresión del CaCU, se realizó otro análisis mediante el uso del software GeneSpring. En este análisis primero

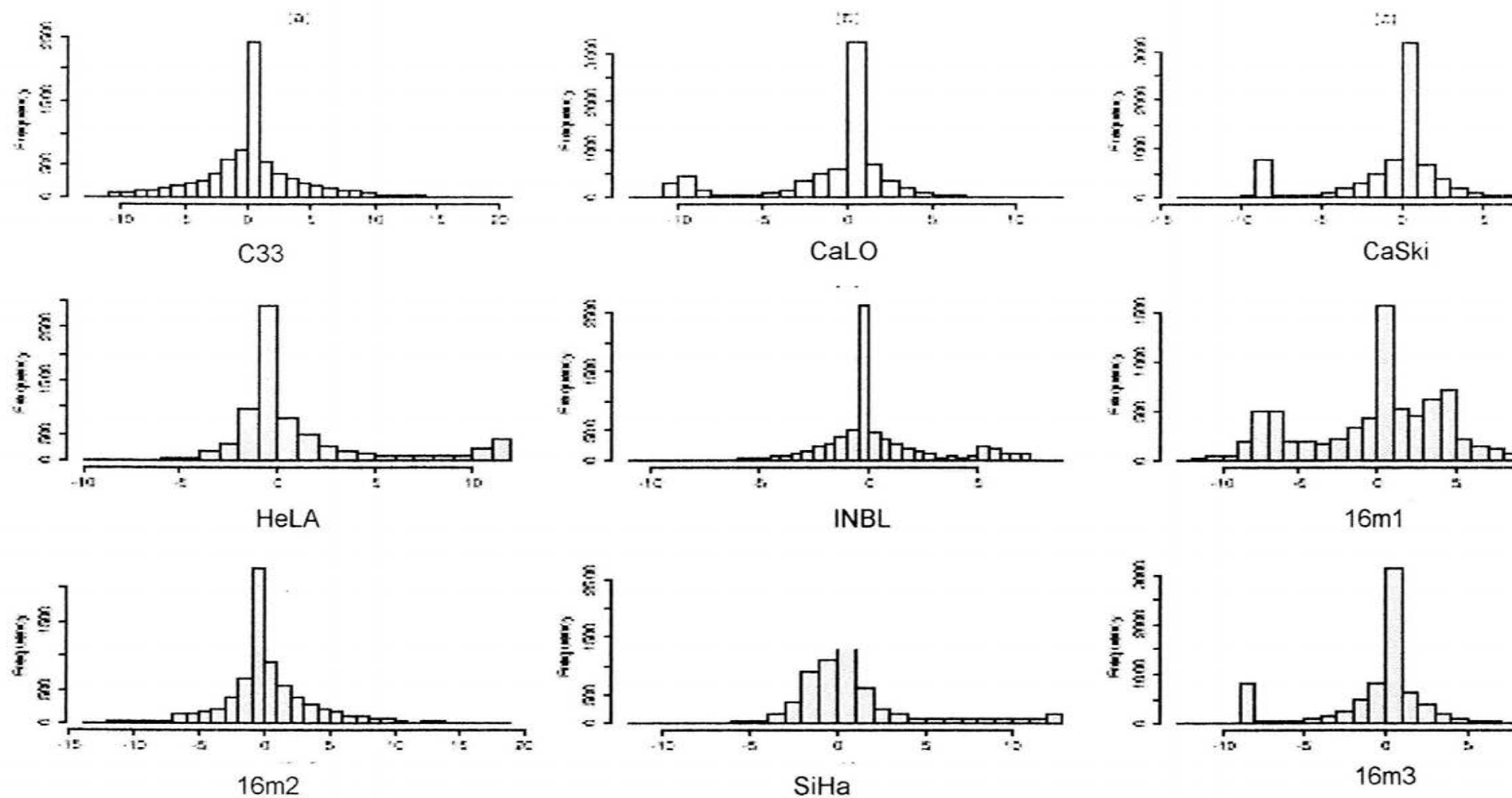


Figura 11.

Histogramas de todas las muestras estudiadas mostrando que la distribución de los datos no es normal.

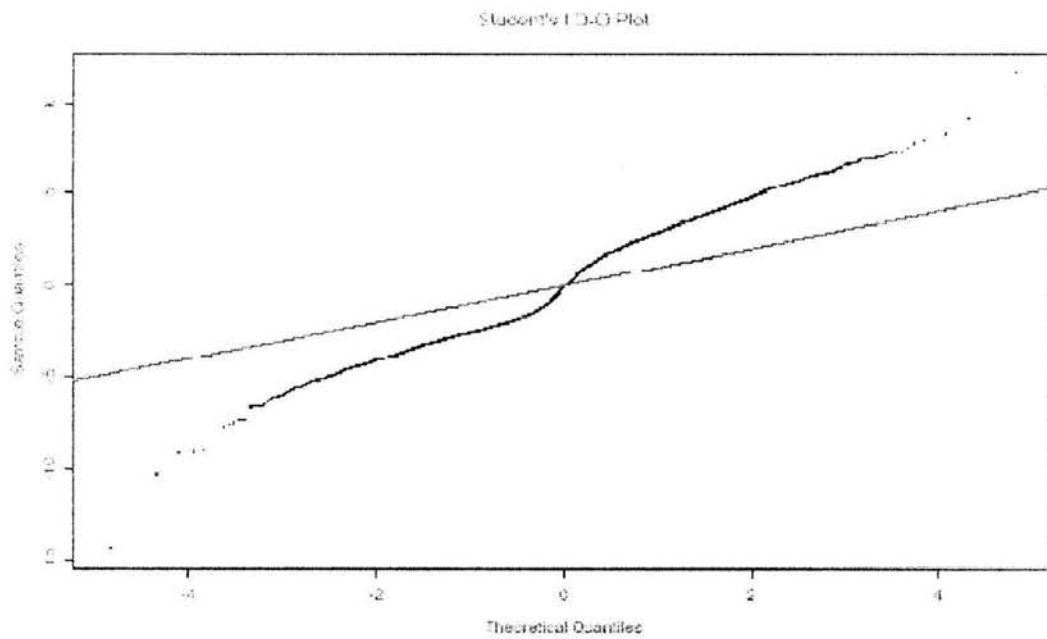


Figura 12.

QQ plot demostrando que la líneas no es recta y que los puntos que se alejan son genes diferencialmente expresados que pueden fungir como posibles genes marcadores

se realizó un corte de los datos a analizar con el fin de sólo utilizar aquellos genes que estuvieran sobreexpresados más de 10 veces con respecto al normal. Esto se hizo con el fin de eliminar falsos positivos. De los 477 genes que se trabajaron, se obtuvieron las relaciones generadas por la mediana de todas las muestras y líneas celulares VPH16+, entre la mediana de las líneas celulares VPH18 y entre los valores de C33 y la mediana de las muestras normales. Ver Figura 13. Claramente se puede apreciar el color amarillo en C33 que no se observa en las otras muestras, sugiriendo que estos genes se pueden expresar diferencialmente cuando se encuentra presente VPH. Por otro lado, el color negro se presenta en las muestras con VPH18 indicando que algunos de estos genes pudieran expresarse solamente en este tipo viral. Con el fin de determinar los valores de expresión de estos mismos datos, los datos fueron graficados de manera lineal (Ver Figura 14.) corroborando nuevamente que sí existen diferencias de expresión entre las células tumorales y las normales.

6. Procesos biológicos alterados.

Para determinar el comportamiento de las células tumorales, se realizó un análisis utilizando la base de datos pública y disponible en la red de Gene Ontology. Este programa computacional agrupa a los diferentes genes pertenecientes a una misma ruta metabólica o proceso biológico para determinar los cambios de expresión que se generan en los diversos procesos celulares. Primero se realizó un análisis exploratorio en donde se representó gráficamente la sumatoria de las medianas de los niveles de expresión de grupos de genes pertenecientes a una misma ruta (ver anexo 3). Con el fin de poder observar los cambios específicos de estos genes en cada una de las muestras se realizaron agrupamientos jerárquicos de cada uno de los grupos obtenidos del Gene Ontology Consortium.

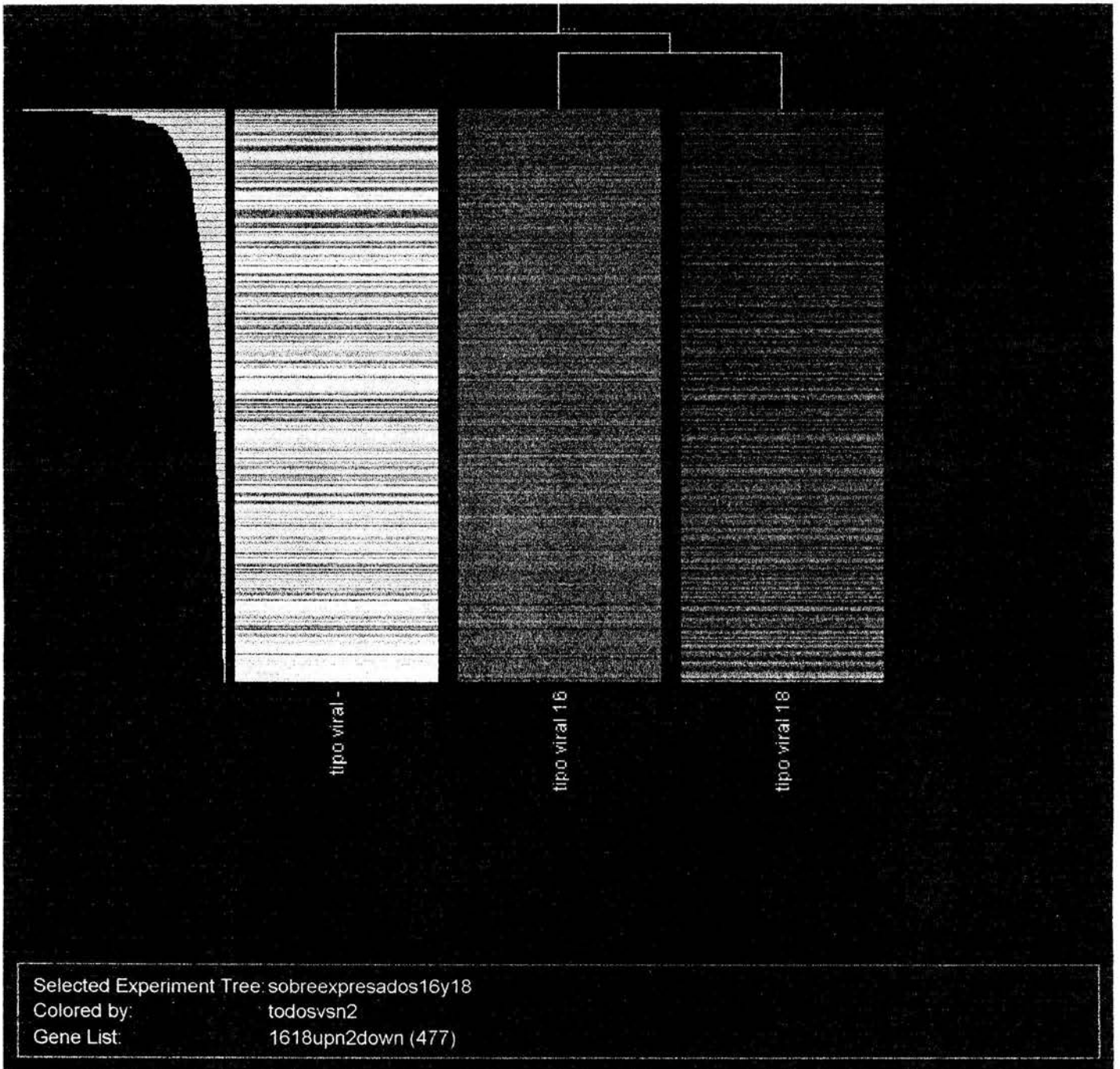


Figura 13.

Dendrograma que muestra los 477 genes más sobreexpresados en todas las muestras, y el comportamiento que tienen en estas; VPH 16, VPH18, tejido normal y C33 (tumoral VPH-)

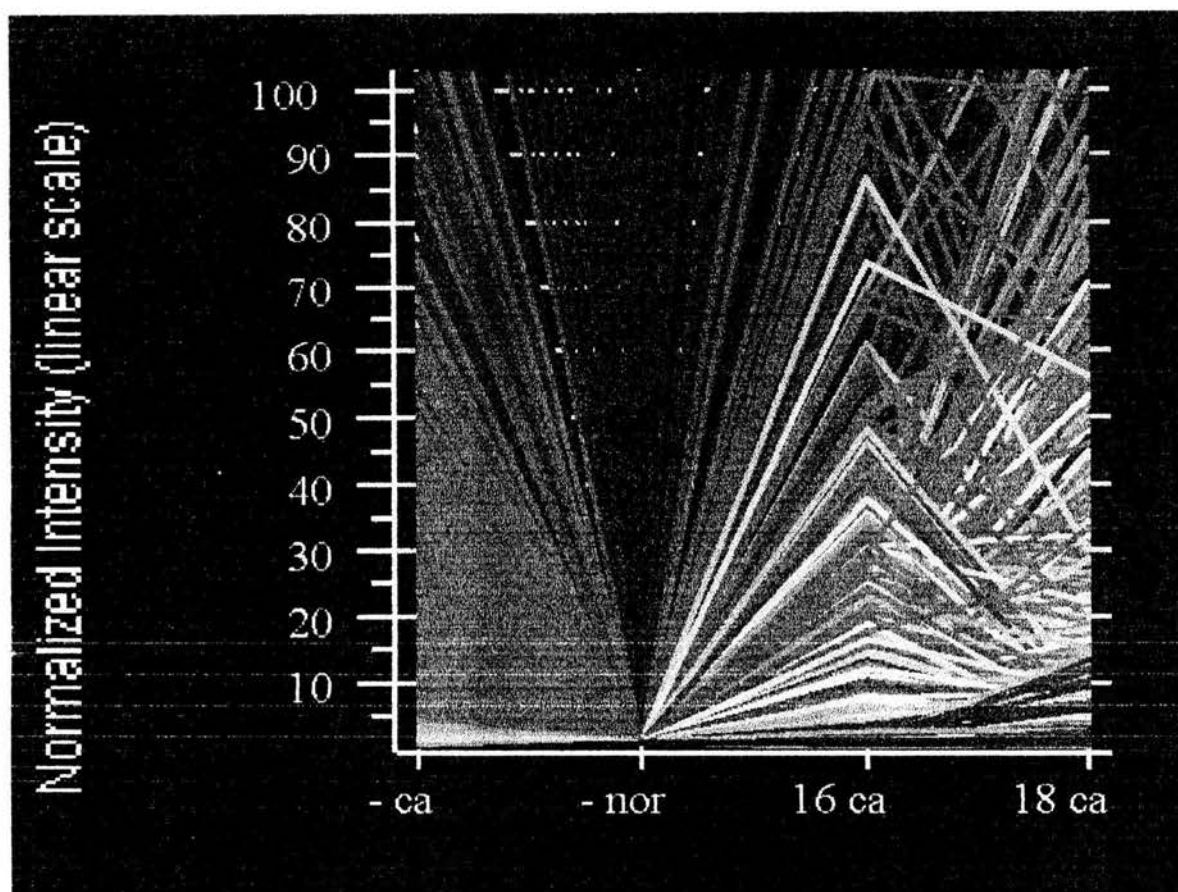


Figura 14.

Grafica que muestra los genes con más expresión en C33, muestras con VPH16 y VPH18 con niveles basales de expresión en las muestras normales. En el eje de las "Y" se muestra el nivel de expresión y en el eje de las "X" se encuentran las muestras.

Como un sondeo inicial se determinaron diferentes rutas metabólicas, procesos y componentes celulares. El primer análisis consistió en graficar los diferentes agrupamientos de genes generados por el *Gene Ontology*. Se obtuvo inicialmente la mediana para cada gen en todas las muestras y se sumaron las medianas para obtener una gráfica de barras que representa el comportamiento global de la enfermedad en dicha ruta metabólica. Los procesos y componentes celulares estudiados fueron: a) componente epitelial, genes que intervienen en proliferación; b) componente estromal, metabolismo de fármacos, ciclo celular; y c) componente tumoral. Es importante mencionar que todos los procesos celulares estudiados se encuentran principalmente sobreexpresados (ver anexo 3). Por otro lado, con el fin de hacer más visibles estas diferencias y poder determinar los cambios de expresión en cada una de las muestras estudiadas, se realizaron análisis de agrupamiento jerárquico para cada una de las rutas metabólicas y procesos celulares identificados por el Gene Ontology (Figura 15, 16 y 17). Es importante notar que los genes que componen al grupo de proliferación, interferón e interleucinas presentan tanto sobreexpresión como supresión génica de los genes que la componen.

Una de las rutas estudiadas, la de Wnt, se encontró sobreexpresada en todos los genes que la componen (Figura 18). Esta ruta participa en diferentes eventos durante el desarrollo y la embriogénesis y se sabe que genera tumores al ser activada de manera anómala. Los factores Wnt son secretados y se encuentran distribuidos a lo largo de todo el reino animal y controlan virtualmente cada decisión del desarrollo. En ausencia de la señal de Wnt las células no estimuladas regulan los niveles de β -catenina por un complejo que fosforila a esta proteína marcándola para su subsiguiente ubiquitinación y degradación en el citoplasma. Sin embargo, cuando las Wnt secretadas están presentes, que parece ser lo que sucede en CaCU, permiten la acumulación de beta catenina, lo cual hace que esta proteína se trasloque al núcleo, donde se une a diferentes proteínas de unión a DNA que sirven como factores de transcripción para muchos otros genes.

Cluster de Genes de Proliferación

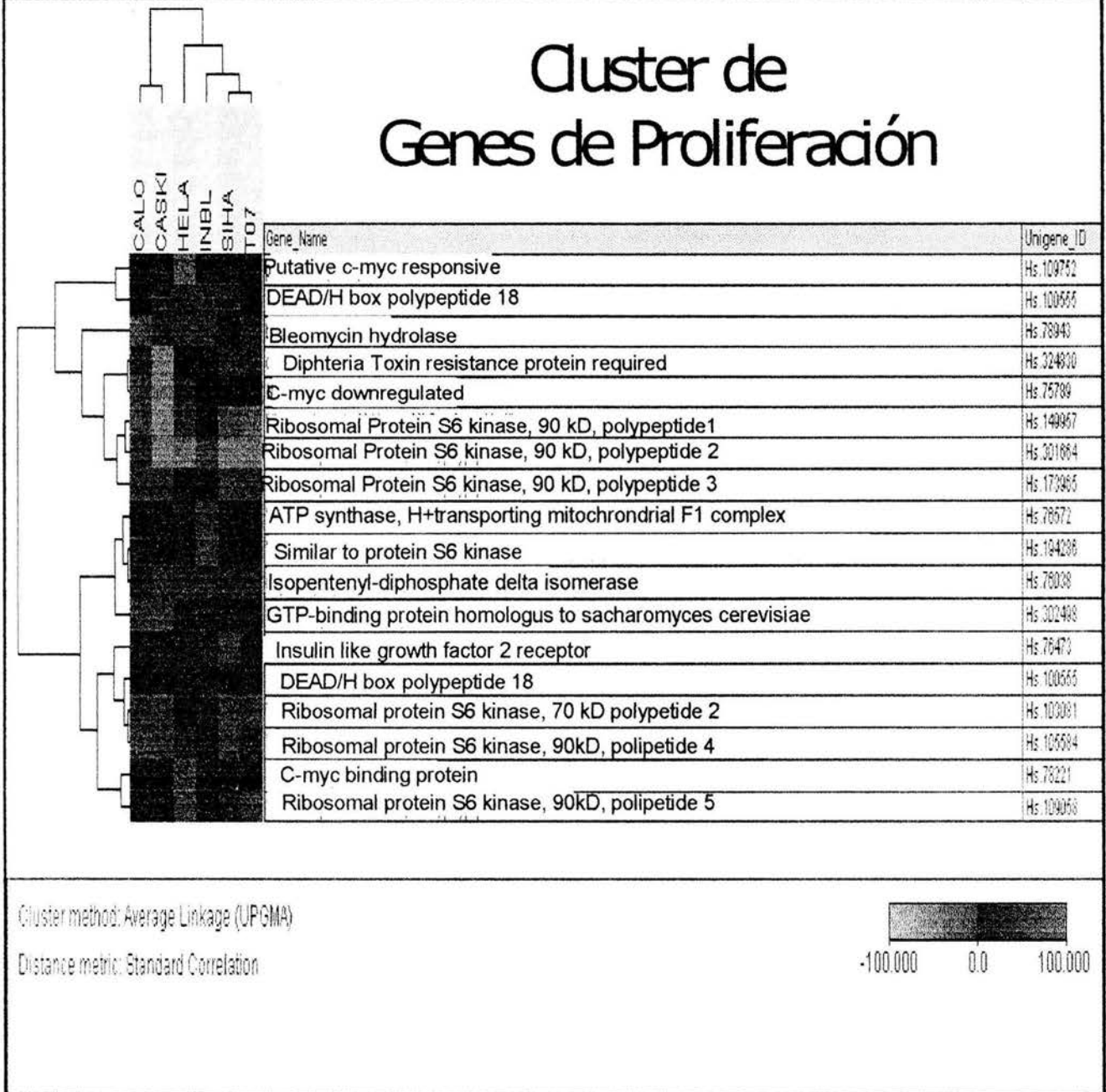
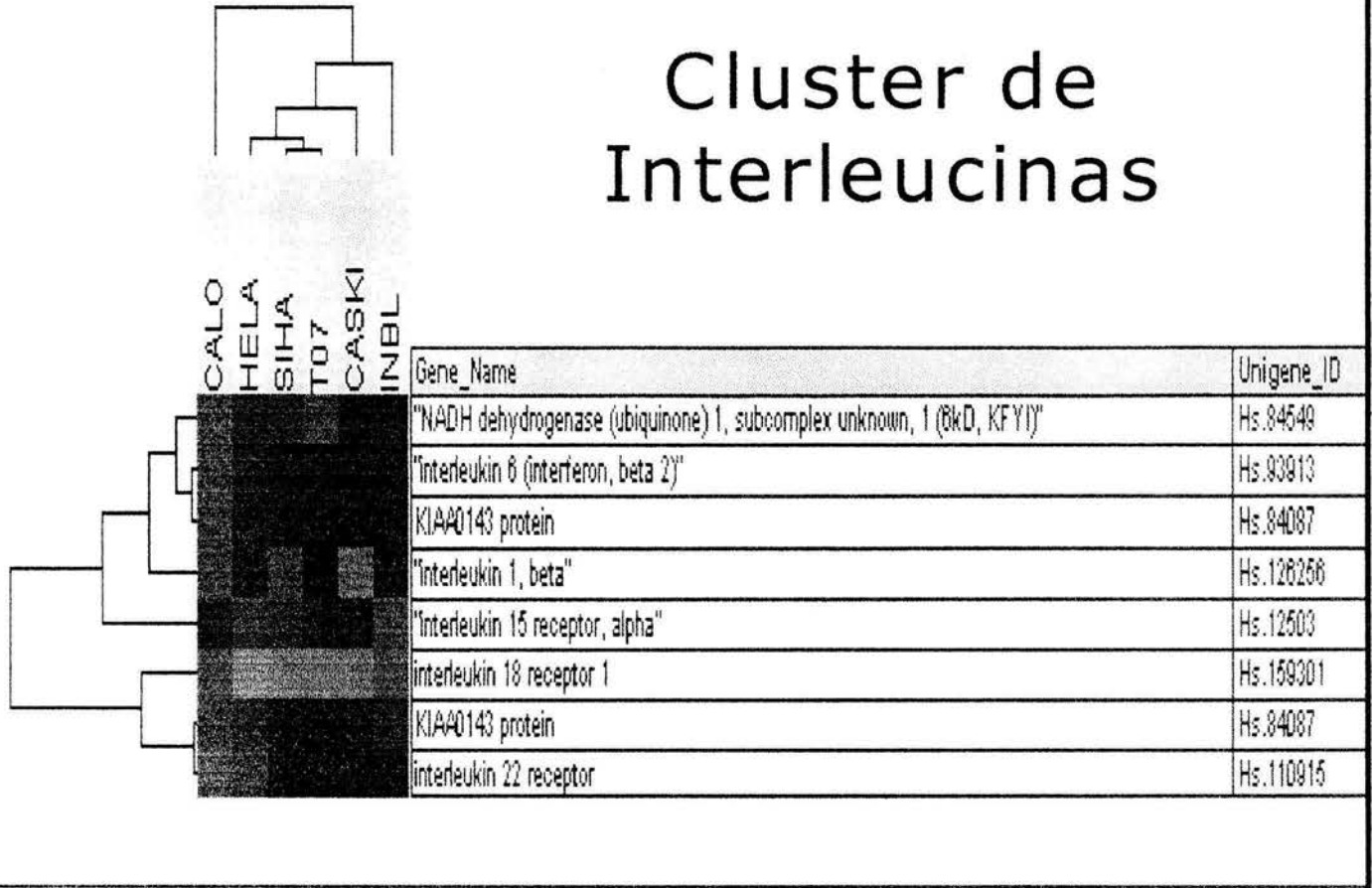


Figura 15.

Cluster jerárquico de genes pertenecientes al proceso celular de Proliferación

Cluster de Interleucinas



Cluster method: Average Linkage (UPGMA)

Distance metric: Standard Correlation

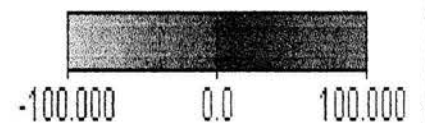


Figura 16.

Cluster jerárquico de genes pertenecientes o regulados por interleucinas.

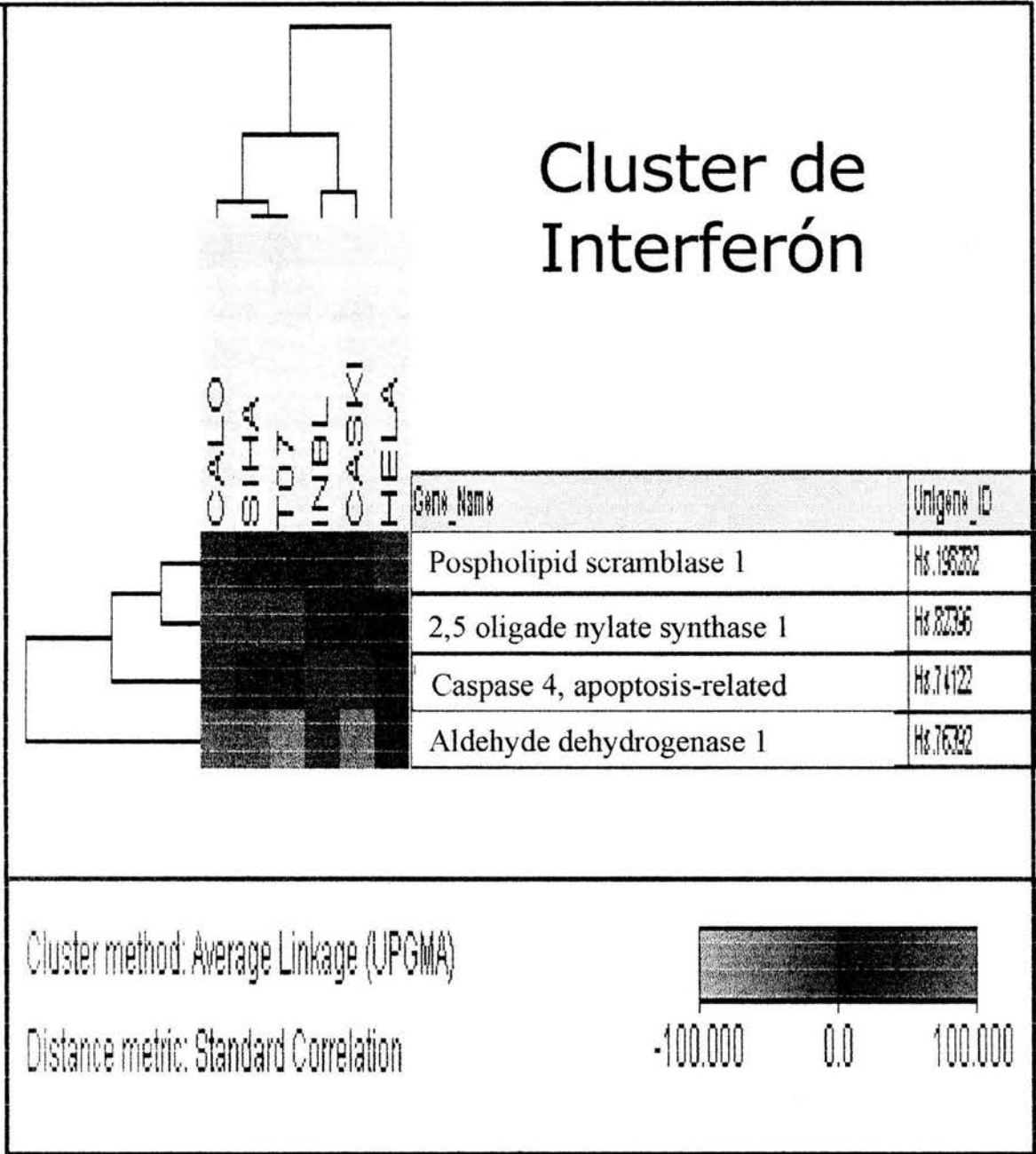


Figura 17.

Ejemplo de agrupamiento jerárquico de interferones o genes regulados por éstos.

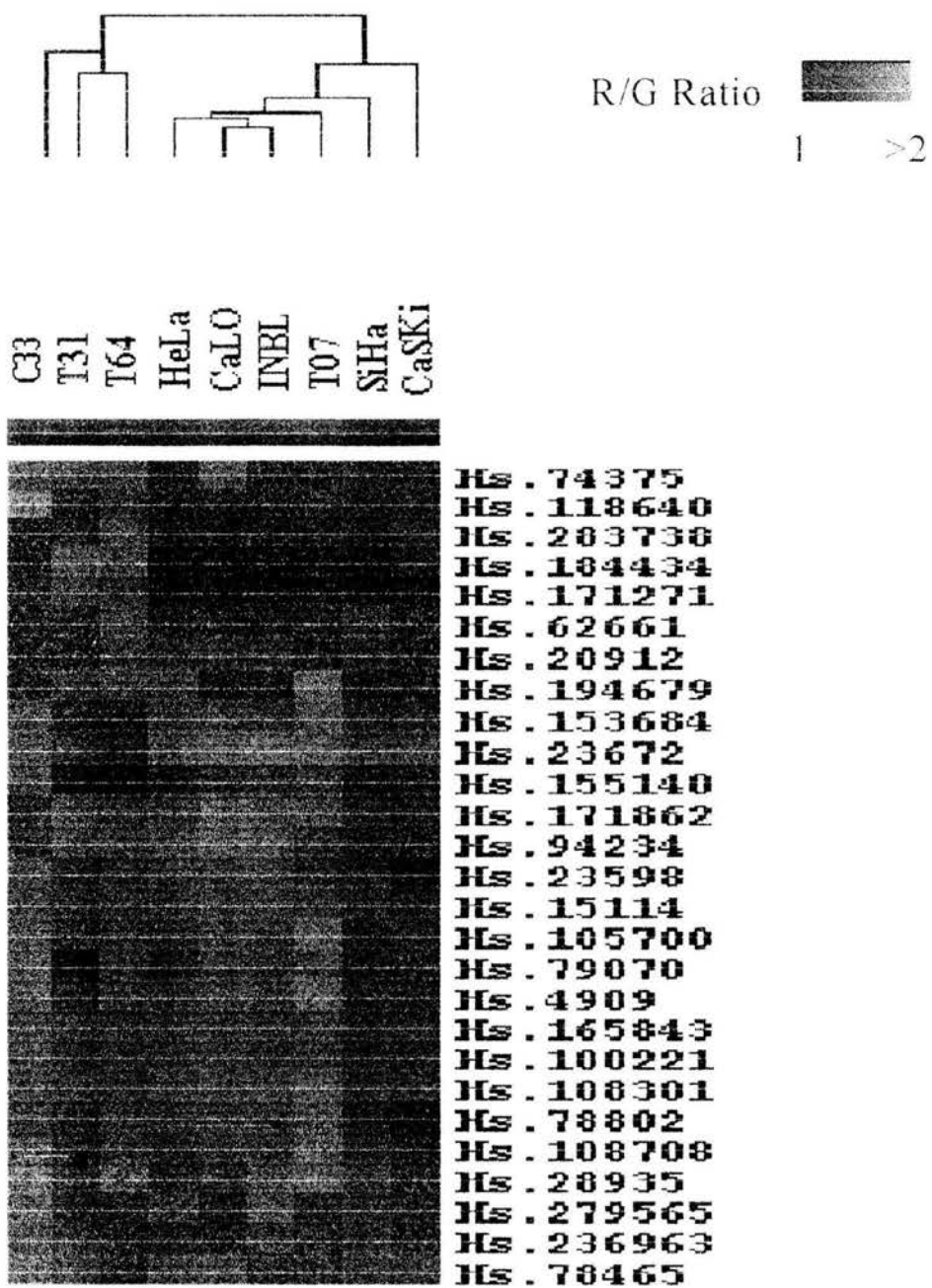


Figura 18.

Agrupamiento jerárquico de genes pertenecientes a la ruta de señalización de Wnt.

DISCUSION

Aunque la sobreexpresión de algunos componentes de esta ruta han sido previamente relacionados con CaCU, la ruta en si, no había sido del todo relacionada a este tipo de cáncer. Con el fin de validar la sobreexpresión se tomaron estratégicamente 4 genes que se encuentran distribuidos en diferentes niveles de la misma. El primer gen analizado es el homólogo de *Frizzled Drosophila 2* (FZD2), cuyo producto proteico se localiza en la superficie celular, un gen de expresión en citoplasma, también se estudió la cinasa sintetizadora de glucógeno 3-beta (GSK3B); y por otro lado dos productos génicos que se expresan en el núcleo: C-MYC y el receptor delta activado del proliferador de peroxisomas (PPARD).

Se realizó RT-PCR semicuantitativa para obtener los valores de expresión relativa de los cuatro genes. Las concentraciones relativas de los mARNs fueron expresadas como la razón de las lecturas densitométricas de los tres genes en relación al mRNA de β -actina (Figura 19). Todas las 9 muestras investigadas mostraron un incremento en la expresión de FZD, y en concordancia con esto, las intensidades de las bandas del gel de agarosa estaban incrementadas en un rango de 1.2 veces en todos los tumores con respecto al normal. La mayor diferencia presente entre los dos métodos fue en el gen de GSK3B, ya que en los arreglos todas las muestras estudiadas tuvieron una sobreexpresión de 3; mientras que en la RT-PCR los niveles de sobreexpresión se presentaron en un nivel de 8 veces más. PPPARD y c-MYC también presentaron una sobreexpresión de 1.3 y 3 veces, con respecto a los normales.

Con el fin de corroborar la sobreexpresión, se realizó hibridación *in situ* en arreglos de tejidos en diferentes estadios (normal, LIEBG (lesión intrepitelial de bajo grado), LIEAG (lesión intraepitelial de alto grado) y CaCU. Los resultados fueron positivos en el 90% de

los LIEAGs y el 100% de los CaCUs. FZD2 se expresó en 15 (100%), GSK3B en 12 (80%), c-MYC en 14 (93%) y PPARD en 13 (86%) de los 15 LIEAGs estudiados. Los cuatro genes tuvieron mayor expresión en los 15 casos de CaCU que en los tejidos normales. Los tejidos cervicales normales tuvieron una reacción muy leve (Figura 20).

7. Búsqueda de posibles marcadores moleculares de la enfermedad.

Con el fin de determinar posibles marcadores de la enfermedad como una entidad patológica; es decir, sin importar el tipo de VPH que esté infectando al tejido; se buscaron genes sobreexpresados en las muestras tumorales y líneas celulares y que estuvieran suprimidos o bien que presentaran una expresión basal en tejido normal y la línea celular C33. Esto se determinó utilizando el programa Gene Spring y el programa de *clustering* jerárquico. Por medio de este análisis bioinformático se detectaron 21 genes sobreexpresados y 20 suprimidos. De estos, se seleccionaron seis genes: claudina 1, laminina gama 2, importina 7, matriz metaloproteasa 10, IL6 y NDUFB6, los cuales se sometieron a RT-PCR semicuantitativa e inmunohistoquímica con el fin de validar la expresión (Anexos 4 y 5).

8. Búsqueda de posibles marcadores de invasión.

Para determinar genes involucrados en procesos de invasión, la base datos obtenida se cargó en el programa de internet Fatigo, y posteriormente también se cargó en la base de datos de GO, con el fin de obtener las anotaciones ontológicas de los genes. Después, se creó una base de datos nueva que contenía exclusivamente genes relacionados a invasión, se les realizó agrupamiento jerárquico para ver los diferentes niveles de expresión en las muestras. Se buscaron genes que estuvieran sobreexpresados en todas las muestras tumorales y líneas celulares y de estos se seleccionaron 3 genes que no estuvieran previamente relacionados con cáncer cérvico uterino; matriz metaloproteasa 11

y 12, Catepsina F. Se diseñaron iniciadores para realizar RT-PCR semicuantitativa en estos genes y para sintetizar sondas con el fin de llevar a cabo HIS en diferentes estadios del cáncer. También se realizó Inmunohistoquímica para la búsqueda de la proteína en los tejidos. Los resultados revelaron un incremento en la expresión conforme los diferentes estadios del cáncer avanzan (Anexo 6).

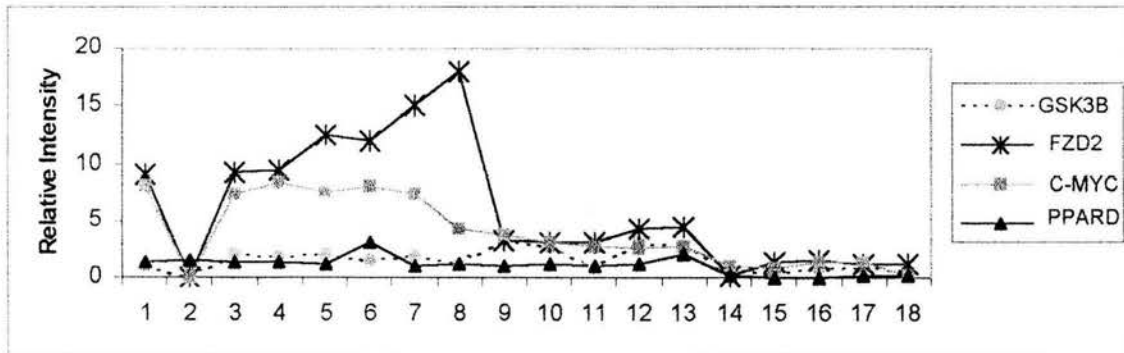
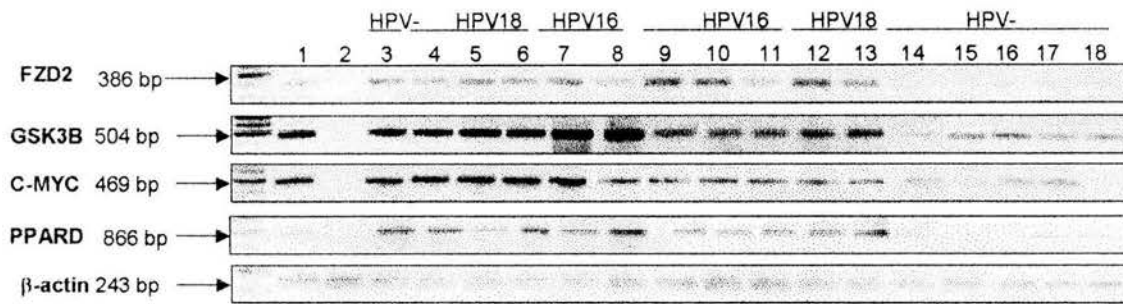


Figura 19.

Análisis densitométrico para obtener los niveles de expresión relativa de los genes GSK3B, FZD2, c-MYC, PPARD. Carril 1) Control positivo, 2) control negativo sin ARN, 3) línea celular C33, 4) HeLa, 5) CaLO, 6) INBL, 7) SiHa, 8) CaSKi, 9) muestra tumoral T16, 10) muestra tumoral T24, 11) muestra tumoral T27, 12) muestra tumoral A21, 13) muestra tumoral A23, 14) muestra de tejido normal N01, 15) muestra de tejido normal N05, 16) muestra de tejido normal N07, 17) muestra de tejido normal N14, 18) muestra de tejido normal N20. Se utilizó un marcador de peso molecular de 100 pb.

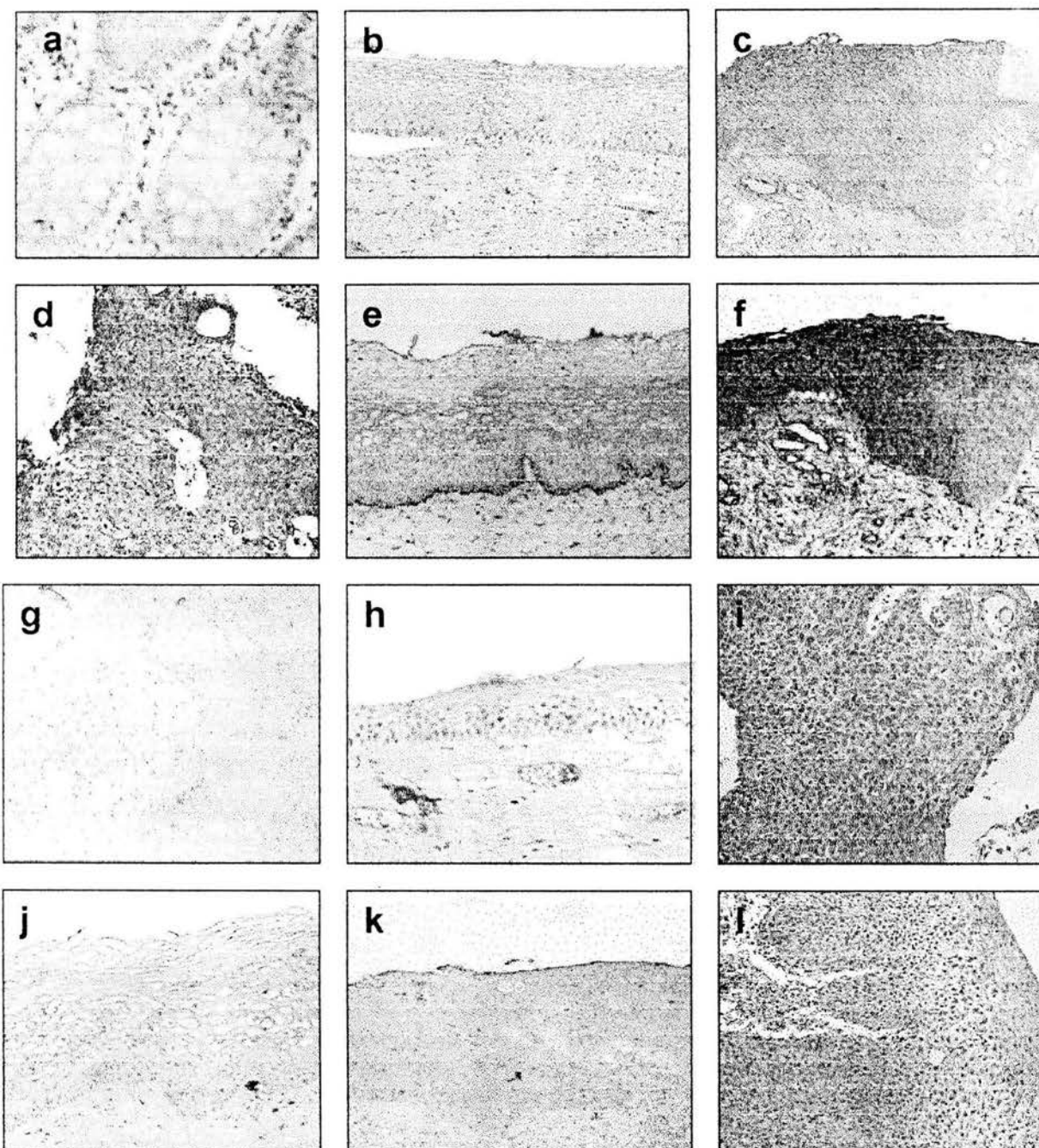


Figura 20.

Ejemplos de hibridación In Situ, a) colon como control positivo para FZD2, b) ejemplo tejido normal con tinción negativa para FZD2, c) ejemplo de LIEAG positivo para FZD2, d) ejemplo de CaCU para FZD2, e) ejemplo de tejido normal con tinción para GSK3B en 2/3 partes del epitelio, f) ejemplo de LIEAG con tinción positiva para GSK3B, g) ejemplo de CaCU hibridado con GSK3B, h) ejemplo de tejido normal con tinción negativa para c-myc, i) ejemplo de CaCU hibridado con c-myc, j) ejemplo de tejido normal con 2/3 partes de tinción para PPARD, k) ejemplo de LIEG con tinción positiva para PPARD, l) ejemplo de CaCU con hibridación positiva para PPARD.

DISCUSION

A nivel mundial, el CaCU está posicionado en el segundo lugar en incidencia de cánceres femeninos; y en México es la primera causa de mortalidad en mujeres. El uso de pruebas moleculares para detección o pronóstico de esta patología es prácticamente inexistente. Actualmente la detección del VPH se lleva a cabo en algunas muestras de pacientes con CaCU. Aunque la infección por el virus es necesaria para la presencia del cáncer, ésta no es suficiente, por lo que la influencia de los procesos celulares y moleculares en el desarrollo de esta neoplasia cobra gran importancia. Las tecnologías de alto rendimiento, como los arreglos, permiten estudiar los cambios de expresión de miles de genes en células tumorales y normales, de ahí que sea una excelente opción aplicar esta tecnología al estudio del CaCU (Anexos 1 y 2).

En este trabajo de tesis se definieron los patrones de expresión del carcinoma cérvico uterino con la metodología de arreglos de cDNA, así mismo se creó una base de datos a partir de nuestros resultados con el fin de condensar la información para análisis posteriores. Por otro lado, se llevaron a cabo diversos análisis bioinformáticos para explorar el impacto potencial de los cambios en la expresión de diversos procesos celulares. Finalmente, se validaron los genes candidatos por RT-PCR semicuantitativa, Hibridación *In Situ* e Inmunohistoquímica.

Debido a que no se conocía ninguna característica que pudiera relacionar al proceso tumoral con la sobreexpresión de la Importina 7, se realizó una búsqueda determinando la región promotora río arriba de este gen. Esto se hizo utilizando la base de datos Source, y posteriormente los posibles factores de transcripción que regulan a este gen fueron identificados mediante el programa TRANSFAC. Se obtuvieron aproximadamente 100 regiones consenso para GR (Receptor de glucocorticoides) en la secuencia promotora de Ipo7. Además de influenciar la expresión de Ipo7, GR también influencia la expresión

temprana de VPH mediante la interacción con el GRE presente en VPH 18 y los tres GREs presentes en VPH16 (ver anexo 4).

Por otro lado otro gen que se demostró sobreexpresado fue la MMP10, y actualmente este trabajo es el primero que asocia a esta proteasa al desarrollo del CaCU (Anexos 4 y 5).

Con el propósito de entender el comportamiento biológico del CaCU, se utilizó la base de datos Gene Ontology para clasificar a los genes en diferentes rutas metabólicas. Cada gen presente en el microarreglo fue asociado con su correspondiente número de Unigene, y tanto los datos de expresión como los números de Unigene se integraron en el website de Gene Ontology. Finalmente para cada uno de los genes se obtuvo el Código de Gene Ontology para procesos biológicos es decir, se obtuvieron las anotaciones funcionales de los genes que componen el arreglo. El programa de Cluster jerárquico y el programa de visualización (Tree View v1.50) fueron utilizados para graficar los análisis de expresión de cada una de las rutas. Se realizaron dos tipos de análisis con los datos obtenidos del GO. Para el primero se hicieron los grupos de: componentes epiteliales, proliferación, componentes estromales, interferones, genes relacionados en metabolismo de fármacos, componentes del ciclo celular y componentes tumorales encontrándose todos los grupos alterados con respecto a su contraparte normal, una explicación mas detallada sobre los hallazgos de este experimento se pueden apreciar en el Anexo 3. Posteriormente, con el fin de determinar las variaciones específicas de la expresión en las diferentes muestras, de cada uno de los agrupamientos de genes se realizó análisis de agrupamiento jerárquico con el fin de hacer más visuales los cambios en la expresión de los genes en las diferentes muestras (Anexo 4). Sin embargo, los resultados mostraron que las rutas y procesos metabólicos estudiados presentaban sobreexpresión y supresión dentro de sus componentes. La ruta metabólica de Wnt se encontró totalmente sobreexpresada probablemente debido al encendido de la misma por la presencia de los Wnt secretados que permiten que beta catenina se acumule en núcleo active diferentes factores de

transcripción así como antiapoptóticos, convirtiendo a esta ruta en un excelente punto de estudio en CaCU (ver anexo 7).

Por otro lado, con el fin de estudiar y empezar a caracterizar a las nuevas líneas celulares derivadas de CaCU, se realizaron estudios de CGH sobre metafases para poder determinar las pérdidas o ganancias de bandas citogenéticas en estas. En el anexo 8 se muestra una concordancia de pérdidas de DNA en el 6q26-q27 y ganancias en 1q31-q32 y 7p13-p14 que pudieran ser específicas para el tipo viral 18. Asimismo, la similitud encontrada entre los desbalances cromosómicos de las líneas celulares y las muestras tumorales, permiten inferir que las líneas celulares representan buenos modelos para el estudio del CaCU; lo cual ya había sido reportado previamente.

CONCLUSIONES

Una de las tendencias actuales en el estudio de la expresión global de las células, es la aplicación de la genómica funcional. En el presente trabajo, se determinaron los transcriptomas parciales tanto, de la célula tumoral como de la célula normal del cervix uterino humano; esto se llevó a cabo utilizando la tecnología de microarreglos de cDNA y empleando herramientas bioinformáticas aplicadas.

Una vez utilizadas estas herramientas, se obtuvieron varias bases de datos, las cuales tuvieron que ser unificadas con el fin de crear la base de datos principal (con más de un millón de valores de intensidad) la cual sirvió de base para los análisis posteriores. El poder llevar cabo la unificación de diversas bases de datos implicó el uso de computadoras de gran capacidad que permitieran utilizar aproximadamente 5 programas de análisis de manera simultánea. Asimismo, se utilizaron algunos análisis estadísticos para identificar la forma de distribución de los datos y así poder utilizar el análisis de normalización adecuado, el cual se realizó posteriormente en base a la distribución no paramétrica de los datos obtenidos en los experimentos. Basados en este tipo de análisis se lograron determinar diferentes patrones de expresión para cada una de las muestras analizadas. Una de las primeras estrategias que fueron llevadas a cabo, fue identificar aquellos genes de las células tumorales que no mostraban ninguna modificación en su patrón de expresión con respecto a su contraparte normal, luego; aquellos genes que presentaran incrementos en su nivel de expresión y posteriormente el grupo de genes que se suprimieran su expresión tanto en las células tumorales con respecto al normal, como en los normales con respecto a las células tumorales.

De esta manera se logró identificar a 450 genes sobreexpresados en las células tumorales y a más del doble de genes suprimidos. Por lo que podemos mencionar que en CaCU existe significativamente un evento de supresión en la regulación genética más que una activación.

Para demostrar lo antes mencionado se decidió validar por otros métodos (inmunohistoquímica e hibridación in situ RNA/RNA en microarreglos de tejido y RT-PCR semicuantitativa) 13 nuevos genes que pueden servir como marcadores de invasividad. Entre ellos se encuentran principalmente proteasas y oncogenes así como citocinas y factores de transcripción. Uno de los hallazgos obtenidos mas importantes en este trabajo fue el identificar la sobreexpresión de los componentes de la ruta metabólica de Wnt, la cual se encuentra activa de manera normal en procesos de desarrollo y se activa de manera anómala en algunos procesos tumorales. Esta ruta se ha visto sobreexpresada en algunos tumores, principalmente en cáncer de colon, demostrando que esta ruta tiene un papel fundaméntela en la tumorigenesis.

Por otro lado, se logró determinar el patrón de expresión de células cervicales infectadas con VPH 16 y VPH 18. Lo anterior, permite dilucidar el comportamiento global de esta enfermedad a nivel de expresión en las células infectadas con VPH de alto riesgo.

Para llegar a estos hallazgos fue necesario llevar a cabo estudios bioinformáticos especializados (Data mining) como el uso de: Gene Ontology Consortium, Gene Spring, Fatigo; TransFac, Source y algunos otros programas computacionales especializados en el análisis de datos obtenidos por microarreglos.

Con el presente estudio se llevo a implementar la tecnología de alto rendimiento de los arreglos de cDNA en el laboratorio de Genómica Oncológica, del Unidad de Investigación Médica en Enfermedades Oncológicas del Hospital de Oncología del CMN Siglo XXI.

Es importante mencionar que las herramientas utilizadas plantean el inicio de una nueva estrategia de investigación en nuestro país en el área de la Oncología Genómica, línea que ha permitido obtener recursos de institucionales y del CONACYT, para la formación de recursos humanos y la vinculación de la investigación clínica con la investigación básica.

BIBLIOGRAFIA

1. World Health Statistics, 1987-1990. *World Health Organization*.
2. Programa de prevención y control del cáncer cervico-uterino. *Secretaría de Salud, México* 1992.
3. Schiffman M, Bauer H, Hoover R, Epidemiologic evidence showing that human papillomavirus infection causes most cervical intraepithelial neoplasia. *J Natl Cancer Inst* 1993; **85**: 958-964.
4. Munoz N. Human papillomavirus and cancer: the epidemiological evidence. *J Clin Virol* 2000; (1-2):1-5.
5. Walboomers J, Jacobs M, Manos M, Bosch F, Kummer K, Shah K, Snijders P, Peto J, Meijer C, Muñoz N. Human Papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J Pathol* 1999; **189**: 12-19.
6. Bosch F, Lorincz A, Muñoz N, Meijer C, Shah K. The causal relation between human papillomavirus and cervical cancer. *J Clin Pathol* 2002; **55**:244-265.
7. de Villiers E. Human pathogenic papillomavirus types: and update. *Curr Top Microbiol Immunol* 1994;**186**:1-12.
8. Ha PK, Califano JA. The role of human papillomavirus in oral carcinogenesis. *Crit Rev Oral Biol Med* 2004;**15**:188-96. Review.
9. Syrjanen S, Puranen M. Human papillomavirus infections in children: the potential role of maternal transmission. *Crit Rev Oral Biol Med* 2000;**11**:259-74. Review.
10. zur Hausen. Papillomaviruses causing cancer: evasion from host-cell control in early events in carcinogenesis. *J Natl Cancer Inst* 2000; **92**:690-698.

11. Munoz N, Bosch F, de Sanjose, S, Herrero R, Castellsague X, Shah K, Snijders P, Meijer C. Epidemiologic Classification of Human Papillomavirus Types Associated with Cervical Cancer. *New Eng J Med* 2003; **348**:518-527.
12. Munoz N. Human papillomavirus and cancer: the epidemiological evidence. *J Clin Virol*; 2000, 1-2:1-5. Review.
13. <http://monographs.iarc.fr/monoeval/allmonos.html>
14. Clifford G, Smith J, Plummer M, Munoz N, Franceschi S. Human papillomavirus types in invasive cervical cancer worldwide: a meta-analysis. *Br J Cancer* 2003; **88**:63-73.
15. Thomas D, Ray R, Kuypers J, et al. Human papillomaviruses and cervical cancer in Bangkok. III. The role of husbands and commercial sex workers. *Am J Epidemiol*, 2001; **153**:740-748.
16. Castellsague X, Bosch FX, Munoz N. The male role in cervical cancer. *Salud Publica Mex.* 2003;45 Suppl 3:S345-53.
17. Matsukura T, Koi Sumiko, Sugase m. Both episomal and integrated form of Human Papillomavirus Type 16 are involved in Invasive Cervical cancers. *Virology* 1989; 172:63-72.
18. Stubenrauch F, Laimins L. Human papillomavirus life cycle: active and latent phases. *Seminars Cancer Biol* 1999; **9**:379-386.
19. Longworth MS, Laimins LA. Pathogenesis of human papillomaviruses in differentiating epithelia. *Microbiol Mol Biol Rev.* 2004;68:362-72. Review.
20. Milde-Langosch K, Riethdorf S. Role of cell-cycle regulatory proteins in gynecological cancer. *J Cell Physiol* 2003;196:224-44. Review.
21. Park T, Fujiwara H, Wright T. Molecular Biology of cervical Cancer and its precursors. *Cancer* 1995;76:1902-1912. Supplement

22. Munger K, Howley P. Human Papillomavirus immortalization and transformation functions. *Virus Res* 2002;89:213-228.
23. Morrison E. Natural History of Cervical Infection with Human Papillomavirus. *Clin Infection Diseases* 1994;18:172-80.
24. Werness B, Levine A, Howley P. Association of human Papillomavirus Types 16 and 18 proteins with p53. *Science* 1990;248:76-79.
25. Liu Z, Ghai J, Ostrow R, Faras A. The expression levels of the human papillomavirus Type 16 E7 correlate with its transforming potential. *Virology* 1995;207:260-270.
26. Barbosa MS. The oncogenic role of human papillomavirus proteins. *Crit Rev Oncog*. 1996;7:1-18. Review.
27. Sang BC, Barbosa MS. Single amino acid substitutions in "low-risk" human papillomavirus (VPH) type 6 E7 protein enhance features characteristic of the "high-risk" VPH E7 oncoproteins. *Proc Natl Acad Sci U S A*. 1992;89:8063-7.
28. Schwarz E, Freese UK, Gissmann L, Mayer W, Roggenbuck B, Stremlau A, zur Hausen H. Structure and transcription of human papillomavirus sequences in cervical carcinoma cells. *Nature*, 314:111-114, 1985.
29. Klaes R, Woerner SM, Ridder R, Wentzensen N, Duerst M, Schneider A, Lotz B, Melsheimer P, von Knebel Doeberitz M. Detection of high-risk cervical intraepithelial neoplasia and cervical cancer by amplification of transcripts derived from integrated papillomavirus oncogenes. *Cancer Res*. 59:6132-6136, 1999.
30. Human Genome Project.
http://www.ornl.gov/sci/techresources/Human_Genome/home.shtml
31. Gillespie D, Spiegelman S. A quantitative assay for DNA-RNA hybrids with DNA immobilised on a membrane. *J Mol Biol* 1965;12:829-842.

32. Lander S. Array of Hope. *Nature Genet* 1999;21:3-4
33. Southern E, Mir K, Schepinov M. Molecular interactions on microarrays *Nature Genet* 1999;21:5-9.
34. Duggan D, Bittner M, Chen Y, Meltzer P, Trent J. Expression profiling using cDNA microarrays. *Nature Genet* 1999;21:10-14.
35. Freeman WM, Robertson DJ, Vrana KE. Fundamentals of DNA hybridization arrays for gene expression analysis. *Biotechniques*. 2000;29:1042-6, 1048-55.
36. Bertucci F, Bernard K, Loriod B, Chang YC, Granjeaud S, Birnbaum D, Nguyen C, Peck K, Jordan BR. Sensitivity issues in DNA array-based expression measurements and performance of nylon microarrays for small samples. *Hum Mol Genet*. 1999;8:1715-22.
37. Hornberg JJ, de Haas RR, Dekker H, Lankelma J. Analysis of multiple gene expression array experiments after repetitive hybridizations on nylon membranes. *Biotechniques*. 2002;33:108, 110, 112-3.
38. Wu TD. Analysing gene expression data from DNA microarrays to identify candidate genes. *J Pathol* 2001;195:53-65.
39. Eickoff B, Korn B, Schick M, Postka A, Van der Bosch J. Normalization of array hybridation experiments in differential gene expression analysis. *Nucleic Acids Res* 1999;27;33.
40. Zien A, Aigner T, Zimmer R, Lengauer T. centralization: a new method for the normalization of gene expression data. *Bioinformatics* 2001;17:323-31.
41. Ramdas L. Sources of nonlinearity in cDNA microarray expression measurements. *Genome Biol Research* 2001;2:47.1-47.7

42. Tseng GC, Oh M, Rohlin L, Liao JC, Wong WH. Issues in cDNA microarray analysis: quality filtering, channel normalization, models of variations and assessments of gene effects. *Nucleic Acid Res* 2002;29:2549-57.
43. Livesey FJ, Furukawa T, Steffen MA, Church GM, Cepko CL. Microarray analysis of the transcriptional network. *Curr Biol* 2000;10:301-310.
44. Jelinski SA, Samson LD. Global response of *Saccharomyces cerevisiae* to an alkylating agent. *Proc Natl Acad Sci* 1999;96:1486-1491.
45. Chan JJ. Profilin expression patterns and isolating differentially expressed genes by cDNA microarray system with colorimetric detection. *Genomics* 1998;51:313-24.
46. Ishii M. Direct comparison of GeneChip and SAGE on the quantitative accuracy in transcript profiling analysis. *Genomics* 2000;68:136-46.
47. Baugh LR, Hill AA, Brown E, Hunter CP. Quantitative analysis of mRNA amplification by *in vitro* transcription. *Nucleic Acid Res* 2001;29:E29.
48. Schadt EE, Li C, Su C, Wong WH. Analyzing high density oligonucleotide gene expression array data. *J Cell Biochem*. 2000;80:192-202.
49. Warford A, Howat W, McCafferty J. Expression profiling by high-throughput immunohistochemistry. *J Immunol Methods*. 2004;290:81-92.
50. Shergill IS, Shergill NK, Arya M, Patel HR. Tissue microarrays: a current medical research tool. *Curr Med Res Opin*. 2004;20:707-12.
51. Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 1995;270:467-70.
52. DeRisi J, Penland L, Brown PO, Bittner ML, Meltzer PS, Ray M, Chen Y, Su YA, Trent JM. Use of a cDNA microarray to analyse gene expression patterns in human cancer. *Nat Genet* 1996;14:457-60.

53. Lashkari DA, DeRisi JL, McCusker JH, Namath AF, Gentile C, Hwang SY, Brown PO, Davis RW. Yeast microarrays for genome wide parallel genetic and gene expression analysis. *Proc Natl Acad Sci U S A*. 1997;94:13057-62.
54. Welford SM, Gregg J, Chen E, Garrison D, Sorensen PH, Denny CT, Nelson SF. Detection of differentially expressed genes in primary tumor tissues using representational differences analysis coupled to microarray hybridization. *Nucleic Acids Res*. 1998;26:3059-65.
55. Chansheob S, Wui Z, Chang H, Je.Ho L. Profiling of differentially expressed genes in human primary cervical cancer by complementary DNA expression array. *Clin Can Res* 1998;4:3045-50.
56. Nees M, Geoghegan JM, Hyman T, Frank S, Miller L, Woodworth CD. Papillomavirus type 16 oncogenes downregulate expression of interferon-responsive genes and upregulate proliferation-associated and NF-kappaB-responsive genes in cervical keratinocytes. *J Virol*. 2001;75:4283-96.
57. Cheng Q, Lau WM, Tay SK, Chew SH, Ho TH, Hui KM. Identification and characterization of genes involved in the carcinogenesis of human squamous cell cervical carcinoma. *Int J Cancer*. 2002;98:419-26.
58. Berger AJ, Baege A, Guillemette T, Deeds J, Meyer R, Disbrow G, Schlegel R, Schlegel R. Insulin-like growth factor-binding protein 3 expression increases during immortalization of cervical keratinocytes by human papillomavirus type 16 E6 and E7 proteins. *Am J Pathol*. 2002;161:603-10.
59. Alazawi W, Pett M, Arch B, Scott L, Freeman T, Stanley MA, Coleman N. Changes in cervical keratinocyte gene expression associated with integration of human papillomavirus 16. *Cancer Res*. 2002;62:6959-65.

60. Wong YF, Selvanayagam ZE, Wei N, Porter J, Vittal R, Hu R, Lin Y, Liao J, Shih JW, Cheung TH, Lo KW, Yim SF, Yip SK, Ngong DT, Siu N, Chan LK, Chan CS, Kong T, Kutlina E, McKinnon RD, Denhardt DT, Chin KV, Chung TK. Expression genomics of cervical cancer: molecular classification and prediction of radiotherapy response by DNA microarray. *Clin Cancer Res* 2003;9:5486-92.
61. Caceres-Cortes JR, Alvarado-Moreno JA, Waga K, Rangel-Corona R, Monroy-Garcia A, Rocha-Zavaleta L, Urdiales-Ramos J, Weiss-Steider B, Haman A, Hugo P, Brousseau R, Hoang T. Implication of tyrosine kinase receptor and steel factor in cell density-dependent growth in cervical cancers and leukemias. *Cancer Res* 2001; 61:6281-9.
62. Monroy-Garcia A, Weiss-Steider B, Hernandez-Montes J, Ortiz-Navarrete VF, Banuelos-Panuco A, Acosta-Araujo A, Diaz-Quinonez A, Lopez-Graniel CM, Herbert G, Granados J, de Leo C, Silva-Lopez RM, Mora-Garcia ML. Identification of two homologous antigenic peptides derived from L1 VPH-16 and 18 proteins specific for the HLA-B*3901 allele. *Arch Virol.* 2002;147:1933-42.
63. Ballman K. Fast Loess for Normalizing Microarray Data. *Statistical Methods for Gene Expression: Microarrays and Proteomics*. Institute for Mathematics and its applications. 2004.

MICROARREGLOS EN ONCOLOGÍA

Mauricio Salcedo, Guelaguetza Vázquez, Alfredo Hidalgo, Carlos Pérez, Patricia Piña, Karla Santillán, Brenda Alatorre, Hugo Arreola, Ricardo López, Carlos Montoya, Bernardo Navarro, Tania Cerón

RESUMEN

Con el objetivo de examinar y descifrar el transcriptoma completo de las células normales y malignas de manera sistemática y con la disponibilidad de la secuencia del genoma humano, ahora es posible gracias al uso de tecnología de alto rendimiento.

En la detección y tamiz de mutaciones de polimorfismo o mutaciones, análisis de alta resolución de aberraciones cromosómicas y los perfiles de expresión génica, el uso de los microarrays o microarreglos, está empezando a proveer una visión más comprensible de las alteraciones genéticas presentes en las células cancerosas. En estos momentos podemos interrogar e interpretar algunos de estos datos prometiendo modificar la manera de como el cáncer debe ser detectado y tratado en el futuro, gracias a la combinación de la bioinformática y las herramientas de análisis de datos. En la presente revisión se muestra una visión muy general del estado de arte referentes a las diversas plataformas tecnológicas de los microarrays aplicados a la Investigación de la Genómica del cáncer.

Palabras Clave: Genoma, microarreglo, genómica, diagnóstico, cáncer.

Microarrays in oncology

ABSTRACT

With the availability of the human genome sequence, the goal of examining and deciphering the entire transcriptome of normal and malignant cells in a systematic way is now possible by using high throughput technology.

The use of microarrays is starting to provide a comprehensive view of genetic alterations in cancer cells by genome-wide polymorphisms screens, high-resolution analysis of chromosomal aberrations and expression profiling. In combination with the development of the bioinformatic and data analysis tools necessary to interrogate and interpret these data. This approach promises to revolutionize how we think about cancer and how we are likely to detect and treat cancer in the future.

The present review shows an overview of the state of the art of different microarrays technologies applied to Cancer Genomics Research.

Key Words: Genome, microarray, genomics, diagnosis, cancer.

ARTÍCULO RECIBIDO EN MAYO DEL 2003 Y ACEPTADO EN JUNIO DEL 2003

El cáncer es uno de los más importantes problemas de salud pública a nivel mundial¹. En los Estados Unidos se estima que para este año se diagnosticarán cerca de un millón y medio de nuevos casos generando un grave problema de salud^{2,3}. En estos momentos, con la gran cantidad de metodologías desarrolladas para la investigación, nos encontramos en un área con enormes oportunidades para descubrir y descifrar aspectos relacionados con las bases moleculares de las neoplasias. El concepto

“descubriendo la ciencia” puede ser ilustrado por el proyecto del genoma humano, es decir, la carrera científica denominada Genómica. Aunque conocemos casi en su totalidad el genoma de la célula humana, esta carrera prosigue y nos conduce a la actual era post-Genómica^{4,5}. Esto involucra la identificación y estudio integrador de los componentes de un sistema sin anticipar la formulación de hipótesis de cómo funcionan estos componentes. Así, un gran flujo de secuencias o “tags” humanas expresadas (ESTs o expression sequence tags) proveen un punto de partida para dilucidar la función de los miles de genes contenidos en el genoma. Para entender de forma detallada el genoma humano se requerirá de la implementación de métodos

Laboratorio de Oncología Genómica, Unidad de Investigación Médica en Enfermedades Oncológicas, Hospital de Oncología, CAMN SXII-IMSS. E-mail: maosa119@yahoo.com

sofisticados de alto rendimiento (high-throughput), así como, para el análisis de su expresión permitiendo el descubrimiento de nuevos genes.

El concepto de "Genómica" está conformado por la genómica estructural, funcional, individual y comparativa. El campo de la genómica funcional incluye la determinación del número de genes, transcritos o proteínas (DNA, RNA mensajero y proteínas) presentes y expresados en una célula-tejido específico, en un momento fisiológico determinado: Transcriptoma y Proteoma, respectivamente.

A la fecha existen dos tipos de sistemas de análisis masivos de expresión: sistemas cerrados y abiertos. El sistema cerrado consiste en determinar la expresión de los genes o secuencias impresas en el array, mientras que el sistema abierto consiste en determinar la expresión de todos los genes que se presentan en una célula, ideal para el transcriptoma. En otras palabras, el sistema cerrado consiste en realizar el estudio por medio de microarreglos (también se les conoce como microarrays, biochips, microhilas, micromatrices), los cuales contienen genes o secuencias conocidas, como los EST o bien genes anotados; mientras que el sistema abierto queda ejemplificado por el uso de la tecnología de análisis de expresión génica en serie o SAGE, tecnología que analiza la totalidad de los genes expresados, permitiendo identificar nuevas secuencias o genes, así como, los niveles de expresión que presentan en un momento determinado de manera cuantitativa⁶.

Un microarreglo o microarray de ácidos nucleicos es una plataforma sólida que puede ser de nylon, vidrio o plástico donde se encuentran inmovilizados generalmente en un área de 1 cm², un grupo limitado y finito de genes o fragmentos de genes (en promedio 50 bases de longitud) que pueden ir de cientos hasta miles, es decir, desde 100 elementos/cm² hasta 25,000 elementos/cm² generando así los microarrays de baja o alta densidad⁷⁻⁹. Estas secuencias deben estar previamente caracterizadas y presentan el acomodo espacial de una matriz matemática. Sobre estas plataformas con las secuencias de DNA, cDNA u oligonucleótidos, se hibrida el DNA o cDNA provenientes de dos estados biológicos diferentes por ejemplo sano vs patológico, previamente marcados con diferentes fluorocromos (fluoresceína o verde, rodamina o rojo; otros fluorocromos usados son los Cy3 Cy5), de esta manera cuando el "arreglo" es leído, se cuantificará la intensidad de la fluorescencia de cada fluorocromo en un escáner especial y se apreciarán diferentes colores: verde cuando hay una sobre-expresión, rojo para supresión y amarillo cuando no hay cambios en la expresión entre la muestra normal y la muestra problema. En este contexto, estas plataformas tecnológicas han sido utilizadas para clasificar diversas patologías basándose en sus patrones de expresión, más que por características morfológicas como en el caso del linfoma de células B¹⁰. Asimismo, estos microarrays están permitiendo monitorear simultáneamente los patrones de expresión de miles de genes durante diversos procesos celulares, como por ejemplo durante la diferenciación. Para ello, los nuevos equipos analíticos

basados en la ultraminiaturización y paralelismo, convergen en la aplicación de nuevos sistemas de fabricación, los cuales provienen de las tecnologías de la informática y la electrónica y se aplican ahora en el diseño de biochips (chips de material biológico de alta densidad de integración), para poder realizar distintos tipos de estudios repetitivos con muestras biológicas simples y por procedimientos analíticos completos sobre muestras individuales (lab-on-a-chip)¹¹. El potencial de estos sistemas, permite lograr en un tiempo corto gran cantidad de información, que debe ser analizada con poderosas técnicas estadísticas y bioinformáticas para extraer conocimiento de utilidad. Estas técnicas ofrecen la capacidad de comparar y relacionar la información genética con una finalidad deductiva, ofreciendo respuestas que van implícitas y que no parecen obvias a la vista de los resultados de los experimentos. Así, todos los sistemas que permiten la integración y análisis de datos provenientes del uso de biochips o microarrays, están justificadas por la necesidad de manejar la información masiva de alto rendimiento con enfoques integrados (genómica funcional, proteómica, expresión génica, genómica comparada, etc.)⁹⁻¹¹.

Esta poderosa tecnología tiene su inicio hace poco más de una década, culminando en la plataforma del conocido GeneChip por la compañía de Affymax (hoy Affymetrix)^{12,13}. El nombre general que recibe esta tecnología es el de biochip. Sin embargo, dependiendo de la macromolécula usada para construir el arreglo, encontramos el DNA chip, chip genómico, RNA chip, los cDNA microarrays, microarrays basados en oligonucleótidos, o los chips de proteínas, entre otros (Figura 1). Actualmente, se cuenta con una gran gama de biochips disponibles comercialmente para casos específicos, por ejemplo los Oncochip, Genechip, linfocchip, nanochip, cardiochips, hepatochips, estafilochips, etc.^{14,15} Sus aplicaciones se han ido incrementando día con día, empleándose entre otras cosas para la:

- i) monitorización de expresión génica,
- ii) detección de mutaciones y polimorfismos,
- iii) secuenciación,
- iv) diagnóstico clínico-detección de microorganismos,
- v) tamizaje en toxicología,
- vi) seguimiento de terapias¹¹.

Uno de los campos en los cuales el uso de los microarrays está siendo esencial y de gran impacto es el área de la oncología.

USO DE LOS MICROARREGLOS EN UNA NUEVA CLASIFICACIÓN DE LOS TUMORES

La transformación neoplásica es el resultado de la acumulación de alteraciones en genes esenciales, relacionados con la proliferación y diferenciación celulares (entre otros la combinación de oncogenes, genes supresores de tumor, falla en los genes de reparación, mecanismos de metilación, apoptosis, entre otros).

Por otro lado, la clasificación de tumores humanos se encuentra variando constantemente, lo que está permitiendo tomar

diferentes opciones de tratamiento. Así mismo, en forma reciente se han publicado gran cantidad de artículos en los cuales se aplican los microarrays en la línea de la onco-hematología¹⁶.

En este contexto, se han aplicado los microarrays para clasificar en cuanto a su expresión génica a las leucemias linfoides y mieloide (leucemias agudas) a partir de muestras de médula ósea. Este estudio demostró con gran certeza que el diagnóstico a nivel molecular coincide con el diagnóstico clínico. Además, cada tipo de leucemia mostró un patrón de expresión génica diferencial. Este estudio sienta las bases para el desarrollo del diagnóstico precoz de este tipo de enfermedades. Con respecto a otra afección onco-hematológica, los linfomas, se ha demostrado en un reporte multi-institucional que es posible realizar nuevas taxonomías tumorales dependiendo de los patrones de expresión génica. Así, fue posible determinar distintos histotipos de linfomas difusos de células grandes tipo B. Este estudio surge de la observación clínica de que, aún cuando histológicamente los tumores son iguales y aunque todos los pacientes reciben el mismo tratamiento, un grupo de ellos fallece (60%) y el otro sobrevive a la enfermedad. Los resultados del estudio mostraron que estas diferencias ante la respuesta al tratamiento tienen su base en los patrones de expresión génica en cada grupo de pacientes. Una expresión génica característica de centros germinales de células B se asocia con un pronóstico más favorable, mientras que el patrón de expresión de células maduras activadas presentes en circulación, se asocia con un mal pronóstico. Así, este trabajo sienta las bases para nuevos criterios pronósticos más certeros que los tradicionales, además de proponer nuevas clasificaciones moleculares del tumor y abre la posibilidad de identificar subtipos de cáncer no detectables con métodos tradicionales^{10,17-21}.

Otros grupos han usado el análisis de los patrones de expresión génica para clasificar a nivel molecular tumores de ovario^{22,23} mama²⁴⁻²⁶, melanomas cutáneos²⁷, próstata²⁸ y adenocarcinomas de pulmón^{29,30}. Estos estudios demuestran la capacidad de combinar el análisis molecular de los perfiles de expresión génica a gran escala, con los métodos clásicos de morfología y clínica para la etapificación de las neoplasias, todo esto para obtener mejores diagnósticos clínicos y para poder predecir la sobrevida de los pacientes³¹.

MARCADORES MOLECULARES ESPECÍFICOS DE TUMOR
Como se mencionó anteriormente, el uso de los microarrays también ha tenido impacto en el estudio de tumores sólidos. Esto está permitiendo obtener los "retratos moleculares de las neoplasias". Un gran esfuerzo se comienza a realizar de manera cotidiana para ayudar a descifrar la información obtenida de muestras tumorales. Así, un gran número de grupos de investigación a nivel internacional, está enfocando sus estudios para obtener grupos de genes que marquen una expresión diferencial entre tejidos sano y enfermo, incluyendo líneas celulares provenientes de neoplasias. Estos estudios incluyen cáncer de cabeza y cuello^{32,33}, cáncer endometrial³⁴, colo-rectal^{35,36}, etc. Sin embargo, a la fecha se cuenta con muy pocos datos para poder dilucidar los mecanismos moleculares de esas neoplasias. Esta situación puede deberse al alto costo que tiene esta tecnología, pero a medida que diversos grupos comiencen a trabajar con estas plataformas, los costos tendrán que disminuir debido a una mayor demanda.

Actualmente, nuestro grupo se encuentra analizando el retrato molecular del cáncer cérvico-uterino (CaCU) dentro de diferentes contextos: Genómica estructural, funcional e individual, así

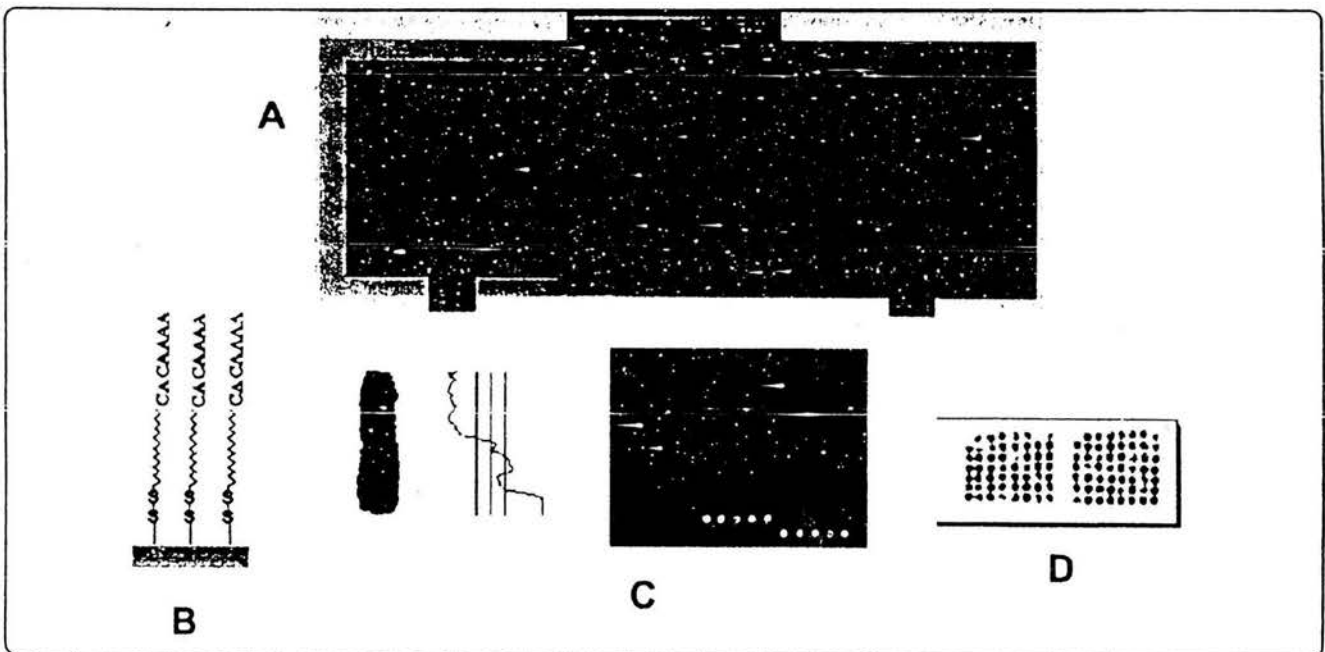


Figura 1. Diversidad de plataformas de microarreglos. Arrays de A, expresión, B, de DNA, C, hibridación genómica comparativa y D, tejidos.

como, la Genómica funcional del virus de papiloma humano de alto riesgo (HPV), tratando de descifrar el "queratinosoma" del cérvix uterino humano, utilizando diversas metodologías de análisis genético global y alto rendimiento. Como se sabe, el factor etiológico más importante en el CaCu es el HPV de alto riesgo tipos 16 y 18, siendo este último el más agresivo clínicamente. Aplicando los microarreglos de cDNA de la plataforma ULTRA Array, nuestro grupo ha logrado identificar los genes alterados transcripcionalmente en esta neoplasia, asimismo, se determinó un patrón de expresión diferencial entre los tipos 16 y 18. Estos resultados son soportados por estudios de hibridación genómica comparativa sobre metafases (HIGM), con lo cual se ha logrado detectar desbalances cromosómicos específicos que coinciden con regiones que se activan o reprimen transcripcionalmente. De esta manera estamos definiendo las firmas genéticas que generan cada tipo de HPV de alto riesgo^{37,38}.

Por otro lado, los biochips han confirmado que la expresión génica en líneas celulares derivadas de cánceres humanos, reflejan en gran medida a la expresión génica de los tejidos de donde provienen, también mostraron que las metástasis surgen a partir de células que modifican su patrón de expresión de manera sutil, con relación a su tumor primario. En su conjunto, estos hallazgos han permitido obtener una visión más clara acerca de las alteraciones genéticas que son necesarias para que un tumor se forme y desarrolle, permitiendo la caracterización genética global de una determinada neoplasia y señalando genes o mecanismos celulares susceptibles de ser utilizados como marcadores pronósticos o diagnósticos, o como blancos adecuados para nuevas terapias.

Un experimento común de microarreglos valora la expresión de entre 1,500 a 10,000 genes en una sola hibridación, generando gran cantidad de resultados que deben ser analizados de manera lógica para poder extraer la mayor información posible de los mismos. Como se ha mencionado anteriormente, estos resultados deben ser interpretados mediante el uso de herramientas bioinformáticas³⁹⁻⁴². Para ilustrar el valor de tales análisis de datos, estas herramientas (por ejemplo mapas autorganizables o SOM), se han aplicado en el estudio de diferenciación hematopoyética utilizando líneas celulares. En ese estudio fue indispensable organizar los genes basándose en su relevancia biológica en grupos o "clusters" usando los paquetes computacionales GENECLUSTER, J-Express y Tree-View disponible en la red. De esta manera se sugirieron nuevas hipótesis acerca de la diferenciación hematopoyética, por ejemplo, enfocándose a ciertos genes y rutas metabólicas involucradas en diferenciación y que puedan ser usadas en las terapias de la leucemia promielocítica aguda. El uso adecuado de las herramientas bioinformáticas, ha permitido caracterizar un nuevo tipo de leucemia basándose en su patrón de expresión génica. En este estudio, se analizó el patrón de expresión de leucemias linfoblásticas agudas que presentaban una translocación que involucra al gen *MILL*, demostrándose que dicho grupo de leucemias constituyen una entidad patológica diferente a las ya descritas⁴⁰. El campo de la bioinformática es uno de los más

dinámicos de las áreas biomédicas, constantemente se desarrollan nuevos algoritmos matemáticos que permiten un análisis más profundo de los datos generados por los microarrays^{39,43}. Un ejemplo de esto lo constituye el trabajo en el cual se definen grupos de genes que permiten predecir el curso clínico de los linfomas difusos de células grandes B. A diferencia del artículo previamente mencionado sobre este tipo de tumor, en éste se utilizó un algoritmo derivado de modelos de inteligencia artificial, denominado "supervised vector machine" o SVM. Este algoritmo toma en cuenta además de los patrones de expresión génica, otras variables tales como, el índice internacional de pronóstico y el tipo de tratamiento. Este tipo de método de predicción con aprendizaje supervisado, permitió identificar patrones de expresión relacionados con la cura de la enfermedad y patrones de expresión presentes en los casos fatales o refractarios a tratamiento⁴⁴⁻⁴⁷. Hablar de los microarreglos o microarrays no es sinónimo de estudios de expresión de RNA, es hablar de una amplia tecnología que puede ser aplicable a diferentes situaciones biológicas, es decir, a la presencia o ausencia de mutaciones o polimorfismos, a expresión proteica, etc.

MICROCHIPS DE POLIMORFISMOS ÚNICOS DE DNA O SNPs: PREDISPOSICIÓN A LA ENFERMEDAD

Recientes evidencias epidemiológicas sugieren que la predisposición genética puede jugar un relevante papel en el desarrollo del cáncer. Por ejemplo, se ha demostrado que en más del 42% del cáncer de próstata tiene influencia genética⁴⁸. En este sentido, más allá de los genes que pudieran estar relacionados con esta neoplasia, la atención se está enfocando a la identificación de polimorfismos de nucleótidos simples o SNPs asociados a la enfermedad y que en este caso, representan el tipo más común de alteración molecular o variación en la secuencia en el DNA del genoma humano. Se sabe que aproximadamente uno de cada mil pares de bases del genoma es polimórfico, y se estima que para finales de este año se tengan descritos más de 2 millones de SNPs, los cuales estarán disponibles en bases de datos públicas compartiendo las bases de datos de SNP y HapMap⁴⁹⁻⁵⁰. Parte importante de los SNPs es que la mayoría de ellos son del tipo silencioso y que no implican cambios funcionales con consecuencias anormales, no tanto aquellos que si caen en regiones codificadoras los cuales cambian la composición de los aminoácidos y las propiedades funcionales de las proteínas. Mayor atención tendrán aquellos que se encuentren en otras regiones como promotores, enhancers, silenciadores, etc.

La identificación de SNPs asociados a fenotipos específicos requerirá de tecnología de gran escala para poder monitorizar una gran cantidad de muestras de DNA y realizar las asociaciones multiparamétricas clínico demográficas mediante algoritmos bien definidos, es decir el parte aguas de la Epidemiología Genómica⁵¹⁻⁵³.

A la fecha ya contamos en el mercado con varios microarrays y tecnología de biochips asociados a SNPs, tal como la

tecnología de primera generación como el Affymetrix SNP array el cual contiene 1,500 SNPs o algunos otros como Genomatrix, Protogene, y el Motorola Biochip Systems. Esto abre la oportunidad de realizar proyectos de gran escala en microarrays en formatos de media densidad, robotizados, multiplexados, rápidos (en horas) y a bajo costo en modelos establecidos como HIV, p53, citocromo P450 y mutaciones en mitocondria humana⁵².

MICROARREGLOS PARA DESBALANCES CROMOSÓMICOS "CGH ARRAY"

Como se ha mencionado, los microarreglos pueden ser de diferentes tipos según la fuente de su impresión DNA o cDNA.

En el caso de los microarreglos de DNA, éstos se han utilizado más recientemente en el estudio de la ganancia y pérdida de regiones cromosómicas en las neoplasias. Esta plataforma deriva de la conocida HGCm. Con las recientes técnicas de alto rendimiento en menos de una década se cuenta ya con una nueva versión de HGC "corregida y aumentada", la HGC en microarreglo o CGH Array. Ahora no sólo contamos con una tecnología que nos permite visualizar todas las alteraciones cromosómicas en un solo ensayo (en 23 estructuras cromosómicas) sino que ahora en el nuevo formato de microchips tenemos en cada spot o clona impresa, fragmentos de DNA de 1,500 pares de bases hasta fragmentos de 2 Mega pares de bases (calcule número de spots en el chip = longitud DNA humano/1,500 o 2Mpb; UCSF, Genomic Array)^{41-45,54,55}.

En el mercado se cuenta actualmente con algunos microarreglos de DNA humano con aprox. 800 "spots" (Genesensorchips, VYSIS). Con la plataforma de CGH Array, de GenoSensor-VYSIS, hemos confirmado nuestros resultados de HGCm y además contamos con más de 20 regiones cromosómicas (amplicones) que potencialmente pueden ser marcadores de invasividad y metástasis (enviado a su publicación).

MICROCHIPS DE ALTO RENDIMIENTO Y SU APLICACIÓN EN PROTEÓMICA "PROTEIN CHIPS"

Las proteínas son las principales moléculas efectoras en las células. Los niveles del RNA mensajero proveen la información sobre la tasa de traducción de los genes a las proteínas o sobre el grado de modificación de mecanismos post-traduccionales, así como, en la complejidad de las numerosas interacciones que las proteínas tiene unas con otras. En este contexto, el entendimiento de los niveles, actividades funcionales y las redes neurales de interacción vienen a ser los retos primarios de la investigación en la era post-genómica.

La clásica técnica de visualización de la proteómica es la electroforesis de doble dimensión, seguida por la purificación del péptido expresado diferencialmente y su cuantificación. Sin embargo, métodos alternos se están desarrollando y así cientos o miles de anticuerpos son microarreglados sobre laminillas y de esta manera determinar en muestras biológicas la presencia de determinados epítomos marcadores. Como en los casos anteriores,

en el mercado también encontramos "Proteinchips" pero en baja oferta y demanda. En otras palabras, la tecnología de los chips de proteínas está aún en fase temprana de desarrollo comparada con la de DNA o RNA, sin embargo, el desarrollo técnico en este campo de investigación será más rápido⁵²⁻⁵⁶.

INTEGRACIÓN DE LA GENÓMICA FUNCIONAL CON HALLAZGOS DE MORFOLOGÍA DE TEJIDOS "TISSUE CHIP"
Es obvio que la información obtenida por la tecnología de los biochips debe ser validada para que pueda tener una interpretación mayor. Así, se ha desarrollado la tecnología de los microarreglos de tejidos (T μ A), la cual ha permitido detectar *in situ* alteraciones en un gran grupo de especímenes humanos mediante las técnicas de inmunohistoquímica o hibridación *in situ* fluorescente. El T μ A provee así, una plataforma tecnológica poderosa para la investigación de la genómica traduccional y funcional^{52,56-58}. Brevemente, los T μ A son portaobjetos clásicos que contienen arreglados hasta más de 1,000 tejidos dispuestos en un arreglo especial y que tiene relaciones sobre información clínica, patológica y demográfica. Estos permiten trabajar desde cientos hasta miles de muestras a un tiempo ("una corrida") de forma masiva y en paralelo. En otras palabras, esta plataforma es ideal para estudiar la significancia clínica de los cientos de genes, transcritos o proteínas obtenidas mediante estudios en microarrays como potenciales marcadores. Dado el alto costo del equipo microarreglador automatizado que se encuentra en el mercado y buscando alternativas a esta tecnología, nuestro grupo ha diseñado recientemente un formato de microarreglos de tejido de baja densidad, manual, económico y capaz de usarse en cualquier laboratorio de investigación biomédica⁵⁹. En este formato es posible aplicar técnicas clásicas de histología, hibridación *in situ* e inmunohistoquímica.

Otras áreas de impacto potenciales de estas plataformas tecnológicas además de la oncología serán: el desarrollo de la medicina preventiva desde una perspectiva genómica (determinar la susceptibilidad hacia determinadas patologías) antes de que aparezcan los primeros síntomas; medicamentos personalizados (farmacogenómica) y el diagnóstico en el *point-of-care* (permitiendo la obtención de datos genómicos individuales en el punto de atención sanitaria).

Finalmente, el conjunto de este tipo de tecnologías, los biochips y la bioinformática, contribuirán en un tiempo razonable a lograr un entendimiento más profundo acerca del transcriptoma completo de la célula normal y transformada y permitirán identificar de manera más rápida y eficiente, usando métodos genéticos de gran escala, el descubrimiento de nuevos genes.

Cabe aclarar que toda esta tecnología es extraordinariamente cara y que lo ideal es la conformación de grupos multidisciplinarios con propuestas afines. A la fecha en nuestra comunidad científica se cuenta ya con diferentes grupos que tienen parte de estas plataformas tecnológicas, ya sea en la generación de microarrays o en la aplicación de éstos, basta decir los grupos de la ENCB-IPN, CINVESTAV, CISEI, UNAM, IMSS entre otros. Asimismo,

proveerá las bases para entender la génesis, progresión y respuesta a las terapias. Sin embargo, la tecnología de microarreglos no lo es todo, es sólo el comienzo y para poder proseguir en estas nuevas líneas de investigación, es necesario la formación de grupos enfocados a la bioinformática, matemática teórica, biocomputación, redes génicas y neurales y al análisis de datos, con los cuales se logre extraer de toda esta minería de datos (Data mining), aquellos con una validez estadística (marcadores potenciales). La combinación de todas estas plataformas creará las bases de decisión para próximos tratamientos individualizados para los pacientes oncológicos^{48,51,59}.

REFERENCIAS

1. Greenlee, T., Murray, T., Bolden, S. y cols. Cancer Statistics. CA Cancer J. Clin. 2000; 50: 7-33.
2. Jemal, A., Murray, T., Samuels, A., y cols. Cancer Statistics. CA Cancer J. Clin. 2003; 53: 5-26.
3. Simmonds, A. Cancer Statistics. Further decrease in mortality rate, increase in persons living with cancer. CA Cancer J. Clin. 2003; 53: 4.
4. Lander, S., Linton, B., Birren, C. y cols. Initial sequencing and analysis of the human genome. Nature 2001; 409: 860-921.
5. Venter, C., Adams M., Myers, E. y cols. The sequence of the human genome. Science 2001; 291: 1304-1351.
6. Velculescu, V., Zhang, L., Vogelstein, B. Y cols. Serial analysis of gene expression. Science 1995; 270:484-87.
7. Okamoto T, Suzuki, T., Yamamoto, N. Microarray fabrication with covalent attachment of DNA using bubble jet technology. Nat Biotechnol. 2000; 18: 438-441.
8. Rose, D. Spotted arrays; technology overview. p13-14 In JA Warrington CR Todd & Wong D (Eds) , Microarrays and Cancer Research. Eaton Publishing, Westborough, MA. 2002.
9. Schena, M., (Ed). Microarray biochip technology. Eaton Publishing, Sunnyvale CA. USA. 2000.
10. Golub, T., Slonim, D., Tamayo, P., Molecular classification of cancer. Class discovery and class prediction by gene expression monitoring. Science 1999; 286: 531-37.
11. www.seis.es/i_s/i_s19/i_s191.htm
12. Lipshultz, R., Fodor, S., Gingeras, T., y cols. High density synthetic oligonucleotide arrays. Nat Genet 1999; 21: 20-24.
13. Fodor, P., Read, J., Pirrung, I. y cols. Light-direct, spatially y addressable parallel chemical synthesis. Science 1991; 251: 767-773.
14. www.ncbi.nlm.nih.gov/SNP/index.html
15. Alexandre I, Hamels S, Dufour S. Colorimetric silver detection of DNA microarrays. Anal Biochem 2001; 295: 1-8.
16. Mayani H, Salcedo M. Microchips de ADN: una nueva herramienta para el diagnóstico molecular del cáncer. Rev Inv Clin 2000; 52: 600-2
17. Khan J., Bittner M, Chen Y y cols. DNA microarray technology: The anticipated impact on the study of human disease. Biochem Biophys Acta. 1423: m17-28.
18. Alizadeh A, Eisen M, Davier R y cols. Distinct types of diffuse large B cell lymphoma identified by gene expression profiling. Nature 2000; 403: 503-511.
19. Ycoh E, Ross M Shurleff W. classification, subtype, discovery and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. Cancer Cell 2002; 1:133-143.
20. Armstrong S, Stauton J, Silverman L., Pieters R. y cols. MML translocations specify a distinct gene expression profile that distinguishes a unique leukemia. Nat. Genet. 2002; 30: 41-47.
21. Shipp M, Ross K, Tamayo P y cols. Diffuse large B-cell lymphoma outcome prediction gene expression profiling and supervised machine learning. Nat. Medicine 2002; 8:68-74.
22. Ono, K., Tanaka, T., Tsumoda, T. y cols. Identification by cDNA microarrays of genes involved in ovarian carcinogenesis. Cancer Res. 2000; 60: 5007-5011
23. Schwuartz, D., Kardia, K., Shedden, R. y cols. Gene expression in ovarian cancer reflects both morphology and biological behavior, distinguishing clear cell from other poor-prognosis ovarian carcinomas. Cancer Res. 2002; 62: 4722-4729.
24. Perou, C., Jeffrey, M., Van de Rijn, M. y cols. Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. Proc. Natl. Acad. Sci. USA. 1999; 96: 9212-9217
25. Perou, C., Sorlie, T., Eidsen, M., Van de Rijn, M. y cols. Molecular portraits of human breast tumors. Nature 2000; 406: 747-752.
26. Sorlie, T., Perou, C., Tibshirani, T. y cols. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc. Natl. Acad. Sci. USA 2001; 98: 10869-10874.
27. Bittner, M., Meltzer, P., Chen, Y. y cols. Molecular classification of cutaneous malignant melanoma by gene expression profiling. Nature 2000; 406: 536-540.
28. Dhanasekaran, S., Barrette, T., Ghosh, D. y col. Delineation of prognostic biomarkers in prostate cancer. Nature 2001; 412: 822-826
29. Bhattacharje, A., Richards, W., Stauton, J. y cols. Classification of human lung carcinoma by mRNA expression profiling reveals distinct adenocarcinoma subclasses. Proc. Natl. Acad. Sci. USA 2001; 98: 13790-13795.
30. Garber, M., Troyanskaya, O., Schluens, K. y cols. Diversity of gene expression in adenocarcinoma of the lung. Proc. Natl. Acad. Sci. USA 2001; 98: 13784-13789.
31. Ochs M, Gowin A. Microarrays in Cancer Research & Applications. Biotechniques 2003; 34:S4-15.
32. Alevizos I, Mahadevappa M, Zhang X. y cols. Oral cancer in vivo gene expression profiling assisted by laser capture microdissection and microarray analysis. Oncogene 2001; 20: 6196-6204
33. Todd R, Gutkind E, Shillito E. Solid tumors microarray analysis of oral cancers. P139-153. In JA Warrington CR Todd & Wong D (Eds). Microarrays and Cancer Research. Eaton Publishing, Westborough, MA. 2002
34. Du F, Mahadevappa M, Warrington J, y cols. Gene expression changes in endometrial cancer. P113-125. In JA Warrington CR Todd & Wong D (Eds). Microarrays and Cancer Research. Eaton Publishing, Westborough, MA. 2002
35. Hegde P, Qi R, Gaspard K. y cols. Identification of tumor markers in model of human colorectal cancer using a 19,200 element complementary DNA microarray. Cancer Res. 2001; 61: 7792-7797
36. Notterman D, Shawer C y Levine A. Tumor biology and microarray

MICRO- ARREGLOS:

monitoreando el
genoma en un chip

GUELAGUETZA VÁZQUEZ, PATRICIA PIÑA,
KARLA VÁZQUEZ, ALEJANDRA MANTILLA Y MAURICIO SALCEDO



El análisis global de la expresión génica abre las puertas para traducir los resultados de la investigación del genoma humano a aspectos de genómica funcional, otorgándonos una visión general del comportamiento de los genes y sus interacciones.

Con el borrador del Proyecto Genoma Humano (PGH), hemos sido testigos en poco tiempo de la particularidad del ADN y de la generación de nueva tecnología denominada de alto rendimiento. Una de las áreas más beneficiadas ha sido la biomedicina y actualmente es posible hablar de la medicina genómica. El nuevo reto en esta era postgenómica es poder determinar la función de los genes y las rutas metabólicas en que se encuentran, a partir de su conocimiento; es decir, obtener una visión global de diversos procesos celulares.

La combinación de investigación y aplicación tecnológica ha creado metodologías que actualmente nos permiten analizar los diferentes niveles de la genética molecular: ADN o el archivo de la información (genoma), ARN o la expresión de la información (transcriptoma) y las proteínas (proteo-

ma). Los datos generados por el PGH estiman aproximadamente 34 mil genes en nuestras células, los cuales podrían producir hasta 90 mil transcritos debido a la presencia de mecanismos tales como el corte y empalme (*splicing* alternativo).¹

Además de los ARN producidos por genes *codificantes* de proteínas, existen también los *no codificantes* que cumplirán funciones regulatorias en diversos procesos; por ejemplo, los ARN de transferencia o los del aparato de corte y empalme. Así podemos decir que este conjunto de genes expresados se denomina *transcriptoma*.

La determinación de la función de los genes involucra el análisis de la expresión génica de un organismo; es decir, el análisis del transcriptoma. Anteriormente la detección de la expre-

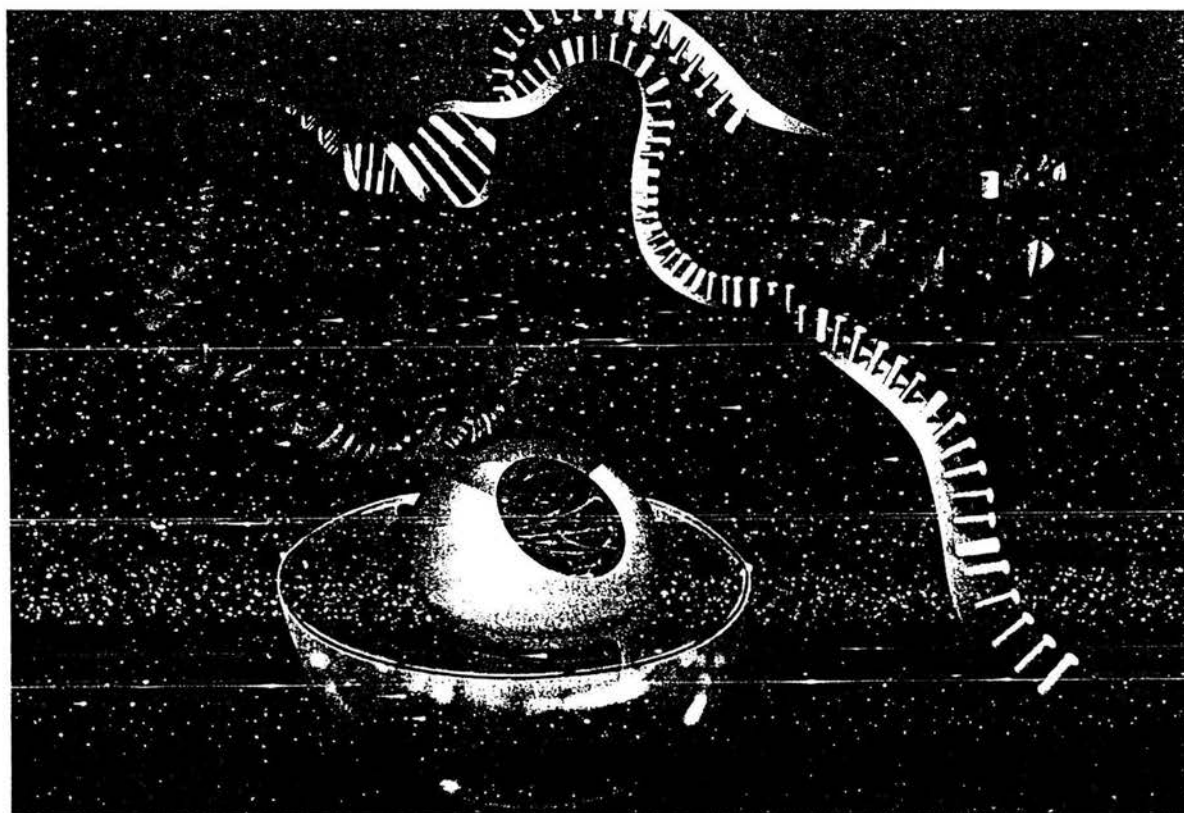


Figura 1. Procesos celulares: del DNA a las proteínas.

1. En el núcleo de la célula se da la transcripción o paso de ADN a ARN, generándose un pre-ARN mensajero (pre-mARN), el cual contiene exones (secuencias que forman parte de la proteína) e intrones (secuencias que son eliminadas para producir un mARN) maduro, el cual será traducido a una proteína. Sin embargo la maquinaria celular de corte y empalme puede generar genes diferentes al incluir o eliminar diferente intrones.

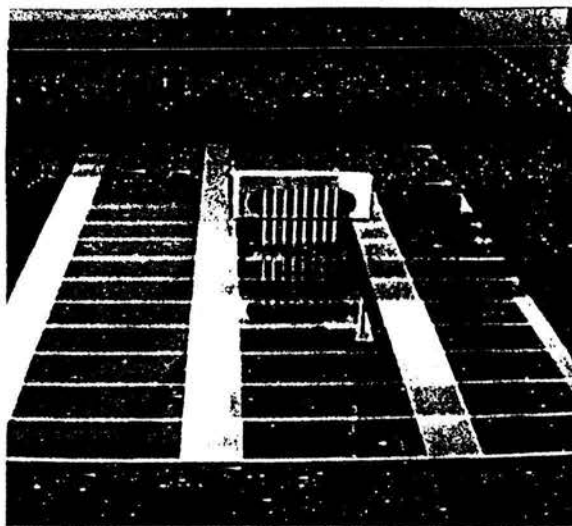


Figura 2. Brazo robótico y cabezal punteador de un microarreglador (tomado de TIGR The Institute for Genomic Research, <http://www.tigr.org/tdb/microarray/images.shtml>)



Tamaño proporcional de un chip de identificación genómica.

sión estaba limitada a un solo gen o a un pequeño grupo de ellos; y no alcanzaba a resolver la gran magnitud del problema, consistente en determinar la función de los genes. Por lo tanto, la obtención de patrones globales de expresión de los procesos biológicos descifrarán poco a poco todos los componentes sobre el funcionamiento de los genes y las redes biológicas funcionales entre ellos.

La expresión génica y los sistemas de análisis global

Actualmente se sabe que los diferentes tipos celulares que conforman el cuerpo humano tienen patrones únicos de expresión específicos para sus funciones fisiológicas en particular. Los agentes internos y externos, pueden modular estos patrones causando estados de alteración fisiológica o patologías. Cuando se obtiene la *fotografía* global de la expresión génica entre diferentes tipos celulares, o bien entre diferentes estados de un tipo celular en particular (por ejemplo: célula normal versus célula tumoral) se podrían identificar los genes candidatos involucrados en procesos celulares normales (figura 1), o bien en el desarrollo de una misma enfermedad. Adicionalmente, la caracterización de genes expresados anormalmente en tejidos tumorales puede generar el descubrimiento de genes que sirvan como posibles marcadores de la enfermedad o de un diagnóstico, pronóstico o intervención terapéutica.²

Para llevar a cabo el análisis global de expresión génica

2. Un gen marcador es aquel cuyo cambio de expresión está directamente asociado a una patología. Es decir, un gen "X" que en un tejido normal presenta un determinado nivel de expresión (por ejemplo, 3), ese mismo gen en la contraparte tumoral de ese tejido presenta otro nivel de expresión (ejemplo: 15). Entonces sabemos que cuando un tumor este presente en ese tejido, el gen "X" se estará sobreescribiendo.

existen dos tipos de sistemas: abierto y cerrado. Se denomina sistema abierto cuando el análisis determina la expresión del total de los genes de la célula en un momento determinado. Una tecnología de tipo abierto es ASEG, o Análisis en Serie de Expresión Génica (SAGE por sus siglas en inglés), la cual ha sido aceptada por el Proyecto del Genoma de la Célula Cancerosa (Por sus siglas en Inglés CGAP, NCCI). El sistema cerrado consiste en la determinación de la expresión de un número limitado y conocido de genes de una célula; en este sistema se encuentran los microarreglos (*microarrays* en inglés) de ADN complementario (cADN) los cuales están cobrando un auge imprevisto en los últimos años. Estos *microarrays* (arreglos o *biochips*) evolucionaron de la idea de Ed Southern desarrollada hace aproximadamente un cuarto de siglo, cuando afirmaba que moléculas de ácidos nucleicos marcadas podían ser utilizadas para *hibridarse* y detectar moléculas de ácidos nucleicos complementarios unidos a un soporte sólido que no permitiera la reasociación (Southern blot). Sin embargo, el primer microarreglo en llevarse a cabo fue el *Dot Blot reverse* que es un microarreglo elaborado manualmente con la tecnología de la época y una cantidad muy pequeña de genes por estudiar. Los arreglos de cADN o *biochips* constan de cientos a miles de secuencias de cADN inmovilizadas en una superficie sólida, del tamaño de media hoja carta, hasta una superficie de 1 cm². Estas secuencias se golean o depositan por duplicado utilizando un brazo robótico y cada gota representa un gen en particular en un proceso de miniaturización (figura 2).

Los arreglos de cADN son de dos tipos: los macro y microarreglos y sus diferencias básicas son: a) el tamaño del punto (en los macroarreglos llega a los 300 micrones, y en el microarreglo es de 100 a 200 micrones), b) el número de puntos inmovilizados (es decir, el número de genes), c) el material del que está hecho el soporte y d) la manera como

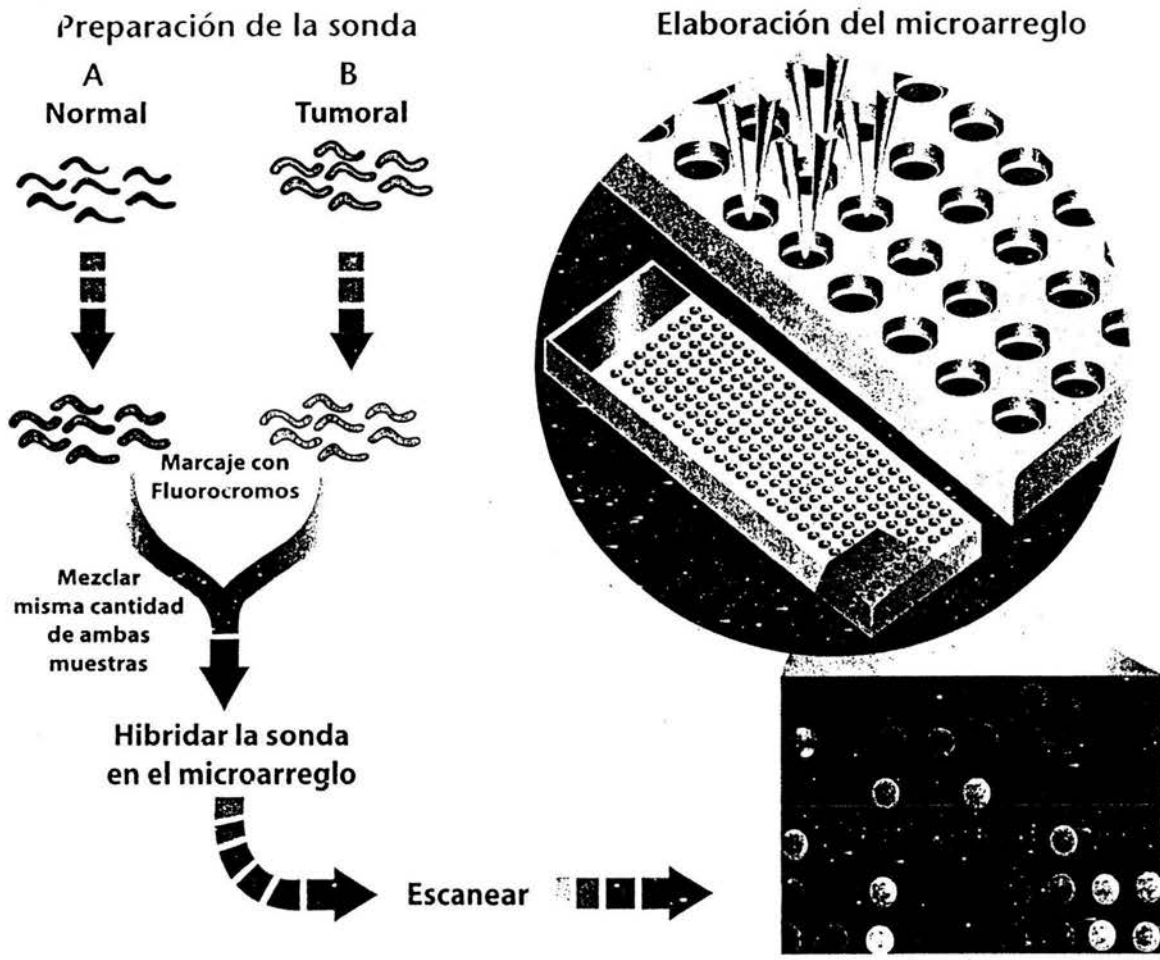


Figura 3. Proceso de hibridación de un microarreglo. (Tomado de Access Excellence, The National Health Museum, <http://www.accessexcellence.org/AB/GG/>)

las sondas son hibridadas. Los macroarreglos generalmente son punteados sobre membranas de nylon o nitrocelulosa, mientras que los microarreglos se colocan sobre vidrio.

Básicamente la aplicación del sistema de *arrays* consiste en realizar un análisis comparativo de la expresión (figura 3); por ejemplo: para el caso de muestras con cáncer se trabaja con el mRNA de las células, se sintetiza el ADN complementario (por la técnica de RT-PCR) y se procede de la siguiente forma: 1) se cuenta con la muestra normal (A) y su contraparte tumoral (B); 2) A se marca con un fluorocromo de un color (por ejemplo, rojo); 3) y en B se marca con otro fluorocromo (verde); 4) se mezclan e hibridan los cADN sobre el arreglo, en un experimento de competencia; 5) se iluminan los puntos con un *scanner* con láser y se determina la intensidad de los fluorocromos sobre cada uno de ellos. Si hay variación se verán puntos rojos y verdes, pero si no hay expresión tampoco habrá color. Si la expresión es similar, ambos cADN hibridados darán un color amarillo, de tal manera, este tipo de análisis, permite en un solo experimento llevar a cabo el equivalente a miles de experimentos individuales.

Podemos encontrar otras plataformas tecnológicas donde el experimento no es competitivo; es decir, donde cada mues-

tra se hibrida en arreglos de manera diferente, los cuales han sido elaborados con cADN inmovilizado. Existen también los arreglos de oligonucleótidos –sintéticos de alta densidad–, los cuales se basan en su síntesis química *in situ* de una secuencia de 25 bases correspondiente a algún gen conocido.

En la actualidad cualquier tipo de los arreglos se venden de forma comercial y su complejidad va desde cientos hasta decenas de miles de genes, incluyendo conocidos y secuen-

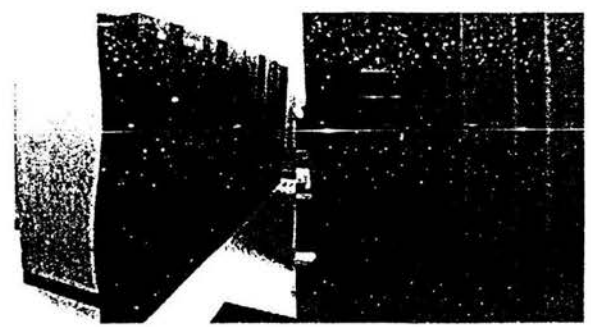


Figura 4. Supercomputadoras en red para análisis de datos de expresión provenientes de laboratorio. (Tomado de US Department of Energy Human Genome Program, <http://www.ornl.gov/hgms/>)

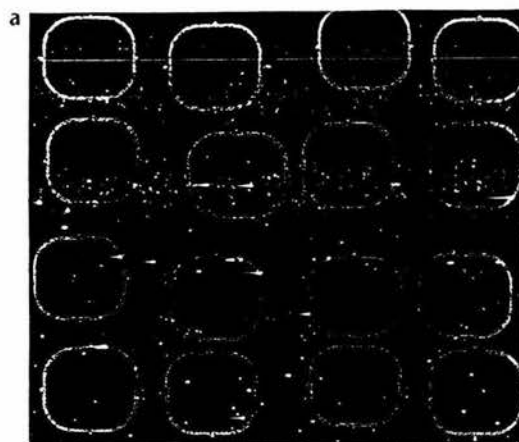
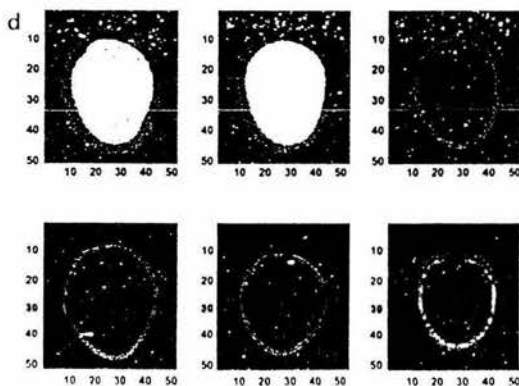
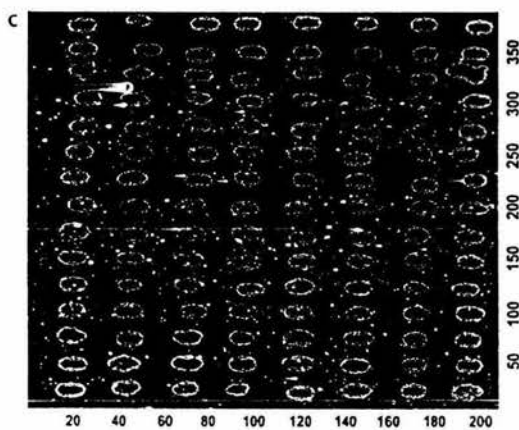
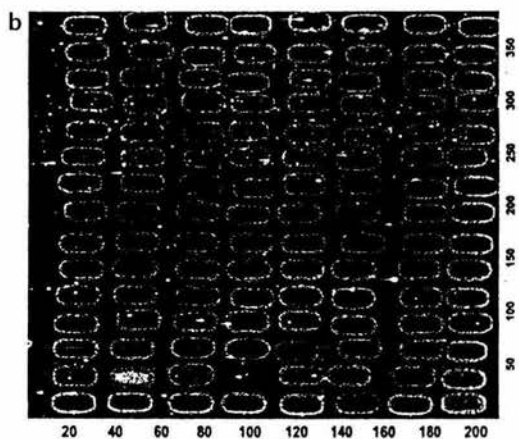


Figura 5. Ejemplo de software de análisis de imágenes.
a), b) y c) Acomodo de la plantilla virtual de lectura sobre el microarreglo, d) Aplicaciones del software



cias de función desconocida llamadas *Secuencias Tag Expresadas* (EST por sus siglas en inglés) lo que nos permite el descubrimiento de nuevos genes.

Flujos de información

Antes de poder decidir sobre el uso de alguna de estas metodologías de análisis global, se debe verificar: a) la pregunta del experimento, b) la disponibilidad del arreglo, c) el equipo de captura de imagen y *software* apropiado, d) el costo y e) una colaboración sólida y multidisciplinaria con un grupo (o grupos) que desarrolle técnicas biomatemáticas y bioinformáticas para el análisis apropiado de los datos (figura 4). Esto último se debe a que un solo estudio puede generar miles de datos que definitivamente no pueden ser analizados a mano o en gráficas simples. La comunidad científica hasta ahora no ha determinado completamente cómo enfrentarse a estas cantidades masivas de datos para explorarlos e interpretarlos. Afortunadamente, en nuestro país están surgiendo grupos con interés en esta materia, los cuales ya cuentan con herramientas disponibles para el procesamiento, almacenaje y análisis de los resultados. Sin embargo, no existe un consenso internacional para comparar los resultados obtenidos por las diferentes técnicas de detección de expresión (por ejemplo: microarreglos, SAGE, chips de oligonucleotidos).

Existe una gran variedad de herramientas que han sido desarrolladas para la captura y el procesamiento de las imágenes de los arreglos. La meta principal es reducir la imagen a puntos de intensidad variable captada independientemente de otros puntos. Aunque suena relativamente sencillo, aún no existe una manera en común para extraer la información de expresión. Muchos grupos todavía están diseñando *softwares* y algoritmos para este propósito (figura 5).

La meta a lograr en este tipo de tecnologías de alto rendimiento es extraer el significado y las implicaciones de los datos, y la única manera de hacerlo es integrando los resultados experimentales con recursos externos como los del Sistema Entrez. Es decir, todos los genes que se encuentran puntuados en un arreglo se encuentran en dirección mediante un número de acceso a esta base; sin embargo, aunque ésta y otras bases de datos son capaces de manejar toda esta información, existen serios problemas para establecerlas como un sistema. En algunos casos es más fácil y factible poder entender los datos, correlacionando los cambios de expresión en procesos y rutas metabólicas conocidas.

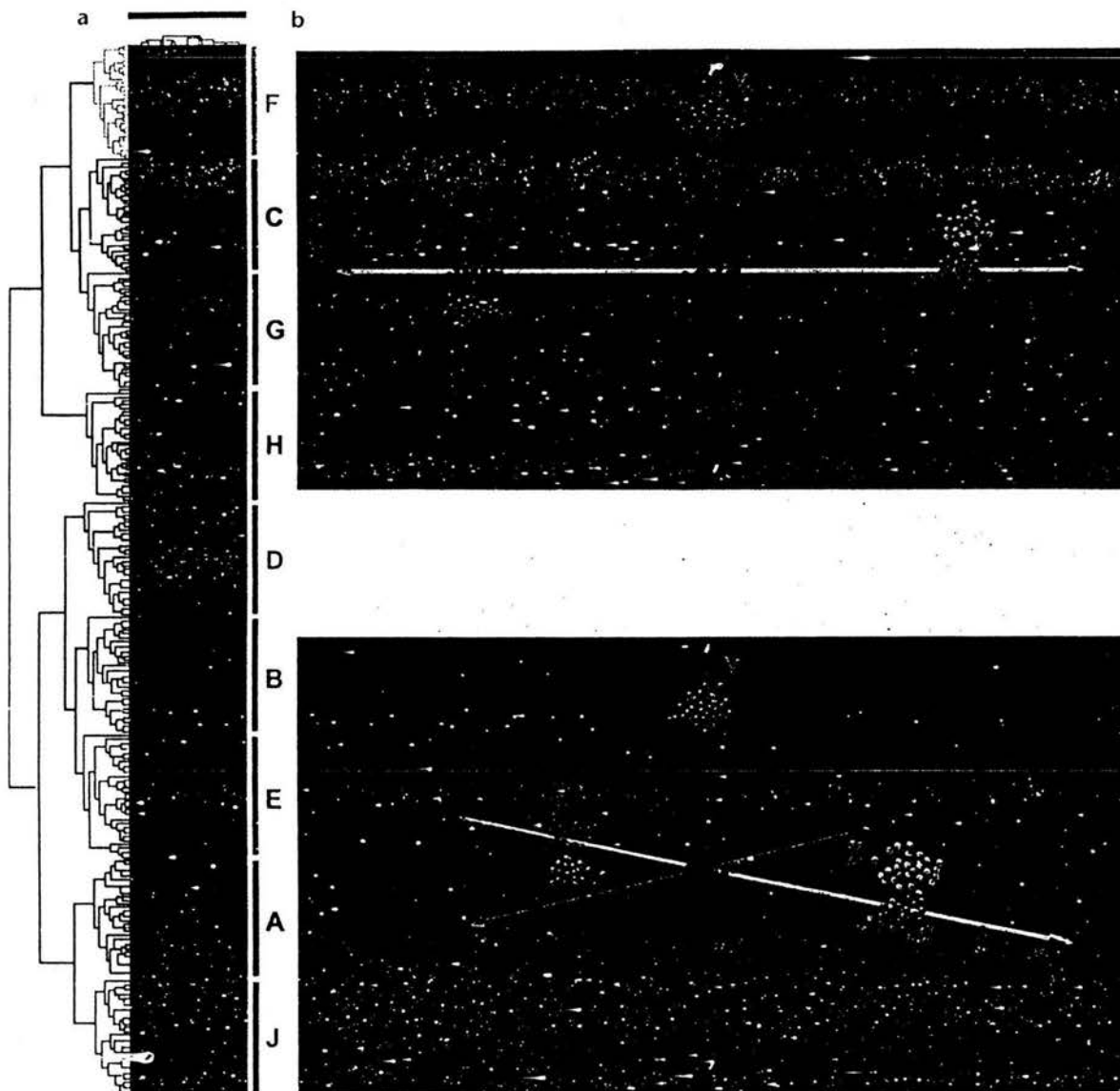


Figura 6. Ejemplos de análisis estadísticos. a) Cluster jerárquico y b) Análisis de Componente Principal, por sus siglas en inglés (PCA). (Tomado de: Quackenbush J. Computational Analysis of Microarray data. Nature Reviews Genetics. 2001. 2:418-427.)

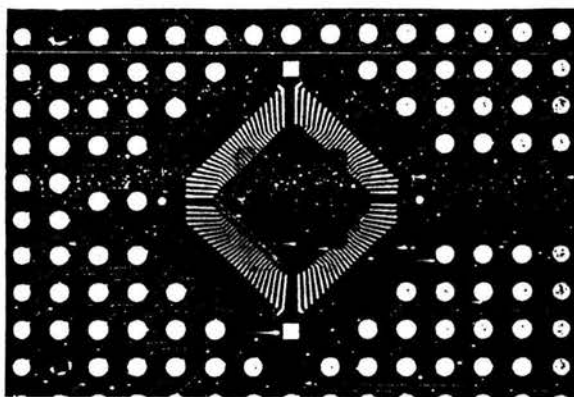
Análisis estadístico

A la fecha se conocen algunos métodos para la normalización de datos entre experimentos. No obstante, todos sufren variabilidad *intra* e *inter* experimental al determinar la expresión génica, haciendo que las comparaciones directas entre experimentos incluyan errores de análisis. Una meta acertada de las bases de datos públicas de expresión sería almacenar datos de diversas tecnologías de análisis de expresión en una forma estandarizada e *interoperable*. Desafortunadamente, los arreglos de *cDNA* no pueden determinar el número de transcritos por célula, en cambio miden la abundancia relativa de los transcritos, lo cual imposibilita la conversión de los datos de expresión a un formato universal, es decir, a una unidad métrica legible, por lo que se distor-

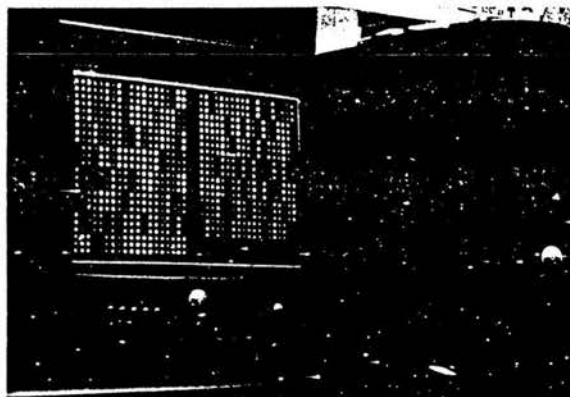
siona la información primaria y no es posible establecer estándares para las bases de datos públicas.

En este mismo contexto de normalización, hasta la fecha se manejan tres formas que son las mayormente aceptadas: normalización de la intensidad total, utilización de técnicas de regresión y normalización mediante la aplicación de estadísticas de la razón y para las tres existen algoritmos ya establecidos.

Suena fácil suponer que el lugar y el momento en que un gen es expresado nos puede decir cuál es su función, pero eso no ocurre en la realidad. De ahí que el primer paso sea organizar los genes con base en su patrón de expresión por medio de *clusters* o grupos. Un *cluster* engloba diferentes algoritmos de clasificación que pueden ser utilizados para desarrollar taxonomías. La mayoría de los



Acercamiento a un chip que contiene información genética.



Interpretación en pantalla de la plantilla virtual de lectura sobre el microarreglo.

clusters son jerárquicos, la clasificación resultante tiene un número de clases y el producto asemeja una clasificación filogenética (figura 6). También existen técnicas de *clustering* no jerárquicas, como el K-means donde simplemente se dividen los *objetos* en diferentes grupos sin tratar de especificar una relación entre ellos.

Los métodos de *clustering* también pueden ser clasificados como: divisivos o aglomerativos. Un método divisivo empezará con todos los elementos en un *cluster* y gradualmente los separará hasta tener *clusters* más pequeños. Los aglomerativos, por su parte comienzan con varios *clusters* que se van fusionando. Finalmente, el *cluster* puede ser *supervisado* o *no supervisado*. El método supervisado utiliza información biológica existente sobre genes específicos que están relacionados funcionalmente para guiar el algoritmo del *cluster*. Aunque la mayoría de los métodos son no supervisados, tenemos que lidiar primeramente con éste.

El Análisis por *cluster* ha sido la herramienta estadística más ampliamente usada. La información obtenida por análisis de expresión a gran escala es una de muchos instrumentos que existen para analizar los datos y es necesario tener cuidado al aplicar esta técnica, pues, aunque los métodos utilizados son objetivos, debido a que los algoritmos están bien definidos y son reproducibles, todavía existe subjetividad al seleccionar diferentes algoritmos, normalizaciones o distancias métricas, porque pueden agrupar los objetos en diferentes *clusters*. Otra manera habitual de analizar los datos de expresión es mediante los *mapas de autoorganización* (Self Organizing Maps, SOM), redes neuronales basadas en *clusters* divisivos.

Aunque todavía no han sido muy desarrollados, algunos métodos computacionales han probado su buen funcionamiento en el procesamiento de información clínica y experimentalmente útil de los datos obtenidos de los perfiles de expresión. Otras técnicas computacionales empleadas actualmente son: redes booleanas, modelaje lineal, Análisis de componente principal (figura 6), modelaje no lineal, redes bayesianas, redes dinámicas, redes Petri, etcétera; sin embargo, aunque las herramientas y técnicas mencionadas anteriormente, así como nuevos algoritmos y softwares están

siendo desarrollados, la decisión de la técnica dependerá del objetivo de nuestro experimento y por lo tanto, los métodos matemáticos y bioinformáticos específicos para llegar a la respuesta adecuada. De ahí que el estudio de la expresión génica se haya convertido en una ciencia multidisciplinaria que incluye disciplinas como: ingeniería, matemática, biología, estadística e informática, favoreciendo así la colaboración entre diferentes grupos de investigación.

Validación

Debido a la inherente variabilidad de los datos de un microarreglo, es importante para los investigadores validar sus descubrimientos utilizando otras técnicas; por ejemplo: Northern Blot, RT-PCR cuantitativa, Hibridación *in situ* RNA-RNA o inmunohistoquímica.

Aplicaciones de los microarreglos

Dado que los microarreglos incluyen EST existe la posibilidad del descubrimiento de nuevos genes. Otros usos podrían ser la observación del resultado después de una perturbación específica y el análisis de la divergencia entre dos células diferentes. La identificación de patrones complejos de expresión utilizando análisis de *cluster*, podría ser otra de las utilidades de esta metodología.

Existen otras áreas de gran potencial de estudio que van desde el examen de los procesos naturales hasta la evaluación de seguridad de fármacos, incluyendo la división celular o envejecimiento, progresión de enfermedades, intervenciones farmacológicas e identificación de carcinógenos. Hasta el momento, el área principalmente trabajada ha sido la de la comparación de perfiles de expresión en tejidos normales con los de su contraparte tumoral. Además de revolucionar la clasificación de patologías, los patrones de expresión y los análisis de *cluster* prometen mejorar las actuales limitaciones en el diagnóstico químico e inmunológico de tejidos (inmunohistoquímico).

Finalmente, estos patrones podrían ser extremadamente útiles para la elaboración de pronóstico e identificación de

La caracterización de genes expresados en tejidos tumorales puede generar el descubrimiento de genes que sirvan como marcadores de diagnóstico, pronóstico o intervención terapéutica

subgrupos en pacientes que puedan responder de manera diferente en una terapia. La genómica funcional y la tecnología de microarreglos se encuentran aún en su infancia; no obstante, basándonos en el rápido progreso que ha tenido esta tecnología en los últimos cinco años, es casi seguro que cada vez será más accesible y útil.

En nuestro grupo, actualmente estamos utilizando esta tecnología para el estudio del cáncer cérvico uterino buscan-

do algunos genes que pueden ser candidatos de pronóstico o diagnóstico para esta patología, además de otro tipo de marcadores. Por otro lado, en conjunto con otros grupos de investigación se exploran nuevas posibilidades de análisis para esta metodología. De esta manera el análisis del transcriptoma humano es ahora una realidad en México. ☺



Consulte las referencias bibliográficas en www.conacyt.mx, en el vínculo *Ciencia y Desarrollo*.

Guelaguetza Vázquez Ortiz es ingeniera en biotecnología ambiental por la Universidad Autónoma de Guadalajara y cursa el programa de doctorado en investigación biomédica en la UNAM. Ha realizado estancias doctorales en España, con apoyo del Conacyt.

Patricia Piña Sánchez es bióloga por la UNAM y maestra en ciencias químico biológicas por el IPN. Ha sido becada por el Conacyt y actualmente trabaja en el Laboratorio de Oncología Genómica del Centro Médico Nacional Siglo XXI.

Karla Vázquez Santillán es bióloga por la UNAM y realizó su tesis en el Laboratorio de Oncología Genómica del Centro Médico Nacional Siglo XXI.

Alejandra Mantilla Morales estudió medicina en la UNAM, cursó la especialidad en Patología en el Hospital General de México y realizó la maestría en ciencias en la UNAM. Es becaria del Conacyt y se encuentra adscrita al Departamento de Patología del Hospital de Oncología del Centro Médico Nacional Siglo XXI.

Mauricio Salcedo Vargas es químico bacteriólogo y parasitólogo por el IPN, maestro en biología molecular y genética y doctor en biología molecular por el CINVESTAV. Fue becario de la Escuela de Medicina de la Universidad de Harvard, donde obtuvo el premio Aida Weiss. Es miembro del SNI, nivel I y es investigador y jefe del Laboratorio de Oncología Genómica del Centro Médico Nacional Siglo XXI.



VE MÁS ALLÁ

LA DICHOSA PALABRA EL JUEGO CON TU IDIOMA

Con Laura García, Pablo Boulosa, Germán Ortega,
Nicolás Alvarado, Eduardo Casar

Sábados 7 de la noche



Revista de Investigación Clínica

Organo Oficial del
Instituto Nacional de Ciencias Médicas y Nutrición
Salvador Zubirán

EDITOR

Dr. Gerardo Gamba

EDITORES ASOCIADOS

Dr. Miguel Herrera

Dr. Fernando Larrea

Dr. Sergio Ponce de León

Dr. Armando Tovar

Dra. Angelina Villasis

EDITOR FORO CLÍNICO

Dr. Eduardo Carrillo

20 de Enero de 2005

Dr. Mauricio Salcedo
Av. San Isidro 91 C-201
Col. Ampl. Petrolera
Del. Azcapotzalco
02470 México DF

04-120

Estimado Dr. Salcedo:

El Comité Editorial ha decidido aceptar para publicación su artículo intitulado: "Análisis de expresión global del cáncer cérvico uterino: rutas metabólicas y genes alterados.", el cual aparecerá publicado próximamente.

Gracias por contribuir con *La Revista*. Esperamos contar nuevamente con su valiosa colaboración.

Atentamente,



Dr. Armando Tovar
Editor Asociado

Análisis de expresión global del Cáncer Cérvico Uterino: rutas metabólicas y genes alterados.

Guelaguetza Vázquez-Ortiz*, Patricia Piña-Sanchez*, Alfredo Hidalgo*, Minerva Lazos**, José Moreno***, Isabel Alvarado****, Fernando Cruz*****, Dulce M. Hernández*, Carlos Pérez-Plascencia*, Mauricio Salcedo*

* Unidad de Investigación Médica en Enfermedades Oncológica, Hospital de Oncología, Centro Médico Nacional Siglo XXI, México

** Departamento de Patología, Hospital General de México, SS, México.

*** Unidad de Investigación Médica en Enfermedades Autoinmunes, Hospital de Especialidades, CMN SXXI-IMSS

**** Departamento de Patología, Hospital de Oncología, Centro Médico Nacional Siglo XXI, México

***** Centro Nacional de Clínica de Displasias, Hospital General de México, SS, México.

Autor Correspondiente:

Dr. Mauricio Salcedo.

Unidad de Investigación Médica en Enfermedades Oncológicas

Hospital de Oncología

Centro Médico Nacional Siglo XXI-IMSS

Av. Cuauhtémoc 330. Col Doctores.

México D. F. C. P. 06720

Phone: +55 56276900 Ext: 22708 Fax: +55 57610952

Email: maosal89@yahoo.com, quelav@yahoo.com

RESUMEN

La infección por Virus de Papiloma de alto riesgo (VPH) es considerada como el factor etiológico más importante del Cáncer Cérvico Uterino (CaCU). Con el fin de determinar el patrón de expresión global e identificar algunos posibles genes marcadores del CaCU, se utilizaron microhileras de ADN que contenían 8000 secuencias que codificaban para transcritos diferentes, para estudiar los perfiles de expresión de 5 líneas celulares derivadas de CaCU, tres muestras tumorales conteniendo VPH16 y tres muestras normales negativas para la presencia de VPH. Se identificaron los niveles de expresión de genes relacionados con diferentes rutas metabólicas. Se llevó a cabo el análisis de agrupamiento jerárquico y posteriormente se confirmó la sobreexpresión de dos genes mediante RT-PCR. Estos dos genes se encontraron sobreexpresados en biopsias tumorales cervicales. Uno de ellos, el gen de IL6, que ha sido previamente reportado en relación con CaCU; así como el gen de la Matriz-metaloproteasa 10 (MMP10) por primera vez relacionado a esta neoplasia.

El análisis de agrupamiento jerárquico, además reveló que las muestras que contienen el mismo tipo viral están asociadas, sugiriendo posibles diferencias en expresión entre tipos virales.

PALABRAS CLAVE: *VPH, Cáncer Cérvico Uterino, microhileras, expresión.*

Global expression analysis in Uterine Cervical Cancer: Metabolic Pathways and altered genes.

ABSTRACT

High risk human papillomavirus (HPV) infection is considered to be the most important etiological factor of Cervical Uterine Cancer. In order to determine the global expression pattern and to identify possible molecular markers of Cervical Cancer, cDNA arrays with probe sets complementary to 8,000 human genes were used to examine the expression profiles among 5 cell lines derived from human cervical cancer, three HPV16(+) tumor samples and three normal cervical tissues HPV(-). The levels of expression of different cellular processes were identified. Hierarchical clustering was performed and the gene expression using RT-PCR was confirmed. Two genes were found to be consistently overexpressed in invasive cervical cancer biopsies; one of them, IL-6 was previously reported to be overexpressed in cervical cancer and one novel gene, MMP10, previously not known to be related to cervical cancer. Hierarchical clustering of the expression data revealed that samples with common HPV type infection grouped together, maybe this could mean that differences between HPV types could be indirectly determined by expression profiles.

KEYWORDS: *HPV, Uterine Cervical Cancer, microarrays, expression.*

INTRODUCCION

El carcinoma cérvico uterino (CaCu), representa la segunda causa de mortalidad en población femenina a nivel mundial¹, y es una de las principales causas de muerte en la población femenina mexicana. Este tumor representa más del 24% de todas las neoplasias malignas reportadas en México, haciendo de esta neoplasia uno de los problemas de salud más graves en México.²

En diferentes reportes de tipo epidemiológico, se han determinado diversos factores de riesgo asociados al desarrollo del CaCu, como el número de parejas sexuales, inicio temprano de vida sexual activa, multiparidad, etc.³⁻⁴ Sin embargo, el factor etiológico más importante para el desarrollo del CaCu es la infección con virus de papiloma humano de alto riesgo (VPH-AR)⁵⁻⁸ principalmente los tipos 16 y 18.⁹

Se ha demostrado que los genomas de los VPH-AR usualmente se integran al genoma de la célula hospedera. Esta integración se realiza en diversos sitios, algunos de ellos cercanos a oncogenes específicos o sitios frágiles, sugiriendo que este evento puede tener un papel importante en el desarrollo de los tumores cervicales¹⁰

La propiedad oncogénica principal de las proteínas virales E6 y E7 es su capacidad de inactivar a las proteínas celulares supresoras de tumor pRB y p53¹¹⁻¹² respectivamente, permitiendo así la transformación e inmortalización de las células infectadas. La inactivación de los genes supresores de tumor, ya sea por mutaciones génicas o por presencia de virus genera la desregulación del ciclo celular y la activación de varios proto-oncogenes.¹³ Estas interacciones no han sido totalmente estudiadas¹⁴ desconociéndose por lo tanto, el comportamiento global del CaCu.

En el presente trabajo fueron utilizadas microhileras de expresión (ADNc) para explorar el comportamiento global de la expresión génica en el CaCu. Los patrones de expresión de las diferentes rutas metabólicas estudiadas permiten observar que estas en general, se encuentran modificadas en el proceso de carcinogénesis cervical con respecto a su contraparte normal.

MATERIAL Y METODOS

Las líneas celulares HeLa (VPH18), SiHa (VPH16) y CaSki (VPH16) se crecieron en el medio de cultivo de Dulbecco. Por otro lado, es importante mencionar que aunque el HPV18 ha sido asociado preferencialmente a adenocarcinomas, también ha sido encontrado en carcinomas escamosos del cérvix.¹⁵ Otras líneas celulares utilizadas en este trabajo: CALO e INBL que presentan HPV18 derivan de: un Carcinoma escamoso invasor del cervix estadio IIB y de un Carcinoma escamoso invasor estadio IVA, respectivamente.¹⁶ En el presente trabajo se obtuvieron ocho tejidos post-mortem del cervix uterino, sin lesión, de mujeres premenopáusicas, cuya muerte no estuvo relacionada a procesos ginecológicos, y sin presencia de miomatosis uterina, del Departamento de Patología del Hospital de Oncología, CMN Siglo XXI. Los tejidos invasores (8 muestras) fueron obtenidos a través de biopsias guiadas por colposcopia de la lesión tumoral de mujeres premenopáusicas en el Centro Nacional de Displasias del Hospital General de México, S.S. Todos los procedimientos descritos han sido evaluados y aprobados por el comité local de ética del Hospital de Oncología del Instituto Mexicano del Seguro Social y las muestras invasoras fueron tomadas bajo el consentimiento de las pacientes previamente informadas. Todas las muestras se dividieron en tres secciones, la parte media se congeló en nitrógeno líquido y se almacenó a -70°C hasta su uso. Los extremos de la biopsia fueron fijados en etanol al 70% e incluidos en parafina. A diferentes cortes histológicos de los extremos de los tejidos fijados y embebidos en parafina se les realizó la tinción de hematoxilina/eosina con el fin de confirmar la presencia de al menos un 80% de células tumorales presentes en la biopsia. Mediante

estos cortes, también se realizó la clasificación del tumor. Todas las muestras invasoras fueron diagnosticadas como lesiones escamosas.

Tanto el ADN como el ARN fueron extraídos de las líneas celulares, así como de los tejidos congelados con Trizol (Life Technologies, Grand Island NY USA). Con el fin de determinar el tipo viral presente en las muestras o bien la ausencia del virus en las muestras normales, el ADN de cada una de estas fue sometida a reacción en cadena de la polimerasa o PCR utilizando los oligonucleótidos consenso MY09/MY11 y secuenciación automatizada de ADN.

El ARN fue tratado con *RNAse-free DNase I* (Ambion Inc). La concentración de ARN total fue determinada por absorbancia y la integridad del ácido nucleico fue confirmada por electroforesis. El ARN mensajero se aisló de HeLa, SiHa, CaSki, CaLO, INBL, de tres muestras normales (N03, N11, N22) y de tres muestras tumorales VPH16 (T07, T31 y T64) mediante el kit Polytract Isolation System II (Promega Madison, WI), desafortunadamente del resto de muestras tumorales y normales no se obtuvo la cantidad de mensajero necesaria; de ahí que éstas fueran utilizadas para realizar RT-PCR. Sin embargo, es importante mencionar que en la actualidad aún no existe publicado ningún método estadístico que permita definir un tamaño muestral específico para un análisis por microarrays; y el número de muestras que utilizamos en este manuscrito es similar al que se utiliza en diversos artículos publicados en donde utilizan la misma tecnología. Los ARNs mensajeros se cuantificaron utilizando el VersaFluor Fluorometer (BioRad). Los ADNc marcados fueron obtenidos por la reacción de transcripción reversa acoplada a PCR o RT-PCR a partir de 800 ng de ARNm en presencia de (α -³³P) dATP utilizando el kit cDNA Labeling Ambion Advantage System (Ambion Inc, USA). Se realizó la hibridación de las microhileras que contienen 8,400 clonas por duplicado (ULTRArray Advantage System, Arrays and Hibridization, Ambion Inc, Austin Tx) siguiendo las instrucciones del fabricante.

La imagen de las microhileras de ADNc se capturó digitalmente mediante el STORM 860 phosphorimager/fluorimager (Molecular Dynamics, Buckinghamshire UK). Una vez capturadas las imágenes, estas fueron analizadas y normalizadas con el paquete ArrayVision software 8.0 rev 4 y LOESS (Imaging Research Inc, USA). Se realizó un corte asignado en los datos, en donde aquellos datos que tuvieran un nivel de expresión menor a 0.5, no fueran tomados en cuenta como sobrexpresados. Las muestras fueron normalizadas contra la mediana de las muestras control (3 muestras normales VPH(-)). La medición de cada gen de las muestras experimentales se hizo dividiendo el valor de muestra entre la mediana de ese gen en su muestra control correspondiente. Con el fin de visualizar los niveles de expresión de los genes en las diferentes muestras, se realizó un agrupamiento o cluster jerárquico utilizando el programa J-Express disponible de manera gratuita en la red (www.molmine.com) y el dendrograma se visualizó con el programa Treeview.

Con el fin de entender el comportamiento biológico del CaCU, se utilizó la base de datos Gene Ontology (<http://www.ebi.ac.uk/GOA/>) con el fin de clasificar a los genes en diferentes rutas metabólicas. Cada gen presente en el microarreglo fue asociado con su correspondiente número de Unigene (<http://www.ncbi.nlm.nih.gov>), y tanto los datos de expresión como los números de Unigene se integraron en el website de Gene Ontology. (<http://www.geneontology.org>) y se procedió a realizar el análisis. Finalmente se obtuvo para cada uno de los genes el Código de Gene Ontology para procesos biológicos es decir, se obtuvieron las anotaciones funcionales de los genes que componen el arreglo. El programa de Hierarchical clustering (Gene Cluster v2.11) y el programa de visualización (Tree View v1.50) fueron utilizados para graficar los análisis de cada una de las rutas (<http://rana.lbl.gov/Eisensoftware.htm>) (Figura 1). Posteriormente se realizó un análisis de medianas. Este se hizo de la siguiente manera: de cada gen a estudiar se obtuvo la

mediana de las muestras tumorales y de las líneas celulares y esta se comparó contra la mediana de las muestras normales VPH(-). Se hizo una sumatoria del total de las medianas de los genes que componen cada ruta metabólica estudiada y así se obtuvo el comportamiento global de la expresión en estas mismas.

Para obtener los ADNc; 200 ng de ARN total de 5 tejidos normales (obtenidos postmortem), las líneas celulares y de 5 muestras tumorales (obtenidas del Centro Nacional de Displasias del Hospital General de México, S.S.) fueron sometidas al uso del estuche de RT-PCR Access System (Promega, Madison, WI). Los oligonucleótidos para los genes de interés fueron diseñados utilizando el programa Biotools de Integrated DNA Technologies (<http://biotools.idtdna.com/gateway>). Para la amplificación de los siguientes transcritos: Matriz metaloproteasa 10 (MMP10), 509 pb TM=60°C (5'-CATTGAGTCTCTCTACGGACCT-3'), (5'-CAGCTATTAGTCTAGGGAAGCC-3') Interleucina 6 (IL6), 621 pb: TM=55°C (5'-AGTTGCCTTCTCCCTGG3') (5'-TGAGGGGCTCTTGGGCAAAT-3'), y β -actin (5'-TGAAGTCTGACGTGGACATC-3') (GTTTCGTTCCCTCATACTGCTCA-3') 243 pb, TM=55°C. Las condiciones de la RT-PCR fueron las siguientes: 48°C durante 45 min y 94°C por 2 min. Seguido por 94°C por 30 seg, (la TM°C específica para cada par de oligonucleótidos) durante 1 min, y 68°C por 2 min en un total de 25 ciclos seguido por un último paso de 68°C durante 7 min. Los productos se visualizaron en geles de garosa al 1.5% teñidos con bromuro de etidio, y las imágenes se capturaron utilizando Eagle Eye System (Stratagene, La Jolla, Ca). Los tamaños de los productos fueron determinados por comparación con un marcador de 100 pb (GeneRuler, MBI Fermentas).

RESULTADOS

Con el fin de estudiar el transcriptoma del cáncer cervical, en el presente trabajo iniciamos con la definición de los patrones de expresión de líneas celulares derivadas de CaCU y de tres muestras de CaCU positivas para VPH16, comparando las medianas de los valores de expresión de las líneas celulares y tejidos enfermos con la mediana de los valores de expresión del cérvix normal negativo para secuencias de VPH. De este análisis, se lograron agrupar los genes estudiados en genes sobreexpresados, suprimidos y aquellos que no tuvieron cambios en los valores de expresión con respecto a tejidos normales.

Con el fin de identificar patrones de expresión génica que se modificaran en las muestras utilizadas, se realizó de primera instancia un análisis de tipo de agrupamiento jerárquico con una matriz de visualización pseudo-coloreada.¹⁷ En dicho agrupamiento o cluster jerárquico, el eje "Y" agrupa genes con niveles de expresión similar y en el eje "X" agrupa muestras con expresión similar. Por otro lado, la matriz coloreada del "cluster" permite que la sobre expresión génica sea fácilmente identificada cuando se comparte entre líneas celulares y muestras (puede visualizarse como un renglón coloreado en rojo). Esto significa que estos genes pueden ser potencialmente utilizados como marcadores. Una propiedad notable de los datos estudiados fue que las muestras y tumores infectados con el mismo tipo de VPH se agrupan en dos ramas distintas en el dendrograma (Figura 1a): es decir, que las líneas celulares SiHa y CaSki (VPH16) y las muestras tumorales VPH16(+) T07, T31 y T64; así como las líneas celulares VPH18(+) (INBL, HeLa, CaLO) fueron segregadas en dos ramas terminales independientes en el cluster.

Por otro lado, se identificaron grupos de genes involucrados en procesos celulares específicos, la variación en sus niveles de expresión podría reflejar las diferencias en la actividad de estos procesos en las diferentes muestras estudiadas. Por ejemplo, en la Figura 1b todos los genes identificados pertenecen a los de las interleucinas. En otro cluster no mostrado, las muestras estudiadas reflejan una sobreexpresión de los genes característicos de células epiteliales, no así en los genes pertenecientes a estroma; haciendo notar que las células estudiadas son de origen epitelial y que las biopsias están

compuestas en su mayoría por células epiteliales.¹⁸ Así, este cluster está constituido por componentes complejos adherentes, moléculas de adhesión célula-célula, e intercambiadores de sodio.¹⁹

En la Figura 1c, se muestra un cluster específico para genes involucrados en metástasis, como las metaloproteasas, las integrinas involucradas en la arquitectura celular así como inhibidores de proteasas.

Debido a que uno de los propósitos de este trabajo era encontrar genes que pudieran estar asociados al CaCU, los 15 genes que estuvieran mayormente sobreexpresados en todas las muestras y los 15 genes suprimidos en todas las muestras se seleccionaron para elaborar una tabla basada en aquellos genes con sobreexpresión o subexpresión >4 veces con respecto al normal. Tabla 1.

Expresión de procesos celulares

Con el fin de estudiar la variación de los niveles de ARNm en diferentes rutas metabólicas, utilizamos la base de datos de Gene Ontology para identificar grupos de genes presentes en nuestras microhileras que estuvieran involucrados en procesos celulares específicos. Estas variaciones se presentan en una gráfica de barras (Figura 2). En general puede observarse que todos los procesos celulares mostrados, se encuentran preferencialmente sobreexpresados, siendo los más afectados los procesos de ciclo celular y los pertenecientes al grupo de metabolismo de fármacos. Cabe mencionar que la línea celular HeLa muestra un patrón de sobreexpresión muy diferente a las otras líneas celulares con VPH18. El resto de las rutas presenta cambios muy discretos.

Confirmación de la sobreexpresión de genes en CaCU por RT-PCR

Para validar la expresión diferencial del gen IL6, previamente identificado como posible marcador de CaCU y MMP10 como un nuevo posible gen marcador, ambos obtenidos a partir del cluster jerárquico, se utilizaron dos diferentes métodos de análisis. El primer método utilizado fue la base de datos "Gene to Tag" que realiza la imagen del experimento de tipo northern blot virtual (<http://www.ncbi.nlm.nih.gov/Unigene/>). En esta aplicación, se hibridaron virtualmente cada uno de los dos transcritos estudiados (IL6 y MMP10) contra todas las librerías de expresión de líneas celulares y tejidos existentes en la base de datos. De manera interesante, los dos genes hibridaron y mostraron sobreexpresión en los tejidos de carcinoma de colon, adenocarcinoma de estómago y carcinoma de mama (resultados no mostrados). Con la información obtenida del experimento anterior se procedió a realizar experimentos de RT-PCR en las células HeLa, CaLO, INBL, SiHa, CaSki, en 5 tejidos normales VPH(-) y en 5 biopsias tumorales del cervix. Los resultados de estos experimentos mostraron que hay sobreexpresión de las líneas celulares y de las muestras tumorales con respecto al epitelio normal (Figura 3).

Discusión

Con el fin de identificar marcadores potenciales y alteraciones moleculares particulares del cáncer cérvico uterino, utilizamos microhileras de ADNc para determinar los niveles de expresión de más de 8,000 genes en cinco líneas celulares derivadas de CaCU, 3 biopsias de CaCU positivas para VHP16 y tres tejidos cervicales normales negativos para VPH.

Como se esperaba, el dendrograma generado por el agrupamiento jerárquico y los niveles de expresión génica de las muestras, dieron como resultado la identificación de dos subgrupos de genes correspondientes al tipo viral presente en las muestras. Otros genes identificados como sobreexpresados (datos no mostrados), son la inosina monofosfato deshidrogenasa, el factor de necrosis tumoral (TNF) y la metiltransferasa 1. Estos sólo se

encontraron sobreexpresados en las muestras con VPH16, por lo que es posible que la expresión de estos y otros genes adicionales permitiera determinar las diferencias específicas entre los tipos virales 16 y 18. Sin embargo, estos resultados son aún preliminares y están pendientes estudios adicionales que permitan establecer su significado funcional.

Con el fin de identificar las diferencias en expresión de procesos celulares específicos de las muestras estudiadas, se identificaron algunos grupos conformados por genes característicos de rutas metabólicas y utilizando una gráfica de barras se mostró los niveles de expresión (sobreexpresión o supresión) de las muestras estudiadas con respecto a su contraparte normal este análisis se logró utilizando la base de GeneOntology. El Gene Ontology Consortium (GO) es un proyecto internacional cuyos resultados están accesibles en la red. Podemos mencionar entre otros aspectos que la base de datos GO muestra un grupo de términos bien definidos y relaciones entre ellos con lo cual podemos interpretar el papel de un gen en particular, producto de un gen o un grupo de productos de genes involucrados en la patogénesis. El análisis GO organiza a los genes en tres ontologías separadas que comprenden procesos biológicos, componentes celulares y funciones moleculares. Una de los factores más importantes es que agrupa diversas bases de datos. Su estructura permite que los genes tengan asignados propiedades en diferentes niveles, dependiendo de cuanto se sabe sobre el producto de ese gen. De ahí que actualmente sea una de las herramientas más importantes para el análisis de datos obtenidos por microhileras.²⁰

El grupo de genes propios de células epiteliales conformado por: Periplakina, Desmoplakina, Junction plakoglobina, y Claudina 4 entre otros; no presenta cambios en los niveles de expresión en las diferentes muestras, sugiriendo que estos transcritos se originaron en células tumorales con características epiteliales similares. El grupo de genes de proliferación y de ciclo celular básicamente están compuestos por productos que son necesarios para la progresión a través del ciclo celular, para procesamiento de ARN y maquinaria de traducción y marcadores para identificar células en proliferación.²¹ En el grupo de proliferación los transcritos se expresan de manera semejante entre las muestras infectadas con los tipos virales 16 y 18, mientras que los de ciclo celular muestran un alto grado de sobreexpresión en las muestras VPH16 y una supresión en las VPH18; excepto en CaLO que muestra un patrón similar a VPH16. Esto ocurre también en el grupo de "metabolismo de fármacos", que esta conformado principalmente por genes que actúan en la regulación del estado "redox" en las células y en transportadores de diferentes compuestos; en este grupo el aumento en la abundancia relativa de estos transcritos en las células VPH16 y en CaLO puede estar indicando una selección para la resistencia a fármacos.²² Este resultado puede sugerir que los genes de respuesta a fármacos son activados en presencia del VPH16 en las pacientes.

El grupo denominado de componente tumoral, esta compuesto por genes tales como: Homólogo de homeobox 2, proteína de tumor D52 y la H2A miembro L, que son genes que han sido frecuentemente mencionados de tener alterada su expresión en diversos tipos de tumores y que el GO les ha asignado esta categoría.

Como se esperaba, en el grupo de transcritos que el programa GeneOntology clasificó como genes pertenecientes procesos estromales tales como: Proteína de membrana epitelial 1, Trombospondina 1, Calponina 3, caldesmonina 1 y el Colágeno tipo IV alfa 1, entre otros; los tejidos tumorales muestran un incremento en su expresión probablemente debido al componente de estroma presente en el tumor. Sin embargo, SiHa y CaSki presentaron el mismo incremento, esto pudiera explicarse debido la presencia de transcritos de un gran grupo de genes de colágeno.

El grupo de "interferón" y algunos genes asociados que son regulados por estos, indican patrones de expresión diferentes entre todas las muestras sugiriendo variación en la

respuesta de la ruta de interferones. Nuestros resultados apoyan a resultados previamente publicados.²³

Por otro lado, con el propósito de identificar posibles marcadores moleculares de esta neoplasia, se analizaron genes que estuvieran sobreexpresados en todas las muestras independientemente del tipo viral presente. Un gran número de esta clase de genes son los ribosomales, cuya sobreexpresión en los tumores ya ha sido previamente reportada y puede reflejar un aumento del metabolismo en células malignas, además de que esta sobreexpresión se presenta en tumores bien diferenciados. Los resultados de los experimentos de expresión mediante RT-PCRs indican y corroboran los niveles de sobreexpresión de IL6 y MMP10 en las líneas celulares y las biopsias previamente mostrados en el experimento de microhilera.

Como se esperaba por reportes previos, uno de los genes sobreexpresados fue el de la IL6; este gen codifica una citocina que, además de regular algunos aspectos de la respuesta inmune, tiene actividad inflamatoria²⁴ y es el principal factor autocrino de crecimiento de algunas neoplasias hematológicas, especialmente aquellas derivadas de linfocitos B.²⁵ Además, se ha sugerido tener una función de mediador central en la inflamación de infecciones del tracto genital femenino.²⁶ Es posible que en CaCU funcione también como factor de crecimiento, lo cual concuerda con el hecho de que algunas líneas celulares de esta neoplasia expresan el receptor de IL6, aunque la relación funcional no se ha probado formalmente, además de que en cuanto a algunas de las líneas celulares utilizadas en este estudio aún no se ha examinado la expresión de IL6R. En el cérvix uterino existen muchos tipos de células capaces de elevar su nivel de expresión como son los linfocitos, macrófagos, fibroblastos, queratinocitos, células endoteliales y células tumorales.²⁷ Además, se ha comprobado que los niveles de IL6 se incrementan en las secreciones cérvico-vaginales de las pacientes con CaCU y su producción incrementa con respecto a la severidad de la neoplasia.²⁸ Por otro lado, existen reportes que indican que esta interleucina principalmente es producida en células tumorales y que además puede funcionar como un factor de crecimiento en líneas celulares cervicales.²⁹ De ahí que la sobreexpresión de este gen valida los experimentos realizados. En cuanto a la sobreexpresión del gen MMP10 encontrada en este trabajo, existen diferentes reportes que relacionan a las MMPs con algunos cánceres,³⁰ sin embargo actualmente no existe ningún reporte que la asocie con CaCU mostrando que este gen puede ser relevante en la carcinogénesis cervical y además indicando que algunos estudios deberían ser dirigidos en esta dirección.

La infección por VPH muestra patrones característicos de expresión con marcadas diferencias con respecto a los tejidos VPH negativos. El análisis mostró diferencias en la expresión de los genes IL6 y MMP10 en los tejidos VPH(+) y VPH(-). Estos resultados podrían contribuir a conocer parcialmente (una fotografía molecular) del proceso invasor del cuello uterino, indicando además, que las alteraciones en citocinas y componentes de matriz extracelular están estrechamente relacionados a la carcinogénesis cervical influenciada por VPHs-AR.

AGRADECIMIENTOS

Al Dr. Alberto Monroy (FES-Zaragoza, UNAM) por proporcionar las líneas celulares CaLO e INBL, al Dr. D. Arenas (CMN SXXI-IMSS) por su apoyo. Este trabajo fue parcialmente apoyado por No. 34686-M y 7114 (Fondos Sectoriales, CONACyT-México). GV, PP, AH, CP fueron becarios de CONACYT, DGEP-UNAM e IMSS.

REFERENCIAS

1. World Health Statistics, 1987-1990. *World Health Organization*.
2. Programa de prevención y control del cáncer cérvico-uterino. *Secretaría de Salud, México* 1992.
3. Brisson J, Morin C, Fortier M, et al. Risk factors for cervical intraepithelial neoplasia: differences between low- and high-grade lesions. *Am J Epidemiol* 1994; 140: 700-10.
4. Castellsague X, Bosch F, Muñoz N. Environmental co-factors in HPV carcinogenesis. *Virus Research* 2002; 89: 191-99.
5. Schiffman M, Bauer H, Hoover R. Epidemiologic evidence showing that human papillomavirus infection causes most cervical intraepithelial neoplasia. *J Natl Cancer Inst* 1993; 85: 958-64.
6. Muñoz N. Human papillomavirus and cancer: the epidemiological evidence. *J Clin Virol* 2000; (1-2): 1-5.
7. Walboomers J, Jacobs M, Manos M, Bosch F, Kummer K, et al. Human Papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J Pathol* 1999; 189: 12-9.
8. Bosch F, Lorincz A, Muñoz N, Meijer C, Shah K. The causal relation between human papillomavirus and cervical cancer. *J Clin Pathol* 2002; 55: 244-65.
9. Lorincz A, Temple G, Kurman R, Jenson A, Lancaster W. Oncogenic association of specific human papilloma virus types with cervical neoplasia. *J Int Cancer Inst* 1987; 79: 671-77.
10. Choo KB, Chen CM, Han CP, Cheng WT, Au LC. Molecular analysis of cellular loci disrupted by papillomavirus 16 integration in cervical cancer: frequent viral integration in topologically destabilized and transcriptionally active chromosomal regions. *J Med Virol* 1996; 49: 15-22.
11. Münger K, Werness BA, Dyson N, Phelps WC, Howley PM. Complex formation of human papillomavirus E7 proteins with the retinoblastoma tumor suppressor gene product. *EMBO J* 1989; 8: 4099-4105.
12. Scheffner M, Werness BA, Huibregtse JM, Levine AJ, Howley PM. The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. *Cell* 1990; 63: 1129-36.
13. Park T, Fujiwara H, Wright T. molecular biology of cervical cancer and its precursors. *Cancer* 1995; 76(Suppl): 1902-13.
14. Milde-Langosch K, Riethdorf S. Role of cell-cycle regulatory proteins in gynecological cancer. *J Cell Physiol* 2003; 196: 224-44.
15. Shyu JS, Chen CJ, Chiu CC, Huang SC, Harn HJ. Correlation of human papillomavirus 16 and 18 with cervical neoplasia in histological typing and clinical stage in Taiwan: an in-situ polymerase chain reaction approach. *J Surg Oncol*. 2001; 78:101-9.
16. Caceres-Cortes J, Alvarado-Moreno J, Waga K, Rangel-Corona R, Monroy-Garcia A. et al. Implication of tyrosine kinase receptor and steel factor in cell density-dependent growth in cervical cancers and leukemias. *Cancer Res* 2001; 61:6281-9.
17. Eisen M, Spellman P, Brown P, Botstein D: Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci* 1998, 95:14863-68.
18. Davies J, Garrod D: Molecular aspects of the epithelial phenotype. *Bioessays* 1997, 19: 699-704.
19. Ross D, Scherf U, Eisen M, Perou C, Rees C, Spellman P, Iyer V., Jeffrey S, Rijn M, Walthman M, Pergamenschikov A, Lee J, Lakhari D, Shalon D, Myers T,

- Weinstein J, Botstein D, Brown P. Systematic variation in gene expression patterns in human cancer cell lines. *Nat genetics* 2000, 24:227-35.
20. Ahn WS, Bae SM, Lee JM, Namkoong SE, Han SJ, Cho YL, Nam GH, Seo JS, Kim CK, Kim YW. Searching for pathogenic gene functions to cervical cancer. *Gynecol Oncol*. 2004, 93:41-8.
 21. Welsh J, Zarrinkar P, Sapinoso L, Kern S, Behling CA. Analysis of gene expression in normal and neoplastic ovarian tissue samples identifies candidate molecular markers of epithelial ovarian cancer. *Proc Natl Acad Sci* 2001, 3: 1176-81.
 22. Moscow J, Schenider E, Ivy S, Cowan K. Multidrug resistance *Chemoter Biol Response Modif* 1997, 17: 139-77.
 23. Nguyen H, Hiscott J, Pitha P. The growing family of interferon regulatory factors. *Cytokine Growth Factor Rev* 1997, 8: 293-312.
 24. Kishimoto T. The biology of Interleukin-6. *Blood* 1998, 74: 1-10
 25. Hoffman-Liebermann B, Liebermann DA. Suppression of c-myc and c-myb is tightly linked to terminal differentiation induce by IL6 or LIF and not growth and no inhibition by myeloid leukemia cells. *Oncogene* 1991, 6: 903-9.
 26. Rochter H, Holley R, Andrews W, Owen J, Miller K. The association of interleukin 6 with clinical and laboratory parameters of acute pelvic inflammatory disease. *Am J Obstet Gynecol* 1999, 181: 940-44.
 27. Espevikn T, Wange S, Faxvaag A, Shalaby M. Regulation of interleukin-2 and interleukin-6 production from T-cells: involvement of interleukin-1 beta and transforming growth factor beta. *Cell Immunol* 1990, 126: 4756-61.
 28. Tjong M, van der Vange N, Kate F, Tjong SP, Ter Shegget J. Increased IL-6 and IL-8 levels in cervicovaginal secretions of patients with cervical cancer. *Gynecol Oncol* 1999, 73: 285-91.
 29. Eustace D, Han X, Gooding R, Riches P, Hyderman E. Interleukin-6 functions as an autocrine growth factor in cervical carcinoma in vitro. *Gynecol Oncol* 1993, 50: 15-9.
 30. Gaiotto MA, Focchi J, Ribalta JL, Stavale JN. Comparative study of MMP-2 (matrix metalloproteinase 2) immune expression in normal uterine cervix, intraepithelial neoplasias, and squamous cells cervical carcinoma. *Am J Obstet Gynecol* 2004; 190: 1278-82.

TABLAS

Tabla 1. Genes mayormente sobreexpresados y suprimidos en el Cáncer Cérico Uterino.

Genes sobreexpresados

Nombre del gen	Unigene ID	tumoral/normal
CDP- diaciglicerol sintetasa	Hs. 152981	4.93
periplakina	Hs.74304	4.87
Proteína Ribosomal S6 cinasa 60 kD, polipeptido 2	Hs. 10081	4.62
Proteína Ribosomal S6 cinasa 70 kD, polipeptido 4	Hs. 105584	4.41
Proteína KIAA0143	Hs. 84087	4.26
2' 5' oligoadenil sintetasa 1	Hs. 82386	4.85
Inhibidor de metaloproteasa 1	Hs. 5831	4.19
Matriz metaloproteasa 9	Hs. 151738	4.48
Integrina alfa 8	Hs. 91296	4.45
Catepsina L2	Hs. 87417	4.44
Matriz metaloproteasa 2	Hs. 367877	4.33
Catepsina B	Hs. 297939	4.28
Proteasa alfa inhibidora de serinas	Hs. 173594	5.62
Matriz metaloproteasa 10	Hs. 2258	4.58
Interleucina 6	Hs. 512234	5.04

Genes suprimidos

Nombre del gen	Unigene ID	tumoral/normal
Miembro C1 de la familia de aldo-keto reductasa	Hs. 306098	0.01
Kiruneninasa	Hs. 169139	0.01
Receptor nuclear de la proteína 1	Hs. 155017	0.01
Grancalcina	Hs. 79381	0.01
Caderina 5 tipo 2, caderina VE	Hs. 76206	0.01
Caderina 6 tipo 2, caderina 2K	Hs. 32963	0.01
Sialoproteína de unión a integrina	Hs. 121552	0.01
Proteasa inhibidora de serina B	Hs. 41072	0.01
Proteasa inhibidora de serina A	Hs. 41072	0.01
Neutrofilo elastasa 2	Hs. 99863	0.01
Receptor de Interleucina 18, 2	Hs 159301	0.01
Homeobox 1B	Hs. 34871	0.01
Proteína Homeobox 5	Hs. 2551	0.01
Proteína Ribosomal S6 cinasa, 90 kD, polipeptido 3	Hs. 173965	0.01
Gen de fusión a ALL 1 del cromosoma 1q	Hs. 75823	0.01

Figura 1. Agrupamientos Jerárquicos a) Patrón de expresión de 8000 genes en 8 muestras. Las columnas indican a las muestras y los renglones a los genes. Cada celda en la matriz representa el nivel de expresión de un transcrito individual comparado con la mediana de ese transcrito en las muestras normales; se indican los niveles de expresión hacia arriba y hacia abajo del normal con rojo y verde respectivamente para ese gen a lo largo de todas las muestras. La saturación del color es proporcional a la magnitud de la diferencia del normal. b) Agrupamiento formado con genes que codifican para Interleucinas o bien que son regulados por estas. c) Cluster enriquecido con genes implicados en metástasis. Las flechas azules señalan a los genes que fueron utilizados para la validación de la expresión por RT-PCR.

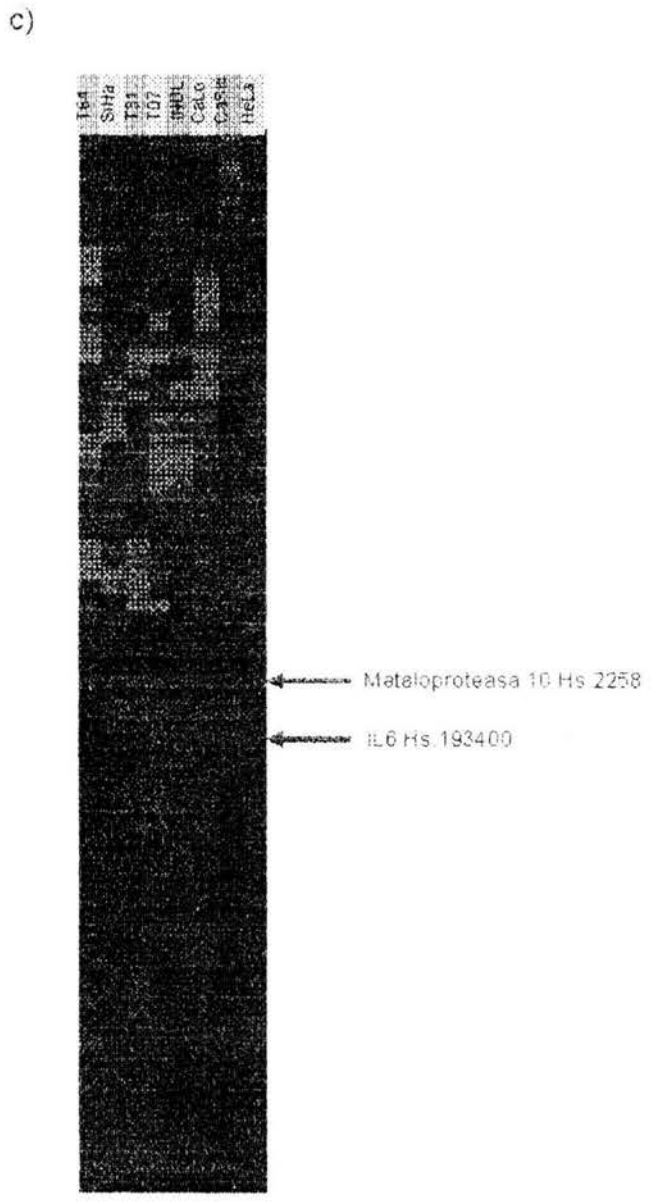
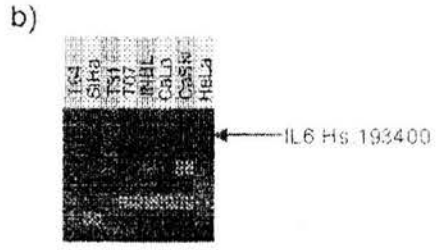
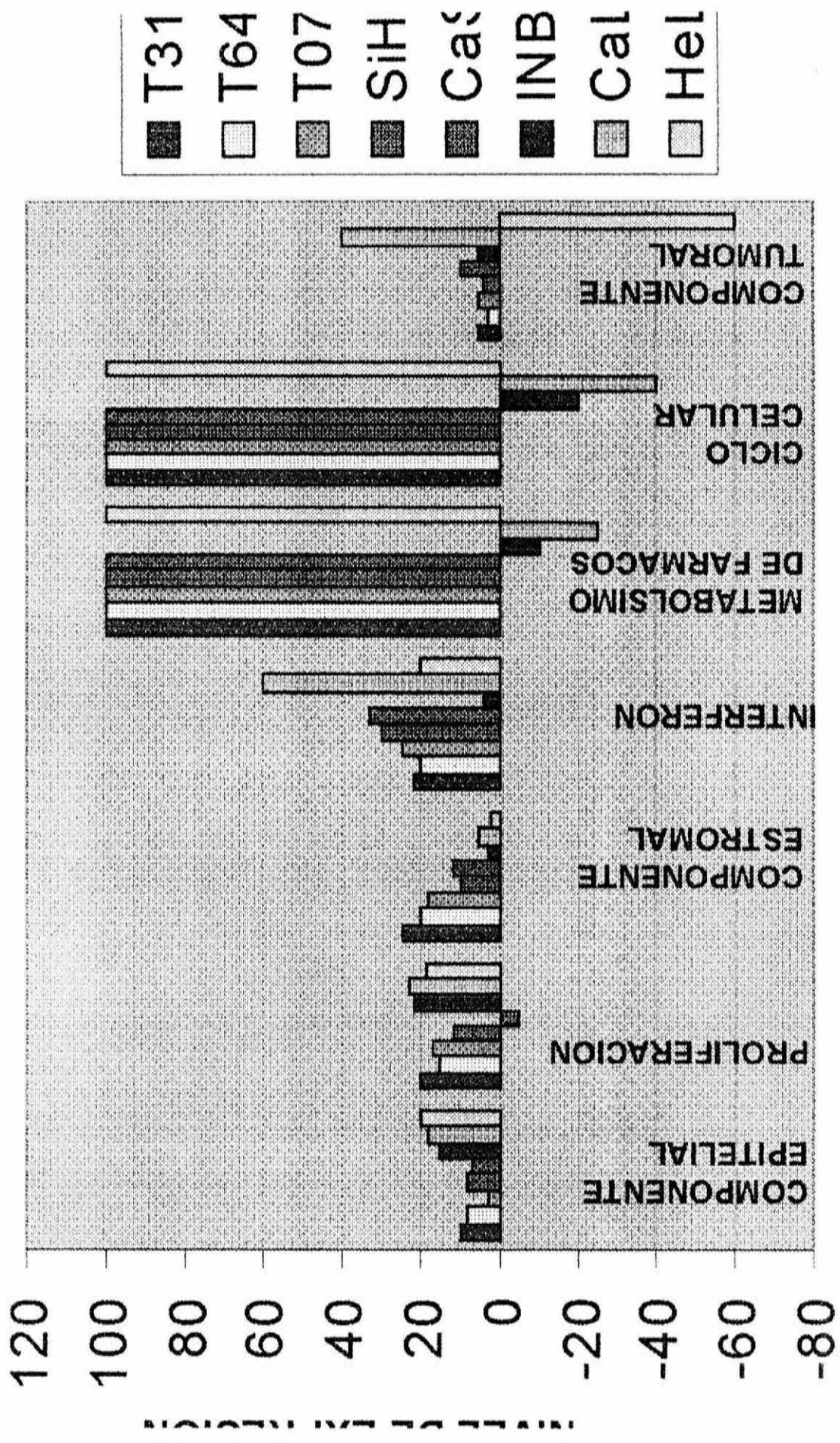


Figura 2. Expresión génica de genes implicados en procesos celulares en Cáncer Cérvico Uterino. La gráfica muestra los niveles de incremento o decremento en la expresión génica de las muestras tumorales; suma de medianas de todas las muestras VPH(+) para cada gen que conforma la ruta metabólica comparada con la mediana de las muestras normales VPH(-). El valor de cero indica que el nivel de expresión no varía con respecto al normal.

RUTAS METABOLICAS



RUTAS

Figura 3. Corroboración de la expresión de IL6 y MMP10 en líneas celulares y tumores cervicales mediante RT-PCR. Cien ng de ARN total fueron utilizados para la RT-PCR. Las muestras utilizadas fueron: 1) control negativo sin RNA, 2) HeLa, 3) INBL, 4) CaLO, 5) SiHa, 6) CaSki, 7) muestra tumoral #T04, 8) muestra tumoral #T12, 9) muestra tumoral #T24, 10) muestra tumoral #T43, 11) muestra tumoral #T45, 12) muestra normal VPH- #N07, 13) muestra normal VPH- #N11, 14) muestra normal VPH- #N20, 15) muestra normal VPH- #N23, 16) muestra normal VPH- #N25. El marcador de peso molecular utilizado fue de 100pb.

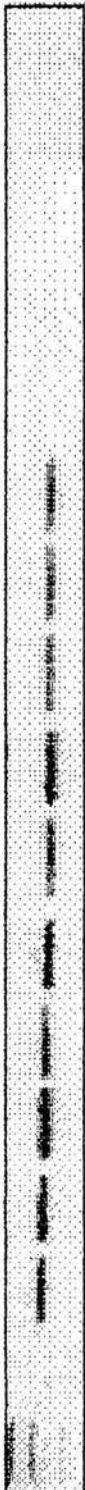
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16

621



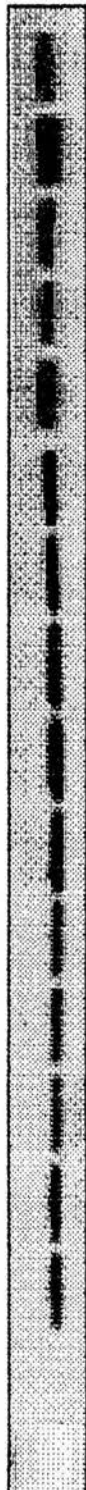
IL6

509



MMP10

243



β -actina

Dear Ms. Guelaguetza Vazquez,

Re: IJGC-2004-00296.R1*-Identification of differentially expressed genes in HPV positive and HPV negative cell lines and tissues

I am pleased to confirm that your paper has been accepted for publication in the International Journal of Gynecological Cancer. You will be receiving the proofs via e-mail two to three weeks prior to publication. The proofs must be returned on schedule or the paper will be delayed in printing.

The Editors thank you very much for your cooperation in the review process. We look forward to more papers from you and your colleagues.

If you are interested in joining the International Gynecologic Cancer Society or the European Society of Gynaecological Oncology, the parent organizations for this journal, the applications are available at <http://www.igcs.org> or <http://www.esgo.org>. Please feel free to email me for more information.

Sincerely yours,

John Kavanagh, M.D.

Editor-in-Chief

International Journal of Gynecological Cancer

Identification of differentially expressed genes in HPV positive and HPV negative cell lines and tissues

G. VÁZQUEZ-ORTÍZ*, S. PEÑUELAS**, V. NOÉ**, C.J. CIUDAD**, P. PIÑA-SANCHEZ*, A. HIDALGO*, C. PÉREZ-PLASCENCIA*, R. LOPEZ-ROMERO*, N. WENTZENSEN*** & M. SALCEDO*.

**Oncology Genomics Laboratory, Oncology Diseases Research Unit, Oncology Hospital, National Medical Center Siglo XXI-IMSS, Mexico.*

***Department of Biochemistry and Molecular Biology, School of Pharmacy, University of Barcelona, Spain.*

****Institute of Molecular Pathology, Department of Pathology, Heidelberg, Germany.*

Running Title: *Expressed genes in HPV infection*

Corresponding Author:

Dr. Mauricio Salcedo

Unidad de Investigación Médica en Enfermedades Oncológicas

Hospital de Oncología

Centro Médico Nacional Siglo XXI-IMSS

Av. Cuauhtemoc 330. Col Doctores.

México D. F. C. P. 06720

Phone: +55 56276900 Ext: 22708 Fax: +55 57610952

Email: maosal89@yahoo.com, guelav@yahoo.com

Abstract

Human Papillomavirus (HPV) infection is the most important risk factor related to the malignant transformation of the human cervical epithelium. High-risk HPVs promote the development of cervical cancer *in vivo* and can immortalize cervical epithelial cells *in vitro*. The HPV 16 and 18 genotypes are the most prevalent in cervical cancers. To identify novel cellular genes that could potentially act as predictive molecular markers for human cervical cancer we employed cDNA arrays to examine the differential gene expression patterns of 6 cervical cell lines, 6 cervical tumor HPV positive samples and 3 normal HPV negative samples. Among the differentially expressed genes in HPV⁺ and HPV⁻ samples, four genes were validated by semi-quantitative RT-PCR. One of the studied genes Lamc2 has been previously implicated in HPV-related oncogenesis *in vivo*. We also studied three novel genes not previously related to cervical cancer (Claudin 1, NDUFB6 and Ipo7). The progress in the identification of characteristic molecular changes in cervical cells due to HPV infection can lead to a better understanding of cervical carcinogenesis. In a long term, it may provide a molecular framework to help to elucidate the biological differences of mechanisms associated with differential HPV type and aggressive prognosis.

KEYWORDS: HPV, cDNA array, expression, cervical cancer.

Introduction

High risk human papillomavirus (Hr-HPVs) infection is the most important etiological factor of cervical cancer (CC) development ⁽¹⁾. Epidemiological and molecular studies identified Hr-HPVs present in more than 99% of all cervical tumors, being HPV16 and HPV18 types the most frequently found in invasive cervical carcinomas in many populations ⁽²⁾. HPV16 and HPV18 seem to be acting differently at some level, based on clinical and experimental data suggesting that HPV18 infection is more aggressive for the patient than with HPV16 ⁽³⁻⁶⁾. Several attempts to determine some of the specific differences between HPV16 and HPV18 infections have been performed.

Previous results from our group regarding cytogenetic abnormalities in cervical tumors analyzed by Comparative Genomic Hybridization (CGH) detected a pattern of chromosomal imbalances that might be specific for HPV18 positive samples ⁽⁷⁻⁸⁾. The role of these genetic changes between these HPV types and their relation with the clinical behaviour is still under evaluation.

Hr-HPVs are usually integrated into the host cell genome in infected cells, concomitantly over-expressing the E6 and E7 viral oncogenes ⁽⁹⁻¹⁰⁾. No preferential integration sites could be identified for these HPV types ⁽¹¹⁾, showing that in this sense there are not specific target-directed insertional recombination between the viral and the cellular DNA.

The main carcinogenic property of E6 and E7 of both Hr-HPV 16 and 18 is their capacity to target pRB and p53 tumor suppressor proteins ⁽¹²⁻¹³⁾. However, their interaction with several other important cell-cycle components ⁽¹⁴⁾ was shown to be important in deregulating several cellular processes. These molecular events have not been fully studied.

The search for potential marker genes in cervical cancer is currently under study ⁽¹⁵⁻¹⁶⁾. However, a comparative screening between HPV⁺ and HPV⁻ samples, to determine genes that might be responsible for its oncogenic potential, has not been carried out.

With the aim to define potential biological markers that might detect the presence of HPV, in this study, we used cDNA arrays to identify differentially expressed genes between HPV positive cell lines, HPV positive cervical cancer tissues and HPV⁻ cervical samples and cell lines.

Material and Methods

HeLa (HPV18), SiHa (HPV16), CaSki (HPV16) and C33 (HPV⁻) cell lines were grown in the recommended media including 10% fetal bovine serum. Cell lines CALO (derived from Invasive squamous carcinoma Stage IIB, HPV18) and INBL (derived from Invasive Squamous carcinoma IV A, HPV18), were established at the FES Zaragoza-UNAM, Mexico⁽¹⁷⁻¹⁹⁾. In a first step of our study, twenty normal cervical tissues from premenopausal women were obtained *post-mortem* at the Pathology Department of the Hospital General de Mexico, S. S. from women who died of non-gynecological related causes. Similar number of invasive lesions was collected through colposcopy guided biopsies of the tumor lesion of pre-menopausal women at the Dysplasias National Center at Mexico City. All the described procedures were evaluated and approved by the local committee of ethics of the Mexican Institute of Social Security; we also obtained written informed consent from the patients. Of the 40 collected samples only 6 normal (HPV negative and non-dysplastic lesions) and 7 invasive samples (HPV positive, invasive lesions with at least 70% of tumoral cells); met the required criteria. Unfortunately, HPV18 tissue samples were not obtained due to its extremely low frequency in this country. All tissue samples were divided in three sections, the central part was snapped frozen in liquid nitrogen and stored at -70°C until nucleic acid extraction, and the extremes from the biopsy were fixed with 70% ethanol overnight and paraffin embedded. Haematoxylin/eosin stained sections from the extremes of the biopsies were analyzed in order to confirm the presence of at least 70% tumor cells. Multiple step sections were made and classified according to the International Federation of Obstetrics and Gynecology (FIGO)⁽²⁰⁾. All the invasive samples were diagnosed as squamous lesions.

Both RNA and DNA were extracted from the tissue samples and the cervical cancer cell lines using the Trizol Reagent (Gibco BRL, Life Technologies, Grand Island NY USA) according to manufacturer's conditions. Each DNA sample was subjected to HPV typing by polymerase chain reaction (PCR) using the MY09/MY11 consensus primers and direct sequencing. To eliminate residual genomic DNA, total RNA was treated with RNase-free DNase I (Ambion Inc). Total RNA concentration of each sample was assessed by absorbency at 260 nm and the nucleic acid integrity was confirmed by agarose gel electrophoresis.

Messenger RNA isolation of each sample (HeLa, SiHa, CaSki, CALO, INBL, and C33), three HPV16⁺ tumor samples (T07, T31, T64) and three HPV⁻ normal samples (N03, N11, N22) was performed using the Polytract Isolation System II (Promega Madison, WI). mRNA was quantified using the VersaFluor Fluorometer (BioRad). Labeled cDNAs were obtained by RT-PCR from 500 ng of mRNA in the presence of (α -³²P) dATP (activity 3000Ci/mmol) using the Ambion Advantage System, following the manufacturer's conditions. Briefly, mRNAs were denatured at 70°C for 10 min and cDNAs were synthesized at 42°C using oligo-dT priming in a final volume of 30 μl . The unincorporated nucleotides were removed by centrifugation using the NucAway Spin Columns (Ambion Inc). The final incorporation of the radiolabel was approximately 1×10^5 cpm for each reaction. ULTRArray Advantage System array blots (Ambion Inc) which contains 8400 genes, were prehybridized at 68°C for at least 90 min before probe addition in Hybridization buffer. Finally, 1.5×10^7 cpm of each labeled cDNA was added to the buffer. Hybridization was performed at 60°C in a rolling bottle overnight. The arrays were washed twice with 2X SSC and 0.5 % SDS at 60°C for 30 min, followed by two stringent washes with 0.5X SSC, 0.5% SDS at the same temperature and time.

Hybridized arrays were scanned on a phosphorimager (STORM). The resulting images were analyzed with the ArrayVision software (Image Research Inc), and data from the

duplicate spots were normalized using Gene Spring software version 5.0 (Silicon Genetics). Values below 0.01 were set to 0.01. Each measurement was divided by the 50.0th percentile of all measurements in the corresponding control samples. A second analysis was performed using as samples: HPV⁻ normal samples, C33, HPV16⁺ (median of normalized expression levels of SiHa, CaSki and three HPV16⁺ invasive samples) and HPV18⁺ (median of normalized expression levels of HeLa, CALO and INBL cell lines) following the above mentioned methodology. Samples were normalized against the median of the control samples (normal HPV⁻ samples). When calculating the mean of the control measurements only those measurements were included that had at least a normalized value of 0.01. Each measurement for each gene in the experimental samples was divided by the median of that gene.

First, all the overexpressed genes (listed at Table 1) were submitted to the "Gene to Tag" database of virtual northern blot at NCBI (<http://www.ncbi.nlm.nih.gov/Unigene/>); these genes were hybridized against all the SAGE libraries available in this site. Claudin 1, Importin 7, Lamc2, NDUFB6 genes showed the highest expression in epithelial tissues; therefore, these were used to validate the expression results.

Two-hundred ng of total RNA from normal tissues, cell lines and tissue samples was reverse transcribed using the RT-PCR Access System (Promega, Madison, WI). Primers for the genes of interest were designed using Integrated DNA technologies Biotools software (<http://biotools.idtdna.com/gateway>) for the amplification of the following gene transcripts: *NADH Dehydrogenase (ubiquinone) 1 beta subcomplex 6 (NDUFB6)*, 328 bp TM=60°C (5'-GTACACTCCGGATGAGAACTG-3' and 5'-CTCCAGAATTGTATCACCAGGG-3'); *Importin7 (lpo7)*, 649 bp TM=60°C (5'-TTGGCCTGATCGAGAAACAGCA-3' and 5'-CTTGGAACATTGCCAGGGCTT-3'); *laminin gamma 2 (lamc2)*, 696 bp TM=56°C (5'-TGTTTAGCTCAGCCCAACGACT-3' and 5'-AGCAGCTGAACCCGTTATGACA-3'); *claudin1 (cldn1)*, 504 bp TM=55°C (5'-GTGGAGGATTTACTCCTATGCC-3' and 5'-CTTGGTGTGGGTAAGAGGT-3'), and β -actin : (5'-TGAAGTCTGACGTGGACATC-3' and GTTCGTTCTCATACTGCTCA-3') 243 bp, TM=55°C. The RT-PCR conditions were: to synthesize the first strand of cDNA 48°C for 45 min and 94°C for 2 min to denature template; and to synthesize second strand and DNA amplification: 94°C for 30 sec, (specific TM°C for each set of primers) for 1 min, and 68°C for 2 min for a total of 25 cycles, followed by a single step at 68°C for 7 min. The amplified products were visualized on 1.5% agarose gels stained with ethidium bromide, and images were captured using the Eagle eye system (Stratagene, La Jolla, Ca). The sizes of the products were determined by comparison with a 100 bp ladder marker (GeneRuler, MBI Fermentas). The intensities of NDUFB6, lpo7, Lamc2, Claudin 1 and β -actin RT-PCR products were analyzed using the MetaView analysing system (version 4.5 Universal Imaging Corp., USA). The relative concentrations of the mRNAs were expressed as the ratio of densitometric readings of the four genes related to β -actin mRNA. All values are presented as the mean of at least 4 determinations. Statistical comparisons among groups were analysed by a Student's t-test. A *P* value of <0.001 was considered as statistically significant.

For the immunohistochemistry (IHC), paraffin blocks of 15 specimens with CC diagnosis and HPV16 infected, were collected at random from the archives of the Pathology Department at Hospital de Oncologia, National Medical Center in Mexico City; while 10 of the normal cervical tissues, HPV- (*post-mortem* obtained) were paraffin embedded. Briefly, the tissue sections were deparaffinized and rehydrated in a graded ethanol series. Endogenous peroxidase inactivation was carried out by incubating the samples in 0.3% hydrogen peroxide in methanol for 30 min, and the slides were preincubated with DAKO Protein Block Serum free Cat X0909 (Carpinteria, Ca) for 30 min at room temperature to prevent unspecific immunoreaction. Excess was decanted and tissues were incubated with

the primary antibody as follows: Claudin 1 (Zymed, Cat. 717800, 1:100) and LamC2 (Chemicon, Cat. MAB19562, 1:100) at 4°C overnight. DAKO Envision System Peroxidase from DAKO (Carpinteria, Ca) was applied and the slides and then counterstained with hematoxylin. Our rough scoring was done as positive or negative staining based in the RT-PCR results. Staining result was considered positive whenever increased stained positive in the majority of the tumor cells was present as compared with staining of the control tissue consisting of histological normal cervical tissues.

Results

To study transcriptional differences between HPV⁺ and HPV⁻ samples, the gene expression profile of normalized HPV16 (median of normalized expression levels of SiHa, CaSki, and three HPV16⁺ invasive samples), and HPV18 (median of normalized expression levels of HeLa, CALO and INBL cell lines) were compared to the normalized median of C33 and 3 normal HPV⁻ tissue samples using the Gene Spring software. We found approximately 100 genes whose expression was at least four-fold higher in HPV positive samples. Although the genes that were identified in this report might not represent the whole picture of cervical carcinogenesis, it was reassuring to confirm that over 90% of the genes identified by this approach in tissues with squamous cell carcinoma of the cervix were similarly identified in cervical carcinoma cell lines by SAGE (data not shown). To eliminate false positive results, out of these 100 genes we chose only those transcripts showing a ten fold expression or higher in any HPV positive tissue compared to basal expression in C33 and normal tissues (Table 1, Figure 1). From the 20 outlier genes, four genes were clearly identified: NADH Dehydrogenase1 beta subcomplex 6 (NDUFB6), which showed a median of 16 fold and 8 fold of overexpression in HPV16 and HPV18 cell lines, respectively; Importin7 (Ipo7) with a median of 8 fold in HPV16⁺ and 4 fold in HPV18⁺, Claudin 1 expressing a median of 12 fold in HPV16⁺ and 10 fold in HPV18; and Laminin gamma 2 (Iamc2) which showed a median of 12 fold overexpression in HPV16⁺ and a basal expression in HPV18⁺. We also observed a set of cellular genes that showed significant decreases in expression; however, these were not characterized in the present study.

Figure 1 shows the changes in expression patterns between normal HPV- samples, C33, HPV16 and HPV18 in all studied specimens. Red lines show the studied genes: a) NDUFB6, b) Ipo7, c) Claudin 1, d) Iamc2. It is worth noting that HPV16 positive samples presented more overexpressed genes than HPV18 positive samples. However, a specific group of genes was only overexpressed in HPV18 (Unpublished observations). Table 1 shows the 20 most overexpressed and the 10 more suppressed genes in HPV positive tissues.

To confirm the differences in expression observed in the array experiments, total RNA of some of the studied samples as well as additional samples, were analyzed by semi-quantitative RT-PCR assays. A representative picture of the amplified products is shown in Figure 2. Computer aided densitometric analysis of the amplified bands showed that Claudin 1 presented overexpression in all HPV⁺ samples ($P < 0.001$, Figure 2A); similar results were recently reported ⁽²¹⁾. Importin 7 showed an important overexpression in HPV16⁺ specimens and a slight overexpression in HPV18 cell lines ($P < 0.001$, Figure 2B). Iamc2 expression increased dramatically in HPV16 cell lines; and it was also overexpressed in HPV16 human cancer specimens ($P < 0.001$, Figure 2C); whereas the NDUFB6 gene (Figure 2D) presented overexpression in the cervical cancer cell lines ($P < 0.0001$) compared with normal basal expression in HPV16 tissue samples.

IHC was successful in 93.33% of the 15 CC samples. The unsuccessful case consisted of one unrepresentative sample. Lamc2 and Claudin 1 immunostaining was positive in 93% of CC cases. All samples consisting of histologically normal cervical tissue did not show any stain for both proteins. (Table 2; Fig. 3). These results were in accordance with RT-PCR results (Fig 2)

Discussion

To identify HPV related tumor-specific molecular alterations in human cervical cancer, we profiled differentially expressed genes in cervical cancer by comparing HPV16 and HPV18 positive tissues and cell lines with HPV⁻ normal samples including the HPV⁻ C33 cell line. The 20 cDNA clones that are consistently upregulated in HPV⁺ human cervical cancer correspond to known genes that can be classified into different functional categories: developmental and physiological processes, intracellular signalling, extracellular functions, membrane and cell adhesion related, whereas some of them are either oncogenes, cyclins or chaperons. In this work we mainly focused on the most overexpressed genes.

We detected some genes that are involved in cell growth and maintenance like NM-003049 and NM-014252 which are solute carriers, NM-002857 which participates in peroxisome biogenesis⁽²²⁾ and NM-002493 (NDUFB6 or Complex I NADH:ubiquinone oxidoreductase).

NDUFB6 is the first multisubunit enzyme complex of the mitochondrial respiratory chain. Some mitochondrial genes were reported to be overexpressed in thyroid carcinomas⁽²³⁾, and these genes present mutations in oral⁽²⁴⁾ and prostatic carcinomas⁽²⁶⁾. In this study, this gene was overexpressed only in the cell lines not in tissues. Given that cell lines present certain characteristics due to its immortalized state⁽²⁶⁾, NDUFB6 upregulation in cell lines might be related to this situation.

Regarding cell death, the Death Associated Protein (DAP) was upregulated upon HPV infection in agreement with a previous report⁽²⁷⁾.

Another metabolic pathway that is also upregulated in the presence of HPV is the intracellular signalling process. Four genes within this process were found to be overexpressed in our study. NM-020184 is a novel gene related to cervical cancer that requires further study. A recent report using computerized data mining and the Cancer Genome Anatomy Project shows that the NM-003785 transcript appears to be upregulated in prostate and cervical cancer⁽²⁸⁾. On the other hand, semiquantitative-RT PCR results demonstrate the upregulation of NM-006391 (Ipo7) and U25029 (GR) in cervical cancer. In this context, it is known that the transport of proteins and large RNAs through the nuclear pore complexes is a regulated process, and that the import of proteins with a nuclear localization signal is performed by the importin-alpha/beta complex⁽²⁹⁾ including Importin 7. According to the present work, this gene was upregulated 3 fold in HPV⁺ specimens. Ipo7 promoter has a GR (glucocorticoid receptor) binding site. GR is cytoplasmic and enters the nucleus after hormone binding by the interaction with the import receptor Ipo7⁽³⁰⁾. GR is a transcription factor that, modulates the expression of glucocorticoid-responsive genes by binding to a specific glucocorticoid response element (GRE)⁽³¹⁾. GR influences HPV early gene expression by the interaction with the GRE present in HPV18 regulatory region and the three GREs present in HPV16⁽³²⁻³⁶⁾. As expected, the GR gene is more overexpressed in HPV16⁺ (median of 4 fold) than in HPV18 cell lines (median of 2 fold) just like the Ipo7 gene (Table 1). Besides the increment of HPV expression, GR activation results in diverse physiological effects depending on the cell type. For example, GR activation in breast

cancer cells, regulates survival signalling through direct transactivation of genes that encode proteins that decrease susceptibility to apoptosis⁽³⁶⁾ showing that this pathway might be relevant to carcinogenesis.

Related to the cell adhesion process, Claudin 1 was overexpressed. The claudin family is the major constituent of the tight junction (TJ) complex. TJ is the structure that controls cell-to-cell adhesion and consequently, regulates paracellular and transcellular transport of solutes across human epithelial and endothelial cells⁽³⁷⁾. Aberrations of tight junction function are an early and key event during the vascular spread of cancer and inflammation⁽³⁸⁾. Claudin 1 (*clnd1*) was described to be downregulated in breast cancer⁽³⁹⁻⁴⁰⁾. This gene is upregulated by β -catenin⁽⁴¹⁾ that belongs to the Wnt signalling pathway, commonly altered in colon and gastric cancers⁽⁴²⁾. Recent results from our group show that genes in the Wnt signaling pathway are overexpressed in cervical cancer, as an event associated to HPV infection (submitted 2004). This overexpression might cause an accumulation of β -catenin in the nucleus, activating many transcription factors like Claudin 1. In the present study, high levels (11 fold) of Claudin 1 mRNA were found in all HPV⁺ samples tested, in keeping with the results obtained by RT-PCR and also with IHC, supporting that Claudin 1 might have a role in cervical cancer. All invasive cancers showed positive immunoreaction for Claudin 1 protein located in the cytoplasm of the cells. It is worth mentioning that in contrast to invasive tissues in the normal cervical tissues this protein was absent. The presence of this gene among the identified markers validates the cDNA microarray approach used in this study.

One cellular anchoring protein mainly found in the basement membrane is Laminin 5. Laminin 5 is composed by three polypeptidic chains that are the products of three distinct genes: *Lama3*, *Lamb3*, and *Lamc2*. It is already known that, in order to metastasize and grow, cancer cells must invade and destroy the basement membrane⁽⁴³⁾ requiring a high activity of these genes during the carcinogenic process. *Lamc2* is the subunit gamma chain of laminin 5 that was reported to be overexpressed in lung cancers⁽⁴⁴⁾, and also methylated in bladder tumors with poor prognosis⁽⁴⁵⁾. In our study all invasive cancers showed *Lamc2* over-expression by RT-PCR as well as clear positive immunoreaction located exclusively in the cytoplasm of the cells, in agreement with the findings of Kohlberger P⁽⁴⁶⁾. As Claudin 1 this protein was not found in normal cervical tissue, and these results are in agreement with those reported by Skyldberg B⁽⁴⁷⁾. *Lamc2* studies have been performed predominantly in cervical or vaginal cancer samples mainly infected with HPV16 or HPV52⁽⁴⁸⁾, and not specific HPV type expression has been described for this gene. Interestingly, our RT-PCR results demonstrates that *Lamc2* is only overexpressed in HPV16⁺ samples, while the HPV18 samples show a basal expression similar to the normal epithelium. This difference might be explained in part because HPV16 and HPV52 are phylogenetically gathered within the same A9 group, and this group probably activates more efficiently this gene than the C group, where HPV18 belongs⁽⁴⁹⁾. Clearly, further studies should be performed in this direction. It is important to mention that *Lamc2* has also already basal expression in normal tissue samples and in the HPV⁻ cell line; suggesting that the overexpression of this transcript might be mostly due to the presence of HPV16.

In summary, HPV16 and 18 infections show a very similar pattern of expression with slight but significant differences. Gene expression profiling indicates differences in the expression profiles of *Lamc2*, *Ipo7* and *NDUFB6* genes in cervical tissues infected with HPV types and in HPV⁻ samples. These genes might provide new insights in the understanding of HPV related cervical cancer biology, indicating that alterations in the mitochondrial inner membrane, IAP family (inhibitor of Apoptosis), extracellular matrix components and signal transduction are closely involved in Hr-HPV cervical carcinogenesis.

Acknowledgments

We are grateful to Dr. Moreno, Dr. Arenas (CMN SXXI-IMSS) and MSc. F Avila (INER, SS) for their helpful support. This work was partially supported by Grants No. 34686-M and 7114 (Fondos Sectoriales, CONACyT-México), FP-2001/043 (IMSS), SAF02-0363 (Spain) and SGR01-0141 (Catalonia). During this work GVO, AH, PP and CPP were recipients of a CONACyT and IMSS fellowship.

References

1. zur Hausen H, de Villiers EM. Human papillomaviruses. *Annu. Rev Microbiol* 1994;**48**:427-47.
2. Walboomers JM, Jacobs MV, Manos MM et al. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J Pathol* 1999; **189**:12-19.
3. Barnes WG, Woodworth S, Waggoner M et al. Rapid dysplastic transformation of human genital cells by human papillomavirus type 18. *Gynecol Oncol* 1990;**38**:343-6.
4. Burger RA, Monk BJ, Kurosaki T et al. Human papillomavirus type 18: association with poor prognosis in early stage cervical cancer. *J Natl Cancer Inst* 1996;**88**:1361-5.
5. Schwarz SM, Darling JR, Shera KA et al. Human papillomavirus and prognosis of invasive cervical cancer: a population-based study. *J Clin Oncol* 2001;**19**:1906-15.
6. Werness BA, Levine AJ, Howley PM. Association of human papillomavirus types 16 and 18 E6 proteins with p53. *Science* 1990;**248**:76-9.
7. Hidalgo A, Schewe C, Petersen S et al. Human papilloma virus status and chromosomal imbalances in primary cervical carcinomas and tumor cell lines. *Eur J Cancer* 2000;**36**:542-8.
8. Hidalgo A, Monroy A, Arana RM et al. Chromosomal imbalances in four new uterine cervix carcinoma derived cell lines. *BMC Cancer* 2003;**20**:8-12.
9. Stoler MC, Rhodes CR, Whitbeck A et al. Human papillomavirus type 16 and 18 gene expression in cervical neoplasias. *Hum Pathol* 1992;**23**:117-28.
10. Hudelist G, Manabí M, Pischinger KI et al. Physical state and expression of HPV DNA in benign and dysplastic cervical tissue: different levels of viral integration are correlated with lesion grade. *Gynecol Oncol* 2004;**92**:873-80.
11. Wentzensen N, Ridder R, Klaes R et al. Characterization of viral-cellular fusion transcripts in a large series of HPV16 and 18 positive anogenital lesions. *Oncogene* 2002;**17**:419-26.
12. Münger K, Werness BA, Dyson N et al. Complex formation of human papillomavirus E7 proteins with the retinoblastoma tumor suppressor gene product. *EMBO J* 1989;**8**:4099-105.
13. Scheffner M, Werness BA, Huibregtse JM et al. The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. *Cell* 1990;**63**:1129-36.
14. Milde-Langosch K, Riethdorf S. Role of cell-cycle regulatory proteins in gynecological cancer. *J Cell Physiol* 2003;**196**:224-44.
15. Chen Y, Miller C, Mosher R et al. Identification of cervical cancer markers by cDNA and tissue microarrays. *Cancer Res* 2003;**15**:1927-35.
16. Cheng Q, Lau WM, Tay SK et al. Identification and characterization of genes involved in the carcinogenesis of human squamous cell cervical carcinoma. *Int J Cancer* 2002;**98**:419-26.
17. Garcia AM, Ortiz-Navarrete VF, Mora-Garcia ML et al. Identification of peptides presented by HLA class I molecules on cervical cancer cells with HPV-18 infection. *Immunol Lett* 1999;**67**:167-77.
18. Caceres-Cortes JR, Alvarado-Moreno JA, Waga K et al. Implication of tyrosine kinase receptor and steel factor in cell density-dependent growth in cervical cancers and leukemias. *Cancer Res* 2001;**61**:6281-9.

19. Lopez-Marure R, Gutierrez G, Mendoza C *et al.* Ceramide promotes the death of human cervical tumor cells in the absence of biochemical and morphological markers of apoptosis. *Biochem Biophys Res Commun* 2002;**293**:1028-36.
20. International Federation of Obstetrics and Gynaecology. News. *J Gynecol Obstet* 1987;**25**:87.
21. Miller C, Mosher R, Zhao X *et al.* Identification of cervical cancer markers by cDNA and tissue microarrays. *Cancer Res* 2003;**63**:1927-35.
22. Snyder W, Faber K, Wenzel T *et al.* Pex19p interacts with Pex3p and Pex10p and is essential for peroxisome biogenesis in *Pichia pastoris*. *Mol Biol Cell* 1999;**10**:1745-61.
23. Máximo V, Preto A, Crespo A *et al.* Core I gene is overexpressed in Hurthle and non-Hurthle cell microfollicular adenomas and follicular carcinomas of the thyroid. *BMC Cancer* 2004;**1**:12-16.
24. Tan DJ, Chang J, Chen WL *et al.* Novel heteroplasmic frameshift and missense somatic mitochondrial DNA mutations in oral cancer of betel quid chewers. *Genes Chromosomes Cancer*. 2003;**37**:186-94.
25. Jerónimo C, Nomoto S, Caballero OL *et al.* Mitochondrial mutations in early stage prostate cancer and bodily fluids. *Oncogene* 2001;**20**:5195-8.
26. Bernard B, Pretet JL, Charlot JF *et al.* Human papillomaviruses type 16+ and 18+ cervical carcinoma cells are sensitive to staurosporine-mediated apoptosis. *Biol Cell* 2003;**95**:17-26.
27. Ahn WS, Bae SM, Lee JM *et al.* Searching for pathogenic gene functions to cervical cancer. *Gynecol Oncol* 2004;**93**:41-8.
28. Brinkmann U, Vasmatzis G, Lee B *et al.* PAGE-1, an X chromosome-linked GAGE-like gene that is expressed in normal and neoplastic prostate, testis, and uterus. *Proc Natl Acad Sci U S A.* 1998;**95**:10757-62.
29. Gorlich D, Dabrowski M, Bischoff FR *et al.* A novel class of RanGTP binding proteins. *J Cell Biol* 1997;**38**:65-80.
30. Freedman ND, Yamamoto KR. Importin 7 and Importin {alpha}/Importin {beta} Are Nuclear Import Receptors for the Glucocorticoid Receptor. *Mol Biol Cell* 2004;**15**:2276-86.
31. Bamberger CM, Bamberger A.M, de Castro M *et al.* Glucocorticoid receptor beta a potential endogenous inhibitor of glucocorticoids action in humans. *J Clin Invest* 1995;**95**:2435-41.
32. Mittal R, Tsutsumi K, Pater A *et al.* Human papillomavirus type 16 expression in cervical keratinocytes: role of progesterone and glucocorticoid hormones. *Obstet Gynecol* 1993;**81**:5-12
33. Khare S, Kumar KU, Tang SC *et al.* Up-regulation of hormone response of human papillomavirus type 16 expression and increased DNA-protein binding by consensus mutations of viral glucocorticoid response elements. *J Med Virol* 1996;**50**:254-62.
34. Medina-Martinez O, Morales-Peza N, Yaniv M *et al.* A single element mediates glucocorticoid hormone response of HPV18 with no functional interaction with AP1 or hbrm. *Virology* 1996;**217**:392-6.
35. Khan MA, Canhoto AJ, Housley PR *et al.* Glucocorticoids stimulate growth of human papillomavirus type 16 (HPV16)-immortalized human keratinocytes and support HPV16-mediated immortalization without affecting the levels of HPV16 E6/E7 mRNA. *Exp Cell Res* 1997;**236**:304-10.

36. Wu W, Chaudhuri S, Brickley DR *et al*. Microarray analysis reveals glucocorticoid-regulated survival genes that are associated with inhibition of apoptosis in breast epithelial cells. *Cancer Res* 2004;**64**:1757-64.
37. Ye L, Martin TA, Parr C *et al*. Biphasic effects of 17-beta-estradiol on expression of occludin and transendothelial resistance and paracellular permeability in human vascular endothelial cells. *J Cell Physiol* 2003;**196**:362-9.
38. Hoerel T, Macek R, Swisshelman R *et al*. Reexpression of the TJ protein CLDN1 induces apoptosis in breast tumor spheroids. *Int J Cancer* 2004; **108**:374-83.
39. Hoevel T, Macek R, Mundigl O *et al*. Expression and targeting of the tight junction protein CLDN1 in CLDN1-negative human breast tumor cells. *J Cell Physiol* 2002;**191**:60-8.
40. Kramer F, White K, Kubbies M *et al*. Genomic organization of claudin-1 and its assessment in hereditary and sporadic breast cancer. *Hum Genet* 2000;**107**:249-56.
41. Miwa N, Furuse M, Tsukita S *et al*. Involvement of claudin-1 in the beta-catenin/Tcf signaling pathway and its frequent upregulation in human colorectal cancers. *Oncol Res* 2000;**12**:469-76.
42. Su TH, Chang JG, Yeh KT, Lin TH, Lee TP *et al*. Mutation analysis of CTNNB1 (beta-catenin) and AXIN1, the components of Wnt pathway, in cervical carcinomas. *Oncol Rep* 2003;**10**:1195-200.
43. Sathyanarayana UG, Padar A, Suzuki M *et al*. Aberrant promoter methylation of laminin-5-encoding genes in prostate cancers and its relationship to clinicopathological features. *Clin Cancer Res* 2003;**9**:6395-400.
44. Manda R, Kohno T, Niki T *et al*. Differential expression of the LAMB3 and LAMC2 genes between small cell and non-small cell lung carcinomas. *Biochem Biophys Res Commun* 2000;**275**:440-5.
45. Sathyanarayana UG, Maruyama R, Padar A *et al*. Molecular detection of noninvasive and invasive bladder tumor tissues and exfoliated cells by aberrant promoter methylation of laminin-5 encoding genes. *Cancer Res* 2004;**64**:1425-30.
46. Kohlberger P, Beneder Ch, Horvat R *et al*. Immunohistochemical expression of laminin-5 in cervical intraepithelial neoplasia. *Gynecol Oncol* 2003;**3**:391-4.
47. Skyldberg B, Salo S, Eriksson E *et al*. Laminin-5 as a marker of invasiveness in cervical lesions. *J Natl Cancer Inst* 1999;**21**:1882-7.
48. Habermann JK, Hellman K, Freitag S *et al*. A recurrent gain of chromosome arm 3q in primary squamous carcinoma of the vagina. *Cancer Genet Cytogenet* 2004;**148**:7-13.
49. Human Papillomaviruses Compendium. 1994. Los Alamos National Laboratory Theoretical Biology and Biophysics.

Tables and Figure Legends

Table 1. Differentially expressed genes in HPV positive and HPV negative samples. Most overexpressed and suppressed genes in any HPV⁺ tissue obtained by a gene expression profile of C33, HPV16 (median of normalized expression levels of SiHa, CaSki and three HPV16⁺ invasive samples) and HPV18 (median of normalized expression levels of HeLa, CALO and INBL cell lines) were compared to normalized median of 3 normal HPV⁻ tissue samples by using Gene Spring software. GeneBank access number and functional groups are listed.

HPV, Human Papillomavirus

Table 2. Immunoreactivity for Lamc2 and clnd1 in cervical tissues.

Lamc2, laminin 5 subunit gamma chain;
Clnd1, Claudin 1
CC, Cervical Cancer;

Fig. 1. Differential gene expression in HPV positive and HPV negative cervical samples. cDNA microarray analysis of the expression of three normal HPV negative samples, C33, HPV16 (CaSki, SiHa cell lines and three cancer specimens) and HPV18 (CALO, HeLa, INBL). Data are presented in a graph format. X axis represents samples and Y axis the transcript levels. Each one of the studied genes is present in the four groups (not observed in HPV⁻ and C33). Line (a) represents NDUFB6 gene expression; line (b) Importin 7 gene; line (c) Claudin 1, and line (d) Lamc2. The controls used were beta-2-microglobulin, alpha tubulin, beta actin, GADPH and as negative control yeast cDNA was used.

HPV, Human Papillomavirus

Fig. 2. Relative mRNA expression levels of Claudin 1, Importin 7, Lamc2 and NDUFB6 in cervical tumors and cell lines detected by RT-PCR.

One-hundred ng of total RNA was employed for RT-PCR experiments. A) represents RT-PCR products for Claudin 1 and graphic representation of the relative mRNA expression levels for the used samples; B) Importin 7 expression levels; C) Lamc2 gene expression, and D) NDUFB6 mRNA levels. Values are presented as ratio of densitometric readings of samples to corresponding β -actin samples. The used samples were; lane 1) negative control, 2) normal HPV⁻ tissue #N07, 3) normal HPV⁻ tissue # N11, 4) normal HPV⁻ tissue # N20, 5) C33 cells, 6) SiHa cells, 7) CaSki cells, 8) tumoral HPV16 sample # T04, 9) tumoral HPV16 sample # T24, 10) tumoral HPV16 sample # T43, 11) tumoral HPV16 sample # T45, 12) INBL cells, 13) CALO cells, and 14) HeLa cells. A 100 bp ladder was used as molecular weight marker.

RT, reverse transcription;

Lamc2 antibody

	Total No.	No. Positive	%
Normal Cervical Tissue, HPV-	10	0	0
Cervical Cancer, HPV16	15	14	93.33

Claudin 1 antibody

	Total No.	No. Positive	%
Normal Cervical Tissue, HPV-	10	0	0
Cervical Cancer, HPV16	15	14	93.33

Lamc2, laminin 5 subunit gamma chain;
NDUFB6, Complex I NADH:ubiquinone oxidoreductase;
HPV, Human Papillomavirus

Figure 3. Immunohistochemical detection of Claudin 1 and Lamc2 in normal cervical and CC tissues.

A) Colon carcinoma sample (as positive control for clnd1 staining), B) negative control; a CC sample without primary antibody, C) one normal cervical tissue showing negative staining for clnd1, D) positive immunostaining in a CC tissue for clnd1, E) colon carcinoma specimen was used as positive control for Lamc2 immunostaining, F) normal cervical tissue shows negative immunostaining for lamc2, and G) a CC showing positive immunoreaction for Lamc2. Original magnifications were X40 for A, B, D, E and G; and 10X were used for C and F.

IHC, Immunohistochemistry
Lamc2, laminin 5 subunit gamma chain;
CC, Cervical Cancer;
Clnd1, Claudin 1

GenBank accession no.	Gene Name or Product	Folds of expression in C33	Folds of expression in HPV16+	Folds of expression HPV18+
Cell death				
NM_004394	Death associated Protein	6	14	6
Cell growth and maintenance				
NM_003049	Solute carrier family, member 1	1	24	6
NM_014252	Solute carrier family 25, member 15	1	24	6
NM_002857	peroxisomal farnesylated protein	1	15	6
NM_002493	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 6	1	16	8
Developmental processes				
L26953	regulator of mitotic spindle assembly 1	2	22	8
U25029	Human glucocorticoid receptor alpha	1	16	5
NM_001781	CD69 antigen	1	6	12
Intracellular signaling				
NM_020184	ancient conserved domain protein 4	1	32	12
U25029	Glucocorticoid receptor alpha	1	4	2
NM_006391	RAN binding protein 7 (importin 7)	1	8	4
NM_003785	G antigen, family B, 1 (prostate associated)	6	18	8
Physiological processes				
NM_015839	ficolin (collagen/fibrinogen domain-containing lectin) 2	6	22	12
NM_020299	aldo-keto reductase family 1, member B11	2	12	16
Extracellular				
NM_005562	laminin, gamma 2	1	12	1
NM_002638	protease inhibitor 3, skin-derived (SKALP)	1	10	6
Membrane				
IM_004945	dynamamin 2	1	10	6
Oncogenes				
NM_002005	feline sarcoma viral (v-fps) oncogene homolog	1	4	18
Cell adhesion				
NM_021101	Claudin 1	1	12	10
Cyclins				
NM_031966	cyclin B1	1	26	4
Chaperons				
NM_006431	chaperonin containing TCP1, subunit 2 (beta)	8	35	45

Suppressed Genes

Peptidases				
IM_003937	Kynureninase (L-kynurenine hydrolase)	65	-99	-99
Nucleus				
IM_003439	Activator (Nuclear) Pol II transcription	78	-99	-99
Cytoplasm				
IM_012198	Granulysin	56	-99	-99
Cell adhesion				
IM_001795	VE-cadherin (vascular endothelial cadherin, cadherin 5)	71	-99	-99
IM_004932	K-cadherin (cadherin 6)	88	-99	-99
Uterine				
U057138	Placental thrombin inhibitor			
Elastase				
IM_001972	Elastase neutrophil-specific serine protease	76	-99	-99
Receptors				
IM_003855	interleukin 18 receptor 1	59	-99	-99
Homeobox				
IM_014795	zinc finger homeobox 1B	70	-99	-99
Developmental processes				
IM_003426	Kruppel-like zinc-finger protein 74	90	-99	-99

Normalized intensity (linear scale)

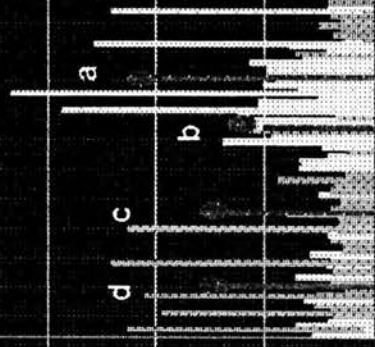
100
90
80
70
60
50
40
30
20
10

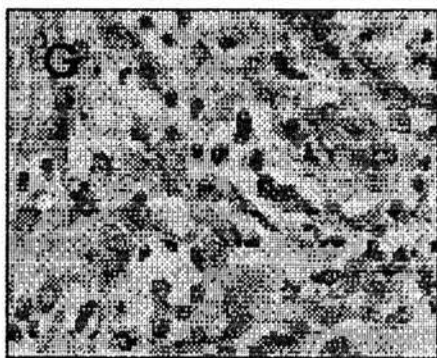
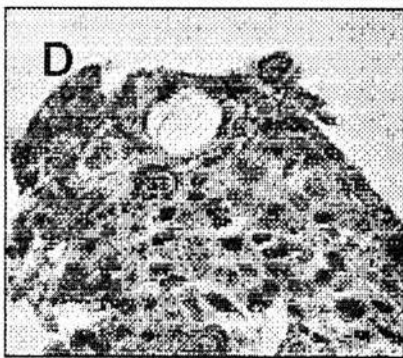
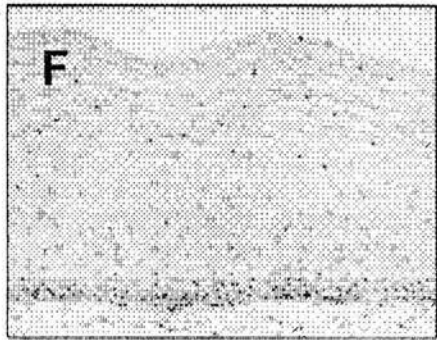
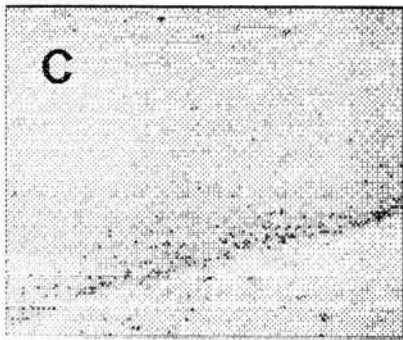
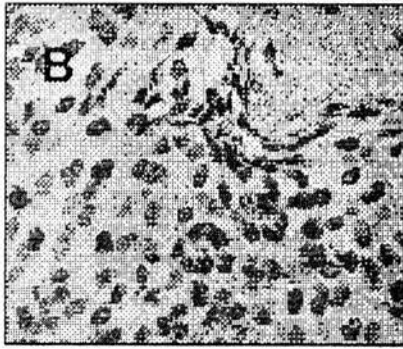
HPV-
Normal samples

C33

HPV16

HPV18





Manuscript Number:

Title: Identification of Genes Involved in HPV-Positive Cervical Cancer by cDNA Arrays

Article Type: Full Length Article

Keywords: Key words: HPV, expression, cDNA arrays, RT-PCR.

Corresponding Author: Dr. Mauricio Salcedo Oncogenomics Laboratory, Oncological Diseases Research Unit, Oncology Hospital,

Other Authors: Guelaguetza Vazquez-Ortiz, Cand; Silvia Peñuelas, Jr, Carlos J Ciudad, PhD; Patricia Piña, MSc; Karla Vazquez, Bach; Alfredo Hidalgo, Cand; Brenda Alatorre, Bsch; José A García, Cand; Fabio Salamanca, Md MSc; Oncogenomics Laboratory, Oncological Diseases Research Unit, Oncology Hospital, National Medical center Siglo XXI-IMSS, México, Department of Biochemistry and Molecular Biology, School of Pharmacy, University of Barcelona, Spain, Department of Biochemistry and Molecular Biology, School of Pharmacy, University of Barcelona, Spain, Oncogenomics Laboratory, Oncological Diseases Research Unit, Oncology Hospital, National Medical Center Siglo XXI-IMSS, México, Oncogenomics Laboratory, Oncological Diseases Research Unit, Oncology Hospital, National Medical Center Siglo XXI-IMSS, México, Oncogenomics Laboratory, Oncological Diseases Research Unit, Oncology Hospital, National Medical Center Siglo XXI-IMSS, México, Oncogenomics Laboratory, Oncological Diseases Research Unit, Oncology Hospital, National Medical Center Siglo XXI-IMSS, México, Laboratory of Theoretical Biology, Research Department, La Salle University, México, Medical Research Unit in Human Genetics, National Medical Center Siglo XXI-IMSS, México

Identification of Genes Involved in HPV Positive Cervical Cancer by cDNA Arrays

Guelaguetza Vázquez-Ortiz^a, Silvia Peñuelas^b, Carlos J. Ciudad^b, Patricia Piña^a, Karla Vazquez^a, Alfredo Hidalgo^a, Brenda Alatorre^a, Jose A. Garcia^c, Fabio Salamanca^d, Mauricio Salcedo^a.

^a*Oncogenomics Laboratory, Oncological Diseases Research Unit, Oncology Hospital, National Medical Center Siglo XXI-IMSS, Mexico.* ^b*Department of Biochemistry and Molecular Biology, School of Pharmacy, University of Barcelona, Spain.* ^c*Laboratory of Theoretical Biology, Research Department, La Salle University, Mexico,* ^d*Medical Research Unit in Human Genetics, National Medical Center Siglo XXI-IMSS, Mexico.*

Guelaguetza Vazquez: guelav@yahoo.com , Patricia Piña: mzactli@hotmail.com , Karla Vazquez: sotaliak@hotmail.com, Alfredo Hidalgo: alfhm@yahoo.com, Fernando Cruz-Talonia, Brenda Alatorre: calili@hotmail.com, Mauricio Salcedo: maosal89@yahoo.com. Unidad de Investigación Médica en Enfermedades Oncológicas, Hospital de Oncología, Centro Médico Nacional Siglo XXI-IMSS, Av. Cuauhtémoc 330. Col Doctores. México D. F. C. P. 06720, Phone: +55 56276900 Ext: 22708 Fax: +55 57610952

Silvia Peñuelas: cciuudad@ub.edu, Carlos J. Ciudad: cciuudad@ub.edu. Departamento de Bioquímica y Biología Molecular, Facultad de Farmacia, Universidad de Barcelona, Av. Diagonal 643, E-08028 Barcelona, España, Phone +34-93-403-4455, Fax: +34-93-402-4520.

Jose A. Garcia jgarcia@ci.ulsal.mx, Laboratorio de Biología Teórica, Departamento de Investigación, Universidad La Salle, Benjamín Franklin No.47 C.P. 06140 México D.F. Col. Condesa, phone +55 5278 9500
Fernando Cruz Talonia, maosal89@yahoo.com, Centro Nacional de Clínica de Displasias, Hospital General de México, SS, Dr. Balmis 34, Col. Doctores, phone +55 56708451

Fabio Salamanca fasalam@prodigy.net.mx, Unidad de Investigación Médica en Genética Humana, Hospital de Pediatría, Centro Médico Nacional Siglo XXI-IMSS, Av. Cuauhtémoc 330. Col Doctores. México D. F. C. P. 06720, Phone: +55 56276900 Ext: 21945 Fax: +55 57610952

Running Title: *Expressed genes in HPV infection.*

Corresponding Author:

Dr. Mauricio Salcedo Vargas

Unidad de Investigación Médica en Enfermedades Oncológicas

Hospital de Oncología

Centro Médico Nacional Siglo XXI-IMSS

Av. Cuauhtémoc 330. Col Doctores.

México D. F. C. P. 06720

Phone: +55 56276900 Ext: 22708 Fax: +55 57610952

E-mail: guelav@yahoo.com, maosal89@yahoo.com

Background. One of the most frequent malignancies in women worldwide is the Uterine Cervical Carcinoma. High risk human papillomavirus (HPV) infection is considered to be the most important etiological factor of uterine cervical cancer. To identify novel cellular genes that could potentially act as predictive molecular markers for human cervical cancer we employed cDNA arrays.

Methods. We used cDNA arrays to examine the expression profiles of 5 cell lines derived from human cervical cancer, three HPV+ tumor samples and three normal (HPV-) epithelium tissues. Data normalization was performed and the top overexpressed genes were obtained. Hierarchical cluster were performed and in order to validate some of the differentially expressed genes between normal and carcinogenic samples, semi-quantitative RT-PCR was used in tissue samples.

Results. Four genes were demonstrated to be consistently overexpressed in invasive cervical cancer biopsies; three novel genes not previously related to cervical cancer : MMP10, Lamc2 and Claudin 1 and in addition, the overexpression of IL6 was corroborated.

Conclusions. The progress in the identification of characteristic molecular changes in cervical cells due to carcinogenesis and HPV infection can lead to a better understanding of the disease. cDNA arrays are starting to provide new possible molecular markers of prognostic and diagnostic. In a long term, this technology might help to elucidate the biological differences of mechanisms associated with different HPV type infection and aggressive prognosis.

Key words: HPV, expression, cDNA arrays, RT-PCR.

Introduction

One of the most frequent malignancies in women worldwide is the Uterine Cervical Carcinoma (CC), both in incidence and mortality (1). It is well known that infection by genital HPVs (mainly types 16, 18, 31 and 33) is the primary causal agent for the CC development and its precursors; this is besides of the epidemiological factors such as first sexual intercourse, smoking etc. HPV 16 and 18 types are present at least in the 65 % of all invasive CC worldwide (2-4), and in human cell lines derived from this malignancy (5-7). The detection of the infection with the known high risk HPV (Hr-HPV) types and its correlation with the presence of cervical cancer, have been lead to consider the infection as a marker for cervical carcinoma development (8). However, the malignant transformation of cervical epithelial cells requires a long latency periods, it is believed that could be up to 10 years (9-10). Nevertheless, Hr-HPV types have been detected in healthy normal women (11), indicating that additional factors and cellular events are required for cervical carcinogenesis.

The HPV genome encode for less than 10 proteins which main function is to regulate viral transcription and replication during cell cycle. Therefore, interactions of viral and cell proteins are essential for the productive life cycle (12). Continuous expression of E6 and E7 oncoproteins of the high risk HPV types appears to be necessary for maintaining the malignant phenotype of the cervical cancer patients interacting with several cellular factors (13, 14).

On the other hand, the colossal development of cDNA arrays or biochips has provided a powerful approach to screen global gene expression (15). The biological diversity of cell tumors is due to the variation in transcription of the genes. The patterns of expression of different breast tumors have provided a distinctive molecular portrait that have been used for the classification of tumor subtypes by differences in their gene expression patterns (16). The expression pattern of CC has not been yet fully analyzed. In a first report using a 588 gene array, the altered expression of several genes associated to apoptosis and cell adhesion were detected in CC (17). Transfection of normal keratinocytes with the viral oncogenes E6 and E7, altered the expression of Interferon responsive genes, genes stimulated by NFK-kB of cell cycle and DNA synthesis (18). A differential display and 44 gene microarray detected that NADH4 and ribosomal protein S12 altered its expression in the normal cervical tissue that is adjacent to the malignant one (19).

We used cDNA arrays to survey the alterations in gene expression between HPV positive and HPV negative samples as a reflection of cellular gene alterations, to identify and evaluate genes that are up-regulated in cervical cancer.

Materials and Methods

Cell lines and Tissue Collection .Hela (HPV18), SiHa (HPV16) and CaSki (HPV16) cell lines were grown in the recommended media including 10% fetal bovine serum. The cell lines designated CALO (derived from Invasive squamous carcinoma Stage IIB, HPV18) and INBL (derived from Invasive Squamous carcinoma IV A, HPV18), were established at the FES Zaragoza-UNAM, Mexico (20-22). Eight normal cervical tissues from premenopausal women were obtained *post-mortem* at the Pathology Department of the Hospital General de México, S. S. from women who died of non-gynecological related causes. The invasive lesions (8 samples) were collected through colposcopy guided biopsies of the tumor lesion of pre-menopausal women at the National Centre of Dysplasias at Mexico City. All the described procedures have been evaluated and approved by the local committee of ethics of the Mexican Institute of Social Security and samples were taken after inform consent from the patients. All tissue samples were divided in three sections, the central part was snapped frozen in liquid nitrogen and stored at -70°C until nucleic acid extraction, and the extremes from the biopsy were fixed with 70% ethanol overnight and paraffin embedded. Haematoxilin/eosin stained sections from the extremes of the biopsies were analyzed in order to confirm the presence of at least 70% tumoral cells. Multiple step sections were made and classified according to the International Federation of Obstetrics and Gynaecology (FIGO). All the invasive samples were diagnosed as squamous lesions.

Nucleic Acids Extraction and Array hybridization. Both RNA and DNA were obtained from the tissue samples and the cervical cancer cell lines using the Trizol Reagent (Gibco BRL, Life Technologies, Grand Island NY USA) according to manufacturer's conditions. Each DNA sample was subjected to HPV typing by polymerase chain reaction (PCR) using the MY09/MY11 consensus primers and direct sequencing.

To eliminate residual genomic DNA, Total RNA was treated with RNase-free DNase I (Ambion Inc, Austin Texas USA). Total RNA concentration of each sample was assessed

by absorbancy and the nucleic acid integrity was confirmed by agarose gel electrophoresis.

Messenger RNA isolation of each sample (HeLa, SiHa, CaSki, CALO, INBL), three HPV16+ tumor samples (T07, T31, T64) and three HPV- normal samples (N03, N11, N22) was performed using the Polyattract Isolation System II (Promega Madison, WI USA). The mRNAs were quantified using the VersaFluor Fluorometer (BioRad Laboratories Inc, USA). Labeled cDNAs were obtained by RT-PCR from 500 ng of mRNA in the presence of (α - ^{32}P) dATP (activity 3000Ci/mmol) using the Ambion Advantage System, following the manufacturers conditions. Briefly, mRNAs were denatured at 70°C for 10 min and cDNAs were synthesized at 42°C using oligo-dT priming in a final volume of 30 μl . The unincorporate nucleotides were removed by centrifugation using the NucAway Spin Columns (Ambion Inc, Austin Texas USA). The final incorporation of the radiolabel was approximately 1×10^8 cpm for each reaction. ULTRArray Advantage System array blots (Ambion Inc. Austin Texas USA) which contains 8400 genes, were prehybridized at 68°C for at least 90 min prior to probe addition in Hybridization buffer. Finally, 1.5×10^7 cpm of each labeled cDNA was added to the buffer. Hybridization was performed at 60°C in a rolling bottle overnight. The arrays were washed twice with 2X SSC and 0.5 % SDS at 60°C for 30 min; followed by two stringent washes with 0.5X SSC, 0.5% SDS at same temperature and time.

Array analysis. Hybridized arrays were scanned on a STORM 860 phosphorimager/fluorimager (Amersham Biosciences, Buckinghamshire UK). The resulting images were analysed with the ArrayVision software 8.0 rev 4 (Imaging Research Inc, USA), and data from the duplicate spots were normalized using Gene Spring software version 5.0 (Silicon Genetics, USA). Values below 0.01 were set to 0.01. Each measurement was divided by the 50.0th percentile of all measurements in that sample. Samples were normalized against the median of the control samples (normal HPV-samples). When calculating the mean of the control measurements only those measurements were included that had at least a normalized value of 0.01. Each measurement for each gene in the experimental samples was divided by the median of that gene's measurements in the corresponding control samples. The 20 more upregulated genes were submitted to Hierarchical clustering by using the J-Express software (MolMine, www.molmine.com) and visualized by Treeview software.

RT-PCR Analysis. Two-hundred ng of Total RNA from normal tissues, cell lines and tissue samples were reverse transcribed using the RT-PCR Access System (Promega, Madison, WI USA). Primers for some of the genes of interest were designed using Integrated DNA technologies' Biotoools software (<http://biotoools.idtdna.com/gateway>) for the amplification of the following gene transcripts: *Laminin gamma 2 (Lamc2)*, 696 bp TM=56°C (5'-TGTTTAGCTCAGCCCAACGACT-3') (5'-AGCAGCTGAACCCGTTATGACA-3'); *Claudin1 (cldn1)*, 504 bp TM=55°C (5'-GTGGAGGATTTACTCCTATGCC-3') (5'-CTTGGTGTGGGTAAGAGGT-3'), *Matrix Metalloproteinase 10 (MMP10)*, 509 bp TM=60°C (5'-CATTCACTCTCTACGGACCT-3'), (5'-CAGCTATTAGTCTAGGGAAGCC-3') *Interleukin-6 (IL6)*, 621 bp: TM=55°C (5'AGTTGCCTTCTCCCTGG3') (5'TGAGGGCTCTTGGGCAAAT-3') (23), and β -actin : (5'-TGAAGTCTGACGTGGACATC-3') (GTTCGTTCCCTCATACTGCTCA-3') 243 bp, TM=55°C. The one-step RT-PCR conditions were 48°C for 45 min to synthesize the first strand of cDNA, 94°C for 2 min to denature template, and to synthesize second strand and DNA amplification: 94°C for 30 sec, (specific TM°C for each set of primers) for 1 min, 68°C for 2 min for a total of 25 cycles and followed by a single step at 68°C for 7 min. The products were visualized on 1.5% agarose gel stained with ethidium bromide, and images were captured using the Eagle eye system (Stratagene, La Jolla, Ca USA). The sizes of the products were determined by comparison with a 100 bp marker (GeneRuler 100 bp DNA Ladder, MBI Fermentas Life Sciences, Burlington ON CAN). The intensities of IL6, MMP10, Lamc2, Claudin 1 and β -actin RT-PCR products were analyzed using the MetaView analysing system (version 4.5 Universal Imaging Corp, USA). The relative concentrations of the mRNAs were expressed as the ratio of densitometric readings of the four genes related to β -actin mRNA. All values are presented as the mean of at least 4 determinations. Statistical comparisons among groups were analysed by a Student's t-test. A *P* value of <0.001 was considered as statistically significant.

Results

Gene expression profile of HPV negative tissue samples and HPV positive samples. To study transcriptional differences between HPV+ and HPV- samples, the gene expression profile of normalized SiHa, CaSki, HeLa, CALO and INBL cell lines and three HPV16+ invasive samples were compared to the normalized median of 3 normal HPV- tissue samples using the Gene Spring Software. With this criteria, we found approximately 50

genes expressed at least four-fold higher in HPV+ , we also observed a set of cellular genes that showed significant decreases in expression; however, these were not characterized in the current study.

To make the data results more apparent we used a hierarchical clustering analysis with a pseudo colour visualization matrix (24) (Figure 1). The aim purpose of the clustering was to show the overexpressed genes with similar levels of expression in cancer tissues and in the cell lines derived from cervical cancer. This graph, demonstrated the upregulated expression of several genes in all samples; meaning that these genes could be potentially employed as diagnostic markers.

In order to eliminate false positives the top 20 up-regulated genes in all malignant samples were selected to determine a table, this was based on more than 10 fold of overexpression (Table 1). From this group, four genes were identified: the Matrix Metalloproteinase 10 (MMP10), Interleukin 6 (IL6), Claudin 1, and Laminin gamma 2 (Lamc2).

Figure 1 shows the expression pattern of all overexpressed HPV+ samples, blue arrows show the studied genes.

Table 1 shows the expression levels of 20 selected transcripts that show a ten fold or higher expression in the HPV positive tissues.

RT-PCR analysis. To confirm the expression differences observed in the hierarchical cluster, total RNA of some of the studied samples as well as additional samples, were analyzed by semi-quantitative single step RT-PCR assays. A representative picture of the amplified products is shown in Figure 2. Computer aided densitometric analysis of the amplified bands showed that Claudin 1 presents overexpression in all HPV+ samples ($P<0.001$); similar results have been recently reported (25). Lamc2 expression increases dramatically in HPV16 cell lines, and it is also overexpressed in HPV16 human cancer specimens ($P<0.001$); whereas the IL6 gene presents overexpression in all HPV+ samples ($P<0.001$) as well as the mmp10 gene ($P<0.001$).

Discussion

The understanding of the molecular differences induced by HPV might help to elucidate mechanisms associated with the clinical prognosis of the cancer. We have employed cDNA arrays to monitor the expression levels of more than 8,000 genes in five cell lines derived from CC, three malignant tumors and three normal HPV- none dysplastic cervical

tissue samples. In the present study, cDNA microarrays were used to detect possible molecular markers and molecular alterations in human squamous cervical cancer, this is the reason by why we only focused on genes that are overexpressed in all samples. A large number of these, are the ribosomal genes; the overexpression of such genes in tumors have been previously noted and may reflect a higher metabolic rate within malignant cells, besides, this overexpression is correlated with tumors that are histologically well differentiated (26).

Other genes that were overexpressed in all samples and were detected by hierarchical cluster were: IL-6, MMP10 Claudin 1 and Lamc2. Interestingly, our RT-PCRs results corroborated that the levels of expression of these genes were upregulated in tissue biopsies of invasive cervical cancer in contrast to the normal tissue. It is likely that as the carcinoma spread, tissues undergoing molecular changes that are reflected in the alterations of their gene expression.

As expected one of the overexpressed genes was IL-6. This gene codifies to a cytokine that can regulate immune and inflammatory responses (27) and it is suggested that may be a central mediator of inflammation in female genital infections (28). Its production is regulated by a variety of stimuli. It is now known that various viruses induce IL-6 in fibroblasts, even HIV induces its production. It is important to note that cervical cancer frequently develops in close association with infection with various sexually transmitted agents like HPV. Besides, in the uterine cervix several cell types are responsible for elevating IL-6 such as lymphocytes, macrophages, fibroblasts, keratinocytes, endothelial cells and tumor cells (29). It is mainly produced in tumor and/or epithelial cells and can act as growth factor for cervical carcinoma cell lines (30). Moreover, IL-6 levels are increased in cervicovaginal secretions of patients with uterine cervical cancer, and its production is related to the severity of cervical neoplasia (31), meaning that it has an increased expression in primary invasive cervical carcinomas as compared to CIN or normal cervix (32). A variety of peptide factor such as $TNF\alpha$, IL-2, $IFN\beta$ and PDGF cause an increase in intracellular cyclic AMP level which also induce IL-6 production (33). A mutation of p53 makes that IL-6 becomes overexpressed, in renal, ovarian, fallopian tubes and oral squamous carcinoma; suggesting that a lack of function of this gene by the presence E6 viral protein could cause the same situation (34-35). Wild type of p53 has a strong repressive effect on the IL-6 promoter (36), although IL-6 does not have a p53 binding site, it has a p53 consensus element. It also has been postulated that in normal conditions p53 acts as repressor by protein-protein interaction with C/EBP family members or the TATA

box, leading to repress the promoter of the IL-6 (37-38). IL-6 exerts diverse biological effects depending on target cell types. It can modulate VEGF production at the transcriptional level in cervical cancer cells. This transcript is overexpressed in our results too (data not shown). VEGF is an endothelial cell-specific mitogen, and is the link between tissue hypoxia and a compensatory angiogenic response in underperfused tumors (39). High microenvironmental levels of IL-6 are an important stimulator of VEGF in cervical cancer cell, promoting angiogenesis. The presence of this gene among the identified markers validates the cDNA microarray approach used in this study.

The Claudin family is the major constituent of the tight junction (TJ) complex. TJ are the structures that control cell-to-cell adhesion and consequently, regulate paracellular and transcellular transport of solutes across human epithelial and endothelial cells (40). Aberrations of tight junction function are an early and key event during the vascular spread of cancer and inflammation (41).

Claudin 1 (*clnd1*) has been described downregulated in breast cancer (42-43). This gene is upregulated by β -catenin (44). β -catenin belongs to the Wnt signalling pathway, which is commonly altered in colon and gastric cancers (45). Recent results from our group show that genes in the Wnt signaling pathway are overexpressed in cervical cancer, as an event associated to HPV infection (submitted 2004). This overexpression might cause an accumulation of β -catenin in the nucleus, activating many transcription factors like Claudin 1. In the present study, high levels of Claudin 1 mRNA were found in all HPV+ samples tested.

One cellular anchoring protein mainly found in the basement membrane is Laminin 5. Laminin 5 is composed by three polypeptidic chains that are the products of three distinct genes: *Lama3*, *Lamb3*, and *Lamc2*. It is already know that, in order to metastasize and grow, cancer cells must invade and destroy the basement membrane (46) requiring a high activity of these genes during the carcinogenic process. *Lamc2* is the subunit gamma chain of laminin 5 and has been reported to be overexpressed in lung cancers (47), and also methylated in bladder tumors with poor prognosis (48). This gene has also been proposed as possible molecular marker of invasiveness in cervical cancer (49-50). Nevertheless, these studies have been performed predominantly in cervical or vaginal cancer samples mainly infected with HPV16 or HPV52 (51), and no specific HPV type expression has been described for this gene. Interestingly, our results show that *Lamc2* is only overexpressed in HPV16+ samples, while the HPV18 samples show a basal expression similar to the normal epithelium. This difference might be explained in part

because HPV16 and HPV52 are phylogenetically gathered in the same A9 group, this group probably activates more efficiently this gene than the C group, where HPV18 belongs (52). Clearly, further studies in this direction should be performed. It is important to mention that Lamc2 also has already basal expression in normal tissue samples; suggesting that the overexpression of this transcript might be mostly due to the presence of HPV16.

Matrix metalloproteinase 10 was also found overexpressed in this work. There are several reports showing the relation between cervical cancer and some MMPs (53-55), nevertheless, no reports exist about cervical cancer and MMP10. It is known that the augmentation of the expression and activation of matrix metalloproteinases (MMPs) is associated with tumor invasion and metastasis and MMP10 is a metalloproteinase related to those substrates including proteoglycans and fibronectin (56). This led us to assume some similarities in its degrading behaviour with other MMPs. It has been reported that cell-cell contact between cervical carcinoma cells and peripheral stromal fibroblasts augments the production and activation of some MMPs, and therefore the subsequent imbalance between MMPs and TIMPs (tissue inhibitors of metalloproteinases) may result in the progression of invasiveness of cervical carcinoma cells in vivo (57), showing that this import pathway might be relevant in carcinogenesis and further studies should be done in this direction.

In summary, HPV infection shows a very characteristic pattern of expression with markedly differences with normal HPV negative tissues. Gene expression profiling indicates differences in the expression profiles of Lamc2, IL6, Claudin 1 and MMP10 genes in HPV+ and HPV- cervical tissues. These genes might provide a new insight in the understanding of HPV related cervical cancer biology, indicating that alterations in extracellular matrix components, cytokines, Proteases, and signal transduction are closely involved in Hr-HPV cervical carcinogenesis.

Acknowledgments

We are grateful to Dr. J. Moreno, Dr. D. Arenas (CMN SXXI-IMSS) and MSc. F Avila (INER, SS) for their helpful support. This work was partially supported by Grants No. 34686-M and 7114 (Fondos Sectoriales, CONACyT-México), FP-2001/043 (IMSS).

SAF02-0363 (Spain) and SGR01-0141 (Catalonia). During this work GVO, AH, and PP were recipients of a CONACyT and IMSS fellowship.

References

1. Bosch X, Manos M, Muñoz N, Sherman M, Jansen A, Schiffman M, Moreno V, Kurman C. Prevalence of human Papillomavirus in Cervical Cancer: a worldwide Perspective (IBSCC) Study Group. *J Natl Cancer Inst* 1995,87:796-802.
2. Walboomers J, Jacobs M, Manos M, Bosch X, Kummer A, Meijer L, Muñoz N. Human Papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J Pathol* 1999,189:12-19.
3. Herrero R, Hildesheim A, Bratti, Hutchinson M, Balmaceda I, Plummer M, Schiffman D. Population-based study of human papillomavirus infection and cervical neoplasia in rural Costa Rica. *J Natl.Cancer Inst* 2000,92:464-474.
4. Villa L, Sichero L, Rahal P, Ferenczy A, Rohan T, Franco D. Molecular variants of human papillomavirus types 16 and 18 preferentially associated with cervical neoplasia. *J Gen Virol* 2000,81:2959-2968.
5. Meissner D. Nucleotide sequences and further characterization of human papillomavirus DNA present in the CaSki, SiHa and HeLa cervical carcinoma cell lines. *J Gen Virol* 1999,80:1725-1733.
6. zur Hausen H. Papillomaviruses causing cancer: evasion from host-cell control in early events in carcinogenesis. *J Natl Cancer Inst* 2000,92:690-698.
7. Cullen A, Reid R, Campion, Lorincz. Analysis of the physical state of different human papillomavirus DNAs in intraepithelial and invasive cervical neoplasm. *J Virol* 1991,65:606-612.
8. Klaes S, Woerner S, Ridder R, Wentzensen N, Duerst M, Schneider A, Lotz B, Melsheimer P, Doeberitz M. Detection of high risk cervical intraepithelial neoplasia and cervical cancer by amplification of transcripts derived from integrated papillomavirus oncogenes. *Cancer Res* 1999,59:6132-6136.
9. Bos A, van Ballegooijen M, van Orortmarsen, van Marle M, Habbena J, Lynge R. Non-progression of cervical intraepithelial neoplasia estimated from population-screening data. *Br J Cancer* 1997,75:124-130.
10. Ho G, Bierman R, Beardsley L, Chang C, Burk L. Natural History of cervicovaginal papillomavirus infection in young women. *N Engl J Med* 1998,338:423-428.

11. Kjaer S, van der Brule A., Bocj J, Poll P, Engholm G, Sherman M, Walboomers J, Meijer F. Determinants for genital human papillomavirus (HPV) infection in 1000 randomly chosen young Danish women with normal Pap smear: are there different risk profiles for oncogenic and non oncogenic HPV types?. *Cancer Epidemiol Biomark Prev* 1997;6:799-805.
12. McBride A, Romanczuk H. Howley. The papillomavirus E2 regulatory proteins. *J Biol Chem* 1991;266:18411-14.
13. Scheffner N, Huibregste J, Viestra. D, Howley P. The HPV-16 E6 and E6-AP complex functions as ubiquitin protein ligase in ubiquitination of p53. *Cell* 1993;75:495-505.
14. Werness B, Levine J, Howley P. Association of human papillomavirus types 16 and 18 E6 proteins with p53. *Science* 1990;248:76-79.
15. DeRisi J, Penland L, Brown P, Britter M. Meltzer P, Chen Y, Trent J. Use of a cDNA microarray to analyse gene expression patterns in human cancer. *Nat Genet* 1996;4:457-460.
16. Brown P, Botstein R. Exploring the new world of the genome with DNA microarrays. *Nature Genet* 1999;21 (suppl):33-37.
17. Perou C, Sorlie T, Eisen M, Rijn M, Jeffrey S, Rees C, Pollack J, Ross, Johnsen H, Akdien L, Fluge O, Pergamenschikov A, Williams C, Zhu S, Lenning P, Brown P, Botstein D. Molecular portrait of human breast tumours. *Nature* 2000;406:747-752.
18. Shim C, Zhang, Hun C, Lee. Profiling of differentially expressed genes in human primary cervical cancer by complementary DNA expression array. *Clin Cancer Res* 1998;4:3045-3050.
19. Nees M, Geoghegan J, Hyman T, Frank S, Miller L, Woodworth R. Papillomavirus type 16 oncogenes downregulate expression of Interferon responsive genes and upregulate proliferation associated and NF-B-Responsive Genes in Cervical Keratinocytes. *J Virol* 2001;75:4283-4296.
20. Garcia AM, Ortíz-NavarreteVF, Mora-García ML, Flores-Borja F, Díaz-Quinonez A, Isibasi-Araujo A, Trejo-Becerril C, Chacon-Salinas R, Hernandez-Montes J, Granados-Arreola J, de Leo C, Weiss-Steider B. Identification of peptides presented by HLA class I molecules on cervical cancer cells with HPV-18 infection. *Immunol Lett* 1999;67:167-177.
21. Caceres-Cortes JR, Alvarado-Moreno JA, Waga K, Rangel-Corona R, Monroy-Garcia A, Rocha-Zavaleta L, Urdiales-Ramos J, Weiss-Steider B, Haman A, Hugo

- P, Brousseau R, Hoang T. Implication of tyrosine kinase receptor and steel factor in cell density-dependent growth in cervical cancers and leukemia. *Cancer Res* 2001;61:6281-6289.
22. Lopez-Marure R, Gutierrez G, Mendoza C, Ventura JL, Sanchez L, Reyes Maldonado E, Zentella A, Montano LF. Ceramide promotes the death of human cervical tumor cells in the absence of biochemical and morphological markers of apoptosis. *Biochem Biophys Res Commun* 2002;293:1028-1036.
 23. Monroy G, Trejo B, Charon S, Hernández-Montes J, and Granados P. Identification of peptides presented by HLA class I molecules on cervical cancer cells with HPV infection. *Immunol Lett* 1999;67:167-177.
 24. Eisen M, Spellman P, Brown P, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci* 1998;95:14863-14868.
 25. Chen Y, Miller C, Mosher R, Zhao X, Deeds J, Morrissey M, Bryant B, Yang D, Meyer R, Cronin F, Gostout BS, Smith-McCune K, Schlegel R. Identification of cervical cancer markers by cDNA and tissue microarrays. *Cancer Res* 2003;63:1927-1935.
 26. Cheng Q, Lau W, Chew SH, Ho TH, Tay SK, Hui KM. Identification of molecular markers for the early detection of human squamous cell carcinoma of the uterine cervix. *Br J Cancer* 2002;86:274-281.
 27. Kishimoto T. The biology of Interleukin-6. *Blood* 1998;74:1-10.
 28. Rochter H, Holley R, Andrews W, Owen J, Miller K. The association of interleukin 6 with clinical and laboratory parameters of acute pelvic inflammatory disease. *Am J Obstet Gynecol* 1999;181:940-944.
 29. Espevik T, Wange S, Faxvaag A, Shalaby M. Regulation of interleukin-2 and interleukin-6 production from T-cells: involvement of interleukin-1 beta and transforming growth factor beta. *Cell Immunol* 1990;126:4756.
 30. Eustace D, Han X, Gooding R, Riches P, Hyderman E. Interleukin-6 functions as an autocrine growth factor in cervical carcinoma in vitro. *Gynecol Oncol* 1993;50:15-19.
 31. Tjong M, van der Vange N, Kate F, Tjong SP, ter Schegget J, Burger MP, Out TA. Increased IL-6 and IL-8 levels in cervicovaginal secretions of patients with cervical cancer. *Gynecol Oncol* 1999;73:285-291.
 32. Wei L, Kuo M, Chen Ch, Cheng C, Hsieg C. Interleukin-6 in cervical cancer: the relationship with vascular endothelial growth factor. *Gynecol Oncol* 2001;82:49-56.

33. Woodworth C, Simpson S. Comparative lymphokine secretion by cultured normal human cervical keratinocytes, papillomavirus-immortalized and carcinoma cell lines. *Am J Pathol* 1993,142:1544-1555.
34. Iglesias M, Plowman G, Woodworth C. Interleukin-6 and interleukin-6 soluble receptor regulate proliferation of normal, human papillomavirus immortalized and carcinoma derived cervical cells in vitro. *Am J Pathol* 1995,146:944-52.
35. Angelo L, Talpaz M, Kurzrock R. Autocrine Interleukin-6 production in renal cell carcinoma: evidence for the involvement of p53. *Cancer Res* 2002,62:932-40.
36. Sabthanum U, Ray A, Sehgal P. Repression of the interleukin 6 promoter by p53 and the retinoblastoma susceptibility gene product. *Proc Natl Acad Sci* 1991,88:7605-7609.
37. Margulies L, Sehgal P. Modulation of the human interleukin 6 promoters and transcription factor C/EBP β activity by p53 species. *J Biol Chem* 1993,268:15096-15100.
38. Seto E, Usheva A, Zambetti G, Momand J, Horikoshi N, Weinmann R, Levine AJ, Shenk T. Wild type p53 binds to the TATA binding protein and represses transcription. *Proc Acad Sci* 1992,89:12028-12032.
39. Leung D, Cachiane G, Kuang W, Goedel D, Ferrara N. Vascular endothelial growth factor is a secreted angiogenic mitogen. *Science* 1989,246:1306-1309.
40. Kramer F, White K, Kubbies M, Swisshelm K, Weber BH. Genomic organization of claudin-1 and its assessment in hereditary and sporadic breast cancer. *Hum Genet* 2000;107:249-256.
41. Ye L, Martin TA, Parr C, Parr C, Harrison GM, Mansel RE, Jiang WG. Biphasic effects of 17-beta-estradiol on expression of occludin and transendothelial resistance and paracellular permeability in human vascular endothelial cells. *J Cell Physiol* 2003;196:362-369.
42. Hoerel T, Macek R, Swisshelman R, Kubbies M. Reexpression of the TJ protein CLDN1 induces apoptosis in breast tumor spheroids. *Int J Cancer* 2004;108:374-383.
43. Hoerel T, Macek R, Mundigl O, Swisshelm K, Kubbies M. Expression and targeting of the tight junction protein CLDN1 in CLDN1-negative human breast tumor cells. *J Cell Physiol* 2002;191:60-68.

44. Miwa N, Furuse M, Tsukita S, Niikawa N, Nakamura Y, Furukawa Y. Involvement of claudin-1 in the beta-catenin/Tcf signaling pathway and its frequent upregulation in human colorectal cancers. *Oncol Res* 2000;12:469-476.
45. Su TH, Chang JG, Yeh KT, Lin TH, Lee TP, Chen JC, Lin CC. Mutation analysis of CTNNB1 (beta-catenin) and AXIN1, the components of Wnt pathway, in cervical carcinomas. *Oncol Rep* 2003;10:1195-1200.
46. Sathyanarayana UG, Padar A, Suzuki M, Maruyama R, Shigematsu H, Hsieh JT, Frenkel EP, Gazdar AF. Aberrant promoter methylation of laminin-5-encoding genes in prostate cancers and its relationship to clinicopathological features. *Clin Cancer Res* 2003;9:6395-6400.
47. Manda R, Kohno T, Niki T, Yamada T, Takenoshita S, Kuwano H, Yokota J. Differential expression of the LAMB3 and LAMC2 genes between small cell and non-small cell lung carcinomas. *Biochem Biophys Res Commun* 2000;275:440-445.
48. Sathyanarayana UG, Maruyama R, Padar A, Suzuki M, Bondaruk J, Sagalowsky A, Minna JD, Frenkel EP, Grossman HB, Czerniak B, Gazdar AF. Molecular detection of noninvasive and invasive bladder tumor tissues and exfoliated cells by aberrant promoter methylation of laminin-5 encoding genes. *Cancer Res* 2004;64:1425-1430.
49. Skyldberg B, Salo S, Eriksson E, Aspenblad U, Moberger B, Tryggvason K, Auer G. Laminin-5 as a marker of invasiveness in cervical lesions. *J Natl Cancer Inst* 1999;21:1882-1887.
50. Kohlberger P, Beneder Ch, Horvat R, Horvat R, Leodolter S, Breitenacker G. Immunohistochemical expression of laminin-5 in cervical intraepithelial neoplasia. *Gynecol Oncol* 2003;3:391-394.
51. Habermann JK, Hellman K, Freitag S, Heselmeyer-Haddad K, Hellstrom AC, Shah K, Auer G, Ried T. A recurrent gain of chromosome arm 3q in primary squamous carcinoma of the vagina. *Cancer Genet Cytogenet* 2004;148:7-13.
52. Human Papillomaviruses Compendium. 1994. Los Alamos National Laboratory Theoretical Biology and Biophysics.
53. Sheu BC, Lien HC, Ho HN, Lin HH, Chow SN, Huang SC, Hsu SM. Increased expression and activation of gelatinolytic matrix metalloproteinases is associated with the progression and recurrence of human cervical cancer. *Cancer Res* 2003;63:6537-6542.

54. Woo JH, Lim JH, Kim HY, Suh SI, Min do S, Chang JS, Lee YH, Park JW, Kwon TK. Resveratrol inhibits phorbol myristate acetate-induced matrix metalloproteinase-9 expression by inhibiting JNK and PKC delta signal transduction. *Oncogene* 2004;23:1845-1853.
55. Gaiotto MA, Focchi J, Ribalta JL, Stavale JN, Baracat EC, Lima GR, Guerreiro da Silva ID. Comparative study of MMP-2 (matrix metalloproteinase 2) immune expression in normal uterine cervix, intraepithelial neoplasias, and squamous cells cervical carcinoma. *Am J Obstet Gynecol* 2004;190:1278-1282.
56. Muller, D, Quantin B, Gesnel M, Millon-Collard R, Abecassis J, Breathnach R. The collagenase gene family in humans consists of at least four members. *Biochem J* 1988;253:187-192.
57. Sato T, Sakai T, Noguchi Y, Takita M, Hirakawa S, Ito A. Tumor-stromal cell contact promotes invasion of human uterine cervical carcinoma cells by augmenting the expression and activation of stromal matrix metalloproteinases. *Gynecol Oncol* 2004;92:47-56.

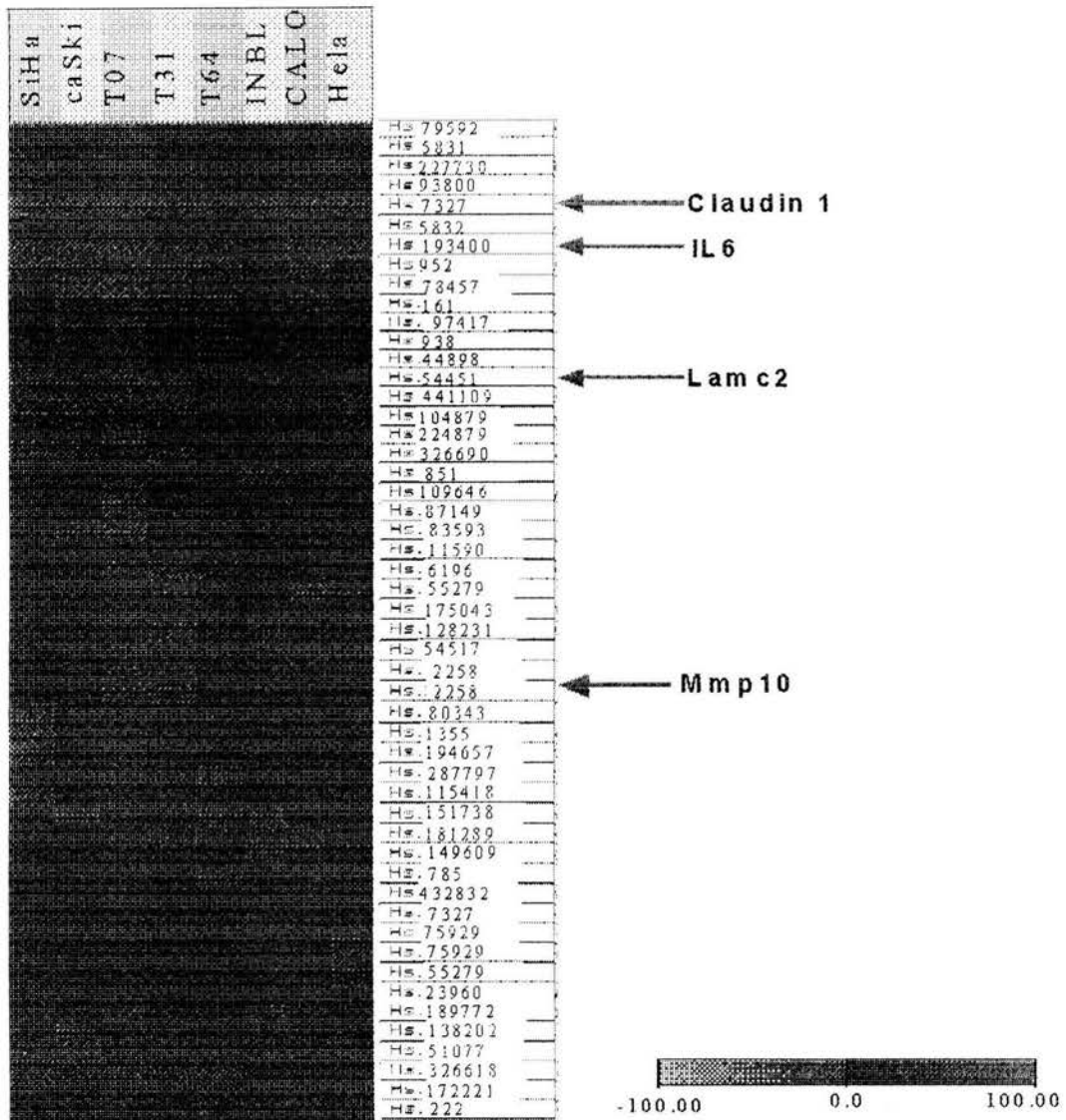


Figure 1. Gene cluster of the most overexpressed genes in cervical cancer.

Expression patterns of the 50 most overexpressed genes in cervical cancer. Rows represent individual genes, columns represent individual samples. Each cell in the matrix represents the expression level of a single transcript compared by the normal transcript. The red and green colours indicate transcript levels above and below the normal for that gene across all samples, respectively. Color saturation is proportional to magnitude of difference from the normal mean. Blue flagged genes were used for the validation of the expression by semi-quantitative RT-PCR.

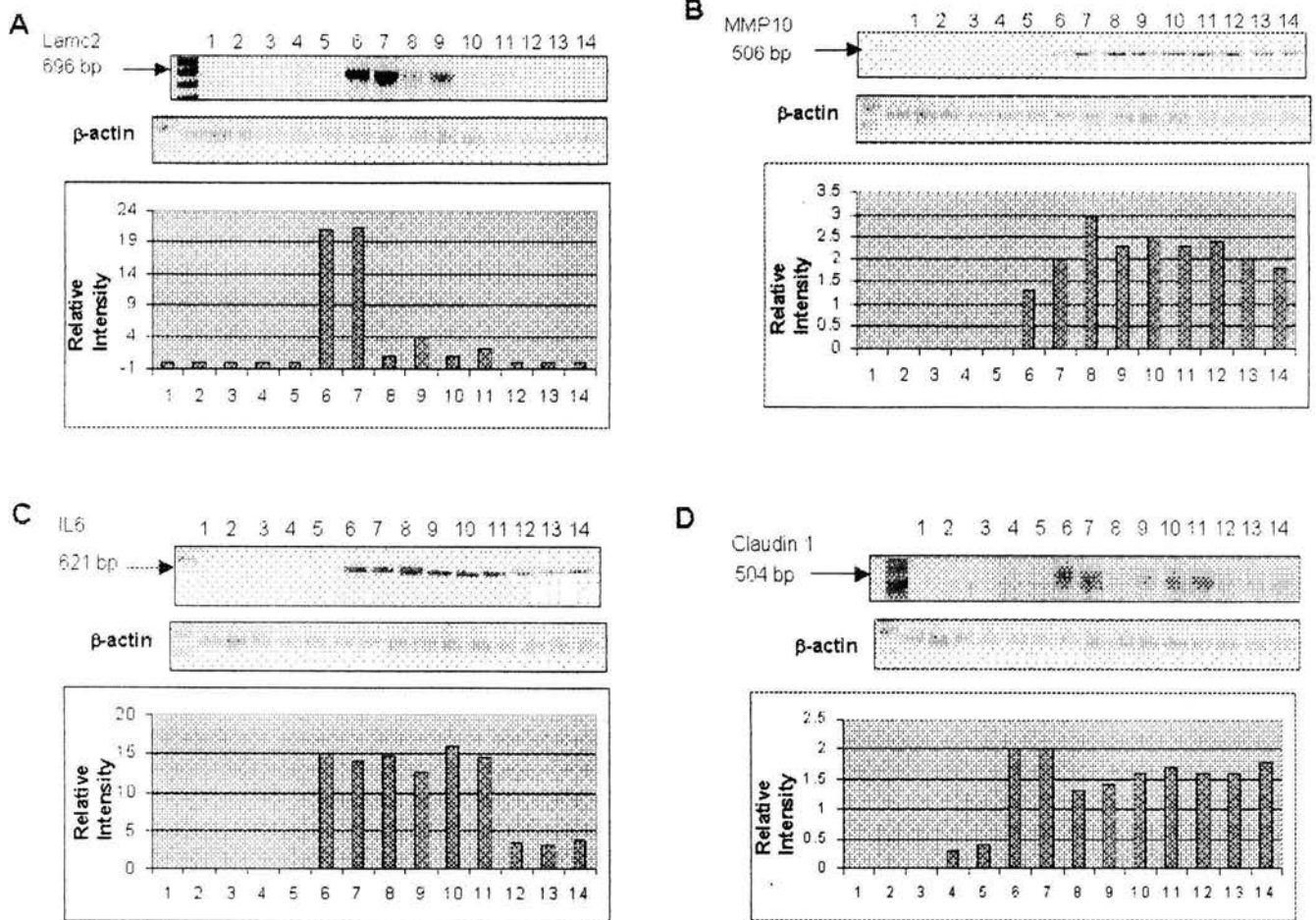


FIGURE 2. Relative mRNA expression levels of Lamc2, MMP210, IL6 and Claudin1 in cervical tumors detected by RT-PCR. One hundred of total RNA were employed for RT-PCR experiments A) Represents RT-PCR products for Lamc2 and graphic representation of the relative mRNA expression levels for the used samples, B) MMP10 expression levels, C) IL6 gene expression and D) Claudin 1 mRNA levels. Values are presented as ratio of densitometric readings of samples to corresponding β -actin samples. The used samples were; line 1) negative control, 2) normal HPV- tissue #N07, 3) normal HPV- tissue # N11, 4) normal HPV- tissue # N20, 5) normal HPV-tissue #N23, 6) SiHa, 7) CaSki, 8) tumoral sample # T04, 9) tumoral sample # T24, 10) tumoral sample # T43, 11) tumoral sample # T45, 12) INBL, 13) CALO and 14) HeLa. A 100 bp ladder was used as molecular weight marker.

UniGene Cluster ID	Gene Name or Product
Hs.512807	Aldo-keto reductase family 7, member A2
Hs.78457	Solute carrier family 25 (mitochondrial carrier, ornithine transporter) member 15
Hs.952	Solute carrier family 10 (sodium/bile acid cotransporter family), member 1
Hs.44898	Similar to cadherin 12, type 2 preproprotein;
Hs.441109	Peroxisomal biogenesis factor 19
Hs.104879	Serine (or cysteine) proteinase inhibitor, clade B (ovalbumin), member 9
Hs.109646	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 6, 17kDa
Hs.87149	Integrin, beta 3 (platelet glycoprotein IIIa, antigen CD61)
Hs.512234	IL6: Interleukin 6 (interferon, beta 2)
Hs.6196	Integrin-linked kinase
Hs.7327	Claudin 1
Hs.55279	Serine (or cysteine) proteinase inhibitor, clade B (ovalbumin), member 5
Hs.54517	Ficolin (collagen/fibrinogen domain containing lectin) 2 (hucolin)
Hs.54451	Laminin, gamma 2
Hs.194657	Cadherin 1, type 1, E-cadherin (epithelial)
Hs.115418	Cadherin 16, KSP-cadherin
Hs.189772	Chaperonin containing TCP1, subunit 2 (beta)
Hs.2258	Matrix metalloproteinase 10 (stromelysin 2)
Hs.222	Integrin, alpha 9
Hs.1355	Cathepsin E

Table 1. Top 20 overexpressed genes in HPV+ transformed cells. Overexpressed genes at >10 folds in any HPV+ tissue obtained by a gene expression profile of SiHa, CaSki HeLa, CALO, INBL and three HPV16+ invasive samples) and compared to normalized median of 3 normal HPV- tissue samples by using Gene Spring software. UniGene ID are listed.

Overexpression of Cathepsin F, Matrix Metalloproteinases 11 and 12 in Cervical Cancer

Guelaguetza Vázquez-Ortiz¹, Patricia Piña-Sánchez¹, Karla Vázquez¹, Alfonso Dueñas², Lucia Taja², Patricia Mendoza¹, José A. García³, Mauricio Salcedo^{1§}.

¹ Oncogenomics Laboratory, Oncology Research Unit, Oncology Hospital, National Medical Center SXXI-IMSS.

² Division of Basic Research, INCAN, SS.

³ Laboratory of Theoretical Biology, Research Department, La Salle University, Mexico.

§ Corresponding author: Mauricio Salcedo

E-mail addresses:

GVO: guelav@yahoo.com

PPS: patricia_1307@yahoo.com.mx

KV: sotaliak@hotmail.com

AD: aldueñas@prodigy.net.mx

LT: ltaja_chayeb@yahoo.com

PM: pamelomx@yahoo.com

JAG: jgarcia@ci.ulsal.mx

MS: maosal89@yahoo.com

Abstract

Background: Cervical carcinoma (CC) is one of the most common cancers among women worldwide and the first cause of death among the Mexican female population. CC progression shows a continuum of neoplastic transitions until invasion. Matrix metalloproteinases (MMPs) and cathepsins play a central role on the enhancement of tumor-induced angiogenesis, cell migration, proliferation, apoptosis and connective tissue degradation. MMPs -2 and -9 expression has been widely studied in cervical cancer, nevertheless, no other metalloproteinases or cathepsins have been yet related with the progression and/or invasion of this cancer.

Methods: Three HPV18 CC cell lines, two HPV16 CC cell lines and three HPV16 tumor CC tissues were compared with three morphologically normal, HPV negative cervical specimens by cDNA arrays. Overexpression of selected genes was confirmed by semiquantitative reverse transcription-PCR. In situ hybridization and protein expression of selected genes was further studied using two tissue microarrays, one consisting of 10 HSIL and 15 CC and the other one consisted of 15 normal cervical and 10 LSIL tissues.

Results: By cDNA arrays *TIMP1*, *Integrins alpha 1 and 4*, *cadherin 2 and 11*, *Cathepsins F, B L2*, *MMP 9, 10 11 and 12* were upregulated and *Cathepsin S, L, H and C*, *Cadherins 3 and 4*, *TIMP3*, *MMP 13*, *Elastase 2* and *Integrin beta 8* were found to be downregulated. Semiquantitative RT-PCR gave consistent results with the cDNA array findings for all three genes selected to be studied (*CTSF*, *MMP11* and *MMP12*). In situ hybridization of three genes showed overexpression in all the HSIL and CC. Two selected proteins were detected in LSIL, HSIL and CC by immunohistochemistry.

Conclusions: Several novel and previously undetected CC promoting genes have been identified. Increased transcription of these genes may result in overexpression of proteins, such as *CTSF*, *MMP11* and *MMP12* which may contribute to the pathogenesis of CC.

Background

Cervical carcinoma (CC) is one of the most common cancers among women worldwide and the first cause of death among the Mexican female population[1]. High-risk human papillomavirus (HPV) infection is considered to be the most important risk factor associated with the development of this tumor, and is present in the 99.7% of invasive cervical tumors worldwide[2]. CC progression shows a continuum of neoplastic transitions from Low grade squamous intraepithelial lesion (LSIL), High grade squamous intraepithelial lesion (HSIL), to invasive cervical cancer. Tumor invasion and metastasis are key steps in the progression of malignant tumors; this involves extracellular matrix (ECM) and basement membrane degradation[3]. Although several classes of proteolytic enzymes are involved in these processes[4], Matrix metalloproteinases (MMPs) play a central role on the enhancement of tumor-induced angiogenesis, cell migration, proliferation, apoptosis and connective tissue degradation[5-6]. Different MMP family members have been identified, including collagenases (MMP-1, -8 and -13), gelatinases (MMP-2 and -9), stromelysins (MMP-3, -10, -11), membrane associated (MMP-14, -15, -16, -17, -23, -24, -25) and other kind of MMPs like the Metalloelastase (MMP-12)[7]. MMPs -2 and -9 have been widely studied. Increased mRNA and protein levels of both MMP-2 and MMP-9 have been detected in breast, colon, pancreatic and cervical cancers[8].

Human macrophage metalloelastase (MMP-12) has been identified in alveolar macrophages of cigarette smokers as an elastolytic MMP[9]. It can degrade elastin, and other substrates, such as type IV collagen, fibronectin, laminin, gelatin, vitronectin, entactin, heparin, and chondroitin sulphates[10]. MMP-12 is overexpressed by macrophages in atherosclerotic lesions[11] and in intestinal ulcerations[12]. It has been demonstrated that besides macrophages, transformed epithelial cells can express MMP-12 in skin and vulvar cancers; MMP12 expression levels correlate with epithelial dedifferentiation and histological aggressiveness[13-14].

Stromelysin 3 ST3 (MMP-11) is another protease that can modulate cancer progression by remodelling extracellular matrix; it cleaves α 1-antitrypsin and IGF-BP1[15]. Normal ST3 expression is associated with embryogenesis and wound healing but in stressed epithelial cells; its expression is detected in the vicinity fibroblasts[16]. Malignant epithelial cells depend on environmental stromal cells including fibroblastic, endothelial and inflammatory cells to develop primary and secondary tumors. ST3 expression is observed in the area which surrounds malignant epithelial tumour cells and sometimes in tumoral cells of esophageal, oral, papillary thyroid, colorectal, skin and ovarian carcinomas[17-22]. Thus, ST3 gene expression has been reported to be associated with tumor progression[23].

Cathepsins are cysteine proteases from the papain family that have been characterized to be responsible for the terminal protein breakdown in lysosomes. Actually the human Cathepsin family is composed by Cathepsins B, C, F, H, K, L, O, S V, W and X[24]. The recently described Cathepsin F has been demonstrated; by Northern blot analysis, to be ubiquitously expressed in several tissues showing the higher expression in skeletal muscle and testis. This transcript was also found in several cancer cell lines showing that this enzyme could be involved in degradative processes during tumor progression[25].

In the present study, we first investigated the mRNA expression levels of 8 000 genes including MMPs and cathepsins by using cDNA microarrays in CC samples and cell lines. Selected genes were corroborated to be up-regulated by using semiquantitative RT-PCR. Besides, RNA and protein expression of selected genes were further examined using *in situ* hybridization and IHC applied on Tissue Microarrays (TMAs)[26]. Several genes that were either up- or down-regulated in CC were detected, showing that these genes might have an active or passive role in the malignant progression of CC.

Methods

Cell lines and Tissue Samples for cDNA arrays and RT-PCR.

HeLa (HPV18), SiHa (HPV16) and CaSki (HPV16) cell lines were grown in the recommended media. The cell lines designated CaLO (derived from Invasive squamous carcinoma Stage IIB, HPV18) and INBL (derived from Invasive Squamous carcinoma IV A, HPV18), were established at the FES Zaragoza-UNAM, Mexico[27]. Fresh biopsies of dysplastic lesions were taken during colposcopy from 20 patients at the Dysplasias National Center in Mexico City. Twenty two normal cervical tissues from pre-menopausal women were freshly dissected immediately after death at the Pathology Department of the Hospital General de Mexico, S. S. from women who died of non-gynecological related causes. All the described procedures have been evaluated and approved by the local committee of ethics of the Mexican Institute of Social Security; we also obtained written informed consent from the patients. All tissue samples were divided in three sections, the central part was snapped frozen in liquid nitrogen and stored at -70°C until nucleic acid extraction, and the extremes from the biopsies were fixed with 70% ethanol overnight and paraffin embedded at the Department of Pathology, Oncology Hospital, National Medical Center SXXI, Mexico. Serial Haematoxylin/Eosin stained sections from the extremes of the biopsies were inspected for representativity of the tissue. Carcinoma samples were considered as representative when at least 70% of all cells in the tissue section consisted of cancer cells and the normal tissue samples; when at least 80% of the sample consisted of microscopically ectocervical normal tissue besides HPV absence. All the invasive samples were diagnosed as squamous lesions.

Tissue samples for ISH and IHC.

Paraffin blocks of 35 patients: 10 HSIL, 15 CC, and 10 LSIL; were randomly collected from the archives of the Dysplasias National Center in Mexico City. Fifteen of the normal cervical tissues (*post-mortem* obtained) were formalin buffered fixed and paraffin embedded for preparation of the TMA blocks. The TMAs were constructed as described previously[28]. One of the TMAs contained 10 HSIL tissues and CC samples; the other was made with the 15 normal cervical tissues and 10 LSIL samples. The TMAs did not include the investigated samples by the cDNA arrays or by RT-PCR.

Nucleic Acids Isolation.

Both RNA and DNA were isolated from the tissue biopsies and the cervical cancer cell lines using the Trizol Reagent (Gibco BRL, Life Technologies, Grand Island NY USA) according to manufacturer's conditions. Each DNA sample was subjected to HPV typing by polymerase chain reaction (PCR) using the MY09/MY11 consensus primers and direct sequencing.

Potentially contaminating DNA was removed by treating with RNase-free DNase I (Ambion Inc. Austin Texas USA). The resulting RNA concentration was measured spectrophotometrically and the quality of both nucleic acids was confirmed on agarose gels.

cDNA probe preparation.

cDNA was synthesized according to Ambion Advantage System User Manual. From all the obtained samples only those which met the criteria (squamous cell carcinoma, HPV type 16 or 18, sufficient RNA available, at least 70% cancer cells in the biopsy) were hybridized. These samples were: cell lines HeLa, SiHa, CaSki, CaLO and INBL, three HPV16+ tumor samples (T07, T31, T64) and three HPV- normal samples (N03, N11, N22);

unfortunately no HPV18+ tumor samples were included for array hybridization due to RNA insufficient RNA quantity. The cDNA probes were obtained by RT-PCR from 500 ng of RNA in the presence of (α - ^{32}P) dATP (3000Ci/mmol NEN). Briefly, RNAs were denatured at 70°C for 10 min and cDNAs were synthesized at 42°C using oligo-dT priming in a final volume of 30 μl . The labelled cDNAs were purified by spin column chromatography. The final incorporation of the radiolabel was approximately 1×10^9 cpm for each reaction.

Hybridization and Array Analysis.

ULTRArray Advantage System array blots (Ambion Inc. Austin Texas USA) which contain 8400 genes, were prehybridized at 68°C for at least 90 min before probe addition in Hybridization buffer. Finally, 1.5×10^7 cpm of each labeled cDNA was added to the buffer. Hybridization was performed at 60°C in a rolling bottle overnight. The arrays were washed twice with 2X SSC and 0.5 % SDS at 60°C for 30 min; followed by two stringent washes with 0.5X SSC, 0.5% SDS at the same temperature and time. Finally damp arrays were sealed in plastic wrap and exposed to imaging plates (BASMP 2040S; Fuji, Nakamura, Japan) for 24 hours, which were then scanned with a STORM 860 phosphorimager/fluorimager (Amersham Biosciences, Buckinghamshire UK) to obtain 16-bit images. ArrayVision software (Image Research Inc) was used for the analysis. The artifacts were eliminated and the intensity of each spot was analyzed after background subtraction. Then, data was normalized and analyzed by using limma bioconductors and marray package[29-30] in all genes presented in the array. The average of the normal array (three normal samples) was obtained. Frequency of positive cases was obtained by comparing the average normal array with each individual tumor or cell line array. The threshold frequency was set to 40% and the intensity ratio (tumor;normal) threshold values of 4 up-regulation and -1 for down-regulation present in all samples were used in attempt to detect significant expression changes.

RT-PCR Analysis.

Estimation of the reliability of the cDNA results for the three selected genes was performed by semiquantitative RT-PCR on four normal samples, 5 cell lines, 3 HPV16 and 2 HPV18 tumors; displaying a clear change in their expression in respect to the normal, independently of the HPV type present. Two-hundred ng of total RNA from all samples were reverse transcribed using the RT-PCR Access System (Promega, Madison, WI). Primers for the genes of interest were: Cathepsin F (CTSF) 396 bp, $T_M = 67^\circ\text{C}$, (5'-GTGCTGATCAGAGTGGCTGCTGC-3' and 5'-AGTTTCCYGGACATGGAGAGGGAC-3'); Matrix Metalloproteinase 12 (MMP12), 370 bp $T_M = 55^\circ\text{C}$ (5'-TCACGAGATTGGCCATTCTT-3' and 5'-TCTGGCTTCAATTCATAAGC-3')[31]; and Matrix Metalloproteinase 11 (MMP11/ STMY3), 399 bp, $T_M = 66^\circ\text{C}$, (5'-CCATGGCAGTTGGTGCAGGAGCAG-3' and 5'-TGCAGTCATCTGGGCTGAGACTCA-3'); and β -actin: (5'-TGAAGTCTGACGTGGACATC-3' and GTTCGTTCTCATACTGCTCA-3') 243 bp, $T_M = 55^\circ\text{C}$ primers were designed using Integrated DNA technologies Biotools software (<http://biotools.idtdna.com/gateway>). The RT-PCR conditions were: for the first strand synthesis of cDNA 48°C for 45 min and 94°C for 2 min to denature template; and for second strand synthesis and DNA amplification: 94°C for 30 sec, (specific $T_M^\circ\text{C}$ for each set of primers) for 1 min, and 68°C for 2 min for a total of 24 cycles, followed by a single step at 68°C for 7 min. The products were visualized on 1.5% agarose gels stained with ethidium bromide, and signals were quantified by densitometry using MetaView analysing system (version 4.5 Universal Imaging Corp., USA). MMP11, MMP12 and CTS expression was standardized to β -actin expression assessed from the same cDNA in separate PCR reactions and run in parallel on separate gels. The standardized mean of

each triplicate PCR was then expressed relative to the levels in β -actin cDNA. Statistical comparisons among groups were analyzed by a Kruskal Wallis-test. A *P* value of <0.05 was considered as statistically significant.

In Situ Hybridization.

Briefly, five-micron tissue sections were obtained from TMA paraffin blocks, deparaffinized and rehydrated in a graded ethanol series (100, 90, 70, and 30%), and transferred to PBS solution (137 mM NaCl, 2.7 mM KCl, 4.3 mM Na₂HPO₄, 1.4 mM KH₂PO₄) for 10 min. Tissues were treated with DNase solution (1 μ g/ml) for 10 min at 37°C and washed three times in PBS. Endogenous peroxidase inactivation was carried out by incubating the samples in hydrogen peroxide in methanol for 3% hydrogen peroxide (H₂O₂) for 40 min. Sense and antisense probes for CTSF, MMP11 and 12 genes were generated by single-strand PCR using specific cDNA obtained from SiHa cells RNA as template and labeled with biotin-16-dUTP (Roche). Each tissue was covered with 50 μ l of hybridization cocktail and a coverslip. The hybridization cocktail consisted of 50% formamide, 10% dextran sulfate, 2X SSC (20X SSC: 3 M NaCl, 300 mM Na₃C₆H₅O₇), PBS, 2% SDS, 100 μ g/ml sonicated salmon sperm and 50 ng of the dUTP-biotin-labeled probe (sense or antisense). Both the RNA of the tissue and probe were denatured at 65°C for 10 min. After hybridization, the coverslips were soaked off in Tris-buffered saline with Tween buffer 1X (TBST 10X: 500 nmol/L TrisHCL, pH 7.6, 3 mol/L NaCl, 1% Tween 20), and the tissues were incubated at 55°C for 20 min in stringent wash solution. The horseradish peroxidase labelled antibody (SA-HRP) (GenPoint System from DAKO) was immediately applied on the tissues and incubated for 15 min in a humidifier chamber. The sections were washed in TBST 1X. Biotinyl tyramide was applied on the tissue sections for signal amplification, for 15 min at room temperature, and washed in TBST. A second SA-HRP step was carried out and the color reaction was developed with 0.06% diaminobenzidine (DAB) in 3% H₂O₂. Finally the slides were washed, haematoxylin/counterstained, dehydrated in graded ethanol and mounted. Negative controls for ISH were carried out using the sense probes or a treatment with RNase solution (100 mg/ml for 30 min at 37°C) prior ISH. Levels of expression were scored semi-quantitatively in tissue sections, under the light microscope. Scores were obtained by estimating average signal intensity (scale of 0 to 3) and the proportion of epithelial cells showing a positive signal (0-100%).

Immunohistochemical assays.

Briefly, the tissue sections were deparaffinized and rehydrated in a graded ethanol series. Endogenous peroxidase inactivation was carried out as described in IHC, and the slides were preincubated with DAKO Protein Block Serum free Cat X0909 (DAKO, Carpinteria, CA) for 30 min at 37°C to prevent non specific immunoreaction. Excess was decanted and tissues were incubated with the primary antibody as follows MMP 11 (Biomedica, Cat V10221, Lot 10441, 1:100), MMP12 (R&D Systems, Cat MAB917, Lot AGE022051, 1:100) at 37°C for 30 min. DAKO Envision System Peroxidase from DAKO was applied and the slides and then counterstained with hematoxylin. A rough scoring was done to quantitate the intensity of the staining by two authors (GVO and PPS). Levels of MMP11 and MMP12 expression were semi-quantitatively scored in tissue sections, under the light microscope. Scores were obtained by estimating average signal intensity (scale of 0 to 3). Staining was considered (+++), when all cells presented positive staining, (++) result was considered when at least half of the analyzed cells presented positive staining, and (+) when a weak staining was present.

Results

Changes in cDNA Expression in CC.

The analysis of the three normal cervical tissues by limma and marray package showed that the expression profiles were similar. Thus, the data of the histologically normal tissue samples were pooled to generate an average normal tissue array. The data from the cell lines and tissue arrays were compared with the average normal tissue arrays. The frequency of increased or decreased gene expression changes was determined by comparing the average normal tissue array with each of the cell lines and tissue arrays. The cut-off values were 4 fold over the normal for up-regulation, and -1 fold for down regulation or suppressed genes. Several genes were over expressed (*TIMP1*, *Integrins alpha 1 and 4*, *cadherin 2 and 11*, *Cathepsins F, B L2*, *MMP 9, 10 11 and 12*), and 10 genes were down-regulated (*Cathepsin S, L, H and C*, *Cadherins 3 and 4*, *TIMP3*, *MMP 13*, *Elastase 2 and Integrin beta 8*) Table 1.

Verification of cDNA Expression Array Results by RT-PCR.

The gene expression profile findings in cDNA arrays were confirmed by semiquantitative RT-PCR for selected genes (*CTSF*, *MMP11*, *MMP12*, and β -*actin*; Fig. 1). In general, these analyses showed consistent significant results with those obtained by cDNA arrays. All 8 investigated samples using cDNA arrays showed increased *CTSF* expression, and in line with this, the tumor band intensities in agarose gels as compared with the control band intensity were increased from an average of 1.2 to 7 fold in all tumors for *CTSF* in the RT-PCR analyses. All studied samples had increased *MMP11* expression in cDNA microarray analysis, and presents an average of 6 times increased expression in RT-PCR. The greatest difference between the two methods was found in *MMP12* expression. All studied samples showed increased *MMP12* expression in a cDNA array analysis (average of 15 folds), whereas samples in RT-PCR analyses showed an average of about six times higher expression than the normal controls. All comparisons were statistically significant.

ISH could be performed on 90% of the HSIL and 100% of CC cases. *CTSF* was expressed in 9 (90%), *MMP11* in 9 (90%), and *MMP12* in 9 (90%) of the 10 HSIL studied in the array. A strong (++++) but focal signal was observed in all CC for the three genes (100%), and the expression in this tissues were higher than in the LSIL (++, +) and normal tissues. The three genes stained in 90% of all LSIL samples (+) while, the signal was weaker or undetectable in the histological normal cervical tissues (Table 2, Fig. 2). Assessment of the sections at high magnification and correlation with serial sections stained only by H&E, revealed that the *CTS*, *MMP11* and *MMP12* expression was usually confined to the cytoplasm of a subset of epithelial tumor cells that showed morphological features of squamous differentiation.

Immunostaining signals for *MMP11* and *MMP12* proteins were observed in all the LSIL samples, in 80% of HSIL and all CC samples. As expected, there was a coincidence in the signal location between immunodetection and ISH assays. There was not immunostaining detection in the normal cervical tissues for any protein (Table 2, Figure 3).

Discussion

In order to identify novel genes that might be associated with CC development, we used a gene expression array based on the cDNA expression of more than 8 000 genes and subsequently studied mRNA and protein expression of three upregulated genes by using ISH and IHC on a tumor tissue array. In the gene expression analysis, 14 genes were found to be expressed only in tumor tissue, and a total of 24 genes showed significant changes in their expression levels. Yet, the molecular genetic mechanisms of how these genes might be involved in the genesis of CC are still being studied. The proteins encoded by the identified genes with altered expression, have variable functions. The main categories include: a) proteolysis/peptidolysis (*TIMP1*, *Cathepsins L2, B, F, S, L, H, C*, *MMP12*, *Elastase 2 neutrophil and TIMP3*); b) protein binding (*Integrin alpha 9*, *Serine B member 9*, *Integrin beta 8*); c) extracellular matrix (*TIMP1*, *TIMP3*, *MMPs 9, 10, 11, 12 and 13*); it is worth to note that all these categories are involved in metastatic processes. To validate the overexpression of CTSF, MMP11 and MMP12 in CC, semiquantitative-RT-PCR was performed. As expected, an up-regulation was detected CC tissues and cell lines. The ISH results of the selected genes (CTSF, MMP11 and MMP12) were in line with results obtained by RT-PCR. It is worth important to note that ISH results show increased expression through different stages of cervical carcinogenesis, except in one case of LSIL. The lack of expression in some HSIL tissues might indicate a possible regression of the disease in these cases, nevertheless, a clinical surveillance is still in process to accept or reject this proposal. All transcripts were detected in CC, suggesting that these genes are essentially active in the invasive processes of cervical cancer.

Besides, MMP11 and MMP12 protein expression was also investigated by IHC and the obtained results corroborated the previous ISH results.

Many normal biological processes, such as reproduction, fetal development and wound healing, are critically dependent on controlled degradation of extracellular matrix (ECM) macromolecules. However, excessive degradation of matrix components also occurs in pathologic tissue destruction like in cancer[32]. MMP11 expression in the immediate vicinity or in the cancer cells has been associated with some human carcinomas. Nevertheless, MMP11 is a consistently active partner of invading cancer cells, but its function differs throughout cancer progression, being a tumor enhancer or repressor in several processes which leads to local or distal invasion and it favours cancer survival in the stromal environment)[33]. Our study demonstrates MMP11 expression in cervical precancerous lesions, with significantly higher expression in invasive carcinomas compared to the histologically normal cervix tissues analyzed. In the current study, positive immunoreaction in cervical tissues was detected in the cytoplasmic region of epithelial tumor cells which differs from other tumoral tissues where MMP11 expression is restricted to stromal cells which surrounds the neoplastic ones)[34]. MMP11 expression has been reported in preinvasive bronchial lesions and in carcinomas. ST3 expression in bronchial lesions was shown to start with dysplasia and carcinoma *in situ* though it was more frequent in invasive lesions just as it seems to be happening in CC[35]. Epithelial expression of MMP11 in intraepithelial squamous lesions suggest that this protein expression just like in squamous lung carcinomas may be related to the progression of phenotypic alterations acquired early during the malignant transformation pathway of cervical epithelium and maintained after invasion[36]. It has been previously reported that when MMP11 is increased in tumorigenesis this is not a result of increased neoangiogenesis or cancer cell proliferation but from decreased cancer cell death through apoptosis and necrosis. These data suggest that MMP11 might have varied roles and may act trough different pathways at different stages of cervical tumor development and progression which need to be further studied in CC.

Two other genes that were found to be overexpressed in this study were MMP9 and TIMP1 (Tissue inhibitor of Metalloproteinase 1, Table 1). Interestingly, MMP9 up-regulation in cervical cancer was previously reported by Van Trappen[37], and the combined overexpression of MMP9, MMP11 and TIMP1 has been associated with the invasive features of some cancers[38]. In our work, these proteins are clearly associated with CC but further studies should be done to analyze the impact of MMP9, MMP11 and TIMP1 expression on CC progression.

The biological role of MMP12 in tumor progression is not clearly understood but it has been discussed to be involved in the degradation of components of the basement membrane[39]. Besides, previous reports have shown that transformed epithelial cells express MMP12 in skin and vulvar cancers and its expression correlates with epithelial dedifferentiation and invasive aggressiveness. In oral verrucous and squamous cell cancer the absence of MMP-12 from epithelial cells has been reported to be a good prognostic marker of non-invasive oral carcinoma[40].

This is the first report demonstrating MMP12 expression in cervical dysplasias and carcinomas. The expression of MMP12 was detected in most of LSIL, HSIL, and all CC samples. By immunohistochemistry assays, MMP12 protein was detected in similar patterns as in ISH. The immunostaining was detected in the cytoplasmic region of the transformed epithelial cells. Our study suggests that the expression pattern of MMP12 might correlate with the histopathological features of cervical carcinoma.

Only little data is available on Cathepsin F. Nevertheless, the analysis of Cathepsin F expression in human tissues has revealed that it is expressed in most normal tissues; suggesting a general role for this enzyme in the lysosomal protein catabolism that takes place in all cell types. Following this, it has been demonstrated that Cathepsin F expression levels in normal tissues exhibit a large variability depending on the tissue type. However, it is remarkable, that the comparison of several human cancer cell lines respect to its normal counterpart shows an increment in its expression levels, suggesting that this enzyme might play an important role in carcinogenesis. Our results are in agreement with a previous study which revealed high levels of Cathepsin F in HeLa cells[25], but our study is the first report which relates Cathepsin F with cervical cancer. Due to a lack of availability of a specific Cathepsin F antibody, we did not performed IHC assays for this protein; however, ISH assays present a similar pattern just like MMP11 and 12.

Conclusions

In conclusion, our results show that CTSF, MMP11 and 12 are expressed only in dysplastic epithelial cells suggesting all of these proteins could be used as "potential" progression markers for cervical cancer. However, the role of these genes in the pathogenesis of CC needs to be confirmed in further studies, and the molecular mechanisms involved need to be studied in detail.

List of abbreviations:

HPV: Human Papilloma Virus, ISH: In Situ Hybridization, IHC: Immunohistochemistry, CTSF: Cathepsin F.

Competing interests:

None declared.

Authors' contributions

GVO performed all molecular assays, designed the study, data collection and analysis, and drafted the manuscript.

PPS carried out the IHC and in the tissue block collection, participated in drafting the manuscript.

KV and PM were involved in fresh tissue collection and preparation of tissue blocks.

AD, LT, were involved in carrying out the tissue analysis and data acquisition.

JAG was involved in statistical analysis and writing and revising the manuscript critically.

MS conceived and designed the study, coordinated and managed the study, performed data analysis.

All authors read and approved the final manuscript.

Acknowledgements

We are grateful to Dr. Moreno, Dr. Arenas (CMN SXXI-IMSS) and César Enrique Barba MD for their helpful support. This work was partially supported by Grants No. 34686-M and 7114 (Fondos Sectoriales, CONACyT-México). During this work GVO, PP, KV and PM were recipients of a CONACyT and IMSS fellowship.

References

1. Mexican Ministry of Health Registro Histopatológico de Neoplasias Malignas, Compendio de mortalidad y morbilidad. Secretaría de Salud, México 1998.
2. Walboomers J, Jacobs M, Manos M, Bosch X, Kummer A, Shah K, Snijders P, Peto J, Meijer C and Muñoz N Human Papilloma Virus is a necessary cause of invasive cervical cancer worldwide. *J Pathol* 1999, 189:12-19.
3. Liotta, L. A., and Kohn, E. C. The microenvironment of the tumor-host interface. *Nature (Lond.)*, 411: 375–379, 2001.
4. Egeblad, M., and Werb, Z. New functions for the matrix metalloproteinases in cancer progression. *Nature Rev* 2002, 2:163–176.
5. McCawley LJ, Matrisian LM. Matrix metalloproteinases: multifunctional contributors to tumor progression: a review. *Mol Med Today* 2000; 6:149–156.
6. Nagase H, Woessner JF Jr. Matrix metalloproteinases: a review. *J Biol Chem* 1999; 274:21 491–21 494.
7. McCawley LJ, Matrisian LM. Matrix metalloproteinases:they're not just for matrix anymore!. *Curr Opin Cell Biol* 2001; 13:534-540.
8. Van Trappen PO, Ryan A, Carroll M, Lecoecur C, Goff L, Gyselman VG, Young BD, Lowe DG, Pepper MS, Shepherd JH, Jacobs IJ. A model for co-expression pattern analysis of genes implicated in angiogenesis and tumour cell invasion in cervical cancer. *Br J Cancer* 2002;87:537-544.
9. Shapiro SD, Kobayashi DK, Ley TJ. Cloning and characterization of a unique elastolytic metalloproteinase produced by human alveolar macrophages. *J Biol Chem* 1993, 268:23 824–23 829.
10. Gronski TJ Jr, Martin RL, Kobayashi DK, *et al.* Hydrolysis of a broad spectrum of extracellular matrix proteins by human macrophage metalloelastase. *J Biol Chem* 1997, 272:12189–12194.
11. Curci JA, Liao S, Huffman MD, Shapiro SD, Thompson RW. Expression and localization of macrophage elastase (matrix metalloproteinase-12) in abdominal aortic aneurysms. *J Clin Invest* 1999, 102:1900–1910.
12. Vaalamo M, Karjalainen-Lindsberg ML, Puolakkainen P, Kere J, Saarialho-Kere U. Distinct expression profiles of stromelysin-2 (MMP-10), collagenase-3 (MMP-13), macrophage metalloelastase (MMP-12), and tissue inhibitor of metalloproteinases-3 (TIMP-3) in intestinal ulcerations. *Am J Pathol* 1998, 152:1005–1014.
13. Kerkela E, Ala-aho R, Jeskanen L, *et al.* Expression of human macrophage metalloelastase (MMP-12) by tumor cells in skin cancer. *J Invest Dermatol* 2000, 114:1113–1119.
14. Kerkela E, Ala-aho R, Klemi P, Grenman S, Shapiro SD, Kahari VM, Saarialho-Kere U. Metalloelastase (MMP-12) expression by tumour cells in squamous cell carcinoma of the vulva correlates with invasiveness, while that by macrophages predicts better outcome. *J Pathol* 2002, 198:258-269.
15. Noel A, Boulay A, Kebers F, Kannan R, Hajitou A, Calberg-Bacq CM, Basset P, Rio MC, Foidart JM. Demonstration in vivo that stromelysin-3 functions through its proteolytic activity. *Oncogene* 2000, 19:1605-1612.
16. Boulay A, Masson R, Chenard MP, El Fahime M, Cassard L, Bellocq JP, Sautes-Fridman C, Basset P, Rio MC. High cancer cell death in syngeneic tumors developed in host mice deficient for the stromelysin-3 matrix metalloproteinase. *Cancer Res* 2001, 61:2189-2193.

17. Yamashita K, Tanaka Y, Mimori K, Inoue H, Mori M. Differential expression of MMP and uPA systems and prognostic relevance of their expression in esophageal squamous cell carcinoma. *Int J Cancer* 2004, 110:201-207.
18. Soni S, Mathur M, Shukla NK, Deo SV, Ralhan R. Stromelysin-3 expression is an early event in human oral tumorigenesis. *Int J Cancer* 2003, 107:309-316.
19. Wasenius VM, Hemmer S, Kettunen E, Knuutila S, Franssila K, Joensuu H. Hepatocyte growth factor receptor, matrix metalloproteinase-11, tissue inhibitor of metalloproteinase-1, and fibronectin are up-regulated in papillary thyroid carcinoma: a cDNA and tissue microarray study. *Clin Cancer Res* 2003, 9:68-75.
20. Wlodarczyk J, Stolte M, Mueller J. E-cadherin, beta-catenin and stromelysin-3 expression in de novo carcinoma of the colorectum. *Pol J Pathol* 2001, 52:119-124.
21. Thewes M, Worret WJ, Engst R, Ring J. Stromelysin-3 (ST-3): immunohistochemical characterization of the matrix metalloproteinase (MMP)-11 in benign and malignant skin tumours and other skin disorders. *Clin Exp Dermatol* 1999, 24:122-126.
22. Mueller J, Brebeck B, Schmalfeldt B, Kuhn W, Graeff H, Hofler H. Stromelysin-3 expression in invasive ovarian carcinomas and tumours of low malignant potential. *Virchows Arch* 2000, 437:618-624.
23. Basset P, Bellocq JP, Lefebvre O, Noel A, Chenard MP, Wolf C, Anglard P, Rio MC. Stromelysin-3: a paradigm for stroma-derived factors implicated in carcinoma progression. *Crit Rev Oncol Hematol* 1997, 26:43-53.
24. Wex T, Levy B, Wex H, Bromme D. Human cathepsins W and F form a new subgroup of cathepsins that is evolutionary separated from the cathepsin B- and L-like cysteine proteases. *Adv Exp Med Biol* 2000, 477:271-280.
25. Santamaria I, Velasco G, Pendas AM, Paz A, Lopez-Otin C. Molecular cloning and structural and functional characterization of human cathepsin F, a new cysteine proteinase of the papain family with a long propeptide domain. *J Biol Chem* 1999, 274:13800-13809.
26. Cross SS, Hamdy FC, Deloulme JC, Rehman I. Expression of S100 proteins in normal human tissues and common cancers using tissue microarrays: S100A6, S100A8, S100A9 and S100A11 are all overexpressed in common cancers. *Histopathology* 2005, 46:256-269.
27. Caceres-Cortes JR, Alvarado-Moreno JA, Waga K, Rangel-Corona R, Monroy-Garcia A, Rocha-Zavaleta L, Urdiales-Ramos J, Weiss-Steider B, Haman A, Hugo P, Brousseau R, Hoang T. Implication of tyrosine kinase receptor and steel factor in cell density-dependent growth in cervical cancers and leukemias. *Cancer Res* 2001, 61:6281-6289.
28. Hidalgo A, Piña P, Guerrero G, Salcedo M. A simple method for the construction of small format tissue array. *J Clin Pathol* 2003, 56:144-146.
29. Smyth, G. K. Limma: linear models for microarray data. *Bioinformatics and Computational Biology Solutions using R and Bioconductor*, R. Gentleman, V. Carey, S. Dudoit, R. Irizarry, W. Huber (eds.), Springer, New York, Chapter 23. *In press*
30. Smyth, G. K. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Statistical Applications in Genetics and Molecular Biology* Smyth, G. K. 2004; 3,33.
31. Hoon N, Hong K, Hui S, Kang S, Young K, S Ho. MMP expression profiling in recurrent stage IB lung cancer. *Oncogene* 2004, 23:845-851.
32. Boulay A, Masson R, Chenard MP, El Fahime M, Cassard L, Bellocq JP, Sautes-Fridman C, Basset P, Rio MC. High cancer cell death in syngeneic tumors

- developed in host mice deficient for the stromelysin-3 matrix metalloproteinase. *Cancer Res* 2001, 61:2189-2193.
33. Basset P, Bellocq J, Wolf C, Stoll I, Hutin P, Limacher J, Podhajcer O, Chernard M, Rio M, P Chambon. A novel metalloproteinase gene specifically expressed in stromal cells of breast carcinomas. *Nature* 1990, 348:699-703.
 34. Rouyer N, Wolf C, Chenard MP, Rio M, Chabon P, Bellocq J, Basset P. Stromelysin-3 gene expression in human cancer: an overview. *Invasion Metastasis* 1994, 95:41269-41275.
 35. Basset P, Bellocq J, Lefebvre O, Noel A, Chenard M, Wolf C, Anglard P, Rio m. Stromelysin 3: a paradigm for stroma derived factors implicated in carcinoma progression. *Crit Rev Oncol Hematol* 1997, 26:43-53.
 36. Van Trappen P, Ryan A, Carroll M, Lecoecur C, Goff L, Gyselman V, Young BD, Lowe DG. A model of co-expression pattern analysis of genes implicated in angiogenesis and tumor cell invasion in cervical cancer. *Brit J Can* 2002, 87:537-544.
 37. Veli-Matti Wasenius, Hemmer S, Kettunen E, Sakari Knuutila, Fanssila K, J Heikki. Hepatocyte Growth Factor Receptor, Matrix Metalloproteinase-11, Tissue Inhibitor of Metalloproteinase-1, and Fibronectin Are Up-Regulated in Papillary Thyroid Carcinoma: A cDNA and Tissue Microarray Study. *Clin Can Res* 2003, 9:68-75.
 38. Shipley M, Wesselchmidt R, Kobashi K, Ley T, Shapiro S. Metalloelastase is required for macrophage mediated proteolysis and matrix invasion in mice. *Proc Natl Acad Sci USA* 1996, 93:3942-3946.
 39. Impola U, Uitto VJ, Hietanen J, Hakkinen K, Zhang L, Larjava H, Isaka K, U Saarialho-Kere. V Differential expression of matrilysin-1 (MMP-7), 92 kD gelatinase (MMP-9), and metalloelastase (MMP-12) in oral verrucous and squamous cell cancer. *J Pathol* 2004, 202:14-22.
 40. Chapman H, Riese RJ, and Guo-Ping S. Emerging roles for cysteine proteases in human biology. *Ann Rev Physiol* 1997, 59: 63-88.

Tables and Figures.

Table 1. Summary of altered gene expression in cervical cancer by cDNA array.

Ratio (T/N), intensity ratio of the corresponding signals between tumor and normal cervical tissue.

Table 2. Summary of the *In situ* hybridizations and immunohistochemical stainings results of cervical tissue microarrays.

Figure 1. Relative mRNA expression levels of CTSF, MMP11 and MMP12 in cervical tumors detected by RT-PCR.

One hundred of total RNA was employed for RT-PCR experiments. The used samples were; line a) positive control, line b) negative control, line 1) normal HPV- tissue #N07, 2) normal HPV- tissue # N11, 3) normal HPV- tissue # N20, 4) normal HPV- tissue #N23, 5) HeLa, 6) CaLO, 7) INBL, 8) CaSki, 9) SiHa, 10) tumoral sample # T04, 11) tumoral sample # T24, 12) tumoral sample # T43, 13) tumoral sample # T45, 14) Tumoral sample T63. A 100 bp ladder was used as molecular weight marker. Values in the graph are presented as ratio of densitometric readings of samples to corresponding beta actin samples.

Figure 2. *In situ* hybridization of cervical tissue microarrays with CTSF, MMP11 and MMP12 probes. A) Breast cancer (as positive control of CTSF), B) normal cervical epithelium negative for CTSF, C) HSIL with positive immunostaining, D) CC tissue positive for CTSF hybridization, E) Breast cancer (as positive control for MMP11), F) normal cervical epithelium, negative for MMP11 hybridization, G) as in C, HSIL positive for MMP11 hybridization, H) positive hybridization for MMP11 in CC, I) Breast cancer as positive control for MMP12, J) MMP12 negative hybridization in normal epithelium, K) as in C and G HSIL positive for MMP12, and L) Positive hybridization for MMP12 in CC. Original magnifications were 40 X for A, C, E, I and L; and 10X were used for B, D, F, G, H, J, and K.

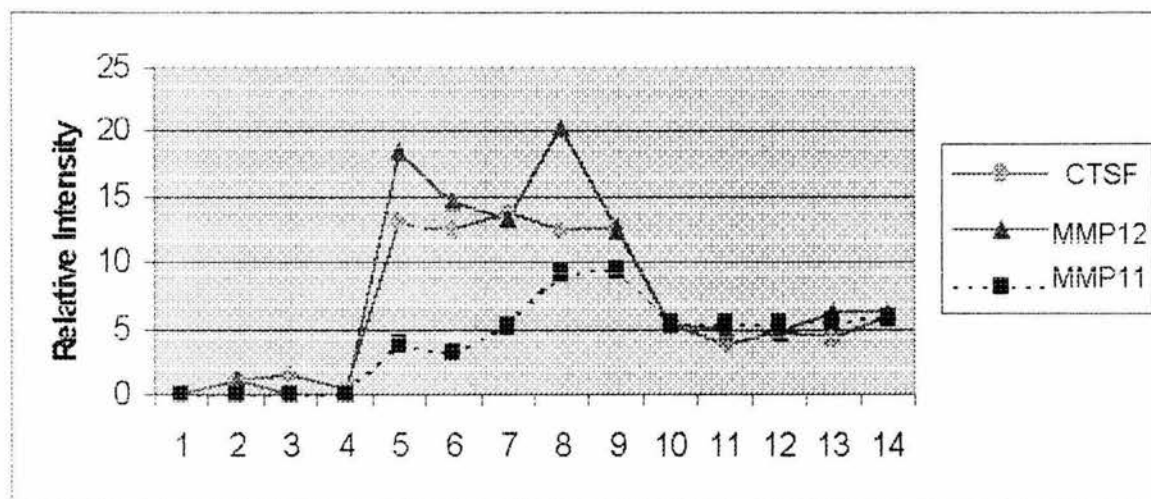
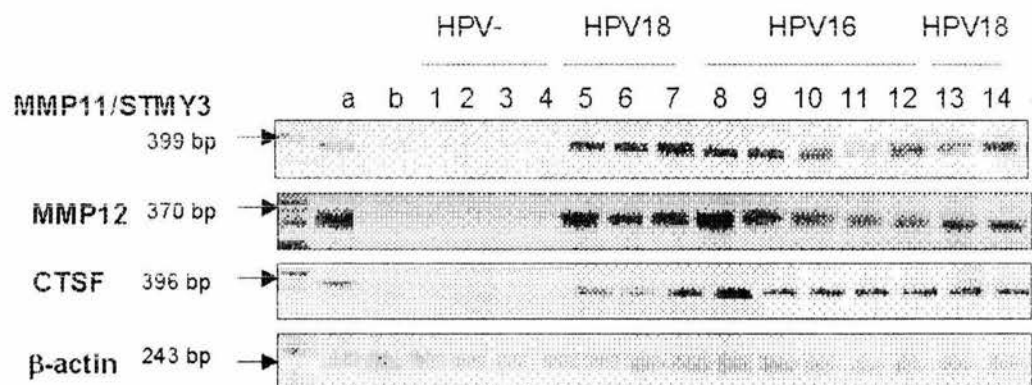
Figure 3. Immunohistochemical detection of MMP11 and MMP12 by cervical tissue microarrays.

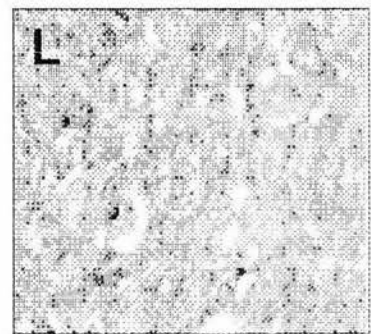
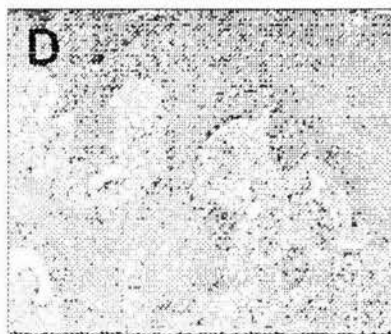
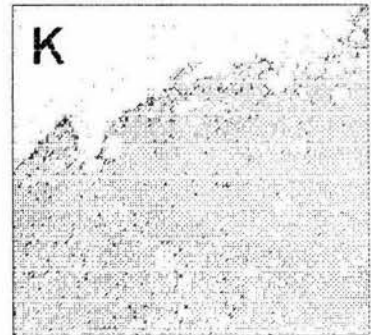
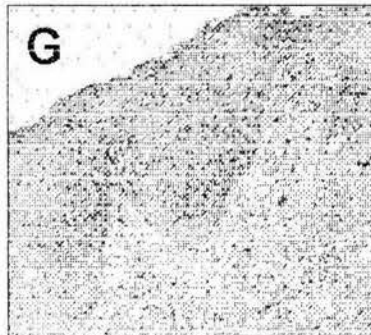
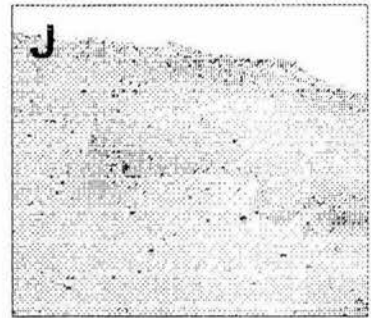
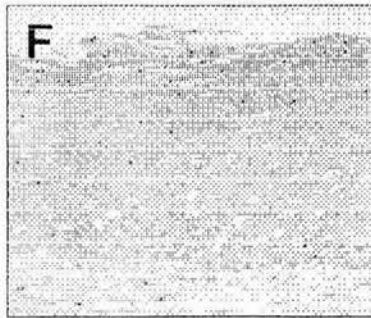
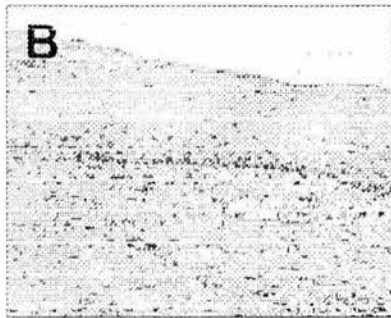
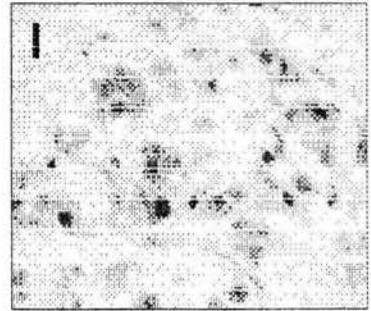
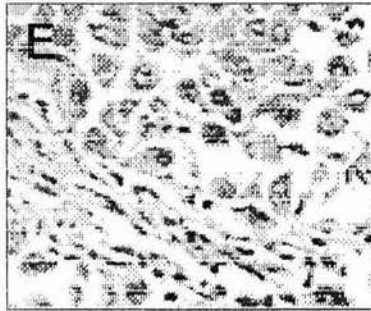
- A) Breast cancer sample (as positive control for MMP11 staining), B) negative control; a CC sample without primary antibody, C) one normal cervical tissue showing negative staining for MMP11, D) LSIL which presents positive immunostaining for MMP11, E) HSIL positive staining for MMP11 F) a CC showing positive immunoreaction for MMP11, G) Positive control for MMP12, H) as in B, negative control for MMP12, I) normal cervical tissue shows negative immunostaining for MMP12, J) LSIL showing positive immunoreaction for MMP12, K) HSIL showing immunodetection for MMP12 and L) a CC showing positive immunoreaction for MMP12. Original magnifications were 40 X for A, C, D, E, F, G, J, and K; and 10X were used for B, H, I, and L.

Gene Name	Locus	Halah	Ca ₂ O ₂ N	INBL/n	SiH ₂ n	CaSk/n	Ratio (sample/normal)			
							T07/n	T31/n	T64/n	T84/n
tissue inhibitor of metalloproteinase 1	Hs.446641	6.78	4.02	3.57	3.27	4.12	3.56	4	4.23	
integrin, alpha 9	Hs.222	7.32	6.2	3.28	4.29	3.25	3.58	3.69	3.47	
cadherin 2, type 1, N-cadherin	Hs.334131	6.48	3.56	3.98	5.02	4.12	4.68	4.37	5.65	
cathepsin L2	Hs.87417	8.69	4.89	3.61	3.56	4.28	3.98	3.74	4.36	
matrix metalloproteinase 12	Hs.1695	5.32	8.75	7.52	8.67	5.98	5.47	4.32	4.89	
integrin, alpha 4	Hs.145140	8.11	3.01	4.68	4.25	3.21	3.65	4.51	4.36	
seine B member 9	Hs.104879	5.1	3.59	3.67	4.56	4.23	3.1	3.54	4.18	
seine, proteinase inhibitor, clade B member 3	Hs.227948	6.31	4.2	3.71	4.95	3.64	3.71	3.69	3.41	
cathepsin B	Hs.297939	5.65	3.6	3.42	5.31	4.65	3.98	3.58	4.12	
cathepsin F	Hs.11590	8.06	8.21	5.95	6.58	6.35	5.32	4.65	5.32	
matrix metalloproteinase 10	Hs.2258	7.32	4.87	4.52	3.35	3.98	3.82	3.98	4.69	
matrix metalloproteinase 11 (stromelysin 3)	Hs.143751	6.29	7.65	5.68	5.79	6.12	5.12	4.98	5.46	
matrix metalloproteinase 9	Hs.151738	5.02	6.23	4.68	3.62	4.28	4.68	3.87	3.48	
cadherin 11, type 2	Hs.443435	6.04	3.21	3.48	3.11	4.11	4.69	4.85	3.82	
cathepsin S	Hs.181301	0.45	0.65	0.64	0.35	0.64	0.56	0.32	0.67	
cathepsin L	Hs.418123	0.23	0.34	0.51	0.62	0.65	0.31	0.29	0.68	
cathepsin H	Hs.114931	0.36	0.89	0.32	0.84	0.36	0.42	0.59	0.23	
cathepsin C	Hs.128065	0.98	0.23	0.65	0.16	0.98	0.68	0.56	0.48	
matrix metalloproteinase 13	Hs.2936	0.27	0.45	0.32	0.62	0.32	0.67	0.23	0.26	
tissue inhibitor of metalloproteinase 3	Hs.245188	0.36	0.56	0.42	0.51	0.86	0.71	0.51	0.32	
cadherin 3, type 1, P-cadherin (placental)	Hs.191842	0.15	0.64	0.53	0.12	0.12	0.59	0.36	0.65	
cadherin 4, type 1, R-cadherin (retinal)	Hs.376792	0.87	0.34	0.61	0.45	0.68	0.87	0.41	0.64	
integrin, beta 8	Hs.355722	0.46	0.21	0.24	0.62	0.87	0.36	0.32	0.65	
elastase 2, neutrophil	Hs.99863	0.32	0.62	0.31	0.32	0.69	0.36	0.46	0.47	

		% Positivity in ISH											
		CTSF				MMP11				MMP12			
	Total of cases	-	(+)	(++)	(+++)	-	(+)	(++)	(+++)	-	(+)	(++)	(+++)
Normal tissues	15	93.4	6.6	-	-	86.7	13.3	-	-	86.7	13.3	-	-
LSIL	10	10	90	-	-	10	90	-	-	10	90	-	-
HSIL	10	10	10	80	-	10	20	70	-	10	20	70	-
CC tissues	15	-	-	-	100	-	-	-	100	-	-	-	100

		% Positivity in IHC							
		MMP11				MMP12			
	Total of cases	-	(+)	(++)	(+++)	-	(+)	(++)	(+++)
Normal tissues	15	15	0	0	0	15	0	0	0
LSIL	10	0	10	0	0	0	9	1	0
HSIL	10	0	2	8	0	0	0	8	0
CC tissues	15	0	0	5	10	0	0	1	14





Manuscript Draft

Manuscript Number:

Title: Wnt signaling pathway expression is altered in cervical cancer

Article Type: Original Research Report

Section/Category:

Keywords: Keywords, cervical cancer, cDNA arrays, cluster, semiquantitative RT-PCR

Corresponding Author: Dr Guelagueta Vazquez, PhD Cand

Corresponding Author's Institution: Oncogenomics Laboratory, Oncology Diseases Research Unit, Oncology Hospital, National Medical Center siglo XXI-IMSS, Mexico

First Author: Guelagueta Vazquez, PhD Cand

Order of Authors: Guelagueta Vazquez, PhD Cand; Patricia Filla, MGS; Ricardo Lopez, PhD; Karla Vazquez, Biol; Brenda Alatorre, Biol; Carlos Perez-Plascencia, PhD cand; Federico Avila, PhD Cand; Sandra Orozco, PhD; Mauricio Salcedo, PhD

Manuscript Region of Origin:

Abstract: Objective. Because of its widespread activation in diverse human tumor diseases, the Wnt pathway has gained considerable and growing interest in tumor research over recent years. Although some reports of specific Wnt genes and cervical cancer exists, no global studies have been reported. Our objective was to determine the expression profile of Wnt pathway in Cervical Cancer. Methods. In order to determine the expression profile of Wnt pathway in Cervical Cancer we used cDNA arrays to examine the expression profiles in 5 cell lines derived from human cervical cancer, 3 tumor samples and 3 normal (HPV -) epithelium tissues. We used a Gene Ontology program to assign each of the genes that were present in the array into known cellular pathways. After this, Hierarchical Cluster was performed and we analyzed and confirmed the gene expression using RT-PCR and RNA-RNA In situ hybridization. Results. All cell lines and tumoral cells showed overexpression in Wnt signaling pathway by cDNA arrays and semiquantitative RT-PCR. All 20 normal samples tissues showed a weak expression by In situ hybridization and by the immunostaining, while in the 30 tissue samples of SILMG and CC tissues an increment in the signal is showed. Conclusions. We found that Wnt signaling pathway is overexpressed in cervical transformed epithelial cells by the possible activation of the canonical signaling.

Wnt signaling pathway expression is altered in Cervical Cancer

Guelaguetza Vázquez-Ortiz,^a Patricia Piña-Sánchez,^a Ricardo López,^a Karla. Vázquez,^a Brenda. Alatorre,^a Carlos Perez-Plasencia,^a Federico Avila,^b Sandra Orozco,^c and Mauricio Salcedo^a

^a Oncogenomics Laboratory, Oncology Diseases Research Unit, Oncology Hospital, National Medical Center Siglo XXI-IMSS, Mexico.

^b *Department of Chronic-Degenerative Diseases, Lung Cancer Laboratory, National Institute Respiratory Diseases, INER, S.S., Mexico.*

^c *Medical Research Unit in Neurological Diseases, Specialty Hospital, National Medical Center Siglo XXI-IMSS, Mexico*

Keywords, *Cervical cancer, cDNA arrays, cluster, semiquantitative RT-PCR*

Corresponding Author:

Dr. Mauricio Salcedo

Unidad de Investigación Médica en Enfermedades Oncológicas

Hospital de Oncología

Centro Médico Nacional Siglo XXI-IMSS

Av. Cuauhtemoc 330. Col Doctores.

México D. F. C. P. 06720

Phone: +55 56276900 Ext. 22708 Fax: +55 57610952

Email: guelav@yahoo.com, maosal89@yahoo.com,

Abstract

Objective. Because of its widespread activation in diverse human tumor diseases, the Wnt pathway has gained considerable and growing interest in tumor research over recent years. Although some reports of specific Wnt genes and cervical cancer exists, no global studies have been reported. Our objective was to determine the expression profile of Wnt pathway in Cervical Cancer.

Methods. In order to determine the expression profile of Wnt pathway in Cervical Cancer we used cDNA arrays to examine the expression profiles in 5 cell lines derived from human cervical cancer, 3 tumor samples and 3 normal (HPV -) epithelium tissues. We used a Gene Ontology program to assign each of the genes that were present in the array into known cellular pathways. After this, Hierarchical cluster was performed and we analyzed and confirmed the gene expression using RT-PCR and RNA-RNA In Situ hybridization.

Results. All cell lines and tumoral cells showed overexpression in Wnt signaling pathway by cDNA arrays and semiquantitative RT-PCR. All 20 normal samples tissues showed a weak expression by In situ hybridization and by the immunostaining, while in the 30 tissue samples of SILHg and CC tissues an increment in the signal is showed.

Conclusions. We found that Wnt signaling pathway is overexpressed in cervical transformed epithelial cells by the possible activation of the canonical signaling.

Introduction

Cervical carcinoma (CC) is one of the most common cancers among women worldwide. High-risk human papillomavirus (HPV) infection is considered to be the most important risk factor associated with the development of this tumor^[1]. Nevertheless, other molecular alterations that are needed for cervical tumorigenesis remain unknown. Wnt signaling pathway is involved in various differentiation events during embryonic development and leads to tumor formation when aberrantly activated. Secreted Wnt factors are found throughout the animal kingdom and control virtually every developmental decision^[2]. In the absence of Wnt signal, unstimulated cells regulate β -catenin levels by a multiplex complex which phosphorylates this protein marking it for subsequent ubiquitination and degradation in the cytoplasm. But when the secreted Wnts are present, this stimulation leads to beta-catenin accumulation and then beta-catenin translocates into the nucleus where it engages DNA-binding proteins of the T cell factor/lymphoid enhancing factor (TCF/LEF) family. In the absence of a Wnt signal, these DNA binding proteins occupy target genes and repress their transcription. But, beta-catenin presence transiently converts TCF factors into transcriptional activators^[3-4] for many gene targets. Connections between Wnt signalling and human disease have been solidified by studies of malignancies such as colorectal^[5], gastric^[6], ovarian^[7], and hepatocellular^[8] carcinomas. In all of these cases, the common denominator is the activation of gene transcription by an accumulation of beta-catenin in the nucleus due to mutations in exons 3, 4 and 6 of the gene. CC reports have shown cytoplasmic and nuclear expression of beta-catenin without mutations on the previously mentioned exons^[9-10] meaning that beta-catenin accumulation in the nucleus might be due to another reasons. On the basis of these observations, we examined the pattern of expression of almost all Wnt genes in uterine cervical carcinoma tissues, cell lines and normal cervical tissues by using cDNA arrays. We found an upregulation of all Wnt studied genes; we verified the overexpression of four selected genes FZD2, GSK3B, C-MYC and PPARD by semiquantitative RT-PCR, In situ

hybridization and immunohistochemical assays. It is worth to note that we didn't find an association of this upregulation with HPV infection.

Material and Methods

HeLa (HPV18), SiHa (HPV16), CaSki (HPV16) and C33 (HPV⁻) cell lines were grown in the recommended media including 10% fetal bovine serum. Cell lines CALO (derived from Invasive squamous carcinoma Stage IIB, HPV18) and INBL (derived from Invasive Squamous carcinoma IV A, HPV18), were established at the FES Zaragoza-UNAM, Mexico [11-12]. Fresh colposcopy tumor guided biopsies were taken from 7 patients at the Dysplasias National Center in Mexico City. Twenty seven normal cervical tissues from pre-menopausal women were obtained two hours post-mortem at the Pathology Department of the Hospital General de Mexico, S. S. from women who died of non-gynecological related causes. All the described procedures have been evaluated and approved by the local committee of ethics of the Mexican Institute of Social Security; we also obtained written informed consent from the patients. All tissue samples were divided in three sections, the central part was snapped frozen in liquid nitrogen and stored at -70°C until nucleic acid extraction, and the extremes from the biopsy were fixed with 70% ethanol overnight and paraffin embedded. Multiple steps Haematoxylin/eosin stained sections from the extremes of the biopsies were inspected for representativity of the tissue. Carcinoma samples were considered as representative when at least 70% of all cells in the tissue section consisted of cancer cells and the normal tissue samples when the entire sample consisted of microscopically normal cervical tissue in all sections examined besides HPV absence. All the invasive samples were diagnosed as squamous lesions.

Both RNA and DNA were extracted from the tissue samples and the cervical cancer cell lines using the Trizol Reagent (Gibco BRL, Life Technologies, Grand Island NY USA) according to manufacturer's conditions. Each DNA sample was subjected to HPV typing by polymerase chain reaction (PCR) using the MY09/MY11 consensus primers and direct sequencing. Potentially contaminating DNA was removed by treating with RNase-free DNase I (Ambion Inc. Austin Texas USA). The resulting RNA concentration was measured by using a Spectrophotometer and the quality of both nucleic acids was confirmed in agarose gels.

cDNA was synthesized according to Ambion Advantage System User Manual. The samples were: cell lines HeLa, SiHa, CaSki, CALO, C33 and INBL, three HPV16+ tumor samples (T07, T31, T64) and three HPV⁻ normal samples (N03, N11, N22); unfortunately no HPV18+ tumor samples were included for array hybridization due to RNA amount was not enough for this type of experiment.

The cDNA probes were obtained by RT-PCR from 500 ng of RNA in the presence of (α -³³P) dATP (3000Ci/mmol NEN). Briefly, RNAs were denatured at 70°C for 10 min and cDNAs were synthesized at 42°C using oligo-dT priming in a final volume of 30 μ l. The labelled cDNAs were purified by spin column chromatography. The final incorporation of the radiolabel was approximately 1×10^8 cpm for each reaction. ULTRArray Advantage System array blots (Ambion Inc) which contains 8400 genes, were prehybridized at 68°C for at least 90 min before probe addition in Hybridization buffer. Finally, 1.5×10^7 cpm of each labeled cDNA was added to the buffer. Hybridization was performed at 60°C in a rolling bottle overnight. The arrays were washed twice with 2X SSC and 0.5 % SDS at 60°C for 30 min, followed by two stringent washes with 0.5X SSC, 0.5% SDS at the same temperature and time. Hybridized arrays were scanned on a phosphorimager (STORM). The resulting images were analyzed with the ArrayVision software (Image Research Inc, UK) to calibrate relative ratios and confidence intervals used for significance

determinations. Radioactive intensity was processed and then the data were imported into an Excel (Microsoft) database, with the corresponding gene names for analysis. For each gene, its relative fold change in expression was the ratio of median expression levels of each sample versus the average of the normal cervical samples. All genes were included in the analyses. In order to find a metabolic pathway entirely overexpressed in cervical cancer we classify the gene profiles into different pathways, cellular process analysis was carried out using Gene Ontology (GO) just as previously reported [13]. Each gene was associated with its corresponding current gene entry in UniGene (<http://www.ncbi.nlm.nih.gov>). Finally, each gene was queried for available GO Code for biological and cellular processes. Almost all studied pathways did not show a pattern of entire overexpression or suppression in all the genes which conforms the different cellular pathways, but Wnt signalling pathway presented it. Once we obtained the table which contains all the Wnt signaling components present in the array and also the folds of expression of the samples. Hierarchical clustering (Gene Cluster v2.11) and display programs (Tree View v1.50) were used to visualize the analysis (<http://rana.lbl.gov/Eisensoftware.htm>) [14-15]. In order to verify the upregulation of the Wnt pathway, four essential genes present at different levels of the signaling pathway were selected to perform other kind of analysis. The first selected gene was the Homolog of Frizzled Drosophila 2 (FZD2), which product protein is located on cell surface, a cytoplasm gene: Glucogen Synthase Kinase 3-beta (GSK3B) was also studied and two gene products which are expressed in the nucleus: C-MYC and Peroxisome Proliferators Activated Receptor Delta (PPARD); were chosen.

Estimation of the reliability of the cDNA results for the four selected genes was performed by conventional RT-PCR on five normal samples, 6 cell lines, 2 HPV16 and 2 HPV18 tumors. Two-hundred ng of total RNA from all samples were reverse transcribed using the RT-PCR Access System (Promega, Madison, WI). Primers for the genes of interest were: FZD2 386 bp, TM= 60°C, (5'- CTGAGCTCAAGTTCTTCCTGTG -3' and 5'- CAGGAAGTGGTAGTTGAGGTAG -3'); GSK3B, 504 bp TM=60°C (5'- GTECTATAGGGCACCAGAGTTG -3' and 5'- GTAGCCAGAGGTGGATTACTTG -3'); C-MYC, 469 bp, TM=60°C, (5'- AGCTTGTACCTGCAGGATCTGA -3' and 5'- TGACACTGTCCAACCTTGACCCT -3'); PPARD, 86 bp, TM=60°C, (5'- GGTTACCCCTTCTCAAGTATGGC -3' and 5'- CTCTGAAGACAGGAGACAGAAG -3') and β -actin: (5'-TGAAGTCTGACGTGGACATC-3' and GTTCGTTCTCATACTGCTCA-3') 243 bp, TM=55°C primers were designed using Integrated DNA technologies Biotools software (<http://biotools.idtdna.com/gateway>). The RT-PCR conditions were: to synthesize the first strand of cDNA 48°C for 45 min and 94°C for 2 min to denature template; and to synthesize second strand and DNA amplification: 94°C for 30 sec, (specific TM°C for each set of primers) for 1 min, and 68°C for 2 min for a total of 24 cycles, followed by a single step at 68°C for 7 min. The products were visualized on 1.5% agarose gels stained with ethidium bromide, and the band intensities were captured using the Eagle eye system (Stratagene, La Jolla, Ca) and analyzed by the MetaView analysing system (version 4.5 Universal Imaging Corp., USA). The relative concentrations of the mRNAs were expressed as the ratio of densitometric readings of the three genes related to β -actin mRNA.

Statistical comparisons among groups were analysed by a Kruskal Wallis-test. A P value of <0.05 was considered as statistically significant.

Paraffin blocks of 15 SILHg and 15 CC were randomly collected from the archives of the Dysplasias National Center in Mexico City. Twenty of the normal cervical tissues (*post-mortem* obtained) were formalin buffered fixed and paraffin embedded for posterior use. For *In Situ* Hybridizations experiments (ISH), 5 μ m tissue sections were cut from paraffin blocks, deparaffinized and rehydrated in a graded ethanol series (100, 90, 70, and 30%), and transferred to PBS solution (137 mM NaCl, 2.7 mM KCl, 4.3 mM Na₂HPO₄, 1.4 mM

KH₂PO₄) for 10 min. Tissues were treated with DNase solution (1 µg/ml) for 10 min at 37°C and washed three times in PBS. Endogenous peroxidase inactivation was carried out by incubating the samples in hydrogen peroxide in methanol for 3% hydrogen peroxide (H₂O₂) for 40 min. Sense and antisense probes for FZD2, GSK3B, PPARD and C-MYC genes were generated by single-strand PCR using specific cDNA obtained from SiHa cells using RNA as template and labeled with biotin-16-dUTP (Roche). Each tissue was covered with 50 µl of hybridization cocktail and a coverslip. The hybridization cocktail consisted of 50% formamide, 10% dextran sulfate, 2X SSC (20X SSC: 3 M NaCl, 300 mM Na₃C₆H₅O₇), PBS, 2% SDS, 100 µg/ml sonicated salmon sperm and 50 ng of the dUTP-biotin-labeled probe (sense or antisense). Both the RNA of the tissue and probe were denatured at 65°C for 10 min and incubated at 42°C for 10 min. After hybridization, the coverslips were soaked off in Tris-buffered saline with Tween buffer 1X (TBST 10X: 500 nmol/L TrisHCL, pH 7.6, 3 mol/L NaCl, 1% Tween 20), and the tissues were incubated at 55°C for 20 min in stringent wash solution. The horseradish peroxidase labeled antibody (SA-HRP) was immediately applied on the tissues and incubated for 15 min in a humidifier chamber. The sections were washed in TBST 1X. Biotinyl tyramide (GenPoint System from DAKO) was applied on the tissue sections for signal amplification, for 15 min at room temperature, and washed in TBST. A second SA-HRP step was carried out and the color reaction was developed with 0.06% diaminobenzidine (DAB) in 3% H₂O₂. Finally, the slides were washed, haematoxylin/counterstained, dehydrated in graded ethanol and mounted. Negative controls for ISH were carried out using the sense probes or a treatment with RNase solution (100 mg/ml for 30 min at 37°C) prior ISH. For Immunohistochemistry assays (IHC), briefly, the tissue sections were deparaffinized and rehydrated in a graded ethanol series. Endogenous peroxidase inactivation was carried out as described in ISH, and the slides were preincubated with DAKO Protein Block Serum free Cat X0909 (Carpinteria, CA, USA) for 30 min at 37°C to prevent non specific immunoreaction. Excess was decanted and tissues were incubated with the primary antibody as follows C-MYC (Santa Cruz), at 37°C for 30 min. DAKO Envision System Peroxidase (DAKO) was applied and the slides and then counterstained with hematoxylin. A rough qualitative scoring was done without any attempts to quantitate the intensity of the staining by two authors (GVO and PPS). Staining result was considered positive whenever increased stained positive in the majority of the tumor cells was present as compared with staining of the control tissue consisting of histological normal cervical tissues.

Results

To identify gene expression profile related to Wnt signaling process in human cervical tissues, we identified specific transcripts which belong to this pathway in normal and cervical cancer samples using the ULTRArray Advantage System. After obtaining the expression ratio for each gene in tumor tissues, cervical cell lines and normal samples, values were globally normalized and transformed as described in Materials and Methods. The mean expression values were then used to identify the differentially expressed genes in this pathway. Surprisingly, the Wnt signaling pathway showed an overexpression in all studied samples. Hierarchical clustering and TreeView were employed to show the total gene overexpression in this pathway. Figure 1 shows the significantly up-regulated Wnt genes, at least 1 fold in nine tumoral samples; including C33 cell line which is HPV negative. These genes were classified into Wnt signalling pathway, following the criteria of the Gene Ontology Consortium.

The gene expression profile findings in cDNA arrays were confirmed by semiquantitative RT-PCR for selected genes (FZD2, GSK3B, C-MYC, PPARD, and β -

actin; Fig. 2). In general, these analyses showed results consistent with those obtained by cDNA arrays. All 9 investigated samples using cDNA arrays showed increased FZD expression, and in line with this, the tumor band intensities in agarose gels as compared with the control band intensity were increased from an average of 1.2 fold in all tumors for FZD in the RT-PCR analyses. The greatest difference between the two methods was found in GSK3B expression. All studied samples had increased GSK3B expression of about 3 folds in cDNA microarray analysis, whereas an average of 8 times increased expression is present in RT-PCR. C-MYC also presents overexpression and in RT-PCR analyses an average of about three folds higher expression than the normal controls is present. PPARD show an average of overexpression of about 1.3 fold in tumoral samples of RT-PCR analyses. Comparisons using Kruskal Wallis test were statistically significant <0.05 except for GSK3B, (Normal HPV- tissues media Vs HPV+ tissue samples media). In general, the ISH results showed a positivity of 90% for SILHgs and 100% of CC cases. FZD2 was expressed in all cases, GSK3B in 80%, C-MYC in 93% and PPARD in 86% of the 15 SILHg samples. The four genes were also expressed in all the CC cases, but, a higher expression was observed respect to the normal tissues. Cervical tissues consisting of histologically normal cervical tissue did not stain or stained only weakly (Table 1, Fig. 3). For C-MYC immunostaining experiments a positivity of 86% was present in the SILHg and 100% for the CC samples. In all CC samples was present a strong immunoreaction while the signal in SILHgs was lower. Normal cervical tissues did not immunostained for the protein or immunostained weakly in the basal cells of epithelium (Table 1, Figure 3). Positive staining for C-MYC was localized in nucleus and this protein did not stained in stromal cells.

Discussion

In order to identify altered metabolic pathways in cervical cancer we used a gene expression array based on 8 000 genes; and subsequently we classified the genes into different pathways. Wnt signaling pathway was found to be entirely overexpressed, and this upregulation was validated by using ISH and IHC in a specific gene. As expected, an up-regulation was detected in CC tissues and cell lines. The ISH results of the selected genes (FZD, GSK3B, C-MYC and PPARD) were in line with results obtained by RT-PCR. It is worth important to note that ISH results showed increased expression through different stages of cervical carcinogenesis.

Wnts are secreted glycoproteins which are produced by different cell types and mostly act in a paracrine fashion^[16]. There are 19 Wnt proteins known, and Wnt1, Wnt3 and Wnt8 activates the canonical Wnt pathway. In CC, Wnt signaling might be activated by Wnt1 due to its overexpression showed in this work.

Wnts binds to 7 transmembrane receptors called frizzelds (FZDs) which are essential for signal transmission^[17-19]. FZD2 overexpression has been reported in head & neck, breast and colon carcinomas showing a possible involvement of this receptor in tumorigenesis^[19-21]. In our results, we are showing an increment of FZD2 expression throughout different kind of lesions of the cervix.

The molecular events which take place immediately after activation of frizzelds remain largely elusive, nevertheless, after activation of Wnt signaling, Dishevelled (DVL) becomes recruited to the plasma membrane. This protein which has been reported to be overpressed in CC^[22], also interacts with other proteins, such as Dapper, Frodo, Daam1, Stbm, PAR-1, and Nkd^[23-24]. The functional role of most of these components in mammalian development or in cancer has not been analyzed yet, but our results show them clearly overexpressed in human cervical cancer. One of the validated genes in this study was serine/Threonine kinase GSK3B which is the next transactivated gene following

the pathway. This protein binds to and phosphorylates several proteins, as a negative regulator of Wnt signalling, GSK3B would qualify as a potential tumor suppressor. In absence of canonical signaling, GSK3B phosphorylates β -catenin in four amino-acid-terminal residues ^[26], targeting it for binding the B-TrCP and its subsequent ubiquitination and degradation by proteasome ^[26-27]. Our study demonstrates GSK3B expression in cervical precancerous lesions and a significantly higher expression in invasive carcinomas compared to the histologically normal cervix tissues analyzed, it is worth to note that this has not been previously reported in CC. Probably meaning that canonical signaling pathway might be turned on at early stages of cervical carcinogenesis, nevertheless further studies should be performed on this direction.

The β -catenin degradation complex consists on GSK3B, Axin and APC. Axin and APC form a complex that allows GSK3B to phosphorylate β -catenin. Axin could present mutations and these could truncate the protein in a manner that eliminates the β -catenin binding sites. In CC 3 polymorphisms were previously described, but these do not alter the protein expression pattern showing their limited role in this kind of tumorigenesis ^[28]. With the canonical signaling present and no bounds between GSK3B and Axin, phosphorylation of β -catenin does not occur at cytosolic levels, and this protein accumulates and translocate to the nucleus ^[29]. In the nucleus; β -catenin transactivation domain ^[30] is an essential cofactor for TCF/LEF transcription factors (TCF1, LEF1, TCF3 and TCF4) ^[31-32] which activates the Peroxisome proliferator activated receptor delta (PPARD) and other TCF target genes. PPARD gene was studied in this report and as expected, it showed an increment in the expression from normal to tumoral tissues.

A way by which Wnt signaling could lead to accumulation of transformed cells in CC is by the inhibition of apoptosis. Expression of the anti-apoptotic gene survivin it is known to be upregulated by APC and the analysis of its promoter revealed also TCF-4 binding sites ^[33-34]. This gene is also overexpressed in CC according to our results. On the other hand, Acetyltransferase CBP (CREB binding protein) acetylates β -catenin which transactivates c-MYC locus ^[35]. In CC, C-MYC transactivation is also performed by viral protein E6 HPV ^[36], and by E2F-1 ^[37] this might partially explain the great overexpression of this transcript in our results. Through C-MYC, Wnt signaling may promote progression of cells through the cell cycle. This is due to C-MYC upregulation leads to repression of the cyclin-dependent kinase (CDK) inhibitor p21CIP and finally this stimulates G1/S progression ^[38]. It has been extensively demonstrated the important role of C-MYC (overexpression and amplification) in the development of the uterine cervix carcinomas ^[39-41]. This could partially explain that epithelial expression of C-MYC in intraepithelial squamous lesions might be related with the progression of phenotypic alterations acquired early during the malignant transformation pathway of cervical epithelium and maintained after invasion. When the Wnt signal is not present TCF/LEF binds to Grg (Groucho) transcriptional repressor ^[42] which helps to actively repress transcription of TCF target genes. In conclusion, our results show that Wnt signaling pathway is overexpressed in cervical transformed epithelial cells by the possible activation of the canonical signaling. Thus the role of these genes in the pathogenesis of CC needs to be confirmed in other studies, and the molecular mechanisms involved need to be studied in detail.

Acknowledgements

We are grateful to Dr. J Moreno, and Dr. D Arenas (CMN SXXI-IMSS) for their helpful support. This work was partially supported by Grants No. 34686-M and 7114 (Fondos

Sectoriales. CONACyT-México) and FOFOI-IMSS. During this work GVO, PP, RL, KV, and BA were recipients of a CONACyT and IMSS fellowship.

References

- [1] Walboomers J, Jacobs M, Manos M, Bosch X, Kummer A, Shah K, Snijders P, Peto J, Meijer C and Muñoz N Human Papilloma Virus is a necessary cause of invasive cervical cancer worldwide. *J Pathol* 1999;189:12-19.
- [2] Logan CY, Nusse R. The wnt signaling pathway in development and disease. *Annu Rev Cell Dev Biol*. 2004;20:781-810.
- [3] Polakis P. Wnt signaling and cancer. *Genes Dev* 2000;14:1837-51.
- [4] Sancho E, Battle E, Clevers H. Signaling Pathways in intestinal development and cancer. *Annu rev Cell Biol* 2004;20:695-723.
- [5] Seidler H, Utsuyama M, Nagoka S, Takemura T, Kitagawa M and Hirokawa K. Expression of Wnt signaling components possibly influences the biological behavior of colorectal cancer in different age groups. *Exp Mol Pathol* 2004;76:224-33.
- [6] Ebert MP, Yu J, Hoffmann J, Rocco A, Rocken C, Kahmann S, Muller O, Korc M, Sung JJ, Malfertheiner P. Loss of beta-catenin expression in metastatic gastric cancer. *J Clin Oncol*. 2003;21:1708-14.
- [7] Schwartz DR, Wu R, Kardia SL, Levin AM, Huang CC, Shedden KA, Kuick R, Misek DE, Hanash SM, Taylor JM, Reed H, Hendrix N, Zhai Y, Fearon ER, Cho KR. Novel candidate targets of beta-catenin/T-cell factor signaling identified by gene expression profiling of ovarian endometrioid adenocarcinomas. *Cancer Res*. 2003;63:2913-22.
- [8] Chung EJ, Sung YK, Farooq M, Kim Y, Im S, Tak WY, Hwang YJ, Kim YI, Han HS, Kim JC, Kim MK. Gene expression profile analysis in human hepatocellular carcinoma by cDNA microarray. *Mol Cells*. 2002;14:382-7.
- [9] Shinohara A, Yokoyama Y, Wan X, Takahashi Y, Mori Y, Takami T, Shimokawa K, Tamaya T. Cytoplasmic/nuclear expression without mutation of exon 3 of the beta-catenin gene is frequent in the development of the neoplasm of the uterine cervix. *Gynecol Oncol* 2001;82:450-5.
- [10] Veeman, M.T., Axelrod, J.D. and Moon, R.T. A second canon. Functions and mechanisms of beta-catenin-independent Wnt signaling. *Dev Cell* 2003;5:367-377
- [11] Caceres-Cortes JR, Alvarado-Moreno JA, Waga K *et al*. Implication of tyrosine kinase receptor and steel factor in cell density-dependent growth in cervical cancers and leukemias. *Cancer Res* 2001;61:6281-9.
- [12] Lopez-Marure R, Gutierrez G, Mendoza C *et al*. Ceramide promotes the death of human cervical tumor cells in the absence of biochemical and morphological markers of apoptosis. *Biochem Biophys Res Commun* 2002;293:1028-36.
- [13] Pletcher SD, Macdonald SJ, Marguerie R, Certa U, Stearns SC, Goldstein DB, *et al*. Genome-wide transcript profiles in aging and calorically restricted *Drosophila melanogaster*. *Curr Biol* 2002;12:712-23.
- [14] Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A* 1998;95:14863-8.
- [15] Nakamura T, Shoji A, Fujisawa H, Kamatani N. Cluster analysis and association study of structured multilocus genotype data. *J Hum Genet*. 2005;11:115-31.

- [16] Cadigan, K.M. and Nusse, R. Wnt signaling: a common theme in animal development. *Genes Dev* 1997;11:3286-3305.
- [17] Bhanot, P., Brink, M., Samos, C.H., Hsieh, J.C., Wang, Y., Macke, J.P., Andrew, D., Nathans, J. and Nusse, R. A new member of the frizzled family from *Drosophila* functions as a Wingless receptor. *Nature* 1996;382:225-230.
- [18] Tamai, K., Semenov, M., Kato, Y., Spokony, R., Liu, C., Katsuyama, Y., Hess, F., Saint-jeanette, J.P. and HE, X. LDL-receptor-related proteins in Wnt signal transduction. *Nature* 2000;407: 530-535.
- [19] Rhee CS, Sen M, Lu D, Wu C, Leoni L, Rubin J, Corr M, Carson DA. Wnt and frizzled receptors as potential targets for immunotherapy in head and neck squamous cell carcinomas. *Oncogene*. 2002;21:6598-605.
- [20] Milovanovic T, Planutis K, Nguyen A, Marsh JL, Lin F, Hope C, Holcombe RF. Expression of Wnt genes and frizzled 1 and 2 receptors in normal breast epithelium and infiltrating breast carcinoma. *Int J Oncol* 2004;25:1337-42.
- [21] Holcombe RF, Marsh JL, Waterman ML, Lin F, Milovanovic T, Truong T. Expression of Wnt ligands and Frizzled receptors in colonic mucosa and in colon carcinoma. *Mol Pathol* 2002;55:220-6.
- [22] Okino K, Nagai H, Hatta M, Nagahata T, Yoneyama K, Ohta Y, Jin E, Kawanami O, Araki T, M Emi. Up-regulation and overproduction of DVL-1, The human counterpart of the *Drosophila* dishevelled gene, in cervical squamous cell carcinoma. *Oncol Rep* 2003;10:1219-1223.
- [23] Rousset, R, Mack, J.A., Wharton, K.A, JR., Axelrod, J.D., Cadigan, K.M., Fish, M.P, Nusse, R. and Scott, M.P. Naked cuticle targets dishevelled to antagonize Wnt signal transduction. *Genes Dev* 2001;15:658-671.
- [24] Sun, TQ, Lu B, Feng, J.J., Reinhard C, Jan YN, Fantl WJ and Williams LT. PAR-1 is a Dishevelled-associated kinase and a positive regulator of Wnt signalling. *Nat Cell Biol* 2001;3:628-636.
- [25] Masatoshi K, Shigetsugu H, Michiko S, Masaki M, Noriko I, Kimihiko H, Ikuo N, Akira K, Kei-ichi, K Nakayama. An F-box protein, FWD1, mediates ubiquitin-dependent proteolysis of β -catenin. *EMBO J* 1999;18:2401-2410.
- [26] Liu Ch, Kato Y, Zhang Z, Viet Minh Do, Yankner B, X He: β -Trcp couples β -catenin phosphorylation-degradation and regulates *Xenopus* axis formation. *PNAS* 1999;96:6273-6278.
- [27] Yanagawa S, Matsuda Y, Lee J, Matsubayashi H, Sese S, Kadowaki T, Akinori I. Casein kinase I phosphorylates the Armadillo protein and induces its degradation in *Drosophila*. *EMBO* 2002;21:1733-1742.
- [28] Su T, Chang J, Lin T, Lee T, Chen J, C Lin: Mutational analysis of CTNNB1 and Axin1, the components of Wnt pathway, in cervical carcinomas. *Oncol Rep* 2003;10:1195-200.
- [29] Giles R, van Es J, H Clevers. Caught up in a Wnt storm: Wnt signalling in cancer. *Biochim Biophys Acta* 2003;1653:1-24.
- [30] Fagotto F, Glück U, B Gumbiner. Nuclear localization signal-independent and importin/karyopherin-independent nuclear import of β -catenin. *Curr Biol* 1998;8:181-198.

- [31] Wetering M, Cavallo R, Dooijes D, van Beest M, van Es J, Loureiro J, Ypma A, Hursh D, Jones T, Bejsovec A, Peifer M, Mortin M, H Clevers. Armadillo Coactivates Transcription Driven by the Product of the Drosophila Segment Polarity Gene *dTCF*. *Cell* 1997;88:789-799.
- [32] Van de Wetering M, Oosterwegel M, Dooijes D, H Clevers. Identification and cloning of TCF-1, a T lymphocyte-specific transcription factor containing a sequence-specific HMG box. *EMBO J* 1991;10:123-132.
- [33] Zhang T, Otevrel T, Gao Z., Ehrlich SM, Fields JZ and Boman BM. Evidence that APC regulates survivin expression: a possible mechanism contributing to the stem cell origin of colon cancer. *Cancer Res* 2001;61:8664-8667.
- [34] Kim PJ., Plescia J, Clevers H, Fearon ER and Altieri DC. Survivin and molecular pathogenesis of colorectal cancer. *Lancet* 2003;362:205-209.
- [35] Wolf D, Rodova M, Miska E, Calvet J, T Kouzarides. Acetylation of β -Catenin by CREB-binding Protein (CBP). *J Biol Chem* 2002;277:25562-25567.
- [36] Kinoshita T, Shirasawa H, Shino Y, Moriya H, Desbarats L, Eilers M, B Simizu. Transactivation of prothymosin alpha and c-myc promoters by human papillomavirus type 16 E6 protein. *Virology* 1999;232:53-61.
- [37] Gu W, Bhatia K, Magrath IT, Dang CV, Dalla-Favera R. Binding and suppression of the Myc transcriptional activation domain p107. *Science* 1994;264:251-254.
- [38] Van De Wetering, M., Sancho, E., Verweij, C., De Lau, W., Oving, I., Hurlstone, A., Van Der Horn, K., Batlle, E., Coudreuse, D., Haramis, A.P., Tjonn-Pon-Fong M, Moerer P, Van Den Born M, Soete G, Pals S, Eilers M, Medema R. and Clevers H. The beta-catenin/TCF-4 complex imposes a crypt progenitor phenotype on colorectal cancer cells. *Cell* 2002;111:241-250.
- [39] Ocadiz R, Saucedo R, Cruz M, Graef A, Gariglio P. High correlation between molecular alterations of the c-myc oncogene and uterine cervix carcinoma. *Cancer Res* 1987;47:4173-4177.
- [40] Solinas-Toldo S, Dürtz M, Lichter P. Specific chromosomal imbalances in human papillomavirus transfected cells during progression toward immortality. *Proc Natl Acad Sci* 1997;94:3854-3859.
- [41] Hidalgo A, Shewe C, Petersen S. Human papillomavirus status and chromosomal imbalances in primary cervical carcinomas and tumor cells lines. *Eur J Cancer* 2000;36:542-548.
- [42] Brantjes H, Roose J, van de Wetering M, H Clevers. All Tcf HMG box transcription factors interact with Groucho-related co-repressors. *Nucl Acids Res* 2001;29:1410-1419.

Tables and Figures

Table 1. Summary of the *In situ* hybridizations and immunohistochemical stainings results of cervical tissues.

Figure 1. Gene cluster of the Wnt Signaling Pathway.

Expression pattern of 27 Wnt signaling genes in cervical cancer. Rows represent individual genes, columns represent individual samples. Each cell in the matrix represents the expression level of a single transcript compared by the normal transcript. The red color indicates the transcript levels above the normal for that gene across all samples, respectively. Color saturation is proportional to magnitude of difference from the normal mean.

Figure 2. Relative mRNA expression levels of FZD, GSK3B, C-MYC and PPARD in cervical tumors detected by RT-PCR.

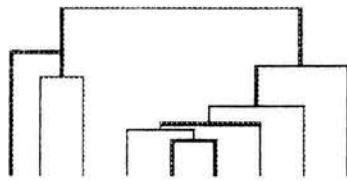
One hundred of total RNA was employed for RT-PCR experiments. The used samples were: line 1) positive control, line 2) negative control, line 3) C33, 4) HeLa, 6) CaLO, 6) INBL, 7) CaSki, 8) SiHa, 9) tumoral sample #T01, 10) tumoral sample #T03, 11)) tumoral sample #T11, 12) tumoral sample #T04, 13) tumoral sample #T10, 14) Normal sample #N15, 15) Normal sample N20, 16) Normal sample #N18, 17) Normal sample #N23 and 18) Normal sample #N30. A 100 bp ladder was used as molecular weight marker. Values in the graph are presented as ratio of densitometric readings of samples to corresponding β -actin samples.

Figure 3. *In situ* hybridization and Immunohistochemical staining of cervical tissues. Capital letters show *In situ* hybridizations for FZD2, GSK3B, C-MYC and PPARD probes. A) Example of positive control (FZD2 in colon cancer), B) Example of negative control (FZD2 for Cervical cancer), C) normal tissue negative for FZD2 hybridization, D) SILHg tissue positive for FZD2 hybridization, E) Positive hybridization for FZD2 in CC, F) normal cervical epithelium, positive for GSK3B hybridization, G) normal cervical epithelium, positive for GSK3B hybridization, H) positive hybridization for GSK3B in SILHg, I) positive hybridization for GSK3B in CC, J) C-MYC negative hybridization in normal epithelium, K) as in D and H SILHg positive for C-MYC, L) Positive hybridization for C-MYC in CC, M) as in C normal tissue negative for PPARD hybridization, N) as in D, H, and K SILHg positive for PPARD and O) positive hybridization for PPARD in CC.

Lower case show results of immunostainings for C-MYC a) Breast cancer sample (as positive control for C-MYC staining), b) negative control; a CC sample without primary antibody, c) one normal cervical tissue showing a weak staining for C-MYC, d) SILHg positive staining for C-MYC and e) CC tissue showing positive immunoreaction for C-MYC.

Original magnifications were X40 for A, B, F and a; and 10X were used for C, D, E, G, H, I, J, K, L, M, N, O, b, c, d and e.

	Total of cases	% Positivity in ISH			% of Positivity in IHC	
		FZD2	GSK3B	C-MYC	PPARD	C-MYC
Normal tissues	20	2	3	3	1	4
SiLHg	15	15	12	14	13	13
CC tissues	15	15	15	15	15	15

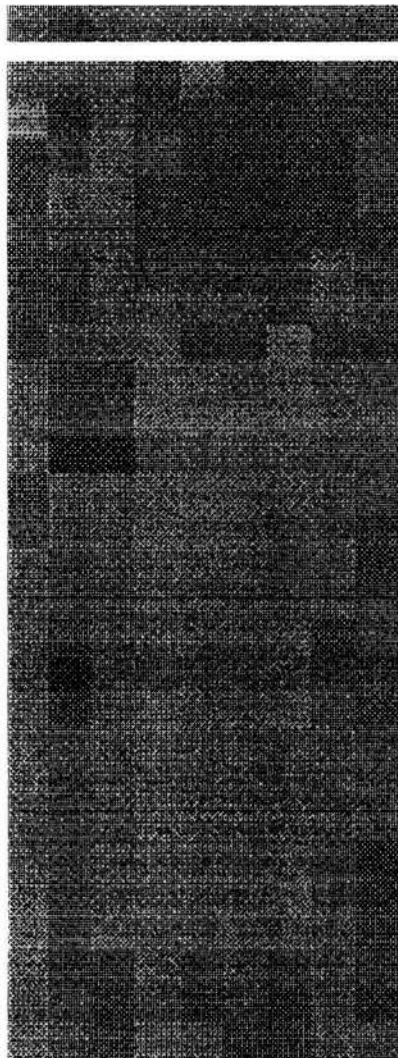


R/G Ratio

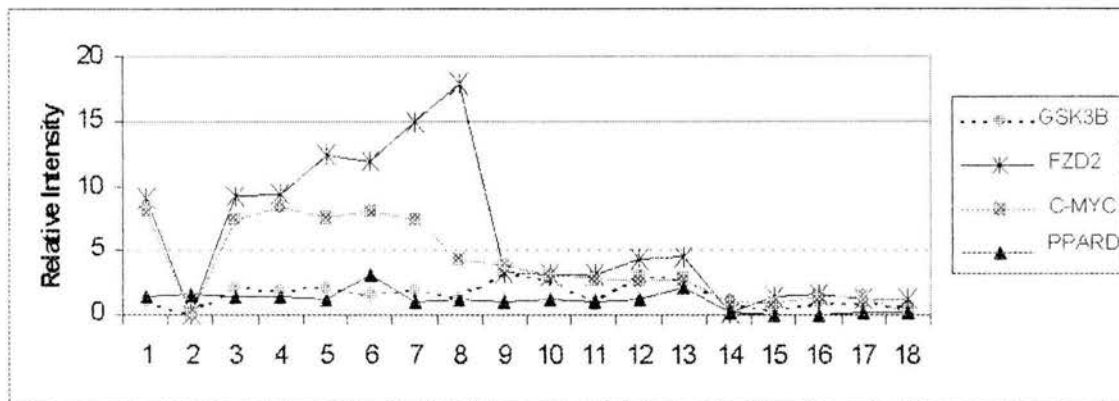
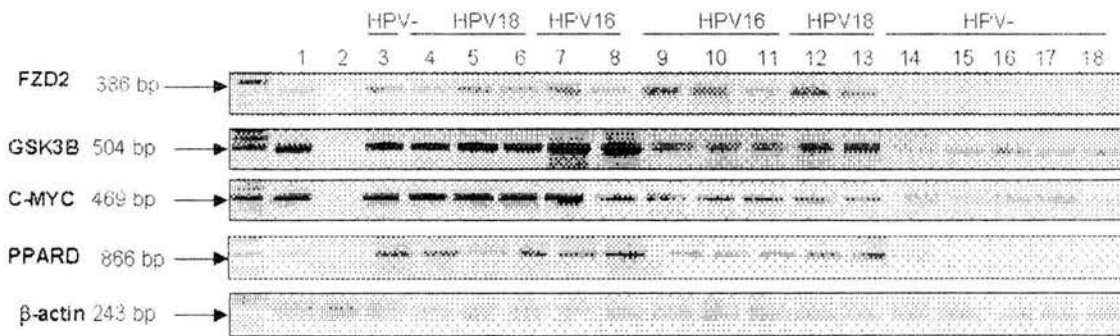


1 >2

C33
T31
T64
HeLa
CaLO
INBL
T07
SiHa
CaSKI



Hs . 74375
Hs . 118640
Hs . 283738
Hs . 184434
Hs . 171271
Hs . 62661
Hs . 20912
Hs . 194679
Hs . 153684
Hs . 23672
Hs . 155140
Hs . 171862
Hs . 94234
Hs . 23598
Hs . 15114
Hs . 105700
Hs . 79070
Hs . 4909
Hs . 165843
Hs . 100221
Hs . 108301
Hs . 78802
Hs . 108708
Hs . 28935
Hs . 279565
Hs . 236963
Hs . 78465



Chromosomal imbalances in four new uterine cervix carcinoma derived cell lines

Alfredo Hidalgo¹, Alberto Monroy², Rosa Ma Arana³, Lucía Taja⁴,
Guelaguetza Vázquez¹ and Mauricio Salcedo*¹

Address: ¹Laboratorio de Oncología Genómica, Unidad de Investigación Médica en Enfermedades Oncológicas, Hospital de Oncología, Centro Médico Nacional Siglo XXI-IMSS, Mexico, ²Laboratorio de Diferenciación Celular y Cáncer, FES Zaragoza, UNAM, Mexico, ³Servicio de Genética, Hospital General de México, SS, Mexico and ⁴División de Investigación Básica, Instituto Nacional de Cancerología, SS, Mexico

Email: Alfredo Hidalgo - alfhm@yahoo.com; Alberto Monroy - albertomon@yahoo.com; Rosa Ma Arana - aranat@prodigy.net.mx; Lucía Taja - ltaja_chayeb@yahoo.com; Guelaguetza Vázquez - guelav@yahoo.com; Mauricio Salcedo* - maosal89@yahoo.com

* Corresponding author

Published: 20 March 2003

Received: 20 January 2003

BMC Cancer 2003, 3:8

Accepted: 20 March 2003

This article is available from: <http://www.biomedcentral.com/1471-2407/3/8>

© 2003 Hidalgo et al; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.

Abstract

Background: Uterine cervix carcinoma is the second most common female malignancy worldwide and a major health problem in Mexico, representing the primary cause of death among the Mexican female population. High risk human papillomavirus (HPV) infection is considered to be the most important risk factor for the development of this tumor and cervical carcinoma derived cell lines are very useful models for the study of viral carcinogenesis. Comparative Genomic Hybridization (CGH) experiments have detected a specific pattern of chromosomal imbalances during cervical cancer progression, indicating chromosomal regions that might contain genes that are important for cervical transformation.

Methods: We performed HPV detection and CGH analysis in order to initiate the genomic characterization of four recently established cervical carcinoma derived cell lines from Mexican patients.

Results: All the cell lines were HPV18 positive. The most prevalent imbalances in the cell lines were gains in chromosomes 1q23-q32, 3q11.2-q13.1, 3q22-q26.1, 5p15.1-p11.2, this alteration present as a high copy number amplification in three of the cell lines, 7p15-p13, 7q21, 7q31, 11q21, and 12q12, and losses in 2q35-qter, 4p16, 6q26-qter, 9q34 and 19q13.2-qter.

Conclusions: Analysis of our present findings and previously reported data suggest that gains at 1q31-q32 and 7p13-p14, as well as losses at 6q26-q27 are alterations that might be unique for HPV18 positive cases. These chromosomal regions, as well as regions with high copy number amplifications, coincide with known fragile sites and known HPV integration sites. The general pattern of chromosomal imbalances detected in the cells resembled that found in invasive cervical tumors, suggesting that the cells represent good models for the study of cervical carcinoma.

Background

Cervical carcinoma stands as the first cause of death among the Mexican female population with 14 deaths per 100,000 women with 15 years old or more, representing

34.2 % of all new female cancer cases reported [1]. High risk human papillomavirus (HPV) infection is considered to be the most important risk factor associated with the

Table 1: Chromosomal imbalances detected in hpv18 positive cell lines

Cell line	Tumor stage	DNA Losses	DNA Gains	Total
CALO	II-B	1pter-p32, 2q35-qter, 4pter-p15.2, 4q32-qter, 5q33-qter, 6q24-qter, 8pter-p22, 8q24.2-qter, 9q32-qter, 11q12, 12p13, 12q22-qter, 13q12, 15q11.2-q12, 16p, 16q22-qter, 17, 18q11, 19, 20, 21, 22. Total:21	1p31-p12, 1q22-q31, 2p15, 2q21-q32, 3p22, 3q11.2-q26.3, 4q26-q31.2, 5p, 5q11.2-q23, 6q22, 7p15-p12, 7q21-q31, 9p23-p21, 11q14-q22, 12p11.2, 12q15-q21, 13q21-q31, 18q12. Total: 18	39
VIPA	II-A	1p32-pter, 2p24, 2q35-qter, 4p15.3-pter, 4q33-qter, 5q33-qter, 6q25-qter, 7p22, 7q33 qter, 8p21-pter, 9q32-qter, 10p13-pter, 10q24-qter, 11q12, 12p13, 12q22-qter, 14q31-qter, 16p, 16q13-qter, 17, 18q22, 19, 20p13, 20q, 21q22, 22. Total:26	1p12-p22, 1q23-q32, 2q21-q32.3, 3p22, 3p13-p12, 3q13.1-q13.3, 3q22-q25, 4q13, 4q22-q31.3, 5p, 5q11.2-q12, 5q14-q23, 6p12-p21.1, 6q12-q16, 7p12-p21, 7q21-q31, 8q21.1-q22, 9p21-p22, 10q21, 11p13, 11q14-q22, 12p11.2, 12q14-q21, 13q14-q31, 14q12-q13, 15q21. Total: 26	52
INBL	IV-A	2q34-qter, 4p, 4q32-qter, 6q24-qter, 8p22, 13q12, 18p21-pter, 19p13.2-pter. Total: 8	1q12-q41, 3q12-q26.3, 5p, 6p24-p22, 6p21.1-p11.1, 7p14-p11.1, 7q21-q31, 11p15, 11q21, 12q12, 15q22-q24. Total: 11	19
ROVA	IV-A	1p35-pter, 2p25, 2q35-qter, 3q29-qter, 4p16, 5q35, 6q26-qter, 7p22, 8p23, 8q24.1-qter, 9q33-qter, 10q26, 12p13, 12q24.1-qter, 16p, 16q22-qter, 17p12-pter, 17q23-qter, 18q22-qter, 19, 20q13.1-qter, 21q22, 22q. Total: 23	1p12-p31, 1q22-q32, 2p11.2-p16, 2q21-q33, 3p12-p22, 3q13.1-q26.3, 4p12-p15.2, 4q12-q31.3, 5p12-p15.1, 5q11.2-q31, 6p22-p23, 6p12-p, 6q12-q15, 6q22, 7p12-p15, 7q21-q31, 8p22, 8q13, 9p21-p23, 9q21, 11p14, 11q14-q24, 12p12, 12q12-q21, 13q14-q31, 14q11.2-q24, 15q15-q21, 15q24, 18q11.2-q12. Total: 28	51
Average		19.5	20.7	ANCA:40.2

All tumors were diagnosed as squamous cell carcinomas. ANCA: Average Number of Chromosomal Aberrations, obtained dividing the total number of observed alterations between the total number of cases.

development of this tumor, and is present in 99.7% of the invasive cervical tumors worldwide [2].

Comparative Genomic Hybridization (CGH) is a method employed in cancer genomics that allows the detection of DNA gains or losses at the genome level in a single hybridization experiment, indicating cytogenetic regions that might be involved in the transformation process. CGH has detected a specific pattern of chromosomal imbalances associated with particular stages of cervical transformation, and with different biological behaviors [3-10]. In this paper, we evaluated the presence of HPV DNA and analyzed the pattern of chromosomal imbalances using CGH in four cell lines established from tumor explants of Mexican patients. The establishment and research use of two of these cell lines has been previously reported [11,12]. Further genomic characterization of these lines will open new possibilities for understanding cervical carcinoma, since the coincidence between the chromosomal imbalances present in these cell lines and patterns found in cervical tumors indicate that they are good models for the study of cervical cancer.

Methods

The cell lines were established from stages IIA and IVA squamous cell cervical carcinoma explants from Mexican females (Table 1) at the National University of Mexico, as

previously described [11,12]. Cell lines were designated CALO, INBL, VIPA and ROVA.

HPV detection was done using the L1 consensus primers MY09/MY11. After denaturation at 94°C for 5 minutes, 100 ng of DNA were subjected to 40 cycles of 94°C for 1 min, 55°C for 2 min and 73°C for 3 min, with a final extension step of 7 min at 72°C. The amplicon was labeled using the Big Dye kit (Applied Biosystems, Foster City, CA) and sequenced with an ABI 373 automated sequencer. BLAST sequence comparison was done in order to know the HPV type.

All CGH procedures were performed using reagents and kits from Vysis Inc. (Downers Grove, IL), following the manufacturer's instructions. Cell line DNA was labeled using the CGH nick translation kit, and the hybridization mixture was prepared according to the CGH reagent kit, consisting of 200 ng of spectrum-green labeled cell line DNA, 100 ng of spectrum-red labeled normal male reference DNA and 20 µg of human Cot-1 DNA. Hybridization was performed over 3 days at 37°C on normal male metaphase spreads (Vysis, Downers Grove, IL).

Digital images were collected using the Smart Capture (Vysis) software with a charged coupled device camera mounted on an epifluorescence microscope. An average of 15 metaphases per case were analysed using the Vysis

CGH software using a >1.2 ratio threshold for detecting DNA gains and <0.8 for losses, these thresholds were chosen based on previous normal VS. normal hybridizations. Sex chromosomes and heterochromatic areas (centromeric and paracentromeric regions of chromosomes 1, 9, 16, p arms of acrocentric chromosomes) were excluded from the analysis. A "reverse painting" experiment was used as a control for the CGH results. CGH data from the cell lines is available at the NCI and NCBI Spectral Karyotyping (SKI) and CGH Database <http://www.ncbi.nlm.nih.gov/sky/skyweb.cgi>.

Results and Discussion

HPV DNA was detected in all the cell lines. Sequence analysis identified HPV18 in all the samples. All the cell lines presented chromosomal imbalances. A line summary ideogram of the chromosomal imbalances in the cell lines is presented in Figure 1. A complete description of the imbalances detected in each cell line is presented in Table 1. On the average, 19.5 DNA losses and 20.7 gains were detected in the samples, with an average number of chromosomal aberrations of 40.2 alterations per case, ranging from 52 alterations in VIPA to 19 alterations in INBL. The most prevalent imbalances were gains in chromosomes 1q23-q32, 3q11.2-q13.1, 3q22-q26.1, 5p15.1-p11.2. This alteration is present as a high copy number amplification in three of the cell lines, 7p15-p13, 7q21, 7q31, 11q21, and 12q12, and losses in 2q35-qter, 4p16, 6q26-qter, 9q34 and 19q13.2-qter. Comparison between chromosomal imbalances in these cell lines and those found in cervical tumors; indicate the presence of common genetic alterations, suggesting that they represent their tissue of origin. These common regions include gains in 3q22-q26.2 and 5p, and losses on 2q35-qter, 4q32-qter and 18q. Some of these alterations have been detected in early stages of cervical transformation and are conserved in advanced tumors or have been associated with the presence of metastases and with a worse clinical behavior [6,7,10], making the cells good models for the study of genes involved in the maintenance of the transformed phenotype. Compared with the imbalances previously found on HPV18 positive cases [9], there is coincidence for the DNA gains at 1q31-q32, 3q22-q26.3, 5p14 and 7p13-p14, as well as in DNA losses at 2q33-qter and 6q26-q27. Differences among the cell lines and data from invasive cervical tumors were also found. For example, we did not find deletions of 3p or 11q regions in the cell lines. This situation might be explained by the enrichment or selection of a particular cell clone during the cell line establishment.

A common observation in uterine cervix carcinomas is the integration of HPV DNA into the cellular genome. However, the consequences of this integration event are only poorly understood. Differences in the frequency of HPV18 and HPV16 integration have been detected, 72%

of the HPV16 tumors presented viral integration, while HPV18 positive tumors had viral integration in 100% of the cases [13,14]. This is a possible explanation for the observed HPV18 incidence in the studied cell lines, since the integration event might represent a growth advantage for the cell in culture. Analysis of our present findings and previously reported data suggest that gains at 1q31-q32 and 7p13-p14, as well as losses at 6q26-q27 are alterations that are unique for HPV18 positive cases. HPV integration at the 1q31-q32 has been found in nasal epithelial cells immortalized with HPV16 DNA and also contains the short tandem repeat region F13B [15]. The NEK7 (NIMA (never in mitosis gene a)-related kinase 7, 1q31.3) gene is located in this region and plays an important role in the control of mitosis initiation [16]. The role of this gene in human cancer has not been explored. Regarding the 7p15-p13 region, a lymphocyte fragile site has been detected at 7p13. high copy number amplification of this region was also found by us in primary cervical carcinomas [9] and gain of 7p was also found in HPV transfected cells during the immortalization process [4]. The PE5L gene, (7p13-p14) presents a sequence homology with the HPV18 E5 gene, making this gene a possible target for recombination and viral integration [17]. However, alignment of the PE5L sequence against the November 2002 freeze of the Human Genome working draft, showed the best matching score on 7p11.2 (not shown). The public availability of viral-cellular fusion transcripts or DNA sequences, combined with data from the human genome, will certainly help to better define areas of viral integration for further investigation.

The lost cytogenetic region 6q26 comprises the fragile site FRA6E, the third most frequently observed common fragile site. Deletion of the 6q26-q27 region has been found in approximately 44% of all tumor types analyzed by CGH [18], and in cervical carcinoma loss of heterozygosity was found 39% of the cases [19]. An interesting gene located in this region is the mitogen-activated protein kinase kinase kinase 4 (MAP3K4), a major mediator of environmental stresses that activate the CSBP2 MAPK pathway, which, in turn, regulates the expression of TNF and several cytokines, and controls the initiation of the G2/M checkpoint after UV radiation [20].

We also found high copy number amplifications in regions 7q21, 7q31, 11q21 and 11q12. Region 7q21 contains the rare, folic acid type FRA7E fragile site and a cellular-HPV fusion transcript has been detected at this area, with no sequence match to a known gene [21]. 7q31 has been reported to be the site of a chromosomal breakpoint in cervical carcinoma cell lines [22] and contains the c-Met oncogene, whose over-expression has been correlated to the diameter of the primary tumor, deep cervical stromal invasion, presence of metastatic lymph node and

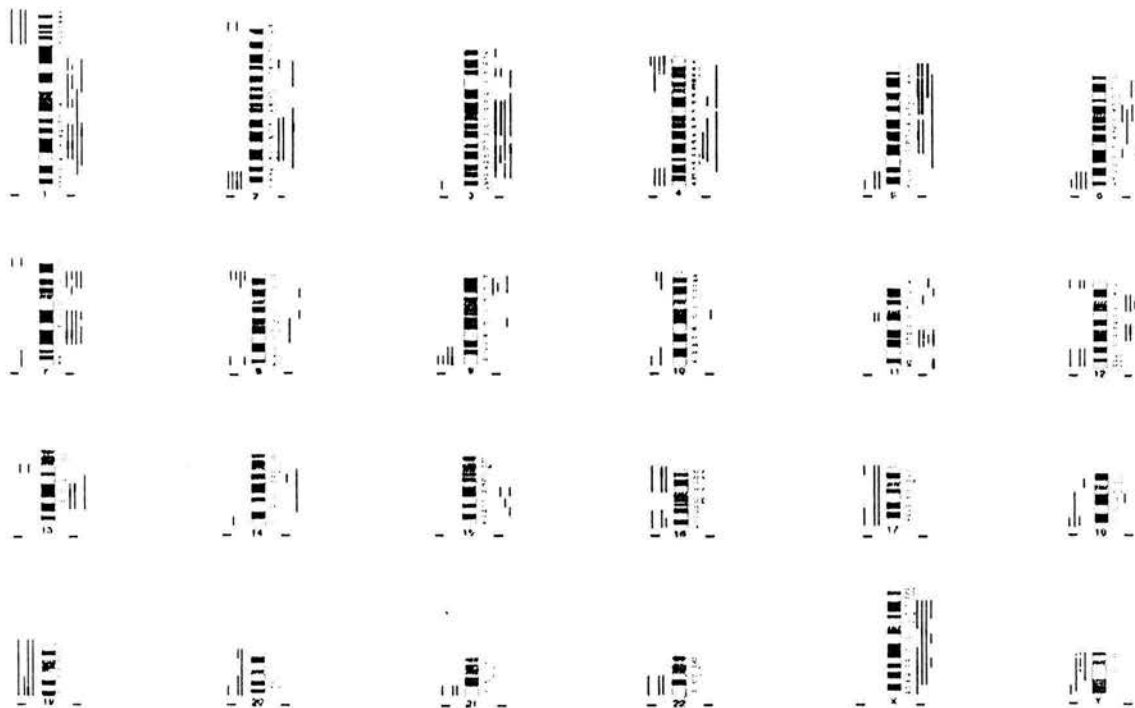


Figure 1

CGH ideograms of the genetic imbalances detected in the cell lines. The lines at the right of the chromosome ideogram represent DNA gains, lines at the left, DNA losses. Thick lines represent high copy number amplifications or multi copy deletions (ratio >1.4 or <0.6). The mark closest to the ideogram represents the cell lines CALO, followed by VIPA, INBL and J-ROVA.

number of metastatic lymph nodes [23]. The 11q21-q22 region was also found in our cell lines, this region has been found amplified in some cervical carcinoma cell lines and over-expression of the cIAP1 gene, an apoptosis inhibitor and possible target of the amplification, was related to a poorer overall survival and local recurrence-free survival in cervical carcinoma [24].

It is interesting to notice that almost all of the possible HPV18 related sites of chromosomal imbalance, as well as the regions with high copy number amplifications in our cell lines correspond to the location of fragile sites in the human genome. Since HPV integration preferentially seems to target these kind of fragile sites, it is possible that HPV18 is integrated into these altered chromosomal regions in our cell lines. Recently, a comprehensive charac-

terization of cervical cancer cell lines gave very useful information about specific chromosomal alterations, DNA gains, losses and clusters of chromosome break-points in cervical cancer cell lines and their relation to HPV presence and integration [24]. Currently, such a characterization of our newly established cell lines is being carried out.

Conclusions

High throughput genetic analysis often require vast amounts of tissue that sometimes is not available, making the cell lines the model of choice to overcome this problem. Here, we report the presence of HPV18 and the pattern of chromosomal imbalances detected in four newly established uterine cervix carcinoma derived cell lines. The pattern of genetic imbalances in these cell lines sug-

gest that they represent good models for cervical carcinoma and HPV18 infection, providing another useful cellular model for the discovery and functional analysis of genomic alterations involved in cervical cancer when this viral type is present.

List of abbreviations

CGH: Comparative Genomic Hybridization. HPV: Human Papilloma virus. SKI: Spectral karyotyping

Competing interests

None declared.

Authors contributions

AH: Carried out the CGH experiments and data analyses, writing and designing paper and figures. AM: Cell line establishment. RMA: CGH metaphase analyses. LT: Cell cultures and DNA extraction. GV: HPV typing. MS: Design and coordination of the study.

Acknowledgments

This work was partially funded by grants from the CONACyT (34686-M) and IMSS (FP-2001/043), AH, GV and LT are recipients of a CONACyT scholarship and AH and GV receive a DGEP-UNAM and IMSS scholarships.

References

- Mexican Ministry of Health Registro Histopatológico de Neoplasias Malignas, Compendio de mortalidad y morbilidad. Secretaría de Salud, México 1998.
- Walboomers J, Jacobs M, Manos M, Bosch X, Kummer A, Shah K, Snijders P, Peto J, Meijer C and Muñoz N Human Papilloma Virus is a necessary cause of invasive cervical cancer worldwide. *J Pathol* 1999, **189**:12-19
- Heselmeyer K, Schröck E, Du Manoir S, Blegen H, Shah K, Steinbeck R, Auer G and Ried T Gain of chromosome 3q defines the transition from severe dysplasia to invasive carcinoma of the uterine cervix. *Proc Natl Acad Sci USA* 1996, **93**:479-484
- Solinas-Toldo S, Dürst M and Lichter P Specific chromosomal imbalances in human papillomavirus transfected cells during progression toward immortality. *Proc Natl Acad Sci USA* 1997, **94**:3854-3859
- Kirchhoff M, Rose H, Laub B, Maahr J, Gerdes T, Lundsteen C, Bryndorf T, Kryger N, Christensen L, Aage S and Philip J Comparative genomic hybridization reveals a recurrent pattern of chromosomal aberrations in severe dysplasia/carcinoma *in situ* of the cervix and advanced-stage cervical carcinoma. *Genes Chrom Cancer* 1999, **24**:144-150
- Allen D, White D, Hutchins A, Scully J, Fabrizio S, Garland S and Armes Progressive genetic aberrations detected by comparative genomic hybridization in squamous cell cervical cancer. *Br J Cancer* 2000, **83**:1659-1663
- Dellas A, Torhorst J, Jiang F, Proffit J, Schultheiss E, Hotzgreve W, Sauter G, Mihatsch M and Moch H Prognostic value of genomic alterations in invasive cervical squamous cell carcinoma of clinical stage IB detected by comparative genomic hybridization. *Cancer Res* 1999, **59**:3475-3479
- Kirchhoff M, Rose H, Petersen B, Maahr J, gerdes T, Philip J and Lundsteen C Comparative genomic hybridization reveals non-random chromosomal aberrations in early pre-invasive cervical lesions. *Cancer Genet Cytogenet* 2001, **129**:47-51
- Hidalgo A, Schewe C, Petersen S, Salcedo M, Gariglio P, Schlüns K, Dietel M and Petersen I Human papilloma virus status and chromosomal imbalances in primary cervical carcinomas and tumor cell lines. *Eur J Cancer* 2000, **36**:542-548
- Umayahara K, Numa F, Suehiro Y, Sakata A, Nawata S, Ogata H, Suminami Y, Sakamoto M, Sasaki K and Kato H Comparative Genomic hybridization detects genetic alterations during early stages of cervical cancer progression. *Genes Chromosomes Cancer* 2002, **33**:98-102
- Caceres-Cortes J, Alvarado-Moreno J, Waga K, Rangel-Corona R, Monroy-Garcia A, Rocha-Zabaleta L, Urdiales-Ramos J, Weiss-Steider B, Haman A, Hugo P, Brousseau R and Hoang T Implication of tyrosine kinase receptor and steel factor in cell density-dependent growth in cervical cancers and leukemias. *Cancer Res* 2001, **61**:6281-6289
- Monroy-García A, Ortiz-Navarrete V, Mora-García M, Flores-Borja F, Diaz-Quñones A, Isibasi-Araujo A, Trejo-Becerril C, Chacón-Salinas R, Hernández-Montes J, Granados Arreola J, de Leo C and Weiss-Steider B Identification of peptides presented by HLA class I molecules on cervical cancer cells with HPV-i8 infection. *Immunol Let* 1999, **67**:167-177
- Cullen A, Reid R., Campion M. and Lorincz A Analysis of the physical state of different human papillomavirus DNAs in intraepithelial and invasive cervical neoplasia. *J Virol* 1991, **65**:606-612
- Badaracco G, Venuti A, Sedati A and Marcante ML HPV16 and HPV18 in genital tumors: Significantly different levels of viral integration and correlation to tumor invasiveness. *J Med Virol* 2002, **67**:574-82
- Debiec-Rychter M, Zukowski K, Wang CY and Wen WN Chromosomal characterizations of human nasal and nasopharyngeal cells immortalized by human papillomavirus type 16 DNA. *Cancer Genet Cytogenet* 1991, **52**:51-61
- Minoguchi S, Minoguchi M and Yoshimura A Differential control of the NIMA-related kinases, Nek6 and Nek7, by serum stimulation. *Biochem Biophys Res Commun* 2003, **301**:899-906
- Geisen C, Delius H, Lichter P and Kahn T A Transcribed human sequence related to the mouse HCl and the human papillomavirus type 18 E5 genes is located at chromosome 7p13-14. *Hum Mol Genet* 1995, **4**:1337-1345
- Knuutila S, Aalto Y, Autio K, Bjorkqvist AM, El-Rifai W, Hemmer S, Huhta T, Kettunen E, Kiuru-Kuhlefelt S, Larramendy ML, Lushnikova T, Monni O, Pere H, Tapper J, Tarkkanen M, Varis A, Wasenius VM, Wolf M and Zhu Y DNA copy number losses in human neoplasms. *Am J Pathol* 1999, **155**:683-694
- Mazurenko N, Attaleb M, Gritsko T, Semjonova L, Pavlova L, Sakharova O and Kisseljov F High resolution mapping of chromosome 6 deletions in cervical cancer. *Oncol Rep* 1999, **6**:859-863
- Takekawa M, Posas F and Saito H A human homolog of the yeast Ssk2/Ssk22 MAP kinase kinase kinases, MTK1, mediates stress-induced activation of the p38 and JNK pathways. *EMBO J* 1997, **16**:4973-4982
- Wentzensen N, Ridder R, Klaes R, Vinokurova S, Schaefer U and Doberitz MK Characterization of viral-cellular fusion transcripts in a large series of HPV16 and 18 positive anogenital lesions. *Oncogene* 2002, **21**:419-426
- Thein A, Trkova M, Fox M and Parrington J The application of comparative genomic hybridization to previously karyotyped cervical cancer cell lines. *Cancer Genet Cytogenet* 2000, **116**:59-65
- Baykal C, Ayhan A, Al A, Yuce K and Ayhan A Overexpression of the c-Met/HGF receptor and its prognostic significance in uterine cervix carcinomas. *Gynecol Oncol* 2003, **88**:123-129
- Imoto I, Tsuda H, Hirasawa A, Miura M, Sakamoto M, Hirohashi S and Inazawa J Expression of cIAP1, a target for 11q22 amplification, correlates with resistance of cervical cancers to radiotherapy. *Cancer Res* 2002, **62**(17):4860-6

Pre-publication history

The pre-publication history for this paper can be accessed here:

<http://www.biomedcentral.com/1471-2407/3/8/prepub>