

**DETERMINACIÓN Y ANÁLISIS DE LA INTEGRACIÓN DEL
VIRUS DEL PAPILOMA HUMANO 16 EN EL GENOMA DE PACIENTES
DIAGNOSTICADOS CON CÁNCER Y SU POSIBLE RELACIÓN CON LA
ETIOLOGÍA DE LA ENFERMEDAD**

Nicole Díaz Moreno

**UNIVERSIDAD ICESI
FACULTAD DE CIENCIAS NATURALES
BIOLOGÍA
Cali
2016**

**DETERMINACIÓN Y ANÁLISIS DE LA INTEGRACIÓN DEL
VIRUS DEL PAPILOMA HUMANO 16 EN EL GENOMA DE PACIENTES
DIAGNOSTICADOS CON CÁNCER Y SU POSIBLE RELACIÓN CON LA
ETIOLOGÍA DE LA ENFERMEDAD**

Nicole Díaz Moreno

**TRABAJO DE GRADO PARA OPTAR AL TÍTULO DE
PREGRADO EN BIOLOGÍA**

DIRECTOR ANDRÉS CASTILLO, PhD

CO-DIRECTOR JULIO CESAR OSORIO, ESTUDIANTE PhD

Cali

2016

AGRADECIMIENTOS

Principalmente al Dr. Andrés Castillo por su paciencia y compromiso con el arte de enseñar la ciencia y por permitirme participar en este proyecto.

A Julio C. Osorio por su entrega, enseñanza y apoyo a lo largo del proyecto.

A los Dres. Carlos Valderrama y Leonardo Herrera por sus enseñanzas y acompañamiento en el proceso académico.

A mi familia por todo su apoyo, amor y comprensión durante este proceso académico.

Y finalmente a todas las personas que de alguna manera contribuyeron a mi crecimiento académico y personal.

CONTENIDO

1. INTRODUCCION	12
2. DESCRIPCIÓN DEL TRABAJO.....	13
2.1. Planteamiento de la pregunta y justificación	13
2.2. Marco teórico y estado del arte	13
2.3. Objetivos	20
2.4. Metodología.....	20
2.4.1 Metodología utilizada en el laboratorio	20
2.4.2 Metodología <i>in silico</i>	22
2.4.3 Matriz de marco lógico.....	26
2.5. Resultados	27
2.5.1 Resultados de laboratorio	27
2.5.2 Resultados del estudio <i>in silico</i>	30
2.6. Discusión.....	35
2.7. Conclusiones:.....	38
2.8. Perspectivas.....	38
2.9. Bibliografía	39
3. ANEXOS	42

Lista de Tablas

Tabla 1.	Descripción de la población del estudio por género y edad.....	28
Tabla 2.	Sitios anatómicos con eventos de integración del VPH16	29
Tabla 3.	Amplificación del oncogén E2 del VPH16.....	29
Tabla 4.	Cuantificación del número de copias del oncogén E6 en muestras positivas para VPH16	30
Tabla 5.	Cuantificación del número de copias del oncogén E2 en muestras positivas para VPH16	30
Tabla 6.	Categorías agrupadas por tipos histopatológicos	32
Tabla 7.	Categorías agrupadas por sitio anatómico.	33
Tabla 8.	Variables genómicas presentes en los sitios de integración del VPH16.....	33

Lista de Graficas

- Gráfica 1.** Distribución cromosómica de los sitios de integración para VPH16, VPH18 y otros genotipos 31
- Gráfica 2.** Proporción de sitios de integración del VPH16 en genes sobre el total de genes por cromosoma.....34

Lista de Figuras

Figura 1. Organización genómica del VPH.	14
Figura 2. Ciclo de vida de VPH en infección cervical.	15
Figura 3. Evento de integración	15
Figura 4. Gráfica de amplificación por PCR tiempo real.....	19
Figura 5. Determinación del estado físico de VPH por el ensayo de PCR en tiempo real.	22
Figura 6. Flujograma de metodología <i>in silico</i>	23
Figura 7. Resultado de electroforesis para amplificación de muestras de ADN con el uso de cebadores PCO3/PCO4 para la amplificación de β -globina	27
Figura 8. Caso de curvas de amplificación y curvas de Melt positivas para E6 y β -globina	28

Lista de Anexos

Anexo 1. Acta de aprobación ética.....	42
Anexo 2. Muestras y sitios anatómicos correspondientes.....	44
Anexo 3. Artículos seleccionados para la recolección de eventos de integración	45
Anexo 4. Sitios de integración viral presentes en regiones exónicas.	46
Anexo 5. Proporción de sitios de integración de VPH16 en genes sobre el total de genes por cromosoma	47

Abreviaturas

VPH: Virus del Papiloma Humano

PCR: Reacción en cadena de la polimerasa (*del inglés Polymerase Chain Reaction*)

ORF: Marco de lectura abierto (*del inglés Open Reading Frame*).

LCR: Larga región de control (*del inglés Long Control Region*)

RESUMEN

Introducción: En los últimos años se ha venido observando a nivel mundial un aumento en el número de diagnósticos de cáncer asociado a infecciones por el VPH. A pesar de los actuales tratamientos profilácticos aun una importante proporción de pacientes infectados por VPH están muriendo de cáncer, lo que está llevando a una concienciación global de la importancia de identificar todos los factores implicados en el desarrollo del cáncer para un mejor entendimiento con fines preventivos, predictivos y terapéuticos.

Objetivo: Determinar y analizar la integración del VPH16 en el genoma de pacientes diagnosticados con cáncer y su posible relación con la etiología de la enfermedad.

Metodología: Determinar el estado físico del genoma del VPH16 en muestras de cáncer por medio de la técnica PCR en tiempo real, y evaluar *in silico* sitios de integración del VPH 16, por medio del programa computacional *Genome Browser* de la Universidad de California en Santa Cruz (UCSC).

Resultados: Se determinó la presencia de VPH16 en 5 de 56 muestras de cáncer. En las cinco muestras positivas para VPH16, el genoma del virus se encontraba integrado en el genoma del hospedero. Para una de estas muestras se determinó también genoma viral en estado episomal. En el estudio *in silico* se encontró que el 50% de los 532 eventos de integración de VPH16 ocurrían dentro de unidades transcripcionales en el genoma del hospedero.

Conclusión: Los eventos de integración del genoma del VPH16 en regiones codificantes del genoma del hospedero pueden inducir a una inestabilidad genómica en el hospedero por las mutaciones de inserción que se pueden estar generando.

Palabras clave: Virus del Papiloma Humano (VPH), PCR en tiempo real, sitios de integración, variables genómicas, etiología.

SUMMARY

Introduction: In recent years it has been observed a worldwide increase in the number of diagnoses of cancer associated with HPV infections. Despite current prophylactic treatments, even a significant proportion of patients infected with HPV are dying of cancer, which is leading to a global awareness of the importance of identifying all the factors involved in the development of cancer for a better understanding for preventive, predictive, and therapeutic purposes.

Objective: To determine and analyze the HPV16 integration in the genome of patients diagnosed with cancer and their possible relationship to the etiology of the disease.

Methodology: Determine the physical state of the HPV16 genome in cancer samples using the real-time PCR technique, and evaluate *in silico* HPV16 integration sites through the computer program Genome Browser at the University of California Santa Cruz (UCSC).

Results: The presence of HPV16 was determined in 5 of 59 cancer samples. In the five HPV16 positive samples, all of them, the virus genome was integrated into the host genome, wherein one sample showed viral genome in episomal states. To *in silico* study, it was found that 50% of 532 HPV16 viral genomes integration events occurred within transcriptional units in the host genome.

Conclusion: the integration events of HPV16 genome in the coding regions of the host genome can induce genomic instability in the host by insertion mutations. This finding could explain the importance of the integration of the viral genome to cancer development.

Keywords: Human Papillomavirus (HPV), real-time PCR, integration sites, genomic variables, etiology.

1. INTRODUCCION

Los agentes infecciosos son sospechosos de jugar papeles causales en una variedad de tumores malignos humanos. El impacto en la salud pública de los efectos oncogénicos de estas infecciones es considerable. La infección se estima que es responsable de aproximadamente el 17,8% de todos los casos incidentes de cáncer en todo el mundo, que representan el 26,3% de todos los tumores malignos en los países en vías de desarrollo económico y el 7,7% en los países desarrollados (Parkin, 2006).

Dentro de estos agentes infecciosos se encuentran los Virus del Papiloma Humano (VPH), responsables de la infección viral de transmisión sexual más común alrededor del mundo (Xu et al., 2013) determinado como el principal agente etiológico del cáncer de cuello uterino (Nobel assembly, 2008) , asociado a cáncer de vulva, vagina y ano en mujeres, y cáncer de ano y pene en hombres (Muños, 2003; zur Hausen, 2002; Walboomers, 1999; Schiffman, 2003) así como lesiones benignas, y recientemente reportado en diferentes tipos de cáncer en cavidad oral (Castillo, 2006) y en sistema aero-digestivo (Castillo, 2011) en estudios realizados en diferentes poblaciones adultas alrededor del mundo.

La evaluación de la causalidad de estos agentes infecciosos como carcinógenos humanos es difícil dada su naturaleza ubicua, la longitud sustancial de tiempo entre la infección y el evento de cáncer, la naturaleza de cofactores, y la rareza de malignidad entre los infectados (Castillo, 2014). Por lo tanto, un problema central para el epidemiólogo es definir la historia natural de la infección e identificar aquellos factores que están relacionados con el desarrollo de cáncer, por lo que se requeriría de biomarcadores informativos del agente (por ejemplo, carga viral), del huésped (como características del ADN relacionados a los eventos de integración), y de otras exposiciones oncogénicos (tales como el consumo de tabaco) para la comprensión de las interacciones virales-humana y el desarrollo del fenotipo tumorigénico, así como para el desarrollo de intervenciones (Castillo, 2014).

Por lo anterior con el presente estudio, se buscó determinar por medio de la técnica PCR en tiempo real el estado físico y carga viral del genoma de VPH16 en muestras de cáncer embebidas en parafina, e identificar a partir de un estudio *in silico* las regiones moleculares del genoma humano donde ocurren los eventos de integración del VPH16 y su posible relación con la etiología del cáncer.

2. DESCRIPCIÓN DEL TRABAJO

2.1. Planteamiento de la pregunta y justificación

En la última década varios estudios sobre el VPH16 han relacionado su infección con diferentes tipos de cáncer, tales como: cáncer de cérvix, cáncer orofaríngeo, cáncer anogenital, entre otros. En el 2008, el Dr. Harald zur Hausen fue merecedor del premio Nobel de medicina al determinar que los genes oncogénicos E6 y E7 del VPH16 y VPH18 participan en la transformación carcinogénica en el cáncer de cérvix. Sin embargo, la agencia internacional para la investigación del cáncer de la organización mundial de la salud (IARC, por sus siglas en inglés) declaró que aunque la expresión de la oncoproteínas virales son factores de riesgo necesarios aun no son suficientes para el desarrollo de la enfermedad. Por lo tanto, es necesario investigar qué otros factores del ciclo infectivo del virus podrían estar participando en el desarrollo de la enfermedad, como son, su carga viral y la integración en el genoma del huésped.

2.2. Marco teórico y estado del arte

Los VPH son virus de ADN icosaédricos sin envoltura con un diámetro de 55 nm, pertenecientes a la familia Papillomaviridae. Este virus epiteliotrópico contiene una molécula de ADN circular de doble cadena de aproximadamente 8000 pares de bases de longitud contenidos en una cápside compuesta de 72 capsómeros pentaméricos. Aunque su ADN es de doble cadena, sólo una hebra contiene marcos de lectura abierta (*ORF*, por sus siglas en inglés) que se transcriben. En total el ADN viral contiene ocho *ORFs* de los cuales seis hacen parte de los genes tempranos del virus (E1, E2, E4, E5, E6 y E7) que codifican proteínas necesarias para la replicación del virus, la transformación celular y el control de la transcripción viral y dos *ORFs* de los genes tardíos (L1 y L2) que codifican proteínas de la cápside vírica, participando así en el ensamblaje del virión. Adicionalmente cuentan con una región de control larga ("*Long Control Region*", *LCR*) que contiene promotores, sitios de iniciación de replicación del genoma viral y elementos de regulación transcripcional que actúan en *cis*. La *Figura 1* muestra la organización del genoma del VPH16, el tipo de VPH más fuertemente relacionada con el cáncer de cuello uterino (Prabhu & Wilson, 2013).

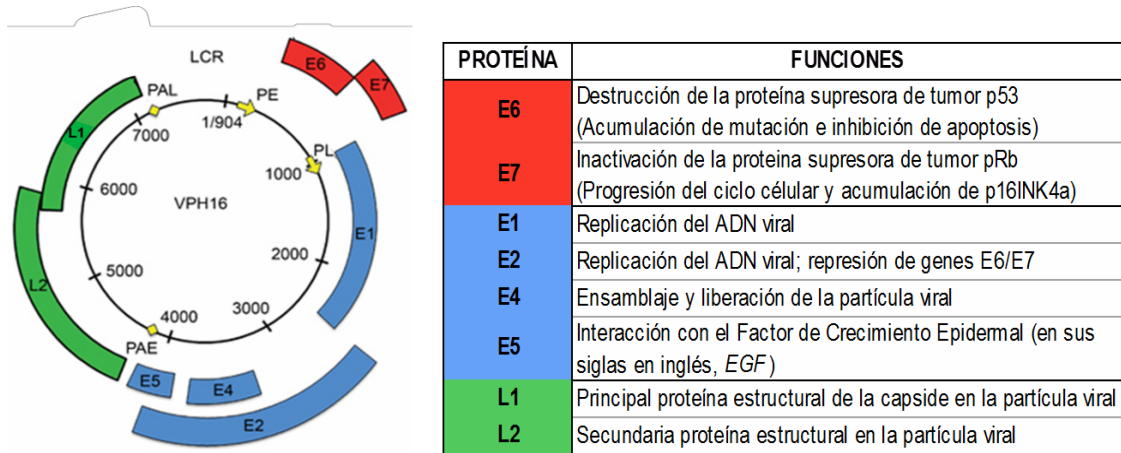


Figura 1. Organización genómica del VPH. Se muestra la distribución de los genes codificantes y sus proteínas. De acuerdo a los marcos abiertos de lectura se distribuye en: genes tempranos E1, E2, E4, E5, E6 y E7, los cuales codifican proteínas necesarias para la replicación viral, transformación celular y control de la transcripción viral; y genes tardíos L1 y L2, que codifican proteínas de la cápside vírica. El virus cuenta también con una larga región control (*LCR*, por sus siglas en inglés) que contiene promotores, sitios de iniciación de replicación del genoma viral y elementos de regulación transcripcional que actúan en *cis*. (Adaptada de: Castillo, 2014).

El genotipo VPH16 se detecta en alrededor de un 82% en pacientes con cáncer cervical invasivo, en donde, la expresión de sus oncogenes virales E6 y E7 son importantes para el mantenimiento del fenotipo tumorigénico y para el mantenimiento del fenotipo transformado (Nobel assembly, 2008).

La infección por VPH en células cervicales inicia en las células epiteliales basales a través de micro-abrasiones en el epitelio cervical, donde los genes tempranos de VPH (E1, E2, E4, E5, E6 y E7) son expresados por el genoma episomal del virus (Fig. 2). Los genes tardíos de las proteínas L1 y L2 se expresan en las capas epiteliales superiores generando la envoltura viral y de esta manera terminando el ciclo infeccioso al liberarse los viriones formados. La infección con algunos de los genotipos de VPH de alto riesgo como VPH16 puede inducir el progreso a la neoplasia intraepitelial cervical de alto grado. Esta progresión de las lesiones a cáncer invasivo y micro invasivo está asociada con la integración del genoma de VPH16 en los cromosomas del huésped (zur Hausen, 2002).

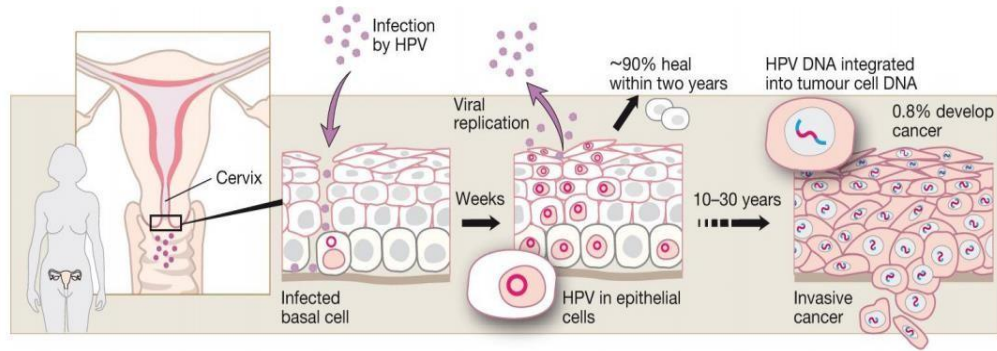


Figura 2. Ciclo de vida de VPH en infección cervical. La infección por VPH inicia por micro-abrasiones epiteliales que permiten la entrada del virus hasta la células madre basales. Algunas células se propagan lateralmente a medida que se dividen, y algunas células migran hacia las capas de células supra basales de diferenciación, en el que se activan los genes virales, se forman copias de ADN viral y las proteínas de la cápside. Las partículas de virus son liberados, difundiendo así la población de células infectadas o algunas pueden integrarse en el genoma humano, evento asociado a la progresión a cáncer en menos del 10% de las personas infectadas (Nobel assembly, 2008).

Dentro del ciclo infeccioso del VPH, genotipos malignos como el VHP16 pueden integrarse en el genoma de la célula huésped (*Fig. 3*) perdiéndose parte del genoma viral, particularmente el marco de lectura abierto de los genes E2, E4, E5 y parte de L2 (zur Hausen, 2002).

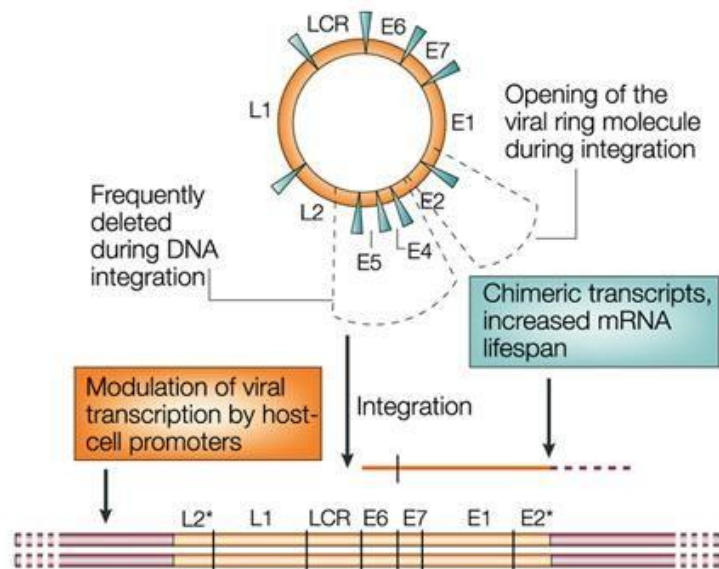


Figura 3. Evento de integración. La molécula circular se abre a menudo dentro del marco de lectura abierto de E2, interrumpiendo la continuidad de ese gen. Parte de E2 y los marcos de lectura abiertos que son adyacentes a E2 -E4, E5 y parte de L2 - se eliminan regularmente después de la integración (genes parciales están representados por un asterisco). Los transcritos virales que abarcan E6 y E7 están presentes y la transcripción puede ser modulada flanqueando los promotores de la célula huésped (zur Hausen, 2002).

Los genes virales E6 y E7 y sus respectivas proteínas poseen un rol muy importante en la transformación maligna. Estos genes se expresan en las células infectadas y son capaces de inmortalizar varios tipos de células humanas en cultivo de tejidos, presentándose una mayor eficiencia cuando se expresan juntas. Estudios sobre el papel oncogénico de estas proteínas han revelado que E6 interactúa con p53, y E7 interactúa con la proteína Rb inhibiendo la actividad de estos supresores de tumores, induciendo así la proliferación, inmortalización y transformación maligna de las células infectadas (Zaldívar-Lelo de Larrea et al., 2011).

En detalle, la degradación de p53 y la proteína proapoptótica BAK, por parte de la oncoproteína E6, resulta en la resistencia a la apoptosis y un aumento en la inestabilidad cromosómica. Además, se presenta la activación de la telomerasa y la inhibición de la degradación de las quinasas de la familia Src que parecen cumplir funciones importantes en la estimulación del crecimiento y podrían contribuir al fenotipo transformado por VPH. El inhibidor de quinasa dependiente de ciclina INK4A que codifica a la proteína p16 parece contrarrestar estas funciones. Por su parte, E7 interactúa y degrada a la proteína RB que también regula positivamente a INK4A, lo que libera el factor de transcripción E2F de la inhibición por Rb. La actividad E2F resultante podría conducir a la apoptosis en células que expresan E7. Por otra parte, E7 estimula la fase S de los genes ciclina A y ciclina E38, y parece bloquear la función de la ciclina dependiente de los inhibidores de quinasa WAF1 que codifica para las proteínas CIP1 y p21, y KIP1 que codifica la proteína p27 (Zaldívar-Lelo de Larrea et al., 2011; zur Hausen, 2002). Mediante la inducción de la amplificación de los centriolos, E7 también induce aneuploidía, lo que contribuye a la tumorigénesis. En la actualidad, es difícil asignar un papel para otras proteínas tempranas de VPH (tal como E1, E2 y E4) en el proceso de conversión maligna, mientras que las dos proteínas estructurales L1 y L2 no se expresan en células precancerosas y malignas, pero son importantes para el ciclo infectivo del virus, y en especial L1 ha sido utilizada para el desarrollo de vacunas profilácticas (zur Hausen, 2002).

En la actualidad se han aislado más de 120 tipos de virus pertenecientes a la familia *Papillomaviridae* (Prabhu & Wilson, 2013), cuya transmisión es principalmente sexual (Xu et al., 2013). Dos grandes grupos se han identificado según el tipo de tejido que infecten: cutáneos y de mucosa, siendo este último grupo el que predomina en el área genital. Igualmente según su patología se han dividido en VPH de alto y bajo riesgo. Dentro de los VPH de alto riesgo se han clasificado un total de 20 genotipos, siendo el VPH16 y VPH18 los más importantes ya que son asociados con el desarrollo del 70% de los cánceres cervicales y del 50% de cánceres orales y orofaríngeos. Por su parte, los VPH de bajo riesgo son causantes de lesiones benignas o verrugas genitales. Adicionalmente, los tipos de alto riesgo, se asocian a cáncer de vulva, vagina y ano en mujeres, y cáncer de ano y pene en hombres (Muños, 2003; zur Hausen, 2002; Walboomers, 1999; Schiffman, 2003) y, recientemente, se ha reportado la

presencia de VPH de alto riesgo oncológico en diferentes tipos de cáncer en cavidad oral (Castillo, 2006) y en sistema aero-digestivo (Castillo, 2011).

Recientemente otros factores posiblemente relacionados con el desarrollo de los cánceres asociados a la infección por VPH se han comenzado a estudiar, como son: una respuesta inmune ineficiente por parte del hospedero que impide eliminar al virus como también un aumento en la inestabilidad del genoma del huésped por los eventos de integración del VPH. Así, se ha postulado que la infección persistente por VPH16 y su carga viral puede conducir a la pérdida de estabilidad genómica en la célula huésped no solamente por la expresión de las oncoproteínas virales E6 y E7, sino también por la alteración de la expresión o función de genes del hospedero por eventos de inserción mutacional o a cambios de perfiles de metilación genómicos como consecuencia de los eventos de integración del genoma viral en el genoma huésped (Kadaja et al., 2007). Por ejemplo, las mujeres con infección persistente por VPH de alto riesgo tienen 300 veces más probabilidades de desarrollar una neoplasia de alto grado, así mismo los receptores de trasplantes de órganos y pacientes infectados por el VIH también tienen una mayor propensión a enfermedades malignas relacionadas con el VPH, especialmente el cáncer de cuello uterino entre las mujeres y el cáncer anal en los hombres, y tener cargas virales más altas e infecciones con genotipos de VPH inusuales (Nobel assembly, 2008).

En cuanto a la alteración de los genes celulares, en carcinoma de cabeza y cuello de células escamosas (HNSCC: por sus siglas en inglés) y carcinoma de cuello uterino con VPH positivo, se han detectado integraciones recurrentes en genes asociados con el ciclo celular como: *MYC*, *RAS*, *RAD51B*, *NR4A2*, y *TP63* además de *p53* y *p16INK4* (Kadaja et al., 2007).

Estudios realizados por Kumar et al 2015, han señalado que la metilación de los sitios ricos en dinucleótidos CG es un proceso epigenético importante en la modulación transcripcional del VPH (Kumar Gupta & Kumar, 2015) y se ha propuesto que la metilación del ADN es una defensa adaptativa en contra de virus y otros patógenos, ya que permite el silenciamiento de secuencias de ADN foráneas (Badal et al., 2003). De esta forma se puede encontrar metilación en genomas virales tras ser integrados dentro del genoma del hospedero (Larsson et al., 2014). Además, cambios en los perfiles de metilación del ADN de VPH han sido relacionados con procesos de tumorigénesis (Wentzensen et al., 2004). El entendimiento de las bases moleculares de la metilación de los sitios CG durante la transformación celular será muy importante para entender el desarrollo de la enfermedad (Kumar Gupta & Kumar, 2015).

Otros estudios han señalado que los eventos de integración del VPH ocurren dentro de regiones genómicas del huésped con secuencias repetitivas, específicamente en secuencias Alu (Carmody et al., 1996), (Vosa et al., 2012) y (Li et al., 2013), y en sitios hipersensibles a DNasa I (regiones de la cromatina sensibles a la escisión por la enzima ADNasa debido a la descondensación de la cromatina) en células epiteliales normales (Christiansen et al., 2015).

Métodos para la detección del VPH

La detección del VPH en células anormales es un importante factor de riesgo de progresión a cáncer (Pett & Coleman, 2007) y por tal razón, varios métodos han sido desarrollados para determinar la presencia viral, la carga viral y la integración del genoma viral con especial énfasis en la región del genoma del huésped donde este ocurre. Algunas técnicas son:

- **Hibridación *in situ* fluorescente (FISH, por sus siglas en inglés):** es una técnica útil para aplicar en los reordenamientos genómicos como la integración de los genomas virales en el genoma humano. En la técnica usa una sonda sintetizada a partir del ADN genómico viral, la cual se hibrida en los cromosomas en metafase de los pacientes. La observación de la señal fluorescente en la hibridación permite localizar los puntos de ruptura de VPH (integración) en los cromosomas. La prueba también permite estimar el número de copias del VPH (Hopman et al, 2004; Mincheva et al, 1987).
- **Amplificación de transcritos oncogenes de papilomavirus (APOT, por sus siglas en inglés):** es un ensayo que permite la diferenciación de los ARNm de VPH transcritos por los genomas de VPH integrado y episomal. La técnica amplifica transcritos de fusión de VPH-humano que pueden ser secuenciados para permitir la identificación de secuencias humanas adyacentes para el extremo 3' de los transcritos de oncogenes virales (Klaes et al., 1999).
- **PCR para sitios de restricción (RS-PCR):** este análisis utiliza una serie de cebadores específicos para VPH y para secuencias de reconocimiento de varias enzimas de restricción. La base de esta técnica es que los sitios de restricción son una ocurrencia común en el genoma humano y como tal, si uno de los sitios de restricción específicos se encuentra adyacente al ADN del VPH integrado, esto permitirá una detección más eficiente del VPH por PCR (Ragin et al., 2004).
- **E2 PCR:** la integración de VPH16 con frecuencia resulta en la interrupción del gen E2 (Baker et al., 1987). Un método simple de detectar la interrupción de E2 es el uso de PCR de solapamiento que cubren todo el gen E2 (Collins et al., 2009). Si E2 se interrumpe entonces una o más de las reacciones de PCR fallará. Es importante tener en cuenta que este método no detecta los eventos de integración en el que el gen E2 mantiene intacto o eventos de integración que han dado lugar a secuencias de VPH concatémico (Collins et al., 2009). El conocimiento de la interrupción de E2 es biológicamente relevante debido a su importante papel regulador (Collins et al., 2009, Baker et al., 1987).
- **Detección de secuencias de virus del papiloma integrados (DIPS, por sus siglas en inglés):** esta técnica se basa en el corte de ADN humano localizado en el extremo 5' con respecto al ADN de VPH usando endonucleasas de restricción seguido por la ligación a un adaptador de secuencia conocida en el extremo cohesivo que se forma. Las construcciones ligadas se amplifican primero en una forma lineal usando un cebador específico del VPH, y luego se amplifican de una manera

exponencial utilizando un cebador de VPH y un cebador específico de adaptador. Estos productos se separan entonces mediante electroforesis, las bandas son cortadas, purificadas y secuenciadas para identificar tanto el VPH y la secuencia humana de los amplificadas (Luft et al., 2001).

- PCR cuantitativa (Q-PCR):** en esta técnica se utilizan los mismos componentes que una PCR convencional solo que se agrega fluorescencia a cada reacción antes de la amplificación. La acumulación del producto de PCR se puede monitorear en tiempo real mientras se lleva a cabo la amplificación, y se analizan ciclo a ciclo los cambios de la señal fluorescente generada durante las tres etapas del PCR (Desnaturalización, alineamiento, y extensión). Entre mayor sea la cantidad de ADN de la muestra, menos serán los ciclos necesarios para obtener una señal fluorescente detectable. Esta amplificación por PCR tiempo real sigue una cinética exponencial en donde se pueden distinguir tres fases: geométrica, lineal y estacionaria (*Figura 4*). Actualmente los agentes reporteros más comunes utilizados para ensayos de tiempo real son: *SYBR Green*, un intercalante de bases de ADN, y sondas TaqMan, que son fragmentos de ADN específicos que tienen unido un fluoróforo y un quencher (represor o agente bloqueador) (ThermoFisher, 2014). De forma general las aplicaciones del PCR en Tiempo Real se dividen en dos grandes grupos: aplicaciones cuantitativas (Expresión génica, carga viral, y transgénicos) y aplicaciones cualitativas (análisis de polimorfismos, y análisis de presencia-ausencia de patógenos) (ThermoFisher, 2014).

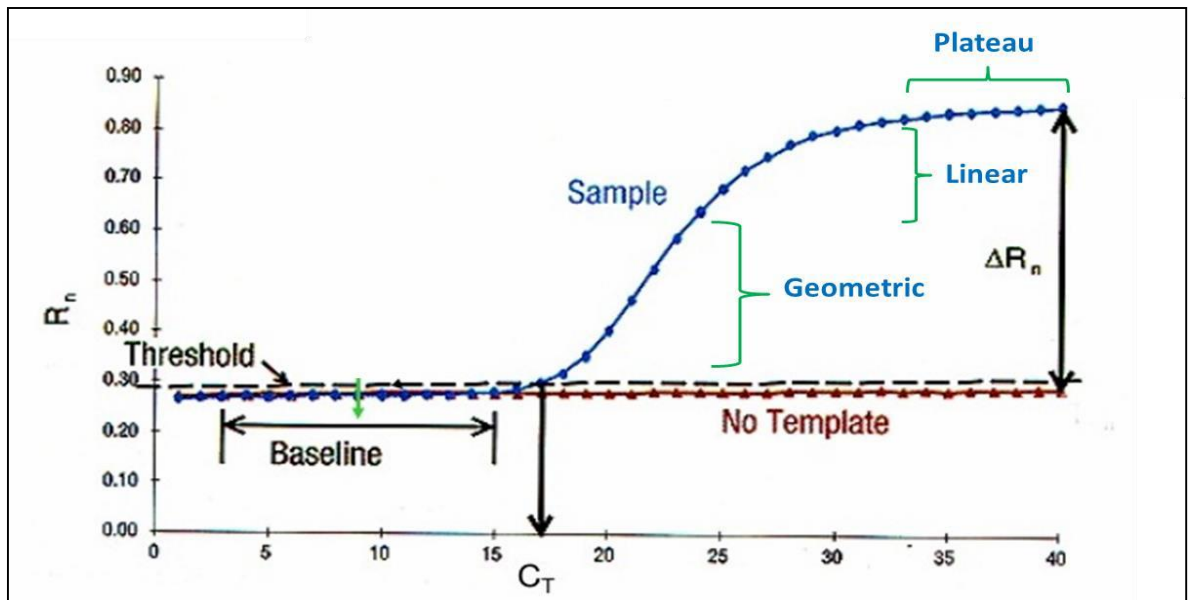


Figura 4. Gráfica de amplificación por PCR tiempo real. R_n : Magnitud de la señal. *Threshold* (Umbral de detección): Promedio de la desviación estándar de la fluorescencia. *Baseline* (Línea base): Los ciclos iniciales de la PCR en los cuales el cambio en la señal de fluorescencia es mínimo. C_T : Numero de ciclos en el cual la fluorescencia alcanza el *threshold* fijado. *No template* (Muestra sin molde): Muestra que no contiene templado. ΔR_n : Intensidad de la emisión de la fluorescencia (ThermoFisher, 2014).

2.3. Objetivos

Objetivo general

Determinar y analizar la integración del VPH16 en el genoma de pacientes diagnosticados con cáncer y su posible relación con la etiología de la enfermedad.

Objetivos específicos

1. Determinar la integración del VPH16 en el genoma de pacientes diagnosticados con cáncer por la técnica de PCR en tiempo real.
2. Analizar los sitios de integración de VPH16 y su posible relación con el desarrollo del cáncer, por medio del uso de herramientas bioinformáticas.

2.4. Metodología

2.4.1 Metodología utilizada en el laboratorio

2.4.1.1 Procedencia de las Muestras

Se analizaron 59 cortes de tejido de cáncer oral embebidas en bloques de parafina y fijadas en formalina del banco de muestras del proyecto titulado: “Prevalencia y caracterización genotípica del Virus del Papiloma Humano en alteraciones potencialmente malignas y cáncer oral en cinco regiones de Colombia. Estudio multicéntrico” que cuenta con la aprobación del Comité Institucional de Revisión de Ética Humana de la Facultad de Salud de la Universidad del Valle, acta 05-014 (*Anexo 1*). Las muestras corresponden a diferentes sitios anatómicos de la cavidad oral de un total de 37 pacientes (*Anexo 2*).

2.4.1.2 Extracción de ADN

Se realizó la extracción de ADN a partir de seis cortes (5 a 10 μm) de tejido embebidos en parafina. Se adicionó el reactivo Xileno y se centrifugó con el fin de eliminar la parafina. Se realizaron tres lavados con etanol para la eliminación de los residuos de Xileno. Se incubaron las muestras por 10 minutos a 37°C para eliminar por evaporación los residuos de etanol. Posteriormente se incubaron las muestras con proteinasa K a 56°C con el fin de lisar las proteínas. Se transfirieron las muestras a una columna de extracción para que por medio de centrifugación se pudiera recuperar el ADN. Este procedimiento de extracción se realizó según las recomendaciones y protocolos del kit QIAAMP DNA FFPE TISSUE de Qiagen.

2.4.1.3 Análisis de integridad

Se realizó un análisis de integridad del ADN de las muestras por espectrofotometría utilizando un equipo NanoDrop 2000 (*Thermo Scientific*). La pureza se determinó con la relación de DO 260/280 igual o cercana a 1,8. Posterior a su cuantificación las muestras de ADN fueron almacenadas a -20°C hasta su posterior uso.

2.4.1.4 Análisis de degradación

Se realizó una prueba de PCR punto final para la amplificación de un fragmento de 110 pb del gen de β -globina, para una reacción final de 25 μ l que contenía 1 μ l de ADN muestra. Adicionalmente, se amplificó como control positivo la muestra 7377A positiva para VPH16. La pareja de cebadores comerciales para β -globina es: PC03: 5`ACA CAA CTG TGT TCA CTA GC3`, y PC04: 5`CAA CTT CAT CCA CGT TCA CC3`. Las condiciones de amplificación fueron: desnaturalización del templado dsDNA a 95°C durante 5 minutos, alineamiento de los cebadores sobre el templado a 53°C durante 1 minuto, y extensión de la nueva cadena a 72°C durante 6 minutos. Posteriormente se realizó una corrida electroforética del ADN en gel de agarosa al 2% durante 30 minutos a 100 voltios. Se visualizaron las bandas por medio de un fotodocumentador adaptado a un transiluminador UV utilizando el bromuro de etidio como agente intercalante del ADN.

2.4.1.5 Detección del VPH16

La detección del ADN del VPH16 se realizó por la técnica de PCR en tiempo real mediante la amplificación del gen E6 viral utilizando la siguiente pareja de cebadores: E6F: 5'GAG AAA CTG CAA TGT TTC AGG ACC3' y E6R: 5'TGT ATA GTT GTT TGC AGC TCT GTG C3'. Igualmente, se amplificó el gen E2 viral utilizando la siguiente pareja de cebadores: E2F: 5'AAC GAA GTA TCC TCT CCT GAA ATT ATT AG3' y E2R: 5'CCA AGG CGA CGG CTT TG3'.

La amplificación de las muestras de ADN se realizaron en el equipo *7500 Fast Real-Time PCR System (Applied Biosystems)* del laboratorio de Biomedicina #3 de la Facultad de Medicina de la Universidad Icesi y las condiciones fueron: molécula fluorescente *SYBR Green kit PCR* (Qiagen, Hilden, Alemania), 590 μ l; cebadores 29.5 μ l c/u; agua milliQ 472 μ l; y 59 μ l de ADN. El volumen adecuado de cada reacción (20 μ l x 1) fue transferido a cada pocillo de la placa óptica. Después de sellar la placa y disponerlo en el equipo, se establecieron las siguientes condiciones de amplificación: 2 min a 50°C, 2 min a 95°C. Se realizaron 40 ciclos con 95°C durante 15s, y 56°C durante 15s y una extensión a 72°C por 1 min.

Adicionalmente se realizó un ensayo de disociación con temperaturas de Melt para el cual se agregaron las siguientes condiciones: 15s a 95°C, 1 min a 60°C, 30s a 95°C y 15s a 60°C. Las curvas de disociación con temperaturas de Melt representan un proceso de validación posterior, con la finalidad de corroborar que sólo un producto de amplificación se está generando en la reacción y que la fluorescencia registrada corresponde al producto en particular y no a productos inespecíficos o dímeros de primer; garantiza especificidad (ThermoFisher, 2014).

2.4.1.6 Cuantificación relativa del número de copias virales para E6 y E2

Se realizó por el método de cuantificación relativa Ct comparativo, determinando el número de copias virales de VPH en las muestras que resultaron positivas para la amplificación de E6 y E2 por PCR en tiempo real. El método consistió en la determinación del delta Ct del gen blanco viral E6 o E2, y posteriormente, su número de copias se normalización utilizando como control endógeno o de referencia al gen humano β -globina.

El equipo de PCR en tiempo real se programó para realizar una cuantificación relativa que es un método comparativo del Ct (ΔCt), mediante la siguiente fórmula: $2^{(-\Delta\Delta\text{Ct})}$, donde $\Delta\Delta\text{Ct} = \Delta\text{Ct blanco} - \Delta\text{Ct referencia}$. Para esto se utilizaron dos blancos experimentales (oncogenes virales E6 y E2) y blanco normalizador o control endógeno (β -globina). Los valores de Ct fueron obtenidos a partir de las curvas de amplificación de los genes E6, E2 y β -globina, respectivamente.

2.4.1.7 Determinación de la integración del genoma del VPH16

Para determinar el estado físico del genoma del VPH16 se utilizaron los números de copias virales para gen E6 y E2, calculados por PCR en tiempo real, bajo el supuesto de que el gen E2 es interrumpido en el evento de integración viral, y comparando las proporciones entre el gen E6 y E2, ver *Figura 5*.

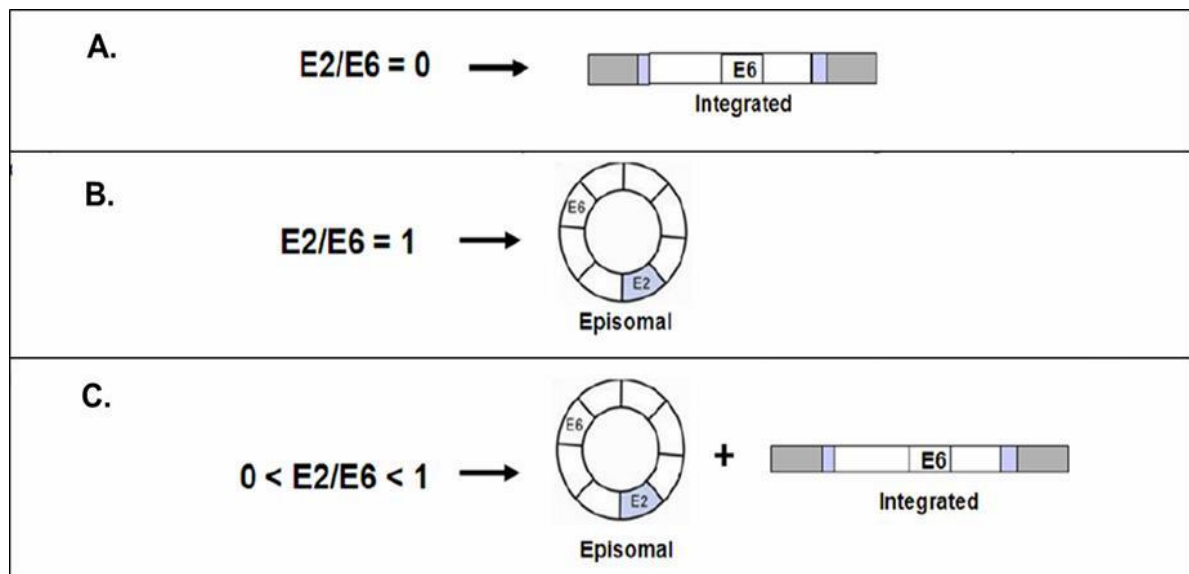


Figura 5. Determinación del estado físico de VPH por el ensayo de PCR en tiempo real. (A) ADN viral integrado: la proporción E2/E6 es igual a cero porque el gen E2 es interrumpido como resultado de la integración y solo el número de copias del gen E6 puede ser determinado. (B) ADN viral episomal: la proporción E2/E6 es igual a 1 porque los genes E2 y E6 tienen equivalente número de copias. (C) ADN viral mixto: la proporción E2/E6 es mayor a cero y menor a 1 porque el gen E2 tiene menor número de copias que el gen E6 (Adaptada de: Castillo 2014).

2.4.2 Metodología *In silico*

La *Figura 6* describe la metodología general llevada a cabo *In silico*.

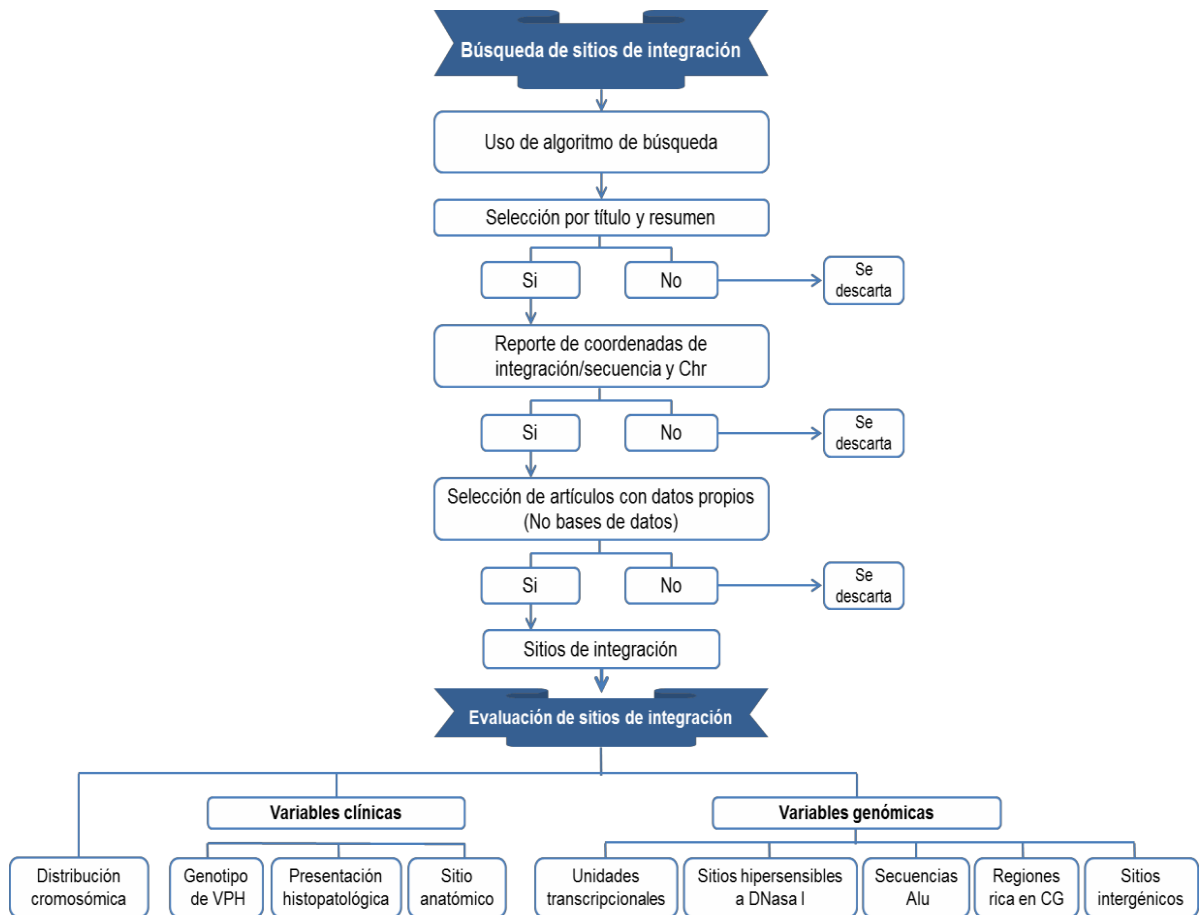


Figura 6. Flujograma de metodología *In silico*. La metodología *In silico* consta de dos componentes: el primero es la búsqueda y obtención de los sitios de integración reportados para diferentes genotipos de VPH, y el segundo es la evaluación de los sitios de integración de forma descriptiva (Distribución cromosómica y variables clínicas) y por medio de la herramienta bioinformática *Genome Browser* (Variables genómicas).

2.4.2.1 Búsqueda de sitios de integración del VPH

Se seleccionó la base de datos *PubMed* de *NCBI* para la búsqueda de artículos con reporte de sitios de integración para el VPH usando el siguiente algoritmo “**((HPV) OR human papillomavirus) AND ((cancer) OR carcinoma) AND ((integration) OR breakpoint)**” construido a partir de palabras clave en el idioma inglés y confirmadas como descriptores de la salud en *MeSH* de *NCBI* (<http://www.ncbi.nlm.nih.gov/mesh>).

Las referencias fueron almacenadas en un solo archivo del programa *EndNoteX6*. El algoritmo se puso a prueba en Febrero 20 del 2016. A continuación, fueron seleccionados los artículos por lectura del título y el resumen. Se incluyeron los estudios de infección por VPH en humanos, sin distinción de edad o sexo, y se excluyeron los estudios que hacían referencia a otros organismos, vacunas, solo expresión de genes, y estudios no experimentales.

Además, se descartaron los artículos que no reportaran los sitios de integración viral, ya sea como coordenadas cromosómicas o secuencias nucleotídicas adyacentes a la integración viral.

Las coordenadas de integración se presentan de la forma “*chr (#cromosoma): #base nucleotídica*” y representan el número nucleotídico donde ocurrió la integración viral en determinado cromosoma humano. Las secuencias nucleotídicas debían contar con mínimo 20 nucleótidos para poder ser evaluadas por la herramienta bioinformática *Genome Browser* (<https://genome.ucsc.edu/>). Otro criterio de inclusión se enfocó en que los artículos reportaran datos propios. Se seleccionaron artículos con reporte de sitios de integración pertenecientes a sus estudios y no bases de datos o recopilación de otras investigaciones para excluir redundancia de datos.

2.4.2.2 Evaluación de sitios de integración del VPH

En el caso donde el sitio de integración viral fue reportado como coordenadas cromosómicas se utilizó las herramienta bioinformáticas de *Genome Browser* para la evaluación de las variables. En cuanto al reporte del sitio de integración en forma de secuencias nucleotídicas se realizó un alineamiento de bases entre la secuencia dada y el genoma de referencia humano GRCh38-2013 con la herramienta *BLAT* del *Genome Browser* para obtener la coordenada cromosómica y posterior evaluación. La evaluación de los sitios de integración en el genoma humano consistió en la revisión de la distribución cromosómica de los sitios, la identificación de 3 variables clínicas y la caracterización molecular de los nucleótidos exactos del sitio de integración del VPH en el genoma humano.

Distribución cromosómica de los sitios de integración

Para la distribución de los sitios de integración por cromosoma se tuvo en cuenta, además de la frecuencia de sitios de integración por cromosoma, la proporción de sitios de integración por tamaño cromosómico en pares de bases para discriminar si la frecuencia de sitios de integración se debía a la mayor exposición de pares de bases a la integración o la integración era independiente del tamaño del cromosoma. El estudio incluyó los 22 cromosomas autosómicos y 2 cromosomas sexuales. La fórmula utilizada fue:

$$\text{Proporción} = \% \text{Sitios de integración} / \% \text{Pares de bases}$$

Donde %Sitios de integración es el número de sitios de integración de cada cromosoma sobre el total de sitios de integración en todos los cromosomas, y %Pares de bases es el número de pares de bases de cada cromosoma sobre el total de pares de bases de todos los cromosomas.

Variables clínicas

- *Genotipo de VPH*: Los sitios de integración seleccionados comprenden infección por genotipos de VPH de alto riesgo.
- *Presentación histopatológica*: Hace referencia a tejidos con algún estadio tumoral benigno o maligno.

- *Sitio anatómico*: Comprende sitios anatómicos asociados a cáncer de cabeza y cuello, cáncer cervico-uterino y otros tipos de cánceres anogenitales.

Variables genómicas

- *Unidades transcripcionales*
Se utilizó la herramienta bioinformática *Gencode v22- Genome Browser* para identificar los genes presentes en las posiciones donde hay integración de VPH. Los genes reportados están contenidos en *RefSeq* una completa, integrada y no redundante base de datos de secuencias de referencia del Centro Nacional de Información Biotecnológica (*NCBI*). Adicionalmente, se identificó si la región del gen en las determinadas posiciones eran intrones o exones.
- *Sitios de hipersensibilidad a DNasa I*
Se utilizó la herramienta bioinformática *DNase I Clusters- Genome Browser* que muestra las agrupaciones de hipersensibilidad a DNasa I derivadas de ensayos experimentales.
- *Secuencias Alu*
Se utilizó la herramienta bioinformática *Repeat Masker-Genome Browser* encargada de generar una anotación detallada de las repeticiones presentes en la secuencia de consulta, entre ellos los elementos cortos intercalados nucleares (SINE) que incluyen Alu.
- *Regiones ricas en CG*
Se utilizó la herramienta bioinformática *CpG Islands- Genome Browser* para identificar las regiones en donde el dinucleótido CG está presente en niveles significativamente más altos de lo que es típico para el genoma en su conjunto.
- *Sitios intergénicos*
Se utilizó la herramienta bioinformática *Gencode v22- Genome Browser*.

2.4.2.3 Análisis de variables

Se realizó un análisis descriptivo de los datos obtenidos de cada una de las variables con relación al genotipo de VPH.

2.4.2.4 Presentación de resultados

La distribución de los sitios de integración por cromosoma se presentó en diagramas de barras, donde el eje X representa los cromosomas y el eje Y la proporción de sitios de integración por tamaño cromosómico. Los diagramas de barras incluyeron la distribución para cada genotipo de VPH estudiado. Los datos de las variables clínicas y genómicas fueron tabulados en tablas resumen con el

número de la variable, nombre de la variable, y frecuencias de sitios de integración con sus respectivos porcentajes, para cada genotipo.

2.4.3 Matriz de marco lógico

OBJETIVO GENERAL			
Determinar y analizar la integración del VPH16 en el genoma de pacientes diagnosticados con cáncer y su posible relación con la etiología de la enfermedad.			
OBJETIVOS ESPECÍFICOS	ACTIVIDADES	INDICADORES	SUPUESTOS
Determinar la integración del VPH16 en el genoma de pacientes diagnosticados con cáncer por la técnica de PCR en tiempo real.	Extracción de ADN Análisis de pureza Análisis de degradación Detección del VPH Cuantificación relativa	Valores de NanoDrop Amplificación de β -globina Amplificación de E6 y/o E2 Ct comparativo	Buenas relaciones de pureza Amplificación de β -globina Amplificación de E6 y/o E2 Obtención de cargas virales
Analizar los sitios de integración de VPH16 y su posible relación con el desarrollo del cáncer, por medio del uso de herramientas bioinformáticas.	Obtener por medio de búsqueda bibliográfica los sitios de integración del VPH. Realizar la distribución cromosómica de los sitios de integración. Identificar diferentes variables clínicas. Evaluar los sitios de integración en un programa bioinformático para determinar la presencia de las variables genómicas.	Sitios de integración del VPH con sus respectivas coordenadas cromosómicas. Tablas de distribución de sitios de integración por cromosoma, variables clínicas y variables genómicas. Gráficas de apoyo.	Disponer de las coordenadas cromosómicas o secuencias nucleotídicas de los sitios de integración. Disponer de datos clínicos para todos los sitios de integración obtenidos. Encontrar presencia de las diferentes variables genómicas en los sitios de integración.

2.5. Resultados

2.5.1 Resultados de laboratorio

Las 59 muestras evaluadas por espectrofotometría mostraron una relación de pureza 260/230 entre 1,7 y 1,8. Por medio de PCR en punto final para β -globina se determinó si las muestras podrían ser amplificadas (*Fig. 7*).

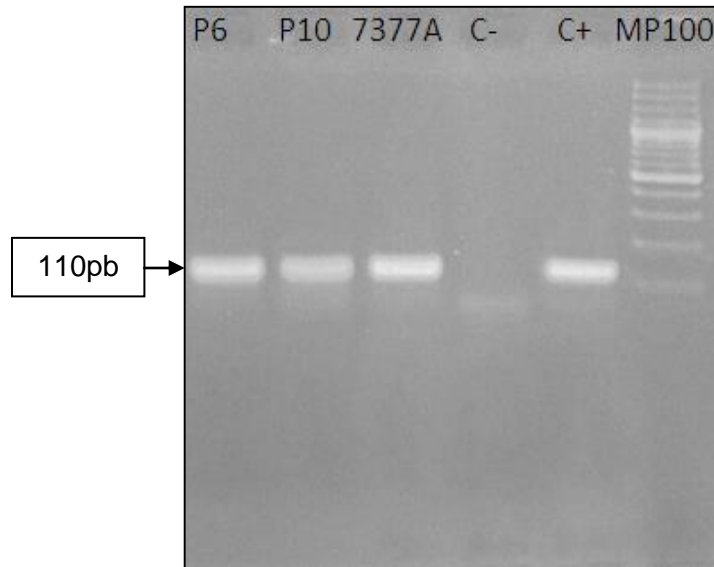


Figura 7. Resultado de electroforesis para amplificación de muestras de ADN con el uso de cebadores PCO3/PCO4 para la amplificación de β -globina. P6 y P10 representan muestras de ADN que muestran la amplificación de la banda de 110 pb de β -globina. 7377A representa un control positivo para VPH16. El marcador de peso es de 100pb.

De las 59 muestras evaluadas no fue posible la amplificación del gen de β -globina para 3 muestras. Esas muestras disminuyeron el número de pacientes con cáncer oral escamocelular de 37 a 35.

Por medio de PCR en tiempo real se determinó que de las 56 muestras de ADN que pasaron el filtro técnico de amplificación 5 fueron positivas para el oncogén E6 del VPH 16. Un ejemplo de las curvas de amplificación y curvas de Melt positivas para E6 y β -globina resultantes se puede observar en la *Figura 8*.

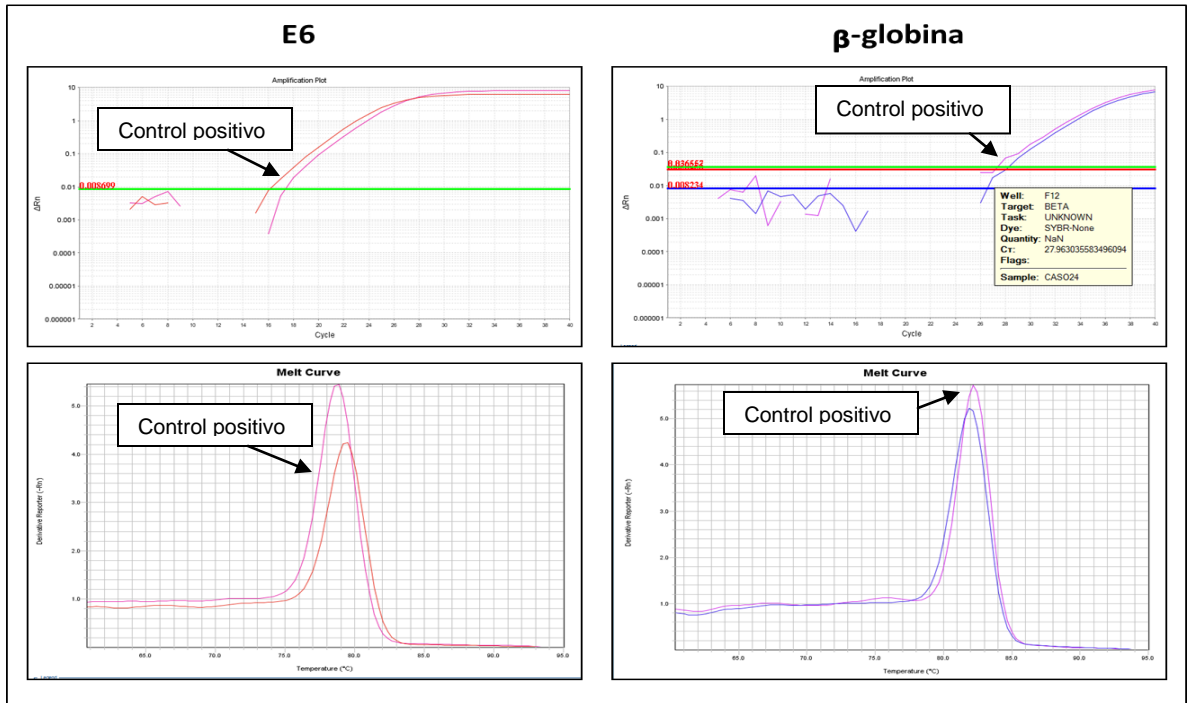


Figura 8. Caso de curvas de amplificación y curvas de Melt positivas para E6 y β-globina. Las gráficas superiores representan las curvas de amplificación para E6 y β-globina; el eje Y representa la fluorescencia (log) y el eje X el número de ciclos de la PCR. Las gráficas inferiores corresponden a las curvas de Melt que confirman la amplificación positiva, donde el pico de la curva corresponde a la temperatura de disociación del producto de amplificación; el eje Y representa la fluorescencia y el eje X la temperatura (ThermoFisher, 2014).

La *Tabla 1* resume los resultados de detección del oncogén E6 del VPH16 con descripción del género de los pacientes a los que pertenecían las muestras. Solo en las mujeres se encontró presencia del VPH16 en un porcentaje del 26%. La *Tabla 2* describe los diferentes sitios anatómicos correspondientes a las muestras de ADN de pacientes con cáncer oral escamoso celular con amplificación del oncogén E6. De los diferentes sitios anatómicos dos muestras amplificaron desde ADN de labio, dos desde paladar y una desde el piso de la boca.

Tabla 1. Descripción de la población del estudio por género y edad.

CARACTERISTICA	VPH 16	
	POSITIVO (%)	NEGATIVO (%)
Genero		
Femenino (19)	5 (26)	14 (74)
Masculino (16)	0	16 (100)
Total (35)	5 (14)	30 (86)
Promedio de edad		
Mujeres	64 ± 9	62 ± 9

Tabla 2. Sitios anatómicos con eventos de integración del VPH16. Se representa el porcentaje por sitio anatómico con amplificación del oncogén E6.

LOCALIZACIÓN (n)	VPH 16	
	POSITIVO (%)	NEGATIVO (%)
Lengua (17)	0	17 (57)
Labio (4)	2 (40)	2 (7)
Maxilar (1)	0	1 (3)
Mucosa yugal (2)	0	2 (7)
Paladar (4)	2 (40)	2 (7)
Piso de la boca (6)	1 (20)	5 (17)
Mandíbula (1)	0	1 (3)
TOTAL (35)	5 (100)	30 (100)

Se evaluó también la presencia del oncogén E2 en las 5 muestras positivas para el oncogén E6 de VPH 16 (Tabla 3).

Tabla 3. Amplificación del oncogén E2 del VPH16. El CT es el número de ciclos requeridos para que la señal de amplificación alcance el *threshold*. La PCR en tiempo real relativa requiere que los genes se evalúen al mismo tiempo, esto incluye al gen β -globina que es el gen de referencia con el cual se relativiza la expresión génica de E6 y E2.

No. Caso	CT		
	Beta-globina	E6	E2
4	31,34	33,95	No detectado
6	29,39	34,53	No detectado
9	35,99	36,18	No detectado
37	26,99	36,86	No detectado
41	27,88	17,44	24,97

Por último se cuantificó el número de copias de los genes E6 y E2 presentes en las 5 muestras con amplificación del oncogén E6 y en la única muestra que amplificó el gen E2, respectivamente, por medio de la cuantificación relativa por el método de Ct comparativo o $\Delta\Delta Ct$ (Tabla 4 y 5). En las tablas 3,4 y 5, se observa que solo una de las muestras evaluadas tiene señal para el gen E2 del VPH16, por lo tanto bajo el supuesto de la relación de amplificación entre los genes E2 y E6 en la cuantificación relativa, la muestra 41 presentó un estado físico mixto ($E2/E6 = 2,93/24,42 = 0,12$), es decir, presentaba tanto copias virales episomales como integradas. Las demás muestras con una relación igual a cero presentaban copias virales totalmente integradas en el ADN humano.

Tabla 4. Cuantificación del número de copias del oncogén E6 en muestras positivas para VPH16. Los valores de carga viral son representados por los valores de cuantificación relativa ($2^{(-\Delta\Delta Ct)}$). La muestra número 41 contiene el mayor número de copias de VPH 16 integrado.

No. Caso	ΔCt	$\Delta\Delta Ct$	No. copias virales
Ref-7377A	-5,83	0,00	1,00000
4	2,61	8,44	0,00288
6	5,14	10,97	0,00050
9	0,19	6,02	0,01541
37	9,87	15,70	0,00002
41	-10,44	-4,61	24,42015

Tabla 5. Cuantificación del número de copias del gen E2 en muestras positivas para VPH16. Los valores de carga viral son representados por los valores de cuantificación relativa ($2^{(-\Delta\Delta Ct)}$). La muestra número 41 contiene el mayor número de copias de VPH 16 episomal.

No. Caso	ΔCt	$\Delta\Delta Ct$	No. copias virales
Ref-7377A	-1,36	0,00	1,00000
41	-2,91	-1,55	2,92817

2.5.2 Resultados del estudio *in silico*

2.5.2.1 Búsqueda de sitios de integración

Se obtuvieron 16 artículos de los cuales 11 presentaron los sitios de integración en coordenadas cromosómicas y 5 en forma de secuencias, reportando en total 667 sitios de integración del VPH.

Del total de 667 sitios de integración cuatro fueron eliminados por presentarse como secuencias de ADN menores a 20 nucleótidos; 39 se eliminaron por ubicación errónea es decir, discordancia entre el cromosoma reportado en el artículo y el cromosoma arrojado por *Genome Browser*; y 13 se eliminaron por redundancia, resultando un total final de 611 sitios de integración viral por VPH.

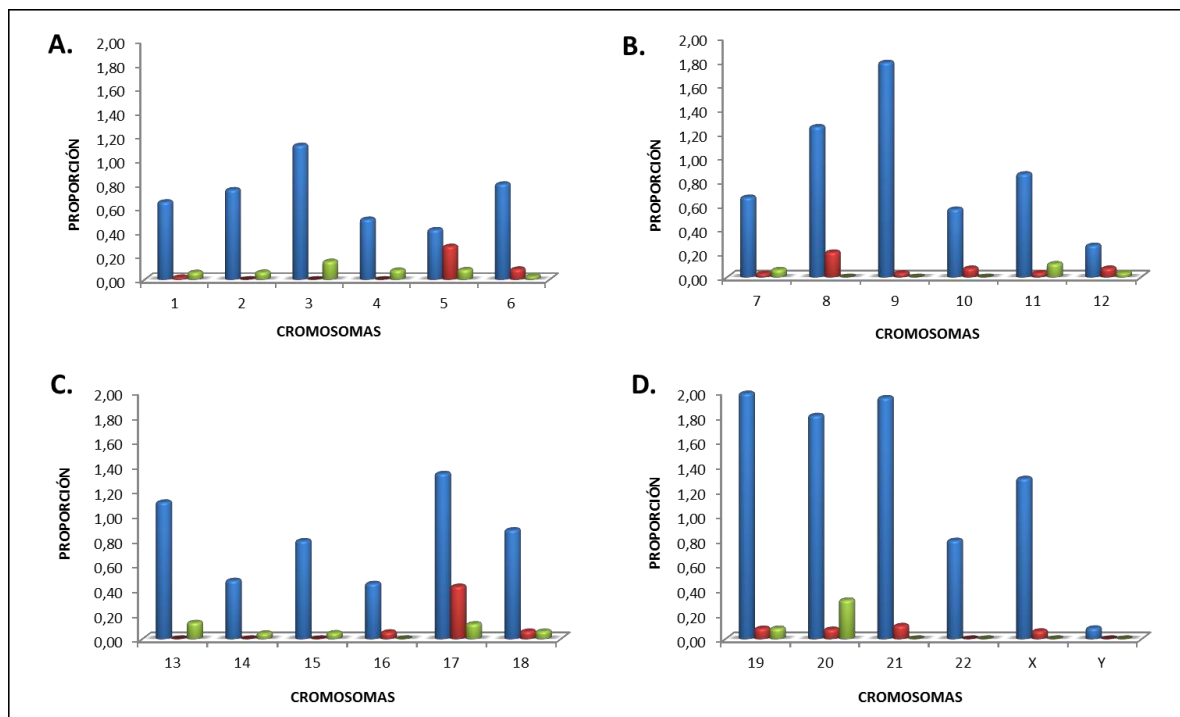
De esta forma, se evaluaron 16 artículos con reporte de 611 eventos de integración del VPH (Anexo 3), obteniendo como resultado información cualitativa asociada al VPH y el hospedero.

2.5.2.2 Distribución cromosómica de los sitios de integración

Se encontró el total de 611 sitios de integración distribuidos en los 22 cromosomas autosómicos y los 2 cromosomas sexuales, de los cuales 532 correspondían a VPH16. En el estudio se tuvo en cuenta la proporción de sitios de integración por tamaño de cromosoma, representada en la *Gráfica 1* para los VPH16, VPH18 y otros VPHs.

Para el VPH16 se obtuvo que el cromosoma 19 con 23 sitios de integración, es el cromosoma con mayor proporción de sitios de integración por tamaño de cromosoma, mientras que el cromosoma Y, con sólo 1 sitio de integración, es el cromosoma con menor proporción de sitios de integración.

La distribución de sitios de integración por tamaño cromosómico para el VPH18 y otros genotipos de VPH indica solo tendencias en la proporción para los cromosomas 5, 17 y 20.



Gráfica 1. Distribución cromosómica de los sitios de integración para VPH 16, VPH 18 y otros genotipos. Los resultados incluyen las 3 categorías de genotipos evaluados: VPH 16 (Azul), VPH 18 (Rojo) y otros genotipos o VPH Otros (Verde). La distribución de los sitios de integración se muestra en 4 grupos de a 6 cromosomas (A-D) para un total de 24 cromosomas humanos. El eje Y representa la proporción de sitios de integración por cromosoma (Proporción = %Sitios de integración / % Pares de bases, donde: % Sitios de integración es el número de sitios de integración de cada cromosoma sobre el total de sitios de integración en todos los cromosomas, y % Pares de bases es el número de pares de bases de cada cromosoma sobre el total de pares de bases de todos los cromosomas) y el eje X los cromosomas correspondientes. Las 4 gráficas mantienen una escala de 0 a 2 en la proporción para hacerlas comparables.

2.5.2.3 Genotipos de VPH

Se obtuvo un total de 7 genotipos de VPH asociados con desarrollo oncogénico: 16, 18, 38, 58, 33, 36, y 26. Para el VPH16 se encontraron 532 sitios de integración (87%), VPH 18 con 41 sitios de integración (7%) y Otros VPHs 38 sitios de integración (6%).

2.5.2.4 Presentación histopatológica

Se obtuvo un total de 6 presentaciones histopatológicas: una de carácter tumoral (Neoplasia intraepitelial cervical escamosa); un tumor maligno de epitelio glandular (adenocarcinoma); tres tipos de carcinoma (carcinoma mixto adenoescamoso, carcinoma de células escamosas, y carcinoma endometrial del cuerpo uterino); y una presentación de células sin crecimiento descontrolado (queratinocitos cervicales).

En la *Tabla 6* se agruparon las presentaciones histopatológicas en: carcinoma de células escamosas por su alta frecuencia de sitios de integración; queratinocitos cervicales por su estado ni canceroso ni tumoral; neoplasia intraepitelial cervical escamosa como estado no maligno; y otras.

Tabla 6. Categorías agrupadas por tipos histopatológica

#	PRESENTACIÓN HISTOPATOLÓGICA	SITIOS DE INTEGRACIÓN n(%)	VPH 16 n(%)	VPH 18 n(%)	VPH Otros n(%)
1	Carcinoma de células escamosas	541 (89)	483(89)	39(7)	19(4)
2	Queratinocitos cervicales	26 (4)	26(100)	0	0
3	Neoplasia intraepitelial cervical escamosa	22 (4)	22(100)	0	0
4	Otras	22 (3)	1(5)	2(9)	19(86)
TOTAL		611 (100)			

El carcinoma de células escamosas fue la presentación histopatológica con el mayor porcentaje de sitios de integración, tanto para el VPH16 como para el VPH18. Los queratinocitos cervicales y las neoplasias intraepiteliales solo presentan sitios de integración para el VPH16 en comparación con los demás genotipos de VPH. El 50% de sitios de integración presentados para la categoría Otros está representado principalmente por el carcinoma endometrial del cuerpo uterino y el VPH38.

2.5.2.5 Sitio anatómico

Se obtuvo un total de 12 sitios anatómicos (cérvix, amígdala, orofaringe, cavidad oral, útero, ano, hipofaringe, lengua, base de la lengua, laringe, base bucal, fosa amigdalina) que fueron agrupadas de acuerdo a su cercanía anatómica (*Tabla 7*).

Tabla 7. Categorías agrupadas por sitio anatómico.

# SITIO ANATÓMICO	SITIOS DE INTEGRACIÓN n(%)	VPH 16 n(%)	VPH 18 n(%)	VPH Otros n(%)
1 Cervicouterino	374 (61)	325 (87)	21 (6)	28 (7)
2 Orofaringeo	162 (26)	160 (99)	0	2 (1)
3 Oral	59 (10)	31 (52)	20 (34)	8 (14)
4 Ano	11 (2)	11 (100)	0	0
5 Laringe	5 (1)	5 (100)	0	0
TOTAL	611 (100)			

Los sitios anatómicos cervicouterinos representaron el mayor porcentaje de sitios de integración para todos los VPH hallados en el estudio. Los sitios anatómicos pertenecientes a la categoría Orofaringeo representan el segundo mayor porcentaje para el VPH16. Para el VPH18 y otros genotipos VPH se observó una tendencia por los sitios anatómicos cervicouterinos seguidos por los orales. Los sitios anatómicos Ano y Laringe solo se encuentran en el VPH16.

2.5.2.6 Variables genómicas

Las variables genómicas fueron evaluadas en los sitios de integración del VPH16. La *Tabla 8* representa el resumen de la presencia de variables genómicas.

Tabla 8. Variables genómicas presentes en los sitios de integración del VPH 16. Punto de ruptura = sitios de integración del VPH.

#	VARIABLES GENÓMICAS	PUNTO DE RUPTURA n(%)
1	Unidades transcripcionales	268 (50)
2	Sitios de hipersensibilidad a DNasa I	93 (18)
3	Secuencias Alu	53 (10)
4	Regiones ricas en CG	5 (1)
5	Sitios intergénicos	113 (21)
	TOTAL	532 (100)

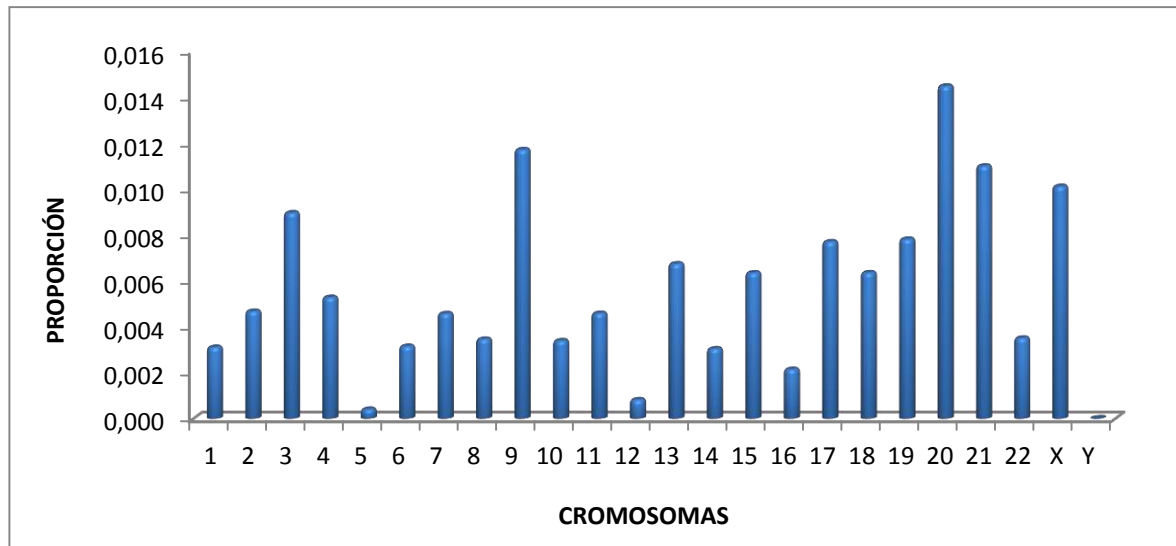
La variable Unidades transcripcionales fue la variable genómica con mayor porcentaje de sitios de integración del VPH16 (50%) en donde los 268 sitios se dividieron en 21 eventos de integración en regiones exónicas (correspondientes a 19 genes) (*Anexo 4*) y 247 eventos de integración en regiones intrónicas (correspondientes a 163 genes). De los 182 genes presentes en los sitios de integración viral, 4 presentaban eventos de integración tanto en regiones exónicas como en regiones intrónicas, por lo tanto, se presentó un total de 178 genes involucrados en los eventos de integración viral; además, de los 178 genes, tres

fueron genes supresores de tumores (TP63, PRBM1 y FHIT) y cuatro fueron oncogenes (ERBB4, MECOM, MLLT1y USP4).

Los sitios intergénicos y sitios de hipersensibilidad a DNasa I estuvieron presentes en aproximadamente el 20% de los sitios de integración. Las secuencias Alu se presentaron en el 10% de los sitios de integración y las regiones ricas en CG solo estuvieron presentes en solo cinco (1%) de los sitios de integración incluidos en el estudio.

Para estimar el impacto del evento de integración en regiones codificantes del genoma, se evaluó la proporción de sitios de integración del VPH16 en genes sobre el total de genes presentes en cada cromosoma (*Gráfica 2, Anexo 5*). El número de genes totales por cromosoma fue obtenido de *NCBI Map Viewer*. El siguiente es un ejemplo de la fórmula empleada para el cálculo de la proporción para el cromosoma 1 (Cr.1):

$$\frac{(\# \text{ Sitios de integración del VPH16 en genes del Cr.1})}{(\# \text{ Genes del Cr.1})} = \frac{(15)}{(4778)} = 0.003$$



Gráfica 2. Proporción de sitios de integración del VPH16 en genes sobre el total de genes por cromosoma. El eje Y representa la proporción de sitios de integración del VPH16 en genes sobre el total de genes presentes en cada cromosoma (# sitios de integración en genes/ # Genes totales en el cromosoma) y el eje X representa los cromosomas correspondientes. El cromosoma 20 y 9 presentan la mayor proporción, seguidos por los cromosomas 21, X, y 3.

2.6. Discusión

En el presente estudio se logró la determinación del estado físico de integración del genoma del VPH16 por medio de la técnica de PCR en tiempo real. Adicionalmente, se analizaron las regiones del genoma del huésped que se ven afectadas por el evento de integración viral. Lo anterior se realizó con el objetivo de entender cómo el evento de integración del VPH16 puede por sí solo, tener un papel importante en el desarrollo del cáncer.

El uso de la técnica de PCR en tiempo real nos permitió detectar y determinar la integración del genoma del VPH16 en muestras de tejidos embebidos en parafina obtenidas de pacientes diagnosticados con cáncer oral. Las muestras fueron fijadas con formalina lo que genera la fragmentación del ADN, por lo cual es necesario el uso de técnicas con una alta sensibilidad y especificidad. Además, la técnica de PCR en tiempo real presenta una mayor ventaja si se compara frente al uso de la PCR convencional ya que no es necesario un análisis post-PCR como la electroforesis en gel de agarosa.

Para la detección se utilizó la molécula reportero *SYBR Green* que permitió realizar un ensayo de disociación por temperaturas de Melt como método de validación (ThermoFisher, 2014), y resulta más económico que el método por sondas TaqMan.

Con relación al porcentaje de detección y determinación del estado de integración del genoma del VPH bajo la premisa de la pérdida parcial del gen viral E2 en el evento de integración, se encontró que 5 de 35 (14%) muestras de pacientes con cáncer oral escamoso celular tenían presencia del VPH16 en estado integrado, y en uno de ellos genomas virales episomales. Este resultado concuerda con lo presentado en el 2005 por la Agencia Internacional para la Investigación del Cáncer, en un análisis sistemático de estudios alrededor del mundo, basados en PCR, donde se reportó una presencia del VPH16 en un 16% de 2642 casos estudiados de cáncer de la cavidad oral (Kreimer et al., 2005).

En cuanto a la identificación de la carga viral, se obtuvo que la muestra # 41 (Tabla 4 y 5) con un estado físico mixto para el VPH16 fue la muestra con un mayor número de copias, 25 veces por encima de las demás muestras, resultado que concuerda con algunos estudios en cáncer cervical asociado a VPH que documentan un mayor número de copias virales en estado episomal (Shukla et al., 2014). Algunos investigadores han sugerido que pacientes afectados por la infección del VPH con cargas virales bajas son más propensas a eliminar la infección por parte de su sistema inmune que los que presentan mayor carga viral (Daponte et al., 2008). Sin embargo, los umbrales de carga viral clínicos deben ser

definidos y los métodos utilizados deben ser estandarizados y validados en poblaciones más grandes.

Por su parte, en el estudio *in silico* de las regiones genómicas humanas afectadas por los eventos de integración del VPH16, se obtuvo como resultado una recopilación de 532 sitios de integración para este genotipo con una alta proporción en carcinomas de células escamosas. Similar hallazgo se observó para el VPH18, sin embargo, al analizar sitios de integración reportados para neoplasias intraepiteliales cervicales y queratinocitos cervicales solo se encontró reportado para el VPH16.

Con relación a las demás presentaciones histopatológicas, el único sitio de integración para carcinoma adenoescamoso pertenecía a infección por VPH16, los dos sitios de integración para adenocarcinomas correspondían a infección por VPH18, y los restantes 19 sitios de integración de carcinoma endometrial del cuerpo uterino pertenecían al VPH38. Kazemian et al., 2015 reportaron que 32 de 168 muestras con cáncer de endometrio presentaron integración de VPH38 aunque con algunos problemas en metodología (Kazemian et al., 2015) lo que hace necesario más estudios en este tipo de cáncer.

Para los sitios anatómicos cervico-uterinos, orofaringe, ano y laringe con los eventos de integración del VPH se encontró una mayor proporción de eventos de integración por parte del VPH16. Cabe destacar que aunque la presencia de VPH18 en cavidad oral de los reportes revisados fue del 34%, hay resultados particulares, como el reportado por Giovannelli et al en el 2002, donde el VPH18 fue el genotipo más frecuente (86%) detectado en todas las muestras que presentaban lesiones orales, mientras que otros tipos de VPH (VPH 16 ,33 y 31) fueron observados en bajos porcentajes (Giovannelli et al., 2002), por lo tanto podría tenerse en cuenta su evaluación en próximos estudios en cavidad oral.

Otra de las relaciones estudiadas entre el VPH16 y su hospedero fue la distribución de los sitios de integración en el total de los 24 cromosomas humanos. Al determinar la proporción de sitios de integración por tamaño cromosómico, se encontró una mayor proporción de sitios de integración por pares de bases en los cromosomas pequeños 19, 20 y 21, seguidos por el cromosoma 9, lo que podría sugerir que la integración no es un evento aleatorio como lo indican algunos autores (Wentzensen et al., 2004), (Vojtechova et al., 2015). Por su parte, en los cromosomas sexuales X y Y se observó una marcada diferencia donde el cromosoma Y sólo contó con 1 sitio de integración del VPH16.

Para explicar la preferencia de los eventos de integración viral en ciertos cromosomas, se identificó, para cada uno, el número de sitios de integración en unidades transcripcionales. Así, nuestros resultados sugieren que los cromosomas 9, X, 20 y 21 presentan una alta frecuencia de regiones génicas afectadas, acorde al número de genes totales por cromosoma.

Los anteriores resultados muestran que los cromosomas 9, 20 y 21 tienen una posible relevancia en el evento de integración ya que presentan las mayores proporciones de sitios de integración de acuerdo a su tamaño y número de genes totales.

Un resultado importante en el análisis de las variables genómicas fue que los eventos de integración del VPH16 ocurren con una mayor tendencia en regiones codificantes o unidades transcripcionales. En las regiones codificantes, los eventos de integración viral ocurrieron en una mayor proporción en regiones intrónicas que en exónicas, nuestros resultados concuerdan con lo reportado por Christiansen et al. 2015. En total 178 genes codificantes fueron afectados por el evento de integración del VPH16. Entre estos genes cabe resaltar la presencia de DIAPH2 que puede sufrir supresión homocigótica tras la integración promoviendo la inestabilidad cromosómica a través de la desalineación de las cromátidas hermanas durante la metafase (Cheng et al. 2011); TP63, un regulador crítico de la diferenciación epitelial relacionado con tumores, al que se le ha determinado relación con el proceso de integración del VPH (Wentzensen et al); COX412, subunidad del citocromo c-oxidasa con aumento de su expresión tras la integración de VPH; así como FGFR3, HNRNPD, REV1, y KLF5, genes relacionados a la patogénesis de VPH cuyas regiones exónicas se encontraban en las posiciones de punto de ruptura por parte de Virus del Papiloma Humano 16. Como se ha indicado en otros estudios (Rusan et al., 2015), esta relación entre las integraciones de VPH y las alteraciones que pueden conducir a la pérdida o ganancia de función en los genes claves del cáncer sugiere que la integración en unidades transcripcionales contribuye a la tumorigénesis.

Adicionalmente se encontró que los sitios de integración coincidían con sitios hipersensibles a DNasa I en un 18%, que son sitios que han sido atribuidos como regiones disponibles para el evento de integración (I. K. Christiansen et al., 2015) y asociados a la inestabilidad genómica que es una característica de los cánceres humanos, incluido el 5% causada por el Virus del Papiloma Humano (VPH) (Akagi et al., 2014) y , secuencias repetitivas Alu en un 9.49%, resultado que coincide con lo reportado por Zhu y colegas donde identificaron que entre el 83.3% y 92%

de los eventos de integración se encontraron en elementos repetitivos, específicamente en un 14% para secuencias Alu (Hu z., 2015). En cuanto a las regiones ricas en dinucleótidos CG, sólo un sitio de integración contenía una isla CpG, lo anterior posiblemente ocurre debido a que la metilación del ADN puede actuar como barrera que impide al virus integrarse en estas regiones (V. Badal et al., 2003). Por último, las regiones intergénicas presentaron un 22% de sitios de integración del VPH 16, las cuales a pesar que no tienen regiones codificantes se proponen que presentan elementos móviles y regiones repetidas que podrían estar siendo afectados por el evento de integración. Adicionales estudios en este tema son necesarios.

2.7. Conclusiones:

- ◆ Se detectó 14% de ADN del VPH16 en muestras de cáncer oral embebidas en parafina por la técnica de PCR en tiempo real. Además, se logró determinar su estado físico de integración. Lo anterior es muy importante ya que permite utilizar una técnica molecular de alta especificidad y sensibilidad para determinar el estado de integración viral en muestras embebidas en parafina y fijadas con formalina que se encuentran almacenadas y disponibles en bancos de muestras patológicas en clínicas y hospitales.
- ◆ Los cromosomas 9,19, 20 y 21 presentaron mayor número de eventos de integración de VPH16 en proporción a su tamaño. Los cromosomas 9,X,20,21 presentaron mayor número de eventos de integración de VPH16
- ◆ Las regiones moleculares del genoma del hospedero relacionadas a los eventos de integración del VPH16 mostraron que el 50% de los sitios de integración evaluados ocurrieron en unidades transcripcionales y solo el 1% en regiones ricas en dinucleótidos CG.

2.8. Perspectivas

- ◆ Son necesarios adicionales estudios para entender cómo la integración del ADN del VPH16 en regiones codificantes puede afectar la expresión de genes codificantes y la función de las proteínas cuyos genes fueron afectados.
- ◆ Son necesarios adicionales estudios para entender cómo la integración del ADN del VPH16 en regiones no codificantes o intergénicas puede afectar la estabilidad cromosómica y su relación con los elementos móviles del genoma.

2.9. Bibliografía


- Akagi, K., Li, J., Broutian, T. R., Padilla-Nash, H., Xiao, W., Jiang, B., . . . Gillison, M. L. (2014). Genome-wide analysis of HPV integration in human cancers reveals recurrent, focal genomic instability. *Genome Research*, *24*(2), 185-199. doi: 10.1101/gr.164806.113
- Badal, V., Chuang, L. S., Tan, E. H., Badal, S., Villa, L. L., Wheeler, C. M., . . . Bernard, H. U. (2003). CpG methylation of human papillomavirus type 16 DNA in cervical cancer cell lines and in clinical specimens: genomic hypomethylation correlates with carcinogenic progression. *J Virol*, *77*(11), 6227-6234.
- Carmody, M. W., Jones, M., Tarraza, H., & Vary, C. P. (1996). Use of the polymerase chain reaction to specifically amplify integrated HPV-16 DNA by virtue of its linkage to interspersed repetitive DNA. *Mol Cell Probes*, *10*(2), 107-116. doi: 10.1006/mcpr.1996.0015
- Castillo, A., Aguayo, F., Koriyama, C., Torres, M., Carrascal, E., Corvalan, A., . . . Akiba, S. (2006). Human papillomavirus in esophageal squamous cell carcinoma in Colombia and Chile. *World J Gastroenterol*, *12*(38), 6188-6192.
- Castillo, A., Koriyama, C., Higashi, M., Anwar, M., Bukhari, M. H., Carrascal, E., . . . Akiba, S. (2011). Human papillomavirus in upper digestive tract tumors from three countries. *World J Gastroenterol*, *17*(48), 5295-5304. doi: 10.3748/wjg.v17.i48.5295
- Castillo, A. (2014). *Human Papillomavirus and Carcinogenesis in the Upper Aero-Digestive Tract*. *Colombia Medica* *42*(2):233-242.
- Christiansen, I. K., Sandve, G. K., Schmitz, M., Durst, M., & Hovig, E. (2015). Transcriptionally active regions are the preferred targets for chromosomal HPV integration in cervical carcinogenesis. *PLoS ONE*, *10*(3), e0119566. doi: 10.1371/journal.pone.0119566
- Daponte, A., Tsezou, A., Oikonomou, P., Hadjichristodoulou, C., Maniatis, A. N., Pournaras, S., & Messinis, I. E. (2008). Use of real-time PCR to detect human papillomavirus-16 viral loads in vaginal and urine self-sampled specimens. *Clin Microbiol Infect*, *14*(6), 619-621. doi: 10.1111/j.1469-0691.2008.01974.x
- Das, P., Thomas, A., Mahantshetty, U., Shrivastava, S. K., Deodhar, K., & Mulherkar, R. (2012). HPV genotyping and site of viral integration in cervical cancers in Indian women. *PLoS ONE*, *7*(7), e41012. doi: 10.1371/journal.pone.0041012
- Giovannelli, L., Campisi, G., Lama, A., Giambalvo, O., Osborn, J., Margiotta, V., & Ammatuna, P. (2002). Human papillomavirus DNA in oral mucosal lesions. *J Infect Dis*, *185*(6), 833-836. doi: 10.1086/339193
- Hu, Z., Zhu, D., Wang, W., Li, W., Jia, W., Zeng, X., . . . Ma, D. (2015). Genome-wide profiling of HPV integration in cervical cancer identifies clustered

- genomic hot spots and a potential microhomology-mediated integration mechanism. *Nat Genet*, 47(2), 158-163. doi: 10.1038/ng.3178
- Kadaja, M., Sumerina, A., Verst, T., Ojarand, M., Ustav, E., & Ustav, M. (2007). Genomic instability of the host cell induced by the human papillomavirus replication machinery. *The EMBO Journal*, 26(8), 2180-2191. doi: 10.1038/sj.emboj.7601665
- Kazemian, M., Ren, M., Lin, J. X., Liao, W., Spolski, R., & Leonard, W. J. (2015). Possible Human Papillomavirus 38 Contamination of Endometrial Cancer RNA Sequencing Samples in The Cancer Genome Atlas Database. *J Virol*, 89(17), 8967-8973. doi: 10.1128/jvi.00822-15
- Kreimer, A. R., Clifford, G. M., Boyle, P., & Franceschi, S. (2005). Human papillomavirus types in head and neck squamous cell carcinomas worldwide: a systematic review. *Cancer Epidemiol Biomarkers Prev*, 14(2), 467-475. doi: 10.1158/1055-9965.epi-04-0551
- Kumar Gupta, A., & Kumar, M. (2015). HPVbase--a knowledgebase of viral integrations, methylation patterns and microRNAs aberrant expression: As potential biomarkers for Human papillomaviruses mediated carcinomas. *Sci Rep*, 5, 12522. doi: 10.1038/srep12522
- Li, H., Yang, Y., Zhang, R., Cai, Y., Yang, X., Wang, Z., . . . Zhu, B. (2013). Preferential sites for the integration and disruption of human papillomavirus 16 in cervical lesions. *J Clin Virol*, 56(4), 342-347. doi: 10.1016/j.jcv.2012.12.014
- Lillsunde Larsson, G., Helenius, G., Sorbe, B., & Karlsson, M. G. (2014). Viral load, integration and methylation of E2B53 and 4 in human papilloma virus (HPV) 16-positive vaginal and vulvar carcinomas. *PLoS ONE*, 9(11), e112839. doi: 10.1371/journal.pone.0112839
- Matos, L. L., Miranda, G. A., & Cernea, C. R. (2015). Prevalence of oral and oropharyngeal human papillomavirus infection in Brazilian population studies: a systematic review. *Braz J Otorhinolaryngol*, 81(5), 554-567. doi: 10.1016/j.bjorl.2015.04.001
- Muñoz, N., Bosch, F. X., de Sanjose, S., Herrero, R., Castellsague, X., Shah, K. V., . . . Meijer, C. J. (2003). Epidemiologic classification of human papillomavirus types associated with cervical cancer. *N Engl J Med*, 348(6), 518-527. doi: 10.1056/NEJMoa021641
- Nobel, Assembly. (2008). The discoveries of human papilloma viruses that cause cervical cancer. http://www.nobelprize.org/nobel_prizes/medicine/laureates/2008/advanced-medicineprize2008.pdf.
- Parkin, D. M. (2006). The global health burden of infection-associated cancers in the year 2002. *Int J Cancer*, 118(12), 3030-3044. doi: 10.1002/ijc.21731
- Pett, M., & Coleman, N. (2007). Integration of high-risk human papillomavirus: a key event in cervical carcinogenesis? *J Pathol*, 212(4), 356-367. doi: 10.1002/path.2192
- Prabhu, S. R., & Wilson, D. F. (2013). Human papillomavirus and oral disease - emerging evidence: a review. *Aust Dent J*, 58(1), 2-10; quiz 125. doi: 10.1111/adj.12020
- Rusan, M., Li, Y. Y., & Hammerman, P. S. (2015). Genomic Landscape of Human Papillomavirus-Associated Cancers. *Clinical cancer research : an official*

- Journal of the American Association for Cancer Research*, 21(9), 2009-2019. doi: 10.1158/1078-0432.CCR-14-1101
- Schiffman, M., & Castle, P. E. (2003). Human papillomavirus: epidemiology and public health. *Arch Pathol Lab Med*, 127(8), 930-934. doi: 10.1043/1543-2165(2003)127<930:hpeaph>2.0.co;2
- Shukla, S., Mahata, S., Shishodia, G., Pande, S., Verma, G., Hedau, S., . . . Bharti, A. C. (2014). Physical state & copy number of high risk human papillomavirus type 16 DNA in progression of cervical cancer. *Indian J Med Res*, 139(4), 531-543.
- Vosa, L., Sudakov, A., Remm, M., Ustav, M., & Kurg, R. (2012). Identification and analysis of papillomavirus E2 protein binding sites in the human genome. *J Virol*, 86(1), 348-357. doi: 10.1128/jvi.05606-11
- Walboomers, J. M., Jacobs, M. V., Manos, M. M., Bosch, F. X., Kummer, J. A., Shah, K. V., . . . Munoz, N. (1999). Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J Pathol*, 189(1), 12-19. doi: 10.1002/(sici)1096-9896(199909)189:1<12::aid-path431>3.0.co;2-f
- Wentzensen, N., Vinokurova, S., & von Knebel Doeberitz, M. (2004). Systematic review of genomic integration sites of human papillomavirus genomes in epithelial dysplasia and invasive cancer of the female lower genital tract. *Cancer Res*, 64(11), 3878-3884. doi: 10.1158/0008-5472.can-04-0009
- Xu, B., Chotewutmontri, S., Wolf, S., Klos, U., Schmitz, M., Dürst, M., & Schwarz, E. (2013). Multiplex Identification of Human Papillomavirus 16 DNA Integration Sites in Cervical Carcinomas. *PLoS ONE*, 8(6), e66693. doi: 10.1371/journal.pone.0066693
- zur Hausen, H. (2002). Papillomaviruses and cancer: from basic studies to clinical application. *Nat Rev Cancer*, 2(5), 342-350. doi: 10.1038/nrc798
- Zaldívar-Lelo de Larrea, Guadalupe, Martín-Molina, Francisco, Sosa-Ferreya, Carlos Francisco, Ávila-Morales, Javier, Lloret-Rivas, Miguel, & Vega-Malagón, Genaro. (2011). Cáncer cervicouterino y virus del papiloma humano. *Gaceta Mexicana de Oncología*, 14(03), 288-294.

3. ANEXOS

Anexo 1. Acta de aprobación ética. Este proyecto tuvo aprobación del Comité Institucional de Revisión de Ética Humana (CIREH) de la Facultad de Salud de la Universidad del Valle con acta de aprobación N° 05-014.

Comité Institucional de Revisión de Ética Humana	
Facultad de Salud	<small>Universidad del Valle</small>

ACTA DE APROBACIÓN N° 05-014

Proyecto: PREVALENCIA Y CARACTERIZACIÓN GENOTÍPICA DEL VIRUS DEL PAPILOMA HUMANO EN ALTERACIONES POTENCIALMENTE MALIGNAS Y CÁNCER ORAL EN 5 REGIONES DE COLOMBIA. ESTUDIO MULTICENTRO

Sumetido por: SANDRA AMAYA/DORA ORDOÑEZ/NATALIA ARAGON/LUIS EDUARDO BRAVO/GLORIA INES LAFAURIE/MARIA ROSA BUENAHORA/SANDRA J. PERDOMO/DANIELA GUTIERREZ/CARLOS MARTIN ARDILA/EFRAIN ALVAREZ/FARITH GONZALEZ/GLORIA ARANZAZU/CARLOS TASAMA

Código Interno: **Fecha en que fue sometido:**

El Consejo de la Facultad de Salud de la Universidad del Valle, ha establecido el Comité Institucional de Revisión de Ética Humana (CIREH), el cual está regido por la Resolución 008430 del 4 de octubre de 1993 del Ministerio de Salud de Colombia por la cual se establecen las normas científicas, técnicas y administrativas para la investigación en salud; los principios de la Asamblea Médica Mundial expresados en su Declaración de Helsinki de 1964, última revisión en 2002; y el Código de Regulaciones Federales, título 45, parte 46, para la protección de sujetos humanos, del Departamento de Salud y Servicios Humanos de los Institutos Nacionales de Salud de los Estados Unidos 2000.

Este Comité certifica que:

- Sus miembros revisaron los siguientes documentos del presente proyecto:

<input checked="" type="checkbox"/> Resumen del proyecto	<input checked="" type="checkbox"/> Protocolo de investigación
<input checked="" type="checkbox"/> Formulario de consentimiento informado	<input checked="" type="checkbox"/> Instrumento de recolección de datos
<input type="checkbox"/> Folleto del investigador (si aplica)	<input checked="" type="checkbox"/> Cartas de las instituciones participantes
<input type="checkbox"/> Resultados de evaluación por otros comités (si aplica)	
- El presente proyecto fue evaluado y aprobado por el Comité;
- Según las categorías de riesgo establecidas en el artículo 11 de la Resolución N° 008430 de 1993 del Ministerio de Salud, el presente estudio tiene la siguiente **Clasificación de Riesgo**:
 SIN RIESGO RIESGO MÍNIMO RIESGO MAYOR DEL MÍNIMO
- Que las **medidas** que están siendo tomadas para proteger a los sujetos humanos son adecuadas;
- La forma de obtener el **consentimiento** informado de los participantes en el estudio es adecuada;
- Este proyecto será **revisado nuevamente** en la próxima reunión plenaria del Comité; sin embargo, el Comité puede ser convocado a solicitud de algún miembro del Comité o de las directivas institucionales para revisar cualquier asunto relacionado con los derechos y el bienestar de los sujetos involucrados en este estudio;
- Informará** inmediatamente a las directivas institucionales:
 - Todo descuido de los investigadores a las solicitudes del Comité;
 - Cualquier suspensión o terminación de la aprobación por parte del Comité;
- Informará** inmediatamente a las directivas institucionales toda información que reciba acerca de:
 - Lesiones a sujetos humanos;

Calle 4B 36 -00 edificio Decanato Teléfono: 5185677 email: eticasalud@univalle.edu.co



- Problemas imprevistos que involucren riesgos para los sujetos u otras personas.
- b. Cualquier cambio o modificación a este proyecto que haya sido revisado y aprobado por el Comité.
 9. El presente proyecto ha sido **aprobado** por un periodo de **1 año** a partir de la fecha de aprobación. Los proyectos de duración mayor a un año, deberán ser sometidos nuevamente con todos los documentos para revisión actualizados.
 10. El **investigador principal** deberá informar al Comité:
 - a. Cualquier cambio que se proponga introducir en este proyecto. Estos cambios no podrán iniciarse sin la revisión y aprobación del Comité excepto cuando sean necesarios para eliminar peligros inminentes para los sujetos.
 - b. Cualquier problema imprevisto que involucre riesgos para los sujetos u otros.
 - c. Cualquier evento adverso serio dentro de las primeras 24 horas de ocurrido, al secretario(a) y al presidente (Anexo 1).
 - d. Cualquier cancelamiento nuevo respecto al estudio, que pueda afectar la tasa riesgo/beneficio para los sujetos participantes.
 - e. cualquier decisión tomada por otros comités de ética.
 - f. La terminación prematura o suspensión del proyecto explicando la razón para esto.
 - g. El investigador principal deberá presentar un informe al final del año de aprobación. Los proyectos de duración mayor a un año, deberán ser sometidos nuevamente con todos los documentos para revisión actualizados.

Firma:  Fecha: 03 07 2014
Nombre: FLORENCIA VELASCO DE M.
Capacidad representativa: PRESIDENTA Teléfono: 5185677

CERTIFICACIÓN DE LA FACULTAD DE SALUD DE LA UNIVERSIDAD DEL VALLE

Por medio de la presente, certifico que la Facultad de Salud de la Universidad del Valle aprueba el proyecto arriba mencionado y respeta los principios, políticas y procedimientos de la Declaración de Helsinki de la Asamblea Médica Mundial, de la Resolución 8430 de 1993 del Ministerio de Salud y de la reglamentación vigente en investigación de la Universidad del Valle.

Firma:  Fecha: 03 07 2014
Nombre: MAURICIO PALACIOS
Capacidad representativa: VICEDECANO DE LA FACULTAD DE SALUD Teléfono: 5185680

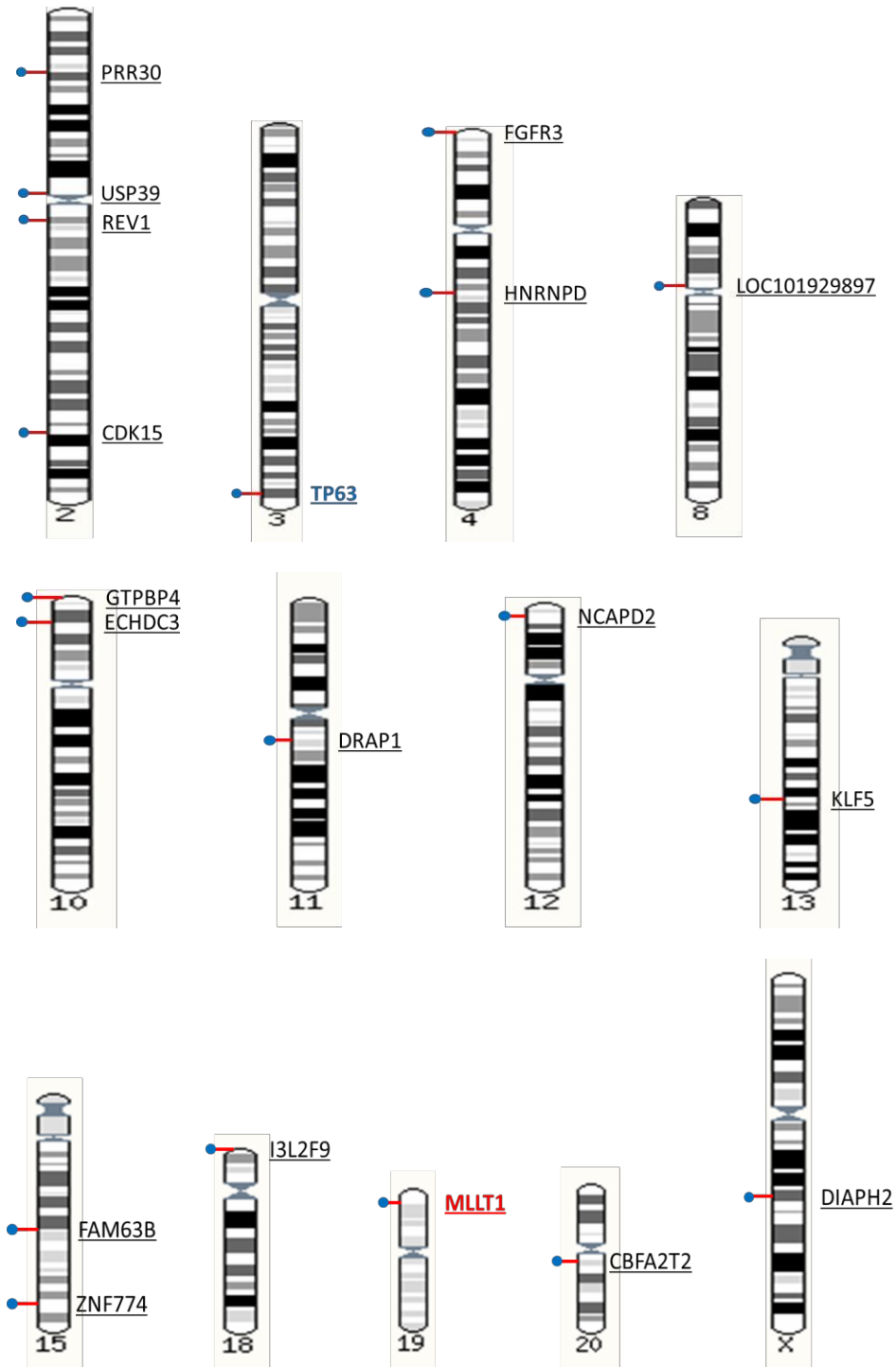
Anexo 2. Muestras y sitios anatómicos correspondientes. 59 muestras de cáncer oral pertenecientes a 37 pacientes.

No. DE PACIENTE	SITIO ANATÓMICO	EDAD	SEXO
1	LENGUA	61	M
2	LENGUA LENGUA -AGT	76	F
3	PISO DE BOCA	59	F
4	LENGUA	73	M
5	LENGUA (base)	48	M
6	MUCOSA YUGAL IZQUIERDA Borde Superior LIBRE Nódulo región bucal izquierda Borde anterior y lesión	53	F
7	PALADAR BLANDO DERECHO	58	M
8	PALADAR BLANDO DURO Posterior Posterior lateral	80	F
9	REGION SUBMANDIBULAR	80	F
10	PISO DE BOCA Amígdala palina izquierda Tumor	64	M
11	PISO DE BOCA BASE DE LENGUA	56	F
12	LENGUA	52	F
13	LENGUA	68	F
14	LENGUA LENGUA	68	F
15	PISO BOCA Y LENGUA Lesion borde profundo	56	F
16	LENGUA	60	M
17	PALADAR	48	M
18	PISO DE BOCA	68	F
19	LENGUA	54	F
20	LENGUA PALADAR BLANDO	70	F
21	VESTIBLO ORAL MEJILLA MAX INF	NR	F
22	LENGUA - BASE DE LENGUA LESION LENGUA CON BORDE BASE CON COMPROMISO DE GLANDULA	33	M
23	LENGUA	58	M
24	LENGUA LENGUA	73	M
25	Extensión a piso de boca Borde resección Tumor Borde superior	49	F
26	LENGUA PISO DE BOCA Borde superior	NR	F
27	LENGUA	81	M
28	PALADAR	52	M
29	LABIO Piel mucosa	57	F
30	LABIO	67	M
31	LENGUA Fragmento	53	M
32	PALADAR BLANDO BASE DE LENGUA	68	F
33	PISO DE BOCA Borde profundo	83	F
34	MAXILAR IZQUIERDO	27	F
35	LABIO	69	M
36	LENGUA	NR	F
37	LABIO	78	M

Anexo 3. Artículos seleccionados para la recolección de eventos de integración. Se seleccionó un total de 16 artículos que abarcaban desde el año 1997 al año 2016. Al final se seleccionaron 611 eventos de integración.

#	ARTÍCULO	AUTOR	REVISTA	AÑO DE PUBLICACIÓN	EVENTOS DE INTEGRACIÓN
1	Complex genomic rearrangement within the 12q15 multiple aberration region induced by integrated human papillomavirus 18 in a cervical carcinoma cell line.	P. Lazo, et al.	MOLECULAR CARCINOGENESIS	1997	1
2	Genomic instability of the host cell induced by the human papillomavirus replication machinery	M. Ustav, et al.	The EMBO Journal	2007	1
3	Characterization of Naturally Occurring HPV16 Integration sites isolated from cervical keratinocytes under noncompetitive conditions	N. Coleman, et al.	Cancer research	2008	26
4	Episomal and integrated human papillomavirus type 16 loads and anal intraepithelial neoplasia in HIV-seropositive men	F. Coutlée, et al.	AIDS	2010	11
5	HPV genotyping and site of viral integration in cervical cancers in Indian women	R. Mulherkar, et al.	PLOS ONE	2012	38
6	Loss of gene function as a consequence of human papillomavirus DNA integration	M. Dürst, et al.	International Journal of Cancer	2012	12
7	Mate pair sequencing of oropharyngeal squamous cell carcinomas reveals that HPV integration occurs much less frequently than in cervical cancer	D. Smith, et al.	Journal of Clinical Virology	2013	4
8	Genome-wide analysis of HPV integration in human cancers reveals recurrent, focal genomic instability	M. Gillison, et al.	Genome Research	2013	89
9	Multiplex Identification of Human Papillomavirus 16 DNA Integration Sites in Cervical Carcinomas	E. Schwarz, et al.	PLOS ONE	2013	70
10	Characterization of HPV and host genome interactions in primary head and neck cancers	M. Parfenov, et al.	PLOS ONE	2014	103
11	Comprehensive Analysis of HPV16 Integration in OSCC Reveals No Significant Impact of Physical Status on Viral Oncogene and Virally Disrupted Human Gene Expression	N. Olthof, et al.	PLOS ONE	2014	46
12	NGS-based approach to determine the presence of HPV and their sites of integration in human cancer genome	A. Dutt, et al.	British Journal of Cancer	2015	39
13	Comprehensive mapping of the human papillomavirus (HPV) DNA integration sites in cervical carcinomas by HPV capture technology	Y. Liu, et al.	Oncotarget	2015	116
14	Possible Human Papillomavirus 38 Contamination of Endometrial Cancer RNA Sequencing Samples in The Cancer Genome Atlas Database	M. Kasemian, et al.	Journal of virology	2015	19
15	Viral load, gene expression and mapping of viral integration sites in HPV16-associated HNSCC cell lines	E. Speel, et al.	International Journal of Cancer	2015	10
16	Chromosomal copy number alterations and HPV integration in cervical precancer and invasive cancer	N. Wentzensen, et al.	Carcinogenesis	2016	26
TOTAL DE SITIOS DE INTEGRACIÓN					611

Anexo 4. Sitios de integración viral presentes en regiones exónicas. Se presentan los 21 sitios de integración viral presentes en regiones exónicas de 19 unidades transcripcionales (genes). Los genes USP39 y TP63 presentaron 2 sitios de integración viral, cada uno. **TP63** (Cr.3) es un gen supresor de tumor y **MLLT1** (Cr. 19) es un oncogén.



Anexo 5. Proporción de sitios de integración del VPH16 en genes sobre el número de genes por cromosoma. El cromosoma 20 y 9 presentan la mayor proporción, seguidos por los cromosomas 21, X y 3.

CROMOSOMA	SITIOS DE INTEGRACIÓN	No. GENES TOTALES	No. GENES EN SITIOS DE INTEGRACIÓN	PROPORCIÓN
1	32	4778	15	0,003
2	36	3600	17	0,005
3	44	2780	25	0,009
4	19	2251	12	0,005
5	15	2393	1	0,000
6	27	2828	9	0,003
7	21	2598	12	0,005
8	36	2003	7	0,003
9	49	2132	25	0,012
10	15	2042	7	0,003
11	23	2806	13	0,005
12	7	2398	2	0,001
13	25	1327	9	0,007
14	10	1951	6	0,003
15	16	1721	11	0,006
16	8	1845	4	0,002
17	22	2326	18	0,008
18	14	938	6	0,006
19	23	2420	19	0,008
20	23	1243	18	0,014
21	18	726	8	0,011
22	8	1129	4	0,004
X	40	1973	20	0,010
Y	1	496	0	0,000