

Identificación de variantes genéticas en el gen *AHR* que puedan contribuir a la susceptibilidad a cáncer de cabeza y cuello en población colombiana.

Nathalia Andrea Trujillo Pelayo

Trabajo de Grado para Optar el título de Magister en Ciencias Básicas Biomédicas

Directora

Laura Fernanda Cifuentes Cardona

Doctora en Ciencias Biológicas

Codirectora

Clara Inés Vargas Castellanos

Magister en Biología, Fellow Genética Clínica

Universidad Industrial de Santander

Facultad de Salud

Departamento de Ciencias Básicas

Maestría en Ciencias Básicas Biomédicas

Bucaramanga

2023

## Dedicatoria

*En memoria de quienes han sido mi gran tesoro en la vida  
y quienes ahora en la distancia me acompañan...  
mi Madre, **URANIA**.  
Mis abuelas, Abigail, Clara y Magdalena.  
Mi tío Hernando, quién luchó contra esta enfermedad llamada Cáncer,  
sin salir vencedor.  
Luchita.*

### **Agradecimientos**

No hay palabras para expresar a Dios la gratitud que siento al finalizar otra etapa en mi vida profesional. A El quién ha sido siempre mi fortaleza para continuar y aunque no ha sido fácil, siempre me sostiene y ha dispuesto para mí, personas maravillosas que de una u otra forma han hecho parte de este proceso. A mi gran ángel, mi querida madre, quién desde siempre creyó en mí, en mis capacidades, quien fue y sigue siendo mi motor y mi sustento, quién mantiene ese deseo vivo en mí de seguir formándome profesionalmente. Hoy echo de menos su presencia, pero sé que desde el cielo celebra con orgullo este nuevo logro.

A mis queridas docentes y tutoras, la Doctora Laura Fernanda Cifuentes Cardona y la Dra. Clara Inés Vargas Castellano, para quiénes tengo un gran afecto, respeto y admiración. No sólo me acogieron como su estudiante, también me brindaron su apoyo emocional y junto con sus enseñanzas, consejos y guía he logrado culminar este proyecto y son para mí un gran ejemplo a seguir dentro de este lindo campo de la Genética.

A los Doctores Sandra Guauque Olarte y Cristian Fong, quienes me acogieron, enseñaron y guiaron durante mi proceso de aprendizaje, afianzando en mí nuevos conocimientos en el área de la bioinformática y nuevas técnicas de laboratorio. Un agradecimiento especial al Dr. Álvaro

Herrera Hernández y todos los pacientes participantes de esta investigación, la cual no habría sido posible sin su amable colaboración.

Agradezco a mi familia, que con todo su amor me han acompañado y apoyado siempre, especialmente mi padre, mis hermanas, mi sobrino y mi amor bonito que me hacen ser mejor cada día, manteniendo el legado de fortaleza y entrega de mamá. Gracias mil, a esas amistades que llenan el alma con palabras de aliento y motivación, Ximena, Liliana, Alejandra, Shirley, Valentina y Mary, quienes me brindaron su apoyo incondicional y han sido parte de este sueño que hoy se hace realidad.

Finalmente, agradezco a la Universidad Industrial de Santander, institución que me ha brindado la oportunidad de formarme profesionalmente, dando continuidad a mis estudios de posgrado; a cada uno de mis docentes del programa de Maestría en Ciencias Básicas Biomédicas, quienes con sus enseñanzas y rigor ayudaron en mi proceso de formación como Magister; a la Universidad Cooperativa de Colombia – CONADI (INV2085), por la financiación de este proyecto.

## Tabla de Contenido

<b>INTRODUCCIÓN</b> .....	<b>18</b>
<b>1. OBJETIVOS</b> .....	<b>22</b>
<b>1.1 Objetivo general</b> .....	<b>22</b>
<b>1.2 Objetivos específicos</b> .....	<b>22</b>
<b>2. MARCO REFERENCIAL</b> .....	<b>23</b>
<b>2.1 Marco teórico</b> .....	<b>23</b>
2.1.1 Carcinoma escamocelular de cabeza y cuello .....	24
2.1.1.1 Definición .....	24
2.1.1.2 Epidemiología.....	25
2.1.1.3 Progresión oncogénica del HNSCC .....	27
2.1.1.4 Factores de riesgo y etiología del HNSCC .....	28
HNSCC - VPH positivo .....	28
HNSCC - VPH negativo .....	30
2.1.2 Gen <i>AHR</i> – consumo de tabaco – desarrollo de HNSCC .....	32
2.1.2.1 Gen <i>AHR</i> .....	33
2.1.2.2 Receptor Aril de Hidrocarburos (AhR) .....	34
2.1.2.3 Vías de señalización celular del AhR. ....	36
2.1.2.4 Rol del AhR en la tumorigénesis. ....	39

<b>2.2 Estado del Arte .....</b>	<b>42</b>
<b>2.3 Marco legal .....</b>	<b>42</b>
2.3.1 Tratamiento de datos personales .....	42
2.3.2 Clasificación del riesgo de la investigación .....	43
<b>3. METODOLOGÍA .....</b>	<b>44</b>
<b>3.1 Fase I: análisis de variantes en el gen <i>AHR</i> reportadas en repositorios de pacientes cáncer y su potencial impacto. ....</b>	<b>44</b>
3.1.1 Variantes en repositorios genómicos y cáncer .....	44
3.1.2 Integración de datos de repositorios, análisis y potencial impacto .....	45
<b>3.2 Fase II: identificación de variantes genéticas en el gen <i>AHR</i> en pacientes colombianos diagnosticados con cáncer de cabeza y cuello.....</b>	<b>49</b>
3.2.1 Población de estudio.....	49
3.2.1.1 Criterios de inclusión.....	50
3.2.1.2 Criterios de exclusión .....	50
3.2.2 Información de variables .....	51
3.2.3 Recolección de muestras .....	52
3.2.4 Extracción de ADN .....	53
3.2.5 Análisis del gen <i>AHR</i> .....	53
3.2.6 Análisis computacional .....	56
<b>3.3 Fase III: variantes en el gen <i>AHR</i>, determinación de su posible impacto en la estructura y función del receptor de Aril hidrocarburos (<i>AhR</i>) .....</b>	<b>56</b>

**4. RESULTADOS Y DISCUSIÓN .....61**

**4.1 Fase I: análisis de variantes en el gen *AHR* reportadas en repositorios de pacientes cáncer y su potencial impacto. ....61**

**4.2 Fase II: identificación de variantes genéticas en el gen *AHR* en pacientes colombianos diagnosticados con cáncer de cabeza y cuello.....61**

4.2.1 Caracterización sociodemográfica y clínica de los pacientes estudiados.....61

4.2.2 Alteraciones de secuencia observadas en el gen *AHR*.....67

4.2.2.1 Análisis de variantes alélicas «reportadas» para el gen *AHR*. ....69

4.2.2.2 Análisis de variantes alélicas «no reportadas» para el gen *AHR*. ....83

**4.3 Fase III: variantes en el gen *AHR*, determinación de su posible impacto en la estructura y función del receptor de Aril hidrocarburos (*AhR*) .....84**

4.3.1 Análisis *in silico* de las alteraciones en la región codificante .....84

4.3.2 Análisis *in silico* de las alteraciones en la región no codificante .....90

**4.4 Discusión de resultados.....96**

**5. CONCLUSIONES.....103**

**REFERENCIAS BIBLIOGRÁFICAS.....105**

**APÉNDICE.....129**

**Lista de Tablas**

	<b>PAG.</b>
<b>Tabla 1.</b> Clasificación de neoplasias malignas de Cáncer de cabeza y cuello, según clasificación CIE10.	50
<b>Tabla 2.</b> Cebadores de PCR específicos para <i>AHR</i>	54
<b>Tabla 3.</b> Protocolo de amplificación de PCR	129
<b>Tabla 4.</b> Condiciones de Termociclado - Técnica de PCR	129
<b>Tabla 5.</b> Caracterización sociodemográfica de la población de estudio	63
<b>Tabla 6.</b> Caracterización clínica de la población de estudio	66
<b>Tabla 7.</b> Alteraciones de secuencia identificadas en el gen <i>AHR</i> en la muestra caracterizada	69
<b>Tabla 8.</b> Alteraciones de secuencias en el gen <i>AHR</i> : variantes reportadas y significancia clínica	70
<b>Tabla 9.</b> Frecuencias alélicas de la variante [-459A>G]	73
<b>Tabla 10.</b> Frecuencias alélicas de la variante [c.908+33G>T]	76
<b>Tabla 11.</b> Frecuencias alélicas de la variante [c.132T>C]	79
<b>Tabla 12.</b> Frecuencias alélicas de la variante [c.1661G>A]	82
<b>Tabla 13.</b> Variantes de significado clínico desconocido halladas en el intrón 2 y el exón 11 del gen <i>AHR</i>	83
<b>Tabla 14.</b> Análisis de predicción in silico para la variante <i>AHR</i> [c.132T>C] y su significancia clínica	86

VARIANTES GENÉTICAS Y SUSCEPTIBILIDAD AL DESARROLLO DE HNC	9
<b>Tabla 15.</b> Análisis de predicción in silico para la variante <i>AHR</i> [c.1661G>A] y su significancia clínica	88
<b>Tabla 16.</b> Análisis in silico de la variante rs2066853 [c.1661G>A] mediante la herramienta Human Splicing Finder (HSF)	130
<b>Tabla 17.</b> Pacientes analizados, portadores de la variante p.Arg554Lys (rs2066853).	89
<b>Tabla 18.</b> Análisis de predicción in silico para la variante <i>AHR</i> [-459A>G] y su significancia clínica	92
<b>Tabla 19.</b> Análisis in silico de la variante rs7796976 [-459A>G] mediante la herramienta Human Splicing Finder (HSF)	130
<b>Tabla 20.</b> Análisis de la variante rs7796976 [-459A>G] mediante la herramienta HaploReg V4.1: unión de proteínas en experimentos ChIP-Seq (ENCODE Project Consortium)	131
<b>Tabla 21.</b> Análisis de la variante rs7796976 [-459A>G] mediante la herramienta HaploReg V4.1: motivos regulatorios alterados (Position Weight Matrix)	132
<b>Tabla 22.</b> Pacientes analizados, portadores de la variante [-459A>G] (rs7796976).	93
<b>Tabla 23.</b> Análisis de predicción in silico para la variante <i>AHR</i> [c.908+33G>T] y su significancia clínica	95
<b>Tabla 24.</b> Análisis in silico de la variante rs2074113 [c.908+33G>T] mediante la herramienta Human Splicing Finder (HSF)	132
<b>Tabla 25.</b> Análisis de la variante rs2074113 [c.908+33G>T] mediante la herramienta HaploReg V4.1: motivos regulatorios alterados (Position Weight Matrix)	133
<b>Tabla 26.</b> Pacientes analizados, portadores de la variante c.908+33G>T (rs2074113).	96

### Lista de Figuras

	<b>PAG.</b>
<b>Figura 1.</b> Sitios anatómicos del desarrollo de HNSCC	25
<b>Figura 2.</b> Tasa de incidencia de HNC en el mundo en 2020	26
<b>Figura 3.</b> Progresión oncogénica del HNSCC	28
<b>Figura 4.</b> Desarrollo de HNSCC - VPH negativo	31
<b>Figura 5.</b> Vista genómica del gen <i>AHR</i>	34
<b>Figura 6.</b> Dominios funcionales de la proteína AhR (proteínas bHLH-PAS)	35
<b>Figura 7.</b> Ligandos AhR y función reguladora del receptor	36
<b>Figura 8.</b> Rutas de señalización del receptor Aril hidrocarburos (AhR)	38
<b>Figura 9.</b> Vía de señalización de AhR en tumores	41
<b>Figura 10.</b> Información recabada para cada una de las variantes identificadas, de acuerdo al repositorio genómico explorado	46
<b>Figura 11.</b> Representación gráfica: análisis de variantes en el gen <i>AHR</i> reportadas en repositorios de pacientes cáncer y su potencial impacto	48
<b>Figura 12.</b> Número de pacientes estudiados de acuerdo con la localización anatómica del tumor	64
<b>Figura 13.</b> Antecedentes de cáncer en parientes de primer y segundo grado de pacientes con HNSCC	67
<b>Figura 14.</b> Representación gráfica del gen <i>AHR</i> y las variantes identificadas en el estudio	68
<b>Figura 15.</b> Cromatograma de la alteración c.155A>G [-459A>G]	72
<b>Figura 16.</b> Cromatograma de la alteración c.1521+33G>T [c.908+33G>T]	75

<b>Figura 17.</b> Cromatograma de la alteración c.745T>C [c.132T>C]	78
<b>Figura 18.</b> Cromatograma de la alteración c.2274G>A [c.1661G>A]	81
<b>Figura 19.</b> Representación esquemática de los dominios de <i>AHR</i>	85

**Lista de Apéndices**

	<b>PÁG.</b>
<b>APÉNDICE A.</b> Estado del Arte: Polymorphisms Associated to Genetic Susceptibility and Head and Neck Cancer Risk: A Systematic Review.	129
<b>APÉNDICE B.</b> Tabla 3. Protocolo de amplificación de PCR	129
<b>APÉNDICE C.</b> Tabla 4. Condiciones de Termociclado - Técnica de PCR	129
<b>APÉNDICE D.</b> <i>AHR</i> gene variants in patients with cancer and their presence in the Latin American population	130
<b>APÉNDICE E.</b> Tabla 16. Análisis in silico de la variante rs2066853 [c.1661G>A] mediante la herramienta Human Splicing Finder (HSF)	130
<b>APÉNDICE F.</b> Tabla 19. Análisis in silico de la variante rs7796976 [-459A>G] mediante la herramienta Human Splicing Finder (HSF)	130
<b>APÉNDICE G.</b> Tabla 20. Análisis de la variante rs7796976 [-459A>G] mediante la herramienta HaploReg V4.1: unión de proteínas en experimentos ChIP-Seq (ENCODE Project Consortium)	131
<b>APÉNDICE H.</b> Tabla 21. Análisis de la variante rs7796976 [-459A>G] mediante la herramienta HaploReg V4.1: motivos regulatorios alterados (Position Weight Matrix)	132
<b>APÉNDICE I.</b> Tabla 24. Análisis in silico de la variante rs2074113 [c.908+33G>T] mediante la herramienta Human Splicing Finder (HSF)	132
<b>APÉNDICE J.</b> Tabla 25. Análisis de la variante rs2074113 [c.908+33G>T] mediante la herramienta HaploReg V4.1: motivos regulatorios alterados (Position Weight Matrix)	133

VARIANTES GENÉTICAS Y SUSCEPTIBILIDAD AL DESARROLLO DE HNC	13
<b>APÉNDICE K.</b> Aval otorgado por el Comité de Ética de la Universidad Industrial de Santander (CEINCI) para el desarrollo de la presente investigación	134
<b>APÉNDICE L.</b> Aval otorgado por el Comité de Ética de la Universidad Cooperativa de Colombia, sede Pasto, para el desarrollo de la presente investigación	136
<b>APÉNDICE M.</b> Consentimiento informado de los individuos participantes del presente estudio.	137
<b>APÉNDICE N.</b> Instrumento de recolección de datos, tipo encuesta.	148

### Glosario

<b>ABREVIATURA</b>	<b>SIGNIFICADO</b>
ADN	Ácido desoxirribonucleico
<i>AHR</i>	Receptor Aril de hidrocarburos
<i>AHRR</i>	Represor del receptor Aril de hidrocarburo
AIP	Proteína interactiva del receptor Aril de hidrocarburo
ARNT	Translocador nuclear del <i>AHR</i>
B(A)P	Benzo(a)pireno
bHLH-PAS	Basic helix loop helix-PER-ARNT-SIM
BLAST	Basic Local Alignment Search Tool
COSMIC	The Catalogue of Somatic Mutations in Cancer
CYP5	Citocromo P450
EDP	Fenotipo extremo discordante
EPHX	Hidroxilasas epóxido microsomales
EBV	Virus de Epstein-Barr
FATHMM	Functional Analysis through Hidden Markov Models
GLOBOCAN	Observatorio Mundial del Cáncer
GnomAD	Genome Aggregation Database
GSTS	Glutation-S-transferasa
GWAS catalog	The Catalog of human genome-wide association studies
HGNC	HUGO Gene Nomenclature Committee
HNC	Cáncer de cabeza y cuello
HNSCC	Carcinoma escamo celular de cabeza y cuello
HAH	Hidrocarburos aromáticos halogenados
HSF	Human Splicing Finder

HSP-90	Proteína homodimérica de shock térmico de 90 kDa
IARC	International Agency for Research on Cancer
ICGC	The International Cancer Genome Consortium
NATS	N-acetiltransferasa
NF-κB	Factor nuclear proinflamatorio-κB
OMS	Organización Mundial de la Salud
P23	Prostaglandina E sintasa
PAHS	Hidrocarburos aromáticos policíclicos
Polyphen2	Polymorphism Phenotyping v2
Provean	Protein Variation Effect Analyzer
SIFT	Sorting Intolerant from Tolerant
SRC	Tirosina quinasa
TCDD	2,3,7,8-tetraclorodibenzo-p-dioxina
TCGA	The Cancer Genome Atlas
VADS	Vías aerodigestivas superiores
VEP	Ensembl Variant Effect Predictor
VPH	Virus del Papiloma Humano
XRE	Elementos sensibles a xenobióticos

## Resumen en español

**Título:** Identificación de variantes genéticas en el gen *AHR* que puedan contribuir a la susceptibilidad a cáncer de cabeza y cuello en población colombiana<sup>1</sup>

**Autor:** Nathalia Andrea Trujillo Pelayo<sup>2</sup>

**Palabras claves:** cáncer de cabeza y cuello, carcinoma escamocelular, gen *AHR*, consumo de tabaco, receptor Aril de Hidrocarburos, enzimas metabolizadoras de xenobióticos.

### Descripción:

El cáncer de cabeza y cuello (HNC) representa el séptimo cáncer más prevalente en el mundo. Colombia, para 2020, reportó 3.148 nuevos casos y 1343 muertes por esta neoplasia. El 90% de estos tumores son carcinomas de células escamosas (HNSCC), siendo su principal factor de riesgo el consumo de tabaco. El receptor Aril de hidrocarburos (AhR), codificado por el gen *AHR*, induce la activación de enzimas metabolizadoras de xenobióticos; es posible que diferencias en la actividad de estas enzimas, sean causadas por variantes genéticas sugiriendo que diferencias interindividuales podrían contribuir con susceptibilidad al cáncer. El objetivo fue identificar variantes en el gen *AHR* en pacientes con HNSCC y su posible efecto sobre la estructura y función de AhR.

Este estudio incluyó muestras de sangre de 23 pacientes con HNSCC y mediante secuenciación Sanger, se analizó la región codificante del gen *AHR* y sus límites intrón-exón. Se determinó la presencia y localización de variantes usando BLAST. Se realizó análisis in silico con herramientas de predicción funcional para establecer el posible efecto de las variantes sobre la estructura y función del receptor.

Se reporta para Colombia variantes genéticas en el gen *AHR* en pacientes con HNSCC. Se identificaron 7 alteraciones de secuencia, de las cuáles 3, no se encuentran reportadas en bases de datos y literatura científica. De las variantes detectadas, una corresponde a la región no traducida 5' (-459A>G, 5'UTR); 3 de tipo intrónicas (c.908+33G>T; c.253+175T>C; c.66-165delT); 2 de región codificante (c.132T>C; c.1661G>A) y una variante de la región no traducida 3' (c.2333C>T, 3'UTR). Sólo 2 de las variantes identificadas se encuentran reportadas en ClinVar como "benignas". Este estudio hace un primer acercamiento a la identificación de SNPs en el gen *AHR* en población colombiana y contribuye al entendimiento de la susceptibilidad genética al cáncer de cabeza y cuello en Colombia.

---

<sup>1</sup> Trabajo de Grado de Maestría

<sup>2</sup> Facultad de Salud, Departamento de Ciencias Básicas. Director: Laura Cifuentes C. Codirector: Clara Vargas C.

## Resumen en inglés

**Title:** Genetic variants in the *AHR* gene were identified and may contribute to susceptibility to head and neck cancer in Colombian population<sup>3</sup>

**Author:** Nathalia Andrea Trujillo Pelayo<sup>4</sup>

**Key words:** Squamous cell carcinoma of the head and neck, *AHR* gene, Tobacco consumption, Aryl Hydrocarbon Receptor, Xenobiotic metabolizing enzymes.

### Description:

Head and neck cancer (HNC) represents the 7<sup>th</sup> most common cancer by incidence worldwide. In Colombia, 3.148 new cases and 1343 deaths were reported in 2020 (Sung et al, 2020). The 90% of these tumors are squamous cell carcinomas (HNSCC) and the strongest risk factor for developing HNSCC is tobacco consumption (Johnson et al, 2020). The Aryl hydrocarbon receptor (AhR), encoded by the *AHR* gene, induces the activation of xenobiotic-metabolizing enzymes. Differences in the activity of these enzymes may be caused by genetic variants in *AHR*, which suggest the contribution of interindividual differences to cancer susceptibility. We aimed to identify genetic variants in the *AHR* gene in Colombian smoker patients with HNSCC and their possible effect on the receptor structure and function.

The study included blood samples from 23 HNSCC patients. The *AHR* gene coding sequence and its exon–intron boundaries were sequenced through Sanger’s method. We determined the presence and localization of variants using BLAST. In silico analysis was performed with functional prediction tools to establish the possible effects of the variants on the protein structure and function.

We reported for Colombian population genetic variants in the *AHR* gene in patients with HNSCC. A total of 7 sequence alterations were identified. However, 3 variants are not reported in genomic databases and not found in the scientific literature. The variants detected in *AHR* corresponded to a 5' prime UTR variant (-459A>G, 5'UTR); 3 intronic type alterations (c.908+33G>T; c.253+175T>C; c.66-165delT); 2 coding region variants (c.132T>C; c.1661G>A) and 1 variant located in the 3' prime UTR region (c.2333C>T, 3'UTR). We found 2 variants reported in ClinVar as "benign". The findings represent a first approach to identify SNPs in the *AHR* gene in Colombian population. The study contributes to the understanding of genetic susceptibility to HNC in Colombia.

---

<sup>3</sup> Master’s thesis

<sup>4</sup> Facultad de Salud, Departamento de Ciencias Básicas. Director: Laura Cifuentes C. Codirector: Clara Vargas C.

## Introducción

El cáncer de cabeza y cuello (HNC), que en su mayoría se presenta como carcinoma de células escamosas de cabeza y cuello (HNSCC), corresponde a una neoplasia maligna que compromete principalmente las estructuras de cavidad oral, faringe y laringe, y en conjunto representa el séptimo tipo de cáncer más prevalente y una de las causas de muerte más frecuente en el mundo (Stelow y Mills, 2005; Bray et al, 2018; Sung et al, 2021). De acuerdo con la “International Agency for Research on Cancer” (IARC GLOBOCAN), para el año 2020, se registraron 931.931 nuevos casos de HNC y 467.125 muertes asociadas con esta patología en el mundo, originándose principalmente en países en vía de desarrollo (Sung et al, 2021). Para América del Sur y Centro América, para este mismo año, fueron reportados 51.694 nuevos casos de HNC y 25.521 muertes por esta neoplasia (Sung et al, 2021). A su vez, Colombia registró 3.148 casos nuevos de HNC en la población, con al menos 1.343 muertes atribuidas a este tipo de cáncer (Instituto Nacional de Cancerología, 2018; Bray et al, 2018; Sung et al, 2021).

Actualmente, se estima que entre el 70% y 90% de casos de HNSCC están relacionados con consumo de tabaco en cualquiera de sus formas (fumadores pasivos y fumadores activos), atribuyendo su desarrollo principalmente al tiempo de exposición a los xenobióticos contenidos en el cigarrillo (Shopland, 1995; Klein and Grandis, 2010; United States Department of Health and Human Services, 2014; Fernández 2019; Johnson et al, 2020). Más de 60 carcinógenos

producto del tabaco han sido identificados, incluidos varios ligandos del receptor Aril de hidrocarburos (AhR) como el benzo(a)pireno (B(a)P) y el 2,3,7,8-tetraclorodibenzo-p-dioxina (TCDD) (Hecht, 2003; Kasai et al, 2006). Así mismo, diferentes estudios han determinado que los hidrocarburos aromáticos policíclicos (HAP) también desempeñan un papel en el HNSCC alterando el destino celular, promoviendo la inmortalización celular y contribuyendo a la carcinogénesis mediante la activación crónica de la vía de señalización del receptor *AHR* (Zhang et al, 2007; Bohonowych y Denison, 2007; Stanford et al, 2016; Yang et al. 2019). En Colombia, alrededor del 9.8% de la población nacional entre 18 y 69 años es consumidora de tabaco y al menos 34.800 muertes al año son atribuibles a este consumo (Instituto Nacional de Cancerología, 2019; Ministerio de salud y Protección Social, 2019).

A pesar de que los factores ambientales contribuyen de forma importante a la carcinogénesis (Yang et al. 2019), el riesgo de cáncer también está fuertemente influenciado por diferencias genéticas que conllevan a variaciones en la susceptibilidad. Varios de los genes codificadores de las proteínas citocromo P450 (CYPs), las hidroxilasas epóxido microsomales (EPHX), la glutatión-S-transferasa (GSTs) y la N-acetiltransferasa (NATs) implicadas en el metabolismo de carcinógenos son polimórficos, lo cual podría estar relacionado con las diferencias en la susceptibilidad individual al desarrollo de cáncer (Perera y Weinstein, 2000; Miller et al, 2002; Norppa, 2003; Fernández et al. 2019).

Diversos estudios han reportado diferencias interindividuales en la inducibilidad por hidrocarburos aromáticos de la enzima del metabolismo xenobiótico CYP1A1 (Fukushima et al, 2004; Chuang et al. 2012; Brunotto et al. 2014). La expresión del gen de la citocromo P450 1A1

(CYP1A1) es casi totalmente dependiente de la actividad del receptor AhR y está altamente inducida por su activación (Murray, Patterson y Perdew, 2014; Yang et al. 2019), por lo que, ¿es posible que diferencias en la actividad transcripcional de la enzima CYP1A1 atribuidas al receptor, puedan ser causadas por polimorfismos genéticos en el gen *AHR*, que codifica para AhR, afectando la inducibilidad de genes diana que codifican para proteínas implicadas en vías regulatorias críticas para la detoxificación de productos derivados del humo del tabaco?

Como ya se ha descrito, el consumo de tabaco representa uno de los principales factores de riesgo para el desarrollo de HNC. Sin embargo, no todas las personas expuestas a los carcinógenos del tabaco desarrollan cáncer a lo largo de su vida (Sellers, 1992); lo cual sugiere que la susceptibilidad a esta patología puede deberse a las variaciones heredadas en la habilidad de metabolizar procarcinógenos del humo del tabaco (McKay et al, 2011).

Estudios previos en Colombia, han analizado polimorfismos de riesgo en los genes CYP1A1, CYP2E1, GSTM1 y GSTT1, encontrando una frecuencia importante en la población de algunas variantes asociadas con aumento en el riesgo de desarrollar cánceres ligados al consumo de tabaco (Torres et al. 2004; Cajas, Hoyos, Sierra, 2005; Ramírez, Rangel y Escobar, 2009; Castaño-Molina et al. 2009, 2010; Paz-Egas et al. 2016). Sin embargo, en población colombiana no existen reportes acerca de la caracterización del gen *AHR* como marcador de susceptibilidad a cáncer de cabeza y cuello.

De esta manera, se hace fundamental determinar si existen diferencias genéticas que puedan afectar la producción de mayor o menor cantidad de metabolitos durante las vías de detoxificación y posiblemente sean de importancia en el aumento de susceptibilidad al desarrollo de cáncer de cabeza y cuello (Garte et al, 2001). Por lo anterior, el propósito de la

presente investigación fue identificar variantes en el gen *AHR* que pudieran contribuir a futuros estudios de susceptibilidad al desarrollo de cáncer de cabeza y cuello, en presencia de tabaquismo. Estos resultados facilitarán la comprensión de la etiología del cáncer de cabeza y cuello en Colombia.

## 1. Objetivos

### 1.1 Objetivo General

Identificar variantes genéticas en el gen *AHR* que puedan contribuir a la susceptibilidad al desarrollo de cáncer de cabeza y cuello.

### 1.2 Objetivos Específicos

1. Describir las variantes genéticas reportadas para el gen *AHR* en repositorios de datos genómicos de pacientes con cáncer.
2. Identificar las variantes genéticas germinales presentes en el gen *AHR* en pacientes colombianos con carcinoma escamocelular de cabeza y cuello.
3. Evaluar el efecto de las variantes identificadas sobre la función del receptor Aril hidrocarburos mediante análisis de predicción *in silico*.

## 2. Marco Referencial

### 2.1 Marco Teórico

Durante el siglo XXI, el cáncer se ha constituido como una barrera en el aumento de expectativa de vida en todos los países del mundo. Según estimaciones de la Organización Mundial de la Salud (OMS) en 2019, el cáncer fue la primera o segunda causa de muerte antes de los 70 años en 112 de 183 países (Sung et al, 2021).

Por su parte, el tabaquismo es responsable de al menos el 30% de todas las muertes por cáncer, aumentando el riesgo de cáncer de cavidad oral, faringe, laringe, pulmón, páncreas, útero, cuello uterino, ovario (mucinoso), riñón, vejiga, estómago, colon y recto (Secretan et al, 2009); constituyéndose en un problema de salud pública dada la persistencia de los factores de riesgo en la población. Por lo anterior, la capacidad de control de este tipo de cánceres depende principalmente de dos aspectos fundamentales, prevención y control, es decir, el manejo de los factores de riesgo y el diagnóstico precoz (Rocha, 2009).

La región de cabeza y cuello está constituida por un conjunto de órganos y cavidades que con base en su anatomía y fisiología pueden ser agrupadas bajo el nombre de vías aerodigestivas superiores (VADS), las cuales se conforman de una región digestiva que incluye la cavidad oral,

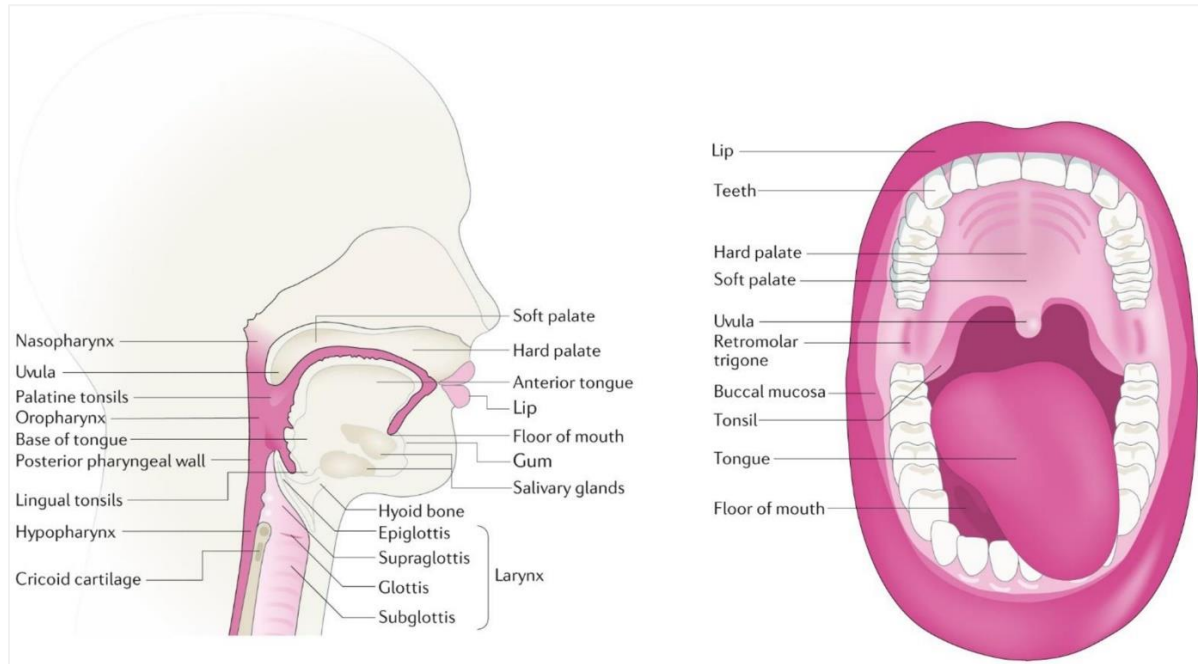
orofaringe, hipofaringe, esfínter esofágico superior, esófago cervical, conductos excretores de glándulas salivales y oído medio; y una región respiratoria formada por las fosas nasales, senos paranasales, nasofaringe, laringe y tráquea (Gallegos, 2006; Cárcamo, 2018) [Fig. 1 Esquema de cabeza y cuello]. Las VADS se encuentran recubiertas por una mucosa donde se originan el 85% de los cánceres de esta región, de los cuales cerca del 90% representan carcinomas de tipo escamo celular (HNSCC), el 15% restante constituye tumores en glándula tiroides, piel, huesos del esqueleto facial, cartílagos y partes blandas (Cárcamo, 2018). En conjunto, el HNC es el séptimo cáncer más común por incidencia en el mundo (Sung et al, 2021).

### ***2.1.1 Carcinoma escamocelular de cabeza y cuello***

**2.1.1.1 Definición.** Los carcinomas de células escamosas de cabeza y cuello (HNSCC, por sus siglas en inglés) son tumores malignos que surgen a partir de las células escamosas planas que componen la delgada capa de tejido que reviste las superficies húmedas y mucosas del interior de las estructuras de la cabeza y el cuello, entre estas, la cavidad oral (encía, paladar duro, lengua y suelo de boca), faringe y laringe (Fig. 1 Esquema de cabeza y cuello). En general, HNSCC representa la neoplasia más común de la región de cabeza y cuello, y su desarrollo se ha correlacionado con la exposición a carcinógenos derivados del tabaco, el consumo excesivo de alcohol o ambos (Instituto Nacional de Cancerología, 2019; Johnson et al. 2020).

**Figura 1.**

*Sitios anatómicos del desarrollo de HNSCC.*



*Nota:* El carcinoma de células escamosas de cabeza y cuello (HNSCC) surge del epitelio mucoso de la cavidad oral (labios, mucosa bucal, paladar duro, lengua anterior, piso de la boca y triángulo retromolar), nasofaringe, orofaringe (amígdalas palatinas, amígdalas linguales, base de lengua, paladar blando, úvula y pared faríngea posterior), hipofaringe (la parte inferior de la garganta, que se extiende desde el hueso hioides hasta el cartílago cricoides) y laringe. Los HNSCC asociados con el Virus del Papiloma Humano (VPH), surgen principalmente de las amígdalas palatinas y linguales de la orofaringe, mientras que los HNSCC asociados con el tabaco surgen principalmente en la cavidad oral, la hipofaringe y la laringe. Tomado de Johnson et al. 2020.

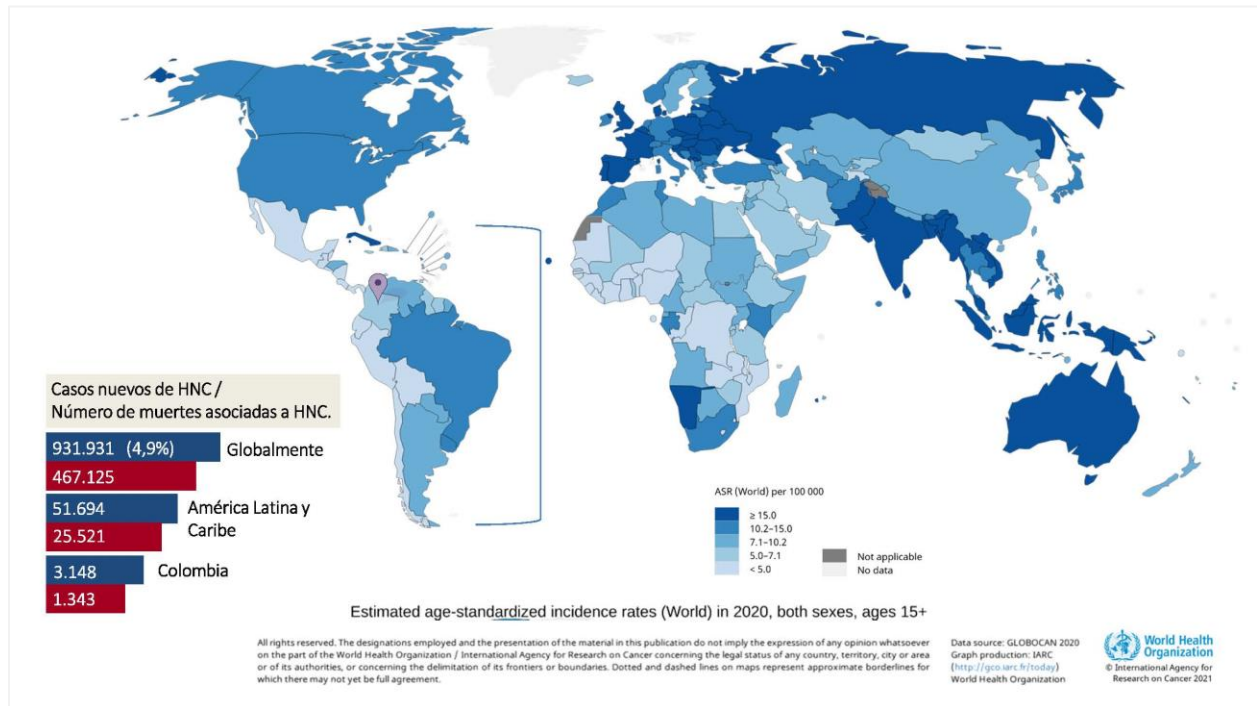
**2.1.1.2 Epidemiología.** El cáncer de cabeza y cuello representa en conjunto el séptimo tipo de cáncer más prevalente en el mundo con 931.931 nuevos casos y 467.125 muertes en 2020 (Sung et al, 2021) (Fig. 2). De acuerdo con el Observatorio Mundial del Cáncer (GLOBOCAN) la incidencia de HNSCC sigue incrementando y se prevé que para el año 2030 aumente en un 30%,

es decir, alrededor de 1.08 millones de casos nuevos al año (Bray et al, 2018; Ferlay et al, 2019; Ferlay et al, 2020).

En general, la incidencia de HNSCC es más elevada en hombres que en mujeres, es decir, los hombres presentan entre dos y cuatro veces más riesgo de desarrollar la enfermedad. Por otra parte, la edad media de diagnóstico de HNSCC no asociado con virus es de 66 años, mientras que la edad media de diagnóstico de cáncer de orofaringe asociado con VPH y cáncer de nasofaringe asociado con el virus de Epstein-Barr (EBV) es de ~53 años y ~50 años, respectivamente (Johnson et al. 2020, p. 2.).

**Figura 2.**

*Tasa de incidencia de HNC en el mundo en 2020.*



*Nota:* tasa de incidencia estandarizada estimada por edad en el mundo para el año 2020, ambos géneros, mayores de 15 años. En los cuadrantes azules se observa el número de nuevos casos de HNC reportados, mientras que en los cuadrantes rojos se presenta el número de muertes registradas por esta patología en el

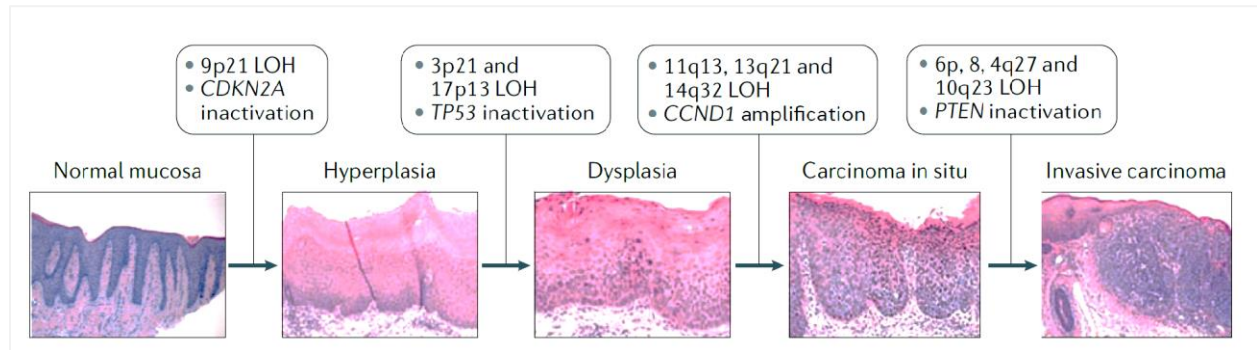
mundo, América Latina - el Caribe y Colombia, respectivamente. Tomado de GLOBOCAN 2020, gráfica adaptada de: IARC (<http://gco.iarc.fr/today>) World Health Organization.

**2.1.1.3 Progresión oncogénica del HNSCC.** La progresión histológica del HNSCC es la consecuencia de una serie de alteraciones genéticas que promueven la malignización, dotando a las células epiteliales de ventajas proliferativas e invasivas (Roesch et al, 2007). Estos eventos oncogénicos que acontecen en el HNSCC consisten principalmente en aberraciones cromosómicas, activación de oncogenes e inactivación de genes supresores de tumores (González, Gil y Ruiz, 2008; Johnson et al. 2020).

Uno de los eventos iniciales que conllevan a la carcinogénesis oral corresponden a lesiones leucoplásicas y alteraciones genéticas precoces (generalmente inactivación de genes supresores de tumores, como *CDKN2A* y *TP53*, que codifican p16INK4A y p53, respectivamente), que involucran la adquisición de una ventaja proliferativa de la célula alterada sobre el resto de las células del epitelio de la mucosa oral, generándose un campo precanceroso expansivo, fases de hiperplasia a displasia (leve, moderada y grave), que va progresivamente reemplazando al epitelio oral, debido a la pérdida de las regiones 3p21 y 17p13, de TP53; posteriormente, se presenta una transición de displasia a carcinoma in situ que implica la pérdida de las regiones 11q13, 13q21 y 14q32; en última instancia la progresión a carcinoma invasivo, esta capacidad invasiva adquirida, es probablemente el resultado de diversos eventos oncogénicos tardíos, como inactivación de *PTEN*, pérdidas de los loci cromosómicos 4q27, 6p, 8p, 10q23, 18q y otras alteraciones genómicas relacionadas con la pérdida de expresión y función de las moléculas de adhesión (Roesch et al, 2007; Olivier, Hollstein y Hainaut, 2010; Squarize et al, 2013; Johnson et al. 2020) (Fig. 3).

**Figura 3.**

*Progresión oncogénica del HNSCC*



*Nota:* modelo de orden histológico de la progresión de HNSCC. La hiperplasia de células epiteliales de la mucosa es seguida por displasia, y el carcinoma in situ precede al desarrollo de carcinoma invasivo. Se indican eventos genéticos específicos durante cada etapa de progresión. Tomada de Johnson et al. 2020. Imágenes de histopatología de hiperplasia, displasia, carcinoma in situ y carcinoma invasivo de Springer Nature Limited. Imagen histopatológica de mucosa normal de R. Jordan, Universidad de California San Francisco.

**2.1.1.4 Factores de riesgo y etiología del HNSCC.** El desarrollo de HNSCC es un proceso multifactorial asociado a una diversa gama de factores de riesgo los cuales incluyen el consumo de tabaco, el consumo de alcohol, la exposición a contaminantes ambientales, factores genéticos y la infección con agentes virales (VPH y EBV) (Hashibe et al, 2007; IARC 2020; Johnson et al. 2020).

**HNSCC - VPH positivo.** La infección por Virus del Papiloma Humano (VPH) es un factor de riesgo etiológico conocido para HNSCC y representa el 70% de los cánceres orofaríngeos (Isayeva, Maswahu y Brandwein, 2012; Stein et al, 2015), siendo el VPH-16 el tipo causal

primario. Sin embargo, otros VPH de alto riesgo (VPH-18, VPH-31, VPH-33 y VPH-52) pueden detectarse en un porcentaje menor de pacientes (Michaud et al, 2014).

El genoma de este virus, que logra integrarse al genoma del huésped, consta de varios marcos de lectura abierta, siete genes tempranos (E1-E7) y dos tardíos (L1 y L2). Estos últimos codifican proteínas estructurales de la cápside viral, mientras que los genes E1-E5 codifican proteínas involucradas en la replicación y regulación viral y los genes E6 y E7 codifican oncoproteínas virales esenciales para la transformación oncogénica de la célula huésped (Johnson et al. 2020). La proteína E6 promueve la ubiquitinación y degradación proteasomal de p53 (Scheffner et al, 1993). Por otra parte, la proteína E7 se une a la proteína del retinoblastoma (RB1), reguladora del ciclo celular, promoviendo la inactivación de RB1 y la liberación de factores de transcripción E2F (Dyson et al, 1989).

La consecuencia molecular de la expresión de estas oncoproteínas virales es la entrada al ciclo celular (regulación positiva de p16INK4A) y la inhibición de la apoptosis mediada por p53, que permite la replicación del virus. Además de E6 y E7, E5 también favorece la capacidad carcinogénica al impulsar la progresión del ciclo celular (Venuti et al, 2011; Estevao et al, 2019).

Finalmente, la incidencia de infección por VPH en la cavidad oral y orofaríngea se ha establecido en alrededor de 7%, aumenta de acuerdo con la edad, número de parejas sexuales, la actividad sexual de cualquier tipo, y es mayor en hombres que en mujeres (Montero, 2018). Generalmente, la detección de la expresión de p16INK4A permite la clasificación de tumores orofaríngeos como VPH positivos (Johnson et al. 2020).

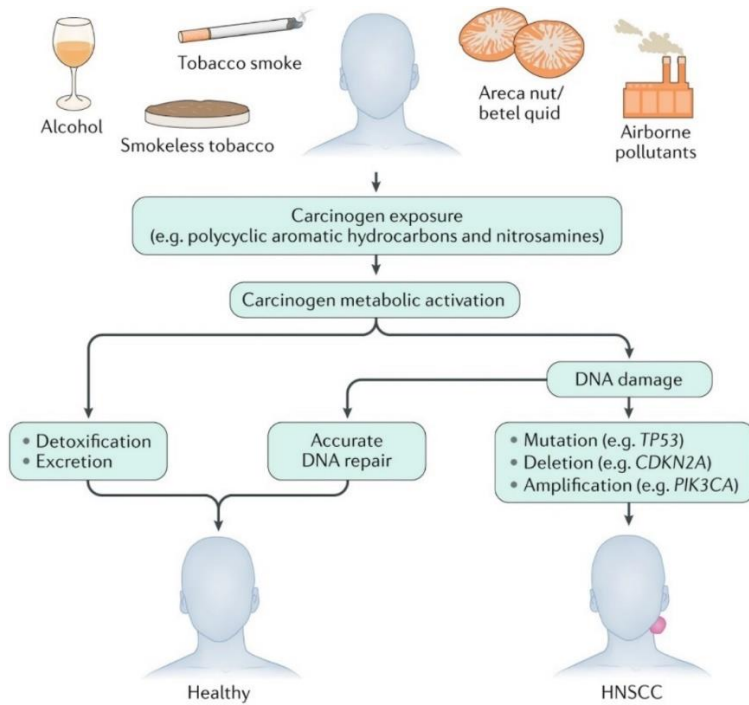
***HNSCC - VPH negativo.*** Varios factores de riesgo asociados al HNSCC VPH negativo muestran una prevalencia geográfica, cultural y/o habitual, donde el consumo de tabaco representa el principal factor de riesgo para el desarrollo de esta neoplasia. De acuerdo con literatura reportada, el tabaco consta de más de 5000 sustancias químicas diferentes, de las cuales docenas de ellas tienen actividad cancerígena (Johnson et al. 2020). Se cree que los hidrocarburos aromáticos policíclicos (PAH, por sus siglas en inglés), como el benzo(a)pireno, las dioxinas y las nitrosaminas, como la 4-(metilnitrosamina)-1-(3-piridil)-1-butanona (NNK) y la N-nitrosornicotina (NNN), son las sustancias químicas más responsables de los efectos cancerígenos del tabaco (Hoffmann y Hoffmann, 1997; Hecht, 1999; Hecht, 2003; Warnakulasuriya y Straif, 2018; Johnson et al. 2020).

La activación metabólica de agentes cancerígenos derivados del tabaco se da mediante enzimas de detoxificación y vías que facilitan su remoción. Sin embargo, muchos de los metabolitos derivados de los ligandos del tabaco (PAHs y nitrosaminas) también pueden formar aductos covalentes de ADN que, de no ser reparados, pueden originar variantes genéticas y otras alteraciones (desde mutaciones puntuales hasta reordenamientos cromosómicos) que pueden afectar genes de control de la división, proliferación y crecimiento celular, contribuyendo a la carcinogénesis, la cual probablemente depende del equilibrio entre la activación metabólica, la desintoxicación y la reparación del ADN (Johnson et al. 2020). A su vez, la inflamación de tejidos expuestos por el uso de productos de tabaco, y la producción local de citocinas, quimiocinas y factores de crecimiento también podrían ejercer un rol fundamental en la promoción de proliferación celular, angiogénesis y, finalmente, el desarrollo de cáncer (Rubín et al, 1998; Zhu et al, 2013; Johnson et al. 2020).

Además del consumo de tabaco, la ingesta excesiva de alcohol es otro factor de riesgo para el desarrollo de HNSCC, a saber, el alcohol es metabolizado a acetaldehído favoreciendo la formación de aductos de ADN (Brooks y Theruvathu, 2005). También se ha determinado que el alcohol podría servir como solvente para carcinógenos, aumentando la exposición de las células epiteliales a estas sustancias (Pai y Westra, 2009), por lo que el efecto sinérgico del consumo de alcohol y tabaco incrementa el riesgo, 35 veces más, a la carcinogénesis (Talamini et al, 2002; Johnson et al. 2020) (Fig.4).

**Figura 4.**

*Desarrollo de HNSCC - VPH negativo*



*Nota:* factores en el desarrollo de HNSCC – VPH negativo. Exposición a carcinógenos como PAH, dioxinas y nitrosaminas están asociados con un fuerte aumento del riesgo de HNSCC. La activación de enzimas metabolizadoras de xenobióticos regula la detoxificación y excreción de metabolitos reactivos evitando la generación de aductos de ADN. La acumulación de alteraciones en genes supresores de tumores como TP53

y CDKN2A o vías de señalización (PI3K–AKT–mTOR y RAS–genes de la ruta MAPK) se asocia con el inicio, la progresión y el mal pronóstico de HNSCC - VPH negativo (Johnson et al. 2020).

Por otra parte, la exposición a contaminantes ambientales cancerígenos es un factor de riesgo para HNSCC, especialmente en aquellas regiones geográficas, donde la polución es cada vez mayor (Mishra, y Meherotra, 2014; Wong y Lui, 2014). En poblaciones asiáticas, el desarrollo de cáncer de células escamosas de la cavidad oral se asocia con la masticación de la nuez de Areca o nuez de betel, incluso mezclas que incluyen la nuez de areca, hoja de betel, cal y/o tabaco, de acuerdo con las costumbres de la población (Zhang et al, 2018; IARC, 2020). Otros factores de riesgo incluyen el envejecimiento, la mala higiene bucal y las dietas carentes de vegetales (Guha et al, 2007; Freedman et al, 2008; Mazul et al, 2018).

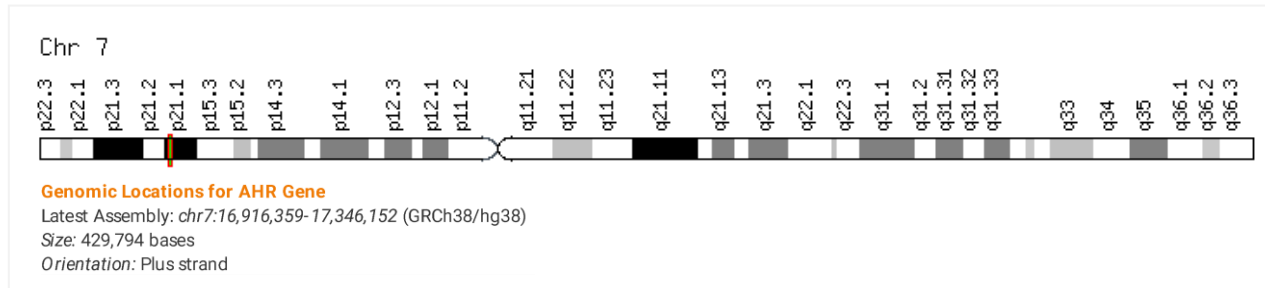
Finalmente, factores genéticos también podrían contribuir al riesgo de HNSCC. Diversos estudios han evaluado la asociación de variantes en genes implicados en el metabolismo de carcinógenos (polimorfismos en la citocromo P450 1A1: CYP1A1; GSTM1; Ile462Val) y en la inmunidad (polimorfismos en CTLA4: rs231775 y rs4553808; IL10: 1082A>G), encontrando un mayor riesgo asociado al cáncer de cabeza y cuello (Cadoni et al, 2012; Niu et al, 2015; Wang et al, 2016; Fang et al, 2018). En consecuencia, una capacidad reducida para metabolizar sustancias cancerígenas y una inmunidad alterada pueden contribuir al desarrollo de HNSCC.

### ***2.1.2 Gen AHR – consumo de tabaco – desarrollo de HNSCC***

Estudios de epidemiología genética (Negri et al, 2009) y estudios familiares (Delahaye et al, 2015), proveen evidencia de un importante papel de la susceptibilidad genética en la patogénesis de cánceres de tracto aerodigestivo superior asociados al consumo de tabaco (Harris, 1989).

La mayoría de los carcinógenos del tabaco se metabolizan en el organismo por vía de mecanismos enzimáticos complejos que implican reacciones de activación y de detoxificación (Johnson et al. 2020). La activación es realizada por enzimas de fase I, y consiste principalmente en reacciones de oxidación catalizadas por las enzimas del citocromo P450 (CYPs) o por hidroxilasas epóxido microsomales (EPHX) (Lanciego, 2013). La detoxificación (desactivación) es realizada por enzimas de fase II, especialmente enzimas de la glutatión-S-transferasa (GSTs) y de la N-acetiltransferasa (NATs), las cuales favorecen la eliminación de reactivos intermedios conjugándolos con moléculas endógenas (Lanciego, 2013). Los hidrocarburos aromáticos policíclicos (PAHs), uno de los principales derivados de la combustión del cigarrillo, estimulan la actividad de CYP1A1 por activación de la transcripción del gen CYP1A1, esta inducción del gen esta mediada por la activación del receptor Aril de Hidrocarburos (AhR), codificado por el gen *AHR* (Mimura y Kuriyama, 2003).

**2.1.2.1 Gen *AHR*.** Se encuentra localizado en el brazo corto del cromosoma 7 en la posición 21 (Fig. 5) y codifica para el Receptor Aril de hidrocarburos (AhR), implicado en la regulación de las respuestas biológicas a los hidrocarburos aromáticos planos (GeneCards, the human gene database). De acuerdo con las anotaciones de Gene Ontology (GO), el gen *AHR* presenta actividades de: (i) factor de transcripción de unión al ADN, permitiendo que las células puedan adaptarse a condiciones variables a partir de la detección de compuestos del medio ambiente, alimentación, microbioma y metabolismo celular, y que presenta un rol fundamental en el desarrollo, la inmunidad y el cáncer (Ema et al 1994; MacPherson et al, 2013; Gomez et al, 2018; Sadik et al, 2020); (ii) heterodimerización de proteínas (GeneCards, the human gene database).

**Figura 5.***Representación de la localización genómica del gen AHR*

*Nota:* Gen *AHR* en localización genómica: 7p21.1 según HGNC, Entrez Gene, Ensembl. Bandas citogenéticas de acuerdo con Ensembl, localizaciones según GeneLoc (y/o Entrez Gene, si son diferentes). Tomado de GeneCards, the human gene database.

**2.1.2.2 Receptor Aril de hidrocarburos (AhR).** El receptor de hidrocarburos de arilo, o AhR, es un factor de transcripción nuclear perteneciente a la familia de los factores de transcripción con dominio bHLH-PAS (basic helix loop helix-PER-ARNT-SIM) (Fig. 6) (Hoffman et al, 1991; Hogenesch et al, 1997).

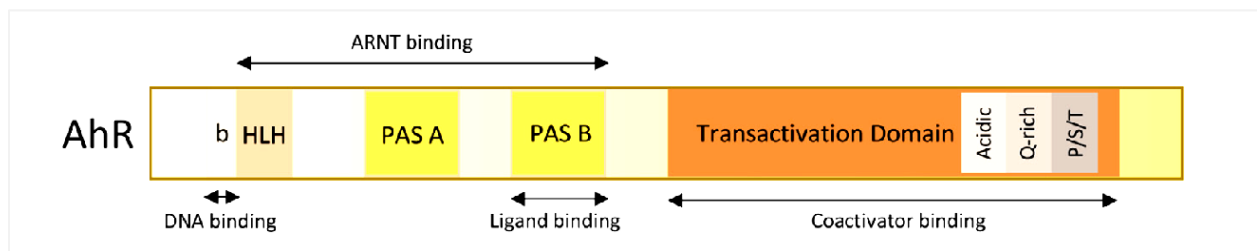
El AhR, conocido también como receptor de dioxinas, desempeña un papel fundamental en la regulación de los efectos tóxicos de contaminantes industriales como los PAHs mediante la inducción de la transcripción al unirse a elementos de respuesta xenobióticos (XRE) (Ema et al 1994; MacPherson et al, 2013; Murray, Patterson y Perdew, 2014; Gomez et al, 2018). Además, participa en una variedad de procesos biológicos, incluida la angiogénesis, la hematopoyesis, el metabolismo de fármacos y lípidos, la motilidad celular y la modulación inmunitaria (Puga, Xia y Elferink, 2002).

En mamíferos, el AhR presenta tres dominios funcionales:

- Un primer dominio altamente conservado bHLH que consiste de 4 a 6 aminoácidos en el extremo amino terminal, el cual permite la dimerización con el ARNT, la unión con el ADN y la interacción con la Hsp90 (proteína de shock térmico 90) (Vázquez, Rubio y Espinosa, 2016; Larigot et al, 2018).
- Un segundo dominio PAS que comprende dos repeticiones estructurales, PAS A y PAS B, cada uno conteniendo aproximadamente 110 aminoácidos. Este dominio, también se encuentra involucrado en la dimerización con ARNT (PAS A) y permite la unión del ligando (PAS B) (Hao et al, 2011; Wu et al, 2013).
- El tercer dominio, C-terminal presenta tres subdominios: (i) uno enriquecido con residuos ácidos (glutamato/aspartato), (ii) otro enriquecido con glutamina (Q-rich), (iii) un tercero enriquecido con serina, treonina y prolina (S/T/P). Coactivadores y correpresores interactúan con el AhR a través de este dominio, fundamental en la transactivación (Kewley, Whitelaw y Chapman, 2004; Larigot et al, 2018) (Fig. 6).

**Figura 6.**

*Dominios funcionales del receptor AhR (proteína bHLH-PAS).*

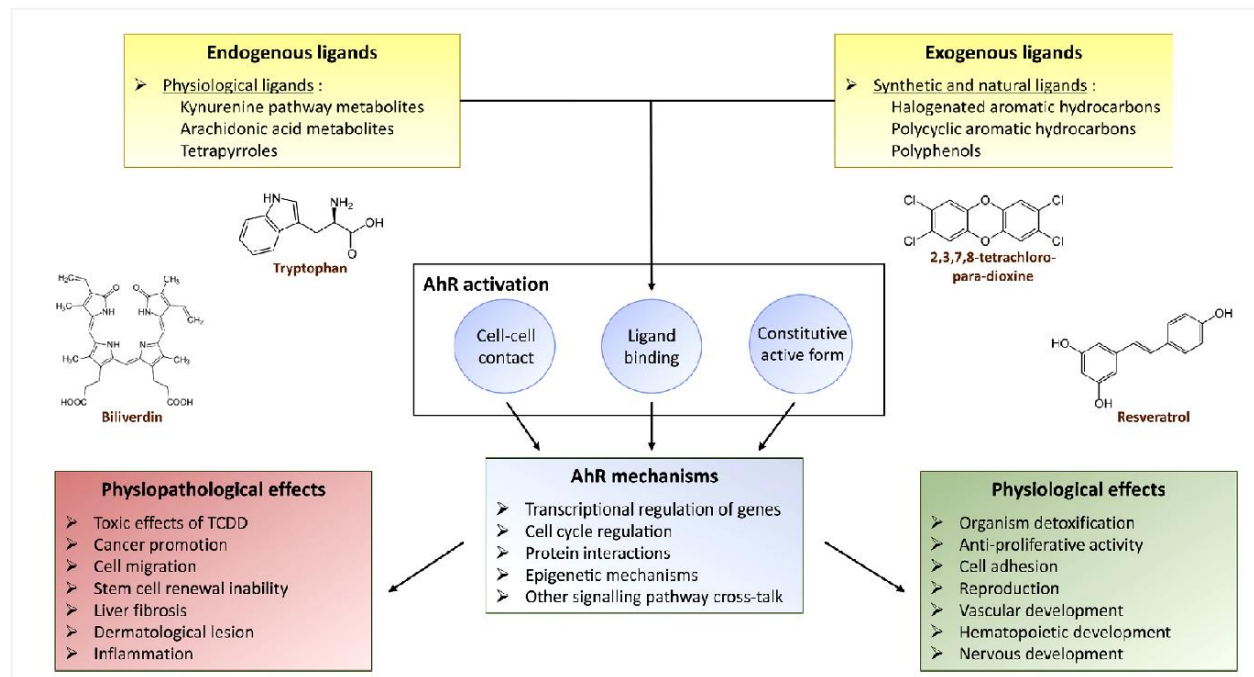


*Nota:* El AhR contiene 1) un dominio de unión al ADN (bHLH); 2) un dominio de dimerización PAS que comprende dos repeticiones estructurales A y B (de unión a ligandos); 3) un dominio de transactivación C-terminal que presenta tres subdominios (Larigot et al, 2018).

**2.1.2.3 Vías de señalización celular del AhR.** La activación del AhR implica la unión de un ligando. Diversos ligandos han sido descritos para el AhR (Larigot et al, 2018), principalmente xenobióticos, de naturaleza hidrofóbica y con estructura plana (Vázquez, Rubio y Espinosa, 2016). Estos ligandos pueden ser de tipo endógeno o exógenos, estos últimos, de origen natural como los metabolitos secundarios de plantas o producto de combustiones incompletas (PAH), o de origen sintético como los hidrocarburos aromáticos halogenados (HAH) o el 2,3,7,8-tetraclorodibenzop-dioxina (TCDD) (Wei et al, 1998; Adachi et al, 2001; Denison y Nagy, 2003; Larigot et al, 2018) (Fig. 7).

**Figura 7.**

*Ligandos AhR y función reguladora del receptor*



*Nota:* Relación funcional entre los ligandos AhR (endógenos y exógenos: de origen natural o sintético), funciones reguladoras del receptor, y sus efectos fisiológicos y fisiopatológicos. Tomado de: Larigot et al, 2018. Dado el rol fundamental del AhR en la fisiología y fisiopatología humana como se observa en la

imagen, los agonistas o antagonistas del receptor podrían desempeñarse como agentes de objetivo terapéutico para el cáncer (Yang et al, 2019).

En su vía canónica de señalización celular, el receptor AhR se encuentra de manera inactiva en el citosol formando un complejo multiproteico, con la proteína homodimérica HSP-90, la prostaglandina E sintasa (p23), con actividad chaperona frente a HSP-90, la tirosina quinasa SRC y la proteína moduladora de la actividad del receptor de Aril hidrocarburos (AIP) (Meyer et al, 1998). Tras su activación, a partir de una amplia gama de ligandos, el AhR se transloca al núcleo y se dimeriza con la proteína del translocador nuclear (ARNT). Finalmente, el complejo AhR-ARNT se une al ADN interactuado con los promotores de elementos sensibles a xenobióticos (XRE) (Ema et al 1994; MacPherson et al, 2013; Gomez et al, 2018), induciendo la activación transcripcional de genes diana, como el citocromo P450 (CYP450) que desempeña un papel fundamental en el metabolismo de contaminantes ambientales y compuestos aromáticos involucrados en la tumorigénesis, entre estos los carcinógenos del tabaco (Fig. 8a) (Zhou et al, 2009; Martano et al, 2018; Larigot et al, 2018; Rothhammer y Quintana, 2019).

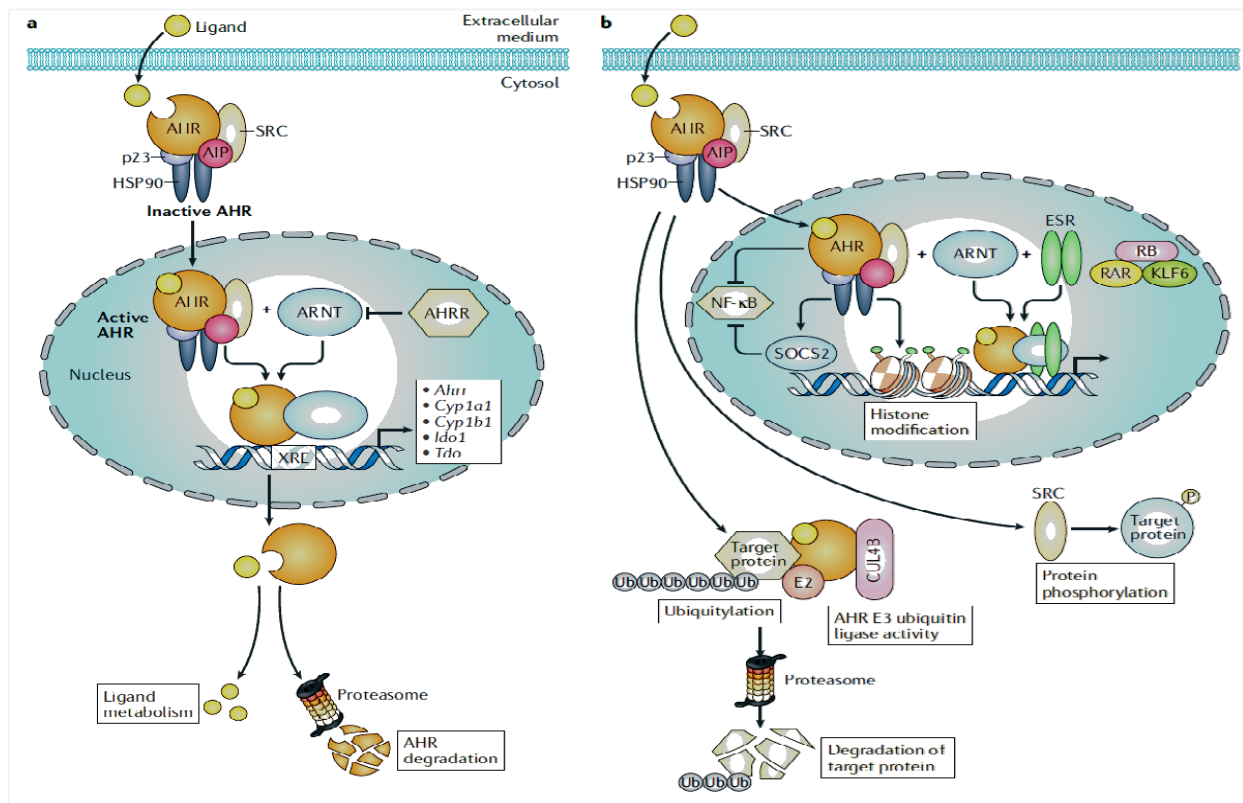
El AhR, además de su ruta clásica, presenta mecanismos alternativos de correlación con factores de transcripción adicionales favoreciendo su unión a sitios específicos en el ADN e interactuando de manera directa o indirecta con distintas vías de señalización, por ejemplo la ruta del factor nuclear proinflamatorio- $\kappa$ B (NF- $\kappa$ B) (Rothhammer y Quintana, 2019); A su vez, el receptor también se desempeña como una ubiquitina-ligasa E3, promoviendo la degradación enzimática de proteínas diana; incluso, la activación de AhR inducida por agonistas, genera cascadas de fosforilación de diversas proteínas diana, entre estas, la indolamina 2,3-dioxigenasa 1 (IDO1) (Bessede et al, 2014) y miembros de las vías del ácido araquidónico o de los leucotrienos

(Matsumura, 2009), tras la liberación de SRC del complejo multiproteico-AhR (Rothhammer y Quintana, 2019) (Fig. 8b).

Finalmente, la regulación de transcripción de genes dependientes de AhR está dada por la activación del AhRR (Represor del Receptor Aril de hidrocarburos) suprimiendo la actividad de AhR al unirse al ARNT y XRE (complejo AhRR-ARNT) (Hahn, 2002; Hahn, Allan y Sherr, 2009). Esta regulación negativa sobre el receptor y la degradación enzimática de la proteína evita sobreestimulación por agonistas favoreciendo un control temporal de la señalización.

**Figura 8.**

*Rutas de señalización del receptor Aril hidrocarburos (AhR).*



*Nota:* <sup>a</sup> Ruta clásica de señalización del AhR. El AhR inactivo forma un complejo con la proteína HSP90, AIP (proteína moduladora de la actividad del receptor de Aril hidrocarburos), p23 y SRC. Este complejo

mantiene inactivo el receptor y evita su degradación proteasomal favoreciendo el estado de afinidad por sus ligandos. Tras la unión del agonista, el complejo de chaperonas - AhR se traslada al núcleo, donde el receptor se une al ADN para controlar la expresión génica. <sup>b</sup>. Mecanismos adicionales de señalización del AhR. *AHRR*, represor AHR; *ARNT*, translocador nuclear AHR; *CUL4B*, complejo cullin4B -ubiquitina ligasa; *ESR*, receptor de estrógeno; *KLF6*, factor Krüppel tipo 6; *P*, fosforilación; *RAR*, receptor de ácido retinoico; *RB*, proteína de retinoblastoma; *SOCS2*, supresor de señalización de citoquinas 2; *Ub*, ubiquitinación; *XRE*, elemento de respuesta xenobiótica (Rothhammer y Quintana, 2019).

**2.1.2.4 Rol del AhR en la tumorigénesis.** Diversos estudios epidemiológicos y experimentales en animales sugieren que la disfunción del receptor AhR podría estar asociada con el desarrollo de cáncer (Yang et al, 2019). La expresión constitutivamente activa del receptor de dioxinas, durante su vía clásica, puede afectar la regulación de la actividad enzimática, conduciendo a la activación metabólica de carcinógenos (productos intermedios o finales del TCDD y el BaP), generando daño del ADN y favoreciendo la tumorigénesis (Romagnolo et al, 2016). Así mismo, el AhR también desempeña un rol importante durante los procesos de:

(i) proliferación: el AhR puede interactuar con algunos factores de crecimiento, como el factor de crecimiento de fibroblastos 9, para favorecer el crecimiento tumoral. Los promotores de estos factores presentan elementos de respuesta a AhR, al mismo tiempo que, como mitógenos, probablemente promuevan la proliferación de células cancerígenas (Wang et al. 2009; Ueng et al. 2010).

(ii) apoptosis: AhR media una amplia resistencia a la apoptosis, particularmente cuando es activado por su ligando TCDD, el cuál activa una serie de factores favorables a la supervivencia,

como EGFR, TGF $\alpha$ , ERK, PI3K, Akt y ETK quinasas, y al mismo tiempo inhibe la función de factores inductores de la apoptosis, como la escisión de PARP (Bekki et al, 2015). Además, el AhR es capaz de interferir con la activación de la vía de las caspasas tanto extrínseca (mediada por receptores) como intrínseca (mediada por mitocondrias), esta última, se ve afectada por la activación de AhR a través de la inducción de Bcl-2 y el reclutamiento de p53 (Bekki et al, 2015). Lo anterior puede relacionarse con la evolución de AhR como un factor favorable a la supervivencia contra el estrés ambiental, que a su vez puede favorecer la evasión del proceso de apoptosis inducido por estrés (Bekki et al, 2015).

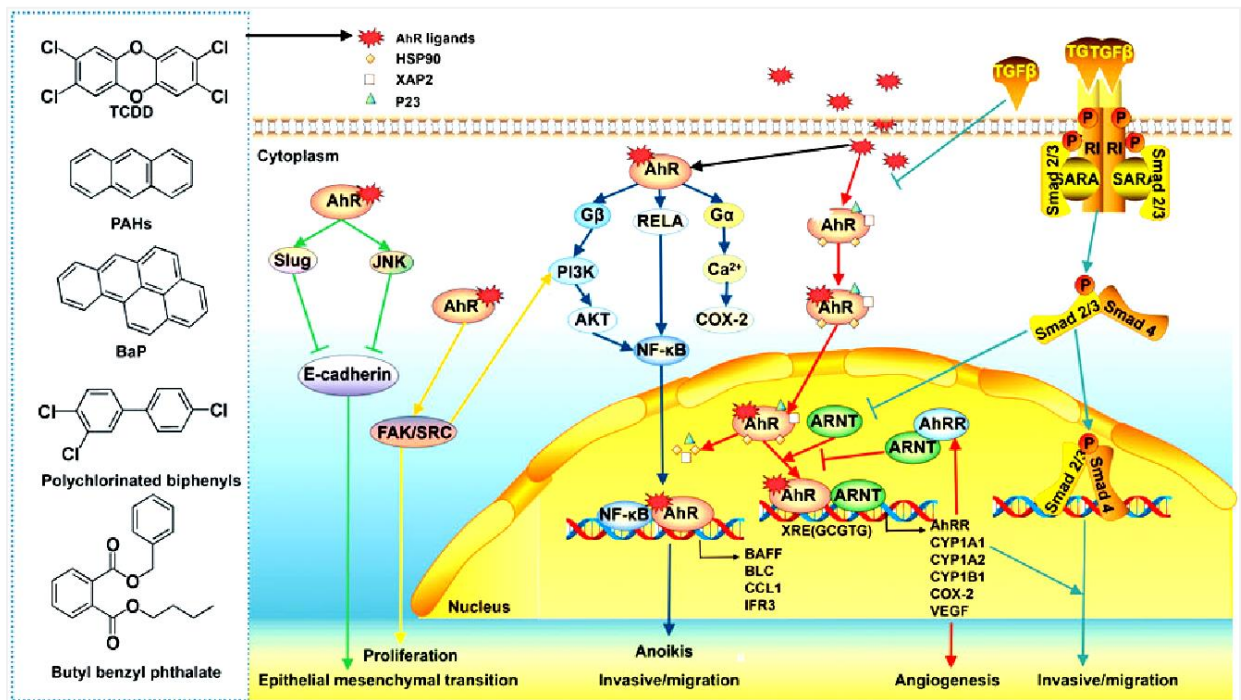
(iii) adhesión y migración celular (Feng et al, 2013; Tummala et al, 2014).

La participación de AhR en la adhesión y migración celular ha sido evaluada mediante estudios in vivo de tratamientos con xenobióticos (Larigot et al. 2018). La exposición de células humanas a TCDD estimula cambios morfológicos como la aparición de lamelipodios, que favorecen la migración celular (Larigot et al. 2018) y la reorganización del citoesqueleto por la redistribución de actina y la activación de las quinasas FAK y Src (Diry et al. 2006; Tomkiewicz et al, 2013; Wei et al, 2018). Conjuntamente, también se presentan cambios en la expresión génica como la activación de JNKs (Diry et al. 2006), PI3K/Akt (Popolo et al, 2017), TGF-  $\beta$  (Gagliani et al, 2015), NF- $\kappa$ B (Galle-Treger et al, 2016) y la regulación a la baja de E-cadherina, esta última, característica de la transición epitelial-mesenquimal (EMT) que se desencadena por factores de transcripción como Slug, un gen objetivo directo del receptor AhR (Ikuta y Kawajiri, 2006; Pierre et al. 2014) (Fig. 9).

Finalmente, el AhR también podría actuar como gen supresor de tumores bajo ciertas condiciones, por ejemplo, Iu et al, encontró que el receptor puede inhibir la inflamación, el estrés oxidativo y la apoptosis causada por el humo del tabaco mediante el mecanismo de regulación del gen ReIB, el cuál codifica para un miembro de la familia de factores de transcripción del factor NF-κB, sugiriendo la función anticancerígena del receptor (Iu et al, 2017). De acuerdo con lo anterior, se hace necesario profundizar sobre los posibles mecanismos celulares y/o moleculares que permitan determinar la contribución del AhR a la progresión o prevención del cáncer (Yang et al, 2019).

**Figura 9.**

*Vía de señalización de AhR.*



*Nota:* el AhR está involucrado en la regulación de una variedad de vías de transducción de señales, como JNK/Slug, FAK/Src, PI3K/Akt, NF-κB y TGF- β. JNK: Jun N-terminal kinase; FAK: Focal Adhesion

Kinase; PI3K: phosphatidylinositol 3-kinase; Akt: protein kinase; NF- $\kappa$ B: nuclear factor kappa-light-chain-enhancer of activated B cells; P: phosphorylation. Tomado de: Yang et al, 2019.

## **2.2 Estado del arte**

Este apartado se presenta como Apéndice A al presente trabajo. Artículo por someter: Polymorphisms Associated to Genetic Susceptibility and Head and Neck Cancer Risk: A Systematic Review.

## **2.3 Marco Legal**

Esta investigación no tuvo ninguna orientación comercial y durante su ejecución se preservaron los principios de beneficencia, no maleficencia, autonomía y justicia.

### ***2.3.1 Tratamiento de datos personales***

En cumplimiento de la Resolución N° 1227 de 2013 expedida por la Universidad Industrial de Santander, por la cual se aprueba el manual de procedimientos administrativos para el tratamiento de datos personales que está en concordancia con el Decreto 1377 de 2013 (Presidencia de la República de Colombia, Decreto 1377, 2013), derogado parcialmente por el Decreto 1081 de 2015 (Presidencia de la República de Colombia, Decreto 1081, 2015), por el cual se reglamenta parcialmente la Ley 1581 de 2012, ésta última, por la cual se dictan disposiciones generales para la protección de datos personales, se aplica el Derecho Fundamental contemplado en el artículo 15 de la Constitución Política de Colombia (Const., 1991), esto es la protección de los datos

personales que permite el reconocimiento, la actualización y rectificación de información de orden privado.

Conforme lo anterior, durante el desarrollo de esta investigación, todas las muestras del estudio fueron colectadas bajo consentimiento informado de los participantes, para uso de forma anónima y responsable. Los principios constitucionales que rigen el actuar investigativo del presente estudio corresponden a: la dignidad humana, el buen nombre, la intimidad y la privacidad, la legalidad, finalidad, libertad, veracidad o calidad, transparencia, de acceso y circulación restringida, seguridad y confidencialidad, de acuerdo con la protección de los datos personales de las personas participantes- Derecho de Habeas Data, bajo la normatividad correspondiente: La Constitución Política de Colombia de 1991 (Const., 1991), Resolución 8430 de 1993 (Ministerio de Salud, Resolución 8430, 1993), Ley estatutaria 1266 de 2008, Decreto Reglamentario 1727 de 2009 (Ministerio de Hacienda y Crédito Público, Decreto 1727, 2009), Decreto 2952 de 2010 (Presidencia de la República de Colombia, Decreto 2952, 2010), Ley Estatutaria 1581 de 2012, Decreto Reglamentario 1377 de 2013 (Presidencia de la República de Colombia, Decreto 1377, 2013).

### ***2.3.2 Clasificación del riesgo de la investigación***

Este estudio es de riesgo mínimo de acuerdo con la Resolución 8430 de 1993 del Ministerio de Salud de la República de Colombia. De modo general, esta investigación no representó un riesgo, moderado o máximo, para el investigador o las comunidades presentes en el rango geográfico de muestreo y no generó impactos negativos sobre la población. El manejo de residuos

de laboratorio se realizó según lo estipulado por el Decreto 351 de 2014 (Ministerio de Salud y Protección Social, Decreto -351, 2014).

Los investigadores del presente estudio se encuentran certificados en buenas prácticas clínicas según Resolución 2378 de 2008 del Ministerio de la Protección Social.

### 3. Metodología

#### **3.1 Fase I: descripción de las variantes genéticas reportadas para el gen *AHR* en repositorios de datos genómicos de pacientes con cáncer.**

Con el fin de describir variantes genéticas en el gen *AHR*, se realizó una compilación de las variantes reportadas para el gen mediante búsquedas en bases de datos genómicos de pacientes con cáncer, con especial interés en población latinoamericana.

##### ***3.1.1 Variantes en repositorios genómicos***

La identificación de variantes en el gen *AHR* se llevó a cabo mediante la exploración de tres repositorios de datos genómicos de cáncer: The Cancer Genome Atlas (TCGA) (<https://portal.gdc.cancer.gov/>), the International Cancer Genome Consortium (ICGC) (<http://icgc.org/>) y el Catálogo de Variantes Somáticas en Cáncer (COSMIC)

(<https://cancer.sanger.ac.uk/cosmic>); además de los portales The GWAS catalog (<https://www.ebi.ac.uk/gwas/>) and the Genome Aggregation Database (GnomAD) v2.1 (<https://gnomad.broadinstitute.org/>). Datos obtenidos de estudios de líneas celulares no fueron incluidos.

Para esta última base de datos, el proceso de búsqueda se dividió en dos pasos. En primer lugar, se realizó una exploración completa utilizando toda la base de datos de GnomAD v.2.1.1. A continuación, una segunda búsqueda involucró sólo variantes reportadas en el conjunto de datos denominado GnomAD v.2.1.1 (Non Cancer). Tras remover duplicados, se determinó un número total de variantes únicas por conjunto de datos, las cuáles fueron agrupadas como GnomAD “Cancer” y GnomAD “Non Cancer”.

### ***3.1.2 Integración de datos de repositorios, análisis y potencial impacto***

Una vez efectuadas las búsquedas e identificación de las variantes, se realizó la integración del conjunto de datos para cada repositorio mediante el software Excel, conteniendo información relevante de cada una de las variantes (Fig. 10). Una lista completa de variantes en *AHR* reportadas para cada base de datos (GnomAD, TCGA, ICGC, COSMIC y GWAS), antes del filtrado, se encuentra contenida en el Apéndice D del presente trabajo.

Después de filtrar y eliminar duplicados, permanecieron un número de variantes únicas de todo el conjunto de datos, las cuáles fueron analizadas de acuerdo con: tipo de consecuencia, sitio

primario reportado y score de impacto sobre la función del receptor AhR, para población en general y población latinoamericana.

**Figura 10.**

*Información recabada para cada una de las variantes identificadas, de acuerdo al repositorio genómico explorado.*



Nota: El gráfico representa los datos extraídos para las variantes reportadas en el gen *AHR* por cada base de datos estudiada. El conjunto final de datos se organizó en una matriz del SW Excel, para su posterior análisis. Gráfico de autoría propia, imagotipos tomados de: <https://gnomad.broadinstitute.org/>; <https://cancer.sanger.ac.uk/cosmic/>; <http://icgc.org/>; <https://www.ebi.ac.uk/gwas/>; <https://portal.gdc.cancer.gov/>, respectivamente.

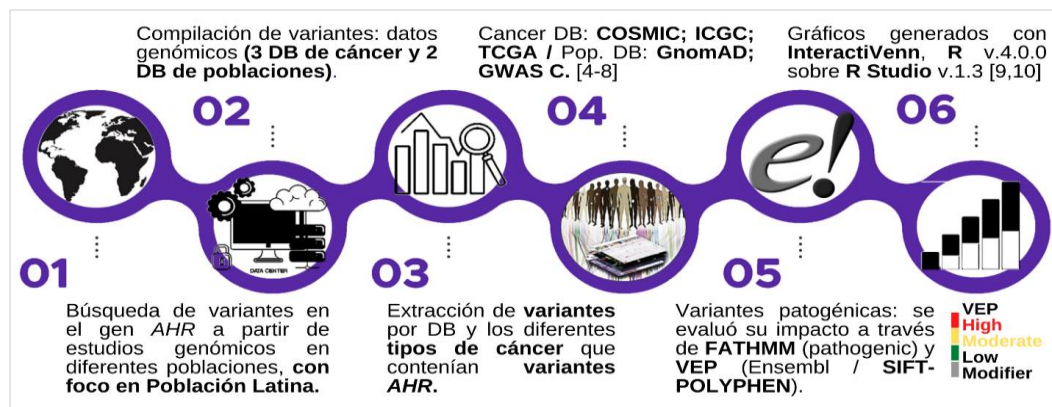
La evaluación de impacto de cada variante se realizó mediante el uso de herramientas de predicción *in silico*, vinculadas a las bases de datos exploradas, permitiendo su clasificación de acuerdo con un score de impacto funcional sobre la proteína, así:

- (i) GnomAD reporta scores asociados a las herramientas Polyphen2 (score: 0.0 – 0.49 con predicción “Benign”; score: 0.50 – 0.89 con predicción “Possibly damaging”; score: 0.90 – 1.00 con predicción “Probably damaging”) (Adzhubei et al. 2010; Adzhubei, Jordan, & Sunyaev, 2013) y SIFT (score: 0.0 – 0.05 con predicción “Damaging”; score: 0.051 – 1.00 con predicción “Tolerated”) (Kumar, Henikoff & Ng, 2009; Vaser et al. 2016).
- (ii) ICGC emplea SIFT y Mutation Assessor (score: 0.0 – 0.90 con predicción “Neutral”; score: 0.91 – 2.00 con predicción “Low”; score: 2.01 – 3.60 con predicción “Medium”; score: < 3.60 con predicción “High”) (Reva, Antipin, & Sander, 2007; 2011) para clasificar las variantes.
- (iii) COSMIC reporta scores de impacto de acuerdo con FATHMM (score: < 0.0 con predicción “Damaging”; score: > 0.0 con predicción “Tolerated”) (Shihab et al. 2013).
- (iv) TCGA utiliza la herramienta de predicción de Ensembl, denominada VEP (score: de acuerdo con el tipo de variante, con predicción “High”, “Moderate”, “Low”, “Modifier”) (McLaren et al. 2016; Hunt et al. 2021) Esta última herramienta permite evaluar el impacto funcional de variantes codificantes y no codificantes.

Si bien, cada repositorio genómico reportaba al menos un score asociado a una herramienta de predicción in silico, se evaluó el impacto de cada variante única identificada, con todos los programas anteriormente mencionados para el conjunto total de datos. Esto permitió predecir el posible impacto funcional de todas las variantes encontradas. Para identificar la presencia de variantes patogénicas de *AHR* asociadas con cáncer en el conjunto de datos, se realizó un filtrado inicial utilizando la herramienta VEP. Sólo fueron seleccionadas las variantes observadas con un impacto funcional “alto”. Posteriormente, para ampliar el análisis de variantes, también fueron consideradas aquellas que presentaban un score de impacto funcional "Moderado" y "Patogénico" de acuerdo con VEP y FATHMM, respectivamente. La visualización de datos se creó a través de InteractiVenn (Heberle et al, 2015) y R (R Core Team 2019 R, 2019) v.4.0.0 mediante R Studio (RStudio Team RStudio , 2019) v.1.3. 1073. Los gráficos se generaron con el software R Studio (Fig. 11).

**Figura 11.**

*Representación gráfica: análisis de variantes en el gen AHR reportadas en repositorios de pacientes cáncer y su potencial impacto.*



Nota: Desarrollo metodológico durante la fase de estudio: análisis de variantes en *AHR* reportadas en bases de datos de pacientes cáncer y evaluación de su potencial impacto. Gráfico de autoría propia.

### **3.2 Fase II: identificación de variantes genéticas germinales presentes en el gen *AHR* en pacientes colombianos con carcinoma escamocelular de cabeza y cuello.**

Esta fase de la investigación comprendió un estudio de tipo descriptivo, a partir del cual se identificaron variantes genéticas germinales en el gen *AHR* y se predijo para las variantes encontradas su efecto en la función del receptor Aril de hidrocarburos mediante análisis *in silico*.

#### ***3.2.1 Población de estudio***

El presente trabajo analizó un grupo de pacientes colombianos diagnosticados con carcinoma escamo celular de cabeza y cuello (HNSCC, por sus siglas en inglés) a partir de muestras colectadas en instituciones prestadoras de salud: Clínica Ambulatoria del Hospital Universitario de Santander-HUS y la Fundación Foscal Internacional-Fosunab, de la ciudad de Bucaramanga y el Laboratorio Patólogos Asociados, de la ciudad de Pasto.

Pacientes con estudio patológico confirmado de HNSCC en cavidad oral, oro-/naso-/hipofaringe o laringe, tratados en estos centros de salud fueron invitados a participar. Un total de 23 pacientes fueron incluidos en la investigación, obteniendo el consentimiento informado por escrito de todos los participantes. Este proyecto fue avalado por el Comité de Ética de la Universidad Cooperativa de Colombia – Campus Pasto, acta de inicio 2085 del 15 enero de 2018 y el Comité de Ética de la Universidad Industrial de Santander – CEINCI, Bucaramanga, acta de aprobación N. 1 del 7 de febrero de 2020.

**3.2.1.1 Criterios de inclusión.**

- Pacientes hombres y mujeres mayores de dieciocho años nacidos en Colombia.
- Pacientes diagnosticados con HNSCC en cavidad oral, faringe o laringe correspondiente a los códigos C01 - C06 /C09 - C14 /C32 de acuerdo con la clasificación CIE10 (Ver Tabla 1).
- Fumadores al momento del diagnóstico o hasta 12 meses antes del diagnóstico (los fumadores serán definidos por el consumo de al menos un paquete (20 unidades) de cigarrillos al año (Nebert et al, 2016).

**Tabla 1.**

*Clasificación de neoplasias malignas de Cáncer de cabeza y cuello, según clasificación CIE10.*

<b>Códigos CIE10</b>	<b>Neoplasias Malignas</b>
(C01)	Neoplasias malignas de la base de la lengua
(C02)	Neoplasias malignas de otras partes y sin especificar de la lengua
(C03)	Neoplasias malignas de la encía
(C04)	Neoplasias malignas de la base de la boca
(C05)	Neoplasias malignas del paladar
(C06)	Neoplasias malignas de otras partes y sin especificar de la boca
(C09)	Neoplasia maligna de amígdala
(C10)	Neoplasias malignas de la orofaringe
(C11)	Neoplasias malignas de la nasofaringe
(C12)	Neoplasias malignas del seno piriforme
(C13)	Neoplasias malignas de la hipofaringe
(C14)	Neoplasias malignas de otras partes del labio, la cavidad oral y la faringe
(C32)	Neoplasias malignas de laringe

Nota: Información tomada de la Edición electrónica de la CIE-10-ES, 2022.

**3.2.1.2 Criterios de exclusión.**

- Pacientes que por su condición (fallecimiento, imposible ubicación, trastorno psiquiátrico), no fuera posible tomar la muestra y/o completar la información de la encuesta.

- Pacientes con diagnóstico de HNSCC asociado a VPH de acuerdo con historia clínica
- Pacientes con antecedentes de otros tipos de cáncer.
- Pacientes que no deseen participar en el estudio.
- Mujeres en estado de embarazo.

### ***3.2.2 Información de variables***

Todos los participantes del estudio completaron una encuesta administrada presencialmente por el investigador, la cual incluía caracterización sociodemográfica, factores de riesgo asociados con el HNC (información detallada sobre consumo de tabaco, alcohol, exposición a agentes ambientales, antecedentes clínicos), antecedentes familiares, diagnóstico y localización. Las variables sociodemográficas fueron definidas de la siguiente manera: (i) Edad: referida a la edad actual del paciente y al momento de su diagnóstico. (ii) Género: hombre o mujer. (iii) Lugar de origen. (iv) Grupo étnico: de acuerdo con el autorreconocimiento, los pacientes fueron agrupados: “Afrodescendiente”, “Mestizo”, “Indígena”, “Otro”. (v) Nivel educativo. (vi) Régimen de salud y estrato socioeconómico.

De acuerdo con los factores de riesgo asociados con la patología de estudio se recolectó la siguiente información: (i) Consumo de tabaco: edad de inicio, edad en la que dejó el consumo de tabaco, tipo de tabaco fumado, número de cigarrillos consumidos al día y exposición habitual en ambientes con humo. (ii) Consumo de alcohol: edad de inicio, edad en la que dejó el consumo de alcohol, frecuencia de consumo de bebidas alcohólicas, consumo de alcohol y tabaco simultáneamente. (iii) Exposición a agentes ambientales: exposición frecuente al sol, promedio en

horas de exposición al sol, razón de la exposición. (iv) Antecedentes clínicos: referido a otras afecciones como “Hepatitis”, “VIH”, “Hipertensión Arterial”, “Diabetes”, “Tuberculosis”, “Otra”, “Ninguna”.

Las variables definidas dentro de la presentación clínica de la patología y los antecedentes familiares fueron: (i) presentación clínica: tipo de cáncer, presencia de signos y síntomas en el momento del diagnóstico. (ii) Historia familiar de cáncer: referido a la existencia de uno o más familiares con cualquier tipo de cáncer. (iii) Grado de parentesco: “Parentesco de primer grado (padres, hermanos e hijos)”, “otros parientes (tíos, primos, sobrinos, abuelos, nietos).

Según el diagnóstico y localización de la neoplasia las variables analizadas fueron: (i) Región anatómica del cáncer. (ii) Clasificación histológica del tumor. (iii) Grado de invasión. (iv) Gradación de acuerdo a la diferenciación celular: “Grado I (Bien diferenciado)”, “Grado II (Moderadamente diferenciado)”, “Grado III (Carcinoma anaplásico o pobremente diferenciado)”, “Grado IV”. (v) Clasificación TNM: T: (Tamaño del tumor), N: (Nódulos linfáticos implicados), M (Metástasis a distancia). (vi) Estadío. Estos datos se registraron de acuerdo con la AJCC Cancer Staging Manual, 8ª edición (Mahul et al. 2017). La historia clínica de cada paciente fue revisada para la recopilación y validación de los datos, previo consentimiento de los participantes.

### ***3.2.3 Recolección de muestras***

Una muestra de sangre periférica, correspondiente a 4 ml, fue obtenida de cada paciente participante de este estudio mediante el procedimiento de venopunción siguiendo los parámetros

de toma de muestra de sangre de la OMS (World Health Organization, 2010). Todas las muestras se colectaron en tubos Vacutainer con EDTA, y fueron almacenadas a 4°C para posteriores procedimientos de extracción de ADN. Las mismas, fueron obtenidas siguiendo las consideraciones éticas y legales establecidas por la Universidad Cooperativa de Colombia, la Universidad Industrial de Santander y la ley colombiana para la investigación en humanos.

### ***3.2.4 Extracción de ADN***

A partir de las muestras colectadas, se realizó extracción de ADN genómico de linfocitos de sangre periférica mediante el kit de extracción Wizard® Genomic DNA Purification de Promega (Promega Corporation, 2019). La pureza y concentración de las muestras de ADN fueron determinadas por espectrofotometría mediante el uso del equipo NanoDrop 2000 Spectrophotometer (Thermo Scientific) y evaluadas mediante corridas electroforéticas en gel de agarosa al 1%.

### ***3.2.5 Análisis del gen AHR***

Los aislados de ADN fueron amplificados para 17 fragmentos de PCR de tamaños que oscilaban entre los 233 y 844 pb que amplifican la región codificante (11 exones) y los límites exón-intrón del gen *AHR*. Dada la longitud de los exones 10 y 11 del gen, estos se dividieron así: (i) Exón 10: dos fragmentos, denominados 10.1 y 10.2; (ii) Exón 11: seis fragmentos, denominados 11.1, 11.2, 11.3, 11.4, 11.5, 11.6 (Tabla 2).

**Tabla 2.**

*Cebadores de PCR específicos para AhR.*

Primer	Secuencia	% GC	Tm (50 M Na+)	Concentración (nmoles)	Tamaño (pb)
AHR_1 F	GGG TGC TCC TGC TAT TCA G	58	58	42.8	819
AHR_1 R	CAG GAT CGC AGG GAT TTG	56	60	37.0	819
AHR_2 F	ACC CAT TCC CTG TTT TTC TT	40	57	45.1	580
AHR_2 R	CAT CAT TTC AGT GGC TGC AT	45	57	36.9	580
AHR_3 F	CAG TGG ATG CCA GTA TTT CA	45	57	36.9	350
AHR_3 R	CGC AGT TCA GAT GAA AGT GG	50	59	38.7	350
AHR_4 F	GGC TGT GTT TGT GAA ATG TGA	43	60	35.3	299
AHR_4 R	AAT GGC CAA CTG TGT CCT AAA	43	59	32.9	299
AHR_5 F	AAT CAG TCC TTT TGT TGT ATT CCT	33	57	42.2	233
AHR_5 R	CCA GCC TCC TTT CAC AGA CT	55	59	36.4	233
AHR_6 F	CAG AAC ACA GAC TCC AGT TTA GAA	42	58	30.8	352
AHR_6 R	GGT TGC CTG ATT TGG GAA TA	45	60	30.7	352
AHR_7 F	CAG CCA TAA TGG AGC CTT TT	45	59	37.6	426
AHR_7 R	GGC TAC ACT GGA AGA ATG TAC G	50	59	35.5	426
AHR_8 F	TGC AGA AAC TAG CGT AAA ACC A	41	60	33.3	250
AHR_8 R	ATG CCG TTT TCT TCT GAC AT	40	57	41.1	250
AHR_9 F	TGT CTT TGG GGA TAA AGG AA	40	57	35.8	500
AHR_9 R	TGC CCA AGA TAA AAC CCT AGA	43	59	35.3	500
AHR_10.1 F	GCA CCA GTC ATG AGA ATA ATT GAG	42	60	39.6	844
AHR_10.1 R	TGT TCC TGT ACC ATA CAG CTT GA	44	60	35.4	844
AHR_10.2 F	TGA CTT AAC GGA TGA AAT CCT G	41	59	39.2	800
AHR_10.2 R	TGC TGT CTC CAT GAA TGC TG	50	61	41.5	800
AHR_11.1 F	TCC AAT AAG TTG CAT CAC CA	40	58	35.0	713
AHR_11.1 R	AAA CAG TGA AGG AGG AAA GCA	43	59	30.5	713
AHR_11.2 F	CTT TGA GTT TTG AGC TAC TGG ATT C	40	60	38.1	696
AHR_11.2 R	ATT GAA ATG GGT GCT TTG CT	40	60	36.1	696
AHR_11.3 F	TGT TGC TAT GTG CCT TAT GTT G	41	59	42.6	825
AHR_11.3 R	AGT TTG GGC GGC AGA GTA AG	55	62	38.8	825
AHR_11.4 F	CAA CCA CAT AGT TCG TTT ACC TTC	42	59	36.7	773
AHR_11.4 R	GCT CTG CAA CCT CCG ATT C	58	61	38.7	773
AHR_11.5 F	TGA TGG AAT CCC AAA CTG AGA	43	60	34.0	841
AHR_11.5 R	CCA CAA TGT AGC AGT GCA CTA AG	48	60	38.0	841
AHR_11.6 F	ACT CTC TTC CAC ATG TTA CTG GA	44	58	40.6	600
AHR_11.6 R	TTC CTT TAC GTG GAC CAG AA	45	58	36.9	600

Nota: información relevante de cada cebador (primer), incluidas sus secuencias, el contenido de GC, temperatura de anillamiento y longitud (pb).

La amplificación se realizó a partir de la técnica de Reacción en Cadena de la Polimerasa (PCR) utilizando master mix MyFiTM Mix (Bioline®) 1X, 0.4  $\mu$ M de mezcla de cebadores y 50 ng/ $\mu$ L de ADN genómico, en 30  $\mu$ L de volumen final de reacción (Apéndice B. Tabla 3. Protocolo de amplificación de PCR). El programa de PCR consistió de un período de denaturación inicial, 94 °C por 2 min; 30 ciclos de: denaturación, 94 °C por 45 s; hibridación, 60 s a una temperatura de 58 - 63 °C dependiendo de los cebadores empleados; elongación, 72 °C por 60 s; y una extensión final a 72 °C por 10 min (Apéndice C. Tabla 4. Condiciones de Termociclado).

La presencia e integridad de los productos obtenidos de PCR se verificaron mediante electroforesis convencional en gel de agarosa a una concentración de [1.5%], teñido con GelRed® y un tiempo de corrida: (i) Precorrida: 46 voltios durante 10 min; (ii) Corrida final: 120 voltios durante 45 min. El tamaño de la banda esperada se confirmó por comparación con el marcador de peso molecular HyperLadder™ 100bp (Bioline).

Finalmente, los fragmentos amplificados fueron secuenciados por medio del método de Sanger en un analizador genético ABI 3130 (Applied Biosystems) de manera bidireccional, empleando los mismos cebadores de la amplificación por cada muestra para cada uno de los fragmentos; el servicio de secuenciación fue proveído por MACROGEN (Seoul, South Korea). Las secuencias obtenidas se procesaron y analizaron con el programa ChromasPro versión 2.1.10 (<http://technelysium.com.au/ChromasPro.html>).

### 3.2.6 Análisis Computacional

Para identificar las variantes en el gen *AHR*, presencia y ubicación de las alteraciones, las secuencias obtenidas fueron comparadas con las secuencias de referencia: (i) *Homo sapiens* PAC clone RP4-658N5 from 7p21, complete sequence, DNA - NCBI Reference Sequence: AC003075.2; (ii) *Homo sapiens* Aryl hydrocarbon receptor (*AHR*), mRNA - NCBI Reference Sequence: NM\_001621.5, a partir de la herramienta de alineamiento de secuencias de tipo local: BLAST (Basic Local Alignment Search Tool) - NCBI.

Las variantes halladas fueron nombradas de acuerdo con la nomenclatura aprobada por Human Genetic Variation Society (HGVS) (<https://varnomen.hgvs.org/>). Se determinaron las frecuencias alélicas de cada una de las variantes identificadas y se compararon con las frecuencias reportadas para cada población en las siguientes bases de datos: (i) dbSNP (<http://www.ncbi.nlm.nih.gov/SNP/>); (ii) 1000 Genomes Project Phase 3 (<https://www.internationalgenome.org/category/phase-3/>); (iii) GnomAD - ExAC (Ensembl: <https://www.ensembl.org/info/genome/variation/species/populations.html>).

### 3.3 Fase III: evaluación del efecto de las variantes identificadas sobre la función del receptor Aril hidrocarburos mediante análisis de predicción *in silico*.

Las variantes genómicas de regiones codificantes y no codificantes identificadas se evaluaron a través de un conjunto de herramientas de predicción *in silico* para evaluar su posible efecto sobre la estructura y función de AhR.

Las variantes codificantes fueron analizadas mediante los programas:

- (i) VEP: Ensembl Variant Effect Predictor, determina el posible efecto de las variantes (SNP, inserciones, deleciones, CNV o variantes estructurales) en genes, transcritos y secuencias de proteínas, así como regiones reguladoras. (*score: de acuerdo con el tipo de variante, con predicción “High”, “Moderate”, “Low”, “Modifier”*) (McLaren et al. 2016; Hunt et al. 2021) (<https://www.ensembl.org/Tools/VEP>).
- (ii) SIFT: Sorting Intolerant from Tolerant, predice si una sustitución de aminoácidos podría afectar la función de la proteína en función de la homología de la secuencia y las propiedades físicas de los aminoácidos. SIFT, permite identificar que tan conservado es el sitio en la proteína donde se encuentra la variante. SIFT puede ser usado para polimorfismos no sinónimos y mutaciones sin sentido. (*score: 0.0 – 0.05 con predicción “Damaging”; score: 0.051 – 1.00 con predicción “Tolerated”*) (Kumar, Henikoff & Ng, 2009; Vaser et al. 2016) (<https://sift.bii.a-star.edu.sg/>).
- (iii) Polyphen2: Polymorphism Phenotyping v2, establece si la variante afecta la estructura tridimensional de la proteína. Este programa determina el posible efecto de variantes codificantes no sinónimas mediante procedimientos de alineación de secuencias múltiples de alta calidad y una clasificación probabilística de acuerdo con el método de “machine learning” (*score: 0.0 – 0.49 con predicción “Benign”; score: 0.50 – 0.89 con predicción “Possibly damaging”; score: 0.90 – 1.00 con predicción “Probably damaging”*) (Adzhubei et al. 2010; Adzhubei, Jordan, & Sunyaev, 2013) (<http://genetics.bwh.harvard.edu/pph2/>).

- (iv) FATHMM: Functional Analysis through Hidden Markov Models (v2.3), permite predecir las consecuencias funcionales de variantes codificantes y no codificantes. De acuerdo con las necesidades del investigador, el análisis de variantes codificantes cuenta con tres herramientas distintas: (a) Enfermedad hereditaria, opción que permite discriminar entre mutaciones causantes de enfermedades y polimorfismos neutrales; (b) Cáncer, esta opción permite distinguir entre mutaciones conductoras que promueven el cáncer y otros polimorfismos de línea germinal; (c) Específico de la enfermedad, permite identificar variantes potencialmente relevantes desde el punto de vista funcional para la enfermedad de interés. Por otra parte, la evaluación de variantes no codificantes cuenta con dos herramientas de análisis: (d) FATHMM-MKL; (e) FATHMM-XF, estos algoritmos integran anotaciones funcionales de ENCODE con HMM basados en nucleótidos. *(Para variantes de región codificante: score: < 0.0 con predicción “Damaging”; score: > 0.0 con predicción “Tolerated”. Para variantes de región no codificante: p-values in the range [0, 1] - score: < 0.5 con predicción “neutral or benign”; score: > 0.5 con predicción “Deleterious”)* (Shihab et al. 2013) (<http://fathmm.biocompute.org.uk/>)
- (v) Mutation Taster (Mutation T@sting): el programa utiliza modelos de Random Forest para las predicciones. El “*Tree Vote*” indica cuántos árboles de decisión del Random Forest sugieren nocividad, frente a cuántos sugieren una alteración de tipo benigna. Mutation Taster integra diferentes fuentes de datos como GnomAD y

ExAC para el análisis del efecto de las variantes de empalme, el programa usa MaxEntScan como modelo de predicción. Además, Mutation Taster proporciona información sobre la enfermedad que causa la variante evaluada. (*score: de acuerdo con el tipo de variante y el modelo asociado: 100 a 300 Random Forest trees, con predicción “deleterious”; “benign”*) (Steinhaus et al, 2021) (<https://www.genecascade.org/MutationTaster2021/>).

- (vi) Provean: Protein Variation Effect Analyzer, predice el impacto de variantes, tales como sustituciones, inserciones y deleciones de uno o varios aminoácidos, sobre la función biológica de una proteína. Este programa utiliza el principio de alineamientos con secuencias homólogas para determinar la posición de residuos y aminoácidos, evolutivamente conservados, y por lo tanto funcionalmente importantes. (*score: < -2.5 con predicción “Deleterious”; score: > -2.5 con predicción “Neutral”*) (Choi y Chan, 2015) (<http://provean.jcvi.org/index.php>).

Por otra parte, las variantes no codificantes fueron evaluadas a partir de las herramientas de predicción ya mencionadas, VEP y FATHMM, Mutation Taster (Mutation T@sting) además de los programas Human Splicing Finder (HSF), que a su vez corre el algoritmo MaxEntScan.

- (i) HSF: Human Splicing Finder, herramienta para predecir los efectos de las mutaciones en las señales de empalme o para identificar motivos de empalme en cualquier secuencia humana. Esta herramienta calcula valores consenso para sitios de splicing y puntos de ramificación. Entre más alto sea este valor, mayor

probabilidad de que sea un sitio auténtico (Desmet, et al. 2009) (<https://hsf.genomnis.com/>).

- (ii) MaxEntScan, Este método modela secuencias de sitios de empalme utilizando la distribución de máxima entropía y generaliza distintos modelos probabilísticos, como los modelos de matriz de peso y los modelos de Markov no homogéneos. Este programa asigna puntajes a secuencias de 9 pb (sitios de splicing 5': 3 bases exónicas y 6 bases intrónicas alrededor de los 5'SS, dada la alta conservación nucleotídica de estas posiciones en el contexto del empalme con elementos canónicos); De manera similar se asignan puntajes para los sitios de splicing 3' (secuencias de 23 pb), donde se utilizan 3 bases exónicas y 20 bases intrónicas (teniendo en cuenta las longitudes variables del tracto de polipirimidina). Cuanto más alto sea el puntaje, mayor la probabilidad de que la secuencia evaluada sea un sitio verdadero de splicing; por el contrario, cuanto más bajo el puntaje, este se interpreta como un sitio de empalme "más débil". (Yeo & Burge 2004) (<http://genes.mit.edu/burgelab/maxent/Xmaxentscan scoreseq.html>).

Conjuntamente, se usó HaploReg v4.133 para explorar qué anotaciones funcionales del genoma no codificante entre los resultados previamente publicados o conjuntos de variantes de los proyectos Roadmap Epigenomics, ENCODE, 1000 Genomes Project, y el efecto de los SNP en la expresión de los estudios de eQTL, colocan con ciertos SNPs identificados en el presente estudio. HaploReg es una herramienta que permite predecir el impacto de variantes no codificantes en los fenotipos clínicos (Ward y Kellis, 2011) (<http://compbio.mit.edu/HaploReg>).

Finalmente, el análisis in silico de cada variante identificada mediante las herramientas previamente expuestas, permitieron determinar las alteraciones de secuencia, probablemente reguladoras, para futuros estudios de asociación.

## **4. Resultados y discusión**

### **4.1 Fase I: descripción de las variantes genéticas reportadas para el gen *AHR* en repositorios de datos genómicos de pacientes con cáncer.**

Este apartado se presenta como Apéndice D al presente trabajo. Artículo por someter: *AHR* gene variants in patients with cancer and their presence in the Latin American population.

### **4.2 Fase II: identificación de las variantes genéticas germinales presentes en el gen *AHR* en pacientes colombianos con carcinoma escamocelular de cabeza y cuello.**

#### ***4.2.1 Caracterización sociodemográfica y clínica de los pacientes estudiados***

En este estudio se analizaron 23 pacientes con carcinoma escamocelular de cabeza y cuello de acuerdo con los criterios de inclusión descritos previamente (Tabla 1). La distribución geográfica de los pacientes fue la siguiente: 6 pacientes del Departamento de Nariño,

correspondientes al municipio de Pasto; 14 pacientes del Departamento de Santander, correspondientes a los municipios de Bucaramanga, San Gil y Socorro; 3 pacientes del Departamento de Norte de Santander, correspondientes al municipio de Cúcuta y el corregimiento de Bábeca (1 paciente, residente en zona rural), quienes recibían tratamiento médico en la ciudad de Bucaramanga. La edad media de los pacientes estudiados fue de  $57 \pm 14$  años con un predominio de hombres (61%,  $n = 14$ ), sobre el número de mujeres afectadas (39%,  $n = 9$ ). Distribuidos por rango etario, el 60,8% ( $n = 14$ ) del total de pacientes eran mayores de 50 años, mientras que sólo el 8,7%, correspondiente a 2 pacientes evaluados, tenían una edad menor o igual a 30 años. Por lo que se refiere al grupo étnico, el total de los participantes (100%) se auto-reconoce como mestizo (Tabla 5).

En cuanto a las características clínicas (Tabla 6), la localización del tumor primario de origen escamocelular más frecuentemente reportado fue la lengua, representando el 43,5% de los pacientes analizados ( $n = 10$ ); Laringe: 6 pacientes (26,1%); Piso de la boca: 2 pacientes (8,7%); Amígdala: 2 pacientes (8,7%); Cavidad oral: 1 paciente (4,4%); Lengua y piso de la boca: 1 paciente (4,4%); Paladar: 1 paciente (4,4%) (Fig. 12).

Al relacionar el género con el consumo de tabaco, como uno de los factores de riesgo asociado al desarrollo de HNSCC, se observó que la edad de inicio de consumo de tabaco es más temprana en hombres (18 años o menos) que, en mujeres representando el 43,5% del total de los pacientes. Así, mientras que el tiempo de exposición a procarcinógenos del tabaco en las mujeres varía entre 1 – 5 años o menos (30,4%;  $n = 7$ ), en los hombres asciende a los 11 - 20 años o más (52,2%;  $n = 12$ ).

**Tabla 5.**

*Caracterización sociodemográfica de la población de estudio.*

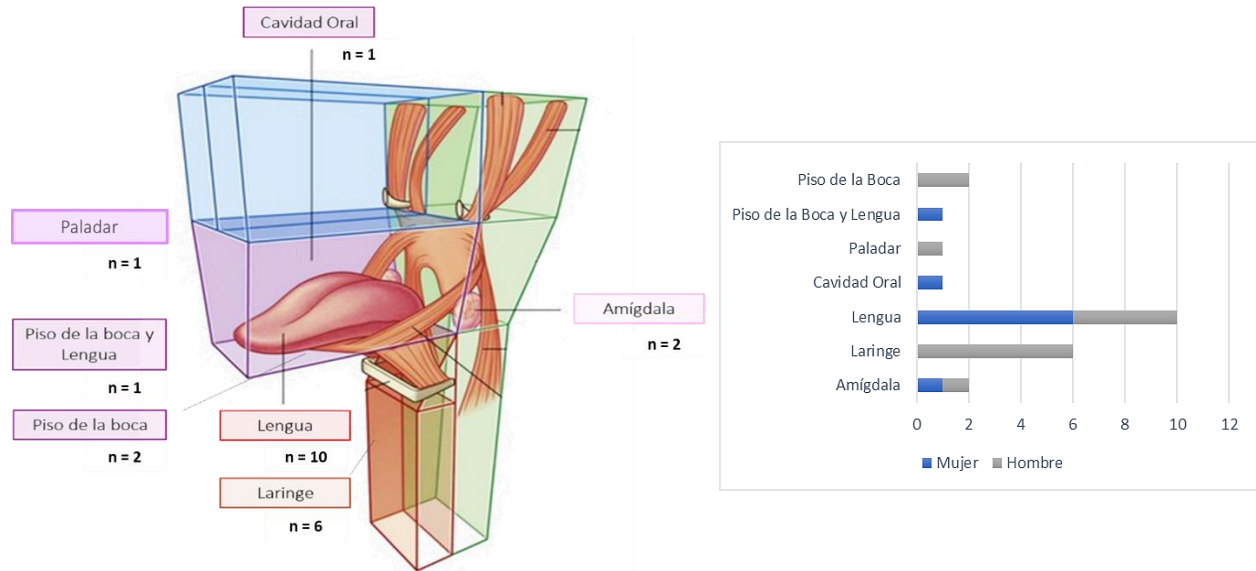
	Mujeres n (%)	Hombres n (%)	Total n (%)
<b>Total</b>	9 (39%)	14 (61%)	23 (100%)
<b>Edad (años)</b>			
18 - 30	2 (8,7%)		2 (8,7%)
31 - 40	1 (4,35%)		1 (4,35%)
41 - 50	1 (4,35%)	5 (21,7%)	6 (26,1%)
51 - 60		5 (21,7%)	5 (21,7%)
61 y más	5 (21,7%)	4 (17,4%)	9 (39,1%)
<b>Grupo Étnico</b>			
Afrodescendiente			
Mestizo	9 (39,1%)	14 (60,9%)	23 (100%)
Indígena			
Otro:			
<b>Estado Civil</b>			
Soltero	3 (13,1%)	4 (17,4%)	7 (30,4%)
Casado	6 (26,1%)	7 (30,4%)	13 (56,5%)
Unión Libre		1 (4,35%)	1 (4,35%)
Divorciado		1 (4,35%)	1 (4,35%)
Viudo		1 (4,35%)	1 (4,35%)
<b>Nivel Educativo</b>			
Ninguno		1 (4,35%)	1 (4,35%)
Primaria	3 (13,1%)	4 (17,4%)	7 (30,4%)
Secundaria		6 (26,1%)	6 (26,1%)
Técnica	4 (17,4%)	1 (4,35%)	5 (21,7%)
Universitario	1 (4,35%)	1 (4,35%)	2 (8,7%)
Posgrado	1 (4,35%)	1 (4,35%)	2 (8,7%)
<b>Régimen de Salud</b>			
Contributivo	6 (26,1%)	7 (30,4%)	13 (56,5%)
Subsidiado	3 (13,1%)	7 (30,4%)	10 (43,5%)
Especial			
Ninguno			
<b>Actividad Laboral</b>			
Estudiante	2 (8,7%)		2 (8,7%)
Hogar	4 (17,4%)		4 (17,4%)
Trabajador Empleado	2 (8,7%)	3 (13,1%)	5 (21,7%)
Trabajador Independiente		8 (34,8%)	8 (34,8%)
Cesante	1 (4,35%)	1 (4,35%)	2 (8,7%)
Pensionado		2 (8,7%)	2 (8,7%)
<b>Estrato Socioeconómico</b>			
Estrato 1 - 2	6 (26,1%)	5 (21,7%)	11 (47,8%)
Estrato 3 - 4	3 (13,1%)	8 (34,8%)	11 (47,8%)
Estrato 5 - 6		1 (4,35%)	1 (4,35%)

Nota: Caracterización sociodemográfica de la población total de estudio, 23 pacientes con HNSCC.

En cuanto al consumo de tabaco, distribuido por número de cigarrillos y/o tabacos consumidos por día, el 21,7% (n = 5) del total de las mujeres reportaron un consumo de 1 – 3 cigarrillos diarios; el 8,7% (n = 2) dijo haber consumido de 10 – 20 cigarrillos por día, es decir de medio a un paquete completo; el 8,7% restante (n = 2) presentó un consumo de más de 20 cigarrillos cada día. Por otra parte, el 47,9% (n = 11) del total de los hombres reportaron un consumo de 10 – 20 cigarrillos o más cada día; el 8,7% (n = 2) consumió de 1 – 3 cigarrillos diarios, y sólo uno de ellos (4,4%) presentó un consumo de más de 3 y menos de 10 cigarrillos en un día (Tabla 6).

**Figura 12.**

*Número de pacientes estudiados de acuerdo con la localización anatómica del tumor.*



Nota: Representación esquemática de los sitios anatómicos de aparición del tumor y distribución de pacientes afectados por éstas neoplasias. Gráfico modificado de <https://clinicalgate.com/head-and-neck/>

Al analizar otros factores de riesgo asociados al HNSCC, tales como consumo de alcohol y exposición al sol se observó que el 34,8% (n = 8) del total de los pacientes dijeron haber consumido alcohol y tener exposición prolongada al sol con relación a sus labores, de manera ocasional, además del consumo de tabaco. De estos, el 26,1% correspondió a hombres y el 8,7% a mujeres. En cuanto al consumo de alcohol y tabaco en concomitancia, sin exposición al sol, el 34,8% (n = 8) de los pacientes (26,1% hombres; 8,7% mujeres) reportaron dicho consumo en simultáneo. Finalmente, el 13,1% (n = 3) de los pacientes restantes afirmaron haber presentado exposición al sol por ocio y un 4,4% reportó haber consumido alcohol en algún momento de su vida antes del diagnóstico. Sólo 3 de los participantes del estudio no refirieron consumo de alcohol y/o exposición prolongada al sol (13,1%) (Tabla 6).

En cuanto a la presencia de una lesión en el sitio de inicio del cáncer, el 65,2% (n = 15) de los pacientes afirmaron haber presentado una lesión inicial. En general, el 43,4% (n = 10) del total de los pacientes estudiados reportaron antecedentes familiares de cáncer en parientes de primer o segundo grado, siendo el cáncer de mamá el más frecuente (33%; n = 5 casos reportados); seguido de cáncer de cabeza y cuello (20%; n = 3 casos reportados); leucemia ((20%; n = 3 casos reportados), y otros menos frecuentes como cáncer de tiroides, hígado, ovario y útero (6,7%; n = 1 caso reportado, asociado a cada neoplasia) (Fig. 13).

**Tabla 6.**

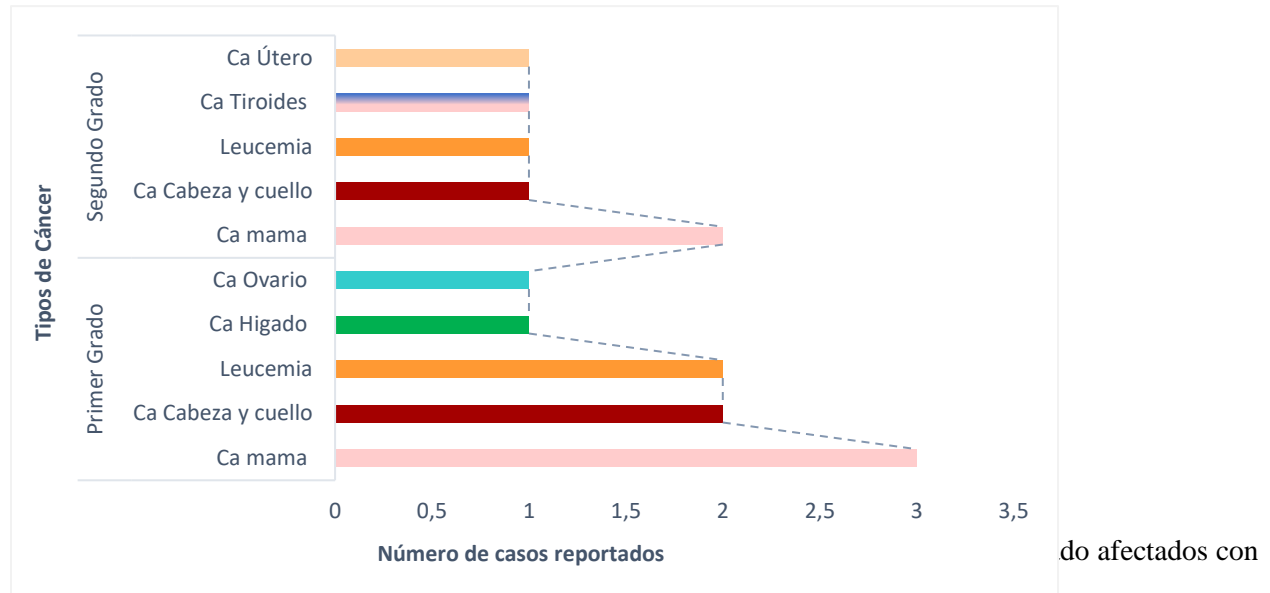
*Caracterización clínica de la población de estudio.*

	Mujeres n (%)	Hombres n (%)	Total n (%)
<b>Total</b>	9 (39%)	14 (61%)	23 (100%)
<b>Consumo de Tabaco</b>			
Edad de Inicio de Consumo de Tabaco			
18 años o menos	3 (13,1%)	10 (43,5%)	13 (56,5%)
19 años o más	6 (26,1%)	4 (17,4%)	10 (43,5%)
Tiempo de Consumo de Tabaco			
Menos de 1 Año	2 (8,7%)		2 (8,7%)
1 - 5 Años	5 (21,7%)		5 (21,7%)
6 - 10 Años		2 (8,7%)	2 (8,7%)
11 - 20 Años	1 (4,4%)	3 (13,1%)	4 (17,4%)
21 Años o más	1 (4,4%)	9 (39,1%)	10 (43,5%)
Número de Cigarrillos y/o Tabaco consumidos al día			
1 - 3 cigarrillos	5 (21,7%)	2 (8,7%)	7 (30,4%)
Más de 3 y menos de 10 cigarrillos		1 (4,4%)	1 (4,4%)
10 - 20 cigarrillos	2 (8,7%)	8 (34,8%)	10 (43,5%)
Más de 20 cigarrillos	2 (8,7%)	3 (13,1%)	5 (21,7%)
<b>Otros Factores de Riesgo</b>			
Consumo de Alcohol	1 (4,4%)		1 (4,4%)
Tabaco + Alcohol en simultáneo	2 (8,7%)	6 (26,1%)	8 (34,8%)
Exposición al Sol	1 (4,4%)	2 (8,7%)	3 (13,1%)
Consumo de Alcohol + Exposición al Sol	2 (8,7%)	6 (26,1%)	8 (34,8%)
Ninguno	3 (13,1%)		3 (13,1%)
<b>Lesión en Sitio de Inicio del Cáncer</b>			
Sí	8 (34,8%)	7 (30,4%)	15 (65,2%)
No	1 (4,4%)	7 (30,4%)	8 (34,8%)
<b>Antecedentes familiares de Cáncer</b>			
Pariente de Primer Grado	3 (13,1%)	2 (8,7%)	5 (21,7%)
Pariente de Segundo Grado	4 (17,4%)	1 (4,4%)	5 (21,7%)
Ninguno	2 (8,7%)	11 (47,8%)	13 (56,5%)

Nota: Caracterización clínica de la población total de estudio, 23 pacientes con HNSCC. De acuerdo con el Instituto Nacional del Cáncer del NIH, se incluyen como parientes de primer grado: progenitores, hermanos y hermanas, hijos. Además, se denotan como parientes de segundo grado: tíos, abuelos, nietos, sobrinos o medio hermanos de una persona.

**Figura 13.**

*Antecedentes de cáncer en parientes de primer y segundo grado de pacientes con HNSCC*



algún tipo de cáncer, de acuerdo con los antecedentes familiares de los participantes del estudio.

#### 4.2.2 Alteraciones de secuencia observadas en el gen AHR

El gen *AHR* fue dividido en 17 fragmentos cubriendo un total de 9901 pares de bases, que incluía la región codificante (11 exones) y los límites exón-intrón del gen para ser secuenciados mediante secuenciación Sanger.

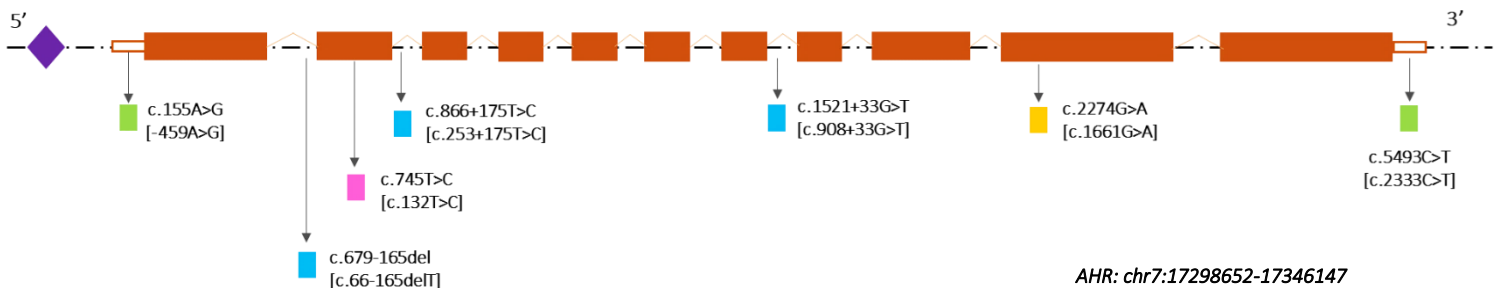
De los 23 pacientes caracterizados para este gen, se identificaron un total 7 alteraciones de secuencia, identificadas mediante el uso de la herramienta Human BLAT Search - UCSC Genome Browser (Kent, 2002; Kent et al, 2002) (<https://genome.ucsc.edu/cgi-bin/hgBlat>). De estas, 4 variantes (57%) se encontraron reportadas en las bases de datos: (i) dbSNP (<https://www.ncbi.nlm.nih.gov/snp/>); (ii) LOVD v.3.0 - Leiden Open Variation Database (<https://www.lovd.nl/3.0/search>); (iii) Varsome, (<https://varsome.com/>). Las 3 alteraciones

restantes (43%) no se encuentran reportadas en estas bases de datos ni tampoco se hallaron en la literatura científica.

Como se observa en la Tabla 7, de las siete variantes identificadas, una corresponde a la región no traducida 5' (-459A>G, 5'UTR); tres alteraciones de tipo intrónicas (c.908+33G>T; c.253+175T>C; c.66-165delT); dos variantes de región codificante (c.132T>C, alteración de tipo sinónima; c.1661G>A, variante de tipo missense que representa un cambio de aminoácido en la posición 554 de Arginina por Lisina); una variante localizada en la región no traducida 3' (c.2333C>T, 3'UTR, variante no reportada previamente) (Figura 14). La alteración de secuencia más frecuente en la población de estudio fue [-459A>G], encontrada en 20 de los 23 pacientes. La variante menos frecuente fue [c.132T>C] presente en 1 sólo paciente. Finalmente, sólo dos (28.6%) de las siete variantes identificadas se encuentran reportadas en ClinVar como “variantes benignas”.

**Figura 14.**

*Representación gráfica del gen AHR y las variantes identificadas en el estudio*



Nota: representación esquemática de las variantes identificadas de acuerdo con su ubicación en el gen. En verde: variantes 5'UTR y 3'UTR; En azul: variantes intrónicas; En rosa: variante sinónima; En naranja: variante de cambio de sentido.

**Tabla 7.**

*Alteraciones de secuencia identificadas en el gen AHR en la muestra caracterizada.*

Fragmento	Posición	Nomenclatura	Alelos	Cambio de A.A	dbSNP	Clasificación de la Variante	Nº. pacientes
Exón 1	GRCh38/HG38: chr7:17298806	<b>-459A&gt;G</b>	Homocigosis	N/A	rs7796976 A/G	Variante 5' UTR	<b>16/23</b> (69.5%)
			Heterocigosis				<b>4/23</b> (17.4%)
Intrón 7	GRCh38/hg38: chr7:17334147	<b>c.908+33G&gt;T</b>	Heterocigosis	N/A	rs2074113	Variante intrónica	<b>9/23</b> (39.1%)
Exón 2	GRCh38/hg38: chr7:17310002	<b>c.132T&gt;C</b>	Heterocigosis	p.Asn44Asn	rs17779352 T/C	Variante codificante de tipo sinónima	<b>1/23</b> (4.3%)
Exón 10 (10.1)	GRCh38/hg38: chr7:17339486	<b>c.1661G&gt;A</b>	Heterocigosis	p.Arg554Lys	rs2066853	Variante codificante de cambio de sentido	<b>6/23</b> (26.1%)
Intrón 2	-	<b>c.253+175T&gt;C</b>	Heterocigosis	N/A	No reportada	Variante intrónica	<b>1/23</b> (4.3%)
Intrón 1	-	<b>c.66-165delT</b>		N/A	No reportada	Variante intrónica	<b>3/23</b> (13.1%)
Exón 11 (11.5)	-	<b>c.2333C&gt;T</b>	Heterocigosis	-	No reportada	Variante 3' UTR	<b>7/23</b> (30.4%)

Nota: En blanco, variantes detectadas en la muestra de estudio y previamente reportadas. Las bases de datos consultadas fueron (i) dbSNP (<https://www.ncbi.nlm.nih.gov/snp/>); (ii) LOVD v.3.0 (<https://www.lovd.nl/3.0/search>); (iii) Varsome, (<https://varsome.com/>). \*Llamado de variantes de acuerdo con la nomenclatura HGVS. En gris, variantes no reportadas. N/A: No aplica

**4.2.2.1 Análisis de variantes alélicas «reportadas» para el gen AHR.** De acuerdo con el UCSC Genome Browser (Human BLAT Search), 4 de las 7 variantes halladas se encontraban previamente reportadas en la base de datos dbSNP, identificadas con los Id rs7796976 [-459A>G]; rs2074113 [c.908+33G>T]; rs17779352 [c.132T>C]; rs2066853 [c.1661G>A]. También se encontraron en las bases de datos LOVD v.3.0 y Varsome. Según su importancia clínica, 2 de las 4 variantes, no han sido reportadas en ClinVar (rs7796976 y rs2074113), mientras que las dos

restantes se encuentran clasificadas como variantes “benignas” (rs17779352 y rs2066853) (Tabla 8).

**Tabla 8.**

*Alteraciones de secuencias en el gen AHR: variantes reportadas y significancia clínica.*

VARIANTES DEL GEN AHR					SIGNIFICANCIA CLINICA				
Nomenclatura HGVS	Cambio de A.A	dbSNP	Clasificación de la Variante	N°. pacientes	CLINVAR	PATHOMAN	VAR SOME	CLINGEN	Citas
NM_001621.5: c.-459A>G	N/A	rs7796976	Variante 5 prima UTR	20/23	No reportada	Variante de significado clínico incierto	Benigna	Indeterminada	7
NM_001621.5: c.908+33G>T	N/A	rs2074113	Variante intrónica	9/23	No reportada	Benigna	Benigna	Indeterminada	4
NM_001621.5: c.132T>C	p.Asn44Asn	rs17779352	Variante codificante de tipo sinónima	1/23	Benigna	Probablemente Benigna	Benigna	Indeterminada	0
NM_001621.5: c.1661G>A	p.Arg554Lys	rs2066853	Variante codificante de cambio de sentido	6/23	Benigna	Variante de significado clínico incierto	Benigna	Indeterminada	65

Nota: Nomenclatura aprobadas por HGVS (Human Genome Variation Society) de acuerdo a la entrada en GenBank no. NM\_001621.5. \*Varsome: significancia clínica de las variantes de tipo “Benigna”, según la clasificación del ACMH. N/A: No aplica.

A continuación, se describen cada una de las variantes identificadas:

**AHR [-459A>G].** Esta alteración de secuencia fue encontrada en el 87% (n = 20) del total de los pacientes estudiados, tanto en homocigosis (n = 16 pacientes) como en heterocigosis (n = 4 pacientes). Variante de la región no traducida 5' (5'UTR) localizada en el exón 1 del gen, e identificada bajo el RefSNP: rs7796976. En la figura 15 se muestra el cromatograma de un paciente con esta alteración.

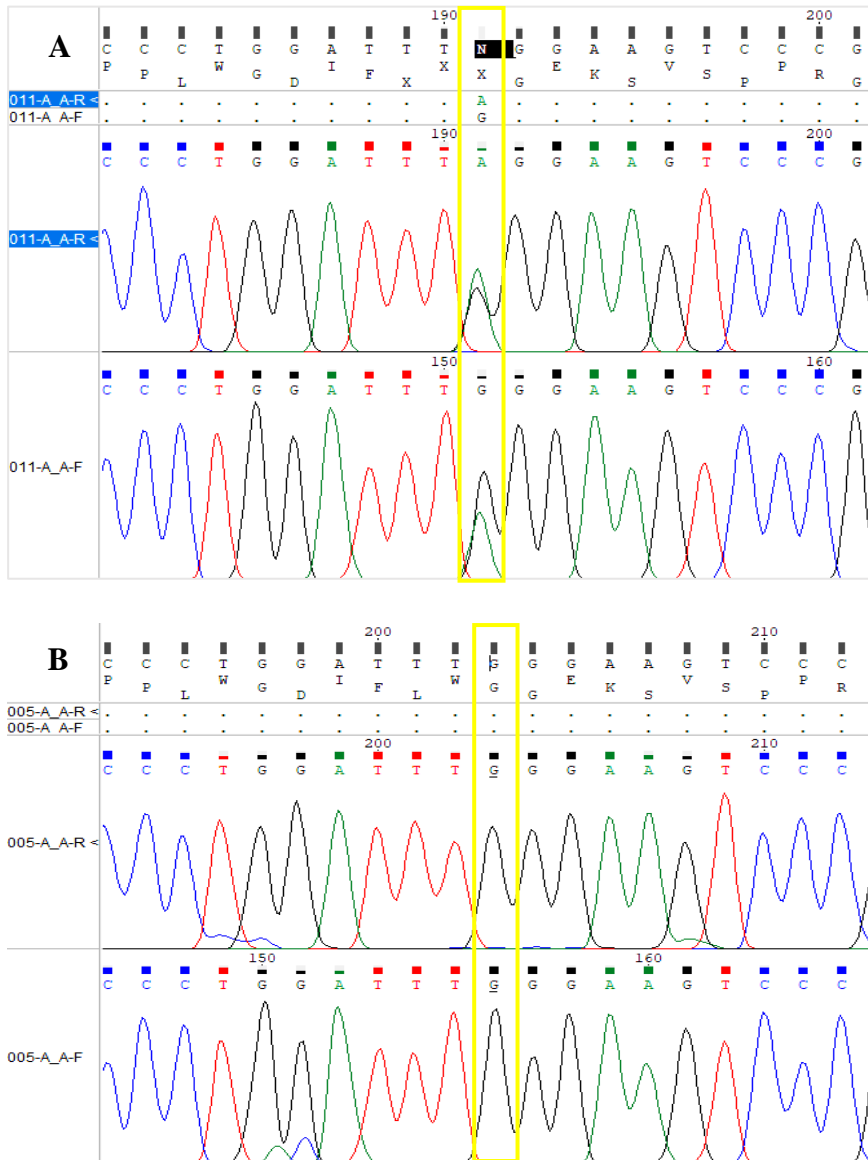
De acuerdo con las bases de datos dbSNP y Ensembl Genome browser, la variante -459A>G presenta una distribución global, como resultado de los estudios del proyecto 1000Genomes, ExAC y GnomAD. Se han encontrado reportes de esta alteración en población asiática: Korea (Kim et al, 2007); población caucásica: España (Figuroa et al, 2008), Alemania (Prager et al, 2016); población mezclada, Canadá (Ng et al, 2010). Para Latinoamérica, se encuentran registros de la frecuencia de esta variante en las poblaciones de Medellín, Colombia; Lima, Perú; Puerto Rico; de ancestría mexicana habitantes de Los Ángeles, CA (Ensembl Genome browser). En la tabla 9, se observan las frecuencias alélicas y genotípicas de la variante en la población de estudio y su distribución en el mundo. No se encuentran datos reportados de esta alteración en otras bases de datos como GWAS Catalog, TCGA, ICGC, COSMIC, y el HapMap Project.

A partir del cálculo de frecuencias para la población muestreada, representada por 23 pacientes, se encontró que la variante [-459A>G] exhibió un valor de frecuencia alélica menor de 0.22 para su alelo «A». Lo anterior se relaciona con el valor de frecuencia alélica menor (MAF, por sus siglas en inglés) reportado por dbSNP para el alelo A (MAF: A= 0.2210/1107), lo que indica que para la variante [-459A>G], el alelo -459A tiene una frecuencia de 22% en la población de estudio (1107 de 5008 individuos) de la fase III del proyecto 1000Genomes (Tabla 9), respaldando a su vez, el gran número de pacientes identificados en el presente estudio portando la alteración, donde el genotipo homocigoto «GG» exhibió la frecuencia más alta en la población (0.70), además del genotipo heterocigoto «AG» que se observó en el 13% de la población estudiada (0.13) (Tabla 9). Conforme a los datos de expresión y regulación génica en tejidos específicos contenidos en el proyecto Genotype-Tissue Expression (GTEx), esta variante se relaciona con la

expresión génica de *AHR* en 54 tejidos, principalmente en cultivos primarios de fibroblastos (RNA-seq of 17382 samples, 948 donors) (<http://varso.me/dkNJ>).

**Figura 15.**

















*Cromatograma de la alteración -459A>G.*



Nota: Cromatograma, identificación de la variante -459A>G en heterocigosis y homocigosis, respectivamente. Alteración en recuadro amarillo. Visualización mediante el SW. ChromasPro V.2.1.10.

**Tabla 9.**

*Frecuencias alélicas de la variante [-459A>G].*

rs7796976					
Allele Frequencies					
1000 Genomes Project Phase 3			GnomAD (Ensembl); ExAC		
ALL	A: 0.221 (1107) G: 0.779 (3901)		ALL	A: 0.206 (29422) G: 0.794 (113530)	
AFRICAN	A: 0.084 (111) G: 0.916 (1211)		AFRICAN / AFR-AME	A: 0.100 (4212) G: 0.900 (37762)	
AMERICAN	A: 0.251 (174) G: 0.749 (520)		LATINO	A: 0.306 (4180) G: 0.694 (9462)	
EAST ASIAN	A: 0.329 (332) G: 0.671 (676)		EAST ASIAN	A: 0.331 (1028) G: 0.669 (2080)	
EUROPEAN	A: 0.222 (223) G: 0.778 (783)		NON FINNISH EUR.	A: 0.241 (15555) G: 0.758 (48851)	
SOUTH ASIAN	A: 0.273 (267) G: 0.727 (711)		FINNISH	A: 0.209 (2178) G: 0.791 (8250)	
Sample Population: rs7796976			SOUTH ASIAN	A: 0.274 (831) G: 0.726 (2197)	
			AMISH	A: 0.209 (188) G: 0.791 (710)	
Genotype	Genotype Frequencies	Allele Frequencies	ASHKENAZI JEWISH	A: 0.239 (792) G: 0.761 (2526)	
AA	0.13 (13%)	A: 0.22  G: 0.78	OTHER	A: 0.213 (458) G: 0.787 (1692)	
AG	0.17 (17%)				
GG	0.70 (70%)				
Patients (n = 23)					

Nota: Frecuencias alélicas y genotípicas de la variante rs7796976 presente en la población muestreada. Además, se muestran las frecuencias alélicas para la misma variante en diferentes poblaciones del mundo, reportadas por las bases de datos: proyecto 1000Genomes, ExAC y GnomAD.

***AHR [c.908+33G>T]***. Esta alteración de secuencia fue encontrada en el 39% (n = 9) del total de los pacientes estudiados, en heterocigosis. Variante de tipo no codificante ubicada en el intrón 7 del gen, e identificada bajo el RefSNP: rs2074113. En la figura 16 se muestra el cromatograma de un paciente con esta alteración.

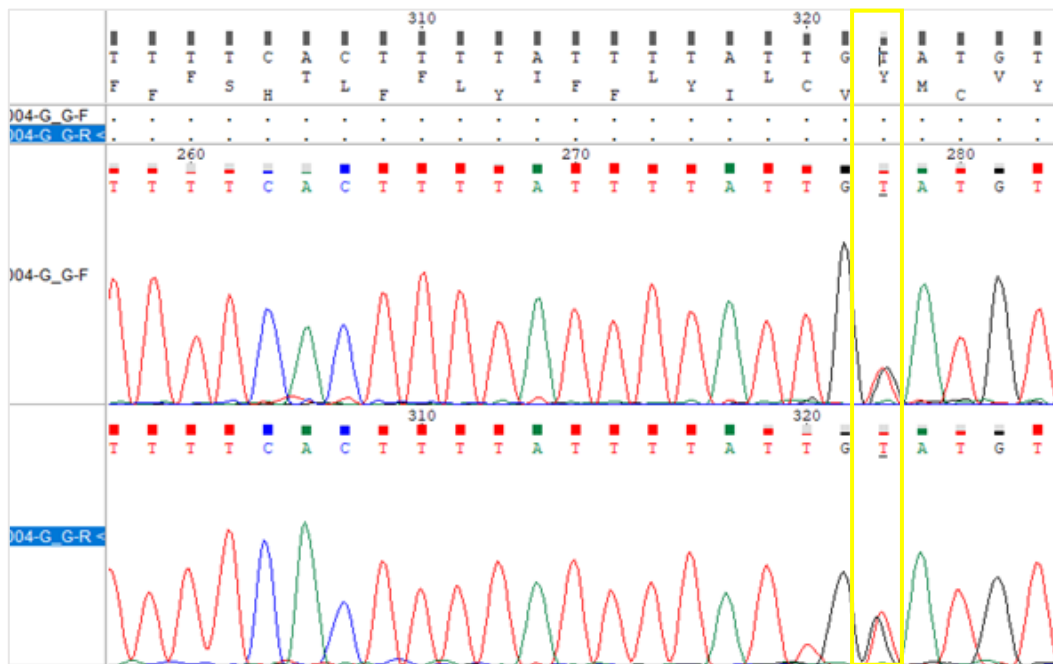
De acuerdo con las bases de datos dbSNP y Ensembl Genome browser, Se han encontrado reportes de la variante c.908+33G>T en población asiática: China (Chen et al, 2009), Korea (Kim et al, 2007) y población caucásica: España (Figueroa et al, 2008). Para Latinoamérica, se encuentran registros de la frecuencia de esta variante en las poblaciones de Medellín, Colombia; Lima, Perú; Puerto Rico; de ancestría mexicana habitantes de Los Ángeles, CA (Ensembl Genome browser). Sin embargo, de manera general esta alteración presenta una distribución global como resultado de los estudios del proyecto 1000Genomes, ExAC y GnomAD (Tabla 10). No se encuentran datos reportados de esta alteración en otras bases de datos como GWAS Catalog, TCGA y el HapMap Project.

De acuerdo con el cálculo de frecuencias para la población muestreada, la variante [c.908+33G>T] presentó un valor de frecuencia alélica menor de 0.20 para su alelo «T», lo cual se relaciona con el valor de MAF reportado por dbSNP para este alelo (MAF: T= 0.179), indicando una frecuencia de 18% del alelo c.908+33T en la población de la fase III del proyecto 1000Genomes (Tabla 10).

Por otra parte, no se encontraron portadores homocigotos de la alteración «TT» en la población de estudio. Sin embargo, el 39% de los pacientes analizados presentó la variante en estado de heterocigosis, mientras que el genotipo silvestre «GG» se observó en el 61% de la población (Tabla 10). Actualmente, esta variante no se encuentra reportada en el proyecto GTE<sub>x</sub>.

**Figura 16.**







*Cromatograma de la alteración [c.908+33G>T].*



Nota: Cromatograma, identificación de la variante c.908+33G>T en heterocigosis. En el recuadro se observa la alteración. Visualización mediante el programa ChromasPro versión 2.1.10.

**Tabla 10.**

*Frecuencias alélicas de la variante [c.908+33G>T].*

rs2074113			
Allele Frequencies			
1000 Genomes Project Phase 3		GnomAD (Ensembl); ExAC	
ALL	G: 0.821 (4114) T: 0.179 (894)		ALL
AFRICAN	G: 0.848 (1121) T: 0.152 (201)		AFRICAN / AFR-AME
AMERICAN	G: 0.863 (599) T: 0.137 (95)		LATINO
EAST ASIAN	G: 0.642 (647) T: 0.358 (361)		EAST ASIAN
EUROPEAN	G: 0.896 (901) T: 0.104 (105)		NON FINNISH EUR.
SOUTH ASIAN	G: 0.865 (846) T: 0.135 (132)		FINNISH
Sample Population: rs2074113			SOUTH ASIAN
Genotype	Genotype Frequencies	Allele Frequencies	AMISH
GG	0.61 (61%)	G: 0.80  T: 0.20	ASHKENAZI JEWISH
GT	0,39 (39%)		OTHER
TT	0.00 (0%)		
Patients (n = 23)			

Nota: Frecuencias alélicas y genotípicas de la variante rs2074113 presente en la población muestreada. Además, se muestran las frecuencias alélicas para la misma variante en diferentes poblaciones del mundo, reportadas por las bases de datos: proyecto 1000Genomes, ExAC y GnomAD.

*AHR [c.132T>C]*. Esta alteración se encontró en un solo individuo del total de los pacientes estudiados. Variante codificante de tipo sinónima (p.Asn44Asn), identificada bajo el RefSNP: rs17779352. De acuerdo con la posición aminoacídica «44» situada en el exón 2 del gen, la variante se encuentra en la región del dominio basic helix-loop-helix-Per-ARNT-Sim de la proteína AhR, en un sitio específico de unión al ADN (NCBI Reference Sequence: NP\_001612.1). En la figura 17 se muestra el cromatograma de un paciente con esta alteración.

En consecuencia, con los resultados de los estudios del proyecto 1000Genomes, ExAC y GnomAD publicados en las bases de datos dbSNP y Ensembl Genome browser, la variante [c.132T>C] presenta una distribución global con reportes en población caucásica: Italia (Ames et al, 2018), Alemania (Prager et al, 2016); población mezclada, Canadá (Ng et al, 2010). Para Latinoamérica, se encuentran registros de las frecuencias de esta variante en las poblaciones de Medellín, Colombia; Lima, Perú; Puerto Rico; de ancestría mexicana habitantes de Los Ángeles, CA (Ensembl Genome browser). En la tabla 11, se observan las frecuencias alélicas y genotípicas de la variante en la población de estudio y su distribución en el mundo. No se encuentran datos reportados de esta alteración en otras bases de datos como GWAS Catalog, TCGA y el HapMap Project.

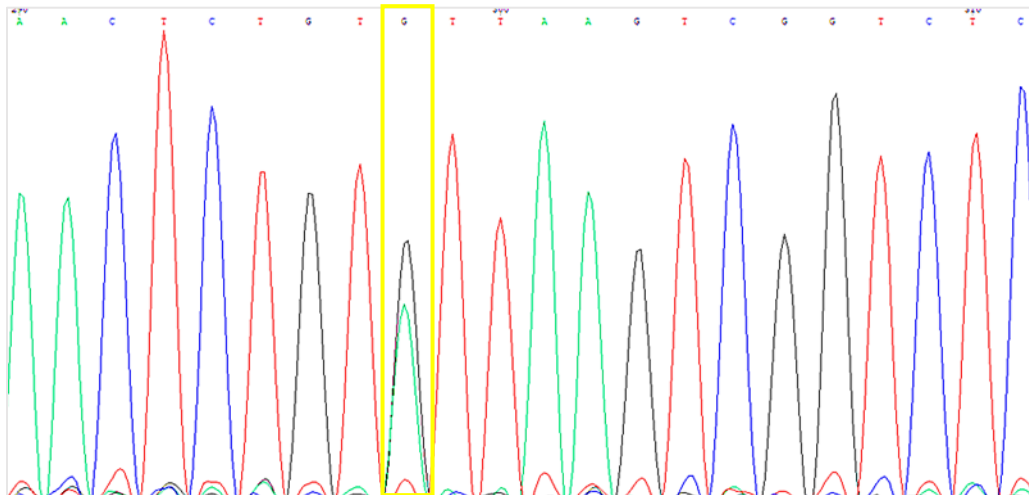
La variante [c.132T>C] exhibió un valor de frecuencia alélica menor de 0.02 para su alelo «C» (Tabla 11), la cual fue similar a las frecuencias observadas en la población de estudio de la

fase III del proyecto 1000Genomes reportado por dbSNP (MAF: C= 0.053), lo que indica que para la variante rs17779352, el alelo [c.132C] tiene una frecuencia de 5% en la población global (Tabla 11). De acuerdo con las frecuencias genotípicas, se observó un sólo portador de la variante en heterocigosis «TC» representando el 4% del total de los pacientes analizados, mientras que el 96% de la población restante no exhibió la alteración, es decir, presentaron un genotipo homocigoto del alelo ancestral «TT» (Tabla 11).

Conforme a los datos contenidos en el proyecto GTEx, la variante [c.132T>C] se relaciona con la expresión y regulación génica de *AHR* en 54 tejidos específicos, principalmente en cultivos primarios de fibroblastos (<http://varso.me/dkQE>).

### Figura 17.


*Cromatograma de la alteración [c.132T>C].*



Nota: Cromatograma, identificación de la variante c.132T>C en heterocigosis. En el recuadro se observa la alteración. Visualización mediante el programa ChromasPro versión 2.1.10.

**Tabla 11.**

*Frecuencias alélicas de la variante [c.132T>C].*

rs17779352					
Allele Frequencies					
1000 Genomes Project Phase 3			GnomAD (Ensembl); ExAC		
ALL	T: 0.947 (4745) C: 0.053 (263)		ALL	T: 0.932 (133487) C: 0.068 (9741)	
AFRICAN	T: 0.980 (1295) C: 0.020 (27)		AFRICAN / AFR-AME	T: 0.971 (40796) C: 0.030 (1242)	
AMERICAN	T: 0.968 (672) C: 0.032 (22)		LATINO	T: 0.961 (13113) C: 0.039 (537)	
EAST ASIAN	T: 0.957 (965) C: 0.043 (43)		EAST ASIAN	T: 0.969 (3033) C: 0.031 (97)	
EUROPEAN	T: 0.926 (932) C: 0.074 (74)		NON FINNISH EUR.	T: 0.912 (58899) C: 0.088 (5649)	
SOUTH ASIAN	T: 0.901 (881) C: 0.099 (97)		FINNISH	T: 0.855 (8933) C: 0.145 (1513)	
Sample Population: rs17779352			SOUTH ASIAN	T: 0.887 (2699) C: 0.113 (343)	
			AMISH	T: 0.847 (761) C: 0.153 (137)	
Genotype	Genotype Frequencies	Allele Frequencies T: 0.98 C: 0.02	ASHKENAZI JEWISH	T: 0.967 (3214) C: 0.033 (110)	
TT	0.96 (96%)		OTHER	T: 0.948 (2039) C: 0.053 (113)	
TC	0.4 (4%)				
CC	0.0 (0%)				
Patients (n = 23)					

Nota: Frecuencias alélicas y genotípicas de la variante rs17779352 presente en la población muestreada. Además, se muestran las frecuencias alélicas para la misma variante en diferentes poblaciones del mundo, reportadas por las bases de datos: proyecto 1000Genomes, ExAC y GnomAD.

**AHR [c.1661G>A].** Esta alteración fue encontrada en el 26% (n = 6) de los pacientes analizados. Variante codificante, de tipo missense o cambio de sentido, localizada en el exón 10 del gen, e identificada bajo el RefSNP: rs2066853. AHR- G1661A causa una substitución de Arginina por Lisina en el subdominio ácido del Dominio de Transactivación (TAD) del receptor AhR. En la figura 18 se muestra el cromatograma de un paciente con esta alteración.

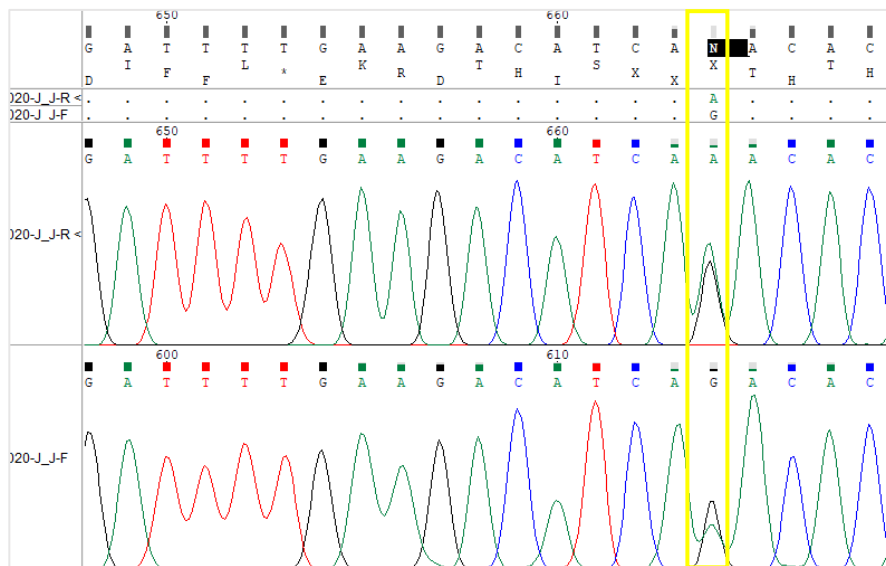
Según los resultados del proyecto 1000Genomes, ExAC y GnomAD contenidos en las bases de datos dbSNP y Ensembl Genome browser, la variante [c.1661G>A] presenta una distribución global; su frecuencia ha sido evaluada en población asiática: China (Zhang et al, 2002; Chen et al, 2009; Gu et al, 2012), Tailandia (Sangrajrang et al, 2009), India (Budhwar et al, 2018), Korea (Kim et al, 2007); población caucásica: España (Figuerola et al, 2008), Alemania (Prager et al, 2016); Francia (Cauchi et al, 2001); Suecia (Simonsson et al, 2016); población mezclada, Estados Unidos (Le et al, 2003), Canadá (Ng et al, 2010); población latina: México (Pérez et al, 2014; Sierra et al, 2018). Además, para Latinoamérica, se encuentran registros de la frecuencia de esta variante en las poblaciones de Medellín, Colombia; Lima, Perú; Puerto Rico; de ancestría mexicana habitantes de Los Ángeles, CA (Ensembl Genome browser). En la tabla 9, se observan las frecuencias alélicas y genotípicas de la variante en la población de estudio y su distribución en el mundo. No se encuentran datos reportados de esta alteración en otras bases de datos como GWAS Catalog, TCGA y el HapMap Project.

De acuerdo con el valor de frecuencia alélica menor reportado por dbSNP para el alelo «A» (MAF: A= 0.2708), el alelo c.1661A se presenta con una frecuencia de 27% en la población de estudio (1356 de 5008 individuos) de la fase III del proyecto 1000Genomes. Por otra parte, la frecuencia del alelo «A» en población latina, según datos reportados en GnomAD, presenta un valor de 0.153, similar a la frecuencia observada en la población de estudio (MAF: A= 0.13). Finalmente, sólo el 26% del total de los pacientes analizados fueron portadores de la variante en heterocigosis «GA», mientras el 74% de individuos restantes no exhibieron la alteración (genotipo en homocigosis para el alelo silvestre) (Tabla 12).

Conforme a los datos de expresión y regulación génica del proyecto GTEx, esta variante se relaciona con la expresión génica de *AHR* en 54 tejidos, principalmente en nervio tibial (<http://varso.me/37C8>).

**Figura 18.**

*Cromatograma de la alteración [c.1661G>A]*



Nota: Cromatograma, identificación de la variante c.1661G>A en heterocigosis. En el recuadro se observa la alteración. Visualización mediante el programa ChromasPro versión 2.1.10.

**Tabla 12.**

*Frecuencias alélicas de la variante [c.1661G>A]*

rs2066853			
Allele Frequencies			
1000 Genomes Project Phase 3		GnomAD (Ensembl); ExAC	
ALL	G: 0.729 (3652) A: 0.271 (1356)	G: 73% A: 27%	ALL
AFRICAN	G: 0.538 (711) A: 0.462 (611)	G: 54% A: 46%	AFRICAN / AFR-AME
AMERICAN	G: 0.827 (574) A: 0.173 (120)	G: 83% A: 17%	LATINO
EAST ASIAN	G: 0.630 (635) A: 0.370 (373)	G: 63% A: 37%	EAST ASIAN
EUROPEAN	G: 0.892 (897) A: 0.108 (109)	G: 89% A: 11%	NON FINNISH EUR.
SOUTH ASIAN	G: 0.854 (835) A: 0.146 (143)	G: 85% A: 15%	FINNISH
Sample Population: rs2066853			SOUTH ASIAN
Genotype	Genotype Frequencies	Allele Frequencies	AMISH
GG	0.74 (74%)	G: 0.87  A: 0.13	AMISH
AG	0.26 (26%)		ASHKENAZI JEWISH
AA	0.0 (0%)		OTHER
Patients (n = 23)			OTHER
			OTHER

Nota: Frecuencias alélicas y genotípicas de la variante rs2066853 en la población de estudio, y su distribución global de acuerdo con el proyecto 1000Genomes, ExAC y GnomAD.

**4.2.2.2 Análisis de variantes alélicas «no reportadas» para el gen AHR.** De las 7 alteraciones de secuencia halladas en los 23 pacientes analizados se encontraron 3 (43%) variantes no reportadas previamente, por lo que han sido consideradas en el presente estudio como variantes de importancia clínica desconocida. En consecuencia, se identificaron 2 variantes de región no codificante, localizadas en los intrones 1 y 2, que involucran un cambio de un único nucleótido [AHR-c.253+175T>C] y una delección [AHR-c.66-165delT], respectivamente. Además, se encontró 1 variante localizada en la región no traducida 3' (3'UTR), que involucra también, un cambio de un único nucleótido [AHR-c. 2333C>T] (Tabla 13).

**Tabla 13.**

*Variantes de significado clínico desconocido halladas en intrones 1 y 2 y el exón 11 del gen AHR.*

Nomenclatura	Alelos	Clasificación de las Variantes	N°. pacientes	ID de la muestra	Edad de diagnóstico	Localización del Tumor	Cancer familiar	
							HNC	Otros Canceres
[c.253+175T>C]	Heterocigosis	Variante intronica	1/23	CaHN015	54	SCC - Lengua	N/A	N/A
[c.66-165delT]		Variante intronica	3/23	CaHN001	66	SCC - Lengua	N/A	N/A
				CaHN022	30	SCC - Lengua	N/A	N/A
				CaHN024	47	SCC - Laringe	N/A	Hígado, Leucemia
				CaHN010	53	SCC - Piso de la boca	1 caso	N/A
[c.2333C>T]	Heterocigosis	Variante 3 prima UTR	7/23	CaHN012	48	SCC - Lengua	N/A	N/A
				CaHN014	30	SCC - Lengua y piso de la boca	N/A	Pulmón, Leucemia
				CaHN015	54	SCC - Lengua	N/A	N/A
				CaHN019	65	SCC - Laringe	N/A	N/A
				CaHN020	43	SCC - Laringe	N/A	N/A
				CaHN023	53	SCC - Amígdala	N/A	Mama, Útero

Nota: Edad de diagnóstico (Age of diagnosis) dada en años; (SCC) Carcinoma escamocelular, por sus siglas en inglés. N/A: no aplica.

### **4.3 Fase III: evaluación del efecto de las variantes identificadas sobre la función del receptor Aril hidrocarburos mediante análisis de predicción *in silico*.**

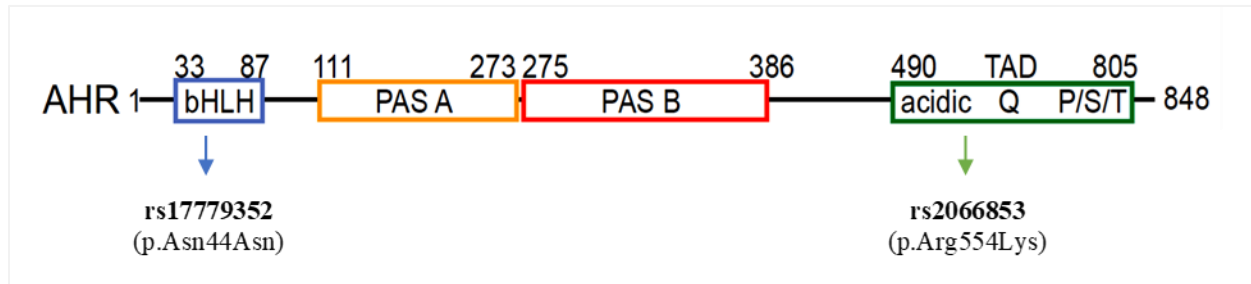
Para evaluar el posible impacto de las alteraciones de secuencia identificadas en este estudio sobre la función de la proteína, se realizó un análisis computacional de cada variante mediante el uso de herramientas de predicción *in silico*. Además, se determinó su importancia clínica de acuerdo con las bases de datos ClinVar, PathoMAN, Varsome y ClinGen.

#### ***4.3.1 Análisis *in silico* de las alteraciones en la región codificante***

***AHR c.132T>C [rs17779352]***. Esta variante localizada en el exón 2 del gen, representa un cambio de un único nucleótido en la posición genómica chr7:17310002T>C (GRCh38/hg38), resultando en un cambio de tipo sinónimo, sin alteración de la secuencia aminoacídica que conforma la proteína (p.Asn44Asn). Está alteración se ubica en el dominio basic helix-loop-helix del receptor (Fig. 19) (Schulte et al, 2017), donde la posición «44» corresponde a uno de los sitios específicos de unión al ADN (NP\_001612.1).

**Figura 19.**

*Representación esquemática de los dominios de AHR.*



Nota: Los números en los límites de cada dominio corresponden a los aminoácidos de la proteína humana AhR. Variante sinónima rs17779352, localizada en el dominio basic helix-loop-helix del receptor; variante missense rs2066853, ubicada en el dominio de transactivación de la proteína. Modificado de Schulte et al, 2017.

A partir del análisis de predicción in silico, todos los algoritmos utilizados coincidieron en la clasificación del efecto de la variante rs17779352 sobre la proteína. Según la herramienta VEP, la alteración se consideró de riesgo bajo “Low”. De igual forma, Mutation Taster, reportó la alteración como “benigna” según su modelo de predicción Random forest: Tree vote = 2|99, lo que significa que 2 árboles de decisión en Random Forest han indicado nocividad y 99 árboles de decisión han indicado una alteración benigna. Además, No se prevé que tenga consecuencias que alteren el empalme (HSF score: N/A). En la tabla 14 se muestra la clasificación y puntajes (scores) obtenidos de cada programa para la variante en cuestión. De acuerdo con los resultados obtenidos, la alteración se encuentra reportada en diferentes bases de datos, entre estas ClinVar, PathoMAN y Varsome con significancia clínica: «Benigna» (Tabla 14).

Con relación al valor de conservación nucleotídica mediante el método de PhastCons, que estima la probabilidad de que cada nucleótido pertenezca a un elemento conservado, basado en el

alineamiento múltiple (Felsenstein y Churchill, 1996; Raney et al, 2014), se encontró que el nucleótido «T» de la variante c.132T>C es altamente conservado, con un score de PhastCons100way de [1.000], evidencia de conservación estructural y funcional (Tabla 14).

**Tabla 14.**

*Análisis de predicción in silico para la variante AHR [c.132T>C] y su significancia clínica.*

Fragmento	Posición	Nomenclatura	Cambio de A.A.	dbSNP	Clasificación de la Variante	
Exón 2	GRCh38/hg38: chr7:17310002	<b>[c.132T&gt;C]</b>	p.Asn44Asn	rs17779352 T/C	Variante codificante de tipo sinónimo	
<b>IMPACTO DE PATOGENICIDAD (SCORES)</b>						
Predicción VEP (Impacto)	Predicción SIFT (Impacto)	Predicción Mutation Taster (Impacto)	Predicción PROVEAN (Impacto)	Predicción HSF (Impacto)	Scores de Conservación	
Baja "Low"	Tolerada [Score=1]	Benigna [Model: without_aae; Tree vote <sub>2</sub>  99]	Neutral [Score=0.00]	No significant impact on splicing signals.	PhastCons 100way: 1.000	PhyloP 100way: 0.669
<b>SIGNIFICANCIA CLÍNICA</b>						
CLINVAR	PATHOMAN	VARSome	CIVIC	CLINGEN		
Benigna	Probablemente benigna	Benigna [ACMH classification]	No reportada	Indeterminada		

Nota: \*Nomenclatura aprobada por HGVS (Human Genome Variation Society).

Finalmente, esta alteración de secuencia se encontró en una paciente del departamento de Santander con diagnóstico de Carcinoma escamocelular de lengua a la edad de 63 años, quien a su vez reportó antecedente familiar de cáncer de cabeza y cuello, para el caso concreto madre de la paciente con desarrollo de cáncer de lengua.

**AHR c.1661G>A [rs2066853].** Variante localizada en la posición genómica chr7:17339486G>A (GRCh38/hg38) del exón 10 del gen; representa un cambio de un único

nucleótido que genera una alteración de secuencia de tipo “missense” o de sentido erróneo, que resulta en la sustitución de Arginina por Lisina en la posición 554 (p.Arg554Lys) de la proteína del receptor Aril de hidrocarburos. Esta variante se ubica en la región acídica del dominio de transactivación del receptor (TAD) (Schulte et al, 2017), lugar de unión de coactivadores al factor de transcripción AhR, permitiendo que se lleve a cabo la transcripción de un gen o conjunto de genes de manera efectiva (Spiegelman y Heinrich, 2004) (Fig. 19).

A partir de los análisis de predicción *in silico*, la variante rs2066853 fue clasificada como alteración de efecto “benigno” o “tolerada” sobre la estructura y función de la proteína, conforme con los resultados obtenidos por todos los algoritmos evaluados. La herramienta VEP, consideró la variante de riesgo moderado “Moderate” (variante no disruptiva que podría cambiar la eficacia de las proteínas), en consonancia con la ontología de secuencia (SO) [SO:0001583]. A su vez, SIFT y Polyphen la reportaron como “Tolerada”, con scores [1.00 y 2.19], respectivamente. Mutation Taster, catalogó la alteración como “benigna”: Tree vote = 1|99, lo que significa que 1 árbol de decisión en Random Forest indica nocividad, mientras que 99 árboles de decisión han indicado una alteración de tipo benigna. En la tabla 15 se muestra la clasificación y puntajes (scores) obtenidos de cada programa para la variante en cuestión.

Por otra parte, una vez realizada la corrida con el programa HSF, se observó que la variante c.1661G>A presenta un impacto de alteración de secuencias auxiliares, específicamente 8 alteraciones significativas de motivos ESE (Exonic Splicing Enhancer), de los cuales, 2 corresponderían a nuevos sitios creados por la variante (matrices para identificar los motivos alterados: RESCUE ESE, y ESE\_Tra2); mientras que los 6 restantes corresponden a motivos

suprimidos por la alteración: (matrices para identificar los motivos alterados: RESCUE ESE; EIE; ESE\_9G8; ESE\_ASF) (Apéndice E. Tabla 16). Además, de acuerdo con la herramienta UCSC Genome Browser (ENCODE annotation data), el SNP rs2066853 se localiza dentro de la región enhancer (EH38E2536893) del gen *AHR*. EH38E2536893 es un elemento regulador en CIS distal (CREs) ubicado a 2 kb del sitio de inicio de transcripción (TSS), y contiene diversos patrones de modificación de histonas (H3K36me3, H3K4me1, H4K20me1) (Kent et al, 2002; Rosenbloom et al, 2013).

Con relación al valor de conservación nucleotídica mediante el método de PhastCons, se encontró que el nucleótido «G» de la variante c.1661G>A no es muy conservado, con un score de PhastCons100way de [0.325]. La alteración se encuentra reportada en ClinVar y Varsome con significancia clínica «Benigna». Sin embargo, PathoMAN la reporta como Variante de Significado Incierto (VUS) (Tabla 15).

**Tabla 15.**

*Análisis de predicción in silico para la variante AHR [c.1661G>A] y su significancia clínica.*

Fragmento		Posición		Nomenclatura		Cambio de A.A		dbSNP		Clasificación de la Variante		
Exón 10		GRCh38/hg38: chr7:17339486		[c.1661G>A]		p.Arg554Lys		rs2066853		Variante codificante de tipo missense o cambio de sentido		
IMPACTO DE PATOGENICIDAD												
Predicción VEP (Impacto)		Predicción SIFT (Impacto)		Predicción Mutation Taster (Impacto)		Predicción PolyPhen (Impacto)		Predicción PROVEAN (Impacto)		Predicción FATHMM (Impacto)		
Moderado "Moderate"		Tolerada [Score=1]		Benigna [Model: without_aae; Tree vote 99]		Benigna [Score=0]		Neutral [Score=1.21]		Tolerada [Score=2.19]		
										Alteración de secuencias auxiliares: Alteración significativa de motivos ESE / ESS		
										PhastCons 100way: 1.000		
										PhyloP 100way: 0.669		
SIGNIFICANCIA CLÍNICA												
CLINVAR			PATHOMAN			VAR SOME			CIVIC		CLINGEN	
Benigna			Variante de significado clínico incierto (VUS)			Benigna [ACMH classification]			No reportada		Indeterminada	

Nota: \*Nomenclatura aprobada por HGVS (Human Genome Variation Society).

Esta alteración se encontró en 6 pacientes, 3 del departamento de Nariño, 2 del departamento de Santander y 1 paciente en Norte de Santander; En la tabla 17 se describen los pacientes con esta variante.

**Tabla 17.**

*Pacientes analizados, portadores de la variante p.Arg554Lys (rs2066853).*

Lugar de origen	ID de la muestra	Edad de diagnóstico	Localización del Tumor	Cáncer familiar	
				Parentesco	Tipo de cáncer
Pasto, Nariño	CaHN006	40	SCC - Paladar	N/A	N/A
Pasto, Nariño	CaHN007	43	SCC - Laringe	N/A	N/A
Pasto, Nariño	CaHN026	49	SCC - Lengua	Tía Prima	Tiroides Linfoma
Bucaramanga, Santander	CaHN022	30	SCC - Lengua	N/A	N/A
Bucaramanga, Santander	CaHN023	53	SCC - Amígdala	Tía Tía	Mama Útero
Cúcuta, Norte de Santander	CaHN024	47	SCC - Amígdala	Abuelo Madre	Hígado Leucemia

Nota: Edad de diagnóstico (Age of diagnosis) dada en años; (SCC) Carcinoma escamocelular, por sus siglas en inglés. N/A: no aplica.

De acuerdo con lo anterior, no se podría descartar la evaluación de la variante rs2066853 mediante futuros estudios in vivo, que permitan determinar su posible efecto sobre la estructura y función de la proteína. Dada la localización de la variante, es pertinente mencionar que a la actividad transcripcional del receptor también contribuyen cofactores transcripcionales, de los cuáles, algunos se unen al AhR en la región del TAD (Kumar et al. 2001), facilitando la unión del complejo AhR/ARNT a la región promotora de genes diana, por lo que dicha alteración podría generar algún efecto sobre las propiedades del TAD y la regulación de la transcripción (Aftabi et al, 2016).

#### *4.3.2 Análisis in silico de las alteraciones en la región no codificante*

**AHR -459A>G [rs7796976]**. Esta variante representa un cambio de un único nucleótido en la posición genómica chr7:17298806A>G (GRCh38/hg38) y se localiza en la región no traducida cinco prima del gen (5 prime UTR variant), región de importancia para la regulación de la expresión génica.

De acuerdo con los resultados obtenidos a partir de los análisis de predicción in silico, la variante rs7796976 fue clasificada, según su efecto sobre la estructura y función de la proteína, de tipo “benigna”. La herramienta VEP, consideró el SNP de riesgo “Modifier” (variante no codificante o variante que afecta genes no codificantes, donde las predicciones son dificultosas o no hay evidencia de impacto), en consonancia con la ontología de secuencia (SO) [SO:0001623]. A su vez, FATHMM la reportó como “Benigna (con alta confianza)”, con un score de [0.010100]. Mutation Taster, catalogó la alteración como “benigna”: Tree vote = 1|299, lo que significa que 1 árbol de decisión en Random Forest indicó nocividad, mientras que 299 árboles de decisión indicaron una alteración de tipo benigna. En la tabla 18 se muestra la clasificación y puntajes (scores) obtenidos de cada programa para la variante en cuestión.

La alteración, también fue evaluada mediante el uso de los programas HSF, HaploReg y la herramienta UCSC Genome Browser. Una vez realizados los análisis, se encontró que la variante -459A>G presenta un impacto de alteración de secuencias auxiliares, específicamente alteración significativa de motivos ESE/ESS (Exonic Splicing Enhancer / Silencer) según la herramienta de

predicción HSF. En total se observaron 4 posibles alteraciones de inactivación de ESE y 7 alteraciones de activación ESS. De las primeras (ESE), 1 corresponde a un nuevo sitio creado por la variante, mientras que los 3 restantes corresponden a motivos suprimidos por la alteración (Matrices para la identificación de las alteraciones: EIE; ESE\_9G8; RESCUE ESE). De las segundas (ESS), 2 corresponden a motivos suprimidos por la alteración (matrices usadas para la identificación de las alteraciones: Fas ESS; ESS\_hnRNPA1); los otros 5 coinciden con nuevos sitios creados por la variante (matrices usadas para la identificación de las alteraciones: Sironi\_motif2; Fas ESS; IIE) (Apéndice F. Tabla 19).

A partir del análisis del UCSC Genome Browser (ENCODE annotation data), el SNP rs7796976 se localiza dentro de la región enhancer (EH38E2536847) del gen *AHR*. EH38E2536847 es un elemento regulador en CIS próximo (CRE) ubicado dentro de los 2 kb del sitio de inicio de transcripción (TSS) y contiene diversos patrones de modificación de histonas (H3K4me3, H3K27ac, H3K9ac) (Kent et al, 2002; Rosenbloom et al, 2013).

Además de los CREs (EH38E2536847 y los patrones de modificación de histonas) identificados con UCSC Genome Browser, la herramienta HaploReg v4.1 permitió evaluar otras anotaciones funcionales del genoma no codificante que se colocan con el rs7796976, entre éstas, se encontró que la ubicación genómica del SNP incluye un sitio potencial de unión de distintos factores de transcripción (POLR2A; CTCF; TAF1; TBP; NF- $\kappa$ B) en diversas líneas celulares, según datos reportados de los experimentos ENCODE ChIP-Seq (Apéndice G. Tabla 20). También se analizó el efecto del rs7796976 sobre los motivos reguladores de acuerdo con experimentos de unión de Factores de Transcripción previamente reportados (ENCODE TF binding experiments), y a partir de una matriz de pesos posicionales (PWM), se observó que el

SNP rs7796976 (A/G) presenta un efecto en 6 de los 11 motivos reguladores de NF-Kb (<http://compbio.mit.edu/encode-motifs/>) (Apéndice H. Tabla 21).

Con relación al valor de conservación nucleotídica mediante el método de PhastCons, se encontró que el nucleótido «G» de la variante -459A>G no es conservado, con un score de PhastCons100way de [0.432]. La alteración no se encuentra reportada en ClinVar, CIVIC o ClinGen; PathoMAN reporta este SNP como Variante de significado incierto (VUS) (Tabla 18).

**Tabla 18.**

*Análisis de predicción in silico para la variante AHR [-459A>G] y su significancia clínica*

Fragmento		Posición	Nomenclatura		Cambio de A.A	dbSNP	Clasificación de la Variante		
Exon 1		GRCh38/HG38: chr7:17298806	[-459A>G]		N/A	rs7796976 A/G	Variante 5 prime UTR		
IMPACTO DE PATOGENICIDAD									
Predicción VEP (Impacto)	Predicción SIFT (Impacto)	Predicción Mutation Taster (Impacto)	Predicción PolyPhen (Impacto)	Predicción PROVEAN (Impacto)	Predicción FATHMM (Impacto)	Predicción HSF (Impacto)		Scores de Conservación	
"Modifier"	N/A	Benigna [Model: 5utr; Tree vote  299]	N/A	N/A	Benigna [Score=0.010100]	Alteración de secuencias auxiliares: Alteración significativa de motivos ESE / ESS		PhastCons 100way: 0.432	PhyloP 100way: 0.579
SIGNIFICANCIA CLÍNICA									
CLINVAR		PATHOMAN			VAR SOME		CIVIC	CLINGEN	
No reportada		Variante de significado clínico incierto (VUS)			Benigna [ACMH classification]		No reportada	Indeterminada	

Nota: \*Nomenclatura aprobada por HGVS (Human Genome Variation Society)

Esta alteración se encontró en 20 pacientes, 4 del departamento de Nariño y 16 del departamento de Santander; En la tabla 22 se describen los pacientes con esta variante.

**Tabla 22.**

*Pacientes analizados, portadores de la variante -459A>G (rs7796976).*

Lugar de Nacimiento	ID de las muestras	Edad de diagnóstico	Localización del Tumor	Cáncer familiar	
				Parentesco	Tipo de cáncer
Pasto, Nariño	CaHN004	63	SCC - Laringe	N/A	N/A
Pasto, Nariño	CaHN006	40	SCC - Paladar	N/A	N/A
Pasto, Nariño	CaHN007	43	SCC - Laringe	N/A	N/A
Pasto, Nariño	CaHN026	49	SCC - Lengua	Tía Prima	Tiroides Linfoma
San Gil, Santander	CaHN009	73	SCC - Lengua	Madre Hermana Hija	Hígado Mama Mama
Bucaramanga, Santander	CaHN010	53	SCC - Piso de la boca	Tío	Cabeza y cuello
Bucaramanga, Santander	CaHN011	53	SCC - Amígdala	N/A	N/A
Bucaramanga, Santander	CaHN012	48	SCC - Lengua	N/A	N/A
Socorro, Santander	CaHN013	63	SCC - Lengua	Madre	Cabeza y cuello
Bucaramanga, Santander	CaHN014	30	SCC - Lengua y Piso de la boca	Primo Sobrino	Pulmón Leucemia
Bucaramanga, Santander	CaHN016	55	SCC - Lengua	Padre Hermano Hermana	Cabeza y cuello Leucemia Mama
Bucaramanga, Santander	CaHN017	33	SCC - Lengua	Abuela	Mama
Bucaramanga, Santander	CaHN018	77	SCC - Cavidad oral	Madre	Ovario
Bucaramanga, Santander	CaHN019	65	SCC - Laringe	N/A	N/A
Bucaramanga, Santander	CaHN020	43	SCC - Laringe	N/A	N/A
Bucaramanga, Santander	CaHN021	83	SCC - Lengua	N/A	N/A
Bucaramanga, Santander	CaHN022	30	SCC - Lengua	N/A	N/A
Bucaramanga, Santander	CaHN023	53	SCC - Amígdala	Tía Tía	Mama Útero
Cúcuta, Norte de Santander	CaHN015	54	SCC - Lengua	N/A	N/A
Cúcuta, Norte de Santander	CaHN024	47	SCC - Laringe	Abuelo Madre	Hígado Leucemia
Bábeca, Norte de Santander	CaHN025	58	SCC - Piso de la boca	N/A	N/A

Nota: Edad de diagnóstico (Age of diagnosis) dada en años; (SCC) Carcinoma escamocelular, por sus siglas en inglés. N/A: no aplica.

**AHR c.908+33G>T [rs2074113]**. Polimorfismo de un solo nucleótido localizado en la posición genómica chr7:17334147G>T (GRCh38/hg38) de la región no codificante del gen (intrón 7); región de importancia durante el proceso de traducción (splicing).

De acuerdo con los análisis de predicción *in silico*, la variante rs2074113 fue clasificada de tipo “benigna” según su efecto sobre la función de la proteína. La herramienta VEP, consideró el SNP de riesgo “Modifier” (variante no codificante, donde las predicciones son dificultosas o no hay evidencia de impacto), en consonancia con la ontología de secuencia (SO) [SO:0001627]. A su vez, Mutation Taster, catalogó la alteración como “benigna”: Tree vote = 0|100, lo que significa que ningún árbol de decisión en Random Forest indicó nocividad, mientras que 100 árboles de decisión indicaron una alteración de tipo benigna. FATHMM reportó la variante como “Benigna”, con un score de [0.267377]. En la tabla 23 se muestra la clasificación y puntajes (scores) obtenidos de cada programa para la variante en cuestión.

La alteración, también fue evaluada mediante el uso de los programas HSF, MaxEntScan, UCSC Genome Browser y HaploReg. Conforme con los resultados del HSF, la alteración c.908+33G>T presentó un impacto potencial de alteración de splicing, mediante la activación de un sitio de empalme críptico (New Donor splice site) [Score Reference: 49.25 | Score Mutation: 76.39]. Lo anterior se correlacionó con los datos obtenidos por el algoritmo MaxEntScan, predicción: potencial alteración de splicing (New Donor splice site) [Score Reference: -3.31 | Score Mutation: 4.34] (Apéndice I. Tabla 24).

A partir del análisis del UCSC Genome Browser (ENCODE annotation data), se encontró que la ubicación genómica del SNP se superpone posicionalmente con marcas activas (H3K27Ac)

en 7 tipos de células diferentes de ENCODE. Por otra parte, HaploReg v4.1, permitió evaluar el efecto del rs2074113 sobre motivos reguladores de acuerdo con experimentos de unión de Factores de Transcripción (ENCODE TF binding experiments), y a partir de una matriz de pesos de posición (PWM), se observó que el SNP (G/T) podría presentar un efecto en diferentes motivos reguladores, entre estos, Foxp1; Hoxb13; Irf\_disc1; TEF; Zfp410 (Apéndice J. Tabla 25).

Con relación al valor de conservación nucleotídica mediante el método de PhastCons, se encontró que el nucleótido «G» de la variante c.908+33G>T no es conservado, con un score de PhastCons100way de [0.010]. La alteración no se encuentra reportada en ClinVar, CIVIC o ClinGen; mientras que PathoMAN y Varsome la reportan como Variante Benigna (Tabla 23).

**Tabla 23.**

*Análisis de predicción in silico para la variante AHR [c.908+33G>T] y su significancia clínica*

Fragmento		Posición		Nomenclatura		Cambio de A.A		dbSNP		Clasificación de la Variante		
Intron 7		GRCh38/hg38: chr7:17334147		[c.908+33G>T]		N/A		rs2074113		Variante intrónica		
<b>IMPACTO DE PATOGENICIDAD</b>												
Predicción VEP (Impacto)	Predicción SIFT (Impacto)	Predicción Mutation Taster (Impacto)	Predicción PolyPhen (Impacto)	Predicción PROVEAN (Impacto)	Predicción FATHMM (Impacto)	Predicción HSF (Impacto)		Scores de Conservación				
"Modifier"	N/A	Benigna [Model: without_aae; Tree vote@ 100]	N/A	N/A	Benigna [Score=0.267377]	Nuevo sitio donador de splicing: Activación de un sitio donador criptico. Alteración potencial de splicing (HSF; MaxEnt)		PhastCons 100way: 0.010	PhyloP 100way: 0.251			
<b>SIGNIFICANCIA CLÍNICA</b>												
CLINVAR			PATHOMAN			VAR SOME			CIVIC		CLINGEN	
No reportada			Benigna			Benigna [ACMH classification]			No reportada		Indeterminada	

Esta alteración se encontró en 9 pacientes, 1 del departamento de Norte de Santander, 2 del departamento de Nariño y 6 del departamento de Santander; En la tabla 26 se describen los pacientes con esta variante

**Tabla 26.**

*Pacientes analizados, portadores de la variante c.908+33G>T (rs2074113).*

Lugar de nacimiento	ID de la muestra	Edad de diagnóstico	Localización del Tumor	Cáncer familiar	
				Parentesco	Tipo de cáncer
Pasto, Nariño	CaHN007	43	SCC - Laringe	N/A	N/A
Pasto, Nariño	CaHN026	49	SCC - Lengua	Tía Prima	Tiroides Linfoma
Bucaramanga, Santander	CaHN011	53	SCC - Amígdala	N/A	N/A
Bucaramanga, Santander	CaHN012	48	SCC - Lengua	N/A	N/A
Bucaramanga, Santander	CaHN017	33	SCC - Lengua	Abuela	Mama
Bucaramanga, Santander	CaHN018	77	SCC - Cavidad oral	Madre	Ovario
Bucaramanga, Santander	CaHN022	30	SCC - Lengua	N/A	N/A
Bucaramanga, Santander	CaHN023	53	SCC - Amígdala	Tía Tía	Mama Ovario
Cúcuta, Norte de Santander	CaHN024	47	SCC - Laringe	Abuelo Madre	Hígado Leucemia

Nota: Edad de diagnóstico (Age of diagnosis) dada en años; (SCC) Carcinoma escamocelular, por sus siglas en inglés. N/A: no aplica.

#### 4.4 Discusión de resultados

Múltiples factores como la susceptibilidad genética, los factores ambientales y la inmunidad, representan un rol fundamental en la patogénesis del cáncer (Li et al, 2020). Diversos estudios han demostrado el rol del AhR como un “sensor ambiental” que favorece la activación del metabolismo de los xenobióticos (Kawajiri y Fujii-Kuriyama, 2017; Guarnieri, Abruzzo y Bolotta, 2020) y un regulador fundamental de la diferenciación de las células T, que a su vez se relaciona con la vigilancia inmunológica del cáncer, lo que permite inferir que el gen *AHR* podría estar implicado en el proceso de la carcinogénesis (Murray, Patterson y Perdew, 2014; Chen y

Mellman, 2017). Sin embargo, no sólo la actividad protumorigénica del *AHR* ha sido evaluada, también se encuentran reportes sobre su función antitumorigénica, por lo que el papel del gen en la transformación maligna aún es objeto de estudio (Murray, Patterson y Perdew, 2014).

Esta investigación, permitió la identificación de variantes genéticas en *AHR* y a su vez, evaluar su posible efecto sobre la estructura y función de la proteína (AhR). Este estudio comprendió un total de 23 pacientes analizados, fumadores con diagnóstico de carcinoma escamocelular de cabeza y cuello (HNSCC). Tras el análisis de la secuenciación genómica del *AHR* para cada paciente, se logró identificar 7 alteraciones de secuencia, entre éstas, 4 variantes consideradas polimórficas, que se comparten con otras poblaciones de Latinoamérica, y 3 variantes nuevas, no reportadas previamente, 2 de ellas localizadas en regiones intrónicas y 1 en la región no traducida del gen (3'UTR).

Se ha descrito previamente que variantes en el gen *AHR* podrían estar involucradas en el efecto deletéreo del tabaquismo en distintas enfermedades, como enfermedad cardiovascular, hipertensión y cáncer (Gambier et al, 2006; Kim et al, 2007; Pérez et al, 2014; Huang et al, 2015; Moorthy et al, 2015; Prager et al, 2016). Los resultados obtenidos, en este estudio, para el polimorfismo rs2066853 [c.1661G>A] fueron consistentes con aquellos encontrados en la literatura (Prager et al, 2016; Gu et al, 2012; Li et al, 2015; Li et al, 2020) y de acuerdo con los análisis in silico, la variante no representa un impacto potencial o dañino sobre la estructura y función de la proteína, lo cual fue consistente con la importancia clínica (benigna) informada en las bases de datos de Varsome y ClinVar. Por lo tanto, se podría especular que la presencia de esta variante no contribuye de forma significativa al riesgo de desarrollar cáncer y se podría considerar

como variante benigna. No obstante, las bases de datos PathoMAN, Civic y ClinGen reportan esta variante como de significado clínico incierto, o indeterminada, por lo que los resultados del polimorfismo rs2066853 deben ser interpretados con precaución y confirmados mediante estudios posteriores de análisis in vivo.

El sitio rs2066853, localizado en el dominio de transactivación (TAD) del gen *AHR*, da lugar a una sustitución de aminoácidos no sinónimos que podría influir en la función de la proteína AhR. Estudios previos reportan que la variante R554K puede causar la pérdida de una parte de la estructura helicoidal de la proteína influyendo sobre las propiedades del TAD de la proteína AhR (Gu et al, 2012; Aftabi et al, 2016); además, otra investigación reveló que la variante alélica (A) del rs2066853 representaba una mayor capacidad de inducción de la transcripción de CYP1A1, fundamental durante el metabolismo y bioactivación de compuestos ambientales en metabolitos altamente cancerígenos (Murray, Patterson y Perdew, 2014; Smart y Daly, 2000). Sin embargo, otra investigación encontró que este polimorfismo no altera la función del receptor AhR en la regulación de la transcripción impulsada por CYP1A1 (Celius y Matthews, 2010), el hecho de que la presencia de esta variante pueda afectar el proceso de activación o detoxificación, de manera bidireccional, de los procarcinógenos regulados por la CYP1A1 podría sugerir que la alteración alélica (A) y los genotipos (AA y GA) de este polimorfismo puedan asociarse con una susceptibilidad indeterminada (aumento o disminución) al cáncer (Li et al, 2020).

Aunque los análisis in silico realizados en este estudio indican que la variante rs2066853 es benigna, no se podría descartar su posible relevancia clínica dada su ubicación dentro del dominio de transactivación del receptor, necesario para la función completa de AhR y la

acumulación nuclear después de la unión del ligando, lo que finalmente podría conllevar a una afectación sobre la función de la proteína. Además, se ha encontrado evidencia contradictoria sobre si esta variante altera la función de AhR en la regulación de la transcripción impulsada por CYP1A1.

Por otra parte, el resultado encontrado a partir de la herramienta USCS Genome Browser, esta alteración corresponde a un potenciador de empalme exónico (Exonic splicing enhancer – ESE), localizada dentro de un elemento regulador en cis, por lo que la presencia de esta variante podría causar alteración significativa de motivos ESE, que a su vez tendría un efecto sobre la regulación de la transcripción, lo que también se corroboró con los datos obtenidos del software Human Splicing Finder (Apéndice E). Además, diversos estudios han revelado que los genotipos alterados (AA o GA) del polimorfismo *AHR* rs2066853 confieren un mayor riesgo estadísticamente significativo de glioma (Gu et al, 2012), cáncer de pulmón (Chen et al, 2009), leucemia aguda (Guo et al, 2005) y cáncer de mama (Sangrajrang et al, 2009), mientras que otros estudios sugieren un riesgo reducido de cáncer de pulmón (Pérez et al, 2014; Budhwar et al, 2018) y cáncer de mama (Long et al, 2006; Pérez et al, 2014) en individuos con los genotipos (AA o GA) del sitio rs2066853. Finalmente, a partir de los resultados presentados en este estudio y de acuerdo con la literatura reportada, la variante rs2066853 se clasifica como probablemente patogénica. Se sugiere realizar estudios in vivo para determinar el posible efecto de esta variante sobre el receptor y su posible asociación con una susceptibilidad al cáncer.

Referente a la variante rs17779352 [c.132T>C], alteración de tipo sinónima, la evaluación de este polimorfismo con las distintas herramientas de análisis in silico, permitieron predecir que

esta alteración no genera cambio sobre la estructura o función de la proteína (Tabla 14), clasificada como benigna o neutral, y que además dichos resultados son consistentes con la importancia clínica (benigna) reportada en las diferentes bases de datos de Varsome, ClinVar y PathoMAN. Este hallazgo probablemente refleja que este polimorfismo no representa un papel fundamental en el desarrollo de cáncer. Por lo anterior, teniendo en cuenta los resultados de este estudio y de acuerdo con la literatura reportada, la variante rs17779352 se clasifica como benigna. Sin embargo, no existen reportes publicados que lo soporten, por lo que se sugiere confirmar mediante futuros análisis in vivo.

Con respecto al polimorfismo rs7796976 [-459A>G], se encontró que la variante alélica (G) podría generar un impacto de alteración de motivos ESE/ESS (Exonic Splicing Enhancer / Silencer) según la herramienta de predicción HSF (Apéndice F). Además, el análisis de UCSC Genome Browser (ENCODE annotation data), permitió determinar la localización de esta alteración dentro de un elemento regulador en cis próximo al sitio de inicio de transcripción (TSS), que a su vez contiene diversos patrones de modificación de histonas (H3K4me3, H3K27ac, H3K9ac), lo que podría representar una alteración de las características reguladoras de la región promotora del gen, patrones de modificación de histonas pueden presentar alteraciones específicas promoviendo la tumorigénesis. Se ha reportado que en HNSCC negativo para el virus del papiloma humano (VPH), alrededor del 13% de los casos tienen una desregulación del epigenoma, debido a mutaciones recurrentes de K36M y NSD1 que alteran el proceso de metilación de H3K36 con niveles reducidos de H3K36me3 y H3K36me2 promoviendo el desarrollo de cáncer (Papillon-Cavanagh et al, 2017).

De acuerdo con la ubicación genómica del polimorfismo, este incluye un sitio potencial de unión con distintos factores de transcripción (TF). Al analizar el efecto del rs7796976 sobre los motivos reguladores de los TF, se logró observar que la variante rs7796976 (A/G) presenta un efecto en 6 de los 11 motivos reguladores de NF- $\kappa$ B (Apéndice H). Existe evidencia de que AhR y NF- $\kappa$ B interactúan modulando funcionalmente la actividad del otro (Tian, Rabson, Gallo, 2002; Tian, 2009). La inducción de AhR por unión de ligando, tales como dioxinas o compuestos derivados del humo de tabaco, podría favorecer la unión AhR - NF- $\kappa$ B, promoviendo resistencia a la anoikis y constituyendo un mecanismo potencial de supervivencia de las células cancerígenas durante la progresión a la metástasis (Fig. 9). Sin embargo, de acuerdo a los análisis *in silico* realizados en este estudio, la variante rs7796976 fue clasificada como benigna de acuerdo con las herramientas de VEP, Mutation Taster y FATHMM. Además, dicha alteración se encuentra reportada con significancia clínica benigna por Varsome y por su alta frecuencia de presentación en la población se define como un polimorfismo genético. Teniendo en cuenta lo anterior y literatura reportada (Nebert et al, 2016; Li et al, 2020; Aarts et al, 2021), la variante rs7796976 se clasifica como benigna. Sin embargo, se sugieren nuevas investigaciones que permitan evaluar los mecanismos biológicos y los fenotipos clínicos para determinar el efecto de este polimorfismo de *AHR* en el desarrollo del cáncer.

Respecto a la variante rs2074113 [c.908+33G>T], esta se reporta como alteración de tipo benigna, es decir, sin efecto sobre la estructura y función del receptor. Lo anterior en relación con los resultados encontrados para tres de las cinco herramientas de predicción *in silico* evaluadas. Sin embargo, los programas HSF y MaxEnt, permitieron determinar que la presencia de este polimorfismo podría favorecer un nuevo sitio donante de empalme afectado el splicing del mRNA

al activar un sitio críptico que podría conllevar a la delección de secuencias exónicas o a la retención de secuencias intrónicas en el mRNA maduro (Vreeswijk et al, 2009) (Apéndice I).

Estos resultados deben interpretarse con precaución debido al escaso número de estudios publicados que evalúen esta alteración. Un estudio previo considera que esta variante, podría potencialmente producir empalmes alternativos (Kim et al, 2007), pero otra investigación, por el contrario, manifiesta que esta variante no podría no interferir en el empalme dada su distancia de las uniones intrón/exón donde se detectan los sitios de empalme (Ng et al, 2010). Aun no es muy claro el efecto de esta alteración sobre la regulación del gen *AHR*, por lo que teniendo en cuenta la aproximación *in silico* de la herramienta Human Splicing Finder (HSF) que indica la activación de sitios de empalme crípticos, a causa de la variante rs2074113G>T, esta alteración se clasifica en el presente estudio como probablemente benigna y se sugiere el desarrollo de futuros estudios *in vivo* que permitan explorar el efecto de esta variante sobre la célula.

Con relación a las variantes no reportadas se logró determinar el efecto de estas alteraciones de secuencia sobre la estructura y función de la proteína AhR mediante análisis *in silico*, y de esta manera inferir su importancia clínica. La variante [c.66-165delT], localizada en el intrón 2 del gen, se consideró de tipo benigno y sin modificación o anulación de posibles sitios de empalme de acuerdo con las herramientas de predicción *in silico* evaluadas. Por lo anterior se podría inferir que esta variante no representa un efecto dañino sobre la función de la proteína, por lo cual se clasifica en el presente estudio como variante de tipo benigna. Sin embargo, esta aproximación *in silico*, sólo podría ser confirmada mediante evaluación con estudios *in vivo*.

Las variantes [c.253+175T>C] y [c.2333C>T], localizadas en el intrón 1 y la región 3'UTR, respectivamente, no representan un cambio estructural o funcional sobre el receptor de

acuerdo con los análisis *in silico*, sin embargo, mediante la herramienta HSF, se logró determinar que dichas alteraciones sí podrían favorecer la alteración significativa de motivos de secuencias ESE / ESS, promoviendo un efecto sobre la regulación de la transcripción. Por lo anterior, las variantes [c.253+175T>C] y [c.2333C>T] se clasifican en este estudio como variantes de significado incierto, sugiriendo la realización de futuros estudios *in vivo* para determinar el efecto de las mismas sobre la célula.

Finalmente, futuros estudios de asociación genética sobre las variantes identificadas y el riesgo de cáncer deben realizarse permitiendo validar los hallazgos obtenidos en esta investigación. Diferentes estudios publicados, como evidencia acumulada, demuestran que la actividad de AhR podría influir durante la progresión del cáncer (Kolluri, Jin, Safe, 2017). En general el *AHR* podría explorarse como un marcador de susceptibilidad para cáncer en función de su actividad biológica funcional, siendo escasos los trabajos informados sobre este gen.

Los hallazgos obtenidos en este estudio representan un primer acercamiento en la identificación de SNPs en el gen *AHR* en la población colombiana y contribuye a la comprensión de la susceptibilidad genética al desarrollo de cáncer de cabeza y cuello en Colombia.

## 5. Conclusiones

- Se identificaron 19 variantes en el gen *AHR*, de las cuáles, siete correspondieron a alteraciones de alto impacto, reportadas para cáncer de Cabeza y Cuello en diferentes repositorios

de cáncer en el mundo. Sin embargo, ninguna de estas variantes fue reportada para población latina, por lo que la baja representación de latinoamericanos en estas bases de datos genómicos indica la necesidad de detectar alteraciones en *AHR* para esta población.

- Se reporta para Colombia, alteraciones de secuencia en el gen *AHR*. Se identificaron siete variantes en el gen *AHR* en pacientes colombianos con carcinoma escamocelular de cabeza y cuello. Las variantes *AHR* c.253+175T>C; *AHR* c.66-165delT; *AHR* c.2333C>T, no han sido reportadas previamente en bases de datos o literatura científica en el mundo.

- De acuerdo con los análisis de predicción *in silico* desarrollados y la literatura reportada, las Variantes identificadas en este estudio: rs7796976, rs17779352 se clasificaron Benignas, ya que no representan un efecto sobre la función del receptor AhR.

- Futuros estudios *in vivo*, o de asociación que involucren las variantes rs2074113 y rs2066853 deben realizarse permitiendo validar los hallazgos obtenidos en esta investigación. Diferentes estudios publicados, como evidencia acumulada, demuestran que la actividad de AhR podría influir durante la progresión del cáncer.

- Los hallazgos obtenidos en este estudio representan un primer acercamiento en la identificación de SNPs en el gen *AHR* en población colombiana y contribuye a la comprensión del desarrollo de cáncer de cabeza y cuello. Dada la importancia del gen *AHR* podría explorarse como un marcador de susceptibilidad para cáncer en función de su actividad biológica funcional.

## Referencias

### Referencias Bibliográficas

- Aarts JMMJG, Alink GM, Franssen HJ, Roebroeks W. (2021). Evolution of Hominin Detoxification: Neanderthal and Modern Human Ah Receptor Respond Similarly to TCDD. *Mol Biol Evol.* 38(4):1292-1305. doi: 10.1093/molbev/msaa287.
- Adachi, J., et al. (2001). Indirubin and indigo are potent aryl hydrocarbon receptor ligands present in human urine. *The Journal of Biological Chemistry.* 276: 31475-31478.
- Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., ... & Sunyaev, S. R. (2010). A method and server for predicting damaging missense mutations. *Nature methods*, 7(4), 248-249.
- Adzhubei, I., Jordan, D. M., & Sunyaev, S. R. (2013). Predicting functional effect of human missense mutations using PolyPhen-2. *Current protocols in human genetics*, 76(1), 7-20.
- Aftabi, Y., Colagar, A. H., & Mehrnejad, F. (2016). An in silico approach to investigate the source of the controversial interpretations about the phenotypic results of the human AhR-gene G1661A polymorphism. *Journal of Theoretical Biology*, 393, 1-15.

- Ames J, Warner M, Mocarelli P, et al. (2018). AHR gene-dioxin interactions and birthweight in the Seveso Second Generation Health Study. *International Journal of Epidemiology*. 47(6):1992-2004. DOI: 10.1093/ije/dyy165. PMID: 30124847; PMCID: PMC6280946.
- Bekki K, Vogel H, Li W, Ito T, Sweeney C, Haarmann-Stemmann T, Matsumura F, Vogel CF. (2015). The aryl hydrocarbon receptor (AhR) mediates resistance to apoptosis induced in breast cancer cells. *Pestic Biochem Physiol*. 120:5–13.
- Bohonowych JE, Denison MS. (2007). Persistent Binding of Ligands to the Aryl Hydrocarbon Receptor. *Toxicological Sciences*. 98(1):99–109. DOI: 10.1093/toxsci/kfm085 [PubMed: 17431010]
- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 68(6):394-424. doi: 10.3322/caac.21492. Epub 2018 Sep 12.
- Brooks, P. J. & Theruvathu, J. A. (2005). DNA adducts from acetaldehyde: implications for alcohol-related carcinogenesis. *Alcohol* 35, 187–193.
- Brunotto, M., Zarate, A. M., Bono, A., Barra, J. L., & Berra, S. (2014). Risk genes in head and neck cancer: a systematic review and meta-analysis of last 5 years. *Oral oncology*, 50(3), 178-188.
- Budhwar S, Bahl C, Sharma S, Singh N, Behera D. (2018). Role of Sequence Variations in AhR Gene Towards Modulating Smoking Induced Lung Cancer Susceptibility in North Indian Population: A Multiple Interaction Analysis. *Curr Genomics*. 19:313–326. doi: 10.2174/1389202918666170915160606.
- Cadoni, G. et al. (2012). A review of genetic epidemiology of head and neck cancer related to polymorphisms in metabolic genes, cell cycle control and alcohol metabolism. *Acta Otorhinolaryngol. Ital*. 32, 1–11.

Cajas, N.; Hoyos, L. S.; Sierra, H. (2005). Efecto De Los Polimorfismos Gstt1, Gstm1 Y Cyp2e1 En La Susceptibilidad Genética Para El Desarrollo De Cáncer De Pulmón En El Departamento Del Cauca. *Rev. Fac. Cienc. Salud Univ. Cauca*, 7, 9-14.

Cárcamo M. (2018). Epidemiología y generalidades del tumor de cabeza y cuello. [REV. MED. CLIN. CONDES. 29(4) 388-396]

Castaño-Molina, Eduardo, Santacoloma, Mario, Arango, Lázaro, & Camargo, Mauricio. (2009). CYP1A1, CYP2E1 y riesgo a cáncer gástrico en una población colombiana de alta incidencia. *Acta Biológica Colombiana*, 14(3), 205-211.

Castaño-Molina, Eduardo, Santacoloma, Mario, Arango, Lázaro, & Camargo, Mauricio. (2010). Cáncer gástrico y genes detoxificadores en una población colombiana. *Revista colombiana de Gastroenterología*, 25(3), 252-260.

Cauchi S, Stucker I, Solas C, Laurent-Puig P, Cenee S, Hemon D, et al. (2001). Polymorphisms of human aryl hydrocarbon receptor (AhR) gene in a French population: relationship with CYP1A1 inducibility and lung cancer. *Carcinogenesis*. 22:1819–1824. doi: 10.1093/carcin/22.11.1819.

Celius T, Matthews J. (2010). Functional analysis of six human aryl hydrocarbon receptor variants in human breast cancer and mouse hepatoma cell lines. *Toxicology*. 277:59–65. doi: 10.1016/j.tox.2010.08.015.

Chen D, Tian T, Wang H, Liu H, Hu Z, Wang Y, Liu Y, Ma H, Fan W, Miao R, Sun W, Wang Y, Qian J, Jin L, Wei Q, Shen H, Huang W, Lu D. (2009). Association of human aryl hydrocarbon receptor gene polymorphisms with risk of lung cancer among cigarette smokers in a Chinese population. *Pharmacogenet Genomics*. 19(1):25-34. doi: 10.1097/FPC.0b013e328316d8d8. PMID: 18818557.

- Chen DS, Mellman I. (2017). Elements of cancer immunity and the cancer-immune set point. *Nature*. 541:321–330. doi: 10.1038/nature21349.
- Choi, Y., Chan, A.P. (2015). PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels, *Bioinformatics*, Volume 31, Issue 16, Pages 2745–2747, <https://doi.org/10.1093/bioinformatics/btv195>
- Chuang CY, Chang H, Lin P, Sun SJ, Chen PH, Lin YY, Sheu GT, Ko JL, Hsu SL, Chang JT. (2012). Up-regulation of osteopontin expression by Aryl hydrocarbon receptor via both ligand-dependent and ligand-independent pathways in lung cancer. *Gene*. 492:262–269.
- Delahaye-Sourdeix, M., Anantharaman, D., Timofeeva, M. N., Gaborieau, V., Chabrier, A., Vallée, M. P., ... & McKay, J. D. (2015). A rare truncating BRCA2 variant and genetic susceptibility to upper aerodigestive tract cancer. *Journal of the National Cancer Institute*, 107(5), djv037.
- Denison, M. & Nagy, S. (2003). Activation of the aryl hydrocarbon receptor by structurally diverse exogenous and endogenous chemicals. *Annual Review of Pharmacology and Toxicology*. 43: 309-334.
- Desmet, F. O., Hamroun, D., Lalande, M., Collod-Bérout, G., Claustres, M., & Bérout, C. (2009). Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic acids research*, 37(9), e67. <https://doi.org/10.1093/nar/gkp215>
- Diry M., Tomkiewicz C., Koehle C., Coumoul X., Bock K.W., Barouki R., Transy C. (2006). Activation of the dioxin/aryl hydrocarbon receptor (AhR) modulates cell plasticity through a JNK-dependent mechanism. *Oncogene*. 25:5570–5574. [PubMed] [Google Scholar]
- Dyson, N., Howley, P. M., Munger, K. & Harlow, E. (1989). The human papilloma virus-16 E7 oncoprotein is able to bind to the retinoblastoma gene product. *Science* 243, 934–937.

- Ema M, Ohe N, Suzuki M, Mimura J, Sogawa K, Ikawa S, Fujii-Kuriyama Y. (1994). Dioxin binding activities of polymorphic forms of mouse and human arylhydrocarbon receptors. *J Biol Chem.* 4;269(44):27337-43. PMID: 7961644.
- Estevao, D., Costa, N. R., Gil da Costa, R. M. & Medeiros, R. (2019). Hallmarks of HPV carcinogenesis: the role of E6, E7 and E5 oncoproteins in celular malignancy. *Biochim. Biophys. Acta Gene Regul. Mech.* 1862, 153–162.
- Fang, M., Huang, W., Mo, D., Zhao, W. & Huang, R-. (2018). Association of five Snps in cytotoxic T-lymphocyte antigen 4 and cancer susceptibility: evidence from 67 studies. *Cell Physiol. Biochem.* 47, 414–427.
- Felsenstein, J., Churchill, G.A. (1996). A Hidden Markov Model approach to variation among sites in rate of evolution., *Molecular Biology and Evolution*, Volume 13, Issue 1, Pages 93–104, <https://doi.org/10.1093/oxfordjournals.molbev.a025575>
- Feng SL, Cao ZH, Wang XM. (2013). Role of aryl hydrocarbon receptor in cancer. *BBA-Rev Cancer.* 1836:197–210.
- Ferlay, J. et al. (2019). Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. *Int. J. Cancer* 144, 1941–1953.
- Ferlay, J. et al. (2020). *Global Cancer Observatory: Cancer Today*. Lyon, France: International Agency for Research on Cancer (accessed 18 September 2020). IARC <https://gco.iarc.fr/today>.
- Fernández-Mateos, J., Seijas-Tamayo, R., Adansa Klain, J. C., Pastor Borgoñón, M., Pérez-Ruiz, E., Mesía, R., ... & Cruz-Hernández, J. J. (2019). Genetic susceptibility in head and neck squamous cell carcinoma in a Spanish population. *Cancers*, 11(4), 493.
- Figuroa JD, Malats N, García-Closas M, Real FX, Silverman D, Kogevinas M, Chanock S, Welch R, Dosemeci M, Lan Q, Tardón A, Serra C, Carrato A, García-Closas R, Castaño-Vinyals G,

- Rothman N. (2008). Bladder cancer risk and genetic variation in AKR1C3 and other metabolizing genes. *Carcinogenesis*. Oct;29(10):1955-62. doi: 10.1093/carcin/bgn163. Epub 2008 Jul 16. PMID: 18632753; PMCID: PMC2556968.
- Freedman, N. D. et al. (2008). Fruit and vegetable intake and head and neck cancer risk in a large United States prospective cohort study. *Int. J. Cancer* 122, 2330–2336.
- Fukushima-Uesaka, Hiromi Sai, Kimie Maekawa, Keiko Koyano, Satoru Kaniwa, Nahoko Ozawa, Shogo Kawamoto, Manabu Kamatani, Naoyuki Komamura, Kazuo Kamakura, et al. (2004). Genetic Variations of the AHR Gene Encoding Aryl Hydrocarbon Receptor in a Japanese Population. *Drug metabolism and pharmacokinetics*. 19. 320-6. 10.2133/dmpk.19.320.
- Gagliani N, Vesely MCA, Iseppon A, Brockmann L, Xu H, Palm NW, de Zoete MR, Licona-Limon P, Paiva RS, Ching T, et al. (2015). TH17 cells transdifferentiate into regulatory T cells during resolution of inflammation. *Nature*. 523:221–225.
- Galle-Treger L, Suzuki Y, Patel N, Sankaranarayanan I, Aron JL, Maazi H, Chen L, Akbari O. (2016). Nicotinic acetylcholine receptor agonist attenuates ILC2-dependent airway hyperreactivity. *Nat Commun*. 7: 13202.
- Gallegos-hernández JF. (2006). El cáncer de cabeza y cuello. Factores de riesgo y prevención. *Cir Ciruj*. 74(4):287–93.
- Gambier N, Marteau JB, Batt AM, et al. (2006). Interaction between CYP1A1 T3801C and AHR G1661A polymorphisms according to smoking status on blood pressure in the Stanislas cohort. *J Hypertens*. 24: 2199–2205.
- Garte S, et al. (2001). Metabolic gene polymorphism frequencies in control populations. *Cancer Epidemiol Biomarkers Prev*. 10: 1239-48.

GeneCards – the human gene database [www.genecards.org](http://www.genecards.org). Safran M, Rosen N, Twik M, BarShir R, Iny Stein T, Dahary D, Fishilevich S, and Lancet D. (2022). The GeneCards Suite Chapter, Practical Guide to Life Science Databases pp 27-56 [PDF]

Gomez A, Bindesbøll C, Satheesh SV, Grimaldi G, Hutin D, MacPherson L, Ahmed S, Tamblyn L, Cho T, Nebb HI, Moen A, Anonsen JH, Grant DM, Matthews J. (2018). Characterization of TCDD-inducible poly-ADP-ribose polymerase (TIPARP/ARTD14) catalytic activity. *Biochem J.* 11;475(23):3827-3846. doi: 10.1042/BCJ20180347. PMID: 30373764; PMCID: PMC6292455.

González-Moles, M. A., Gil-Montoya, J. A., & Ruiz-Ávila, I. (2008). Bases moleculares de la cancerización de cavidad oral. *Avances en odontoestomatología*, 24(1), 55-60.

Gu A, Ji G, Jiang T, Lu A, You Y, Liu N, et al. (2012). Contributions of aryl hydrocarbon receptor genetic variants to the risk of glioma and PAH-DNA adducts. *Toxicol Sci.* 128:357–364. doi: 10.1093/toxsci/kfs158.

Guarnieri T, Abruzzo PM, Bolotta A. (2020). More than a cell biosensor: aryl hydrocarbon receptor at the intersection of physiology and inflammation. *Am J Physiol Cell Physiol.* 318:C1078–C1c82. doi: 10.1152/ajpcell.00493.2019.

Guha, N. et al. (2007). Oral health and risk of squamous cell carcinoma of the head and neck and esophagus: results of two multicentric case-control studies. *Am. J. Epidemiol.* 166, 1159–1173.

Guo J, Zhang J, Yin L, Pu Y. (2005). A Case-Control Study on the Relation of Aryl Hydrocarbon Receptor(Ahr)Polymorphisms with Susceptibility to Adult Acute Leukemia in Chinese HAN Population. *J Environ Occup Med.* 22:388–391.

Hahn, M.E. (2002). Aryl hydrocarbon receptors: diversity and evolution, *Chem. Biol. Interact.* 141 131e160.

- Hahn, M.E., Allan, L.L., Sherr, D.H. (2009). Regulation of constitutive and inducible AHR signaling: complex interactions involving the AHR repressor, *Biochem. Pharmacol.* 77:485-497, <https://doi.org/10.1016/j.bcp.2008.09.016>.
- Hashibe M, Brennan P, Benhamou S, et al. (2007). Alcohol drinking in never users of tobacco, cigarette smoking in never drinkers, and the risk of head and neck cancer: pooled analysis in the International Head and Neck Cancer Epidemiology Consortium. *J Natl Cancer Inst.* 16;99(10):777-89. doi: 10.1093/jnci/djk179. Erratum in: *J Natl Cancer Inst.* 2008 Feb 6;100(3):225. Fernandez, Leticia [added]. PMID: 17505073.
- Hao, N., Whitelaw, M., Shearwin, K., Dodd, I. & Chapman-Smith, A. (2011). Identification of residues in the N-terminal PAS domains important for dimerization of ARNT and AhR. *Nucleic Acids Research.* 39: 3695-3709.
- Harris C.C. (1989). Interindividual variation among humans in carcinogen metabolism, DNA adduct formation and DNA repair. *Carcinogenesis.* 10: 1563-66
- Heberle, H.; Meirelles, G. V.; da Silva, F. R.; Telles, G. P.; Minghim, R. (2015). InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. *BMC Bioinformatics;* 16:169.
- Hecht, SS. (1999). Tobacco smoke carcinogens and lung cancer. *J. Natl Cancer Inst.* 91, 1194–1210.
- Hecht SS. (2003). Tobacco carcinogens, their biomarkers and tobacco-induced cancer. *Nat Rev Cancer.* 3(10):733–44. DOI: 10.1038/nrc1190 [PubMed: 14570033]
- Hoffman, E., et al. (1991). Cloning of a factor required for activity of the Ah (dioxin) receptor. *Science.* 252: 954-958.
- Hoffmann, D. & Hoffmann, I. (1997). The changing cigarette, 1950-1995. *J. Toxicol. Env. Health* 50, 307–364.

- Hogenesch, J., et al. (1997). Characterization of a subset of the basic-helixloop- helix-PAS superfamily that interacts with components of the dioxin signaling pathway. *The Journal of Biological Chemistry*. 272: 8581-8593.
- Huang S, Shui X, He Y, et al. (2015). AhR expression and polymorphisms are associated with risk of coronary arterial disease in Chinese population. *Sci Rep*.5:8022.
- Hunt, S. E., Moore, B., Amode, R. M., Armean, I. M., Lemos, D., Mushtaq, A., ... & Cunningham, F. (2021). Annotating and prioritizing genomic variants using the Ensembl Variant Effect Predictor—A tutorial. *Human mutation*.
- Ikuta T., Kawajiri K. (2006). Zinc finger transcription factor Slug is a novel target gene of aryl hydrocarbon receptor. *Exp. Cell Res*. 312:3585–3594. [PubMed] [Google Scholar]
- International Agency for Research on Cancer. (2020). List of Classifications by cancer sites with sufficient or limited evidence in humans, Volumes 1 to 127. IARC Monographs On The Identification Of Carcinogenic Hazards To Humans. IARC <https://monographs.iarc.fr/agents-classified-by-the-iarc/>.
- Instituto Nacional de Cancerología (2018). Cifras y Estimaciones de Cancer en el Mundo: GLOBOCAN e IARC. Octubre de 2018, tomado de: [https://www.cancer.gov.co/sites/default/files/boletin-prensa/archivo/boletin\\_globocan.pdf](https://www.cancer.gov.co/sites/default/files/boletin-prensa/archivo/boletin_globocan.pdf); <https://gco.iarc.fr/today/data/factsheets/populations/170-colombia-fact-sheets.pdf>
- Instituto Nacional de Cancerología. (2019). Hoja informativa: Que entre colores y sabores el tabaco no te enamore. Bogotá, D.C. Tomado de: [www.cancer.gov.co](http://www.cancer.gov.co)
- Isayeva, T., Li, Y., Maswahu, D. & Brandwein-Gensler, M. (2012). Human papillomavirus in non- oropharyngeal head and neck cancers: a systematic literature review. *Head Neck Pathol*. 6, S104–S120.

- Iu M, Zago M, Rico de Souza A, Bouttier M, Pareek S, White JH, Hamid Q, Eidelman DH, Bagloli CJ. (2017). RelB attenuates cigarette smoke extract-induced apoptosis in association with transcriptional Regulation of the aryl hydrocarbon receptor. *Free Radic Biol Med.* 108:19–31.
- Johnson, D. E., Burtness, B., Leemans, C. R., Lui, V. W. Y., Bauman, J. E., & Grandis, J. R. (2020). Head and neck squamous cell carcinoma. *Nature reviews Disease primers*, 6(1), 1-22.
- Kasai A, Hiramatsu N, Hayakawa K, Yao J, Maeda S, Kitamura M. (2006). High levels of dioxin-like potential in cigarette smoke evidenced by in vitro and in vivo biosensing. *Cancer Res.* 66(14):7143–50. DOI: 10.1158/0008-5472.CAN-05-4541 [PubMed: 16849560]
- Kawajiri K, Fujii-Kuriyama Y. (2017). The aryl hydrocarbon receptor: a multifunctional chemical sensor for host defense and homeostatic maintenance. *Exp Anim.* 66:75–89. doi: 10.1538/expanim.16-0092.
- Kent WJ. (2002). BLAT - the BLAST-like alignment tool. *Genome Res.* 12(4):656-64.
- Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. (2002). The human genome browser at UCSC. *Genome Res.* 12(6):996-1006.
- Kewley, R., Whitelaw, M. & Chapman-Smith, A. (2004). The mammalian basic helixloop-helix/PAS family of transcriptional regulators. *The International Journal of Biochemistry and Cell Biology.* 36: 189-204.
- Kim JH, Kim H, Lee KY, Kang JW, Lee KH, Park SY, Yoon HI, Jheon SH, Sung SW, Hong YC. (2007). Aryl hydrocarbon receptor gene polymorphisms affect lung cancer risk. *Lung Cancer.* 56(1):9-15. doi: 10.1016/j.lungcan.2006.11.010. Epub Dec 15. PMID: 17174437.
- Klein, J. D., & Grandis, J. R. (2010). The molecular pathogenesis of head and neck cancer. *Cancer biology & therapy*, 9(1), 1-7.

- Kolluri SK, Jin UH, Safe S. (2017). Role of the aryl hydrocarbon receptor in carcinogenesis and potential as an anti-cancer drug target. *Arch Toxicol.* 91:2497–2513. doi: 10.1007/s00204-017-1981-2.
- Kumar, M. B., Ramadoss, P., Reen, R. K., Vanden Heuvel, J. P. y Perdew, G. H. (2001). The Q-rich Subdomain of the Human Ah Receptor Transactivation Domain Is Required for Dioxin-mediated Transcriptional Activity. *J Biol Chem* 276(45): 42302-42310.
- Kumar, P., Henikoff, S., & Ng, P. C. (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nature protocols*, 4(7), 1073-1081.
- Lanciego Diaz. (2013). Estudio Genotípico Del Citocromo P450 En Una Cohorte De Pacientes Con Cáncer De Pulmón En El Área Sanitaria De Mérida. Universidad De Extremadura. Tesis Doctoral. Pp 14
- Larigot, L., Juricek, L., Dairou, J., & Coumoul, X. (2018). AhR signaling pathways and regulatory functions. *Biochimie open*, 7, 1-9.
- Le Marchand L, Donlon T, Kolonel LN, Henderson BE, Wilkens LR. (2005). Estrogen metabolism-related genes and breast cancer risk: the multiethnic cohort study. *Cancer Epidemiol Biomarkers Prev.* 14:1998–2003. doi: 10.1158/1055-9965.EPI-05-0076.
- Li Y, Qin HZ, Song Q, Wu XD, Zhu JH. (2015). Lack of association between the aryl hydrocarbon receptor rs2066853 polymorphism and breast cancer: A meta-analysis on Ahr polymorphism and breast cancer. *Genet Mol Res.* 14:16162–16168. doi: 10.4238/2015.December.8.5.
- Li H, Luo L, Wang D, Duan J, Zhang R. (2020). Lack of association between multiple polymorphisms in aryl hydrocarbon receptor (AhR) gene and cancer susceptibility. *Environ Health Prev Med.* 5;25(1):79. doi: 10.1186/s12199-020-00907-z. PMID: 33278884; PMCID: PMC7718691.

- Long JR, Egan KM, Dunning L, Shu XO, Cai Q, Cai H, et al. (2006) Population-based case-control study of AhR (aryl hydrocarbon receptor) and CYP1A2 polymorphisms and breast cancer risk. *Pharmacogenet Genomics*. 16:237–243. doi: 10.1097/01.fpc.0000189803.34339.ed.
- MacPherson L, Tamblyn L, Rajendra S, Bralha F, McPherson JP, Matthews J. (2013). 2,3,7,8-Tetrachlorodibenzo-p-dioxin poly(ADP-ribose) polymerase (TiPARP, ARTD14) is a mono-ADP-ribosyltransferase and repressor of aryl hydrocarbon receptor transactivation. *Nucleic Acids Res*. 1;41(3):1604-21. doi: 10.1093/nar/gks1337. Epub 2012 Dec 28. PMID: 23275542; PMCID: PMC3562000.
- Mahul, B. A., Stephen, E., Frederick, L. G., David, R. B., Robert, K., Mary, K. W., ... & Daniel, C. S. (2017). *AJCC cancer staging manual*. Eight Edition: Springer Nature.
- Martano M, Stiuso P, Facchiano A, De Maria S, Vanacore D, Restucci B, Rubini C, Caraglia M, Ravagnan G, Lo Muzio L. (2018). Aryl hydrocarbon receptor, a tumor grade-associated marker of oral cancer, is directly downregulated by polydatin: A pilot study. *Oncol Rep*. 40(3):1435-1442. doi: 10.3892/or.2018.6555. Epub 2018 Jul 10.
- Mazul, A. L., Shivappa, N., Hébert, J. R., Steck, S. E., Rodriguez-Ormaza, N., Weissler, M., Olshan, A. F., & Zevallos, J. P. (2018). Proinflammatory diet is associated with increased risk of squamous cell head and neck cancer. *International journal of cancer*, 143(7), 1604–1610. <https://doi.org/10.1002/ijc.31555>
- McKay JD, Truong T, Gaborieau V, et al. (2011). A genome-wide association study of upper aerodigestive tract cancers conducted within the INHANCE consortium. *PLoS Genet*. 7(3).
- McLaren, W., Gil, L., Hunt, S. E., Riat, H. S., Ritchie, G. R., Thormann, A., ... & Cunningham, F. (2016). The ensembl variant effect predictor. *Genome biology*, 17(1), 1-14.

Meyer BK Pray-Grant MG, Vanden Heuvel JP and Perdew GH. (1998). Hepatitis B virus X-associated protein 2 is a subunit of the unliganded aryl hydrocarbon receptor core complex and exhibits transcriptional enhancer activity. *Mol Cell Biol* 18: 978-88.

Michaud, D. S. et al. (2014). High-risk HPV types and head and neck cancer. *Int. J. Cancer* 135, 1653–1661.

Ministerio de Salud y Protección Social. (2019). Las señales del humo del tabaco. Ministerio de Salud y Protección Social, Bogotá, D.C. Tomado de: [www.minsalud.gov.co](http://www.minsalud.gov.co)

Miller PD, et al. (2002). Combination of the variant genotype of GSTP1, GSTM1 and, p53 are associated with an increased lung cancer risk. *Cancer Res.* 62: 2819-23.

Mishra, A. & Meherotra, R. (2014). Head and neck cancer: global burden and regional trends in India. *Asian Pac. J. Cancer Prev.* 15, 537–550.

Mimura, J., & Fujii-Kuriyama, Y. (2003). Functional role of AhR in the expression of toxic effects by TCDD. *Biochimica et Biophysica Acta - General Subjects.* 1619(3): 263–268.

Montero, P. H. (2018). Cáncer de Cabeza y Cuello asociado a Virus Papiloma Humano: Prevención, diagnóstico y tratamiento. *Revista Médica Clínica Las Condes*, 29(4), 419-426.

Moorthy B, Chu C, Carlin DJ. (2015). Polycyclic aromatic hydrocarbons: from metabolism to lung cancer. *Toxicol Sci.*145:5–15.

Murray IA, Patterson AD, Perdew GH. (2014). Aryl hydrocarbon receptor ligands in cancer: friend and foe. *Nat Rev Cancer.* 14:801–814.

Nebert, J., L. F., Zhang, G., Wilson, K. M., Jiang, Z., Butler, R., Gluckman, J. L., Pinney, S. M., & Nebert, D. W. (2016). Head-and-neck squamous cell carcinoma risk in smokers: no

association detected between phenotype and AHR, CYP1A1, CYP1A2, or CYP1B1 genotype. *Human genomics*, 10(1), 39. <https://doi.org/10.1186/s40246-016-0094-y>

Negri E, Boffetta P, Berthiller J, et al. (2009). Family history of cancer: pooled analysis in the International Head and Neck Cancer Epidemiology Consortium. *Int J Cancer*. 124(2):394–401.

Ng CH, Janoo-Gilani R, Sipahimalani P, Gallagher RP, Gascoyne RD, Connors JM, Weber JP, Lai AS, Leach S, Le ND, Brooks-Wilson AR, Spinelli JJ. (2010). Interaction between organochlorines and the AHR gene, and risk of non-Hodgkin lymphoma. *Cancer Causes Control*. 21(1):11-22. doi: 10.1007/s10552-009-9429-5. Epub 2009 Sep 26. PMID: 19821039.

Niu, Y. M. et al. (2015). Increased risks between Interleukin-10 gene polymorphisms and haplotype and head and neck cancer: a meta-analysis. *Sci. Rep.* 5, 17149.

Norppa H. (2003). Genetic susceptibility, biomarkers response, and cancer. *Mutat Res*. 544:339-48.

Olivier M, Hollstein M, Hainaut P. (2010). TP53 mutations in human cancers: origins, consequences, and clinical use. *Cold Spring Harb Perspect Biol*. 2: a001008.

Organización Mundial de la Salud. (2010). OMS/SIGN: carpeta de material sobre seguridad de las inyecciones y los procedimientos conexos. Organización Mundial de la Salud, Ginebra, Suiza. ISBN 978 92 4 359925 0

Organización Mundial de la Salud. (2022). Edición electrónica de la CIE-10. 4ª Edición. Tomado de <https://eciemaps.mscbs.gob.es/ecieMaps/browser/metabuscaador.html>

Pai, S. I. & Westra, W. H. (2009). Molecular pathology of head and neck cancer: implications for diagnosis, prognosis, and treatment. *Annu. Rev. Pathol.* 4, 49–70.

Papillon-Cavanagh S, Lu C, Gayden T, Mikael LG, Bechet D, Karamboulas C, Ailles L, Karamchandani J, Marchione DM, Garcia BA, et al. (2017). Impaired H3K36 methylation defines a subset of head and neck squamous cell carcinomas. *Nat Genet.* 49(2):180–5.

Paz-Egas C, Fong C, Cifuentes, L. (2016). Frecuencia de polimorfismos en genes implicados en detoxificación de carcinógenos del tabaco en una población colombiana. *Journal of Basic And Applied Genetics.* 27. 178.

Perera FP & Weinstein IB. (2000). Molecular epidemiology: Recent advances and future directions. *Carcinogenesis*, 21:517-24.

Pérez-Morales R, Méndez-Ramírez I, Moreno-Macias H, Mendoza-Posadas AD, Martínez-Ramírez OC, Castro-Hernández C, et al. (2014). Genetic susceptibility to lung cancer based on candidate genes in a sample from the Mexican Mestizo population: a case-control study. *Lung.* 192:167–173. doi: 10.1007/s00408-013-9536-7.

Pierre S., Chevallier A., Teixeira-Clerc F., Ambolet-Camoit A., Bui L.-C., Bats A.-S., Fournet J.-C., Fernandez-Salguero P., Aggerbeck M., Lotersztajn S., Barouki R., Coumoul X. (2014). Aryl hydrocarbon receptor-dependent induction of liver fibrosis by dioxin. *Toxicol. Sci. Offic. J. Soc. Toxicol.* 137:114–124. [PubMed] [Google Scholar]

Prager, M., Büttner, J., Grunert, P., Ellinghaus, D., & Büning, C. (2016). A promoter variant within the aryl hydrocarbon receptor gene is associated with an epithelial barrier defect in smokers with Crohn's disease. *Inflammatory Bowel Diseases*, 22(10), 2356-2368.

Popolo A, Pinto A, Daglia M, Nabavi SF, Farooqi AA, Rastrelli L. (2017). Two likely targets for the anti-cancer effect of indole derivatives from cruciferous vegetables: PI3K/Akt/mTOR signalling pathway and the aryl hydrocarbon receptor. *Semin Cancer Biol.* 46:132–137.

Puga A, Xia Y, Elferink C. (2002). Role of the aryl hydrocarbon receptor in cell cycle regulation. *Chem Biol Interact.* 20;141(1-2):117-30. doi: 10.1016/s0009-2797(02)00069-8. PMID: 12213388.

Ramírez-Clavijo, S., Rangel, N., & Escobar Espinosa, S. B. (2009). Polimorfismos de los genes P53, CYP1A1, CYP1B1, GSTM1, GSTT1 y GSTP1 en cáncer de seno familiar en una población colombiana (Doctoral dissertation, Universidad del Rosario).

Raney BJ, Dreszer TR, Barber GP, Clawson H, Fujita PA, Wang T, Nguyen N, Paten B, Zweig AS, Karolchik D, Kent WJ. (2014). Track Data Hubs enable visualization of user-defined genome-wide annotations on the UCSC Genome Browser. *Bioinformatics.* 1;30(7):1003-5. Epub 2013 Nov 13.

Reva, BA, Antipin, YA, and Sander, C. (2007). Determinants of protein function revealed by combinatorial entropy optimization. *Genome Biol.*; 8(11): R232.

Reva, B., Antipin, Y., & Sander, C. (2011). Predicting the functional impact of protein mutations: application to cancer genomics. *Nucleic acids research*, 39(17), e118-e118.

Rocha Buelvas A. (2009). Cáncer oral: el papel del odontólogo en la detección temprana y control. *Rev Fac Odontol UnivAntioq.* 21(1): 112-121

Roesch-Ely M, Nees M, Karsai S, et al. (2007). Proteomic analysis reveals successive aberrations in protein expression from healthy mucosa to invasive head and neck cancer. *Oncogene.* 26:54-64.

Romagnolo DF, Daniels KD, Grunwald JT, Ramos SA, Propper CR, Selmin OI. (2016). Epigenetics of breast cancer: modifying role of environmental and bioactive food compounds. *Mol Nutr Food Res.* 60:1310–1329.

Rosenbloom KR, Sloan CA, Malladi VS, Dreszer TR, Learned K, Kirkup VM, Wong MC, Maddren M, Fang R, Heitner SG, Lee BT, Barber GP, Harte RA, Diekhans M, Long JC, Wilder SP, Zweig AS, Karolchik D, Kuhn RM, Haussler D, Kent WJ. (2013). ENCODE data in the UCSC Genome Browser: year 5 update. *Nucleic Acids Res.* 41(Database issue):D56-63.

Rothhammer V, & Quintana FJ. (2019). The aryl hydrocarbon receptor: an environmental sensor integrating immune responses in health and disease. *Nature Reviews Immunology* volume 19, pages 184–197

Rubin Grandis, J. et al. (1998). Levels of TGF- $\alpha$  and EGFR protein in head and neck squamous cell carcinoma and patient survival. *J. Natl Cancer Inst.* 90, 824–832.

Sadik A, Somarribas Patterson LF, Öztürk S, et al. (2020). IL4I1 Is a Metabolic Immune Checkpoint that Activates the AHR and Promotes Tumor Progression. *Cell.* 3;182(5):1252-1270.e34. doi: 10.1016/j.cell.2020.07.038. Epub 2020 Aug 19. PMID: 32818467.

Sangrajrang S, Sato Y, Sakamoto H, Ohnami S, Laird NM, Khuhaprema T, et al. (2009). Genetic polymorphisms of estrogen metabolizing enzyme and breast cancer risk in Thai women. *Int J Cancer.* 125:837–843. doi: 10.1002/ijc.24434.

Secretan B, Straif K, Baan R, et al. (2009). A review of human carcinogens – Part E: tobacco, areca nut, alcohol, coal smoke, and salted fish. *Lancet Oncol.* 10(11):1033-1034.

Sellers T.A, Potter J.D, Baily-Wilson J.E, Rich S.S, Rotschild H, and Elston R.C. (1992). Lung cancer detection and prevention evidence for an interaction between smoking and genetic predisposition. *Cancer Res. (Suppl.)* 52: 2694s-97s

Simonsson M, Veerla S, Markkula A, Rose C, Ingvar C, Jernström H. (2016). CYP1A2--a novel genetic marker for early aromatase inhibitor response in the treatment of breast cancer patients. *BMC Cancer.* 16:256. doi: 10.1186/s12885-016-2284-3. PMID: 27029552; PMCID: PMC4815192.

- Shihab, H. A., Gough, J., Cooper, D. N., Stenson, P. D., Barker, G. L., Edwards, K. J., ... & Gaunt, T. R. (2013). Predicting the functional, molecular, and phenotypic consequences of amino acid substitutions using hidden Markov models. *Human mutation*, 34(1), 57-65.
- Shopland D. (1995). Tobacco Use and Its Contribution to Early Cancer Mortality with a Special Emphasis on Cigarette Smoking. *Environmental Health Perspectives* 103(8):131–42. [PubMed: 8741773]
- Schulte KW, Green E, Wilz A, Platten M, Daumke O. (2017). Structural Basis for Aryl Hydrocarbon Receptor-Mediated Gene Activation. *Structure*. 25(7):1025-1033.e3. doi: 10.1016/j.str.2017.05.008. Epub 2017 Jun 9. PMID: 28602820.
- Sierra-Martinez M, Hernandez-Cadena L, Garcia-Sanchez JR, Acosta-Altamirano G, Palacios-Reyes C, Alonso-Themann PG, et al. (2018). Predictive polymorphisms for breast cancer in postmenopausal Mexican women. *J Cancer Res Ther*. 14:640–646. doi: 10.4103/0973-1482.179109.
- Smart J, Daly AK. (2000). Variation in induced CYP1A1 levels: relationship to CYP1A1, Ah receptor and GSTM1 polymorphisms. *Pharmacogenetics*. 10:11–24. doi: 10.1097/00008571-200002000-00003.
- Spiegelman, B. M., & Heinrich, R. (2004). Biological control through regulated transcriptional coactivators. *Cell*, 119(2), 157-167.
- Squarize, C. H. et al. (2013). PTEN deficiency contributes to the development and progression of head and neck cancer. *Neoplasia* 15, 461–471.
- Stanford E, Ramirez-Cardenas A, Wang Z, Novikov O, Alamoud K, Koutrakis P, Mizgerd JP, Genco C, Kukuruzinska M, Monti S, et al. (2016). Role for the aryl hydrocarbon receptor and

diverse ligands in oral squamous cell carcinoma migration and tumorigenesis *Mol. Cancer Res.*, 14, pp. 696-706

Steinhaus, R., Proft, S., Schuelke, M., Cooper, D.N., Schwarz, J.M., Seelow, D. (2021). MutationTaster2021, *Nucleic Acids Research*, Volume 49, Issue W1, 2 July, Pages W446–W451, <https://doi.org/10.1093/nar/gkab266>

Stelow E & Mills S. (2005). Squamous Cell Carcinoma Variants of the Upper Aerodigestive Tract. *Am J Clin Pathol.* 124(Suppl 1): S96-S109

Stein, A. P. et al. (2015). Prevalence of human papillomavirus in oropharyngeal cancer: a systematic review. *Cancer J.* 21, 138–146.

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, 71(3), 209-249.

Talamini, R. et al. (2002). Combined effect of tobacco and alcohol on laryngeal cancer risk: a case-control study. *Cancer Causes Control.* 13, 957–964.

Tian Y, Rabson AB, Gallo MA. (2002). Ah receptor and NF-kappaB interactions: mechanisms and physiological implications. *Chem Biol Interact.* 141:97–115.

Tian Y. (2009). Ah receptor and NF-kappaB interplay on the stage of epigenome. *Biochem Pharmacol.* 77:670–680.

Tomkiewicz C, Herry L, Bui LC, Metayer C, Bourdeloux M, Barouki R, Coumoul X. (2013). The aryl hydrocarbon receptor regulates focal adhesion sites through a non-genomic FAK/Src pathway. *Oncogene.* 32: 1811–1820.

- Torres, M. M., Acosta, C. P., Sicard, D. M., & Groot de Restrepo, H. (2004). Susceptibilidad genética y riesgo de cáncer gástrico en una población del Cauca. *Biomédica*, 24(2), 153-162.
- Tummala KS, Gomes AL, Yilmaz M, Grana O, Bakiri L, Ruppen I, Ximenez- Embun P, Sheshappanavar V, Rodriguez-Justo M, Pisano DG, et al. (2014). Inhibition of de novo NAD(p) synthesis by oncogenic URI causes liver tumorigenesis through DNA damage. *Cancer Cell*. 26: 826–839.
- Ueng TH, Chang YL, Tsai YY, Su JL, Chan PK, Shih JY, Lee YC, Ma YC, Kuo ML. (2010). Potential roles of fibroblast growth factor-9 in the benzo(a)-pyrene-induced invasion in vitro and the metastasis of human lungadenocarcinoma. *Arch Toxicol*. 84:651–660.
- United States Department of Health and Human Services. (2014). A Report of the Surgeon General. US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health. Reducing the health consequences of smoking: 50 years of progress.
- Vaser, R., Adusumalli, S., Leng, S. N., Sikic, M., & Ng, P. C. (2016). SIFT missense predictions for genomes. *Nature protocols*, 11(1), 1-9.
- Vázquez-Gómez, G., Rubio-Lightbourn, J., & Espinosa-Aguirre, J. J. (2016). Mecanismos de acción del receptor de hidrocarburos de arilos en el metabolismo del benzo [a] pireno y el desarrollo de tumores. *TIP. Revista especializada en ciencias químico-biológicas*, 19(1), 54-67.
- Venuti, A. et al. (2011). Papillomavirus E5: the smallest oncoprotein with many functions. *Mol. Cancer* 10, 140.
- Vreeswijk, M. P., Kraan, J. N., van der Klift, H. M., Vink, G. R., Cornelisse, C. J., Wijnen, J. T., ... & Devilee, P. (2009). Intronic variants in BRCA1 and BRCA2 that affect RNA splicing can be reliably selected by splice-site prediction programs. *Human mutation*, 30(1), 107-114.

- Wang C-K, Hang H, Chen P-H, Chang JT, Kuo Y-C, Ko J-L, Lin P. (2009). Aryl hydrocarbon receptor activation and overexpression upregulated fibroblast growth factor-9 in human lung adenocarcinomas. *Int J Cancer*. 125:807–815.
- Wang, Y., Yang, H., Duan, G. & Wang, H. (2016). The association of the CYP1A1 Ile462Val polymorphism with head and neck cancer risk: evidence based on a cumulative meta-analysis. *Onco Targets Ther*. 9, 2927–2934.
- Ward LD, Kellis M. (2012). HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res*. 40(Database issue):D930-4. doi: 10.1093/nar/gkr917. Epub 2011 Nov 7. PMID: 22064851; PMCID: PMC3245002.
- Warnakulasuriya, S. & Straif, K. (2018). Carcinogenicity of smokeless tobacco: evidence from studies in humans & experimental animals. *Indian. J. Med. Res*. 148, 681–686.
- Wei, Y., Hellenberg, H., Rannug, U. & Rannug, A. (1998). Rapid and transient induction of CYP1A1 gene expression in human cells by the tryptophan photoproduct 6-formylindolo [3, 2] carbazole. *Chemico-Biological Interactions*. 110: 39-55.
- Wei T, Liu L, Zhou X. (2018). Cortex dictamni extracts inhibit over-proliferation and migration of rat airway smooth muscle cells via FAK/p38/ Bcl-2 signaling pathway. *Biomed Pharmacother*. 102:1–8.
- Wong, I. C., Ng, Y. K. & Lui, V. W. (2014). Cancers of the lung, head and neck on the rise: perspectives on the genotoxicity of air pollution. *Chin. J. Cancer* 33, 476–480.
- Wu, D., Potluri, N., Kim, Y. & Rastinejad, F. (2013). Structure and dimerization properties of the aryl hydrocarbon receptor PAS-A domain. *Molecular Cell Biology*. 33: 4346-4356.

Yang, T., Feng, Y. L., Chen, L., Vaziri, N. D., & Zhao, Y. Y. (2019). Dietary natural flavonoids treating cancer by targeting aryl hydrocarbon receptor. *Critical reviews in toxicology*, 49(5), 445-460.

Zhang DS, Lin GF, Ma QW, Shen JH. (2002). Nonassociation of aryl hydrocarbon receptor genotypes with susceptibility to bladder cancer in Shanghai population. *Acta Pharmacol Sin.* 23(2):188-92. PMID: 11866883.

Zhang L, Wu R, Dingle RW, Gairola CG, Valentino J, Swanson HI. (2007). Cigarette smoke condensate and dioxin suppress culture shock induced senescence in normal human oral keratinocytes. *Oral oncology.* 43(7):693–700. DOI: 10.1016/j.oraloncology.2006.08.008 [PubMed: 17070097]

Zhang, L. W., Li, J., Cong, X., Hu, X. S., Li, D., Wu, L. L., ... & Kerr, A. R. (2018). Incidence and mortality trends in oral and oropharyngeal cancers in China, 2005–2013. *Cancer epidemiology*, 57, 120-126.

Zhou S, Liu J and Chowbay B. (2009). Polymorphism of human cytochrome P450 enzymes and its clinical impact. *Drug Metab Rev* 41: 89-295.

Zhu, X. et al. (2013). Prognostic role of epidermal growth factor receptor in head and neck cancer: a meta-analysis. *J. Surg. Oncol.* 108, 387–397.

### **Referencias Legales**

Constitución Política de Colombia [Const]. (1991). Art. 15. Julio 7 de 1991 (Colombia).

Resolución 8430 de 1993 [Ministerio de Salud]. Por la cual se establecen las normas científicas, técnicas y administrativas para la investigación en salud. Octubre 4 de 1993.

Ley 1266 de 2008 [Congreso de la República]. Por la cual se dictan las disposiciones generales del hábeas data y se regula el manejo de la información contenida en bases de datos personales, en especial la financiera, crediticia, comercial, de servicios y la proveniente de terceros países y se dictan otras disposiciones. Diciembre 31 de 2012. Diario Oficial N° 47.219.

Resolución 2378 de 2008 [Ministerio de la Protección Social]. Por la cual se adoptan las Buenas Prácticas Clínicas para las instituciones que conducen investigación con medicamentos en seres humanos. Junio 27 de 2008.

Decreto Reglamentario 1727 de 2009 [Ministerio de Hacienda y Crédito Público]. por el cual se determina la forma en la cual los operadores de los bancos de datos de información financiera, crediticia, comercial, de servicios y la proveniente de terceros países, deben presentar la información de los titulares de la información. Mayo 15 de 2009.

Decreto Reglamentario 2952 de 2010 [Presidencia de la República de Colombia]. por el cual se reglamentan los artículos 12 y 13 de la Ley 1266 de 2008. Agosto 6 de 2010.

Ley 1581 de 2012 [Congreso de la República]. Por la cual se dictan disposiciones generales para la protección de datos personales. Octubre 17 de 2012. Diario Oficial N°48.587.

Decreto 1377 de 2013 [Presidencia de la República de Colombia]. Por el cual se reglamenta parcialmente la Ley 1581 de 2012, Derogado Parcialmente por el Decreto 1081 de 2015. Junio 27 de 2013.

Decreto 351 de 2014 [Ministerio de Salud y Protección Social]. Por el cual se reglamenta la gestión integral de los residuos generados en la atención en salud y otras actividades. Febrero 19 de 2014.

Decreto 1081 de 2015 [Presidencia de la República de Colombia]. Por medio del cual se expide el Decreto Reglamentario Único del Sector Presidencia de la República. Mayo 26 de 2015.

### **Referencias Reglamentarias**

Resolución 1227 de 2013 [Universidad Industrial de Santander]. Por la cual se aprueba el Manual de Procedimientos Administrativos para el Tratamiento de Datos Personales. Agosto 22 de 2013.

**Apéndice**

**Apéndice A. Estado del arte: Polymorphisms Associated to Genetic Susceptibility and Head and Neck Cancer Risk: A Systematic Review (Documento anexo).**

**Apéndice B. Tabla 3. Protocolo de amplificación de PCR**

	[Stock]	[]/Individuo	Vol/Individuo
Master mix MyFiTM Mix (Bioline®)	2X	1X	15 µL
Mezcla de Primers	20 µM	0.4 µM	0.6 µL
H <sub>2</sub> O			12.4 µL
ADN	50 ng/µL	50 ng/µL	2 µL
Volumen Final de Reacción			30 µL

**Apéndice C. Tabla 4. Condiciones de Termociclado - Técnica de PCR**

Periodo de Incubación	30 Ciclos			Extensión Final	Período Final
	Denaturación	Anillamiento	Extensión		
94°C	94°C	58°C	72°C	72°C	4°C
2 min	45 s	60 s	60 s	10 min	Infinito

Nota: La temperatura de anillamiento o Annealing varía (58°C – 63°C) de acuerdo con el fragmento del gen *AHR* a amplificar.

**Apéndice D. AHR gene variants in patients with cancer and their presence in the Latin American population (Documento Anexo).**

**Apéndice E. Tabla 16. Análisis in silico de la variante rs2066853 [c.1661G>A] mediante la herramienta Human Splicing Finder (HSF)**

Gen	Variante	Motivo	Signal	Interpretation	Associated transcripts	Algorithme/Matrix	Position	Sequence
AHR	rs2066853 [c.1661G>A]	-ESE	Alteration of auxiliary sequences	Significant alteration of ESE / ESS motifs	ENST00000242057 ENST00000463496 ENST00000642825	RESCUE ESE (ESE Site Broken)	chr7:17339481	CATCAG
		-ESE				EIE (ESE Site Broken)	chr7:17339482	ATCAGA
		-ESE				ESE_9G8 (ESE Site Broken)	chr7:17339483	TCAGAC
		-ESE				EIE (ESE Site Broken)	chr7:17339483	TCAGAC
		+ESE				RESCUE ESE (New ESE Site)	chr7:17339484	CAAACA
		-ESE				EIE (ESE Site Broken)	chr7:17339484	CAGACA
		+ESE				ESE_Tra2 (New ESE Site)	chr7:17339486	AACAC
		-ESE				ESE_ASF (ESE Site Broken)	chr7:17339486	GACACAT

Nota: la variante rs2066853 podría dar lugar a la omisión del exón a través de la inactivación de ESE o la activación de ESS. (+): la mutación creó un nuevo sitio; (-): el motivo fue abolido por la mutación. Algoritmos y matrices utilizados para identificar los motivos fueron: RESCUE ESE hexamers; EIEs; ESE\_9G8; ESE\_Tra2; ESE\_ASF

**Apéndice F. Tabla 19. Análisis in silico de la variante rs7796976 [-459A>G] mediante la herramienta Human Splicing Finder (HSF)**

Gen	Variante	Motivo	Signal	Interpretation	Associated transcripts	Algorithme/Matrix	Position	Sequence
AHR	rs7796976 [-459A>G]	-ESE	Alteration of auxiliary sequences	Significant alteration of ESE / ESS motifs	ENST00000242057 ENST00000463496	EIE (ESE Site Broken)	chr7:17298801	GATTTA
		+ESE				Sironi_motif2 (New ESS Site)	chr7:17298802	ATTTGGG
		-ESE				Fas ESS (ESS Site Broken)	chr7:17298803	TTTAGG
		+ESE				Fas ESS (New ESS Site)	chr7:17298803	TTTGGG
		+ESE				Sironi_motif2 (New ESS Site)	chr7:17298803	TTTGGGA
		+ESE				IIE (New ESS Site)	chr7:17298804	TTGGGA
		-ESE				ESE_9G8 (ESE Site Broken)	chr7:17298805	TAGGAA
		-ESE				ESS_hnRNPA1 (ESS Site Broken)	chr7:17298805	TAGGAA
		+ESE				EIE (New ESE Site)	chr7:17298805	TGGGAA
		+ESE				IIE (New ESS Site)	chr7:17298805	TGGGAA
		-ESE				RESCUE ESE (ESE Site Broken)	chr7:17298806	AGGAAG

Nota: la variante rs7796976 podría dar lugar a la omisión del exón a través de la inactivación de ESE o la activación de ESS. (+): la mutación creó un nuevo sitio; (-): el motivo fue abolido por la mutación.

Algoritmos y matrices utilizados para identificar los motivos fueron: RESCUE ESE hexamers; EIEs; ESE\_9G8; Sironi\_motif2; Fas ESS; IIEs; ESS\_hnRNPA1.

**Apéndice G. Tabla 20.** *Análisis de la variante rs7796976 [-459A>G] mediante la herramienta*

*HaploReg V4.1: unión de proteínas en experimentos ChIP-Seq (ENCODE Project Consortium)*

Cell ID	Protein	GM19193	POL2
A549	POL2	H1-hESC	POL2
A549	POL2	H1-hESC	POL2
GM12878	POL2	H1-hESC	TAF1
GM12878	POL2	H1-hESC	TAF1
GM12878	POL2	H1-hESC	TAF7
GM12891	POL2	H1-hESC	TBP
GM12891	TAF1	HCT-116	POL2
GM12892	POL2	HCT-116	POL24H8
GM12892	POL2	HEK293(b)	POL2
GM12892	TAF1	HeLa-S3	HAE2F1
GM12892	TAF1	HeLa-S3	POL2
GM15510	NFKB	HeLa-S3	POL2
GM15510	POL2	HeLa-S3	POL2
GM18526	POL2	HepG2	POL2
GM18951	POL2	HepG2	POL2
		HepG2	POL2
		HepG2	TBP
		MCF10A-Er-Src	POL2
		MCF10A-Er-Src	POL2
		PBDE	POL2

Nota: sitios potenciales de unión de la variante rs7796976 a distintos factores de transcripción (POLR2A; CTCF; TAF1; TBP; NF-kB) en diversas líneas celulares, según datos reportados de los experimentos ENCODE CHIP-Seq.

**Apéndice H. Tabla 21.** *Análisis de la variante rs7796976 [-459A>G] mediante la herramienta HaploReg V4.1: motivos regulatorios alterados (Position Weight Matrix)*

Position Weight Matrix ID (Library from <a href="#">Kheradpour and Kellis, 2013</a> )	Strand	Ref	Alt	Match on:
				Ref: GCTCCGCAGGCGGGAAGCACCCCTGGATTTAGGAAGTCCCGGGAGCAGCGCGGCGGCACC Alt: GCTCCGCAGGCGGGAAGCACCCCTGGATTGGGAAGTCCCGGGAGCAGCGCGGCGGCACC
BCL_disc6	+	1.2	13.2	GGGRAADYCC
ELF1_known1	+	11.7	8.9	RDWASAGGAARW
Ik-1_1	+	9	12.5	NHKYGGGAATRY
Ik-1_2	+	10.7	12.9	KYTRGRRR
Ik-2_1	+	9.1	11.4	VDBYGGGAWNNM
NF-kappaB_disc1	+	-9.4	2.6	KGGRAWTYCC
NF-kappaB_known1	+	1	13	GGGGAKYCCC
NF-kappaB_known2	-	11.6	15.7	GGRRAWYCCC
NF-kappaB_known3	-	5.2	17.1	KGRRADTCCCY
NF-kappaB_known4	+	1.3	13	DKGGGAMTYCCBN
NF-kappaB_known5	+	-12.1	-0.2	SGRRRTTCC
PU.1_known1	+	12.5	10.1	AGGAAGT
STAT_known13	-	10.1	9.8	RRGAARBB
STAT_known5	-	7.3	8.5	VRGAABNN

Nota: matriz de pesos de posición (PWM). La variante rs7796976 (A/G) podría presentar un efecto en 6 de los 11 motivos reguladores de NF-Kb que interactúa con el receptor AhR.

**Apéndice I. Tabla 24.** *Análisis in silico de la variante rs2074113 [c.908+33G>T] mediante la herramienta Human Splicing Finder (HSF)*

Gen	Variante	Signal	Interpretation	Associated transcripts	Algorithme/ Matrix	Position	Sequence	Variation
AHR	rs2074113 [c.908+33G>T]	New Donor splice site	Activation of a cryptic Donor site. Potential alteration of splicing	ENST00000242057 ENST00000463496 ENST00000642825	HSF Donor site (matrix GT)	chr7:17334143	ATTGGATGT> ATTGTATGT	49.25>76.39 (55.11%)
AHR	rs2074113 [c.908+33G>T]	New Donor splice site	Activation of a cryptic Donor site. Potential alteration of splicing	ENST00000242057 ENST00000463496 ENST00000642826	MaxEnt Donor site	chr7:17334143	ATTGGATGT> ATTGTATGT	-3.31>4.34 (231.12%)



Nota: la variante rs2074113 podría dar lugar a la activación de nuevos sitios de empalme crípticos. Algoritmos y matrices utilizados para identificar nuevos sitios de splicing fueron: HSF Donor site; MaxEnt Donor site.

**Apéndice J. Tabla 25.** *Análisis de la variante rs2074113 [c.908+33G>T] mediante la herramienta HaploReg V4.1: motivos regulatorios alterados (Position Weight Matrix)*

Position Weight Matrix ID (Library from <a href="#">Kheradpour and Kellis, 2013</a> )	Strand	Ref	Alt	Match on:
				Ref: AGTGAGACTTTTTCACCTTTTATTTTATTGGATGTACATTATGTTTCAGTAAGTCTCTCT Alt: AGTGAGACTTTTTCACCTTTTATTTTATTGTATGTACATTATGTTTCAGTAAGTCTCTCT
Foxp1	+	-21	-9.1	TTATTTTRTRTTKKYKWKWT
Hoxa10	-	11.5	12.6	WDHWTTTATKRYVKH
Hoxa9	-	11.6	10.8	WHHWTTTATKRCYBY
Hoxb13	-	16.8	13	BDDWTTTATTRGNYB
Hoxd10	-	17.5	17.7	DHHHDTTWTATKRNVBH
Irf_disc1	-	12.5	0.6	YBATTGGYYR
Sox_18	-	6.4	11.5	HDWTBWATTGWTHHB
Sox_2	-	6.3	12	DNDWNNDYHATTGTTTHDHVDD
TEF	-	2	13.7	TTRYDWDWVYAW
Zfp410	+	11.4	5.2	DNDDATGGGATGKMDDN

Nota: matriz de pesos de posición (PWM). La variante rs2074113 podría presentar un efecto en diferentes motivos reguladores, entre estos, Foxp1; Hoxb13; Irf\_disc1; TEF; Zfp410.

**Apéndice K.** *Aval otorgado por el Comité de Ética de la Universidad Industrial de Santander (CEINCI) para el desarrollo de la presente investigación*

4110

Bucaramanga,

Estudiante  
**NATHALIA ANDREA TRUJILLO PELAYO**  
 Investigadora principal  
 Maestría en Ciencias Básicas Biomédicas  
 Escuela Medicina  
 Facultad de Salud  
 Universidad Industrial de Santander

Asunto: Aval Comité de Ética proyecto, "Identificación de variantes genéticas en el gen AHR que puedan contribuir a la susceptibilidad a cáncer de cabeza y cuello en población colombiana".

Cordial Saludo. El Comité de Ética en Investigación Científica de la Universidad Industrial de Santander (CEINCI-UIS) en reunión realizada del día 7 de febrero de 2020, según consta en el Acta N° 1°, evaluó los ajustes realizados al proyecto del asunto y al respecto conceptúa:

En consideración a que el proyecto cumple con todos los requerimientos del CEINCI-UIS, el Comité acuerda por consenso, **APROBAR** el documento en digital y el consentimiento informado en su última versión.

Se solicita emplear las estrategias que considere necesario, para verificar que el consentimiento informado ha sido comprendido por los participantes. De otra parte, adoptar los mecanismos necesarios para garantizar la confidencialidad de la información recabada. Todo ello amparado en lo reglamentado en la Ley Estatutaria 1581 de 2012 del Congreso de la República de Colombia, por la cual se dictan disposiciones para la protección de datos personales, en la Resolución de Rectoría 1227 del 22 de agosto de 2013, sobre el tratamiento de datos personales. Además, recomendamos tener presente los criterios y procedimientos establecidos por el laboratorio de genética humana de la UIS y el laboratorio de Biología molecular de la UCC de Pasto, para el acceso a la información consignada en las historias clínicas de los pacientes. Igualmente realizar los trámites necesarios en las instituciones para acceder a la información y a los participantes. Finalmente, socializar los resultados generados en este proyecto en las instancias correspondientes.

Se solicita que se remita al correo del Comité, información de las siguientes circunstancias, cuando lleguen a ocurrir:

- Reporte de mala práctica científica por parte de cualquier miembro del equipo investigador.
- Notificación previa de las modificaciones realizadas al protocolo.
- Reporte de cualquier eventualidad que usted considera deba conocer el CEINCI-UIS.

Elaboró César Hastamorir, revisó Francisco Espinel y aprobó José Luis Osma Rueda



4110

- Informe de avance sobre los aspectos éticos según guía e instructivo anexo. Este informe debe enviarse a la mitad del desarrollo de la investigación y al finalizar la misma según el cronograma establecido en el Formato FIN 65.
- El informe debe ser enviado al correo: [ceinci.seguimientos@uis.edu.co](mailto:ceinci.seguimientos@uis.edu.co)

En el momento del seguimiento se verificará el cumplimiento de las consideraciones éticas.

En nombre del CEINCI-UIS le ofrecemos el apoyo que usted considere necesario, para la aplicación y salvaguarda de los asuntos éticos durante la investigación.

**JOSE LUIS OSMA RUEDA**  
 Presidente  
 CEINCI- UIS

**FRANCISCO ESPINEL CORREAL**  
 Secretario  
 CEINCI- UIS

Copia: Profesoras Laura Fernanda Cifuentes Cardona (Escuela de Biología), directora del trabajo de investigación; Clara Inés Vargas Castellanos (Departamento de Ciencias Básicas), codirectora del trabajo de investigación.

Profesor Fernando Rodríguez Sanabria, Coordinador de la Maestría en Ciencias Básicas Biomédicas.  
 Archivo Comité de Ética en Investigación Científica.

**Apéndice L.** *Aval otorgado por el Comité de Ética de la Universidad Cooperativa de Colombia, sede Pasto, para el desarrollo de la presente investigación*



Subcomité de Bioética  
Universidad Cooperativa de Colombia – Pasto  
Calle 18 No. 47-150 Teléfono: 7336706

### **Memorando al Investigador Principal**

**Memorando No.** SCBE11-17

**ANT.:** Acta del Consejo del Subcomité de Bioética

**MAT.:** Remite Acta de Evaluación.

**Fecha:** 02 de Agosto de 2017.

**LAURA FERNANDA CIFUENTES CARDONA**  
Investigador Principal

**INVESTIGADOR PRINCIPAL**  
**Presente**

De nuestra consideración:

Adjunto envío a Ud., Consentimientos y Acta de Evaluación No. 02 del 04 de mayo de 2016 que certifica que el Consejo del Subcomité de Bioética de la Universidad Cooperativa de Colombia - Pasto, **APROBÓ** el proyecto de investigación denominado "IDENTIFICACIÓN DE VARIANTES GENÉTICAS EN EL GEN AHR Y SU ASOCIACIÓN CON CÁNCER DE CABEZA Y CUELLO EN LA POBLACIÓN COLOMBIANA"

**Apéndice M. Consentimiento informado de los individuos participantes del presente estudio.**



**Versión 02**

**Código:** \_\_\_\_\_

**CONSENTIMIENTO PARA PARTICIPAR EN UNA INVESTIGACIÓN  
Apartado para el participante**

A Usted se le ha solicitado participar en el proyecto de investigación "IDENTIFICACIÓN DE VARIANTES GENÉTICAS EN EL GEN AHR Y SU ASOCIACIÓN CON CÁNCER DE CABEZA Y CUELLO EN LA POBLACIÓN COLOMBIANA". El propósito de este documento es proveer información para considerar, o no, participar en el mismo. Su consentimiento debería estar basado en el entendimiento de la naturaleza de los riesgos del tratamiento, mecanismos o procedimientos. Por favor realice preguntas si hay alguna cosa que no entiende. Su participación es voluntaria y no tendrá efecto sobre el proyecto de investigación si decide no participar.

Así mismo, son requeridos una serie de datos personales que se solicitan en el formulario de Inscripción en nuestra base de datos. Antes de que usted decida entregar esta información es importante que lea y firme el presente consentimiento a cerca de los detalles del manejo que recibirán sus datos personales.

**INVESTIGADORES RESPONSABLES DEL ESTUDIO**

**Principal(es):**

1. Laura Fernanda Cifuentes Cardona. Docente. Universidad Cooperativa de Colombia sede Pasto. Facultad de Odontología.
2. Nathalia Andrea Trujillo Pelayo. Estudiante de Maestría en Ciencias Básicas Biomédicas. Universidad Industrial de Santander.  
Correo electrónico: nathalia.trujillo@correo.uis.edu.co

**Coinvestigador(es):**

\_\_\_\_\_

UNIVERISIDAD INDUSTRIAL DE SANTANDER  
 COMITÉ DE ÉTICA U I FEB 2020



3. Clara Inés Vargas Castellanos. Docente Universidad Industrial de Santander. Facultad de Salud.
4. Álvaro Herrera. Médico. Cirujano de Cabeza y Cuello. Hospital Universitario de Santander; Clínica Foscal Internacional.
5. Cristian Javier Fong Reales. Docente Universidad Cooperativa de Colombia sede Pasto. Facultad de Odontología.
6. Pedro Hidalgo. Docente Universidad Cooperativa de Colombia sede Pasto. Facultad de Odontología.
7. Sandra Milena Guauque Olarte. Universidad Cooperativa de Colombia. Sede Pasto. Facultad de odontología.
8. Constanza Duque Vélez. Universidad Cooperativa de Colombia. Sede Medellín. Facultad de Odontología.

#### **FUENTES DE APOYO**

CONADI- Comité Nacional para el Desarrollo de la Investigación de la Universidad Cooperativa de Colombia (UCC)

#### **LUGAR DONDE SE DESARROLLARÁ EL ESTUDIO**

Se realizará en diferentes instituciones que reciben pacientes oncológicos en las ciudades de Bucaramanga y Pasto.

#### **OBJETIVO DEL ESTUDIO**

- Identificar variantes genéticas que puedan contribuir con la susceptibilidad al desarrollo de cáncer de cabeza y cuello asociado al consumo de cigarrillo

#### **JUSTIFICACIÓN DEL ESTUDIO**



el cáncer de cabeza y cuello corresponde a una neoplasia (tumor maligno), con un origen predominante de tipo escamoso-celular, y compromete principalmente las estructuras de cavidad oral, faringe y laringe. Se ha establecido el consumo de tabaco como principal factor de riesgo para este tipo de cáncer.

A pesar de que los factores ambientales contribuyen de forma importante al desarrollo de cáncer, el riesgo de presentarlo también está fuertemente influenciado por diferencias genéticas que llevan a variaciones en la susceptibilidad. Cada individuo presenta unidades de información genética (genes), entre las que se encuentra el gen AHR, implicado en la detoxificación de los componentes (carcinógenos) del tabaco, que puede presentar variantes y generar diferencias genéticas entre individuos, lo cual puede relacionarse con una susceptibilidad al desarrollo de cáncer. La presente investigación busca mediante el análisis del gen AHR, identificar variantes que puedan estar asociadas con un mayor riesgo al desarrollo de cáncer de cabeza y cuello, en presencia de tabaquismo. Así mismo, los resultados de este trabajo permitirán generar estrategias de prevención y diagnóstico temprano en personas fumadoras, ya que será posible identificar dentro de este grupo cuales presentan un riesgo aumentado de desarrollar esta neoplasia asociada al consumo de tabaco.

## ELEGIBILIDAD

### Casos

Se contactará a los pacientes por medio de su médico tratante y se les realizará una encuesta para completar los datos y ampliar la información necesaria.

### Criterios de inclusión

- Hombres y mujeres, nacidos en Colombia, mayores de dieciocho años
- Con diagnóstico histopatológico de carcinoma escamocelular en cavidad oral, faringe o laringe
- Diagnóstico correspondiente a los códigos C01, C02, C03, C04, C05, C06, C10, C11, C12, C13, C14 y C32 de la clasificación CIE10
- Fumadores\*



\* Los fumadores serán definidos por el consumo diario de una cantidad igual o mayor a 60mg de tabaco, esta cantidad estará determinada por el número de cigarrillos consumidos diariamente multiplicado por la cantidad de tabaco que se registra en la cajetilla.

#### **Criterios de exclusión**

- Pacientes que no deseen participar en el estudio
- Mujeres en estado de embarazo
- Pacientes que por su condición (fallecimiento, imposible ubicación, trastorno psiquiátrico), no sea posible tomar la muestra y/o completar la información de la encuesta.

#### **Controles**

Individuos sanos que cumplan con los siguientes criterios:

- Individuos que no hayan presentado ningún tipo de cáncer durante su vida
- Personas fumadoras\*, que fumen actualmente o hasta 12 meses antes del momento del enrolamiento
- Los controles serán pareados con los casos de acuerdo con edad, sexo y hábito de fumar

#### **Procedimiento**

Los participantes del estudio firmarán un consentimiento informado. Se realizará una toma de muestra de sangre, la cual será realizada en centros médicos por personal de la salud entrenado. A partir de la muestra de sangre se extraerá el material genético (DNA) que permitirá identificar variantes genéticas que puedan favorecer una susceptibilidad al desarrollo de cáncer de cabeza y cuello asociado al consumo de cigarrillo en la población colombiana. Se tomarán datos de la historia clínica de cada paciente (casos) relacionados a resultados de sus exámenes diagnósticos, histopatológicos. Cada participante llenará una encuesta sobre historia familiar oncológica y exposición a agentes ambientales. Los datos obtenidos de cada participante se usarán hasta finalizado el proyecto o en subsecuentes



estudios solo si el participante lo acepta. Este estudio tiene una duración estimada de 12 meses.

**La toma de muestra de personas residentes en la ciudad de Bucaramanga y el área metropolitana será realizada en las instalaciones del laboratorio de genética humana, ubicado en la Carrera 32 No. 29-31. Facultad de Salud de la Universidad Industrial de Santander. Contacto: Nathalia Andrea Trujillo Pelayo**  
**La toma de muestras en la ciudad de Pasto será realizada en el laboratorio de Biología molecular, ubicado en la Calle 18 No. 47 -150B Torobajo, Universidad Cooperativa de Colombia- campus Pasto. Contacto: Laura Fernanda Cifuentes Cardona**

Los resultados de estas pruebas serán puestos a disposición de su médico si Usted lo requiere y el asesoramiento genético se realizará en el estudio sin costo alguno para Usted.

**Banco de Especímenes (muestras)**

En este estudio a partir de su sangre se extraerá material genético (ADN), si acepta que su muestra se utilice en estudios futuros, los datos serán codificados y se eliminará todo rastro que pueda relacionar los resultados del nuevo proyecto con su identidad. Por lo tanto, aunque existe la posibilidad de revelar más información genética, esta información no podrá asociarse con usted para que no afecte en ninguna manera su confidencialidad.

**Riesgos**

Existe un riesgo mínimo asociado a la toma de muestra de sangre. Para minimizar el riesgo durante la toma de muestra de sangre, esta se hará acorde a la guía para extracción de sangre de la Organización Mundial de la Salud.  
 En casos aislados, estos riesgos pueden incluir infecciones, dolor en el sitio de toma de la muestra, hematomas o mareo. Para evitarlos se seguirán las estrategias de reducción de riesgo de la guía que incluyen toma de muestras por personal entrenado, uso de jeringas nuevas y estériles, informarse sobre uso de

UNIVERSIDAD INDUSTRIAL DE SANTANDER  
 COMITÉ DE ÉTICA 07 FEB 2020



anticoagulantes y preguntar al paciente sobre posible alergia al látex o alcohol antes de iniciar el procedimiento.

#### **Beneficios**

Los pacientes analizados se informarán de los resultados del estudio, se les llevará cabo consejería genética respecto al nivel de riesgo encontrado para ellos y sus familias.

Analizar el perfil genético de este gen implicado en la detoxificación de carcinógenos del tabaco en la población, permitirá identificar las variantes genéticas que puedan estar asociadas con el desarrollo de cáncer de cabeza y cuello, y a futuro, permitirá implementar estrategias de prevención y protocolos de diagnóstico temprano.

#### **Costos asociados al Estudio**

No existe costo alguno para los participantes involucrados en el estudio.

#### **Tratamiento Médico/Psicológico**

No existen tratamientos médico o psicológico asociados al estudio.

#### **Compensación por la participación**

No existen compensaciones económicas por participar en el estudio.

#### **Confidencialidad**

Durante el desarrollo de este estudio, los investigadores se comprometen a actuar con responsabilidad y proteger la privacidad, custodiando la información de los participantes mediante el uso de bases de datos y archivos físicos, estableciendo lo siguiente para el tratamiento de datos personales:



- Los datos recolectados para la investigación son de uso confidencial, tanto la información suministrada en las encuestas, como la que sea tomada de las historias clínicas. No se almacenarán datos personales de los pacientes incluidos en el estudio.
- La información obtenida será guardada en medios físicos tales como documentos y magnéticos como discos de almacenamiento externo, en dos ejemplares que serán resguardados en sitios diferentes cada uno y custodiados por el investigador principal quien se encargará de tomar las medidas necesarias de seguridad para eliminar cualquier tipo de acceso a terceras personas no vinculadas con el trabajo de investigación.
- Dado que la presente investigación tiene una finalidad científica, la información será conservada en una base de datos que permita su análisis, en la cual se adaptarán las medidas adecuadas para la supresión de identidad de los pacientes (uso de códigos anónimos).
- En ninguna circunstancia se traspasarán los datos personales de los participantes a personas ajenas. Con lo anterior queda establecido que no habrá sesión de datos a terceras personas. Lo anterior con excepción de las entidades públicas o administrativas en ejercicio de sus funciones legales o por orden judicial que así lo requieran.

Sus datos personales siempre le pertenecerán. Como dueño o titular de estos tiene derecho a solicitarle al responsable del registro que:

- Sean modificados
- Sean eliminados del registro
- Se le entregue personalmente una copia del registro que evidencie las modificaciones o eliminación de sus datos personales de la base de datos
- Sea informado por el responsable de la investigación, respecto del uso que le ha dado a sus datos personales.

**Libertad para salir del estudio**

UNIVERSIDAD INDUSTRIAL DE SANTANDER  
 COMITÉ DE ÉTICA U 7 FEB 2020



Los participantes tendrán libertad de retirar su consentimiento y salir del estudio en el momento que deseen y por tanto retirar la información que haya aportado en la investigación. Los participantes tienen la libertad de no responder alguna(s) pregunta(s) en particular si así lo desean.

**Libertad para ser removido del estudio**

Los investigadores tendrán la libertad de remover a un participante del estudio.

**El presente trabajo se ampara a las consideraciones éticas de:**

No maleficencia (no se causará ningún daño físico ni mental a los pacientes del estudio), Beneficencia (los resultados obtenidos pueden servir para dar inicio a otras investigaciones derivadas de los hallazgos de este estudio), Justicia (los pacientes serán tratados con respeto y equidad y se empleará un lenguaje adecuado) y Autonomía (sólo se incluirán en el estudio pacientes que participen voluntariamente).



**CONSENTIMIENTO VOLUNTARIO**  
 Apartado para archivar el investigador

Todos los aspectos mencionados en las anteriores páginas me han sido explicados, y todas las preguntas que he tenido han sido contestadas. Yo he sido motivado(a) a preguntar acerca de todos los aspectos de este estudio.

El participante podrá requerir en cualquier momento del estudio copia completa de la política de tratamiento de datos personales que rige esta investigación. Para tal motivo, el participante podrá contactar al investigador principal (**Nathalia Andrea Trujillo Pelayo: Correo electrónico nathatrujillo@gmail.com;** ) y una copia de dicho documento será enviado por medio electrónico o físico, según lo requiera el paciente.

Cualquier pregunta que yo tenga acerca de los derechos de los participantes del estudio será contestada por el investigador principal (**Nathalia Andrea Trujillo Pelayo: Correo electrónico nathatrujillo@gmail.com** / **Laura Fernanda Cifuentes Cardona: Correo electrónico lauracifuentesc@campusucc.edu.co** o en su defecto por el Comité de ética de la Universidad Industrial de Santander (Carrera 19 N° 35-02, oficina 245, UIS Bucarica) o el Sub-Comité de Bioética de la Universidad Cooperativa de Colombia – sede Pasto (Calle 18 No. 47-150

En pleno uso de mis facultades mentales, de manera libre, voluntaria y consciente, autorizo mi participación en el proyecto titulado **“IDENTIFICACIÓN DE VARIANTES GENÉTICAS EN EL GEN AHR Y SU ASOCIACIÓN CON CÁNCER DE CABEZA Y CUELLO EN LA POBLACIÓN COLOMBIANA”**

Habiendo leído lo anterior doy mi consentimiento para participar en las actividades anteriormente detalladas y autorizo que se consulte mi historia clínica para obtener los datos que sean necesarios en el desarrollo de la investigación. (Una copia de este consentimiento me será entregada).

UNIVERSIDAD INDUSTRIAL DE SANTANDER  
COMITÉ DE ÉTICA

17 FEB 2020



En el siguiente cuadro que se presentan a continuación coloque SI o NO dentro del

A. Confirmo haber leído y entendido la información acerca del manejo de mis datos personales.	SI	NO
---	----	----

cuadro.

Además de la investigación a la que usted está consintiendo en este estudio, indique si da su permiso para guardar en un biobanco las muestras de ADN extraídas a partir de su sangre para investigaciones futuras en el campo de la Genética de poblaciones y oncología. Las muestras en futuros estudios serán codificadas y cualquier vínculo con usted será destruido. Las muestras pueden mantenerse después de terminado el presente estudio.

**POR FAVOR INDIQUE**

SI consiento que mi muestra sea guardada para futuros estudios de investigación.

NO consiento que mi muestra sea guardada para futuros estudios de investigación.

Con mi firma acepto y entiendo las condiciones del estudio: "Identificación De Variantes Genéticas en el Gen AHR y su Asociación con Cáncer de Cabeza y Cuello en la Población Colombiana" y autorizo integralmente este consentimiento informado. Si el participante de la investigación no supiera firmar, imprimirá su huella digital y a su nombre firmará otra persona que él designe.

\_\_\_\_\_

**Nombre del participante          Firma del Participante          Lugar y fecha**



---

**Testigo 1**

Nombre:  
Dirección:  
Relación con el participante:

---

**Testigo 2**

Nombre:  
Dirección:  
Relación con el participante:

**Bibliografía**

University at Buffalo. The State University of New York. Research Subjects Protection. [Online] cited 2007. Available from: [http://www.research.buffalo.edu/rsp/irb/health\\_sciences/Informed\\_consent\\_template.cfm](http://www.research.buffalo.edu/rsp/irb/health_sciences/Informed_consent_template.cfm)

7 FEB 2020



**Apéndice N. Instrumento de recolección de datos, tipo encuesta.**

**INSTRUMENTO DE RECOLECCIÓN DE DATOS**

**PROYECTO:** IDENTIFICACIÓN DE VARIANTES GENÉTICAS EN EL GEN AHR Y SU ASOCIACIÓN CON CÁNCER DE CABEZA Y CUELLO EN LA POBLACIÓN COLOMBIANA

CARACTERIZACION SOCIO-DEMOGRAFICA			
No.:	Fecha(dd/mm/aaaa):	Identificación:	Género: F _____ M _____
Nombres Completos:		Apellidos:	
Fecha de Nacimiento: (dd/mm/aaaa) ___/___/___	Lugar de Nacimiento: (Municipio – Dpto.)	Edad: _____ años	Ocupación: _____ _____
Grupo Étnico: Afrodescendiente _____ Mestizo _____ Indígena _____ Otro: _____ Cual _____	Estado Civil: Soltero(a) _____ Casado(a) _____ Unión Libre _____ Divorciado(a) _____ Viudo(a) _____	Nivel Educativo: Ninguno _____ primaria _____ secundaria _____ Técnica _____ Universitario _____ Posgrado _____	Régimen de Salud: Contributivo _____ Subsidiado _____ Especial _____ Ninguno _____
Estrato: 0 ___ 1 ___ 2 ___ 3 ___ 4 ___ 5 ___ 6 ___	Si es mujer: ¿Se encuentra en estado de embarazo? SI ___ NO ___	Para uso de los investigadores: Caso _____ Control _____	

CONSUMO DE TABACO		
1. ¿A lo largo de su vida a consumido tabaco (cigarrillos, habanos, tabaco de mascar, pipa, etc.)?	SI _____	NO _____
Si su respuesta es <b>NO</b> pase a la pregunta número 9, si su respuesta es <b>SI</b> continúe:		
2. ¿Qué edad tenía cuando empezó a fumar?	_____ años	
3. ¿Continúa fumando en la actualidad?	SI _____	NO _____
4. ¿Si ya dejo de fumar, indique hace cuánto tiempo?	Meses: _____	Años: _____
5. ¿Durante cuánto tiempo ha fumado o fuma?	Meses: _____	Años: _____
6. ¿Cuántos cigarrillos fuma o fumaba al día?: _____		
7. ¿De que marca?: _____		
8. ¿De que tipo?: _____		
CONSUMO DE ALCOHOL		
9. ¿consume bebidas alcohólicas?	SI _____	NO _____
Si su respuesta es <b>NO</b> pase a la pregunta número 13. Si su respuesta es <b>SI</b> continúe:		
10. ¿Qué edad tenía cuando empezó a consumir bebidas alcohólicas?	_____ Años	

11. Frecuencia con la que consume o consumía bebidas alcohólicas:	Menos de 5 veces al año: ____ De 5 a 12 veces al año: ____ De 1 a 3 veces al mes: ____ De 1 a 3 veces por semana: ____ más de 3 veces por semana: ____
12. ¿Cuándo consume bebidas alcohólicas usualmente fuma?	SI ____ NO ____

<b>OTROS:</b>	
13. ¿Presenta algunas de las siguientes patologías? Indique con una X Hepatitis ____ VIH ____ Hipertensión Arterial ____ Diabetes ____ Tuberculosis ____ Otra ____ Ninguna ____	
14. ¿Se expone o se expuso frecuentemente al sol?	Si: ____ NO: ____
15. ¿Si contesto que SI especifique el promedio de cuantas horas diarias se expone o se exponía al sol?	Horas al día: _____
16. ¿Razón por la cual se expone al sol?	Trabajo: ____ Recreación: ____ Otra(cual): _____
17. Tras la exposición al sol usted se quema:	Siempre: ____ Habitualmente: ____ Algunas Veces: ____ Raramente: ____ Nunca: _____

<b>CASO</b>	
18. ¿Usted ha sido diagnosticado con algún tipo de cáncer?	SI: ____ NO: ____
Si su respuesta fue SI por favor conteste las siguientes preguntas:	
19. ¿Cuál fue la fecha en la cual le diagnosticaron el cáncer?	Mes: ____ Año: ____
20. ¿Qué cáncer padece?: _____	
21. ¿Tiene antecedentes familiares de cáncer?	SI: ____ NO: ____
22. Si tiene Antecedentes Familiares de cáncer, especifique:	
Parentesco	Tipo de Cáncer
_____	_____
_____	_____
_____	_____
_____	_____
23. ¿ Cuándo le diagnosticaron cáncer usted era fumador activo?: SI ____ NO ____	
24. ¿ Cuándo le diagnosticaron cáncer usted ingería bebidas alcohólicas?: SI ____ NO ____	

25. ¿ Cuándo le diagnosticaron cáncer usted se exponía frecuentemente al sol?: SI \_\_\_ NO \_\_\_

26. ¿ Previo al diagnóstico del cáncer usted había presentado alguna lesión en el sitio donde inicio el cáncer?: SI \_\_\_ NO \_\_\_

DIAGNOSTICO Y LOCALIZACION (Para ser diligenciado por los Investigadores)	
Localización anatómica del cáncer (lugar):	
Clasificación histogénica del tumor:	Carcinoma Escamocelular: _____ Otro: _____ Cual: _____
Grado de invasión:	Carcinoma Insitu: _____ Carcinoma Escamocelular superficialmente Invasivo o Microinvasivo: _____ Carcinoma Escamocelular infiltrante o invasivo: _____ Otro: _____
Gradación de acuerdo a la diferenciación celular:	Grado I (Bien diferenciado): _____ Grado II (Moderadamente diferenciado): _____ Grado III (Carcinoma anaplásico o pobremente diferenciado): _____ Grado IV: _____
CLASIFICACIÓN TNM	
T:(Tamaño del tumor); N: (Nódulos linfáticos implicados); M(Metástasis a distancia) T _____ N: _____ M: _____ NO reportado _____	Localización anatómica de las metástasis: _____
<b>Estadio</b> Estadio I: _____ II: _____ III: _____ IV: _____ No Reportado: _____	

Basada en:

- Cuestionario de Identificación de los Trastornos debidos al Consumo de Alcohol. OMS-OPS. 2001.
- Londoño, Constanza y cols. Cuestionario para la clasificación de consumidores de cigarrillo (C4) para jóvenes. Revista Diversitas - Perspectivas en Psicología. 2011;7(2).
- OMS. ASSIST V3.0. [14/07/2014]. Disponible en: [http://www.who.int/substance\\_abuse/activities/assist\\_spanish.pdf](http://www.who.int/substance_abuse/activities/assist_spanish.pdf)
- Bello S, Flores A, Bello M, Chamorro H. Diagnóstico y tratamiento psicosocial del tabaquismo. Rev Chil Enf Respir. 2009;25:218-30.
- Cuevas B, Zenteno R, Sampieri CL. Identificación de factores y marcadores moleculares de riesgo para cáncer oral. Tesis de maestría en salud pública. Universidad Veracruzana. 2008:1-138. [05/08/2014] Disponible en: <http://www.uv.mx/msp/files/2012/11/BETZAI DACUEVASCORDOBA.pdf>