



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO  
PROGRAMA DE MAESTRÍA Y DOCTORADO EN CIENCIAS MÉDICAS,  
ODONTOLÓGICAS Y DE LA SALUD  
INSTITUTO NACIONAL DE CANCEROLOGÍA  
CIENCIAS MÉDICAS – CIRUGÍA ONCOLÓGICA**

**CARACTERIZACIÓN MOLECULAR DEL CARCINOMA BUCAL DE CÉLULAS  
ESCAMOSAS EN PACIENTES MEXICANOS**

## **TESIS**

**PARA OPTAR POR EL GRADO DE  
MAESTRÍA EN CIENCIAS MÉDICAS**

### **PRESENTA:**

**DR. JAVIER LÓPEZ GÓMEZ**

### **TUTOR PRINCIPAL:**

**DRA. CLAUDIA MARÍA GARCÍA CUELLAR  
INSTITUTO NACIONAL DE CANCEROLOGÍA**

### **MIEMBROS DEL COMITÉ TUTOR:**

**DR. DAVID CANTÚ DE LEÓN  
INSTITUTO NACIONAL DE CANCEROLOGÍA  
DR. DAVID VELÁZQUEZ FERNÁNDEZ  
INSTITUTO NACIONAL DE CIENCIAS DE LA SALUD Y NUTRICIÓN SALVADOR ZUBIRÁN**

**CIUDAD DE MÉXICO, ABRIL DEL 2021**



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

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

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

## **AGRADECIMIENTOS**

A la Universidad Nacional Autónoma de México, que me ha permitido desarrollarme profesionalmente la mitad de mi vida (preparatoria con una beca del 100%, licenciatura, especialidad, subespecialidad, alta especialidad en medicina y maestría).

Al Consejo Nacional de Ciencia y Tecnología por otorgarme una beca para estudiar mi maestría con el número CVU849199 y solicitud 2019-000002-01NACF-13474.

A mi tutora, la Dra. Claudia María García Cuellar, quien siempre estuvo pendiente de mi aprendizaje, así como a los miembros del comité tutor.

Al Instituto Nacional de Cancerología, que me otorgó todas las comodidades y herramientas para desempeñar mi proyecto de investigación: al Dr. David Cantú de León en el área de investigación, y en el área clínica, al Dr. Martín Granados García, jefe del servicio de cabeza y cuello.

A mi madre, Matilde Gómez Tazón, quien tuvo la oportunidad de enterarse del inicio de este proyecto profesional antes de su fallecimiento.

Al resto de mi familia, seres queridos y amigos que siempre me han apoyado para realizar mis aspiraciones personales y profesionales.

**GRACIAS**

## ÍNDICE DE CONTENIDOS

LISTA DE ABREVIATURAS.....	7
RESUMEN.....	8
ABSTRACT .....	9
1. INTRODUCCIÓN .....	11
1.1 Epidemiología.....	11
1.2 Factores de riesgo.....	12
1.3 Carcinogénesis.....	13
1.3.1 Funciones de los genes más comúnmente reportados .....	15
1.3.2 Firmas mutacionales asociadas .....	21
1.4 Presentación clínica del cáncer bucal.....	22
1.5 Tratamiento .....	23
1.6 Pronóstico.....	24
2. PREGUNTA DE INVESTIGACIÓN .....	24
3. PLANTEAMIENTO DEL PROBLEMA .....	24
4. JUSTIFICACIÓN .....	24
5. HIPÓTESIS .....	25
6. OBJETIVOS .....	25
7. METODOLOGÍA.....	25
7.1 Diseño del estudio.....	25
7.2 Criterios de inclusión .....	25
7.3 Criterios de exclusión .....	26
7.4 Criterio de eliminación.....	26
7.5 Tamaño de la muestra .....	26
7.6 Variables analizadas .....	26
7.7 Criterios operacionales.....	26
7.8 Procedimientos clínicos.....	28
7.9 Procedimientos en laboratorio.....	29
7.9.1 Extracción de DNA .....	29

7.9.2 Análisis bioinformático de los datos de secuenciación .....	30
7.9.2.1 Preprocesamiento .....	30
7.8.2.2 Identificación y análisis de variantes .....	31
7.9 Determinación de VPH y VEB a través de PCR .....	32
7.10 Análisis estadístico .....	32
8. RESULTADOS .....	33
8.1 Análisis descriptivo .....	33
8.1.1 Datos demográficos .....	33
8.1.2 Datos de estadificación .....	33
8.1.3 Datos histopatológicos .....	33
8.1.4 Datos de tratamiento .....	34
8.1.5 Datos pronósticos .....	34
8.2 Análisis descriptivo molecular .....	37
8.2.1 Identificación de variantes patogénicas .....	37
8.2.2. Identificación de mutaciones conductoras .....	43
8.2.3 Identificación de vías moleculares .....	43
8.2.4 Carga mutacional tumoral (TMB) .....	47
8.2.5 Firmas mutacionales .....	48
8.3 Factores de riesgo .....	49
8.3.1 Datos de tabaquismo y alcohol .....	49
8.3.2 Resultados de PCR VPH .....	50
8.3.3 Resultados de PCR VEB .....	50
8.4 Análisis de supervivencia .....	50
8.4.1 Variables patogénicas .....	50
8.4.2 Firmas mutacionales .....	52
8.4.3 Carga tumoral .....	55
8.4.4 Factores de riesgo .....	56
9. DISCUSIÓN .....	566
10. CONCLUSIONES .....	62

11. APORTACIONES .....	63
12. PERSPECTIVAS .....	63
13. BIBLIOGRAFÍA .....	64

### ÍNDICE DE TABLAS

Tabla 1. Resumen de los principales estudios de secuenciación en cáncer bucal .....	14
Tabla 2. Funciones en genes mutados en cáncer bucal .....	<b>¡Error! Marcador no definido.</b>
Tabla 3. Mililitros considerados como unidad de bebida de acuerdo con el tipo de bebida .....	27
Tabla 4. ECOG: Eastern Cooperative Oncology Group .....	28
Tabla 5. Criterios utilizados para el filtrado de variantes .....	31
Tabla 6. Secuencias utilizadas para la determinación de VPH y VEB en las muestras de cáncer bucal .....	32
Tabla 7. Características clínico-patológicas de los pacientes .....	35
Tabla 8. Principales funciones de las variantes patogénicas encontradas en el análisis de secuenciación de carcinoma bucal .....	38
Tabla 9. Mutaciones conductoras .....	43
Tabla 10. Vías de señalización alteradas en pacientes con cáncer bucal .....	44
Tabla 11. Número total de mutaciones somáticas encontradas en el genoma de los pacientes con cáncer bucal .....	48
Tabla 12. División de pacientes de acuerdo con la carga mutacional en alta y baja .....	48
Tabla 13. Resultados de determinación de VEB en pacientes con cáncer bucal .....	50
Tabla 14. Análisis de supervivencia (Kaplan Meier y regresión de Cox) para supervivencia global y libre de enfermedad de variantes patogénicas .....	51
Tabla 15. Análisis de supervivencia (Kaplan Meier y regresión de Cox) para supervivencia global y libre de enfermedad de firmas mutacionales .....	53
Tabla 16. Análisis multivariado de Cox, para supervivencia global y libre de enfermedad en los grupos de carga tumoral.....	55
Tabla 17. Análisis de supervivencia (Kaplan Meier y regresión de Cox) para supervivencia global y libre de enfermedad de factores de riesgo.....	56

## ÍNDICE DE FIGURAS

Figura 1. Incidencia y mortalidad por cáncer en ambos sexos a nivel mundial en 2020.....	11
Figura 2. Incidencia y mortalidad por cáncer en ambos sexos en México en 2020 .....	12
Figura 3. Porcentaje de afección por localización anatómica de las muestras de cáncer de cabeza y cuello del TCGA .....	14
Figura 4. Firmas mutacionales de sustituciones simples descritas en diferentes tipos de cáncer. En el cuadro rojo se señalan las firmas encontradas en cáncer de cabeza y cuello .....	22
Figura 5. Figura representativa del electroferograma obtenido de las muestras de cáncer bucal. .	29
Figura 6. Supervivencia global (se separan las curvas, mostrando una clara tendencia a la peor supervivencia global en el cáncer avanzado) .....	35
Figura 7. Supervivencia libre de enfermedad (debido a los pocos casos de recurrencia no se observa en el gráfico la separación de las curvas de forma adecuada).....	35
Figura 8. Variantes patogénicas.....	42
Figura 9. Análisis de enriquecimiento de las vías alteradas en cáncer bucal de acuerdo con el número de mutaciones conductoras encontradas .....	45
Figura 10. Análisis de enriquecimiento de las funciones moleculares alteradas en cáncer bucal de acuerdo con el número de mutaciones conductoras encontradas .....	46
Figura 11. Análisis de enriquecimiento de los procesos biológicos alterados en cáncer bucal de acuerdo con el número de mutaciones conductoras encontradas .....	47
Figura 12. Mapa de calor de la distribución genómica del tumor según el perfil de firmas mutacionales encontradas. SB-firmas de sustitución simple .....	49
Figura 13. Supervivencia global (Kaplan Meier) en relación con la variante patogénica FAT1 .....	51
Figura 14. Análisis de supervivencia global y libre de enfermedad de acuerdo con la carga tumoral .....	55

## LISTA DE ABREVIATURAS

**IARC:** International Agency for Research on Cancer

**OMS:** Organización Mundial de la Salud

**DNA:** Ácido desoxirribonucleico

**RNA:** Ácido ribonucleico

**VIH:** Virus de la Inmunodeficiencia Humana

**AJCC:** American Joint Committee on Cancer

**INCAN:** Instituto Nacional de Cancerología

**DIN:** DNA Integrity Number

**ECOG:** Eastern Cooperative. Oncology Group

**PCR:** Reacción en Cadena de la Polimerasa

**TCGA:** The Cancer Genome Atlas

**ICGC:** International Cancer Genome Consortium

**VPH:** Virus del papiloma humano

**VEB:** Virus Epstein Barr

**MMR:** Mismatch Repair

**QT:** Quimioterapia

**RT:** Radioterapia

**TMB:** Tumor Mutational Burden



## RESUMEN

Dentro de las neoplasias malignas de cabeza y cuello, la cavidad bucal es la región más frecuentemente afectada y más del 90% de esas neoplasias corresponde a carcinomas de células escamosas. Los principales factores de riesgo son el consumo de tabaco y la ingesta de alcohol. En México, cerca del 70% de estas neoplasias son detectadas en etapas localmente avanzadas; únicamente el 25%, se detecta en etapas tempranas; y metastásicos en el 4% de los casos. La resección quirúrgica es el tratamiento de elección y, en muchas ocasiones, el manejo requiere quimioterapia y radioterapia. En cuanto a factores pronósticos, el más importante es la presencia de metástasis en los ganglios linfáticos, seguido de la etapa clínica.

A nivel molecular se han reportado las siguientes mutaciones a través del proceso de carcinogénesis bucal en los siguientes genes: *P53* (76-80%), *FAT1* (30-40%), *NOTCH1* (20-22%), *CASP8* (4-34%) y *PIK3CA* (6-20%). En poblaciones asiáticas (Japón, Taiwán, Malasia y China) se encuentran tasas más bajas de mutación en *P53* (35% –62%) y más elevadas de *FAT1*. En las poblaciones de India, Arabia y Taiwán se han encontrado diferencias con respecto a los genes clásicos descritos en población caucásica, estos resultados sugieren que las diferentes poblaciones presentan diferentes tipos de mutaciones. Por lo que es importante estudiar alteraciones genéticas en poblaciones con diferente origen étnico, para lograr encontrar firmas mutacionales asociadas al desarrollo de cáncer bucal

Las firmas mutacionales relacionadas con cáncer de cabeza y cuello, principalmente en sustituciones simples (SBS), son las siguientes: SBS 1, relacionada con la desaminación de 5-metilcitosina y que tiene que ver con la edad del paciente; SBS 2 y 13, con actividad en la vía APOBEC asociada con inflamación; y SBS 5, relacionada con la edad y el tabaquismo. En cavidad oral únicamente están reportadas las firmas SBS 4 relacionada a consumo tabaco y la firma 5.

Debido a la alta incidencia de cáncer de cavidad bucal, se realizó un estudio prospectivo para Identificar mutaciones en carcinoma bucal de células escamosas mediante secuenciación masiva en pacientes mexicanos y asociarlas a pronóstico, se incluyeron 20 pacientes con cáncer bucal de células escamosas, de los cuales se realizó la recolección de datos demográficos, estadificación, tratamiento y pronóstico. Se hizo la toma de muestra del tumor y de tejido adyacente, de los cuales se extrajo DNA y se realizó secuenciación del exoma para la identificación de variantes así como las vías de señalización afectadas. Además se determinó la presencia de VPH y VEB mediante PCR y se correlacionó la carga tumoral con la supervivencia global y libre de enfermedad.

Se encontró el gen *P53* se encuentra mutado en un 47% de las muestras, porcentaje similar al de las poblaciones asiáticas (35-62%), comparado con poblaciones caucásicas (76-80%). Los genes *NOTCH1* (26%), *RASA1* (11%) y *CASP8* (5%) se encontraron mutados en intervalos reportados en estudios caucásicos, asiáticos, árabes e indios. El gen *FAT1* se encontró mutado en porcentaje menor (21%) que en las poblaciones caucásicas y asiáticas, el cual codifica para proteínas de adhesión en epitelios. Los genes *KMT2D* y *FBXW7*, reportados mutados

frecuentemente en estudios caucásicos y que en nuestra muestra representaron el 5% de las mutaciones, tienen funciones en los receptores de beta-globina y estrógenos (*KMT2D*) y degradación de la ciclina E por ubiquitinación (*FBXW7*). El gen *EPHA2*, reportado mutado frecuente en Taiwán, representó el 11%. No se encontró mutación en el gen *PIK3CA*, el cual se encuentra frecuentemente mutado en los estudios reportados. Dentro de los genes que se hallaron frecuentemente mutados (>11%) y que no están reportados en la literatura, se encuentran los siguientes: *PTPRK*, *BIRC6*, *DDX3X*, *KMT2C*, *NIN*, *NXF1* y *SMARCA4*.

Por otro lado encontramos que las vías que más se encuentran afectadas en la mayoría de los pacientes es la vía de señalización de Wnt/ $\beta$ catenina y la vía de *P53*, que correlacionan con las funciones de los genes *FAT1* y *P53*, los cuales se encuentran mutados en este estudio. Las firmas mutacionales encontradas son, principalmente, las vías asociadas al envejecimiento (SBS1), la vía de inflamación vía APOBEC (SBS 13 y 2), el tabaquismo (SBS 5, 4 y 29) y la vía de reparación de desajustes o mismatch repair (SBS 6, 15, 20 y 21). La mayoría de los pacientes que tienen la firma afectada de inflamación y APOBEC, tiene una fuerte correlación con la infección por VEB y/o VPH. Sin embargo, en este trabajo se encontró que la presencia de estos virus no tiene impacto aparente en el pronóstico. Con base en el análisis estadístico de regresión de Cox encontramos que la presencia de mutación en el gen *FAT1* se asocia a una menor supervivencia global y una carga tumoral baja, podría asociarse a una menor supervivencia global y mayor recurrencia.

Conclusiones: con base a nuestros resultados podemos concluir que el perfil mutacional de nuestra población es diferente a lo reportado en otras poblaciones, que las vías que están afectadas en la mayoría son la vía de señalización Wnt/ $\beta$ catenina y la vía *P53*, la mayoría de los pacientes que presentan las firmas relacionadas a APOBEC presentan infección por VEB. que existe una asociación entre la mutación de *FAT1* y una menor supervivencia global y una tendencia a mejor supervivencia global y libre de enfermedad en pacientes con carga tumoral alta.

## ABSTRACT

Within the head and neck malignant neoplasms, the oral cavity is the most frequently affected region and more than 90% of these neoplasms correspond to squamous cell carcinomas. The main risk factors are tobacco consumption and alcohol intake. In Mexico, about 70% are in locally advanced stages, only 25% are in early stages, and 4% are metastatic. Surgical resection is the treatment of choice and in many cases the management requires chemotherapy and radiotherapy. Regarding prognostic factors, the most important is the presence of metastases in the lymph nodes, followed by the clinical stage.

The oral carcinogenesis process is characterized by mutations in the following genes: *P53* (76-80%), *FAT1* (30-40%), *NOTCH1* (20-22%), *CASP8* (4-34%) and *PIK3CA* (6- 20%). In Asian populations (Japan, Taiwan, Malaysia and China) there are lower mutation rates in *P53* (35% -62%)

and higher in *FAT1*. In the populations of India, Arab and Taiwan differences have been found with respect to genes classics described in the Caucasian population.

Mutational signatures related to head and neck cancer mainly in simple substitutions are: SBS 1 related to deamination of 5-methylcytosine and which has to do with the patient's age, SBS 2 and 13: with activity in the APOBEC pathway associated with inflammation, SBS 5: related to age and smoking. In the oral cavity, only the signature SBS 4 and 5 are reported. Regarding insertions and deletions, they are: ID 1 and ID 2 defective DNA repair (mismatch repair) and ID 8 repair of double-stranded breaks by joining non-homologous ends.

A prospective study was carried out in which 20 patients with squamous cell oral cancer were included, from which demographic data, staging, treatment and prognosis were collected, a tumor sample was taken as well as adjacent tissue, from the which DNA was extracted and sent to exome sequencing for the identification of variants as well as affected signaling pathways, the presence of HPV and EBV was determined by PCR and the tumor burden was correlated with global and disease-free survival.

It was found that the *P53* gene mutated in 47% this percentage is similar to the Asian populations (35-62%) compared to caucasian populations (76-80%), the *NOTCH1* gene (26%), *RASA1* (11%) and *CAPS8* (5%) were found in ranges reported in caucasian, asian, arab and indian studies. The *FAT1* gene was found in a lower percentage (21%) than caucasian and asian populations, this gene encodes adhesion proteins in epithelia. The *KMT2D* and *FBXW7* gene, reported as frequent in caucasian studies and which in our sample represented 5%, these genes have functions in beta-globin and estrogen receptors (*KMT2D*) and degradation of cyclin E by ubiquitination (*FBXW7*). The *EPHA2* gene reported as frequent in Taiwan represented 11% in our study. No mutation was found in the *PIK3CA* gene, which is frequently mutated in the reported studies. Among the genes that were frequently mutated in our sample (> 11%) and that are not reported in the literature are: *PTPRK*, *BIRC6*, *DDX3X*, *KMT2C*, *NIN*, *NXF1* and *SMARCA4*.

The pathways that are affected in most patients are the Wnt / $\beta$ cadherin signaling pathway and the *P53* pathway, which coincide with the functions of the mutated *FAT1* and *P53* genes. The mutational signatures found are mainly the pathways associated with aging (SBS1), inflammation pathway via APOBEC (SBS 13 and 2), smoking (SBS 5, 4 and 29) and mismatch repair (SBS 6, 15, 20 and 21). Most of the patients who have the signature affected by inflammation and APOBEC have EBV and / or HPV infection. However, the presence of these viruses has no apparent impact on prognosis. We found that the presence of *FAT1* mutations is associated with lower overall survival and low tumor burden could be associated with lower overall survival and higher recurrence.

# 1. INTRODUCCIÓN

## 1.1 Epidemiología

De acuerdo con los datos de la Agencia Internacional para la Investigación en Cáncer (IARC por sus siglas en inglés) de la Organización Mundial de la Salud (OMS) del 2020, se estima que a nivel mundial, el cáncer bucal representa el lugar 16 en incidencia y mortalidad con 377,713 casos y 177,757 muertes al año. Las mayores tasas de incidencia se encuentran en Sri Lanka, India, Pakistán, Bangladesh, Hungría y Francia (figura 1). En México, se reportaron 1,500 casos y 586 muertes al año, con lo que ocupa el lugar 21 en incidencia y el 24 en mortalidad, y representa el 2% del total de las neoplasias. Es importante mencionar que la estadística, en general, del cáncer y, en particular, del cáncer bucal, está subestimada (figura 2) (1, 2). Dentro de las neoplasias malignas de cabeza y cuello, la cavidad bucal es la región más frecuentemente afectada, y más del 90% de esas neoplasias corresponde a carcinomas de células escamosas (3).

El cáncer bucal es más común en los hombres y generalmente ocurre después de la quinta década de la vida. Debido a la cancerización de campo, se estima que el 1,5% tendrá otro primario sincrónico en la cavidad bucal o en el tracto aero-digestivo (laringe, esófago o pulmón). Los tumores metacrónicos se desarrollan de un 10% a un 40% en la primera década después del tratamiento del tumor inicial (1).

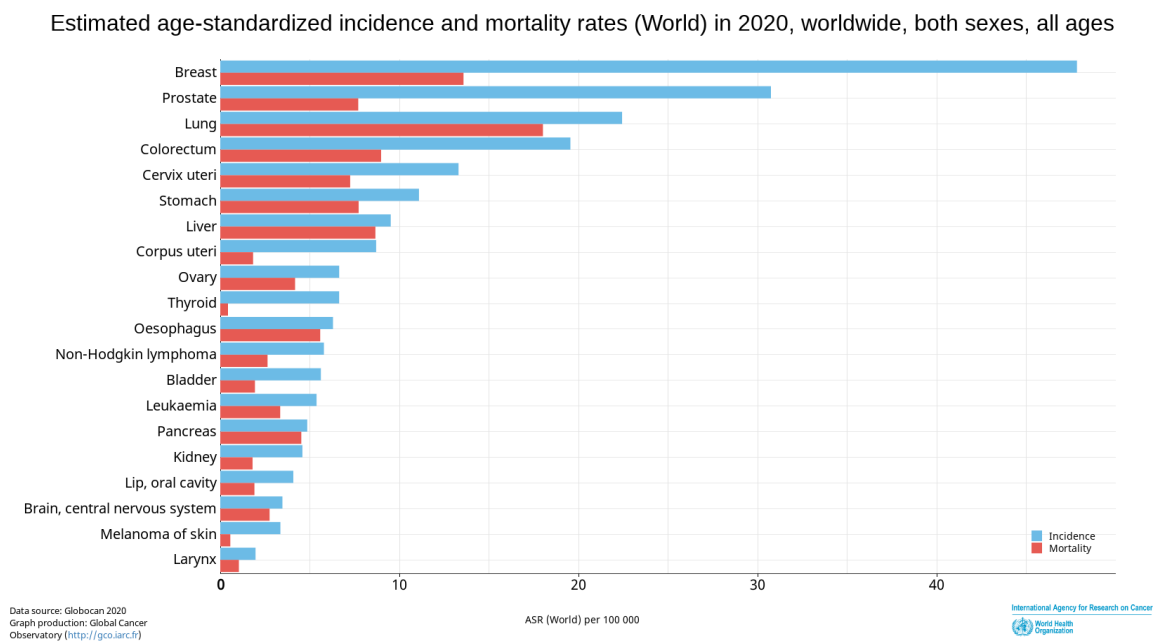


Figura 1. Incidencia y mortalidad por cáncer en ambos sexos a nivel mundial en 2020

Estimated age-standardized incidence and mortality rates (World) in 2020, Mexico, both sexes, all ages

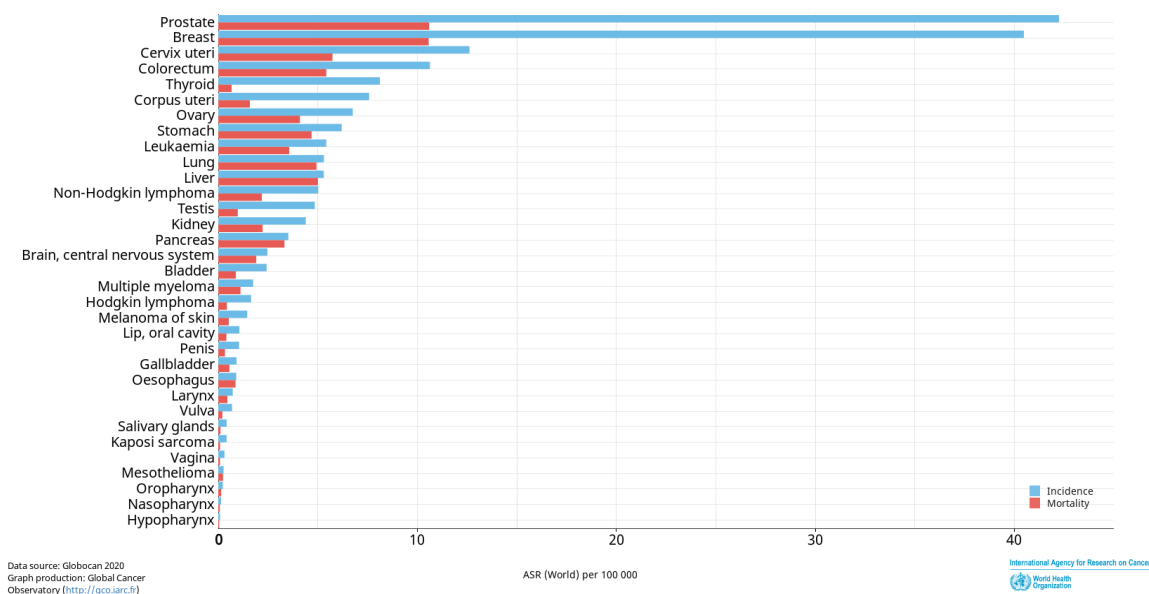


Figura 2. Incidencia y mortalidad por cáncer en ambos sexos en México en 2020

## 1.2 Factores de riesgo

Los principales factores de riesgo relacionados con el desarrollo del cáncer bucal de células escamosas son el consumo de tabaco y la ingesta de alcohol (4-9). Los efectos del tabaco y del alcohol están determinados por la intensidad del hábito, y la exposición simultánea a estos factores tiene un efecto sinérgico. El tabaco representa el principal factor de riesgo y el 65% de los cánceres está relacionado directamente a este hábito. En todas sus presentaciones, los efectos carcinogénicos son atribuidos a nitrosaminas y a los hidrocarburos aromáticos policíclicos que forman aductos con el DNA y pueden conducir a mutaciones. En cuanto al alcohol, el 43% se relaciona directamente al consumo, y su efecto genotóxico es atribuido al acetaldehído, su principal metabolito. Adicionalmente, el consumo de alcohol juega un papel importante como solvente de los carcinogénicos que se encuentran en el tabaco, favoreciendo, entre otros eventos, la producción de especies reactivas de oxígeno. Otras condiciones que se relacionan con aumento del riesgo son las siguientes: pacientes con VIH, postransplantados, con enfermedades inflamatorias intestinales, con infecciones periodontales, condiciones de pobreza, exposición a productos de la industria como solventes, residuos de madera, metal y cemento, y muy importante en Asia, el consumo de nuez de areca y betel que generan inflamación crónica la cual puede generar mutaciones (1, 10-15).

El virus del papiloma humano (VPH) es un virus epiteliotrópico que infecta las células de las capas basales de las membranas de piel y mucosas. Su transmisión es por sexo oral y su carcinogénesis es debido a sus proteínas E6 que degradan P53 y E7; actúa inhibiendo la acción de RB favoreciendo la proliferación celular. La prevalencia de VPH en personas sanas en cavidad oral

es de 4-6.9%; en los pacientes con cáncer bucal también es baja en lo reportado a nivel mundial. En México es del 5%, en su gran mayoría por el serotipo 16. Sin embargo, esto no parece tener un papel en la estadificación, tratamiento y pronóstico como ocurre en orofaringe (16-18).

El virus del Epstein Barr (VEB) tiene un tropismo hacia células B y a las células del epitelio. Está presente en el 90% de la población mundial, se transmite por la saliva y la gran mayoría de las personas se infecta en la infancia. VEB se ha relacionado con el cáncer de nasofaringe, el linfoma NK, el linfoma de Hodgkin y las enfermedades linfoproliferativas. Este virus es detectado en, aproximadamente, el 45% de los casos de cáncer bucal y su presencia se asocia a un aumento del RR (riesgo relativo) de 2.5 a 5. Se ha sugerido la interacción de VEB con infecciones periodontales y coinfecciones con VPH como mecanismos para desarrollar la carcinogénesis. No obstante, su rol específico hasta el momento se desconoce (18, 19).

### 1.3 Carcinogénesis

El proceso de carcinogénesis bucal se caracteriza por la presencia de múltiples eventos, tanto genéticos como epigenéticos, que modifican las funciones de oncogenes y de genes supresores tumorales, dando como resultado alteraciones en diversas vías de señalización, lo cual conduce a un fenotipo capaz de originar una proliferación incrementada, pérdida de la adhesión celular y capacidad para infiltrar tejidos adyacentes, destruyéndolos y diseminándose a sitios distantes (20-22). La acumulación y combinación de dichas alteraciones moleculares, que en gran parte son atribuidas a los principales factores de riesgo, juegan un papel importante en la carcinogénesis bucal. Actualmente, la secuenciación masiva ha permitido identificar alteraciones moleculares en distintos tumores, tales como cáncer de mama y el cáncer de pulmón (23, 24).

Respecto al cáncer bucal, de acuerdo con el Atlas del Genoma del Cáncer (TCGA, por sus siglas en inglés) se han estudiado, hasta al momento, 528 muestras de carcinoma epidermoide en el área de cabeza y cuello, de las cuales el 60.6% corresponde a cáncer bucal y el resto a otras localizaciones (figura 3). Los principales genes que se encontraron mutados son los siguientes: *P53* (76-80%), *FAT1* (30-40%), *NOTCH1* (20-22%), *CASP8* (4-34%) y *PIK3CA* (6-20%). Cabe mencionar que la gran mayoría de la población estudiada en la base de datos del TCGA corresponde a población caucásica. Se han realizado otros estudios con secuenciación del genoma en poblaciones distintas; si bien las mutaciones previamente mencionadas se mantienen, se ha detectado diferencias importantes en porcentajes y la presencia de otros genes afectados (25-33).

En poblaciones asiáticas (Japón, Taiwán, Malasia y China) se encuentran tasas más bajas de mutación en *P53*: 35%–62% y más elevadas de *FAT1*. En las poblaciones de India se encontraron mutaciones en *USP9X*, *KMT2B*, *ARID2*, *UNC13C* y *TRPM3* (las cuales representan menos del 6% en la base del TCGA). En el estudio de población árabe se encontraron mutaciones en los genes *CLTCL1* (15%), *TRPM2* (10%) y *OSMR* (10%), amplificaciones de *FOSL1* (15%), *RELA* (10%), *GDF5* (10%), *MDM2* (5%) y *BAG1* (5%), todas estas mutaciones representando menos del 5% en el TCGA.

En este estudio destaca el hecho de que el 70% de las muestras tenía presente el VEB. En un estudio de Taiwán se encontró que el gen *P53* presentaba tasas más bajas de mutaciones; por el contrario, *CASP8* y *FAT1* presentaron las tasas más altas comparadas con el estudio de la TCGA (tabla 1) (25-33).

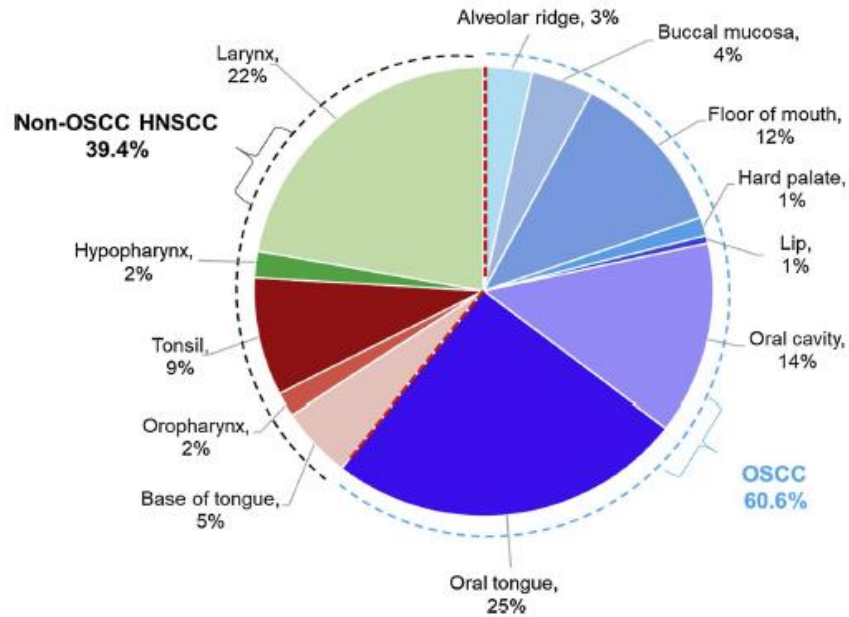


Figura 3. Porcentaje de afectación por localización anatómica de las muestras de cáncer de cabeza y cuello del TCGA

Tabla 1. Resumen de los principales estudios de secuenciación en cáncer bucal					
AUTOR	PACIENTES (Número)	POBLACIÓN	FACTOR DE RIESGO	SUBSITIOS	GENES MUTADOS
Agrawal et al. (2011)	25*	Caucásica	Tabaco 76%	No especificados	<i>TP53, NOTCH1, CDKN2A, PIK3CA, HRAS, FBXW7, PIK3AP1, RIMBP2, SI, NRXN2, NRXN3, RASA1, RXP3</i>
Stransky et al. (2011)	38*	Caucásica	Tabaco 95% y VPH 5%	No especificados	<i>TP53, NOTCH1, CDKN2A, FAT1, CASP8, HRAS, PIK3CA, SYNE1, MLL2, CSMD3, SI, PRDM</i>
Lui et al. (2013)	24	Caucásica	No especificado	No especificados	Solo se informan los genes en las vías <i>PI3K, MAPK</i> y <i>JAK / STAT</i> , pero no hay información detallada sobre si las mutaciones se encuentran en cáncer bucal

					u otras localizaciones en cabeza y cuello.
<b>Pickering et al. (2013)</b>	40	Caucásica	Tabaco 77%	Lengua 71.4%, piso de boca 19%, otros 9.5%	<i>TP53, FAT1, PIK3CA, CASP8, NOTCH1, HRAS, TP63, NOTCH2, BRAF</i>
<b>ICGC (2013)</b>	50	Indio	Tabaco 96%, alcohol 48%, nuez de areca 36%, VPH 26%	Gingivo-bucal 100%	<i>TP53, FAT1, CASP8, NOTCH1, HRAS, MLL2, USP9X, MLL4, UNC13C, ARID2, TRPM3, PCLO, FAT3, SMG1, SYNE2, EP300</i>
<b>TCGA (2015)</b>	172*	153 blancos, 11 negros, 2 asiáticos y 6 desconocidos	Tabaco 73% y VPH 7%	Lengua 44%, mucosa 28%, Piso 15%, encía 4%, paladar 3% y labio 1%	<i>TP53, FAT1, CDKN2A, NOTCH1, PIK3CA, CASP8, MLL2 (KMT2D), HRAS, NSD1, FBXW7, AJUBA, HLA, NFE2L2</i>
<b>Al-Hebshi et al. (2016)</b>	20	Árabe	Shammah 75%, tabaquismo 30% y VEB 70%	Lengua 45%, encía 20%, piso 15%, labio 10% y otros 10%	<i>CDKN2A, TP53, CASP8, CSMD3, PIK3CA, TP63, FAT1, HRAS, CRB1, NOTCH3, CLTCL1, OSMR, TRPM2</i>
<b>Su et al. (2017)</b>	120	Taiwán	Alcohol 89%, nuez de betel 79% y Tabaco 50%	Mucosa oral 40%, Lengua 26.7%, labio 10%, encía 9.2% y otros 14.1%	<i>TP53, FAT1, NOTCH1, CASP8, PIK3CA, CDKN2A, HRAS, EPHA2, ELAVL1, CHUK, RPTN, ASXL1</i>
<p>*Estudios que incluyeron otras localizaciones anatómicas en cabeza y cuello  VPH Virus del Papiloma Humano  VEB Virus del Epstein Bar  ICGA: Consorcio Internacional del Genoma del Cáncer  TCGA: Atlas del genoma del cáncer  Los genes resaltados en negrillas no están reportados en los estudios caucásicos</p>					

### 1.3.1 Funciones de los genes más comúnmente reportados

En la siguiente tabla se resumen las funciones de los diferentes genes que se han reportado como alterados en el cáncer bucal.



<b>Tabla 2. Funciones en genes mutados en cáncer bucal</b>	
<b>GEN</b>	<b>RESUMEN</b>
<b>ENCONTRADOS EN VARIAS POBLACIONES</b>	
<b>P53</b>	La proteína codificada responde a diversos estímulos celulares para regular la expresión de genes, induciendo paro del ciclo celular, apoptosis, senescencia, reparación del ADN o cambios en el metabolismo.
<b>FAT1</b>	El producto génico es un miembro de la superfamilia de caderinas, un grupo de proteínas integrales de membrana. Se expresa en niveles elevados en varios epitelios fetales. Su producto probablemente funciona como molécula de adhesión y/o receptor de señalización, por lo que participa en los procesos de desarrollo y la comunicación celular.
<b>NOTCH1 y 2</b>	Los miembros de esta familia de proteínas transmembrana de tipo I constan de múltiples repeticiones del factor de crecimiento epidérmico (EGF). La señalización de Notch es una vía de señalización intercelular que regula las interacciones entre células físicamente adyacentes mediante la unión de los receptores de la familia Notch a sus ligandos afines.
<b>CASP8</b>	Caspasa, que juega un papel central en la fase de ejecución de la apoptosis celular mediante Fas.
<b>PIK3CA</b>	La proteína codificada por este gen representa la subunidad catalítica, que usa ATP para fosforilar PtdIns, PtdIns4P y PtdIns. Se ha involucrado en el cáncer cervicouterino.
<b>CDKN2A</b>	Codifica una proteína que funciona como un estabilizador de la proteína supresora de tumores P53, ya que puede interactuar y secuestrar la ubiquitina-proteína ligasa MDM2 E3, una proteína responsable de la degradación de P53, controlando G1 del ciclo celular.
<b>FBXW7</b>	Este gen codifica proteínas que funcionan en la ubiquitinación dependiente de la fosforilación. Esta proteína se une directamente a la ciclina E y probablemente se dirige a la ciclina E para la degradación mediada por ubiquitina. Las mutaciones en este gen se detectan en líneas celulares de cáncer de mama y de ovario.
<b>HRAS</b>	Los productos codificados por estos genes funcionan en vías de transducción de señales. Estas proteínas pueden unirse a GTP y GDP y tienen actividad GTPasa intrínseca. Los defectos en este gen están implicados en una variedad de cánceres, incluido el cáncer de vejiga, el cáncer de tiroides folicular y el carcinoma bucal de células escamosas.
<b>PIK3AP1</b>	Función desconocida. Se expresa en bazo y ganglios linfáticos, principalmente.
<b>RIMBP2</b>	Función desconocida. Se expresa en cerebro y glándulas adrenales.

<b>NRXN2 Y 3</b>	Este gen codifica un miembro de la familia de genes de la neurexina. Los productos de estos genes funcionan como moléculas de adhesión celular y receptores en el sistema nervioso de los vertebrados.
<b>RASA1</b>	La proteína codificada por este gen se encuentra en el citoplasma y es parte de la familia GAP1 de proteínas activadoras de GTPasa. El producto génico estimula la actividad GTPasa de RAS p21 normal pero no su contraparte oncogénica. Actuando como un supresor de la función RAS, la proteína mejora la actividad GTPasa intrínseca débil de las proteínas RAS, dando como resultado la forma inactiva de RAS unida a GDP, lo que permite el control de la proliferación y diferenciación celular. Se asocia con carcinomas de células basales.
<b>RXFP3</b>	Función desconocida.
<b>SYNE1</b>	Este gen codifica una repetición de espectrina que contiene una proteína expresada en el músculo esquelético y liso, así como en los linfocitos de sangre periférica que se localiza en la membrana nuclear.
<b>KMT2D (MLL2, MLL4)</b>	La proteína codificada por este gen es una histona metiltransferasa que metila la posición Lys-4 de la histona H3. Es parte de un gran complejo de proteínas llamado ASCOM, que se ha demostrado que es un regulador transcripcional de los genes receptores de beta-globina y estrógeno.
<b>CSMD3</b>	Función desconocida. Se expresa en testículo y cerebro.
<b>SI</b>	Este gen codifica una enzima sacarasa-isomaltasa que se expresa en el borde en cepillo intestinal. Este complejo es esencial para la digestión de los carbohidratos de la dieta, incluidos el almidón, la sacarosa y la isomaltosa.
<b>PRDM</b>	La proteína codificada por este gen es un factor de transcripción que contiene dedos de zinc de tipo C2H2. También contiene un dominio regulador positivo, que se ha encontrado en varios otros factores de transcripción con dedos de zinc, incluidos los implicados en la diferenciación de células B y la supresión de tumores.
<b>TP63</b>	Este gen codifica un miembro de la familia de factores de transcripción P53. Los dominios funcionales de las proteínas de la familia P53 incluyen un dominio de transactivación N-terminal, un dominio de unión al ADN central y un dominio de oligomerización. El corte y empalme alternativo de este gen y el uso de promotores alternativos dan como resultado múltiples variantes de transcripción que codifican diferentes isoformas que varían en sus propiedades funcionales. Estas isoformas funcionan durante el desarrollo y mantenimiento de la piel, la regulación de las células stem, progenitoras adultas, el desarrollo del corazón y el envejecimiento prematuro.
<b>BRAF</b>	Este gen codifica una proteína que pertenece a la familia RAF de proteínas quinasas de serina/treonina. Esta proteína juega un papel en la regulación de la vía de señalización MAP quinasa/ERK, que afecta

	la división, diferenciación y secreción celular. Las mutaciones en este gen son causantes de cáncer identificadas con mayor frecuencia en el melanoma, linfoma no Hodgkin, cáncer colorrectal, carcinoma de tiroides, cáncer pulmonar de células no pequeñas y leucemia de células pilosas.
<b><i>NSD1</i></b>	Este gen codifica una proteína que contiene un dominio SET, 2 motivos LXXLL, 3 señales de translocación nuclear (NLS), 4 regiones de dedos de homeodominio vegetal (PHD) y una región rica en prolina. La proteína codificada potencia la transactivación del receptor de andrógenos (AR).
<b><i>AJUBA</i></b>	Función no conocida. Expresado en piel.
<b><i>HLA-A</i></b>	HLA-A pertenece a los parálogos de cadena pesada HLA clase I. Esta molécula de clase I es un heterodímero que consta de una cadena pesada y una cadena ligera (microglobulina beta-2). La cadena pesada está anclada en la membrana. Las moléculas de clase I juegan un papel central en el sistema inmunológico al presentar péptidos derivados del lumen del retículo endoplásmico para que puedan ser reconocidos por las células T citotóxicas. Se expresan en casi todas las células.
<b><i>NFE2L2</i></b>	Este gen codifica un factor de transcripción que es miembro de una pequeña familia de proteínas de cremallera de leucina básica (bZIP). El factor de transcripción codificado regula genes que contienen elementos de respuesta antioxidante (ARE) en sus promotores; muchos de estos genes codifican proteínas implicadas en respuesta a lesiones e inflamación que incluyen la producción de radicales libres.
<b>POBLACIÓN INDIA</b>	
<b><i>USP9X</i></b>	Este gen es miembro de la familia de la peptidasa C19 y codifica una proteína que es similar a las proteasas específicas de ubiquitina. Aunque este gen se encuentra en el cromosoma X, escapa a la inactivación de X. Las mutaciones en este gen se han asociado con el síndrome de Turner.
<b><i>KMT2B</i></b>	Este gen codifica una proteína que contiene múltiples dominios que incluyen un dedo de zinc CXXC, tres dedos de zinc PHD, dos dominios ricos en FY y un dominio SET. El dominio SET es un dominio C-terminal conservado que caracteriza a las proteínas de la familia MLL (leucemia de linaje mixto). Este gen se expresa de forma ubicua en tejidos adultos.
<b><i>ARID2</i></b>	Este gen codifica un miembro de la familia de proteínas de unión al DNA que contiene el dominio interactivo rico en AT (ARID). Tienen funciones en el patrón embrionario, la regulación de genes del linaje celular, el control del ciclo celular, la regulación transcripcional y la modificación de la estructura de la cromatina. Esta proteína funciona como una subunidad del factor asociado a polibromo y BRG1 o complejo de remodelación de cromatina PBAF (SWI/SNF-B) que facilita la activación transcripcional dependiente de ligando por receptores nucleares. Las mutaciones en este gen están asociadas con carcinomas hepatocelulares.

<b>UNC13C</b>	Función no conocida. Se ha encontrado expresado en cerebro y tiroides principalmente.
<b>TRPM3</b>	El producto de este gen pertenece a la familia de canales de potencial receptor transitorio (TRP). Los canales TRP son canales selectivos de cationes importantes para la señalización y homeostasis del calcio celular.
<b>PCLO</b>	La proteína codificada por este gen es parte de la matriz citoesquelética presináptica, que participa en el establecimiento de zonas sinápticas activas y en el tráfico de vesículas sinápticas.
<b>FAT3</b>	Función no descrita. Expresado en cerebro y próstata.
<b>SMG1</b>	Este gen codifica una proteína involucrada en la desintegración del RNAm (NMD) mediada sin sentido como parte del complejo de vigilancia del RNAm. La proteína tiene actividad quinasa y se cree que funciona en la NMD fosforilando el regulador de la proteína 1 transcripta sin sentido.
<b>SYNE2</b>	La proteína codificada por este gen es una proteína de la membrana externa nuclear que se une a la F-actina citoplasmática. Esta unión une el núcleo al citoesqueleto y ayuda a mantener la integridad estructural del núcleo.
<b>EP300</b>	Este gen codifica la proteína coactivadora transcripcional celular p300 asociada a adenovirus E1A. Funciona como histona acetiltransferasa que regula la transcripción a través de la remodelación de la cromatina y es importante en los procesos de proliferación y diferenciación celular. Interviene en la regulación del gen cAMP al unirse específicamente a la proteína CREB fosforilada. Este gen también se ha identificado como coactivador de HIF1A (factor alfa inducible por hipoxia 1) y, por tanto, desempeña un papel en la estimulación de genes inducidos por hipoxia como el VEGF. Los defectos en este gen pueden desempeñar un papel en el cáncer epitelial.
<b>POBLACIÓN ÁRABE</b>	
<b>CRB1</b>	Este gen codifica una proteína que es similar a la proteína de las migas de Drosophila y se localiza en el segmento interno de los fotorreceptores de mamíferos.
<b>NOTCH3</b>	Este gen codifica el tercer homólogo humano descubierto de la muesca de la proteína de membrana de Drosophilamelanogaster tipo I, pero quedan por determinar interacciones precisas entre estos ligandos y los homólogos de muesca humanos.
<b>CLTCL1</b>	Este gen es miembro de la familia de las cadenas pesadas de la clatrina y codifica una proteína principal de la capa poliédrica de las fosas y vesículas recubiertas.

<b>TRPM2</b>	La proteína codificada por este gen forma un canal catiónico tetramérico que es permeable al calcio, sodio y potasio y está regulado por ADP-ribosa intracelular libre. La proteína codificada se activa por estrés oxidativo y confiere susceptibilidad a la muerte celular.
<b>OSMR</b>	Este gen codifica un miembro de la familia de receptores de citocinas de tipo I. La proteína codificada se heterodimeriza con el transductor de señal de interleucina 6 para formar el receptor de oncostatina M de tipo II y con el receptor de interleucina 31 A para formar el receptor de interleucina 31, y así transduce los eventos de señalización inducidos por oncostatina M e interleucina 31. Las mutaciones en este gen se han asociado con amiloidosis cutánea localizada primaria familiar.
<b>FOSL1</b>	La familia de genes Fos consta de 4 miembros: FOS, FOSB, FOSL1 y FOSL2. Estos genes codifican proteínas de cremallera de leucina que pueden dimerizarse con proteínas de la familia JUN, formando así el complejo de factor de transcripción AP-1. Como tal, las proteínas FOS han sido implicadas como reguladoras de la proliferación, diferenciación y transformación celular.
<b>RELA</b>	NF-kappa-B es un factor de transcripción ubicuo involucrado en varios procesos biológicos. Se mantiene en el citoplasma en un estado inactivo por inhibidores específicos. Tras la degradación del inhibidor, NF-kappa-B se mueve hacia el núcleo y activa la transcripción de genes específicos. NF-kappa-B se compone de NFKB1 o NFKB2 unido a REL, RELA o RELB.
<b>GDF5</b>	Este gen codifica un ligando secretado de la superfamilia de proteínas TGF-beta (factor de crecimiento transformante beta). Los ligandos de esta familia se unen a varios receptores de TGF-beta que conducen al reclutamiento y activación de los factores de transcripción de la familia SMAD que regulan la expresión génica. Esta proteína regula el desarrollo de numerosos tipos de tejidos y células, incluidos cartílagos, articulaciones, grasa parda, dientes y el crecimiento de axones y dendritas neuronales.
<b>MDM2</b>	Este gen codifica una ubiquitina ligasa E3 localizada en el núcleo. La proteína codificada puede promover la formación de tumores dirigiéndose a proteínas supresoras de tumores, como p53, para la degradación proteasómica. Este gen en sí mismo está regulado transcripcionalmente por p53.
<b>BAG1</b>	El oncogén BCL2 es una proteína de membrana que bloquea un paso en una vía que conduce a la apoptosis o muerte celular programada. La proteína codificada por este gen se une a BCL2 y se conoce como antagonista asociado a BCL2. Mejora los efectos antiapoptóticos de BCL2 y representa un vínculo entre los receptores del factor de crecimiento y los mecanismos antiapoptóticos.
<b>POBLACIÓN DE TAIWÁN</b>	
<b>EPHA2</b>	Este gen pertenece a la subfamilia del receptor de efrina de la familia de las proteínas tirosina quinasa. Se han implicado en la mediación de eventos del desarrollo, particularmente en el sistema nervioso.

<b>ELAVL1</b>	La proteína codificada por este gen es un miembro de la familia ELAVL de proteínas que se unen selectivamente a elementos ricos en AU (ARE). Al unirse, desempeñan un papel en la estabilización de los RNAm que contienen ARE. Se expresa, en gran medida, en muchos cánceres y podría ser potencialmente útil en el diagnóstico, pronóstico y terapia del cáncer.
<b>CHUK</b>	Este gen codifica un miembro de la familia de las proteínas quinasas de serina/treonina. La proteína codificada, un componente de un complejo de proteínas activadas por citocinas que es un inhibidor del factor de transcripción esencial complejo NF-kappa-B, fosforila los sitios que desencadenan la degradación del inhibidor a través de la vía de ubiquitinación, activando así el factor de transcripción.
<b>RPTN</b>	Función no descrita. Expresado en la piel.
<b>ASXL1</b>	Este gen codifica una proteína de unión a cromatina necesaria para la determinación normal de la identidad del segmento en el embrión en desarrollo. La proteína es un miembro del grupo de proteínas Polycomb, que son necesarias para el mantenimiento de la represión estable de los loci homeóticos y otros. Se cree que la proteína interrumpe la cromatina en áreas localizadas, mejorando la transcripción de ciertos genes mientras reprime la transcripción de otros genes. La proteína codificada por este gen funciona como un coactivador dependiente de ligando para el receptor de ácido retinoico en cooperación con el coactivador del receptor nuclear 1.
<a href="https://www.ncbi.nlm.nih.gov/gene">https://www.ncbi.nlm.nih.gov/gene</a> (34)	

### 1.3.2 Firmas mutacionales asociadas

Todo proceso mutacional genera componentes de daño y modificaciones al DNA, DNA de reparación y de replicación. Esto, a su vez, produce una firma mutacional característica que incluye sustituciones simples de bases, sustituciones dobles o inserciones y deleciones. Se han desarrollado modelos matemáticos para generar los diferentes tipos de firmas, y se han estudiado 84, 729,690 mutaciones somáticas a través de 4,645 lecturas del genoma completo y de 19,184 lecturas del exoma de los tipos más comunes de cáncer (35).

De las firmas que produjo el modelo matemático, se encontraron 49 de sustitución simple (SBS), 11 de doble sustitución (DBS) y 17 de inserción-delección (Indel) en los diferentes tipos de cáncer. Este proyecto de investigación se enfocó en la descripción de las firmas de sustituciones simples, ya que han sido las firmas más estudiadas y reportadas y las que cuentan con mayor evidencia científica sobre sus implicaciones (35). En el área de cabeza y cuello, las sustituciones simples que se han reportado (de mayor a menor asociación) son las siguientes: SBS 1, relacionada con desaminación de 5-metilcitosina y con el envejecimiento; SBS 2, con actividad en la vía APOBEC asociada con inflamación; SBS 5, relacionada con el envejecimiento y el tabaquismo; y SBS 13, con

actividad APOBEC e inflamación. Otras formas encontradas, pero de menor relevancia en el cáncer bucal son las siguientes: SBS 3, con alteraciones en la reparación de DNA, BRCA 1 y 2; SBS 4, relacionada a consumo tabaco; SBS 7, relacionada a exposición a luz UV; SBS 18, que está relacionada con daño por especies reactivas de oxígeno; y SBS 16, 17, 33 y 40, cuyas funciones no han sido identificadas. Específicamente en cavidad oral, únicamente están reportadas la firma SBS 4 y 5 (35) (figura 4).

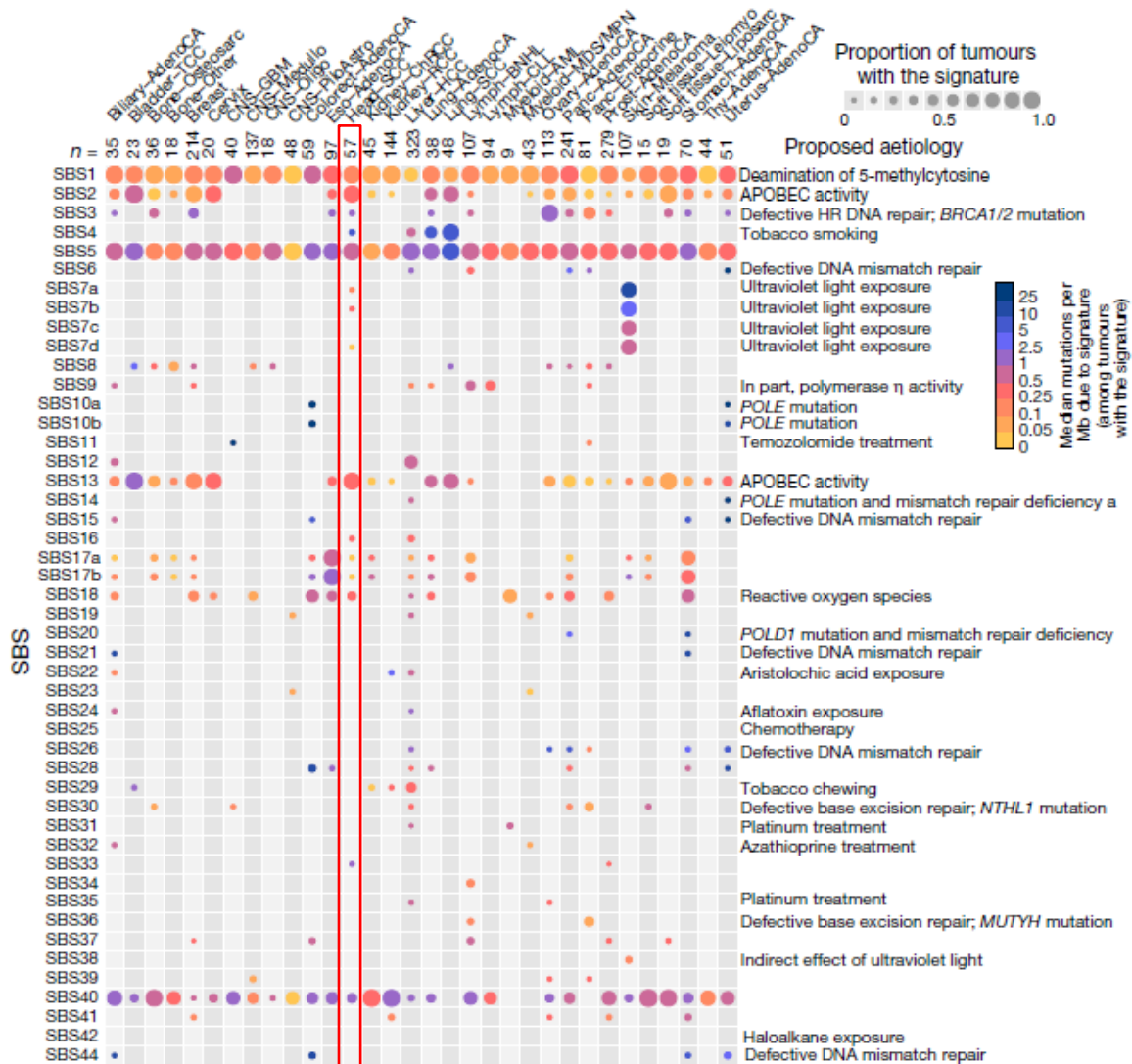


Figura 4. Firmas mutacionales de sustituciones simples descritas en diferentes tipos de cáncer. En el cuadro rojo se señalan las firmas encontradas en cáncer de cabeza y cuello

#### 1.4 Presentación clínica del cáncer bucal

La cavidad bucal se divide en subsitios anatómicos para su estudio oncológico, los cuales, de acuerdo con la AJCC (American Joint Committee on Cancer), son los siguientes: lengua móvil

(30 a 40%), piso de la boca (20%), labio (20 a 30%), encía (10%), mucosa oral (<10%), trigono retromolar (5%) y paladar duro (<5%) (36, 37).

La forma en que se presenta este cáncer es con úlceras, que progresan a tumores en los distintos subtipos anatómicos, las cuales se acompañan de dolor y sangrado. En los tumores avanzados se agrega trismus, inmovilidad de la musculatura lingual, afección de nervios, piel y adenopatías en cuello. Para su estudio es elemental la exploración física del área de cabeza y cuello. Se requiere de biopsia y estudios de imagen como tomografía axial computada y/o resonancia magnética, para la valoración del cuello se puede utilizar el ultrasonido (1).

Una vez que se efectúa la exploración y los estudios complementarios, se requiere estadificar al paciente, para lo cual se toma en cuenta el tamaño tumoral con la profundidad de invasión, la presencia de afección metastásica en ganglios y de metástasis a distancia. Los tumores tempranos son todos aquellos que están limitados al sitio anatómico con un tamaño menor de 4cm y con una profundidad de invasión menor a 10mm en ausencia de ganglios linfáticos afectados. La etapa localmente avanzada se caracteriza por tumores mayores de 4cm o con profundidad mayor de 10mm, por los que afectan tejidos anatómicos adyacentes o por la presencia metastásica de ganglios linfáticos. La etapa más avanzada se diagnostica cuando hay evidencia de enfermedad en órganos distantes; el más afectado es el pulmón (36, 37).

En Estados Unidos de América, el 29% de los pacientes se diagnostica en etapas tempranas; el 48%, en etapas localmente avanzadas; y el 19%, en etapa metastásica. Los datos en México señalan que cerca del 70% se encuentra en etapas localmente avanzadas al momento del diagnóstico; únicamente el 25%, en etapas tempranas; y el 4%, en etapa metastásica (36-38).

## **1.5 Tratamiento**

Para el tratamiento del cáncer es esencial la estadificación del paciente y la valoración de un equipo multidisciplinario de especialistas en oncología para asegurar el mejor resultado posible (cirujano oncólogo, oncólogo médico, radiooncólogo, odontólogo, rehabilitador del habla y deglución, nutrición, patología, radiología, etc.). Para el inicio del tratamiento se deben tomar en cuenta múltiples factores como las características del tumor y condiciones propias del paciente como comorbilidades y edad (1).

La resección quirúrgica es el tratamiento de elección para el cáncer bucal. Permite una estadificación patológica y proporciona información sobre el estado de los márgenes, la diseminación del tumor y las características histopatológicas. Esto dará la pauta para determinar si el paciente requiere un manejo posterior con radioterapia o quimio-radioterapia concomitante (1).

De modo general, se puede decir que las etapas tempranas se pueden tratar con cirugía únicamente; en algunas ocasiones, se agrega radioterapia. En las etapas localmente avanzadas, el manejo es trimodal (cirugía, quimioterapia y radioterapia); y en la etapa metastásica, el tratamiento es con manejo sistémico como quimioterapia, terapia blanca (cetuximab) o inmunoterapia (1).



## 1.6 Pronóstico

El factor más importante para el pronóstico es la presencia de metástasis en los ganglios linfáticos, seguido de la etapa clínica. Otros factores importantes son la profundidad de invasión del tumor primario, los márgenes positivos de resección quirúrgica, la invasión perineural y la extensión extracapsular nodal (1). En cuanto la presencia de mutaciones, el único estudio que reporta supervivencia global, es la base de datos de TCGA en donde se encontraron los siguientes genes con impacto: *FAT1* ( $p = 0.00764$ ), *NOTCH1* ( $p= 0.0380$ ) y *CASP8* ( $p= 0.0378$ ) (30).

La supervivencia de manera general es 66.2% a 5 años, en la etapa temprana es de 85.1%; en la etapa localmente avanzada, de 66.8%; y en la metastásica, de 40.1%. En la serie del Memorial Sloan-Kettering Cancer Center (1960–2005), que representa la serie con mejor pronóstico publicado al momento, la supervivencia global a 5 años es para la etapa I, de 78.5%; para la etapa II, de 68.4%; para la etapa III, de 64.5%; y para la etapa IV, de 34.5% (1, 38).

## 2. PREGUNTA DE INVESTIGACIÓN

¿Las mutaciones encontradas en el cáncer bucal de células escamosas en una muestra de pacientes de población mexicana serán diferentes a las muestras de pacientes reportadas en distintas poblaciones y se podrán asociar con el pronóstico de los pacientes?

## 3. PLANTEAMIENTO DEL PROBLEMA

El cáncer bucal de células escamosas en México representa el 2% de las neoplasias malignas. Es de gran importancia clínica debido a que el 70% de los pacientes se diagnostica en etapas localmente avanzadas, lo cual dificulta el tratamiento y reduce la supervivencia a 5 años hasta en un 50%. Por ello es necesario caracterizar e identificar nuevos marcadores moleculares que optimicen el diagnóstico temprano, pronóstico y tratamiento de esta enfermedad.

## 4. JUSTIFICACIÓN

La carcinogénesis bucal se caracteriza por la presencia de diversos eventos a nivel genético. Sin embargo, las alteraciones moleculares del cáncer bucal de células escamosas no han sido totalmente identificadas en pacientes de la población mexicana y la información que existe hasta el momento es limitada y escasa, por lo que será de gran utilidad realizar estudios de secuenciación masiva en las diferentes etapas clínicas del cáncer bucal, que permitan definir marcas moleculares específicas y que, al conjuntarse con el análisis de variables clínicas, puedan ser utilizadas para el diagnóstico, tratamiento y/o pronóstico de este tipo de cáncer. Además, este estudio hará posible

identificar nuevas alteraciones en genes que participan en el carcinoma bucal, contribuyendo al conocimiento de esta enfermedad en pacientes de la población mexicana.

## 5. HIPÓTESIS

Se identificaran mutaciones específicas para la población mexicana en el carcinoma bucal de células escamosas, las cuales se relacionan con el pronóstico de los pacientes.

## 6. OBJETIVOS

**General:** Caracterizar a nivel molecular el carcinoma bucal de células escamosas mediante secuenciación masiva en una muestra pacientes mexicanos y asociarlas a pronóstico.

**Particulares:**

- a) Recolectar muestras de pacientes con carcinoma bucal de células escamosas (tejido tumoral, tejido sano adyacente y sangre) y realizar la extracción del DNA.
- b) Elaborar base de datos clínico-patológicos de los pacientes con cáncer bucal.
- c) Realizar la secuenciación de exoma, identificar cambios moleculares de un nucleótido, inserciones y deleciones, y determinar las firmas mutacionales.
- d) Identificar las vías de señalización afectadas en relación con los cambios moleculares encontrados.
- e) Determinar la infección por VPH y VEB.
- f) Asociar las "huellas moleculares" e infección por patógenos con las características clínico-patológicas.

## 7. METODOLOGÍA

### 7.1 Diseño del estudio

El presente proyecto de maestría se realizó como un estudio piloto de dos años, el cual forma parte del estudio aprobado por el comité de ética e investigación del Instituto Nacional de Cancerología con el folio 018/086/IBI CEI/1316/18, que se desarrolla en el Servicio de Cabeza y Cuello del Instituto Nacional de Cancerología, planificado para un período de cuatro años.

**Tipo de estudio:** Estudio prospectivo, longitudinal, observacional, descriptivo y analítico.

**Tipo de población:** Se incluyeron, de forma consecutiva, pacientes con diagnóstico de cáncer bucal de células escamosas.

### 7.2 Criterios de inclusión

- a) Individuos mayores de 18 años con cáncer bucal
- b) Diagnóstico histológico de cáncer bucal de células escamosas

- c) Pacientes que no hayan recibido tratamiento oncológico previo de cualquier tipo (cirugía, quimioterapia y/o radioterapia)
- d) Que acepten participar en el estudio mediante la firma de carta consentimiento.

### 7.3 Criterios de exclusión

- a) Pacientes con muestra insuficiente del tumor o del tejido sano (menor 0.5cm)
- b) Pacientes que hayan perdido seguimiento

### 7.4 Criterio de eliminación

- a) Pacientes con muestra de DNA con un DIN menor a 6

A cada paciente se le realizó una ficha de recolección de datos (características demográficas, clínicas y antecedentes de consumo de tabaco y alcohol). Posteriormente, se obtuvo mediante una biopsia de tejido proveniente del tumor y de tejido sano y se correlacionó con indicadores del comportamiento biológico de la enfermedad posteriores al inicio del tratamiento: supervivencia libre de enfermedad y supervivencia global.

### 7.5 Tamaño de la muestra

Para el estudio piloto se contemplaron 20 pacientes *por factibilidad económica y conveniencia*, por lo que no se calculó el tamaño de muestra con métodos estadísticos.

### 7.6 Variables analizadas

- a) Dependientes:** supervivencia global y supervivencia libre de enfermedad.
- b) Independientes:** edad, género, nivel socioeconómico, subsitio, estadificación del paciente (AJCC clínica y/o patológica), profundidad de invasión, permeación linfovascular, infiltración perineural, márgenes quirúrgicos, tratamiento, grado, dispersión tumoral, desnutrición, escala ECOG, metástasis regionales o a distancia, consumo de tabaco y alcohol, mutaciones identificadas, firmas mutacionales, carga tumoral y presencia de VPH y VEB.

### 7.7 Criterios operacionales

- a) Supervivencia global:** porcentaje de pacientes que están vivos en un periodo determinado de tiempo a partir del diagnóstico.
- b) Supervivencia libre de enfermedad:** porcentaje de pacientes que están libres de cáncer en un periodo determinado de tiempo a partir del término del tratamiento.

**c) Recurrencia:** cuando hay evidencia clínica, radiológica o patológica de cáncer, una vez que transcurrió un tiempo sin enfermedad.

**d) Persistencia:** presencia de tumor por clínica, imagen o patología después de un tratamiento oncológico con intento curativo.

**e) Progresión:** la enfermedad crece más de un 20% de las lesiones que se miden o hay presencia de nuevas lesiones que no estaban en estudios de imagen previos.

**f) Consumo de tabaco:** para el análisis del consumo de tabaco, los pacientes se clasificaron en tres categorías: 1) No fumadores (aquellos que nunca han fumado regularmente), 2) Fumadores (individuos que han fumado por lo menos 1 cigarrillo por día durante el último año) y 3) Exfumadores (individuos que, por lo menos, hayan fumado un cigarrillo al día durante un año y que no han fumado en el último año).

Para los pacientes fumadores y exfumadores, se registró la edad de inicio del hábito y se calculó el tiempo durante el cual han mantenido el hábito (meses). En los exfumadores se registró el número de años que fumaron y el número de años que han estado libres del hábito.

Posteriormente, se calculó el número de paquetes (20 cigarrillos por paquete) consumidos por año (número de cigarrillos por día x 365/20 x años de consumo).

**g) Consumo de alcohol:** la cantidad de alcohol que se ingiere por semana se cuantificó en unidades de bebida (UB), de acuerdo con el tipo de bebida que es consumida y tomando en cuenta el contenido de etanol de cada una de ellas (tabla 3).

Tabla 3. Mililitros considerados como unidad de bebida de acuerdo con el tipo de bebida			
Tipo de bebida	Mililitros	Tipo de bebida	Mililitros
vino	125 ml	tequila	30 ml
ron	30 ml	mezcal	30 ml
licor	30 ml	cerveza	330 ml
brandy	30 ml	pulque	330 ml

Se clasificó a los pacientes en tres categorías: 1) No bebedores (aquellos que no toman alcohol o que toman menos de una copa por semana), 2) Bebedores (individuos que han ingerido más de una UB a la semana) y 3) Exbebedores (individuos que no han ingerido más de una UB por semana durante el año previo al examen bucal). El grupo de exbebedores que refieran más de 10 años sin el hábito, se reclasificó como no bebedores.

La Unidad de Bebida estándar (UB) en México se considera de 13 gramos y se calcula multiplicando los ml de bebida por el porcentaje de alcohol/100. Para transformar los mililitros en gramos, se multiplica por 0.79 (1 UB = 13 gramos). Se clasificará como moderado en mujer y hombre menores de 65 años, menos de 3 UB al día; en mayores de 65 años, menos de 2 UB al día. El consumo excesivo en hombre, más de 14 UB a la semana o 4 UB de una ocasión; y en mujeres,

más de 7 UB a la semana o más de 3 UB de una ocasión. Todos los demás consumos se clasificaron como consumo leve.

**h) ECOG:** Eastern Cooperative Oncology Group. Esta escala sirve para valorar el estado funcional del paciente que guarda relación con el pronóstico (tabla 4).

<b>Tabla 4. ECOG: Eastern Cooperative Oncology Group</b>	
<b>ECOG 0</b>	El paciente se encuentra totalmente asintomático y es capaz de realizar un trabajo y actividades normales de la vida diaria.
<b>ECOG 1</b>	El paciente presenta síntomas que le impiden realizar trabajos arduos, aunque se desempeña normalmente en sus actividades cotidianas y en trabajos ligeros. Solo permanece en la cama durante las horas de sueño nocturno.
<b>ECOG 2</b>	El paciente no es capaz de desempeñar ningún trabajo. Se encuentra con síntomas que le obligan a permanecer en la cama durante varias horas al día, además de las de la noche, pero que no superan el 50% del día. El individuo satisface la mayoría de sus necesidades personales solo.
<b>ECOG 3</b>	El paciente necesita estar encamado más de la mitad del día por la presencia de síntomas. Necesita ayuda para la mayoría de las actividades de la vida diaria como, por ejemplo, el vestirse.
<b>ECOG 4</b>	El paciente permanece encamado el 100% del día y necesita ayuda para todas las actividades de la vida diaria como, por ejemplo, la higiene corporal, la movilización en la cama e incluso la alimentación.
<b>ECOG 5</b>	Fallecido.

**i) Dispersión tumoral WPOI-5:** Cuando en la pieza de patología se detectan islas tumorales dispersas en forma de hebras, a más de 1mm de distancia del tumor primario.

## 7.8 Procedimientos clínicos

A los pacientes con cáncer bucal de células escamosas que acudieron por primera vez a la consulta externa de Servicio de Cabeza y Cuello del INCan, se les realizó un examen bucal y un interrogatorio para obtener una ficha de identificación con los datos socioeconómicos y las características clínicas de la lesión. Se solicitaron estudios de imagen y se realizó una toma de biopsia de la lesión, datos con los que se estadificó de acuerdo con criterios establecidos por la AJCC y se registró la historia de consumo de tabaco y/o alcohol, así como las demás variables independientes.

En el presente proyecto se obtuvieron las siguientes muestras: tejido proveniente del tumor y tejido sano adyacente (0.5 cm aproximadamente), así como una muestra de sangre periférica de los pacientes (4 ml aproximadamente por paciente la cual se almaceno). Las muestras obtenidas se transportaron al laboratorio de carcinogénesis y medio ambiente del INCan, en donde se realizó la

extracción de DNA y preparación para secuenciación, además de la determinación de patógenos. El procesamiento y diagnóstico histológico de las muestras obtenidas se realizó en el laboratorio de histopatología del INCan.

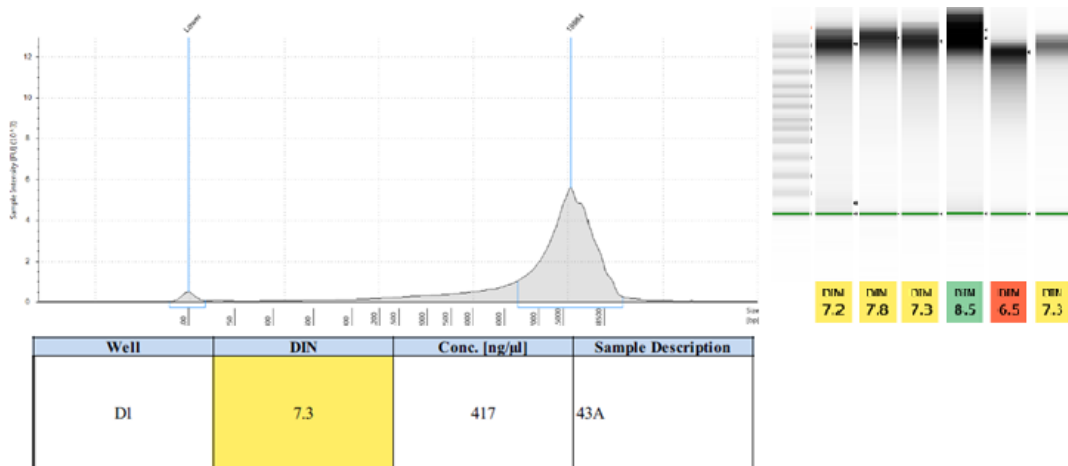
## 7.9 Procedimientos en laboratorio

### 7.9.1 Extracción de DNA

La extracción de DNA se llevó a cabo utilizando el kit *Promega wizard genomic DNA purification kit*.

**Procesamiento de muestras:** En una balanza analítica se pesaron 20 mg de cada uno de los tejidos, con una hoja de bisturí cortando el tejido en pequeños fragmentos, los cuales se suspendieron en 300  $\mu$ l de solución lítica celular y se sometieron a disgregación en un homogenizador de tejidos en un tubo para microcentrifuga de 1.5 ml. Se agregaron 20  $\mu$ l de proteinasa k (20 mg/ml) y se incubaron a 65° C en un termoblock durante 1 hora. Se dejó reposar la muestra a temperatura ambiente y se agregaron 2  $\mu$ l de solución RNAsa. Se incubó a 37° C durante 30 minutos. Se dejó enfriar a temperatura ambiente y se agregaron 200  $\mu$ l de solución precipitadora de proteínas. Se agitó en el vórtex y se colocó en hielo durante 5 minutos. Se centrifugó 5 minutos a 13 000 rpm. Se removió el sobrenadante, se transfirió a un tubo para microcentrifuga de 1.5 ml con 300  $\mu$ l de isopropanol y se mezcló suavemente. Se centrifugó por 5 minutos a 13 000 rpm y se retiró el sobrenadante. Se agregaron 300  $\mu$ l de etanol al 70% y se agitó varias veces el tubo para lavar el DNA. Se centrifugó por 5 minutos y se aspiró el etanol usando una pipeta de secuencia. Se colocaron los tubos destapados sobre papel absorbente limpio durante 15 minutos. Se agregaron 100  $\mu$ l de agua libre de DNAsas y RNAsas y se incubó a 65° C por una hora. Finalmente, se llevó a cabo la cuantificación (ng/ $\mu$ l) del DNA utilizando el nanodrop, se obtuvo la concentración de 50 ng/microL y se realizó la medición del DIN (DNA Integrity number) por medio de electroforesis capilar (figura 5). Solo las muestras con un DIN superior a 6 fueron utilizadas para secuenciación. El DNA obtenido se almacenó a -20° C para realizar los estudios moleculares posteriores.

A)



B)

Tumor		Tejido adyacente		Tumor		Tejido adyacente	
Muestra	DIN	Muestra	DIN	Muestra	DIN	Muestra	DIN
CSTE_1	7.9	CSTE_1	7.6	CSTE_11	7.4	CSTE_11	7.9
CSTE_2	8.9	CSTE_2	8.5	CSTE_12	7.5	CSTE_12	6.8
CSTE_3	9	CSTE_3	7.4	CSTE_13	6	CSTE_13	7.2
CSTE_4	7.4	CSTE_4	8.9	CSTE_14	7.8	CSTE_14	7.3
CSTE_5	7	CSTE_5	6.7	CSTE_15	8.5	CSTE_15	6.5
CSTE_6	7.1	CSTE_6	6.1	CSTE_16	7.3	CSTE_16	6
CSTE_7	7	CSTE_7	7.2	CSTE_17	6.8	CSTE_17	6.3
CSTE_8	7.2	CSTE_8	8.1	CSTE_18	7.3	CSTE_18	7.2
CSTE_9	7.5	CSTE_9	7.2	CSTE_19	6.8	CSTE_19	6.2
CSTE_10	6.2	CSTE_10	7.1	CSTE_20	6.3	CSTE_20	6.9

Figura 5. A) Figura representativa del electroferograma de DNA obtenido de las muestras de cáncer bucal. B) Se muestra el DIN obtenido de las muestras tanto de tejido tumoral como adyacente.

## Preparación de bibliotecas y secuenciación del exoma

El DNA genómico se fragmentó mecánicamente y se prepararon las bibliotecas mediante la adición de código de barras de secuenciación indexados con Sure-Select XT Target Enrichment System para Illumina Paired-End Multiplexed Sequencing Library (Agilent Technologies, Santa Clara, CA, EE. UU.). Las librerías precapturadas se enriquecieron con el kit Sure-Select Human All Exon V6 (Agilent Technologies) para regiones codificantes y sitios de unión de 20,000 genes humanos, siguiendo las instrucciones del fabricante. Posteriormente, se capturaron las bibliotecas y secuenciaron en el instrumento NovaSeq 6000 (Illumina).

### 7.9.2 Análisis bioinformático de los datos de secuenciación

El análisis bioinformático se realizó en dos partes: 1) el preprocesamiento y 2) la identificación y análisis de variantes (anotación y filtrado). Para implementar el flujo de trabajo, se consideró que el enriquecimiento de la región codificante se realizó mediante el método de captura de híbridos con el kit SureSelect All Exon V6 (Agilent Technologies, Santa Clara C.A., USA). También la plataforma de secuenciación, Illumina NovaSeq 6000.

#### 7.9.2.1 Preprocesamiento

Se obtuvo la calidad de las secuencias crudas de los archivos fastq con el programa FastQc. Posteriormente, con MultiQc se generó un reporte único de todas las muestras. Las secuencias fueron alineadas con bwa-mem, utilizando como referencia el genoma hg 19. Las siguientes herramientas de GATK fueron utilizadas para el resto del preprocesamiento: 'SortSam' para ordenar las secuencias alineadas del primer cromosoma hasta los cromosomas sexuales; 'Sam Format

*Converter* para convertir las secuencias en archivos binarios; *Mark Dupliucates* identificó los duplicados de PCR que actúan como artefactos en la identificación de variantes; *Base Recalibrator* y *Apply BQSR* para recalibrar el score de la calidad de las bases secuenciadas; y, por último, *Mutect2* para el llamado de variantes (39-42).

### 7.8.2.2 Identificación y análisis de variantes

#### a. Identificación de variantes

Los archivos derivados del llamado de variantes fueron anotados utilizando wAnnotar (<http://wannovar.wglab.org/>). Con los archivos resultantes se consolidó un documento único del que se extrajeron las variantes patogénicas y las variantes de significado clínico incierto (VUS). Para la identificación de dichas variantes se considera una frecuencia alélica (AF) >0.05 y que la lectura mutada estuviera presente por lo menos dos veces en cada sentido de las lecturas. Finalmente, se filtraron los genes conductores utilizando la lista propuesta por Intogen (43, 44) (tabla 5).

Tabla 5. Criterios utilizados para el filtrado de variantes	
Variantes patogénicas	VUS
Localizadas en la región exónica, 5'UTR, 3'UTR, ncRNA o de splicing	
Clasificadas patogénicas por ClinVar (45)	Frecuencia <0.001 gnomAD ( <a href="https://gnomad.broadinstitute.org/">https://gnomad.broadinstitute.org/</a> ) (46) y ESP6500 <a href="https://evs.gs.washington.edu/EVS/">https://evs.gs.washington.edu/EVS/</a> (47).
	Predicción de alelo deletéreo en por los menos dos de los tres predictores: SIFT, PolyPhen-1 y Mutational Taster (48-50).

#### b. Análisis de variantes

##### Carga mutacional tumoral

La carga mutacional tumoral (TMB) se definió como el número de mutaciones somáticas, codificantes, de sustitución de bases y deleciones presentes por megabase en el genoma examinado. Las variantes germinales conocidas reportadas en dbSNP (151) fueron descartadas. Para calcular la TMB por megabase, el número total de mutaciones contada se dividió entre el tamaño de la región de codificación del territorio objetivo. A su vez, se dividieron las muestras en dos grupos: carga tumoral alta y baja, tomando un punto de corte de 10.3%, respecto al puntaje más alto; esto, basado en la metodología del Memorial Sloan Kettering Cancer Center (51).

##### Firmas mutacionales

Se realizó el análisis de las firmas mutacionales V2 y V3. Para ambas versiones las variantes germinales conocidas y reportadas en dbSNP (151) fueron descartadas. Para las firmas V2 se utilizó R (versión 3.5.1 Feather spray) con el paquete deconstructSigs desarrollado por Rosenthal. El análisis de firmas mutacionales V3 se realizó con "SigProfiler Bioinformatic Tool" de Cosmic (52, 53) y se agruparon de acuerdo con la similitud de firmas mutacionales (agrupamiento no supervisado).



## Identificación de mutaciones conductoras

Se identificaron las mutaciones conductoras utilizando el programa en línea *Cancer genome interpreter*.

## Análisis de enriquecimiento

Para el análisis de enriquecimiento de las diversas vías de señalización, se utilizaron los genes conductores filtrados, obtenidos del paso de identificación de variantes con la herramienta de anotación PANTHER (Protein Analysis Through Evolutionary Relationships) Version 16.0 (54). A la par, se obtuvieron las funciones moleculares y los procesos biológicos alterados.

## 7.9 Determinación de VPH y VEB a través de PCR

La presencia de VPH o de VEB se determinó mediante reacción en cadena de la polimerasa en tiempo real (qRT-PCR), utilizando el DNA de las muestras a una concentración de 100ng. Las reacciones de PCR se llevaron a cabo en un volumen final de 20  $\mu$ l utilizando el kit RealQ Plus 2x Master Mix Green, High ROX (Ampliqon). En primer lugar, se utilizaron los oligonucleótidos universales Gp5 y Gp6 para determinar positividad a VPH, debido a que detecta diferentes tipos de VPH (6, 11, 16, 18, 31, 33 y 325). Las muestras que resultaron positivas se sometieron a determinación de VPH16 y VPH18. La determinación de VEB se realizó con un solo juego de oligos. Como control positivo para VPH se utilizaron 20 ng de DNA de las líneas celulares SIHA (VPH 16) y HELA (VPH 18) y para VEB una muestra de cáncer gástrico proveniente del archivo de patología previamente determinada como positiva para este virus. Para verificar la integridad del DNA obtenido, se llevó a cabo la amplificación de un fragmento del gen  $\beta$ -globina, caracterizado por ser un gen altamente conservado en las células humanas. Los oligonucleóticos utilizados se encuentran en la tabla 6.

Tabla 6. Secuencias utilizadas para la determinación de VPH y VEB en las muestras de cáncer bucal		
Gen	Secuencia	Temperatura
Gp5-Forward Gp6-Reverse	5'-TTTGTTACTGTGGTAGATACTA-3' 5'-GAAAAATAAACTGTAAATCATATTC-3'	50°C
VPH-16-Forward VPH-16-Reverse	5'-CCGGACAGAGCCCATTACAAT-3' 5'-ACGTGTGTGCTTTGTACGCAC-3'	60°C
VPH-18-Forward VPH-18-Reverse	5'-GACTCAGAGGAAGAAAACGATGAAA-3' 5'-GTGACGTTGTGGTTCCGGCT-3'	60°C
EBV-Forward EBV-Reverse	5'-CGGAAGCCCTCTGGACTTC-3' 5'-CCCTGTTTATCCGATGGAATG-3'	60°C
$\beta$ -Globina-Forward $\beta$ -Globina-Reverse	5'-GACAGGTACGGCTGTCATCA-3' 5'-TAGATGGCTCTGCCCTGACT-3'	60°C

## 7.10 Análisis estadístico

Para describir las características de los casos incluidos, se analizaron las variables mediante medidas de tendencia central y de dispersión. Las mutaciones identificadas, las firmas mutacionales, los factores de riesgo y la carga tumoral se asociaron con la supervivencia mediante análisis de Kaplan Meier y regresión de Cox. Las pruebas estadísticas se realizaron para un nivel de error alfa de 0.05. El análisis estadístico se realizó con el paquete SPSS v27 2021.

## **8. RESULTADOS**

### **8.1 Análisis descriptivo**

#### **8.1.1 Datos demográficos**

Se incluyeron en este estudio 20 pacientes con el diagnóstico de cáncer bucal de células escamosas en un periodo de enero del 2019 a febrero del 2021 con una media de 13.4 meses (1.6–25.5 meses), de los cuales 12 son masculinos (60%) y 8 femeninos (40%), con una media de edad de 54 años (21–86 años). En relación con el grado escolar, 9 tienen primaria concluida (45%); 5, licenciatura (25%); 3, bachillerato y/o carrera técnica (15%); 2 pacientes son analfabetas (10%) y un paciente tiene la secundaria concluida (5%). En cuanto al estado nutricional, 7 están en desnutrición (IMC menor de 18) (35%) y 13 en adecuado índice de masa corporal (65%). El estado funcional, de acuerdo con el ECOG, fue en su gran mayoría ECOG 1 en 19 pacientes (95%) y un paciente con ECOG de 3 (5%). Todas las características demográficas y clínico-patológicas de los pacientes se resumen en la tabla 7.

#### **8.1.2 Datos de estadificación**

La afectación por subsitio de acuerdo con la AJCC fue el siguiente: lengua móvil, 9 (45%); encía superior, 5 (25%); encía inferior, 4 (20%); piso de la boca, uno (5%); y trígono retromolar, uno (5%). No se incluyeron pacientes con afección en mucosa oral y paladar duro. Referente a la etapa clínica, un paciente fue diagnosticado con cáncer *in situ* (5%); uno, en etapa I (5%); uno, en etapa II (5%), 3 en etapa III (15%), 9, en etapa IVa (45%); y 5, en etapa IVb (25%). No se presentaron pacientes en etapa metastásica (etapa clínica IVc). De acuerdo con la etapa clínica categorizada, se encontró que solo 3 pacientes fueron diagnosticados en etapa temprana 3 (15%) y la mayoría de los casos (85%) fueron diagnosticados en etapa avanzada. En cuanto a la afección metastásica ganglionar 11, (55%) fueron clasificados como N0 y 9 (45%) con N (+).

#### **8.1.3 Datos histopatológicos**

Se encontró una media de profundidad de invasión (DOI) de 7.3 mm (3 – 130mm). En 4 pacientes el DOI no fue valorable (20%), en 7 pacientes se encontró permeación linfovascular presente (35%), estuvo ausente en 10 (50%) y en 3 pacientes no fue valorable (15%), en 5 pacientes

se presentó infiltración perineural (25%), en 12 pacientes, ausente (60%) y no valorable en 3 pacientes (15%). En cuanto al grado histológico, 15 pacientes presentaron un grado moderado (75%); 3, bien diferenciado (15%); y 2, pobremente diferenciado (10%). Los márgenes quirúrgicos en los pacientes que se llevaron a cirugía fueron negativos en 5 pacientes (41.6%), positivos en 4 (33.3%) y cercanos en 3 (25%). El tipo de dispersión tumoral WPOI-5 en los pacientes que se trataron con cirugía está presente en 10 (83.3%) y ausente en 2 (16.6%).

#### **8.1.4 Datos de tratamiento**

En cuanto al tipo de tratamiento, 5 pacientes tuvieron únicamente cirugía (25%); 6 pacientes, cirugía seguido de radioterapia (30%); 4 fueron tratados mediante quimioterapia con radioterapia concomitante (20%); un paciente, con cirugía seguido de QT/RT concomitante (5%); uno con QT/RT concomitante seguido de cirugía (5%); uno con tratamiento con fines paliativos (5%) y 2 pacientes no iniciaron tratamiento alguno (10%).

#### **8.1.5 Datos pronósticos**

Se encontró recurrencia de la enfermedad en 3 pacientes (15%) con una media de 5.2 meses (2.5 – 10.3 meses); persistencia de la enfermedad pese a tratamiento con intento curativo, en 3 pacientes (17.6%); y progresión de la enfermedad durante el tratamiento, en 4 pacientes (23.5%). Todos los pacientes con persistencia y/o progresión de la enfermedad han fallecido. Para el término del seguimiento, 12 permanecen vivos (60%) y 8 fallecieron (40%). La supervivencia global a 2 años se calculó en 57%. En cuanto a etapa clínica categorizada, en etapa temprana, todos los pacientes continúan vivos; en etapa avanzada, fue de 8.2 meses (IC N/A), el 49% en este grupo está vivo a un año, con una  $p=0.193$  (figura 6). La supervivencia libre de enfermedad a un año fue de 67%: en etapa temprana, de 10.3 meses (IC N/A) y en etapa avanzada, de 2.5 meses (IC N/A) con  $p=0.225$  (figura 7).

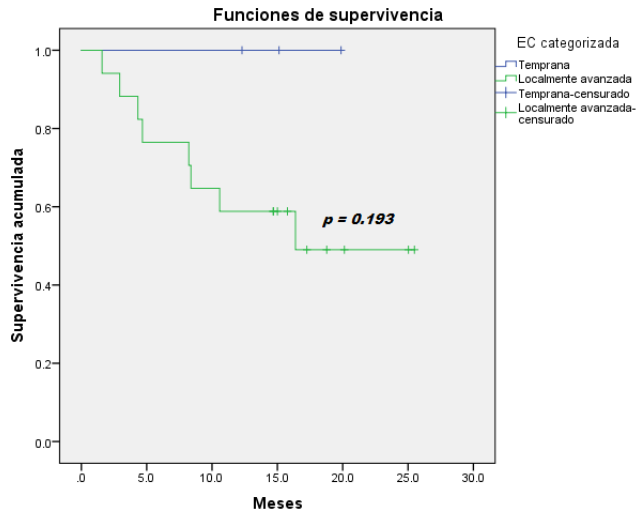


Figura 6

Figura 6. Supervivencia global (se separan las curvas, mostrando una clara tendencia a la peor supervivencia global en el cáncer avanzado)

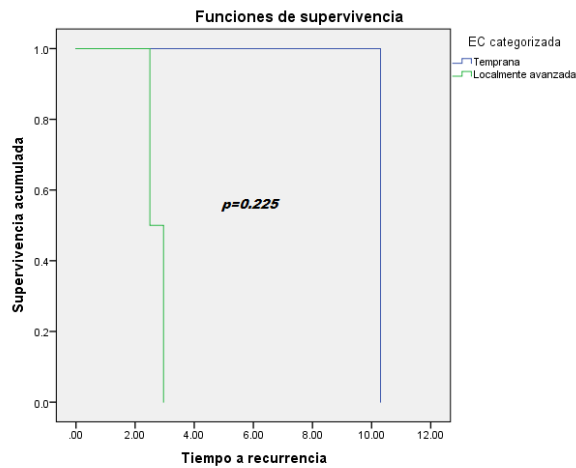


Figura 7

Figura 7. Supervivencia libre de enfermedad (debido a los pocos casos de recurrencia no se observa en el gráfico la separación de las curvas de forma adecuada)

Tabla 7. Características clínico-patológicas de los pacientes					
N = 20 Pacientes					
<b>Edad</b>	54 años (21 – 86)	NA	<b>Subsitio AJCC</b>		
<b>Sexo</b>			Lengua móvil	9	45
<b>Masculino</b>	12	60	Encía inferior	4	20
<b>Femenino</b>	8	40	Encía superior	5	25
<b>Seguimiento</b>	13.4 meses (1.6 – 25.5)	NA	Piso de boca	1	5

<b>Escolaridad</b>			Trígono retromolar	1	5
Analfabeta	2	10	Mucosa oral	0	0
Primaria	9	45	Paladar duro	0	0
Secundaria	1	5	<b>Tamaño Tumoral</b>		
Bachillerato y/o carrera técnica	3	15	Tis	1	5
Licenciatura	5	25	T1	1	5
Posgrado	0	0	T2	1	5
<b>Desnutrición</b>			T3	3	15
Sí	7	35	T4a	11	55
No	13	65	T4b	3	15
<b>ECOG</b>			<b>Ganglios linfáticos</b>		
1	19	95	N0	11	55
3	1	5	N1	3	15
<b>Tabaquismo</b>			N2a	0	0
Sí	15	75	N2b	2	10
No	5	25	N2c	1	5
Actual fumador	1	6.6	N3a	0	0
<b>Índice tabáquico</b>	6.06 (+/- 6.08)	NA	N3b	3	15
<b>Tiempo de Fumar</b>			<b>Etapa Clínica</b>		
<10	7	50	0	1	5
11-20	3	21.4	I	1	5
21-30	2	14.2	0	0	5
31-40	0	0	III	3	15
>40	2	14.2	IVa	9	45
<b>Alcoholismo</b>			IVb	5	25
Sí	13	65	IVc	0	0
No	7	35	<b>Profundidad de Invasión DOI</b>	7.3 (3 – 130mm)	20
<b>Bebedor actual</b>	3	23	4		
<b>Unidades de Bebida Estándar a la semana</b>			<b>Permeación linfovascular</b>		
<1	2	15.3	Si	7	35
1-3	4	30.7	No	10	50
4-5	2	15.3	No valorable	3	15
>5	5	38.4	<b>Infiltración perineural</b>		
<b>Grado de consumo</b>			Sí	5	25
Ninguno	6	30	No	12	60
Leve	2	10	No valorable	3	15
Moderado	6	30	<b>Grado histológico</b>		
Excesivo	5	25	Bien diferenciado	3	15
<b>Virus del Epstein Barr</b>			Moderadamente diferenciado	15	75

Negativo	11	55	Pobremente diferenciado	2	10
Positivo	6	30	<b>Márgenes quirúrgicos</b>		
No determinado	3	15	Negativo	5	41.6
<b>Virus del Papiloma Humano</b>			Positivo	4	33.3
Negativo	17	85	Cercanos	3	25
Positivo	1	5	No aplica	8	40
No determinado	2	10	<b>Dispersión tumoral WPOI-5</b>		
<b>Tratamiento</b>			Sí	10	83.3
Cirugía	5	25	No	2	16.6
Cirugía + RT	6	30	No valorable	8	40
QT/RT	4	20	<b>Recurrencia</b>	3 5.2 meses (2.5- 10.3)	15
Sin inicio de Tx	2	10	<b>Persistencia de la enfermedad</b>		
Paliativo (Cx, QT, RT, QT/RT)	1	5	No	14	82.3
Cirugía + QT/RT	1	5	Sí	3	17.6
QT/RT + Cirugía	1	5	<b>Progresión de la enfermedad</b>		
<b>Supervivencia global a 1 año</b>			No	13	76.4
Etapas tempranas	Todos vivos	100	Sí	4	23.5
Etapas avanzadas	8.2 meses (IC N/A)	NA	No aplica	3	15
<b>Supervivencia libre de enfermedad a 1 año</b>			<b>Estado del paciente</b>		
Etapas tempranas	10.3 meses (IC N/A)	NA	Vivo	12	60
Etapas avanzadas	2.73 meses (IC N/A)	NA	Muerto	8	40
<b>NA: No aplica</b>					
<b>N/A: No alcanzado</b>					

## 8.2 Análisis descriptivo molecular

### 8.2.1 Identificación de variantes patogénicas

Las variantes patogénicas detectadas en, al menos, 11% de los pacientes con cáncer bucal, abarcaron 12 genes alterados. Se destaca que *P53* se encontró mutado en el 47% los pacientes, seguido de *NOTCH1*, en el 26% de los pacientes; *FAT1*, en el 21%; y *PTPRK*, en el 16%. Los siguientes genes se encontraron mutados en el 11% de los pacientes: *BIRC6*, *DDX3X*, *EPHA2*, *FBXW7*, *KMT2C*, *KMT2D*, *NIN*, *NXF*, *RASA1* y *SMARCA4* (figura 8). En la tabla 8 se encuentran los genes mutados en, al menos, 11% de los pacientes y se señalan las funciones conocidas. Además, se indica si el gen ha sido reportado en estudios previos de cáncer bucal.

En el 5% de los pacientes se encontraron variantes patogénicas en los genes que se enlistan a continuación: *ABCB1, ADGRB1, AJUBA, ALK, APC, ARHGAP35, ARHGEF10L, ARID1A, ARID2, ASXL1, ATRX, BMPR2, CASP8, CDKN2A, CEBPA, CHEK2, CUP1, CREBBP, DCC, DOT1L, EFTUD2, EML4, EP300, EPHA3, FANCA, FAS, FAT3, FAT4, FLT3, GRIN2A, HGF, HNF1A, HSP90AA1, KEAP1, KMT2A, KMT2B, LATS1, LPAR4, MAML2, MAPK1, MYH11, NCOR1, NCOR2, NIPBL1, NSD1, NTRK1, NTRK3, PBRM1, PEG3, PML, PPM1D, PPP3CA, PPP6C, PREX2, PRF1, RASA2, RBM15, RGD3, RNF213, RPS6KA3, SETBP1, SETD2, SMARCA1, SPEN, STAT3, STK11, TBX3, TET1, TGFB2, TNC, TRIM33, U2AF2, USP8, WNK2, ZFX y ZNF521* (figura 8).

<b>Tabla 8. Principales funciones de las variantes patogénicas encontradas en el análisis de secuenciación de carcinoma bucal</b>				
<b>Gen</b>	<b>Frecuencia %</b>	<b>Pacientes afectados</b>	<b>Función</b>	<b>Riesgo para cáncer bucal de células escamosas (tabla 1)</b>
<i>TP53</i>	47	CSTE_2 CSTE_3 CSTE_6 CSTE_10 CSTE_11 CSTE_17 CSTE_18 CSTE_19 CSTE_20	La proteína codificada responde a diversos estímulos celulares para regular la expresión de genes, induciendo paro del ciclo celular, apoptosis, senescencia, reparación del DNA o cambios en el metabolismo.	Descrito en los siguientes estudios: Agrawal et al. (25) Stransky et al. (26) Pickering et al. (28) ICGC (29) TCGA (30) Al-Hebshi et al. (31) Su et al. (32)
<i>NOTCH1</i>	26	CSTE_4 CSTE_6 CSTE_15 CSTE_17 CSTE_18	Los miembros de esta familia de proteínas transmembrana de tipo I que consta de múltiples repeticiones del factor de crecimiento epidérmico (EGF) y un dominio intracelular que consta de múltiples tipos de dominios diferentes. La señalización de Notch es una vía de señalización intercelular que regula las interacciones entre células físicamente adyacentes mediante la unión de los receptores de la familia Notch a sus ligandos afines.	Descrito en los siguientes estudios: Agrawal et al. (25) Stransky et al. (26) Pickering et al. (28) ICGC (29) TCGA (30) Su et al. (32)
<i>FAT1</i>	21	CSTE_1 CSTE_6 CSTE_11 CSTE_14	El producto génico es un miembro de la superfamilia de cadherinas, un grupo de proteínas integrales de membrana. Se expresa en niveles elevados en varios epitelios fetales. Su producto probablemente funciona como molécula de adhesión y / o receptor de señalización, y es probable que sea	Descrito en los siguientes estudios: Stransky et al. (26) ICGC (29) TCGA (30) Al-Hebshi et al. (31) Su et al. (32)

			importante en los procesos de desarrollo y la comunicación celular.	
<i>PTPRK</i>	16	CSTE_1 CSTE_19 CSTE_20	La proteína codificada por este gen es un miembro de la familia de la proteína tirosina fosfatasa (PTP). Se sabe que las PTP son moléculas de señalización que regulan una variedad de procesos celulares que incluyen el crecimiento celular, la diferenciación, el ciclo mitótico y la transformación oncogénica. Se descubrió que la expresión de este gen está estimulada por TGF-beta 1, que puede ser importante para la inhibición de la proliferación de queratinocitos.	<b>No descrito en los estudios previos</b>
<i>BIRC6</i>	11	CSTE_1 CSTE_19	Codifica una proteína que inhibe la apoptosis al facilitar la degradación de proteínas apoptóticas por ubiquitinación.	<b>No descrito en los estudios previos</b>
<i>DDX3X</i>	11	CSTE_1 CSTE_9	La proteína codificada por este gen es un miembro de la gran familia de proteínas DEAD-box, tiene actividad RNAhelicasa dependiente de ATP. Las funciones nucleares incluyen la regulación transcripcional, el ensamblaje de mRNP, el empalme de pre-mRNA y la exportación de mRNA. En el citoplasma, se cree que esta proteína está involucrada en la traducción, la señalización celular y la replicación viral.	<b>No descrito en los estudios previos</b>
<i>EPHA2</i>	11	CSTE_1 CSTE_9	Este gen pertenece a la subfamilia del receptor de efrina de la familia de las proteínas tirosina quinasa. Se han implicado en la mediación de eventos del desarrollo, particularmente en el sistema nervioso.	Descrito en los siguientes estudios: Su et al. (32)
<i>FBXW7</i>	11	CSTE_6 CSTE_7	Este gen codifica proteínas que funcionan en la ubiquitinación dependiente de la fosforilación. Esta proteína se une directamente a la ciclina E y probablemente se dirige a la ciclina E para la degradación mediada por ubiquitina. Las	Descrito en los siguientes estudios: Agrawal et al. (25) TCGA (30)



			mutaciones en este gen se detectan en líneas celulares de cáncer de mama y de ovario.	
<i>KMT2C</i>	11	CSTE_1 CSTE_16	Este gen es miembro de la familia de la leucemia mieloide / linfocida o de linaje mixto (MLL) y codifica una proteína nuclear, la cual es miembro del complejo ASC-2 / NCOA6 (ASCOM), que posee actividad de metilación de histonas y participa en la coactivación transcripcional.	<b>No descrito en ningún estudio previo</b>
<i>KMT2D</i> ( <i>MLL2</i> , <i>MLL4</i> )	11	CSTE_1 CSTE_6	La proteína codificada por este gen es una histona metiltransferasa que metila la posición Lys-4 de la histona H3. La proteína codificada es parte de un gran complejo de proteínas llamado ASCOM, que se ha demostrado que es un regulador transcripcional de los genes receptores de beta-globina y estrógeno.	Descrito en los siguientes estudios: Stransky et al. (26) TCGA (30)
<i>NIN</i>	11	CSTE_1 CSTE_4	Este gen codifica una de las proteínas importantes para la función del centrosoma. Esta proteína es importante para colocar y anclar los extremos negativos de los microtúbulos en las células epiteliales.	<b>No descrito en ningún estudio previo</b>
<i>NXF1</i>	11	CSTE_1 CSTE_6	Este gen es un miembro de una familia de genes de factores de exportación de RNA nuclear. Los dominios LRR y similares a NTF2 son necesarios para la actividad de exportación.	<b>No descrito en ningún estudio previo</b>
<i>RASA1</i>	11	CSTE_14 CSTE_16	La proteína codificada por este gen se encuentra en el citoplasma y es parte de la familia GAP1 de proteínas activadoras de GTPasa. Actúa como un supresor de la función RAS, la proteína mejora la actividad GTPasa intrínseca débil de las proteínas RAS dando como resultado la forma inactiva de RAS unida a GDP, lo que permite el control de la proliferación y diferenciación celular. Las	Descrito en los siguientes estudios: Agrawal et al. (25)

			mutaciones que conducen a cambios en los sitios de unión de cualquiera de las proteínas se asocian con carcinomas de células basales.	
SMARCA4	11	CSTE_1 CSTE_16	La proteína codificada por este gen es miembro de la familia de proteínas SWI / SNF. Los miembros de esta familia tienen actividades de helicasa y ATPasa y se cree que regulan la transcripción de ciertos genes al alterar la estructura de la cromatina alrededor de esos genes. La proteína codificada es parte del gran complejo de remodelación de cromatina dependiente de ATP SNF / SWI, que se requiere para la activación transcripcional de genes normalmente reprimidos por la cromatina. Además, esta proteína puede unirse a BRCA1, así como regular la expresión de la proteína tumorigénica CD44. Las mutaciones en este gen causan el síndrome de predisposición al tumor rabdoide tipo 2.	<b>No descrito en ningún estudio previo</b>

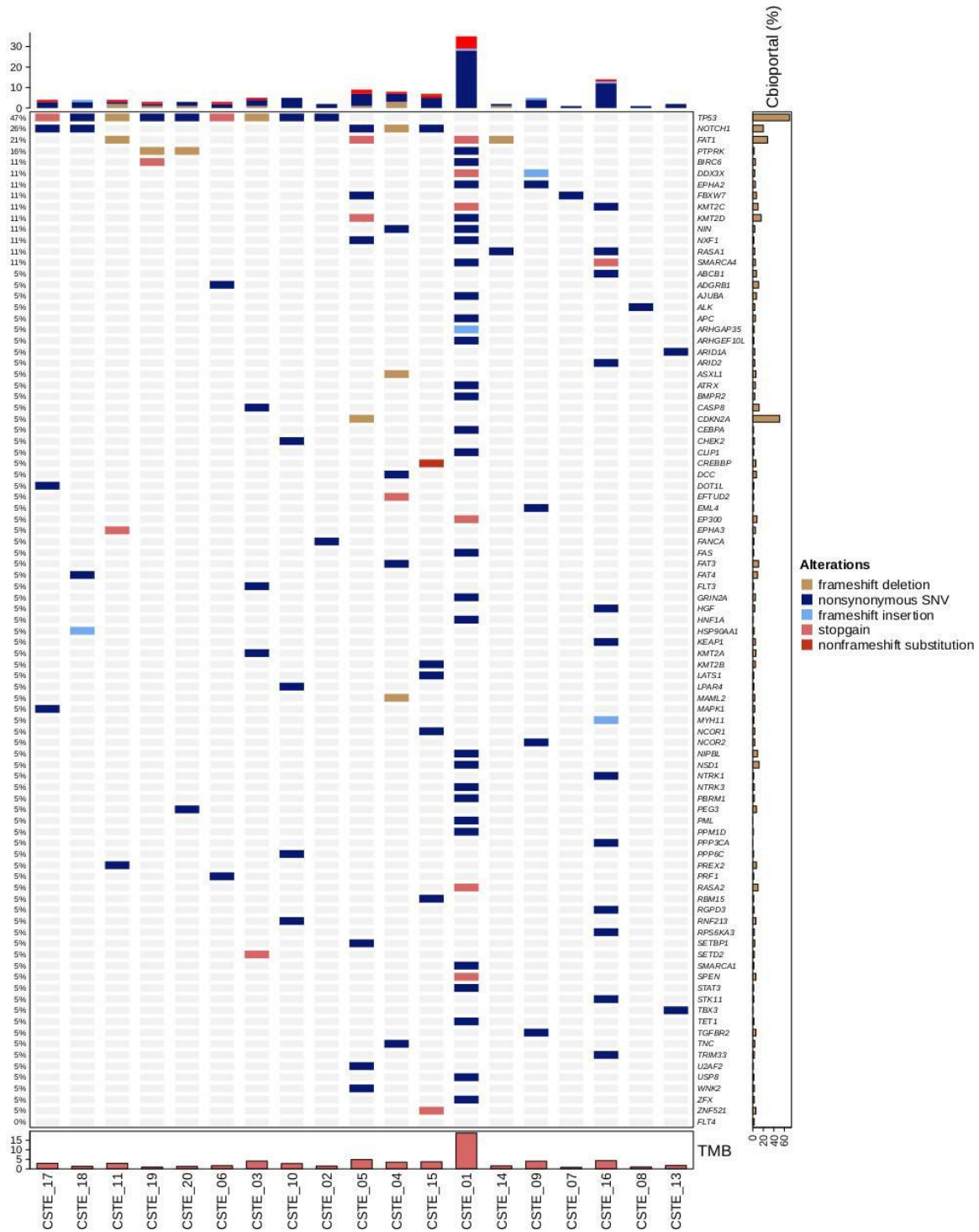


Figura 8. Variantes patológicas. En amarillo, variantes de delección de nucleótidos con cambio de marco de lectura, azul marino se muestran variantes no sinónimas, en azul claro variantes con inserción de nucleótidos con cambio de marco de lectura, en rosa variantes con generación de un codón de paro y en rojo variantes de sustitución sin cambio de marco de lectura.

### 8.2.2. Identificación de mutaciones conductoras

De acuerdo con el análisis diferencial entre el tejido adyacente y tumoral, se encontraron un total de 9398 genes mutados en las 20 muestras de los pacientes. Entre los genes mutados, se identificaron 181 mutaciones conductoras. En todos los pacientes se encontraron, al menos, 2 mutaciones conductoras, mientras que el número máximo de mutaciones conductoras en un paciente fue de 44 (tabla 9).

Paciente	Número de genes conductores	Paciente	Número de genes conductores
CSTE_1	44	CSTE_11	5
CSTE_2	4	CSTE_12	5
CSTE_3	12	CSTE_13	4
CSTE_4	18	CSTE_14	6
CSTE_5	14	CSTE_15	8
CSTE_6	4	CSTE_16	13
CSTE_7	2	CSTE_17	6
CSTE_8	3	CSTE_18	5
CSTE_9	7	CSTE_19	7
CSTE_10	8	CSTE_20	6

### 8.2.3 Identificación de vías moleculares

De acuerdo con el análisis de enriquecimiento, se identificaron 44 vías de señalización molecular afectadas, siendo las más frecuentes la vía de señalización Wnt/ $\beta$ catenina, seguido de la vía *P53* y vías de angiogénesis, incluyendo la vía del factor de crecimiento derivado de las plaquetas (PDGF). Adicionalmente, se encontró alteración de la vía del receptor del factor de crecimiento epitelial (EFG), apoptosis, del factor de crecimiento transformador beta, de la vía de la e-caderina, la vía de las citocinas, la vía del factor de crecimiento fibroblástico y la vía de retroalimentación en asa de *P53*. El resto de las vías, así como el número de genes alterados en cada una, se describen en la tabla 10 (figura 9).

<b>Tabla 10. Vías de señalización alteradas en pacientes con cáncer bucal</b>			
<b>Vía</b>	<b>Número de genes afectados</b>	<b>Vía</b>	<b>Número de genes afectados</b>
Guía de axones mediada por netrin	3	Vía de señalización de VEGF	3
Guía de axones mediada por Slit / Robo	1	Regulación de la transcripción por el factor de transcripción bZIP	2
Vía de señalización de la apoptosis	6	Vía de señalización del receptor Toll	2
Vía de señalización JAK / STAT	2	Vía de señalización Ras	2
Angiogénesis	8	Activación de células T	4
Vía de la enfermedad de Alzheimer-presenilina	3	Vía de señalización FGF	5
Vía de señalización de interleucina	4	Bucles de retroalimentación de la vía p53 1	1
Vía de la secretasa amiloidea de la enfermedad de Alzheimer	1	Vía de señalización de TGF-beta	6
Vía de señalización de interferón-gamma	3	Vía de señalización de FAS	2
Vía de señalización de integrina	4	Vía de señalización Endothelin	2
Cascada de señalización de la proteína cinasa B de la vía de la insulina / IGF	2	Vía de señalización del receptor EGF	7
Ruta de la insulina / IGF-proteína cinasa activada por mitógenos / cascada de la cinasaMAP	3	Enfermedad de Parkinson	2
Inflamación mediada por la vía de señalización de quimiocinas y citocinas	5	Vía de señalización cinasa PI3	3
Respuesta de hipoxia a través de la activación de HIF	3	Sistema de reloj circadiano	1
Vía farmacodinámica de la nicotina	1	Vía de señalización PDGF	8
Enfermedad de Huntington	5	Vía de señalización Notch	2
Vía p53	9	Vía de señalización e-caderina	5
Bucles de retroalimentación de la vía p53 2	5	Vía de señalización del receptor de dopamina	1
Vía p53 por privación de glucosa	2	Activación de células B	3
Vía de señalización de proteína G heterotrimérica - Vía mediada por Gi alfa y Gs alfa	2	Señalización estimulada por angiotensina II a través de proteínas G y beta-arrestina	1
Vía de señalización de WNT	13	Mapa de señalización CCKR	3
Vía de señalización de Hedgehog	2	Vía del receptor de la hormona liberadora de gonadotropina	8

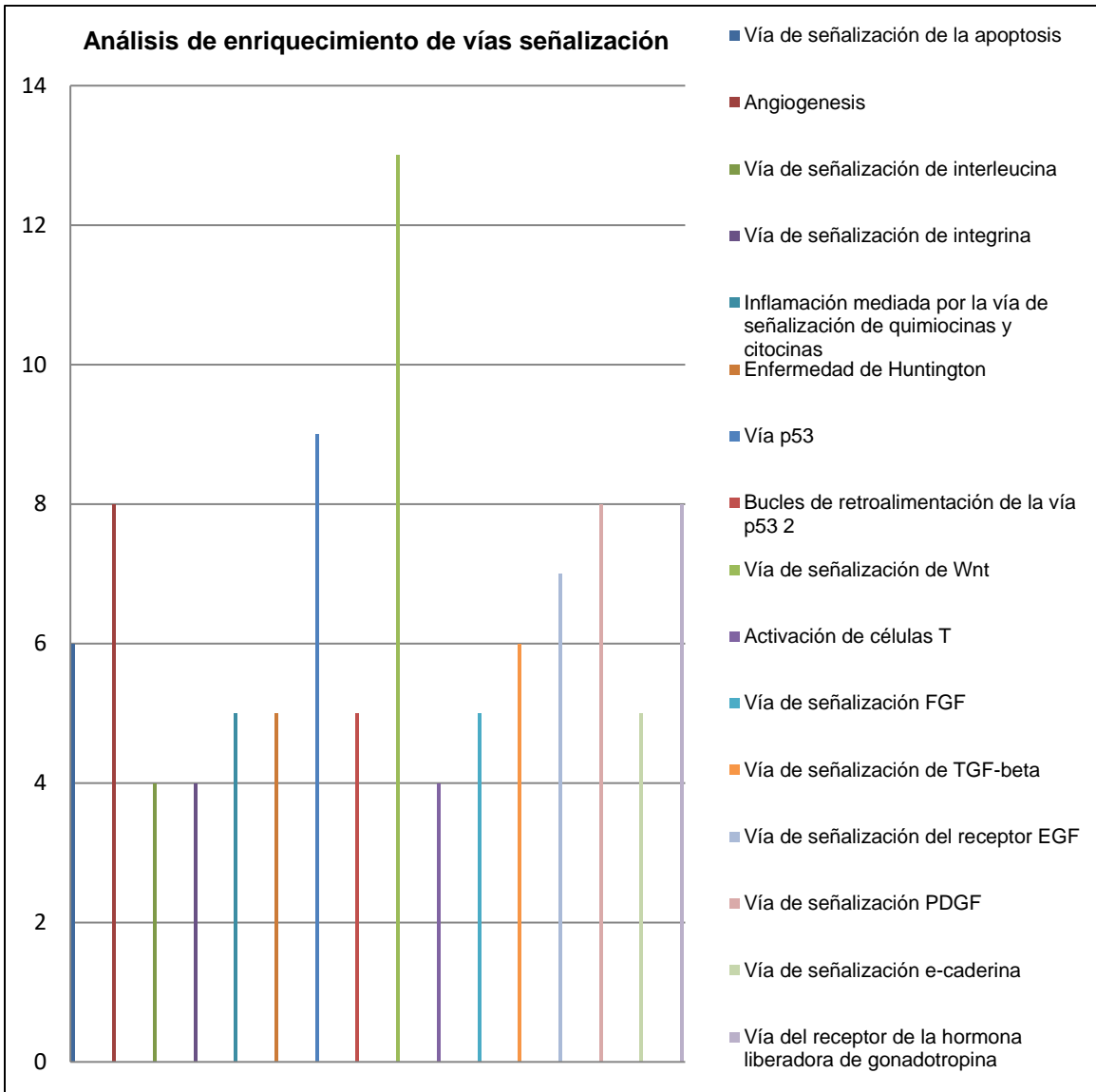


Figura 9. Análisis de enriquecimiento de las vías alteradas en cáncer bucal de acuerdo con el número de mutaciones conductoras encontradas

Dentro de las funciones moleculares alteradas en los pacientes con cáncer bucal, se encontraron mutaciones conductoras que participan en uniones celulares (binding), actividad catalítica, transducción molecular, actividad de transportadores, actividad de adaptadores moleculares y en regulación de la traducción (figura 10).

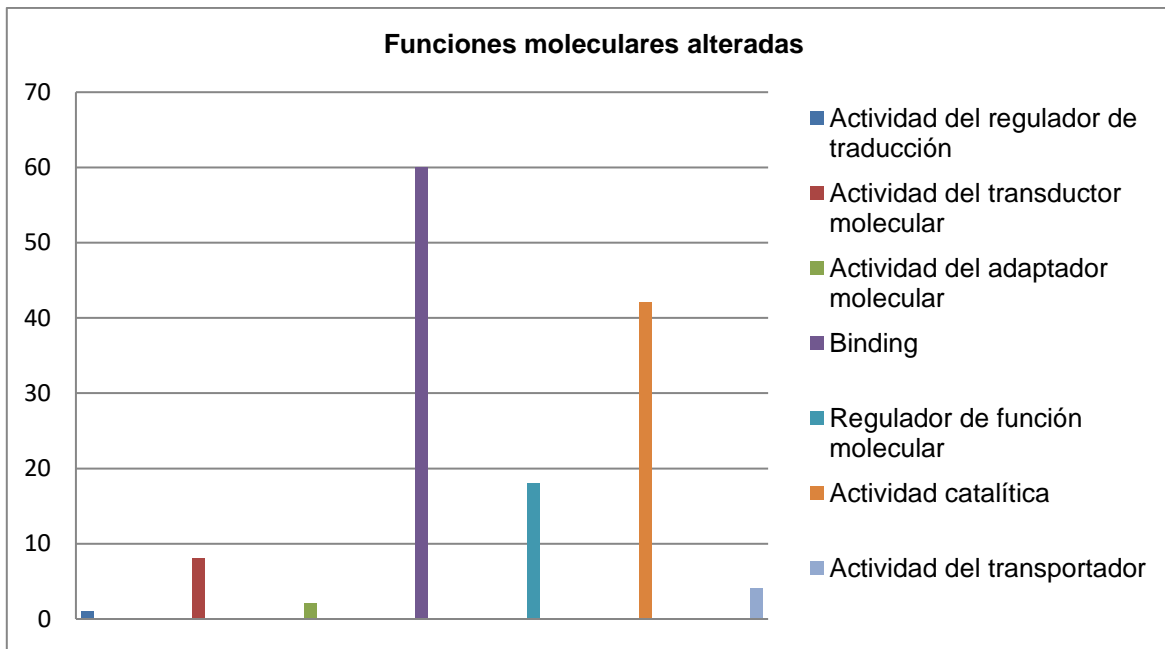


Figura 10. Análisis de enriquecimiento de las funciones moleculares alteradas en cáncer bucal de acuerdo con el número de mutaciones conductoras encontradas

En cuanto a los procesos biológicos alterados, se encontraron que los genes conductores participan en procesos celulares, metabólicos, regulación biológica, respuesta a estímulos, señalización, desarrollo, localización, locomoción, adhesión biológica, sistema inmune y fase biológica (figura 11).

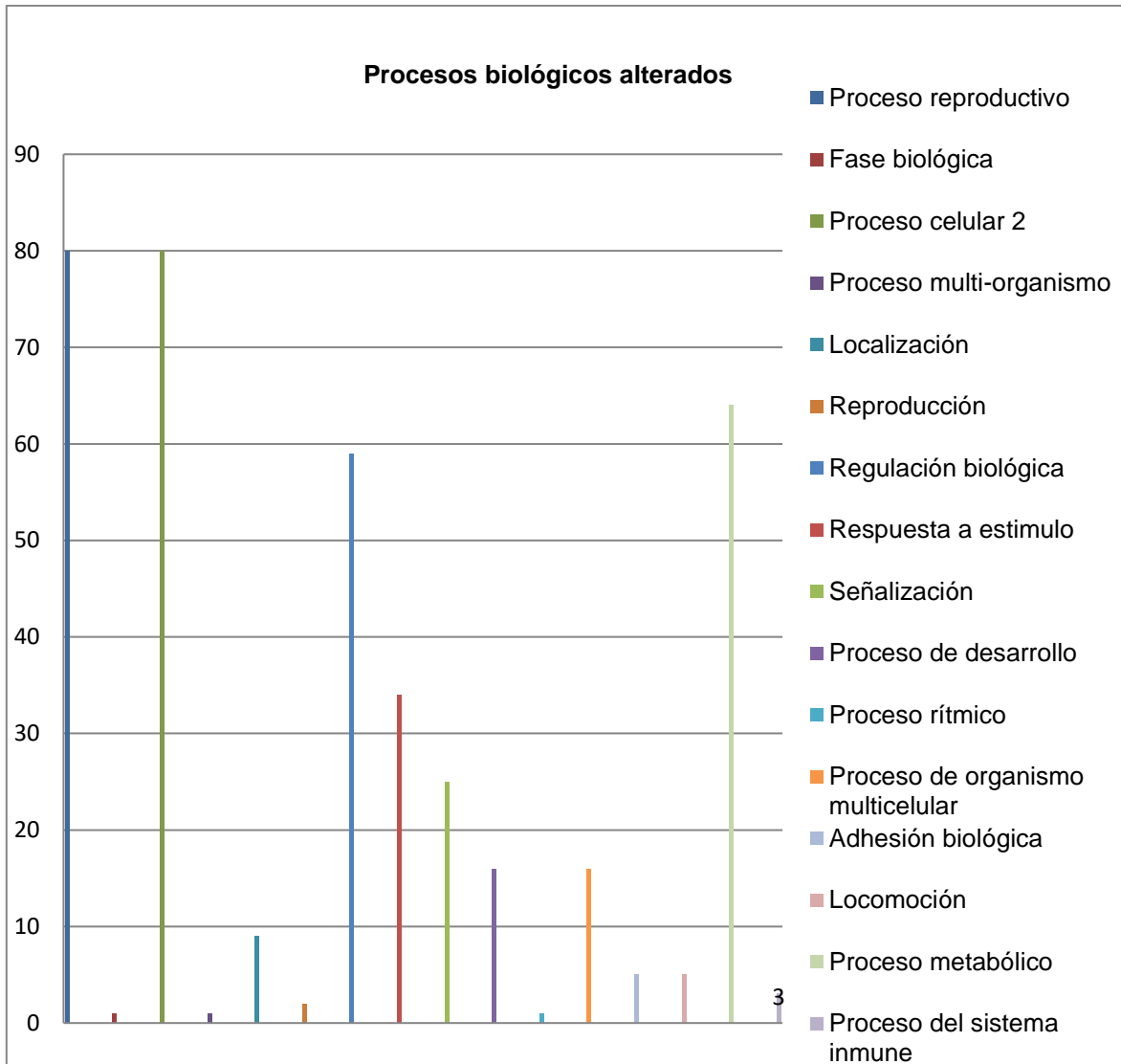


Figura 11. Análisis de enriquecimiento de los procesos biológicos alterados en cáncer bucal de acuerdo con el número de mutaciones conductoras encontradas

### 8.2.4 Carga mutacional tumoral (TMB)

En cuanto al número total de mutaciones somáticas que se encuentra en el DNA de los pacientes con cáncer bucal, se determinó que la mediana es de 3.29 mutaciones/Mb, mientras que se encontró un máximo de 18.78 mutaciones/Mb de DNA en el paciente 1. Por el contrario, el índice mínimo de mutación fue de 0.8 mutaciones/Mb en el paciente 4 (tabla 11).



Muestra	Frecuencia	TMB Mb	Muestra	Frecuencia	TMB Mb
CSTE_1	1127	18.78	CSTE_11	171	2.85
CSTE_2	89	1.48	CSTE_12	124	2.06
CSTE_3	246	4.1	CSTE_13	105	1.75
CSTE_4	211	3.51	CSTE_14	95	1.58
CSTE_5	292	4.86	CSTE_15	223	3.71
CSTE_6	101	1.68	CSTE_16	259	4.31
CSTE_7	48	0.8	CSTE_17	172	2.86
CSTE_8	61	1.01	CSTE_18	82	1.36
CSTE_9	243	4.05	CSTE_19	57	0.95
CSTE_10	166	2.76	CSTE_20	77	1.28

**TMB: Tumor Mutational Burden**

Para su estudio y correlación con la sobrevida, la carga tumoral se dividió en carga mutacional alta y baja con base al estudio del Memorial Sloan Kettering Cancer Center (47) (tabla 12 y figura 12).

Carga tumoral	Pacientes afectados
Carga tumoral alta	CSTE_01, CSTE_03, CSTE_04, CSTE_05, CSTE_09, CSTE_10, CSTE_11, CSTE_12, CSTE_15, CSTE_16 y CSTE_17
Carga tumoral baja	CSTE_02, CSTE_06 CSTE_07, CSTE_08, CSTE_13, CSTE_14, CSTE_18, CSTE_19 y CSTE_20

### 8.2.5 Firmas mutacionales

Se analizaron únicamente firmas mutacionales de sustituciones simples. La firma más representada fue la SBS 1 en 18 pacientes (90%), seguida de la firma SBS 13 en 12 pacientes (60%) y la firma SBS 3 en 11 pacientes (55%). Adicionalmente, se encontraron las firmas SBS 5, 6 y 7 en 6 pacientes (30%), así como la firma SBS 2, hallada en 5 pacientes (25%).

Otras firmas menos frecuentes fueron las siguientes: SBS 15 y 24, en 4 pacientes (20%); SBS 29 y 30, en 3 pacientes (15%); y las firmas SBS 4, 20, 21 y 27, en 2 pacientes (10%). Las firmas menos frecuentes que se encontraron en un solo paciente (5%) fueron SBS 9, 10, 12, 16, 17, 19 y 22. No se encontraron las siguientes firmas: SBS 8, 11, 14, 18, 23, 25, 26 y 28.

Posteriormente, se distribuyeron todos los pacientes para definir grupos basado en la composición de las firmas mutacionales en 3 grupos: el grupo 1, formado por 2 pacientes (10%); el grupo 2, por 11 (55%); y el grupo 3, por 7 (35%) (figura 13).

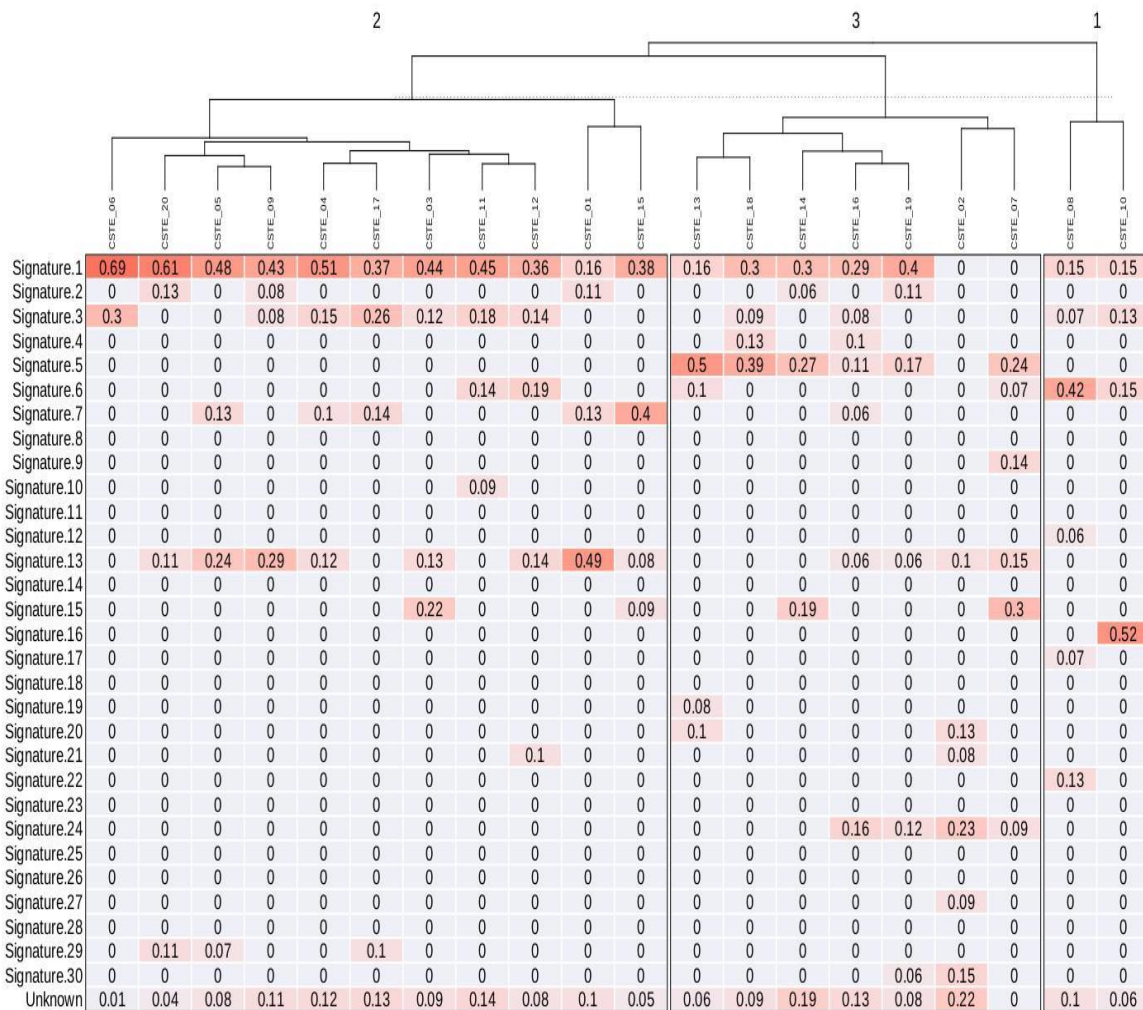


Figura 12. Mapa de calor de la distribución genómica del tumor según el perfil de firmas mutacionales encontradas. SB-firmas de sustitución simple

### 8.3 Factores de riesgo

#### 8.3.1 Datos de tabaquismo y alcohol

En cuanto al tabaquismo, 15 pacientes (75%) se consideraron como fumadores, de los cuales, 1 (6.6%) continuaba fumando, a pesar del diagnóstico de cáncer bucal. El índice tabáquico fue de 6.06 (+/- 6.08): 7 pacientes (50%) tienen menos de 10 años fumando; 3 (21.4%), 11 a 20 años; 2 (14.2%), 21 a 30 años; y 2 (14.2%), más de 40 años de hábito tabáquico. El consumo de alcohol se documentó en 13 pacientes (65%), 7 no tienen antecedentes de su ingesta (35%) y 3 pacientes son bebedores actuales, pese al diagnóstico (23%). El consumo por unidades estándar de bebida fue más de 5 UB a la semana en 5 pacientes (38.4%); 1 a 3 UB a la semana, en 4 (30.7%); 4 a 5 UB a la semana, en 2 (15.3%); y menos de 1 UB a la semana, en 2 (15.3%). Por grado de consumo se catalogó como leve, 2 (10%); moderado, 6 (30%); y excesivo, 5 (25%).

### 8.3.2 Resultados de PCR VPH

Del total de las muestras, solo 18 contaban con DNA suficiente para realizar la determinación por qRT-PCR de VPH. La presencia de VPH fue observada únicamente en un paciente CSTE\_9, lo cual representó el 5.5% de la muestra y se determinó que correspondió al tipo 16.

### 8.3.3 Resultados de PCR VEB

De las 20 muestras secuenciadas, solo 17 fueron correspondientes al tumor y 18 de tejido adyacente tenían suficiente DNA para analizar la presencia de VEB, la cual fue observada en el tejido tumoral de 6 pacientes (35.3%). De estos, solo 2 presentaron positividad para VEB en el tejido adyacente correspondiente (muestra 3 y 9) y 2 presentaron positividad para VEB únicamente en el tejido adyacente (11.1%) (muestras 4 y 18) (tabla 13).

Muestra	Tumor	Tejido adyacente	Muestra	Tumor	Tejido adyacente
CSTE_1	Negativo	Negativo	CSTE_11	Positivo	Negativo
CSTE_2	Negativo	Negativo	CSTE_12	Positivo	Negativo
CSTE_3	Positivo	Positivo	CSTE_13	Negativo	Negativo
CSTE_4	Negativo	Positivo	CSTE_14	Negativo	Negativo
CSTE_5	Positivo	Negativo	CSTE_15	Positivo	Negativo
CSTE_6	Negativo	Negativo	CSTE_16	Negativo	Negativo
CSTE_7	Negativo	Negativo	CSTE_17	Negativo	Negativo
CSTE_8	Negativo	Negativo	CSTE_18	Sin material suficiente	Positivo
CSTE_9	Positivo	Positivo	CSTE_19	Sin material suficiente	Sin material suficiente
CSTE_10	Negativo	Negativo	CSTE_20	Sin material suficiente	Sin material suficiente

## 8.4 Análisis de supervivencia

### 8.4.1 Variables patogénicas

Se analizó la supervivencia con todas las variables patogénicas encontradas en el análisis. Se encontró que únicamente la mutación de *FAT1* tiene impacto en la supervivencia global en el análisis multivariado RR 73.2 (1.4 – 3,602)  $p=0.031$ , mientras que para el análisis de Kaplan Meier, se encontró solo una tendencia ( $p=0.290$ ). La presencia de esta variante patogénica se asoció a 8.3

meses (IC 95% 0 – 17.1) vs. 18.3 meses (IC 95% 14 – 22.7) (figura 13). Para supervivencia libre de enfermedad, la única mutación que mostró una tendencia fue *FBXW7* RR 30.7 (0 – 90.2) ( $p=0.574$ ). Para el resto de las variantes, no se encontraron diferencias estadísticamente significativas ( $p<0.05$ ) o tendencias claras (tabla 14).

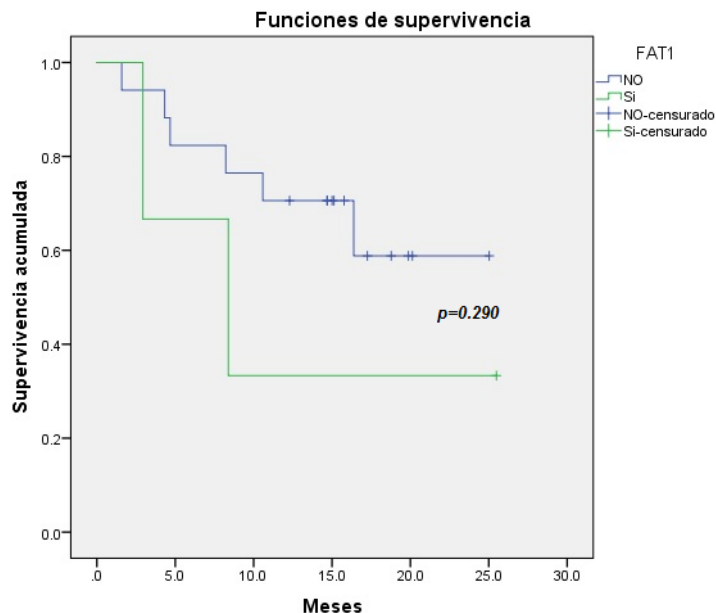


Figura 13. Supervivencia global (Kaplan Meier) en relación con la variante patogénica FAT1. Se observan curvas separadas; sin embargo, la p resultó no significativa.

Tabla 14. Análisis de supervivencia (Kaplan Meier y regresión de Cox) para supervivencia global y libre de enfermedad de variantes patogénicas							
SUPERVIVENCIA GLOBAL (SG)				SUPERVIVENCIA LIBRE DE ENFERMEDAD (SLE)			
GEN	MEDIANA MESES (IC 95%)	P	REGRESIÓN DE COX (IC 95%)	GEN	MEDIANA MESES (IC 95%)	P	REGRESIÓN DE COX (IC 95%)
<i>TP53</i>	Ausente 18.5 (13–24) Presente 16.8 (10.6-22.9)	0.630	RR 0.05 (0.05–6.4) $p=0.656$	<i>TP53</i>	NA		NA
<i>NOTCH1</i>	Ausente 17.8 (12.9-22.7) Presente 12.1 (10.6-15.7)	0.924	RR 0.006 (NA) $p=0.698$	<i>NOTCH1</i>	Ausente 2.5 (NA) Presente 6.6 (0-13.8)	0.157	RR 0 (NA) $p=0.438$
<i>FAT1</i>	<b>Ausente 18.3 (14-22.7)</b> <b>Presente 8.3 (0-17.1)</b>	<b>0.290</b> <b>++</b>	<b>RR 73.2 (1.4-3,602) <math>p=0.031</math></b> <b>**</b>	<i>FAT1</i>	NA		NA

<b>PTPRK</b>	Ausente 17.4 (12.9-21.8) Presente 18.4 (0 -29.7)	0.727	RR 0.001 (NA) p=0.905	<b>PTPRK</b>	NA		NA
<b>BIRC6</b>	Ausente 17.7 (13.4-22) Presente 14.9 (0.2-29.5)	0.727	NA	<b>BIRC6</b>	NA		NA
<b>DDX3X</b>	NA		RR 0 (NA) p=0.988	<b>DDX3X</b>	NA		NA
<b>EPHA2</b>	NA		NA	<b>EPHA2</b>	NA		NA
<b>FBXW7</b>	Ausente 17.6 (13-22.2) Presente 14.6 (9 -20.3)	0.981	RR 0.004 (NA) p=0.759	<b>FBXW7</b>	<b>Ausente 6.4 (0-14) Presente 2.9 (2.9-10)</b>	<b>0.808 ++</b>	<b>RR 30.78 (0- 90.20) p=0.547 ++</b>
<b>KMT2C</b>	NA		NA	<b>KMT2C</b>	NA		NA
<b>KMT2D (MLL2, MLL4)</b>	Ausente 17.4 (12.9–21.8) Presente 18 (7.7–28.3)	0.981	NA	<b>KMT2D (MLL2, MLL4)</b>	Ausente 6.4 (0–14) Presente 2.9 (NA)	0.808	NA
<b>NIN</b>	Ausente 17.7 (13.4–22) Presente 15 (0.6–29.4)	0.790	NA	<b>NIN</b>	NA		NA
<b>NXF1</b>	Ausente 17.4 (12.9–21.8) Presente 18 (7.7–28.3)	0.981	NA	<b>NXF1</b>	Ausente 6.4 (0–14) Presente 2.9 (NA)	0.808	NA
<b>RASA1</b>	Ausente 17.9 (13.4–22.3) Presente 8.3 (7.1–15.8)	0.774	RR 0.04 (0.001–3.9) p=0.176	<b>RASA1</b>	NA		NA
<b>SMARCA4</b>	NA		NA	<b>SMARCA4</b>	NA		NA
<b>NA (No Alcanzado)</b> <b>RR (Riesgo relativo)</b> <b>++ las p más cercana a &lt;0.05</b> <b>** P &lt;0.05</b>							

#### 8.4.2 Firmas mutacionales

En relación con el análisis de supervivencia, de acuerdo con las firmas mutacionales encontradas a partir del análisis de secuenciación, se observó que la presencia de la firma SBS 10 se asoció a una menor supervivencia global de 2.9 meses (IC NA) comparada con su ausencia, que resulta en 18.5 meses (14.3-22.6)  $p=0.004$ . En el mismo sentido, la firma SBS 19 se asoció a una menor supervivencia global de 1.5 meses (IC NA) comparada con su ausencia, que alcanza los 18.6 meses

(14.5-22.6)  $p < 0.001$ . En ambas firmas no se encontraron diferencias en el análisis multivariado. En el resto de firmas y sus agrupaciones por grupos no se encontraron diferencias estadísticas significativas y tampoco tendencias (tabla 15).

Tabla 15. Análisis de supervivencia (Kaplan Meier y regresión de Cox) para supervivencia global y libre de enfermedad de firmas mutacionales							
SUPERVIVENCIA GLOBAL (SG)				SUPERVIVENCIA LIBRE DE ENFERMEDAD (SLE)			
FIRMA	MEDIANA MESES (IC 95%)	P	REGRESIÓN DE COX (IC 95%)	FIRMA	MEDIANA MESES (IC 95%)	P	REGRESIÓN DE COX (IC 95%)
SBS1	NA		NA	SBS1	NA		NA
SBS2	Ausente 17.4 (12.6-22.1) Presente 17.8 (9.5-26.1)	0.972	RR 0.08 (0.004-1.8) $p=0.119$	SBS2	NA	0.157	NA
SBS3	Ausente 17 (10.7-23.3) Presente 15.2 (11.2-19.2)	0.793	RR 0.1 (0.008-1.4) $p=0.097$	SBS3	NA		RR 66.6 (NA) $P=0.607$
SBS4	NA		RR 0.3 (0.001-169.7) $p=0.734$	SBS4	NA		NA
SBS5	Ausente 19 (14.3-23.6) Presente 11.7 (5.9-17.5)	0.385	RR 2.4 (0.3 – 18.6) $p=0.374$	SBS5	NA		NA
SBS6	Ausente 17.5 (12.7-22.2) Presente 14.1 (7.4-20.9)	0.842	RR 0.6 (0.06-5.8) $p=0.683$	SBS6	NA		NA
SBS7	Ausente 16.8 (11.8-21.9) Presente 16.3 (0-33.1)	0.655	RR 0.09 (0.003 – 2.5) $p=0.160$	SBS7	Ausente 2.7 (2.2 – 3.1) Presente 10.3 (NA)	0.225	RR 0.03 (0 – 2,243) $p=0.547$
SBS9	NA		RR 0.1 (0 – 2,161) $p=0.663$	SBS9	NA		NA
SBS10	<b>Ausente 18.5 (14.3 – 22.6)</b> <b>Presente 2.9 (NA)</b>	<b>0.004</b> <b>++</b>	NA	SBS10	NA		NA
SBS12	NA		NA	SBS12	NA		NA
SBS13	Ausente 13 (7.7 – 18.2) Presente 19.1 (14.1 – 24.2)	0.405	RR 1.5 (0.19 – 12.1) $p=0.694$	SBS13	Ausente 6.6 (0 – 13.8) Presente 2.5 (NA)	0.157	RR 66.6 (NA) $p=0.607$

<b>SBS15</b>	Ausente 17.9 (13.1 – 22.7)) Presente 13.5 (8.4 – 18.6)	0.725	RR 2.5 (0.24-27.11) p=0.433	<b>SBS15</b>	Ausente 6.6 (0 – 13.8) Presente 2.5 (NA)	0.157	RR 66.6 (NA) p=0.607
<b>SBS16</b>	NA		RR 0.2 (0.002 – 45.3) p=0.638	<b>SBS16</b>	NA		NA
<b>SBS17</b>	NA		RR 0 (NA) p=0.366	<b>SBS17</b>	NA		NA
<b>SBS19</b>	<b>Ausente 18.6 (14.5 – 22.6)</b> <b>Presente 1.5 (NA)</b>	<b>&lt;0.001 ++</b>	NA	<b>SBS19</b>	NA		NA
<b>SBS20</b>	Ausente 18.1 (13.8 – 22.4) Presente 13.3 (0 – 28.6)	0.602	RR 0.8 (0 – 7,649) p=0.975	<b>SBS20</b>	NA		NA
<b>SBS21</b>	NA		RR 3 (.001 – 169.7) p=0.734	<b>SBS21</b>	NA		NA
<b>SBS22</b>	NA		NA	<b>SBS22</b>	NA		NA
<b>SBS24</b>	Ausente 17 (12.3 – 21.7) Presente 19.8 (11 – 28.6)	0.546	RR 1 (0 – 4,500) p= 1	<b>SBS24</b>	NA		NA
<b>SBS27</b>	NA		NA	<b>SBS27</b>	NA		NA
<b>SBS29</b>	Ausente 17.3 (12.6 – 22) Presente 16.3 (NA)	0.712	RR 0.6 (0.4-9.9) p=0.757	<b>SBS29</b>	Ausente 2.7 (2.2 – 3.1) Presente 10.3 (NA)	0.225	RR 0.03 (NA) p=0.547
<b>SBS30</b>	Ausente 17.4 (12.8 – 22) Presente 18.1 (7 – 29.1)	0.786	NA	<b>SBS30</b>	NA		NA
<b>Supervivencia global y libre de enfermedad de acuerdo con el agrupamiento de los pacientes por firmas mutacionales</b>							
<b>GRUPO 1</b>	NA		RR 0 (NA) p=0.990	<b>GRUPO 1</b>	NA		NA
<b>GRUPO 2</b>	Ausente 18.2 (11.9 – 24.6) Presente 16.3 (5.7 – 27)	0.652	RR 0.9 (0.2 – 4) p=0.956	<b>GRUPO 2</b>	NA		NA
<b>GRUPO 3</b>	Ausente 18.4 (13.5-23.3) Presente 16.3 (8.7 – 23.8)	0.709	NA	<b>GRUPO 3</b>	NA		NA
<b>NA (No Alcanzado)</b> <b>RR (Riesgo relativo)</b> <b>++ p &lt; 0.05</b>							

### 8.4.3 Carga tumoral

De acuerdo con el análisis de supervivencia y carga tumoral, se encontró una tendencia a mayor supervivencia global de 18.3 meses (IC 95% 12.7 – 23.8) en los pacientes con carga tumoral alta, comparado con los pacientes con carga tumoral baja de 16.6 meses (IC 95% 10.3 - 22.9) ( $p=0.715$ ) (figura 14A). En cuanto a la supervivencia libre de enfermedad, en el grupo de carga tumoral alta se determinaron 6.4 meses (IC 95% 0 - 14) y para el grupo de carga tumoral baja, de 2.9 meses (IC 95% NA) ( $p=0.808$ ) (figura 14B). En el análisis multivariado se encontró una tendencia de menor supervivencia global en el grupo de baja carga RR 1.29 (IC 95% 0.3 – 5.1) ( $p=0.716$ ) así como mayor recurrencia RR 1.41 (IC 95% 0.08 – 23.5) ( $p=0.808$ ) (tabla 16).

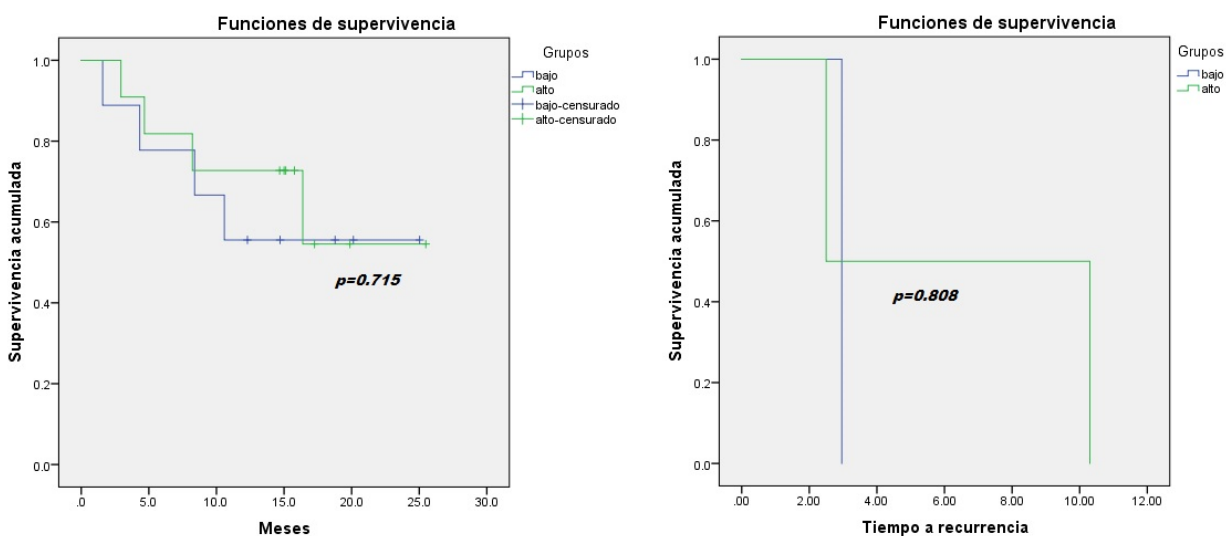


Figura 14. Análisis de supervivencia global y libre de enfermedad de acuerdo con la carga tumoral. A) Supervivencia global por carga tumoral (no significativo las curvas no se separan de forma evidente). B) Supervivencia libre de enfermedad por carga tumoral (debido a los pocos casos de recurrencia no se observa en el gráfico la separación de las curvas de forma adecuada)

Tabla 16. Análisis multivariado de Cox, para supervivencia global y libre de enfermedad en los grupos de carga tumoral					
SUPERVIVENCIA GLOBAL (SG)			SUPERVIVENCIA LIBRE DE ENFERMEDAD (SLE)		
GRUPO DE CARGA TUMORAL	REGRESIÓN DE COX (IC 95%)	P	GRUPO DE CARGA TUMORAL	REGRESIÓN DE COX (IC 95%)	P
BAJA	RR 1.29 (0.3 -5.1)	0.716	BAJA	RR 1.41 (0.08 – 23.5)	0.809
ALTA	NA	NA	ALTA	NA	NA

NA (No Alcanzado)  
RR (Riesgo Relativo)



#### 8.4.4 Factores de riesgo

En cuanto a la determinación de la supervivencia asociada a los diferentes factores de riesgo para cáncer bucal, no se encontraron diferencias significativas ( $p > 0.05$ ) entre los pacientes con hábito tabáquico, alcoholismo, ni en la presencia de VPH o de VEB (tabla 17).

Tabla 17. Análisis de supervivencia (Kaplan Meier y regresión de Cox) para supervivencia global y libre de enfermedad de factores de riesgo							
SUPERVIVENCIA GLOBAL (SG)				SUPERVIVENCIA LIBRE DE ENFERMEDAD (SLE)			
FACTOR DE RIESGO	MEDIAN A MESES (IC 95%)	P	REGRESIÓN DE COX (IC 95%)	FACTOR DE RIESGO	MEDIANA MESES (IC 95%)	P	REGRESIÓN DE COX (IC 95%)
TABAQUISMO	Ausente 17.5 (8.8-26.1) Presente 16.3 (11.2-18)	0.959	RR 0.8 (0.15-4.8) P=0.866	TABAQUISMO	Ausente 2.5 (2.5-2.5) Presente 6.6 (0-13.8)	0.157	NA
ALCOHOLISMO	Ausente 19.4 (12.9-26) Presente 16.3 (11.6-22)	0.542 ++	RR 1.6 (0.28-9.4) P=0.580++	ALCOHOLISMO	Ausente 10.3 (10.3-10.3) Presente 2.5 (2.2-3.1)	0.225	RR 30.7 (0-21.2) P=0.547
VEB	Ausente 18.5 (12.9-24) Presente 16.3 (4.1-28.6)	0.868	RR 1.1 (0.31-3.8) P=0.879	VEB	Ausente 6.6 (0-13.8) Presente 2.5 (2.5-2.5)	0.157	RR 0.1 (0-13.7) P=0.607
VPH	NA	0.638	RR 0.9 (0.22-4.1) P=0.954	VPH	NA	NA	NA
<b>NA (No Alcanzado)</b> <b>RR (Riesgo Relativo)</b> <b>++ las p más cercana a &lt;0.05</b>							

## 9. DISCUSIÓN

El cáncer bucal de células escamosas en México representa el 2% de las neoplasias malignas. Sin embargo, el 70% de los pacientes se diagnostica en etapas localmente avanzadas, lo cual dificulta el tratamiento y reduce la supervivencia a 5 años hasta en un 50%. Por ello es necesario caracterizar e identificar nuevos marcadores moleculares que optimicen el diagnóstico temprano, pronóstico y tratamiento de esta enfermedad. Por esta razón, en este trabajo se analizaron muestras de carcinoma bucal proveniente de tejido tumoral fresco, a las cuales se les realizó secuenciación

del exoma para determinar las variables patogénicas, las firmas mutacionales y la carga tumoral de los pacientes, a fin de analizar si dichas alteraciones podrían estar asociadas con la supervivencia.

Los datos demográficos de la muestra que se estudiaron coinciden con los resultados publicados en la literatura (1). Una de las características que se destaca en este estudio, es el porcentaje de desnutrición (IMC menor de 18), el cual alcanza un 35%. Todos los casos se presentaron en etapas avanzadas; la AJCC, en su última edición, lo considera como un factor pronóstico con evidencia nivel 3, por lo que se sugiere investigar más esta asociación en estudios futuros. La desnutrición en los pacientes podría deberse a los síntomas que provoca el tumor, específicamente, en las etapas clínicas avanzadas (trismus, dolor, etc.), lo que dificulta la deglución e induce la presencia de mediadores de inflamación que generan anorexia.

En cuanto a la estadificación, de acuerdo con la AJCC, comparando con diversos estudios previos (25-33), se reportaron más subsitios afectados, debido a que se tiene un caso de piso de boca y otro más de trígono retromandibular. Por otra parte, la lengua móvil fue el substio más comúnmente afectado (45% de los casos), tal como se ha reportado previamente (25-33). Los principales factores pronósticos en este cáncer son la etapa clínica y la afectación ganglionar, a diferencia de los estudios de secuenciación consultados (25-33), en los que no se especifica la etapa clínica. En este trabajo se reporta desde cáncer *in situ* (etapa clínica 0) hasta la etapa localmente avanzada irresecable IVb.

En México, el cáncer bucal se diagnostica de forma tardía (17), lo cual también se observó en este estudio, en el que la gran mayoría (85%) se encontraron en etapa avanzada (III, IVa y IVb), lo cual repercute en el pronóstico y en la calidad de vida. Es relevante mencionar que no se tuvo ninguna muestra de etapa IVc (metástasis a distancia). Por otro lado, el 45% de los pacientes presentó afección ganglionar, lo que explica, en parte, las diferencias encontradas en el pronóstico, específicamente, la baja supervivencia global (10.6 meses) y la corta supervivencia libre de enfermedad (2.7 meses) en los casos avanzados.

De los criterios histopatológicos, destaca la media de profundidad de invasión encontrada, que es de 7.3mm. Esto es importante, ya que se sabe que una profundidad de invasión de más de 4mm se asocia a mayor riesgo de metástasis ganglionares (1), y este resultado concuerda con el alto porcentaje de afección ganglionar observado (45%).

Por otro lado, respecto a los desenlaces oncológicos, la recurrencia, la persistencia y la progresión coinciden con la literatura (15, 18 y 24% respectivamente) e indican un comportamiento agresivo de la enfermedad (1). La supervivencia global en las etapas tempranas es del 100% con un solo caso de recurrencia a los 10.3 meses, presentado en una paciente que tenía un DOI de 9mm y un patrón de dispersión agresivo, y que no recibió adyuvancia con radioterapia (se encuentra viva y libre de enfermedad al momento).

Si bien el seguimiento de la mayoría de los pacientes alcanza más de un año, también se encuentran dos pacientes en los que el seguimiento no superó los dos meses, lo cual puede atribuirse al estado avanzado de la enfermedad, que no les permitió iniciar un tratamiento oncológico.

Incluso, uno de estos dos pacientes tuvo un ECOG de 3, que representa un pobre estado funcional, lo cual pone en contexto la importancia del diagnóstico temprano de esta enfermedad.

En el presente trabajo no se observó asociación entre la supervivencia y los principales factores de riesgo, específicamente, en relación con el consumo de tabaco, a pesar de que un porcentaje elevado (75%) de pacientes tiene antecedentes de tabaquismo; de los cuales, en la mitad, el consumo ha sido durante más de 10 años, lo que representa una exposición prolongada a diversos carcinogénicos. Los estudios en diferentes poblaciones han reportado desde un 30% de tabaquismo hasta un 96% en el caso de los caucásicos, los cuales no se asociaron a supervivencia (25-33). Adicionalmente, la población estudiada no consume nuez de areca o de betel, las cuales han sido consideradas como uno de los principales factores de riesgo para este tipo de neoplasia en Oriente (1, 33). Por otra parte, más de la mitad de los pacientes presentaron antecedentes de ingesta de alcohol (65%), de los cuales la mayoría fueron catalogados con un consumo moderado con más de 5 UB a la semana; solamente un estudio reportó un porcentaje mayor (89%) (32). A pesar de tener un efecto carcinogénico directo y también sinérgico con el tabaco (1, 8-15), no se encontró asociado al pronóstico en este estudio.

La prevalencia del VPH (5%) y la falta de asociación con la supervivencia en cavidad bucal coinciden con la literatura (7, 10, 16-18). El caso corresponde a una femenina de 54 años con cáncer localizado en la lengua y etapa clínica I, la cual se encuentra viva al momento. En cuanto a la prevalencia de VEB, es inferior a la reportada (35 Vs 45%) (18-19); así mismo, no se encontró que esté asociado a supervivencia.

En la búsqueda de describir las variantes patogénicas presentes en el cáncer bucal, se identificó a *P53* como el principal gen mutado en la mitad de los pacientes, lo cual coincide con la segunda vía de señalización más afectada que se encontró en este análisis. Este porcentaje es similar a lo descrito en poblaciones asiáticas (35-62%), pero menor comparado con poblaciones caucásicas, en las cuales es de hasta 80% (25-33). Si bien se sabe que *P53* es un gen que controla diversos procesos celulares relacionados a la carcinogénesis (20-22), de acuerdo con el análisis de supervivencia, su alteración no tiene impacto ni en la supervivencia global ni en la libre de enfermedad, por lo que probablemente el impacto de *P53* en el cáncer bucal esté más asociado a la iniciación y promoción de la neoplasia.

En este estudio, los genes *NOTCH1*, *RASA1* y *CAPS8*, que codifican proteínas de tipo factor de crecimiento epitelial, intervienen en el control del crecimiento y diferenciación celular mediada por *RAS* y la apoptosis, respectivamente (34). Se encontraron mutados en frecuencias similares a lo reportado en estudios caucásicos, asiáticos, árabes e indios (25-33); sin embargo, no se encontró correlación con la supervivencia. En contraste, genes como *EPHA2*, *KMT2D*, *FBXW7* y *CDKN2A* se reportaron como frecuentemente mutados (25-33), en el presente estudio representaron solo 11% para *EPHA2* y el 5% para los genes restantes; ninguno de estos mostró tener impacto en la supervivencia. Es importante destacar que no se encontró mutación en el gen *PIK3CA*, el cual se encuentra frecuentemente reportado en cáncer bucal en diferentes estudios (25-33).

Por otra parte, dentro de los genes que se encontraron frecuentemente mutados en la muestra (>11%), y que no han sido reportados en estudios previos de cáncer bucal, se encuentran *PTPRK* (inhibición de la proliferación de queratinocitos), *BIRC6* (inhibición de apoptosis), *DDX3X* (regulación de la transcripción), *KMT2C* (familia leucemia mieloide/linfoide, metilación de histonas y coactivación transcripcional), *NIN* (función en microtúbulos de las células epiteliales), *NXF1* (exportación del RNA nuclear) y *SMARCA4* (regulación de la transcripción por medio de remodelación de la cromatina asociado a *BRC1* y el síndrome de tumor rabdoide 2). La mayoría de estos genes, a excepción de *KMT2C* y *SMARCA4*, no han sido reportados en otros tipos de cáncer (34).

Contrariamente, genes que se sabe que participan directamente en el proceso carcinogénico a través de vías de proliferación celular, inhibición de apoptosis, etc., como son *APC*, *CHEK2*, *CREBBP*, *PPM1D*, *MAPK1*, *STAT3*, *TGFBR2*, se encontraron mutados solamente en 5% de los pacientes. Esto podría deberse a la participación de dichos genes en las etapas tempranas del desarrollo de cáncer, y a que la mayoría de los pacientes de este trabajo se encuentran en etapas avanzadas de la enfermedad, por lo que el perfil genético podría ser distinto al estudiado al inicio de dicha patología.

Los resultados obtenidos sugieren que existen importantes diferencias moleculares en los tumores de pacientes mexicanos comparado con series de pacientes estudiadas en distintas poblaciones, los cuales deberán estudiarse más a fondo. No obstante, las diferencias encontradas también podrían atribuirse al diseño de estudio, así como al número de muestras incluidas, por lo que será importante llevar a cabo este tipo de análisis a partir de un número mayor de pacientes.

Un hallazgo importante fue la asociación entre la variante patogénica *FAT1* y el acortamiento en la supervivencia global de los pacientes, a pesar de que este gen se encontró en tasas menores (21%), en comparación con los diferentes estudios caucásicos y asiáticos que lo han reportado (25-33%). Específicamente, de los 4 pacientes que presentaron esta mutación, 3 fallecieron y todos presentaron una etapa clínica avanzada, lo que sugiere que este gen, el cual es un supresor tumoral (34), podría estar involucrado en la progresión de la enfermedad. De hecho, coincide con la vía de señalización más frecuentemente alterada en este análisis, que es la vía Wnt/ $\beta$ catenina, en la que *FAT1* puede participar y mediar procesos de proliferación, así como la pérdida de adhesión celular por su interacción con beta-catenina (34); y a su vez, correlaciona con el proceso molecular más alterado en este análisis, que es el de unión celular (binding). En cuanto a tratamientos novedosos que actúen en la vía Wnt al momento no hay ninguno aprobado, sin embargo, hay tratamientos experimentales, 18 estudios clínicos fase 1 y 3 fase dos, en los que se tratan tumores diversos, solo uno incluye tumores de cabeza y cuello. (56) En relación con la supervivencia libre de enfermedad, solamente se encontró una tendencia a mayor recurrencia en presencia de *FBXW7*, que está involucrado en la ubiquitinación dependiente de la fosforilación actuando en la ciclina E; se ha visto

afectado en cáncer de mama y ovario (34). Sin embargo, sería importante incrementar el tamaño de la muestra para observar si la tendencia logra significancia estadística.

Otras de las vías de señalización que se encontraron alteradas en este trabajo son la angiogénesis, el receptor del EGF, la apoptosis, así como el factor de crecimiento beta y el factor de crecimiento derivado de las plaquetas, que pueden relacionarse con los estados más avanzados de la enfermedad determinados en esta muestra de estudio.

Diversos trabajos que estudian tumores de cabeza y cuello han reportado la presencia de firmas mutacionales específicas SBS1, 2, 3, 4, 5, 7, 13 y 18. En relación con los tumores de cavidad bucal, se conocen las firmas SBS 4 y 5 (35). En este análisis se encontraron representadas todas las firmas descritas previamente para el cáncer de cabeza y cuello, así como las firmas SBS 12, 16, 17, 19 y 27 que no están descritas para esta neoplasia. A pesar de esto, ninguna firma correlacionó con el pronóstico de la enfermedad.

En este contexto, se determinó que el 90% de las muestras presentó la firma SBS1 relacionada con envejecimiento, como era esperado, debido a que esta firma se encuentra presente en prácticamente todos los tumores de la edad adulta (35). En el 60% de la muestra se encontró la firma SBS 13, que junto con SBS 2 —la cual también está representada en un 25%—, se relacionan con la actividad de APOBEC debido a inflamación; pero que también destacan por ser firmas relacionadas con la presencia de VEB, específicamente en cáncer gástrico (18, 19 y 56). La mayoría de los pacientes con firma APOBEC tienen infección por VEB (4 pacientes), lo que sugiere que VEB podría estar involucrado en la carcinogénesis de estos tumores, a pesar de que su presencia no mostró diferencias en cuanto al pronóstico. (57)

El tabaquismo es otro de los factores de riesgo presentes en esta población, ya que el 75% de los pacientes fueron o son fumadores recurrentes. En relación con las firmas de tabaquismo, se encontró que el 30% de los pacientes tiene la firma SBS 5; el 15%, la firma SBS 29; y un 10%, la firma SBS 4, la cual se esperaba encontrar más representada, por ser una firma altamente reportada en el cáncer bucal (35), y esto pudiera atribuirse a la presencia de otros factores de riesgo (VEB).

Por otra parte, se encontraron distintas firmas relacionadas con los defectos de la reparación del DNA. Específicamente, la firma SBS 3 se presentó en la mitad de la muestra (55%) y está asociada con la reparación defectuosa del DNA mediado por los genes *BRCA1* y 2 (35). De manera más representada, se encontraron las firmas con afectación en la vía mismatch repair (MMR), que son SBS 6 (30%), SBS 15 (20%), SBS 20 (10%) y SBS 21 (10%); adicionalmente, la firma SBS 30 (15%), relacionada a la reparación deficiente por escisión de bases y mutación en *NTHL1*. Cabe mencionar que no se encontraron mutaciones en los genes conductores de estas vías de reparación, las cuales podrían estudiarse en el cáncer bucal.

Otras firmas que se encontraron, pero que no tienen relación con los factores de riesgo estudiados son las siguientes: SBS 7 (30%), relacionada con exposición a luz ultravioleta; SBS 24 (20%), con las aflatoxinas; SBS 9 (5%), asociada con actividad de polimerasa; SBS 10 (5%),

mutación POLE, la cual no se encontró; y SBS 22 (5%), por exposición a ácido aristolóquico de plantas.

Finalmente, otro de los aspectos a resaltar en este estudio es el análisis de la carga mutacional, en el que, si bien no se obtuvieron diferencias estadísticamente significativas, se observó una tendencia a una mejor supervivencia y menor recurrencia en el grupo de carga tumoral alta. En una revisión reciente de 1,662 muestras de múltiples tipos de cáncer, incluyendo cabeza y cuello del Memorial Sloan Kettering Cancer Center, demostraron que la carga mutacional alta tiene afectación de la vía MMR (como se demostró en las firmas mutacionales encontradas en las muestras) y este grupo produce mayor número de neoantígenos tumorales que son expuestos a las células presentadoras de antígenos, que facilitan el reconocimiento inmunológico como extraños y, con esto, el desarrollo de una respuesta inmunitaria antitumoral; lo que explicaría el mejor pronóstico de estos pacientes. (47) En cuanto a la vía MMR actualmente está aprobado el tratamiento con Pembrolizumab en pacientes con cáncer colorectal que presenten mutaciones en esta vía ya que ha demostrado mejorar el pronóstico en tumores recurrentes y metastásicos; por lo que una vía de investigación en los cánceres de cavidad bucal podría incluir este fármaco (58).

Cada muestra estudiada de diferentes poblaciones tiene un perfil molecular diferente y el perfil de pacientes mexicanos no es la excepción. Es importante señalar que, como limitantes de este estudio, para poder tener resultados contundentes en cuanto a las asociaciones con la supervivencia, se encuentran, en primer lugar, el tamaño de muestra, el cual debe incrementarse, pese a que se alcanzó significancia estadística en algunos análisis y se observaron tendencias.

Otro de los aspectos a considerar es la obtención de más muestras de etapas tempranas. Si bien se trabajó con 3 pacientes de etapas iniciales, incrementar el número de muestras de cáncer *in situ* en etapas I y II, permitiría distinguir los procesos moleculares a lo largo de la carcinogénesis. Sin embargo, es un aspecto difícil a nivel clínico, ya que la mayoría de los pacientes en México llegan en etapas avanzadas, principalmente en el INCan. Finalmente, el seguimiento de la población de estudio es otro de los aportes que deben ser considerados con la finalidad de tener precisión en cuanto a los índices de recurrencia, la supervivencia libre de enfermedad, la calidad de vida, etc.

Los datos obtenidos en este trabajo aportan evidencia de las diferencias que pueden existir en el perfil molecular de los pacientes de una población específica. Por lo tanto, es necesario realizar más estudios que permitan consolidar los hallazgos de este estudio, con la finalidad de encontrar marcadores moleculares que tengan una utilidad clínica en el diagnóstico, tratamiento o pronóstico. Además, apoyan fuertemente la importancia de continuar con la línea de investigación en diferentes áreas como son el análisis de factores de riesgo con atención especial en VEB y factores que producen inflamación en la cavidad oral, nuevos tratamientos dirigidos a las vías afectadas (vía Wnt, P53, MMR y APOBEC) y biomarcadores para el pronóstico (carga tumoral).

## 10. CONCLUSIONES

1. Este estudio representa la primera secuenciación de una muestra de pacientes mexicanos y latina. Es el primero que incorpora otros subsitios anatómicos y reporta las etapas clínicas (excepto metastásica).

2. Los factores de riesgo conocidos para cáncer bucal, específicamente tabaco y alcohol, no son un mecanismo asociado a la supervivencia de los pacientes con cáncer bucal.

3. Se detectó la presencia de VPH en 5% de la población y de un 35% de infección por VEB. Sin embargo, ninguno de los dos patógenos se asoció con el pronóstico.

4. Se encontraron diferencias en cuanto a los genes mutados previamente, reportados en estudios de otras poblaciones. Las más importantes fueron el porcentaje menor de afectación de los genes como *P53* (47%), *CDKN2A* y *CASP8* y la ausencia de mutaciones en *PIK3CA*.

5. Se determinaron variantes patogénicas como *PTPRK*, *BIRC6*, *DDX3X*, *KMT2C*, *NIN*, *NXF1*, *SMARCA4*, *ABCB1*, *ADGRB1*, *APC*, *ARHGAP35*, *ARHGEF10L*, *ARID1A*, *ATRX*, *BMP2*, *CASP8*, *CDKN2A*, *CEBPA*, *CHEK2*, *CUP1*, *CREBBP*, *DCC*, *DOT1L*, *EFTUD2*, *EML4*, *EPHA3*, *FANCA*, *FAS*, *FAT4*, *FLT3*, *GRIN2A*, *HGF*, *HNF1A*, *KEAP1*, *KMT2A*, *KMT2B*, *LATS1*, *LPAR4*, *MAML2*, *MAPK1*, *MYH11*, *NCOR1*, *NCOR2*, *NIPBL1*, *NTRK1*, *NTRK3*, *PBRM1*, *PML*, *PPP3CA*, *PPP6C*, *PREX2*, *PRF1*, *RASA2*, *RBM15*, *RGPD3*, *RNF213*, *RPS6KA3*, *SETBP1*, *SETD2*, *SMARCA1*, *SPEN*, *STAT3*, *STK11*, *TBX3*, *TET1*, *TNC*, *TRIM33*, *U2AF2*, *USP8*, *WNK2*, *ZFX* y *ZNF521*, que no han sido reportadas en otras poblaciones.

6. Las vías que están afectadas en la mayoría de los pacientes son la vía de señalización Wnt/ $\beta$ catenina y la vía *P53*, que coinciden con las funciones de los genes *FAT1* y *P53* respectivamente y que se encuentran mutados.

7. Existe una asociación entre la mutación de *FAT1* y una menor supervivencia global.

8. Las firmas mutacionales encontradas son principalmente las vías asociadas a envejecimiento (SBS1), vía de inflamación vía APOBEC (SBS 13 y 2), tabaquismo (SBS 5, 4 y 29) y MMR (SBS 6, 15, 20 y 21).

9. La mayoría de los pacientes que presentan las firmas relacionadas a APOBEC presentan infección por VEB.

10. Se encontró una tendencia a mejor supervivencia global y libre de enfermedad en pacientes con carga tumoral alta.

## 11. APORTACIONES

- Se evidencia que una muestra de pacientes con carcinoma de cavidad bucal en población mexicana, tiene un perfil mutacional diferente al reportado para otros grupos de pacientes en el mundo.
- Permite continuar la línea de investigación en las siguientes áreas: factores de riesgo, diagnóstico, tratamiento dirigido y factores pronósticos.

## 12. PERSPECTIVAS

Con los resultados descritos, surgen nuevas incógnitas. A continuación, se enumeran algunos puntos que sería conveniente tomar en cuenta para la realización de futuros estudios:

- a) Aumentar el tamaño de muestra a un mínimo de 30 (por teorema de límite central), así mismo se requiere aumentar el seguimiento para realizar la validación de los resultados obtenidos.
- b) Incluir más subsitios afectados (mucosa oral y paladar duro).
- c) Incluir pacientes en etapa clínica IVc (metástasis a distancia) y aumentar los casos en etapas tempranas.
- d) Analizar en las firmas mutacionales las inserciones–delecciones y sustituciones dobles.
- e) Investigación sobre factores de riesgo no descritos como infección por VEB, infecciones periodontales, uso de prótesis dentales y factores que provoquen inflamación.
- f) Investigación de nuevos tratamientos sistémicos que actúen sobre las vías afectadas por los genes mutados encontrados (vía Wnt, P53, MMR y APOBEC). Como se describió en la sección de discusión la única vía que tiene un medicamento aprobado con impacto en la supervivencia es para la vía MMR y el uso de Pembrolizumab (en pacientes con cáncer colorectal con tumores recurrentes y metastásicos). (55, 57)
- g) Investigación y validación de biomarcadores que ayuden al pronóstico y a elegir un tratamiento adyuvante (carga mutacional, mutaciones en *FAT1*, firmas mutacionales)



### 13. BIBLIOGRAFÍA

- (1) Montero PH, Patel SG. Cancer of the oral cavity. *Surg Oncol Clin N Am*. 2015 Jul;24(3):491-508. doi: 10.1016/j.soc.2015.03.006. Epub 2015 Apr 15. PMID: 25979396; PMCID: PMC5018209.
- (2) Globocan 2018. <https://gco.iarc.fr/>
- (3) Lambert R, Sauvaget C, de Camargo Cancela M, Sankaranarayanan R. Epidemiology of cancer from the oral cavity and oropharynx. *Eur J Gastroenterol Hepatol*. 2011 Aug;23(8):633-41. doi: 10.1097/MEG.0b013e3283484795. PMID: 21654320.
- (4) Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. *CA Cancer J Clin*. 2011 Mar-Apr;61(2):69-90. doi: 10.3322/caac.20107. Epub 2011 Feb 4. Erratum in: *CA Cancer J Clin*. 2011 Mar-Apr;61(2):134. PMID: 21296855.
- (5) Aksu G, Karadeniz A, Saynak M, Fayda M, Kadehçi Z, Kocaelli H. Treatment results and prognostic factors in oral tongue cancer: analysis of 80 patients. *Int J Oral Maxillofac Surg*. 2006 Jun;35(6):506-13. doi: 10.1016/j.ijom.2006.01.006. Epub 2006 Feb 28. PMID: 16503396.
- (6) Rogers SN, Brown JS, Woolgar JA, Lowe D, Magennis P, Shaw RJ, Sutton D, Errington D, Vaughan D. Survival following primary surgery for oral cancer. *Oral Oncol*. 2009 Mar;45(3):201-11. doi: 10.1016/j.oraloncology.2008.05.008. Epub 2008 Jul 31. PMID: 18674959.
- (7) Miller CS, Johnstone BM. Human papillomavirus as a risk factor for oral squamous cell carcinoma: a meta-analysis, 1982-1997. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*. 2001 Jun;91(6):622-35. doi: 10.1067/moe.2001.115392. PMID: 11402272.
- (8) Canova C, Richiardi L, Merletti F, Pentenero M, Gervasio C, Tanturri G, Garzino-Demo P, Pecorari G, Talamini R, Barzan L, Sulfaro S, Franchini G, Muzzolini C, Bordin S, Pugliese GN, Macrì E, Simonato L. Alcohol, tobacco and genetic susceptibility in relation to cancers of the upper aerodigestive tract in northern Italy. *Tumori*. 2010 Jan-Feb;96(1):1-10. PMID: 20437850.
- (9) Gillison ML, Zhang Q, Jordan R, Xiao W, Westra WH, Trotti A, Spencer S, Harris J, Chung CH, Ang KK. Tobacco smoking and increased risk of death and progression for patients with p16-positive and p16-negative oropharyngeal cancer. *J Clin Oncol*. 2012 Jun 10;30(17):2102-11. doi: 10.1200/JCO.2011.38.4099. Epub 2012 May 7. PMID: 22565003; PMCID: PMC3397696.
- (10) Ribeiro KB, Levi JE, Pawlita M, Koifman S, Matos E, Eluf-Neto J, Wunsch-Filho V, Curado MP, Shangina O, Zaridze D, Szeszenia-Dabrowska N, Lissowska J, Daudt A, Menezes A, Bencko V, Mates D, Fernandez L, Fabianova E, Gheit T, Tommasino M, Boffetta P, Brennan P, Waterboer T. Low human papillomavirus prevalence in head and neck cancer: results from two large case-control studies in high-incidence regions. *Int J Epidemiol*. 2011 Apr;40(2):489-502. doi: 10.1093/ije/dyq249. Epub 2011 Jan 11. PMID: 21224273.
- (11) Kulkarni SS, Kulkarni SS, Vastrad PP, Kulkarni BB, Markande AR, Kadakol GS, Hiremath SV, Kaliwal S, Patil BR, Gai PB. Prevalence and distribution of high risk human papillomavirus (HPV) Types 16 and 18 in Carcinoma of cervix, saliva of patients with oral squamous cell carcinoma and in

the general population in Karnataka, India. *Asian Pac J Cancer Prev.* 2011;12(3):645-8. PMID: 21627358.

(12) Hashibe M, Sturgis EM. Epidemiology of oral-cavity and oropharyngeal carcinomas: controlling a tobacco epidemic while a human papillomavirus epidemic emerges. *Otolaryngol Clin North Am.* 2013 Aug;46(4):507-20. doi: 10.1016/j.otc.2013.05.001. Epub 2013 Jun 25. PMID: 23910467.

(13) Chang JS, Hsiao JR, Chen CH. ALDH2 polymorphism and alcohol-related cancers in Asians: a public health perspective. *J Biomed Sci.* 2017 Mar 3;24(1):19. doi: 10.1186/s12929-017-0327-y. PMID: 28253921; PMCID: PMC5335829.

(14) Chi AC, Day TA, Neville BW. Oral cavity and oropharyngeal squamous cell carcinoma--an update. *CA Cancer J Clin.* 2015 Sep-Oct;65(5):401-21. doi: 10.3322/caac.21293. Epub 2015 Jul 27. PMID: 26215712.

(15) Radoi L, Luce D. A review of risk factors for oral cavity cancer: the importance of a standardized case definition. *Community Dent Oral Epidemiol.* 2013 Apr;41(2):97-109, e78-91. doi: 10.1111/j.1600-0528.2012.00710.x. Epub 2012 Aug 11. PMID: 22882534.

(16) Gillison ML, Broutian T, Pickard RK, Tong ZY, Xiao W, Kahle L, Graubard BI, Chaturvedi AK. Prevalence of oral HPV infection in the United States, 2009-2010. *JAMA.* 2012 Feb 15;307(7):693-703. doi: 10.1001/jama.2012.101. Epub 2012 Jan 26. PMID: 22282321; PMCID: PMC5790188.

(17) González-Ramírez I, Irigoyen-Camacho ME, Ramírez-Amador V, Lizano-Soberón M, Carrillo-García A, García-Carrancá A, Sánchez-Pérez Y, Méndez-Martínez R, Granados-García M, Ruiz-Godoy L, García-Cuellar C. Association between age and high-risk human papilloma virus in Mexican oral cancer patients. *Oral Dis.* 2013 Nov;19(8):796-804. doi: 10.1111/odi.12071. Epub 2013 Feb 4. PMID: 23379359.

(18) Núñez-Acurio D, Bravo D, Aguayo F. Epstein-Barr Virus-Oral Bacterial Link in the Development of Oral Squamous Cell Carcinoma. *Pathogens.* 2020 Dec 18;9(12):1059. doi: 10.3390/pathogens9121059. PMID: 33352891; PMCID: PMC7765927.

(19) Drop B, Strycharz-Dudziak M, Kliszczewska E, Polz-Dacewicz M. Coinfection with Epstein-Barr Virus (EBV), Human Papilloma Virus (HPV) and PolyomaBK Virus (BKPyV) in Laryngeal, Oropharyngeal and Oral Cavity Cancer. *Int J Mol Sci.* 2017 Dec 19;18(12):2752. doi: 10.3390/ijms18122752. PMID: 29257122; PMCID: PMC5751351.

(20) Williams HK. Molecular pathogenesis of oral squamous carcinoma. *Mol Pathol.* 2000 Aug;53(4):165-72. doi: 10.1136/mp.53.4.165. PMID: 11040937; PMCID: PMC1186964.

(21) Esteller M. Epigenetics in cancer. *N Engl J Med.* 2008 Mar 13;358(11):1148-59. doi: 10.1056/NEJMra072067. PMID: 18337604.

(22) Little MP. Cancer models, genomic instability and somatic cellular Darwinian evolution. *Biol Direct.* 2010 Apr 20;5:19; discussion 19. doi: 10.1186/1745-6150-5-19. PMID: 20406436; PMCID: PMC2873266.

- (23) Ma D, Jiang YZ, Liu XY, Liu YR, Shao ZM. Clinical and molecular relevance of mutant-allele tumor heterogeneity in breast cancer. *Breast Cancer Res Treat.* 2017 Feb;162(1):39-48. doi: 10.1007/s10549-017-4113-z. Epub 2017 Jan 16. PMID: 28093659.
- (24) Campbell JD, Lathan C, Sholl L, Ducar M, Vega M, Sunkavalli A, Lin L, Hanna M, Schubert L, Thorner A, Faris N, Williams DR, Osarogiagbon RU, van Hummelen P, Meyerson M, MacConaill L. Comparison of Prevalence and Types of Mutations in Lung Cancers Among Black and White Populations. *JAMA Oncol.* 2017 Jun 1;3(6):801-809. doi: 10.1001/jamaoncol.2016.6108. PMID: 28114446; PMCID: PMC5464986.
- (25) Agrawal N, Frederick MJ, Pickering CR, Bettgowda C, Chang K, Li RJ, Fakhry C, Xie TX, Zhang J, Wang J, Zhang N, El-Naggar AK, Jasser SA, Weinstein JN, Treviño L, Drummond JA, Muzny DM, Wu Y, Wood LD, Hruban RH, Westra WH, Koch WM, Califano JA, Gibbs RA, Sidransky D, Vogelstein B, Velculescu VE, Papadopoulos N, Wheeler DA, Kinzler KW, Myers JN. Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in NOTCH1. *Science.* 2011 Aug 26;333(6046):1154-7. doi: 10.1126/science.1206923. Epub 2011 Jul 28. PMID: 21798897; PMCID: PMC3162986.
- (26) Stransky N, Egloff AM, Tward AD, Kostic AD, Cibulskis K, Sivachenko A, Kryukov GV, Lawrence MS, Sougnez C, McKenna A, Shefler E, Ramos AH, Stojanov P, Carter SL, Voet D, Cortés ML, Auclair D, Berger MF, Saksena G, Guiducci C, Onofrio RC, Parkin M, Romkes M, Weissfeld JL, Seethala RR, Wang L, Rangel-Escareño C, Fernandez-Lopez JC, Hidalgo-Miranda A, Melendez-Zajgla J, Winckler W, Ardlie K, Gabriel SB, Meyerson M, Lander ES, Getz G, Golub TR, Garraway LA, Grandis JR. The mutational landscape of head and neck squamous cell carcinoma. *Science.* 2011 Aug 26;333(6046):1157-60. doi: 10.1126/science.1208130. Epub 2011 Jul 28. PMID: 21798893; PMCID: PMC3415217.
- (27) Lui VW, Hedberg ML, Li H, Vangara BS, Pendleton K, Zeng Y, Lu Y, Zhang Q, Du Y, Gilbert BR, Freilino M, Sauerwein S, Peyser ND, Xiao D, Diergaarde B, Wang L, Chiosea S, Seethala R, Johnson JT, Kim S, Duvvuri U, Ferris RL, Romkes M, Nukui T, Kwok-Shing Ng P, Garraway LA, Hammerman PS, Mills GB, Grandis JR. Frequent mutation of the PI3K pathway in head and neck cancer defines predictive biomarkers. *Cancer Discov.* 2013 Jul;3(7):761-9. doi: 10.1158/2159-8290.CD-13-0103. Epub 2013 Apr 25. PMID: 23619167; PMCID: PMC3710532.
- (28) Pickering CR, Zhang J, Yoo SY, Bengtsson L, Moorthy S, Neskey DM, Zhao M, Ortega Alves MV, Chang K, Drummond J, Cortez E, Xie TX, Zhang D, Chung W, Issa JP, Zweidler-McKay PA, Wu X, El-Naggar AK, Weinstein JN, Wang J, Muzny DM, Gibbs RA, Wheeler DA, Myers JN, Frederick MJ. Integrative genomic characterization of oral squamous cell carcinoma identifies frequent somatic drivers. *Cancer Discov.* 2013 Jul;3(7):770-81. doi: 10.1158/2159-8290.CD-12-0537. Epub 2013 Apr 25. PMID: 23619168; PMCID: PMC3858325.
- (29) India Project Team of the International Cancer Genome Consortium. Mutational landscape of gingivo-buccal oral squamous cell carcinoma reveals new recurrently-mutated genes and molecular

subgroups. *Nat Commun.* 2013;4:2873. doi: 10.1038/ncomms3873. PMID: 24292195; PMCID: PMC3863896.

(30) Cancer Genome Atlas Network. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature.* 2015 Jan 29;517(7536):576-82. doi: 10.1038/nature14129. PMID: 25631445; PMCID: PMC4311405.

(31) Al-Hebshi NN, Li S, Nasher AT, El-Setouhy M, Alsanosi R, Blancato J, Loffredo C. Exome sequencing of oral squamous cell carcinoma in users of Arabian snuff reveals novel candidates for driver genes. *Int J Cancer.* 2016 Jul 15;139(2):363-72. doi: 10.1002/ijc.30068. Epub 2016 Mar 18. PMID: 26934577; PMCID: PMC5821466.

(32) Su SC, Lin CW, Liu YF, Fan WL, Chen MK, Yu CP, Yang WE, Su CW, Chuang CY, Li WH, Chung WH, Yang SF. Exome Sequencing of Oral Squamous Cell Carcinoma Reveals Molecular Subgroups and Novel Therapeutic Opportunities. *Theranostics.* 2017 Feb 26;7(5):1088-1099. doi: 10.7150/thno.18551. PMID: 28435450; PMCID: PMC5399578.

(33) Chai AWY, Lim KP, Cheong SC. Translational genomics and recent advances in oral squamous cell carcinoma. *Semin Cancer Biol.* 2020 Apr;61:71-83. doi: 10.1016/j.semcancer.2019.09.011. Epub 2019 Sep 19. PMID: 31542510.

(34) <https://www.ncbi.nlm.nih.gov/gene>

(35) AlexandrovLB, Kim J, Haradhvala NJ, Huang MN, Tian Ng AW, Wu Y, Boot A, Covington KR, Gordenin DA, Bergstrom EN, Islam SMA, Lopez-Bigas N, KlimczakLJ, McPherson JR, Morganello S, Sabarinathan R, Wheeler DA, Mustonen V; PCAWG Mutational Signatures Working Group, Getz G, RozenSG, Stratton MR; PCAWG Consortium. The repertoire of mutational signatures in human cancer. *Nature.* 2020 Feb;578(7793):94-101. doi: 10.1038/s41586-020-1943-3. Epub 2020 Feb 5. PMID: 32025018; PMCID: PMC7054213.

(36) Beggan C, Fives C, O'Leary G, Sheahan P, Heffron CC, Feeley L. Pattern of invasion and lymphovascular invasion in squamous cell carcinoma of the floor of the mouth: an interobserver variability study. *Histopathology.* 2016 Dec;69(6):914-920. doi: 10.1111/his.13014. Epub 2016 Aug 30. PMID: 27271979.

(37) AJCC Cancer Staging Manual (8th edition). Springer International Publishing: American Joint Commission on Cancer; 2017 [cited 2016 Dec 28].

(38) Surveillance, Epidemiology, and End Results Program. <https://seer.cancer.gov/statfacts/html/oralcav.html>

(39) Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. Available online at: <http://www.bioinformatics.babraham.ac.u>

(40) Ewels P, Magnusson M, Lundin S, Källér M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics.* 2016 Oct 1;32(19):3047-8. doi: 10.1093/bioinformatics/btw354. Epub 2016 Jun 16. PMID: 27312411; PMCID: PMC5039924.

- (41) Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009 Jul 15;25(14):1754-60. doi: 10.1093/bioinformatics/btp324. Epub 2009 May 18. PMID: 19451168; PMCID: PMC2705234.
- (42) Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, Jordan T, Shakir K, Roazen D, Thibault J, Banks E, Garimella KV, Altshuler D, Gabriel S, DePristo MA. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *CurrProtoc Bioinformatics*. 2013;43(1110):11.10.1-11.10.33. doi: 10.1002/0471250953.bi1110s43. PMID: 25431634; PMCID: PMC4243306.
- (43) Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010 Sep;38(16):e164. doi: 10.1093/nar/gkq603. Epub 2010 Jul 3. PMID: 20601685; PMCID: PMC2938201.
- (44) Gonzalez-Perez, A., Perez-Llamas, C., Deu-Pons, J., Tamborero, D., Schroeder, M. P., Jene-Sanz, A., Santos, A., & Lopez-Bigas, N. (2013). IntOGen-mutations identifies cancer drivers across tumor types. *Nature Methods*, 10(11), 1081–1084. <https://doi.org/10.1038/nmeth.2642>
- (45) Landrum MJ, Lee JM, Benson M, Brown GR, Chao C, Chitipiralla S, Gu B, Hart J, Hoffman D, Jang W, Karapetyan K, Katz K, Liu C, Maddipatla Z, Malheiro A, McDaniel K, Ovetsky M, Riley G, Zhou G, Holmes JB, Kattman BL, Maglott DR. ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res*. 2018 Jan 4;46(D1):D1062-D1067. doi: 10.1093/nar/gkx1153. PMID: 29165669; PMCID: PMC5753237.
- (46) Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alföldi, J., Wang, Q., Collins, R. L., Laricchia, K. M., Ganna, A., Birnbaum, D. P., Gauthier, L. D., Brand, H., Solomonson, M., Watts, N. A., Rhodes, D., Singer-Berk, M., England, E. M., Seaby, E. G., Kosmicki, J. A., ... MacArthur, D. G. (2019). Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. *BioRxiv*, 531210. <https://doi.org/10.1101/531210>
- (47) Exome Variant Server.(n.d.). Retrieved January 28, 2020, from <https://evs.gs.washington.edu/EVS/>
- (48) Adzhubei IA, Schmidt S, Peshkin L, RamenskyVE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. A method and server for predicting damaging missense mutations. *Nat Methods*. 2010 Apr;7(4):248-9. doi: 10.1038/nmeth0410-248. PMID: 20354512; PMCID: PMC2855889.
- (49) Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res*. 2003 Jul 1;31(13):3812-4. doi: 10.1093/nar/gkg509. PMID: 12824425; PMCID: PMC168916.
- (50) Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: mutation prediction for the deep-sequencing age. *Nat Methods*. 2014 Apr;11(4):361-2. doi: 10.1038/nmeth.2890. PMID: 24681721.
- (51) SamsteinRM, Lee CH, Shoushtari AN, Hellmann MD, Shen R, JanjigianYY, Barron DA, Zehir A, Jordan EJ, Omuro A, KaleyTJ, Kendall SM, Motzer RJ, Hakimi AA, Voss MH, Russo P, Rosenberg

- J, Iyer G, BochnerBH, Bajorin DF, Al-Ahmadie HA, Chaff JE, Rudin CM, RielyGJ, Baxi S, Ho AL, Wong RJ, Pfister DG, Wolchok JD, Barker CA, Gutin PH, Brennan CW, Tabar V, MellingeroffIK, DeAngelis LM, Ariyan CE, Lee N, Tap WD, Gounder MM, D'AngeloSP, Saltz L, StadlerZK, Scher HI, Baselga J, Razavi P, Klebanoff CA, Yaeger R, Segal NH, Ku GY, DeMatteo RP, Ladanyi M, Rizvi NA, Berger MF, Riaz N, Solit DB, Chan TA, Morris LGT. Tumor mutational load predicts survival after immunotherapy across multiple cancer types. *Nat Genet.* 2019 Feb;51(2):202-206. doi: 10.1038/s41588-018-0312-8. Epub 2019 Jan 14. PMID: 30643254; PMCID: PMC6365097.
- (52) Rosenthal, R., McGranahan, N., Herrero, J., Taylor, B. S., & Swanton, C. (2016). deconstructSigs: delineating mutational processes in single tumors distinguishes DNA repair deficiencies and patterns of carcinoma evolution. *Genome Biology*, 17(1), 31. <https://doi.org/10.1186/s13059-016-0893-4>
- (53) Bergstrom, E. N., Barnes, M., Martincorena, I., & Alexandrov, L. B. (2020). Generating realistic null hypothesis of cancer mutational landscapes using SigProfilerSimulator. *BioRxiv*, 2020.02.13.948422. <https://doi.org/10.1101/2020.02.13.948422>
- (54) Mi H, Ebert D, Muruganujan A, Mills C, Albou LP, Mushayamaha T, Thomas PD. PANTHER version 16: a revised family classification, tree-based classification tool, enhancer regions and extensive API. *Nucleic Acids Res.* 2021 Jan 8;49(D1):D394-D403. doi: 10.1093/nar/gkaa1106. PMID: 33290554; PMCID: PMC7778891.
- (55) de Roda Husman AM, Walboomers JM, van den Brule AJ, Meijer CJ, Snijders PJ. The use of general primers GP5 and GP6 elongated at their 3' ends with adjacent highly conserved sequences improves human papillomavirus detection by PCR. *J Gen Virol.* 1995 Apr;76 ( Pt 4):1057-62. doi: 10.1099/0022-1317-76-4-1057. PMID: 9049358.
- (56) Jung YS, Park JI. Wnt signaling in cancer: therapeutic targeting of Wnt signaling beyond  $\beta$ -catenin and the destruction complex. *Exp Mol Med.* 2020 Feb;52(2):183-191. doi: 10.1038/s12276-020-0380-6. Epub 2020 Feb 10. PMID: 32037398; PMCID: PMC7062731.
- (57) Bobrovnitchaia I, Valieris R, Drummond RD, Lima JP, Freitas HC, Bartelli TF, de Amorim MG, Nunes DN, Dias-Neto E, da Silva IT. APOBEC-mediated DNA alterations: A possible new mechanism of carcinogenesis in EBV-positive gastric cancer. *Int J Cancer.* 2020 Jan 1;146(1):181-191. doi: 10.1002/ijc.32411. Epub 2019 Jun 7. PMID: 310900
- (58) André T, Shiu KK, Kim TW, Jensen BV, Jensen LH, Punt C, Smith D, Garcia-Carbonero R, Benavides M, Gibbs P, de la Fouchardiere C, Rivera F, Elez E, Bendell J, Le DT, Yoshino T, Van Cutsem E, Yang P, Farooqui MZH, Marinello P, Diaz LA Jr; KEYNOTE-177 Investigators. Pembrolizumab in Microsatellite-Instability-High Advanced Colorectal Cancer. *N Engl J Med.* 2020 Dec 3;383(23):2207-2218. doi: 10.1056/NEJMoa2017699. PMID: 33264544.