

CARACTERIZACIÓN DE VARIANTES GENÉTICAS DE LÍNEA GERMINAL EN EL
EXOMA DE PACIENTES DIAGNOSTICADOS CON ADENOCARCINOMA
GÁSTRICO EN EL DEPARTAMENTO DE NARIÑO

Presentado por:

STEPHANY CAROLINA ROSERO ROJAS

UNIVERSIDAD DE NARIÑO
FACULTAD DE CIENCIAS EXACTAS Y NATURALES
DEPARTAMENTO DE BIOLOGÍA
SAN JUAN DE PASTO

2020

CARACTERIZACIÓN DE VARIANTES GENÉTICAS DE LÍNEA GERMINAL EN EL
EXOMA DE PACIENTES DIAGNOSTICADOS CON ADENOCARCINOMA GÁSTRICO
EN EL DEPARTAMENTO DE NARIÑO

Presentado por:

STEPHANY CAROLINA ROSERO ROJAS

Directoras:

CAROL YOVANNA ROSERO GALINDO M.Sc. Ph.D.
UNIVERSIDAD COOPERATIVA DE COLOMBIA

EDITH MARIELA BURBANO ROSERO M.Sc. Ph.D.
UNIVERSIDAD DE NARIÑO

Trabajo de Grado presentado como requisito parcial para optar al título de Magister en Ciencias
Biológicas

Modalidad Trabajo de Investigación

UNIVERSIDAD DE NARIÑO
FACULTAD DE CIENCIAS EXACTAS Y NATURALES
DEPARTAMENTO DE BIOLOGÍA
SAN JUAN DE PASTO

2020

NOTA DE RESPONSABILIDAD

“Las ideas y conclusiones aportadas en este trabajo de grado, son responsabilidades exclusivas de los autores”

Artículo primero del acuerdo No. 324 de octubre 11 de 1966, emanado por el Honorable Consejo Superior de la Universidad de Nariño.

NOTA DE ACEPTACIÓN

Carol Yovanna Rosero Galindo, PhD.
Firma de la Directora del Trabajo de Grado

Edith Mariela Burbano Rosero, PhD.
Firma de la Directora del Trabajo de Grado

Diana Carolina Muñoz Lasso, PhD.
Firma del Jurado

Luz Estela Lagos Mora, MSc.
Firma del Jurado

Mauricio Corredor Rodríguez, PhD.
Firma del Jurado

San Juan de Pasto, 30 de septiembre de 2020

DEDICATORIA

Dedico este trabajo a mi familia, el motor de mi vida y mi todo; gracias por confiar en mí, su infinito amor me ha dado la valentía para continuar día a día. En especial a mi hermano Junior, cambiaste la percepción de mi existencia, te seguiré honrando en la vida con mis logros, sé que te sentirías bastante orgulloso de mí.

AGRADECIMIENTOS

A Dios, mi guía y mi luz en cada instante.

A mis padres y a mi hermana, con su apoyo las metas son más sencillas de alcanzar, hacen mis sueños realidad, admiro su tenacidad y ganas de verme triunfar.

A mi directora Carol Rosero, por enseñarme que existen diferentes maneras de llegar a una meta y creer en mi misma fue una de ellas. La oportunidad que me brindo para hacer parte de este proyecto contribuyó significativamente a la culminación de mi logro. Gracias profe por su amistad, sé que su paciencia, palabras, consejos y regaños fueron bastantes oportunos, su sabiduría siempre forja mi formación académica.

A mi codirectora Mariela Burbano, sus aportes y sugerencias, fueron claves para lograr un buen desempeño. Gracias por su atención, apoyo e infinita paciencia, sus palabras de ánimo en momentos difíciles siempre me impulsaron a continuar con mi trabajo

A Liz Mejía, nada de esto sería posible sin ti, infinitas gracias por enseñare el mundo de la bioinformática, por la orientación que me brindaste, fuiste mi guía en las largas jornadas de trabajo.

A los jurados Diana Carolina Muñoz, Luz Estela Lagos y Mauricio Corredor, por su oportuna atención, apoyo y aportes en el desarrollo del trabajo de grado.

Al Dr Marino Coral, por su apoyo y su experiencia como gastroenterólogo en la ejecución del proyecto.

A los pacientes que hicieron parte de esta investigación, quienes, con su actitud, y su colaboración desinteresada hicieron que fuera posible el desarrollo de la misma.

A mis profesores, que con sus enseñanzas forjaron cada día mi formación académica. Al profesor Jhon Jairo Calderón, Guillermo Castillo y en especial a Luz Estela Lagos, gracias por apoyarme en los momentos difíciles, por los ánimos y las palabras para salir adelante.

A mis compañeros de la maestría, gracias por su amistad y compañía, hicieron que las arduas jornadas de estudio fueran más llevaderas, están siempre en mi corazón.

A Esteban Durango, que la vida nos lleve de la mano alcanzando más logros juntos, tu amor me llena de energía para continuar adelante.

A Mario Eraso, tu valiosa amistad y ayuda siempre aportan a mis logros.

A la Universidad Cooperativa de Colombia, por la financiación de esta investigación enmarcada en el proyecto “Perfil genético de cáncer gástrico en Nariño: Estudio de casos en línea somática y germinal”. Y permitirme procesar las muestras en el laboratorio de Biología Molecular. En general a todas las personas que me colaboraron en las actividades realizadas.

RESUMEN

Introducción: El cáncer gástrico (CG) es un problema de salud pública a nivel mundial. En Colombia la incidencia de la enfermedad es alta, particularmente en la zona Andina del departamento de Nariño. La etiología del CG esporádico se explica por múltiples factores, incluido el componente genético que en su mayoría abarca variantes somáticas que explican el riesgo al desarrollo de la patología, no obstante, existe una menor proporción de mutaciones germinales que podrían contribuir a la carcinogénesis gástrica.

Objetivo: Caracterizar variantes genéticas de línea germinal en el exoma de pacientes diagnosticados con CG de tipo esporádico en Nariño.

Métodos: Secuenciación de exomas a partir de cuatro muestras de pacientes con CG. Se realizó el análisis bioinformático que incluyó: control de calidad de las lecturas, alineación con el genoma de referencia hg19, llamado, anotación y priorización de variantes deletéreas de línea germinal. Se realizaron redes de interacción y un análisis funcional de vías de los genes con variantes perjudiciales.

Resultados: Se identificaron 1.597.012 SNPs y 203.133 INDELs. De estas variantes genéticas, 52 SNPs y 12 inserciones se reportaron como raras y deletéreas, las cuales se presentaron en 62 genes, como: *APC*, *CTBP*, *LAMB3*, *MET*, *PTGER3*, *TCFLT2*, *CYP2E1*, *MAPK1*, *MSH3*, entre otros. Se reportaron 5 SNPs con cierto grado de patogenicidad en los genes: *ACADS* (g.121175696) y *BRCA1*(g.41249297) en la muestra SP03 y, *FLNC* (g.128498402), *MCCC2* (g.70936895) y *ACSS2* (g. 33509608) en la muestra S07.

Conclusión: Las variantes genéticas raras y poco frecuentes podrían estar relacionadas con el riesgo del desarrollo de la enfermedad, siendo necesaria una validación posterior de estos resultados.

PALABRAS CLAVES: Neoplasias, línea germinal, exoma, variantes genéticas.

ABSTRACT

Introduction: Gastric cancer (GC) is a public health problem worldwide. In Colombia the incidence of the disease is high, particularly in the Andean area of the department of Nariño. The etiology of sporadic GC is explained by multiple factors, including the genetic component that mostly encompasses somatic variants that explain the risk for the development of the pathology; however, exists a lower proportion of germinal mutations that could contribute to gastric carcinogenesis.

Objective: To characterize germinal genetic variants in the exome of patients diagnosed with sporadic GC in Nariño.

Methods: Exome sequencing from four samples of GC patients. The bioinformatics analysis was carried out that included: quality control of the readings, alignment with the hg19 reference genome, calling, annotation and prioritization of deleterious germinal variants. Interaction networks and a functional analysis of pathways of genes with deleterious variants were performed.

Results: 1,597,012 SNPs 203,133 INDELs were identified. Of these genetic variants, 52 SNPs and 12 insertions were reported as rare and deleterious, which occurred in 62 genes, such as: *APC*, *CTBP*, *LAMB3*, *MET*, *PTGER3*, *TCFLT2*, *CYP2E1*, *MAPK1*, *MSH3*, among others. 5 SNPs with a certain degree of pathogenicity were reported in the genes: *ACADS* (g.121175696) and *BRCA1* (g.41249297) in the SP03 sample and, *FLNC* (g.128498402), *MCCC2* (g.70936895) and *ACSS2* (g. 33509608) in sample S07.

Conclusion: The rare and infrequent genetic variants could be related to the risk of developing the disease, requiring a subsequent validation of these results.

KEYWORDS: Neoplasms, germline, exome, genetic variants.

TABLA DE CONTENIDO

1.INTRODUCCIÓN	22
2.JUSTIFICACIÓN	25
3.MARCO TEÓRICO	27
3.1.Aspectos generales del Cáncer Gástrico.....	27
3.2.Etiología y patogénesis del Cáncer Gástrico.....	27
3.3.Factores de riesgo del Cáncer Gástrico.....	29
3.3.1.Polimorfismos genéticos.....	30
3.4.Susceptibilidad genética al desarrollo del Cáncer Gástrico.....	30
3.4.1.Estudios de polimorfismos en genes implicados en el desarrollo de Cáncer Gástrico en Colombia.....	31
3.4.2.Estudios de polimorfismos en genes implicados en el desarrollo de cáncer gástrico en Nariño	34
3.5.Una nueva visión: Estudios genómicos en Cáncer Gástrico	35
3.5.1.Secuenciación de siguiente generación.....	35
3.5.1.1. <i>Plataformas de secuenciación de siguiente generación</i>	37
3.5.1.1.1. <i>Secuenciación de Illumina</i>	39
3.5.2.Secuenciación del exoma completo	39
4.ESTADO DEL ARTE DE LA INVESTIGACIÓN	42
5.OBJETIVOS	45
5.1.Objetivo general.....	45
5.2.Objetivos específicos.....	45
6.METODOLOGÍA	46
6.1.Zona de estudio	46

6.2.Diseño del estudio.....	47
6.2.Tipo de estudio.....	47
6.2.2. Tamaño de la muestra	47
6.3.Consideraciones éticas	48
6.4.Toma de muestras de sangre	48
6.5.Extracción de ADN genómico de sangre humana.....	48
6.5.1.Control de calidad.....	49
6.5.2.Muestra	49
6.6.Secueñaci3n Ex3mica	49
6.7.Análisis de los datos	51
6.8.Calidad de los datos.....	52
6.9.Alineamiento de secuencias	52
6.10.Llamado de variantes	53
6.10.1.Llamado de variantes con GATK:	54
6.10.2.Llamado de variantes con <i>freebayes</i>	55
6.11.Anotaci3n de variantes.....	56
6.12.Priorizaci3n de variantes.	57
6.13.Redes de interacciones funcionales entre genes.....	58
6.14.Reporte de genes que han sido estudiados en al c3ncer g3strico	59
7.RESULTADOS.....	60
7.1.Caracterizaci3n de la poblaci3n.....	60
7.2.Control de calidad, cantidad y pureza de las muestras	60
7.3.Calidad de las lecturas.....	61
7.4.Alineamiento de secuencias	62
7.5.Llamado de variantes	63

7.6.Clasificación de las variantes genómicas	65
7.7.Vías de interacciones entre genes	70
7.7.1. Vías para genes con presencia de SNPs	70
7.7.2. Vías para genes con presencia de inserciones	72
7.8.Análisis funcional	74
7.8.1. Rutas de genes con presencia de SNPs.....	74
7.8.2. Rutas de genes con presencia de inserciones.....	78
8.DISCUSIÓN	88
9.CONCLUSIONES	96
10.PERSPECTIVAS	98
11.APORTES.....	99
12.REFERENCIAS BIBLIOGRÁFICAS.....	100
13.ANEXOS	118

LISTA DE TABLAS

Tabla 1. Plataformas de secuenciación de siguiente generación. Se especifican los tamaños máximos de lectura, el número de lecturas máximo por corrida y el total de Gb alcanzado en los modelos más modernos de cada plataforma, sus aplicaciones o usos extendidos en estudios de cáncer y exómica.....	38
Tabla 2. Datos clinicopatológicos de cuatro pacientes diagnosticados con cáncer gástrico del departamento de Nariño.	60
Tabla 3. Cuantificación de las muestras en Nanodrop 2000..	61
Tabla 4. Estadísticas de alineamiento	62
Tabla 5. Datos generales resultantes del llamado de variantes con HaplotypeCaller de las cuatro muestras analizadas.....	63
Tabla 6. Datos generales resultantes del llamado de variantes con Freebayes de las cuatro muestras analizadas.....	63
Tabla 7. Variantes únicas y compartidas de los llamadores HaplotypeCaller y Freebayes para las cuatro muestras analizadas.....	64
Tabla 8. Lista de genes con SNPs e INDELs deletéreos de las muestras analizadas. Se presenta el reporte de las variantes genéticas en las bases de datos dbSNP y ClinVar, así como, el reporte de los genes con presencia de estas variantes asociados a neoplasia maligna del estómago en la plataforma DisGenet.....	83

LISTA DE FIGURAS

Figura 1. Factores que determinan el riesgo de CG.....	28
Figura 2. Acontecimientos en la patogénesis del Cáncer gástrico.....	29
Figura 3. Zonas de incidencia de CG en Nariño.....	46
Figura 4. Lectura de extremo emparejado de Illumina.....	50
Figura 5. Flujo de trabajo general para el análisis de secuenciación de exoma para mapeo, control de calidad, detección, anotación, filtrado y evaluación de INDELs y SNPs	51
Figura 6. Electroforesis en gel de agarosa al 1,5% de. M: marcador de tamaño molecular de 1 Kb. Condiciones de corrida: 100 V, 140 en buffer TAE 1X, usando cámara de electroforesis horizontal Scie Plas Ltd.	61
Figura 7. Calidad de contenido de GC por secuencia de cuatro muestras de pacientes diagnosticados con cáncer gástrico del departamento de Nariño. GC (Guanina y Citosina).....	62
Figura 8. Total de SNPs resultantes del llamado de HaplotypeCaller y Freebayes en las cuatro muestras analizadas.....	64
Figura 9. Total de INDELs resultantes del llamado de HaplotypeCaller y Freebayes en las cuatro muestras analizadas.....	65
Figura 10. Porcentaje de SNPs de origen germinal de cuatro pacientes con cáncer gástrico del departamento de Nariño. A. Variantes germinales por posición genómica. B. Variantes germinales exónicas clasificadas por su impacto a nivel molecular..	66

Figura 11. Porcentaje de INDELs de origen germinal de cuatro pacientes con cáncer gástrico del departamento de Nariño. A. Variantes germinales por posición genómica. B. Variantes germinales exónicas clasificadas por su impacto a nivel molecular	67
Figura 12. Número y porcentaje de SNPs raros, deletéreos y probablemente causantes de enfermedad en cuatro muestras estudiadas... ..	69
Figura 13. Número y porcentaje de inserciones raras y probablemente perjudiciales en las cuatro muestras estudiadas.....	69
Figura 14. Red de genes con presencia de mutaciones deletéreas para la muestra SP03 usando la base de datos STRING. Visualización de confianza alta (0,7).	70
Figura 15. Red de genes con presencia de mutaciones deletéreas para la muestra SP05 usando la base de datos STRING. Visualización de confianza alta (0,7).	71
Figura 16. Red de genes con presencia de mutaciones deletéreas para la muestra SP06 usando la base de datos STRING. Visualización de confianza alta (0,7).	71
Figura 17. Red de genes con presencia de mutaciones deletéreas para la muestra SP07 usando la base de datos STRING. Visualización de confianza alta (0,7).	72
Figura 18. Red de genes con presencia de inserciones perjudiciales para la muestra SP03 usando la base de datos STRING. Visualización de confianza alta (0,7).	72
Figura 19 . Red de genes con presencia de inserciones perjudiciales para la muestra SP05 usando la base de datos STRING. Visualización de confianza alta (0,7).	73
Figura 20. Red de genes con presencia de inserciones perjudiciales para la muestra SP05 usando la base de datos STRING. Visualización de confianza alta (0,7).	73

Figura 21. Red de genes con presencia de inserciones perjudiciales para la muestra usando la base de datos STRING. Visualización de confianza alta (0,7).	74
Figura 22. Red agrupada funcionalmente de genes con presencia de SNPs deletéreos para SP03 enriquecida con la base de datos GO y KEGG con términos como nodos vinculados en función de su nivel de puntuación $\kappa=(\geq 0,3)$	75
Figura 23. Red agrupada funcionalmente de genes con presencia de SNPs deletéreos para SP05 enriquecida con la base de datos GO y KEGG con términos como nodos vinculados en función de su nivel de puntuación $\kappa=(\geq 0,3)$	75
Figura 24. Red agrupada funcionalmente de genes con presencia de SNPs deletéreos para SP06 enriquecida con la base de datos GO y KEGG con términos como nodos vinculados en función de su nivel de puntuación $\kappa=(\geq 0,3)$	76
Figura 25. Red agrupada funcionalmente de genes con presencia de SNPs deletéreos para SP07 enriquecida con la base de datos GO y KEGG con términos como nodos vinculados en función de su nivel de puntuación $\kappa=(\geq 0,3)$	77
Figura 26. Red agrupada funcionalmente de genes con presencia de inserciones perjudiciales para SP03 enriquecida con la base de datos GO y KEGG con términos como nodos vinculados en función de su nivel de puntuación $\kappa=(\geq 0,3)$	78
Figura 27. Red agrupada funcionalmente de genes con presencia de inserciones perjudiciales para SP06 enriquecida con la base de datos GO y KEGG con términos como nodos vinculados en función de su nivel de puntuación $\kappa=(\geq 0,3)$	79
Figura 28. Red agrupada funcionalmente de genes con presencia de inserciones perjudiciales para SP07 enriquecida con la base de datos GO y KEGG con términos como nodos vinculados en función de su nivel de puntuación $\kappa=(\geq 0,3)$	79

LISTA DE ANEXOS

Anexo 1. Aval del comité de ética del proyecto “PERFIL GENÉTICO DE CÁNCER GÁSTRICO EN NARIÑO: ESTUDIO DE CASOS EN LÍNEA SOMÁTICA Y GERMINAL”	118
Anexo 2. Formato de consentimiento informado.	120
Anexo 3. Autorización para salida de muestras biológicas por parte del Instituto Nacional de Vigilancia de Medicamentos y Alimentos (INVIMA) adscrito al Ministerio de Salud y Protección Social.....	127
Anexo 4. Protocolo de extracción de ADN a partir de muestras de sangre usando el kit Ultra Clean Blood spin (Mobio).....	129
Anexo 5. Variantes puntuales deletéreas y probablemente causantes de enfermedad de cuatro muestras pacientes con cáncer gástrico del departamento de Nariño.	130
Anexo 6. Inserciones con impacto moderado y alto de cuatro pacientes con cáncer gástrico del departamento de Nariño.....	155

LISTA DE SIGLAS Y ABREVIATURAS

ADN: Ácido desoxirribonucleico

ARN: Ácido Ribonucleico

BQSR: Recalibración del nivel de calidad de la base

BWA: Burrows-Wheeler Aligner

C: Citosina

CG: Cáncer gástrico

CYP: Citocromo P450

ddNTP: Didesoxinucleótido

dNTP: Desoxirribonucleótidos

Fb: Freebayes

FS: FisherStrand

G: Guanina

GATK: Genome Analysis Tool Kit

Gb: pares bases

GC: Contenido de guanina y citosina

GO: Gene Ontology

GRC: Genome Reference Consortium

GST: Glutathión S-Transferasas

HC: HaplotypeCaller

HTS: Secuenciación de alta capacidad

IL-1 β : Citoquinas proinflamatorias Interleuquina-1 β

INDEL: Inserciones, Deleciones

INVIMA: Instituto Nacional de Vigilancia de Medicamentos y Alimentos

KEGG: Kyoto Encyclopedia of Genes and Genomes

MAF: Frecuencia del alelo menor

MI: Metaplasia intestinal

MQ: RMSMappingQuality

MQRankSum: MappingQualityRankSumTest

NSG: Secuenciación de siguiente generación

OMS: Organización Mundial de la Salud

PCR: Reacción en cadena de la polimerasa

QD: QualByDepth

SAM: Sequence Alignment Map

SMRT: Single Molecular Real Time sequencing

SNPs: Single nucleotide polymorphism

SOR: StrandOddsRatio

SV: Structural Variants

TD: tetramerización

TNF: Factor de necrosis tumoral

TNF α : Factor de Necrosis Tumoral- α

UCC: Universidad Cooperativa de Colombia

UVP: Geles a luz ultravioleta con un sistema de foto documentación

VCF: Variant Call Format

WES: Secuenciación de exoma completo

GLOSARIO DE TÉRMINOS

ADENOCARCINOMA: Carcinoma que tiene su origen en células que constituyen el revestimiento interno de las glándulas de secreción externa.

ANGIOGÉNESIS: Crecimiento de vasos sanguíneos nuevos que los tumores necesitan para crecer.

APOPTOSIS: Proceso de muerte celular programada.

CARCINOGENÉISIS: Conjunto de fenómenos que determinan la aparición y desarrollo del cáncer.

DISPLASIA: Anormalidad en el aspecto de las células debido a alteraciones en el proceso de maduración de las mismas.

ENHANCERS: Región corta del ADN eucariota que puede unirse con proteínas (factores de transcripción) para aumentar los niveles de transcripción de genes en un grupo de genes.

EPIGENÉTICA: Mecanismos modificadores de cromosomas que cambian la plasticidad fenotípica de una célula u organismo.

EPITELIO GLANDULAR: Tipo de tejido encargado de tapizar y cubrir los órganos asociados con la secreción de sustancias.

ETIOLÓGICO: Ciencia centrada en el estudio de la causalidad de la enfermedad.

EXÓN: Región de un gen que codifica aminoácidos.

HAPLOTIPO: Conjunto de variaciones del ADN, o polimorfismos, que tienden a ser heredados juntos.

HIPOMÓRFICAS: Efecto menor en el gen en presencia de una mutación

HISTOPATOLÓGICA: Estudio de las células y el tejido enfermos bajo un microscopio.

INTRON: Es una parte del gen que no codifica ningún aminoácido.

METAPLASIA: Transformación citológica de un epitelio maduro en otro que puede tener un parentesco próximo o remoto.

METILACIÓN GÉNICA: Proceso epigenético que participa en la regulación de la expresión génica.

MITÓGENOS: Son factores que actúan en el ciclo celular estimulando la división celular.

MUTACIÓN INTRÓNICA: Mutación localizada en un intrón. Puede generar un sitio de corte y empalme (splicing) alternativo que compita con los sitios de splicing normales durante el procesamiento del ARN.

NANOPORO: Es un pequeño agujero del orden de 1 nanómetro en su diámetro interno, se pueden crear por una proteína formadora de poros o como un agujero en materiales sintéticos como silicona o grafeno.

NEOPLASIA: Formación anormal en alguna parte del cuerpo de un tejido nuevo de carácter tumoral, benigno o maligno.

OLIGONUCLEÓTIDOS: Secuencia corta de ADN o ARN, con cincuenta pares de bases o menos.

POLIMORFISMO: Frecuencia de uno de los alelos en una población superior al 1%.

TRANSCRIPTOMAS: Conjunto de todas las moléculas de ARN presentes en una célula o grupo de células en un momento determinado.

XENOBIÓTICOS: Sustancia química en el interior de un organismo que no se produce naturalmente.

ZONA HOMOPOLIMÉRICA: Repeticiones en tándem del mismo aminoácido, de manera que abundan las proteínas que tienen un mismo aminoácido seguido varias veces.

1. INTRODUCCIÓN

El cáncer gástrico (GC), es una enfermedad heterogénea que se desarrolla en la capa interna del estómago, con una diseminación probable a otras partes del cuerpo como el hígado, pulmón, huesos y ganglios linfáticos. A nivel mundial, el cáncer de estómago es la sexta neoplasia con mayor incidencia, en 2018 se reportaron alrededor de un millón de nuevos casos y 783.000 muertes (Yoshida *et al.*, 2020).

En Colombia, la patología es el tercer tipo de cáncer más común con 7.419 nuevos casos y es la primera causa de muerte por cáncer con 5.505 fallecimientos (Bray *et al.*, 2018). No obstante, el riesgo y la incidencia de la patología es mayor en algunas regiones del país, tal es el caso del departamento de Nariño donde la incidencia de CG es una de las más altas (26.1 y 14.8 casos/100000 habitantes en hombres y mujeres, respectivamente) (Pardo & Cendales, 2015).

Al igual que otros tipos de tumores, el desarrollo de la enfermedad se explica en parte por la interacción de diferentes factores de riesgo, entre los que se encuentran: una edad avanzada de 60 años, el sexo masculino, el grupo sanguíneo tipo A, el tabaquismo, la infección por virus de Epstein-Barr, las dietas con un alto contenido de carnes rojas y ahumadas, alimentos procesados o conservados, alto contenido de sal, nitratos y nitrosaminas, la baja ingesta de frutas y vegetales verdes, además, la epigenética es un factor a tener en cuenta (Delgado *et al.*, 2017). Además, la infección crónica por la bacteria *Helicobacter pylori* (*H. pylori*), es considerada el principal factor de riesgo en el desarrollo de la patología, siendo catalogado por la Organización Mundial de la Salud (OMS) a través de la Agencia Internacional para la Investigación del Cáncer como carcinógeno de categoría I en humanos (Rivas *et al.*, 2017).

Aunado a este proceso, la carcinogénesis gástrica se relaciona con la acumulación de alteraciones genéticas que juegan un papel preponderante en la patología. Por lo cual, en los últimos años se ha realizado una búsqueda intensiva de variantes en genes relacionados con la historia clínica del paciente o en genes cuya relación con la enfermedad no haya sido aún descrita, para este proceso se está implementando la utilización de la secuenciación de alto rendimiento (*High-throughput sequencing*, HTS) (Rivas *et al.*, 2017).

La HTS y la implementación de programas bioinformáticos han supuesto una revolución en el abordaje del diagnóstico oncológico, ya que amplían el espectro de variantes genéticas que

contribuyen al aumento del riesgo de la enfermedad (Calabria *et al.*, 2016; Meyerson *et al.*, 2010). Una de las técnicas de HTS que está siendo ampliamente utilizada y se limita al análisis de las regiones del genoma que codifican proteínas es la secuenciación de exoma completo (*Whole Exome Sequencing* (WES)), donde se estima que ocurren 85% de las mutaciones responsables del cáncer. (Meyerson *et al.*, 2010; Sastre, 2014; Tetreault *et al.*, 2015).

La aplicación de WES en la detección de alteraciones genéticas proporcionaría una información más completa sobre la biología molecular del cáncer relacionada con: 1. Mutaciones somáticas: adquiridas por exposición a factores ambientales (infección por *H. pylori*) y localizadas solo en el ADN de las células tumorales, y 2. Mutaciones de la línea germinal: heredadas a la descendencia y que resultan en susceptibilidad al cáncer. Aunque, los dos tipos de mutaciones contribuyen al desarrollo de la enfermedad, la mayoría de alteraciones genéticas en cáncer son de línea somática (Peralta *et al.*, 2015).

Aproximadamente 10% de las variantes genéticas patógenas identificadas en los tejidos con cáncer esporádico son de origen germinal, sin embargo, este tipo de mutaciones pasan desapercibidas debido a la falta de una fuerte historia familiar de cáncer, incidencia específica del sexo en ciertas neoplasias, penetrancia incompleta, mutaciones hipomórficas, aparición de mutaciones de novo, formas familiares de expresión tardía, o regulación epigenética de la expresión génica (Welander *et al.*, 2014; Zhang *et al.*, 2016; Konnick & Pritchard, 2016). Este ultimo aspecto es importante si se tiene en cuenta que las células cancerosas humanas albergan tanto anormalidades epigenéticas globales, como numerosas alteraciones genéticas. Estas en conjunto pueden interactuar en todas las etapas del desarrollo del cáncer, promoviendo la progresión de la enfermedad. Además, se sugiere que las alteraciones epigenéticas pueden ser los eventos iniciadores clave en algunas formas de cáncer (Sharma *et al.*, 2010).

Por lo anterior, es necesario incorporar estudios de caracterización de exoma por medio de los cuales se pueda obtener y analizar datos de secuenciación a gran escala de casos de cáncer con el fin de descubrir variantes raras de línea germinal que influyan en la susceptibilidad al desarrollo de la patología (Lu *et al.*, 2014).

De esta manera, con esta investigación se evaluaron mutaciones puntuales que tuvieran predisposición al CG aplicando HTS, además, se realizó la implementación de programas

bioinformáticos de análisis de datos, que permitieron detectar de forma directa y eficaz variantes genéticas involucradas en el aumento de riesgo al desarrollo de la enfermedad.

2. JUSTIFICACIÓN

El CG es una enfermedad que involucra diferentes factores, entre los que se incluyen un conjunto distinto de mutaciones claves en la evolución de la enfermedad, por lo que es necesario identificar variantes genéticas que contribuyen al riesgo de desarrollo de la misma, tanto mutaciones somáticas que pueden ayudar a la progresión de la carcinogénesis, como mutaciones de línea germinal que pueden revelar los genes de predisposición al cáncer (Chang *et al.*, 2013).

El cáncer de manera frecuente se distingue de tipo “hereditario” (familiar) o esporádico (por agente exógeno). Tales divisiones surgieron inicialmente de la identificación de mutaciones raras de línea germinal altamente penetrantes que causan cáncer "hereditario", tales como mutaciones en el gen *CDH1* implicadas en el desarrollo de CG difuso (Garziera *et al.*, 2013); Lu *et al.*, 2014). Estas mutaciones raras son importantes en familias particulares, pero no explican muchos casos encontrados en la población.

Se cree que las variaciones de la línea germinal de la secuencia de ADN representan un aspecto clave de la predisposición a la mayoría de los rasgos complejos, como el cáncer. Las variantes genéticas particulares pueden alterar el perfil de expresión génica y afectar la función de los genes, lo que conduce a un mayor riesgo de susceptibilidad a una variedad de enfermedades, incluido el cáncer (Hu & Chen, 2012).

Es importante destacar que, a diferencia de los estudios que determinan la contribución de los genes basados en estimaciones de riesgo de cáncer entre los miembros de una familia, la exclusión de individuos estrechamente relacionados permitiría generar estimaciones menos probables de sesgo por el ambiente compartido (Lu *et al.*, 2014) y determinar que en realidad el cáncer de tipo “esporádico” presenta un componente de carácter hereditario, lo cual se demostraría de forma objetiva identificando mutaciones en genes concretos en la línea germinal (Zhang *et al.*, 2018).

La prevalencia de mutaciones de línea germinal en los genes de susceptibilidad genética a CG en pacientes del departamento de Nariño se desconoce, siendo necesario estudios que evalúen la prevalencia real de estas mutaciones en el CG carente de antecedentes familiares en esta población; posiblemente, el número de mutaciones detectadas en pacientes con CG esporádico

podría ser similar al de las mutaciones detectadas en pacientes con historia familiar con CG. Los esfuerzos dirigidos a definir mejor el papel patogénico de las variantes de la línea germinal implicarían ampliar el número de casos del CG y a futuro descubrir posibles dianas terapéuticas novedosas para posibles tratamientos efectivos (Urbini *et al.*, 2018).

Esta investigación pretende evaluar mutaciones puntuales que tengan predisposición al CG, este propósito exige la aplicación de nuevas tecnologías de alto rendimiento como la secuenciación de siguiente generación (NSG) o actualmente también denominada *Highthroughput sequencing* (HTS), que ha supuesto una revolución en el abordaje del diagnóstico oncológico, proporcionando una información más completa sobre la biología molecular del tumor. Además, el desarrollo de estas nuevas tecnologías, junto con la implementación de programas bioinformáticos de análisis de datos, permiten evaluar múltiples variantes genéticas dentro de una gran cantidad de genes que pueden promover un aumento del riesgo al desarrollo de la enfermedad (Calabria *et al.*, 2016).

El presente estudio brinda información por primera vez de la caracterización de variantes germinales en exoma de pacientes con diagnóstico de CG esporádico en el departamento de Nariño, para ello esta investigación obtuvo datos genómicos, complementariamente, se estableció un flujo de trabajo bioinformático con el cual se realizó la identificación de las variantes para esta patología. Además, los resultados fueron sometidos a publicación y a futuro servirán de base para el estudio del CG en la región desde la perspectiva de las ciencias genómicas.

La caracterización desde una perspectiva molecular de los pacientes con cáncer, permitió la identificación de posibles marcadores genéticos tempranos de la enfermedad que en el mediano o largo plazo se podrían implementar para la realización de una estratificación individual del riesgo, diagnóstico y tratamiento precoz de la enfermedad.

3. MARCO TEÓRICO

3.1. Aspectos generales del Cáncer Gástrico

El CG se constituye como la sexta neoplasia con mayor incidencia en el mundo con 5,7% de casos nuevos y es el cuarto tipo de cáncer más letal luego del cáncer de pulmón, seno y colorrectal, cada año se registran 783,000 muertes (Bray *et al.*, 2018).

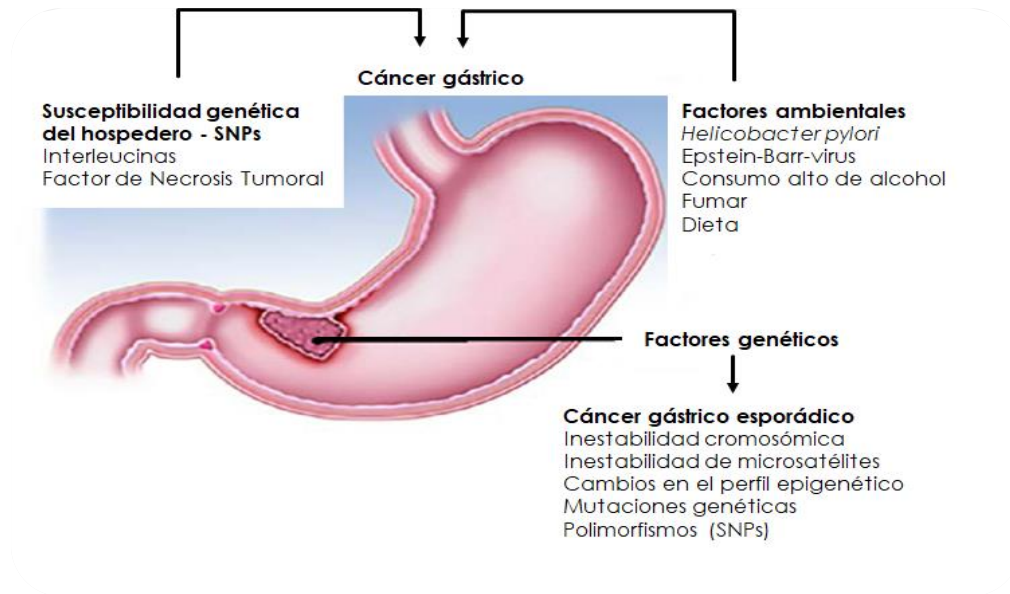
Esta patología es un tumor, originado en el epitelio glandular de la mucosa gástrica que invade la pared interna del estómago. La progresión del CG consiste en varios procesos que incluyen la proliferación y la diferenciación celular, la angiogénesis y la degradación de la matriz extracelular (Cumbal, 2010). La enfermedad es el resultado final de un largo proceso carcinogénico multifactorial que se desarrolla en distintas etapas y en el que intervienen un elevado número de factores ambientales y genéticos (Marín, 2013).

3.2. Etiología y patogénesis del Cáncer Gástrico

El CG es un proceso complejo y de etiología multifactorial, implica componentes genéticos, la dieta, el estilo de vida y la infección por *H. pylori* que juegan un papel preponderante en el desarrollo de la enfermedad (Cheng *et al.*, 2016).

El desarrollo de CG por la infección crónica de *H. pylori* depende de la cepa bacteriana y la duración de la infección, la presencia de la bacteria en el estómago induce en las células epiteliales gástricas la síntesis de citocinas proinflamatorias: interleucina 1 (IL1), factor de necrosis tumoral (TNF), ciclooxigenasa-2 (COX-2). Además se forman especies reactivas de oxígeno y óxido nítrico con capacidad para inducir la transformación a neoplasia o la proliferación maligna de linfocitos que da lugar a la formación de linfomas, estos dos procesos mediados por el estímulo de expresión de oncogenes que aumentan la proliferación celular y alteran la estructura del ADN como la capacidad de reparación del mismo (Figura 1) (Ramos *et al.*, 2018). Como en otros tipos de cáncer, los factores ambientales genéticos y epigenéticos en combinación contribuyen a la tumorigénesis y la progresión del CG (Kang *et al.*, 2014).

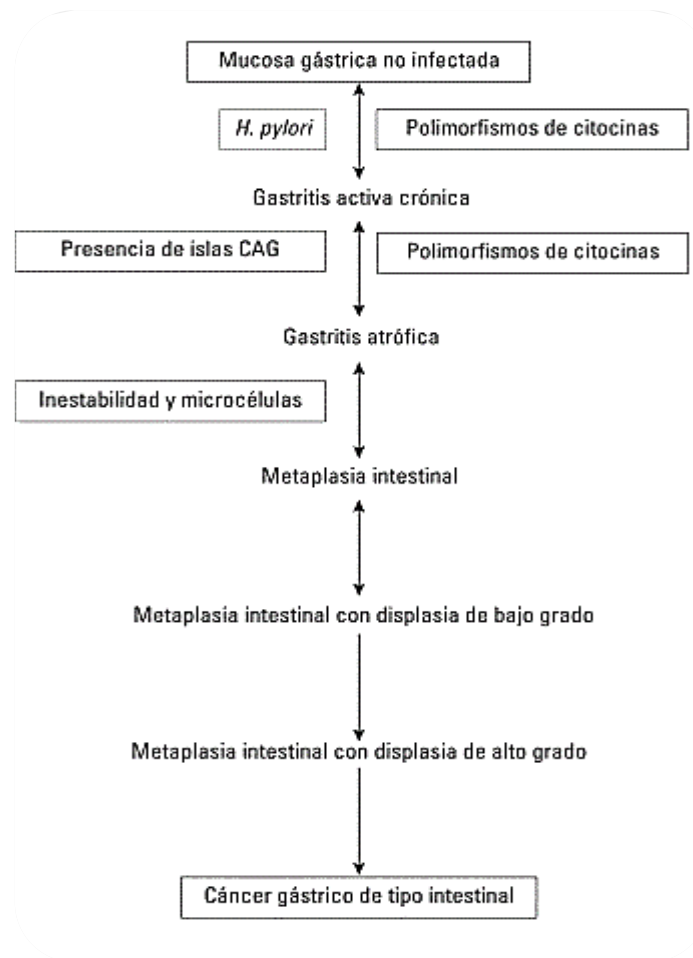
Figura 1. Factores que determinan el riesgo de CG. Fuente: Tomado y modificado de Ramos *et al.* (2018).



El cáncer de estómago se puede subdividir en dos entidades histopatológicamente bien diferenciadas según la clasificación de Lauren: el tipo intestinal y el tipo difuso. El tipo intestinal se caracteriza por la formación de estructuras tubulares que imitan las glándulas intestinales, suele ser la forma predominante en las regiones con una incidencia elevada de CG (por tal razón se conoce como “tipo epidémico”). El tipo difuso es menos diferenciado, careciendo de estructura glandular e invade con más frecuencia la pared gástrica, se observa con igual frecuencia en todo el mundo, aparece a una edad más temprana y conlleva un pronóstico menos favorable (Cebrián *et al.*, 2016).

El desarrollo del cáncer gástrico de tipo intestinal corresponde a un modelo multifásico introducido en su mayor parte por Correa. (2011). El inicio y la progresión del CG de tipo intestinal se caracterizan por la inflamación de la mucosa gástrica producida por la infección crónica de *H. pylori* que se considera el principal agente etiológico de este tipo de cáncer de estómago. Otros factores que pueden inducir o favorecer la inflamación son la dieta y los efectos de las sales biliares. En algunos pacientes el proceso inflamatorio conduce al desarrollo de una gastritis atrófica (con pérdida de tejido glandular), seguida de metaplasia intestinal, displasia, CG precoz y, finalmente, CG avanzado (Figura 2). Por otro lado, se ignora si el CG de tipo difuso sigue una progresión histopatológica similar a la del cáncer de tipo intestinal (Cebrián *et al.*, 2016).

Figura 2. Acontecimientos en la patogénesis del Cáncer gástrico.
Fuente: Tomado de Cebrián *et al.* (2016).



3.3. Factores de riesgo del Cáncer Gástrico

Según la Organización Mundial de la Salud, un factor de riesgo es cualquier rasgo, característica o exposición de un individuo que aumente su probabilidad de sufrir una enfermedad o lesión. Aunque no se conocen con exactitud las causas del desarrollo del CG, se han reportado múltiples factores de riesgo asociados, entre ellos una edad mayor de 60 años, el sexo masculino, el grupo sanguíneo tipo A, el tabaquismo, la infección por *H. pylori* o por virus de Epstein-Barr, el abuso de tabaco, el consumo de alcohol, y las dietas con un alto contenido de carnes rojas, ahumadas, alimentos procesados o con conservantes, alto contenido de sal, nitratos y nitrosaminas, así como baja ingesta de frutas y vegetales verdes. En los países en desarrollo la dieta influye en el desarrollo de cáncer en hasta 20% de los casos (Delgado *et al.*, 2017). Mientras que los países

desarrollados, que basan su alimentación en una dieta Mediterránea se asocian con un menor riesgo de mortalidad por CG (Schwingshackl & Hoffmann, 2015).

Adicionalmente, los factores genéticos juegan un papel preponderante en la carcinogénesis gástrica, estos se relacionan con variantes alélicas específicas o polimorfismos en una gran variedad de genes, que pueden modificar el efecto de la exposición ambiental y su interacción gen-ambiente (Rosero *et al.*, 2016).

3.3.1. Polimorfismos genéticos

Dos individuos diferentes comparten alrededor de 99.9% de las secuencias del ADN, mientras que, 0,1% restante constituye la variabilidad genética y fenotípica. A estas variaciones genéticas se les conoce como polimorfismos, que representan diferentes formas en las secuencias de ADN y son considerados como tal cuando la frecuencia de uno de sus alelos en la población es superior a 1% (Checa, 2007; Ramírez *et al.*, 2013).

Los polimorfismos constituyen uno de los principales factores genéticos involucrados en la susceptibilidad y la predisposición genética a las enfermedades. La susceptibilidad, se define como la presencia de determinadas variaciones en las secuencias del ADN y/o la combinación de una serie de ellas (haplotipos) en un individuo, que pueden incrementar el riesgo de desarrollar una determinada enfermedad. Mientras que la predisposición genética, es el aumento de la susceptibilidad que se asocia con un mayor riesgo de padecer la patología (Rosero *et al.*, 2016).

Existen diversos polimorfismos como inserciones, deleciones, cambios en el número de secuencias repetidas y SNP's (*single nucleotide polymorphism*) que se distribuyen con mayor frecuencia en el genoma humano. Actualmente se han descrito más de 10 millones de SNP's, con una distribución de aproximadamente un SNP por cada 100-300 nucleótidos (Checa, 2007; Ramírez *et al.*, 2013).

3.4. Susceptibilidad genética al desarrollo del Cáncer Gástrico

La susceptibilidad genética se define como la presencia de determinadas variaciones en las secuencias del ADN y/o la combinación de una serie de ellas (haplotipos) en un individuo, que pueden incrementar el riesgo de desarrollar una determinada enfermedad (Rosero *et al.*, 2016).

La susceptibilidad genética puede ser crítica en muchos de los pasos de la carcinogénesis gástrica. Esta susceptibilidad se debe a la presencia de variantes alélicas de genes involucrados en: 1) factores de protección de la mucosa gástrica (mucinas); 2) factores inmunogenéticos como la respuesta inflamatoria inducida por citocinas (como IL-1 β y TNF α); 3) variabilidad intrínseca en las proteínas involucradas en los mecanismos de reparación de ADN (como XRCC1: X-ray Repair Cross-Complementing Protein), y 4) variabilidad en la respuesta a la desintoxicación de compuestos carcinógenos llevada a cabo por las enzimas metabólicas (Torres *et al.*, 2004).

3.4.1. Estudios genéticos relacionados con el desarrollo de cáncer gástrico en Colombia

En Colombia, el CG es la primera causa de muerte por tumores malignos y el tercero en incidencia, teniendo en cuenta ambos sexos, por lo cual esta neoplasia se ha convertido en un problema de salud pública con fuerte impacto a nivel social y económico (Bray *et al.*, 2018). Los departamentos con mayor incidencia de CG son aquellos ubicados en regiones montañosas de la geografía colombiana en contraste con los de zonas costeras. La diferencia en la incidencia de la patología entre las poblaciones, puede ser dada por variaciones en la susceptibilidad genética, la virulencia de las cepas de *H. pylori* circulantes, los factores ambientales y aspectos relacionados con el estilo de vida. Para enfrentar esta problemática, el Plan Nacional para el Control del Cáncer (Ministerio de Salud y Protección Social, 2012), propuso, el control del riesgo y la detección temprana con el fin de disminuir la carga de la enfermedad. En este sentido, en el país cobró importancia la búsqueda de biomarcadores que han permitido identificar aquellos individuos de mayor riesgo dentro de una población, con el fin de implementar estrategias de prevención (Yáñez, 2013).

En este aspecto, las investigaciones, dirigen sus esfuerzos a la búsqueda de asociación de polimorfismos genéticos o variantes alélicas con capacidad de hacer susceptible a su portador al desarrollo de lesiones precursoras de cáncer o CG. En Colombia se han explorado principalmente algunos polimorfismos en los genes:

- Interleuquina-1 β (IL-1 β) e interleuquina-1Ra (receptor antagónico) genes polimórficos que pueden originar la sobreproducción de la proteína codificada lesionando los tejidos del estómago. Las citoquinas proinflamatorias Interleuquina-1 β y anti-inflamatoria

interleuquina1Ra (receptor antagónico) que participan en la respuesta inducida por la infección de *H. pylori* (Arango *et al.*, 2010; Badía Martínez & García, 2011; (Martínez, *et al.*, 2014).

- Interleuquina-10 (*IL-10*), con polimorfismos que producen niveles reducidos de la citocina IL-10 muestran asociaciones contrastantes con las lesiones neoplásicas gástricas (Martínez *et al.*, 2014).
- Factor de Necrosis Tumoral- α (*TNF α*), gen que contiene varios polimorfismos de un solo nucleótido distribuidos principalmente en la región promotora que tienen relevancia funcional, es decir, los alelos de estos polimorfismos afectan los niveles de expresión génica. Una variante de referencia es la sustitución de Guanina (G) por Adenina (A) en la posición -308 (Torres *et al.*, 2004).
- Gen supresor de tumores *p53*, encargado de inducir la respuesta de la célula ante el daño del ADN, deteniendo el ciclo celular en caso de mutación, activando enzimas de reparación del ADN, iniciando el proceso de senescencia celular y activación de la apoptosis. Las mutaciones en la secuencia del gen pueden originar mutaciones no sinónimas o errores en el marco de lectura por la inserción, delección o desplazamiento de nucleótidos: ejemplo, mutación c.358A>G en el exón 4 y variantes que se albergan en los exones 9 y 10 del dominio TD (tetramerización) (Rosero., 2019).
- Genes de enzimas Glutación S-Transferasas (GST), relacionados con los mecanismos de desintoxicación y toxificación, a través de la conjugación de glutatión reducido (GSH) con sustratos como productos farmacéuticos y contaminantes ambientales. Los genes de GST se regulan positivamente en respuesta al estrés oxidativo e inexplicablemente se sobreexpresan en muchos tumores, lo que genera problemas durante la quimioterapia contra el cáncer (Torres *et al.*, 2004; Castaño *et al.*, 2010).

- Genes que codifican las enzimas del Citocromo P450 (CYP), como los alelos *CYP1A1*2A* y *CYP1A1*2C*, polimorfismos de desintoxicación química que han sido asociados con aumento del riesgo de neoplasias (Castaño *et al.*, 2010).

Sin embargo, los estudios que han analizado la susceptibilidad genética al CG en Colombia son pocos. Arango *et al.* (2010) observaron la prevalencia del gen *IL-1 β* en enfermedades gástricas benignas y malignas, encontrando que el genotipo CC en la región polimórfica -31, se encontró con mayor frecuencia en los pacientes con enfermedades benignas. Torres *et al.* (2004) determinaron que la frecuencia del polimorfismo G/A del *TNF- α* en las personas infectadas con *H. pylori* fue de 18% en la población con CG y de 7% en el grupo control. Además, la prevalencia de *GSTM1-0* (polimorfismo nulo) fue alta en los pacientes con CG (65,2%) al ser comparada con la del grupo control (37,5%), permitiendo a los investigadores establecer asociación entre este polimorfismo y la enfermedad. Castaño *et al.* (2010) por su parte, no encontraron asociación entre CG y *GSTM1-0* pero si con el polimorfismo *GSTT1-0* al igual que con el *GSTP1-val*, siendo este último, asociado también con riesgo para el desarrollo de cáncer. Además, demostraron la asociación de CG con el alelo C2 de *CYP2E1*5A*. La presencia simultánea del alelo C2 y el hábito de fumar, aumenta la asociación con CG, ocurriendo lo mismo con el alelo *CYP1A1*2C*.

Los resultados encontrados por Badía Martínez & García, (2011) sugieren que la respuesta pro-inflamatoria es un factor relacionado por la infección por *H. pylori* en la población estudiada; así, el polimorfismo *IL-1 β -511* es un factor relacionado con CG y ulcera duodenal.

Martínez *et al.* (2014) estudiaron varios polimorfismos (*1 β -511*, *1RN*, *10-819*, *10-1082*, *TNF- α -308*) y solo encontraron asociación entre el genotipo *IL-1 β -511TT* y metaplasia intestinal (MI).

El gen *p53* fue estudiado por Cardona *et al.* (2007) en pacientes afectados por neoplasia gástrica primaria, encontrando mayor frecuencia de la variante normal *Arg*.

Suarez (2016) estudió otros polimorfismos que podrían correlacionarse con la susceptibilidad a desarrollar CG, como: rs12155758 (*PSCA*), rs4072037, rs2070803 (*MUC1*), rs3813867, rs2031920 (*CYP2E1*), rs707718, rs2286965, rs2241057 y rs3768641 (*CYP26B1*); estos aunados a las interacciones con algunos factores clínico-patológicos, se asociarían con el

riesgo a desarrollar la enfermedad, lo cual podría explicar, en parte, la aparición del CG en la población colombiana.

Galvis *et al.*, (2017), evaluaron la presencia de los SNPs C677T en *MTHFR*, mutaciones en este gen son los factores de riesgo más comunes para tener niveles elevados de homocisteína, y A592C y T819C del gen *IL-10* y factores ambientales para establecer su posible influencia en la susceptibilidad de padecer de CG, los autores concluyeron que aunque no observaron diferencias significativas en las variantes genéticas al evaluar los factores ambientales, detectaron distintos patrones entre los pacientes del grupo de casos y grupo control, lo cual probablemente explicaría la ausencia de la enfermedad en individuos sanos.

3.4.2. Estudios genéticos relacionados en el desarrollo de cáncer gástrico en Nariño

En Nariño son pocos los estudios realizados en esta temática, en el año 2013 se evaluaron dos marcadores de predisposición genética, el Factor de Necrosis Tumoral (*TNF α*) y el Citocromo P450 (*CYP1A1*) en casos con CG, registrándose elevadas frecuencias de los alelos de riesgo (*TNF α* =40%, *CYP1A1*=50%), 60% de susceptibilidad genética, el consumo de habas, ingesta de carne roja y hábito alcohólico fueron factores cercanos a la significancia ($p<0.10$), mientras que los factores genéticos-familiares evaluados no parecen relacionarse ($p>0.10$) con la prevalencia de los alelos “raros”. Además, el estudio sugiere la asociación del hábito tabáquico con una baja variabilidad en el *TNF α* (Paz, 2013).

Posteriormente, con el desarrollo del proyecto de investigación “Diagnóstico molecular de polimorfismos asociados al desarrollo de CG en Nariño” y la ejecución del trabajo de Maestría “Relación entre polimorfismos de cuatro genes y patrón de respuesta inflamatoria en pacientes con gastritis crónica y CG en una población de Nariño”, se llevó a cabo la caracterización a nivel genético de pacientes con CG (Casos) y pacientes con gastritis crónica no atrófica (Controles), mediante la evaluación de la frecuencia de los polimorfismos genéticos 511 C/T del gen *IL-1*, -342 C/G del gen *TP53*, -521 G/A del gen *EGFR* y -370 A/C del gen *CDH1* C/T y su relación con variables tanto sociodemográficas como patrones de respuesta inflamatoria. Pese a describirse el papel fundamental que desempeñan estos biomarcadores en la susceptibilidad genética, se reportó no asociación con el desarrollo de CG en la población de estudio. No obstante, los autores

recomiendan, la identificación de un mayor número de polimorfismos en genes involucrados en procesos claves en el desarrollo de esta patología (Montenegro & Carlosama, 2016).

En otra investigación determinaron que los polimorfismos de los genes *IL-1 β -511* (OR=0,7) y *TNF- α -308* (OR no determinado por fijación del alelo normal G) no están asociados con las lesiones precursoras de CG observadas en la población de Tumaco-Nariño, pese a describirse el papel fundamental que desempeñan estos biomarcadores en la susceptibilidad genética propia del hospedero en otras regiones geográficas del mundo (Rosero *et al.*, 2019).

Rosero *et al.* (2019) realizaron la secuenciación de exoma en muestras tumorales y tejido fresco de dos pacientes con CG intestinal y uno con difuso. Los autores identificaron 90 SNPs patogénicos en 61 genes, entre estos *XRCC1*, *IL4R*, *TP53*, *HNF1A*, *HNF1A*, *FLT3*, *APOB*, *BRCA2*, *CDH11*, *COL5A3*, *PKHD1* y *MUC6*. Los genes *MUC6* y *MEOX2*, presentaron indels patogénicos en los tres tumores. *MUC6*, *PKHD1* y *COL5A3* estuvieron frecuentemente mutados. Entre las principales vías alteradas se identificaron la vía del ciclo celular y adhesión focal. Los individuos con CG de tipo intestinal presentaron variantes patogénicas en genes como *ATP4A*, *GSTP1* y *CYP2E1* relacionados con la respuesta a la exposición de factores ambientales. En el paciente con el histotipo difuso, se identificaron variantes en los genes *FAT4* y *MTHFR* que pueden estar relacionadas con un peor pronóstico, característica de este tipo de cáncer.

3.5. Una nueva visión: Estudios genómicos en Cáncer

3.5.1. Secuenciación de siguiente generación en cáncer y exómica

La técnica de secuenciación de siguiente generación NGS o HTS es un conjunto de nuevas tecnologías posteriores a la secuenciación de Sanger descrito inicialmente en el año 2007, las cuales se basan en la preparación de la muestra de ADN, su secuenciación y el posterior análisis bioinformático de los datos obtenidos (Metzker, 2010).

Esta tecnología ha sido una verdadera revolución en comparación con los métodos de secuenciación tradicionales, en los que uno o unos pocos fragmentos relativamente cortos de ADN, previamente amplificados por la reacción en cadena de la polimerasa (PCR), podrían secuenciarse (Kamps *et al.*, 2017). La secuenciación genómica se ha vuelto mucho menos costosa y más eficiente, desarrollándose incluso más rápido que la tecnología informática comparable como

predice la Ley de Moore. Esto ha llevado a que la secuenciación de próxima generación (NGS) se utilice cada vez más para estudiar casi todos los tipos de neoplasias malignas, y el CG no es una excepción (Katona & Rustgi, 2017).

La secuenciación masiva permite detectar mutaciones puntuales, inserciones, deleciones, variaciones en el número de copias y translocaciones. También es de notable importancia en el estudio del cáncer el uso de la NGS en la detección de variantes en muestras tumorales y sanguíneas (Calabria *et al.*, 2016).

La gran cantidad de datos obtenidos de estos estudios recientes de NGS ha ayudado a definir el panorama genético del CG, un enfoque contemporáneo para el desarrollo de subtipos moleculares de CG con base genómica, además de aclarar nuevas mutaciones impulsoras del CG, que pueden conducir a nuevas perspectivas sobre la terapéutica (Katona & Rustgi, 2017). La aplicación de tecnologías de secuenciación de próxima generación (NGS) a escala de exoma/genoma en la investigación del cáncer ha acelerado el descubrimiento de mutaciones somáticas; contrario a esto el progreso en la identificación de la variación de la línea germinal asociada con el riesgo de cáncer es menos claro (Rotunno *et al.*, 2020).

Diferentes investigaciones han utilizado NGS para caracterizar mejor la genómica del cáncer gástrico, entre las que destacan un informe de *The Cancer Genome Atlas* (TCGA), así como un estudio de *Asian Cancer Research Group* (ACRG). El estudio TCGA evaluó 295 pacientes primarios sin tratamiento previo adenocarcinomas gástricos de múltiples centros participantes, donde se realizaron múltiples análisis genómicos incluyendo la secuenciación del exoma completo (The Cancer Genome Atlas Research Network, 2014). El estudio del ACRG examinó 300 adenocarcinomas gástricos primarios de un solo centro en Seúl, Corea del Sur (Cristescu *et al.*, 2015). Este estudio usó 49 adenocarcinomas gástricos que previamente se sometieron a un estudio con secuenciación del genoma completo 13 combinados con 251 muestras adicionales, y luego usó una combinación de perfiles de expresión génica, secuenciación dirigida de genes de interés, así como microarrays de número de copias en todo el genoma. Además de estos dos grandes estudios, también se han realizado varios estudios más pequeños que han utilizado NGS para caracterizar mejor la genómica del CG (Weren *et al.*, 2018; Cai *et al.*, 2019; Setia *et al.*, 2020; Suzuki *et al.*, 2020).

3.5.1.1. Plataformas de secuenciación de siguiente generación para el cáncer y la exómica

Las tecnologías para secuenciar el ADN se han desarrollado de forma acelerada y constituyen uno de los avances tecnológicos más importantes. Al inicio, prácticamente toda la secuenciación, incluyendo la más sofisticada y automatizada a nivel mundial estaba basada en la técnica desarrollada por Sanger en la década de los 70's (Sanger, 1977; Heather & Chain, 2016). Esta técnica nombrada “secuenciación dideoxi”, usa químicos similares a los monómeros de las cadenas de ADN, los desoxirribonucleótidos (Dntp). La precisión, solidez y facilidad de su uso llevaron a la secuencia de Sanger a ser la tecnología más utilizada para secuenciar el ADN (Heather & Chain, 2016).

Los secuenciadores automáticos basados en este método utilizan dNTP's fluorescentes marcados diferencialmente para determinar la naturaleza y orden de los nucleótidos. Hasta la fecha, esta tecnología sigue siendo ampliamente utilizada y es el estándar de oro de la secuenciación; sin embargo, ha alcanzado su máxima optimización en costo y capacidad, y presenta la limitación de no permitir el análisis de regiones extensas (>1000bp por reacción) (Shendure & Ji, 2008).

Por lo anterior, se han desarrollado métodos conocidos como de siguiente generación, los cuales tienen mayor capacidad, y permiten analizar millones de pares bases (Gb) por reacción, con costos por base decenas de veces menores que la secuenciación tradicional. Recientemente se ha desarrollado una amplia variedad de aplicaciones basadas en esta tecnología, como la secuenciación de genomas y exomas completos, la caracterización de transcriptomas y el mapeo de regiones de interacción ADN-proteína, entre otras (Silva, 2011)

Con el desarrollo y la mejora de la nueva tecnología de secuenciación, la secuenciación de próxima generación se ha aplicado cada vez más en la investigación de enfermedades como la genómica del cáncer. Esta tecnología se utiliza para identificar mutaciones nuevas y raras de cáncer en regiones específicas como el exoma, detectar portadores de mutaciones de cáncer familiar y proporcionar el fundamento molecular para la terapia dirigida adecuada. En comparación con la secuenciación tradicional, NGS tiene muchas ventajas, como la capacidad de secuenciar

completamente todos los tipos de mutaciones para una gran cantidad de genes (de cientos a miles) en una sola prueba a un costo relativamente bajo (Guan *et al.*, 2012).

En la actualidad, el desarrollo de diversas plataformas de siguiente generación con diferentes características particulares han permitido el desarrollo de numerosos estudios en cáncer y en regiones codificantes (Tabla 1).

Tabla 1.

Plataformas de secuenciación de siguiente generación. Se especifican los tamaños máximos de lectura, el número de lecturas máximo por corrida y el total de Gb alcanzado en los modelos más modernos de cada plataforma, sus aplicaciones o usos extendidos en estudios de cáncer y exómica.

Generación	Plataforma	Tecnología	Tamaño máximo de lectura (pb)	Lecturas por corrida (ML)	GB Totales	Total de estudios con las plataformas	
						Exómica	Cáncer
1ª	Sanger	Longitud didesoxinucleótidos	~ 1000	-	-	53.300	198.000
2ª	Roche 454 (no se comercializa)	Pirosecuenciación	~ 700	1	14	5.380	80.200
	Illumina/ SOLiD	Ligación y codificación por dos bases	2 x 75	100	320	55.000	149.000
	Solexa	Secuenciación por síntesis	2 x 150	6.000	1800	2.220	11.400
	PGM/ Ion Proton	Tecnología de semiconductores	2.000	80	5	2.600	16.100
3ª	Helicos (no se comercializa)	Secuenciación individual con moléculas fluorescentes	70	20 (por celda o canal)	20 (por celda o canal)	1.070	13.200
	PacBio	Secuenciación de una única molécula l en tiempo real	30.000	0,55 (por celda o canal)	1 (por celda o canal)	3.390	12.200
	Nanopore	Bioporos	200.000	1.250	12.000 (Máximo o teórico)	3.190	26.300

3.5.1.1.1. Secuenciación de Illumina y su relación con exómica y cáncer

La secuenciación por medio de Illumina se caracteriza básicamente por la ejecución de los siguientes procesos:

- a. Amplificación de fragmentos de ADN para la generación del clúster (se realiza por PCR en puente).
- b. Detección de bases en la secuenciación mediante etiquetas fluorescentes (Rubio *et al.*, 2019).

Ya que una gran cantidad de variantes oncogénicas se encuentran dentro de los exones (regiones codificantes de proteínas), que comprenden aproximadamente el 1-2% del genoma, la secuenciación del exoma ofrece una secuenciación de alta calidad con un enfoque rentable y eficiente. La secuenciación de la tecnología *Illumina* ha sido el método preferido para grandes proyectos, como *The Cancer Genome Atlas* (TCGA), porque ofrece una atractiva combinación de tiempo y precio de procesamiento eficientes y se centra en la facilidad de interpretación regiones genómicas (Jeon *et al.*, 2019).

La secuenciación a través de *Illumina* permite expandir el contenido dirigido para incluir regiones no traducidas (UTR) y microARN para una visión más completa de la regulación génica. Además, las bibliotecas de ADN se pueden preparar en tan solo un día y solo requieren cuatro a cinco Gb de secuenciación por exoma. *Illumina* ofrece varias opciones de preparación de bibliotecas, secuenciación y análisis de datos para secuenciar exomas de cáncer. Los flujos de trabajo de preparación de bibliotecas optimizados y las configuraciones de kits flexibles se adaptan a varios diseños de estudio. Los sistemas *Illumina* ofrecen una calidad de datos líder en la industria; de hecho, aproximadamente el 90% de los datos de secuenciación del mundo se generan utilizando la química de secuenciación por síntesis (SBS) de *Illumina* (Illumina, 2017).

3.5.2. Secuenciación del exoma completo como herramienta para el cáncer

En algunas ocasiones la secuenciación de genomas es poco práctica e innecesaria, por tal razón se está optando por alternativas más eficaces en la búsqueda de mutaciones que podrían causar riesgo en el desarrollo de enfermedades como la secuenciación del exoma completo o WES

por sus siglas en inglés *Whole-Exome-Sequencing*, la tecnología facilita el descubrimiento de alteraciones genéticas raras en el exoma, región codificante del genoma que representa alrededor de 2% del material genético humano y en donde se encuentran aproximadamente 85% de las mutaciones causales conocidas (Suárez, 2015).

En los últimos años, WES se ha utilizado para comparar el ADN tumoral y el ADN normal para identificar variantes específicas del cáncer. El WES del ADN tumoral requiere una mayor profundidad de secuencia que el ADN normal para identificar variantes que están presentes solo en las células tumorales. Estos análisis pueden proporcionar información sobre genes con variantes impulsoras (aquellas con un efecto sobre el desarrollo del cáncer) versus variantes pasajeras (aquellas sin efecto sobre el desarrollo de la enfermedad). La identificación de variantes potencialmente dañinas mediante WES puede proporcionar información sobre posibles nuevas vías de diagnóstico y tratamiento (Malhotra *et al.*, 2014).

Se ha demostrado que las mutaciones en las regiones de secuencias de ADN codificantes y no codificantes influyen en el desarrollo del cáncer. Los cambios de ácidos nucleicos en el exoma pueden resultar en cambios de aminoácidos en las secuencias de proteínas. Los cambios de aminoácidos conducen a una actividad debilitada de los supresores de tumores, como *APC* en el cáncer colorrectal, *VHL* en el cáncer de células renales o *BRCA* en el cáncer de mama. Los cambios en el número de copias en los reguladores del ciclo celular, como *TP53* y *RB1*, así como el mal funcionamiento de los mecanismos de reparación, incluidos los sistemas de reparación de la recombinación homóloga y el desajuste del ADN, predisponen a las células al desarrollo del cáncer. La actividad de estos sistemas de reparación se puede controlar midiendo la carga mutacional del tumor (Bartha & Györfy, 2019).

La ventaja de WES se centra en que es una alternativa más económica que la secuenciación del genoma completo, es útil para detectar variantes de alta penetrancia, la cobertura media de las regiones secuenciadas es mayor, lo que aporta mayor fiabilidad de las variantes detectadas, además, secuenciar únicamente el genoma codificante conlleva una mayor simplicidad del análisis bioinformático y, por lo tanto, representa un enfoque clínicamente factible para el diagnóstico del paciente. De esta forma, la secuenciación del exoma completo proporciona una herramienta versátil para monitorear simultáneamente múltiples cambios genómicos diferentes en el tejido

tumoral (Suárez, 2015). En contraste, las desventajas de WES se asocian a que se reduce a una parte del genoma, lo que implica que no se detecten variantes fuera de los exones y los intrones cercanos lo que hace que ciertas alteraciones patogénicas ubicadas en *enhancers* u otros elementos reguladores de la expresión de genes se pasen por alto (Esteban , 2016).

Un aspecto importante a tener en cuenta en WES es la cobertura, LaDuca *et al.* (2017) evaluaron 100 muestras de WES detectando 153.300 variantes patogénicas y una cobertura de 98,6%, el rendimiento diagnóstico en su estudio fue similar a las pruebas basadas en panel para las enfermedades mendelianas, no obstante, los autores señalan que pese a que la cobertura es mayor en las regiones codificantes esta ha mostrado variabilidad base por base y gen por gen dentro y entre genes candidatos; las razones se centran en la orientación incompleta de regiones exónicas, baja eficiencia de captura, baja profundidad de secuenciación, mayor tamaño de gen y alto contenido de GC dentro de la región objetivo. A pesar de la mejora en la cobertura media en todo el exoma con kits de captura más nuevos y con mayores profundidades de secuenciación, la cobertura de genes individuales aun no es uniforme. Es importante evaluar la idoneidad de la cobertura de genes diana cuando se analizan los datos del exoma completo para los genes de una enfermedad de estudio y la necesidad de mejorar la tecnología actual antes de que la secuenciación del exoma completo pueda desplegarse ampliamente para el diagnóstico clínico (Kong *et al.*, 2018).

4. ESTADO DEL ARTE DE LA INVESTIGACIÓN

Múltiples factores explican en parte la evolución al CG, la interacción de agentes ambientales con los factores del hospedero (polimorfismos responsables de la predisposición genética), modulan la base molecular de la carcinogénesis gástrica, que contribuye en cierta medida el desarrollo de neoplasias (Sanz *et al.*, 2005).

Sin embargo, como lo reporta Lichtenstein *et al.* (2000), las contribuciones de variantes germinales que implican los rasgos recesivos y las combinaciones de genes en el cáncer esporádico son difíciles de determinar a partir de estudios familiares.

Garziera *et al.* (2013), realizaron la caracterización de variantes germinales del gen *CDH1* en pacientes con CG esporádico, con el fin demostrar que el desarrollo de CG intestinal (esporádico) presentaba predisposición genética heredada y no sólo estaba relacionado a factores ambientales, los investigadores identificaron que la mutación intrónica IVS12 (c.1937-13T.C) condujo a un transcripto de *CDH1* aberrante con delección en el exón 11, la variante polimórfica 59UTR (254G.C) induce disminución en la transcripción. Además, se encontraron varias sustituciones silenciosas y polimórficas en el grupo de estudio. Ellos concluyeron que la identificación de otros factores de riesgo, como marcadores que permitan la detección temprana de tumores podría ser útil para la detección precoz de individuos con riesgo a CG.

Lu *et al.* (2014), cuantificaron el papel de las mutaciones genéticas de la línea germinal en la susceptibilidad a cáncer. Para esto, los autores estimaron la importancia de las variantes genéticas comunes usando la heredabilidad de matriz (h^2 g) en doce tipos de cáncer. Para el CG observaron un efecto positivo, pero no significativo h^2 g= 0.11 (rango de 0-0.27). El estudio demostró que el cáncer esporádico presenta un componente hereditario.

Por otra parte, Sahasrabudhe *et al.* (2017), determinaron mutaciones diferentes a las variantes genéticas más estudiadas para el gen *CDH1*, asociadas con alto riesgo de CG heredado. Los autores colectaron muestras de sangre periférica y analizaron secuencias de exoma completo en cáncer esporádico y de tipo heredado, identificando mutaciones en los genes *PALB2*, *BRCA1* encargados de la regulación de la recombinación homóloga del ADN, estas mutaciones se encontraron en 2 de 31 pacientes con CG hereditario y en 9 de 331 pacientes con CG esporádico.

Weren *et al.* (2018), evaluaron mutaciones de línea germinal en los genes *CTNNA1*, *MAP3K6* o *MYD88*, nuevos genes candidatos que predisponen al GC, para este objetivo evaluaron muestras de 286 pacientes con CG mediante NGS. Los autores no encontraron variantes deletéreas de línea germinal en *MYD88*, pero se observaron de forma recurrente en *CTNNA1* (n=2) y *MAP3K6* (n = 3). Las variantes deletéreas en *MAP3K6* ocurrieron con frecuencia en la población general. El estudio concluye que *MAP3K6* ya no debería considerarse un gen de predisposición a GC; las variantes deletéreas de *CTNNA1* se confirmaron como una causa poco frecuente de susceptibilidad a GC; las mutaciones de la línea germinal *MYD88* son una causa muy rara de susceptibilidad a GC.

Wang *et al.* (2019), realizaron un amplio análisis genómico de línea germinal en muestras de 83 pacientes y la secuenciación del ARN se realizaron en 57 muestras de pacientes hispanos/latinos con CG. Se realizaron análisis comparativos utilizando datos de pacientes asiáticos y blancos publicados previamente por *The Cancer Genome Atlas* (TCGA). Los investigadores determinaron que los tumores de pacientes hispanos/latinos se enriquecieron significativamente para el subtipo genómicamente estable (según lo definido por TCGA), en comparación con los asiáticos y los blancos (65% frente a 21% frente a 20%, $p < 0,001$). Los autores concluyeron que los pacientes hispanos/latinos con CG poseen paisajes genómicos únicos. Esto incluye una alta tasa de mutaciones de la línea germinal de *CDH1* que pueden explicar parcialmente sus fenotipos clínicos agresivos.

Avanesyan *et al.* (2020), sugirieron que existe un vínculo entre las mutaciones de la línea germinal *BRCA1/2* y un mayor riesgo de CG. Los investigadores realizaron un cribado endoscópico de tumores malignos de estómago en 120 portadores de la mutación *BRCA1* con el fin de evaluar la probabilidad de detectar la enfermedad tumoral. Los resultados no revelaron casos de CG en la primera visita. No obstante, un portador de la mutación en el gen *BRCA1* desarrolló CG 4 años después de la primera visita para examen endoscópico. Como conclusión, los datos justifican más estudios sobre las relaciones entre el *BRCA1/2* y el CG.

Setia *et al.* (2020), realizaron una caracterización genómica y morfológica de cánceres gástricos de aparición temprana y tradicional; para esto realizaron un análisis genómico de 81 pacientes con CG que tenían 50 años o menos, un análisis de correlación morfológico-molecular

e identificaron variantes de posible origen de la línea germinal a partir de datos de NGS. Los autores encontraron una tasa más alta de mutaciones en *CDH1* (22,2%) en CG de aparición temprana frente al 11,4% de CG tradicional ($p = 0,0042$); pero una tasa similar de mutaciones en *TP53* (63% de los CG de aparición temprana frente a 56,6% a los CG tradicionales; $p = 0,2674$). Observaron variantes de posible origen de la línea germinal en los genes *CDH1*, *APC*, *ATM*, *NBN* y *MUTYH*, en el 9,9% de los cánceres. Los autores concluyen que el CG temprano tiene distintas alteraciones genómicas, como las mutaciones *CDH1*, pero comparte con el CG tradicional una alta frecuencia de mutaciones *TP53*. Posibles alteraciones de la línea germinal en los genes de susceptibilidad al cáncer de penetrancia moderada pueden estar implicados en la patogenia de estos cánceres.

Suzuki *et al.* (2020), realizaron un análisis transétnico a escala genómica de 531 casos de GC (319 asiáticos y 212 no asiáticos). Los autores determinaron que había una subclase distinta de GC con una de mutación asociada al alcohol y una fuerte especificidad asiática, casi todas las cuales eran atribuibles al comportamiento de consumo de alcohol, el hábito de fumar y el alelo *ALDH2* defectuoso (específico de Asia). Además, encontraron que variantes de *CDH1* de línea germinal eran frecuentes (7,4%) entre los GC japoneses, la mayoría de las cuales se atribuyeron a algunas variantes recurrentes de un solo nucleótido compartidas por japoneses y coreanos, lo que sugiere la existencia de eventos ancestrales comunes entre los asiáticos orientales.

Los estudios relacionados a variantes de línea germinal son muy pocos en relación a los estudios de línea somática, Rotunno *et al.* (2020), realizaron una revisión sistemática de la literatura de los estudios de susceptibilidad genética al cáncer que utilizaron tecnologías NGS a una escala de exoma/genomas. Señalaron que varios estudios utilizaron un enfoque de análisis de candidatos e identificaron posibles variantes o genes de la línea germinal relacionados con el cáncer en una pequeña fracción de los casos de cáncer secuenciados. Esta revisión destaca el progreso en la identificación de la variación de la línea germinal relacionada con el cáncer hasta la fecha pese a que existen menos estudios en comparación a la línea somática. Estos hallazgos apuntan al potencial sin explotar y racialmente diversas, combinando resultados entre estudios y expandiéndose más allá de un enfoque de análisis de candidatos para avanzar en el descubrimiento de la variación genética que explica la heredabilidad inexplicable del cáncer.

5. OBJETIVOS

5.1.Objetivo general

Caracterizar variantes genéticas de línea germinal en el exoma de pacientes diagnosticados con adenocarcinoma gástrico de tipo esporádico en el departamento de Nariño.

5.2.Objetivos específicos

Identificar variantes genéticas de línea germinal en el exoma de pacientes con cáncer gástrico.

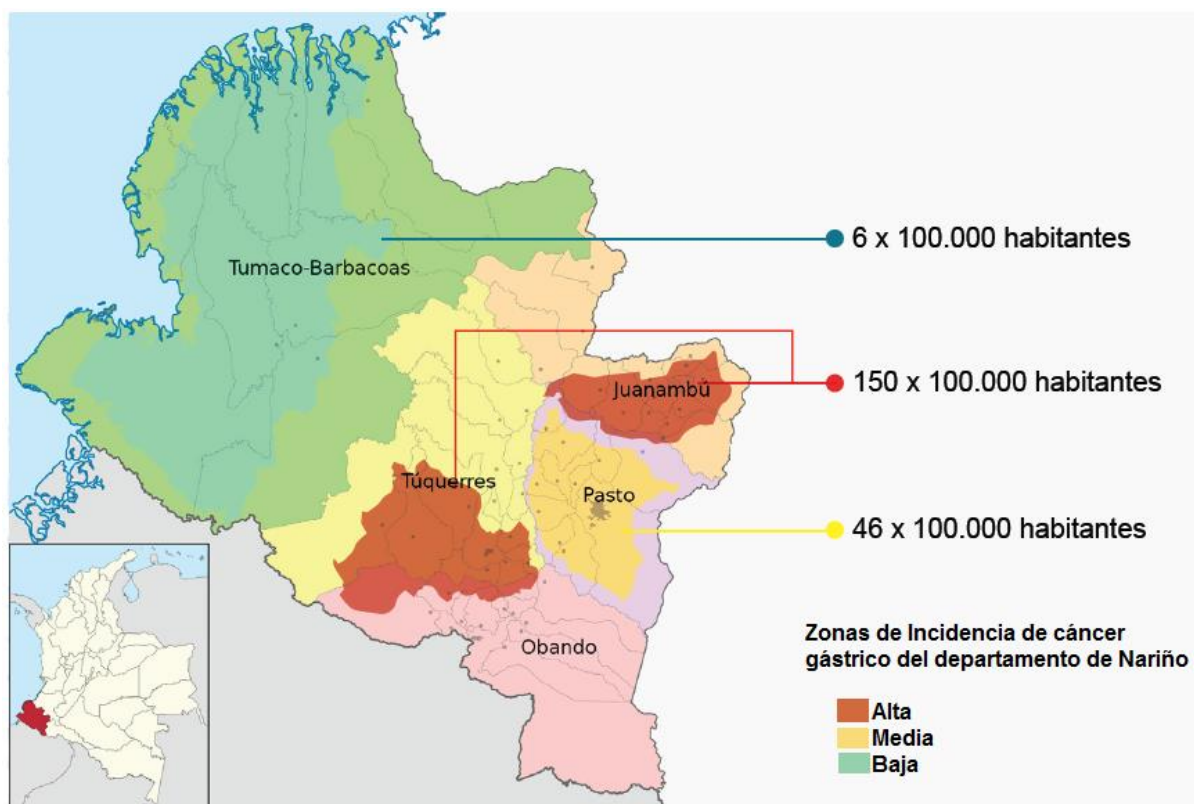
Establecer las variantes genéticas de línea germinal asociadas a cáncer gástrico para la población de estudio.

6. METODOLOGÍA

6.1. Zona de estudio

El presente estudio se realizó en el Departamento de Nariño, territorio que se caracteriza por presentar tres zonas definidas de cáncer gástrico: La región Andina, con una alta incidencia de CG (150 casos por cada 100.000 habitantes), alta prevalencia de lesiones precursoras de malignidad (gastritis crónica atrófica, metaplasia intestinal y displasia) e infección por *H. pylori* (Figura 3). Una zona intermedia en Pasto y municipios circunvecinos (Incidencia: 46 por 100.000 habitantes), con una menor prevalencia de lesiones precursoras de malignidad y alta prevalencia de infección por la bacteria. Por el contrario, la zona costera se caracteriza por la alta prevalencia de infección por *H. pylori* y baja incidencia de GG (6 casos por cada 100.000 habitantes) (Bedoya, *et al*, 2018). Este fenómeno se conoce como “enigma Nariñense” (Figura 3).

Figura 3. Zonas de incidencia de CG del departamento de Nariño. Fuente: Tomado y modificado de Bedoya *et al*, (2018).



6.2. Diseño del estudio

6.2.1. Tipo de estudio

Estudio cuantitativo, observacional de tipo transversal analítico, en casos con datos de secuenciación de exoma completo de pacientes diagnosticados con CG a partir de muestras de sangre periférica (Otzen & Manterola, 2017). Para el desarrollo de la presente investigación se incluyeron cuatro muestras de individuos diagnosticados con cáncer gástrico tipo intestinal procedentes de diferentes regiones del departamento de Nariño.

6.2.2. Tipo de muestreo

Muestreo por conveniencia (Otzen & Manterola, 2017). Fueron considerados dentro del estudio todos los pacientes que cumplieron con los siguientes criterios de inclusión y exclusión:

Criterios de inclusión.

- ✓ Pacientes mayores de 18 años provenientes del departamento de Nariño.
- ✓ Pacientes con manifestaciones histopatológicas de adenocarcinoma gástrico.
- ✓ No tener antecedentes de familiares de primer o segundo grado afectados por CG.
- ✓ No presentar ni haber presentado infección por *H. pylori*.

Criterios de exclusión.

- ✓ Pacientes con intervención de cirugía gástrica.
- ✓ Tratamiento anti-*H. pylori*.
- ✓ Infección por VIH.
- ✓ Presencia de otro tipo de cáncer.

6.3.Consideraciones éticas

Esta investigación hace parte del macroproyecto titulado “Perfil genético de cáncer gástrico en Nariño: estudio de casos en línea somática y germinal” que cuenta con el aval del acta de aprobación N° SCBE06-17 del Subcomité de Bioética de la Universidad Cooperativa de Colombia-Pasto, Anexo 1. Previo a la toma de muestras de sangre, los pacientes fueron informados verbalmente y por escrito sobre los objetivos, protocolos y características del estudio, ítems descritos en el consentimiento informado (Anexo 2). Las muestras fueron enviadas a Corea del Sur a la empresa MacroGen, Inc con autorización previa del Instituto Nacional de Vigilancia de Medicamentos y Alimentos (INVIMA) adscrito al Ministerio de Salud y Protección Social (Radicado N° 201824000498311), Anexo 3.

6.4.Toma de muestras de sangre

La elección de pacientes, obtención de muestras de sangre y diagnóstico de los participantes estuvo a cargo de un médico gastroenterólogo que hizo parte del equipo de coinvestigadores del macroproyecto mencionado anteriormente, este procedimiento se llevó a cabo por punción de la vena media cubital en condiciones asépticas, basales y en ayunas. Se recolectó una alícuota de 5 mL de sangre en una sola venopunción en tubos (*No high altitude, Vacutainer*) tapa lila con anticoagulante (EDTA). Para reducir la probabilidad de hemólisis y la formación de microcoágulos (S.A, Colombiana de Salud, 2014)

Las muestras fueron transportadas en posición vertical en neveras de icopor con geles y pilas refrigerantes (4°C – 8°C) para conservar la integridad de las muestras, hasta su posterior almacenamiento y procesamiento en el laboratorio de Biología Molecular de la Universidad Cooperativa de Colombia (UCC) de la ciudad de Pasto (S.A, Colombiana de Salud, 2014).

6.5.Extracción de ADN genómico de sangre humana

El ADN genómico fue obtenido a partir de sangre periférica total mediante el kit de extracción DNA isolation Kit Ultra Clean Blood Spin (MOBIO, Mobio Technologies Inc, San Diego, EU), siguiendo las instrucciones del fabricante (Anexo 4). Las muestras de ADN se almacenaron en un freezer a una temperatura de -20 °C, previo el análisis exómico.

6.5.1. Control de calidad

Previo al envío de las muestras al laboratorio donde se realizó el proceso de la secuenciación (Macrogen), se comprobó que su calidad, concentración y pureza fueran las adecuadas. En primer lugar, se realizó el proceso de electroforesis en un gel de agarosa a 1.5 %, en buffer TAE 1X, al cual se le adicionó 3 µL del agente intercalante Gel Red. Las muestras se prepararon con 2 µL de buffer de carga, 5 µL de ADN, además, se usaron 3 µL de marcador de tamaño molecular de 1 Kb (Bioline, Meridian Life Science, Inc, Memphis, US). La corrida electroforética se realizó en una cámara de electroforesis (Scie Plas Ltd) con buffer TAE 1X a 100 V durante 140 min. La visualización de las bandas se realizó por exposición de los geles a luz ultravioleta con un sistema de foto documentación (UVP).

Posteriormente, se realizó la cuantificación de ADN en un NanoDrop 2000 (Thermo Scientific) por triplicado y se realizó un promedio con el fin de reducir cualquier error que se pueda haber producido en la lectura, como blanco para la medición se utilizó la solución de resuspensión del ADN.

6.5.2. Muestra

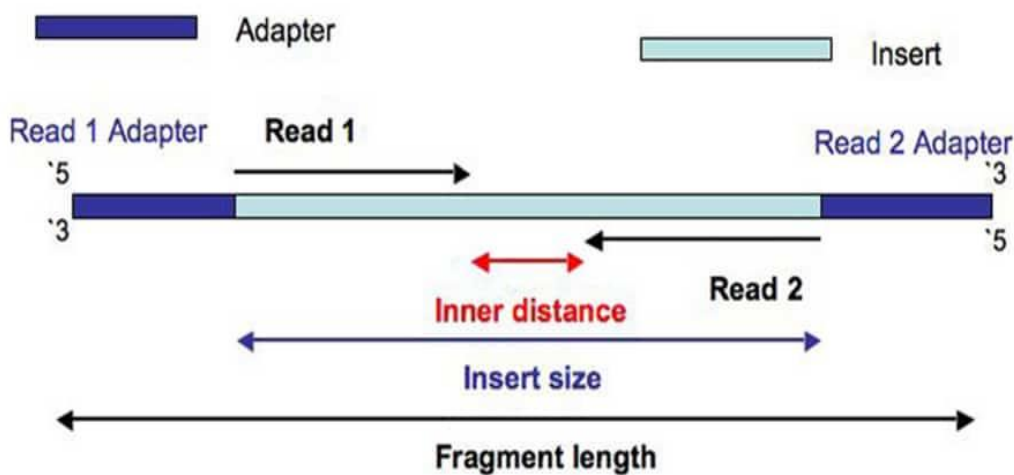
La concentración de ADN enviada para la secuenciación fue superior a 50 ng/µL (relación A260/280 de $\geq 1,8$). Las muestras se enviaron a Corea del Sur a la empresa Macrogen, Inc. con número de pedido HN00105933, para esto se contó con autorización previa del Instituto Nacional de Vigilancia de Medicamentos y Alimentos (INVIMA) adscrito al Ministerio de Salud y Protección Social, cumpliendo con los protocolos y normas reglamentarias establecidas para la exportación de muestras biológicas y/o genéticas usadas en el desarrollo de protocolos de investigación (Anexo 3).

6.6. Secuenciación Exómica

La secuenciación del exoma completo se llevó a cabo por captura de las regiones codificantes con el kit Agilent SureSelect V5. La preparación de la muestra y la secuenciación de las regiones exónicas se llevó a cabo en un equipo Illumina HiSeq – 2500 de Illumina con una lectura de 100-150pb modo *Paired-end* o llamadas lecturas de extremo emparejado. En el proceso de secuenciación, el ADN genómico se corta en millones de fragmentos cortos llamados lecturas

o “reads”. Las lecturas individuales se emparejan para crear extremos emparejados, las lecturas presentan una secuencia de adaptador única en ambos extremos de la lectura llamados: Adaptador de lectura 1 y Adaptador de lectura 2. La Lectura 1, llamada lectura directa, se extiende a lo largo de la hebra de ADN desde el Adaptador de lectura 1 en la dirección 5' - 3' hacia la Lectura 2. Por su parte, La Lectura 2, denominada lectura inversa, se extiende a lo largo de la cadena de ADN inversa desde el adaptador de lectura 2 en la dirección 5' - 3' hacia la Lectura 1. Hay una secuencia de ADN insertada entre la Lectura 1 y 2, denominada secuencia interna. La longitud de esta secuencia se mide como la Distancia interior. Por definición, el Insertar es la unión de la Lectura 1, la secuencia de Distancia interior y la Lectura 2. La longitud del Insertar es el tamaño del inserto. Un solo fragmento entonces incluiría el "Adaptador de lectura 1", la Lectura 1, la "Secuencia interna", la Lectura 2 y el "Adaptador de lectura 2". La longitud de este "Fragmento" es solo la "Longitud del fragmento" (Figura 4) (Casey, 2019).

Figura 4. Lectura de extremo emparejado de Illumina. Fuente: Casey (2019).



Se estableció y se solicitó a la compañía una cobertura media de 100X. Con 100X la cobertura fue mayor o igual a 4X (cobertura mínima aceptable) en 95 % de las secuencias, de esta forma los datos obtenidos tuvieron suficiente calidad para desarrollar los objetivos planteados en esta investigación (Gobernado, 2014).

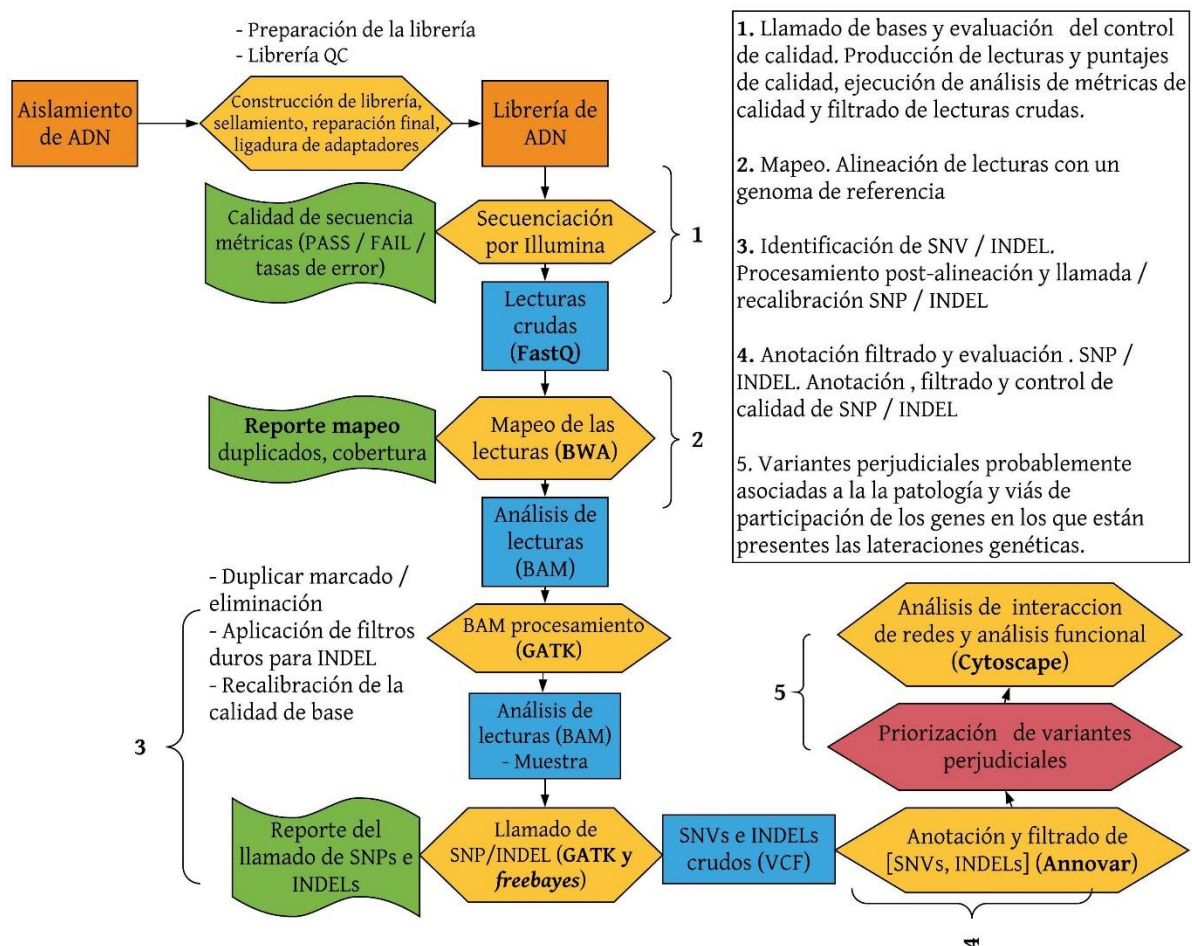
El ADN para WES se prepara utilizando protocolos basados en PCR para apuntar de forma selectiva a las regiones que se van a secuenciar. El ADN genómico purificado se fragmenta, se liga con adaptadores, se amplifica por PCR, se captura y luego se amplifica posteriormente

mediante PCR utilizando cebadores indexados. Debido a que los cebadores de PCR se aparean más fácilmente a algunas regiones del genoma que a otras, algunas regiones se amplificarán muchas veces mientras que otras se amplificarán con poca frecuencia. Para limitar el impacto del sesgo de la PCR, estas muestras normalmente se secuencian a una profundidad de lectura promedio de 100X. Esta es la profundidad requerida para minimizar el número de sitios con puntos de datos insuficientes para realizar una llamada de nucleótidos de alta confianza (Illumina, 2017).

6.7. Análisis de los datos

Para el análisis de los datos se siguió el flujo de trabajo que se resume en la Figura 5.

Figura 5. Flujo de trabajo general para el análisis de secuenciación de exoma para mapeo, control de calidad, detección, anotación, filtrado y evaluación de INDELs y SNPs. Fuente: Tomado y modificado de Tetreault *et al.* (2015).



6.8. Calidad de los datos

El proceso de análisis de datos (Figura 5), empezó cuando fueron recibidas de forma cruda las lecturas (reads) de la secuenciación en formato *FastQ*. De cada muestra se obtuvieron dos archivos, en donde la calidad de las lecturas se expresó en base logarítmica según la escala de *Phred*, la cual contribuyó a facilitar la evaluación de la secuencia (Bao *et al.*, 2014). De esta manera, una calidad Q20 representó una posibilidad de 0.01 de haber leído una base de forma errónea, Q30 una posibilidad de error de 0.001, etc (Gobernado, 2014).

Con el programa FastQC (www.bioinformatics.babraham.ac.uk/projects/), se exploraron parámetros de calidad como: total de secuencias procesadas, filtradas y su longitud, porcentaje de guanina (G) y citosina (C) total en los datos, calidad de la secuencia por base, puntuaciones de calidad por secuencia, contenido de bases en la secuencia, contenido de GC por base y por secuencia, contenido de bases que no se obtuvieron durante la lectura marcadas con N, distribución de la longitud, niveles de duplicación y la sobrerrepresentación de las secuencias.

Posteriormente, se realizaron operaciones de limpieza de las secuencias para corregir los sesgos técnicos y hacer que los datos sean adecuados para el análisis. Para este paso se usó el paquete BBMap, específicamente la opción “FilterByTile.sh” para aumentar la calidad de las bibliotecas en sesgos, o ayudar a salvar bibliotecas con problemas importantes de calidad posicional (como burbujas de celdas de flujo). Además, se realizó un recorte de adaptadores que se introdujeron al preparar la librería a través de la opción “BBDuk” (Bushnell, 2014).

6.9. Alineamiento de secuencias

El secuenciamiento por medio de Illumina generó una gran cantidad de lecturas (*reads*) cortas que posteriormente fueron alineadas con un genoma de referencia para obtener la secuencia de ADN original. En este estudio para el alineamiento se utilizó por su rapidez, precisión y acceso gratuito el paquete de software BWA (*Burrows-Wheeler Aligner*) v. 0.7.10 (Li & Durbin, 2010), que indexa el genoma de referencia y mapea secuencias de baja divergencia (alta similaridad), frente a dicho genoma. El algoritmo que se empleó fue BWA-MEM, usado para secuencias mayores a los 100 pb y lecturas paired- end (García, 2016).

El genoma de referencia usado fue el GRCh37 o hg19, publicado por *Genome Reference Consortium* (GRC), que aporta una cobertura cercana a 99.8% del genoma humano completo (Amigo, 2012).

- Procesamiento del archivo “*.bam”

El alineamiento de los archivos .fq con BWA generó un único archivo en formato “*.sam” (*Sequence Alignment Map*), este archivo fue ordenado y transformado al formato binario de SAM “*.bam” utilizando las herramientas *Picard* del programa GATK (*Genome Analysis Tool Kit*) (DePristo *et al.*, 2011), posteriormente se creó un código hash basado en los grupos de reads (RG), y se produjo un resumen de las métricas de alineación del archivo BAM.

Complementariamente, mediante las herramientas de *Picard* se procedió a marcar los duplicados de PCR del archivo “*.bam” mediante la opción “*MarkDuplicates*”, ya que durante los múltiples procesos de PCR realizados para el proceso de secuenciación, se pueden generar duplicados que contribuyen a la profundidad de lectura de una forma ficticia. Estos duplicados por lo general no son eliminados, pero son identificados y marcados con un “*FLAG*” en los archivos SAM o BAM para que los programas llamadores de variantes no los tomen en cuenta a la hora de establecer los genotipos para cada variante.

Finalmente, se generó un archivo BAM de índice “*.bai” con la opción “*BuildBamIndex*” de *Picard*, esta herramienta crea un archivo de índice para la entrada BAM que permite la búsqueda rápida de datos en un archivo BAM, como un índice en una base de datos. Culminada la etapa de procesamiento de la muestra se contó con los *reads* mapeados, alineados y filtrados.

6.10. Llamado de variantes

En el llamado de variantes las lecturas se utilizaron para identificar aquellos sitios en los cuales el exoma de estudio difería del genoma de referencia. Una variación en la secuencia de ADN de un solo nucleótido se denomina SNP (*Single Nucleotide Polymorphism*), mientras que la inserción o delección de unos pocos nucleótidos se denomina INDEL (*Insertions, Deletions*) o variantes estructurales (*Structural Variants*; SV). Estas variaciones que suelen ser las más frecuentes se suele volcar a un archivo de formato tipo VCF (por sus siglas *Variant Call Format*) introducido por el consorcio responsable del proyecto 1000 Genomas (Danecek *et al.*, 2011).

Para la identificación de todas las variaciones presentes en cada exoma, se utilizaron dos programas: *Genome Analysis Tool Kit* o GATK, versión 4.0.2.1 (DePristo *et al.*, 2011) y *freebayes*, v1.3.1-1 (Garrison & Marth, 2012).

6.10.1. Llamado de variantes con GATK:

“*HaplotypeCaller*” de GATK (desarrollado por el *Broad Institute, Cambridge, Massachusetts*. <http://www.broadinstitute.org/gatk/>), disponible en la red de manera libre (Poplin *et al.*, 2017), es un programa capaz de llamar SNPs e INDELs de manera simultánea vía un ensamblado *de-novo* local, lo que implica que cuando el programa encuentra una región que muestra señales de variación, descarta la información existente del mapeo y vuelve a montar completamente las lecturas en esa región. Este procedimiento le confiere una mayor precisión a la hora de descubrir variantes especialmente en regiones difíciles, como por ejemplo, cuando contienen diferentes tipos de variantes cerca unas de otras. Además, al ser capaz de manejar información perteneciente a organismos no diploides, se vuelve de utilidad para encontrar variantes con frecuencias menores a la típicas heterocigotas (es decir, posibles mosaicismos). El resultado entonces del proceso mencionado es la obtención de un archivo de tipo VCF para cada una de las muestras secuenciadas, que incluye una lista de variantes.

Los pasos que se realizaron para el llamado de variantes fueron los siguientes:

1. Llamado de variantes con la opción “*HaplotypeCaller*”.
2. Extracción de SNPs e Indels del conjunto de llamados a través de la opción “*SelectVariants*”.
3. Aplicación de parámetros de filtro rígido o riguroso debido a que, algunas variantes fueron artificios de las técnicas utilizadas por el mapeo y/o el proceso de secuenciación. Para SNPs, se aplicaron los siguientes parámetros:
 - QualByDepth (QD) 2.0: Esta es la variante de confianza (del *QUALcampo*) dividida por la profundidad sin filtrar las muestras sin referencia.
 - FisherStrand (FS) 60.0: Valor p con escala de *Phred* utilizando la prueba exacta de Fisher para detectar el sesgo de la cadena (la variación se observa solo en la cadena directa o inversa) en las lecturas.

- RMSMappingQuality (MQ) 40.0: Media cuadrática de la calidad de mapeo de las lecturas en todas las muestras.
 - MQRankSum: MappingQualityRankSumTest (MQRankSum) -12.5: aproximación z basada en u de la Prueba de suma de rangos de Mann-Whitney para calidades de mapeo (lecturas con bases de referencia frente a aquellas con el alelo alternativo). La prueba de suma de rango de calidad de mapeo no se puede calcular para sitios sin una mezcla de lecturas que muestren alelos alternativos y de referencia, es decir, solo se aplicará a llamadas heterocigotas.
 - ReadPosRankSumTest (ReadPosRankSum) -8.0: Esta es la aproximación z basada en u de la Prueba de suma de rango de Mann-Whitney para la distancia desde el final de la lectura para lecturas con el alelo alternativo.
 - StrandOddsRatio (SOR) 3.0: Evalúa si hay sesgo de cadena en los datos.
4. Aplicación de los parámetros de filtro duro para INDELs son: QD 2.0, FS 200.0, ReadPosRankSum 20.0, SOR 10.0.
 5. Realización de recalibración del nivel de calidad de la base (BQSR): Permite obtener cualidades básicas más precisas, lo que a su vez mejora la precisión de las variantes llamadas. Primero se aplicó la opción “*BaseRecalibrator*” y posteriormente “*BQSR*”.
 6. Obtención de un archivo VCF con las variantes que solo pasaron los filtros, es decir, se eliminan todos los sitios de FILTRO que no sean PASS, esto a través de la opción “*--remove-filtered-all*” de vcftools (Danecek *et al.*, 2011).

6.10.2. Llamado de variantes con *freebayes*

freebayes es un detector de variantes genéticas bayesianas diseñado para encontrar SNPs, INDELs, MNP (polimorfismos de múltiples nucleótidos) y eventos complejos (eventos compuestos de inserción y sustitución) más pequeños que los de alineación de secuencia de lectura corta.

freebayes está basado en haplotipos, en el sentido de que llama variantes basadas en las secuencias literales de lecturas alineadas a un objetivo particular, no su alineación precisa. Este modelo es una generalización directa de llamadores como, PolyBayes, samtools, GATK que detectan variantes basadas en alineaciones. Este método evita uno de los problemas centrales con la detección de variantes basada en la alineación: que secuencias idénticas pueden tener múltiples alineaciones posibles (Garrison & Marth, 2012).

Los pasos que se realizaron para el llamado de variantes fueron los siguientes:

1. El llamado de variantes se realizó mediante la opción “*freebayes*” la cual requiere solo de dos entradas una secuencia de referencia fasta y un archivo de alineación en formato “BAM”, al final se produce un archivo “VCF” que incluye SNPs, INDELs y variantes complejas.
2. Se filtró el archivo de salida utilizando la herramienta “*vcffilter*” de “*vcflib*”. Gracias a este filtro en el campo “*QUAL*” del archivo VCF se estima la probabilidad de existir un polimorfismo en los loci descritos por el registro. Para esta investigación se usó “*QUAL > 20*”, que quiere decir eliminación de cualquier sitio con probabilidad estimada de no ser polimórfico menor en la escala de *phred* 20 (también conocido como 0,01) o probabilidad de polimorfismo $> 0,99$.
3. Al archivo VCF filtrado se le extrajeron los SNPs e INDELs por medio de la opción “*SelectVariants*” de GATK.

Antes de la anotación se realizó la unificación de los archivos VCF para SNPs e INDELs resultantes del llamado de las variantes por medio de los dos llamadores descritos anteriormente, por medio de la opción “*--diff-site*” de VCFtools se compararon los archivos para obtener variantes únicas, compartidas y totales.

6.11. Anotación de variantes

La herramienta bioinformática ANNOVAR fue empleada para anotar funcionalmente las variantes genómicas que sean detectadas, mediante el uso de diferentes bases de datos públicas. El empleo de esta herramienta permitió distintos tipos de anotación: anotación por gen mediante la

base de datos RefSeq que proporcionó información sobre el gen en el que se encuentra la variante, su localización (exón, intrón, splicing, UTR), los diferentes transcritos en los que aparece, el tipo de variante (sinónima, no sinónima entre otras) y el cambio aminoacídico que produce; anotación por filtro, mediante *Exome Aggregation Consortium* (Exac). Además, como predictores in silico de la patogenicidad de los cambios aminoacídicos se usaron herramientas como: *PolyPhen2* y *Meta-Taster*.

6.12. Priorización de variantes.

A partir de la información anotada en las variantes se seleccionaron aquellas con probabilidad patogénica, el proceso se realizó siguiendo algunas recomendaciones de la guía para filtrar y priorizar variantes genéticas: “*A practical guide to filtering and prioritizing genetic variants*” propuesta por (Jalali & Gamielien, 2017).

Las variantes de interés fueron categorizadas de la siguiente manera:

- Se excluyeron aquellas cuya posición y cromosoma no está determinado.
- Se incluyeron variantes exónicas y se excluyeron las clasificadas con función de: RNA no codificante, regiones *upstream* y *downstream* del ADN, regiones intergénicas e intrónicas, regiones UTR 3' y 5'.
- Se incluyeron variantes exónicas no sinónimas, variantes que resultaron en ganancia o pérdida de codones de inicio o finalización de la traducción y/o cambios en el marco abierto de lectura; se excluyeron las clasificadas como sinónimas.
- De acuerdo a la frecuencia de la variante en la población se priorizaron variantes raras (MAF<0.01) según la base de datos del EXAC.
- Para evaluar la severidad del efecto de las mutaciones puntuales se utilizaron los predictores SIFT y *Mutation Taster*, se incluyeron SNPs deletéreos y perjudiciales/causantes automáticas de enfermedad y probablemente perjudiciales/causantes de enfermedad. Para INDELs la predicción se realizó con SnpEff, se seleccionaron INDELs de alto impacto que implicaron cambios de marco, adición/supresión de codones de parada y, las variantes moderadas que produjeron

cambios, eliminación e inserción de codones; las variantes que puntuaron como bajas fueron excluidas por generar cambios sinónimos.

6.13. Redes de interacciones funcionales entre genes

Una vez identificadas las variantes presentes en diferentes genes, se usó el software de código abierto Cytoscape versión 3.8.0 (Shannon *et al.*, 2003), una plataforma de visualización de interacción molecular y rutas biológicas, que integra redes con anotaciones, perfiles genéticos y otros datos. Cytoscape se empleó con el fin de realizar una red de interacción génica por medio del complemento StringApp, aplicación que integra la base de datos STRING que proporciona datos conocidos y previstos de asociaciones gen-gen o proteína-proteína para un gran número de organismos, incluidas las interacciones físicas y funcionales asociaciones con puntuaciones de confianza que cuantifican su fiabilidad. Además de integrar los experimentos disponibles datos y rutas de bases de datos seleccionadas, STRING predice interacciones basadas en análisis de coexpresión, evolución señales a través de genomas, extracción automática de texto del literatura biomédica y transferencia de evidencia basada en ortología de diferentes organismos. STRING está diseñado para realizar grandes redes y proporciona una flexibilidad amplia en términos de análisis y visualización de redes (Doncheva *et al.*, 2019).

Las vías se realizan con el propósito de tener una mejor comprensión del CG, ya que los genes con mutaciones podrían tener mayor influencia en el desarrollo de la enfermedad. Para este estudio, las vías génicas se construyeron con un nivel de confianza de 0,7, lo que expresa una mayor confianza, es decir una asociación respaldada por varios tipos de evidencia. Para el análisis de la red se usó la herramienta “*Analyze Network*”, con el fin de determinar número de nodos, interacciones y componentes conectados.

Posteriormente, con el objetivo de entender las posibles implicaciones que tienen las mutaciones en estos genes se realizó el análisis funcional, considerando principalmente la función biológica y las vías de actuación. Para este propósito, se usaron las herramientas ClueGo y CluePedia de Cytoscape empleando la base de datos actualizada del Gene Ontology (GO) y del *Kyoto Encyclopedia of Genes and Genomes* (KEGG). ClueGo y CluePedia se complementan para facilitar la interpretación biológica comparando genes en términos agrupados funcionalmente en forma de redes y gráficos, además, combina una variedad de criterios de restricción para la

obtención de visualizaciones en diferentes niveles de especificidad (Bindea *et al.*, 2009). La construcción de las vías se realizó usando el coeficiente Kappa con una fuerza de concordancia considerable ($\kappa \geq 0,3$).

6.1.4. Reporte de genes que han sido estudiados en el cáncer gástrico

Finalmente, para determinar si los genes con variantes deletéreas se han estudiado con anterioridad en cáncer se utilizó la plataforma DisGenet, esta contiene una de las mayores colecciones de genes y variantes disponibles públicamente asociadas a enfermedades humanas. La base de datos DisGenet integra información de asociaciones de enfermedades de genes humanos (GDA) y asociaciones de enfermedades variantes (VDA) de varios repositorios, incluidas las enfermedades mendelianas, complejas y ambientales.

Para este estudio se usó una base de datos de DisGenet con información de 3.806 genes relacionados con neoplasia maligna de estómago. Además, se reportó el puntaje de asociación para los genes, este se encuentra entre los valores 0 y 1, 1 indica que todas las publicaciones apoyan la GDA o la VDA, mientras que valores menores a 1 indican que hay publicaciones que afirman que no existe asociación entre el gen o las variantes y la enfermedad. Si el gen o variante no tiene valor, indica que el puntaje no se ha calculado para esta asociación. (Piñero *et al.*, 2019).

7. RESULTADOS

7.1. Caracterización de la población

En el presente estudio se analizó el exoma secuenciado a partir del ADN extraído de 4 muestras de sangre de individuos sin infección por *H. pylori* diagnosticados con CG de tipo intestinal y provenientes de diferentes zonas de riesgo para la patología del departamento de Nariño. La distribución por sexo fue heterogénea (1 mujer, 3 hombres). El rango de edad estuvo entre 54 y 79 años, con una media de 64 años (Tabla 2).

Tabla 2.

Datos clinicopatológicos de cuatro pacientes diagnosticados con cáncer gástrico del departamento de Nariño.

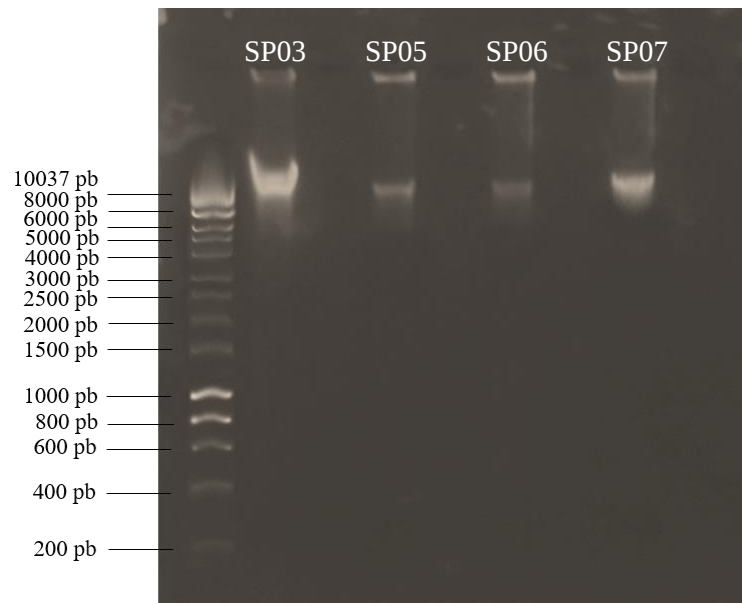
N	ID	Género/Edad	Diagnóstico de infección por <i>H. pylori</i>	Diagnóstico/ Clasificación de Lauren/ Clasificación de Borrmann	Procedencia	Incidencia de CG
1	SP03	M/79	Negativo	Cáncer gástrico/ Intestinal/ Tipo 0	Tumaco bajo	Baja
2	SP05	M/54	Negativo	Cáncer gástrico/ Intestinal/ Tipo 0	Samaniego	Alta
3	SP06	F/62	Negativo	Cáncer gástrico/ Intestinal/ Tipo V	Pasto	Intermedia
4	SP07	M/63	Negativo	Cáncer gástrico/ Intestinal/ Tipo V	Consacá	Intermedia

CG: Cáncer gástrico

7.2. Control de calidad, cantidad y pureza de las muestras

El protocolo de extracción usando el kit (*UltraClean™ Blood DNA Isolation*), permitió la obtención de ADN de buena calidad. No obstante, se evidenció rastros de barrido y bandas tenues en algunas muestras, probablemente debido a la formación de coágulos, que impidieron el paso de la muestra a través de la columna de extracción y un correcto lavado y purificación de las mismas (Figura 6).

Figura 6. Análisis de la calidad del DNA mediante electroforesis en gel de agarosa 1.5%. M: marcador de tamaño molecular de 1 Kb. Condiciones de corrida: 100 V, 140 en buffer TAE 1X, usando cámara de electroforesis horizontal Scie Plas Ltd. Fuente: Esta investigación.



En cuanto a la cuantificación de las muestras se obtuvo una relación A260/280 de 1,8-2, todas las muestras se ajustaron a una concentración de 50 ng/μL para su respectiva secuenciación (Tabla 3).

Tabla 3.

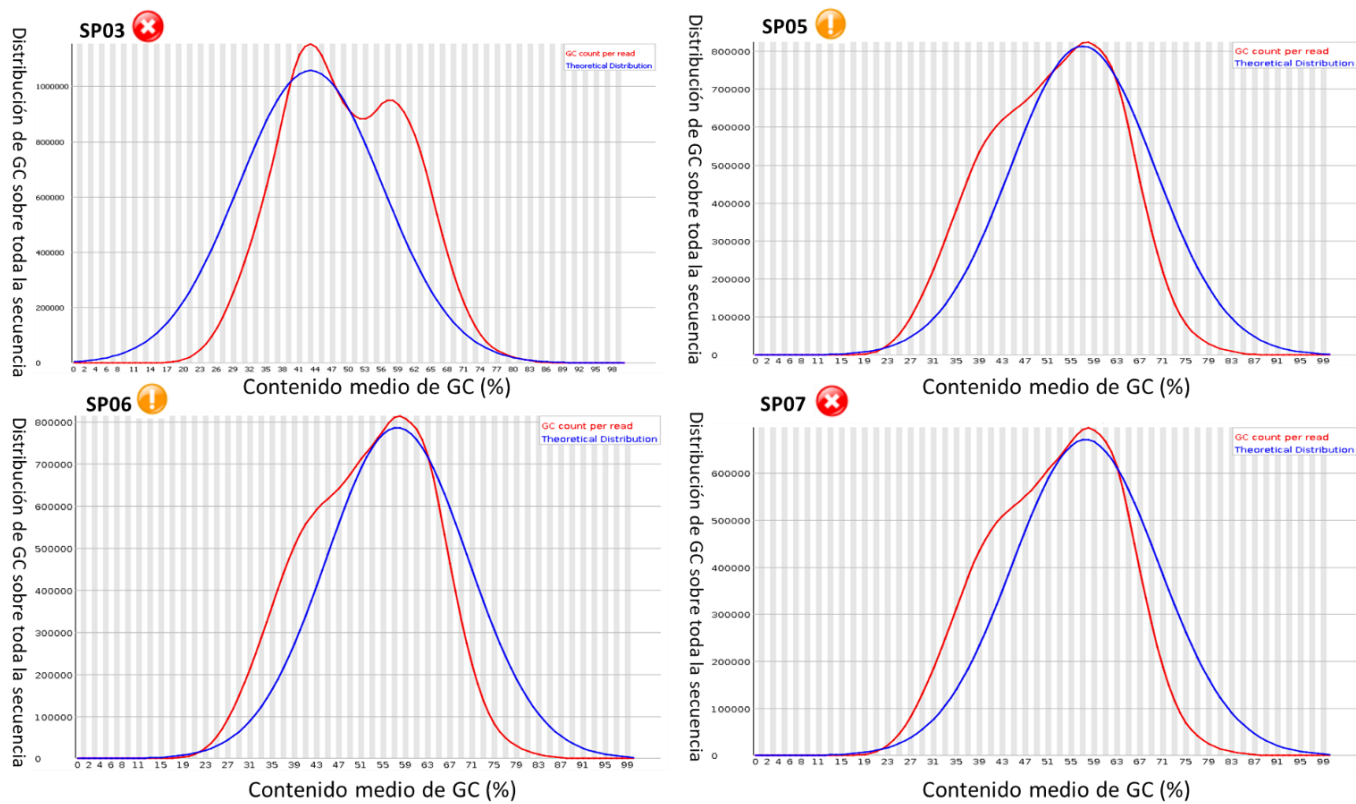
Cuantificación de las muestras en Nanodrop 2000

Muestra	Concentración Nanodrop 2000	Relación 260/280
SP03	135,9 ng/μL	1,91
SP05	124,8 ng/μL	1,81
SP06	76,2 ng/μL	1,81
SP07	110,3 ng/μL	1,86

7.3 Calidad de las lecturas

En general las estadísticas del programa FastQC indicaron una buena calidad de las secuencias respecto a los diferentes parámetros citados anteriormente. No obstante, en todas las muestras se obtuvo un contenido de guanina y citosina (g-c) mayor al esperado (Figura 7).

Figura 7. Calidad de contenido de guanina y citosina por secuencia de las cuatro muestras analizadas.
Fuente: Esta investigación.



7.4 Alineamiento de secuencias

Las lecturas obtenidas para cada paciente fueron mapeadas de manera independiente frente al genoma de referencia humano de acuerdo a lo descrito en métodos. Para las muestras de estudio el porcentaje de mapeo fue mayor a 99% (Tabla 4).

Tabla 4.
Estadísticas de alineamiento

ID	<i>Reads</i> totales	<i>Reads</i> mapeadas	<i>Reads</i> mapeadas (%)
SP03	69.534.672	69.458.184	99,89
SP05	51.400.046	51.245.845	99,7
SP06	50.193.962	50.008.244	99,63
SP07	42.933.884	42.762.148	99,6

7.5 Llamado de variantes

Con Haplotypcaller (HC) se obtuvo un archivo VCF para cada una de las muestras secuenciadas, este contenía el número total de variantes llamadas que oscilo entre 300.000 y 600.000 variantes. Posteriormente, se obtuvieron las mutaciones que pasaron la aplicación de los filtros rigurosos. El mayor número de alteraciones genéticas estuvieron presentes en la muestra SP05 (Tabla 5).

Tabla 5.

Datos generales resultantes del llamado de variantes con HaploTypeCaller de las cuatro muestras analizadas

Muestra	Llamado de variantes total	VARIANTES FILTRADAS	
		SNPs	INDELs
SP03	398.181	349.968	33.825
SP05	616.028	535.388	66.140
SP06	585.603	508.295	62.808
SP07	524.524	455.208	54.087

Con Freebayes (Fb), se obtuvo para cada muestra analizada un archivo VCF que contenía el número total de variantes llamadas que oscilo entre 400.000 y 600.000. Posteriormente, se obtuvieron las variantes que pasaron la aplicación de los filtros rigurosos. Igual que con HC, el mayor número de alteraciones genéticas estuvieron presentes en la muestra SP05 (Tabla 6).

Tabla 6.

Datos generales resultantes del llamado de variantes con Freebayes de las cuatro muestras analizadas

Muestra	Llamado de variantes total	VARIANTES TOTALES QUE PASARON EL FILTRO	
		SNPs totales	INDELs
SP03	415.068	371.324	30.203
SP05	670.365	575.608	67.162
SP06	631.066	542.428	62.135
SP07	449.818	391.016	43.533

A continuación, se muestran los datos resultantes del número de variantes únicas y compartidas llamadas por HC y Fb para los SNPs e INDELs (Tabla 7).

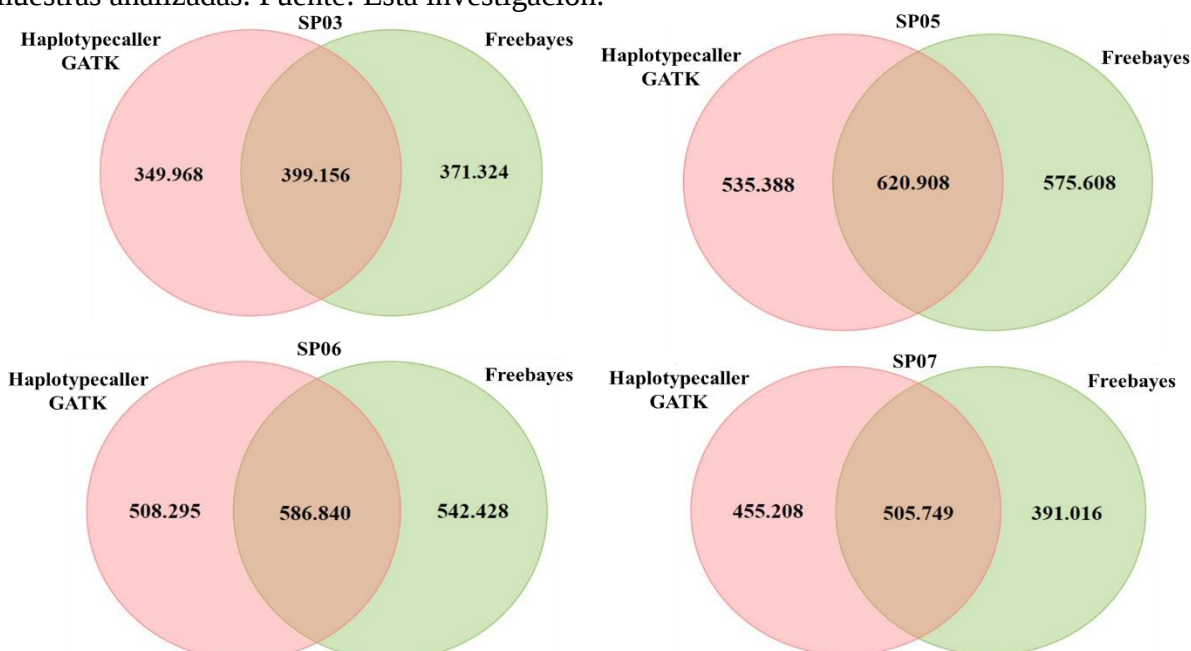
Tabla 7.

Variantes únicas y compartidas de los llamadores HaplotypeCaller y Freebayes para las cuatro muestras analizadas

Tipo de variante	Muestra Llamador de variantes	SP03		SP05		SP06		SP07	
		HC	Fb	HC	Fb	HC	Fb	HC	Fb
SNPs	Únicos	49.188	27.832	85.393	45.193	78.545	44.412	50.541	114.733
	Compartidos	322.136		490.088		463.883			
INDELs	Únicos	4.936	8.558	12.689	11.667	11.846	12.519	7.496	18.050
	Compartidos	25.267		54.473		50.289		36.037	

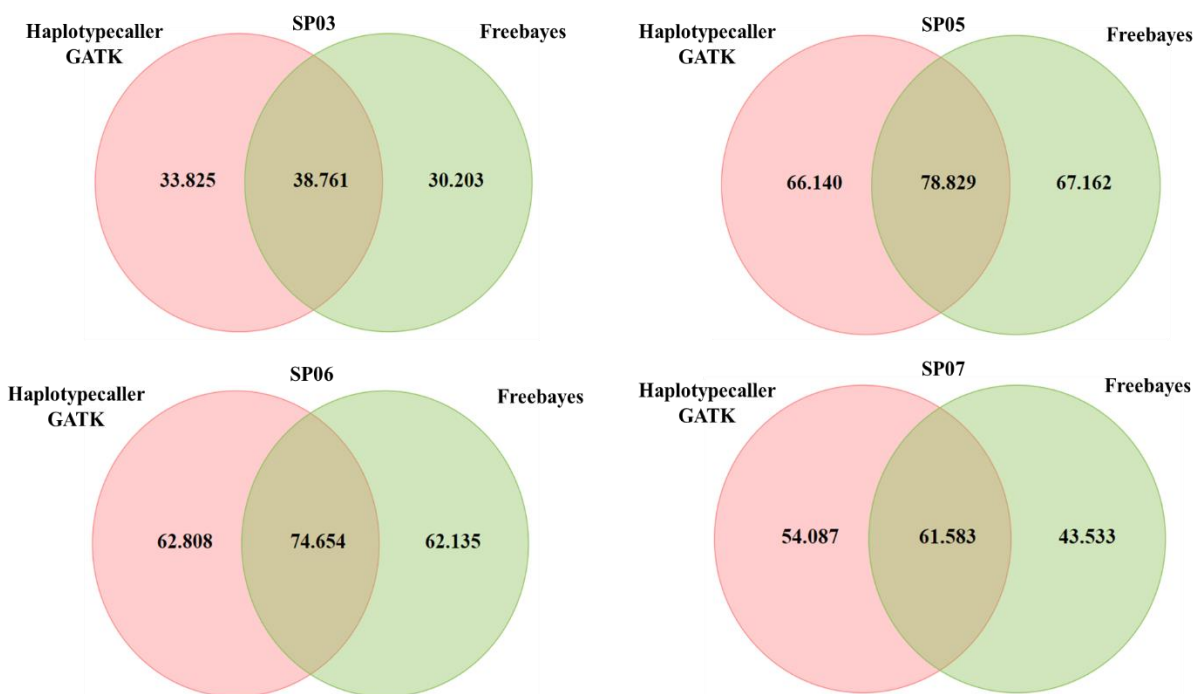
Con el fin de incluir un número mayor de variantes, se unificaron los SNPs llamados por ambos callers para cada una de las muestras, el mismo proceso se realizó para los INDELs. En la Figura 8, se observan los SNPs totales y por cada llamador. La muestra SP05 abarcó un mayor número de alteraciones genéticas, contrario a la muestra SP03 que presentó variantes por debajo de 400.000.

Figura 8. Total de SNPs resultantes del llamado de HaplotypeCaller y Freebayes en las cuatro muestras analizadas. Fuente: Esta investigación.



En la Figura 9, se observan los INDELS totales y por cada llamador. Al igual que HC, la muestra SP05 abarcó un mayor número de alteraciones genéticas, seguida de la muestra SP06, SP07 y SP03.

Figura 9. Total de INDELS resultantes del llamado de HaplotypeCaller y Freebayes en las cuatro muestras analizadas. Fuente: Esta investigación.

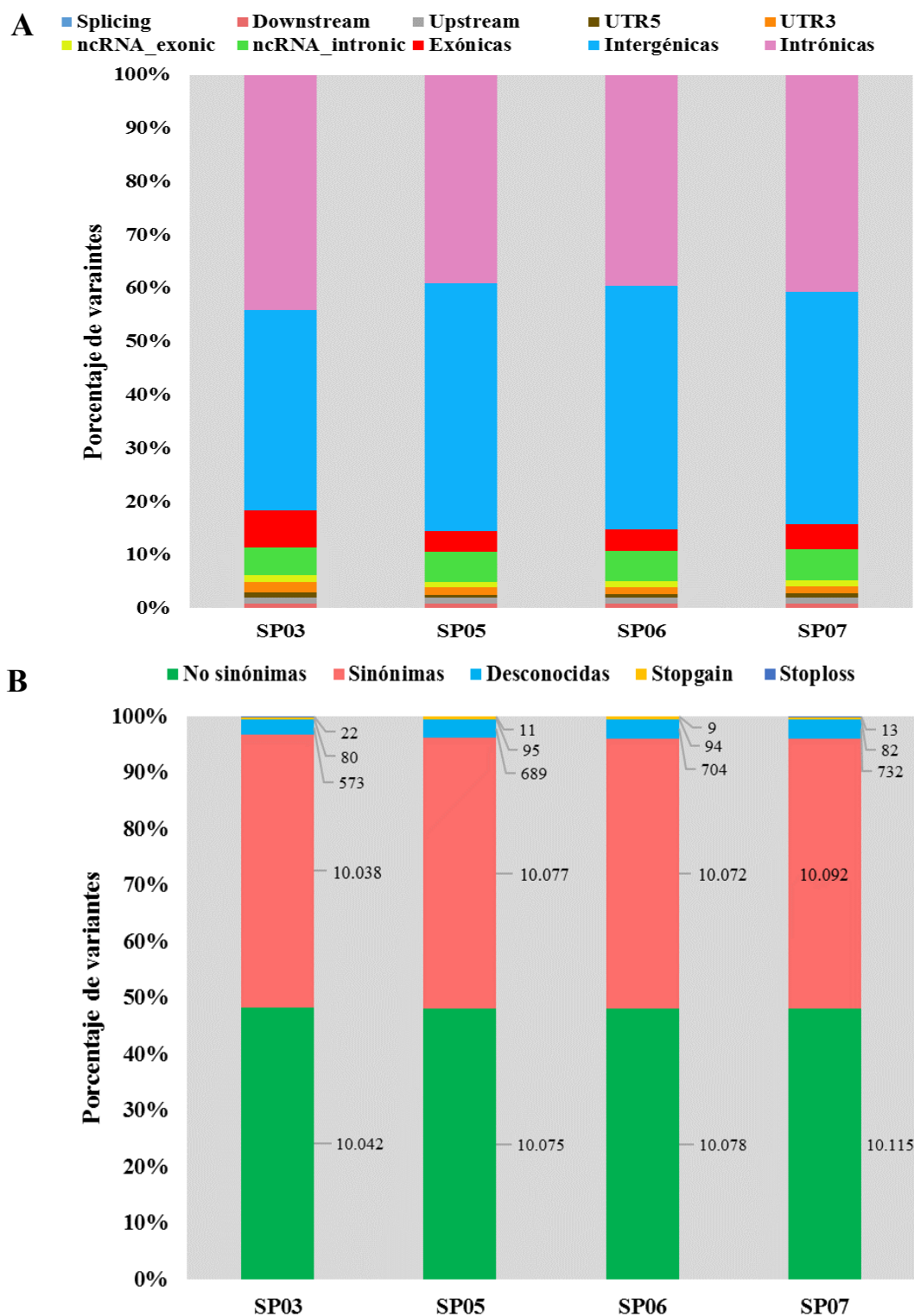


7.6 Clasificación de las variantes genómicas

- De acuerdo a la posición genómica e impacto a nivel molecular.

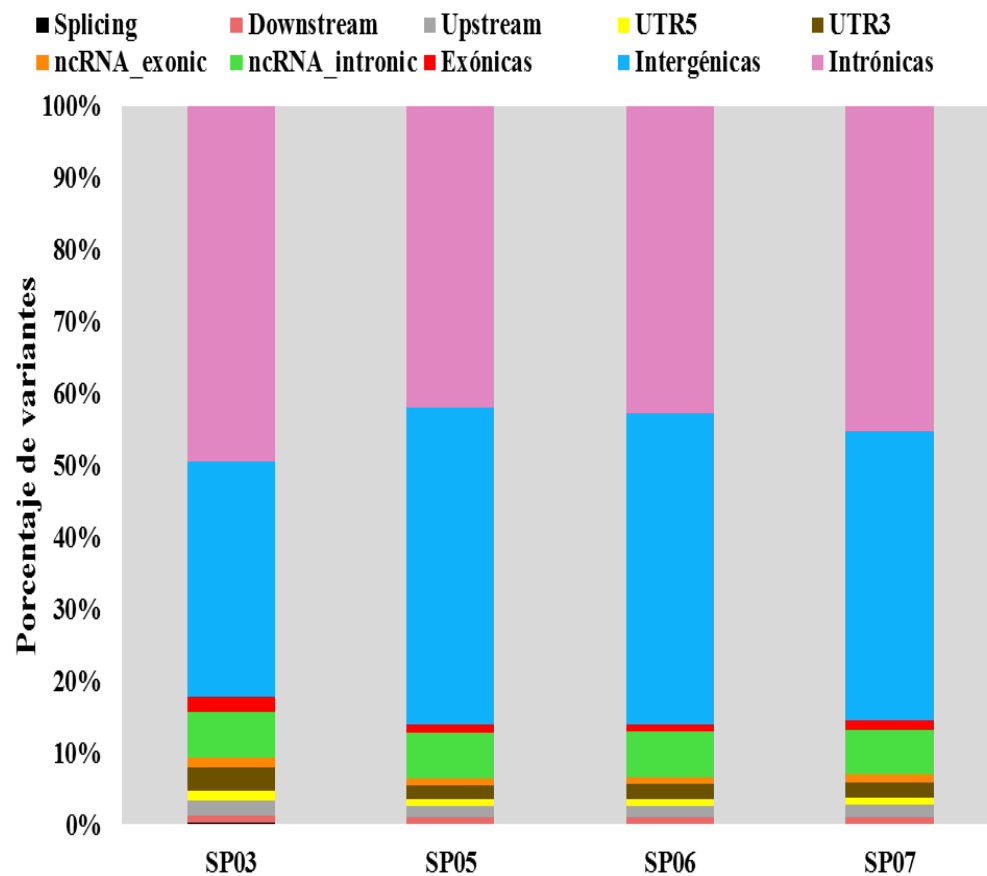
En general, en el archivo VCF unificado de los cuatro exomas se obtuvieron un total de 1.597.012 SNPs de tipo germinal, para cada muestra se anotaron el siguiente número de variantes: 400.138 (SP03), 622.337 (SP05), 588.209 (SP06) y 506.868 (SP07). Las alteraciones genéticas se encontraron ubicadas en diferentes regiones del genoma (Figura 10A). Respecto a la posición genómica solo se tuvieron en cuenta variantes exónicas que representaron más del 3% en cada muestra, no obstante, el mayor porcentaje se observó en la muestra SP03 (Figura 10A). De acuerdo al impacto a nivel molecular, se incluyeron aquellas no sinónimas y de pérdida y ganancia de codón de parada, que abarcaron cerca de 48% en todas las muestras (Figura 10B).

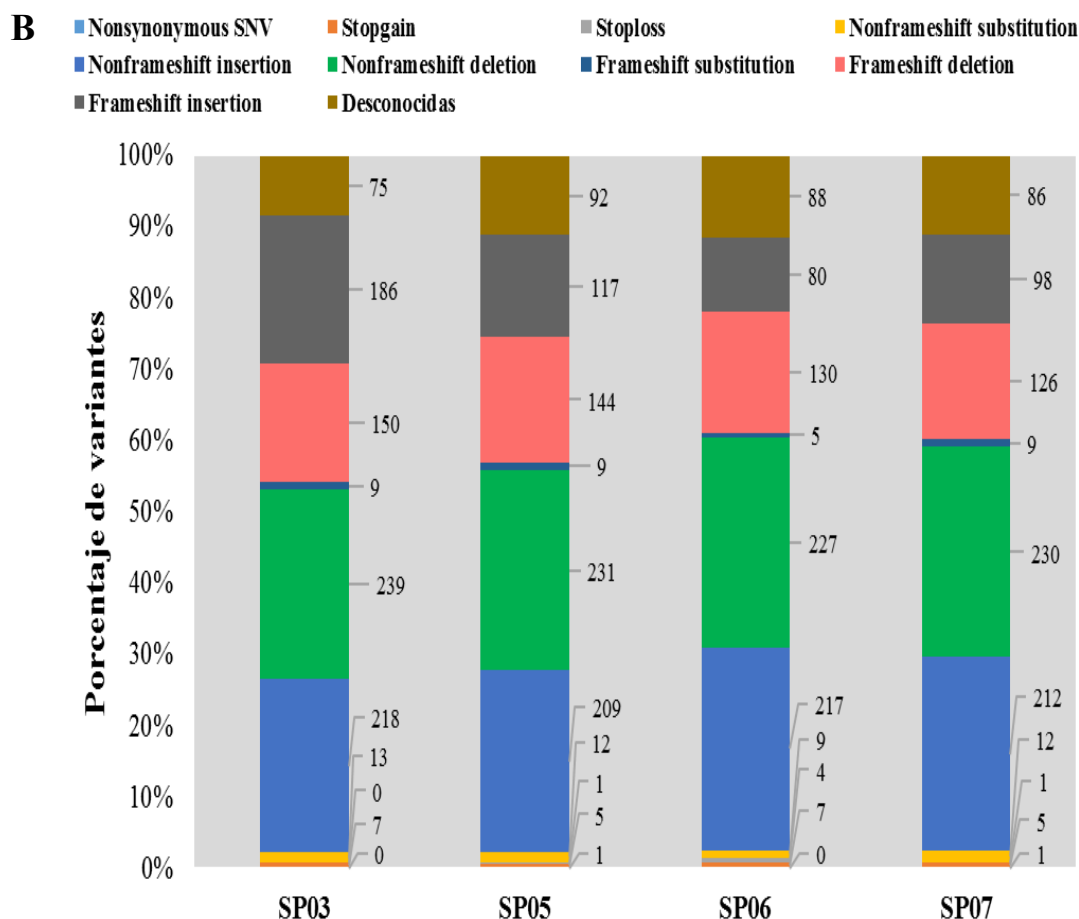
Figura 10. Porcentaje de SNPs de origen germinal de cuatro pacientes con cáncer gástrico. A. Variantes germinales por posición genómica. B. Variantes germinales exónicas clasificadas por su impacto a nivel molecular. Fuente: Esta investigación.



En cuanto a los INDELs se obtuvieron un total de 203.133 variantes genéticas, para cada muestra se anotaron el siguiente número de alteraciones genéticas: 36.663 (SP03), 75.642 (SP06), 72.143 (SP07) y 60.690 (SP08). Respecto a la posición genómica solo se tuvieron en cuenta variantes exónicas, que representaron entre el 1 y 2% para las muestras analizadas, al igual que los SNPs, el mayor porcentaje se observó en la muestra SP03 (Figura 11A). Las delecciones, inserciones, sustituciones, variantes de ganancia y pérdida de codón de parada de tipo exónico representaron 34,7%, las que fueron incluidas en la investigación (Figura 11B).

Figura 11. Porcentaje de INDELs de origen germinal de cuatro pacientes con cáncer gástrico del departamento de Nariño. A. Variantes germinales por posición genómica. B. Variantes germinales exónicas clasificadas por su impacto a nivel molecular. Fuente: Esta investigación.



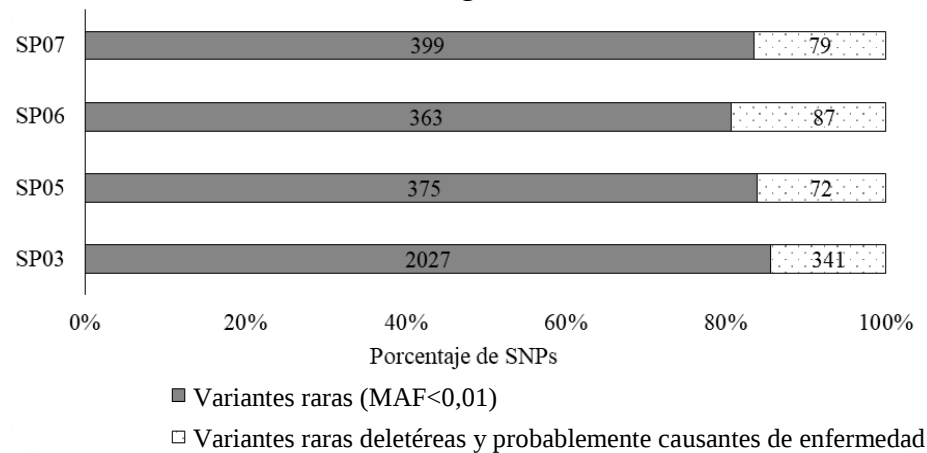


- De acuerdo a la frecuencia del alelo menor (MAF)

Se clasificaron los SNPs obtenidos anteriormente, resultando 15.269 variantes comunes ($MAF > 0,05$), 2.399 de baja frecuencia ($MAF < 0,05$ hasta 0,01) y 3.533 raras ($MAF < 0,01$). De las mutaciones encontradas se seleccionaron solamente las de reporte raro, en la Figura 12 se reporta el total de estas variantes para cada muestra. Posteriormente, de estas alteraciones se tuvieron en cuenta las deletéreas (SIFT) y probablemente causantes de enfermedad (Mutation taster).

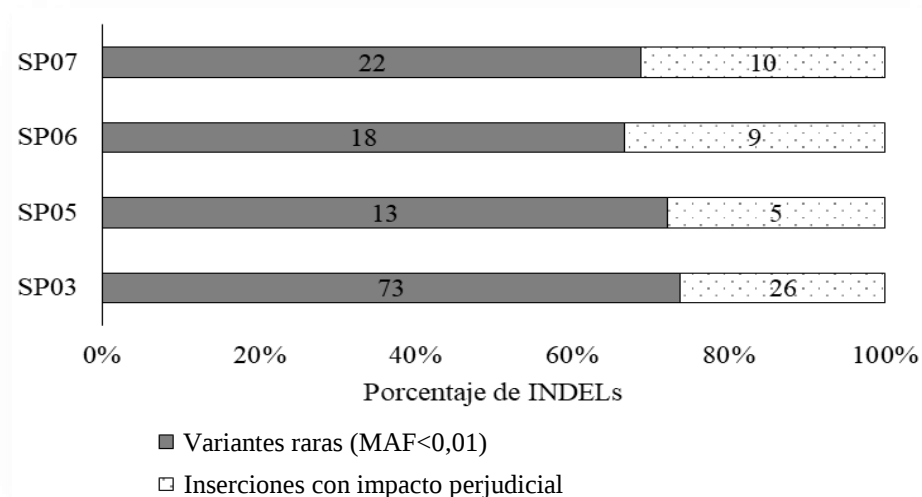
En el anexo 5 se describen las alteraciones genéticas filtradas y los genes que fueron usados en la construcción de las redes. Se encontraron para SP03 el mayor número de variantes (341) presentes en 330 genes y para SP05, SP06 y SP07 un total de 72, 87 y 79 mutaciones puntuales en 72, 85, y 77 genes, respectivamente (Figura 12).

Figura 12. Número y porcentaje de SNPs raros, deletéreos y probablemente causantes de enfermedad en cuatro muestras estudiadas. Fuente: Esta investigación.



Para los INDELS se clasificaron 197 variantes comunes, 57 de baja frecuencia y 138 variantes raras, de estas últimas 3 se clasificaron de alto impacto y 57 moderadas como se evidencia en el anexo 6, en donde se detallan las alteraciones genéticas y los genes que fueron usados para la construcción de redes. Para SP03, SP05, SP06 y SP07 solamente se observaron inserciones probablemente perjudiciales, 26, 5, 9 y 10 variantes, correspondientemente (Figura 13).

Figura 13. Número y porcentaje de inserciones raras y probablemente perjudiciales en las cuatro muestras estudiadas. Fuente: Esta investigación.



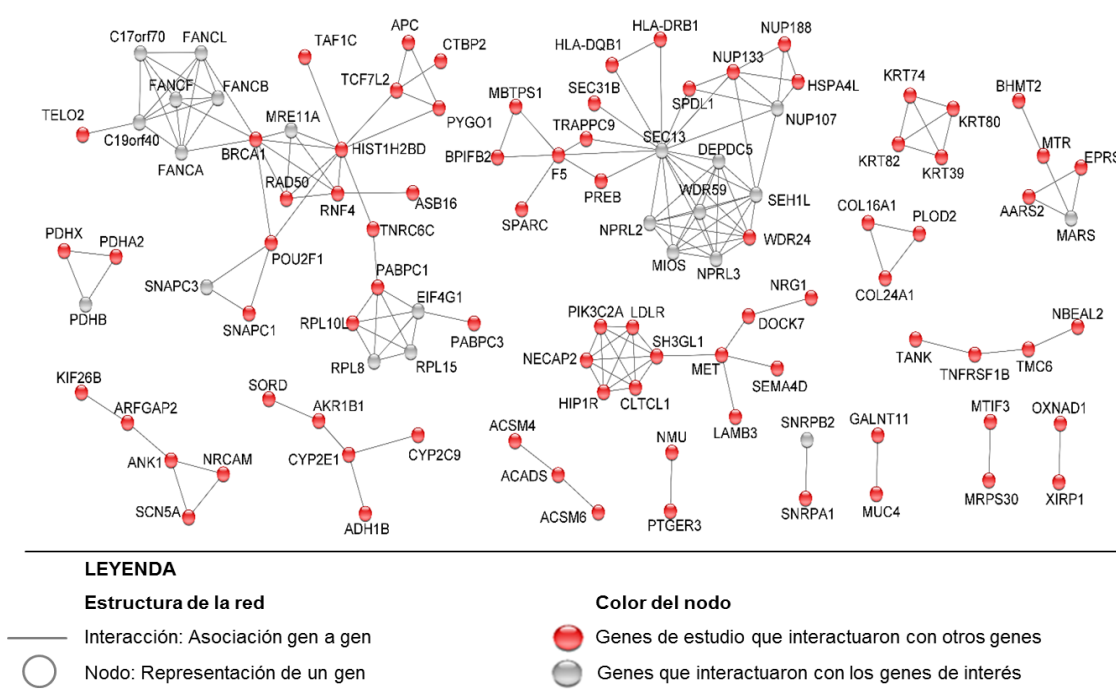
7.7 Vías de interacciones entre genes

En la construcción de las redes de interacción para todas las muestras se tuvieron en cuenta los genes que presentaron SNPs no sinónimos, raros, con predicción deletérea y posiblemente causantes de enfermedad e inserciones con alto o moderado impacto, descritas en el Anexo 5 y 6. Las vías que se describen a continuación están conformadas por componentes: subgrupo de interacción génica, es decir, los genes están representados por nodos en forma de circunferencia unidos por líneas que denotan las interacciones entre estos, los nodos rojos simbolizan los genes candidatos y los nodos en color gris los genes con los cuales interactuaron. Es preciso mencionar que cada una de estas redes presentó un modelo a libre escala donde una gran parte de los nodos de la red poseen pocas interacciones y reducidos nodos presentan un gran número de interacciones.

7.7.1 Vías para genes con presencia de SNPs

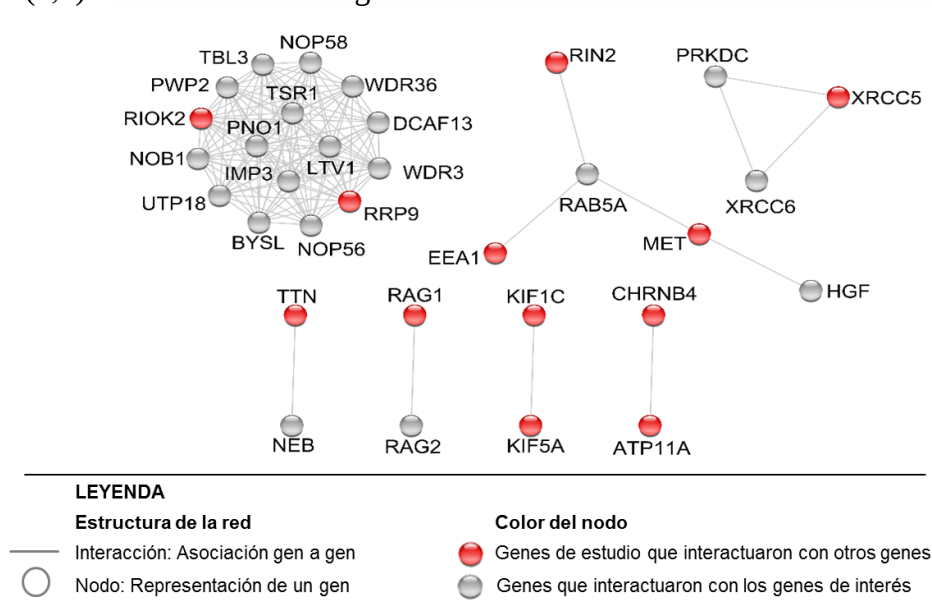
Para SP03 se obtuvo la mayor red de genes, ya que se construyó a partir de 330 genes candidatos con diferentes mutaciones, probablemente perjudiciales, la vía presentó 16 componentes, 103 nodos y 165 interacciones, de los genes de interés 81 presentaron interacción con otros 22 genes que conforman la red, esto según la base de datos de STRING (Figura 14).

Figura 14. Red de genes con presencia de mutaciones deletéreas para la muestra SP03 usando la base de datos STRING. Visualización de confianza alta (0,7). Fuente: Esta investigación.



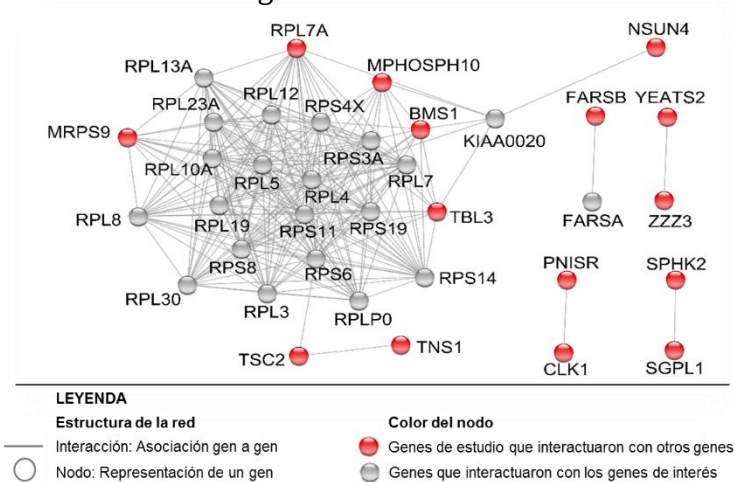
En la vía de interacción de SP05 se presentaron 32 nodos, 131 interacciones y 7 componentes conectados, la red se construyó con 72 genes candidatos, de estos solo en 12 se observó interacción con otros 20 genes (Figura 15).

Figura 45. Red de genes con presencia de mutaciones deletéreas para la muestra SP05 usando la base de datos STRING. Visualización de confianza alta (0,7). Fuente: Esta investigación.



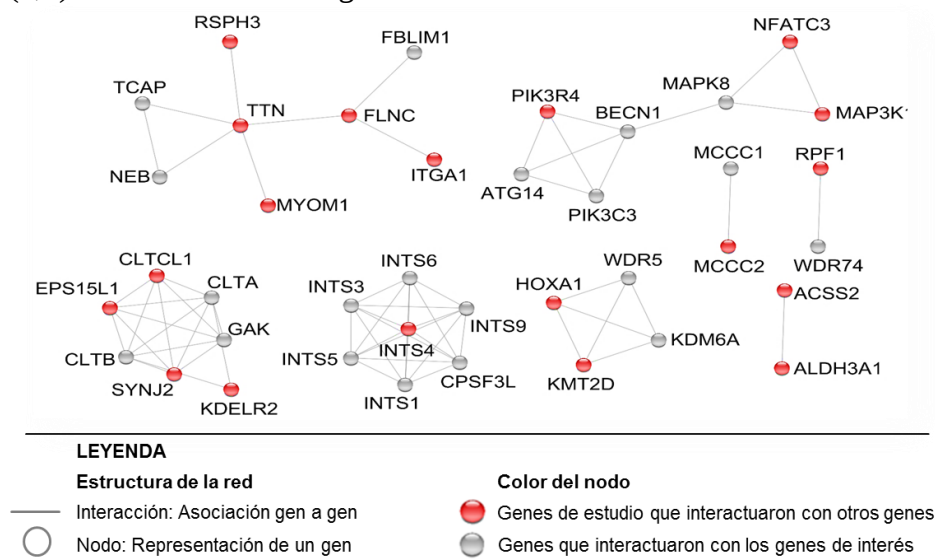
Para SP06 se obtuvo una red a partir de los 85 genes de interés, la vía estuvo conformada por 5 componentes y 328 interacciones entre 36 nodos, de los cuales 15 presentaron cierta interacción con otros 21 genes, los genes con más interacciones se encuentran en el mayor componente (Figura 16).

Figura 56. Red de genes con presencia de mutaciones deletéreas para la muestra SP06 usando la base de datos STRING. Visualización de confianza alta (0,7). Fuente: Esta investigación.



En el caso de SP07 la vía presentó 8 componentes con 39 nodos y 65 interacciones, de los 79 genes con variantes deletéreas 19 tuvieron alguna interacción con otros 20 genes (Figura 17).

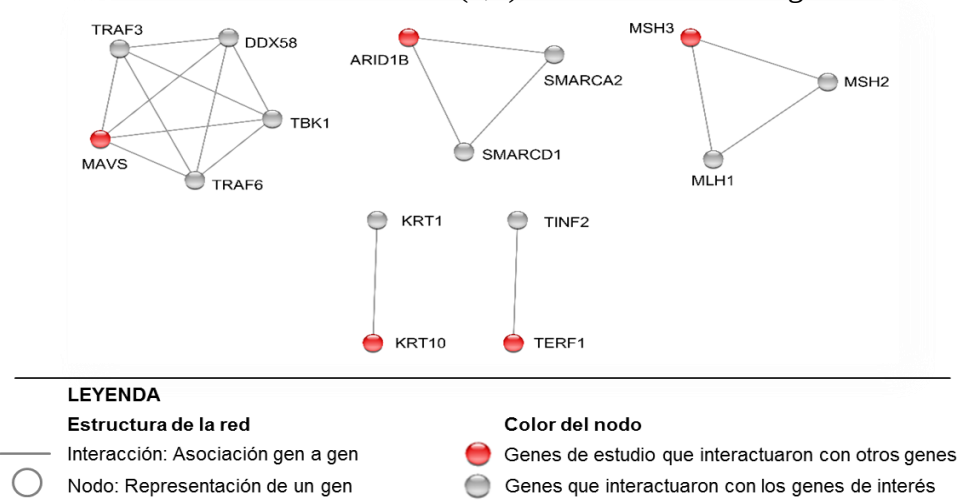
Figura 17. Red de genes con presencia de mutaciones deletéreas para la muestra SP07 usando la base de datos STRING. Visualización de confianza alta (0,7). Fuente: Esta investigación.



7.7.2. Vías para genes con presencia de inserciones

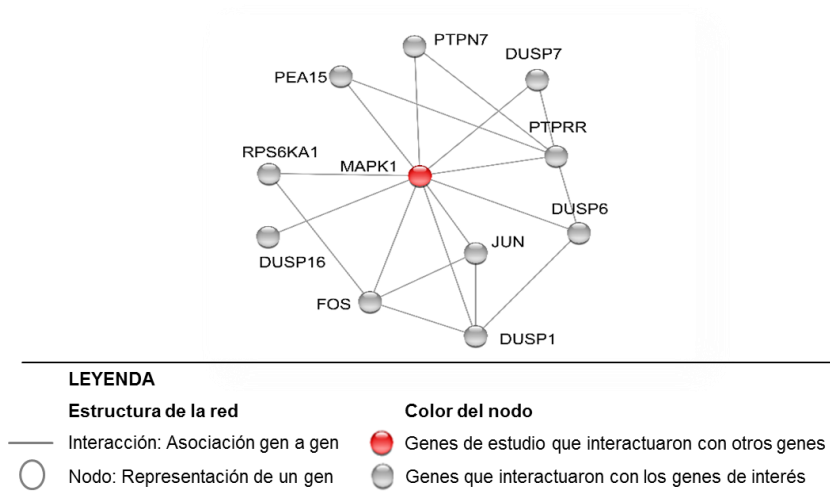
La ruta de SP03 presentó 5 componentes y 18 interacciones entre 15 nodos, de estos genes 5 fueron de interés (Figura 18).

Figura 18. Red de genes con presencia de inserciones perjudiciales para la muestra SP03 usando la base de datos STRING. Visualización de confianza alta (0,7). Fuente: Esta investigación.



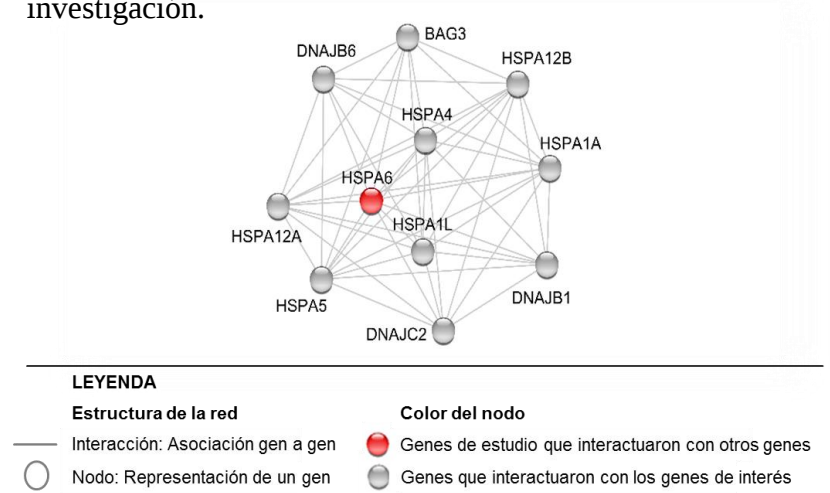
Para SP05 se obtuvo un solo componente, de los 5 genes candidatos mencionados con anterioridad solamente MAPK1 se encontró interactuando en la vía que incluyó otros 10 nodos y un total de 19 interacciones (Figura 19).

Figura 19. Red de genes con presencia de inserciones perjudiciales para la muestra SP05 con una visualización de confianza alta (0,7). Fuente: Esta investigación.



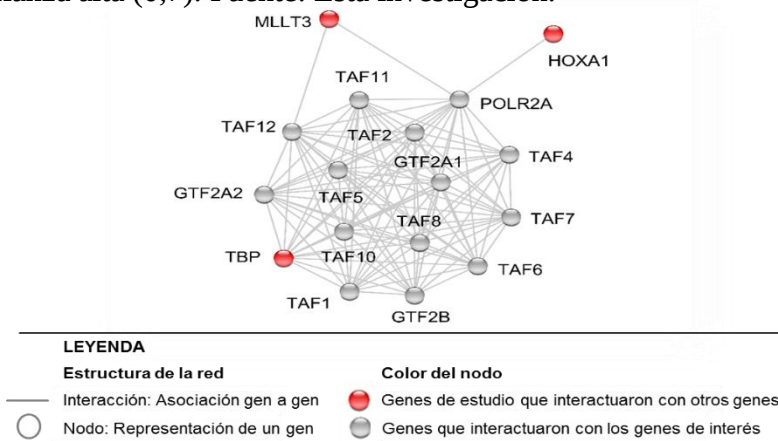
En el caso de SP06 el gen HSPA6 fue el único que apareció en la red de interacción con otros 10 nodos, constituyendo un solo componente y 51 interacciones (Figura 20).

Figura 20. Red de genes con presencia de inserciones perjudiciales para la muestra SP06 usando la base de datos STRING. Visualización de confianza alta (0,7). Fuente: Esta investigación.



Finalmente, para la vía de SP07 se consiguió un componente con 108 interacciones entre los diferentes nodos que constituyeron la red, 3 de los genes candidatos para esta muestra interactuaron con otros 16 nodos (Figura 21).

Figura 21. Red de genes con presencia de inserciones perjudiciales para la muestra SP07 usando la base de datos STRING. Visualización de confianza alta (0,7). Fuente: Esta investigación.



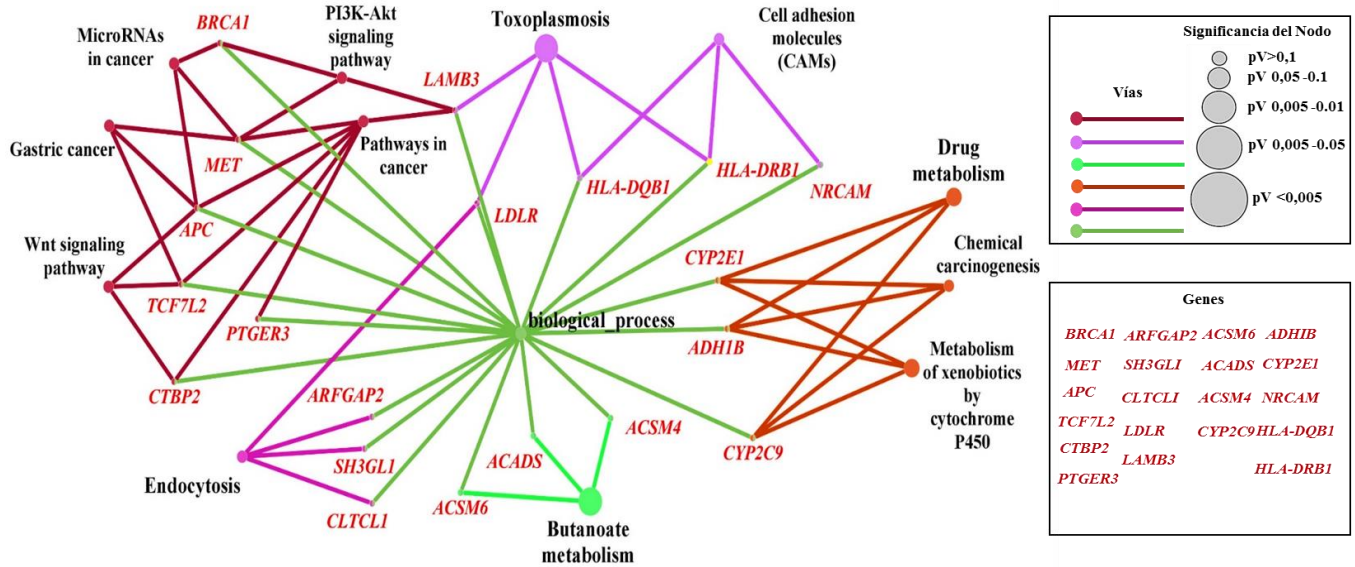
7.8 Análisis funcional

Una vez se identificaron los genes con mayor significancia que interactuaban en las diferentes vías génicas establecidas se procedió a realizar un análisis funcional para facilitar la interpretación biológica de los genes y visualizarlos de manera agrupada. El resultado del análisis se muestra como una red que incluye cada nodo como un término y las líneas de colores iguales definen la pertenencia al grupo.

7.8.1. Rutas de genes con presencia de SNPs

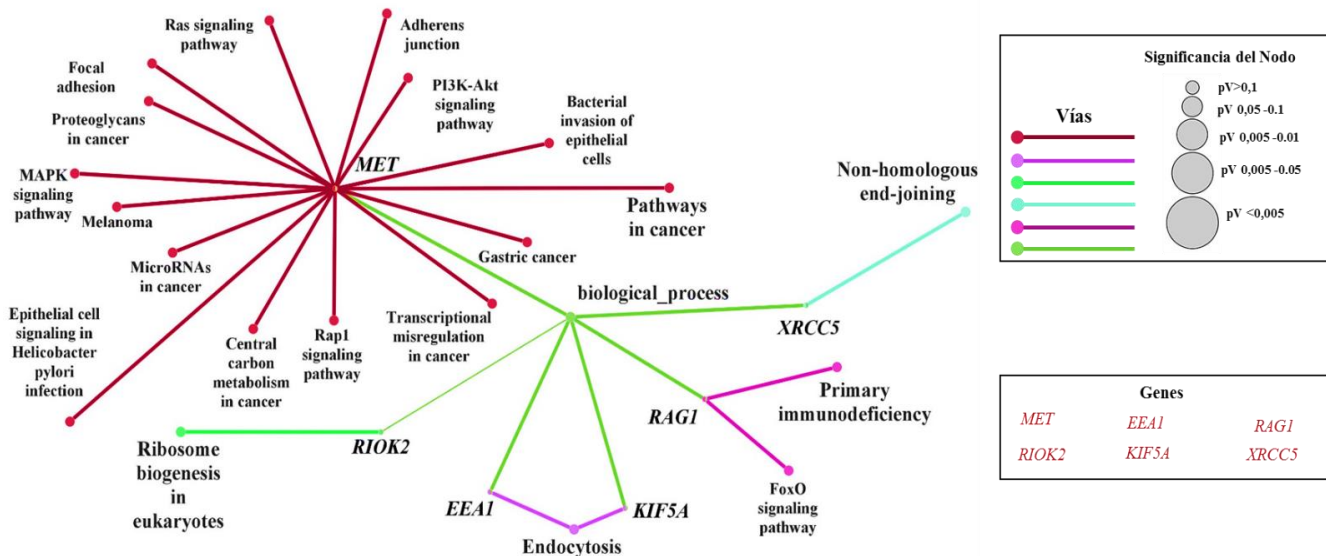
Para SP03 los 81 genes se involucraron con procesos biológicos, “*Pathways in cancer*” (KEGG: 05200) presentó el mayor número de genes, en total 7 (*APC*, *BRAC1*, *CTBP*, *LAMB3*, *MET*, *PTGER3*, *TCFLT2*), algunos de estos están en la vía del CG (KEGG: 05226), y algunas mutaciones podrían afectar su funcionamiento (Figura 22). Las rutas con mayor significancia fueron “*Butanoate Metabolism*” ($p= 0,01$), infección por “*Staphylococcus aureus infection*” ($p= 0,04$), “*Toxoplasmosis*” ($p= 0,04$), los genes *ADH1B*, *CYP2C9*, *CYP2E1* se encontraron involucrados en la vía “*Metabolism of xenobiotics by cytochrome P450*” ($p= 0,05$), “*Drug metabolism*” y “*Chemical carcinogenesis*” ($p=0,04$), rutas involucradas en el proceso de carcinogénesis.

Figura 22. Red agrupada funcionalmente de genes con presencia de SNPs deletéreos para SP03 enriquecida con la base de datos GO y KEGG con términos como nodos vinculados en función de su nivel de puntuación $\kappa(\geq 0,3)$. Fuente: Esta investigación.



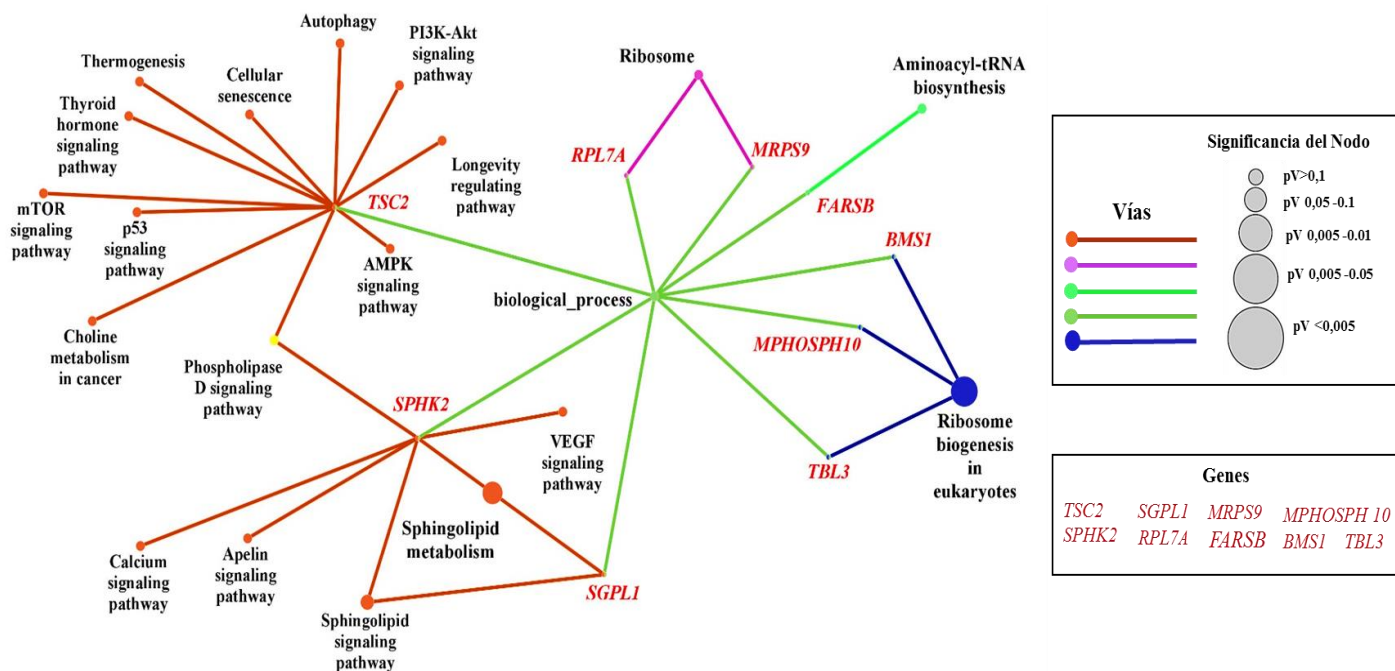
Para SP05 la agrupación se hace alrededor de los procesos biológicos en los que están involucrados cada gen (Figura 23), XRCC5 estuvo presente en la ruta más significativa que fue “Non-homologous end-joining” ($p = 0,01$), las vías que siguieron en mayor importancia fueron “Endocytosis” ($p = 0,02$) y “Primary immunodeficiency” ($p = 0,04$). Cabe resaltar que MET estuvo presente al igual que en la red de SP03.

Figura 23. Red agrupada funcionalmente de genes con presencia de SNPs deletéreos para SP05 enriquecida con la base de datos GO y KEGG con términos como nodos vinculados en función de su nivel de puntuación $\kappa(\geq 0,3)$. Fuente: Esta investigación.



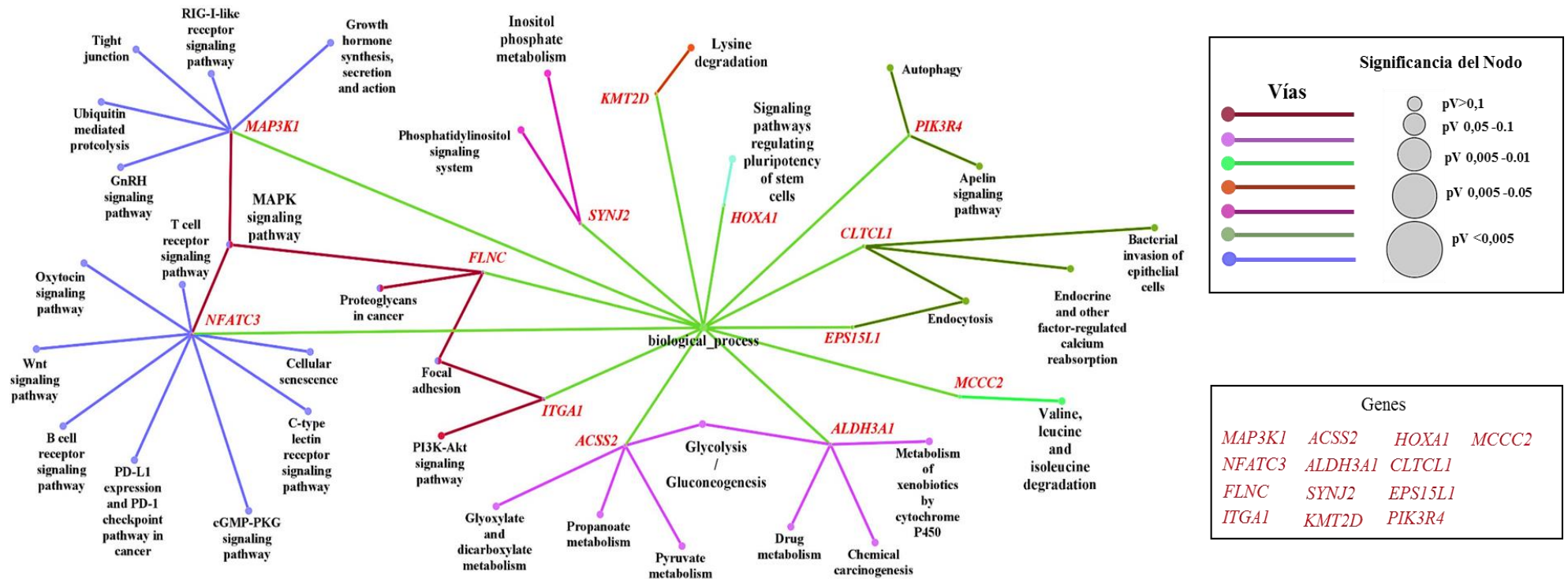
En el caso de SP06 de los genes escogidos para la red, 9 estuvieron involucrados en diferentes rutas, nuevamente el agrupamiento en general se realizó en torno al proceso biológico (Figura 24), *BMS1*, *MPHOSPH10*, *TBL3* se relacionaron con la vía representativa “*Ribosome biogenesis in eukaryotes*” ($p=0,001$). Posteriormente, siguió la vía “*Sphingolipid signaling*” ($p=0,003$) y “*Phospholipase D signaling*” ($p=0,006$). Es importante señalar el gen *TSC2* involucrado en rutas como “*mTor signaling pathway*”.

Figura 6. Red agrupada funcionalmente de genes con presencia de SNPs deletéreos para SP06 enriquecida con la base de datos GO y KEGG con términos como nodos vinculados en función de su nivel de puntuación $\kappa(\geq 0,3)$. Fuente: Esta investigación.



Los genes con mutaciones puntuales de SP07 fueron los que tuvieron mayor presencia en diferentes vías (Figura 25), “*Glycolysis/Gluconeogenesis*” ($p=0,002$) fue la ruta más representativa, seguido de “*MAPK signaling pathway*” ($p=0,003$) relacionada con *FLNC*, *MAP3K1*, *NFATC3*. *FLNC* también estuvo involucrado en vías como “*Proteoglycans in cancer*”; este gen junto *ITGA1* estuvieron presentes en “*Focal adhesion*”. Destacar las vías en las que se involucra el gen *NFATC3* como “*Wnt signaling pathway*” y “*Cellular senescence*”. Además, *HOXA1* se asoció con “*Signaling pathways regulating pluripotency of stem cells*”. Se presentaron al igual que en otras muestras las vías: “*Metabolism of xenobiotics by cytochrome P450*”, “*Chemical carcinogenesis*” y “*Endocytosis*”.

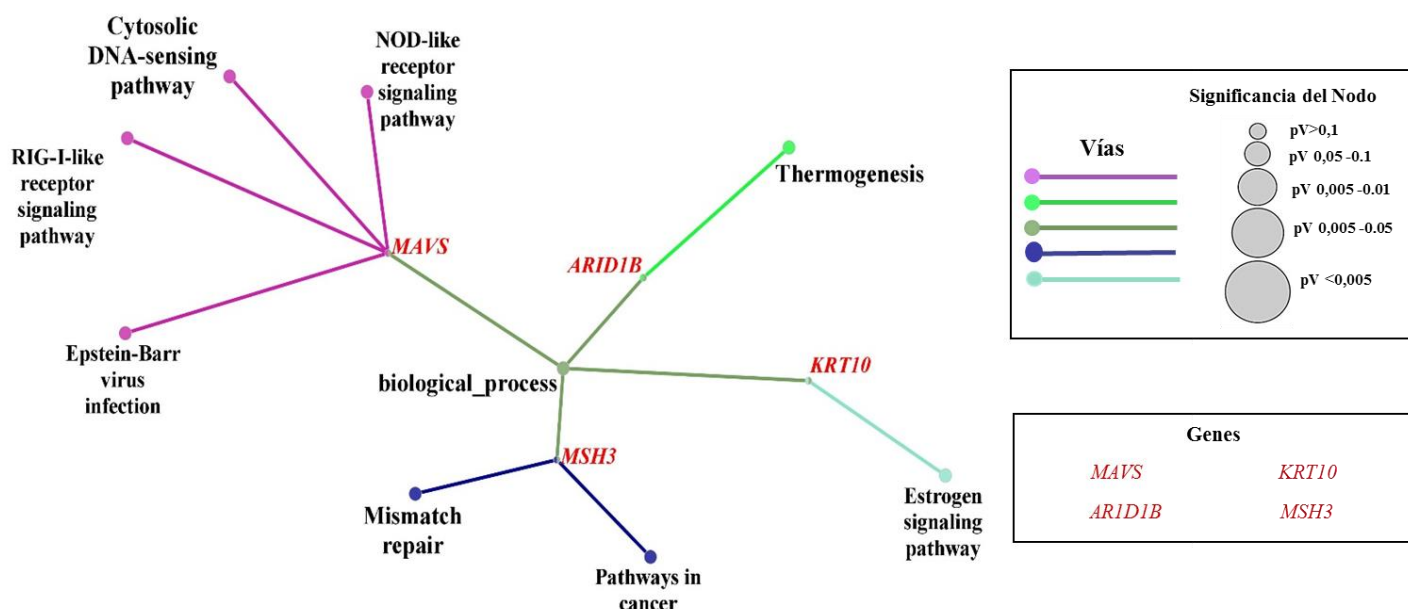
Figura 257. Red agrupada funcionalmente de genes con presencia de SNPs deletéreos para SP07 enriquecida con la base de datos GO y KEGG con términos como nodos vinculados en función de su nivel de puntuación $\kappa=(\geq 0,3)$. Fuente: Esta investigación.



7.8.2. Rutas de genes con presencia de inserciones

La red de SP03 indicó que el gen *MSH3* está involucrado en la vía más significativa “Mismatch repair” ($p=0,006$). Además, este gen hace parte de “Pathways in cancer”. MAVS estuvo presente en rutas de reconocimiento como “Cytosolic DNA-sensing” (0,017). Como se puede apreciar en la Figura 26 *AIRID1* está involucrado en “Hepatocellular carcinoma” así como en otros tipos de cáncer.

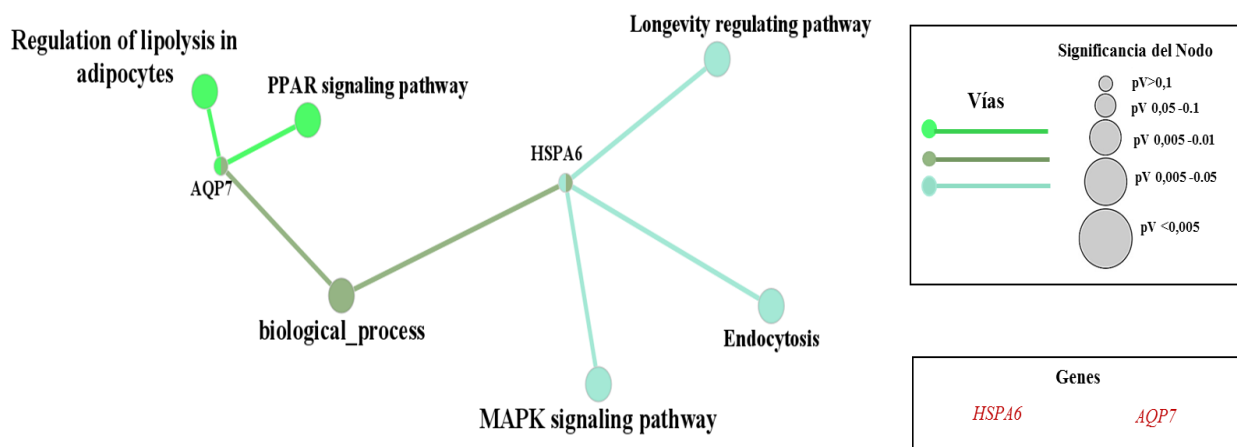
Figura 26. Red agrupada funcionalmente de genes con presencia de inserciones perjudiciales para SP03 enriquecida con la base de datos GO y KEGG con términos como nodos vinculados en función de su nivel de puntuación $\kappa=(\geq 0,3)$. Fuente: Esta investigación.



El análisis funcional para SP05 no se muestra ya que *MAPK1* participa en más de 80 rutas y en distintas enfermedades como el CG.

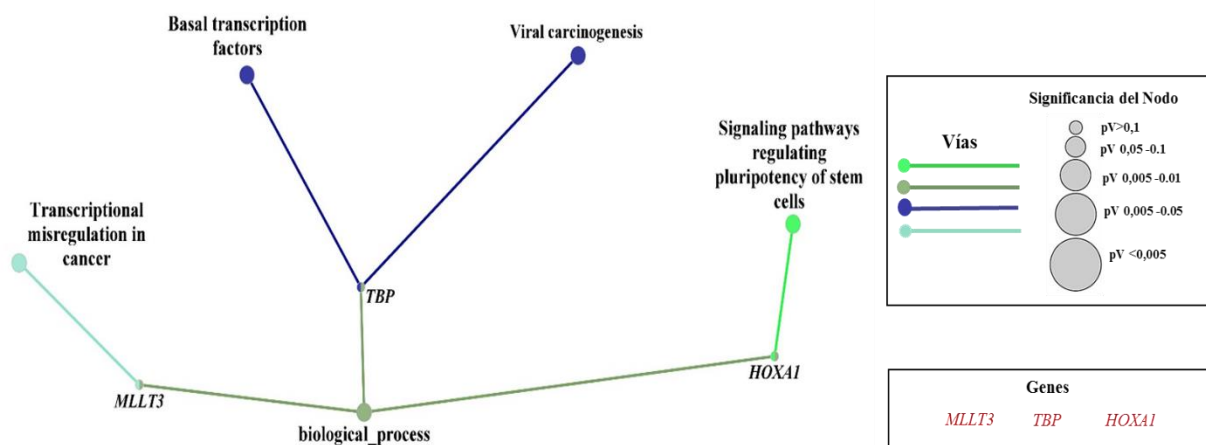
En el caso de la muestra SP06, el gen *HSPA6* estuvo asociado a la ruta “MAPK signaling” (Figura 27), además de la “Endocytosis”, entre otras.

Figura 27. Red agrupada funcionalmente de genes con presencia de inserciones perjudiciales para SP06 enriquecida con la base de datos GO y KEGG con términos como nodos vinculados en función de su nivel de puntuación $\kappa=(\geq 0,3)$. Fuente: Esta investigación.



En SP07 se observó nuevamente al gen *HOXA1* involucrado en “*Signaling pathways regulating pluripotency of stem cells*” como se describió con anterioridad para esta muestra. El gen *MLLT3* se presentó en la ruta “*Transcriptional misregulation in cancer*”. El gen *TBP* se encontró en vías como “*Basal Transcription factors*” y “*Viral carcinogenesis*” ruta asociada con Epstein-Barr (Figura 28).

Figura 28. Red agrupada funcionalmente de genes con presencia de inserciones perjudiciales para SP07 enriquecida con la base de datos GO y KEGG con términos como nodos vinculados en función de su nivel de puntuación $\kappa=(\geq 0,3)$. Fuente: Esta investigación.



- **Variantes compartidas entre las cuatro muestras**

Se encontraron tres SNPs deletereos en dos genes (*PABPC3*, *PRSS3*) y cinco inserciones perjudiciales en cinco genes (*OR2T35*, *LURAP1L*, *MICALCL*, *BCL6B*, *PPM1E*), dos de estas mutaciones no presentaron reportes, por lo tanto, se consideran como variantes nuevas. Los genes *PRSS3*, *BCL6B*, *PPM1E* se asociaron a neoplasias malignas del estómago según la base de datos DisGenet (Tabla 8).

En este estudio con la serie de filtros que se han descrito anteriormente se reportaron 62 genes que presentaron variantes raras y deletéreas, estas incluyeron 52 SNPs y 12 inserciones para las cuatro muestras analizadas (Tabla 8).

7.8.3. Comparación con otros estudios

En este estudio se identificaron 5 variantes entre patogénicas, probablemente patogénicas y con conflicto de patogenicidad en 5 de los 62 genes reportados en las vías de los 4 exomas analizados, con una prevalencia de (8,06%). Las variantes con esta clasificación fueron mutaciones puntuales presentes en los genes: *FLNC*, *ACSS2* y *MCCC2* de la muestra SP07. *ACADS*, y *BRCA1* de la muestra SP03 (Tabla 8). Además, se reportaron 26 genes que han sido asociados a neoplasia maligna del estómago según los datos de la plataforma DisGenet (Tabla 8).

Los resultados de esta investigación presentaron una prevalencia cercana a la reportada por Rodríguez (2019), que identificó seis variantes genómicas patogénicas y probablemente patogénicas en seis de los 59 genes estudiados por los autores (10,2%). Los valores fueron considerables si se tiene en cuenta que, Jurgens *et al.* (2015) realizaron un estudio de exoma completo en 232 individuos de ascendencia europea, africana y latina; identificando dos variantes en dos genes (0,86%).

Por otra parte, es importante señalar que otros estudios en exómica de CG han trabajado con un bajo número de muestras, en Corea del Sur Kang *et al.* (2013) realizaron la secuenciación completa de 4 exomas de pacientes con CG intestinal temprano (CGT) y compararon los resultados con CG avanzado (CGA), reportando mutaciones somáticas independientes en CGT (516 genes) y CG (3104 genes), entre los que estaban *DYRK3*,

GPR116, *MCM10*, etc. Además, Zhang *et al.* (2018) seleccionaron cinco muestras de pacientes chinos para detectar variantes de un solo nucleótido (SNVs) por medio de la secuenciación de exoma completo (WES), identificando 67 variantes entre somáticas y germinales en 35 genes, de los cuales 7 fueron mutaciones candidatas en tumores primarios. Estos casos reflejan la importancia del análisis exómico, aunque se trabaje con pocas muestras, ya que se abarca bastante información con la que se puede llegar a identificar nuevos genes con significado clínico, así como, la detección de nuevas variantes asociadas a la patología que aún no han sido descritas.

Es preciso tener en cuenta que las variantes de línea germinal se han estudiado sobre todo en pacientes o familias con CG difuso, cáncer de tipo hereditario (Donner *et al.*, 2015; Vogelaar *et al.*, 2017; Fewings *et al.*, 2018). Zhu *et al.* (2020) indicaron que las variantes de la línea germinal son una fuente importante de carcinogénesis, pero no están bien estudiadas en la patología debido a las dificultades técnicas previas para llamar con seguridad a estas variantes. Al igual que en esta investigación, los autores estudiaron pacientes no relacionados (n=38) para detectar por WES variantes raras de la línea germinal asociadas a la susceptibilidad tumoral, encontrando 5 variantes frecuentes en los genes *AR*, *POLE*, *ATM*, *BRCA2* y *ALK*.

Algunos estudios que han incluido WES han identificado genes con variantes germinales en pacientes con CG de tipo intestinal, no obstante, se encontró un bajo número de publicaciones que incluyan los genes aquí reportados, tal es el caso del gen *BRCA1*, Sahasrabudhe *et al.* (2017), analizaron mutaciones germinales en el gen *BRCA1* en muestras de pacientes hispanos con diagnóstico de CG no hereditario, encontrando que en 160 casos independientes de América Latina un portador presentó una mutación en el gen *BRCA1* (p.Gly559Valfs).

Otro gen estudiado por medio del análisis genómico es el gen *APC*, este registró una variante que representó el 1,5%, un dato similar se observó en la investigación de Setia *et al.* (2020), al identificar una mutación en el gen supresor de tumores *APC*, con un porcentaje de la alteración del 1,2%. El estudio se realizó en 81 pacientes con CG, de los cuales 23 fueron diagnosticados con CG de tipo intestinal.

Pese a que la mutación del gen *APC* en la presente investigación fue informada como benigna, otras publicaciones reportan en este gen variantes que serían perjudiciales, lo que lo haría un candidato para futuros estudios del cáncer de estómago en la región; Ito *et al.* (2018) en un paciente con CG intermedio una mutación R876 en el gen *APC* (COSM18852) que inserta un codón de terminación que la hace dañina. Los autores señalaron que, aunque no hay informes de que las mutaciones de *APC* estén asociadas con el pronóstico en pacientes con cáncer, es posible que, las mutaciones de *APC* hayan activado la oncogénesis, debido a que el paciente presentó recidiva y mal pronóstico de la patología.

Por otra parte, varias vías reportadas en este estudio ya han sido relacionadas con el CG, Hu & Chen (2012) utilizaron redes de interacción y análisis de grupos para investigar exhaustivamente las vías celulares involucradas en el CG, encontrando rutas como: Inmunodeficiencia primaria (inmunidad celular y humoral), adhesión focal, metabolismo de xenobióticos por citocromo P450 (iniciación y promoción de tumor), metabolismo de las drogas, ribosoma, vía de señalización Wnt. Los autores sugieren que cambios en estas vías pueden estar asociados con el CG.

Tabla. 8

Lista de genes con SNPs e INDELs deletéreos de las muestras analizadas. Se presenta el reporte de las variantes genéticas en las bases de datos dbSNP y ClinVar, así como, el reporte de los genes con presencia de estas variantes asociados a neoplasia maligna del estómago en la plataforma DisGenet.

Muestra SP03 (SNPs)										
Vías	Gen	Posición	Variante	dbSNP		ClinVar		Disgenet		
				Importancia clínica	Consecuencia funcional	Interpretación	Condición	Enfermedad	Asociación	
Metabolismo del butanoato	ACADS	121175696	T>C	rs57443665	Patogénica, probablemente -patogénica	Sin sentido	Patogénica, probablemente -patogénica	Deficiencia de butiril-CoA deshidrogenasa	No se reporta	No se reporta
	ACSM4	7477100	G>A	rs61584783	No se reporta	Sin sentido	No se reporta	No provisto	No se reporta	No se reporta
	ACSM6	96985057	A>C	rs192368492	No se reporta	Sin sentido	No se reporta	No provisto	No se reporta	No se reporta
Carcinogénesis química, Metabolismo de xenobióticos por citocromo P450, Metabolismo de las drogas	ADH1B	100240016	A>G	rs113075608	Benigna	Benigna	Benigna	Inmunodeficiencia grave	Neoplasia maligna del estómago	0,09
	CYP2C9	96708974	A>G	rs2256871	Benigna	Benigna	Benigna	No provisto	No se reporta	No se reporta
	CYP2E1	135352356	A>T	rs28969387	Benigna	Benigna	Benigna	No provisto	Neoplasia maligna del estómago	0,1
Vías en cáncer, cáncer gástrico, Vía de señalización Wnt	APC	112043492	C>A	rs113782655	Benigna	Sin sentido	Benigna	No provisto	Neoplasia maligna del estómago	0,7
Vías en cáncer, cáncer gástrico, MicroRNAs en cáncer	MET	116340223	T>C	rs77523018	Probablemente benigna, benigna	Sin sentido	Benigna	Síndrome predisponente al cáncer hereditario	Neoplasia maligna del estómago	0,7
Vías en cáncer, cáncer gástrico, Vía de señalización Wnt	TCF7L2	114920400	T>C	rs200100323	No se reporta	Sin sentido	No se reporta	No provisto	Neoplasia maligna del estómago	0,02
Vías en cáncer, MicroRNAs en cáncer, vía de señalización de la fosfatidilinositol3-kinasa (PI3K)	BRCA1	41249297	G>T	rs55688530	Interpretaciones conflictivas de patogenicidad	Sin sentido	Significado incierto	Cáncer de mama y ovario familiar 1	Neoplasia maligna del estómago	0,1
Vías en cáncer	PTGER3	71440025	G>A	rs5694	No se reporta	Sin sentido	No se reporta	No provisto	Neoplasia maligna del estómago	0,01
Vías en cáncer, Vía de señalización Wnt	CTBP2	126714948	C>T	rs73375140	No se reporta	Sin sentido	No se reporta	No provisto	Neoplasia maligna del estómago	0,01
Vías en cáncer, vía de señalización de la fosfatidilinositol3-kinasa (PI3K)	LAMB3	209806061	C>G	rs61734502	Significado incierto, benigna	Sin sentido	Benigna	Epidermólisis ampollosa de la unión	Neoplasia maligna del estómago	0,01
Endocitosis, Toxoplasmosis	LDLR	11221357	G>A	rs72658860	No se reporta	Sin sentido	Conflicto patogenicidad	Hipercolesterolemia	No se reporta	0,01

Las Moléculas de Adhesión Celular, Toxoplasmosis	<i>HLA-DQB1</i>	32632632	C>G	rs281862105	No se reporta	Sin sentido	No se reporta	No provisto	Neoplasia maligna del estómago	0,03
	<i>HLA-DRB1</i>	32552081	A>G	rs17887154	No se reporta	Sin sentido	No se reporta	No provisto	Neoplasia maligna del estómago	0,01
Las Moléculas de Adhesión Celular	NRCAM	107830092	T>C	rs111616050	No se reporta	Sin sentido	No se reporta	No provisto	No se reporta	No se reporta
Endocitosis	ARFGAP2	47189728	C>T	rs34662994	No se reporta	Sin sentido	No se reporta	No provisto	Neoplasia maligna del estómago	0,3
	SH3GL1	4363371	G>A	rs139264612	No se reporta	Sin sentido	No se reporta	No provisto	No se reporta	No se reporta
	CLTCL1	19197988	T>A	rs140709833	Probablemente benigna	Sin sentido	Probablemente benigna	No provisto	No se reporta	No se reporta
		19213744	C>T	rs145514163	Probablemente benigna	Sin sentido	Probablemente benigna	No provisto	No se reporta	No se reporta

Muestra SP03 (INDELs)

Vías	Gen	Posición	Variante	dbSNP		ClinVar		Disgenet		
				Importancia clínica	Consecuencia funcional	Interpretación	Condición	Enfermedad	Asociación	
Detección de ADN citosólico, Epstein Barr, Vía de señalización del receptor tipo NOD, Vía de señalización del receptor tipo RIG-I	MAVS	3844992	Ins- AAG	rs61675970	No se reporta	Inserción	No se reporta	No provisto	No se reporta	No se reporta
Reparación de errores de apareamiento del ADN, Vías en cáncer	MSH3	79950708	Ins- IGCAGC GGCCGC AGCGG CC	rs773983405	No se reporta	Inserción	No se reporta	No provisto	Neoplasia maligna del estómago	0,15
Termogénesis	KRT10	38975103	Ins- GCTGCC GCCGCC GTATCC GCCGCC GGAGCT	rs776920005	Probablemente benigna	Inserción	No se reporta	No provisto	No se reporta	No se reporta
Vía de señalización de estrógenos	ARID1B	157100023	Ins-GGA	rs797045287	Probablemente benigna	Inserción	Probablemente benigna	Trastorno del neurodesarrollo	No se reporta	No se reporta

Muestra SP05 (SNPs)

Vías	Gen	Posición	Variante	dbSNP		ClinVar		Disgenet	
				Importancia clínica	Consecuencia funcional	Interpretación	Condición	Enfermedad	Asociación

Inmunodeficiencia primaria, Vía de señalización FoxO	<i>RAG1</i>	36597492	C>A	rs4151033	Benigna	Sin sentido	Benigna	Inmunodeficiencia grave	No se reporta	No se reporta
Endocitosis	<i>EEA1</i> <i>KIF5A</i>	93251098 57969020	C>G C>T	rs150091509 rs748248329	No se reporta No se reporta	Sin sentido Sin sentido	No se reporta No se reporta	No provisto No provisto	No se reporta No se reporta	No se reporta No se reporta
Unión de extremos no homólogos	<i>XRCC5</i>	217001820	G>T	rs192418356	No se reporta	Sin sentido	No se reporta	No provisto	Neoplasia maligna del estómago	0,03
Biogénesis de ribosomas en eucariotas	<i>RIOK2</i>	96512920	C>T	rs144826286	No se reporta	Sin sentido	No se reporta	No provisto	No se reporta	No se reporta
Vía de señalización MAPK, Cáncer gástrico, Vía de señalización RAS, entre otras	<i>MET</i>	116340086	A>G	rs35225896	Probablemente benigna	Sin sentido	Probablemente benigna	Cáncer	Neoplasia maligna del estómago	0,7

Muestra SP05 (INDELs)

Vías	Gen	Posición	Variante	dbSNP		ClinVar		Disgenet		
				Importancia clínica	Consecuencia funcional	Interpretación	Condición	Enfermedad	Asociación	
Vía de señalización MAPK	<i>MAPK1</i>	22221708	Ins-CCG	rs770506771	No se reporta	Inserción	No se reporta	No provisto	Neoplasia maligna del estómago	0,4

Muestra SP06 (SNPs)

Vías	Gen	Posición	Variante	dbSNP		ClinVar		Disgenet		
				Importancia clínica	Consecuencia funcional	Interpretación	Condición	Enfermedad	Asociación	
Vía de señalización mTOR, Vía de señalización p53, entre otras	<i>TSC2</i>	2134233	C>T	rs200411484	Probablemente benigna, significado incierto	Sin sentido	Probablemente benigna, significado incierto	Esclerosis tuberosa 2	No se reporta	No se reporta
Vía de señalización de esfingolípidos, R, Vía de señalización del calcio, R, Vía de señalización de la fosfolipasa D, entre otras	<i>SPHK2</i>	49132605	G>T	rs201378907	No se reporta	Sin sentido	No se reporta	No provisto	No se reporta	No se reporta
Metabolismo de esfingolípidos	<i>SGPL1</i> <i>TBL3</i>	72604246 2024098	A>G G>C	rs188194665 rs181712046	No se reporta No se reporta	Sin sentido Sin sentido	No se reporta No se reporta	No provisto No provisto	No se reporta No se reporta	No se reporta No se reporta
Biogénesis de ribosomas en eucariotas	<i>MPHOSP</i> <i>H10</i> <i>BMS1</i>	71377014 43287148	G>A G>A	rs4852764 rs2272881	No se reporta Benigna	Sin sentido Sin sentido	No se reporta No se reporta	No provisto No provisto	No se reporta Neoplasia maligna del estómago	No se reporta 0,02
Ribosoma	<i>RPL7A</i> <i>MPRS1</i>	136217563 105708941	G>A G>A	rs140349266 rs116793449	No se reporta No se reporta	Sin sentido Sin sentido	No se reporta No se reporta	No provisto No provisto	No se reporta No se reporta	No se reporta No se reporta
Biosíntesis del aminoacil-tRNA	<i>FARSB</i>	223497962	T>C	rs750603882	No se reporta	Sin sentido	No se reporta	No provisto	No se reporta	No se reporta

Muestra SP06 (INDELs)

Vías	Gen	Posición	Variante	dbSNP		ClinVar		Disgenet		
				Importancia clínica	Consecuencia funcional	Interpretación	Condición	Enfermedad	Asociación	
Vía de señalización PPAR, Regulación de lipólisis en adipocitos	<i>AQP7</i>	33395131	Ins-CATCTT CCTC	rs745617679	No se reporta	Inserción	No se reporta	No provisto	No se reporta	No se reporta
Vías de regulación de la longevidad, Vía de señalización MAPK, Endocitosis	<i>HSPA6</i>	161494730	Ins-AGG	-	No se reporta	Inserción	No se reporta	No provisto	No se reporta	No se reporta
Muestra SP07 (SNPs)										
Vías	Gen	Posición	Variante	dbSNP		ClinVar		Disgenet		
				Importancia clínica	Consecuencia funcional	Interpretación	Condición	Enfermedad	Asociación	
Vía de señalización MAPK, Vía de señalización del receptor tipo RIG-I, Vía de señalización GHZ, entre otras	<i>MAP3K1</i>	56111633	T>C	rs572205361	Benigna	Sin sentido	Benigna	46, reversión de sexo XY, tipo 6	Neoplasia maligna del estómago	0,04
Senescencia celular, Vía de señalización Wnt, entre otras	<i>NFATC3</i>	68156591	T>A	rs139500568	No se reporta	Sin sentido	No se reporta	No provisto	No se reporta	No se reporta
Vía de señalización MAPK, Proteoglicanos en cáncer	<i>FLNC</i>	128498402	T>C	rs200502811	Conflicto de patogenicidad	Sin sentido	Conflicto de patogenicidad, probablemente benigna	Miocardiopatía	Neoplasia maligna del estómago	0,02
Adhesión focal, vía de señalización de la fosfatidilinositol3-kinasa (PI3K)	<i>ITGA1</i>	52243230	T>A	A	No se reporta	Sin sentido	No se reporta	No provisto	Neoplasia maligna del estómago	0,01
Glicolisis/gluconeogénesis, Metabolismo del piruvato, entre otras	<i>ACSS2</i>	33509608	T>C	rs59088485	Patogénica	Sin sentido	Patógeno	Paladar hendido no sindrómico	Neoplasia maligna del estómago	0,01
Carcinogénesis química, Metabolismo de xenobióticos por citocromo P450, Metabolismo de las drogas	<i>ALDH3A1</i>	19645458	T>C	rs145697414	No se reporta	Sin sentido	No se reporta	No provisto	No se reporta	No se reporta
Degradación de Valina, Leucina e Isoleucina	<i>MCCC2</i>	70936895	A>T	rs757052602	Patogénica, probablemente patogénica	Sin sentido	Patogénica, probablemente patogénica	Deficiencia de 3-metilcrotonil CoA carboxilasa 2	No se reporta	No se reporta
Endocitosis	<i>EPS15L1</i>	16514698	A>G	rs778253640	No se reporta	Sin sentido	No se reporta	No provisto	No se reporta	No se reporta

Endocitosis, Invasión bacteriana de células epiteliales	<i>CLTCL1</i>	19195771	G>A	rs190351859	Benigna	Sin sentido	Benigna	No provisto	No se reporta	No se reporta
Autofagia, Vía de señalización de la apelina	<i>PIK3R4</i>	130452500	C>T	rs141740732	Benigna	Sin sentido	Benigna	No provisto	Neoplasia maligna del estómago	0,01
Vías de señalización que regulan la pluripotencia de las células madre	<i>HOXA1</i>	27134966	T>G	rs17500494	Significado incierto, benigna, probablemente benigna	Sin sentido	Significado incierto, benigna, probablemente benigna	Trastorno del neurodesarrollo, Síndrome de Bosley-Salih-Alorainy	Neoplasia maligna del estómago	0,01
Degradación de Lisina	<i>KMT2D</i>	49431844	G>A	rs747415183	No se reporta	Sin sentido	No se reporta	No provisto	Neoplasia maligna del estómago	0,02
Vía de señalización del fofatidilinositol	<i>SYNJ2</i>	158495613	G>A	rs140940655	No se reporta	Sin sentido	No se reporta	No provisto	No se reporta	No se reporta

Muestra SP07 (INDELs)

Vías	Gen	Posición	Variante	DbSNP		ClinVar		Disgenet		
				Importancia clínica	Consecuencia funcional	Interpretación	Condición	Enfermedad	Asociación	
Carcinogénesis viral, Factores de transcripción viral	TBP	170871049	Ins-CAA	rs760309917	No se reporta	Inserción	No se reporta	No provisto	No se reporta	No se reporta
Desregulación transcripcional en el cáncer	MLLT3	20414343	Ins-CTG	rs749239229	No se reporta	Inserción	No se reporta	No provisto	No se reporta	No se reporta
Vías de señalización que regulan la pluripotencia de las células madre	HOXA1	27135316	Ins-TGG	rs758237482	Benigna		Benigna	Trastorno del neurodesarrollo	Neoplasia maligna del estómago	0,01

Variantes compartidas entre las cuatro muestras

Vías	Gen	Posición	Variante	DbSNP		ClinVar		Disgenet		
				Importancia clínica	Consecuencia funcional	Interpretación	Condición	Enfermedad	Asociación	
Secreción pancreática	PABPC3	25671168	C>T	rs78552667	No se reporta	Sin sentido	No se reporta	No provisto	No se reporta	No se reporta
	PABPC3	25671027	A>G	rs78826513	No se reporta	Sin sentido	No se reporta	No provisto	No se reporta	No se reporta
Vía de vigilancia del ARNm	PRSS3	33798042	G>C	rs141382822	No se reporta	Sin sentido	No se reporta	No provisto	Neoplasia maligna del estómago	0,01
Transducción olfativa	OR2T35	248802249	Ins-AGA	rs201931676	No se reporta	Inserción	No se reporta	No provisto		
Ninguna	LURAP1L	12775861	Ins-GGC	rs139315731	No se reporta	Inserción	No se reporta	No provisto	No se reporta	No se reporta
Ninguna	MICALCL	12316344	Ins-CTCCTC	rs767104379	No se reporta	Inserción	No se reporta	No provisto	No se reporta	No se reporta
Ninguna	BCL6B	6928019	Ins-CAGCA GCAG	Sin reporte	No se reporta	Inserción	No se reporta	No provisto	Neoplasia maligna del estómago	0,04
Ninguna	PPM1E	56833457	Ins-GAACCC GAACCC	Sin reporte	No se reporta	Inserción	No se reporta	No provisto	Neoplasia maligna del estómago	0,01

8 DISCUSIÓN

El CG es un problema de salud pública (GBD 2017 Stomach Cancer Collaborators, 2020). Nakagawa *et al* (2015) indica que el cáncer es una enfermedad que se desarrolla y evoluciona por diferentes factores, entre los que se encuentran la acumulación de diversos tipos de mutaciones. Los avances de HTS implementaron plataformas como WES para la obtención de grandes cantidades de datos mutacionales en genes o regiones que codifican proteínas para varios tipos de enfermedades incluido el CG. Aunado a este aspecto, con la bioinformática y los enfoques computacionales se ha realizado el análisis exhaustivo de estos datos masivos de distintos perfiles de cáncer que aún no han sido tan explorados (Nakagawa *et al.*, 2015).

No obstante, frente a este avance Zhang *et al.* (2018) enfatizaban la falta de información en la caracterización a nivel genómico del CG, en comparación con otros tipos de cáncer como el de pulmón o colon, además, los autores afirmaban que no existe una comprensión concluyente sobre las alteraciones genéticas inclusive en países desarrollados. Por lo cual, es necesario realizar investigaciones que amplíen el conocimiento del factor genético que tiene un papel preponderante en el desarrollo de la carcinogénesis gástrica, en este sentido en el presente estudio se caracterizan por primera vez las alteraciones genéticas de tipo germinal en pacientes diagnosticados con CG del departamento de Nariño.

En este estudio, las variables edad promedio de 64 años y sexo hombre (Tabla 2) incrementan el riesgo de CG igual a lo reportado en otros estudios (Díaz Serrano, 2017; Rawla & Barsouk, 2019). Respecto al diagnóstico histológico se encontró mayor frecuencia del CG de tipo intestinal, diferente a lo descrito por (Díaz Serrano, (2017) quienes señalan que en la clasificación de los adenocarcinomas gástricos cabe esperar aproximadamente un 54% adenocarcinomas intestinales, 32% de adenocarcinomas difusos y 15% mixtos o indeterminados, según la clasificación de Laurén.

La distribución de g-c por secuencia indicó que existe un sesgo en la calidad de los datos (Figura 6). Clark *et al.* (2011), señalaron que este sesgo en los estudios de secuenciación se debe en gran parte a los primeros pasos de la PCR durante la generación de

la biblioteca, donde la variación del contenido de g-c provocan una amplificación reducida y, por lo tanto, una cobertura de secuenciación más baja. El contenido de g-c afecta la eficacia de la hibridación con los oligonucleótidos y, por lo tanto, puede influir en el enriquecimiento de la diana. La técnica de Illumina muestra una fuerte caída en la profundidad de lectura a medida que el contenido de g-c aumenta de 60% a 80%. A medida que el contenido de g-c descende de 40% a 20%, el rendimiento de Illumina disminuye con una profundidad de lectura menor sobre esos objetivos. Cabe destacar que los autores señalan que este aspecto no necesariamente se correlaciona con la capacidad para identificar variantes. Las variaciones de un solo nucleótido representan las variaciones de secuencia más numerosas en el exoma humano y su identificación precisa y completa es un objetivo principal de la secuenciación del exoma, además el número total de INDELs detectados también se correlaciona con el recuento de lecturas.

Por otra parte, en esta investigación se utilizaron dos llamadores de variantes Haplotypecaller y Freebayes (Tabla 5 y 6) como lo recomiendan Bohnert *et al.*, (2017), quienes determinaron para datos de exoma de línea germinal sensibilidades de 95,3% y precisión de 100% para HaplotypeCaller y de 95,1% a 99,9% para FreeBayes. Gézi *et al.* (2015), señalan que no existe un método individual de llamada de alteraciones genéticas que presenten mejor sensibilidad y precisión, las discrepancias son significativas entre los llamadores de variantes de uso común incluso cuando se aplican al mismo conjunto de datos de secuencia, por lo cual es mejor optar por un llamado conjunto. Chapman, (2013), presentó un método de conjunto automatizado para combinar y evaluar variantes genómicas de múltiples llamadores con el fin de consolidar variantes para producir un conjunto final de llamadas con mejor sensibilidad y especificidad que los métodos individuales, el autor concluye que usar FreeBayes y HaplotypeCaller proporcionan una resolución de haplotipos regionales e individuales, respectivamente, lo que ayuda en sitios con INDELs heterocigotos o SNPs e INDELs poco espaciados. En este estudio se observó que con este método acompañado de los filtros duros de cada llamador se aumenta el número de variantes (Tabla 5 y 6) lo que se traduce en la identificación de un mayor número de mutaciones verdaderas.

Se estudiaron, además, variantes raras en el exoma (Figura 12 y 14). Utilizando como base la hipótesis enfermedad común-variante rara, se filtraron las variantes con una $MAF \leq 1\%$. El gran número de variantes, especialmente SNPs ocurren en MAF por debajo del 5% ($<1\%$ de la población), destacando la dificultad de la anotación de la variante (Lee *et al*, 2015). Las variantes raras presentes en el genoma pueden tener un efecto en diferentes enfermedades y compensar su MAF reducido. En este sentido, la mayoría de mutaciones del genoma humano son de frecuencia menor a 5%. Otro de los aspectos a considerar es el efecto funcional de estas variantes. El efecto de las alteraciones génicas poco frecuentes puede ser deletéreo, afectando la función de un gen y favoreciendo la selección poblacional. Asimismo, también podría tratarse de variantes recientes (Otero Rodríguez, 2018).

En cuanto al análisis funcional de redes fue utilizado el software Cytoscape, Doncheva *et al*, (2019) señala que es un programa mucho más adecuado para trabajar con redes grandes y ofrece una mayor flexibilidad en términos de análisis de red, importación y visualización de datos adicionales. Su uso es ampliamente extendido citándose en más de 51.400 artículos científicos. Por consiguiente, el análisis de ruta basado se ha utilizado con éxito para identificar los mecanismos subyacentes asociados con las enfermedades (Hu & Chen, 2012). En cáncer, el aporte más importante es que facilita la interpretación de resultados mediante la integración de redes que permiten conocer las alteraciones que modulan el funcionamiento de los genes asociados a la vía de la patología (Menon & Elengoe., 2020).

En esta investigación se reportaron 5 SNPs con cierto grado de patogenicidad en los genes: *ACADS* (g.121175696) y *BRCA1*(g.41249297) en la muestra SP03 y, *FLNC* (g.128498402), *MCCC2* (g.70936895) y *ACSS2* (g. 33509608) de la muestra S07 (Tabla 8).

Un gen ampliamente estudiado en diferentes tipos de cáncer es el gen *BRCA1*, Kim *et al*, (2019), indicaron que los portadores de mutaciones de la línea germinal de este gen exhiben un riesgo cuatro a siete veces mayor de desarrollar CG en comparación con la población general.

El gen *ACADS* estuvo presente en la vía con mayor significancia de la muestra SP03: “*Butanoate Metabolism*” junto a los genes *ACSM4* y *ACSM6* (Figura 22), esta vía sería crucial en el estudio de poblaciones con bajo riesgo de CG, ya que el butirato inhibe las histonas desacetilasas para aumentar la expresión de genes que ralentizan el ciclo celular e inducen la apoptosis. Comprender los mecanismos que contribuyen al cambio metabólico de la oxidación del butirato en las células cancerosas es importante para comprender los efectos beneficiosos de la molécula sobre el cáncer (Han *et al.*, 2018).

El gen *FLNC* estuvo implicado en vías como “*Proteoglycans in cancer*” (Figura 25), vía relacionada con la proliferación, adhesión, angiogénesis y metástasis, además, este gen junto (Kanehisa *et al.*, 2019). Slavin *et al.* (2017), reportaron variantes puntuales en este gen, para un paciente diagnosticado con adenocarcinoma gástrico.

En cuanto al gen *MCCC2*, este sería el primer reporte de una variante patogénica en el gen para un paciente con CG, Caltabiano *et al.*, (2019) afirmó que existen pocos estudios sobre *MCCC2* que lo asocien con la patología, no obstante, los autores identificaron la sobreexpresión de *MCCC2* ligada a tejidos tumorales en cáncer de próstata. Las mutaciones de este gen se asociarían con la 3-metilcrotonilglicinuria, un trastorno autosómico recesivo del catabolismo de la leucina.

Para el gen *ACSS2*, Hur *et al.* (2015) afirman que se desconoce el papel del gen en la patología, además no se encuentra su reporte en publicaciones que hayan implementado WES. Hur *et al.* (2015) investigamos la expresión de *ACSS2* en tejidos de CG, concluyendo que la pérdida de expresión de *ACSS2* es un factor de mal pronóstico en GC

Por otra parte, los resultados de análisis funcional obtenidos en este estudio evidencian que la muestra SP03 presentó genes importantes involucrados en la vía “*Pathways in cancer*” y el CG, como *APC*, un gen supresor de tumores que regula la oncoproteína β -catenina que participa en procesos como la proliferación, diferenciación, apoptosis, entre otros. El riesgo de CG asociado a mutaciones en este gen es de 0,5 a 2% (Leoz *et al.*, 2016). El gen *MET* es un protooncogén que codifica un receptor de tirosina quinasa c-MET y es necesario para el desarrollo embrionario y la reparación de tejidos. La activación oncogénica

suprime la apoptosis y promueve la supervivencia, proliferación, migración, diferenciación y angiogénesis de las células (Cota *et al.*, 2017). El gen *TCF7L2* codifica un factor de transcripción que actúa como un efector de la vía de señalización Wnt, por lo que desempeña un papel fundamental en el desarrollo celular y la regulación del crecimiento. Además, la señalización Wnt aberrante está involucrada en la patogénesis de numerosos tipos de cánceres humanos, y particularmente en el desarrollo y progresión del CG (Torres *et al.*, 2016).

En este estudio se mostró la interacción entre la vía “*Metabolism of xenobiotics by cytochrome P450*”, “*Drug metabolism*” y “*Chemical carcinogenesis*” (Figura 22). Los genes presentes en estas vías ya se han reportado como posibles genes asociados con el riesgo de CG. Ziwei, exhibieron la mayor cantidad de mutaciones entre todos los genes analizados.

Un estudio realizado por Hu & Chen, (2012) sugiere que el gen *CYP2C9* puede estar involucrado en el CG, este gen codifica una de las isoformas de epoxigenasa, involucrado en la epoxigenación del ácido linoleico y el producto principal de esta reacción es la leucotoxina que aumenta el estrés oxidativo y los eventos proinflamatorios posteriores, lo que conduce a la progresión de las células tumorales.

Al igual que en este estudio Linhao *et al.* (2018) reportaron un bajo número de mutaciones puntuales en el gen *CYP2E1*, detectando 2 SNPs no sinónimos en las regiones codificantes, no obstante, ninguna de estas alteraciones genéticas se informó en esta investigación. Los autores sugieren que las variantes genéticas de *CYP2E1* entre diferentes poblaciones puede explicarse por los diversos estilos de vida, cultura e historia.

Wang *et al.* (2020) sugirieron el papel oncogénico de *CYP2E1* en CG, el gen codifica para una proteína de interés debido a su capacidad para metabolizar y activar una amplia gama de pequeños sustratos toxicológicos, a compuestos y fármacos más hidrófobos, incluidos los carcinógenos potenciales.

Este estudio reportó una variante benigna en el gen *ADH1B*, en contraste, Rivera *et al.* (2016) reportaron la existencia de mutaciones a nivel de polimorfismos de nucleótido único que hacen que la cantidad de la producción o la oxidación del acetaldehído sea variada

entre los individuos. El acetaldehído se considera tóxico, carcinogénico y mutagénico por sus efectos directos sobre inhibir la reparación y metilación del ADN, además de tener un efecto directo sobre esta molécula.

En la red de SP03 de genes con inserciones probablemente perjudiciales se observó que el gen *MSH3* estuvo presente en vías importantes como “*Pathways in cancer*” y “*Mismatch repair*” (Figura 26), vía relacionada con la reparación de errores de apareamiento y corrección de los desajustes de ADN generados durante la replicación, evitando que las mutaciones se vuelvan permanentes en las células en división (Kanehisa *et al*, 2019). *MAVS* estuvo presente en rutas como “*Cytosolic DNA-sensing*”, familias específicas de receptores de reconocimiento responsables de detectar ADN extraño de microorganismos invasores y generar respuestas inmunes innatas (Kanehisa *et al.*, 2019). En la Figura 26 *AIRID1* hace parte de “*Hepatocellular carcinoma*” así como otros tipos de cáncer, como lo señalo Aso *et al.* (2015), en CG la baja expresión de este gen se relaciona con la progresión de tumores.

Para la muestra SP05, el gen *XRCC5* estuvo presente en la vía “*Non-homologous end-joining*” (Figura 23), que repara roturas de doble hebra en el ADN (Davis & Chen, 2013). *EEA1* y *KIF5A* se presentaron en la vía “*Endocytosis*”, las células carcinógenas presentan una serie de modificaciones relativamente sutiles de los procesos fisiológicos normales que, cuando se combinan, pueden crear fenotipos y comportamientos marcadamente alterados. Durante mucho tiempo se sospechó que la endocitosis es uno de esos procesos fisiológicos que se modifica en el cáncer (Mellman & Yarden, 2013), mientras que “*Primary immunodeficiency*”, se relaciona con trastornos que afectan la inmunidad celular y humoral o los mecanismos de defensa no específicos del hospedero que provocan una mayor susceptibilidad a infecciones, enfermedades auto inmunitarias y malignidad (Haas, 2018); en esta vía se destaca el gen *RAG1* que codifica una proteína involucrada en la inmunidad adaptativa al mediar en la activación de la recombinación de inmunoglobulina en los linfocitos en desarrollo, la actividad anormal promueve malignidades en los linfocitos (Li *et al.*, 2019).

El análisis funcional para los genes con presencia de inserciones indicó que *MAPK1* participa en más de 80 vías y en distintas enfermedades como el CG (Kanehisa *et al.*, 2019).

Se ha informado que *MAPK1* es un oncogén importante en la progresión del CG para promover la proliferación, migración e invasión celular, varios estudios en el gen (Diao *et al.*, 2018; Hu *et al.*, 2018; Wang *et al.*, 2018) han encontrado diferentes mutaciones sin sentido, contrario a lo reportado en esta investigación donde se encontró una inserción deletérea con impacto moderado en el gen (Tabla 8). Luo & Liang, (2020), encontraron una alta expresión de *MAPK1* en la patología, no obstante, cuando la expresión decrece se inhibe la viabilidad, la migración y la invasión de las células del CG, pero se promueve la apoptosis, lo que indica el papel carcinogénico de *MAPK1*.

En el caso de la muestra SP06, los genes *BMS1*, *MPHOSPH10*, *TBL3* se relacionaron con la vía representativa “*Ribosome biogenesis in eukaryotes*” (Figura 24), involucrado con la formación del nucléolo que juega un papel en la regulación de una serie de mecanismos que al alterarse pueden contribuir a la tumorigénesis (Penzo *et al.*, 2019). También se observó, la vía “*Sphingolipid signaling*”, los esfingolípidos tienen funciones opuestas en la regulación de la muerte y supervivencia de las células cancerosas (Ogretmen , 2018). Otra vía a destacar es “*Phospholipase D signaling*” que genera ácido fosfatídico, para regular e influir en procesos celulares como tumorigénesis, migración, proliferación, angiogénesis e invasión (Cho & Han, 2017). Es importante señalar el gen *TSC2* está involucrado en las vías metabólicas del cáncer, además, está alterado en el 2,82% de los pacientes con adenocarcinoma gástrico, en esta investigación solamente se encontró una variante en el gen (Tabla 8) (AACR Project GENIE Consortium, 2017).

En esta misma muestra, el análisis de genes con inserciones perjudiciales evidencia asociación del gen *HSPA6* con la vía “*MAPK signaling*” y “*Endocytosis*” (Figura 27), Kuballa *et al.* (2015). Señala la importancia del estudio de este gen codificador de proteínas implicadas en procesos celulares, incluida la protección del proteoma contra el estrés, el plegamiento y transporte de polipéptidos recién sintetizados.

Los genes con mutaciones puntuales de la muestra SP07 estuvieron presentes en vías de relevancia como “*MAPK signaling pathway*” (Figura 25), vías que sintetizan proteínas como la quinasa activada por mitógenos (MAPK) que participa en la proliferación, diferenciación y migración celular. El gen *ITGA1* se presentó en “*Focal adhesion*”, vía

asociada a la regulación de la expresión genética, proliferación, diferenciación y la supervivencia celular. Por otra parte, *ITGA1* está presente en otras vías de regulación de función celular (Kanehisa *et al.*, 2019). Destacar las vías en las que estuvo el gen *NFATC3* como “*Wnt signaling pathway*” implicada en eventos patológicos como el mantenimiento y la proliferación del subconjunto de células madre cancerosas presentes en la masa tumoral (Orozco *et al.*, 2019), y “*Cellular senescence*” importante en la supresión de células cancerosas (Kanehisa *et al.*, 2019). Se observó que se presentan al igual que en otras muestras como la SP03 vías como “*Metabolism of xenobiotics by cytochrome P450*”, “*Chemical carcinogenesis*” y “*Endocytosis*” (Figura 25).

El gen *HOXA1* presentó mutaciones puntuales e inserciones probablemente perjudiciales para esta muestra y estuvo asociado a “*Signaling pathways regulating pluripotency of stem cells*” (Figura 25 y 28), un gen implicado en la progresión tumoral. Yuan *et al.* (2016) indicaron que mutaciones en el gen podrían alterar la expresión del ARNm y proteínas de *HOXA1*, siendo esta expresión mayor en los tejidos GC que en las mucosas normales adyacentes.

En la figura 28 se observó la vía “*Transcriptional misregulation in cancer*”, en la que ciertos genes pueden desregularizarse por mutaciones lo que conduciría a que las células adquieran una variedad de propiedades alteradas que contribuyen al proceso tumorgénico, el gen presente en esta vía *MLLT3* codifica una proteína involucrada en el crecimiento y mantenimiento celular (Ribeiro *et al.*, 2015). Importante señalar que el gen *TBP* podría ser un componente crítico en la señalización desregulada que ocurre aguas abajo de las lesiones genéticas que causan tumores (Johnson *et al.*, 2003). Además, el gen se encontró en vías como “*Viral carcinogenesis*” asociada con Epstein-Barr, virus que promueve la proliferación celular mediante la modulación de las vías de señalización y escapan del sistema de defensa celular como al bloqueo de la apoptosis (Kanehisa *et al.*, 2019).

9 CONCLUSIONES

Se identificó un mayor número de variantes de nucleótido único (1.597.012 SNPs), las inserciones y deleciones (203.133) estuvieron en una menor proporción en las muestras estudiadas.

Se reportaron 52 mutaciones puntuales y 12 inserciones raras y deletéreas presentes en 62 genes que se obtuvieron al interactuar en las vías presentadas.

Se identificaron 5 mutaciones puntuales entre patogénicas, probablemente patogénicas y con conflicto de patogenicidad, presentes en los genes: *ACADS* (g.121175696) y *BRCA1*(g.41249297) en la muestra SP03 y, *FLNC* (g.128498402), *MCCC2* (g.70936895) y *ACSS2* (g. 33509608) de la muestra S07.

De los genes de estudio 27 han sido asociados a neoplasia maligna del estómago según los datos de la plataforma DisGenet, como *APC*, *BRAC1*, *CTBP*, *LAMB3*, *MET*, *PTGER3*, *TCFLT2*, *CYP2E1*, *MAPK1*, *MSH3*.

Un gran número de genes presentaron variantes perjudiciales, estos genes candidatos hicieron parte de las posibles vías de acción más importantes en el cáncer gástrico para las muestras.

En la muestra SP03 se encontró el mayor número de SNPs e inserciones perjudiciales con respecto a las otras muestras. Los genes *APC*, *MET* y *TCF7L2* se asociaron a las vías en cáncer. La vía del Butirato representada por los genes *ACSM4*, *ACSM6* y *ACADS*, es importante por los efectos beneficiosos de la molécula sobre el cáncer. Los genes *CYP2C9*, *CYP2E1* y *ADH1B* hicieron parte de las vías metabolismo de xenobióticos por citocromo p450, metabolismo de fármacos y carcinogénesis química, vías que tienen un papel central en la carcinogénesis, relacionadas con la iniciación y promoción de tumores.

El gen *MSH3* estuvo presente en las vías en cáncer y de reparación de los desajustes de ADN durante la replicación. El gen *MAVS* se presentó en vías que detectan ADN extraño de microorganismos invasores. Las mutaciones en el gen *AIRID1* podrían ser importantes por su relación con la progresión de tumores.

Los genes *XRCC5*, *EEA1* y *KIF5A*, *RAG1* se relacionaron con vías de reparación del ADN, procesos fisiológicos e inmunidad celular respectivamente. El análisis funcional para los genes con presencia de inserciones indicó que *MAPK1* se relaciona con distintas enfermedades como el CG.

Los SNPs presentes en los genes *BMS1*, *MPHOSPH10*, *TBL3* y *TSC2* pueden contribuir a la tumorigénesis, ya que estos genes hacen partes de vías implicadas en procesos de regulación, supervivencia y proliferación celular.

El gen *HSPA6* se asoció a la vía de señalización MAPK y Endocitosis, este gen con inserciones perjudiciales, se implica en procesos celulares protectores.

Varios genes con mutaciones puntuales estuvieron implicados en vías relacionadas con procesos de proliferación, diferenciación y la supervivencia celular, como la vía de señalización MAPK, proteoglicanos en cáncer, adhesión focal, señalización wnt y senescencia celular. El gen *HOXA1* presentó mutaciones puntuales e inserciones probablemente perjudiciales para esta muestra y estuvo asociado a vías de señalización que regulan la pluripotencia de las células madre, gen implicado en la progresión tumoral.

Los genes *MLLT3* y *TBP* presentaron inserciones perjudiciales, estos genes son candidatos de estudio, ya que están involucrados en el crecimiento, proliferación y mantenimiento celular.

El elevado número de nuevas variantes genéticas encontradas podría ayudar a explicar parte del componente genético del cáncer gástrico. Los resultados aportan nuevas variantes de estudio, que podrían estar relacionadas con la enfermedad.

10 PERSPECTIVAS

Los resultados amplían el conocimiento de la genética del cáncer gástrico en el departamento de Nariño, además, con el aumento del tamaño de la muestra se podría aportar nuevas variantes al estudio que tendrían una posible relación con la patología.

La confirmación de estos hallazgos con la validación y la replicación de los resultados encontrados en muestras de otras poblaciones, así como el diseño de experimentos funcionales.

Realizar un diseño de casos y controles para validar las variantes patogénicas reportadas en este estudio. Esto permitirá evidenciar las variantes genéticas que realmente están involucradas en el desarrollo del cáncer gástrico.

La identificación de estos factores genéticos podría permitir una mejor estratificación del riesgo, ofrecer a los pacientes la posibilidad de una detección temprana, para reducir el riesgo y brindar nuevas opciones de tratamiento.

Hallazgos genéticos en el manejo clínico de los pacientes permitirá en un futuro establecer posibles tratamientos personalizados, incluyendo el asesoramiento genético.

11 APORTES

La identificación por primera vez de las variantes genética de tipo germinal del exoma de cuatro pacientes diagnosticados con cáncer gástrico del departamento de Nariño. Estudio base para el descubrimiento en la región de nuevos biomarcadores que podrían estar relacionados con el desarrollo de la enfermedad.

Resultados de investigación asociados al Trabajo de Grado se presentaron en:

El V Congreso Colombiano de Bioinformática y Biología Computacional (CCBCOL) Ibagué, Construyendo Innovación y Desarrollo. Este evento fue realizado los días 6 al 8 de noviembre de 2019 en la ciudad de Ibagué-Tolima.

El producto fue en calidad de ponente, y se presentó con el nombre de:

“Exómica de cáncer gástrico en pacientes del departamento de Nariño, en búsqueda de marcadores terapéuticos”.

Los resultados de la investigación serán sometidos a publicación en revista del índice Scimago - Scopus.

12 REFERENCIAS BIBLIOGRÁFICAS

- AACR Project GENIE Consortium. (2017). AACR Project GENIE: Powering Precision Medicine through an International Consortium. *Cancer discovery*, 7(8), 818-31. doi:10.1158/2159-8290.CD-17-0151
- Amigo, J. (septiembre de 2012). La bioinformática al servicio de la genómica. Santiago de Compostela, España: Universidad de Santiago de Compostela.
- Arango, M., Jaramillo, C., Montealegre, M., Bohórquez, M., & Delgado, M. (2010). Genotipificación de los polimorfismos -511, -31 y +3954 del gen de la interleucina-1 b humana en una población colombiana con cuadro de dispepsia. *Biomédica*, 30(2), 199-206.
- Aso, T., Uozaki, H., Morita, S., Kumagai, A., & Watanabe, M. (2015). Loss of ARID1A, ARID1B, and ARID2 Expression During Progression of Gastric Cancer. *Anticancer research*, 35(12), 6819-6827.
- Avanesyan, A., Sokolenko, P., Ivantsov, A., Kleshchev, M., Maydin, M..... (2020). Germline Mutation Carriers: Results of Endoscopic Screening and Molecular Analysis of Tumor Tissues. *Pathobiology*. doi: 10.1159/000511323
- Badía Martínez, M., & García, M. (2011). Polimorfismos en genes de baja penetrancia implicados en respuesta inmune e infección por *Helicobacter pylori*. Influencia en el desarrollo del cáncer gástrico. *Medicina Clínica*, 137(12), 555-560. doi:10.1016/j.medcli.2010.09.030
- Bao, R., Huang, L., Andrade, J., Tan, W., Kibbe, W., Jiang, H., & Feng, G. (2014). Review of Current Methods, Applications, and Data Management for the Bioinformatics Analysis of Whole Exome Sequencing. *Cancer informatics*, 13(2), 67-82. doi:10.4137/CIN.S13779

- Bartha, Á., & Györfy, B. (2019). Comprehensive Outline of Whole Exome Sequencing Data Analysis Tools Available in Clinical Oncology. *Cancers*, 11(11), 1725. <https://doi.org/10.3390/cancers11111725>
- Bedoya, Álvaro., Yépes, Y., Calvache, D., Cifuentes, Y., Lucero, N..... (2019). Proyecto Urkunina 5000- Investigación de la prevalencia de lesiones precursoras y del efecto de la erradicación de *Helicobacter pylori* como prevención primaria del cáncer gástrico en el departamento de Nariño. *Revista Colombiana de Cirugía*, 33(4), 345-52.
doi.org/10.30944/20117582.81
- Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., Fridman, W. H., Pagès, F., Trajanoski, Z., & Galon, J. (2009). ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics (Oxford, England)*, 25(8), 1091–1093.
doi.org/10.1093/bioinformatics/btp101
- Bohnert, R., Vivas, S., & Jansen, G. (2017). Comprehensive benchmarking of SNV callers for highly admixed tumor data. *PLoS One*, 12(10), e0186175.
doi:10.1371/journal.pone.0186175
- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R., A Torre, L., & Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, 68(6), 394-424.
doi:10.3322/caac.21492.
- Bushnell, B. (2014). *BBMap: A Fast, Accurate, Splice-Aware Aligner*. United States.
- Cai, H., Jing, C., Chang, X., Ding, D., Han, H. (2019). Mutational landscape of gastric cancer and clinical application of genomic profiling based on target next-generation sequencing. *Journal of Translational Medicine*, 17: 189.
doi: 10.1186/s12967-019-1941-0

- Calabria, I., Pedrola, L., Berlanga, P., Aparisi, M., Sánchez, D., Cañetea, A., . . . Castel, V. (2016). El nuevo reto en oncología: la secuenciación NGS y su aplicación a la medicina de precisión. *Anales de pediatría*, 85(5), 273.e1-273.e7. doi:10.1016/j.anpedi.2016.05.006
- Cardona, D., Castaño, E., & Marín, J. (2007). Cáncer gástrico, tabaquismo, consumo de licor, estrato socioeconómico y polimorfismo en el codon 72 del gen p53 en una población de Manizales. *Biosalud*, 6, 33-44.
- Casey, R.(2019). What are paired-end reads? *The sequencing center*. Disponible en: <https://thesequencingcenter.com/knowledge-base/what-are-paired-end-reads/>
- Castaño, E., Santacoloma, M., Arango, L., & Camargo, M. (2010). Cáncer gástrico y genes detoxificadores en una población colombiana. *Revista Colombiana de Gastroenterología*, 25(3), 252-260.
- Caltabiano, R., Castrogiovanni, P., Barbagallo, I., Ravalli, S., Szychlińska, M..... (2019). Identification of novel markers of prostate cancer progression, potentially modulated by vitamin D. *Applied Sciences*. 9(22), 4923. doi.org/10.3390/app9224923
- Cebrián, A., de la Concha, G., & Fernández, I. (2016). Cáncer gástrico. *Medicine*, 12(3), 118-27. doi:10.1016/j.med.2016.02.002
- Chang, V., Federman, N., Martinez-Agosto, J., Tatishchev, S., & Nelson, S. (2013). Whole Exome Sequencing of Pediatric Gastric Adenocarcinoma Reveals an Atypical Presentation of Li-Fraumeni Syndrome. *Pediatr Blood Cancer*, 60(4), 570-574.
- Chapman, B. (2013). An automated ensemble method for combining and evaluating genomic variants from multiple callers. *Blue Collar Bioinformatics: Community built tools for biological data analysis*, 4. Obtenido de <http://bcb.io/2013/02/06/an-automated-ensemble-method-for-combining-and-evaluating-genomic-variants-from-multiple-callers/>

- Checa, M. (2007). Polimorfismos genéticos: Importancia y aplicaciones. *Rev Inst Nal Enf Resp Mex*, 20(3), 213-221.
- Cheng, X., Lin, J., & Tu, S. (2016). Etiology and Prevention of Gastric Cancer. *Gastrointestinal tumors*, 3(1), 25-36. doi:10.1159/000443995
- Cho, J., & Han, J.-S. (2017). Phospholipase D and Its Essential Role in Cancer. *Molecules and cells*, 11, 805-813. doi:10.14348/molcells.2017.0241
- Clark, M., Chen, R., Lam, H., Karczewski, K., Chen, R., Euskirchen, G., . . . Snyder, M. (2011). Performance comparison of exome DNA sequencing technologies. *Nature biotechnology*, 29(10), 908-914. doi:10.1038/nbt.1975
- Correa, P. (2011). Cáncer gástrico: una enfermedad infecciosa. *Rev Colomb Cir.*, 26, 111-117.
- Cota, B., Lima, K., Murad, A., Xavier, M., Álvares, M., & Coelho, L. (2017). Expression of the c-MET, HGF and VEGF biomarkers in intestinal and diffuse gastric cancer in the Brazilian population: a pilot study for the standardization of the quantitative PCR technique. *Applied Cancer Research*, 37(18). doi:10.1186/s41241-017-0029-5
- Cumbal, N. (2010). Estudio de los polimorfismos en las regiones -31(T>C), -511(C>T) Y +3954(C>T) del gen IL-1 β de interleuquina-1 β mediante PCR-RFLP en población ecuatoriana con cáncer gástrico y presencia de *Helicobacter pylori*. Sangolqui, Ecuador: Escuela Politécnica del Ejército.
- Cristescu, M. (2015). Genetic reconstructions of invasion history. *Molecular Ecology*. 24(9), 2212-2225. doi.org/10.1111/mec.13117
- Danecek, P., Auton, A., Abecasis, G., Albers, C., Banks, E., DePristo, M., . . . 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics (Oxford, England)*, 27(15), 2156-2158. doi:10.1093/bioinformatics/btr330

- Davis, A., & Chen, D. (2013). DNA double strand break repair via non-homologous end-joining. *Translational cancer research*, 2(3), 130-143. doi:10.3978/j.issn.2218-676X.2013.04.02
- Delgado, N., Casas, P., Torres, J., Bustos, A., Santiago, E., Marín, M., & Sánchez, J. (2017). Factores de riesgo asociados a adenocarcinoma gástrico de patrones histológicos de tipo intestinal y difuso en población adulta del occidente de México. 153, 173-178.
- DePristo, M., Banks, E., Poplin, R., Garimella, K., Maguire, J., Hartl, C., . . . Daly, M. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature genetics*, 43(5), 491-498. doi:10.1038/ng.806
- Diao, L., Wang, S., & Sun, Z. (2018). Long noncoding RNA GAPLINC promotes gastric cancer cell proliferation by acting as a molecular sponge of miR-378 to modulate MAPK1 expression. *OncoTargets and therapy*, 11, 2797-2804. doi:10.2147/OTT.S165147
- Díaz Serrano, A. (2017). *Caracterización molecular del cáncer gástrico HER2 positivo y su aplicabilidad práctica para el tratamiento personalizado del cáncer gástrico*. Universidad Complutense de Madrid, Facultad de Medicina. Madrid: Memoria para optar al grado de doctor.
- Doncheva, N., Morris, J., Gorodkin, J., & Jensen, L. (2019). Cytoscape StringApp: análisis de red y visualización de datos proteómicos. *J. Proteome Res*, 18(2), 623-632. doi:10.1021/acs.jproteome.8b00702
- Donner, I., Kiviluoto, T., Ristimäki, A., Aaltonen, L. A., & Vahteristo, P. (2015). Exome sequencing reveals three novel candidate predisposition genes for diffuse gastric cancer. *Familial cancer*, 14(2), 241–246. <https://doi.org/10.1007/s10689-015-9778-z>
- Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G., . . . Turner, S. (2009). Real-time DNA sequencing from single polymerase molecules. *Science*, 323(5910), 133-138. doi:10.1126/science.1162986

- Esteban, C. (2016). Identificación y caracterización de nuevos genes de predisposición al cáncer colorrectal familiar. *Tesis doctoral*. Barcelona, España: Universidad de Barcelona.
- Fewings, E., Larionov, A., Redman, J., Goldgraben, M. A., Scarth, J., Richardson, S., Brewer, C., Davidson, R., Ellis, I., Evans, D. G., Halliday, D., Izatt, L., Marks, P., McConnell, V., Verbist, L., Mayes, R., Clark, G. R., Hadfield, J., Chin, S. F., Teixeira, M. R., ... Tischkowitz, M. (2018). Germline pathogenic variants in PALB2 and other cancer-predisposing genes in families with hereditary diffuse gastric cancer without CDH1 mutation: a whole-exome sequencing study. *The lancet. Gastroenterology & hepatology*, 3(7), 489–498. [https://doi.org/10.1016/S2468-1253\(18\)30079-7](https://doi.org/10.1016/S2468-1253(18)30079-7)
- Galvis, F., Rivero, E., & Cardenas, D. (2017). Polimorfismos en los genes MTHFR e IL-10 y factores de riesgo en pacientes con cáncer gástrico del municipio de San José de Cúcuta, Colombia. *Correo Científico Médico de Holguín*, 21(2), 487-500.
- García-Prieto, I. (2016). *Analisis integrado de datos de secuenciacion masiva y su aplicacion en el diagnostico de enfermedades raras*. Master en bioinformática y biología computacional, Instituto de Salud Carlos III, Málaga, España.
- Garrison, E., & Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. *arXiv*, 1207.3907, 1-9.
- Garziera, M., Canzonieri, V., Cannizzaro, R., Geremia, S., Caggiari, L., De Zorzi, M., . . . De Re, V. (2013). Identification and characterization of CDH1 germline variants in sporadic gastric cancer patients and in individuals at risk of gastric cancer. *PloS one*, 8(10), e77035. doi:10.1371/journal.pone.0077035
- GBD 2017 Stomach Cancer Collaborators. (2020). The global, regional, and national burden of stomach cancer in 195 countries, 1990–2017: a systematic analysis for the Global Burden of Disease study 2017. *Lancet Gastroenterol Hepatol*, 5, 42-54. doi:10.1016/S2468-1253(19)30328-0

- Gézi, A., Bolgár, B., Marx, P., Sarkozy, P., Szalai, C., & Antal, P. (2015). VariantMetaCaller: automated fusion of variant calling pipelines for quantitative, precision-based filtering. *BMC Genomics*, 875, 1-15. doi:10.1186/s12864-015-2050-y
- Gobernado, I. (Junio de 2014). Secuenciación de exoma completo en trastorno bipolar autosómico dominante: afectación del gen *Period3-Ritmo Circadiano*. *Tesis doctoral*. Alcalá de Henares, España: Universidad de Alcalá.
- Guan, Y. F., Li, G. R., Wang, R. J., Yi, Y. T., Yang, L., Jiang, D., Zhang, X. P., & Peng, Y. (2012). Application of next-generation sequencing in clinical oncology to advance personalized treatment of cancer. *Chinese journal of cancer*, 31(10), 463–470. doi.org/10.5732/cjc.012.10216
- Haas, O. (2018). Primary Immunodeficiency and Cancer Predisposition Revisited: Embedding Two Closely Related Concepts Into an Integrative Conceptual Framework. *Frontiers in immunology*, 9, 3136. doi:10.3389/fimmu.2018.03136
- Han, A., Bennett, N., Ahmed, B., Whelan, J., & Donohoe, D. (2018). Butyrate decreases its own oxidation in colorectal cancer cells through inhibition of histone deacetylases. *Oncotarget*, 9(43), 27280-27292. doi:10.18632/oncotarget.25546
- Heather, J., & Chain, B. (2016). The sequence of sequencers: The history of sequencing DNA. *Genomics*, 1-8.
- Hu, K., & Chen, F. (2012). Identification of significant pathways in gastric cancer based on protein-protein interaction networks and cluster analysis. *Genetics and molecular biology*, 35(3), 701-708. doi:10.1590/S1415-47572012005000045
- Hu, L., Wu, H., Wan, X., Liu, L., He, Y., Zhu, L., . . . Zhu, Z. (2018). MicroRNA-585 suppresses tumor proliferation and migration in gastric cancer by directly targeting MAPK1. *Biochem Biophys Res Commun*, 499(1), 52-58. doi:10.1016/j.bbrc.2018.03.116.

- Hur, H., Kim, Y. B., Ham, I. H., & Lee, D. (2015). Loss of ACSS2 expression predicts poor prognosis in patients with gastric cancer. *Journal of surgical oncology*, 112(6), 585–591. <https://doi.org/10.1002/jso.24043>
- Illumina. (2017). An introduction to Next-Generation Sequencing Technology. Pub. No. 770-2012-008-B.
- Ito T, Matoba R, Maekawa H, Sakurada M, Kushida T, Orita H, Wada R, Sato K. Detection of gene mutations in gastric cancer tissues using a commercial sequencing panel. *Mol Clin Oncol*. 2019 Nov;11(5):455-460. doi: 10.3892/mco.2019.1926. Epub 2019 Sep 24. PMID: 31620276; PMCID: PMC6787944.
- Jalali, M., & Gamielien, J. (2017). A practical guide to filtering and prioritizing genetic variants. *Biotechniques*, 62(1), 18-30. doi:10.2144/000114492
- Jeon, S., Park, J.M., Kim, J., Kim, J., Kim, Y., & Kim, S. (2019). Comparison of the MGISEQ-2000 and Illumina HiSeq 4000 sequencing platforms for RNA sequencing. *Genomics & Informatics*, 17. doi:10.5808/GI.2019.17.3.e32
- Johnson, S., Dubeau, L., Kawalek, M., Dervan, A., Schönth, A., Chi, D., & Johnson, D. (2003). Increased Expression of TATA-Binding Protein, the Central Transcription Factor, Can Contribute to Oncogenesis. *Molecular and cellular biology*, 23(9), 3043–3051. doi:10.1128/mcb.23.9.3043-3051.2003
- Jurgens, J., Ling, H., Hetrick K, Pugh E, Schiettecatte F, Doheny K, et al. Assessment of incidental findings in 232 whole-exome sequences from the Baylor-Hopkins Center for Mendelian Genomics. *Genetics in Medicine*, 17(10), 782-788.
- Kamps, R., Brandão, R., van den Bosch, B., D. C. Paulussen, A., Xanthoulea, S., J. Blok, M., & Romano, A. (2017). Next-Generation Sequencing in Oncology: Genetic Diagnosis, Risk Prediction and Cancer Classification. *International journal of molecular sciences*, 18(2), 1-57. doi:10.3390/ijms18020308

- Kanehisa, M., Sato, Y., Furumichi, M., Morishima, K., & Tanabe, M. (2019). New approach for understanding genome variations in KEGG. *Nucleic acids research*, 47(D1), 590-595.
doi:10.1093/nar/gky962
- Kang, G., Hwang, W. C., Do, I. G., Wang, K., Kang, S. Y., Lee, J., Park, S. H., Park, J. O., Kang, W. K., Jang, J., Choi, M. G., Lee, J. H., Sohn, T. S., Bae, J. M., Kim, S., Kim, M. J., Kim, S., Park, C. K., & Kim, K. M. (2013). Exome sequencing identifies early gastric carcinoma as an early stage of advanced gastric cancer. *PloS one*, 8(12), e82770. <https://doi.org/10.1371/journal.pone.0082770>
- Kang, C., Song, J.-J., Lee, J., & Kim, M. (2014). Epigenetics: An emerging player in gastric cancer. *World journal of gastroenterology*, 20(21), 6433-6447.
doi:10.3748/wjg.v20.i21.6433
- Katona, B; Rustgi, A. (2017). Gastric Cancer Genomics: Advances and Future Directions. *Cellular and Molecular Gastroenterology and Hepatology*, 3(2), 211-217. doi: 10.1016/j.jcmgh.2017.01.003
- Kim, H., Hwang, I., Young Min, H., Bang, Y.-J., & Kim, W. (2019). Clinical significance of BRCA1 and BRCA2 mRNA and protein expression in patients with sporadic gastric cancer. *Oncology letters*, 17(5), 4383-4392. doi:10.3892/ol.2019.1013
- Kong, S., Lee, I.-H., Liu, X., Hirschhorn, J., & Mandl, K. (2018). Measuring Coverage and Accuracy of Whole Exome Sequencing in Clinical Context. *Kong, S. W., Lee, I. H., Liu, X., Hirschhorn, J. N., & Mandl, K. D. (2018). Measuring coverage and accuracy of whole-exome sequencing in clinical context. Genetics in medicine : official journal of the American College of Medical Genetics*, 20(12), 1617-1616.
doi:10.1038/gim.2018.51

- Konnick, E., & Pritchard, C. (2016). Germline, hematopoietic, mosaic, and somatic variation: interplay between inherited and acquired genetic alterations in disease assessment. *Genome medicine*, 8(1), 100. doi:10.1186/s13073-016-0350-8
- Kuballa, P., Baumann, A.-L., Mayer, K., Bär, U., Burtscher, H., & Brinkmann, U. (2015). Induction of heat shock protein HSPA6 (HSP70B0) upon HSP90 inhibition in cancer cell lines. *FEBS letters*, 589(13), 1450-1458. doi:10.1016/j.febslet.2015.04.053
- LaDuca, H., Farwell, K., Vuong, H., Mei Lu, H., Mu, W., Shahmirzadi, L., . . . Chao, E. (2017). Exome sequencing covers >98% of mutations identified on targeted next generation sequencing panels. *PLoS One*, 12(2). doi:10.1371/journal.pone.0170843
- Leoz, M., Sánchez, A., Carballal, S., Ruano, L., Ocaña, T., Pellisé, M., . . . Moreira, L. (2016). Síndromes de predisposición a cáncer gástrico y cáncer pancreático. 39(7), 481-493. doi:10.1016/j.gastrohep.2015.11.009
- Li, H., & Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics (Oxford, England)*, 26(5). doi:10.1093/bioinformatics/btp698
- Li, S., Yuan, Y., Xiao, H., Dai, J., Ye, Y., Zhang, Q., . . . Wang, G. (2019). Discovery and validation of DNA methylation markers for overall survival prognosis in patients with thymic epithelial tumors. *Clinical Epigenetics*, 11(38), 1-12. doi:10.1186/s13148-019-0619-z
- Linhao, Z., Yongjun, E., Fanglin, N., Mengdan, Y., Jing, Y..... (2018). Polymorphisms of drug-metabolizing enzyme *CYP2E1* in Chinese Uygur population, *Medicine*: 97(7), e9970 doi: 10.1097/MD.0000000000009970
- Lichtenstein, P., Holm, N., Verkasalo, P., Iliadou, A., Kaprio, J., Koskenvuo, M., . . . Hemminki, K. (2000). Environmental and Heritable Factors in the Causation of Cancer — Analyses of Cohorts of Twins from Sweden, Denmark, and Finland. *Survey of ophthalmology*, 45(2), 167-168. doi:10.1016/s0039-6257(00)00165-x

- Lu, Y., Ek, W., Whiteman, D., Vaughan, T., Spurdle, A., Easton, D., . . . Macgregor, S. (2014). Most common ‘sporadic’ cancers have a significant germline genetic component. *Human molecular genetics*, 23(22), 6112-6118. doi:10.1093/hmg/ddu312
- Luo, M., & Liang, C. (2020). LncRNA LINC00483 promotes gastric cancer development through regulating MAPK1 expression by sponging miR-490-3p. *Biological research*, 53(1), 14. doi:10.1186/s40659-020-00283-6
- Malhotra, A., Levine, S., Allingham-Hawkins, D. (2014.) Whole exome sequencing for cancer – is there evidence of clinical utility? . *Advances in Genomics and Genetics*, 4, 115-128. doi.org/10.2147/AGG.S58809
- Marín, F. (2013). Análisis de la variabilidad genética de los genes MUC y TFF de protección de la mucosa gástrica en relación al cáncer gástrico y las lesiones precursoras (tesis doctoral). Barcelona, España: Universidad Autónoma de Barcelona.
- Martínez, T., Hernández, G., Bravo, M., Trujillo, E., Pérez, J., Robayo, J., & Camorlinga, M. (2014). Lesiones preneoplásicas gástricas en pacientes colombianos: asociación de polimorfismos genéticos interleucinas 1B-511, 1RN, 10-819, 10-1082, factor de necrosis tumoral- α -308 y anticuerpos inmunoglobulina G hacia cagA de *Helicobacter pylori*. *Revista Colombiana de cancerología*, 18(1), 8-17.
- Orozco, E., Sanchez-Fernandez, A., Ortiz-Parra, I., & Ayala-San Nicolas, M. (2019). WNT Signaling in Tumors: The Way to Evade Drugs and Immunity. *Frontiers in immunology*, 10, 22. doi:10.3389/fimmu.2019.02854
- Mellman, I., & Yarden, Y. (2013). Endocytosis and Cancer. *Cold Spring Harbor perspectives in biology*, 5(12), a016949. doi:10.1101/cshperspect.a016949
- Metzker, M. (2010). Sequencing technologies-the next generation. *Nature*, 461, 31-46.

- Meyerson, M., Gabriel, S., & Getz, G. (2010). Advances in understanding cancer genomes through second-generation sequencing. *Nature reviews. Genetics*, 11(10), 685-696. doi:10.1038/nrg2841
- Montenegro, F., & Carlosama, Y. (2016). Relación entre polimorfismos de cuatro genes y patrón de respuesta inflamatoria en pacientes con gastritis crónica y cáncer gástrico en una población de Nariño. *Trabajo de grado Maestría*. Pasto: Universidad CES.
- Nakagawa, H., Wardell, C., Furuta, M., Taniguchi, H., & Fujimoto, A. (2015). Cancer whole-genome sequencing: present and future. *Oncogene*, 34(49), 5943-5950. doi:10.1038/onc.2015.90
- Ogretmen, B. (2018). Sphingolipid metabolism in cancer signalling and therapy. *Cancer*, 18(1), 33-50. doi:10.1038/nrc.2017.96
- Otero Rodríguez, T. (2018). *Identificación de variantes genéticas poco frecuentes en diabetes mellitus tipo 2 mediante secuenciación masiva de amplicones*. Universidad Politécnica de Valencia, TRABAJO FIN DE MÁSTER EN BIOTECNOLOGÍA BIOMÉDICA, Valencia.
- Otzen, T., & Manterola, C. (2017). Técnicas de Muestreo sobre una Población a Estudio. *Int. J. Morphol*, 35(1), 227-232.
- Pardo, C., & Cendales, R. (2015). *Incidencia, mortalidad y prevalencia de cáncer en Colombia, 2007-2011* (Primera edición ed., Vol. 1). Bogotá: Instituto Nacional de Cancerología.
- Paz, C. (2013). incidencia de factores genetico-ambientales en la susceptibilidad al cáncer gástrico en pacientes del municipio de Pasto. *Tesis de pregrado*. Pasto: Universidad de Nariño.
- Penzo, M., Montanaro, L., Treré, D., & Derenzini, M. (2019). The Ribosome Biogenesis—Cancer Connection. *Cells*, 8(1), 55. doi:10.3390/cells8010055

- Peralta-Rodriguez, R., Valdivia, A., Mendoza, M., Rodríguez, J., Marrero, D., Paniagua, L., . . . Salcedo, M. (2015). Los genes del cáncer. *Rev Med Inst Mex Seguro*, 178-187.
- Piñero, J., Ramírez-Anguita, J. M., Saüch-Pitarch, J., Ronzano, F., Centeno, E., Sanz, F., & Furlong, L. I. (2020). The DisGeNET knowledge platform for disease genomics: 2019 update. *Nucleic acids research*, 48(D1), D845–D855. <https://doi.org/10.1093/nar/gkz1021>
- Poplin, R., Ruano-Rubio, V., DePristo, M., Fennell, T., Carneiro, M., Van der Auwera, G., . . . Banks, E. (2017). Scaling accurate genetic variant discovery to tens of thousands of samples. *bioRxiv*, 1-22. doi:10.1101/201178
- Ramírez, J., Vargas, G., Tovilla, C., & Fragoso, J. (2013). *Gaceta Médica de México*, 149, 220-228.
- Ramos, M., Ribeiro, U., Juliana, V., Zilberstein, B., Cecconello, I., & Eluf-Neto, J. (2018). Risk factors associated with the development of gastric cancer - case-control study. *Revista da Associacao Medica Brasileira (1992)*, 64(7), 611-619. doi:10.1590/1806-9282.64.07.611
- Rawla, P., & Barsouk, A. (2019). Epidemiology of gastric cancer: global trends, risk factors and prevention. *Rawla, P., & Barsouk, A. (2019). Epidemiology of gastric cancer: global trends, risk factoPrzegląd gastroenterologiczny*, 14(1), 26-38. doi:10.5114/pg.2018.80001
- Ribeiro, D., Liehr, T., Emerenciano, M., Meyer, C., Marschalek, R., Pombo-de-Oliveira, M., . . . Macedo, M. (2015). Molecular studies reveal a MLL-MLLT3 gene fusion displaced in a case of childhood acute lymphoblastic leukemia with complex karyotype. *Cancer Genet*, 208(4), 143-147. doi:10.1016/j.cancergen.2015.02.002
- Rivas, C., Lopez, Y., Arredondo, L., & Castillo, G. (2017). Genetic Alterations in Gastric Cancer Associated with *Helicobacter pylori* infection. *Front Med (Lausanne)*, 1-12.

- Rivera, X., Cobos, O., & Remes, J. (2016). Los efectos carcinogénicos del acetaldehído. Una visión actual. *Gaceta Mexicana de Oncología*, 15(4), 231-239. doi:10.1016/j.gamo.2016.07.007
- Rodríguez, L. (2019). Hallazgos secundarios en la secuenciación del exoma según las recomendaciones del Colegio Americano de Genética Médica y Genómica (ACMG) en una muestra de la población colombiana. Tesis para otorgar al título de Master en epidemiología. Universidad del Rosario. Bogotá, Colombia.
- Rosero, C., Corredor, M., & Mejía, L. (2016). Polimorfismos en genes implicados en el desarrollo de cáncer gástrico: revisión. *Gastroenterology*, 152(5), 983-986.
- Rosero, C., Mejía, L., & Corredor, M. (2019). Variantes patogénicas de TP53 relacionadas con cáncer, *Journal of Basic and Applied Genetics*(2), 27-40. doi:10.35407/bag.2019.xxx.02.03
- Rosero, S., Chaleal, J., Pazos, A., & Rosero, C. (2020). Polimorfismos IL1B-511 y TNF-A-308 en una población infectada con *Helicobacter pylori* de una zona de bajo riesgo de cáncer gástrico en Nariño-Colombia. *Infectio*, 24(2), 81-87. doi:10.22354/in.v24i2.837.
- Rotunno, M., Barajas, R., Clyne, M., Hoover, E., Simonds, N..... (2020). A Systematic Literature Review of Whole Exome and Genome Sequencing Population Studies of Genetic Susceptibility to Cancer. *American Association for Cancer Research*, 29:1519–34. doi: 10.1158/1055-9965.EPI-19-1551
- S.A, Colombiana de Salud. (Marzo de 2014). Manual para la toma, transporte y conservación de muestras de laboratorio clínico. Colombia.
- Sahasrabudhe, R., Lott, P., Bohorquez, M., Toal, T., Estrada, A., Suarez, J., . . . Prieto, R. (2016). Germline Mutations in *PALB2*, *BRCA1*, and *RAD51C*, Which Regulate DNA Recombination Repair, in Patients With Gastric Cancer. *Gastroenterology*, 152(5), 983-986. doi:10.1053/j.gastro.2016.12.010.

- Sanger, F., Nicklen, S., & Coulson, A. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A*, 5463–5467.
- Sanz, J., Blasco, A., Arrinda, J., & Olmedilla, G. (2005). PATOLOGÍA GÁSTRICA: Lesiones precursoras de c áncer gástrico. Revisión. 7º Congreso Virtual Hispanoamericano de Anatomía Patológica, 1-29.
- Sastre, L. (2014). Exome sequencing: what clinicians need to know. *Advances in Genomics and Genetics*, 4(4), 15-27. doi:10.2147/AGG.S39108
- Schwingshackl, L., & Hoffmann, G. (2015). Adherence to Mediterranean diet and risk of cancer: an updated systematic review and meta- analysis of observational studies. *Cancer Med*, 933-1947.
- Setia, N., Wang, C., Lager, A., Maron, S., Shroff, S.....(2020). Morphologic and molecular analysis of early-onset gastric cáncer. *American Cancer Society*, 127(1): 103-114 doi.org/10.1002/cncr.33213
- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., & Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research*, 13(11), 2498–2504. <https://doi.org/10.1101/gr.1239303>
- Sharma, S., Kelly, T., & Peter , A. (2010). Epigenetics in cancer. *Carcinogenesis*, 27-36.
- Shendure, J., & Ji, H. (2008). Next-generation DNA sequencing. *Nat Biotechnol*, 1135-1145.
- Silva Zolezzi, I. (2011). Genómica y medicina. *Educ. quím*, 22(1), 15-27.
- Slavin, T., Neuhausen, SL, Rybak, C., Solomon, I., Nehoray, B., Blazer, K., Niell-Swiller, M., Adamson, AW, Yuan, YC, Yang, K., Sand , S., Castillo, D., Herzog, J., Wu, X., Tao, S., Chavez, T., Woo, Y., Chao, J., Mora, P., Horcasitas, D.,... Weitzel, J. (2017). Genetic Gastric Cancer Susceptibility in the International Clinical Cancer Genomics Community Research Network. *Cancer genetics*, 216-217, 111-119. <https://doi.org/10.1016/j.cancergen.2017.08.001>

- Suárez Moya, A. (2017). Microbioma y secuenciación masiva. *Revista Española de Quimioterapia*, 1-7.
- Suarez Olaya, J. (2016). *Análisis genético y clínico-patológico del carcinoma gástrico en una muestra de pacientes colombianos*. Universidad del Tolima, Ibagué.
- Suárez, A. (Septiembre de 2015). Utilización de herramientas genómicas en el estudio de la arquitectura molecular de caracteres de importancia económica en el ganado ovino. *Tesis doctoral*. León, España: Universidad de León.
- Suzuki, K., Katoh, H., Komura, D., Kakiuchi, M., Tagashira, A..... Defined lifestyle and germline factors predispose Asian populations to gastric cancer. (2020). *Science Advances*. 6(19), 1-13. DOI: 10.1126/sciadv.aav9778
- Tetreault, M., Bareke, E., Nadaf, J., Alirezaie, N., & Majewski, J. (2015). Whole-exome sequencing as a diagnostic tool: current challenges and future opportunities. *Expert review of molecular diagnostics*, 15(6), 749-760. doi:10.1586/14737159.2015.1039516
- The Cancer Genome Atlas Research Network. (2014). Comprehensive molecular characterization of gastric adenocarcinoma. *Nature*, 513(7517):202-9. doi: 10.1038/nature13480. Epub 2014 Jul 23.
- Torres, K., Labrador, L., Valderrama, E., & Chiurillo, M. (2016). *TCF7L2* rs7903146 polymorphism is associated with gastric cancer: A case-control study in the Venezuelan population. *World journal of gastroenterology*, 22(28), 6520-6526. doi:10.3748/wjg.v22.i28.6520
- Torres, M., Acosta, C., Sicard, D., & Groot, H. (2004). Susceptibilidad genética y riesgo de cáncer gástrico en una población del Cauca. *Biomédica*, 24, 153-162.
- Vogelaar, I. P., van der Post, R. S., van Krieken, J., Spruijt, L., van Zelst-Stams, W. A., Kets, C. M., Lubinski, J., Jakubowska, A., Teodorczyk, U., Aalfs, C. M., van Hest, L. P., Pinheiro, H., Oliveira, C., Jhangiani, S. N., Muzny, D. M., Gibbs, R. A., Lupski, J. R., de Ligt, J., Vissers,

- L., Hoischen, A., ... Hoogerbrugge, N. (2017). Unraveling genetic predisposition to familial or early onset gastric cancer using germline whole-exome sequencing. *European journal of human genetics : EJHG*, 25(11), 1246–1252. <https://doi.org/10.1038/ejhg.2017.138>
- Urbini, M., Nannini, M., Astolfi, A., Indio, V., Vicennati, V., De Luca, M., . . . Pantaleo, M. (2018). Whole Exome Sequencing Uncovers Germline Variants of Cancer-Related Genes in Sporadic Pheochromocytoma. *International journal of genomics*(6582014), 1-9. doi:10.1155/2018/6582014
- Wang, R.-Y., Chen, X.-W., Zhang, W.-W., Jiang, F., Liu, M.-Q., & Shen, X.-B. (2020). CYP2E1 changes the biological function of gastric cancer cells via the PI3K/Akt/mTOR signaling pathway. *Molecular medicine reports*, 21(2), 842-850. doi:10.3892/mmr.2019.10890
- Wang, S. C., Yeu, Y., Hammer, S., Xiao, S., Zhu, M., Hong, C., Clemenceau, J. R., Yoon, L. Y., Nassour, I., Shen, J., Agarwal, D., Reznik, S. I., Mansour, J. C., Yopp, A. C., Zhu, H., Hwang, T. H., & Porembka, M. R. (2020). Hispanic/Latino Patients with Gastric Adenocarcinoma Have Distinct Molecular Profiles Including a High Rate of Germline CDH1 Variants. *Cancer research*, 80(11), 2114–2124. <https://doi.org/10.1158/0008-5472.CAN-19-2918>
- Wang, X., Liu, B., Wen , F., & Song, Y. (2018). MicroRNA-454 inhibits the malignant biological behaviours of gastric cancer cells by directly targeting mitogen-activated protein kinase 1. *Oncology Reports*, 39, 1494-1504. doi:10.3892/or.2017.6171
- Welander, J., Andreasson, A., Juhlin, C., Wyuiseman, R., Bäckdahl, M., Höög, A., . . . Söderkvist, P. (2014). Rare germline mutations identified by targeted next-generation sequencing of susceptibility genes in pheochromocytoma and paraganglioma. *The Journal of clinical endocrinology and metabolism*, 99(7), E1352-E1360. doi:10.1210/jc.2013-4375
- Weren, R., van der Post, R., Vogelaar, I., van Krieken, K., Spruijt, L..... (2018). Role of germline aberrations affecting CTNNA1, MAP3K6 and MYD88 in gastric cancer susceptibility. *Journal of Medical Genetics*, 55(10):669-674.

doi: 10.1136/jmedgenet-2017-104962

Yáñez-Gutiérrez, C. M. (2013). Los genes como marcadores de riesgo en cáncer gástrico:revisión de estudios en población colombiana. *Respuestas*, 18(2), 61-73.

Yoshida, N., Hisashi, D., Yano, T., Horimatsu, T., Uedo, N., Ya, Y., . . . Yokoi, C. (2020). Early gastric cancer detection in high-risk patients: a multicentre randomised controlled trial on the effect of second-generation narrow band imaging. *Gut*, *gutjnl-2019-319631*, 1-9. doi:10.1136/gutjnl-2019-319631

Yuan, C., Zhu, X., Han, Y., Song, C., Liu, C., Lu, S., Zhang, M., Yu, F., Peng, Z., & Zhou, C. (2016). Elevated HOXA1 expression correlates with accelerated tumor cell proliferation and poor prognosis in gastric cancer partly via cyclin D1. *Journal of experimental & clinical cancer research*: CR, 35, 15. <https://doi.org/10.1186/s13046-016-0294-2>

Zhang, J., Nichols, K., & Downing, J. (2016). Germline Mutations in Predisposition Genes in Pediatric Cancer. *The New England journal of medicine*, 374(14), 1391. doi:10.1056/NEJMc1600338.

Zhang, J., Qiu, W., Liu, H., Qian, C., Liu, D., Wang, H., . . . Shen, Z. (2018). Genomic alterations in gastric cancers discovered via whole-exome sequencing. *BMC cancer*, 18(1), 1270. doi:10.1186/s12885-018-5097-8.

Zhu Z, Fu H, Wang S, Yu X, You Q, Shi M, Dai C, Wang G, Cha W, Wang W. (2020). Whole-exome sequencing identifies prognostic mutational signatures in gastric cancer. *Ann Transl Med*, 8(22), 1484.

doi: 10.21037/atm-20-6620

13 ANEXOS

Anexo 1. Aval del comité de ética del proyecto “PERFIL GENÉTICO DE CÁNCER GÁSTRICO EN NARIÑO: ESTUDIO DE CASOS EN LÍNEA SOMÁTICA Y GERMINAL”



Subcomité de Bioética
Universidad Cooperativa de Colombia – Pasto
Calle 18 No. 47-150 Teléfono: 7338706

Acta de Evaluación Proyecto de Investigación

ACTA No. SCBE06-17

Proyecto de Investigación No. 6

En Pasto, al 15^{avo} día del mes de Junio de 2017, el H. Consejo del Subcomité de Bioética de la Universidad Cooperativa de Colombia - Pasto, con asistencia de sus miembros permanentes:

Odontóloga, MSP, EDU *Ana Cristina Mafla Chamorro* como Presidente (E); Odontólogo, Esp. Gerencia y Auditoría en Salud *Fernando Dávila Narváez* como Secretario; Microbióloga Industrial, Mg. Microbiología, Doctora en Ciencias Agrarias *Ana María Rey Obando*; Fisioterapeuta, Esp. y Mg. en Intervención Integral en el Deportista, Mg. Gerontología, Estudiante de Doctorado en Proyectos *Giovana Rosas Estrada* y Abogado, Esp. en Derecho Laboral, Mg. Derecho del Trabajo *Hugo Armando Medina Chávez* han revisado los documentos presentados por CAROL YOVANNA ROSERO GALINDO, a saber:

1. Formato del Proyecto “Perfil genético de cáncer gástrico en Nariño: Estudio de casos en línea somática y germinal”.
2. Formato de Consentimiento Informado del Proyecto “Perfil genético de cáncer gástrico en Nariño: Estudio de casos en línea somática y germinal”.
3. Formato de Consentimiento para tratamiento y almacenamiento de datos personales del Proyecto “Perfil genético de cáncer gástrico en Nariño: Estudio de casos en línea somática y germinal”.
4. El Currículo Vitae (link) de los investigadores registrado en la plataforma CvLAC – Colciencias.

Con los anteriores documentos el Subcomité ha considerado que el proyecto de investigación no presenta reparos éticos y se exponen las siguientes observaciones:

- 1) Los miembros del Subcomité de Bioética declararon no tener conflicto de interés.
- 2) El diseño se ajusta a las normas de Investigación en Seres Humanos.
- 3) La razón beneficio fue estimada aceptable.
- 4) Los formatos de consentimientos cumplen con los requisitos exigidos.
- 5) Los antecedentes curriculares de los Investigadores, garantizan la ejecución del Proyecto de Investigación dentro de los marcos éticamente aceptables.

En consecuencia, el Subcomité de Bioética de la Universidad Cooperativa de Colombia - Pasto aprueba por unanimidad de sus miembros el proyecto de investigación "Perfil genético de cáncer gástrico en Nariño: Estudio de casos en línea somática y germinal". Dicho estudio se llevará a cabo en el Departamento de Nariño y será realizado bajo la supervisión de CAROL YOVANNA ROSERO GALINDO de la Universidad Cooperativa de Colombia - Pasto como Investigadora Principal.



Ana Cristina Mafla Chamorro
Presidente (E)



Fernando Dávila Narváez
Secretario

C/C. Investigadora Principal. Secretaría SCBE.

Anexo 2. Formato de consentimiento informado.

FORMATO DE CONSENTIMIENTO INFORMADO DE MUESTRAS PARA ESTUDIOS DE EVALUACIÓN GENÉTICA Y BANCO DE ESPECÍMENES (adultos)

PERFIL GENÉTICO DE CÁNCER GÁSTRICO EN NARIÑO: ESTUDIO DE CASOS EN LÍNEA SOMÁTICA Y GERMINAL

*Sólo si Usted está legalmente autorizado como representante de una persona que va a ser parte en este estudio, su permiso será requerido y también el asentimiento (o acuerdo) de su familiar o custodiado. Cuando la palabra “Usted” aparece en este formato de Consentimiento, esta se refiere a la persona para quien Usted está legalmente autorizado representar.

A Usted se le dará el tiempo suficiente para leer y entender la información provista en este formato de Consentimiento y Autorización. Por favor pida al responsable del grupo de estudio explicar las palabras que Usted no entiende. Usted puede mantener una copia no firmada de este Consentimiento y formato de Autorización para pensar acerca de su decisión o discutir con su familia, amigos o su médico antes de decidir ser parte de este estudio.

Antecedentes/Propósito:

A usted se le ha pedido que participe en esta investigación porque ha sido diagnosticado con cáncer gástrico (CG). El cáncer gástrico (CG) es la segunda neoplasia maligna más frecuente en el mundo, en Colombia constituye la principal causa de muerte por cáncer en hombres y la tercera en mujeres y en la región Andina de Nariño su incidencia es una de las más altas del mundo. El CG es el resultado final de un largo proceso multifactorial que se desarrolla en distintas etapas y en el que intervienen un elevado número de factores ambientales y genéticos. En Colombia y particularmente en Nariño, se han adelantado estudios de polimorfismos en genes candidatos de desarrollo de CG, sin embargo, la mayoría de ellos no demuestran asociación con la enfermedad. En este sentido, la secuenciación del exoma permite una visión detallada de las regiones codificantes del genoma, proporcionando una perspectiva global para el estudio de variaciones genéticas, comunes, raras y nuevas, sin tener que limitarse a un subconjunto de genes. En un contexto genético, cada tumor contiene variantes heredadas (línea germinal) y tumor-específicas (somáticas), por tanto, la secuenciación y la comparación del ADN normal con el ADN del tumor de un individuo afectado, teóricamente permitiría una identificación y una separación precisa de las alteraciones de la línea germinal de los cambios somáticos, favoreciendo la interpretación de dichas variantes en el desarrollo de la enfermedad y un tratamiento adecuado del paciente. Este estudio se plantea con el objetivo principal de caracterizar el exoma de línea somática y

de línea germinal de pacientes diagnosticados con adenocarcinoma gástrico de tipo esporádico en el departamento de Nariño. Para ello se analizará el tejido germinal (sangre) y somático (tumor) de los pacientes incluidos en el estudio. Con los resultados de esta investigación se pretende caracterizar a nivel molecular al paciente con cáncer, con el fin de generar conocimiento sobre las variantes implicadas en el desarrollo de la enfermedad, que en el mediano o largo plazo permita la identificación de potenciales biomarcadores para el diagnóstico temprano y un tratamiento adecuado de la patología.

Objetivos del estudio (general y específicos):

General

Caracterizar el exoma de línea somática (tumor) y de línea germinal (sangre) de pacientes diagnosticados con adenocarcinoma gástrico de tipo esporádico en el departamento de Nariño.

Específicos

1. Identificar variantes como polimorfismos de un solo nucleótido (SNPs), pequeñas inserciones o deleciones (Indels) y otras variantes estructurales en el exoma de pacientes con cáncer gástrico.
2. Determinar las variantes genéticas de línea germinal y las variantes somáticas en el exoma del paciente con cáncer gástrico.
3. Establecer las variantes genéticas germinales y somáticas implicadas en la carcinogénesis gástrica para la población de estudio.
4. Proponer variantes genéticas somáticas de importancia terapéutica para el tratamiento personalizado del paciente con cáncer gástrico.

Sujetos que participarán en el estudio (población y muestra, tipo de muestreo):

Se implementará un muestreo por conveniencia, para la obtención de tejido tumoral proveniente de biopsias gástricas y una muestra pareada de sangre periférica de 6 pacientes con diagnóstico certero de cáncer gástrico.

Procedimiento del estudio:

Se van incluir en el estudio un total de 6 pacientes diagnosticados con adenocarcinoma gástrico y se colectarán de cada paciente, una muestra de tejido tumoral fresco proveniente de una biopsia gastroscópica y una muestra de sangre periférica a partir de venopunción.

Las muestras de ADN extraídas a partir del tejido gástrico tumoral y de sangre periférica, serán sometidas al proceso de secuenciación de las regiones codificadoras de proteínas (exomas) y posteriormente las lecturas o “reads” obtenidas serán computacionalmente alineadas a una secuencia de referencia para el procesamiento secundario de los datos que implica el llamado de variantes y el filtrado usando criterios aplicados al control de calidad, la anotación de variantes y la identificación de posibles variantes implicada en el desarrollo de cáncer gástrico. Además, se identificarán aquellas variantes de línea somática (propias del tumor) y aquellas de tipo germinal (de nacimiento). Finalmente, se obtendrá un reporte enfocado en aquellos cambios genéticos de importancia y relevancia clínica con el cáncer gástrico, en cuanto a predisposición y tratamiento de la enfermedad.

Las muestras colectadas de cada paciente, se comenzarán a analizar a partir de la estandarización de los procesos, probablemente en el primer semestre de 2018.

Investigación Genética

Sus muestras de tejido tumoral y sangre periférica podrán ser evaluadas para factores genéticos y la información obtenida podrá revelar información genética sobre Usted. Esta información se mantendrá con absoluta confidencialidad.

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

Banco de Especímenes (muestras)

Además de la investigación a la que Usted está consintiendo en este estudio, la Doctora Carol Yovanna Rosero, solicita su permiso para guardar en un banco su biopsia gástrica tumoral y su muestra sanguínea para investigación futura. Estas muestras, tomadas de su cuerpo, serán capaces de ser vinculadas nuevamente con Usted. Sus especímenes pueden ser enviados a un banco de tejidos y ADN. Las muestras pueden mantenerse por un largo tiempo, posiblemente más de 50 años. La información sobre Usted puede ser compartida con otros investigadores que mantendrán la información confidencial, como se indica en este consentimiento.

POR FAVOR INDIQUE EL CONSENTIMIENTO CON UNA INICIAL DEL NOMBRE

Mis especímenes (muestras) pueden ser guardados para investigaciones futuras, aunque el propósito de la investigación no se conozca en este momento.

_____ Consiento que mis muestras sean guardadas para futuros estudios de investigación.

_____ No consiento que mis muestras sean guardadas para futuros estudios de investigación.

Riesgos:

Con base en lo consignado en la resolución 8430 de 1993, la presente investigación implica un riesgo menor que el mínimo como resultado del procedimiento para la toma de muestras. La gastroscopia, puede incluir complicaciones propias del procedimiento la mayoría de ellas leves y sin repercusión alguna, entre las que se encuentran: Nauseas; Distensión o dolor abdominal, que suelen ser frecuentes pero leves y tolerables, también reacciones no deseadas a la medicación administrada; y una pequeña hemorragia, que puede ocurrir tras la toma de muestra de biopsia que cede espontáneamente.

En el caso de la toma de muestra sanguínea, puede producirse un leve dolor, ardor o mareo en el paciente. Como riesgos menos comunes, también puede presentarse un mínimo hematoma en la zona del pinchazo, por lo que será conveniente que después se realice presión sobre la zona puncionada, desmayos e infecciones en el área de punción.

Beneficios:

En cualquier estudio de investigación médica ciertos beneficios se pueden derivar. Estos beneficios incluyen la posibilidad a largo plazo, de acceder a un asesoramiento genético sobre la enfermedad que padece, así como un asesoramiento genético familiar. Además, la información aprendida durante este estudio puede ayudar a otros en el futuro, en el sentido de que permitirá caracterizar desde una perspectiva molecular a los pacientes con cáncer. Los resultados obtenidos de variantes genéticas para línea germinal y somática pueden permitir la identificación de potenciales dianas terapéuticas, así como dar información de posibles genes de predisposición al cáncer gástrico, que en el mediano o largo plazo se puedan implementar para prevención, diagnóstico y tratamiento precoz de la enfermedad.

De los resultados de esta investigación, son potenciales beneficiarios las comunidades que estén en alto riesgo de cáncer gástrico.

Participación Voluntaria:

Su participación en este estudio es completamente voluntaria y Usted puede negarse a participar o dejar de participar en cualquier momento, sin sanción o pérdida de beneficios a los cuales Usted normalmente tendría derecho. La decisión sobre si desea o no participar en el estudio no afectará a la atención que recibe en su Entidad Prestadora de Salud a la que se encuentra afiliado o su relación con la Universidad Cooperativa de Colombia.

Alternativas:

- ☐ Si Usted decide no participar en este estudio de investigación, Usted seguirá recibiendo su atención médica habitual y no se le realizará ninguno de los procedimientos anteriormente mencionados con fines de investigación.
- ☐ Si Usted decide no participar en este estudio, sus muestras no serán usadas.
- ☐ Si Usted decide no participar en este estudio, su información médica no será utilizada para este estudio.

Costos/Pagos:

- ☐ No hay ningún costo para Usted o a su aseguradora por participar en este estudio. A Usted no se le pagará por participar en este estudio.

Preguntas:

- Si Usted tiene alguna pregunta acerca de la investigación, por favor póngase en contacto con el investigador (a) al (315) 4155508
- Si Usted tiene alguna pregunta sobre sus derechos como sujeto de investigación, por favor póngase en contacto con el Subcomité de Bioética al 7336706.

Confidencialidad de los Registros y autorización para usar/compartir Información de Salud Protegida para Investigación:

Si Usted está de acuerdo en participar en esta investigación, la información de salud identificable sobre Usted será usada y compartida con otras personas involucradas en esta investigación. Para que Usted pueda estar en esta investigación necesitamos su permiso para recopilar y compartir esta información. La Ley protege su derecho a la privacidad con respecto a esta información.

Al firmar este formulario de consentimiento al final, significa que ha leído esta sección y autoriza usar y/o compartir su Información de Salud Protegida, como se explica a continuación. Su firma también significa que Usted ha recibido una copia de la información del Subcomité de Bioética sobre prácticas de privacidad. La información sobre prácticas de privacidad debería ser dada a cada sujeto/padre en el momento del consentimiento sea obtenido para participar en el estudio.

La información individual de salud identificable, según la Ley de privacidad, es considerada como cualquier información de su historia clínica, u obtenida de este estudio, que se puede ser asociada con Usted, y se refiere a su pasada, presente o futura salud física o mental o condición. Esto se conoce como Información de Salud Protegida.

Su Información de Salud Protegida se mantendrá confidencial. Su identidad no será revelada en ninguna publicación o presentación de resultados de esta investigación.

¿Por qué es necesario usar/compartir su Información de Salud Protegida con otros?

La razón principal para usar y compartir su Información de Salud es llevar a cabo la investigación como se describe en este formulario de consentimiento. Su información también puede ser compartida con personas y organizaciones que hacen que la investigación se está haciendo correctamente, y para informar inesperados o malos efectos colaterales que Usted pueda tener.

Además, podemos estar obligados por Ley a revelar la Información de Salud Protegida sobre Usted, por ejemplo, si un juez requiere dicha revelación en una demanda, o si Usted nos dice de su intención de hacerse daño a sí mismo o a los demás.

¿Qué Información de Salud Protegida sobre Usted será utilizada o compartida con otros como parte de esta investigación?

Nosotros podemos usar y compartir resultados de pruebas, cuestionarios y entrevistas. También pudiéramos utilizar y compartir la información de su historia clínica y de investigación. Sólo recopilaremos la información necesaria para la investigación.

¿Quién estará autorizado a usar y/o compartir su Información de Salud Protegida?

Los investigadores, su personal de apoyo y el personal de la Universidad Cooperativa de Colombia participante en la investigación usarán su Información de Salud Protegida. Además, el Subcomité de Bioética responsable de proteger los derechos de los sujetos de investigación, y otro personal de la Universidad Cooperativa de Colombia quien supervise la forma en que se realiza la investigación podrá tener acceso a su Información de Salud Protegida.

Los investigadores y su personal determinarán si su Información de Salud Protegida será utilizada o compartida con otras personas fuera de la Universidad Cooperativa de Colombia para fines directamente relacionados con la realización de la investigación.

¿Con quién la Información de Salud Protegida se compartirá?

Su Información de Salud Protegida podrá ser compartida con:

- Las Entidades patrocinadoras de este estudio: Universidad Cooperativa de Colombia, Universidad del Valle y Universidad de Chile.

Todos los esfuerzos razonables se utilizarán para proteger la confidencialidad de su información de salud protegida. Sin embargo, no todos los individuos o grupos tienen que cumplir con la ley de privacidad. Por lo tanto, una vez que su información de salud protegida se divulgue (dejada en la Universidad Cooperativa de Colombia), la ley de privacidad pudiera no protegerla.

¿Durante cuánto tiempo su información protegida de salud será utilizada o compartida con otros?

No hay una fecha programada en la que esta información será destruida o ya no se utilizará. Esto se debe a que la información que se recopila con fines de investigación sigue siendo utilizada y analizada durante muchos años y no es posible determinar cuándo se completará.

¿Puedo retirar la autorización para recopilar/ usar / compartir su información médica protegida?

Usted tiene el derecho de retirar su permiso (revocar autorización) para que nosotros usemos y compartamos su información de salud, poniendo su solicitud por escrito al investigador a

cargo del estudio. Esto significa que no se recogerá más información privada sobre la salud. Una vez que se revoca la autorización, es posible que ya no participe en esta actividad de investigación, pero la atención médica estándar y cualquier otro beneficio al que tenga derecho no se verán afectados. La revocación de su autorización sólo afecta a los usos y el intercambio de información obtenida después de que se haya recibido su solicitud por escrito, pero no de información obtenida antes de ese momento.

Incluso después de retirar su permiso, la Universidad Cooperativa de Colombia puede continuar utilizando y compartir la información necesaria para la integridad del estudio; por ejemplo, información sobre un mal o inesperado efecto secundario que experimentó relacionado con el estudio.

¿Puede tener acceso a su información de salud?

Al final del estudio, Usted tiene derecho a ver y copiar información de salud sobre Usted de acuerdo con las políticas de la Universidad Cooperativa de Colombia; sin embargo, su acceso pudiera ser limitado mientras el estudio está en progreso.

Consentimiento para participar en la investigación y autorización para usar y compartir información personal de salud.

Por la presente doy mi consentimiento para participar en este estudio de investigación y estoy de acuerdo en que mi información personal de salud puede ser recopilada, utilizada y compartida por los investigadores y el personal para el estudio de investigación descrito en este formulario. Yo recibiré una copia firmada de este formulario de consentimiento.

Firma del Participante

Fecha

Firma del Representante Legal Autorizado (*si aplica*)

Fecha

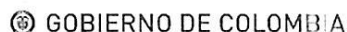
Relación con el participante (*si aplica*)

Firma de la Persona que obtiene

El Consentimiento/Autorización

Fecha

Anexo 3. Autorización para salida de muestras biológicas por parte del Instituto Nacional de Vigilancia de Medicamentos y Alimentos (INVIMA) adscrito al Ministerio de Salud y Protección Social



Al contestar por favor cite estos datos:

Radicado No.: **201824000498311**

Fecha: **02-05-2018**

Página 1 de 2

Bogotá D.C.,

Señores

UNIVERSIDAD COOPERATIVA DE COLOMBIA

MARITZA RONDÓN RANGEL

Rectora y Representante Legal

Carrera 42 49 - 95

Medellín, Antioquia

ASUNTO: Radicado N° 201842300527092. Salida de muestras biológicas de origen humano con fines de investigación.

Respetada Doctora Rondón,

En atención al radicado del asunto, en el cual solicita autorización para la salida del país de muestras biológicas de origen humano con fines de investigación que serán utilizados en el marco de los proyectos "Exoma de cáncer gástrico en Nariño: identificación de biomarcadores para el tratamiento de la enfermedad" y "Perfil genético de cáncer gástrico en Nariño: estudios de caso en línea somática y germinal", me permito informarle lo siguiente:

Una vez revisada la información suministrada, se encuentra que el grupo de investigación "Grupo Interdisciplinario de Investigación en Salud-Enfermedad-GIISE (COL 0045347)" se encuentra registrado en la base ScienTI, y CUMPLE con los lineamientos del Ministerio para el proceso de autorización.

Por lo anterior, teniendo en cuenta el cumplimiento de los requisitos definidos en la Resolución 8430 de 1993 y el artículo 5 de la Resolución 3823 de 1997, se expide AUTORIZACIÓN PARA LA SALIDA DE MUESTRAS BIOLÓGICAS, hacia:

División de secuenciación NGS de Macrogen
División NGS 10F world Meridian Venture Center #60-24
Gasam-dong Geumchun-gu Seoul 153-781 Corea del Sur,
Teléfono +82-2-2113-7352
Jiyoung Kim

Carrera 13 No.32-76 - Código Postal 110311, Bogotá D.C

Teléfono: (57-1)3305000 - Línea gratuita: 018000952525 Fax: (57-1)3305050 - www.minsalud.gov.co



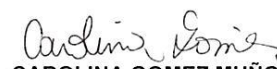
Al contestar por favor cite estos datos:

Radicado No.: **201824000498311**

Fecha: **02-05-2018**

Página 2 de 2

Cordialmente,


CAROLINA GOMEZ MUÑOZ
Directora de Medicamentos y Tecnologías en Salud

Elaboró:  psolano
Revisó/Aprobó:  Jhernandez
Ruta electrónica: \\MINSP\VM81\DirMed\Tec\COMPONENTES ANATOMICOS\TRAMITES\MUESTRAS BIOLOGICAS A\INVESTIGACION\UNIVERSIDAD COOPERATIVA
DE COLOMBIA\2018\527092 Autorizacion salida de muestras.docx
Archivar: Muestras 2018

Anexo 4. Protocolo de extracción de ADN a partir de muestras de sangre usando el kit Ultra Clean Blood spin (Mobio)

Se añadirán 200 μ L de sangre conservada en medio EDTA en un tubo de 2 mL y se adicionarán 10 μ L de proteinasa K. Posteriormente se agregarán 200 μ L de solución B1 y se homogenizará la solución en un equipo vórtex durante 15 seg.

Se incubará la muestra a 65 °C por 10 min y se centrifugará brevemente para recoger el lisado. Luego se adicionarán 200 μ L de solución B2 y se agitará por inversión durante 15 seg. Posterior a la homogenización se centrifugará brevemente para tomar el sobrenadante.

El lisado se añadirá a un spin filter de 2 mL y se centrifugará durante 1 min por 13.000 r.p.m. Luego se transferirá el SPIN FILTER a un tubo de 2 mL. Se agregó 50 μ L de solución B3 al spin filter y se centrifugará por 30 seg a 13.000 r.p.m. Se retirará el spin filter y se descartará el flujo.

Posteriormente se colocará el spin filter de nuevo en el mismo tubo de 2 mL, y se adicionará 500 μ L de solución B4 al SPIN FILTER. Se centrifugará por 30 seg a 13.000 r.p.m.

Se removerá el SPIN FILTER y descartará el flujo; se colocará nuevamente el SPIN FILTER en el mismo tubo de 2 mL, se centrifugará durante 30 seg a 13.000 r.p.m. para secar la membrana del SPIN FILTER.

Con cuidado se removerá el SPIN FILTER y se transferirá a un nuevo tubo de 2 mL, en el cual se adicionará de 100-200 μ L de solución B5 e incubará por 5 min a 65 °C. Finalmente se centrifugará por 1 min a 13.000 r.p.m, se retirará la unidad del SPIN FILTER cerca del tubo, y el DNA se conservará en el tubo para su posterior amplificación.

Anexo 5. Variantes puntuales deletéreas y probablemente causantes de enfermedad de cuatro pacientes con cáncer gástrico del departamento de Nariño.

Varaintes SP03. (Ref: Nucleotido de referencia; Alt: Nucleotido alternativo; D: variante causa de enfermedad, es decir, probablemente perjudicial; A: variante enfermedad que causa automáticamente, es decir, se sabe que es perjudicial)

n	Crom	Posición	Ref	Alt	Gen	Nombre del gen	EXAC	SIFT		Mutation Taster		Genotipo
							Frecuencia	Puntaje	Predicción	Puntaje	Predicción	
1	1	1179822	A	G	<i>C1QTNF12</i>	Adipolina	0.0001	0.001	Deletérea	1.000	D	0/1
2	1	32158384	G	A	<i>COL16A1</i>	Cadena De Colágeno Alfa-1 (Xvi)	0.0005	0.025	Deletérea	0.999	D	0/1
3	1	206821856	G	A	<i>DYRK3</i>	Quinasa 3 Regulada Por Fosforilación De Tirosina De Especificidad Dual	0.0005	0.0	Deletérea	1	D	0/1
4	1	16778393	G	A	<i>NECAP2</i>	Proteína 2 Asociada A La Capa De Unión A La Oreja De Adaptina	0.0007	0.004	Deletérea	0.999	D	0/1
5	1	110038772	T	A	<i>CYB561D1</i>	Miembro De La Familia Del Citocromo B561 D1	0.0007	0.001	Deletérea	0.999	D	0/1
6	1	27874606	A	C	<i>AHDC1</i>	Proteína 1 Que Contiene El Motivo De Unión Al Adn Del Gancho At	0.0007	0.0	Deletérea	1	D	0/1
7	1	16273559	C	T	<i>ZBTB17</i>	Proteína 17 Que Contiene Dedo De Zinc Y Dominio Btb	0.0009	0.001	Deletérea	0.980	D	0/1
8	1	220152884	G	A	<i>EPRS</i>	Glutamato / Prolina Bifuncional - Trna Ligasa	0.0010	0.025	Deletérea	1	D	0/1
9	1	46425062	C	T	<i>MAST2</i>	Serina / Treonina-Proteína Quinasa 2 Asociada A Microtúbulos	0.0010	0.0	Deletérea	1.000	D	0/1
10	1	934735	A	C	<i>HES4</i>	Factor De Transcripción 4 De Bhlh De La Familia Hes	0.0011	0.009	Deletérea	1.000	D	0/1
11	1	167341211	C	T	<i>POU2F1</i>	Dominio Pou, Clase 2, Factor De Transcripción 1	0.0012	0.002	Deletérea	0.975	D	0/1
12	1	26513940	T	C	<i>CNKSR1</i>	Potenciador Del Conector Del Supresor De Quinasa De Ras 1	0.0013	0.023	Deletérea	0.531	D	0/1
13	1	71440025	G	A	<i>PTGER3</i>	Subtipo Ep3 Del Receptor De Prostaglandina E2	0.0014	0.0	Deletérea	0.568	D	0/1
14	1	245849857	C	T	<i>KIF26B</i>	Proteína De Tipo Quinesina Kif26B	0.0014	0.0	Deletérea	1	D	0/1
15	1	62483586	C	T	<i>PATJ</i>	Proteína Similar A Inad	0.0015	0.042	Deletérea	0.949	D	0/1
16	1	46083226	C	T	<i>NASP</i>	Proteína De Esperma Autoantigénica Nuclear	0.0016	0.007	Deletérea	0.946	D	0/1
17	1	38265622	A	G	<i>MANEAL</i>	Proteína Similar A La Glicoproteína Endo-Alfa-1,2-Manosidasa	0.0019	0.004	Deletérea	1.000	D	0/1
18	1	1387776	C	T	<i>ATAD3C</i>	Familia De Atpasa, Dominio Aaa Que Contiene 3C	0.0019	0.006	Deletérea	1.000	D	0/1
19	1	209806061	C	G	<i>LAMB3</i>	Subunidad Beta-3 De Laminina	0.0020	0.024	Deletérea	0.874	D	0/1

n	Crom	Posición	Ref	Alt	Gen	Nombre	Frecuencia (EXAC)	Puntaje	Predicción	Puntaje	Predicción	Genotipo
20	1	46827337	T	C	<i>NSUN4</i>	5-Metilcitosina Rna Metiltransferasa	0.0020	0.005	Deletérea	0.947	D	0/1
21	1	86515808	C	T	<i>COL24A1</i>	Cadena De Colágeno Alfa-1 (Xxiv) Miembro 1B De La Superfamilia	0.0021	0.048	Deletérea	0.993	D	0/1
22	1	12254015	T	C	<i>TNFRSF1B</i>	Del Receptor Del Factor De Necrosis Tumoral	0.0022	0.004	Deletérea	1.000	D	0/1
23	1	215847490	G	T	<i>USH2A</i>	Usherin	0.0029	0.002	Deletérea	0.994	D	0/1
24	1	175306758	G	A	<i>TNR</i>	Tenascina-R	0.0030	0.013	Deletérea	1.000	D	0/1
25	1	229622236	G	A	<i>NUP133</i>	Proteína Del Complejo De Poros Nucleares	0.0034	0.015	Deletérea	1	D	0/1
26	1	97544543	G	T	<i>DPYD</i>	Dihidropirimidina Deshidrogenasa [Nadp (+)]	0.0034	0.004	Deletérea	1.000	D	0/1
27	1	111854859	C	T	<i>CHIA</i>	Quitinasa Ácida De Mamífero Succinil-CoA: 3-Cetoácido	0.0036	0.016	Deletérea	0.973	D	1/1
28	1	40235432	G	A	<i>OXCT2</i>	Coenzima A Transferasa 2, Mitochondrial	0.0036	0.0	Deletérea	0.748	D	0/1
29	1	59978048	C	G	<i>FGGY</i>	Fggy Que Contiene Dominio De Carbohidrato Quinasa	0.0036	0.002	Deletérea	1.000	D	0/1
30	1	54059857	T	C	<i>GLIS1</i>	Proteína De Dedos De Zinc	0.0038	0.008	Deletérea	1	D	0/1
31	1	155583296	T	C	<i>MSTO1</i>	Homólogo 1 De Misato De Proteína	0.0047	0.012	Deletérea	1.000	D	0/1
32	1	43687934	A	C	<i>CFAP57</i>	Proteína Asociada A Cilios Y Flagelos 57	0.0048	0.008	Deletérea	0.584	D	0/1
33	1	219384945	C	G	<i>LYPLAL1</i>	Proteína 1 Similar A La Lisofosfolipasa	0.0050	0.009	Deletérea	0.999	D	0/1
34	1	169555582	C	T	<i>F5</i>	Factor De Coagulación V	0.0058	0.019	Deletérea	1	D	0/1
35	1	237060372	A	G	<i>MTR</i>	Metionina Sintasa	0.0063	0.033	Deletérea	0.546	D	0/1
36	1	146757092	A	C	<i>CHD1L</i>	Similar A Proteína 1 De Unión A Adn De Cromodominio-Helicasa	0.0066	0.049	Deletérea	0.984	D	0/1
37	1	879481	G	C	<i>SAMD11</i>	Similar A Proteína 1 De Unión A Adn De Cromodominio-Helicasa	0.0068	0.002	Deletérea	1.000	D	0/1
38	1	10683166	G	A	<i>PEX14</i>	Proteína De Membrana Peroxisomal Pex14	0.0075	0.048	Deletérea	1.000	D	0/1
39	1	23763938	T	G	<i>ASAP3</i>	Arf-Gap Con Dominio Sh3, Repetición Ank Y Proteína 3 Que Contiene Dominio Ph	0.0092	0.001	Deletérea	1.000	D	1/1
40	1	63010649	G	A	<i>DOCK7</i>	Dedicador De La Proteína 7 De Citocinesis	6.075e-05	0.013	Deletérea	0.945	D	0/1
41	1	204199605	C	T	<i>PLEKHA6</i>	Dominio De Homología De Pleckstrina Que Contiene A6	7.467e-05	0.008	Deletérea	1.000	D	0/1
42	2	233408302	A	T	<i>CHRNA7</i>	Subunidad Gamma Del Receptor De Acetilcolina	0.0001	0.003	Deletérea	0.974	D	0/1

n	Crom	Posición	Ref	Alt	Gen	Nombre	Frecuencia (EXAC)	Puntaje	Predicción	Puntaje	Predicción	Genotipo
43	2	47250579	C	T	<i>TTC7A</i>	Proteína De Repetición Tetratricopeptídica 7ª	0.0004	0.001	Deletérea	0.966	D	0/1
44	2	264985	G	T	<i>ACP1</i>	Fosfotirosina Proteína Fosfatasa De Bajo Peso Molecular	0.0015	0.022	Deletérea	1	D	0/1
45	2	27354687	C	T	<i>PREB</i>	Proteína De Unión Al Elemento Regulador De Prolactina	0.0015	0.017	Deletérea	0.964	D	0/1
46	2	127453543	T	C	<i>GYPC</i>	Glicoforina-C	0.0023	0.002	Deletérea	0.747	D	0/1
47	2	102828620	T	C	<i>IL1RL2</i>	Tipo 2 Del Receptor De Interleucina-1	0.0025	0.004	Deletérea	0.920	D	0/1
48	2	28050468	A	G	<i>RBKS</i>	Ribocinasa	0.0026	0.035	Deletérea	1.000	D	0/1
49	2	202357645	A	C	<i>C2CD6</i>	Proteína 6 Que Contiene El Dominio Dependiente De Calcio C2	0.0034	0.0	Deletérea	1	D	0/1
50	2	120861698	C	A	<i>EPB41L5</i>	Proteína 5 Similar A La Banda 4.1	0.0044	0.001	Deletérea	1	D	0/1
51	2	132021178	G	A	<i>POTEE</i>	Miembro E De La Familia Del Dominio De Anquirina Pote	0.0051	0.001	Deletérea	1	D	1/1
52	2	190611304	A	T	<i>ANKAR</i>	Ankyrin Y Armadillo Repiten Que Contiene	0.0054	0.04	Deletérea	1	D	0/1
53	2	108875270	G	A	<i>SULT1C3</i>	Sulfotransferasa 1C3	0.0054	0.0	Deletérea	0.998	D	0/1
54	2	162088034	C	T	<i>TANK</i>	Activador De Nf-Kappa-B Asociado A Miembros De La Familia Traf	0.0056	0.003	Deletérea	0.994	D	0/1
55	2	70504399	A	G	<i>PCYOX1</i>	Prenilcisteína Oxidasa 1	0.0064	0.021	Deletérea	1	D	0/1
56	2	206480353	C	A	<i>PARD3B</i>	Partición 3 Homólogo B Defectuoso	0.0065	0.002	Deletérea	0.792	D	0/1
57	2	24236236	G	A	<i>MFSD2B</i>	Proteína 2B Que Contiene El Dominio De La Superfamilia Facilitadora Principal	0.0090	0.024	Deletérea	1	D	0/1
58	2	160019901	C	T	<i>TANC1</i>	Proteína Tanc1	0.0097	0.008	Deletérea	1	D	0/1
59	2	241622124	G	A	<i>AQP12B</i>	Acuaporina-12B	0.0098	0.023	Deletérea	0.843	D	0/1
60	2	132021702	A	G	<i>POTEE</i>	Miembro E De La Familia Del Dominio De Anquirina Pote	0.0099	0.005	Deletérea	1	D	0/1
61	2	27466336	G	A	<i>CAD</i>	Proteína Cad	1.083e-05	0.001	Deletérea	1	D	0/1
62	2	112944754	C	T	<i>FBLN7</i>	Fibulina-7	5.053e-05	0.035	Deletérea	0.999	D	0/1
63	2	95541447	C	T	<i>TEKT4</i>	Tektin-4	8.261e-06	0.0	Deletérea	1	D	0/1
64	2	114257215	C	T	<i>FOXD4L1</i>	Caja De Horquilla D4 Como 1	9.352e-05	0.0	Deletérea	1	D	0/1
65	3	145809622	A	G	<i>PLOD2</i>	Procolágeno-Lisina, 2-Oxoglutarato 5-Dioxigenasa 2	0.0002	0.0	Deletérea	1	D	0/1
66	3	47038837	C	T	<i>NBEAL2</i>	Proteína 2 Similar A La Neurobeachina	0.0002	0.036	Deletérea	1.000	D	0/1
67	3	184100205	G	T	<i>CHRD</i>	Chordin	0.0002	0.024	Deletérea	0.999	D	0/1
68	3	51812788	G	A	<i>IQCF6</i>	Motivo Iq Que Contiene F6	0.0004	0.012	Deletérea	0.984	D	0/1

n	Crom	Posición	Ref	Alt	Gen	Nombre	Frecuencia (EXAC)	Puntaje	Predicción	Puntaje	Predicción	Genotipo
69	3	160156708	T	G	TRIM59	Proteína 59 Que Contiene Un Motivo Tripartito	0.0011	0.022	Deletérea	0.973	D	0/1
70	3	195935346	C	T	ZDHHC19	Probable Palmitoiltransferasa Zdhhc19	0.0011	0.021	Deletérea	0.861	D	0/1
71	3	195513740	C	G	MUC4	Mucina-4	0.0014	0.049	Deletérea	1.000	D	0/1
72	3	16313222	G	A	OXNAD1	Proteína 1 Que Contiene El Dominio De Unión A Nad De Oxidorreductasa	0.0014	0.0	Deletérea	1.000	D	0/1
73	3	50331107	C	T	HYAL3	Hialuronidasa-3	0.0017	0.008	Deletérea	0.970	D	0/1
74	3	128181787	C	T	DNAJB8	Miembro 8 De La Subfamilia B Del Homólogo De Dnaj	0.0023	0.024	Deletérea	0.998	D	0/1
75	3	42705960	C	T	ZBTB47	Proteína 47 Que Contiene Dedo De Zinc Y Dominio Btb	0.0026	0.013	Deletérea	0.726	D	0/1
76	3	74350869	T	C	CNTN3	Contactin-3	0.0026	0.011	Deletérea	0.997	D	0/1
77	3	111831909	G	C	C3orf52	Proteína Transmembrana Inducida Por Tpa	0.0029	0.014	Deletérea	1	D	0/1
78	3	49723867	G	T	MST1	Proteína Similar Al Factor De Crecimiento De Hepatocitos	0.0030	0.004	Deletérea	0.983	D	0/1
79	3	69171286	C	T	LMOD3	Leiomodin-3	0.0032	0.027	Deletérea	1	D	0/1
80	3	121527826	A	G	IQCB1	Proteína 1 Que Contiene El Motivo De Unión A Calmodulina Iq	0.0032	0.012	Deletérea	1.000	D	0/1
81	3	123695759	A	C	ROPN1	Ropporin-1A	0.0038	0.004	Deletérea	0.982	D	0/1
82	3	51980236	G	A	PARP3	Poli [Adp-Ribosa] Polimerasa 3	0.0046	0.0	Deletérea	1	D	0/1
83	3	142141436	G	A	XRN1	Exoribonucleasa 1 5'-3'	0.0050	0.001	Deletérea	1	D	0/1
84	3	39227755	G	A	XIRP1	Proteína 1 Que Contiene Repetición De Unión A Actina Xin	0.0056	0.004	Deletérea	1	D	0/1
85	3	37144415	G	T	LRRFIP2	Proteína 2 Que Interactúa Sin Vuelo Repetida Rica En Leucina	0.0058	0.003	Deletérea	1	D	0/1
86	3	52422300	C	T	DNAH1	Cadena Pesada 1 De Dineína, Axonemal	0.0072	0.001	Deletérea	1	D	0/1
87	3	108572533	C	A	TRAT1	Adaptador Transmembrana Asociado Al Receptor De Células T 1	0.0072	0.045	Deletérea	0.870	D	0/1
88	3	58110191	C	T	FLNB	Filamin-B	0.0091	0.012	Deletérea	1.000	D	0/1
89	3	38674699	G	A	SCN5A	Proteína Del Canal De Sodio Tipo 5 Subunidad Alfa	0.0095	0.0	Deletérea	0.999	D	0/1
90	3	69168526	A	G	LMOD3	Leiomodin-3	1.656e-05	0.003	Deletérea	1	D	0/1
91	3	49759678	C	G	GMPPB	Manosa-1-Fosfato Guaniltransferasa Beta	8.294e-06	0.001	Deletérea	1	D	0/1
92	4	128722949	C	T	HSPA4L	Proteína 4L De 70 Kda De Choque Térmico	0.0007	0.025	Deletérea	1	D	0/1

n	Crom	Posición	Ref	Alt	Gen	Nombre	Frecuencia (EXAC)	Puntaje	Predicción	Puntaje	Predicción	Genotipo
93	4	128861116	C	T	<i>MFSD8</i>	Proteína 8 Que Contiene El Dominio De La Superfamilia Facilitadora Principal	0.0008	0.004	Deletérea	1	D	0/1
94	4	169100461	T	C	<i>ANXA10</i>	Anexina A10	0.0010	0.011	Deletérea	1.000	D	0/1
95	4	100240016	A	G	<i>ADH1B</i>	Alcohol Deshidrogenasa 1B, Polipéptido Beta	0.0015	0.0	Deletérea	1	D	0/1
96	4	113359752	G	A	<i>ALPK1</i>	Alfa-Proteína Quinasa 1	0.0015	0.015	Deletérea	1.000	D	0/1
97	4	2954057	A	T	<i>NOP14</i>	Proteína Nucleolar 14	0.0030	0.002	Deletérea	1.000	D	0/1
98	4	113199517	G	A	<i>TIFA</i>	Proteína Que Interactúa Con Traf Con Proteína A Que Contiene El Dominio Fha	0.0032	0.011	Deletérea	0.999	D	0/1
99	4	96762427	C	G	<i>PDHA2</i>	Subunidad Alfa Del Componente Piruvato Deshidrogenasa E1, Forma Específica De Testículo, Mitocondrial	0.0039	0.001	Deletérea	1.000	D	0/1
100	4	2514846	C	T	<i>RNF4</i>	E3 Ubiquitina-Proteína Ligasa Rnf4	0.0041	0.0	Deletérea	1	D	0/1
101	4	436170	G	T	<i>ZNF721</i>	Proteína De Dedo De Zinc 721	0.0048	0.019	Deletérea	0.693	D	0/1
102	4	25849329	T	C	<i>SEL1L3</i>	Homólogo 3 De La Proteína Sel-1	0.0055	0.018	Deletérea	1	D	0/1
103	4	56475330	T	C	<i>NMU</i>	Neuromedina-U	0.0065	0.01	Deletérea	0.678	D	0/1
104	4	151207098	C	G	<i>LRBA</i>	Proteína De Anclaje De Tipo Beige Y Sensible A Lipopolisacáridos	0.0066	0.01	Deletérea	0.967	D	0/1
105	4	129792810	G	A	<i>JADE1</i>	Proteína Jade-1	0.0087	0.046	Deletérea	0.999	D	0/1
106	4	186435443	C	T	<i>PDLIM3</i>	Proteína 3 Del Dominio Pdz Y Lim	0.0091	0.032	Deletérea	1	D	0/1
107	4	40356155	A	G	<i>CHRNA9</i>	Subunidad Alfa-9 Del Receptor De Acetilcolina Neuronal	2.471e-05	0.021	Deletérea	1	D	0/1
108	4	121774637	G	C	<i>PRDM5</i>	Proteína 5 Del Dedo De Zinc Del Dominio Pr	9.062e-05	0.02	Deletérea	1.000	D	0/1
109	5	140589681	A	G	<i>PCDHB12</i>	Protocadherina Beta-12	0.0001	0.0	Deletérea	0.986	D	0/1
110	5	121799193	G	C	<i>SNCAIP</i>	Sinfilina-1	0.0005	0.002	Deletérea	0.647	D	0/1
111	5	89970013	C	T	<i>ADGRV1</i>	Receptor 98 Acoplado A Proteína G	0.0005	0.017	Deletérea	0.735	D	0/1
112	5	49695734	T	C	<i>EMB</i>	Embigin	0.0006	0.006	Deletérea	0.611	D	0/1
113	5	32089103	G	A	<i>PDZD2</i>	Dominio Pdz Que Contiene 2	0.0006	0.0	Deletérea	0.813	D	0/1
114	5	118454596	T	C	<i>DMXL1</i>	Proteína 1 Similar A Dmx	0.0006	0.036	Deletérea	0.885	D	0/1
115	5	151051255	T	C	<i>SPARC</i>	Sparc	0.0007	0.019	Deletérea	1	D	0/1
116	5	169018056	A	G	<i>SPDL1</i>	Proteína Delgadas	0.0008	0.009	Deletérea	0.593	D	0/1
117	5	68616225	G	C	<i>CCDC125</i>	Proteína 125 Que Contiene El Dominio En Espiral	0.0008	0.039	Deletérea	1	D	0/1
118	5	44813282	A	G	<i>MRPS30</i>	Proteína Ribosómica Mitocondrial S30	0.0010	0.007	Deletérea	0.999	D	0/1

n	Crom	Posición	Ref	Alt	Gen	Nombre	Frecuencia (EXAC)	Puntaje	Predicción	Puntaje	Predicción	Genotipo
119	5	41842798	C	A	<i>OXCT1</i>	Succinil-CoA: 3-Cetoácido Coenzima A Transferasa 1, Mitochondrial	0.0010	0.024	Deletérea	0.821	D	0/1
120	5	10286468	T	C	<i>CMBL</i>	Homólogo De Carboximetilnbutenolidasa	0.0012	0.002	Deletérea	1	D	0/1
121	5	112043492	C	A	<i>APC</i>	Proteína De Poliposis Coli Adenomatosa	0.0015	0.0	Deletérea	1	D	0/1
122	5	131915574	C	T	<i>RAD50</i>	Proteína De Reparación Del Adn Rad50	0.0018	0.049	Deletérea	1	D	0/1
123	5	147692077	T	C	<i>SPINK7</i>	Inhibidor De La Serina Proteasa Kazal-Tipo 7	0.0025	0.036	Deletérea	0.997	D	0/1
124	5	89979485	C	T	<i>ADGRV1</i>	Receptor 98 Acoplado A Proteína G	0.0035	0.015	Deletérea	0.991	D	0/1
125	5	122725768	G	A	<i>CEP120</i>	Proteína Centrosomal De 120 Kda	0.0046	0.046	Deletérea	0.908	D	0/1
126	5	78375222	C	T	<i>BHMT2</i>	S-Metilmetionina - Homocisteína S- Metiltransferasa Bhmt2	0.0046	0.016	Deletérea	1.000	D	0/1
127	5	150666907	A	G	<i>SLC36A3</i>	Transportador 3 De Aminoácidos Acoplado A Protones	0.0047	0.0	Deletérea	1.000	D	0/1
128	5	34850389	G	A	<i>TTC23L</i>	Dominio De Repetición Tetratricoptide 23	0.0084	0.0	Deletérea	0.883	D	0/1
129	5	110097362	G	T	<i>SLC25A46</i>	Familia De Portadores De Solutos 25 Miembro 46	0.0098	0.029	Deletérea	1.000	D	0/1
130	5	150444571	A	C	<i>TNIP1</i>	Proteína 1 Que Interactúa Con Tnfaip3	1.648e-05	0.004	Deletérea	1.000	D	0/1
131	5	102260725	C	T	<i>PAM</i>	Monooxigenasa Alfa-Amidante De Peptidil-Glicina	1.65e-05	0.009	Deletérea	1	D	0/1
132	6	32548032	T	G	<i>HLA-DRB1</i>	Antígeno De Histocompatibilidad Hla De Clase Ii	0.0001	0.022	Deletérea	0.942	D	0/1
133	6	49753706	T	G	<i>PGK2</i>	Fosfoglicerato Quinasa 2	0.0002	0.0	Deletérea	1.000	D	0/1
134	6	99956505	C	T	<i>USP45</i>	Peptidasa 45 Específica De Ubiquitina	0.0004	0.0	Deletérea	1	D	0/1
135	6	159660951	T	A	<i>FNDC1</i>	Proteína 1 Que Contiene El Dominio De Fibronectina Tipo Iii	0.0019	0.0	Deletérea	0.659	D	0/1
136	6	42075126	G	A	<i>C6orf132</i>	Proteína No Caracterizada C6Orf132	0.0020	0.0	Deletérea	0.774	D	0/1
137	6	136560617	G	A	<i>MTFR2</i>	Regulador De Fisión Mitochondrial 2	0.0023	0.0	Deletérea	1.000	D	0/1
138	6	44269171	G	A	<i>AARS2</i>	Alanina - Trna Ligasa, Mitochondrial	0.0023	0.009	Deletérea	0.865	D	0/1
139	6	54025545	C	G	<i>MLIP</i>	Proteína Que Interactúa Con Lmna Muscular	0.0025	0.04	Deletérea	0.875	D	0/1
140	6	56966676	G	A	<i>ZNF451</i>	E3 Sumo-Proteína Ligasa Znf451	0.0025	0.004	Deletérea	1	D	0/1
141	6	26158456	C	T	<i>HIST1H2BD</i>	Histona H2B Tipo 1-D	0.0033	0.019	Deletérea	0.868	D	0/1
142	6	90471392	T	G	<i>MDN1</i>	Midasina	0.0046	0.011	Deletérea	1.000	D	0/1

n	Crom	Posición	Ref	Alt	Gen	Nombre	Frecuencia (EXAC)	Puntaje	Predicción	Puntaje	Predicción	Genotipo
143	6	13321321	A	C	<i>TBC1D7</i>	Miembro 7 De La Familia Del Dominio Tbc1	0.0049	0.001	Deletérea	1.000	D	0/1
144	6	26393058	C	T	<i>BTN2A2</i>	Miembro A2 De La Subfamilia 2 De Butirofilina	0.0083	0.01	Deletérea	1	D	0/1
145	6	116953552	C	T	<i>RSPH4A</i>	Proteína 4 Homóloga A De La Cabeza De Radio Radial	0.0093	0.024	Deletérea	0.884	D	0/1
146	6	32632632	C	G	<i>HLA-DQB1</i>	Antígeno De Histocompatibilidad Hla De Clase Ii, Cadena Dq Beta 1	2.521e-05	0.0	Deletérea	0.981	D	0/1
147	6	43608008	G	A	<i>MAD2L1BP</i>	Proteína De Unión A Mad2L1	6.689e-05	0.035	Deletérea	1.000	D	0/1
148	6	157405929	C	T	<i>ARID1B</i>	Proteína 1B Que Contiene El Dominio Interactivo Rico En At	8.238e-06	0.006	Deletérea	1.000	D	0/1
149	7	107830092	T	C	<i>NRCAM</i>	Molécula De Adhesión Celular Neuronal	0.0002	0.004	Deletérea	0.989	D	0/1
150	7	127992585	G	A	<i>PRRT4</i>	Proteína Transmembrana 4 Rica En Prolina	0.0002	0.01	Deletérea	0.993	D	0/1
151	7	22985710	G	A	<i>FAM126A</i>	Hyccin	0.0005	0.012	Deletérea	1	D	0/1
152	7	134135581	T	A	<i>AKR1B1</i>	Aldosa Reductasa	0.0006	0.0	Deletérea	1	D	0/1
153	7	74193642	G	A	<i>NCF1</i>	Factor 1 Del Citosol De Neutrófilos	0.0010	0.029	Deletérea	1	D	0/1
154	7	116340223	T	C	<i>MET</i>	Receptor Del Factor De Crecimiento De Hepatocitos	0.0020	0.002	Deletérea	0.977	D	0/1
155	7	150905167	G	A	<i>ABCF2</i>	Miembro 2 De La Subfamilia F Del Casete De Unión De Atp	0.0026	0.0	Deletérea	0.998	D	0/1
156	7	76029798	C	T	<i>SSC4D</i>	Proteína Del Grupo B Que Contiene El Dominio Rico En Cisteína Del Receptor Scavenger	0.0027	0.0	Deletérea	1.000	D	0/1
157	7	107323951	A	T	<i>SLC26A4</i>	Pendrina	0.0031	0.005	Deletérea	0.977	D	0/1
158	7	117385978	G	T	<i>CTTNBP2</i>	Proteína 2 De Unión A Cortactina	0.0033	0.004	Deletérea	0.996	D	0/1
159	7	44116207	C	A	<i>POLM</i>	Polimerasa Mu De Adn / Arn Dirigida Por Adn	0.0043	0.002	Deletérea	1.000	D	0/1
160	7	76796995	A	T	<i>CCDC146</i>	Dominio De Bobina Enrollada Que Contiene 146	0.0065	0.007	Deletérea	1	D	0/1
161	7	151800228	C	T	<i>GALNT11</i>	Polipéptido N-Acetilgalactosaminiltransferasa 11	0.0087	0.038	Deletérea	1	D	0/1
162	7	99445207	A	G	<i>CYP3A43</i>	Citocromo P450 3A43	2.475e-05	0.007	Deletérea	0.997	D	0/1
163	7	107875036	G	A	<i>NRCAM</i>	Molécula De Adhesión Celular Neuronal	3.307e-05	0.0	Deletérea	1	D	0/1
164	8	86048151	G	C	<i>LRRCC1</i>	Repetición Rica En Leucina Y Proteína 1 Que Contiene El Dominio De Espiral Enrollada	0.0002	0.0	Deletérea	1	D	0/1

n	Crom	Posición	Ref	Alt	Gen	Nombre	Frecuencia (EXAC)	Puntaje	Predicción	Puntaje	Predicción	Genotipo
165	8	41561956	T	C	ANK1	Ankyrin-1; Adjunta Proteínas De Membrana Integrales A Elementos Citoesqueléticos	0.0003	0.0	Deletérea	1	D	0/1
166	8	88885262	T	C	DCAF4L2	Proteína 2 Similar Al Factor 4 Asociada A Ddb1 Y Cul4	0.0016	0.002	Deletérea	0.750	D	0/1
167	8	140744338	C	T	TRAPPC9	Subunidad 9 Del Complejo De Partículas De Proteína De Tráfico	0.0023	0.032	Deletérea	1.000	D	0/1
168	8	37822931	G	A	ADRB3	Receptor Adrenérgico Beta-3	0.0045	0.0	Deletérea	1	D	0/1
169	8	101717220	C	T	PABPC1	Proteína De Unión A Poliadenilato 1	0.0047	0.001	Deletérea	1	D	0/1
170	8	27680615	T	A	PBK	Proteína Quinasa Originada Por Linfocitos T Asesinos Activados Por Linfocinas	0.0054	0.022	Deletérea	1.000	D	0/1
171	8	32621631	G	A	NRG1	Pro-Neuregulin-1	0.0076	0.023	Deletérea	1.000	D	0/1
172	8	67786585	A	G	MCMDC2	Proteína 2 Que Contiene El Dominio De Mantenimiento Del Minicromosoma	0.0086	0.001	Deletérea	1.000	D	0/1
173	8	41522407	G	C	ANK1	Ankyrin-1	0.0093	0.006	Deletérea	1	D	0/1
174	8	106815578	G	A	ZFPM2	Proteína De Dedo De Zinc Zfp2	1.703e-05	0.02	Deletérea	1.000	D	0/1
175	9	131730791	C	A	NUP188	Homólogo De Nucleoporina Nup188	0.0003	0.034	Deletérea	1.000	D	0/1
176	9	91991853	C	A	SEMA4D	Semaforina-4D	0.0007	0.032	Deletérea	1	D	0/1
177	9	20944696	G	A	FOCAD	Focadhesin	0.0012	0.035	Deletérea	0.999	D	0/1
178	9	33943505	G	A	UBAP2	Proteína 2 Asociada A Ubiquitina	0.0019	0.021	Deletérea	1.000	D	0/1
179	9	95896475	C	G	NINJ1	Ninjurin-1	0.0021	0.014	Deletérea	0.977	D	0/1
180	9	100109681	A	C	CCDC180	Dominio De Bobina Enrollada Que Contiene 180	0.0037	0.016	Deletérea	0.714	D	0/1
181	9	33941702	C	T	UBAP2	Proteína 2 Asociada A Ubiquitina	0.0060	0.005	Deletérea	1.000	D	0/1
182	9	136030688	A	G	GBGT1	Globósido Alfa-1,3-N-Acetilgalactosaminiltransferasa 1	0.0078	0.0	Deletérea	1.000	D	0/1
183	9	100085147	C	T	CCDC180	Dominio De Bobina Enrollada Que Contiene 180	5.769e-05	0.01	Deletérea	0.943	D	0/1
184	9	72365700	C	A	PTAR1	Repetición De La Subunidad Alfa De La Proteína Preniltransferasa Que Contiene 1	9.977e-06	0.027	Deletérea	1	D	0/1
185	10	114920400	T	C	TCF7L2	Factor De Transcripción Tipo 7 2	0.0002	0.035	Deletérea	0.925	D	0/1
186	10	3200257	C	A	PITRM1	Proteasa De Presencia De Secuencia, Mitocondrial	0.0004	0.002	Deletérea	0.764	D	0/1
187	10	90438283	C	A	LIPF	Lipasa F, Tipo Gástrico	0.0005	0.043	Deletérea	0.982	D	0/1
188	10	18940092	T	C	NSUN6	Metiltransferasa Putativa Nsun6	0.0005	0.047	Deletérea	1.000	D	0/1
189	10	126714948	C	T	CTBP2	Proteína 2 De Unión Al Terminal C	0.0011	0.007	Deletérea	0.999	D	0/1
190	10	102740026	C	T	SEMA4G	Semaforina-4G	0.0012	0.031	Deletérea	1	D	0/1

n	Crom	Posición	Ref	Alt	Gen	Nombre	Frecuencia (EXAC)	Puntaje	Predicción	Puntaje	Predicción	Genotipo
191	10	96985057	A	C	ACSM6	Acil-Coenzima A Sintetasa Acsm6 , Mitochondrial	0.0016	0.021	Deletérea	1	D	0/1
192	10	135093287	C	T	TUBGCP2	Componente 2 Del Complejo Gamma-Tubulina	0.0022	0.0	Deletérea	0.687	D	0/1
193	10	102269182	A	C	SEC31B	Proteína De Transporte De Proteínas Sec31B	0.0024	0.001	Deletérea	1	D	0/1
194	10	102824352	G	A	KAZALD1	Proteína 1 Que Contiene El Dominio Inhibidor De Serina Proteasa De Tipo Kazal	0.0028	0.013	Deletérea	1.000	D	0/1
195	10	105362790	G	A	SH3PXD2A	Proteína 2A Que Contiene El Dominio Sh3 Y Px	0.0032	0.0	Deletérea	1.000	D	0/1
196	10	135352356	A	T	CYP2E1	Citocromo P450 2E1	0.0043	0.001	Deletérea	0.952	D	0/1
197	10	93425	A	G	TUBB8	Cadena De Tubulina Beta-8;	0.0046	0.047	Deletérea	1.000	D	0/1
198	10	18254599	C	T	SLC39A12	Transportador De Zinc Zip12	0.0047	0.004	Deletérea	1.000	D	0/1
199	10	31750140	A	C	ZEB1	Homeobox 1 De Encuadernación E-Box Con Dedos De Zinc	0.0051	0.05	Deletérea	0.000	A	0/1
200	10	96708974	A	G	CYP2C9	Citocromo P450 2C9	0.0067	0.001	Deletérea	0.908	D	0/1
201	10	134912226	G	A	ADGRA1	Receptor A1 Acoplado A Proteína De Adhesión G	0.0067	0.0	Deletérea	1.000	D	1/1
202	10	88478529	G	A	LDB3	Proteína 3 De Unión Al Dominio Lim	0.0070	0.025	Deletérea	1.000	D	0/1
203	10	135038335	G	A	KNDC1	Proteína Muy Amable	5.801e-05	0.002	Deletérea	1.000	D	0/1
204	10	93416	G	A	TUBB8	Cadena De Tubulina Beta-8	6.018e-05	0.005	Deletérea	1.000	D	0/1
205	11	2997254	G	A	NAP1L4	Tipo 4 De Proteína 1 De Ensamblaje De Nucleosomas	0.0003	0.004	Deletérea	1.000	D	0/1
206	11	47355553	G	A	MYBPC3	Proteína C De Unión A Miosina, De Tipo Cardíaco	0.0004	0.021	Deletérea	0.960	D	0/1
207	11	5020948	C	T	OR51L1	Receptor Olfatorio 51L1	0.0006	0.0	Deletérea	0.968	D	0/1
208	11	794509	C	T	SLC25A22	Portador 1 De Glutamato Mitochondrial	0.0007	0.019	Deletérea	1	D	0/1
209	11	281073	C	T	NLRP6	Proteína 6 Que Contiene Dominios Nacht, Lrr Y Pyd	0.0007	0.017	Deletérea	0.984	D	0/1
210	11	114393757	C	T	NXPE1	Miembro 1 De La Familia Nxpe	0.0010	0.0	Deletérea	0.992	D	0/1
211	11	71169583	G	A	NADSYN1	Nad (+) Sintetasa Dependiente De Glutamina	0.0013	0.001	Deletérea	1	D	0/1
212	11	17190678	G	A	PIK3C2A	Subunidad Alfa Que Contiene El Dominio C2 De Fosfatidilinositol 4-Fosfato 3-Quinasa	0.0016	0.034	Deletérea	0.995	D	0/1
213	11	59225172	G	A	OR4D6	Receptor Olfatorio 4D6	0.0020	0.004	Deletérea	0.831	D	0/1
214	11	47189728	C	T	ARFGAP2	Proteína 2 Que Activa La Gtpasa Del Factor De Ribosilación De Adp	0.0020	0.009	Deletérea	1	D	0/1

n	Crom	Posición	Ref	Alt	Gen	Nombre	Frecuencia (EXAC)	Puntaje	Predicción	Puntaje	Predicción	Genotipo
215	11	34981990	G	A	<i>PDHX</i>	Componente De La Proteína X Piruvato Deshidrogenasa, Mitochondrial	0.0020	0.003	Deletérea	1	D	0/1
216	11	114393091	C	A	<i>NXPE1</i>	Miembro 1 De La Familia Nxpe	0.0022	0.001	Deletérea	0.581	D	0/1
217	11	59562950	C	T	<i>STX3</i>	Sintaxina -3	0.0027	0.0	Deletérea	1	D	0/1
218	11	78369684	C	A	<i>TENM4</i>	Teneurina-4	0.0039	0.0	Deletérea	1	D	0/1
219	11	549957	C	G	<i>LRRC56</i>	Proteína 56 Que Contiene Repeticiones Ricas En Leucina	0.0041	0.026	Deletérea	0.993	D	0/1
220	11	116633548	T	C	<i>BUD13</i>	Homólogo De Bud13	0.0043	0.022	Deletérea	0.998	D	0/1
221	11	61724453	C	A	<i>BEST1</i>	Bestrofina-1	0.0045	0.023	Deletérea	1.000	D	0/1
222	11	36597492	G	A	<i>RAG1</i>	Proteína 1 Activadora De Recombinación V (D) J	0.0045	0.001	Deletérea	1.000	D	0/1
223	11	5411028	C	T	<i>OR51M1</i>	Receptor Olfatorio 51M1	0.0051	0.008	Deletérea	1	D	0/1
224	11	64677270	C	T	<i>ATG2A</i>	Homólogo A De La Proteína 2 Relacionada Con La Autofagia	0.0053	0.009	Deletérea	0.998	D	0/1
225	11	61511036	G	T	<i>DAGLA</i>	Diacilglicerol Lipasa Alfa Específica De Sn1	0.0073	0.003	Deletérea	1	D	1/1
226	11	17656445	G	A	<i>OTOG</i>	Otogelin	0.0077	0.011	Deletérea	1	D	0/1
227	11	130784708	G	A	<i>SNX19</i>	Clasificación Nexin-19	8.24e-06	0.0	Deletérea	1	D	0/1
228	11	76892492	C	T	<i>MYO7A</i>	Miosina Viia No Convencional	8.697e-05	0.009	Deletérea	0.999	D	0/1
229	12	13102500	G	C	<i>GPRC5D</i>	Receptor Acoplado A Proteína G Clase C Grupo 5 Miembro D	0.0001	0.046	Deletérea	0.999	D	0/1
230	12	133502050	G	A	<i>ZNF605</i>	Proteína De Dedos De Zinc 605	0.0001	0.01	Deletérea	0.987	D	0/1
231	12	119588910	T	G	<i>SRRM4</i>	Proteína 4 De Matriz Repetitiva De Serina / Arginina	0.0002	0.0	Deletérea	0.694	D	0/1
232	12	100603920	G	T	<i>ACTR6</i>	Homólogo De Proteína 6 Relacionada Con Actina Arp6	0.0003	0.027	Deletérea	1.000	D	0/1
233	12	121175696	T	C	<i>ACADS</i>	Acil-CoA Deshidrogenasa Específica De Cadena Corta, Mitochondrial	0.0006	0.0	Deletérea	1	D	0/1
234	12	52567488	G	A	<i>KRT80</i>	Queratina, Tipo Ii Citoesquelético 80	0.0008	0.001	Deletérea	0.996	D	0/1
235	12	80839309	G	A	<i>PTPRQ</i>	Fosfatidilinositol Fosfatasa Ptpqr Desmetilasa 5A Específica De Lisina; Histona Desmetilasa Que Desmetila Específicamente 'Lys-4' De La Histona H3,	0.0010	0.0	Deletérea	1.000	D	0/1
236	12	404858	C	G	<i>KDM5A</i>	Proteína Fam234B	0.0020	0.005	Deletérea	1.000	D	0/1
237	12	13208728	C	A	<i>FAM234B</i>	Proteína Fam234B	0.0021	0.037	Deletérea	0.928	D	0/1
238	12	52795099	C	G	<i>KRT82</i>	Queratina, Hb2 Cuticular De Tipo Ii	0.0022	0.0	Deletérea	0.991	D	0/1
239	12	52962038	C	T	<i>KRT74</i>	Queratina, Tipo Ii Citoesquelético 74	0.0045	0.001	Deletérea	1.000	D	0/1
240	12	50368200	C	G	<i>AQP6</i>	Acuaporina-6;	0.0047	0.0	Deletérea	0.997	D	0/1

n	Crom	Posición	Ref	Alt	Gen	Nombre	Frecuencia (EXAC)	Puntaje	Predicción	Puntaje	Predicción	Genotipo
241	12	133378465	G	A	<i>GOLGA3</i>	Miembro 3 De La Subfamilia A De Golgin	0.0062	0.028	Deletérea	0.995	D	0/1
242	12	7477100	G	A	<i>ACSM4</i>	Acil-Coenzima A Sintetasa Acsm4 , Mitochondrial	0.0093	0.001	Deletérea	0.976	D	0/1
243	12	80615951	T	C	<i>OTOGL</i>	Proteína Similar A La Otogelina	3.385e-05	0.0	Deletérea	0.997	D	0/1
244	12	123340155	G	A	<i>HIP1R</i>	Proteína Relacionada Con La Proteína 1 Que Interactúa Con La Huntingtina	4.776e-05	0.003	Deletérea	1.000	D	0/1
245	12	54577616	G	A	<i>SMUG1</i>	Adn Glicosilasa De Uracilo Monofuncional Selectivo Monocatenario	8.945e-06	0.003	Deletérea	0.597	D	0/1
246	13	33591385	C	G	<i>KL</i>	Klotho	0.0004	0.001	Deletérea	1	D	0/1
247	13	75936334	T	G	<i>TBC1D4</i>	Miembro 4 De La Familia Del Dominio Tbc1	0.0008	0.01	Deletérea	1	D	0/1
248	13	28014173	T	C	<i>MTIF3</i>	Factor De Iniciación De La Traducción If-3, Mitochondrial	0.0081	0.009	Deletérea	0.606	D	0/1
249	13	25670797	C	G	<i>PABPC3</i>	Proteína 3 De Unión A Poliadenilato	1.647e-05	0.002	Deletérea	0.999	D	0/1
250	14	23742047	C	T	<i>RNF212B</i>	Proteína 212B Del Dedo Anular	0.0001	0.003	Deletérea	1	D	0/1
251	14	62242879	C	T	<i>SNAPC1</i>	Subunidad 1 Del Complejo Proteico Activador Del Snrna	0.0002	0.039	Deletérea	1.000	D	0/1
252	14	59112382	G	C	<i>DACT1</i>	Homólogo 1 De Dapper	0.0007	0.027	Deletérea	1	D	0/1
253	14	70634568	T	C	<i>SLC8A3</i>	Intercambiador De Sodio / Calcio 3	0.0007	0.007	Deletérea	1	D	0/1
254	14	47120402	C	T	<i>RPL10L</i>	Proteína Ribosómica 60S Similar A L10	0.0017	0.026	Deletérea	1.000	D	0/1
255	14	90650548	A	G	<i>KCNK13</i>	Miembro 13 De La Subfamilia K Del Canal De Potasio	0.0035	0.004	Deletérea	1	D	0/1
256	14	50101724	G	T	<i>DNAAF2</i>	Proteína Kintoun	0.0042	0.0	Deletérea	1.000	D	0/1
257	14	20389039	A	G	<i>OR4K5</i>	Receptor Olfatorio 4K5	0.0088	0.005	Deletérea	0.595	D	0/1
258	14	57749913	G	A	<i>AP5M1</i>	Subunidad Del Complejo Ap-5 Mu-1	3.302e-05	0.038	Deletérea	1.000	D	0/1
259	14	57085372	G	A	<i>TMEM260</i>	Proteína Transmembrana 260	8.238e-06	0.021	Deletérea	0.729	D	0/1
260	15	70969452	T	C	<i>UACA</i>	Autoantígeno Uveal Con Dominios En Espiral Y Repeticiones De Anquirina	0.0002	0.036	Deletérea	1.000	D	0/1
261	15	70980114	C	T	<i>UACA</i>	Autoantígeno Uveal Con Dominios En Espiral Y Repeticiones De Anquirina	0.0002	0.008	Deletérea	1	D	0/1
262	15	68119245	C	G	<i>SKOR1</i>	Correpresor Transcripcional 1 De La Familia Ski	0.0010	0.012	Deletérea	0.528	D	0/1
263	15	86077097	A	G	<i>AKAP13</i>	Proteína De Anclaje De La A-Quinasa 13	0.0010	0.034	Deletérea	0.940	D	0/1

n	Crom	Posición	Ref	Alt	Gen	Nombre	Frecuencia (EXAC)	Puntaje	Predicción	Puntaje	Predicción	Genotipo
264	15	77325272	C	T	<i>PSTPIP1</i>	Proteína 1 Que Interactúa Con Prolina-Serina-Treonina Fosfatasa	0.0012	0.036	Deletérea	1.000	D	0/1
265	15	55838504	G	A	<i>PYGO1</i>	Homólogo 1 De Pygopus	0.0013	0.038	Deletérea	0.544	D	0/1
266	15	74481600	C	T	<i>STRA6</i>	Receptor Para La Captación De Retinol Stra6	0.0017	0.004	Deletérea	0.684	D	0/1
267	15	45365618	G	T	<i>SORD</i>	Sorbitol Deshidrogenasa	0.0066	0.001	Deletérea	1	D	0/1
268	15	101827876	G	A	<i>SNRPA1</i>	Ribonucleoproteína A Nuclear Pequeña U2	2.485e-05	0.01	Deletérea	0.998	D	0/1
269	16	4659684	G	C	<i>UBALD1</i>	Proteína 1 Que Contiene Un Dominio Similar A Uba	0.0006	0.04	Deletérea	0.975	D	0/1
270	16	240000	C	T	<i>LUC7L</i>	Proteína De Unión A Arn Putativa Tipo Luc7 1	0.0006	0.002	Deletérea	1.000	D	0/1
271	16	71482227	T	G	<i>ZNF23</i>	Proteína 23 De Dedos De Zinc; Puede Estar Involucrado En La Regulación Transcripcional.	0.0007	0.001	Deletérea	1	D	0/1
272	16	736848	G	A	<i>WDR24</i>	Proteína Del Complejo Gator Wdr24	0.0008	0.048	Deletérea	1	D	0/1
273	16	1557367	G	A	<i>TELO2</i>	Homólogo Tel2 De La Proteína Reguladora De La Longitud De Los Telómeros	0.0008	0.005	Deletérea	1	D	0/1
274	16	82033344	G	A	<i>SDR42E1</i>	Familia 42E De Deshidrogenasa / Reductasa De Cadena Corta, Miembro 1	0.0012	0.002	Deletérea	1.000	D	0/1
275	16	15802687	G	T	<i>MYH11</i>	Miosina-11	0.0019	0.008	Deletérea	1	D	0/1
276	16	74955911	G	C	<i>WDR59</i>	Proteína Del Complejo Gator Wdr59	0.0024	0.0	Deletérea	1	D	0/1
277	16	3569909	G	T	<i>CLUAP1</i>	Proteína 1 Asociada A Clusterina	0.0038	0.009	Deletérea	1.000	D	0/1
278	16	85144023	C	T	<i>FAM92B</i>	Proteína Fam92B	0.0059	0.002	Deletérea	1.000	D	0/1
279	16	31485534	G	C	<i>TGFB1I1</i>	Proteína Transcripta 1 Inducida Por Factor De Crecimiento Transformante Beta-1	0.0073	0.023	Deletérea	0.544	D	0/1
280	16	15123857	T	C	<i>PDXDC1</i>	Proteína 1 Que Contiene El Dominio Descarboxilasa Dependiente De Piridoxal	0.0080	0.024	Deletérea	0.864	D	0/1
281	16	27642466	C	T	<i>KIAA0556</i>	Proteína Kiaa0556	3.86e-05	0.0	Deletérea	1.000	D	0/1
282	16	68382293	C	T	<i>PRMT7</i>	Proteína Arginina N-Metiltransferasa 7	4.129e-05	0.019	Deletérea	1.000	D	0/1
283	16	84125429	C	T	<i>MBTPS1</i>	Proteasa Del Sitio 1 Del Factor De Transcripción Unido A Membrana Subunidad C De La Arn Polimerasa	5.766e-05	0.0	Deletérea	1	D	0/1
284	16	84214941	T	G	<i>TAF1C</i>	I Del Factor De Unión A La Proteína De Unión A Caja Tata	5.936e-05	0.0	Deletérea	0.999	D	0/1

n	Crom	Posición	Ref	Alt	Gen	Nombre	Frecuencia (EXAC)	Puntaje	Predicción	Puntaje	Predicción	Genotipo
285	16	1816337	C	T	MAPK8IP3	Proteína 3 Que Interactúa Con La Quinasa C-Jun-Amino-Terminal	8.538e-06	0.002	Deletérea	1.000	D	0/1
286	17	11738061	T	A	DNAH9	Cadena Pesada 9 De Dineína, Axonemal	0.0001	0.0	Deletérea	1.000	D	0/1
287	17	26962107	T	G	KIAA0100	Proteína Kiaa0100	0.0002	0.012	Deletérea	1	D	0/1
288	17	76047236	C	T	TNRC6C	Proteína Del Gen 6C Que Contiene Repeticiones De Trinucleótidos	0.0002	0.001	Deletérea	1	D	0/1
289	17	79516352	C	T	FAAP100	Proteína 100 Asociada Al Complejo Central De La Anemia De Fanconi	0.0004	0.012	Deletérea	0.965	D	0/1
290	17	56700270	C	T	TEX14	Serina / Treonina-Proteína Quinasa Inactiva Tex14	0.0004	0.0	Deletérea	0.992	D	0/1
291	17	41249297	G	T	BRCA1	Proteína De Susceptibilidad Al Cáncer De Mama Tipo 1	0.0009	0.0	Deletérea	0.727	D	0/1
292	17	76117158	G	A	TMC6	Proteína 6 Similar A Un Canal Transmembrana;	0.0015	0.022	Deletérea	0.987	D	0/1
293	17	40940414	C	T	WNK4	Serina / Treonina-Proteína Quinasa Wnk4	0.0031	0.05	Deletérea	0.972	D	0/1
294	17	67012450	G	A	ABCA9	Miembro 9 De La Subfamilia A Del Casete De Unión A Atp	0.0059	0.002	Deletérea	0.943	D	0/1
295	17	41004681	C	T	AOC3	Amino Oxidasa Primaria De Membrana	0.0063	0.0	Deletérea	1.000	D	0/1
296	17	4646655	G	A	ZMYND15	Proteína 15 Que Contiene El Dominio Mynd De Dedo De Zinc	0.0064	0.003	Deletérea	0.921	D	0/1
297	17	42254255	C	T	ASB16	Repetición De Anquirina Y Proteína 16 De Caja Socs	0.0067	0.02	Deletérea	0.955	D	0/1
298	17	74473752	G	A	RHBDF2	Proteína Romboide Inactiva 2	0.0073	0.003	Deletérea	1.000	D	0/1
299	17	17919397	C	T	DRC3	Proteína 1 Similar Al Sustrato De Quinasa Del Receptor Del Factor De Crecimiento Epidérmico 8	0.0094	0.001	Deletérea	1	D	0/1
300	17	37262133	A	T	PLXDC1	Proteína 1 Que Contiene El Dominio De Plexina	1.655e-05	0.002	Deletérea	0.983	D	0/1
301	17	39122660	A	G	KRT39	Queratina, Tipo I Citoesquelético 39	8.247e-05	0.008	Deletérea	0.958	D	0/1
302	18	43224125	G	A	SLC14A2	Transportador De Urea 2	0.0012	0.025	Deletérea	0.981	D	0/1
303	18	21136233	G	A	NPC1	Proteína Niemann-Pick C1	0.0039	0.041	Deletérea	0.726	D	0/1
304	18	11889467	C	T	MPPE1	Metalophosphoesterase 1	0.0086	0.003	Deletérea	1	D	0/1
305	19	16024649	G	A	CYP4F11	Filoquinona Omega-Hidroxilasa Cyp4F11	0.0002	0.008	Deletérea	0.995	D	0/1
306	19	4363371	G	A	SH3GL1	Endofilina-A2	0.0003	0.005	Deletérea	0.924	D	0/1
307	19	17338741	A	C	OCEL1	Dominio De Occludina / Ell Que Contiene 1	0.0005	0.003	Deletérea	0.665	D	0/1
308	19	15535510	C	G	WIZ	Protein Wiz	0.0010	0.005	Deletérea	0.500	D	0/1

n	Crom	Posición	Ref	Alt	Gen	Nombre	Frecuencia (EXAC)	Puntaje	Predicción	Puntaje	Predicción	Genotipo
309	19	11221357	G	A	<i>LDLR</i>	Receptor De Lipoproteínas De Baja Densidad	0.0012	0.041	Deletérea	1	D	0/1
310	19	47258942	G	A	<i>FKRP</i>	Proteína Relacionada Con Fukutina	0.0016	0.011	Deletérea	0.130	A	0/1
311	19	50880867	C	T	<i>NR1H2</i>	Receptor De Oxiesteroles Lxr-Beta	0.0016	0.0	Deletérea	0.730	D	0/1
312	19	5733870	G	A	<i>CATSPERD</i>	Delta De La Subunidad Proteica Asociada Al Esperma Del Canal Catiónico	0.0018	0.013	Deletérea	1	D	0/1
313	19	45997446	G	T	<i>RTN2</i>	Reticulon-2	0.0020	0.004	Deletérea	1	D	0/1
314	19	51582885	C	T	<i>KLK14</i>	Calicleína-14	0.0021	0.036	Deletérea	0.640	D	0/1
315	19	9226055	G	T	<i>OR7G1</i>	Receptor Olfatorio 7G1	0.0037	0.0	Deletérea	0.982	D	0/1
316	19	51875296	T	C	<i>NKG7</i>	Proteína Nkg7	0.0053	0.0	Deletérea	0.672	D	0/1
317	19	39914002	G	A	<i>PLEKHG2</i>	Miembro 2 De La Familia G Que Contiene El Dominio De Homología De Pleckstrina	0.0066	0.0	Deletérea	0.807	D	0/1
318	19	55556574	G	C	<i>RDH13</i>	Retinol Deshidrogenasa 13	0.0083	0.011	Deletérea	0.782	D	0/1
319	19	8180480	C	A	<i>FBN3</i>	Fibrilina-3; Fibrilina-3	1.651e-05	0.009	Deletérea	0.604	D	0/1
320	20	42344717	T	G	<i>MYBL2</i>	Proteína B Relacionada Con Myb	0.0003	0.001	Deletérea	1.000	D	0/1
321	20	30232603	G	A	<i>COX4I2</i>	Subunidad 4 De La Citocromo C Oxidasa, Isoforma 2, Mitocondrial	0.0006	0.011	Deletérea	1.000	A	0/1
322	20	34571917	C	T	<i>CNBD2</i>	Proteína 2 Que Contiene El Dominio De Unión A Nucleótidos Cíclicos	0.0006	0.0	Deletérea	1.000	D	0/1
323	20	47266647	C	A	<i>PREX1</i>	Proteína Intercambiadora 1 De Rac Dependiente De Fosfatidilinositol 3,4,5- Trifosfato	0.0009	0.0	Deletérea	1	D	0/1
324	20	49508585	C	G	<i>ADNP</i>	Proteína Homeobox Neuroprotectora Dependiente De La Actividad	0.0012	0.012	Deletérea	0.719	D	0/1
325	20	34220563	A	C	<i>CPNE1</i>	Copine-1	0.0015	0.001	Deletérea	0.991	D	0/1
326	20	60774295	G	A	<i>MTG2</i>	Gtpasa 2 Asociada A Ribosoma Mitocondrial	0.0018	0.024	Deletérea	1	D	0/1
327	20	3452041	C	T	<i>ATRN</i>	Attractin	0.0023	0.041	Deletérea	0.818	D	0/1
328	20	43936836	G	C	<i>RBPJL</i>	Supresor De Proteína De Unión Recombinante De Proteína Similar A La Calvicie	0.0032	0.044	Deletérea	1	D	0/1
329	20	43374694	T	C	<i>KCNK15</i>	Miembro 15 De La Subfamilia K Del Canal De Potasio	0.0076	0.001	Deletérea	1	D	0/1
330	20	32378988	C	T	<i>ZNF341</i>	Proteína De Dedo De Zinc 341	4.302e-05	0.004	Deletérea	0.893	D	0/1
331	20	31608391	G	C	<i>BPIFB2</i>	Miembro 2 De La Familia B Que Contiene Pliegues De Bpi	7.725e-05	0.023	Deletérea	0.977	D	0/1
332	21	19666590	T	C	<i>TMPRSS15</i>	Enteropeptidasa	0.0011	0.006	Deletérea	1	D	0/1
333	21	44483184	A	G	<i>CBS;CBSL</i>	Cystathionine Beta-Synthase Like	0.0012	0.001	Deletérea	0.909	A	0/1

n	Crom	Posición	Ref	Alt	Gen	Nombre	Frecuencia (EXAC)	Puntaje	Predicción	Puntaje	Predicción	Genotipo
334	21	46067036	T	C	<i>KRTAP10-11</i>	Proteína Asociada A La Queratina 10-11	0.0084	0.0	Deletérea	1	D	0/1
335	22	29537940	G	A	<i>KREMEN1</i>	Proteína De Kremen 1	0.0005	0.025	Deletérea	0.996	D	0/1
336	22	19197988	T	A	<i>CLTCL1</i>	Cadena Pesada De Clatrina 2	0.0008	0.001	Deletérea	1.000	D	0/1
337	22	19213744	C	T	<i>CLTCL1</i>	Cadena Pesada De Clatrina 2	0.0008	0.003	Deletérea	1	D	0/1
338	22	29706885	C	T	<i>GAS2L1</i>	Proteína 1 Similar A Gas2	0.0008	0.014	Deletérea	0.982	D	0/1
339	22	24036138	G	A	<i>RGL4</i>	Proteína Relacionada Con Ral-Gds Miembro 1 De La Subfamilia D Del	0.0029	0.002	Deletérea	0.832	D	0/1
340	X	48822589	A	C	<i>KCND1</i>	Canal Dependiente De Voltaje De Potasio	0.0067	0.033	Deletérea	0.981	D	1/1
341	X	108638592	G	A	<i>GUCY2F</i>	Guanilil Ciclasa 2 Retiniana	2.294e-05	0.016	Deletérea	0.988	D	1/1
Variantes SP05. (Ref: Nucleotido de referencia; Alt: Nucleotido alternativo; D: variante causa de enfermedad, es decir, probablemente perjudicial; A: variante enfermedad que causa automáticamente, es decir, se sabe que es perjudicial)												
n	Crom	Posición	Ref	Alt	Gen	Nombre del gen	EXAC Frecuencia	SIFT Puntaje	Predicción	Mutation Taster Puntaje	Predicción	Genotipo
1	1	248616465	C	T	<i>OR2T2</i>	Receptor Olfatorio 2T2	0.0004	0.045	Deleterea	0.988	D	0/1
2	1	45500064	C	T	<i>ZSWIM5</i>	Dedo De Zinc Tipo Swim Que Contiene 5	0.0007	0.012	Deleterea	1	D	0/1
3	1	17567149	G	A	<i>PADI1</i>	Proteína-Arginina Deiminasa Tipo 1	0.0051	0.011	Deleterea	0.987	D	0/1
4	1	230492801	C	T	<i>PGBD5</i>	Proteína 5 Derivada De Elementos Transponibles Piggybac	0.0068	0.019	Deleterea	1	D	0/1
5	1	211654619	G	A	<i>RD3</i>	Proteína Rd3	0.0096	0.004	Deleterea	0.999	D	0/1
6	1	74957853	C	T	<i>TNNI3K</i>	Serina / Treonina-Proteína Quinasa Tnni3K	2.473e-05	0.0	Deleterea	0.999	D	0/1
7	1	154184830	G	A	<i>C1orf43</i>	Proteína No Caracterizada C1Orf43	4.942e-05	0.039	Deleterea	0.997	D	0/1
8	1	205568353	A	G	<i>MFSD4A</i>	Dominio De Superfamilia De Facilitador Principal Que Contiene 4	9.064e-05	0.006	Deleterea	0.999	D	0/1
9	2	179419727	C	T	<i>TTN</i>	Titina	0.0004	0.019	Deleterea	0.971	D	0/1
10	2	24345324	C	T	<i>PFN4</i>	Profilin-4	0.0023	0.028	Deleterea	0.996	D	0/1
11	2	27721143	G	A	<i>GCKR</i>	Proteína Reguladora De Glucoquinasa	0.0023	0.004	Deleterea	1.000	D	0/1
12	2	220079136	A	G	<i>ABCB6</i>	Miembro 6 De La Subfamilia B Del Casete De Unión A Atp, Mitochondrial	0.0046	0.004	Deleterea	1	D	0/1
13	2	61412677	C	G	<i>AHSA2</i>	Activador Del Homólogo 2 De Atpasa De Proteína De Choque Térmico De 90 Kda	0.0054	0.002	Deleterea	1.000	D	0/1
14	2	153484907	C	T	<i>FMNL2</i>	Proteína 2 Similar A La Formina	1.656e-05	0.02	Deleterea	0.998	D	0/1

n	Crom	Posición	Ref	Alt	Gen	Nombre	Frecuencia (EXAC)	Puntaje	Predicción	Puntaje	Predicción	Genotipo
15	2	217001820	G	T	<i>XRCC5</i>	Proteína 5 De Complemento Cruzado De Reparación De Rayos X;	2.474e-05	0.002	Deleterea	1	D	0/1
16	2	220089872	C	T	<i>ATG9A</i>	Proteína 9A Relacionada Con La Autofagia	3.313e-05	0.044	Deleterea	1	D	0/1
17	3	183495377	A	C	<i>YEATS2</i>	Proteína 2 Que Contiene El Dominio Yeats	0.0002	0.0	Deleterea	0.999	D	0/1
18	3	48369830	A	G	<i>SPINK8</i>	Inhibidor De La Serina Proteasa Kazal-Tipo 8	0.0015	0.0	Deleterea	0.847	D	0/1
19	3	132277907	G	A	<i>ACAD11</i>	Miembro 11 De La Familia De Las Acil-CoA Deshidrogenasa	0.0093	0.0	Deleterea	1	D	0/1
20	3	184910078	G	A	<i>EHHADH</i>	Enzima Bifuncional Peroxisomal; Enoil-CoA Hidratasa Y 3-Hidroxiacil CoA Deshidrogenasa	0.0097	0.004	Deleterea	0.665	D	0/1
21	3	51972149	C	T	<i>RRP9</i>	Proteína 2 Que Interactúa Con Arn Nucleolar Pequeño U3	2.041e-05	0.0	Deleterea	1	D	0/1
22	3	47455402	T	C	<i>SCAP</i>	Proteína Que Activa La Escisión De La Proteína De Unión Al Elemento Regulador De Esteroles	2.478e-05	0.016	Deleterea	0.977	D	0/1
23	4	48563497	C	T	<i>FRYL</i>	Proteína Similar A Un Homólogo Peludo	0.0022	0.004	Deleterea	1.000	D	1/1
24	4	6304053	C	T	<i>WFS1</i>	Wolframin	5.875e-05	0.036	Deleterea	1	D	0/1
25	4	38119753	C	T	<i>TBC1D1</i>	Miembro 1 De La Familia Del Dominio Tbc1	9.884e-05	0.0	Deleterea	1	D	0/1
26	4	190876272	G	A	<i>FRG1</i>	Proteína Frg1	9.916e-05	0.011	Deleterea	1	D	0/1
27	5	176002580	C	T	<i>CDHR2</i>	Miembro 2 De La Familia Relacionado Con La Cadherina	0.0003	0.005	Deleterea	0.868	D	0/1
28	5	56219010	T	C	<i>MIER3</i>	Proteína 3 De Respuesta Temprana De Inducción De Mesodermo	0.0004	0.004	Deleterea	0.996	D	0/1
29	5	96512920	C	T	<i>RIOK2</i>	Serina / Treonina-Proteína Quinasa Rio2	0.0006	0.002	Deleterea	1	D	0/1
30	6	7542160	C	G	<i>DSP</i>	Desmoplakin	0.0050	0.009	Deleterea	0.946	D	0/1
31	6	152129399	T	C	<i>ESR1</i>	Receptor De Estrógenos	0.0098	0.001	Deleterea	1.000	D	0/1
32	6	109766471	T	A	<i>MICAL1</i>	[F-Actina] -Monoxigenasa Mical1	0.0098	0.044	Deleterea	0.999	D	0/1
33	7	44556373	C	T	<i>NPC1L1</i>	Proteína 1 Similar A C1 De Niemann-Pick	0.0009	0.002	Deleterea	0.949	D	0/1
34	7	116340086	A	G	<i>MET</i>	Receptor Del Factor De Crecimiento De Hepatocitos	0.0018	0.006	Deleterea	0.997	D	0/1
35	7	100086599	C	T	<i>NYAP1</i>	Adaptador 1 De Fosfoinositido-3-Quinasa Fosforilada En Tirosina Neuronal	0.0024	0.0	Deleterea	0.989	D	0/1

n	Crom	Posición	Ref	Alt	Gen	Nombre	Frecuencia (EXAC)	Puntaje	Predicción	Puntaje	Predicción	Genotipo
36	8	144650859	C	G	<i>MROH6</i>	Maestro Caliente Como Miembro Repetido De La Familia 6	0.0043	0.013	Deleterea	1	D	0/1
37	8	22473078	A	G	<i>CCAR2</i>	Proteína 2 Reguladora Del Ciclo Celular Y La Apoptosis	8.252e-05	0.04	Deleterea	0.995	D	0/1
38	9	140002934	C	T	<i>MAN1B1</i>	Manosil-Oligosacárido 1,2-Alfa-Manosidasa Del Retículo Endoplásmico	0.0002	0.0	Deleterea	1	D	0/1
39	10	27529313	G	A	<i>ACBD5</i>	Proteína 5 Que Contiene El Dominio De Unión A Acil-CoA	0.0034	0.005	Deleterea	0.994	D	0/1
40	11	18425256	G	C	<i>LDHA</i>	Lactato Deshidrogenasa A	0.0026	0.008	Deleterea	1	D	0/1
41	11	5529995	A	G	<i>UBQLN3</i>	Ubiquilina-3	0.0026	0.0	Deleterea	1	D	1/1
42	11	110451011	G	A	<i>ARHGAP20</i>	Proteína Activadora De Rho Gtpasa 20	0.0040	0.0	Deleterea	0.596	D	0/1
43	11	36595760	C	A	<i>RAG1</i>	Proteína 1 Activadora De Recombinación V (D) J	0.0074	0.0	Deleterea	0.998	D	0/1
44	11	11987497	A	G	<i>DKK3</i>	Proteína 3 Relacionada Con Dickkopf	2.764e-05	0.0	Deleterea	1.000	D	1/1
45	12	93251098	C	G	<i>EEA1</i>	Antígeno 1 Del Endosoma Temprano	0.0027	0.025	Deleterea	1.000	D	1/1
46	12	111886063	G	A	<i>SH2B3</i>	Proteína Adaptadora 3 De Sh2B	1.706e-05	0.0	Deleterea	1	D	0/1
47	12	57969020	C	T	<i>KIF5A</i>	Isoforma 5A De La Cadena Pesada De Kinesina	2.69e-05	0.001	Deleterea	1.000	D	0/1
48	12	344275	G	A	<i>SLC6A13</i>	Transportador 2 De Gaba Dependiente De Sodio Y Cloruro	5.766e-05	0.005	Deleterea	0.893	D	0/1
49	12	57593183	G	A	<i>LRP1</i>	Proteína 1 Relacionada Con El Receptor De Lipoproteínas De Densidad Prolongada	8.26e-06	0.011	Deleterea	1	D	0/1
50	12	51072654	C	T	<i>DIP2B</i>	Proteína De Interacción Disco 2 Homólogo B	8.315e-06	0.0	Deleterea	1	D	0/1
51	13	113505454	G	A	<i>ATP11A</i>	Atpasa 1h Probablemente Transportadora De Fosfolípidos	7.414e-05	0.006	Deleterea	1	D	0/1
52	14	20528558	G	A	<i>OR4L1</i>	Receptor Olfatorio 4L1	0.0089	0.0	Deleterea	0.881	D	0/1
53	15	78921602	G	A	<i>CHRNA4</i>	Subunidad Beta-4 Del Receptor De Acetilcolina Neuronal	0.0055	0.0	Deleterea	1	D	0/1
54	15	73552705	T	C	<i>NEO1</i>	Neogenina	1.649e-05	0.025	Deleterea	1	D	0/1
55	15	89861968	G	A	<i>POLG</i>	Subunidad Gamma-1 De La Adn Polimerasa	1.649e-05	0.0	Deleterea	1	D	0/1
56	15	52539168	G	T	<i>MYO5C</i>	Miosina-Vc No Convencional	4.969e-05	0.0	Deleterea	1	D	0/1
57	15	85201233	C	T	<i>NMB</i>	Neuromedina-B	8.845e-06	0.005	Deleterea	0.998	D	0/1
58	16	89774899	G	T	<i>VPS9D1</i>	Dominio Vps9 Que Contiene 1	0.0026	0.011	Deleterea	0.999	D	0/1

n	Crom	Posición	Ref	Alt	Gen	Nombre	Frecuencia (EXAC)	Puntaje	Predicción	Puntaje	Predicción	Genotipo
59	16	3081061	C	T	<i>BICDL2</i>	Adaptador De Carga Tipo Familia Bcd 2	0.0028	0.007	Deleterea	0.903	D	0/1
60	16	19083300	G	C	<i>COQ7</i>	5-Demetoxiubiquinona Hidroxilasa, Mitochondrial	0.0059	0.006	Deleterea	1.000	D	0/1
61	16	68112725	C	T	<i>DUS2</i>	Arnt-Dihidrouridina (20) Sintasa [Similar A Nad (P) +]	0.0096	0.005	Deleterea	0.964	D	0/1
62	17	76201226	C	T	<i>AFMID</i>	Quinurenina Formamidasa	0.0001	0.01	Deleterea	1	D	0/1
63	17	4925475	C	T	<i>KIF1C</i>	Proteína De Tipo Kinesina Kif1C	0.0032	0.001	Deleterea	1	D	0/1
64	17	11572808	A	G	<i>DNAH9</i>	Cadena Pesada 9 De Dineína, Axonemal	0.0056	0.0	Deleterea	1.000	D	0/1
65	17	6544421	G	A	<i>TXNDC17</i>	Proteína 17 Que Contiene El Dominio De Tiorredoxina	0.0075	0.005	Deleterea	1.000	D	0/1
66	17	74273308	C	T	<i>QRICH2</i>	Proteína 2 Rica En Glutamina	5.81e-05	0.0	Deleterea	0.969	D	0/1
67	19	39018329	C	T	<i>RYR1</i>	Receptor 1 De Rianodina	0.0001	0.023	Deleterea	0.658	D	0/1
68	19	2408365	C	T	<i>TMPRSS9</i>	Serina 9 De Proteasa Transmembrana	5.845e-05	0.002	Deleterea	1	D	0/1
69	20	19956311	G	A	<i>RIN2</i>	Interactor 2 Ras Y Rab	0.0049	0.024	Deleterea	0.997	D	0/1
70	20	20585894	C	T	<i>RALGAPA2</i>	Subunidad Alfa-2 De La Proteína Activadora De Ral Gtpasa	0.0096	0.02	Deleterea	1.000	D	0/1
71	21	42830641	C	T	<i>MX1</i>	Proteína De Unión A Gtp Inducida Por Interferón Mx1	7.49e-05	0.0	Deleterea	0.997	D	1/1
72	22	23915670	G	A	<i>IGLL1</i>	Polipéptido 1 Similar A Inmunoglobulina Lambda	0.0002	0.014	Deleterea	1.000	A	0/1

Variantes SP06. (Ref: Nucleotido de referencia; Alt: Nucleotido alternativo; D: variante causa de enfermedad, es decir, probablemente perjudicial)

n	Crom	Posición	Ref	Alt	Gen	Nombre del gen	EXAC	SIFT		Mutation Taster		Genotipo
							Frecuencia	Puntaje	Predicción	Puntaje	Predicción	
1	1	156830911	T	C	<i>NTRK1</i>	Receptor Del Factor De Crecimiento Nervioso De Alta Afinidad	0.0003	0.002	Deleterea	1	D	0/1
2	1	95290119	G	A	<i>SLC44A3</i>	Proteína 3 Similar Al Transportador De Colina	0.0005	0.0	Deleterea	1.000	D	0/1
3	1	100133249	G	C	<i>PALMD</i>	Palmdephin	0.0019	0.011	Deleterea	0.962	D	0/1
4	1	36553608	C	T	<i>TEKT2</i>	Tektin-2	0.0027	0.012	Deleterea	0.990	D	0/1
5	1	151503071	G	A	<i>CGN</i>	Cingulina	0.0044	0.005	Deleterea	1.000	D	0/1
6	1	46810797	A	G	<i>NSUN4</i>	5-Metilcitosina Rna Metiltransferasa Nsun4	0.0065	0.004	Deleterea	1.000	D	1/1
7	1	109801543	A	G	<i>CELSR2</i>	Receptor 2 De Tipo G De Siete Pasos De Cadherina Egf Lag	0.0080	0.013	Deleterea	1.000	D	0/1
8	1	47024454	G	A	<i>MKNK1</i>	Serina / Treonina-Proteína Quinasa 1 Que Interactúa Con Map Quinasa	0.0099	0.013	Deleterea	1	D	0/1

n	Crom	Posición	Ref	Alt	Gen	Nombre	Frecuencia (EXAC)	Puntaje	Predicción	Puntaje	Predicción	Genotipo
9	1	9086402	C	A	SLC2A7	Familia De Portadores De Solutos 2, Miembro Transportador De Glucosa Facilitado 7	1.346e-05	0.045	Deleterea	1.000	D	0/1
10	1	41514431	C	T	SCMH1	Proteína Polycomb Scmh1	1.647e-05	0.008	Deleterea	1	D	0/1
11	1	78097949	T	C	ZZZ3	Proteína 3 Que Contiene Dedos De Zinc Tipo Zz	1.648e-05	0.042	Deleterea	0.847	D	0/1
12	1	241946598	C	T	WDR64	Proteína Que Contiene Repeticiones Wd 64	5.781e-05	0.001	Deleterea	1.000	D	0/1
13	1	222895843	G	A	BROX	Proteína Brox Que Contiene El Dominio Bro1	8.341e-06	0.0	Deleterea	1	D	0/1
14	1	156846291	G	A	NTRK1	Receptor Del Factor De Crecimiento Nervioso De Alta Afinidad	8.52e-06	0.037	Deleterea	1	D	0/1
15	2	232393323	G	A	NMUR1	Receptor 1 De Neuromedina -U;	0.0005	0.0	Deleterea	1	D	0/1
16	2	218699832	T	C	TNS1	Tensina-1	0.0010	0.005	Deleterea	0.999	D	0/1
17	2	109116016	A	G	GCC2	Proteína 2 Que Contiene El Dominio Grip Y En Espiral	0.0028	0.0	Deleterea	1.000	D	0/1
18	2	105708941	G	A	MRPS9	Proteína Ribosómica Mitocondrial S9	0.0041	0.049	Deleterea	1	D	1/1
19	2	71377014	G	A	MPHOSPH10	Proteína De Ribonucleoproteína Nucleolar Pequeña U3 Mpp10	0.0083	0.015	Deleterea	0.988	D	0/1
20	2	103340230	T	C	MFSD9	Dominio De Superfamilia Facilitador Principal Que Contiene 9	2.472e-05	0.003	Deleterea	0.998	D	1/1
21	2	202514844	A	G	MPP4	Miembro 4 De La Subfamilia De Maguk P55	2.886e-05	0.0	Deleterea	1	D	0/1
22	2	223497962	T	C	FARSB	Subunidad Beta De Fenilalanil-Trna Sintetasa	8.242e-06	0.008	Deleterea	1.000	D	0/1
23	2	220039771	C	T	CNPPD1	Proteína Cnppd1	8.298e-06	0.009	Deleterea	1	D	0/1
24	2	201724452	A	T	CLK1	Proteína Quinasa De Especificidad Dual Clk1	8.316e-05	0.023	Deleterea	1.000	D	0/1
25	3	195511804	G	A	MUC4	Mucina-4	0.0011	0.031	Deleterea	0.973	D	0/1
26	3	183432967	G	A	YEATS2	Proteína 2 Que Contiene El Dominio Yeats	0.0040	0.001	Deleterea	1.000	D	0/1
27	3	12861600	G	A	CAND2	Proteína 2 Disociada De Nedd8 Asociada A Cullina	0.0041	0.002	Deleterea	1	D	0/1
28	3	196454939	C	T	PIGX	Proteína De Clase X De Biosíntesis De Fosfatidilinositol-Glicano	0.0079	0.0	Deleterea	1	D	0/1
29	3	127394818	C	T	ABTB1	Repetición De Anquirina Y Proteína 1 Que Contiene El Dominio Btb / Poz	2.702e-05	0.005	Deleterea	1.000	D	0/1
30	3	120130757	C	T	FSTL1	Proteína 1 Relacionada Con Folistatina	3.296e-05	0.005	Deleterea	1.000	D	0/1

n	Crom	Posición	Ref	Alt	Gen	Nombre	Frecuencia (EXAC)	Puntaje	Predicción	Puntaje	Predicción	Genotipo
31	4	187088399	G	T	<i>FAM149A</i>	Proteína Fam149A	0.0054	0.003	Deleterea	1.000	D	0/1
32	4	76809396	G	A	<i>PPEF2</i>	Serina / Treonina-Proteína Fosfatasa Con Ef-Manos 2	0.0098	0.002	Deleterea	1.000	D	0/1
33	4	3478120	C	T	<i>DOK7</i>	Proteína Dok-7	1.677e-05	0.022	Deleterea	1.000	D	0/1
34	5	191665	G	A	<i>LRRC14B</i>	Proteína 14B Que Contiene Repetición Rica En Leucina	0.0002	0.0	Deleterea	1.000	D	0/1
35	5	1213611	G	A	<i>SLC6A19</i>	Transportador De Aminoácidos Neutros Dependiente De Sodio B (0) At1	5.022e-05	0.002	Deleterea	1.000	D	0/1
36	6	99848591	C	T	<i>PNISR</i>	Arginina / Serina-Proteína Rica Pnistr	0.0041	0.0	Deleterea	1	D	0/1
37	6	24520644	G	A	<i>ALDH5A1</i>	Succinato-Semialdehído Deshidrogenasa, Mitocondrial	0.0055	0.008	Deleterea	1	D	0/1
38	7	87517349	A	C	<i>DBF4</i>	Homólogo A De Proteína Dbf4	0.0025	0.018	Deleterea	0.920	D	0/1
39	7	100205640	C	T	<i>PCOLCE</i>	Potenciador 1 Del Procolágeno C-Endopeptidasa	3.306e-05	0.0	Deleterea	0.937	D	0/1
40	9	136217563	G	A	<i>RPL7A</i>	Proteína Ribosomal L7A	0.0036	0.023	Deleterea	1	D	0/1
41	9	2820046	G	C	<i>PUM3</i>	Homólogo 3 De Pumilio	0.0058	0.0	Deleterea	1	D	0/1
42	9	131094512	G	C	<i>COQ4</i>	Proteína De Biosíntesis De Ubiquinona Homólogo Coq4 , Mitocondrial	0.0075	0.006	Deleterea	1.000	D	0/1
43	9	139925596	C	T	<i>FUT7</i>	- Alfa- (1,3) -Fucosiltransferasa 7	1.711e-05	0.001	Deleterea	1	D	0/1
44	9	107367545	G	A	<i>OR13C2</i>	Receptor Olfatorio 13C2	5.775e-05	0.012	Deleterea	0.921	D	0/1
45	10	72604246	A	G	<i>SGPL1</i>	Esfingosina -1-Fosfato Liasa 1	0.0004	0.01	Deleterea	1.000	D	0/1
46	10	94006	C	A	<i>TUBB8</i>	Cadena De Tubulina Beta-8	0.0017	0.0	Deleterea	1.000	D	0/1
47	10	31810674	T	C	<i>ZEB1</i>	Homeobox 1 De Encuadernación E-Box Con Dedos De Zinc	0.0019	0.005	Deleterea	1.000	D	0/1
48	10	43287148	G	A	<i>BMS1</i>	Homólogo De Bms1 De La Proteína De Biogénesis Del Ribosoma	0.0097	0.0	Deleterea	1.000	D	0/1
49	10	116014686	C	T	<i>VWA2</i>	Dominio Del Factor A De Von Willebrand Que Contiene 2	3.324e-05	0.006	Deleterea	0.949	D	0/1
50	10	24669936	C	T	<i>KIAA1217</i>	Homólogo De Proteína De Cola De Hoz	6.712e-05	0.0	Deleterea	1	D	0/1
51	11	60901658	G	A	<i>VPS37C</i>	Proteína 37C Asociada A La Clasificación De Proteínas Vacuolares	7.457e-05	0.0	Deleterea	0.999	D	0/1
52	11	134023065	T	C	<i>NCAPD3</i>	Subunidad D3 Del Complejo De Condensina-2	8.238e-06	0.018	Deleterea	1.000	D	0/1
53	11	77412037	G	A	<i>RSF1</i>	Factor De Remodelación Y Espaciamento 1	8.247e-06	0.01	Deleterea	1.000	D	0/1

n	Crom	Posición	Ref	Alt	Gen	Nombre	Frecuencia (EXAC)	Puntaje	Predicción	Puntaje	Predicción	Genotipo
54	12	122724463	C	T	VPS33A	Proteína 33A Asociada A La Clasificación De Proteínas Vacuolares	0.0002	0.047	Deletérea	1.000	D	0/1
55	12	56721807	A	C	PAN2	Subunidad Catalítica Del Complejo De Desadenilación Pan2 -Pan3 Pan2	0.0010	0.002	Deletérea	1.000	D	0/1
56	12	104144426	C	T	STAB2	Estabilina -2	0.0060	0.0	Deletérea	1.000	D	0/1
57	12	1963174	G	A	CACNA2D4	Subunidad Alfa-2 / Delta-4 Del Canal De Calcio Dependiente Del Voltaje	0.0064	0.004	Deletérea	1	D	0/1
58	12	120995085	C	T	RNF10	Proteína De Dedo Anillo 10	3.319e-05	0.0	Deletérea	1	D	0/1
59	12	57627365	G	A	SHMT2	Serina Hidroximetiltransferasa, Mitocondrial	3.32e-05	0.046	Deletérea	1.000	D	0/1
60	12	108912794	G	A	FICD	Adenosina Monofosfato-Proteína Transferasa Ficd	5.782e-05	0.0	Deletérea	1	D	0/1
61	12	5744378	C	T	ANO2	Anoctamin-2	6.006e-05	0.003	Deletérea	1.000	D	0/1
62	14	91763720	G	A	CCDC88C	Proteína Daple	0.0049	0.0	Deletérea	1	D	0/1
63	14	105411287	G	A	AHNAK2	Nucleoproteína 2 De Ahnak	0.0066	0.015	Deletérea	1.000	D	0/1
64	14	105411286	G	A	AHNAK2	Nucleoproteína 2 De Ahnak	0.0068	0.004	Deletérea	1.000	D	0/1
65	14	92959940	G	A	SLC24A4	Intercambiador 4 De Sodio / Potasio / Calcio	0.0084	0.025	Deletérea	1	D	0/1
66	15	31332488	C	G	TRPM1	Miembro 1 De La Subfamilia M Del Canal Catiónico Potencial Del Receptor Transitorio	0.0063	0.001	Deletérea	1	D	0/1
67	15	56386514	T	A	RFX7	Proteína De Unión A Adn Rfx7	0.0086	0.001	Deletérea	0.967	D	0/1
68	15	72191110	T	C	MYO9A	Miosina-Ixa No Convencional	0.0100	0.042	Deletérea	0.993	D	0/1
69	15	90610884	G	A	ZNF710	Proteína De Dedos De Zinc 710	1.914e-05	0.013	Deletérea	1.000	D	0/1
70	16	88600128	A	C	ZFPM1	Proteína De Dedo De Zinc Zfpm1	0.0002	0.002	Deletérea	1.000	D	0/1
71	16	2024098	G	C	TBL3	Proteína 3 Similar A La Transducina Beta	0.0003	0.011	Deletérea	1	D	0/1
72	16	3707023	C	G	DNASE1	Desoxirribonucleasa-1	0.0059	0.045	Deletérea	0.999	D	0/1
73	16	89703030	C	T	DPEP1	Dipeptidasa 1	2.513e-05	0.0	Deletérea	1.000	D	0/1
74	16	2134233	C	T	TSC2	Tuberina	6.111e-05	0.002	Deletérea	1	D	0/1
75	17	39190973	C	G	KRTAP1-3	Proteína 1-3 Asociada A Queratina	0.0004	0.0	Deletérea	0.816	D	0/1
76	17	39261759	G	A	KRTAP4-9	Proteína 4-9 Asociada A Queratina	0.0025	0.048	Deletérea	0.572	D	0/1
77	17	1538083	A	G	SCARF1	Miembro 1 Del Receptor Eliminador De Clase F	0.0026	0.009	Deletérea	1	D	0/1
78	17	41606033	G	A	ETV4	Variante 4 De Translocación De Ets	0.0076	0.021	Deletérea	1	D	0/1
79	18	42531221	A	G	SETBP1	Set Proteína De Unión 1	8.598e-06	0.011	Deletérea	0.604	D	0/1
80	19	58004950	C	A	ZNF419	Proteína De Dedo De Zinc 419	0.0001	0.0	Deletérea	0.998	D	0/1

n	Crom	Posición	Ref	Alt	Gen	Nombre	Frecuencia (EXAC)	Puntaje	Predicción	Puntaje	Predicción	Genotipo
81	19	49132605	G	T	SPHK2	Esfingosina Quinasa 2	0.0001	0.003	Deleterea	0.964	D	0/1
82	19	53453260	C	G	ZNF816	Proteína De Dedos De Zinc 816	0.0043	0.011	Deleterea	1	D	0/1
83	19	58118371	G	A	ZNF530	Proteína De Dedos De Zinc 530	0.0082	0.0	Deleterea	0.540	D	0/1
84	19	327166	C	T	MIER2	Proteína 2 De Respuesta Temprana De Inducción De Mesodermo	8.245e-06	0.032	Deleterea	0.774	D	0/1
85	20	43353580	A	T	WISP2	Proteína 2 De La Vía De Señalización Inducible Por Wnt1	6.249e-05	0.001	Deleterea	1.000	D	0/1
86	22	41605776	G	C	L3MBTL2	Proteína 2 Similar A Un Tumor Cerebral Maligno Letal (3)	0.0093	0.018	Deleterea	0.591	D	0/1
87	22	26868300	A	G	HPS4	Proteína Del Síndrome De Hermansky-Pudlak 4	8.266e-06	0.022	Deleterea	1.000	D	0/1
88	X	54989791	G	A	PFKFB1	6-Fosfofructo-2-Quinasa / Fructosa-2,6-Bisfosfatasa 1	0.0001	0.013	Deleterea	1.000	D	0/1

Variantes SP07. (Ref: Nucleotido de referencia; Alt: Nucleotido alternativo; D: variante causa de enfermedad, es decir, probablemente perjudicial)

n	Crom	Posición	Ref	Alt	Gen	Nombre del gen	EXAC	SIFT		Mutation Taster		Genotipo
							Frecuencia	Puntaje	Predicción	Puntaje	Predicción	
1	1	111893918	G	A	PIFO	Horquilla De Proteína	1.674e-05	0.024	Deleterea	1.000	D	0/1
2	1	67518668	G	A	SLC35D1	Transportador De Udp-Ácido Glucurónico / Udp-N-Acetilgalactosamina	1.669e-05	0.002	Deleterea	1	D	0/1
3	1	20209071	G	T	OTUD3	Proteína 3 Que Contiene El Dominio Otu	0.0027	0.006	Deleterea	1	D	0/1
4	1	75108782	G	A	ERICH3	Proteína 3 Rica En Glutamato	1.671e-05	0.0	Deleterea	1	D	0/1
5	1	1848197	G	C	CALML6	Proteína 6 Similar A La Calmodulina	0.0041	0.021	Deleterea	0.985	D	0/1
6	1	11771929	G	A	DRAXIN	Draxin	0.0085	0.0	Deleterea	1	D	0/1
7	1	207195520	G	A	C1orf116	Proteína Génica Específicamente Regulada Por Andrógenos	0.0020	0.0	Deleterea	1	D	0/1
8	1	84948653	C	A	RPF1	Factor De Producción De Ribosoma 1	0.0010	0.0	Deleterea	1	D	0/1
9	1	65157122	C	T	CACHD1	Vwfa Y Proteína 1 Que Contiene El Dominio De Caché	0.0006	0.0	Deleterea	1	D	0/1
10	2	76854612	A	G	DUSP13	Isoforma B De Proteína Fosfatasa 13 De Especificidad Dual	0.0014	0.004	Deleterea	1.000	D	0/1
11	2	119057994	C	T	PDZD3	Cofactor Regulador De Intercambio De Na (+) / H (+) Nhe-Rf4	0.0006	0.0	Deleterea	0.617	D	0/1
12	2	83641521	G	A	DLG2	Discos De Gran Homólogo 2	0.0008	0.05	Deleterea	1	D	0/1
13	2	77602487	T	C	INTS4	Subunidad 4 Del Complejo Integrador	0.0019	0.014	Deleterea	1	D	0/1
14	2	52774934	C	T	KRT84	Queratina, Hb4 Cuticular De Tipo Ii	8.238e-06	0.001	Deleterea	0.568	D	0/1

n	Crom	Posición	Ref	Alt	Gen	Nombre	Frecuencia (EXAC)	Puntaje	Predicción	Puntaje	Predicción	Genotipo
15	2	88439500	A	G	<i>C12orf29</i>	Proteína C12Orf29 No Caracterizada	0.0021	0.019	Deleterea	0.999	D	0/1
16	2	49431844	G	A	<i>KMT2D</i>	Histona-Lisina N-Metiltransferasa 2D	5.925e-05	0.0	Deleterea	0.859	D	0/1
17	3	130452500	C	T	<i>PIK3R4</i>	Subunidad 4 Reguladora De Fosfoinositido 3-Quinasa	0.0039	0.0	Deleterea	1	D	0/1
18	3	146171883	G	C	<i>PLSCR2</i>	Scramblase 2 De Fosfolípidos	0.0036	0.013	Deletérea	0.729	D	0/1
19	4	187628419	C	T	<i>FAT1</i>	Protocadherin Fat 1	0.0022	0.034	Deletérea	1	D	0/1
20	4	38775929	A	G	<i>TLR10</i>	Receptor Tipo Toll 10	0.0005	0.0	Deletérea	1	D	0/1
21	5	52243230	T	A	<i>ITGA1</i>	Integrina Alfa-1	0.0002	0.0	Deletérea	1	D	0/1
22	5	70936895	A	T	<i>MCCC2</i>	Cadena Beta De Metilcrotonoil-CoA Carboxilasa, Mitochondrial	0.0002	0.001	Deletérea	1	D	0/1
23	5	56111633	T	C	<i>MAP3K1</i>	Proteína Quinasa 1 Activada Por Mitógeno	0.0005	0.035	Deletérea	1.000	D	0/1
24	5	74807116	C	T	<i>COL4A3BP</i>	Proteína De Unión A Alfa-3 De Colágeno Tipo Iv	0.0064	0.028	Deletérea	1	D	0/1
25	5	176315547	G	A	<i>HK3</i>	Protein Hook Homólogo 3	0.0091	0.013	Deletérea	0.789	D	0/1
26	5	140228208	C	A	<i>PCDHA9</i>	Protocadherina Alfa-9	0.0065	0.01	Deletérea	1	D	1/1
27	5	176931871	C	T	<i>DOK3</i>	Proteína De Acoplamiento 3	0.0014	0.001	Deletérea	0.999	D	0/1
28	6	159398871	C	T	<i>RSPH3</i>	Homólogo De Proteína 3 De Cabeza De Radios Radiales	0.0083	0.006	Deletérea	0.996	D	0/1
29	6	158495613	G	A	<i>SYNJ2</i>	Synaptojanin-2	0.0005	0.001	Deletérea	1.000	D	0/1
30	6	142737048	C	T	<i>ADGRG6</i>	Receptor G6 Acoplado A Proteína G De Adhesión	0.0030	0.038	Deletérea	1	D	0/1
31	6	74440236	G	A	<i>CD109</i>	Antígeno Cd109	6.613e-05	0.0	Deletérea	0.999	D	0/1
32	7	128498402	T	C	<i>FLNC</i>	Filamin-C	0.0012	0.002	Deletérea	1.000	D	0/1
33	7	150784001	C	T	<i>AGAP3</i>	Arf-Gap Con Gtpasa, Repetición Ank Y Proteína 3 Que Contiene El Dominio Ph	0.0001	0.002	Deletérea	1.000	D	0/1
34	7	128441545	G	C	<i>CCDC136</i>	Proteína 136 Que Contiene El Dominio En Espiral	0.0086	0.001	Deletérea	0.992	D	0/1
35	7	150933593	G	A	<i>CHPF2</i>	Condroitín Sulfato Glucuroniltransferasa	0.0025	0.008	Deletérea	1.000	D	0/1
36	7	6502761	G	A	<i>KDELRL2</i>	Receptor 2 De Retención De Proteínas De La Luz Del Er	0.0003	0.007	Deletérea	1	D	0/1
37	7	27134966	T	G	<i>HOXA1</i>	Proteína Homeobox Hox-A1	0.0010	0.049	Deletérea	1.000	D	0/1
38	9	128678097	C	T	<i>PBX3</i>	Factor De Transcripción 3 De Leucemia De Células Pre-B	0.0068	0.014	Deletérea	1	D	0/1
39	9	1056945	C	T	<i>DMRT2</i>	Factor De Transcripción 2 Relacionado Con Dobleex Y Mab-3	0.0005	0.006	Deletérea	1	D	0/1

n	Crom	Posición	Ref	Alt	Gen	Nombre	Frecuencia (EXAC)	Puntaje	Predicción	Puntaje	Predicción	Genotipo
40	9	100767307	A	C	ANP32B	Miembro B De La Familia De Fosfoproteína Nuclear 32 Rica En Leucina Ácida	0.0037	0.014	Deletérea	1	D	0/1
41	9	33797898	C	T	PRSS3	Tripsina-3	9.07e-05	0.049	Deletérea	0.877	D	0/1
42	10	73978772	A	C	HEATR4	Proteína 4 Que Contiene Repeticiones De Heat	0.0059	0.002	Deletérea	0.707	D	0/1
43	11	23354014	C	T	REM2	Proteína De Unión A Gtp Rem 2 Regulador Del Crecimiento Celular	0.0044	0.001	Deletérea	1.000	D	0/1
44	11	54997751	C	T	CGRRF1	Con Proteína 1 De Dominio De Dedo Ring	0.0074	0.014	Deletérea	0.995	D	0/1
45	11	31792893	C	T	HEATR5A	Proteína 5A Que Contiene Repeticiones De Heat	4.977e-05	0.009	Deletérea	1.000	D	0/1
46	12	49917618	C	T	DTWD1	Dominio Dtw Que Contiene 1	0.0005	0.003	Deletérea	1	D	0/1
47	12	90814559	A	G	NGRN	Neugrin	0.0039	0.007	Deletérea	1.000	D	0/1
48	12	86076982	C	T	AKAP13	Proteína De Anclaje De La A-Quinasa 13	2.474e-05	0.002	Deletérea	0.998	D	0/1
49	14	21658490	T	C	IGSF6	Miembro 6 De La Superfamilia De Inmunoglobulinas	8.239e-06	0.008	Deletérea	0.712	D	0/1
50	14	68156591	T	A	NFATC3	Factor Nuclear De Células T Activadas, Citoplasmático 3	0.0074	0.001	Deletérea	1.000	D	0/1
51	14	56703797	A	G	MT1H	Metalotioneína -1H	8.237e-06	0.001	Deletérea	1	D	0/1
52	14	19617246	A	G	SLC47A2	Proteína 2 De Extrusión De Múltiples Fármacos Y Toxinas	9.982e-05	0.0	Deletérea	1	D	0/1
53	15	19645458	T	C	ALDH3A1	Aldehído Deshidrogenasa, Que Prefiere Nadp Dimérico	0.0001	0.004	Deletérea	1	D	0/1
54	15	48917355	C	T	WFIKK2	Proteína 2 Que Contiene El Dominio Wap, Kazal, Inmunoglobulina, Kunitz Y Ntr	0.0045	0.006	Deletérea	0.982	D	0/1
55	15	18647651	T	G	FBXW10	Proteína 10 Que Contiene Repeticiones F-Box / Wd	0.0053	0.011	Deletérea	0.800	D	0/1
56	16	63898308	C	T	CEP112	Proteína Centrosomal 112	0.0008	0.004	Deletérea	0.897	D	0/1
57	16	10487661	G	A	APCDD1	Proteína Apcdd1	8.271e-06	0.002	Deletérea	1	D	0/1
58	16	3071878	C	T	MYOM1	Myomesin-1	0.0086	0.005	Deletérea	1.000	D	0/1
59	17	16514698	A	G	EPS15L1	Sustrato 1 Del Receptor Del Factor De Crecimiento Epidérmico 1	2.477e-05	0.002	Deletérea	1.000	D	0/1
60	17	418954	G	A	SHC2	Proteína 2 Transformadora De Shc	0.0001	0.001	Deletérea	1.000	D	0/1
61	17	17671248	C	G	COLGALT1	Procolágeno Galactosiltransferasa 1	0.0008	0.001	Deletérea	1	D	0/1
62	17	1506220	G	A	ADAMTSL5	Proteína 5 Similar A Adamts	0.0019	0.044	Deletérea	0.995	D	0/1
63	17	27528605	G	A	TRIM54	Motivo Tripartito Que Contiene 54	6.032e-05	0.047	Deletérea	1.000	D	0/1
64	18	96687945	T	G	GPAT2	Licerol-3-Fosfato Aciltransferasa 2, Mitocondrial	6.55e-05	0.004	Deletérea	1	D	0/1

n	Crom	Posición	Ref	Alt	Gen	Nombre	Frecuencia (EXAC)	Puntaje	Predicción	Puntaje	Predicción	Genotipo
65	18	183095749	C	T	<i>PDE1A</i>	Fosfodiesterasa 1A De Nucleótidos Cíclicos 3', 5' Dependiente De Calcio / Calmodulina	0.0003	0.021	Deletérea	1.000	D	0/1
66	19	179641360	G	A	<i>TTN</i>	Titina	0.0063	0.011	Deletérea	1.000	D	0/1
67	19	215815807	C	A	<i>ABCA12</i>	Miembro 12 De La Subfamilia A Del Casete De Unión A Atp	0.0092	0.003	Deletérea	0.918	D	0/1
68	19	196765098	G	A	<i>DNAH7</i>	Cadena Pesada 7 De Dineína, Axonemal	0.0066	0.0	Deletérea	1	D	0/1
69	19	26750782	G	A	<i>OTOF</i>	Otoferlin	0.0084	0.002	Deletérea	1	D	0/1
70	20	7980487	C	T	<i>TMX4</i>	Proteína 4 Transmembrana Relacionada Con Tiorredoxina	0.0001	0.0	Deletérea	1.000	D	0/1
71	20	57599401	C	T	<i>TUBB1</i>	Cadena De Tubulina Beta-1	0.0044	0.0	Deletérea	1	D	0/1
72	20	62178594	C	T	<i>SRMS</i>	Srms De Tirosina-Proteína Quinasa	0.0013	0.0	Deletérea	1.000	D	0/1
73	20	33509608	T	C	<i>ACSS2</i>	Acetil-Coenzima A Sintetasa, Citoplasmática	0.0019	0.023	Deletérea	1	D	0/1
74	21	32639137	T	G	<i>TIAM1</i>	Invasión Del Linfoma T Y Proteína Inductora De Metástasis 1	8.244e-06	0.002	Deletérea	0.998	D	0/1
75	22	39122023	A	G	<i>GTPBP1</i>	Gtpbp1 - Proteína De Unión A Gtp 1	8.244e-06	0.023	Deletérea	1.000	D	0/1
76	22	19195771	G	A	<i>CLTCL1</i>	Cadena Pesada De Clatrina 2	0.0055	0.0	Deletérea	1	D	0/1
77	X	118603925	G	A	<i>SLC25A5</i>	Translocasa 2 De Adp / Atp	0.0004	0.032	Deletérea	1	D	0/1
78	X	118603958	G	A	<i>SLC25A5</i>	Translocasa 2 De Adp / Atp	0.0004	0.035	Deletérea	0.999	D	0/1
79	X	118604060	G	T	<i>SLC25A5</i>	Translocasa 2 De Adp / Atp	0.0025	0.016	Deletérea	1	D	0/1

Anexo 6. Inserciones con impacto moderado y alto de cuatro pacientes con cáncer gástrico del departamento de Nariño.

Inserciones SP03 (Ref: Referencia; Alt: Inserción; M: Impacto moderado de la variante; H: Impacto alto de la variante)									
N	Crom	Posición	Ref	Alt	Gen	Nombre	EXAC	SnEff	Genotipo
							Frecuencia	Predicción	
1	1	240370839	-	CCCCCT	FMN2	Formin-2	0.0063	M	0/1
2	3	49755642	-	CAGCGGCGG	AMIGO3	Proteína 3 Inducida Por Anfoterina	0.0011	M	0/1
3	5	79950708	-	GCAGCGGCCGCAG CGGCC	MSH3	Proteína De Reparación De Errores De Apareamiento De Adn Msh3	1,46E-02	M	0/1
4	6	43970503	-	GCG	C6orf223	Proteína No Caracterizada C6Orf223	0.0084	M	0/1
5	6	82461790	-	CGCCGCCGCCGA	FAM46A	Nucleotidiltransferasa Putativa Fam46A	4,51E-02	M	0/1
6	6	123102398	-	CCTCCTTCCTTCTT CTTCCT	FABP7	Proteína De Unión A Ácidos Grasos, Cerebro	0.0002	M	0/1
7	6	157100023	-	GGA	ARID1B	Proteína 1B Que Contiene El Dominio Interactivo Rico En At	0.0010	M	0/1
8	7	73477970	-	CTCGGAGGA	ELN	Elastina	0.0001	M	0/1
9	8	65493617	-	GGC	BHLHE22	Proteína Básica 22 De Hélice-Bucle- Hélice De Clase E	0.0079	M	0/1
10	8	73921283	-	GAGGAG	TERF1	Factor De Unión A Repetición Telomérico 1	0.0017	M	0/1
11	9	35906598	-	CCC	HRCT1	Terminal Carboxilo Rico En Histidina 1	0.0012	M	0/1
12	10	21805697	-	GGT	SKIDA1	Dominio Ski / Dach Que Contiene 1	0.0005	M	0/1

Inserciones SP03 (Ref: Referencia; Alt: Inserción; M: Impacto moderado de la variante; H: Impacto alto de la variante)									
N	Crom	Posición	Ref	Alt	Gen	Nombre	EXAC	SnPEff	Genotipo
							Frecuencia	Predicción	
13	11	73371884	-	GGTAAGTCT	PLEKHB1	Miembro 1 De La Familia B Que Contiene El Dominio De Homología De Pleckstrina	0.0021	M	0/1
14	14	74727547	-	GAG	VSX2	Homeobox 2 Del Sistema Visual	0.0032	M	0/1
15	15	41853508	-	CCTGAA	TYRO3	Receptor De Tirosina-Proteína Quinasa Tyro3	0.0053	H	0/1
16	15	89876827	-	TGCTGCTGC	POLG	Subunidad Gamma-1 De La Adn Polimerasa	0.0017	M	0/1
17	16	29821423	-	GCGGCA	MAZ	Proteína De Dedo De Zinc Asociada A Myc	0.0025	M	0/1
18	16	67208726	-	CCGGAG	NOL3	Proteína Nucleolar 3; Isoforma 1: Puede Estar Involucrada En El Empalme De Arn	0.0012	M	0/1
19	17	18025226	-	TCACTCCCCCAAG GATA	MYO15A	Miosina-Xv No Convencional	0.0013	M	Z0/1

Inserciones SP03 (Ref: Referencia; Alt: Inserción; M: Impacto moderado de la variante; H: Impacto alto de la variante)									
N	Crom	Posición	Ref	Alt	Gen	Nombre	EXAC	SnpEff	Genotipo
							Frecuencia	Predicción	
20	17	38975103	-	GCTGCCGCCGCCG TATCCGCCGCCG AGCT	KRT10	Queratina, Tipo I Citoesquelético 10	0.0022	M	0/1
21	19	3613264	-	TCGCCC	CACTIN	Cactina	0.0007	M	0/1
22		13988558	-	GAG	NANOS3	Nanos Homólogo 3	0.0036	M	0/1
23	20	3844992	-	AAG	MAVS	Proteína De Señalización Antiviral Mitocondrial	0.0045	M	0/1
24	21	32119309	-	GGT	KRTAP21-2	Proteína 21-2 Asociada A Queratina	0.0034	M	0/1
25	22	21212966	-	CCGCCT	PI4KA	Fosfatidilinositol 4-Quinasa Alfa	0.0057	M	0/1
26	22	28194912	-	TGC	MN1	Activador Transcripcional Mn1	0.0064	M	0/1
Inserciones SP05 (Ref: Referencia; Alt: Inserción; M: Impacto moderado de la variante)									
N	Crom	Posición	Ref	Alt	Gen	Nombre	EXAC	SnpEff	Genotipo
							Frecuencia	Predicción	
1	chrX	57618532	-	GGCGGCGGT	ZXDB	Proteína Ligada Al Cromosoma X Con Dedos De Zinc Zxdb	0.0088	M	1/1
2	chr2	186653522	-	GAT	FSIP2	Proteína 2 Que Interactúa Con La Vaina Fibrosa	0.0090	M	0/1
3	chr11	65366990	-	GGGAAGGGGGCGGCTC GGTCGGGC	MAP3K11	Proteína Quinasa 11 Activada Por Mitógeno	0.0002	M	0/1
4	chr19	38572367	-	GCCACC	SIPA1L3	Proteína 3 Similar A 1 Asociada A La Proliferación Inducida Por Señal	0.0025	M	0/1
5	chr22	22221708	-	CCG	MAPK1	Proteína Quinasa 1 Activada Por Mitógenos	0.0052	M	0/1

Inserciones SP06 (Ref: Referencia; Alt: Inserción; M: Impacto moderado de la variante; H: Impacto alto de la variante)									
N	Crom	Posición	Ref	Alt	Gen	Nombre	EXAC Frecuencia	SnEff Predicción	Genotipo
1	1	161494730	-	AGG	HSPA6	Proteína 6 De 70 Kda De Choque Térmico		M	0/1
2	7	26192217	-	CTG	NFE2L3	Factor 3 Relacionado Con El Factor Nuclear Eritroide 2		M	0/1
3	9	33395131	-	CATCTTCCTC	AQP7	Acuaporina-7		H	0/1
4	13	100622837	-	CGG	ZIC5	Proteína De Dedos De Zinc Zic 5		M	0/1
5	13	112722684	-	CGCACC	SOX1	Factor De Transcripción Sox-1		M	0/1
6	15	76632872	-	GCA	ISL2	Proteína Potenciadora Del Gen De La Insulina Isl-2		M	0/1
7		85383719	-	GGCCCA	ALPK3	Alfa-Proteína Quinasa 3		M	0/1
8	16	28354287	-	GAC	NPIP6	Miembro B6 De La Familia De Proteínas Que Interactúan Con El Complejo De Poros Nucleares		M	0/1
9	17	39240796	-	CTGCTGCATCTCCAGC TGCTGTCGCCCCAGCT GCTGTGTGTCCAGGTG CTGCAGGCCCCAGTGC TGCCAGTCTGTGTGCT GCCAGCCCAC	KRTAP4-7	Proteína 4-7 Asociada A Queratina		M	1/1

Inserciones SP07 (Ref: Referencia; Alt: Inserción; M: Impacto moderado de la variante; H: Impacto alto de la variante)									
N	Crom	Posición	Ref	Alt	Gen	Nombre	EXAC	SnEff	Genotipo
							Frecuencia	Predicción	
1	2	128390861	-	AGGTGG	MYO7B	Miosina VIIb	9,46E-03	M	0/1

2	3	195513398	-	GGTGACAGGAAGAGG GGTGGCATGACCTGTG GATACTGAGGAATTGT CGGTGACAGGAAGAG GGGTGGCGTGACCGGT GGATGCTGAGGAAGTG CT	MUC4	Mucina -4	0.0014	M	0/1
3	6	170871049	-	CAA	TBP	Proteína de unión a caja TATA	0.0020	M	0/1
4	7	27135316	-	TGG	HOXA1	Proteína Homeobox Hox-A1	0.0050	M	0/1
5	11	71249532	-	CCAGTCCAGCTGCTGT AAGCCCTGCTGCTGCT CTTCAGGCTGTGGGTC ATCCTGCTG	KRTAP5-8	Proteína 5-8 asociada a queratina	0.0008	M	0/1
6	14	44974878	-	GAGCCTCTTCTCTAG	FSCB	Proteína de unión a CABYR de vaina fibrosa;	1,73E-02	M	0/1
7	16	15457735	-	GCTGAGGGTAGA	NPIPA5	Miembro de la familia de proteínas de interacción del complejo de poros nucleares A5	6,12E-02	M	0/1
8	17	18576347	-	CGCCGCCGC	ZNF286B	Proteína de dedo de zinc supuesta 286B	0.0027	M	0/1
9	17	39254142	-	CAGCAGCTGGAGATGC AGCATCTGGGGCGG	KRTAP4-8	Proteína 4-8 asociada a queratina	0.0004	M	0/1
10	9	20414343	-	CTG	MLLT3	Proteína AF-9	2.87e-05	M	0/1