



UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS

UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS

FACULTAD DE MEDICINA

SECRETARIA DE DOCENCIA

JEFATURA DE POSGRADO



TESIS:

**CANCEROMA: DETECCIÓN MOLECULAR DE MARCADORES DE
CÁNCER POR SECUENCIACIÓN MASIVA DE TERCERA
GENERACIÓN**

PARA OBTENER EL GRADO DE MAESTRO EN:

MEDICINA MOLECULAR

QUE PRESENTA:

IBT LUIS URIBE ESPEJO GALICIA

CODIRECTORES DE TESIS:

DR. GABRIEL GUILLEN SOLÍS

DR. VERA LUCIA PETRICEVICH

Cuernavaca, Morelos, México a 28 de junio de 2019.

La tesis se realizó en la facultad de medicina de la universidad autónoma del estado de Morelos en el periodo de enero del 2017 al junio del 2019.

COLABORADORES

Directores y coordinadores del proyecto de tesis:

Dr. Gabriel Guillén Solís

Dra. Vera Lucia Petricevich

Comité tutorial:

Dra. Gabriela Castañeda Corral

Dr. Joaquín Sánchez Castillo

Dr. Omar Augusto Tello Terán

Jurado de examen

Dr. Joaquín Sánchez Castillo

Dr. Jesús Santa Olalla Tapia

Dr. Omar Augusto Tello Terán

Dr. Gabriel Guillen Solís

Dra. Vera Lucia Petricevich

Financiamiento:

GENOMAS SA DE C.V

Por la cobertura total de los gastos económicos de la tesis.

CONACTY

Beca nacional de Conacyt número 782222

AGRADECIMIENTOS

Me gustaría que estas líneas sirvieran para expresar mi más profundo y sincero agradecimiento a todas aquellas personas que con su ayuda han colaborado en la realización del presente trabajo, A mis profesores de la maestría por su apoyo y orientación, en especial a los doctores **Joaquín Sánchez Castillo** y **Jesús Santa Olalla Tapia** por el tiempo invertido en mí, a mis directores de tesis, el doctor **Gabriel Guillen Solís** y la doctora **Vera Lucia Petricevich** por la orientación, el seguimiento y la supervisión, pero sobre todo por la motivación y el apoyo recibido por el tiempo que duro la maestría.

Especial reconocimiento merece el interés mostrado por mi trabajo y las sugerencias recibidas por la Máster en Medicina Molecular **Laura C. Zárrega Vargas**, con la que me encuentro en deuda por el ánimo infundido y la confianza en mí depositada.

Quisiera hacer extensiva mi gratitud a mis compañeros y amigos de la Maestría de Medicina Molecular de la facultad de medicina de la **AUEM**.

También quiero dar las gracias al Máster en biotecnología **José Francisco Pulido Baraja** y la empresa **Genomas**, por su apoyo para realizar la secuenciación y estandarización de la técnica.

Un agradecimiento muy especial merece la comprensión, paciencia y el ánimo recibidos de mi familia y amigos. Sobre todos a mis padres **María Florencia Galicia Herrera** y **Uribe Espejo Ruiz** por apoyarme por esta decisión, a mi hermano **José Roberto Espejo Galicia** por apoyarme en los momentos difíciles, a mis hermanos del alma, **Arnulfo Sánchez Félix** y **Estefany Jazmín Sánchez Félix** por sus consejos y convivencia que pasamos juntos.

Por último, a la mujer más especial que tengo el placer de haber conocido, a mi prometida **Ania Itzel López Cobos** por todo el apoyo, tiempo, desvelos, esperanzas y sueños invertido en mis durante el tiempo que duro esta fase de mi vida, por todo esto, lo único que puedo decir es muchas gracias.

A la universidad del Papaloapan por su apoyo y colaboración para usar sus equipos para poder realizar el análisis de los datos.

Gracias a la beca **CONACYT** con número 782222, por su apoyo para realizar la tesis.

GLOSARIO

Biomarcadores genéticos: secuencia de ADN que nos revela una característica diferencial entre dos o más individuos, permitiendo así distinguir fenotipos. En epidemiología un biomarcador permite relacionar un fenotipo con un estado biológico que puede ser beneficioso o perjudicial a la persona.

Cáncer: Proceso de crecimiento y diseminación incontrolados de células que se asocia a cambios genéticos o epigenéticos.

Etiología: Ciencia encargada del estudio y análisis de la causa o procedencia de las enfermedades.

Exoma: Conjunto total de exones de la célula, es decir las regiones de ADN que codifican proteínas o péptidos. El exoma corresponde aproximadamente al 1-2% del genoma.

Genoma: La totalidad de la información genética contenida en una célula u organismo. En células eucariotas el genoma está constituido por ADN nuclear (conjunto de cromosomas) y el ADN extranuclear (ADN mitocondrial).

Incidencia: término en epidemiología que define el número de casos nuevos de una patología en una población en un periodo determinado. Generalmente se expresa como el número absoluto de casos nuevos de la enfermedad en un año o el número de casos nuevos por 100.000 personas por año.

SNP: Deriva del inglés: “Single Nucleotide Polymorphism”. Un SNP es una variación genética entre personas en la cual una secuencia de ADN difiere en un solo nucleótido.

Secuenciación: Conjunto de métodos y técnicas bioquímicas cuya finalidad es determinar el orden de los nucleótidos en el ADN o ARN de una muestra.

ABREVIATURAS

AMP: Asociación de Patología Molecular

ARN-seq: Técnica de secuenciación de todo el ARN en una célula

CDC: Centros para el Control y Prevención de Enfermedades (en inglés: Centers for Disease Control and Prevention)

CNE: Centro Nacional de Epidemiología de España

CISH: Hibridación in-situ cromogénica

FISH: Hibridación fluorescente in situ

HGP: Proyecto del genoma humano (en inglés: Human Genome Project)

ICGC: Consorcio Internacional del Genoma del Cáncer (en inglés: International Cancer Genome Consortium)

IHC: Inmunohistoquímica (en inglés: Immunohistochemistry)

Kb: Kilobase

miR-rSNP: polimorfismos de nucleótido únicos regulador de los microRNA

MLPA: Amplificación múltiple por ligación dependiente de sonda (en inglés: Multiplex ligation-dependent probe amplification)

mRNA: Ácido ribonucleico mensajero

NCI: Instituto nacional del cáncer (en inglés National Cancer Institute)

NGS: Secuenciación de nueva generación (en inglés: Next Generation Sequencing)

rsSNP: polimorfismos de nucleótido únicos no sinónimo

TCGA: Atlas de genoma del cáncer (en inglés: Cancer Genome Atlas)

TLA: Amplificación localizada de un locus (en inglés: Targeted Locus Amplification)

TNM: Clasificación TNM de tumores malignos (en inglés: TNM Classification of Malignant Tumours)

OMIM: Herencia mendeliana del hombre en línea (en inglés: Online Mendelian Inheritance in Man)

OMS: Organización Mundial de la Salud

OPS: Organización Panamericana de la Salud

pb: Par de bases, dos nucleótidos complementarios en cadenas de ADN opuestas

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

rSNP: polimorfismos de nucleótido únicos regulador

SEOM: Sociedad Española de Oncología Médica

SNP: Polimorfismo de nucleótido único

srSNP: polimorfismos de nucleótido únicos estructurales

sSNP: polimorfismos de nucleótido únicos sinónimo

WES: Secuenciación completa del exoma (en inglés: Whole exome sequencing)

WGS: secuenciación completa del genoma (en inglés: Whole genome sequencing)

ÍNDICE

GLOSARIO	4
ABREVIATURAS	5
ÍNDICE DE FIGURAS	11
ÍNDICE DE TABLAS	13
ÍNDICE DE ECUACIONES	14
RESUMEN	15
INTRODUCCIÓN	16
Genética del Cáncer	18
Epidemiología	19
Clasificación de los cánceres	22
Carcinoma	22
Leucemia	23
Sarcoma	23
Linfoma	23
Mielomas	24
Tipos mixtos	24
Factores de riesgo	25
Factores conductuales y dietéticos	25
Factores químicos	26
Factores físicos	27
Factores patogénicos	27
Factores biológicos	28
ETIOLOGÍA	30
Cambios Genéticos	31

Secuenciación para el diagnóstico de SNP	32
Exomas	34
PLANTEAMIENTO DEL PROBLEMA.....	35
OBJETIVOS.....	37
GENERAL.....	37
PARTICULAR	37
HIPÓTESIS	37
DESCRIPCIÓN DEL TIPO DE ESTUDIO.....	37
MATERIAL Y MÉTODOS	38
BÚSQUEDA DE MUESTRAS BIOLÓGICAS.....	38
Determinación del número de muestras para la validación de la técnica	38
OBTENCIÓN DE MUESTRA BIOLÓGICAS	38
Criterios para la obtención de muestra y su procesamiento	39
PROCESAMIENTO DE LA MUESTRA.....	39
ANÁLISIS DE BASES DE DATOS BIOLÓGICAS.....	39
Base de datos de genes asociados a diferentes tipos de cáncer	39
Desarrollo de una base de datos de genes relacionado con cáncer	39
Polimorfismos asociados a cáncer.....	40
Posición cromosómica de cada una de las mutaciones.	40
Diseño	40
Estudio biológico de genes.....	40
Modelo booleano.....	41
Circus plot	41
PCR para la secuenciación.....	41
Confirmación de la PCR	42

Diseño de los oligonucleótidos para la validación	42
SECUENCIACIÓN	44
Preparación de los productos de PCR para las bibliotecas	44
Validación	45
Validación para la técnica	45
Matriz de validación del equipo de secuenciación	46
ANÁLISIS DE LOS DATOS	47
Índice de calidad.....	47
Identificación de variantes	47
Rango de detección	48
Comparación de dos diferentes secuenciaciones	48
ANÁLISIS ESTADÍSTICO MEDIANTE CURVAS ROC	49
RESULTADOS	51
BÚSQUEDA DE MUESTRAS BIOLÓGICAS.....	51
Determinación del número de muestras.....	51
OBTENCIÓN DE MUESTRA BIOLÓGICAS	51
BÚSQUEDA BIBLIOGRÁFICA DE GENES	52
Base de datos de genes relacionados con cáncer	53
Interacciones de los genes con el cáncer.....	53
Análisis de los genes involucrados con SNP relacionados con alguno tipo de neoplasia.....	55
Obtención de polimorfismos relacionado con cáncer	56
SNP asociados con una enfermedad	59
Posición cromosómica de cada una de las mutaciones.	62
PCR.....	64
Diseño de oligonucleótidos para la validación de la técnica.....	65

SECUENCIACIÓN	68
Análisis de las secuenciación y validación de la técnica.	69
Secuenciación	69
Análisis de costos.....	74
Circos Plot.....	75
Modelo booleano.....	76
DISCUSIÓN	77
CONCLUSIONES.....	85
REFERENCIAS BIBLIOGRÁFICAS	87
Anexo 1 Consentimiento informativo	99
Anexo 2 estudio del paciente.....	100

ÍNDICE DE FIGURAS

Figura 1. Esquema de regulación del ciclo celular. Los factores externos e internos son los causantes de direccionar a la célula.....	19
Figura 2. Infrograma de predicción de números de incidencia estimados para el 2040 mediante el algoritmo de la OMS (GLOBLAN)	20
Figura 3. Incidencia de cáncer a nivel mundo en el 2018	21
Figura 4. Estimaciones del número de incidencia de cáncer en México en el 2018, obtenido de los datos de mortalidad de cáncer obtenida por el GCO por el programa Globocan.	21
Figura 5. Infograma de la incidencia geográfica de Cáncer a nivel mundial 2018	22
Figura 6. Infografía de la localización de los canceres con su clasificación histológica en el cuerpo humano.	24
Figura 7. Distribución de casos nuevos de cáncer a nivel mundial por rango de edad ¡Error!	
Marcador no definido.	
Figura 8. Número de genes que expresa en sangre en cada cromosoma.....	52
Figura 9. Porcentaje de genes que se expresa por cromosoma detectable en sangre	53
Figura 10. Mapa de interacciones entre los genes y tipos de cáncer.	55
Figura 11. Análisis de las propiedades que tiene los genes relacionados con cáncer que se expresa en sangre.....	56
Figura 12. Número de SNP estudiado por año validados por análisis Gwas	57
Figura 13. Equipos de secuenciación utilizados para la identificación de SNPs en genomas humanos.....	58
Figura 14. Análisis GWAS por población de estudio.	59
Figura 15. Diagrama de Veen de las mutaciones presente en el genoma humana	60
Figura 16. Posición y efecto que tiene cada uno de los SNP relacionados con cáncer	62
Figura 17. localización cromosómica de los SNP relacionados con 31 cáncer en el genoma humano.	63
Figura 18. Integridad del ADN genómicos de las muestras de sangre corrida en un gel de agarosa del 1.5%.....	64
Figura 19. Barrido de productos de PCR con los oligonucleótidos de genomas.....	65

Figura 20. Gel de Agarosa 2% de los 6 SNP.....	66
Figura 21. Validación de la PCR en formato Multiplex mediante qPCR.....	66
Figura 22 Controles negativos de la reaccion.....	67
Figura 23. Comparación de calidad de las lecturas entre MinION.	70
Figura 24 Análisis de las lecturas que contiene el SNP rs121913327.....	72
Figura 25 Curva ROC para la secuenciación por MinION en comparación con la secuenciación de ilumina por Hiseq.....	74
Figura 26 Circus plot para la interacción de los cánceres con los cromosomas.....	75
Figura 27 Modelo Booleano.....	76

ÍNDICE DE TABLAS

Tabla 1 Factores de riesgo para el desarrollo de cáncer.....	25
Tabla 2 Estadístico de los principalmente factores de riesgo conductuales en la población mexicana en el año 2013 realizado por la OMS.....	26
Tabla 3 Cánceres ocupacionales y las sustancias químicas que produce el cáncer.....	27
Tabla 4 Genes relacionados con algún tipo de cáncer y son utilizados como biomarcadores de cáncer aceptados por la OMS	30
Tabla 5 Clasificación de los SNP	32
Tabla 6 Técnicas de diagnóstico que se han diseñados que se encuentra en mercado.....	35
Tabla 7 Preparación de la reacción de PCR	42
Tabla 8 Parámetros de reacción de la PCR.....	42
Tabla 9 Protocolo de Preparación de la reacción de PCR para la validación.....	43
Tabla 10 Parámetros de reacción de la PCR.....	43
Tabla 11 Preparación de la reacción de qPCR para validación	44
Tabla 12 Parámetros de reacción de la qPCR para validación	44
Tabla 13 Matriz de procesamiento de las muestras	46
Tabla 14 Cuadro 2x2 para el análisis de sensibilidad y especificidad de técnicas de diagnóstico.....	48
Tabla 15 Parámetros que se calcula con la comparación de dos técnicas	49
Tabla 16 Escala de area bajo la curva de ROC.....	50
Tabla 17 Total de genes que se expresa en sangre en el genoma humano.	54
Tabla 18 Cáncer que se ha determinados SNP en estudios Gwas	61
Tabla 19 Oligonucleótidos para PCR y qPCR.....	65
Tabla 20 Microorganismo detectado en la muestra de heces fecales	68
Tabla 21 SNP con relevancia clínica en cáncer que son detectable por el protocolo diseñado por la tesis.....	71
Tabla 22 Resultados detección de SNP	73
Tabla 23 Parámetros que se calcula con la comparación de dos técnicas	73

ÍNDICE DE ECUACIONES

Ecuación 1 Distribución Binomial	51
Ecuación 2 Formula simplificados de la Distribución Binomial.....	51

RESUMEN

El cáncer es un conjunto de enfermedades causadas por el crecimiento y diseminación incontrolado de células por cambios en la regulación de su ciclo, otorgándoles ventajas para su sobrevivencia en comparación de la familia celular de origen, el cambio en regulación de su ciclo es principalmente de carácter genético¹. Para el año 2015 el cáncer era el causante del 16% de las muertes a nivel mundial, siendo la segunda causa de muerte². La detección de cambios en las secuencias ADN ha ayudado a establecer las bases moleculares del cáncer y sus posibles etiologías. Gracias al desarrollo de la tecnología de secuenciación de ADN desde los años 70³, se han generado un gran número de bases de datos^{4,5} y se ha acelerado de manera exponencial el análisis de las variaciones genéticas entre individuos⁶. Las variaciones más estudiadas son los polimorfismos de una sola base (SNP). El análisis de los SNP ha permitido introducir los principios para una medicina personalizada^{7,8}. El estudio de las variaciones genéticas ha tenido un impacto en la oncología⁹, porque se ha utilizado para el diagnóstico preventivo¹⁰. La técnica más utilizada para diagnóstico preventivo es por secuenciación de exomas, debido a su precio relativamente accesible. En el caso de México, la técnica se ha aplicado con éxito, pero hace falta validar un panel de SNPs relacionados con la incidencia de los diversos cánceres para precisar biomarcadores en la población mexicana. Los objetivos de la presente tesis fueron diseñar y validar una técnica de secuenciación de muy alto rendimiento para detectar simultáneamente un gran número de biomarcadores de cáncer. Para este fin, se realizó una minería de datos para definir los SNP relacionados con alguna neoplasia. Se identificaron un total de 57,181 artículos relacionado con SNP, de los cuales solamente 43,780 SNP tienen una importancia clínica, solo 7,628 SNP esta relacionados con algún tipo de cáncer distribuidos en 1468 genes, se diseñó un panel de 4439 SNP de los 7628 a partir de los criterios de la tesis. Se identificaron un total de 3456 SNP, se comparó la SNP detectados con la técnica diseñada con una comercial y se obtuvo una exactitud de 98.06%, lo que permite usar la técnica como una propuesta para el diagnóstico preventivo. La técnica diseñada tiene un costo de procesamiento de muestra de 725 dólares, siendo elevado en comparación con la técnica comercial, dando como resultado una técnica que presenta futuro para su posible aplicación como una técnica de diagnóstico preventivo para población mexicana, que requiere automatización para ser una técnica competitiva, eficaz y robusta para detectar varios biomarcadores de cáncer.

INTRODUCCIÓN

El término cáncer designa agrupadamente un conjunto de enfermedades que son causadas por el crecimiento y diseminación incontrolada de células por alteración en su ciclo celular. El cáncer ocurre mayoritariamente debido a cambios en el genoma celular¹. Dichos cambios genómicos son multifactoriales¹¹⁻¹³, pero la mayoría de los factores convergen en la formación de radicales libres¹⁴, que causan daño genético¹⁵. Se considera que, debido a los nuevos estilos de vida, que conllevan bajo consumo de vegetales, alta ingesta de grasas y carbohidratos, cambios demográficos y ambientales han repercutido en la incidencia de cáncer¹⁶, ocasionado un aumento en más del 50% en las últimas tres décadas¹⁷.

El cáncer es la quinta causa de muerte a nivel mundial, en el 2012 causó 8,201,575 defunciones, lo que representa un 16% de todas las muertes por enfermedad a nivel mundial². El cáncer es una patología que inicia cuando una célula o células presenta daño o alteración en su genoma de manera significativa¹⁸, lo que desencadena un conjunto de señales afectan los mecanismos de regulación del ciclo celular y la célula empieza a proliferar de manera incontrolada^{1,19,20}. Al mismo tiempo, la célula liberan sustancias y realizan cambios estructurales en la superficie de su membrana, limitando la respuesta del sistema inmune para detectar y destruir a la célula, lo cual promueve el proceso de neoplasia²¹. Cuando la célula entra en un estado de proliferación incontrolada, requiere espacio, materia y energía, que las células cancerosas obtienen de las células adyacentes¹⁹, lo que debilita y daña la integridad del tejido donde se manifestó, y en muchas ocasiones el daño es irreparable causando la muerte de las células aledañas.

Muchos factores pueden influir en el desarrollo del cáncer, principalmente factores exógenos, como el hábito de fumar, consumo de alcohol, costumbres de alimentación, infecciones bacterianas y virales, y los factores endógenos que predisponen a la persona a la enfermedad por su estado fisiológico, mental y genético²². Cada factor tiene un efecto diferente dependiendo al tipo de cáncer, en el caso de infecciones por virus del papiloma, predispones a las mujeres a ser más susceptibles al cáncer de matriz y de mama²³.

Para varios tipos de cáncer se ha detectado factores endógenos genéticos sobre todo en cáncer de temprana aparición, como la leucemia²⁴, este tipo de factores genéticos, por ser invariables pueden ser usados como biomarcadores de predisposición al cáncer.

La detección de biomarcadores para la predisposición al cáncer tiene más de 40 años que se realizan²⁵. Los primeros análisis se realizaban por técnicas como PCR y RFLP^{4,5}, la utilización de la PCR fue pionera para el diagnóstico personalizado, pero presentaba la limitante que solo se podría detectar entre 1 a 6 biomarcadores por estudio, esta forma de análisis resultó insuficiente para realizar un diagnóstico de manera precisa³, debido a gran número de variaciones genéticas que se han asociado por lo menos con algún tipo de neoplasias²⁶. En la actualidad las tecnologías de secuenciación de tercera generación (NGS), han empezado a ser utilizados para la detección de SNP asociados a enfermedades genéticas, principalmente en tumores⁶, pero todavía existe muchas limitantes para que la detección de biomarcadores a nivel mundial, una de las principales limitante es que se tiene que desarrollar para cada tipo de población debidos a que los factores étnicos influye en la propensión para desarrollar algún tipo de cáncer, y la razón más importantes es que hasta el momento no existe ningún estudio o herramienta matemática que permite ver el efecto acumulativo que se tiene cuando se presenta varios biomarcadores, lo que ha limitado la utilidad de varios marcadores genéticos para realizar un mejor diagnóstico preventivo²⁵. Para contribuir a que el diagnóstico preventivo se vuelva realidad, proponemos diseñar un panel de biomarcadores para población mexicana, para que posteriormente puede ser usada como herramienta para detección certera y confiable en la población mexicana.

Genética del Cáncer

El cáncer constituye uno de los problemas de salud más importantes del mundo, debido a su alto nivel de mortalidad y costo de recuperación²⁷. El origen de un Cáncer depende del tipo y los factores exógenos y endógenos, pero de manera general una célula neoplásica se produce por una alteración genética, que tiene como consecuencia una reprogramación de sus funciones, lo que desencadena el tumor¹. El origen de la alteración genética puede deberse a muchos factores, por lo que se considera una enfermedad heterogénea, la dificultad que se tiene para determinar el factor detonante de una neoplasia dificultando la identificación de la etiología para cada tipo de cáncer⁸.

La alteración que causa que la célula se desvincule de su función, a una célula neoplásica puede ocurrir en cualquier parte del genoma, pero las alteraciones de mayor impacto ocurren en los genes de reparación de daño al ADN, en los genes responsables de la regulación del ciclo celular, genes de señalización y de la diferenciación celular¹.

Se han clasificados tres tipos de genes donde las alteraciones genéticas tiene un alto impacto para la generación de las células neoplásicas:

Oncogenes: son los genes que regula la proliferación celular, cuando se encuentra una alteración en su secuencia de ADN, la célula no controla de manera eficiente al ingreso a la fase G0 en el ciclo celular, lo que permite que la célula se divida de manera continua e independiente a las células de sus tejidos de origen^{28,29}.

Genes Supresores de Tumores: son los genes que tiene como función controlar el ciclo celular, cuando ocurre una mutación en algún gene supresor de tumor, las proteínas responsables del control del ciclo celular, como las ciclinas, no funcionan de manera adecuada, permitiendo a la célula multiplicarse sin restricción en los puntos normales de regulación³⁰.

Genes de Reparación del ADN: son los genes que tienen la función de asegurar la integridad de la información genética de la célula, las mutaciones en estos genes implican una replicación y reparación del DNA ineficaces, lo que promueve que las mutaciones se acumulen en la célula, aumentando la posibilidad que las mutaciones afecten a los oncogenes y genes supresores de tumor³¹.

Las alteraciones genéticas se van acumulando de generación en generación, cuando la célula tiene un número determinado de alteraciones en su genoma, entra a un punto de regulación, en este punto la célula puede tomar alguna de tres posibles rutas; la primera es la apoptosis³², que es un proceso fisiológico de eliminación celular regulado por una secuencia de cambios que se producen dentro y fuera de la célula, afectando la integridad de la célula e indicando a las células fagocíticas próximas que la célula inició su proceso de apoptosis³³. La apoptosis es un mecanismo de protección de las células vecinas que ayuda a preservar la vida del organismo (Figura 1). El segundo camino que puede tomar la célula es proseguir con su ciclo biológico normal, es tomado cuando los mecanismo celulares determinaron que las alteraciones genómicas presente en la célula no tiene un efecto negativo en la vida de del organismo³⁴, y la tercera ruta, es cuando las alteraciones acumuladas afectan o inhibe los mecanismos de regulación³⁵, debidos a estas alteraciones la célula puede crecer y proliferar a su propio ritmo³⁴, cuando la célula toma el tercer camino, empieza de desencadenar un conjunto de cambios en las células adyacentes para poder obtener los recursos necesarios para su rápido consumo de nutrientes³⁶⁻³⁸.

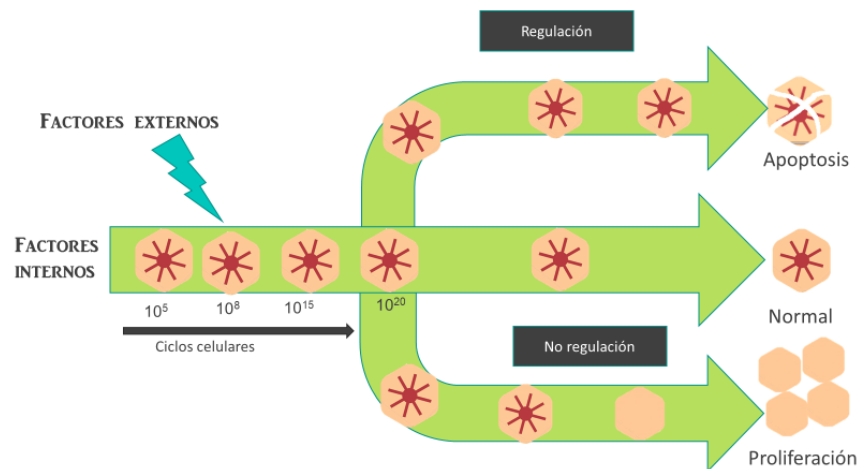


Figura 1. Esquema de regulación del ciclo celular. Los factores externos e internos son los causantes de direccionar a la célula.

Epidemiología

El cáncer es una de las principales patologías que afectan a la población a nivel mundial hoy en día debidos a su amplio distribución³⁹⁻⁴², la OMS reportó que en el 2018 se presentaron 18,078,957 casos nuevos de cáncer en el mundo (Figura 2), el algoritmo de la OMS estima

que la incidencia de Cáncer en el 2040 aumentara alrededor de un 70%, llegando a un estimado de 29,532,994 caso de Cáncer⁴².

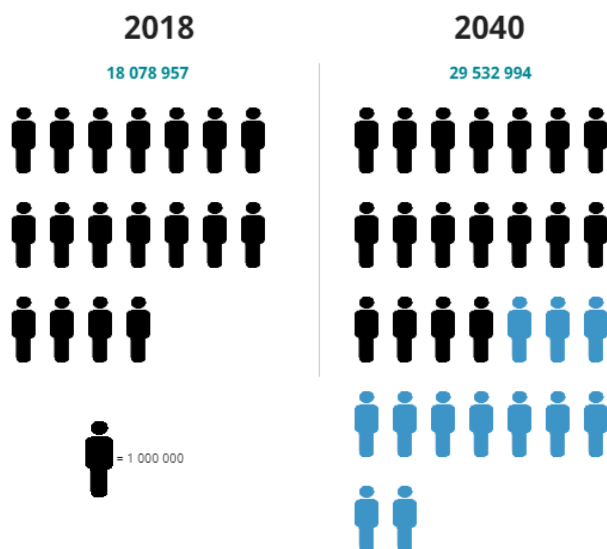


Figura 2. Infrograma de predicción de números de incidencia estimados para el 2040 mediante el algoritmo de la OMS (GLOBLAN).

Los tipos de cáncer con mayor incidencia a nivel mundial son los carcinomas, sobresaliendo los cánceres de pulmón, seno y colorrectal, abarcando aproximadamente el 30% de la incidencia de cáncer (Figura 3). El cáncer tiene el potencial de afectar a cualquier persona sin importar su edad, sexo u origen étnico, pero sí puede haber una mayor incidencia en ciertos grupos de personas que tienen características genéticas o factores que benefician a desarrollarse algunos tipo de cáncer (Figura 5), lo que clarifica la heterogeneidad en la incidencia y mortalidad mundial⁴³. Los factores que se considera que generan la heterogeneidad son; el nivel económico, factores culturales y el origen genético⁴⁴. Los 3 factores son la razón por la que las regiones de África, Asia, América Central y Sudamérica concentran más del 60% de los casos y 70% de las defunciones por cáncer⁴⁵.

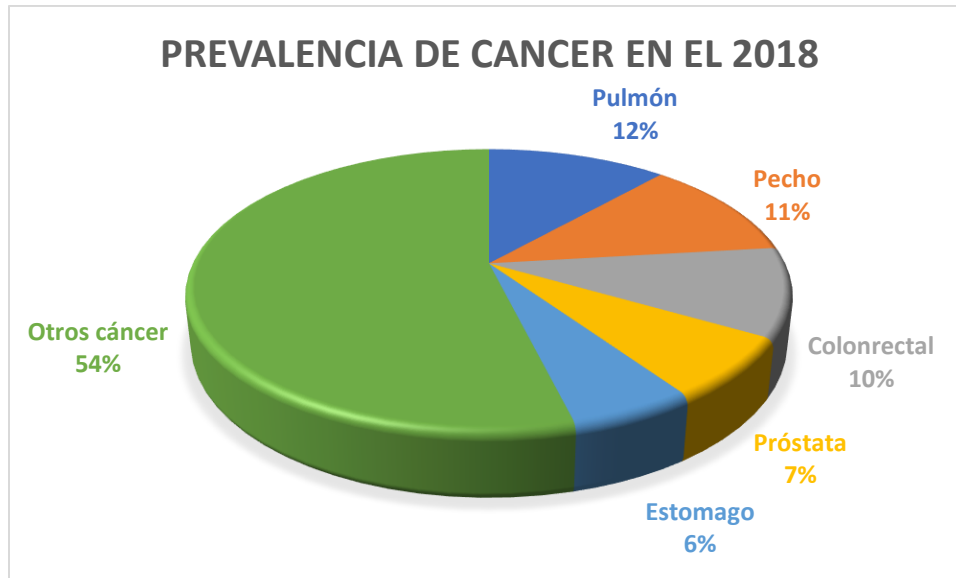


Figura 3. Incidencia de cáncer a nivel mundo en el 2018. Fuente OMS

Lamentablemente México no cuenta con un registro fidedigno de incidencia nacional de cáncer, sólo cuenta con estimaciones obtenidas mediante estudios de interpolación entre la mortalidad reportada y las estadísticas de los países vecinos¹⁷. El tipo de cáncer que se ha reportado con mayor incidencia en México parece tener una tendencia muy similar que el resto de mundo, siendo de mayor incidencia los que se presentan en mama, próstata y en cerviz uterino (Figura 4).

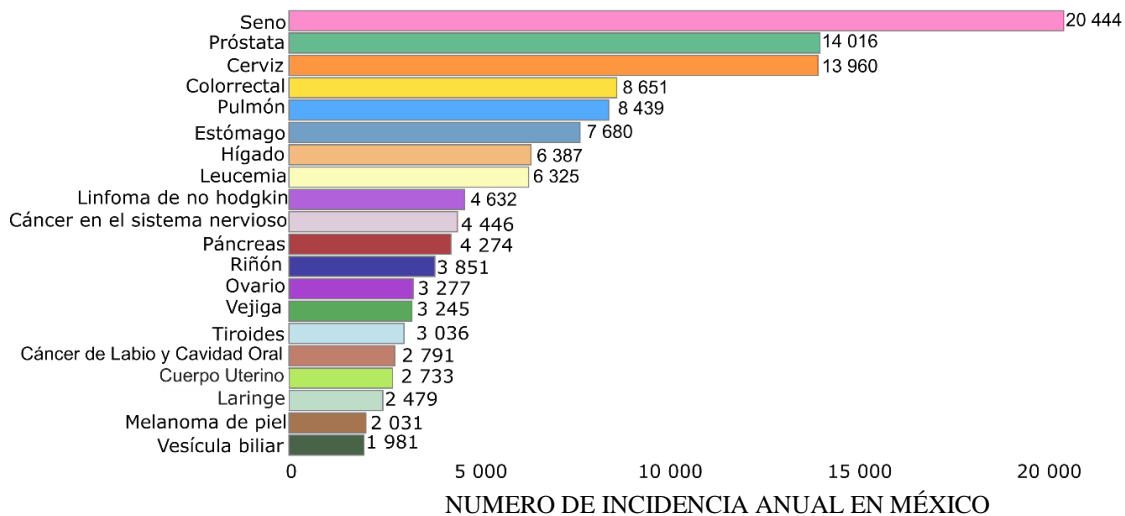


Figura 4. Estimaciones del número de incidencia de cáncer en México en el 2018, obtenido de los datos de mortalidad de cáncer obtenida por el GCO por el programa Globocan.

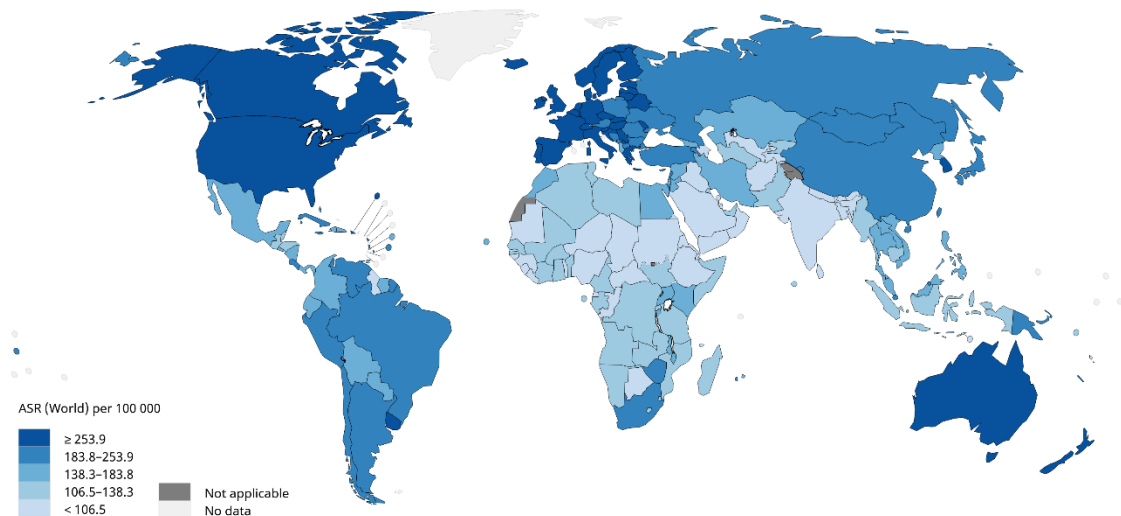


Figura 5. Infograma de la incidencia geográfica de Cáncer a nivel mundial 2018. Fuente OMS

En México se estima que durante el año 2018 se presentaron un total de 147 985 casos nuevos de cáncer, provocando 78 719 defunciones, lo cual le hace una enfermedad con alta tasa de mortalidad. El instituto de cáncer estima que mueren 936 personas por hora, por algún tipo de cáncer, a nivel mundial, de los cuales a México le corresponden 9 casos, lo que lo vuelve un problema de salud en el país⁴⁶.

Clasificación de los canceres

Además de designar a los cánceres por el órgano en el que se desarrolla inicialmente la patología, los canceres son clasificado histológicamente de acuerdo con el tejido en el que inicia la neoplasia (Figura 6). Desde el punto de vista histológico los cánceres son agrupados en 6 categorías: Carcinoma, Leucemia, Melanoma, Sarcoma, Linfoma, Mieloma y Mixto.

Carcinoma

Los carcinomas son los canceres que se originan en las células del epitelio y son el tipo de cáncer más diversificado y frecuente en la población mundial siendo un poco más del 70% de todos los cánceres que se reportan. La alta incidencia de los carcinomas posiblemente es porque las células epitelio están expuestas al contacto directo con los factores de riesgo del medio ambiente, adicionado a ello las células epiteliales tienen una alta tasa de recambio en comparación con los otros tipos celulares. La alta replicación celular aumenta la posibilidad

que la célula adquiera mutaciones y pierda la capacidad de autorregularse volviéndose cancerosa. El carcinoma se subclasifica en adenocarcinoma, que son los cánceres que se desarrollan en el epitelio de órganos y glándulas, y los carcinomas de células escamosas, que son las neoplasias en las células de la capa más externa del epitelio^{47,48}. Entre los carcinomas con mayor incidencia se encuentran el cáncer de mama, el cáncer de próstata, el cáncer de colon, y el cáncer de pulmón⁴⁹.

Leucemia

La leucemia se origina en células de la médula ósea, en este tipo de cáncer se observa un aumento descontrolado del número de leucocitos⁵⁰, que proliferan en forma de células inmaduras, disminuyendo su capacidad de proteger contra patógenos externos, y por lo tanto los pacientes con leucemia presentan infecciones recurrentes, las cuales son responsables de más del 60% de las muertes relacionadas con esta enfermedad⁵¹. La leucemia es el cáncer infantil más común en todo el mundo^{51,52}, y la leucemia linfocítica aguda representa aproximadamente 72% de casos de leucemia infantil⁵³.

Sarcoma

El término Sarcoma se refiere al cáncer que se origina en los tejidos de soporte y conjuntivo, como los huesos, tendones, cartílagos, músculos y grasa⁵⁴. Existen dos tipos principales de sarcoma: sarcomas de tejidos óseos y blandos. Los sarcomas son los cánceres de menor frecuencia en adultos, se estima que los sarcomas son aproximadamente un 1% de los cánceres diagnosticados a nivel mundial en personas mayores de 18 años y es entre el 15-21% en niños⁵⁵.

Linfoma

Los Linfomas se desarrollan en las glándulas o nodos del sistema linfático, una red de vasos, nódulos y órganos (específicamente el bazo, las amígdalas y el timo) que purifican los fluidos corporales y producen glóbulos blancos o linfocitos que combaten las infecciones. Los linfomas también pueden ocurrir en órganos específicos como el estómago, la mama o el cerebro. Estos linfomas se denominan linfomas extranodales. Los linfomas se subclasifican en dos categorías: linfoma de Hodgkin y linfoma no Hodgkin⁵⁶.

Mielomas

El Mieloma es el cáncer que se origina en las células plasmáticas (glóbulos blancos que producen anticuerpos) de la médula ósea. Este es el segundo cáncer sanguíneo más común en América y representa el 1% de casos de cáncer a nivel mundial. En los pacientes que sufren mieloma su cuerpo produce demasiadas células plasmáticas anormales en la médula, aumentando la concentración de proteína M en la médula ósea. El principal perjuicio que causa este tipo de cáncer es que daña a los órganos encargados de limpiar y purificar la sangre como el riñón^{57,58}.

Tipos mixtos

Existen otros tipos de tumores que se originan en tejidos no mencionados previamente, pero por ser poco frecuentes no son muy estudiados (corresponde entre todos a menos del 2% de la incidencia de todas las neoplasias). Los tumores cerebrales (glioma), neuroendocrinos y de células germinativas son cánceres que conforma esta categoría.

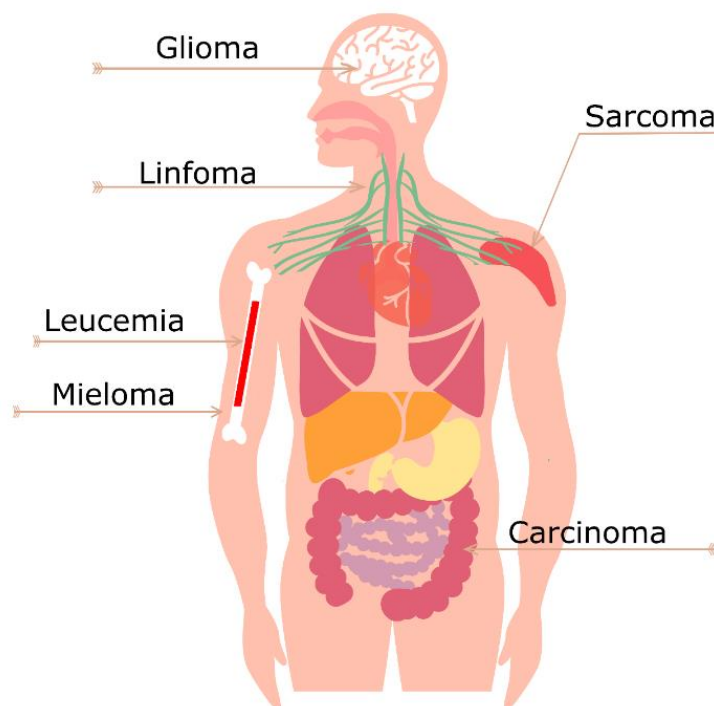


Figura 6. Infografía de la localización de los cánceres con su clasificación histológica en el cuerpo humano.

Factores de riesgos

Todos los seres humanos tenemos la posibilidad de desarrollar alguno tipo de cáncer, la posibilidad depende de los factores de riesgo que aumenta o predispone a la persona a desarrollar la enfermedad, estos factor puede ser exógenos o endógenos^{59,60}.

Los factores exógenos y endógenos, dependiendo de sus origen, se pueden clasificar en 5 grupos⁶¹; factores conductuales y dietéticos, químicos, físicos, biológicos y patológicos (Tabla 1). Actualmente no existe un diagnostico o algoritmo que permita entender la contribución de cada uno de los factores, debidos a que cada factor influyen de manera diferente para el desarrollo de la enfermedad⁶².

Tabla 1 Factores de riesgo para el desarrollo de cáncer.

Grupo	Ejemplos
Físicos	<ul style="list-style-type: none">• Radiaciones ultravioletas
Químicos	<ul style="list-style-type: none">• Amianto• Aflatoxinas• Arsénico• Radicales libres• Benzopireno• Formaldehído
Biológicos	<ul style="list-style-type: none">• Disposición genética• Envejecimiento• Hormonas
Patologías	<ul style="list-style-type: none">• Enfermedades cardiovasculares• Diabetes• Enfermedades respiratorias crónicas• Virus• Bacterias• Parásitos
factores conductuales y dietéticos	<ul style="list-style-type: none">• Índice de masa corporal• Ingesta reducida de frutas y verduras• Falta de actividad física• Consumo de tabaco y alcohol

Fuente: OMS y ONC

Factores conductuales y dietéticos

Los factores conductuales y dietéticos son los factores de riesgo que parecen influir más en el desarrollo de cáncer, estos factores están relacionados con el estilo de vida del paciente; índice de masa corporal, ingesta reducida de frutas y verduras, falta de actividad física,

consumo de tabaco y alcohol. Se estima que más de un tercio de todas las muertes por cáncer son provocados por esta última categoría⁵⁹, el tabaco es el principal factor de riesgo de cáncer reportado (Tabla 2), se estima que es responsable aproximadamente el 22% de las muertes por cáncer en todo el mundo¹¹. Otro factor de riesgo conductual es el consumo de alcohol⁶³, se ha reportados que el alcohol está relacionado con el 3,5% de todas las muertes por cáncer en América, el efecto de alcohol tiende a aumentar el riesgo de cáncer porcentualmente mientras más tiempo se consuma⁶⁴.

Tabla 2 Estadístico de los principalmente factores de riesgo conductuales en la población mexicana en el año 2013 realizado por la OMS.

Factores de riesgos			
Prevalencia	Total	Hombres	Mujeres
Adultos fumadores actuales de tabasco ^a	17.00%	27.0%	8.0%
Jóvenes consumidores actuales de tabasco	19.8%	21.6%	17.7%
Consumo de alcohol en los últimos 12 meses	31.5%	45.1%	18.0%
Actividad física baja en adulto ^b	37.7%	37.1%	38.4%
Obesidad en adulto (IMC ≥ 30)	30.9%	26.1%	35.6%
a: prevalencia estandarizada por edad de consumo de tabaco en adultos (mayores de 15 años)			
b: realiza menos de 60 min de ejercicio diario.			

Factores químicos

Los factores químicos son unos de los principales factores en países en subdesarrollo como México, porque no cuentan con las medidas necesarias para controlar y regular emisión de reactivos peligrosos^{13,65}. Los factores químicos tiene la capacidad de provocar mutaciones y/o alteración en el genoma⁶⁶, estas sustancia tienden a producir un cierto número específico de cáncer, denominados cáncer ocupacional (Tabla 3), porque se desarrolla en tejidos expuesto a exposiciones prolongadas, un claro ejemplo de cancerígenos son el asbesto y los metales pesados⁶⁷.

Tabla 3 Cánceres ocupacionales y las sustancias químicas que produce el cáncer

Sustancia	Tipo de cáncer relacionados	Trabajo relacionado
Hidrocarburos aromáticos policíclicos	Pulmón y Vejiga	Siderúrgica Industria aluminio
Aminas aromáticas	Vejiga	Pigmentos Tinturas Benceno
Benceno	Leucemia	Petroquímica Fabricación de pinturas
Cloruro de vinilo	Hígado y Cerebro	PCV Derivados de madera
Asbestos	Pulmón	Fábricas de productos de asbesto-cemento

Factores físicos

Los factores físicos son todos aquéllos que tienen que ver con la transferencia de energía, el factor físico de mayor impacto en los seres humanos es la radiación, en forma de ondas, como la radiación solar. Lo que es el principal causante de los carcinomas. La radiación también está relacionada con otros tipos de cáncer, como la leucemia, pero en un menor grado⁵⁹. El efecto de la energía radiante se debe principalmente a que los rayos pueden llegar hasta al ADN, provocando cambios estructurales en las bases nitrogenadas. Un ejemplo son las mutaciones causadas por la luz ultravioleta en las células de la dermis⁶⁸, donde la radiación ultravioleta puede inducir transiciones de citocina a timina y la formación de dímeros de pirimidina que se relacionan con mutagénesis asociada a la replicación del ADN⁶⁹

Factores patogénicos

Los canceres asociados a patógenos externos presenta una tasa muy elevada, se estima que 1 de cada 5 cáncer es producido por mutaciones en la células causadas por patógenos⁷⁰. Los patógenos más importante que se asocian con el desarrollo de cáncer son; VPH, VHC, VHB y *Helicobacter pylori*, Estos microorganismos tienen la capacidad de realizar un daño genético a la célula transformándolas en células cancerosas⁷¹ Los tejidos diana que son afectados por este tipo de factor son el epitelial y el gastrointestinal. Un claro ejemplo es *H. pylori*, su presencia esta correlacionada con cáncer de esófago, cáncer estomacal, colorrectal entre otros. El *H. pylori* causa un daño en los tejidos que estimula a las células del sistema

inmune a sintetizar y secretar prostaglandinas^{72,73}, que son unas de las principales sustancias que están relacionadas con la respuesta inflamatoria⁷⁴. Con ello se estimula a las células a una rápida proliferación⁷⁵, lo que aumenta la tasa de recambio de las células, lo que incrementa la probabilidad de adquirir mutaciones y con ello la posibilidad de desarrollar cáncer⁷⁶.

Factores biológicos

Los factores biológicos son los más difíciles de controlar debido a que están presentes durante toda la vida de la persona y su efecto sobre el desarrollo del cáncer es constante. La edad es uno de los factores biológicos más importante para el cáncer, sobre todos el cáncer en tejidos de protección y conjuntivo⁷⁷. La edad biológica es un detonante de cáncer por dos motivos; por exposición, mientras mayor es la persona, mayor ha sido el tiempo de contacto con factores exógenos⁷⁸. Esto puede producir un desequilibrio en los mecanismos de reparación celular, y el daño causado por los radicales libres puede desencadenar mutaciones en la estructura del ADN, aumentando la posibilidad de cáncer^{79,80}. El segundo motivo asociado a la edad, es que durante la vida de la persona, las células, sobre todos los que conforma el tejidos de protección como la piel, se han dividido miles de veces, con lo que las células hijas pueden acumular cientos e incluso miles de mutaciones en su genoma. Esta condición puede contribuir de manera significativa a que las células somáticas o germinales adquieran propiedades de célula madre iniciando el cancer⁸¹.

De manera general los tumores que predominan en edad avanzada son los carcinomas, la distribución de rango de edad de la adquisición de nuevos canceres tiene un comportamiento como campana de Gauss (Figura 7), predominando el rango de edad de personas adulta con edades entre los 54 y 74 años, debido a los efectos mencionados anteriormente.

Otro factor biológico de importancia es la disposición genética, existe mutaciones o/y deleciones en genes que promueven la aparición o aceleran el cáncer. Estas variantes génicas pueden haber estado presentes toda la vida del paciente, por herencia génica, o pueden haber sido adquiridas por una mutación debido al ambiente al que está expuesto el paciente⁸².

En la literatura se han descrito varios genes y regiones genéticas que cuando presentan una mutación aumenta la probabilidad de adquirir o acelerar la aparición de una neoplasia². Por tal motivo los factores genéticos han adquirido un gran interés para el diagnóstico molecular,

un claro ejemplo son las mutación en el gen TP53 (Tabla 4) que se localiza en el brazo corto del cromosoma 17 (17p13.1), el gen codifica para una fosfoproteína llamada p53⁸³, el p53 tiene como función principal el control del ciclo celular principalmente en la fase G1^{84,85} y se le considera un gen supresor de tumor.

La proteína p53 cumple más de una función fisiológica en la célula y su actividad es muy importante para mantener la integridad y estabilidad del ADN genómico^{10,69,86}. Se ha determinado que en por lo menos 40% de los cánceres hay alteración en el gen TP53⁸⁵, la proteína mutada tiene deficiencias que le impiden que cumpla su función de regulación⁸⁷. Diversos estudios han demostrado que las alteraciones genéticas en TP53 no son azarosas⁸⁸, se ha reportado que ciertas mutaciones predisponen a un tipo de neoplasias en específico, por ejemplo, una mutación en el codón 249 de TP53 aumenta la posibilidad de cáncer de hígado, mientras mutaciones causadas por dímeros de timina en su región codificante lo vuelve más susceptible a cáncer de piel⁸⁹, mutaciones en gen tp53 también está ampliamente relacionados con leucemias y cáncer de mama^{12,90-92}.

Existe más genes de importancia clínica con el mismo impacto que el tp53, uno de ellos es el gen BRCA1, que presenta una función muy similar al gen Tp53, que consiste en mantener la integridad del ADN, Sus mutaciones estar relacionadas principalmente con carcinomas, sobresaliendo los canceres de mama, ovario y próstata⁹³,Algunas mutaciones puntuales en BRCA1 se han asociado con un aumento de hasta diez veces en el riesgo relativo de adquirir cáncer⁹⁴.

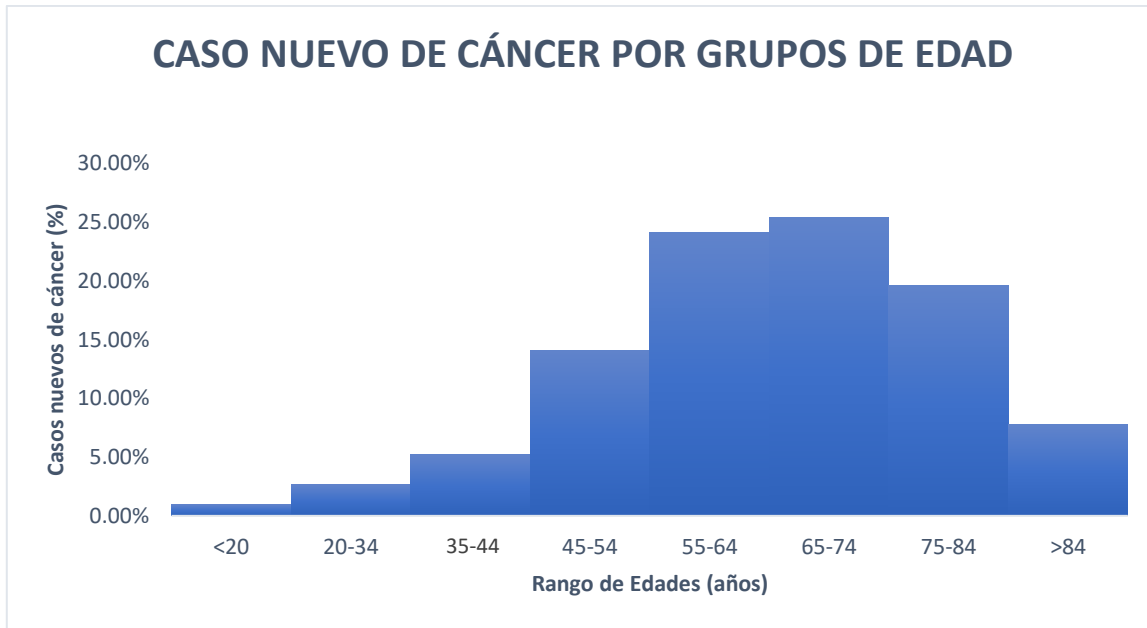


Figura 7 Distribución de casos nuevos de cáncer a nivel mundial por rango de edad

Tabla 4 Genes relacionados con algún tipo de cáncer y son utilizados como biomarcadores de cáncer aceptados por la OMS

Cáncer	Células que afectan	Genes usados como marcadores genéticos
Carcinoma	Células epiteliales	PIK3CA, TP53, MED12 y CDH1
Leucemia	Leucocitos	EBF1, IKZF1, ETV6, TP53 y CDKN2A
Sarcoma	Músculos y Cartílagos	FLI1, ERG, ETV1 y E1AF
Linfoma	células del sistema linfático	IgVH y ETV1
Mielomas	células plasmáticas	ULK4, PSORS1C2, CBX7 y CCND1
Gliomas	Células cerebrales	EGFR, PDGFRA, MET y FGFR2

ETIOLOGÍA

El cáncer es un trastorno crónico, complejo y dinámico⁹⁵, cada cáncer tiene su propia etiología, pero su presencia es atribuido a un mal control del crecimiento celular², debido a la gran diversidad de los factores que influyen en el cáncer, es difícil determinar la causa exacta para la formación de un cáncer⁹⁶, pero todos los cáncer son desencadenados por un cambio o una variación genética que fue adquirida y es heredable a sus progenitoras.

Cambios Genéticos

SNPedia es una base de datos con más de 5000 variaciones genéticas de un solo nucleótido (SNP), relacionado con al menos un tipo de cáncer. Las variaciones o alteraciones genéticas SNPs se encuentran distribuidas por todo el genoma, es decir se puede localizar tanto en zonas codificantes como no codificantes. Los SNPs en región codificante representan el 70% de las variaciones con relevancia clínica, Aunque en los últimos años se ha demostrado un gran impacto de las alteraciones genéticas en regiones no codificantes⁹⁷, se piensa que estas regiones son importantes para la regulación en genes de señalización, proliferación y modificaciones epigenéticas⁹⁷.

Los SNPs son estudiados debido a su facilidad de detección y la posibilidad de realizar una correlación directa con una enfermedad determinada, los SNPs se forman principalmente por mala reparación del ADN y por daño causado por radicales libres⁷⁴. La detección de SNPs puede ayudar a disminuir el contacto con ciertos factores de riesgo o incluso eliminar la posibilidad de su desarrollo mediante medidas médicas, como la extirpación de tejido susceptible, Un ejemplo de esta aplicación es la extirpación de seno, para evitar el cáncer de mama⁹⁸.

Los SNPs se clasifican (Tabla 5) dependiendo de la región donde se ubican y al efecto que ejercen sobre ella⁷⁴. Si el SNP ocurre en una región promotora se denomina de regulación⁹⁹, si el promotor es para un gen que codifica para una proteína se denominan SNP regulador (rSNP), y si es para un gen que no codifica para proteínas se denomina SNP reguladores de los microRNA (miR-rSNP), los polimorfismos reguladores tienen la capacidad de cambiar la expresión génica¹⁰⁰, los SNP localizados en la estructura del mRNA, pero no codifica a proteínas se denomina de tipo estructurales, si el polimorfismo modifica la traducción de un ARNm se denomina SNP de RNA estructurales (srSNP), pero sí afecta a los microRNA se denominan miR-srSNP. Este tipo de alteración afecta la síntesis, degradación y procesamiento del ARN¹⁰¹. Por último, si el SNP se encuentra en región codificante que se expresa a proteínas se le denomina codificante, si la alteración genética produce un cambio en el orden de los aminoácidos se le denomina SNP no sinónimos (nsSNP), en el caso que la mutación no genere ningún cambio se le denomina SNP sinónimos (sSNP). Se estima que los sSNP tienen un mayor impacto en predisponer al paciente a cualquier tipo de cáncer¹⁰².

Tabla 5 Clasificación de los SNP

TIPO	SNP	UBICACIÓN	FUNCIÓN
Reguladores	rSNP	Promotor de genes codificantes de proteínas	Alteran la expresión génica
	miR-rSNP	Promotor de un gen que no codificantes a proteína	Alteran la expresión génica
Estructural	srSNP	PmRNA	Alteran la traducción, la estabilidad, la longitud y la interacción de mRNA
	miR-srSNP	MiRNA	Afectan la estructura, el procesamiento y la función de los miRNA
Codificantes	sSNP	Secuencia codificante	Afectan la estructura y la función de las proteínas o enzimas
	nsSNP	Secuencia codificante	No cambia la configuración de la proteína o enzima

Secuenciación para el diagnóstico de SNP

Las nuevas tecnologías han permitido el avance en ciencias como la genética y metagenómica, las tecnologías de secuenciación son las más desarrolladas para la aplicación en oncología¹⁰³, una de las más importantes es la técnica de microarreglo llamada MammaPrint, que es un panel de 70 genes para el diagnóstico de cáncer de bajo y alto riesgo de mama¹⁰⁴.

Otra técnica de gran impacto en los últimos años es la secuenciación masiva, este tipo de secuenciación tiene como finalidad realizar lecturas a millones de secuencias de ácido nucleico al mismo tiempo, lo que permite detectar hasta miles de SNP, esta forma de detección abrió la puerta a un nuevo campo de estudio denominado medicina personalizada, que permite tomar las medidas óptimas a partir de su perfil genético.

En la actualidad casi todos los estudios de genética para correlacionar con una patología se utiliza la secuenciación masiva, se estudian principalmente los SNP, se han descritos miles de SNP en humanos, se han demostrados que la presencia de estos SNP predispones y/o protegen a la persona a obtener alguna enfermedad, la gran impacto que ha obtenidos por este tipos de estudio, permite que se desarrollaran base de datos especializados para SNP que tiene efecto con enfermedades (SNPedia), incluso se han desarrollados paginas para SNP para una enfermedad en concreto como el cáncer (COSMIS). Gracias a esta información una se hacen posible realizar una búsqueda de SNP en cientos de genes a la vez^{6,105}, proporcionando una cantidad de información muy extensa sobre la genética de un individuo y acortando el tiempo para el diagnóstico^{25,106,107}.

Las nuevas tecnología de secuenciación y las bases de datos que se encuentra disponible en la red es un nuevo mar de posibilidades que se abre para el diagnóstico preventivo de cáncer¹⁰⁸, Aunque la tecnología ya está presente en todo el mundo, la dificultad para realizar un diagnóstico consiste en determinar la técnica más conveniente y diseñar un correcto panel de diagnóstico (determina las alteraciones genéticas que se relacione con la enfermedad y las técnicas que permite realizar la correcta identificación de esta alteraciones), en el caso de la tesis se selección diseña la técnica de secuenciación por Exomas debido que es la técnicas más accesibles para realizar el diagnósticos de SNP relacionados a cáncer.

Exomas

Las técnicas de secuenciación genómicas todavía no son una opción viable para realizar un diagnóstico debido a su alto costo, y su requisito bioinformático para poder procesar toda la información¹⁰⁹, una elección para resolver esta limitante son las técnicas de secuenciación de exomas^{105,110}, que permite secuenciar solo los exones de los genes que se expresan^{28,29}, esta secuencia representa aproximadamente el 1-2% del total de genoma humano¹¹¹.

Se estima que más del 50% de los SNP relacionados con cáncer están presentes en los exones¹¹². A pesar de que se tienen grandes avances en el tratamiento y detección se cree que la carga del cáncer aumentará y esto debido a que, la población mundial va en aumento y la expectativa de vida cada vez es mayor por lo que cada vez habrá más gente que pueda desarrollar cáncer por lo que, la OMS calcula que para el 2025 el cáncer pasará de 14 a 19 millones de personas afectadas por el cáncer y la cifra puede llegar hasta llegar a los 24 millones en el 2035, por dichos motivos se requieren técnicas cada día más exactas, una de las alternativas para realizar un correcto diagnóstico preventivo es mediante el uso de biomarcadores¹¹³, para ser una herramienta útil para los próximos años, se requiere establecer una técnica que sea válida como técnicas de detección de biomarcadores para cáncer, en el mercado ya existe varias técnicas de diagnósticos de secuenciación masiva que cumplieron los criterios mínimos necesarios para su uso clínico (Tabla 6), pero se han validado para poblaciones principalmente americana y europea¹¹⁴, pero no existe o se publicaron un estudio para el diagnóstico genético de biomarcadores para cáncer en población mexicana que cumpla con las reglas para la validación de una técnica de diagnóstico, por lo que se quiere diseñar una técnica de secuenciación como un panel de diagnósticos que cumpla con las características mínimas necesarias para su posible aplicación en población mexicana.

Tabla 6 Técnicas de diagnóstico que se han diseñados que se encuentra en mercado

Producto	Enfermedad	Tipo	Número de genes	Muestras
Foundation One	Cáncer	Panel	236 genes y 19 genes para splice	Tumores solidos
Foundation One heme	Cáncer	Panel	670 genes, de los cuales 265 tiene splice	Tumores sólidos y líquidos
Cancer Complete	Cáncer	Exoma	~20000 genes	Muestra biológica
Cancer Select	Cáncer	120 genes	120 genes	Tumor
Exome Next	Cáncer	Exoma	~20000	Sangre
BRCA2 gene sequencing	Cáncer de Mama	Panel	BRCA1 y BRCA2	Sangre y corte histológico
XomeDX	Cáncer	Exoma	20000 genes	Sangre y saliva
XomeDX Plus	Cáncer	Exoma	>20000 genes	Sangre y saliva
XomeDX Slice	Cáncer	Exoma	Panel abierto	Sangre y saliva
Comprehensive cancer panel	Cáncer	Panel	Panel abierto	Sangre
EGFR Mutation Analysis	NSCLC	Panel	1 gen (EGFR)	Epidermis
NeoSITE Melonoma	Cáncer	Panel	5 genes	Biopsia de tejido
FISH for non-small cell lung Cancer	NSCLC	Panel	2 genes	líquido pleural y líquido peritoneal
Colorectal Cancer panel	Cáncer Colorrectal	Panel	2 genes	Sangre
Mi Profile	CANCER	Panel	47 genes	Sangre
BRCA Analysis	Cáncer de mama y ovario	Panel	BRCA1 y BRCA2	Tejido y sangre
OncoVantage	Tumores Solidos	Panel	34 genes	Tejido biológico
Comprehensive cancer Gene Set Analysis	Cáncer	Panel	42 genes	Sangre y tejido
Gastrointestinal hereditary cancer panel	Cáncer gástrico	Panel	15 genes	Sangre
Cancer panel	Cáncer	Panel	~20000 genes	Sangre
ACE Clinical Exome	Cáncer	Exoma	Panel abierto	Sangre
Exome cancer	Cáncer	Exoma	Panel abierto	Sangre
GEM Cancer Panel	Cáncer	Exoma	562 genes	Sangre y tumor

PLANTEAMIENTO DEL PROBLEMA

Clínicamente se diferencian más de dos centenares de tipos de cáncer, debido a la gran gamma que se tiene, es difícil tener una etiología general para la enfermedad, pero si lo vemos

desde un punto de vista genético, tal diferenciación es más puntual, se puede encontrar secuencias específicas en el genoma que nos permite tener un acercamiento para su etiología.

El genoma humano en su versión GRCh38.p12 está formado por alrededor de 52,000 genes, de los cuales se ha demostrado que menos del 2% son genes que presenta un efecto significativo en la formación y/o desarrollos de por lo menos de algún tipo de cáncer. Este pequeño grupo de genes son esenciales para el diagnóstico preventivo, debido que durante varias décadas el mundo científico ha invertido sus recursos en determinar mutaciones puntuales que aumenta la disponibilidad y/o la resistencia que presenta las personas para desarrollar un tipo específico de cáncer, dado la gran cantidad de información que se han recolectados en las últimas décadas, por el cual se han desarrollados bases de datos que contiene la información de miles de SNP que presenta una correlación con la patología.

Las técnicas que se utilizaron para empezar las bases de datos fueron la PCR y técnicas de inmunoprecipitación, al presente estos métodos son ineficaces debidos a que no se puede detectar más de 10 mutaciones puntuales por reacción¹¹⁵, actualmente las tecnología de secuenciación de tercera generación han aumentado el rendimiento de bases de ADN que puede leer por reacción y sus costo han bajado, lo que permite detectar hasta miles de mutaciones en una sola reacción, lo que habré una ventana de posibilidad en el mundo del diagnósticos²⁵.

Con la información de las bases de datos sobre mutaciones de cáncer y las nuevas técnicas de secuenciación masiva que son más accesibles debido a su disminución de costos, se propone diseñar una técnica de detección de biomarcadores (SNP) especialidades en cáncer ya validados por estudios previos, con la información que se obtenga de la técnica se pretender que puede ser usado para poder tomar medidas preventivas¹¹⁶.

OBJETIVOS

GENERAL

Diseñar, desarrollar y validar una técnica de secuenciación masiva para la detección de SNP, con los parámetros mínimos necesario para su aplicación clínica.

PARTICULAR

- Analizar las bases de datos biológicas con genes reportados que se expresan en las células del tejido sanguíneo
- Determinar los genes y las mutaciones que están asociadas a los diferentes tipos de cáncer reportados
- Diseñar los oligonucleótidos para la detección de biomarcadores de cáncer en células sanguíneas
- Determinar el número de muestras estadísticamente necesarias para la validación de la técnica
- Validar el paquete tecnológico

HIPÓTESIS

La secuenciación por nanoporos diseñada para la detección de los SNP asociados a cáncer cumple con los criterios mínimos necesarios para ser una técnica de detección de biomarcadores de cáncer en formato de panel para su aplicación clínica.

DESCRIPCIÓN DEL TIPO DE ESTUDIO

Tipo de investigación: Descriptivos

Clasificación del estudio: Transversal

Población de estudio: Personas en un rango de edad de 18-65 años

Tipo de variable: Cualitativa

Variable: Polimorfismos de una solo nucleótidos relacionado con cáncer.

MATERIAL Y MÉTODOS

BÚSQUEDA DE MUESTRAS BIOLÓGICAS

Determinación del número de muestras para la validación de la técnica

Determinar el número mínimo de muestras para la validación de una técnica es de suma importancia, porque nos indica la fiabilidad y la confianza de los resultados, por lo general todas las técnicas de secuenciación requiere una confiabilidad y confianza del 95% para poder usarlo como un panel clínico¹¹⁷, para realizarlo se tomó que los polimorfismos tiene una distribución binomial.

OBTENCIÓN DE MUESTRA BIOLÓGICAS

las muestras fueron obtenidas por punción venosa, recolectadas por personal capacitado de la clínica GENO+ S.A de C.V.

Población de estudio

Personas entre 18-65 años que solicitaron un servicio para la detección de SNP relacionados con obesidad y diabetes que donaron su muestra para investigación en la clínica de diagnóstico genético GENO+ S.A de C.V.

Criterios de inclusión de la población de estudio

Muestras sanguíneas que no presente niveles altos de lípidos y/o hemolisis.

Muestras sanguíneas que el paciente dono voluntariamente su muestra sanguínea.

Muestras que fueron etiquetas y preservada adecuadamente.

Criterios de exclusión de la población de estudio

EL volumen de la muestra no es suficiente.

Criterios de eliminación de la población control

La muestra presente problema cuando se realiza la PCR punto final.

Criterios de exclusión de la población control

El paciente llamo a las oficinas de la clínica para pedir que la muestra recolectada no puede ser usadas para investigación clínica.

Criterios para la obtención de muestra y su procesamiento

Las muestras fueron recolectadas por el técnico laboratorista Javier López García encargado del área de flebotomía de la clínica de diagnóstico molecular GENO+ S.A de C.V, antes de la recolección de la muestra sanguínea, el paciente firmo el consentimiento informativo, posteriormente llenará un cuestionario con sus datos antropomórficos y estilo de vida realizado por la empresa GENO+ S.A de C.V.

PROCESAMIENTO DE LA MUESTRA

La muestra se almaceno a -70°C hasta su uso. Las reacciones y preparación de las muestras se realizaron en el laboratorio de GENO+ S.A de C.V.

ANÁLISIS DE BASES DE DATOS BIOLÓGICAS

Base de datos de genes asociados a diferentes tipos de cáncer

Se realizó una base de datos con los genes que se han reportados que son detectable en sangre periférica, estos genes se descargaron de la base de datos de GEO de NCBI (<https://www.ncbi.nlm.nih.gov/geo/>), se procesaron un total de 72 estudios que utilizaron como muestra, sangre periférica por venopunción, de los 72 estudios, 20 fueron de pacientes con obesidad, 20 con diabetes tipo 2, 20 con leucemia, 10 de personas que no presentaba ninguna patología y/o deficiencia y 2 estudios en personas que tiene un consumo de alcohol anual de 7 litros por años (mayor a la media mundial), se seleccionaron todos los genes que se expresaron por lo menos una vez en sangre sin importar la condición y el tipo de estudio.

Desarrollo de una base de datos de genes relacionado con cáncer

Con los genes que se expresan en sangre se realizó un script en el lenguaje phyton, para asociar los genes que se expresa en sangre con alguna patología de neoplasia reportada en las bases de datos especializados en cáncer; intogen (<https://www.intogen.org/search>), Cosmic (<https://cancer.sanger.ac.uk/cosmic>), NIH (<https://www.cancer.gov>) y MalaCards (<http://www.malacards.org/>).

Con los genes obtenidos se realizó una base de datos de los tumores, se seleccionaron los genes que esta correlacionados con al menos 3 tipos de tumores, con estos datos obtenidos

se realizó un mapa de interacción entre los genes, agrupándolos por el tipo de cáncer que se ha reportado, para determinar el impacto que tiene este gen para desarrollar varios tipos de cáncer. Para realizarlo se usó el programa Cytoscape versión 3.0 con la app cy3sbml versión 0.2.7.

Polimorfismos asociados a cáncer

Se identificaron todos los polimorfismos asociados a cáncer en las diferentes bases de datos biológicas especializadas en mutaciones y variaciones de genes relacionados con una o varias patologías: SNPedia (<https://www.snpedia.com/>), dbSNP (<https://www.ncbi.nlm.nih.gov/snp/>), MalaCards (<http://www.malacards.org/>) y GWAS Catalog (<https://www.ebi.ac.uk/gwas/>).

Se seleccionaron todos los SNP que se han determinado que su presencia aumenta o predispone a un fenotipo neoplásico, los criterios de selección para fenotipos es que tuvieran una RM mayor de 1, con un intervalo de confianza del 95% mayor a 1.01.

De todos los SNP con relevancia clínica se seleccionaron solo los SNP con una relación con neoplasia y que pudieran ser detectado por secuenciación por exoma.

Posición cromosómica de cada una de las mutaciones.

Cada una de las mutaciones que fueron seleccionados como posibles biomarcadores, se determinó su localización cromosómica mediante el programa PhenoGram modificando el script en lenguaje RUBY, para adicionar la condición que en los genes que presenta más de 20 mutaciones en su secuencia, se posicione la mutación a una distancia que se encuentre a no mayor de 2500 bases en ambas direcciones del gen.

Diseño

Estudio biológico de genes

Para conocer la función, proceso biológico y el componente celular de los genes que se expresa en sangre y están relacionados con cáncer, se realiza un análisis bioinformático en la base de datos GeneOntology (<http://www.geneontology.org/>) y GORILLA (<http://cbl-gorilla.cs.technion.ac.il/>), obteniendo las funciones, procesos y componente celular está presente los genes relacionados con cáncer.

Modelo booleano

Con la base de datos de SNP seleccionado para el estudio, se modificó la simulación de modelo booleano sincrónico de redes de regulación de genes relacionado a la metástasis en GINsim del consorcio de cáncer de París¹¹⁸, la modificación fue introducir penalizaciones en la función de los genes que son detectable en sangre que presente mutaciones¹¹⁹, para determinar de manera cualitativa si el paciente aparte de tener posibilidad de tener cáncer, también determinar si tiene riesgos a metástasis¹²⁰.

Circus plot

Con los datos recolectados de los SNP, se realizó un análisis Circus plot, para poder realizar una correlación de todos los datos obtenidos de los SNP obtenidos en la tesis con los cromosomas, para su diseño se usó la biblioteca Circus plot de programa R, la intensidad de cada línea representa de manera cualitativa el número de mutaciones para tipo de cáncer con dicha enfermedad.

PCR para la secuenciación

Calidad de la extracción del ADN para la secuenciación.

Como se utilizó una enzima que amplifica directamente del tejido sin la necesidad de purificar el ADN genómico para la PCR, se realizó una lisis celular con el kit Blood-Animal-Plant DNA Preparation de la marca Jena Bioscience con número de catálogo, PP-213XS para asegurar la integridad de las muestras se realizó un gel de agarosa al 2%.

PCR

Para la PCR, se usó el kit Direct PCR Máster de la marca Jena Bioscience con número de catálogo PCR-111S. la reacción de PCR fue preparada como se presenta en la tabla 7, la reacción de PCR se realizó en un termociclador de la marca Axygen II, el protocolo se presenta en la tabla 8. la muestra fue sangre diluida a una relación 1:1 con agua destilada.

Tabla 7 Preparación de la reacción de PCR

Componentes	Volumen	Concentración
Direct PCR máster	10 ul	
Oligonucleótidos	2 ul	10 µM
Muestras	2 ul	10 µM
Agua libre de ADNasa	6 ul	
Total	20 ul	

Tabla 8 Parámetros de reacción de la PCR

Temperaturas	Tiempo	Ciclos
95°C	300s	1
95°C	20s	35
55°C	30s	
72°C	90s	
72°C	120s	1

Nota: los oligonucleótidos que se usaron para hacer la PCR son propiedad intelectual de la empresa geno+, por el motivo que son los oligonucleótidos para el formato multiplex que está en validación.

Confirmación de la PCR

Para confirmar que la PCR funciona, se realizó una validación 6 SNP seleccionados a la AZAR que se encuentra presente en el panel de detección diseñado, mediante PCR y qPCR. Usando como molde o muestra el producto de la PCR diseñada para la secuenciación. Todas las reacciones de PCR y qPCR. se realizó por triplicado, los datos obtenidos se graficaron con el programa Gram Prism versión 7.0, usando como variable dependiente la media aritmética de cada amplificado y como variable independiente el número del ciclo.

Diseño de los oligonucleótidos para la validación

Los 6 pares de oligonucleótidos para la validación fueron diseñados bajo las siguientes condiciones; contenidos de CG entre 40-60%, Tm ajustada por sales entre 63-67°C, con terminación SS en el extremo 3' y que todos los amplificados tenga un tamaño aproximado de 100 pb.

Validación por PCR.

Para la PCR de validación, se usó el kit Direct PCR Máster de la marca Jena Bioscience con número de catálogo PCR-111S. la reacción de PCR fue preparada como se presenta en la Tabla 9, la reacción de PCR se realizó en un termociclador de la marca Axygen II, el protocolo se presenta en la Tabla 10.

Tabla 9 Protocolo de Preparación de la reacción de PCR para la validación

Componentes	Volumen	Concentración final
Direct PCR máster	10 ul	
Oligonucleótidos	2 ul	10 µM
Muestras	2 ul	10 µM
Agua libre de ADNasa	6 ul	
Total	20 ul	

Tabla 10 Parámetros de reacción de la PCR

Temperaturas	Tiempo	Ciclos
95°C	300s	1
95°C	20s	35
60°C	30s	
72°C	60s	
72°C	120s	1

Validación por qPCR.

Para la validación por qPCR, se procesaron Los 6 SNP seleccionados al azar, 3 controles negativos (muestra de ADN de bacterias) y 3 controles sin muestras (la muestra se sustituye por agua destilada).

Se realizó la qPCR con el kit qPCR SybrMaster de la marca Jena Bioscience con número de catálogo PCR-372S. la reacción de PCR fue preparada como se presenta en la tabla 11, la reacción de PCR se realizó en el termociclador PIKO24, el protocolo para el termociclador se presenta en la tabla 12.

Tabla 11 Preparación de la reacción de qPCR para validación

Componentes	Volumen	Concentración
qPCR SybrMaster	10 ul	
primer forward	0.6 ul	10 µM
primer reverse	0.6 ul	10 µM
Muestras	2 ul	15 µM
Agua libre de ADNasa	6.8 ul	
Total	20 ul	

Tabla 12 Parámetros de reacción de la qPCR para validación

Temperaturas	Tiempo	Ciclos
95°C	60s	1
95°C	15s	40
64°C	10s	
72°C	60s	

SECUENCIACIÓN

Se utilizaron dos tipos de secuenciaciones, la primera es la técnica que usaremos de referencia, es por ilumina, es un kit registrado por MacroGen para obtener exomas, el kit comercial es para 99,995 SNP, la muestra que se envió a MacroGen era la muestra con cáncer de tiroides, la muestra enviada fue ADN genómico con alta pureza a una concentración de 70 ng por microlitros, en un volumen de 20 microlitros. La segunda fue por MinION con la técnica propuesta para validar con la técnica de referencia.

Preparación de los productos de PCR para las bibliotecas

Los productos de la PCR, se le realizó una purificación con perlas magnéticas AMPure XP beads de la empresa Agencourt a una concentración 0.6 X, este proceso permite eliminar residuos con un tamaño menores de 50 pb, permitiendo tener un tamaño uniforme para poder tener una mayor eficiencia de la secuenciación.

Preparación de la biblioteca

Para preparar la biblioteca se confirma la concentración de ADN con el cubit versión 2.0, si tenía una concentración mínima de 6 ng/ul y no presentaba degradación, la muestra fueron

procesada para preparar la biblioteca para la secuenciación. la biblioteca para secuenciación para la tesis fue 1D de nanopore, el protocolo de preparación fue la de la casa comercial (https://community.nanoporetech.com/protocols/pcr-barcoding-amplicons-r/v/pbae_9008_v108_revt_18oct2016).

Validación

Para confirmar el uso de una técnica como un panel de genes con aplicación clínica, se requiere confirmar el resultado obtenido con una técnica diferentes, para determinar que realmente tenemos lectura que cubre la región donde está el variante clínico, se seleccionaron un total de 6 SNP al azar, se diseñaron los oligonucleótidos para amplificar unas bases antes y después de las mutaciones, se usaran como molde el producto de la PCR de 3 muestras diferentes para determinar si están presente las secuencias seleccionadas al azar.

Validación para la técnica

La validación se realizará con los parámetros mínimos que pide COFEPRIS y la Asociación de Patología Molecular (AMP), la AMP por la necesidad de establecer criterios mínimos para realizar un panel de genes para su uso clínico, en enero del 2017, se establecieron reglas y recomendaciones que son necesarias para asegurar una correcta técnica de diagnóstico para secuenciación masiva en formato; producto de PCR, WES y WGS para células de tejidos tanto sólido como líquido¹²¹.

Controles

Los control en cualquier técnica es fundamental para confirmar que los resultados con un contraste, en caso del panel de secuenciación que se está diseñando se seleccionó un par de oligonucleótidos que amplifica exclusivamente a bacterias por la región ribosomal 16s para su detección, los oligonucleótidos, se procesaron un total de 4 muestra con una repetición y tres controles negativos (solo oligonucleótidos), la muestra para los control fueron heces fecales de humanos, porque en ella se encuentra tanto muestras genómica de humanos y de bacterias ¹²². Para realizarlo se obtuvo el ADN genómico con el Kit de purificación ADN genómico, viral, bacteriano y de parásitos a partir de materia fecal de Quiagen. Los oligonucleótidos seleccionados son propiedad intelectual de la empresa genomias, por el cual la secuencia no puede ser publicada, pero hibrida en las regiones de 16S rRNA con un tamaño

promedio de 600 pb, estos controles se secuenciarán en la matriz de validación del equipo de secuenciación.

Matriz de validación del equipo de secuenciación

Se diseñó una matriz para determinar el orden como se van a secuenciar cada una de las muestras (Tabla 13), el ordenamiento de los datos fue realizado con el comando de aleatorización del programa SPSS versión 10.4. debido al costo de la secuenciación para validar toda los SNP, se seleccionaron solo los 6 SNP que se usaron en la validación de la PCR para la validación del equipo, solo las muestras de cáncer se realizó la secuenciación por exomas completo, debido al alto costo. En la validación del equipo participaron tres técnicos en dos tiempos diferentes, para poder realizar todos los cálculos mínimos que la técnica es robusta para su uso como una técnica de diagnóstico preventivo, la validación será determinada por tres técnicos, a cada técnico se denominara como TEC1, TEC2 y TEC3, cada corrida tendrá 2 días de diferencia en su preparación.

Tabla 13 Matriz de procesamiento de las muestras

Corrida	Técnico	1	2	3	4	5	6	7	8	9	10	11
1	TEC1	P1 ^a	P17 [*]	P50	P37	P6	P24	P7	P12 [*]	P51	P17 [*]	NTC ^b
2	TEC2	P12 [*]	P30 ^a	P32	P3	P42	P47	P46	P49	P17 [*]	P36	C ^c
3	TEC3	P18	P2	P12 [*]	P38	P17 [*]	P4	P45	P31	P56	P54	C ^c
4	TEC1	P1 ^a	P11	P28	P39	P16	P12 [*]	P44 ^a	P48	P57	P53	C ^c
5	TEC2	P13	P21	P34	P12 [*]	P5 ^a	P27	P17 [*]	P9	P58	P52	NTC ^b
6	TEC3	P22	P17 [*]	P20	P36	P40	P26	P35	P15	P59	P12 [*]	C ^c
7	TEC1	P29	P33	P41	P14	P10	P43	P8	P17 [*]	P19	P55	NTC ^b

*: Son las muestras seleccionadas para determinar la reproducibilidad, robustez y repetitividad de la técnica

a: Muestras seleccionadas para determinar los parámetros de especificidad y sensibilidad

b: Control blanco

c: control de corrida de 16s

ANÁLISIS DE LOS DATOS

Las secuencias que se obtenga se le realizar todos los parámetros necesarios para poder realizar una correcta identificación de variante, para poder realizar un diagnóstico preventivo.

Proceso de las lecturas en bruto

Cualquier trabajo con datos provenientes de tecnología de secuenciación de tercera generación sin importa la técnica utilizada se requiere procesar los datos para poder realizar los análisis bioinformático, para poder realizar el análisis de las lecturas se tiene que transformar los archivos FATS5 que se obtiene al terminar una corrida de secuenciación a los formatos FastQ, que es el adecuados para realizar índice de calidad y poder filtrar apropiadamente las lecturas de baja calidad para calcular el índice de calidad e identificación de las variantes.

Índice de calidad

El Índice de calidad (Q), se define como la probabilidad de que una base de una determinada posición sea un error, este parámetros tiene como finalidad depurar o eliminar de manera significativa las lecturas que posiblemente presente error, esto permite disminuir el número de secuencias procesadas, permitiendo determinar la manera más rápida el procesamiento de datos, no existe un límite inferior que justifique el valor mínimo de error, estos es debido a que depende del error sistemático y aleatorio que presente la técnica. Para calcularlo se utilizó el programa fastQC.

Identificación de variantes

Una vez que se tenga las secuencias depuradas, se realizara la identificación de los SNP de interés para la tesis, para la identificación se usara el programa GATK¹²³, mientras el análisis de los SNP se usara el programa IGV¹²⁴.

Profundidad

La profundidad es uno de los parámetros con mayor importancia en la secuenciación masiva porque indica el número de veces que cada base debe estar presente en las lecturas de la secuenciación, esto es para determinar si la base detectada se encuentra en el genoma del paciente y el resultado no es debido a los errores de tipo 1 y/o tipo 2, para determinar la profundidad no existe una ecuación ya establecida para su cálculo, para la detección de nuevas variaciones genéticas se pide un mínimo de 150 lecturas¹²¹, en el caso de ser variaciones genéticas ya reportada en la literatura se recomienda un mínimo de 30 lecturas para confirmar la variación genética¹²⁵.

Rango de detección

El rango de detección se define como intervalo de todos los SNP de interés detectada por la técnica, este parámetro de suma importancia porque nos indica la verdadera robustez que presenta la técnica. Su cálculo es mediante el análisis de todos los SNP obtenidos de cada muestra.

Comparación de dos diferentes secuenciaciones

Los parámetros como valores predictivos, especificidad y sensibilidad se obtendrán mediante la comparación de los datos de secuenciación de 1 muestra, con tres repeticiones que se le realizó dos tipos de secuenciación, la primera con la técnica diseñada para la tesis y la segunda una técnica de exoma ya comercial de la empresa MacroGen.

Tabla 14 Cuadro 2x2 para el análisis de sensibilidad y especificidad de técnicas de diagnóstico.

		Técnica de referencia		Total
		Presente	Ausente	
Prueba diagnóstica	Positivo	A	B	A+B
	Negativo	C	D	C+D
Total		A+C	B+D	A+B+C+D

Con los datos obtenidos se realizar una tabla de 2x2 para calcular los siguientes parámetros de validación;

Tabla 15 Parámetros que se calcula con la comparación de dos técnicas

Sensibilidad: Es la capacidad que tiene la prueba para detectar variantes clínicas que está realmente presente en el genoma	$\text{Sensibilidad} = \frac{A}{A + C}$
Especificidad: Es la capacidad que tiene la prueba para detectar posiciones de bases sin mutaciones que está realmente presente en el genoma	$\text{Especificidad} = \frac{D}{B + D}$
Exactitud: Se obtiene con la unión de especificidad y sensibilidad y dicta la capacidad que tiene la técnica de dar un resultado verdadero	$\text{Exactitud} = \frac{A + D}{A + B + C + D}$
Valor Predictivo Positivo (VPP): la proporción de individuos que realmente tiene el SNP y son diagnosticado que presenta SNP en la prueba diseñada	$VPP = \frac{A}{A + B}$
Valor Predictivo Negativo (VPN): la proporción de individuos que realmente no tiene el SNP y son diagnosticado que no presenta SNP en la prueba diseñada	$VPN = \frac{D}{C + D}$

ANÁLISIS ESTADÍSTICO MEDIANTE CURVAS ROC

Para calcular la exactitud de un método de diagnóstico y su capacidad para clasificar de forma correcta dentro de diferentes categorías se realiza una curva ROC, la curva ROC es de gran utilidad ya que permite medir de forma pura la capacidad predictiva de un sistema o técnica, la curva ROC se calcula a partir de la sensibilidad y la especificidad de la técnica, se realizó una curva ROC no paramétrica debido a que no se puede determinar la distribución de cada SNP debido que cada mutación presenta una distribución independiente, el objetivo de la tesis para hacer la curva ROC es obtener el área bajo la curva, este dato permite saber la probabilidad que tiene la secuenciación de obtener la base correspondiente de la ubicación del polimorfismo. El valor bajo la curva va desde 0 a 1, mientras más cerca este del 1, se considera más precisa la técnica.

La escala para el área bajo la curva es la siguiente

Tabla 16 Escala de área bajo la curva de ROC

Valor de área bajo la curva	Clasificación
0.9-1.0	Excelente
0.8-0.89	Bueno
0.7-0.79	Razonablemente bueno
0.6-0.69	Pobre
0.5-0.59	Malo

Análisis de curvas ROC Para determinar la exactitud de los métodos diagnósticos se empleó el programa STATA 14®. Dicho programa determino el área bajo la curva, junto con su error típico, su nivel de significación estadística y su intervalo de confianza del 95%.

RESULTADOS

BÚSQUEDA DE MUESTRAS BIOLÓGICAS

Determinación del número de muestras

Para determinar el número mínimo de muestra para la tesis, se utilizó la ecuación de la distribución binomial (Ecuación 1), donde CL es el intervalo de confianza que requerimos para ser usado en prueba de diagnóstico (95%), p es la probabilidad de tener un resultado, n es el número total de eventos y k es el número de evento^{117,121}.

Ecuación 1 Distribución Binomial

$$\sum_{i=k}^n \binom{n}{i} = p^{n-1} (1 - p) = 1 - CL$$

Si partimos de la premisa que k solo puede tomar valores enteros desde 0 a n, la ecuación se puede simplificar (Ecuación 2).

Ecuación 2 Formula simplificados de la Distribución Binomial.

$$n = \frac{\ln(1 - CL)}{\ln(p)}$$

La AMP y FDA determino que para utilizar una técnica de secuenciación a nivel clínico se requiere un intervalo de confianza mínimo del 95%, con un margen de error no superior del 5%, es decir, se debe de tener una precisión del 95% de obtener la secuencia correcta (p=0.95), sustituyendo los datos en la ecuación 2, se determinó que el número mínimo de muestra para la validación de la técnica es de 59 muestras biológicas.

$$n = \frac{\ln(1 - 0.95)}{\ln(0.95)} = 58.4 \approx 59$$

OBTENCIÓN DE MUESTRA BIOLÓGICAS

Se recolectaron un total de 83 muestras sanguíneas de personas de un rango de edad de 18-65 años, ninguna muestra presento hemolisis o lipemia, de las 83 muestra una muestra pertenecía de un paciente de cáncer de tiroides.

Ninguna muestra recolectada fue eliminada, todas las muestras están resguardadas en el laboratorio de Geno+ S.A de C.V, a una temperatura de -70°C .

BÚSQUEDA BIBLIOGRÁFICA DE GENES

Se localizaron un total de 15744 genes que son detectable en sangre periférica reportado en la base de dato de GEO (tabla 17), abarcando el 26.11% de todos los genes que se tiene presente en el genoma humano (comparación con la base de datos GRCH38.p11), los cromosomas 1, 6 y 19 son los cromosomas que más genes expresan en sangre como se muestra en la Figura 9, los cromosomas 1, 6 y 19 abarcan entre ellos el 26.39% de los genes detectables en sangre. Del total de genes que puede expresar un cromosoma detectable en sangre, los cromosomas 6 y 19 sobresalen de la media, introduciendo el cromosoma 17, expresa más del 40% de los genes codificado en su estructura (Figura 10).

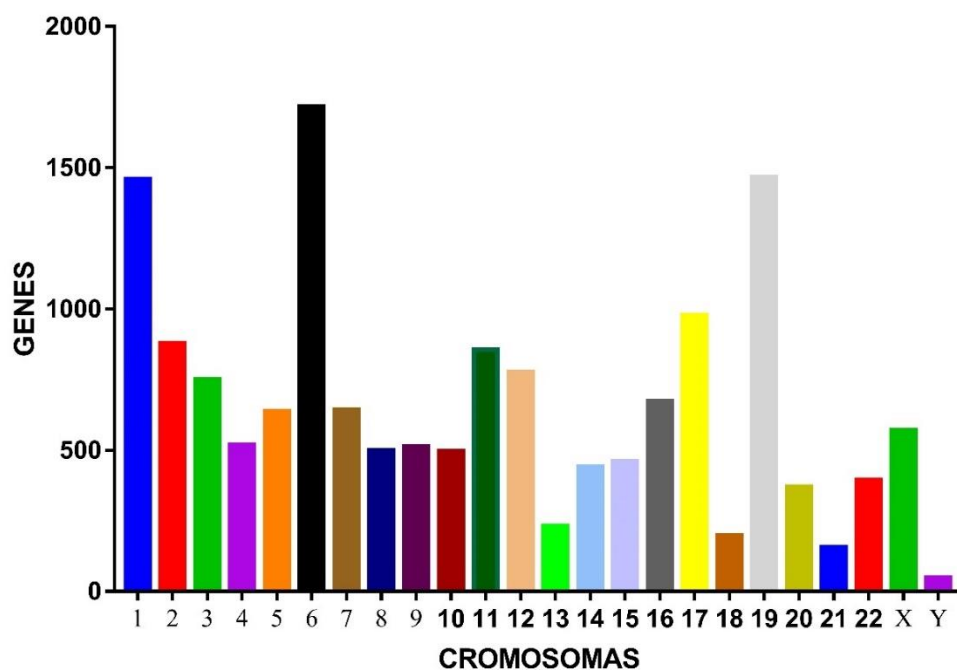


Figura 9. Número de genes que expresa en sangre en cada cromosoma

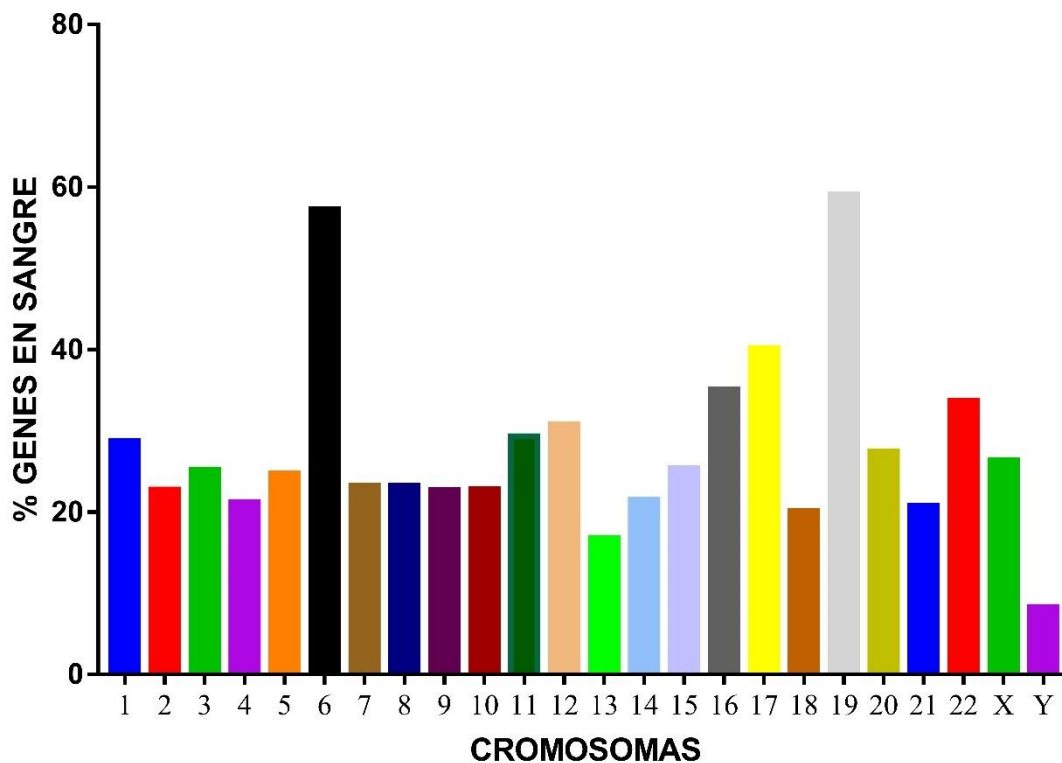


Figura 10. Porcentaje de genes que se expresa por cromosoma detectable en sangre

Base de datos de genes relacionados con cáncer

De los 15744 genes que se expresan en sangre, se determinó que el 9.32% (1468 genes) tiene por lo menos una relación con alguna neoplasia, que se pueden detectar en sangre usando la base de datos reportadas hasta el 28 de octubre del 2017.

Interacciones de los genes con el cáncer

Se determinó un total de 72 genes que interacción con un mínimo de 3 tipos de cáncer, en la Figura 11, se muestra los genes y con los tipos de cáncer que se relacionan, mientras más al centro se encuentre el gen, el efecto que tiene cuando presenta un mal funcionamiento aumenta la probabilidad que se forme una neoplasia, lo importante es que estos genes se pueden usar para buscar SNP que no se han relacionados con cáncer para poder estudiar y determinar el efecto que puede tener la variación genética.

Tabla 17 Total de genes que se expresa en sangre en el genoma humano.

Cromosoma	Tamaño (Mb)	Genes	Genes que se expresa en sangre	Porcentaje de genes que se expresa en sangre
1	248.96	5,078	1460	28.75
2	242.19	3,862	879	22.76
3	198.3	2,971	750	25.24
4	190.22	2,441	519	21.26
5	181.54	2,578	638	24.75
6	170.81	3,000	1717	57.23
7	159.35	2,774	644	23.22
8	145.14	2,152	500	23.23
9	138.4	2,262	513	22.68
10	133.8	2,174	496	22.82
11	135.09	2,920	856	29.32
12	133.28	2,521	777	30.82
13	114.36	1,381	232	16.80
14	107.04	2,055	442	21.51
15	101.99	1,814	461	25.41
16	90.34	1,920	674	35.10
17	83.26	2,432	978	40.21
18	80.37	988	199	20.14
19	58.62	2,481	1466	59.09
20	64.44	1,349	371	27.50
21	46.71	756	157	20.77
22	50.82	1,172	395	33.70
X	156.04	2,158	570	26.41
Y	57.23	577	48	8.32
Mitocondria	0.02	37	1	2.70
Desconocido	165.56	6,445	1	0.02
Total	3253.88	60298	15744	26.11

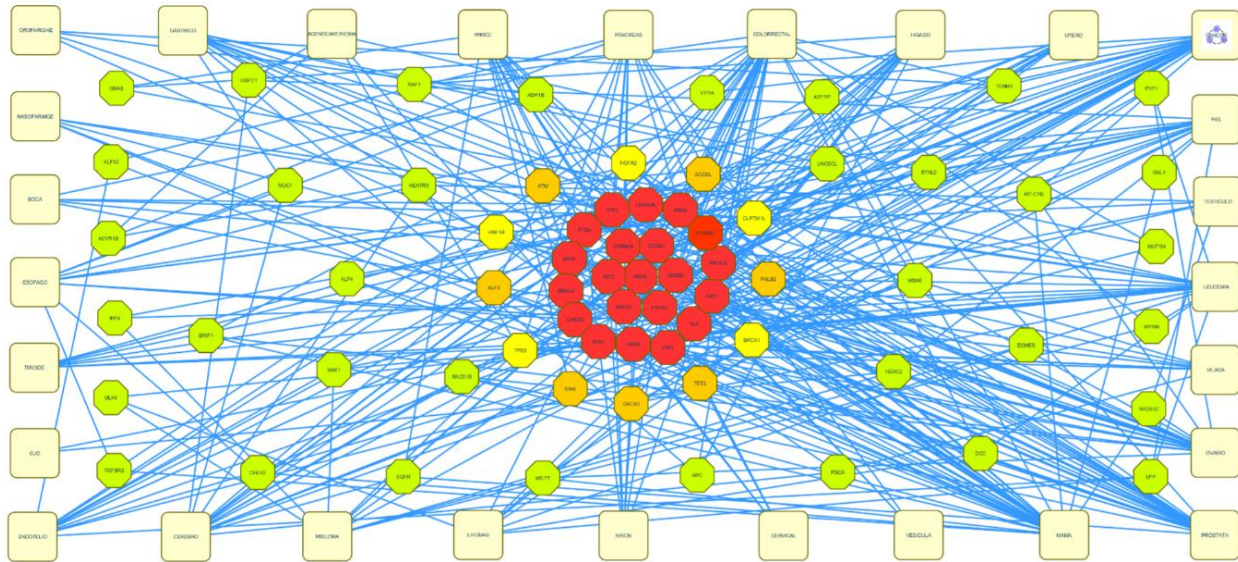


Figura 11. Mapa de interacciones entre los genes y tipos de cáncer, los genes en color verdes están relacionados con tres tipos de cáncer, en color naranja se representan a los genes que se asocian con cuatro tipos de cáncer, en color amarillo con cinco tipos de cáncer y en color rojo con seis o más tipos de cáncer.

Análisis de los genes involucrados con SNP relacionados con alguno tipo de neoplasia.

Con los 1468 genes que se han reportado que están asociados un tipo de cáncer se realizó un estudio de propiedades para poder determinar el efecto, se determinó que 414 genes son de moléculas que tienen una clasificación de acuerdo a su función celular en “Gen Ontology”, de los cuales 40.6% de los genes se clasifican de acuerdo a su función molecular en las rutas de señalización, entre los más importante es la ruta de WNT (Figura 12 Panel A), de igual manera se determinó la clasificación de estos genes en la categoría de procesos biológicos y 873 genes pudieron clasificarse en esta categoría de “Gen Ontology”, en donde alrededor de una cuarta parte (27.9%) de los genes se clasificaron en los procesos de crecimiento celular (Figura 12, Panel B). Adicionalmente, los 1486 genes se clasificaron en “Gen Ontology” en la categoría de componente celular (Figura 12, Panel C), solamente 347 genes están clasificados en esta categoría, de los cuales 17.6% están clasificados en el grupo de proteínas de unión a ADN, estas proteínas tienen un efecto en la regulación de la transcripción y

replicación del ADN. Como podemos observar los genes seleccionados están reportados como genes que tienen un efecto en el crecimiento y en la división celular.

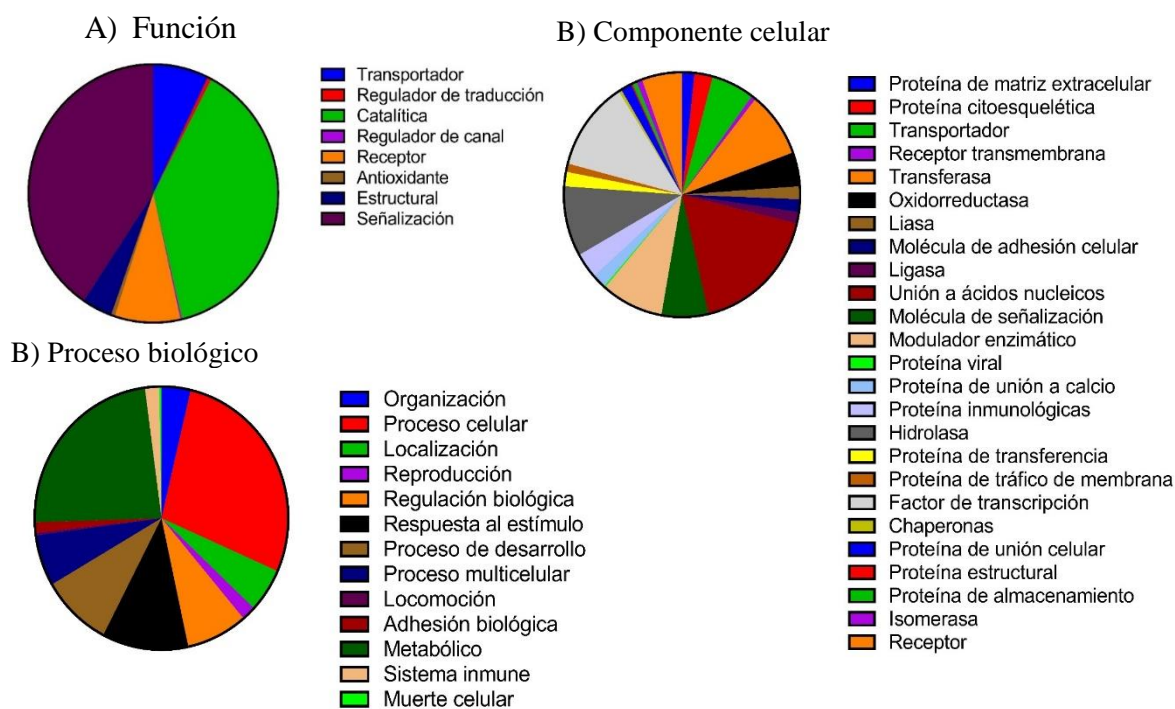


Figura 12. Análisis de las propiedades que tiene los genes relacionados con cáncer que se expresa en sangre. Se analizaron un total de 1468, de los cuales 414 genes tienen una función celular, 873 genes con efecto en el proceso biológico y 347 son componente celular. El panel A son función celular, Panel B son componente celular biológicos y Panel C son procesos biológicos.

Obtención de polimorfismos relacionado con cáncer

Hasta el día 28 de octubre del 2017 se identificaron 57,181 artículos de investigación relacionado con SNP, de los cuales solamente 43,780 SNP reportaban que los SNPs tienen una importancia clínica, el registro de estas bases de datos empezó desde el 2005, cuando se realizó el primer estudio GWAS con los criterios y parámetros que hoy en día son utilizados para realizar estudios de correlación de un fenotipo con una variación genética, la tendencia de los SNP que presenta una relevancia clínica va aumentando por cada año (Figura 13), viendo crecimiento tipo exponencial desde el año 2015.

Por la gran demanda de obtener cientos o incluso miles de secuencias en menos tiempos y más baratas, ha permitido la creación de una gran variedad de tecnologías que permite la obtención de más de una secuencia. Desde el año 2005 hasta la fecha se han usados varios equipos de secuenciación para determinar los SNP con un impacto en el fenotipo humano (Figura 14), las tecnologías más utilizadas para la secuenciación son los microarreglos de AFFLYMETRIX, que están diseñadas para detectar todas las mutaciones del ser humano mediante una hibridación con su cadena complementaria y la tecnología de ilumina, que permite leer miles de secuencia ADN en forma paralelas, lo que disminuye el índice de error de las bases, lo que permite tener un menor número de error durante su lectura, siendo el equipo más utilizado para detectar SNP¹²⁵. Otros equipos de secuenciación son los que pertenece a la empresa Perlegen, que tiene como principal característica la detección de SNP mediante dos haces de luz que permite determinar si presenta hibridación o no¹²⁶, hay otros tipos de secuenciaciones como por nanoporos, pero debidos a su poco tiempo de su diseño, todavía no ha tomado el lugar que le corresponde, pero se estima que en algunos años, sea unos de los equipos con mayor uso debidos a su facilidad de uso y precio económico que tiene.

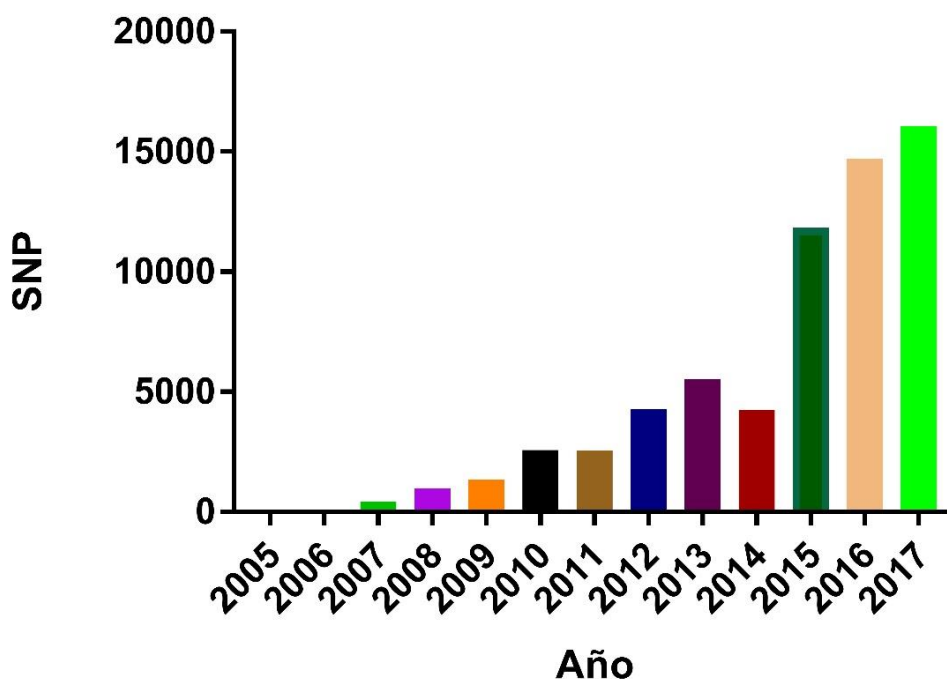


Figura 13. Número de SNP estudiado por año validados por análisis Gwas

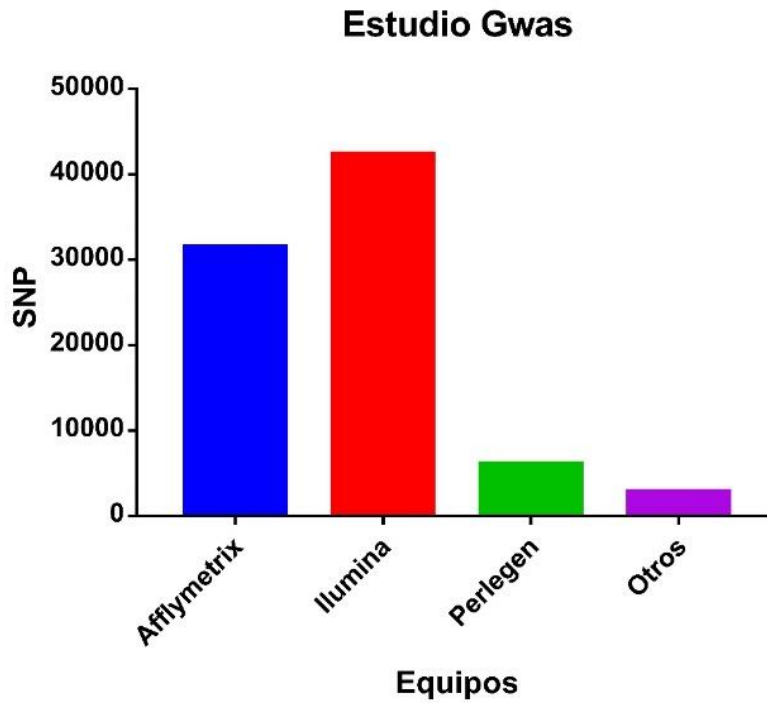


Figura 14. Equipos de secuenciación utilizados para la identificación de SNPs en genomas humanos

La tecnología de secuenciación en sus inicios era costosa tanto en insumos como en su infraestructura, por tal motivo la tecnología solo estaba disponibles para los continentes que presenta un índice de nivel socioeconómico más alto, explicando porque el 48% de los estudios GWAS son de países europeas (Figura 15), seguidos de los continentes de Asia y África con el 21 y 14% respectivamente, los SNP en población América representa solo el 13% de todos los estudios, el 3% de los estudios no tiene reportado de donde son originales sus pacientes, debido principalmente por ser estudios de enfermedades raras, siendo complicado pacientes con la enfermedad de un solo origen étnico.

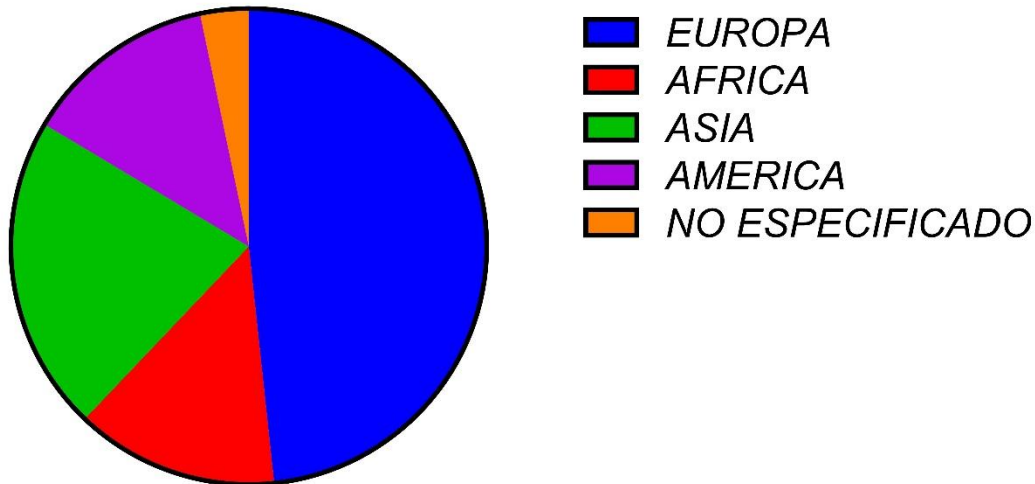


Figura 15. Análisis GWAS por población de estudio.

SNP asociados con una enfermedad

La versión del genomas humano denominado GRCh38.p11, tiene un total de 324,570,586 mutaciones o variaciones en sus genomas, de las cuales se han determinados hasta el 28 de octubre del 2017 un total de 43,780 mutaciones que presenta una alteración en el fenotipo (Figura 16), estos fenotipos puede ser tanto de protección o beneficio para la salud de las personas y mutaciones relacionados con una deficiencia o enfermedad, se tiene un total 17,549 mutaciones que tiene como características predisponer a una enfermedad, que corresponde el 40% de todos los SNP estudiados, siendo 15,903 SNP relacionados con alguna neoplasia. mientras que 32,027 de los SNP tiene un efecto protector a un fenotipo clínico o enfermedad siendo el 73% de los SNP analizados hasta el 28 de octubre del 2017, se presentaron un total de 5,796 SNP que tiene como característica predisponer a una enfermedad y al mismo tiempo tiene un efector protector o característica beneficiosa para la salud de los humanos.

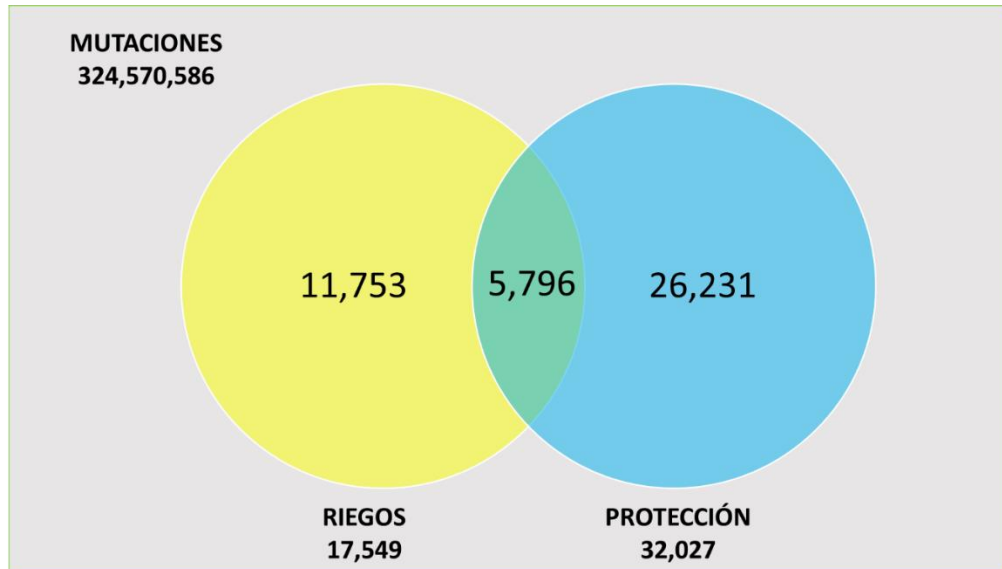


Figura 16. Diagrama de Veen de las mutaciones presente en el genoma humana

De las 43,780 mutaciones con varianza clínica, se seleccionaron todos los SNP relación con neoplasia y que predispone a la enfermedad que se puede detectar en los genes que son detectable en sangre, se obtuvo un total de 7,628 SNP reportados entre 31 tipos de neoplasias (Tabla 18), en 1468 genes estos SNP se encuentra ampliamente distribuida en el genoma, en la Figura 17 se presenta la posición en el genoma de cada SNP, se determinó que la gran mayoría de las mutaciones que se han validados por los estudios GWAS se encuentra en regiones que se transcriben, posiblemente es debido a los costos de secuenciación del genomas completos la mayorías de los estudios se obtuvieron mediante exomas, se tiene un total de 726 SNP que tiene la capacidad de regular la expresión del transcriptos (209 rio abajo, 176 rio arriba, 41 regiones reguladoras y 300 intergenicas), los SNP que están relacionados splice alternativas son 317 (3 splice donador, 227 splice acceptor y 87 mutaciones en los intrones que permite se forme un splice alternativo con la introducción del intrón al transcripto maduro). Se han reportados un total de 2000 mutaciones en los intrones que esta relacionados con factores de transcripción, se tiene un total de 4439 SNP en el transcripto maduro (41 en las regiones UTR y 4398 en los exones).

Se puede detectar un máximo de 4,668 SNP relacionados con una neoplasia, siendo el 61.1% de las mutaciones que se han relacionado con cáncer que tiene un estudio de correlación de la enfermedad con una razón de momio mayor a 1 con un intervalo de confianza del 95%

mayor de 1.01, los datos de SNP se puede descargar en el siguiente enlace;
<https://1drv.ms/x/s!Ao7pR1PQEuAJt2pynB4Y000ydpap>

Tabla 18 Cáncer que se ha determinados SNP en estudios Gwas

CANCER	Subtipos	GENES	SNP	Artículos
Adenocarcinoma	2	9	9	10
Biomarcador de cáncer	1	29	62	4
Cáncer cervical	1	6	21	4
Cáncer colorrectal	2	178	1543	35
Cáncer de boca	1	15	18	2
Cáncer de cerebro	3	78	144	18
Cáncer de cuello uterino	1	9	29	5
Cáncer de esófago	4	69	101	9
Cáncer de hígado	1	21	243	7
Cáncer de huesos	2	9	10	2
Cáncer de laringe	1	6	6	1
Cáncer de mama	1	156	5568	50
Cáncer de ojo	1	3	102	1
Cáncer de orofaringe	1	11	11	1
Cáncer de ovario	4	131	4346	9
Cáncer de páncreas	1	65	91	9
Cáncer de paratiroides	1	1	12	1
Cáncer de piel	5	125	183	11
Cáncer de próstata	1	235	498	31
Cáncer de pulmón	2	100	2167	24
Cáncer de riñón	3	20	100	6
Cáncer de testículo	1	22	45	9
Cáncer de tiroides	1	24	80	5
Cáncer de vejiga	1	33	57	11
Cáncer de vesícula	1	3	15	2
Cáncer gástrico	2	25	86	6
Cáncer nasofaringe	1	16	32	4
Cáncer del endometrio	1	27	66	7
HNSCC	1	65	290	2
Leucemia	4	129	234	20
Linfoma	3	70	100	11
Mieloma múltiple	1	66	121	6

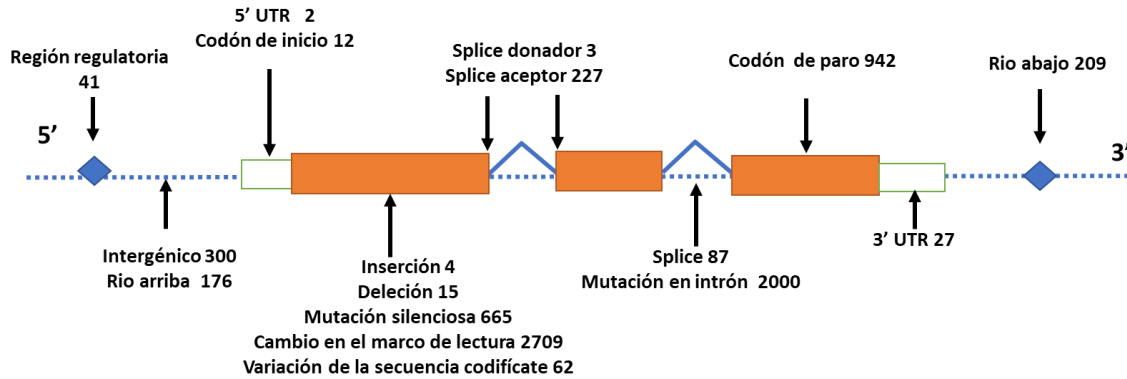


Figura 17. Posición y efecto que tiene cada uno de los SNP relacionados con cáncer

Posición cromosómica de cada una de las mutaciones.

Las 4,668 SNP que se obtuvieron se realizó un mapeo cromosómico para determinar en qué genes se encuentra la mayor cantidad de genes relacionados con cáncer (Figura 18). Los genes con el mayor número de mutaciones relacionados con cáncer están localizados en los cromosomas 7 y 17, posiblemente esto se debe a que estos cromosomas se encuentra los genes BRAF y BRCA respectivamente, que son dos de los genes más relacionados con cáncer, BRAF se asocia con la regulación de la proliferación y ciclo celular, mientras que BRCA produce una proteína de reparación del ADN.

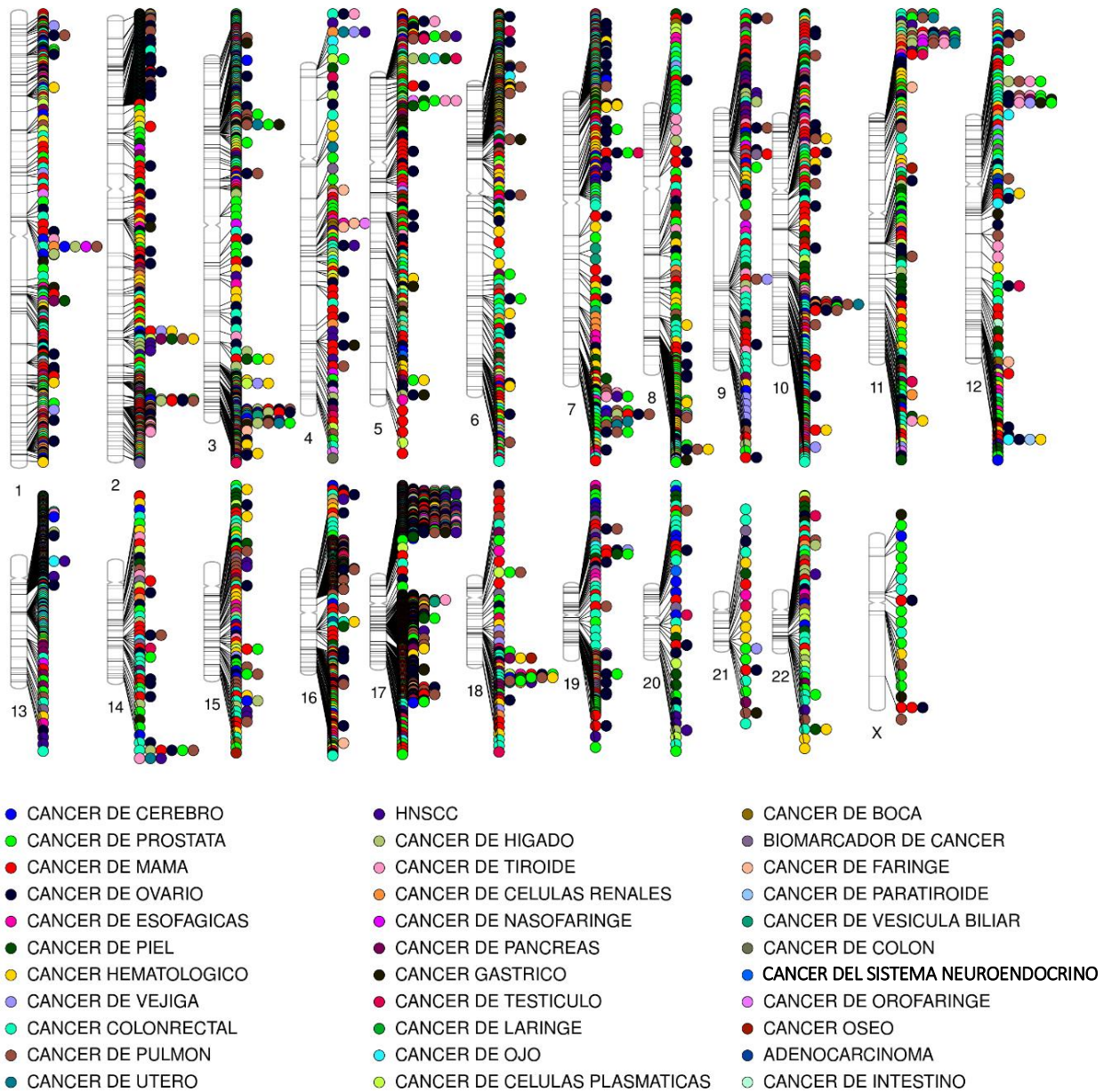


Figura 18. localización cromosómica de los SNP relacionados con 31 cáncer en el genoma humano. Los genes que presenta más de 500 SNP en su estructura se agruparon como SNP en el gen para cada tipo de cáncer.

PCR

Calidad de la extracción del ADN para la secuenciación.

Se realizó la extracción de ADN con 200 ml de sangre de cada una de los pacientes, siguiendo el protocolo de la casa comercial para la extracción de ADN, se tomaron un volumen total de 5 ul de cada una de las muestras y se corrieron en un gel de agarosa al 1.5%, el patrón de bandeo obtenido mediante la electroforesis muestra que no hay presentaba una degradación significativa visible en el gel de agarosa (Figura 19), indicado una buena calidad de las muestras.

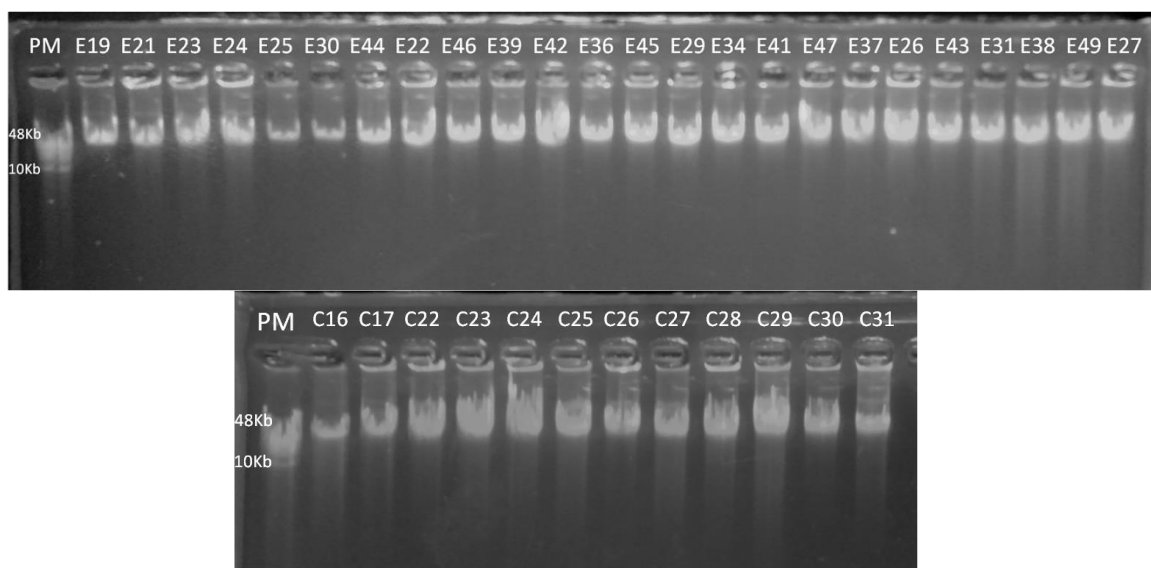


Figura 19. Integridad del ADN genómicos de las muestras de sangre corrida en un gel de agarosa del 1.5%.

PCR

Con los oligonucleótidos diseñados por Geno+, se realizaron la PCR, los oligonucleótidos tienen como objetivos amplificar diversas partes del genoma, produciendo de manera teórica un barrido de producto que abarca desde los 100 a 600 pares de bases de producto. Por tales motivos se espera un barrido de producto en dicho rango, en la Figura 20 se presenta dos muestras con el barrido esperado, indicando que se tiene productos de diferentes tamaños que abarca entre 100 a 1000 bases.

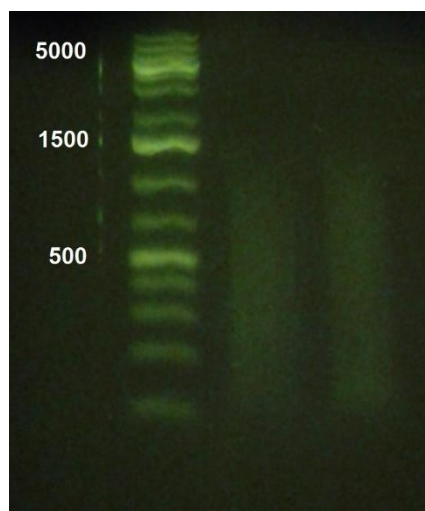


Figura 20. Barrido de productos de PCR con los oligonucleótidos de genomas.

Diseño de oligonucleótidos para la validación de la técnica.

Con los productos de la PCR se realizó una Reamplificación con los oligonucleótidos con los 6 pares de oligonucleótidos para la validación (tabla 19), los productos de PCR no superan las de 200 pb.

Tabla 19 Oligonucleótidos para PCR y qPCR

Gen	SNP	Nombre	Oligonucleótidos	Aplicones
FTO	rs8050136	FTOF	AAAAATTTTCCAAGCATTCCATGAGTCC	170
		FTOR	TAAACTCGGTATTTGATTTCCCTTTTCCC	
UMOD	rs4293393	UMODF	AAACACCCATTCTCATGAGATCAGC	175
		UMODR	GCTATTTAAGTGCAAATTTATTTTCGCCTCC	
APOA5	rs2736428	APOAF	AAACTCACCCACAATCCCCACC	180
		APOAR	GCCTACTGATGCCCATGCTGC	
SLC44A4	rs1130214	SLC44A4F	GGAAAGCAAAGAAATTCAAACATGAGG	176
		SLC44A4R	GGGCGCTGTGGTTTAGGAGG	
DPP10	rs7581057	DPP10F	AACCTTCTCAAGTGCTGTGCCTGC	180
		DPP10R	AGGTTAAGTACCAACAAAGGAAGCTCC	
HIF1A	rs11549465	HIF1AF	AACTGTGCTTTGAGGACTTGCGC	196
		HIF1AR	ATGTTAGCTCCCTATATCCCAATGG	

Validación por PCR.

Los 6 SNP se realizaron por triplicados, para confirmar su tamaño en un gel de agarosa al 2%, siendo la línea roja, la posición del marcador de pesos de 200 pb, indicando que todas las muestras amplificaron para cada SNP.

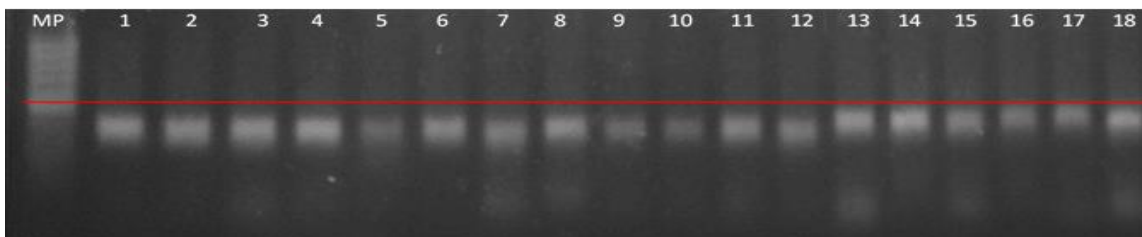


Figura 21. Gel de Agarosa 2% de los 6 SNP, carril 1 al 3 son para el SNP del gen *FTO*, carril 4 al 6 son para *UMOD*, del 7 al 9 son para el SNP del gen *APOA5*, del 10 al 12 son para *SLC44A4*, 13 al 15 para *DPP10* y los carriles 16, 17 y 18 son del gen *HIF1A*.

Validación por qPCR.

Los mismo 6 SNP se diseñaron para poder ser utilizados para PCR y qPCR, para confirmar los productos, cada muestra se ejecutó por triplicado, y se graficó el promedio de cada uno, se obtuvo una curva característica de una muestra positivas, mientras que los controles negativos y los controles sin muestras no presentaron una fluorescencia mayor de 50 unidades, indicando que los oligonucleótidos no forma dinero con ellos mismo y que son selectivos.

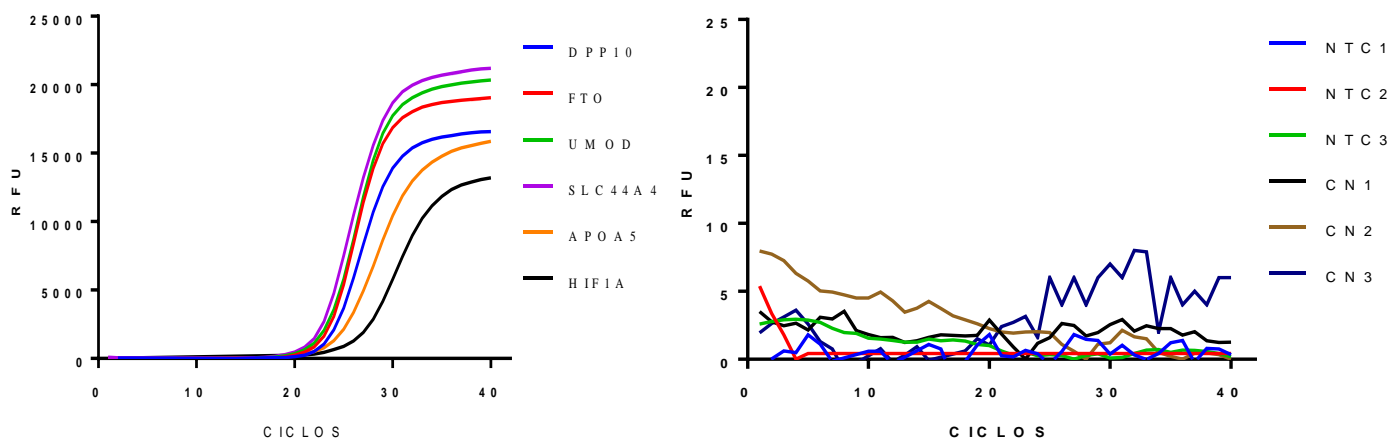


Figura 22. Validación de la PCR en formato Multiplex mediante qPCR. NTC1, NTC2 y NTC3 son reacciones de qPCR sin muestra biológica y los CN1, CN2 y CN3 son reacciones de PCR con muestra biológica de ADN de bacteria

CONTROL DE REACCIÓN

De las 4 muestras y de los dos controles fueron confirmadas por dos equipos diferentes (Figura 23); por qPCR, que permitió ver si existe amplificación exponencial en el termociclador y determinar si los controles negativos no presenten falsos positivos, y por el bioanalyzer, que permite comprobar el tamaño de los productos y la concentración de ADN en la muestra.

Todas las muestras fueron positivas, los controles negativos no presentaron una fluorescencia mayor de 100 RFU, determinado que los oligonucleótidos no realiza dímeros ni amplifica entre ellos mismo, lo que lo convierte en unos excelentes oligonucleótidos para usarlos como control de la secuenciación. El tamaño esperado de los productos de PCR es aproximadamente de 570 pares de bases, en el gel in silico obtenidos por el bioanalyzer se observa que todos tiene un tamaño aproximados de 600 pb, lo que indica que las muestras obtenidas son apropiadas para realizar la secuenciación. Las secuencias de los controles se procesaron con el programa Epi2me de nanopore que clasifica las lecturas obtenida para determinar si los productos de PCR pertenece exclusivamente a bacterias, los datos obtenidos no se presentó ninguna lectura de humanos, solo de bacterias (Tabla 20), lo que indica que el equipo de secuenciación permite la identificación correcta de secuencias.

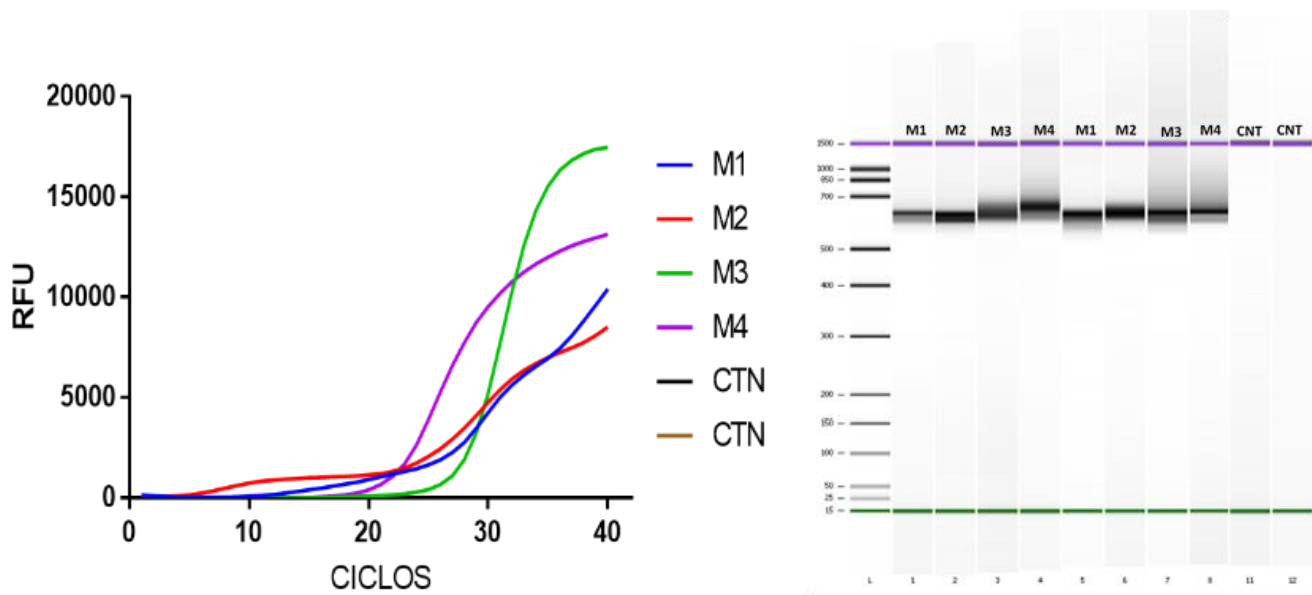


Figura 23 Controles negativos, figura A es el resultado de la qPCR, B es el gel generado por el bioanalyzer, siendo M1, M2, M3 y M4 las muestras de heces fecales y CTN son controles sin muestras de heces fecales (control sin muestra).

Tabla 20 Microorganismo detectado en la muestra de heces fecales

Taxon	Precisión
Lactobacillus oryzae JCM 18671	Especie
Prevotella copri DSM 18205	Especie
Citrobacter freundii ATCC 8090	Especie
Dialister succinatiphilus YIT 11850	Especie
Citrobacter freundii ATCC 8090	Especie
Dialister succinatiphilus YIT 11850	Especie
Lactobacillus paracasei	Especie
Lactobacillus casei	Especie
Lactobacillus zeae	Especie
Lactobacillus rhamnosus	Especie
Lactobacillus camelliae	Especie
Helicobacter canadensis	Especie
Helicobacter ganmani	Especie
Helicobacter japonicus	Especie
Helicobacter rodentium	Especie
Oscillibacter valericigenes	Especie
Helicobacter mesocricetorum	Especie
Prevotella oulorum	Especie
Prevotella buccalis	Especie
Citrobacter murlinae	Especie
Citrobacter braakii	Especie

SECUENCIACIÓN

Se secuenciaron un total de 94 muestras, 90 muestras con el equipo MinION, y 4 por ilumina (3 en el Hiseq del IBT y uno de MacroGen usando el Hiseq). Para preparar las bibliotecas las muestras se cuantificaron con el cubit versión 2.0, se ajustó para tener una concentración mínima de 30 ng/ul.

Los datos obtenidos fueron en formato FAST5, para su análisis fueron transformado a formato FASTQ.

Proceso de las lecturas en bruto

Se compararon la calidad de las lecturas entre el ilumina y el MinION (Figura 24). En el caso de las lecturas de ilumina se utilizó una Q30, mientras que en el MinION se obtuvo una

calidad de Q15, las lecturas de ilumina presenta una mayor calidad las del MinION, estos es debido a la modalidad que ilumina realiza secuenciación en formato pareada, permitiendo realizar la lecturas en ambas hembras de ADN, lo que disminuye el error en la secuenciación, el único inconveniente que presenta este estilo de secuenciación es que están limitado a un rango de tamaño, cosa que no sucede con la secuenciación por MinION, nanopore recomienda en sus protocolos, que para asegurar que se obtuvo una buena corrida de datos una Q12, la secuenciación obtenidas por el MinION cumple con los requisitos recomendados por la empresa para poder utilizar los datos para realizar el análisis informático.

Análisis de las secuenciación y validación de la técnica.

Debido a la poca frecuencia de cada uno de los SNP, no se realizó la búsqueda de las mutaciones puntuales porque posiblemente más del 80% de los SNP no lo podremos localizar de manera adecuada, para solucionar se realizó un script para la búsqueda de 25 bases consecutivas antes y después de la mutación seleccionar, para lograrlo identificar el sitio de mutación, se realizó un script que permite buscar las bases antes y después de la mutación con una similitud mínima de 90%.

Secuenciación

Debido a los millones de datos que se generas por cada muestra y poder simplificar el análisis, se realizó el seguimiento de solo 6 SNP para 82 muestras, de este modo se realizó la validación del equipo de MinION para realizar un panel de secuenciación. Para la muestra de cáncer de tiroides, que se secuencio por dos equipos distinto para determinar la sensibilidad y especificidad, se realizó la búsqueda de todos los SNP posible entre los dos equipos con el script diseñado, para el exoma comercial de MacroGen, se tiene garantizados un mínimo de 90,000 posiciones de SNP, realizando limpieza de los datos se obtuvieron un total de 99,995 SNP, de los cuales se usaron para buscar todos los SNP relacionados, en el panel comercial se tiene presente un total de 3456 SNP en el exomas de los 4439 SNP siendo un poco más de 75% de los SNP que se puede detectar en teoría con el panel de secuenciación, y esta relacionados con un tipo de neoplasia, el número de SNP sea menor en estos paneles comerciales para patologías posiblemente porque el kit fue realizado y estandarizados en el año 2010, y varios SNP todavía no se habían descubierto.

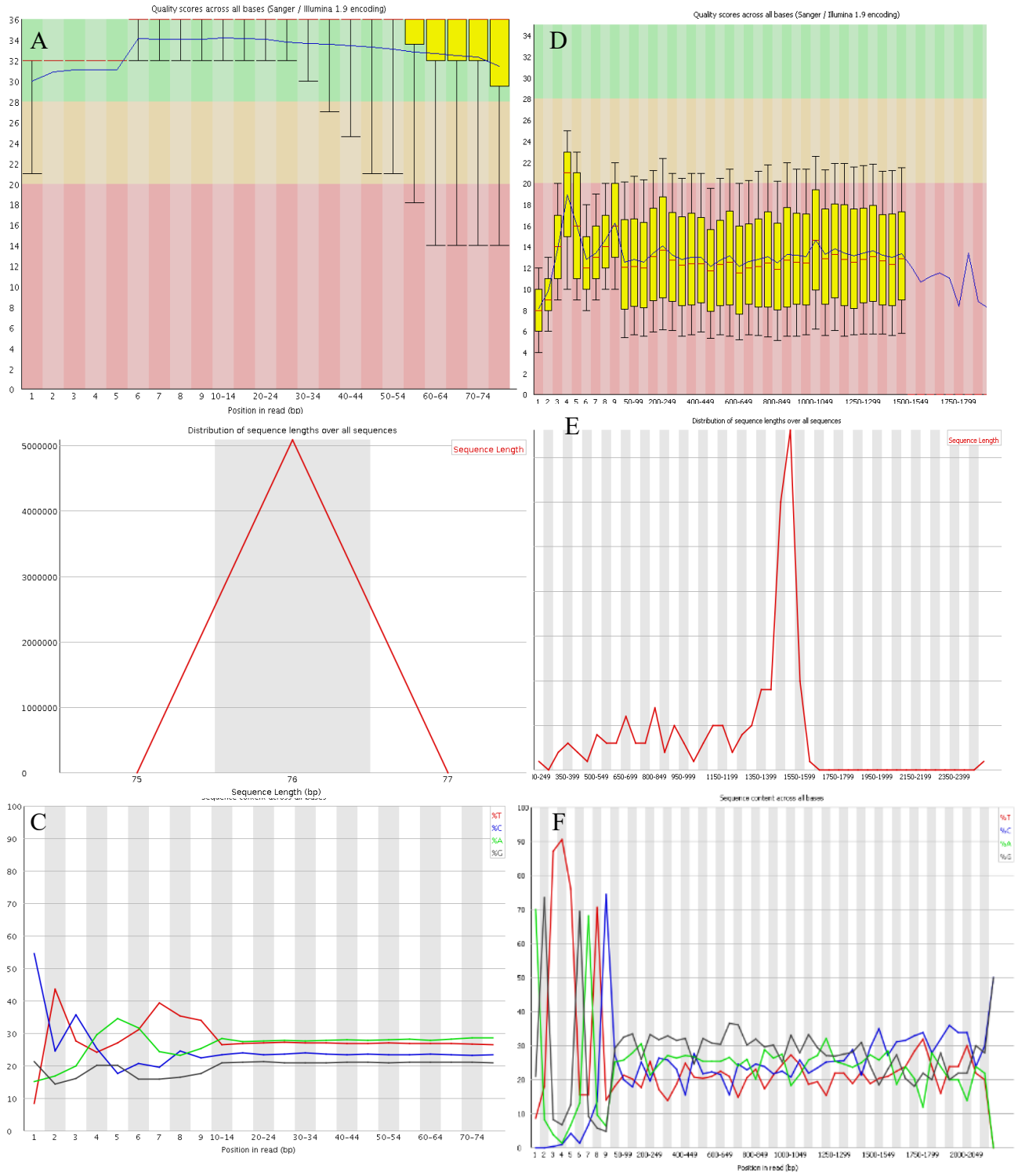


Figura 24. Comparación de calidad de las lecturas entre MinION, la casilla A, B y C pertenece al Illumina, casilla D, F y G son los datos obtenidos por MinION, A y D son los puntajes de calidad de las lectura, B y F muestra el tamaño promedio de las lecturas, C y F muestra el contenido de bases nitrogenado, siendo Timina rojo, Citocina azul, Adenina verde y Guanina negro, todos en porcentaje.

En el panel de secuenciación diseñada para la tesis se encontró en las 83 muestras los 6 SNP, sin importar el técnico o el día en que se realizó, por tal motivo se puede utilizar el MinION como un equipo para la detección de secuencia de ADN para usarlo como un panel de detección de SNP.

Debido a que la lista de todos los SNP que se puede detectar en el MinION con importancia clínica en neoplasia es muy grande (3456 SNP), se muestra en la Tabla 21 un ejemplo de cómo se tiene los datos en la base de datos, para poder visualizar la lista completa, una copia de los datos se colocó en la nube de OneDrive en formato xlsx con el siguiente link: <https://onedrive.live.com/edit.aspx?cid=09e012d05347e98e&page=view&resid=9E012D05347E98E!12687&parId=9E012D05347E98E!12686&app=Excel>

Tabla 21 SNP con relevancia clínica en cáncer que son detectable por el protocolo diseñado por la tesis

TIPO DE CANCER	GEN	SNP	RM	CL de 95 %
Mama	BRCA1	rs80357417	1.4	[1.2-1.5]
Mama	BRCA1	rs80357792	1.11	[1-1.26]
Mama	BRCA1	rs886040078	1.11	[1.02-1.24]
Ovario	BRCA1	rs80357792	1.3	[1.1-1.5]
Ovario	BRCA1	rs886040078	1.6	[1.4-1.67]
Mama	BRCA1	rs80357788	1.6	[1.5-1.66]
Ovario	BRCA1	rs80357788	1.3	[1.2-1.4]
Pulmón	BRCA1	rs397509014	1.0309278	[1-1.08]
Colorrectal	PMS2	rs63750871	1.04	[1.04-1.10]
Mama	BRCA1	rs80358063	1.04	[1.01-1.08]
Ovario	BRCA1	rs80358063	1.04	[1.01-1.08]
Pulmón	BRCA1	rs397508973	1.04	[1.00-1.08]

Para confirmar que el algoritmo funciona correctamente se sometió a un análisis, se seleccionaron un total de 10 SNP que fueron detectados en la secuenciación, fueron procesadas y se recortaron a 51 bases que corresponde a una bases de la mutación y 25 bases en ambos sentidos, a las lecturas se les realizó un blast para determinar si las secuencias no se encuentran repetidas en otra zona del genomas para confirmar que los resultados obtenidos no es un falso positivos,

se determinó que las secuencias localizadas corresponden a un polimorfismo registrado en la base de datos de variaciones genéticas (Ensembl). Se obtuvo que todas las lecturas cumplen con los criterios establecidos y se tiene un nivel de confianza significativo en el procesamiento de los datos. Unos de los SNP que se seleccionaron es el RS121913327 que es una mutación del gen APC, fue localizada en las lecturas del MinION, con una identidad de 96%, las secuencias de manera individual, se le sometió a un blast, se obtuvo que la secuencia tiene una hibridación del 96 con mRNA del gen APC, de igual manera se realizó un CLUSTAL entre la lectura y el SNP seleccionado, se obtuvo una gran similitud entre las lecturas, el CLUSTAL permite ver la posición de la mutación (Figura 25), se utilizó una herramienta de NCBI de la página Emsabl para determinar la posición en el cromosoma de las 51 bases, se determinó que la lectura proviene del brazo largo del cromosomas 5, que corresponde al gen APC, confirmando que con 50 bases de identidad puede confiarse en la detección de regiones donde se presenta la mutación.



Figura 25 Análisis de las lecturas que contiene el SNP rs121913327, A) es el blast de la secuencias, B) clustal entre el SNP contra la lectura recortada que presenta la el rs121913327 y el SNP, en él se puede ver la bases que presenta la variabilidad C) análisis de posición de secuencias con variaciones con el programa ensambl.

Con los datos obtenidos, se confirmó que las lecturas son confiables, se realizó el cuadro 2x2 para determinar la sensibilidad y especificidad de la técnica comparando con la técnica de oro (secuenciación de exomas por MacroGen). En la Tabla 22, se presenta los resultados de la detección de SNP, si no detecto un cambio de bases en la posición de la mutación se tomó como normal, si detecto una mutación se considera SNP.

Tabla 22 Resultados detección de SNP

		Técnica de referencia		
		SNP	No mutación	Total
Técnica diseñada	SNP	161	17	178
	No mutación	50	3228	3278
Total		211	3245	3456

A partir de los datos obtenidos se calcularon todos los parámetros para calcular la robustez de la técnica, los resultados obtenidos se presentan en la Tabla 23.

Tabla 23 Parámetros que se calcula con la comparación de dos técnicas

Parámetro	Valor	Porcentual
Especificidad	0.99476117	99.47%
Sensibilidad	0.76303318	76.30%
Exactitud	0.98061343	98.06%
Valor Predictivo Negativo (VPN)	0.9847468	98.47%
Valor Predictivo Positivo (VPP)	0.90449438	90.44%

La técnica presenta una exactitud del 98%, volviendo la técnica para detección de SNP mediante Nanoporos confiables para uso clínico, aunque esto solo se realizó con dos muestras debido a su alto costo, para poder llevar la técnica a su aplicación clínica se requiere realizarlo con todos los SNP seleccionados para la tesis, pero da un panorama alentador para su traslocación como una técnica de diagnóstico preventivo. Con eso datos se calculó la curva ROC.

Curva ROC.

LA Curva ROC fue calculada para obtener el área bajo la curva y determinar la precisión de la técnica diseñada comparadas con una técnica estándar (Figura 26), el área bajo la curva

fue de 0.9446, es decir que la técnica puede ser usado como una técnica para la detección de SNP en paciente entre 18 a 65 años sin importar el sexo.

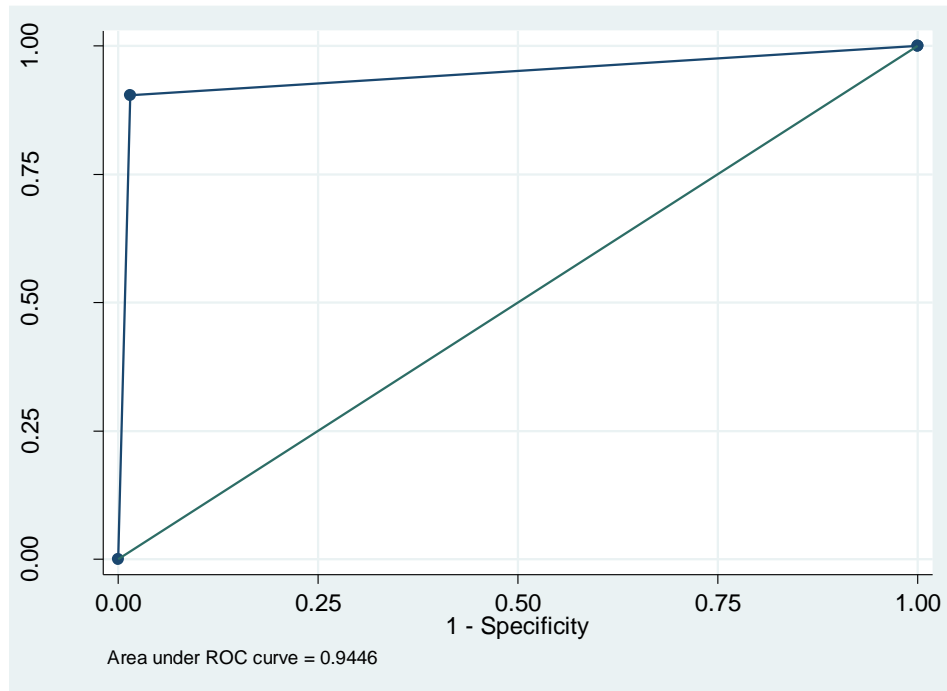


Figura 26 Curva ROC para la secuenciación por MinION en comparación con la secuenciación de ilumina por Hiseq

Análisis de costos

El kit comercial de exomas de MacroGen tiene un precio al mercado de 300 Dólares más IVA, debido al cambio de moneda oscila entre 6500 a 7200 pesos mexicanos, mientras que el panel diseñado tiene un costo de procesamiento de 725 dólares por muestras (13500 a 15000 pesos mexicanos), siendo dos veces más costosas que las se encuentra en la actualidad en el mercado solo realizar la técnicas, el panel que se diseñó todavía falta entrar al proceso de optimización para disminuir los costos de secuenciación y mejorar los tiempos de procesamiento de los datos para poder competir con kits comerciales que domina el mercado de la medicina preventiva mediante análisis de SNP.

Circos Plot

El análisis Circos plot nos permito correlacionar los 31 tipos de cáncer analizados con los cromosomas, no se incluye el cromosoma Y, porque en la base de datos que se generó, no se obtuvo ninguna secuencia del cromosoma y que cumplía con los requisitos establecidos para realizar el análisis, la información que nos permite tener esta base de datos nos permite visualizar de manera más claro los datos obtenido en la minería de datos.

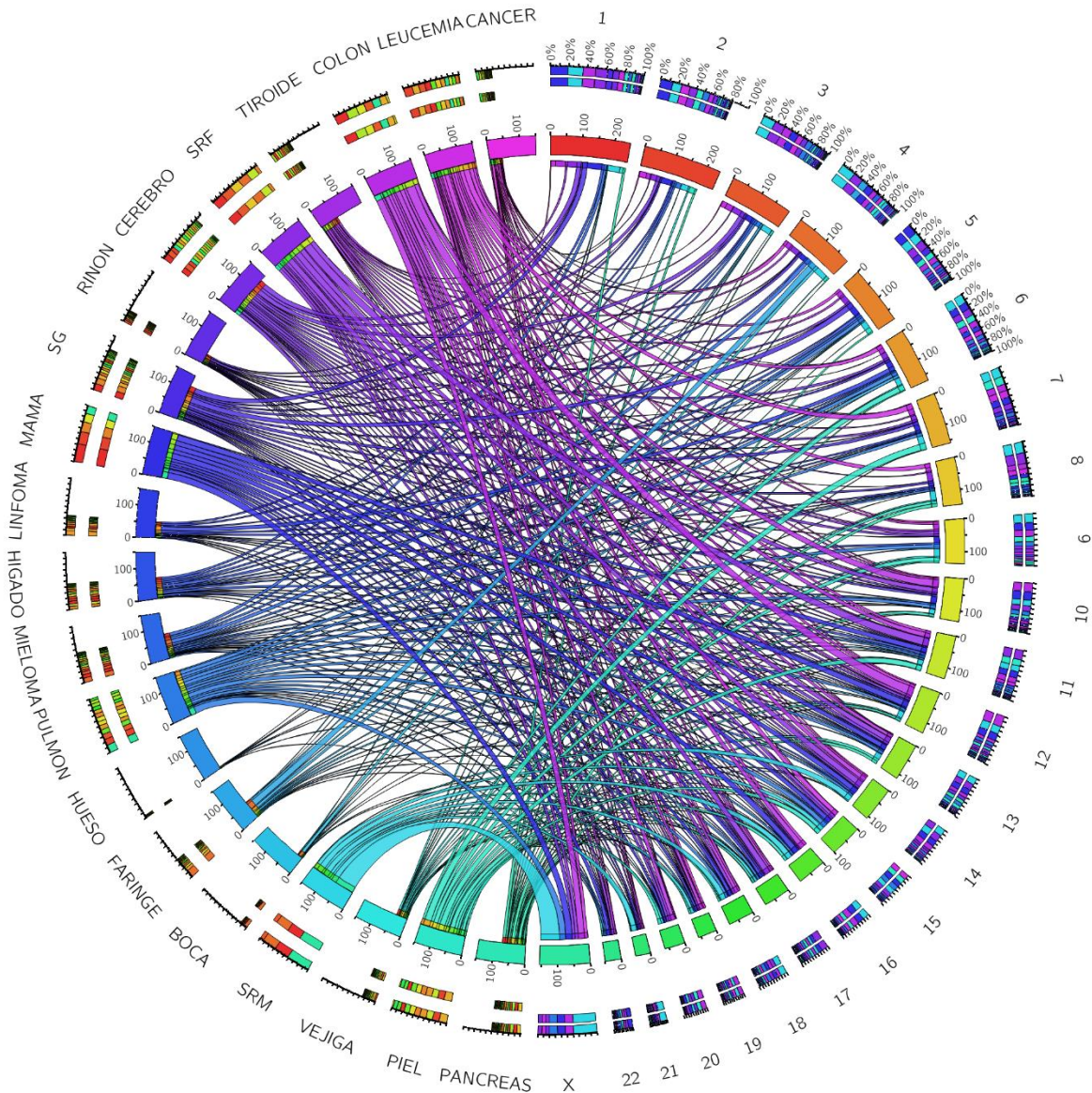


Figura 27 Circos plot para la interacción de los cánceres con los cromosomas, SRM es cáncer del sistema reproductor Masculino y SRF es cáncer del sistema reproductor femenino. El grosos de las líneas indica el número de mutaciones que se relacionada con cada cromosoma.

Modelo booleano

A partir de los datos obtenidos de las bases de datos de cáncer y del modelo de metástasis, se diseñó el siguiente modelo Booleano que tiene como finalidad poder determinar si el paciente mediante simulaciones tiene posibilidad de que la neoplasia que se puede generar tiene un índice alto o bajo de metástasis.

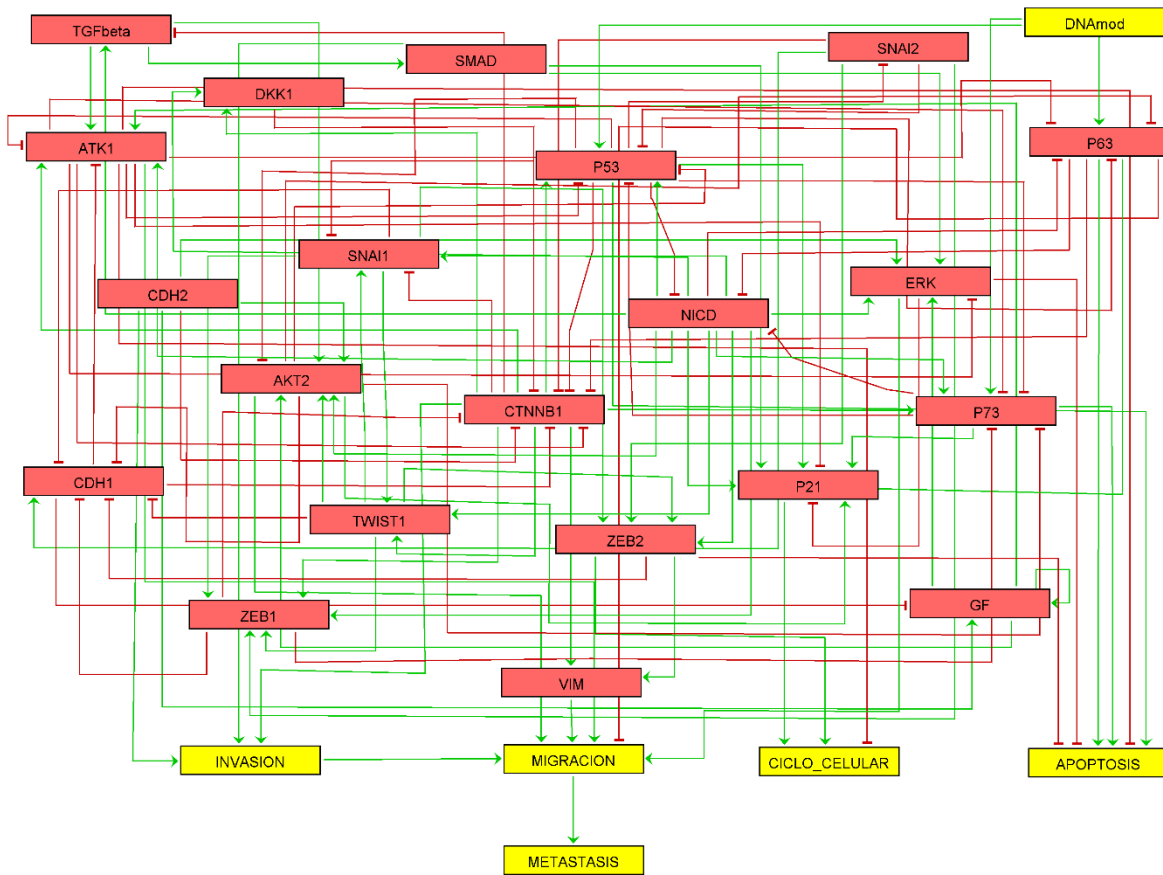


Figura 28 Modelo Booleano, las líneas rojas son efecto negativo para el gen, línea verde son efecto positivo.

DISCUSIÓN

Para diseñar un panel de diagnóstico se deben de cumplir un mínimo de muestras para su validación¹¹⁷, por los parámetros establecidos para la tesis se obtuvo que 59 muestras es lo mínimo que se requiere para tener una confiabilidad de 95% de los datos cuando presenta una distribución binomial, pero para llevar una técnica de biología molecular al mercado debe de pasar por varios filtros o protocolos que requiere un número mayor de muestra¹²⁷, para determinar el nuevo número de muestras que se requiere, se debe de determinar la población objetivo, tiempo y disponibilidad de la muestra, determinar la población objetivo no fue abordado en la tesis, debido que se tendría que realizar una estandarizar para cada tipo de cáncer que se quiere aplicar la técnica, la estandarización para algunos tipos de cáncer en específicos en población mexicana, queda como un trabajo a futuro.

Se recolectaron un total de 83 muestras sanguíneas de personas de un rango de edad de 18 a 65 años, ninguna muestra presento hemolisis o lipemia, las muestras fueron suficientes para confiar que se tiene una muestra representativa para determinar que la secuenciación por MinION tiene futuro para utilizado como un equipo para diseñar un panel de diagnósticos para Cáncer, actualmente se está usando la secuenciación por nanoporos para detectar variaciones genéticas en Cáncer, el grupo de investigación Gardner en el 2018, crearon una técnica para determinar todas las posibles variaciones en los 27 genes de importancia para cáncer, para su validación se utilizó un total de 630 individuos, que presentaba descendencia de varios tipos de carcinomas¹²⁸.

Más del 25% de los genes se encuentra expresados en sangre y es posibles de detectar con técnicas de biología molecular a partir de una biopsia sanguínea, los artículos que por lo general se reporta en la página de GEO, proviene de microarreglos, la minera de datos nos permito determinar que los cromosomas 1,6 y 19 presenta el mayor número de genes que se expresa en sangre. El contenido de genes de cada cromosoma se estima y/o se sospecha que es de modo azaroso debido a que no hay un patrón determinado para establecer el orden y número de genes que tiene cada cromosoma, peros se han reportados islotes o conjunto de genes que está relacionado en sus funciones y se encuentra en el mismo cromosoma, un ejemplo es el brazo corto del cromosoma 6, en él se encuentra todos los genes que perteneciente al complejo mayor de histocompatibilidad¹²⁹, que son expresados de manera

constante por las células del sistema inmune como los linfocitos que se encuentra en circulación en sangre¹³⁰.

De los 15744 genes que se expresan en sangre, se determinó que el 9.32% (1468 genes) tiene por lo menos una relaciona con alguna neoplasia, la mayoría de los genes que afectan tienden a ser genes dianas o genes involucrados en vías de señalización celular, lo que indica la importancia que la célula que entra a un proceso de neoplasia tiene que tener la capacidad de inhibir las señales del organismo para su regulación y crecer de manera independiente^{18,131,132}, unos de los principales genes más involucrados en el desarrollo de una neoplasia son los genes TCF1 y APC, que son genes relacionados con rutas de señalización de WNT, la ruta de WNT regula la capacidad de la célula en proliferar y diferenciarse, por tal motivo casi siempre se encuentra alterada en las neoplasias²⁰, por eso las mutaciones en estos genes tiene una gran impacto en la formación de la patología, también se encontraron otros genes con otras vías de señalización, como el gen NICD, que es un gen importante para la vía de NOTCH. Las alteraciones que se encontraron no fueron solo en señalización, también los genes que tiene que ver con la regulación de metabolismo, esto tiene gran importancia porque en todo proceso de neoplasia, la células tienen que cambiar su rutas metabólicas para poder tener biomasa y energía para su proliferación, el cambio metabólico más importante es el efecto Warburg¹³³, el efecto Warburg permite a las células cancerígenas obtener una fuente rápida de ATP para su rápido crecimiento, mediante la conversión de piruvato a ácido acético, esto no solo permite tener la energía sino que la síntesis de ácido acético permite a la células cancerígenas usarlo como señal para la angiogénesis y metástasis¹³³, lo que lo vuelve de suma importancia para la sobrevivencia del tumor. Para que una célula pueda entrar a un proceso de neoplasia debe adquirir todas las herramientas para empezar un proceso similar a la embriogénesis para diferenciarse y crecer de manera rápida e independiente.

En el estudio se encontró que en el cromosoma 17 tiene un conjunto muy grande de genes involucrados a neoplasia, estos genes están íntimamente relacionados con mantener la integridad del ADN¹³⁴, lo que indica que posiblemente actúen como islotes cromosomaticos, lo que vuelve el cromosoma 17 un blanco estudios. Los genes que más sobresale en el cromosoma 17 para la formación de un tomar son los genes; HER2, TOP2A, TAU, Tp53,

HIC-1 y BRCA1¹³⁵, mutaciones en ellos son asociados con Cáncer de células somáticas, principalmente en sarcomas¹³⁶⁻¹³⁸, como la piel. Estudios demuestran que anomalías en cultivos de tejido tumoral de pacientes con cáncer de mama, los cromosomas con mayor índice de mutaciones son; 3, 7 y 8¹³⁵.

Los 43,780 SNP que se tiene reportado en las bases de datos con importancia clínica que se han registrados desde el 2005 hasta el 28 de octubre del 2017, el 65% de los SNP detectados con importancia clínica se ha publicados en los últimos 3 años, posiblemente estos se deba por dos razones, primero en el año 2015 el grupo de trabajo del doctor Evan Eichler, perfeccionaron una técnica de secuenciación denominada secuenciación en tiempo real de moléculas únicas (SMRT, en sus siglas en inglés), permitiendo encontrar más de 22,000 variaciones genómicas que no era posibles detectar con las técnicas de secuenciación masivas de años, por el índice de error que presenta en regiones repetitivas y zonas ricas en triples enlaces¹³⁹, lo que despertó el interés de muchos investigadores para el uso de nuevas técnicas de detección de secuencias, y la segunda posible razón es que los avances en la tecnología y la promesa de la medicina personalizada o biodirigida, concibió que varias organizaciones relacionadas con el área de la salud como el instituto nacional del cáncer y gubernamentales como la casa blanca durante la presidencia de Obama, apoyara con recursos técnicos y económicos para poder realizar una base de datos de miles de personas para poder tener suficientes datos para poder realizar las correlaciones más exactas y poder tener tratamiento personalizados¹⁴⁰.

De los 1468 genes relacionados al cáncer que se expresan en sangre que son detectables en sangre, de manera teórica se pueden detectar unos 7,628 SNP con relevancia clínica en la patología que fueron reportados hasta el mes de octubre del 2017, siendo un total de 4,668 SNP que se localiza en regiones que se pueden detectar por secuenciación en regiones de codificantes, debido a que la secuenciación del genoma todavía sigue siendo de alto costo, la gran mayoría de las técnicas que se han ocupado para la detección de SNP se realiza por Exomas, explicando porque más del 60% de estas variaciones estudiado y relacionado con cáncer son detectable en transcritos, lo que puede sesgar el impacto que atribuye cada mutación por las limitantes que se tiene por las técnicas que predominan en la actualidad debidos a los costos.

Todavía no se tiene estimado el número de mutaciones requeridas para conducir o adquirir algún tipo de cáncer y tampoco se tiene el conocimiento si el número de mutaciones o tipos de igual para cada población del mundo^{121,141-143}, por tal necesidad se ha establecido un análisis llamado GWAS para tener un acercamiento para ver la posible contribución de cada mutación para el desarrollo, dando unos de los primeros pasos para la medicina personalizadas, lo que permite abrir una nueva ventana de la ciencias de la salud para mejorar el tratamiento y el diagnóstico.

Se tiene caracterizados SNP para 31 tipos de cáncer, estos cánceres, son los cánceres más comunes en población humana como es el de mama y de piel, pero también se han caracterizados los cánceres con la mayor tasa de mortalidad como es el cáncer de pulmón y los gliomas debido a su gran impacto en la sociedad.

Las variantes alélicas que se han reportado en bases de datos que presentaba razón de momios mayor con un intervalo de confianza mayor a 1, fueron usado para realizar el panel de genes, debidos a los pocos estudios que se han realizados en población Américas, solo el 16% de los SNP han sido probados en población del continente, más del 50% de los SNP fueron probados por descendencia europea, abriendo la posibilidad que los SNP relacionados con alguna neoplasia en población occidental no tenga un impacto clínico para la población mexicana. Para solucionar este inconveniente las bases de datos COSMIC y SNPedia solicitan en sus páginas, que los SNP que suban en sus servidores debe de validarse como mínimo en dos poblaciones de diferentes, aumentando la posibilidad de que el efecto de SNP se presente sin importar la étnica de la población, pero no existe ningún estadístico hasta el momento que permita validar el supuesto.

Se encontraron un total de 7,628 SNP relacionados a cáncer, y se determinaron un total de 16,557 correlaciones con alguna neoplasia, algunos SNP por lo menos se encuentra presente en 2 de los 31 tipos de cáncer, un ejemplo de la afirmación anterior el SNP rs397508058, que proviene del gen BRCA2, se ha determinado varias correlaciones mayores de 1.05 en cánceres de pulmón, de mama y de ovario, lo que nos puede indicar que la mutación tiene un efecto muy relevante para la formación de una neoplasia, los genes que se encontraron que presentan estas características son principalmente los genes que están relacionados con la

integridad del ADN, lo que nos permite reafirmar la hipótesis que los genes de regulación tienen un alto impacto en la formación de varios cánceres.

Se determinó que los genes con los mayores números de mutaciones en su estructura son los genes BRCA1, BRCA2 y BRAF, posiblemente las mutaciones que afectan de manera directa las funciones de estos genes aumentan la posibilidad de generar alguna neoplasia debido que componentes de los puntos de regulación del ciclo Biológico de la célula. Los genes de BRCA1 y BRCA2 tienen como función primordial realizar las reparar el ADN cuando sufre alguna mutación, mientras que BRAF normaliza la proliferación y crecimiento celular, cuando estos genes no cumplen su función de la manera más óptimas, pueden contribuir de manera significativa para que la célula pierda o inhiba los mecanismos de regulación celular y se vuelve en una célula neoplásica.

De total de SNP encontrados en bases de datos con importancia clínica, se presentaron 5796 SNP que presenta una dualidad entre un efecto protector y de riesgo a una patología en común, permitiendo formular dos posibles hipótesis para explicar este fenómeno, la primera es que los SNP requiere alguna condición adicional para aumentar o disminuir la probabilidad de presentar un fenotipo, por ejemplos alto niveles de grasas o presentar otros SNP al mismo tiempo, a estas características, hoy en día es un “top topic” de debate, como es que interactúan varios factores que modifica la probabilidad de un fenotipo, la discusión sobre como los demás factores genético interviene de manera significativo para un fenotipo se le denomina fondo genético. La segunda hipótesis que se puede generar de la contradicción de los datos es porque en los estudios realmente no se ha tomado una muestra representativa del universo, esta última hipótesis, es muy controversial, porque no existe realmente un criterio de cuantas muestras es representativas en una población a nivel genético, de manera general se estima que de un 5 al 10% de la población es un buen estimador de una población, el problema que a nivel genético un 5% de una población siendo imposible obtenerlo, solo para el caso de México, se necesitaría tener más de 7 millones de muestras por estudios para tener solo el 5% de la población, superando por miles de veces el número de muestras totales que se tiene por todos los estudios en población mexicana que se tiene reportado hasta el 2019. Todavía no existe un criterio estadístico que nos permita limitar realmente cuanto sería una muestra representativa para un estudio de SNP, pero existe algunos SNP debido a su alta incidencia

y su correlación fuerte con algunas neoplasias, se puede observar su efecto incluso en población pequeñas.

El ser humano tiene reportado en su genomas un poco más de 300,000,000 variaciones genéticas, presentando menos del 0.01% con un efecto en la vida de la persona de manera individual, pero no se ha determinado el efecto que se tiene por la presencia de todos las variaciones en un individuo, tal vez esto explicaría de manera más lineal porque algunas personas o poblaciones son más susceptible a contraer un tipo de cáncer o presenta un bajo pronóstico de recuperación, por ejemplo, en la leucemia linfocítica aguda en la población mundial tiene un 85% de recuperación luego de ser tratada, pero en la población mexicana el índice de recuperación no supera el 60%, permitiendo ver que en la población mexicana a pesar de ser la misma enfermedad, el pronóstico de recuperación es muy diferente^{23,144}, explicando de manera parcial porque algunos SNP presenta una dualidad en algunos casos para un fenotipo individual, dando un poco de peso a la hipótesis que algunos SNP para tener un efecto de protección o de predisposición requiere la presencia de otras variaciones genéticas o condiciones adicional que no se tomaron encuentra en la investigación para poder causar un fenotipo, por este motivo ya no se debe de realizar correlaciones de solo un SNP con un fenotipo a la vez porque no se sabe si el efecto de fondo genético influye de manera significativa a la patología.

Debido a los altos costo que se puede generar para validar cada uno de los SNP para validar un equipo, se seleccionaron un total de 6 SNP de manera azarosa para la tesis, para llevarlo al mercado, queda como perspectiva futura del trabajo.

Los SNP que se seleccionaron para la validación del equipo, demostraron que puede ser detectable sin tener un efecto significativo por las variables personal técnico y tiempo, lo que permite usar el equipo para la detección de varios SNP.

La secuenciación de la empresa MacroGen, reportaron un total de 99,950 SNP, de todos los SNP que ellos reportaran, podemos detectar un total de 3456 SNP relacionados con alguno tipo de cáncer seleccionados para la tesis, siendo el 77.85% de todos los SNP presente en las bases de datos hasta el 2017 relacionado con alguna patología, el porcentaje de los SNP que se encuentra en el panel de MacroGen un poco limitado en comparación al servicio que ofrecen es posible debidos a que la técnica es del 2012, lo que limito la técnicas a solo los

SNP caracterizados hasta esa fecha, para poder aumentar el número de SNP en la técnica de oro, se tiene que buscar técnicas más actualizadas para usarlo como referencia para aumentar el número de SNP que se puede comparar.

En general, las tasas de error parecen ser constantes en todas las secuencias¹⁴⁵, sin importar el SNP, de los 99,995 SNP ofrecidos en el exoma de la MacroGen, solo se detectaron 3456 del panel que se diseñó, comparando la dos técnicas sus datos obtenidos de la detección de la región que se encuentra los SNP, se obtuvo un área bajo la curva de 0.94, volviendo una técnica que cumple todos los parámetros para poder usarlo como una técnica de detección. Dando las pautas para ser postulada como una posible técnica de diagnóstico preventivo.

Los paneles de secuenciación dirigido a la detección de mutaciones relacionados con tumores ya es rutinaria en algunos países de primer mundo¹²¹, sin embargo, en países en desarrollo como México todavía no existe normas y/o protocolos de referencia para usar tecnología de secuenciación para el diagnóstico preventivo, por lo que la contribución del presente trabajo esperar agregar un pequeño bloque para las futuras generaciones para la medicina personalizada que se está abriendo camino en México. La secuenciación no era una técnica disponible al público general, debidos a diversos factores, entre los más importantes es el alto costo e infraestructura que se requiere para poder realizar la secuenciación, pero gracias a los nuevos desarrollos tecnológicos como MinION y MACPIO, permite al personal capacitado en el área realizar la secuenciación de manera más rápida y de menor costo, esta ventaja es muy prometedora para que el diagnóstico molecular sea posibles en instituciones como hospitales o los centros de salud de manera rutinaria¹⁴⁶.

El dispositivo MinION se ha utilizados en varios tipos de cáncer por todo el mundo, entre los más estudiados se encuentran reportados en los canceres de leucemias¹⁴⁷⁻¹⁵⁰, cáncer de pulmón¹⁵¹ y pancreáticos¹⁵²⁻¹⁵⁴, recomendando altamente el uso de MinION debido a la excelente precisión, detección y confiabilidad en la identificación de mutaciones debido a la cantidad de lecturas que se puede obtener, acercándose a los métodos actuales estandarizados para secuenciación de panel de cáncer (illumina). El MinION ha demostrado una detección rápida para SNP, inserciones, deleciones y translocaciones en muchos genes relevantes varios tipos de cáncer de manera confiable¹⁵⁵, entre los más relevantes se encuentran; TP53, EGFR, KRAS, NRAS, NF1, CDKN2A/P16, SMAD4/DPC4 y BCR^{86,151,152,156}. Ciertos

artículos han sugerido que la secuenciación masiva por nanoporos, en su estado actual, está listo para ser implementado en laboratorios de patología hospitalaria para ayudar en la detección primaria¹⁵¹, incluyendo el seguimiento de la evolución de la enfermedad y del tratamiento¹⁵⁷.

Desde el desarrollo de MinION en el 2014, solo se han validados pocos protocolos para su uso en diagnósticos clínicos en poblaciones del occidente, unos de los protocolos más importante son para detección de SNP relacionados con tumores del sistema nervioso^{50,99}, demostrando la eficiencia de usar la secuenciación por nanoporos en el uso de la detecta SNP y/o ARNm que se expresa en la sangre del pacientes que se tiene sospecha una patología, disminuyendo el uso de técnicas invasivas y radiografías, manifestando el verdadero poder que se obtener usando equipos de secuenciación portátil como el MinION para la recolección y análisis de SNP^{99,158}. La precisión y tiempo de respuesta sugieren que la integración de MinION en la línea de diagnóstico preventivo puede mejorar la atención del paciente y las decisiones de tratamiento acelerado.

Para llevar una técnica de biología molecular al mercado debe de pasar por varios filtros o protocolos ya establecidos, cada país tiene sus propias instituciones que rigen y dictan las normas, entre los más importante es determinar la población objetivo, por tiempo y disponibilidad de la muestra, determinar la población objetivo no fue abordado en la tesis, debido a que, si se requiere estandarizar para alguno tipo de cáncer en específico, se requiere trabajar con muestras de paciente con dicha patología, queda como un trabajo a futuro estandarizar la técnica para una patología específica.

Para terminar la tesis dio una pauta para llevar una técnica de diagnósticos preventivo, todavía le falta determinar ser estandarizados en población objetivos para cada tipo de cáncer, al mismo tiempo se requieres realizar una optimización para disminuir el costo de procesamiento debido a que si los costos de procesamiento no disminuyen de manera significativas no tendrá oportunidad de competir con el mercado emergente de los próximos años.

CONCLUSIONES

A partir de los resultados de esta tesis es posible concluir que:

Se identificaron 15,744 genes, reportados hasta octubre del 2017, que se detectan en sangre.

Se identificaron 43,780 SNP con relevancia clínica validada con poblaciones de 4 continentes (excepto Oceanía que no realizó los estudios con los criterios para ser aceptados por páginas especializadas o específica la población de estudio).

Se identificaron 17,549 SNP que predisponen a desarrollar una enfermedad, mientras que 32,027 de los SNP tiene un efecto protector.

Con los resultados obtenidos se determinó que el equipo de secuenciación de la marca Nanoporo “MinION” puede ser usado como un equipo portátil para la detección de SNP^{149,159}.

La secuenciación masiva permite general paneles para realizar un diagnóstico preventivo hasta los 31 tipos de cáncer que se tiene reportado.

La secuenciación por nanoporos proporciona datos de genes de manera rápida y precisa con mejor accesibilidad y conveniencia en comparación con los métodos actuales de NGS, lo que convierte la tecnología de secuenciación por MinION un potencial equipo para el diagnóstico preventivo²⁵.

Los datos obtenidos con el Modelo booleano nos indica que puede ser una herramienta semicuantitativa para determinar si el paciente presenta o tiene alguna posibilidad de presentar una neoplasia o si existe un riesgo de presentar metástasis.

El análisis Circus plot nos permito correlacionar los 31 tipos de cáncer que se seleccionaron para la tesis y el efecto que tiene en conjunto todas las mutaciones de un cromosoma relacionando con la prevalencia de las mutaciones en personas con la neoplasia.

La secuenciación masiva se presenta como una herramienta fundamental para de enfermedades genéticas multifactoriales, pero debido a la cantidad de variables y datos que se procesa, ha surgido la necesidad de apoyarse con la ciencia denominada Big Data, asintiendo en algoritmos para el manejo de un gran volumen de datos de forma apropiada y efectiva.

Los resultados obtenidos en la secuenciación con el equipo de NanoPore se obtuvo una tasa de error del 38,2% después de la llamada de base, lo que se considera altamente inexacto, pero debido al gran número de lecturas que realizar, este error puede bajar hasta un 5%.

Se diseñó un protocolo para la detección de un conjunto de SNP mediante secuenciación por nanoporos usando el equipo MinION.

El área bajo la curva de 0.94 indica que la técnica cumple con los parámetros establecidos por las autoridades especializada para poder validarla con una población determinada.

El costo de procesamientos de la secuenciación es elevado en comparación a las técnicas ya existentes, por lo que se requiere estandarización y optimización para que puede ser usado como un panel preventivo.

REFERENCIAS BIBLIOGRÁFICAS

1. Catherine-Sánchez N. Conociendo y comprendiendo la célula cancerosa: Fisiopatología del cáncer. *Rev Médica Clínica Las Condes*. 2013;24(4):553-562. doi:10.1016/S0716-8640(13)70659-X
2. OMS. Datos y estadísticas de la OMS. OMS. <http://www.who.int/gho/es/>. Published 2016. Accessed March 3, 2017.
3. Esser D, Holze N, Haag J, et al. Interpreting whole genome and exome sequencing data of individual gastric cancer samples. *BMC Genomics*. 2017;18(1):1-15. doi:10.1186/s12864-017-3895-z
4. Cox DR, Kartsonaki C, Keogh RH. Big data: Some statistical issues. *Stat Probab Lett*. 2018;136:111-115. doi:10.1016/J.SPL.2018.02.015
5. Kajatt EA. Lung Cancer: a Review of Current Knowledge, Diagnostic Methods and Therapeutic Perspectives. *Rev Peru Exp Salud Publica*. 2013;30(1):85-92. doi:<http://dx.doi.org/10.1590/S1726-46342013000100017>
6. Morris K V, Mattick JS. The rise of regulatory RNA. *Nat Rev Genet*. 2014;15(6):423-437. doi:10.1038/nrg3722
7. Yang L, Lee MS, Lu H, et al. Analyzing Somatic Genome Rearrangements in Human Cancers by Using Whole-Exome Sequencing. *Am J Hum Genet*. 2016;98(5):843-856. doi:10.1016/j.ajhg.2016.03.017
8. Astolfi A, Urbini M, Indio V, et al. Whole exome sequencing (WES) on formalin-fixed, paraffin-embedded (FFPE) tumor tissue in gastrointestinal stromal tumors (GIST). *BMC Genomics*. 2015;16(1):1-11. doi:10.1186/s12864-015-1982-6
9. Lohr JG, Adalsteinsson VA, Cibulskis K, et al. Whole-exome sequencing of circulating tumor cells provides a window into metastatic prostate cancer. *Nat Biotechnol*. 2014;(April):1-8. doi:10.1038/nbt.2892
10. Nakagawa H, Fujita M, Fujimoto A. Genome sequencing analysis of liver cancer for precision medicine. *Semin Cancer Biol*. 2018. doi:10.1016/j.semcancer.2018.03.004
11. America Cancer Society. Datos y Estadísticas sobre el Cáncer entre los Hispanos/Latinos. *Soc Am Contra El Cáncer*. 2014;40. <http://www.cancer.org/acs/groups/content/@epidemiologysurveillance/documents/document/acspc-036792.pdf>.
12. Aguilar Cordero MJ, Neri Sánchez M, Padilla López C a, Pimentel Ramírez ML, García Rillo A, Sánchez López a M. Risk factors in the development of breast cancer, state of Mexico. *Nutr Hosp*. 2012;27(5):1631-1636. doi:10.3305/nh.2012.27.5.5997
13. Siemiatycki J, Richardson L, Straif K, et al. Listing occupational carcinogens. *Environ Health Perspect*. 2004;112(15):1447-1459. doi:10.1289/ehp.7047
14. Portilla EC, Mñnoza W, Sierra CH. Genes y variantes polimórficas asociadas a la enfermedad cardiovascular. *Rev Colomb Cardiol*. 2014;21(5):318-326.

doi:10.1016/j.rccar.2014.09.003

15. Sanz-Montero A, Gredilla R, Herrero A, Barja de Quiroga G. Efecto del peróxido de hidrógeno en la producción mitocondr de radicales libres en relación con el envejecimiento. *Rev Esp Geriatr Gerontol*. 2004;39(1):29-34. doi:10.1016/S0211-139X(04)74928-8
16. Santisteban S. Revisión Cáncer en el Siglo XXI. *Rev Med PERU*. 2006;020(2):112-118.
17. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-tieulent J, Jemal A. Global Cancer Statistics, 2012. *CA a cancer J Clin*. 2015;65(2):87-108. doi:10.3322/caac.21262.
18. Mondragón-Terán P, López-Hernández LB, Gutiérrez-Salinas J, Suárez-Cuenca JA, Luna-Ceballos RI, Erazo Valle-Solís A. Mecanismos de señalización intracelular en cancer de tiroides. *Cir Cir*. 2016;84(5):434-443. doi:10.1016/j.circir.2016.05.017
19. Estela Herrera-González N, Martínez-García F, Mejía-Jiménez E, Estela N, González H. El efecto Warburg: la mano derecha en el desarrollo del cáncer. *Rev Esp Méd Quir*. 2015;20:171-177. <http://www.medigraphic.com/pdfs/quirurgicas/rmq-2015/rmq152h.pdf>.
20. Valdespino-Gómez VM, Valdespino-Castillo VE. Interacción de las vías de señalización intracelulares participantes en la proliferación celular: potencial blanco de intervencionismo terapéutico. *Cir y Cir (English Ed)*. 2015;83(2):165-174. doi:10.1016/j.circir.2015.04.015
21. Martin S. A Different Path for Harnessing TRAIL. *Cancer Discov*. 2017;7(5):OF7 LP-OF7. <http://cancerdiscovery.aacrjournals.org/content/7/5/OF7.abstract>.
22. Lagunas-Rangel FA. Leucemia mieloide aguda. Una perspectiva de los mecanismos moleculares del cáncer. *Gac Mex Oncol*. 2016;15(3):150-157. doi:10.1016/j.gamo.2016.05.007
23. Organista-Nava J, Gómez-Gómez Y, Ocadiz-Delgado R, et al. The HPV16 E7 oncoprotein increases the expression of Oct3/4 and stemness-related genes and augments cell self-renewal. *Virology*. 2016;499:230-242. doi:10.1016/j.virol.2016.09.020
24. Jiménez-morales S, Hidalgo-miranda A. Leucemia linfoblástica aguda infantil : una aproximación genómica. *Hosp infactil Mex*. 2017;74(1).
25. Santillán-Garzón S, Diego-Álvarez D, Buades C, et al. Diagnóstico Molecular De Enfermedades Genéticas: Del Diagnóstico Genético Al Diagnóstico Genómico Con La Secuenciación Masiva. *Rev Médica Clínica Las Condes*. 2015;26(4):458-469. doi:10.1016/j.rmclc.2015.07.004
26. Lay-Son RG, León PL. Perspectivas actuales sobre el diagnóstico genómico en pediatría. *Rev Chil Pediatría*. 2015;86(1):3-11. doi:10.1016/j.rchipe.2015.04.002
27. Jacques F, Shin H-R, Bray F, Forman D, Mathers C, Maxwell D. Estimates of worldwide burden of cancer in 2008: *Rev Cancer*. 2010;127(1):2893–2917.

doi:10.1002/ijc.25516

28. Price J, Gordon NC, Crook D, Llewelyn M, Paul J. The usefulness of whole genome sequencing in the management of *Staphylococcus aureus* infections. *Clin Microbiol Infect.* 2012;19(9):784-789. doi:10.1111/1469-0691.12109
29. Ellington MJ, Ekelund O, Aarestrup FM, et al. The role of whole genome sequencing in antimicrobial susceptibility testing of bacteria: report from the EUCAST Subcommittee. 2017;23:2-22. doi:10.1016/j.cmi.2016.11.012
30. Hidalgo-Miranda A, Jiménez-Sánchez G. Bases genómicas del cáncer de mama: Avances hacia la medicina personalizada. *Salud Publica Mex.* 2009;51(SUPPL.2):197-207. doi:10.1590/S0036-36342009000800010
31. Jayaraman SS, Rayhan DJ, Hazany S, Kolodney MS. Mutational Landscape of Basal Cell Carcinomas by Whole-Exome Sequencing. *J Invest Dermatol.* 2014;134(1):213-220. doi:10.1038/jid.2013.276
32. Cid RAP, Alejandro MÁ, Miguel AC De, Agustín MJ. Apoptosis, cáncer & Co R.A. *Rev Int Grup Investig en Oncol.* 2012;1(1):23-28.
33. Tang HL, Yuen KL, Tang HM, Fung MC. Reversibility of apoptosis in cancer cells. *Br J Cancer.* 2009;32(100):118-122. doi:10.1038/sj.bjc.6604802
34. I. Zapardiel Gutierrez, Gamiza SH, Carbajoa EP, Fontan JS. Factores moleculares pronósticos relacionados con el control del ciclo celular en el cáncer de mama. Situación actual. *Ginecol Obstet (Lima).* 2009;36(1):19-24. doi:10.1016/j.gine.2008.09.001
35. Wilbur MA, Shih I, Segars JH, Fader AN. Cancer Implications for Patients with Endometriosis. *Semin Reprod Med.* 2017;35(01):110-116. doi:http://dx.doi.org/10.1055/s-0036-1597120
36. Lum F-M, Ng LFP. Cellular and molecular mechanisms of chikungunya pathogenesis. *Antiviral Res.* 2015;120(June):165-174. doi:10.1016/j.antiviral.2015.06.009
37. Hanahan D, Weinberg RA, Francisco S. The Hallmarks of Cancer. *Cell.* 2000;100(1):57-70.
38. Fouad YA, Aanei C. Revisiting the hallmarks of cancer. 2017;7(5):1016-1036.
39. G ITRIAGO L, NICOLÁS SI, CORTES G. CÁNCER EN CHILE Y EL MUNDO : UNA MIRADA EPIDEMIOLÓGICA , PRESENTE Y FUTURO. *Rev medica clinia chile.* 2013;24(4):531-552. doi:10.1016/S0716-8640(13)70195-0
40. Roesch-Dietlena, Cano-Contrerasb, Sánchez-Mazac J, et al. Frecuencia de infección por virus del papiloma humano en pacientes con cáncer del aparato digestivo. *Gastroenterol Mex.* 2018;83(3):253-258. doi:10.1016/j.rgmx.2017.09.003
41. Castrezana R. Geografía del cáncer de mama en México The geography of Mexico breast cancer. *Investig Geográficas Boletín del Inst Geogr.* 2017;2017(93):140-157. doi:10.14350/rig.56879

42. Csendes A. Situación del cáncer gástrico en el mundo y en Chile. *Rev Chil Cirugía*. 2017;69(6):502-507. doi:10.1016/j.rchic.2016.10.014
43. CDC. Cáncer - Por características demográficas - Tasas de cáncer por raza y grupo étnico y sexo. demografía del cáncer. <https://www.cdc.gov/spanish/cancer/dcpc/data/race.htm>. Published 2017.
44. Danhier P, Ba P, Payen VL, et al. Biochimica et Biophysica Acta Cancer metabolism in space and time: Beyond the Warburg effect. 2017;1858:556-572. doi:10.1016/j.bbabi.2017.02.001
45. IARC. Cancer Fact Sheets: All Cancers Excluding Non-Melanoma Skin Cancer. <http://gco.iarc.fr/today/data/pdf/fact-sheets/cancers/cancer-fact-sheets-29.pdf>. Published 2012. Accessed March 12, 2017.
46. INEGI. *ESTADÍSTICA DE CÁNCER. MÉXICO; 2017*. http://www.inegi.org.mx/saladeprensa/aproposito/2017/cancer2017_Nal.pdf.
47. Espejo Romero H, Navarrete Siancas J. Clasificación de los Adenocarcinomas. *Rev Gastroenterol Perú*. 2003;23:199-212. http://www.scielo.org.pe/scielo.php?pid=S1022-51292003000300006&script=sci_arttext.
48. Alejandro Bonilla-Sepúlveda O, Matute-Turízo G, Severiche C. Classification of intrinsic subtypes of breast carcinomas analyzed in a pathology center of Medellin in 2011. *Enero -Junio*. 2015;29(1):35-46.
49. Dong S, Song X-S, Chen G, Liu J. Mixed primary squamous cell carcinoma, follicular carcinoma, and micropapillary carcinoma of the thyroid gland: A case report. *Auris Nasus Larynx*. 2016;43(4):455-459. doi:10.1016/j.anl.2015.10.011
50. Jiménez-Morales S, Hidalgo-Miranda A, Ramírez-Bello J. Leucemia linfoblástica aguda infantil: una aproximación genómica. *Bol Med Hosp Infant Mex*. 2017;74(1):13-26. doi:10.1016/j.bmhmx.2016.07.007
51. Vera AM, Pardo C, Duarte MC, Suárez A. Análisis de la mortalidad por leucemia aguda pediátrica en el Instituto Nacional de Cancerología. *Biomédica*. 2012;32(3):355-364. doi:10.7705/biomedica.v32i3.691
52. Sánchez Salinas A, Monserrat Coll J, Rosique Cortina P, Moraleta Jiménez JM. Leucemias agudas. *Med - Programa Form Médica Contin Acreditado*. 2012;11(21):1268-1279. doi:10.1016/S0304-5412(12)70450-X
53. García L, Cabrero M, del Cañizo C. Leucemias agudas. *Med*. 2016;12(21):1201-1212. doi:10.1016/j.med.2016.10.011
54. Gabriel R, Cano A. Sarcomas: etiología y síntomas Sarcomas: etiology and symptoms. 2018:116-129.
55. Maza-astiazarán C De, González-vergara C, Bieletto-trejo O. Estudio retrospectivo de la incidencia de sarcomas en un hospital general privado de la Ciudad de México. 2018;(1):23-27.

56. Bagot M, Ortonne N. Linfomas cutáneos: clasificación. *EMC - Dermatología*. 2013;47(1):1-6. doi:10.1016/S1761-2896(13)64086-8
57. Luis Cereceda G. Emergencias oncológicas. *Rev Médica Clínica Las Condes*. 2011;22(5):665-676. doi:10.1016/S0716-8640(11)70479-5
58. Páramo Fernández JA. Gammapatías monoclonales de significado incierto. *Rev Clin Esp*. 2008;208(6):267-268. doi:10.1157/13123184
59. World Health Organization W. *Aplicación de Los Conocimientos Control Del Cáncer Guía de La OMS Para Desarrollar Programas Eficaces*. 1st ed. (Consejería de Sanidad de la Comunidad Valenciana, ed.). España; 2007.
60. Ezzati M, Lopez AD, Rodgers A, Murray CJL. *Comparative Quantification of Health Risks*. Vol 2. 1st ed. (Salud OM de la, ed.). Geneva; 2004.
61. Von Bernhardt R, Zanlungo S, Arrese M, Arteaga A, Rigotti A. The metabolic syndrome: from an aggravating condition to a pathogenic risk factor for chronic diseases. *Rev Med Chil*. 2010;138(8):1012-1019. doi:10.4067/S0034-98872010000800012
62. Zuleta MG, Regino WO, Ruiz Lobo X, Gómez Zuleta M, Otero Regino W. Factores de riesgo para cáncer gástrico en pacientes colombianos. *Rev Col Gastroenterol*. 2009;24(2):134-143.
63. Peralta-Zaragoza O, Lagunas-Martínez A, Madrid-Marina V. Factor de crecimiento transformante beta-1: estructura, función y mecanismos de regulación en cáncer. *Salud Pública de México*. 2001;43(4):340-351. http://bvs.insp.mx/rsp/_files/File/2001/num_4/434_9_factor.pdf%5Cnfile:///Users/va_lentinmartinez/Library/Application_Support/Papers2/Articles/2001/Peralta-Zaragoza/Salud_Publica_de_México_2001_Peralta-Zaragoza.pdf%5Cnpapers2://publication/uuid/5C08DECF-5.
64. Nelson DE, Jarman DW, Rehm J, et al. Alcohol-attributable cancer deaths and years of potential life lost in the United States. *Am J Public Health*. 2013;103(4):641-648. doi:10.2105/AJPH.2012.301199
65. Smith AH, Lingas EO, Rahman M. Contamination of drinking-water by arsenic in Bangladesh: A public health emergency. *Bull World Health Organ*. 2000;78(9):1093-1103. doi:10.1590/S0042-96862000000900005
66. Feo O, Martínez M. Cáncer ocupacional: epidemiología y prevención. *Salud los Trab*. 1993;2(109). <http://www.toxnet.com.br/download/cancer-ocupacional-oscar-feo.pdf>.
67. Oller-Arlandis V, Sanz-Valero J. Cáncer por contaminación química del agua de consumo humano en menores de 19 años: una revisión sistemática. *Salud pública de Panama*. 2012;32(6):435-443. doi:10.1590/S1020-49892012001400007
68. Cabrera Morales CM, López-Nevot MA. Efectos de la radiación ultravioleta (UV) en la inducción de mutaciones de p53 en tumores de piel. *Oncol*. 2010;29(7):25-32. doi:10.4321/s0378-48352006000700003

69. Vogelstein B, Lane D, Levine AJ. Surfing the p53 network. *Nature*. 2000;408(6810):307-310. doi:10.1038/35042675
70. Rodríguez González B, Ruiz Lencioni E, Romero Jiménez L, Cutuli de Simón MT. Esporos bacterianos para el tratamiento del cáncer. *Rev Complut Ciencias Vet*. 2012;6(2). doi:10.5209/rev_rccv.2012.v6.n2.41088
71. Cuestas ML. Relaciones íntimas entre los microorganismos y el cáncer. *Rev Argent Microbiol*. 2014;46(3):173-174. doi:10.1016/S0325-7541(14)70068-9
72. Meresman GF, Olivares CN. Implicancia de las prostaglandinas en la fisiopatología de la endometriosis. *Saegre*. 2012;19(3):40-56.
73. Castells A. Implicación de ciclooxigenasa 2 en el cáncer: utilidad de los coxib. *Rev Española Reumatol*. 2016;30(7).
74. Rosero C, Corredor M. Polimorfismos en genes implicados en el desarrollo de cáncer gástrico : revisión A Review of Polymorphisms in Genes Involved in the Development of Gastric Cancer. *Rev Col Gastroenterol*. 2016;31(4):391-402. <http://www.scielo.org.co/pdf/rcg/v31n4/v31n4a09.pdf>.
75. Meza J, Montaña A, Aguayo Á. Las Bases Moleculares del Cáncer. *Acta Univ*. 2006;16(1):40-49.
76. Secretaria De Salud. *PROGRAMA DE ACCIÓN ESPECÍFICO 2007-2012 Otras Enfermedades Transmitidas Por Vector*. 2nd ed. (Salud S de, Dr., eds.). Mexico; 2012. <http://www.cdi.salud.gob.mx:8080/BasesCDI/Archivos/Enfermedadestransmitidasporvector/etv.pdf>.
77. Navarro-Colás S. Fisiopatología, diagnóstico y tratamiento de la insuficiencia pancreática exocrina en el paciente con pancreatitis crónica. *Gastroenterol Hepatol*. 2005;28(Supl 2):33-38. doi:http://dx.doi.org/10.1157/13071383
78. Mendiola AV, Soto Cruz I. Energetic metabolism and cancer metabolismo energético y cáncer. *Rev Espec en Ciencias la Salud*,. 2014;17(2):108-113.
79. Dyson HJ, Wright PE. Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol*. 2005;6(3):197-208. doi:10.1038/nrm1589
80. Martínez-Vicente M, Sovak G, Cuervo AM. Protein degradation and aging. *Exp Gerontol*. 2005;40(8-9):622-633. doi:10.1016/j.exger.2005.07.005
81. Farfán T, Fernando L. Implicancias del papel de las células madre de cáncer en la biología del cáncer . *UVC*. 2013;5(2):176-181.
82. Vallejo Bestard J, Tremps Velázquez E, Blázquez Mañá C, Celma Doménech A, Torres Ramírez I, Morote Robles J. Tumor adenomatoide de epidídimo: El tumor más frecuente de las estructuras paratesticulares. *Anatomía Patológica del Hosp Gen Univ Vall d'Hebron UAB Barcelona*. 2008;32(6):611-617.
83. Soussi T, Dehouche K, Bérout C. p53 website and analysis of p53 gene mutations in human cancer: Forging a link between epidemiology and carcinogenesis. *Hum Mutat*. 2000;15(1):105-113. doi:10.1002/(SICI)1098-1004(200001)15:1<105::AID-

HUMU19>3.0.CO;2-G

84. Rocha C, Lima O, Rogério I, et al. P53 gene: major mutations in neoplasias and anticancer gene therapy Gene. *Cienc Rural*. 2012;42(5):845-853.
85. Rangel-López A, Piña-Sánchez P, Salcedo M. Variaciones genéticas del gen supresor de tumores TP53: relevancia y estrategias de análisis. *Rev Investig Clin*. 2006;58(3):254-264.
86. Minervini C, Cumbo C, Orsini P, et al. TP53 gene mutation analysis in chronic lymphocytic leukemia by nanopore MinION sequencing. *Diagn Pathol*. 2016;11(1):1-9. doi:10.1186/s13000-016-0550-y
87. Hollstein M, Hergenhahn M, Yang Q, Bartsch H, Wang ZQ, Hainaut P. New approaches to understanding p53 gene tumor mutation spectra. *Mutat Res - Fundam Mol Mech Mutagen*. 1999;431(2):199-209. doi:10.1016/S0027-5107(99)00162-1
88. Pfeifer GP. p53 mutational spectra and the role of methylated CpG sequences. *Mutat Res - Fundam Mol Mech Mutagen*. 2000;450(1-2):155-166. doi:10.1016/S0027-5107(00)00022-1
89. Dumaz N, Drougard C, Sarasin A, Daya-Grosjean L. Specific UV-induced mutation spectrum in the p53 gene of skin tumors from DNA-repair-deficient xeroderma pigmentosum patients. *Proc Natl Acad Sci*. 1993;90(22):10529 LP - 10533. doi:10.1073/pnas.90.22.10529
90. Muñoz IC, Prieto AE, Chapa AI, Treviño M, Nández JV. Linfomas, nueva clasificación y tratamiento. Incidencia de las lesiones en el Hospital San José de 1990 a 2004. Correlación clínico-radiológica. *An Radiol Mex*. 2006;2:117-136.
91. Machado I, Navarro S, Llombart-Bosch A. Pautas en el diagnóstico morfológico, inmunohistoquímico y genético de los tumores de células redondas y pequeñas con especial referencia al sarcoma de Ewing/PNET. *Rev Esp Patol*. 2012;45(3):145-156. doi:10.1016/j.patol.2012.04.004
92. Forero-Castro M, Robledo C, Benito R, et al. Genome-wide DNA copy number analysis of acute lymphoblastic leukemia identifies new genetic markers associated with clinical outcome. *PLoS One*. 2016;11(2):1-20. doi:10.1371/journal.pone.0148972
93. González IC, Honrubia VG. Cáncer de mama hereditario más allá de BRCA1/BRCA2. *Genética Médica y Genómica*. 2018;2(02):67-77. revistageneticamedica.com.
94. Yuriko L, Vargas-cely S, Sebastián J. SÍNDROME DE CÁNCER HEREDITARIO DE MAMA Y OVARIO : APLICACIÓN CLÍNICA Hereditary breast and ovarian cancer syndrome: Clinical application. *Red Rev Científicas América Lat*. 2016;67(1):36-47.
95. Macintyre G, Ylstra B, Brenton JD. Sequencing Structural Variants in Cancer for Precision Therapeutics. *Trends Genet*. 2016;32(9):530-542. doi:10.1016/j.tig.2016.07.002

96. Instiuto Nacional del Cáncer. Definición de auxiliar médico - Diccionario de cáncer - National Cancer Institute. <https://www.cancer.gov/espanol/publicaciones/diccionario?cdrid=644325>. Published 2015. Accessed March 14, 2017.
97. Tume L, Cisneros C, Sevillano J, et al. Desregulación de microARN en el cáncer: un enfoque terapéutico y diagnóstico. *Gac Mex Oncol.* 2016;15(5):298-304. doi:10.1016/j.gamo.2016.08.004
98. Rivere AE, Klimberg VS, Bland KI. *Breast Conservation Therapy for Invasive Breast Cancer*. Vol 30. Fifth Edit. Elsevier Inc.; 2018. doi:10.1016/B978-0-323-35955-9.00032-5
99. Ramírez-Bello J, Jiménez-Morales M. Implicaciones funcionales de los polimorfismos de un solo nucleótido (SNP) en genes codificantes de proteínas y no codificantes en enfermedades multifactoriales. *Gac Med Mex.* 2017;153(2):238-250.
100. Ramírez-Bello J, Vargas-Alarcón G, Tovilla-Zárate C, Fragoso JM. Polimorfismos de un solo nucleótido (SNP): Implicaciones funcionales de los SNP reguladores (rSNP) y de los SNP-ARN estructurales (srSNP) en enfermedades complejas. *Gac Med Mex.* 2013;149(2):220-228.
101. Sadee W, Wang D, Papp AC, et al. Pharmacogenomics of the RNA world: Structural RNA polymorphisms in drug therapy. *Clin Pharmacol Ther.* 2011;89(3):355-365. doi:10.1038/clpt.2010.314
102. Pipan V, Zorc M, Kunej T. MicroRNA polymorphisms in cancer: A literature analysis. *Cancers (Basel).* 2015;7(3):1806-1814. doi:10.3390/cancers7030863
103. Franco E De. Genome, Exome, and Targeted Next-Generation Sequencing in Neonatal Diabetes. *Pediatr Clin NA.* 2015;62(4):1037-1053. doi:10.1016/j.pcl.2015.04.012
104. Groenendijk FH, Jager A, Cardoso F, van Deurzen CHM. A nationwide registry-based cohort study of the MammaPrint genomic risk classifier in invasive breast cancer. *Breast.* 2018;38:125-131. doi:10.1016/j.breast.2017.12.015
105. Liao BB, Sievers C, Donohue LK, et al. Adaptive Chromatin Remodeling Drives Glioblastoma Stem Cell Plasticity and Drug Tolerance. *Cell Stem Cell.* 2016;0(0):1171-1176. doi:10.1016/j.stem.2016.11.003
106. Wei X, Dai Y, Yu P, et al. Targeted next-generation sequencing as a comprehensive test for patients with and female carriers of DMD/BMD: a multi-population diagnostic study. *Eur J Hum Genet.* 2014;22(1):110-118. doi:10.1038/ejhg.2013.82
107. Hernan I, Borrs E, De Sousa Dias M, et al. Detection of genomic variations in BRCA1 and BRCA2 genes by long-range PCR and next-generation sequencing. *J Mol Diagnostics.* 2012;14(3):286-293. doi:10.1016/j.jmoldx.2012.01.013
108. Lloyd-jones LR, Holloway A, Mcrae A, et al. The Genetic Architecture of Gene Expression in Peripheral Blood. *Am J Hum Genet.* 2017;100(2):228-237. doi:10.1016/j.ajhg.2016.12.008

109. Abdel-Hafiz HA, Horwitz KB. Role of epigenetic modifications in luminal breast cancer. *Epigenomics*. 2015;7:1-16. doi:10.2217/epi.15.10
110. Patel AP, Tirosh I, Trombetta JJ, et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science (80-)*. 2014;344(6190):1396-1401. doi:10.1126/science.1254257
111. Iyer MK, Niknafs YS, Malik R, et al. The landscape of long noncoding RNAs in the human transcriptome. *Nat Genet*. 2015;47(3):199-208. doi:10.1038/ng.3192
112. Metzker ML. Sequencing technologies - the next generation. *Nat Rev Genet*. 2010;11(1):31-46. doi:10.1038/nrg2626
113. Dashrath K, Chakrapani V, Kumar A, et al. Gene Reports Status and future perspectives of single nucleotide polymorphisms (SNPs) markers in farmed fishes : Way ahead using next generation sequencing. *Gene Reports*. 2017;6:81-86. doi:10.1016/j.genrep.2016.12.004
114. Weimer ET. Clinical validation of NGS technology for HLA : An early adopter ' s perspective. *Hum Immunol*. 2016;77(10):820-823. doi:10.1016/j.humimm.2016.06.014
115. Glotov AS, Kazakov S V, Zhukova EA, et al. Clinica Chimica Acta Targeted next-generation sequencing (NGS) of nine candidate genes with custom AmpliSeq in patients and a cardiomyopathy risk group. *Clin Chim Acta*. 2015;446:132-140. doi:10.1016/j.cca.2015.04.014
116. MSS. *Lineamientos Técnicos Para La Prevención y Control Del Cáncer Cérvico Uterino y de Mama*. 1st ed. (Salud M de, ed.); 2015. http://asp.salud.gob.sv/regulacion/pdf/lineamientos/lineamientos_prevencion_cancer_cervico_uterino_y_de_mama_v3.pdf.
117. Montgomery D. *Introduction to Statistical Quality Control*.; 2009. doi:10.1002/1521-3773(20010316)40:6<9823::AID-ANIE9823>3.3.CO;2-C
118. Shen T, Pajaro-Van de Stadt SH, Yeat NC, Lin JCH. Clinical applications of next generation sequencing in cancer: From panels, to exomes, to genomes. *Front Genet*. 2015;6(JUN):1-9. doi:10.3389/fgene.2015.00215
119. Di Cara A, Garg A, De Micheli G, Xenarios I, Mendoza L. Dynamic simulation of regulatory networks using SQUAD. *BMC Bioinformatics*. 2007;8:1-10. doi:10.1186/1471-2105-8-462
120. Garg A, Di Cara A, Xenarios I, Mendoza L, De Micheli G. Synchronous versus asynchronous modeling of gene regulatory networks. *Bioinformatics*. 2008;24(17):1917-1925. doi:10.1093/bioinformatics/btn336
121. Jennings LJ, Arcila ME, Corless C, et al. Guidelines for Validation of Next-Generation Sequencing–Based Oncology Panels: A Joint Consensus Recommendation of the Association for Molecular Pathology and College of American Pathologists. *J Mol Diagnostics*. 2017;19(3):341-365. doi:10.1016/j.jmoldx.2017.01.011

122. Zhang J, Ding X, Guan R, et al. Evaluation of different 16S rRNA gene V regions for exploring bacterial diversity in a eutrophic freshwater lake. *Sci Total Environ.* 2018;618:1254-1267. doi:10.1016/j.scitotenv.2017.09.228
123. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2009;20:254-260. doi:10.1101/gr.107524.110.20
124. Robinson JT, Thorvaldsdóttir H, Winckler W, et al. Integrative Genome Viewer. *Nat Biotechnol.* 2011;29(1):24-26. doi:10.1038/nbt.1754.Integrative
125. Illumina. Sequencing Coverage Calculation Methods for Human Whole-Genome Sequencing. *illumina.* 2014:1-2.
126. Peacock E, Whiteley P. Perlegen sciences, inc. *Pharmacogenomics.* 2005;6(4):439-442. doi:10.1517/14622416.6.4.439
127. Epidemiología U De, Reynoso-noverón N, Torres-domínguez JA. Epidemiología del cáncer en México: carga global y proyecciones 2000-2020. *Latinoam Med.* 2020;8(1):9-15.
128. Gardner SA, Weymouth KS, Kelly WS, et al. Evaluation of a 27-gene inherited cancer panel across 630 consecutive patients referred for testing in a clinical diagnostic laboratory. 2018:1-10. doi:10.1186/s13053-017-0083-8
129. Fernández MLM, MacDonald A, Aceña I, Sánchez EB, ECEMC GP del, Frías MLM. Análisis de alteraciones cromosómicas estructurales y su distribución por cromosomas en la serie de recién nacidos con defectos congénitos del ECEMC. *Boletín del ECEMC.* 2012;0(1):66-82.
130. Pérez GL, Nicolás FJ, Garrido AR, Raquel Álvarez G. Artículo de revisión Inmunogenética del asma. Participación de los genes HLA en la susceptibilidad al asma. *Acta Pediatr Mex.* 2005;26(1):29-34.
131. Cruz IS. Transducción De Señales Y Cáncer. *Vertientes Rev Espec en Ciencias la Salud.* 2003;6(1):45-50.
132. Valdespino-gómez VM, Valdespino-castillo VE, Márquez-holmberg M. Intracelulares En El Cáncer. *Cir Ciruj.* 2009;77(4):329-339.
133. Quintero Villarreal A, Rodriguez CN. Terapia nutricional en el enfermo oncológico en estado crítico. *Acta Colomb Cuid Intensivo.* 2016;16:1-25. doi:10.1016/j.acci.2016.07.002
134. Rondon M, Caicedo J, Robledo JF. Determinación de anomalías cromosómicas y secuencias de ADN amplificadas en cáncer de mama. *Rev Cienc Salud.* 2006;4(2):7-22. <http://revistas.urosario.edu.co/index.php/revsalud/article/view/565/48>.
135. Orsetti B, Nugoli M, Cervera N, et al. Genomic and expression profiling of chromosome 17 in breast cancer reveals complex patterns of alterations and novel candidate genes. *Cancer Res.* 2004;64(18):6453-6460. doi:10.1158/0008-5472.CAN-04-0756

136. Struski S, Doco-Fenzy M, Cornillet-Lefebvre P. Compilation of published comparative genomic hybridization studies. *Cancer Genet Cytogenet.* 2002;135(1):63-90. doi:10.1016/S0165-4608(01)00624-0
137. Forozan F, Mahlamäki EH, Monni O, et al. Comparative genomic hybridization analysis of 38 breast cancer cell lines: A basis for interpreting complementary DNA microarray data. *Cancer Res.* 2000;60(16):4519-4525.
138. Mousses S, Kononen J, Sauter G, et al. Comprehensive copy number and gene expression profiling of the 17q23 amplicon in human. *Genetics.* 2001;98(10):5711-5716.
139. Chaisson MJP, Huddleston J, Dennis MY, et al. Resolving the complexity of the human genome using single-molecule sequencing. *Nature.* 2015;517(7536):608-611. doi:10.1038/nature13907.Resolving
140. Jones S, Anagnostou V, Lytle K, et al. Personalized genomic analyses for cancer mutation discovery and interpretation. *Sci Transl Med.* 2015;7(283):283ra53-283ra53. doi:10.1126/scitranslmed.aaa7161
141. Ostrow SL, Barshir R, DeGregori J, Yeger-Lotem E, Hershberg R. Cancer Evolution Is Associated with Pervasive Positive Selection on Globally Expressed Genes. *PLoS Genet.* 2014;10(3):16-20. doi:10.1371/journal.pgen.1004239
142. Bahcall O. Rare variant association studies + parent of origin effects. *Nat Genet.* 2014;46(3):219-219. doi:10.1038/ng.2908
143. Martincorena I, Raine KM, Gerstung M, et al. Universal Patterns of Selection in Cancer and Somatic Tissues. *Cell.* 2017. doi:10.1016/j.cell.2017.09.042
144. Baizabal J-M, Furlan-Magaril M, Santa-Olalla J, Covarrubias L. Neural stem cells in development and regenerative medicine. *Arch Med Res.* 2003;34(6):572-588. doi:10.1016/j.arcmed.2003.09.002
145. Sauvage V, Boizeau L, Candotti D, et al. Early MinION™ nanopore single-molecule sequencing technology enables the characterization of hepatitis B virus genetic complexity in clinical samples. *PLoS One.* 2018;13(3):1-15. doi:10.1371/journal.pone.0194366
146. Kenji Fujimoto, Shigetoshi Yano, Naoki Shinojima, Takuichiro Hide JK. MinION rapid sequencing: Review of potential applications in neurosurgery. *Surg Neurol Int.* 2017;8(1):1-7. doi:10.4103/sni.sni
147. Batovska J, Lynch SE, Rodoni BC, Sawbridge TI, Cogan NO. Metagenomic arbovirus detection using MinION nanopore sequencing. *J Virol Methods.* 2017;249(September):79-84. doi:10.1016/j.jviromet.2017.08.019
148. Reed GH, Kent JO, Wittwer CT. High-resolution DNA melting analysis for simple and efficient molecular diagnostics. *Pharmacogenomics.* 2007;8(6):597-608. doi:10.2217/14622416.8.6.597
149. Simpson JT, Workman RE, Zuzarte PC, David M, Dursi LJ, Timp W. Detecting DNA

- cytosine methylation using nanopore sequencing. *Nat Methods*. 2017;14:407. <http://dx.doi.org/10.1038/nmeth.4184>.
150. Jain M, Koren S, Miga KH, et al. Nanopore sequencing and assembly of a human genome with ultra-long reads. *Nat Biotechnol*. 2018;36(4):338-345. doi:10.1038/nbt.4060
 151. Suzuki A, Suzuki M, Mizushima-Sugano J, et al. Sequencing and phasing cancer mutations in lung cancers using a long-read portable sequencer. *DNA Res*. 2017;24(6):585-596. doi:10.1093/dnares/dsx027
 152. Norris AL, Workman RE, Fan Y, Eshleman JR, Timp W. Nanopore sequencing detects structural variants in cancer. *Cancer Biol Ther*. 2016;17(3):246-253. doi:10.1080/15384047.2016.1139236
 153. Cornelis S, Gansemans Y, Deleye L, Deforce D, Van Nieuwerburgh F. Forensic SNP Genotyping using Nanopore MinION Sequencing. *Sci Rep*. 2017;7:41759. <http://dx.doi.org/10.1038/srep41759>.
 154. Euskirchen P, Bielle F, Labreche K, et al. Same-day genomic and epigenomic diagnosis of brain tumors using real-time nanopore sequencing. *Acta Neuropathol*. 2017;134(5):691-703. doi:10.1007/s00401-017-1743-5
 155. Jansen HJ, Liem M, Jong-Raadsen SA, et al. Rapid de novo assembly of the European eel genome from nanopore sequencing reads. *Sci Rep*. 2017;7(1):1-13. doi:10.1038/s41598-017-07650-6
 156. Minervini CF, Cumbo C, Orsini P, et al. Mutational analysis in BCR-ABL1 positive leukemia by deep sequencing based on nanopore MinION technology. *Exp Mol Pathol*. 2017;103(1):33-37. doi:10.1016/j.yexmp.2017.06.007
 157. Ashton PM, Nair S, Dallman T, et al. MinION nanopore sequencing identifies the position and structure of a bacterial antibiotic resistance island. *Nat Biotechnol*. 2014;33:296. <http://dx.doi.org/10.1038/nbt.3103>.
 158. Oh JJ, D M, Park S, et al. Genome-wide detection of allelic genetic variation to predict advanced-stage prostate cancer after radical prostatectomy using an exome SNP chip. *Urol Oncol Semin Orig Investig*. 2015;33(9):385.e7-385.e13. doi:10.1016/j.urolonc.2015.05.016
 159. Liu Y, Kheradmand F, Davis CF, et al. Focused Analysis of Exome Sequencing Data for Rare Germline Mutations in Familial and Sporadic Lung Cancer. *J Thorac Oncol*. 2015;11(1):52-61. doi:10.1016/j.jtho.2015.09.015

Anexo 1 Consentimiento informativo



Cuernavaca, Mor., a de del 2017

Proyecto de investigación "CARCEROMA: detección molecular de marcadores de cáncer por secuencia masiva de tercera generación"

IBT. Luis Uribe Espejo Galicia

Investigador-analista

A través de esta carta, quiero dar mi consentimiento para utilizar la(s) muestras tomadas(s), solo con fines de investigación, mis datos personales no podrán utilizarse por ningún motivo en los documentos que se deriven de esta investigación ni utilizarse con fines comerciales y/o lucrativo.

Atentamente:

Nombre y firma del paciente

Testigo 1

Nombre y firma

Testigo 2

Nombre y firma

Toda investigación médica que se realiza en el laboratorio de Diagnóstico Molecular "Grupo Quae" se apega en forma estricta a los principios de la Declaración de Helsinki. Todos los protocolos son revisados en forma independiente y libre de conflicto de intereses y que cada estudio cumpla con los principios éticos de respeto al individuo, beneficencia (reducir al mínimo los riesgos y búsqueda siempre el bienestar del individuo) y justicia (participación equitativa de todos los sujetos candidatos a un estudio de investigación).

DR. VERA LUCIA PETRICEVICH¹
Facultad de medicina UAEM
Lab Inflamación y Toxicología

DR. GABRIEL GUILLEN SOLIS²
Grupo Quae
Lab de Diagnóstico Molecular

1. Leñero S. N, los Volcanes, 62350 Cuernavaca, Morelos.
2. Calle de la luz núm. 44, col Chapultepec. Dentro del Hospital Morelos.

Anexo 2 estudio del paciente

ESTUDIO DE PACIENTE PARA DISEÑAR LA TECNICA DE CARCEROMA

DATOS DEL PACIENTE

Nombre: _____ Sexo: _____ Edad: _____ Folio: _____

Datos generales

Ocupación: _____ Talla (cm): _____ Peso (Kg): _____

Desendencia: Australia Europea America Africana Asiática Otros: _____

Habitos nocivos

	SI	NO	Cantidad por semana	
Consumo de alcohol:	<input type="checkbox"/>	<input type="checkbox"/>	_____	Litro(s)
Consumo de tabaco:	<input type="checkbox"/>	<input type="checkbox"/>	_____	Cigarro(s)
Consumo de drogas:	<input type="checkbox"/>	<input type="checkbox"/>	_____	U Especificar: _____
Otros:	<input type="checkbox"/>	<input type="checkbox"/>	_____	U Especificar: _____

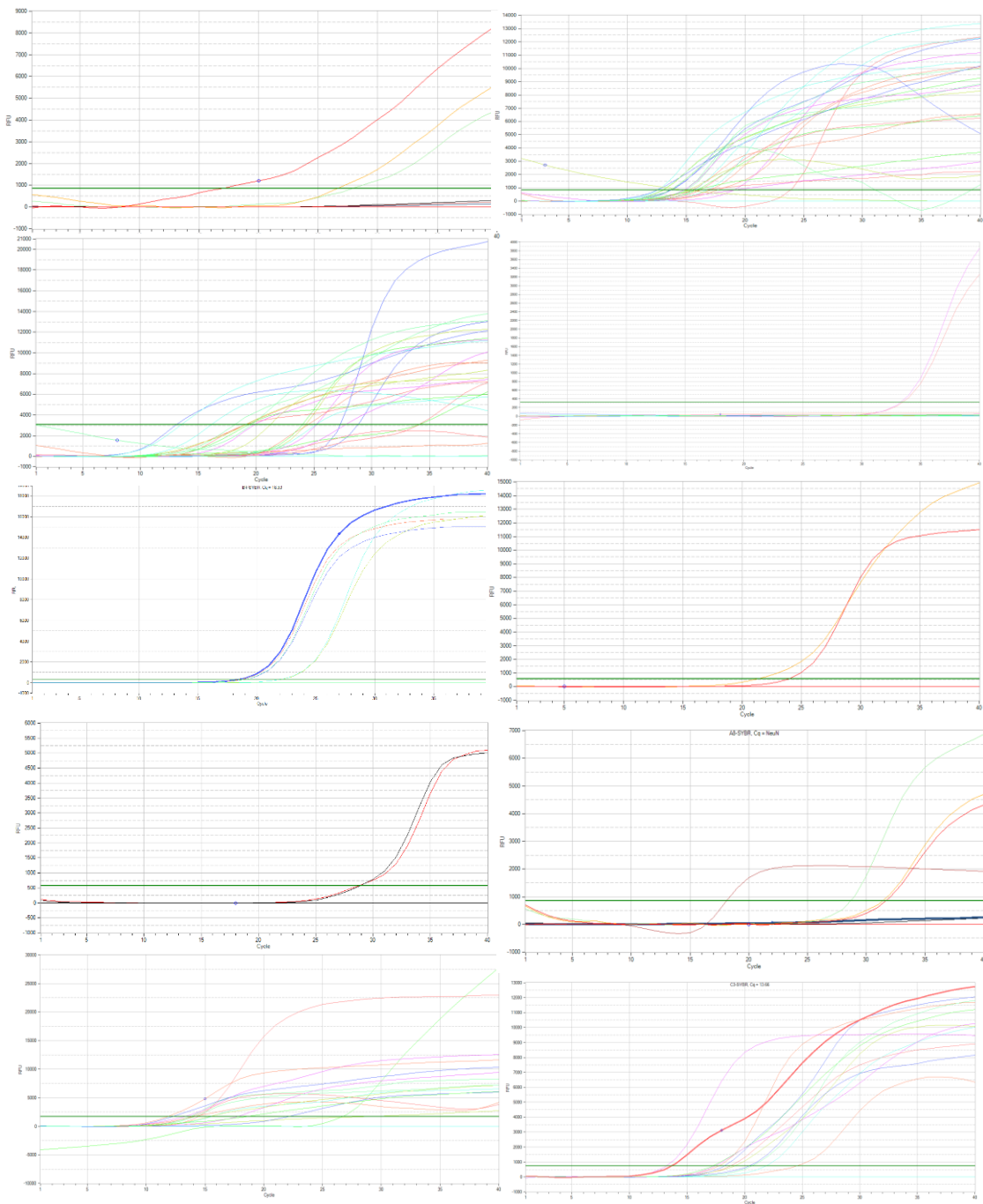
Historial médico

	SI	NO		SI	NO	Especificar
Diabetes	<input type="checkbox"/>	<input type="checkbox"/>	Enfermedades cardiovasculares	<input type="checkbox"/>	<input type="checkbox"/>	_____
Obesidad	<input type="checkbox"/>	<input type="checkbox"/>	Trastornos neurodegenerativos	<input type="checkbox"/>	<input type="checkbox"/>	_____
Epilepsia	<input type="checkbox"/>	<input type="checkbox"/>	Alergia	<input type="checkbox"/>	<input type="checkbox"/>	_____
Hipertensión	<input type="checkbox"/>	<input type="checkbox"/>	Trastornos mentales	<input type="checkbox"/>	<input type="checkbox"/>	_____
Asma	<input type="checkbox"/>	<input type="checkbox"/>	Trastornos de la audición	<input type="checkbox"/>	<input type="checkbox"/>	_____
Osteoporosis	<input type="checkbox"/>	<input type="checkbox"/>	Deficiencia visual	<input type="checkbox"/>	<input type="checkbox"/>	_____
Menopausia	<input type="checkbox"/>	<input type="checkbox"/>	Cáncer	<input type="checkbox"/>	<input type="checkbox"/>	_____
Preeclampsia	<input type="checkbox"/>	<input type="checkbox"/>	Otras enfermedades	<input type="checkbox"/>	<input type="checkbox"/>	_____

Historial Familiar

	SI	NO	Parentesco	Tipo de Cáncer
Familiares con cáncer	<input type="checkbox"/>	<input type="checkbox"/>	_____	_____
			_____	_____
			_____	_____

Anexo 3 datos obtenidos por el piko24





Cuernavaca, Mor., 28 de mayo de 2019.

Dr. David Martínez Duncker Ramírez
Director
Facultad de Medicina
P r e s e n t e:

Estimado Dr. Martínez Duncker, por este conducto me permito informarle que he revisado el trabajo de Tesis **"CANCEROMA: DETECCIÓN MOLECULAR DE MARCADORES DE CÁNCER POR SECUENCIACIÓN MASIVA DE TERCERA GENERACIÓN"**, que para obtener el grado de Maestro en Medicina Molecular me proporcionó el alumno Luis Uribe Espejo Galicia. Le comunico que su contenido es adecuado y suficiente, por lo que de acuerdo al artículo 110 del Reglamento de Posgrado de la UAEM le otorgo:

Dictamen I se aprueba el trabajo de tesis tal como se presenta.

Así mismo, le agradezco la invitación a participar en este programa educativo, reiterando además mi disposición para seguir colaborando en este tipo de actividades. Sin otro particular por el momento, reciba un cordial saludo.

Atentamente



Dr. Joaquín Sánchez Castillo



FACULTAD DE MEDICINA

Secretaría de Docencia

Jefatura de Posgrado

"1919-2019: en memoria del General Emiliano Zapata Salazar"

Cuernavaca, Mor., 28 de mayo de 2019.

Dr. David Martínez Duncker Ramírez
Director
Facultad de Medicina
Presente:

Estimado Dr. Martínez Duncker, por este conducto me permito informarle que he revisado el trabajo de Tesis "CANCEROMA: DETECCIÓN MOLECULAR DE MARCADORES DE CÁNCER POR SECUENCIACIÓN MASIVA DE TERCERA GENERACIÓN", que para obtener el grado de Maestro en Medicina Molecular me proporcionó el alumno Luis Uribe Espejo Galicia. Le comunico que su contenido es adecuado y suficiente, por lo que de acuerdo al artículo 110 del Reglamento de Posgrado de la UAEM le otorgo:

Dictamen I se aprueba el trabajo de tesis tal como se presenta.

Así mismo, le agradezco la invitación a participar en este programa educativo, reiterando además mi disposición para seguir colaborando en este tipo de actividades. Sin otro particular por el momento, reciba un cordial saludo.

Atentamente


Dr. Gabriel Guillén Solís



Cuernavaca, Mor., 28 de mayo de 2019.

Dr. David Martínez Duncker Ramírez
Director
Facultad de Medicina
Presente:

Estimado Dr. Martínez Duncker, por este conducto me permito informarle que he revisado el trabajo de Tesis **"CANCEROMA: DETECCIÓN MOLECULAR DE MARCADORES DE CÁNCER POR SECUENCIACIÓN MASIVA DE TERCERA GENERACIÓN"**, que para obtener el grado de Maestro en Medicina Molecular me proporcionó el alumno Luis Uribe Espejo Galicia. Le comunico que su contenido es adecuado y suficiente, por lo que de acuerdo al artículo 110 del Reglamento de Posgrado de la UAEM le otorgo:

Dictamen I se aprueba el trabajo de tesis tal como se presenta.

Así mismo, le agradezco la invitación a participar en este programa educativo, reiterando además mi disposición para seguir colaborando en este tipo de actividades. Sin otro particular por el momento, reciba un cordial saludo.

Atentamente



Dra. Vera Lucia Petricevich