



Genómica funcional para la descripción de mutaciones germinales en el diagnóstico molecular del cáncer de colon y recto no seleccionado en población colombiana

Juliana Valentina Rodríguez Salamanca

**Universidad del Rosario
Facultad de Ciencias Naturales
Bogotá, Colombia
2023**

Genómica funcional para la descripción de mutaciones germinales en el diagnóstico molecular del cáncer de colon y recto no seleccionado en población colombiana

Juliana Valentina Rodríguez Salamanca

Tesis presentada como requisito para obtener el título de:

Magister en Ciencias Naturales

Director

Dora Janeth Fonseca-Mendoza. Biol, MSc, PhD

Profesora Titular de Carrera

Escuela de Medicina y Ciencias de la Salud

Universidad del Rosario

Co-director

Adrien Morel. Bioq, MSc, PhD

Profesor Principal de Carrera

Escuela de Medicina y Ciencias de la Salud

Universidad del Rosario

**Facultad de Ciencias Naturales
Maestría en Ciencias Naturales
Universidad del Rosario
Bogotá, Colombia
2023**

AGRADECIMIENTOS

Agradezco al Hospital Universitario Mayor Méderi por su financiación a este proyecto y a la Universidad del Rosario por brindarme la oportunidad de llevar a cabo esta investigación en sus instalaciones.

A mis directores, la Dra. Dora Fonseca y Adrien Morel, por su orientación, apoyo y paciencia durante este proceso académico.

Al grupo de investigación CIGGUR y a todos sus miembros por el respaldo brindado durante la ejecución de este proyecto.

A mis padres, hermanas y demás familiares que siempre me han brindado su apoyo incondicional.

A mis amigos Miguel, Melissa, Catalina, Camilo y Nattaly por el cariño y comprensión en este proceso académico.

A los pacientes atendidos en el hospital Méderi y sus familias, que participaron en esta investigación.

CONTENIDO

1. Introducción	8
2. Materiales y métodos	9
2.1. Muestra Poblacional	9
2.2. Secuenciación de Exoma Completo	9
2.3. Análisis Bioinformático	11
2.4. Análisis Genético Poblacional	12
2.5. Análisis Estadístico	12
2.6. Construcción de plásmidos	13
2.7. Transfección en cultivos celulares	13
2.8. Extracción total de ARN y análisis de RT-PCR	13
3. Resultados	14
3.1. Características demográficas y clinicopatológicas	14
3.2. Variantes germinales patogénicas, probablemente patogénicas y Oncodriver identificada en la muestra de pacientes colombianos con CCR	15
3.2.1. Variantes moleculares identificadas de acuerdo con las recomendaciones del ACMG/AMP.	15
3.2.2. Variantes moleculares identificadas con el modelo BoostDM	22
3.3. Ensayo <i>in vitro</i> en variantes de sitios de splicing	28
4. Discusión	29
5. Conclusiones	37
6. Referencias	39

ÍNDICE DE TABLAS

Tabla 1. Panel de genes (206 genes)	10
Tabla 2. Secuencias de primers	13
Tabla 3. Caracterización demográfica y clínico-patológica	14
Tabla 4. Variantes germinales patogénicas y probablemente patogénicas identificadas en pacientes con CCR – Filtro A (ACMG/AMP)	17
Tabla 5. Variantes germinales de significado incierto (VUS) identificada en pacientes con CCR – Filtro A (ACMG/AMP)	17
Tabla 6. Variantes oncdriver germinales identificadas en pacientes con CCR – Filtro B (BoostDM)	23
Tabla 7. Comparación de frecuencias alélicas de variantes patogénicas/probablemente patogénicas y oncdrivers	26

ÍNDICE DE FIGURAS

Figura 1. Esquema metodológico del filtrado de variantes genómicas en pacientes con cáncer colorrectal.	12
Figura 2. A) variantes clasificadas según recomendaciones de la ACMG/AMP. B) penetrancia de los genes con variantes P/PP (filtro a). C) pacientes con 1 o >1 variante <i>oncdriver</i> clasificadas con el modelo boostdm. D) penetrancia de los genes con variantes <i>oncdriver</i> (filtro b).	25
Figura 3. Ensayo funcional en variantes P/PP intrónicas en sitios canónicos de splicing.	29

RESUMEN

El cáncer colorrectal (CCR) es el tercer tipo de cáncer de mayor incidencia a nivel mundial, con altas tasas de mortalidad reportadas anualmente. A pesar de que la secuenciación de próxima generación (NGS) ha permitido caracterizar perfiles genómicos mutacionales en diversas poblaciones, la información específica sobre pacientes colombianos con CCR es limitada. El objetivo de esta investigación es identificar variantes germinales asociadas al CCR en dicha población, utilizando un panel de 206 genes que incluye tanto genes de paneles de diagnóstico clínico como genes candidatos obtenidos de estudios de literatura.

La metodología empleada incluyó dos enfoques de clasificación: uno basado en las recomendaciones de la ACMG/AMP (*American College of Medical Genetics and Genomics and the Association for Molecular Pathology*) para identificar variantes patogénicas y probablemente patogénicas (P/PP), y otro utilizando el modelo de inteligencia artificial BoostDM. Los resultados obtenidos revelaron tasas significativas de variantes patogénicas, con un 12% de pacientes con variantes P/PP y un 65% con variantes “*oncdriver*” identificadas mediante BoostDM. Estos hallazgos sugieren la importancia de utilizar un panel ampliado en la detección de variantes germinales y la consideración de adoptar e indagar en nuevas estrategias de clasificación de variantes. Entre las variantes P/PP, se identificaron tres variantes intrónicas en sitios de splicing en genes candidatos. La validación funcional de estas variantes mediante un ensayo de minigenes demostró la generación de transcritos aberrantes, debido a la alteración en el splicing.

En conclusión, esta investigación proporcionó información valiosa sobre la presencia y frecuencia de variantes patogénicas en pacientes colombianos con CCR, usando un análisis genómico ampliado mediante NGS, utilizando dos enfoques bioinformáticos. Adicionalmente, se logró probar funcionalmente el efecto de tres variantes intrónicas de interés, que demostró la consecuencia molecular de estas y la potencial implicación a nivel de la proteína. En conjunto, este estudio contribuye al conocimiento del perfil genómico de pacientes no seleccionados con CCR en la población colombiana, generando nuevas perspectivas para la aplicación clínica y traslacional que busca la identificación temprana y la aplicación de estrategias que mejoren el pronóstico y supervivencia de los portadores de variantes de interés. Para nuestro conocimiento, esta corresponde a la primera aproximación en el país que aborda esta estrategia en pacientes no seleccionados con CCR.

Palabras clave: Secuenciación de próxima generación (NGS), variantes germinales patogénicas, inteligencia artificial, minigenes, validación funcional, cáncer colorrectal.

ABSTRACT

Colorectal cancer (CRC) is the third most common cancer worldwide, with high mortality rates reported annually. While next-generation sequencing (NGS) has facilitated the characterization of mutational genomic profiles in several populations, our understanding of Colombian patients with CRC is limited. The aim of this research is to identify germline variants associated with CRC in this population, using a panel of 206 genes that includes clinically diagnostic panel genes and candidate genes obtained from literature studies.

The methodology included two classification approaches: one based on ACMG/AMP (*American College of Medical Genetics and Genomics and the Association for Molecular Pathology*) recommendations to identify pathogenic and likely pathogenic variants (P/LP), and the other one using an artificial intelligence model called BoostDM. The results revealed significant rates of pathogenic variants, with 12% of patients having P/PP variants and 65% with oncogene variants identified through BoostDM. These findings suggest the importance of using an extended panel for germline variant detection and considering the adoption of new variant classification strategies. Within the P/LP variants, three intronic variants located at splicing sites in candidate genes were identified. Functional validation of these variants using a minigene assay demonstrated the generation of aberrant transcripts attributed to splicing alterations.

In conclusion, this study provided valuable insights into the prevalence and frequency of pathogenic variants in Colombian patients with CRC through NGS, using two different classification approaches. Additionally, the functional effect of three intronic variants was successfully tested, supporting the hypothesis of protein-level damage induced by these variants. Overall, this study contributes to the genomic knowledge of unselected CRC patients in the Colombian population, producing new perspectives for the clinical and translational application aimed at early identification and the implementation of strategies to improve the prognosis and survival of carriers of pathogenic variants. To our knowledge, this study represents the first approach in the country to address this strategy in unselected CRC patients.

Key words: Next Generation sequencing (NGS), pathogenic germline variants, artificial intelligence, minigenes, functional validation, colorectal cancer.

1. Introducción

El cáncer colorrectal (CCR) es el tercer tipo de cáncer más frecuente en todo el mundo, y constituye la segunda causa principal de muerte por cáncer (Globocan, 2020). Anualmente, el CCR tiene una incidencia de 1.8 millones de casos nuevos y se le atribuyen más de 900.000 muertes a nivel global (Globocan, 2020; Sung et al., 2021). En este contexto, alrededor de 1 de cada 23 hombres y 1 de cada 24 mujeres podrían ser diagnosticados con CCR en algún momento de sus vidas (Ferlitsch et al., 2011). Aunque el CCR es más común en los países desarrollados, las tasas de mortalidad han disminuido gracias a las estrategias de detección temprana, como la colonoscopia y la prueba de sangre oculta en heces (Ahmad et al., 2021). Según Globocan se estima que Colombia presenta una incidencia y tasa de mortalidad similar a la reportada globalmente, con aproximadamente 10.738 nuevos casos y 5.417 muertes reportadas anualmente (Globocan, 2020). En 2022, Colombia registró 44.371 nuevos casos de cáncer, de los cuales 3.851 correspondieron a CCR (*Cuenta de Alto Costo (CAC) /HIGIA, 2022*).

El CCR es una enfermedad compleja en la que las alteraciones genéticas y los factores ambientales desempeñan un rol crucial en su desarrollo y progresión (Lewandowska et al., 2022; Xu et al., 2021). Aunque más del 30% de los pacientes con CCR muestran historia familiar, se han detectado variantes patogénicas (P) y probablemente patogénicas (PP) en genes de predisposición al CCR en aproximadamente del 5 al 16% de los casos (T. J. Jiang et al., 2020; Patel & Ahnen, 2012; Rebutzi et al., 2023). Las variantes germinales patogénicas (VPs) en los genes relacionados con el CCR dan lugar a síndromes hereditarios, como el síndrome de Lynch (*MLH1, MSH2, MSH6, PMS2 y EPCAM*), la poliposis adenomatosa familiar (*APC*), la poliposis asociada a *MUTYH* y otros síndromes menos frecuentes, que en conjunto representan el 5% de todos los casos diagnosticados de CCR (Poliani et al., 2022). Dado que los pacientes que portan VPs en estos genes, están sujetos a estrategias adecuadas de vigilancia, según lo recomendado por las pautas de la Red Nacional Integral del Cáncer (NCCN, por sus siglas en inglés) para el CCR hereditario (National Comprehensive Cancer Network, 2023), se dispone de análisis genéticos de variantes germinales canónicas.

Informes recientes que utilizan secuenciación de próxima generación (NGS, por sus siglas en inglés) para evaluar variantes germinales en múltiples genes relacionados con el cáncer, han identificado VPs no canónicas potencialmente asociadas con el CCR en poblaciones seleccionadas y no seleccionadas (Gong et al., 2019; Poliani et al., 2022; Yurgelun et al., 2017). Por lo tanto, aumentar la tasa de detección de VPs a través de NGS maximizará la identificación de alteraciones moleculares en pacientes con CCR. En consecuencia, además de los genes relacionados con síndromes y trastornos genéticos que confieren un alto riesgo de CCR, se deben analizar otros genes de alta penetrancia tales como: *AKT, ATM, BMPR1A, BRAF, BRCA1/2, CHEK2, CTLA4, KRAS, MYO3A, PI3KCA, PTEN, RAS, SMAD2, SMAD3, TCF7L2, TGFBR2, TP53*, y otros clasificados como genes de penetrancia moderada o menos definida. Esta estrategia podrá potencialmente mejorar la identificación del riesgo conferido por variantes genéticas con la predisposición al CCR (Olkinuora et al., 2021; Rebutzi et al., 2023; Stodolna et al., 2021).

En Colombia, los estudios de mutaciones germinales en CCR se han centrado principalmente en genes sindrómicos como *APC, MLH1* y *TP53* (Alonso-Espinaco et al., 2011; Alvarez-Gomez et al.,

2021; Palacio-Rúa et al., 2014; Palacio Rúa et al., 2013). Sin embargo, la falta de datos genómicos en los pacientes con CCR y la frecuencia de las mutaciones germinales de susceptibilidad al CCR, dificultan la construcción de un perfil genómico de la población colombiana. El objetivo de esta investigación fue identificar variantes moleculares potencialmente relacionadas con la enfermedad en pacientes colombianos con CCR no seleccionado y determinar su frecuencia en genes asociados. El estudio analizó 100 pacientes afectados con CCR, y generó un panel genómico que fue evaluado mediante NGS y que incluyó 206 genes de interés. Adicionalmente, se realizó la validación de mutaciones intrónicas mediante ensayos funcionales de minigenes.

2. Materiales y métodos

2.1. Muestra Poblacional

Este estudio incluyó 100 pacientes que recibieron atención en el servicio de coloproctología del Hospital Méderi en Bogotá. Los pacientes colombianos mayores de 18 años, con una biopsia confirmatoria de cualquier tipo de cáncer colorrectal fueron elegidos y se les invitó a participar en el estudio, previa obtención de su consentimiento informado. Los datos sociodemográficos y clínicos se recopilaron a través de entrevistas y revisión de historias clínicas. Las variables evaluadas incluyeron sexo, edad, comorbilidades (hipertensión, diabetes mellitus, enfermedad pulmonar obstructiva crónica, cáncer y otras), antecedentes familiares de cáncer, hábitos de estilo de vida, pruebas de detección de CCR, altura, peso, edad al diagnóstico, ubicación del tumor, infiltración linfocelular, estadio y metástasis. El tamaño de la muestra se determinó por conveniencia e incluyó a pacientes atendidos en el hospital Méderi desde 2020 hasta 2022. Todos los procedimientos experimentales fueron aprobados por el Comité de Ética de la Universidad del Rosario y se llevaron a cabo de acuerdo con los principios establecidos en la Declaración de Helsinki (DVO005 1607-CV1436).

2.2. Secuenciación de Exoma Completo

Se realizó la extracción de ADN de los pacientes se extrajo utilizando el kit Quick-DNA 96 plus (*Zymo Research*). La calidad y cantidad de ADN se determinó utilizando el sistema *Quantifluor ONE dsDNA* en un instrumento *GloMax Discover* (Promega). La preparación de la biblioteca se realizó utilizando 250 ng de ADN utilizando el kit de preparación de bibliotecas *MGI Easy FS DNA*. Se llevó a cabo la fragmentación enzimática del ADN para obtener fragmentos que oscilaran entre 200 y 400 pb, seguida de la reparación de extremos y la amplificación por PCR. Se capturaron regiones específicas de interés utilizando la sonda de captura de exoma V5 y perlas de estreptavidina. Se emplearon cebadores específicos para el enriquecimiento en la reacción final de PCR. Para la secuenciación, el ADN se circularizó y la biblioteca se desnaturalizó después de la ligación de oligos divididos, seguida de la digestión y purificación utilizando perlas específicas. El ADN circularizado se utilizó para generar DNBs (nanobolas) a través del proceso de amplificación de *rolling circle* (MGI Tech Co, 2022). Los DNBs se cuantificaron y posteriormente se secuenciaron en la plataforma DNBSeg400. Las lecturas obtenidas se asignaron al genoma de referencia hg19 utilizando el software de alineamiento *Burrows-Wheeler* (BWA) y se organizaron con la herramienta SAMtools (<https://github.com/samtools/samtools>). Las lecturas duplicadas se identificaron y eliminaron utilizando el software Picard (<https://broadinstitute.github.io/picard/>). El análisis de cobertura y profundidad se realizó utilizando la herramienta BAMBA, y se consideró un umbral aceptable de

50X. Más del 93% de las bases totales llamadas tenían una puntuación de calidad con escala *Phred* superior a 30 (>Q30).

Se obtuvo un mínimo de 7 Gb de datos crudos, y el porcentaje de lecturas correctamente mapeadas fue >99.99% (56,446,362-101,662,468) por muestra. La eficiencia de mapeo promedio fue >99%, con una profundidad de secuenciación en el objetivo y una cobertura de la región objetivo de >50x y >94%, respectivamente. La fracción promedio del objetivo cubierta con >20x fue >72%. La uniformidad de la cobertura (10x) fue ≥90%, y la fracción promedio del objetivo cubierta con al menos 10x, 20x, 50x y 100x fue >90%, >72%, >40% y >19%, respectivamente. El número promedio de lecturas pareadas que se asignaron al genoma de referencia fue de 72,351,272 (99.99%).

La preparación de la biblioteca y la secuenciación fueron realizadas por Gencell Pharma (Bogotá, Colombia). El panel analizado incluyó 206 genes, que fueron seleccionados por su evidencia según estudios de asociación caso-control en CCR, su relevancia biológica, su participación en los procesos fisiopatológicos asociados con el CCR, derivados del análisis de GWAS y sus roles en la oncogénesis (este grupo fue considerado como genes candidatos). Adicionalmente, se incluyeron los genes incluidos en paneles de diagnóstico de CGC genetics, ICM Atrys division, Mayo Clinic Laboratories, GENDIA-genetics and molecular biology, CD Genomics disease panel, Centogene e Invitae. El panel definitivo quedó conformado por 102 genes de los paneles de diagnóstico establecidos y 104 genes candidatos (Tabla 1).

Tabla 1. Panel de genes (206 genes).

Panel de diagnóstico molecular				Panel de genes candidatos			
<i>AIP</i>	<i>EGFR</i>	<i>NF2</i>	<i>RUNX1</i>	<i>ACTR1B</i>	<i>EDN1</i>	<i>MAMSTR</i>	<i>SBF2</i>
<i>ALK</i>	<i>EPCAM</i>	<i>NOTCH2</i>	<i>SDHA</i>	<i>ALCAM</i>	<i>EIF3H</i>	<i>MGMT</i>	<i>SEMA4A</i>
<i>APC</i>	<i>EPM2A</i>	<i>NOTCH3</i>	<i>SDHAF2</i>	<i>APE1</i>	<i>ERCC1</i>	<i>MRE11</i>	<i>SF3A3</i>
<i>ATM</i>	<i>FAN1</i>	<i>NRAS</i>	<i>SDHB</i>	<i>ARFGEF2</i>	<i>EXO1</i>	<i>MRE11A</i>	<i>SFMBT1</i>
<i>AXIN2</i>	<i>FH</i>	<i>NTHL1</i>	<i>SDHC</i>	<i>ATF1</i>	<i>FAM109A</i>	<i>MYC</i>	<i>SH2B3</i>
<i>BAP1</i>	<i>FLCN</i>	<i>PALB2</i>	<i>SDHD</i>	<i>ATXN2</i>	<i>FANCC</i>	<i>MYO3A</i>	<i>SHROOM2</i>
<i>BARD1</i>	<i>GALNT12</i>	<i>PDGFRA</i>	<i>SMAD4</i>	<i>B9D2</i>	<i>FANCE</i>	<i>NABP1</i>	<i>SLC15A4</i>
<i>BLM</i>	<i>GATA2</i>	<i>PHOX2B</i>	<i>SMARCA4</i>	<i>BMP2</i>	<i>FEN1</i>	<i>NCAPG</i>	<i>SLC6A18</i>
<i>BMPR1A</i>	<i>GPC3</i>	<i>PIF1</i>	<i>SMARCB1</i>	<i>BMP4</i>	<i>FKBP5</i>	<i>NXN</i>	<i>SMAD6</i>
<i>BRAF</i>	<i>GREM1</i>	<i>PIK3CA</i>	<i>SMARCE1</i>	<i>BMP5</i>	<i>FMN1</i>	<i>OGG1</i>	<i>SMAD7</i>
<i>BRCA1</i>	<i>HOXB13</i>	<i>PMS1</i>	<i>STK11</i>	<i>BORA</i>	<i>FUT2</i>	<i>PIAS1</i>	<i>SMAD9</i>
<i>BRCA2</i>	<i>HRAS</i>	<i>PMS2</i>	<i>SUFU</i>	<i>C11orf53</i>	<i>GLI3</i>	<i>PITX1</i>	<i>SMARCD1</i>
<i>BRIP1</i>	<i>KIT</i>	<i>POLD1</i>	<i>TELO2</i>	<i>CABLES2</i>	<i>GNL1</i>	<i>PLCB1</i>	<i>TBX3</i>
<i>CASR</i>	<i>KRAS</i>	<i>POLE</i>	<i>TERC</i>	<i>CCND2</i>	<i>HHIP</i>	<i>PLGLA</i>	<i>TCF7L2</i>
<i>CCND1</i>	<i>MAX</i>	<i>POT1</i>	<i>TERT</i>	<i>CD44</i>	<i>HNF4A</i>	<i>PNKD</i>	<i>TFEB</i>
<i>CDC73</i>	<i>MEN1</i>	<i>PRKAR1A</i>	<i>TGFBR2</i>	<i>CDKN2B</i>	<i>IL12RB1</i>	<i>POLD3</i>	<i>TLE4</i>
<i>CDH1</i>	<i>MET</i>	<i>PTCH1</i>	<i>TMEM127</i>	<i>CHRDL2</i>	<i>KDR</i>	<i>POU5F1B</i>	<i>TMBIM1</i>
<i>CDK4</i>	<i>MITF</i>	<i>PTEN</i>	<i>TP53</i>	<i>COL4A2</i>	<i>KLF5</i>	<i>PRDM1</i>	<i>TMEM59</i>
<i>CDKN1B</i>	<i>MLH1</i>	<i>RAD50</i>	<i>TSC1</i>	<i>COLCA1</i>	<i>LAMA5</i>	<i>PREX1</i>	<i>TNS3</i>
<i>CDKN1C</i>	<i>MLH3</i>	<i>RAD51C</i>	<i>TSC2</i>	<i>COLCA2</i>	<i>LAMC1</i>	<i>PTPN1</i>	<i>TOX2</i>
<i>CDKN2A</i>	<i>MSH2</i>	<i>RAD51D</i>	<i>VHL</i>	<i>COX14</i>	<i>LGR5</i>	<i>PTPN12</i>	<i>TP53BP1</i>
<i>CEBPA</i>	<i>MSH3</i>	<i>RB1</i>	<i>WRN</i>	<i>CRTC3</i>	<i>LIG1</i>	<i>RBBP8</i>	<i>TTC22</i>
<i>CHEK2</i>	<i>MSH6</i>	<i>RBL1</i>	<i>WT1</i>	<i>CTNNB1</i>	<i>LIMA1</i>	<i>RHPN2</i>	<i>VTI1A</i>
<i>CTNNA1</i>	<i>MUTYH</i>	<i>RECQL4</i>	<i>XAF1</i>	<i>DCLRE1C</i>	<i>LIMK2</i>	<i>RPS20</i>	<i>WNT4</i>
<i>DICER1</i>	<i>NBN</i>	<i>RET</i>		<i>DIP2B</i>	<i>LRP1</i>	<i>RTEL1</i>	<i>XRCC2</i>
<i>DIS3L2</i>	<i>NF1</i>	<i>RNF43</i>		<i>DUSP10</i>	<i>MACC1</i>	<i>SATB2-AS1</i>	<i>ZAP70</i>

2.3. Análisis Bioinformático

Los archivos en formato *Variant Call Format* (VCF) se analizaron utilizando el software Varseq (Golden Helix). Incorporamos las siguientes anotaciones de bases de datos: ClinVar (<https://www.ncbi.nlm.nih.gov/clinvar/>), Ensembl (<https://www.ensembl.org/index.html>), RefSeq (<https://www.ncbi.nlm.nih.gov/refseq/>), predicciones funcionales de dbNSFP, dbSNP, REVEL, Ontología de Fenotipos OMIM (<https://www.omim.org/>), Variantes de UniProt (<https://www.uniprot.org/>) y gnomAD v2.1 (<https://gnomad.broadinstitute.org/>).

Para identificar las variantes patogénicas, probablemente patogénicas y las variantes oncodriver en los 206 genes relacionados con el CCR, aplicamos un análisis bioinformático utilizando dos estrategias de filtrado y clasificación: A) clasificación manual siguiendo las recomendaciones de la ACMG/AMP (Hampel et al., 2015) y B) utilizando el sistema de inteligencia artificial BoostDM (Muiños et al., 2021). Para la estrategia A, priorizamos las variantes moleculares con una frecuencia alélica menor (MAF) del 1%, incluyendo aquellas con pérdida de función (LoF) (*nonsense*, *frameshift* y en sitio de splicing) y mutaciones *missense*. El impacto funcional potencial de las variantes de splicing se evaluó utilizando los puntajes de los algoritmos *adaptive boosting* (ADA) y *random forest* (RF) (umbral ≥ 0.6). Para las variantes *missense*, priorizamos aquellas con resultados de patogenicidad *in silico* positivos en al menos 3 de los 6 predictores incorporados en el software VarSeq (SIFT, Polyphen2, Mutation Taster, Mutation Assessor, FATHMM y FATHMM MKL Coding). Las variantes clasificadas como relevantes desde el punto de vista funcional se clasificaron manualmente siguiendo las recomendaciones de la ACMG/AMP. Los datos finales consideraron variantes que cumplían con los criterios de patogénicas y probablemente patogénicas según lo definido por Hampel et al (2015). Para el filtro B, se consideraron todas las variantes con una MAF de menos del 5% y se analizaron utilizando el modelo de inteligencia artificial BoostDM, este análisis se llevó a cabo con la colaboración del Dr. Andrés López de la Universidad de Ecuador. Los datos finales consideraron las variantes clasificadas como mutaciones oncodriver según lo descrito por Muiños et al (2021) (Figura 1).

2.6. Construcción de plásmidos

Se usó ADN genómico de pacientes para amplificar la región que abarcó el exón más cercano a la mutación, junto con 300 pb de las secuencias intrónicas flanqueantes (*up* y *downstream* del exón). Los primers fueron diseñados siguiendo la guía de Putscher et al., 2021 y verificados mediante Primer-BLAST (Tabla 2) (Putscher et al., 2021). La PCR se realizó utilizando Master Mix Q5 de New England Biolabs (Q5® High-Fidelity 2X Master Mix, cat: M049). Los productos de PCR se recombinaron con el vector pSpliceExpress utilizando la mezcla de enzimas Gateway™ BP Clonase™ II (Invitrogen), siguiendo las instrucciones del fabricante. La secuencia de la construcción del vector se confirmó mediante secuenciación de Sanger.

Tabla 2. Secuencias de primers de recombinación.

Gen	Variante	Transcrito	Secuencia de primers	Tamaño del amplicón (pb)
<i>ERCC1</i>	c.702+1G>A	NM_202001.3	GGGGACAAGTTTGTACAAAAAAGCAGGCTCTGAAATGGAGTTTCACTCTTGTTC	868
<i>IL12RB1</i>	c.1791+2T>G	NM_005535.3	GGGGACAAGTTTGTACAAAAAAGCAGGCTTAGGGAAGGGATAGTAACTTTTGG	918
<i>SMAD9</i>	c.781+2T>A	NC_000013.11	GGGGACAAGTTTGTACAAAAAAGCAGGCTAAAGTCTTGCCCTGAAGTCATTAG	664

2.7. Transfección en cultivos celulares

Las líneas celulares HCT-116 y HEK-293 se cultivaron en medio Dulbecco's Modified Eagle's (DMEM), suplementado con un 10% de suero fetal bovino, 100 U/mL de penicilina/estreptomicina a 37°C en un ambiente con un 5% de CO₂. Las células se sembraron a una densidad de 60.000 células por pozo, con triplicados para cada condición experimental. Se transfectó 1 µg de vector en las células y se incubó durante 48 horas utilizando el reactivo de transfección FuGENE® 6 (Promega), según el protocolo del fabricante.

2.8. Extracción total de ARN y análisis de RT-PCR

Se extrajo el ARN total usando TRIzol®. Se usaron 500 ng de ARN para sintetizar ADNc con *SuperScript III cDNA first strand* (Invitrogen). Utilizando primers dirigidos a los exones 2 y 3 de la insulina 2 de rata (cuya secuencia se encuentra en el vector pSpliceExpress y rodean las secuencias intrónicas del proceso de clonación Gateway), se realizó la PCR bajo las siguientes condiciones: 95°C durante 1 minuto, 60°C durante 40 segundos y luego 3 minutos a 72°C durante 30 ciclos. Los productos de RT-PCR se visualizaron en geles de agarosa al 1,2%.

Las bandas obtenidas fueron analizadas mediante densitometría utilizando el software Image Lab (BIO-RAD). Para cada una de las bandas obtenidas en los ensayos de RT-PCR en las versiones *Wild-type* (WT) y mutante (MUT) se realizó la medición de la intensidad, que permitió establecer posteriormente diferencias en los transcritos obtenidos. Para las comparaciones entre los dos grupos experimentales y sus réplicas, se utilizó una prueba t de grupos independientes (t de Student), utilizando el software SPSS V.29 y GraphPad Prism 10. Se considera una diferencia estadísticamente significativa si p<0.05.

Los productos de RT-PCR WT y MUT se clonaron utilizando el Kit de Clonación TOPO® TA para Secuenciación (Invitrogen) siguiendo las instrucciones del fabricante. Luego del proceso de transformación bacteriana usando las células quimiocompetentes *One Shot Top 10*

(<https://www.thermofisher.com/>) se seleccionaron al azar 10 colonias, para las cuales se realizó la secuenciación de Sanger luego de la extracción del ADN plasmídico.

3. Resultados

3.1. Características demográficas y clinicopatológicas

Este estudio incluyó una cohorte de 100 pacientes no seleccionados con un diagnóstico histopatológico de CCR. Las características de los pacientes se resumen en la Tabla 3. Entre estos individuos, las mujeres constituían la mayoría, representando el 55% (n=55) de los pacientes. El 63% (n=63) de los pacientes tenían antecedentes familiares de cáncer, con un 11% (n=11) de los casos correspondientes al CCR. En cuanto a la ubicación del tumor, el CCR del lado derecho era el más prevalente, seguido por el sigmoide y rectal (44% y 29% respectivamente). La muestra analizada mostró una heterogeneidad en cuanto a la histología tumoral y el estadio del cáncer, con una mayor proporción de adenocarcinoma moderadamente diferenciado y estadio II (30% y 36% respectivamente). La mayoría de los pacientes no presentaron metástasis en el momento de su inclusión en el estudio (76%). La edad promedio en el momento del diagnóstico de CCR fue de 65.5 años y el 12% (n=12) de los pacientes fueron diagnosticados antes de los 50 años. El análisis de asociación de variables clínicas con la edad del diagnóstico mostró diferencias estadísticamente significativas con antecedentes familiares de CCR ($p < 0.05$) (Tabla 3).

Tabla 3. Caracterización demográfica y clínico-patológica

Características	Total	Edad del diagnóstico		p-valor	p-valor Corrección de Yates
		< 50	≥ 50		
Número de pacientes	100	12	88		
Género				0.035*	0.073
Hombres	45	2	43		
Mujeres	55	10	45		
Edad del diagnóstico de CCR					
Promedio (SD)	65,5 (12,6)	41,25 (8,4)	68,8 (8,9)		
Antecedentes familiares de cáncer				0.632	0.889
No	32	3	29		
Sí	63	8	55		
Desconocido	5	1	4		
Antecedentes familiares de CCR				0.004*	0.017*
No	81	6	75		
Sí	11	4	7		
Desconocido	8	2	6		
Localización del tumor					
Colon derecho	44	2	42	0.042*	0.085
Sigmoide	29	6	23	0.087	0.171
Recto	17	4	13	0.108	0.232
Todo el colon	1	0	1	0.711	1
Múltiples sitios	9	0	9	0.246	0.533
Estado histológico del tumor					
Adenocarcinoma bien diferenciado	7	0	7	0.311	0.682
Adenocarcinoma moderadamente diferenciado	30	4	26	0.788	1
Adenocarcinoma mal diferenciado	1	0	1	0.711	1
Adenocarcinoma mucinoso	3	1	2	0.248	0.801
Adenocarcinoma infiltrante	3	1	2	0.248	0.801

Múltiples estados histológicos	56	6	50	0.655	0.892
Estadio del cáncer				0.484	0.484
0	2	1	1		
I	14	1	13		
II	36	4	32		
III	32	3	29		
IV	15	2	13		
Desconocido	1	1	0		
Infiltración linfovascular				0.188	0.319
Sí	53	8	45		
No	45	3	42		
Desconocido	2	1	1		
Metástasis				0.719	1
Sí	22	2	20		
No	76	9	67		
Desconocido	2	1	1		
Variantes patogénicas o probablemente patogénicas (ACMG/AMP)				0.140	0.315
Sí	12	3	9		
No	88	9	79		
Variantes Oncodriver (BoostDM)				0.156	0.273
Sí	65	10	55		
No	35	2	33		

*Estadísticamente significativo

3.2. Variantes germinales patogénicas, probablemente patogénicas y Oncodriver identificadas en la muestra de pacientes colombianos con CCR

3.2.1. Variantes moleculares identificadas de acuerdo con las recomendaciones del ACMG/AMP.

Siguiendo el flujo de trabajo bioinformático descrito en la metodología (Figura 1) y aplicando el Filtro A, que identificó mutaciones patogénicas y probablemente patogénicas de acuerdo con las recomendaciones de la ACMG/AMP, nuestro estudio detectó un total de 248 variantes (Figura 2A). De estas, 5 (2%) se clasificaron como patogénicas (P), 8 (3.2%) como probablemente patogénicas (PP), 153 (61.7%) como variantes de significado incierto (VUS) y 82 (33.1%) como probablemente benignas (PB) (Tabla 4, Tabla 5 y Figura 2A). La tasa de detección de variantes P/PP fue del 12% (12 pacientes). Un total del 11% de los pacientes presentaron una sola mutación P/PP en estado heterocigoto, mientras que un paciente fue identificado con dos variantes, *MLH1* c.1039delA e *IL12RB1* c.1791+2T>G, clasificadas como PP y P, respectivamente.

Entre las variantes P/PP identificadas en la población de estudio, el 38.5% (5/13) se encontraban en genes que normalmente no se incluyen en los paneles de diagnóstico de rutina en la práctica clínica. Estas variantes corresponden a *ERCC1* c.702+1G>A, *EXO1* c.1465delA, *IL12RB1* c.1791+2T>G, *OGG1* c.137G>A y *SMAD9* c.781+2T>A (Tabla 4). De las variantes que cumplieron con los criterios de clasificación P/PP, el 61.5% (8/13) se encontraban en genes seleccionados de los paneles de diagnóstico molecular. Cuatro de ellas pertenecen a genes de alta penetrancia (*BMPRI1* c.176T>A, *MLH1* c.1039delA, *MSH6* c.3516_3517delAG y *PTCH1* c.3241G>A) y una pertenece a un gen de penetrancia moderada (*BARD1* c.2229dupT). Para ocho de los genes analizados donde se

encontraron variantes de interés, la penetrancia no ha sido definida claramente (*FLCN*, *NOTCH3*, *NTHL1*, *ERCC1*, *EXO1*, *IL12RB1*, *OGG1* y *SMAD9*) (Tabla 4 y Figura 2B).

En cuanto a los tipos de variantes, observamos 6 de pérdida de función (LoF), 4 *missense* y 3 en sitios canónicos de splicing. De manera interesante, 2 variantes PP eran nuevas y no se habían reportado en bases de datos públicas como gnomAD (*MLH1*: c.1039del, p.Thr347Leufs*20 y *EXO1*:c.1465delA, p.Arg489Glyfs*32) (Tabla 4). Todas las variantes clasificadas como P/PP se confirmaron mediante secuenciación de Sanger.

Tabla 4. Variantes germinales patogénicas y probablemente patogénicas identificadas en pacientes con CCR – Filtro A (ACMG/AMP).

Penetrancia	Panel	Gen	Transcrito	Variante	Proteína	rs	Cigocidad	Tipo	Clasificación ACMG/AMP	Criterios cumplidos	No. pacientes
Alta	Diagnóstico	<i>BMPR1A</i>	NM_004329.3	c.176T>A	p.Leu59Ter	rs1564714834	Het	<i>Missense</i>	Patogénica	PVS1 + PM2 + PM4 + BP1	1
Alta	Diagnóstico	<i>MLH1</i>	NM_000249.4	c.1039delA	p.Thr347Leufs*20	No reportada	Het	<i>LOF</i>	Probablemente patogénica	PVS1 + PM2	1
Alta	Diagnóstico	<i>MSH6</i>	NM_000179.3	c.3516_3517delAG	p.Arg1172Serfs*4	rs398123232	Het	<i>LOF</i>	Patogénica	PVS1 + PM2 + PM4	1
Alta	Diagnóstico	<i>PTCH1</i>	NM_000264.5	c.3241G>A	p.Val1081Met	rs587778629	Het	<i>Missense</i>	Probablemente patogénica	PS3 + PM2 + PP3	1
Menos definida	Diagnóstico	<i>FLCN</i>	NM_144997.7	c.1285delC	p.His429Thrfs*39	rs80338682	Het	<i>LOF</i>	Probablemente patogénica	PVS1 + PM2	1
Menos definida	Diagnóstico	<i>NOTCH3</i>	NM_000435.3	c.1345C>T	p.Arg449Cys	rs762734007	Het	<i>Missense</i>	Probablemente patogénica	PM1 + PM2 + PP2 + PP3	1
Menos definida	Diagnóstico	<i>NTHL1</i>	NM_002528.7	c.244C>T	p.Gln82Ter	rs150766139	Het	<i>LOF</i>	Patogénica	PVS1 + PM2 + PM4	1
Moderada	Diagnóstico	<i>BARD1</i>	NM_000465.4	c.2229dupT	p.Asn744Ter	rs1259296823	Het	<i>LOF</i>	Patogénica	PVS1 + PM1 + PM2 + PM4	1
Menos definida	Candidato	<i>ERCC1</i>	NM_202001.3	c.702+1G>A	-	rs747911302	Het	<i>Splice</i>	Probablemente patogénica	PVS1 + PM2	1
Menos definida	Candidato	<i>EXO1</i>	NM_130398.4	c.1465delA	p.Arg489Glyfs*32	No reportada	Het	<i>LOF</i>	Probablemente patogénica	PVS1 + PM2	1
Menos definida	Candidato	<i>IL12RB1</i>	NM_005535.3	c.1791+2T>G	-	rs554063682	Het	<i>Splice</i>	Patogénica	PVS1 + PM2 + PP3	1
Menos definida	Candidato	<i>OGG1</i>	NM_002542	c.137G>A	p.Arg46Gln	rs104893751	Het	<i>Missense</i>	Probablemente patogénica	PS3 + PM2	1
Menos definida	Candidato	<i>SMAD9</i>	NM_001127217.3	c.781+2T>A	-	rs770716081	Het	<i>Splice</i>	Probablemente patogénica	PVS1 + PM2 + BP4	1

Abreviaturas: Het, heterocigoto; LOF, Pérdida de función.

Tabla 5. Variantes germinales de significado incierto (VUS) identificadas en pacientes con CCR – Filtro A (ACMG/AMP).

Panel	Gen	Transcrito	Variante	Proteína	rs	Tipo	Clasificación ACMG/AMP	Criterios cumplidos	No. pacientes
Diagnóstico	<i>AIP</i>	NM_003977.4	c.296C>T	p.Pro99Leu	rs1388668435	<i>Missense</i>	VUS	PM1 + PM2 + PP3	1
Diagnóstico	<i>ALK</i>	NM_004304.5	c.4425C>A	p.His1475Gln	rs139039449	<i>Missense</i>	VUS	PM2	2
Diagnóstico	<i>APC</i>	NM_000038.6	c.4178T>C	p.Leu1393Pro	No reportada	<i>Missense</i>	VUS	PM2 + PP3 +BP1	1
Diagnóstico	<i>ATM</i>	NM_000051.4	c.295A>G	p.Ser99Gly	rs137882485	<i>Missense</i>	VUS	PM2	1
Diagnóstico	<i>ATM</i>	NM_000051.4	c.610G>A	p.Gly204Arg	rs147915571	<i>Missense</i>	VUS	PM2	1
Diagnóstico	<i>ATM</i>	NM_000051.4	c.1082C>A	p.Thr361Asn	rs1565378803	<i>Missense</i>	VUS	PM2	1

Diagnóstico	ATM	NM_000051.4	c.5039C>T	p.Pro1680Leu	rs587782153	Missense	VUS	PM2 + PP3	1
Diagnóstico	ATM	NM_000051.4	c.6313A>G	p.Arg2105Gly	rs879253983	Missense	VUS	PM1 + PM2 + PP3	2
Diagnóstico	ATM	NM_000051.4	c.7375C>G	p.Arg2459Gly	rs730881383	Missense	VUS	PM1 + PM2 + PP3	1
Diagnóstico	ATM	NM_000051.4	c.7381C>T	p.Arg2461Cys	rs201314561	Missense	VUS	PM1 + PM2 + PP3	1
Diagnóstico	ATM	NM_000051.4	c.7502A>G	p.Asn2501Ser	rs531617441	Missense	VUS	PM1 + PM2	1
Diagnóstico	AXIN2	NM_004655.4	c.733C>T	p.Pro245Ser	rs62640028	Missense	VUS	PM2	1
Diagnóstico	AXIN2	NM_004655.4	c.1419_1421delCCA	p.His474del	rs570443161	LOF	VUS	PM2	1
Diagnóstico	BLM	NM_000057.4	c.2452C>T	p.Arg818Cys	rs1279814185	Missense	VUS	PM1 + PM2	1
Diagnóstico	BRCA1	NM_007294.4	c.5225A>G	p.Asn1742Ser	rs864622104	Missense	VUS	PM1 + PM2	1
Diagnóstico	CHEK2	NM_007194.4	c.1427C>T	p.Thr476Met	rs142763740	Missense	VUS	PM1 + PM2	1
Diagnóstico	CHEK2	NM_007194.4	c.320-5T>A	-	rs121908700	Splice	VUS	PM2 + PP3	1
Diagnóstico	EGFR	NM_005228.5	c.2380C>A	p.Pro794Thr	rs370289230	Missense	VUS	PM2 + PP3 + BP1	1
Diagnóstico	FH	NM_000143.4	c.535G>A	p.Asp179Asn	rs1553341588	Missense	VUS	PM1 + PM2 + PP3	1
Diagnóstico	FLCN	NM_144997.7	c.535C>T	p.Arg179Trp	rs774358971	Missense	VUS	PM1 + PM2 + PP3	1
Diagnóstico	GALNT12	NM_024642.5	c.544C>T	p.His182Tyr	rs1835864292	Missense	VUS	PM2	1
Diagnóstico	HOXB13	NM_006361.6	c.23C>T	p.Thr8Ile	rs747307642	Missense	VUS	PM2	1
Diagnóstico	MLH1	NM_000249.4	c.1514G>A	p.Ser505Asn	rs771044689	Missense	VUS	PM2 + PP3	1
Diagnóstico	MSH1	NM_000251.3	c.128A>G	p.Tyr43Cys	rs17217723	Missense	VUS	PM2 + PP3	1
Diagnóstico	MSH2	NM_000251.3	c.743A>G	p.Lys248Arg	rs1064794704	Missense	VUS	PM2	1
Diagnóstico	MSH2	NM_000251.3	c.2537A>G	p.Gln846Arg	rs140754514	Missense	VUS	PM2	1
Diagnóstico	MSH3	NM_002439.5	c.1622C>T	p.Thr541Ile	rs1178614269	Missense	VUS	PM2 + PP3	1
Diagnóstico	MSH6	NM_000179.3	c.1403G>A	p.Arg468His	rs41295268	Missense	VUS	PM2 + PP3	1
Diagnóstico	MSH6	NM_000179.3	c.2041C>T	p.Leu681Phe	rs1553413412	Missense	VUS	PM2 + PP3	1
Diagnóstico	MSH6	NM_000179.3	c.2653A>C	p.Lys885Gln	rs587782593	Missense	VUS	PM2	1
Diagnóstico	MSH6	NM_000179.3	c.3961A>G	p.Arg1321Gly	rs41295278	Missense	VUS	PM2 + PP3	1
Diagnóstico	MUTYH	NM_001128425.2	c.1276C>T	p.Arg426Cys	rs150792276	Missense	VUS	PM2	1
Diagnóstico	NBN	NM_002485.5	c.643C>T	p.Arg215Trp	rs34767364	Missense	VUS	PM2	1
Diagnóstico	NOTCH2	NM_024408.4	c.5927A>G	p.His1976Arg	No reportada	Missense	VUS	PM1 + PM2 + PP2 + BP1	1
Diagnóstico	NOTCH3	NM_000435.3	c.6464G>A	p.Gly2155Glu	rs567953412	Missense	VUS	PM2 + PP2	1
Diagnóstico	NOTCH3	NM_000435.3	c.6305T>C	p.Val2102Ala	No reportada	Missense	VUS	PM2 + PP2	1
Diagnóstico	NOTCH3	NM_000435.3	c.2935G>A	p.Ala979Thr	rs1286854173	Missense	VUS	PM1 + PM2 + PP2	1
Diagnóstico	PALB2	NM_024675.4	c.3044C>A	p.Thr1015Asn	No reportada	Missense	VUS	PM1 + PM2	1
Diagnóstico	PIF1	NM_001286496.2	c.838C>T	p.Pro280Ser	rs201566855	Missense	VUS	PM2	4
Diagnóstico	PIF1	NM_001286496.2	c.314T>A	p.Leu105His	rs960648006	Missense	VUS	PM2 + PP3	1
Diagnóstico	PMS1	NM_000534.5	c.329C>G	p.Thr110Arg	rs372752293	Missense	VUS	PM2 + PP3	1
Diagnóstico	PMS1	NM_000534.5	c.2311C>A	p.Pro771Thr	rs143686298	Missense	VUS	PM2	1
Diagnóstico	PMS2	NM_000535.7	c.2395C>T	p.Arg799Trp	rs149202766	Missense	VUS	PM2 + PP3 + BS2	2
Diagnóstico	PMS2	NM_000535.7	c.880C>T	p.Arg294Trp	rs563433235	Missense	VUS	PM2	1
Diagnóstico	POLD1	NM_002691.4	c.520C>T	p.Arg174Trp	rs749334182	Missense	VUS	PM2	1
Diagnóstico	POLD1	NM_002691.4	c.2017G>A	p.Glu673Lys	rs61751955	Missense	VUS	PM2	1
Diagnóstico	POLD1	NM_002691.4	c.2052G>C	p.Gln684His	rs144143245	Missense	VUS	PM2	1
Diagnóstico	POLE	NM_006231.4	c.4145G>A	p.Arg1382His	rs143229302	Missense	VUS	PM2	2

Diagnóstico	<i>POLE</i>	NM_006231.4	c.686A>C	p.His229Pro	rs1482513360	Missense	VUS	PM2 + PP3	1
Diagnóstico	<i>POLE</i>	NM_006231.4	c.4952+5G>A	-	rs1555222548	Splice	VUS	PM2	1
Diagnóstico	<i>PTCH1</i>	NM_000264.5	c.3487G>A	p.Gly1163Ser	rs113663584	Missense	VUS	PM2 + PP3 + BP1	1
Diagnóstico	<i>PTCH1</i>	NM_000264.5	c.1664A>G	p.Asn555Ser	rs181192122	Missense	VUS	PM1 + PM2 + PP3 + BP1	1
Diagnóstico	<i>PTCH1</i>	NM_000264.5	c.1808G>A	p.Arg603His	rs199523893	Missense	VUS	PM2 + PP3 + BP1	1
Diagnóstico	<i>RAD50</i>	NM_005732.4	c.113A>G	p.Asn38Ser	rs750480943	Missense	VUS	PM2	1
Diagnóstico	<i>RAD50</i>	NM_005732.4	c.3036+5G>A	-	rs181016343	Splice	VUS	PM2 + BS2	1
Diagnóstico	<i>RAD51C</i>	NM_058216.3	c.659T>C	p.Leu220Pro	No reportada	Missense	VUS	PM1 + PM2 + PP3	1
Diagnóstico	<i>RBL1</i>	NM_002895.5	c.3142C>G	p.Gln1048Glu	rs765000152	Missense	VUS	PM2	1
Diagnóstico	<i>RBL1</i>	NM_002895.5	c.2205A>G	p.Ile735Met	rs78413638	Missense	VUS	PM2	1
Diagnóstico	<i>RBL1</i>	NM_002895.5	c.1991A>T	p.Asp664Val	rs200548883	Missense	VUS	PM2	1
Diagnóstico	<i>RBL1</i>	NM_002895.5	c.781G>A	p.Val261Ile	rs746560730	Missense	VUS	PM2	1
Diagnóstico	<i>RNF43</i>	NM_017763.6	c.2276A>C	p.Tyr759Ser	No reportada	Missense	VUS	PM2	1
Diagnóstico	<i>RUNX1</i>	NM_001754.5	c.998C>G	p.Pro333Arg	rs200104203	Missense	VUS	PM2	1
Diagnóstico	<i>SDHA</i>	NM_004168.4	c.1367C>T	p.Ser456Leu	rs76896145	Missense	VUS	PM2 + PP3	2
Diagnóstico	<i>SDHA</i>	NM_004168.4	c.1396G>A	p.Ala466Thr	rs111387770	Missense	VUS	PM2	2
Diagnóstico	<i>SDHD</i>	NM_003002.4	c.149A>G	p.His50Arg	rs11214077	Missense	VUS	PM2 + PP3 + BS2	3
Diagnóstico	<i>SMARCA4</i>	NM_001128849.3	c.403C>G	p.Pro135Ala	rs150949949	Missense	VUS	PM2 + PP2 + BP1	1
Diagnóstico	<i>TELO2</i>	NM_016111.4	c.1547G>A	p.Arg516Gln	rs141154049	Missense	VUS	PM2	1
Diagnóstico	<i>TERT</i>	NM_198253.3	c.1210C>G	p.Pro404Ala	rs1561213668	Missense	VUS	PM1 + PM2 + PP2 + BP1	1
Diagnóstico	<i>TERT</i>	NM_198253.3	c.316G>C	p.Gly106Arg	rs1751265277	Missense	VUS	PM1 + PM2 + PP2 + BP1	1
Diagnóstico	<i>TSC2</i>	NM_000548.5	c.2161G>A	p.Val721Met	rs768261185	Missense	VUS	PM2 + PP3 + BP1	1
Diagnóstico	<i>WRN</i>	NM_000553.6	c.1192A>G	p.Thr398Ala	rs532554615	Missense	VUS	PM2	1
Diagnóstico	<i>WRN</i>	NM_000553.6	c.1717A>G	p.Thr573Ala	rs150148567	Missense	VUS	PM1 + PM2	1
Diagnóstico	<i>WRN</i>	NM_000553.6	c.4216C>T	p.Arg1406Ter	rs11574410	LOF	VUS	PVS1 + PM1 + PM2 + PM4+ BS2 + BP4	2
Candidato	<i>ALCAM</i>	NM_001627.4	c.1679A>G	p.His560Arg	No reportada	Missense	VUS	PM2	1
Candidato	<i>APEX1</i>	NM_001641.4	c.5C>G	p.Pro2Arg	rs752031447	Missense	VUS	PM1 + PM2	1
Candidato	<i>ATXN2</i>	NM_002973.4	c.57_59delACA	p.Gln28del	rs751377502	LOF	VUS	PM1 + PM2	1
Candidato	<i>C11ORF53</i>	NM_198498.3	c.194T>C	p.Val65Ala	rs371749769	Missense	VUS	PM2	1
Candidato	<i>COL4A2</i>	NM_001846.4	c.349G>A	p.Asp117Asn	rs374976511	Missense	VUS	PM2	1
Candidato	<i>COL4A2</i>	NM_001846.4	c.923G>A	p.Gly308Asp	No reportada	Missense	VUS	PM2 + PP3	1
Candidato	<i>COL4A2</i>	NM_001846.4	c.4064T>C	p.Met1355Thr	rs199956455	Missense	VUS	PM2	1
Candidato	<i>COL4A2</i>	NM_001846.4	c.4561G>A	p.Glu1521Lys	rs182278784	Missense	VUS	PM1 + PM2 + PP3	1
Candidato	<i>CRTC3</i>	NM_022769.5	c.1789A>C	p.Met597Leu	rs777546566	Missense	VUS	PM2	1
Candidato	<i>CTNNB1</i>	NM_001904.4	c.991T>C	p.Tyr331His	No reportada	Missense	VUS	PM1 + PM2 + PP2	1
Candidato	<i>FANCC</i>	NM_000136.3	c.851C>T	p.Ala284Val	rs201281511	Missense	VUS	PM2	2
Candidato	<i>FANCC</i>	NM_000136.3	c.1393_1395del	p.Pro465del	rs587778328	LOF	VUS	PM2 + PM4	1
Candidato	<i>FANCE</i>	NM_021922.3	c.1316G>T	p.Gly439Val	rs776298967	Missense	VUS	PM2	1
Candidato	<i>FEN1</i>	NM_004111.6	c.787C>G	p.Leu263Val	rs148519764	Missense	VUS	PM2	1
Candidato	<i>FMN1</i>	NM_001277313.2	c.1307G>T	p.Gly436Val	rs537060596	Missense	VUS	PM2	1
Candidato	<i>FUT2</i>	NM_000511.6	c.436C>T	p.Arg146Cys	rs201915844	Missense	VUS	PM2 + PP2 + BS2	19
Candidato	<i>IL12RB1</i>	NM_005535.3	c.1295C>T	p.Thr432Met	rs200875188	Missense	VUS	PM2	1

Candidato	KDR	NM_002253.4	c.2555A>T	p.Asp852Val	No reportada	Missense	VUS	PM1 + PM2 + PP3	1
Candidato	LAMA5	NM_005560.6	c.6421C>T	p.Arg2141Cys	rs745369447	Missense	VUS	PM1 + PM2	1
Candidato	LAMA5	NM_005560.6	c.6040G>A	p.Ala2014Thr	rs372254142	Missense	VUS	PM1 + PM2	1
Candidato	LAMA5	NM_005560.6	c.5539C>T	p.Pro1847Ser	No reportada	Missense	VUS	PM1 + PM2	1
Candidato	LAMA5	NM_005560.6	c.5232C>G	p.Ile1744Met	No reportada	Missense	VUS	PM1 + PM2	1
Candidato	LAMA5	NM_005560.6	c.4679G>A	p.Arg1560His	rs137986729	Missense	VUS	PM1 + PM2	1
Candidato	LAMA5	NM_005560.6	c.3671G>A	p.Arg1224His	rs199705213	Missense	VUS	PM2	1
Candidato	LAMA5	NM_005560.6	c.2323C>T	p.Arg775Cys	rs201613863	Missense	VUS	PM1 + PM2	1
Candidato	LAMC1	NM_002293.4	c.1162C>T	p.Arg388Cys	rs769003582	Missense	VUS	PM1 + PM2 + PP3 + BP1	1
Candidato	LGR5	NM_003667.4	c.1909T>G	p.Cys637Gly	No reportada	Missense	VUS	PM2 + PP3	1
Candidato	LIG1	NM_000234.3	c.350G>A	p.Gly117Glu	No reportada	Missense	VUS	PM2	1
Candidato	LIG1	NM_000234.3	c.8G>A	p.Arg3Gln	rs201713810	Missense	VUS	PM2	1
Candidato	LIMA1	NM_016357.5	c.2134G>C	p.Asp712His	No reportada	Missense	VUS	PM2	1
Candidato	LIMA1	NM_016357.5	c.1383C>G	p.His461Gln	rs1474165830	Missense	VUS	PM2	1
Candidato	LIMK2	NM_005569.4	c.308T>C	p.Val103Ala	rs188450615	Missense	VUS	PM1 + PM2	1
Candidato	LRP1	NM_002332.3	c.145G>A	p.Gly49Ser	rs375203123	Missense	VUS	PM2 + PP2 + BP1	1
Candidato	LRP1	NM_002332.3	c.12685G>C	p.Asp4229His	No reportada	Missense	VUS	PM1 + PM2 + PP2 + BP1	1
Candidato	MACC1	NM_182762.4	c.2108A>T	p.Asp703Val	rs574575235	Missense	VUS	PM2	1
Candidato	MRE11	NM_005591.4	c.2048G>A	p.Gly683Glu	rs764705257	Missense	VUS	PM2 + PP3	1
Candidato	MYO3A	NM_017433.5	c.170A>C	p.Asp57Ala	rs146511800	Missense	VUS	PM1 + PM2	1
Candidato	MYO3A	NM_017433.5	c.1145A>G	p.Gln382Arg	rs200715190	Missense	VUS	PM1 + PM2	1
Candidato	MYO3A	NM_017433.5	c.1559C>T	p.Ala520Val	rs72787346	Missense	VUS	PM1 + PM2 + PP3	1
Candidato	MYO3A	NM_017433.5	c.3037C>T	p.Arg1013Cys	rs140154015	Missense	VUS	PM1 + PM2	1
Candidato	MYO3A	NM_017433.5	c.3217A>G	p.Arg1073Gly	rs754613454	Missense	VUS	PM1 + PM2	2
Candidato	NCAPG	NM_022346.5	c.2980G>A	p.Ala994Thr	rs141707867	Missense	VUS	PM2	1
Candidato	NXN	NM_022463.5	c.709G>A	p.Asp237Asn	No reportada	Missense	VUS	PM1 + PM2 + PP3	1
Candidato	OGG1	NM_002542.6	c.137G>A	p.Arg46Gln	rs104893751	Missense	VUS	PM2	1
Candidato	OGG1	NM_002542.6	c.565+5G>A	-	No reportada	Splice	VUS	PM2 + PP3	1
Candidato	PNKD	NM_015488.5	c.323A>G	p.His108Arg	rs142536637	Missense	VUS	PM2 + PP3	1
Candidato	PNKD	NM_015488.5	c.559C>T	p.Arg187Trp	rs375550686	Missense	VUS	PM1 + PM2	1
Candidato	PNKD	NM_015488.5	c.964C>T	p.Arg322Trp	rs751050132	Missense	VUS	PM2 + PP3	1
Candidato	PNKD	NM_015488.5	c.1126C>T	p.Arg376Cys	rs200042352	Missense	VUS	PM2 + PP3	1
Candidato	POU5F1B	NM_001159542.3	c.118G>C	p.Gly40Arg	rs1563683877	Missense	VUS	PM2	1
Candidato	POU5F1B	NM_001159542.3	c.839G>A	p.Arg280Gln	rs751963914	Missense	VUS	PM1 + PM2 + PP3	1
Candidato	PREX1	NM_020820.4	c.4943A>C	p.Tyr1648Ser	No reportada	Missense	VUS	PM2 + PP3	2
Candidato	PTPN1	NM_002827.4	c.707A>G	p.Asp236Gly	rs572308077	Missense	VUS	PM1 + PM2	1
Candidato	RHPN2	NM_033103.5	c.652A>T	p.Ser218Cys	rs577697788	Missense	VUS	PM1 + PM2	1
Candidato	RHPN2	NM_033103.5	c.347G>T	p.Gly116Val	No reportada	Missense	VUS	PM1 + PM2	1
Candidato	RHPN2	NM_033103.5	c.1225+5G>A	-	rs765192413	Splice	VUS	PM2	1
Candidato	RTEL1	NM_001283009.2	c.860C>T	p.Thr287Ile	rs561054389	Missense	VUS	PM1 + PM2	2
Candidato	RTEL1	NM_001283009.2	c.3282G>T	p.Lys1094Asn	rs778086975	Missense	VUS	PM2	1
Candidato	RTEL1	NM_001283009.2	c.3566C>T	p.Pro1189Leu	No reportada	Missense	VUS	PM2	1

Candidato	<i>SBF2</i>	NM_030962.4	c.5192C>T	p.Thr1731Met	rs142041264	Missense	VUS	PM2	1
Candidato	<i>SBF2</i>	NM_030962.4	c.4675T>G	p.Leu1559Val	rs751615833	Missense	VUS	PM1 + PM2	1
Candidato	<i>SBF2</i>	NM_030962.4	c.4592G>A	p.Gly1531Glu	rs1423180682	Missense	VUS	PM1 + PM2	1
Candidato	<i>SFMBT1</i>	NM_016329.4	c.317C>T	p.Pro106Leu	No reportada	Missense	VUS	PM2 + PP3 + BP1	1
Candidato	<i>SH2B3</i>	NM_005475.3	c.232G>A	p.Glu78Lys	rs754838420	Missense	VUS	PM2	1
Candidato	<i>SLC15A4</i>	NM_145648.4	c.964G>T	p.Val322Phe	rs149837040	Missense	VUS	PM2	1
Candidato	<i>SLC15A4</i>	NM_145648.4	c.727G>A	p.Gly243Ser	rs35580712	Missense	VUS	PM2 + BS2	2
Candidato	<i>SLC6A18</i>	NM_182632.3	c.131T>C	p.Phe44Ser	No reportada	Missense	VUS	PM2 + PP3	1
Candidato	<i>SLC6A18</i>	NM_182632.3	c.1486G>A	p.Gly496Arg	rs147278493	Missense	VUS	PM1 + PM2 + PP3 + BS2	1
Candidato	<i>SMAD6</i>	NM_005585.5	c.802C>T	p.Arg268Trp	rs780080560	Missense	VUS	PM1 + PM2 + PP3	1
Candidato	<i>SMAD6</i>	NM_005585.5	c.973G>A	p.Ala325Thr	rs199822239	Missense	VUS	PM2	1
Candidato	<i>SMAD7</i>	NM_005904.4	c.234_251del	p.Ala80_Alal85del	rs556829501	LOF	VUS	PM1 + PM2 + BS2 + BP3	1
Candidato	<i>SMAD9</i>	NM_001127217.3	c.17C>T	p.Pro6Leu	rs764161648	Missense	VUS	PM2	1
Candidato	<i>TCF7L2</i>	NM_030756.5	c.1468T>C	p.Ser490Pro	rs754884239	Missense	VUS	PM2	18
Candidato	<i>TMBIM1</i>	NM_022152.6	c.575T>C	p.Ile192Thr	No reportada	Missense	VUS	PM2	1
Candidato	<i>TOX2</i>	NM_001098797.2	c.1073C>T	p.Pro358Leu	rs374248829	Missense	VUS	PM2	2
Candidato	<i>TP53BP1</i>	NM_005657.4	c.5482C>T	p.His1828Tyr	rs780546857	Missense	VUS	PM1 + PM2 + PP3 + BP1	1
Candidato	<i>TTC22</i>	NM_001114108.2	c.604A>G	p.Met202Val	rs1258636323	Missense	VUS	PM2	1
Candidato	<i>WNT4</i>	NM_030761.5	c.877G>A	p.Gly293Ser	rs202139199	Missense	VUS	PM2 + PP3	1
Candidato	<i>ZAP70</i>	NM_001079.4	c.202A>T	p.Thr68Ser	rs138628302	Missense	VUS	PM1 + PM2	1
Candidato	<i>ZAP70</i>	NM_001079.4	c.1025T>C	p.Ile342Thr	rs142702703	Missense	VUS	PM1 + PM2	2

Abreviaturas: LOF, Pérdida de función, VUS: Variante de significado incierto.

3.2.2. Variantes moleculares identificadas con el modelo BoostDM

El modelo BoostDM, desarrollado por Muiños et al. en 2021, es una metodología basada en el aprendizaje automático diseñada para evaluar el potencial oncogénico de las mutaciones. La aplicación de este modelo facilitó la identificación de 68 variantes en 43 genes clasificadas como *oncdrivers* (Tabla 6), las cuales se encontraron en 65 de los pacientes analizados, lo que resultó en una tasa de detección del 65%. Se identificó, que el 27% de los individuos portaban más de una mutación (Figura 2C). De las variantes identificadas a través de este análisis, el 72% estaban en genes de paneles de diagnóstico de cáncer, mientras que el 28% restante estaban en genes candidatos. Adicionalmente, el 18.3% (8/43) de los genes con variantes *oncdriver* están relacionados con el CCR sindrómico. El 20.5% (14/68) de las variantes del estudio se identificaron en estos genes. La penetrancia se ha definido claramente para 18 de los genes en los que se identificaron mutaciones *oncdriver*, con 14 de alta penetrancia y 4 de penetrancia moderada (Tabla 6, Figura 2D). En cuanto a los tipos de variantes, el 82.3% (56/68) eran mutaciones *missense*, el 13.2% (9/68) eran mutaciones de pérdida de función (LoF), el 1.5% (1/68) eran mutaciones en sitios canónicos de splicing y el 3% (2/68) eran variantes sinónimas. Los genes con mayor número de variantes *oncdriver* fueron *MSH6* (n=5) y *MLH1*, *WRN*, *FANCC*, *KDR* y *TCF7L2* (cada uno con 3 variantes) (Tabla 6). Finalmente, cinco de las variantes *oncdriver* eran nuevas y no se habían reportado previamente en bases de datos públicas (*APC*: c.3663_3665delTTC, *MLH1*: c.1039delA, *RAD51C*: c.659T>C, *CTNNB1*: c.991T>C y *KDR*: c.2555A>T). Todas estas variantes se confirmaron mediante secuenciación de Sanger.

Tabla 6. Variantes *oncodriver* germinales identificadas en pacientes con CCR – Filtro B (BoostDM)

Penetrancia	Panel	Gen	Transcrito	Variante	Proteína	rs	Tipo	No. pacientes
Alta	Diagnóstico	<i>APC</i>	NM_000038.6	c.3663_3665delTTC	p.Ser1223del	No reportada	LOF	1
Alta	Diagnóstico	<i>BMPR1A</i>	NM_004329.3	c.176T>A	p.Leu59Ter	rs1564714834	LOF	1
Alta	Diagnóstico	<i>BRCA2</i>	NM_000059.4	c.8851G>A	p.Ala2951Thr	rs11571769	Missense	11
Alta	Diagnóstico	<i>CDH1</i>	NM_004360.5	c.1710T>C	p.Asn570=	rs202115589	Sinónima	1
Alta	Diagnóstico	<i>MLH1</i>	NM_000249.4	c.1039delA	p.Thr347Leufs*20	No reportada	LOF	1
Alta	Diagnóstico	<i>MLH1</i>	NM_000249.4	c.1514G>A	p.Ser505Asn	rs771044689	Missense	1
Alta	Diagnóstico	<i>MLH1</i>	NM_000249.4	c.1852A>G	p.Lys618Glu	rs35001569	Missense	1
Alta	Diagnóstico	<i>MSH2</i>	NM_000251.3	c.128A>G	p.Tyr43Cys	rs17217723	Missense	1
Alta	Diagnóstico	<i>MSH3</i>	NM_002439.5	c.1622C>T	p.Thr541Ile	rs1178614269	Missense	1
Alta	Diagnóstico	<i>MSH6</i>	NM_000179.3	c.3516_3517delAG	p.Arg1172Serfs*4	rs398123232	LOF	1
Alta	Diagnóstico	<i>MSH6</i>	NM_000179.3	c.1403G>A	p.Arg468His	rs41295268	Missense	1
Alta	Diagnóstico	<i>MSH6</i>	NM_000179.3	c.2041C>T	p.Leu681Phe	rs1553413412	Missense	1
Alta	Diagnóstico	<i>MSH6</i>	NM_000179.3	c.3961A>G	p.Arg1321Gly	rs41295278	Missense	1
Alta	Diagnóstico	<i>MSH6</i>	NM_000179.3	c.3245C>T	p.Pro1082Leu	rs191109849	Missense	1
Alta	Diagnóstico	<i>NTHL1</i>	NM_002528.7	c.244C>T	p.Gln82Ter	rs150766139	LOF	1
Alta	Diagnóstico	<i>PALB2</i>	NM_024675.4	c.2794G>A	p.Val932Met	rs45624036	Missense	1
Alta	Diagnóstico	<i>PMS2</i>	NM_000535.7	c.2395C>T	p.Arg799Trp	rs149202766	Missense	2
Alta	Diagnóstico	<i>POLD1</i>	NM_002691.4	c.2052G>C	p.Gln684His	rs1444143245	Missense	1
Alta	Diagnóstico	<i>POLD1</i>	NM_002691.4	c.2017G>A	p.Glu673Lys	rs61751955	Missense	1
Alta	Diagnóstico	<i>POLE</i>	NM_006231.4	c.686A>C	p.His229Pro	rs1482513360	Missense	1
Alta	Diagnóstico	<i>POLE</i>	NM_006231.4	c.6494G>A	p.Arg2165His	rs5745068	Missense	1
Alta	Diagnóstico	<i>PTCH1</i>	NM_000264.5	c.3241G>A	p.Val1081Met	rs587778629	Missense	1
Alta	Diagnóstico	<i>PTCH1</i>	NM_000264.5	c.1808G>A	p.Arg603His	rs199523893	Missense	1
Alta	Diagnóstico	<i>PTCH1</i>	NM_000264.5	c.3487G>A	p.Gly1163Ser	rs113663584	Missense	1
Moderada	Diagnóstico	<i>ATM</i>	NM_000051.4	c.6313A>G	p.Arg2105Gly	rs879253983	Missense	2
Moderada	Diagnóstico	<i>ATM</i>	NM_000051.4	c.7502A>G	p.Asn2501Ser	rs531617441	Missense	1
Moderada	Diagnóstico	<i>BARD1</i>	NM_000465.2	c.1075_1095del	p.Leu359_Pro365del	rs28997575	LOF	3
Moderada	Diagnóstico	<i>CHEK2</i>	NM_001005735.2	c.1556C>T	p.Thr519Met	rs142763740	Missense	1
Moderada	Diagnóstico	<i>RAD51C</i>	NM_058216.3	c.659T>C	p.Leu220Pro	No reportada	Missense	1
Moderada	Diagnóstico	<i>RAD51C</i>	NM_058216.3	c.859A>G	p.Thr287Ala	rs28363317	Missense	3
Menos definida	Diagnóstico	<i>BLM</i>	NM_000057.4	c.2452C>T	p.Arg818Cys	rs1279814185	Missense	1
Menos definida	Diagnóstico	<i>BLM</i>	NM_000057.4	c.2594A>G	p.Tyr865Cys	rs777842626	Missense	1
Menos definida	Diagnóstico	<i>BRIP1</i>	NM_032043.3	c.517C>T	p.Arg173Cys	rs4988345	Missense	2
Menos definida	Diagnóstico	<i>EGFR</i>	NM_005228.5	c.2380C>A	p.Pro794Thr	rs370289230	Missense	1

Menos definida	Diagnóstico	<i>FH</i>	NM_000143.4	c.535G>A	p.Val179Ile	rs1553341588	<i>Missense</i>	1
Menos definida	Diagnóstico	<i>FLCN</i>	NM_144997.7	c.1285delC	p.His429Thrfs*39	rs80338682	LOF	1
Menos definida	Diagnóstico	<i>FLCN</i>	NM_144997.7	c.535C>T	p.Arg179Trp	rs774358971	<i>Missense</i>	1
Menos definida	Diagnóstico	<i>GPC3</i>	NM_004484.4	c.1285G>A	p.Val429Met	rs11539789	<i>Missense</i>	3
Menos definida	Diagnóstico	<i>NBN</i>	NM_002485.5	c.643C>T	p.Arg215Trp	rs34767364	<i>Missense</i>	1
Menos definida	Diagnóstico	<i>NF1</i>	NM_000267.3	c.3498C>T	p.Gly1166=	rs2066733	Sinónima	1
Menos definida	Diagnóstico	<i>PDGFRA</i>	NM_006206.6	c.2282T>G	p.Leu761Arg	rs148654387	<i>Missense</i>	4
Menos definida	Diagnóstico	<i>PHOX2B</i>	NM_003924.4	c.738_776del	p.Ala248_Ala260del	rs757020181	<i>Missense</i>	1
Menos definida	Diagnóstico	<i>PMS1</i>	NM_000534.5	c.329C>G	p.Thr110Arg	rs372752293	<i>Missense</i>	1
Menos definida	Diagnóstico	<i>RAD50</i>	NM_005732.4	c.113A>G	p.Asn38Ser	rs750480943	<i>Missense</i>	1
Menos definida	Diagnóstico	<i>RAD50</i>	NM_005732.4	c.572C>T	p.Thr191Ile	rs2230017	<i>Missense</i>	1
Menos definida	Diagnóstico	<i>TSC2</i>	NM_000548.5	c.2161G>A	p.Val721Met	rs768261185	<i>Missense</i>	1
Menos definida	Diagnóstico	<i>WRN</i>	NM_000553.6	c.1717A>G	p.Thr573Ala	rs150148567	<i>Missense</i>	1
Menos definida	Diagnóstico	<i>WRN</i>	NM_000553.6	c.1192A>G	p.Thr398Ala	rs532554615	<i>Missense</i>	1
Menos definida	Diagnóstico	<i>WRN</i>	NM_000553.6	c.2500C>T	p.Arg834Cys	rs3087425	<i>Missense</i>	3
Menos definida	Candidato	<i>CRTC3</i>	NM_022769.5	c.1789A>C	p.Met597Leu	rs777546566	<i>Missense</i>	1
Menos definida	Candidato	<i>CTNNB1</i>	NM_001904.4	c.991T>C	p.Tyr331His	No reportada	<i>Missense</i>	1
Menos definida	Candidato	<i>ERCC1</i>	NM_202001.3	c.702+1G>A	-	rs747911302	Splice	1
Menos definida	Candidato	<i>ERCC1</i>	NM_202001.3	c.875G>A	p.Trp292Ter	rs116640350	LOF	1
Menos definida	Candidato	<i>FANCC</i>	NM_000136.3	c.851C>T	p.Ala284Val	rs201281511	<i>Missense</i>	2
Menos definida	Candidato	<i>FANCC</i>	NM_000136.3	c.77C>T	p.Ser26Phe	rs1800361	<i>Missense</i>	3
Menos definida	Candidato	<i>FANCC</i>	NM_000136.3	c.584A>T	p.Asp195Val	rs1800365	<i>Missense</i>	1
Menos definida	Candidato	<i>FUT2</i>	NM_000511.6	c.412C>T	p.Arg138Cys	rs1800022	<i>Missense</i>	2
Menos definida	Candidato	<i>FUT2</i>	NM_000511.6	c.812del	p.Pro271Leufs*16	rs1799761	LOF	1
Menos definida	Candidato	<i>KDR</i>	NM_002253.4	c.2555A>T	p.Asp852Val	No reportada	<i>Missense</i>	1
Menos definida	Candidato	<i>KDR</i>	NM_002253.4	c.2837G>A	p.Arg946His	rs140041720	<i>Missense</i>	1
Menos definida	Candidato	<i>KDR</i>	NM_002253.4	c.2312C>T	p.Thr771Met	rs149745504	<i>Missense</i>	1
Menos definida	Candidato	<i>MRE11</i>	NM_005591.4	c.2048G>A	p.Gly683Glu	rs764705257	<i>Missense</i>	1
Menos definida	Candidato	<i>SH2B3</i>	NM_005475.3	c.232G>A	p.Glu78Lys	rs754838420	<i>Missense</i>	1
Menos definida	Candidato	<i>TBX3</i>	NM_005996.4	c.613C>T	p.His205Tyr	rs749487839	<i>Missense</i>	1
Menos definida	Candidato	<i>TCF7L2</i>	NM_030756.5	c.1466C>G	p.Pro489Arg	rs77673441	<i>Missense</i>	2
Menos definida	Candidato	<i>TCF7L2</i>	NM_030756.5	c.1466C>A	p.Pro489His	rs77673441	<i>Missense</i>	5
Menos definida	Candidato	<i>TCF7L2</i>	NM_030756.5	c.1472C>T	p.Pro491Leu	rs573425555	<i>Missense</i>	1
Menos definida	Candidato	<i>TP53BP1</i>	NM_005657.4	c.5482C>T	p.His1828Tyr	rs780546857	<i>Missense</i>	1

Abreviaturas: LOF, Pérdida de función.

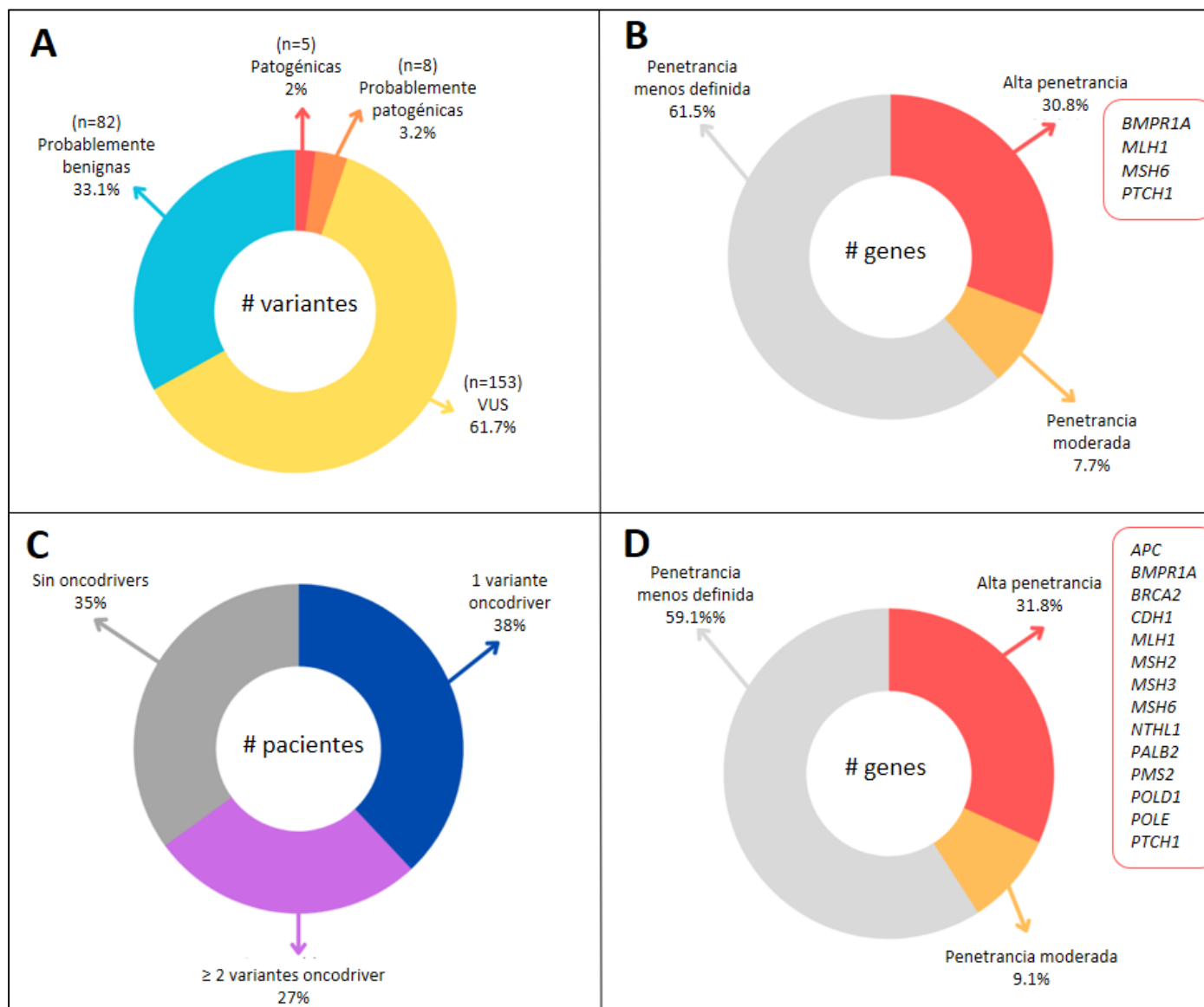


Figura 2. A) Variantes clasificadas según recomendaciones de la ACMG/AMP. B) Penetrancia de los genes con variantes P/PP (Filtro A). C) Pacientes con 1 o >1 variante oncdriver clasificadas con el modelo BoostDM. D) Penetrancia de los genes con variantes oncdriver (Filtro B).

La comparación entre las frecuencias alélicas de las variantes P/PP determinadas por el Filtro A y las variantes *oncdriver* identificadas por el Filtro B, mostró diferencias estadísticamente significativas ($p < 0,05$) en comparación con las descritas para la población global y latinoamericana, según lo establecido en la base de datos de gnomAD (Tabla 7). De las variantes P/PP clasificadas según los criterios del ACMG/AMP, el 46% exhibió frecuencias alélicas superiores a las descritas en la población global, mientras que solo el 23% de ellas diferían de la población latinoamericana. Este hallazgo es relevante ya que demuestra perfiles genómicos distintos y específicos de nuestra población, con mayores frecuencias de variantes P/PP asociadas con el CCR.

De manera similar, el 42,6% de las variantes identificadas como *oncdriver* según el modelo BoostDM mostraron diferencias significativas en las frecuencias alélicas globales, mientras que el 70,6% de los datos poblacionales de nuestros pacientes fueron similares con los de la población latinoamericana (Tabla 7).

Tabla 7. Comparación de frecuencias alélicas de variantes patogénicas/probablemente patogénicas y *oncdrivers*.

Filtro A – Clasificación ACMG/AMP									
Gen	Variante	rs	Frecuencias alélicas ^a	Frecuencias alélicas globales de gnomAD	p-valor	p-valor con corrección de Yates	Frecuencias alélicas latinas de gnomAD	p-valor	p-valor con corrección de Yates
<i>BARD1</i>	c.2229dupT	rs1259296823	0,005	3,97902E-06	< 0.00001	< 0.00001*	2,89218E-05	< 0.00001	< 0.00001*
<i>BMPR1A</i>	c.176T>A	rs1564714834	0,005	NR	NA	NA	NR	NA	NA
<i>ERCC1</i>	c.702+1G>A	rs747911302	0,005	3,97763E-06	< 0.00001	< 0.00001*	NR	NA	NA
<i>EXO1</i>	c.1465delA	No reportada	0,005	NR	NA	NA	NR	NA	NA
<i>FLCN</i>	c.1285delC	rs80338682	0,005	1,22287E-05	< 0.00001	< 0.00001*	NR	NA	NA
<i>IL12RB1</i>	c.1791+2T>G	rs554063682	0,005	0,000102081	< 0.00001	< 0.00001*	0,000120034	< 0.00001	< 0.00001*
<i>MLH1</i>	c.1039delA	No reportada	0,005	NR	NA	NA	NR	NA	NA
<i>MSH6</i>	c.3516_3517delAG	rs398123232	0,005	NR	NA	NA	NR	NA	NA
<i>NOTCH3</i>	c.1345C>T	rs762734007	0,005	NR	NA	NA	NR	NA	NA
<i>NTHL1</i>	c.244C>T	rs150766139	0,005	0,001417324	0,178	0,684	0,000705577	0,0248	0,3523
<i>OGG1</i>	c.137G>A	rs104893751	0,005	0,002165888	0,388476	0,919451	0,001864091	0,31	0,84
<i>PTCH1</i>	c.3241G>A	rs587778629	0,005	9,15182E-05	< 0.00001	< 0.00001*	0,000260191	0,00008	0,06
<i>SMAD9</i>	c.781+2T>A	rs770716081	0,005	7,98433E-05	< 0.00001	< 0.00001*	3,91942E-05	< 0.00001	< 0.00001*
Filtro B - Clasificación BoostDM									
Gen	Variante	rs	Frecuencias alélicas ^b	Frecuencias alélicas globales de gnomAD	p-valor	p-valor con corrección de Yates	Frecuencias alélicas latinas de gnomAD	p-valor	p-valor con corrección de Yates
<i>APC</i>	c.3663_3665delTTC	No reportada	0,005	NR	NA	NA	NR	NA	NA
<i>ATM</i>	c.6313A>G	rs879253983	0,010	0,0000	< 0.00001	< 0.00001*	NR	NA	NA
<i>ATM</i>	c.7502A>G	rs531617441	0,005	0,0002	< 0.00001	< 0.00001*	0,0001	< 0.00001	< 0.00001*
<i>BARD1</i>	c.1075_1095del	rs28997575	0,015	0,0283	0,2665	0,3722	0,033728814	0,143	0,206
<i>BLM</i>	c.2452C>T	rs1279814185	0,005	0,0000	< 0.00001	< 0.00001*	0,0000	< 0.00001	< 0.00001*
<i>BLM</i>	c.2594A>G	rs777842626	0,010	0,0000	< 0.00001	< 0.00001*	NR	NA	NA
<i>BMPR1A</i>	c.176T>A	rs1564714834	0,005	NR	NA	NA	NR	NA	NA
<i>BRCA2</i>	c.8851G>A	rs11571769	0,060	0,0089	< 0.00001	< 0.00001*	0,0378	0,1689	0,2363
<i>BRIP1</i>	c.517C>T	rs4988345	0,010	0,0026	0,5140	0,9698	0,0008	0,0468	0,4327
<i>CDH1</i>	c.1710T>C	rs202115589	0,005	0,0000	< 0.00001	< 0.00001*	0,0001	< 0.00001	< 0.00001*
<i>CHEK2</i>	c.1556C>T	rs142763740	0,005	0,0003	0,0002	0,0824	0,0002	< 0.00001	< 0.00001*
<i>CRTC3</i>	c.1789A>C	rs777546566	0,005	0,0000	< 0.00001	< 0.00001*	0,0000	< 0.00001	< 0.00001*
<i>CTNNB1</i>	c.991T>C	No reportada	0,005	NR	NA	NA	NR	NA	NA
<i>EGFR</i>	c.2380C>A	rs370289230	0,005	NR	NA	NA	NR	NA	NA
<i>ERCC1</i>	c.702+1G>A	rs747911302	0,005	3,97763E-06	< 0.00001	< 0.00001*	NR	NA	NA
<i>ERCC1</i>	c.875G>A	rs116640350	0,005	0,0026	0,4989	0,8245	0,0021	0,0000	0,0507
<i>FANCC</i>	c.851C>T	rs201281511	0,010	0,0000	< 0.00001	< 0.00001*	0,0000	< 0.00001	< 0.00001*
<i>FANCC</i>	c.77C>T	rs1800361	0,005	0,0049	0,9774	0,6302	0,0041	0,3563	0,8890
<i>FANCC</i>	c.584A>T	rs1800365	0,005	0,0030	0,6012	0,8994	0,0027	0,5206	0,9633
<i>FH</i>	c.535G>A	rs1553341588	0,005	NR	NA	NA	NR	NA	NA

FLCN	c.535C>T	rs774358971	0,005	0,0001	< 0.00001	< 0.00001*	0,0001	< 0.00001	< 0.00001*
FLCN	c.1285delC	rs80338682	0,005	1,22287E-05	< 0.00001	< 0.00001*	NR	NA	NA
FUT2	c.412C>T	rs1800022	0,010	0,0061	0,8472	0,7912	0,0049	0,9855	0,6239
FUT2	c.812delC	rs1799761	0,005	0,0015	0,1953	0,7077	0,0007	0,0248	0,3522
GPC3	c.1285G>A	rs11539789	0,020	0,0059	0,0887	0,2160	0,0204	0,6015	0,7896
KDR	c.2555A>T	No reportada	0,005	NR	NA	NA	NR	NA	NA
KDR	c.2837G>A	rs140041720	0,005	0,0002	0,0000	0,0410*	0,0002	< 0.00001	< 0.00001*
KDR	c.2312C>T	rs149745504	0,005	0,0012	0,1282	0,6080	0,0021	0,3780	0,9108
MLH1	c.1039delA	No reportada	0,005	NR	NA	NA	NR	NA	NA
MLH1	c.1514G>A	rs771044689	0,005	0,0000	< 0.00001	< 0.00001*	0,0001	< 0.00001	< 0.00001*
MLH1	c.1852A>G	rs35001569	0,005	0,0034	0,7020	0,822	0,0045	0,8984	0,68
MRE11	c.2048G>A	rs764705257	0,005	0,0000	< 0.00001	< 0.00001*	0,0002	< 0.00001	< 0.00001*
MSH2	c.128A>G	rs17217723	0,005	0,0001	< 0.00001	< 0.00001*	NR	NA	NA
MSH3	c.1622C>T	rs1178614269	0,005	NR	NA	NA	NR	NA	NA
MSH6	c.3516_3517delAG	rs398123232	0,005	NR	NA	NA	NR	NA	NA
MSH6	c.1403G>A	rs41295268	0,005	0,0000	< 0.00001	< 0.00001*	0,0001	< 0.00001	< 0.00001*
MSH6	c.2041C>T	rs1553413412	0,005	NR	NA	NA	NR	NA	NA
MSH6	c.3961A>G	rs41295278	0,005	0,0001	< 0.00001	< 0.00001*	0,0001	< 0.00001	< 0.00001*
MSH6	c.3245C>T	rs191109849	0,005	0,0003	0,0002	0,0811	0,0014	0,1881	0,6980
NBN	c.643C>T	rs34767364	0,005	0,0025	0,4630	0,9860	0,0016	0,2362	0,7600
NF1	c.3498C>T	rs2066733	0,005	0,0012	0,1220	0,5988	0,0006	0,0123	0,2800
NTHL1	c.244C>T	rs150766139	0,005	0,001417324	0,178	0,684	0,000705577	0,0248	0,3523
PALB2	c.2794G>A	rs45624036	0,005	0,0051	0,9800	0,6360	0,0020	0,3466	0,8799
PDGFRA	c.2282T>G	rs148654387	0,020	0,0021	< 0.00001	< 0.00001*	0,0159	0,6200	0,8350
PHOX2B	c.738_776del	rs757020181	0,010	0,0020	0,2580	0,7860	0,0366	0,0410	0,0690
PMS1	c.329C>G	rs372752293	0,010	0,0001	< 0.00001	< 0.00001*	0,0002	0,0000	0,04811*
PMS2	c.2395C>T	rs149202766	0,010	0,0005	0,0043	0,2040	0,0031	0,6230	0,8807
POLD1	c.2052G>C	rs144143245	0,005	0,0004	0,0021	0,1648	0,0003	0,0006	0,1129
POLD1	c.2017G>A	rs61751955	0,005	0,0001	< 0.00001	< 0.00001*	0,0002	0,0000	0,04488*
POLE	c.686A>C	rs1482513360	0,005	0,0000	< 0.00001	< 0.00001*	0,0000	< 0.00001	< 0.00001*
POLE	c.6494G>A	rs64945068	0,005	0,0000	< 0.00001	< 0.00001*	0,0000	< 0.00001	< 0.00001*
PTCH1	c.1808G>A	rs199523893	0,005	0,0001	< 0.00001	< 0.00001*	0,0004	0,0010	0,1310
PTCH1	c.3241G>A	rs587778629	0,005	9,15182E-05	< 0.00001	< 0.00001*	0,000260191	0,00008	0,06
PTCH1	c.3487G>A	rs113663584	0,005	0,0005	0,0052	0,2160	0,0006	0,0099	0,2640
RAD50	c.113A>G	rs750480943	0,005	0,0000	< 0.00001	< 0.00001*	NR	NA	NA
RAD50	c.572C>T	rs2230017	0,005	0,0018	0,2700	0,7980	0,0008	0,0330	0,3860
RAD51C	c.659T>C	No reportada	0,005	NR	NA	NA	NR	NA	NA
RAD51C	c.859A>G	rs28363317	0,015	0,0056	0,0750	0,1919	0,0059	0,0900	0,2190
SH2B3	c.232G>A	rs754838420	0,005	0,0011	0,0910	0,5400	0,0001	< 0.00001	< 0.00001*
TBX3	c.613C>T	rs749487839	0,005	NR	NA	NA	NR	NA	NA
TCF7L2	c.1472C>T	rs573425555	0,005	0,0000	< 0.00001	< 0.00001*	0,0001	< 0.00001	< 0.00001*
TCF7L2	c.1466C>G	rs77673441	0,010	0,0044	0,2310	0,5080	0,0032	0,0890	0,2800
TCF7L2	c.1466C>A	rs77673441	0,030	0,0034	< 0.00001	< 0.00001*	0,0219	0,7300	0,9266
TP53BP1	c.5482C>T	rs780546857	0,005	0,0000	< 0.00001	< 0.00001*	NR	NA	NA
TSC2	c.2161G>A	rs768261185	0,005	0,0000	< 0.00001	< 0.00001*	0,0029	< 0.00001	< 0.00001*
WRN	c.1717A>G	rs150148567	0,005	0,0014	0,1700	0,6790	0,0877	0,0525	0,4490
WRN	c.1192A>G	rs532554615	0,005	0,0000	< 0.00001	< 0.00001*	0,0058	< 0.00001	< 0.00001*
WRN	c.2500C>T	rs3087425	0,015	0,0026	0,0005	0,0054*	1,9674	0,6420	0,8374

a Variantes del filtro A del presente estudio

b Variantes del filtro B del presente estudio

* Estadísticamente significativo

NA: No aplica

NR: No reportado

3.3. Ensayo *in vitro* en variantes de sitios de splicing

Se realizó la validación funcional de tres mutaciones de splicing identificadas en los genes *ERCC1* c.702+1G>A, *SMAD9* c.781+2T>A e *IL12RB1* c.1791+2T>G, ya que corresponden a variantes en genes que generalmente no se consideran en paneles de diagnóstico, por lo que representan factores genéticos innovadores relacionados con el CCR.

La verificación de la formación de los transcritos potencialmente aberrantes, que fue confirmado mediante RT-PCR posterior a la transfección en los modelos celulares HCT-116 y HEK-293, determinó que la mutación *SMAD9* c.781+2T>A conduce a la omisión del exón 3 con la pérdida de 111 pb. La versión mutante no genera un transcrito normal como el obtenido en la forma *Wild-type*, lo que resulta en una pérdida de 37 aminoácidos en la proteína. Este hallazgo se observó en las tres réplicas y en ambas líneas celulares analizadas ($p < 0.05$). La omisión del exón se corroboró mediante secuenciación de Sanger del ADN plasmídico obtenido del proceso de clonación (Figura 3A).

Por otro lado, la mutación *IL12RB1* c.1791+2T>G se asoció con la formación de dos transcritos aberrantes, correspondientes a la retención del intrón 14 de *IL12RB1*, introduciendo 98 pb, y a la omisión del exón 15 del gen (76 pb). Las consecuencias predichas en la proteína indican que estos transcritos conducen al corrimiento en el marco de lectura, p.Thr598Glyfs*21 y p.A573Lfs*21, respectivamente (Figura 3C). Aunque estos transcritos aberrantes también se observaron en la versión WT, se redujeron significativamente en relación con la versión mutante ($p < 0.05$). Estos hallazgos se replicaron en ambos modelos celulares analizados y en las tres condiciones experimentales. Los transcritos obtenidos se verificaron mediante secuenciación de Sanger del ADN plasmídico obtenido de la clonación utilizando el Kit de Clonación TOPO® TA para Secuenciación (Invitrogen) (Figura 3C).

Por último, el análisis funcional de la variante *ERCC1* c.702+1G>A reveló la generación de un transcrito aberrante con una pérdida de 100 pb, correspondiente a la omisión del exón 6 del gen, lo que resulta en una mutación por cambio del marco de lectura (p.Ser201Argfs*1) (Figura 3B). La versión WT genera dos transcritos, uno de los cuales es normal, y el otro corresponde a la omisión del exón. A pesar de que se evidenció diferencia estadísticamente significativa entre los transcritos formados en las versiones WT y mutante en la línea HCT-116 ($p < 0.05$), lo más relevante de este resultado fue la proporción del transcrito de la omisión del exón, tanto en su versión WT como MUT (Figura 3B).

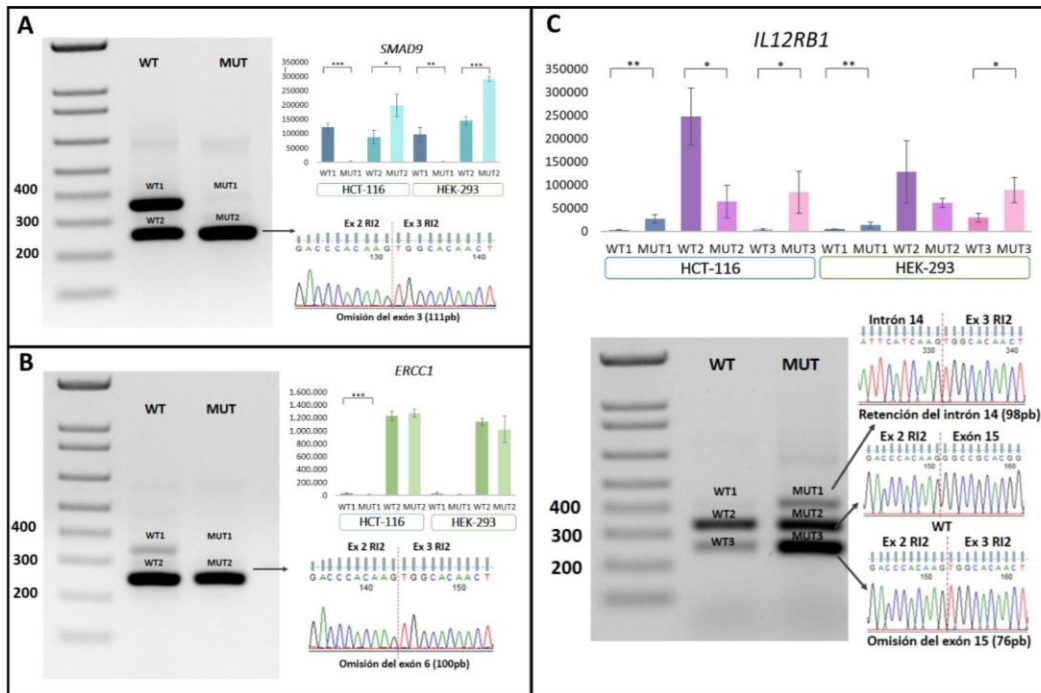


Figura 3. Ensayo funcional en variantes P/PP intrónicas en sitios canónicos de splicing. **A)** Análisis funcional de *SMAD9*:c.781+2T>A en células HCT-116 y HEK-293 transfectadas (WT y MUT), con una omisión de exón confirmada mediante secuenciación de Sanger. **B)** Análisis funcional de *ERCC1*:c.702+1G>A en células HCT-116 y HEK-293 transfectadas (WT y MUT), con una omisión de exón confirmada mediante secuenciación de Sanger. **C)** Análisis funcional de *IL12RB1*:c.1791+2T>G en células HCT-116 y HEK-293 transfectadas (WT y MUT), mostrando una omisión del exón 15 y una retención de intrón 14, confirmados mediante secuenciación de Sanger. Abreviaturas: Ex, exón; MUT, mutante; RI2, exón insulina 2 de rata; WT, tipo salvaje; *Significancia estadística.

4. Discusión

Este estudio presenta, por primera vez, una descripción de la prevalencia y el espectro de las variantes germinales en una cohorte de pacientes colombianos no seleccionados con CCR. Nuestro enfoque implicó el uso de la secuenciación de próxima generación (NGS) para la identificación de variantes moleculares dentro de un panel integral de 206 genes asociados con el cáncer (Tabla 1). Este panel incluye genes que aparecen en paneles de diagnóstico utilizados rutinariamente en la práctica clínica, así como otros posibles genes candidatos identificados a través de una exhaustiva revisión de la literatura. El flujo de trabajo bioinformático implementado permitió la identificación de variantes moleculares clasificadas de acuerdo con los criterios establecidos por ACMG/AMP y el modelo de inteligencia artificial BoostDM (Figura 1).

Nuestra cohorte estuvo compuesta por pacientes no seleccionados, siendo solo el 12% diagnosticado con CCR antes de los 50 años. Esto contrasta con la mayoría de los estudios sobre mutaciones germinales, que suelen centrarse en individuos con cáncer de inicio temprano o sospechosos de tener síndromes de predisposición hereditaria relacionados con el CCR (Djursby et al., 2020; T. J. Jiang et al., 2020; Olkinuora et al., 2021; Poliani et al., 2022; J. Zhang et al., 2017). En estos casos, la susceptibilidad al CCR suele atribuirse principalmente a las variantes germinales, lo

que es especialmente evidente en casos de CCR hereditario o sindrómico, donde el patrón de herencia generalmente es autosómico dominante. Además de la edad de inicio del CCR, nuestro estudio demostró, a diferencia de los hallazgos previamente reportados, una mayor proporción de mujeres afectadas (55%). Esto difiere de la tendencia observada en la mayoría de los estudios en los que los hombres tienden a presentar una mayor incidencia de CCR (Gong et al., 2019; T. J. Jiang et al., 2020; Pearlman et al., 2017; P. L. S. Uson et al., 2022). Nuestro hallazgo podría explicarse por el tipo de muestreo, el cual fue por conveniencia y durante este período más mujeres fueron atendidas en el hospital.

Conocer y comprender cómo funcionan los factores genéticos relacionados con el CCR es de gran importancia, ya que respalda la implementación de estrategias de vigilancia adecuadas, como se recomienda en las directrices internacionales, con el objetivo de reducir la incidencia y la mortalidad por CCR en portadores de variantes P/PP (Poliani et al., 2022). En este contexto, este trabajo propone el uso de NGS para el análisis de múltiples genes, lo que permite la identificación de VPs, tanto canónicas como no canónicas, asociadas con la enfermedad. Las VPs no canónicas en genes de predisposición al cáncer, aparentemente no relacionados con el CCR, han demostrado ser una fuente significativa para identificar la causa molecular en casos no seleccionados (Yurgelun et al., 2017). Además, se ha recomendado el análisis de poblaciones no seleccionadas, como lo hemos estudiado, para evitar subestimar la verdadera prevalencia de factores hereditarios en la población global de CCR (Poliani et al., 2022).

En cuanto a la clasificación de las variantes moleculares, aplicamos dos enfoques. En primer lugar, está el filtro A, que sigue las recomendaciones de ACMG/AMP (Hampel et al., 2015). Este enfoque se utiliza tanto en la investigación como en la rutina clínica de diagnóstico molecular y permitió la detección de 13 variantes P/PP en el 12% de nuestra cohorte (Tabla 4).

Nuestros hallazgos demostraron una tasa de detección más alta en comparación con otros estudios que analizan poblaciones no seleccionadas, como el descrito por Yurgelun et al., 2017, que informó una tasa del 10% (Yurgelun et al., 2017). Quizás el menor número de genes analizados por los autores (n = 29) explique esta diferencia y respalde la ventaja de utilizar paneles ampliados, que potencialmente mejoran el rendimiento diagnóstico (Poliani et al., 2022; Roht et al., 2022). Se ha demostrado que las tasas de detección de VPs en casos con sospecha de CCR familiar o en pacientes diagnosticados antes de los 50 años son sustancialmente más altas, con una tasa del 15,5% al 26,8% (Gong et al., 2019; Pearlman et al., 2017; Poliani et al., 2022; P. S. Uson et al., 2021). El 61,5% de estas variantes P/PP estaban asociadas con genes incluidos en paneles de diagnóstico, mientras que el 38,5% restante provenía de genes candidatos identificados en estudios. Esta proporción sustancial respalda la hipótesis que el análisis de genes actualmente no incluidos en paneles de diagnóstico rutinario de diferentes plataformas de secuenciación, puede aumentar la detección de variantes germinales P/PP relacionadas con el desarrollo del CCR.

En la población analizada, se evidenciaron 4 mutaciones de alta penetrancia en los genes *MLH1*, *MSH6*, *BMPR1A* y *PTCH1*. Este hallazgo es de relevancia ya que la identificación de mutaciones en algunos de estos genes, particularmente *MLH1* y *MSH6* permite la adopción de estrategias de tamizaje, tratamiento y seguimientos en estos pacientes considerados de alto riesgo según las

recomendaciones descritas en las guías NCCN (*National Comprehensive Cancer Network*) (<https://www.nccn.org/patients/guidelines/content/PDF/colorectal-screening-patient.pdf>). Las variantes germinales en los genes MMR (*mismatch repair*) (*MLH1*, *MSH2*, *MSH6* y *PMS2*) están relacionadas con el Síndrome de Lynch, uno de los síndromes de predisposición genética más común, que corresponde del 2-4% de todos los casos de CCR. Se ha determinado que en países de Latinoamérica y del Caribe hay mayor prevalencia de pacientes menores de 50 años afectados con CCR, por lo que se estima que los programas de tamizaje y la identificación de pacientes en alto riesgo que portan mutaciones patogénicas germinales, pueden reducir de manera más efectiva su incidencia (Vaccaro et al., 2019).

Estudios previos han demostrado que la mayoría de las variantes P/PP encontradas en pacientes con CCR se localizan en los genes MMR, cuyos productos se encargan de la reparación de errores de secuencia producidos durante la replicación del DNA. En este contexto, Gong et al., 2019 describió que del 19.1% de variantes P/PP identificados en su estudio, 11.3% eran de los genes MMR. De manera similar Yurgelun et al., 2017 evidenció una alta tasa de mutaciones en estos genes, y otros estudios como el desarrollado por Zhang et al., 2017 ratificaron estos hallazgos con cerca del 50% de variantes genéticas en genes MMR (Gong et al., 2019; Yurgelun et al., 2017; J. Zhang et al., 2017). A diferencia de estos hallazgos previos, en nuestra población se evidenciaron solo dos variantes en los genes MMR (*MLH1:c.1039del* y *MSH6:c.3516_3517delAG*) que correspondieron al 15.4% del total identificado. Este resultado, que demuestra una diferencia en la frecuencia de mutaciones en MMR para la población de estudio, puede ser debido al criterio de selección de la muestra, ya que los estudios con mayor tasa de mutaciones en genes MMR incluyen pacientes con antecedente familiar de CCR, con edad de diagnóstico menor de 50 años y con pólipos. Nuestro estudio analiza casos no seleccionados y por tanto es plausible encontrar mutaciones diferentes a las relacionadas con síndromes hereditarios.

Adicionalmente, mutaciones en otros genes relacionados con CCR sindrómico fueron evidenciadas en nuestro estudio. Una de ellas corresponde a la mutación *BMPRI1* (c.176T>A p. Leu59Ter) que genera un codón de parada prematuro que elimina 473 aminoácidos. Variantes en el gen *BMPRI1* se han asociado con el síndrome de poliposis juvenil, está relacionado con la vía de señalización AKT y es un receptor tipo 1 de la superfamilia TGFβ (Blatter et al., 2020; Howe et al., 2013; Liu et al., 2020; Waite & Eng, 2003). Además de su papel en la activación de las funciones epiteliales del colon, ejerce influencia sobre procesos celulares críticos, como el crecimiento, la diferenciación y la apoptosis (Cross et al., 2018; Howe et al., 2013; Massagué, 1996; Waite & Eng, 2003; Zhao et al., 2018). Adicionalmente, esta mutación se ha reportado previamente en un paciente con poliposis adenomatosa (Ngeow et al., 2013). En este contexto, la mutación identificada afecta la mayor parte de la proteína, incluyendo sus dominios funcionales de activina tipo I y II, así como el dominio Ser/Thr proteína quinasa. En consecuencia, se pierde la funcionalidad de la proteína, su interacción con otras proteínas e incluso su papel en las vías de señalización mencionadas anteriormente.

Finalmente, un paciente presentó una variante heterocigota en un gen que se ha descrito como de alta penetrancia y correlacionado con potencial metastásico de CCR (You et al., 2010). La variante identificada corresponde a un cambio *missense* en el gen *PTCH1* (c.3241G>A p.Val1081Met). *PTCH1*

tiene un rol fundamental en la vía de señalización *Hedgehog*, que está involucrada en la tumorigénesis, en la regulación de la proliferación, la angiogénesis y la renovación de células madre, eventos que se han asociado a procesos oncogénicos relacionados con CCR (Raleigh & Reiter, 2019). El hallazgo de esta variante es de gran relevancia ya que estudios previos han demostrado que *PTCH1* debe ser considerado como un potencial biomarcador que puede discriminar CCR con alto y bajo riesgo de metástasis, lo cual está relacionado inversamente con los niveles de expresión de la proteína *PTCH1* (You et al., 2010). Adicionalmente, esta variante se ha reportado en pacientes con síndrome hereditario de predisposición al cáncer (Chassaing et al., 2016) y contiene evidencia funcional de pérdida de función en la proteína (Chassaing et al., 2016), lo que respalda los criterios de patogenicidad sobre esta mutación.

Para otras variantes P/PP identificadas según las recomendaciones de ACMG/AMP no se tiene la definición exacta de la penetrancia, pero dada su patogenicidad y su clara implicación en vías de señalización moleculares relacionadas con la fisiopatología del CCR son de relevancia en su etiopatología molecular. El 46% de las variantes correspondieron a variantes de pérdida de función (LoF) y alteraron la función de los genes *FLCN*, *NOTCH3*, *NTHL1* y *BARD1*.

BARD1 codifica una proteína que interactúa con la región N-terminal de *BRCA1*. Se conoce que el complejo *BARD1/BRCA1* tiene un papel en la reparación del ADN, la recombinación y el control del ciclo celular (Bork et al., 1997; Gerloff et al., 2012; Q. Jiang & Greenberg, 2015; Leung & Glover, 2011; D. Wu et al., 2023). La variante *BARD1:c.2229dupT* genera un codón de parada prematuro, que afecta la longitud de la proteína (p.Asn744Ter), exactamente en el dominio BCRT 2, el cual apoya principalmente la interacción proteína-proteína, y se ha relacionado con la reparación, recombinación y control del ciclo celular, mecanismos involucrados en el desarrollo de neoplasias malignas (Billing et al., 2018; Peña-Guerrero et al., 2023). Por su parte, el gen *FLCN* está asociado al síndrome de Birt-Hogg-Dubé y aumenta el riesgo de CCR (Sattler et al., 2021). Recientemente, se han detectado variantes patogénicas en este gen, principalmente en pacientes con CCR de inicio temprano (Jirka et al., 2023; Nahorski et al., 2010; Schmidt & Linehan, 2018; van de Beek et al., 2020; Whitworth et al., 2016). La variante *FLCN:c.1285delC* genera un corrimiento en el marco de lectura de la proteína (p.His429Thrs*39), con pérdida del dominio DENN terminal, importante en las vías de señalización AMPK y mTOR (Ramirez Reyes et al., 2021). *NOTCH3* es un receptor de la superficie celular que desempeña un papel importante en la señalización que induce la proliferación, la polaridad/adhesión de las células epiteliales y la apoptosis (Huang et al., 2023; Xiu et al., 2021; Y.-Q. Zhang et al., 2021). El paciente porta la variante *NOTCH3:c.1345C>T* p.Arg449Cys, clasificada como PP. Investigaciones previas han establecido vínculos entre las variantes en este gen y la susceptibilidad al CCR, especialmente en hombres (Alanazi et al., 2021). El gen *NTHL1* está asociado con poliposis adenomatosa familiar y el síndrome de predisposición hereditaria al cáncer, de patrón de herencia autosómica recesiva. Actúa en la reparación de escisión de bases (BER), que es la vía de reparación principal para el daño oxidativo en el ADN (Bessho, 1999; Das et al., 2020; Rebuzzi et al., 2023; Williams & Parsons, 2018). La variante *NTHL1:c.244C>T* p.Gln82Ter induce la pérdida del dominio HhH, importante en la unión al ADN y por tanto en la función normal de la vía BER (Carroll et al., 2021).

De manera innovadora, nuestro estudio identificó que cerca del 40% de las variantes P/PP identificadas mediante los criterios de clasificación de la ACMG/AMP (Filtro A descrito en la metodología) se identificaron en genes no analizados en el ámbito del diagnóstico clínico-molecular. Nuestro hallazgo apoya la necesidad de implementar análisis moleculares que mediante NGS permitan el análisis simultáneo de cientos de genes potencialmente relacionados con CCR.

El 60% de dichas variantes se localizaron en sitios genómicos involucrados con el proceso de splicing y corresponden a las mutaciones *ERCC1* (c.702+1G>A), *IL12RB1* (c.1791+2T>G) y *SMAD9* (c.781+2T>A). Estos genes participan en las vías de señalización de reparación del daño celular y los procesos relacionados con la carcinogénesis (Dai et al., 2015; Ngeow et al., 2015; Ni et al., 2014), por lo que mutaciones en dichos genes pueden afectar estas vías de señalización y así aumentar la predisposición al CCR (Y.-K. Li et al., 2020; Pietrzyk et al., 2015).

Nuestro estudio identificó también mutaciones heterocigotas P/PP en los genes *EXO1* y *OGG1* que participan en procesos de reparación y recombinación (Bolderson et al., 2010; Keijzers et al., 2018; Reilly et al., 2019; Rubenstein et al., 2020; Y. Wu et al., 2001; Y. Zhang et al., 2011). Por lo tanto, si esos genes se ven afectados por mutaciones, se alteran los procesos de replicación y post-replicación, incluida la activación de puntos de control, lo que puede llevar a la inestabilidad genómica y al desarrollo del cáncer (Keijzers et al., 2018; Lai et al., 2016; Moreno et al., 2006).

En conjunto, nuestros hallazgos indican que las variantes germinales P/PP se concentran predominantemente en genes relacionados con las vías de reparación y control del ADN, como *BARD1*, *BMPR1A*, *ERCC1*, *EXO1*, *MLH1*, *MSH6*, *NTHL1* y *OGG1* (60% de las variantes). Estas vías son cruciales para mantener la estabilidad genómica (Wallace et al., 2012). En consecuencia, las variantes germinales en genes vinculados a estas vías de señalización pueden tener efectos perjudiciales que aumentan la predisposición al CCR.

Adicional a la búsqueda de mutaciones germinales patogénicas, mediante estrategias usadas en la rutina de diagnóstico molecular y que permite la identificación de pacientes en alto riesgo de sufrir CCR, usamos un método de inteligencia artificial para tal fin. BoostDM es una metodología basada en el aprendizaje automático para llevar a cabo mutagénesis *in silico* de genes asociados con el desarrollo del cáncer. Este enfoque innovador evalúa sistemáticamente todos los posibles cambios dentro de un gen o proteína para identificar factores que causan cáncer (Hamamoto et al., 2022; Muiños et al., 2021). Las variantes con efectos perjudiciales son categorizadas como "oncdriver". Identificamos un total de 68 variantes en el 65% de los pacientes, y 27% de ellos presentaron más de una variante *oncdriver*. El 72% de estas variantes se encontraban en los genes usados en la rutina clínica, mientras que el 28% restante, se encontraba en los demás genes analizados. Estos hallazgos, al igual que con el filtro anterior, demuestran el impacto en términos de rendimiento diagnóstico con la incorporación de métodos de secuenciación masiva como la NGS, en el análisis de mutaciones relacionadas con cáncer. Recientemente, Garg et al., 2020 indicaron que el uso de NGS incrementa la identificación de mutaciones en cerca del 36% en comparación con el análisis de gen único, lo que supone el claro beneficio del uso de estrategias de análisis molecular masivo (Garg et al., 2020).

Respecto al tipo de variante molecular predominante, el 82% de ellas corresponden a variantes *missense*, hallazgo que discrepa de nuestro resultado según el filtro A, en donde la mayoría de las mutaciones fueron LoF. Estimamos que este hallazgo está relacionado con la diferencia metodológica de los dos filtros, con mayor astringencia para el caso del filtro A, cuyos criterios han sido implementados para la descripción de variantes mendelianas y siguen criterios estrictos de MAF, estudios funcionales y análisis *in silico*, entre otros (Richards et al., 2015). Adicionalmente, la mayor representación de variantes *missense* con el análisis de BoostDM puede ser potencialmente explicado por la omisión en este algoritmo del uso de predictores *in silico* como REVEL (*Rare Exome Variant Ensemble Learner*), un método para predecir la patogenicidad de variantes raras que han ensamblado los puntajes de 13 predictores tales como MutPred, PolyPhen, SIFT, PROVEAN entre otros, haciéndolo altamente astringente (Ioannidis et al., 2016).

Del total de variantes clasificadas como *oncdriver* por BoostDM, se identificó una mayor proporción de cambios en genes relacionados con síndromes de predisposición al CCR, comparado al filtro A. De manera significativa, el 20.6% de las variantes moleculares se encontraron en estos genes, y fue posible su identificación en genes relacionados con el Síndrome de Lynch; *MSH6* (n=5) y *MLH1* (n=3). Como se mencionó previamente, estos hallazgos permiten asesorar a los pacientes y sus familiares siguiendo las guías internacionales, lo que potencialmente mejora la identificación temprana de personas en riesgo y el manejo de los portadores de dichas variantes (Benson et al., 2021). Adicionalmente, este modelo permitió la identificación de otras variantes en genes relacionados con CCR síndromico, tales como *APC*, *BMPR1A*, *MSH2*, *MSH3*, *NTHL1* y *PMS2*. Estos cambios moleculares y los observados en los genes relacionados con las vías de señalización de reparación del daño del ADN demuestran la capacidad de este algoritmo de inteligencia artificial en la detección de variantes *oncdriver* (Reilly et al., 2019; Rubenstein et al., 2020; Williams & Parsons, 2018). Variaciones en estos genes han sido reportadas en grandes cohortes de pacientes con asociaciones altamente significativas con el desarrollo de CCR, tal como es el caso de *MSH2* (OR: 18.1), *MLH1* (OR: 8.6) o *APC* (OR: 49.4), que apoya la importancia de los estudios de identificación de variantes germinales en población no seleccionada para contribuir a la generación de políticas de evaluación genética e interpretación de variantes en el CCR (Fujita et al., 2022).

En la actualidad se ha reconocido la implicación de las variantes sinónimas como causales de fenotipos de enfermedades mendelianas y multifactoriales, incluidas el cáncer (Bailey et al., 2014; Giacoletto et al., 2023; Kaissarian et al., 2022; Lebeuf-Taylor et al., 2019; Lin et al., 2023; Zeng & Bromberg, 2019; Zhou et al., 2021). De manera relevante, el modelo de BoostDM identificó como *oncdriver* dos variantes sinónimas en los genes *CDH1* (c.1710T>C) y *NF1* (c.3498C>T), cada una de ellas en pacientes diferentes. Este hallazgo es interesante y mucho más por involucrar genes relacionados con neurofibromatosis (*NF1*) y cáncer gástrico hereditario difuso (*CDH1*). Aun así, el gen *NF1* es un gen supresor tumoral que contribuye al desarrollo del cáncer y se ha asociado a adenocarcinoma del tracto gastrointestinal (Y. Li et al., 1992). Estudios llevados a cabo por Seminog & Goldacre, 2013, determinaron que los pacientes con mutaciones patogénicas en *NF1* tienen mayor riesgo de cáncer de colon comparado con la población general, lo cual apoya el hallazgo del paciente analizado en nuestro estudio (Seminog & Goldacre, 2013). Es importante resaltar que esta variante en *NF1* tiene una clasificación *oncdriver* para el filtro B, sin embargo, no cumple los criterios

suficientes para ser clasificada como P/PP por la ACMG/AMP, por tal razón este hallazgo no es reportado como una variante asociada a neurofibromatosis.

Por su parte, variantes germinales patogénicas en el gen *CDH1* son las responsables de más del 20% del cáncer gástrico difuso hereditario. Sin embargo, en concordancia con nuestros hallazgos, estudios recientes han establecido que ser portador de una variante en este gen, genera una mayor predisposición de pólipos colorrectales, lo que supone una posible asociación de variantes en *CDH1* y riesgo de CCR. Estos datos son relevantes para la evaluación del riesgo de cáncer y el asesoramiento de pacientes con variantes en el gen (Passi et al., 2023). Los mecanismos moleculares asociados a la patogenicidad de variantes sinónimas han sido explorados, pero se considera que mayoritariamente afecta el splicing, generando transcritos aberrantes (ej: *exon skipping*). Un estudio reciente que analizó este tipo de variantes en más de 3000 muestras de cáncer demostró que cerca del 6-8% de todas las variantes *driver* en oncogenes son sinónimas (Supek et al., 2014). En este contexto, podemos resaltar la importancia de incorporar métodos adicionales de detección de variantes relacionadas al desarrollo del cáncer, dado que los criterios ACMG/AMP pueden excluir estas variantes. Sin embargo, los hallazgos deben analizarse con cautela y ser apoyados por estudios de validación funcional.

Además, el aporte de variantes *oncdriver* en genes usualmente usados en los paneles de diagnóstico, el análisis molecular ampliado con genes candidatos es de impacto, tal como lo demuestran nuestros resultados en los que fueron identificadas variantes en *FANCC*, *KDR* y *TCF7L2*. *FANCC* participa en la reparación del ADN y el proceso de transcripción, y se ha relacionado con un mayor riesgo de CCR (Esteban-Jurado et al., 2016; Laitman et al., 2016; Martínez-Romero et al., 2018). Por otro lado, *KDR* codifica uno de los dos receptores del VEGF. Este es un receptor tirosina quinasa de tipo III y está involucrado en la proliferación, supervivencia, migración y morfogénesis tubular de las células endoteliales, y se ha relacionado como un marcador pronóstico en el CCR (Dong et al., 2012; Toledo et al., 2018). Finalmente, *TCF7L2* desempeña un papel clave en la vía de señalización Wnt, y también se han encontrado variantes en este gen en pacientes con CCR (Lu et al., 2020; Testa et al., 2018). Tomando en conjunto los hallazgos derivados del modelo BoostDM, podemos resaltar una alta capacidad de identificación de mutaciones *oncdriver*, lo que apoya recientes afirmaciones del uso de inteligencia artificial como una herramienta que podría mejorar la precisión y la eficacia para la predicción, el diagnóstico y el tratamiento del cáncer (Sebastian & Peter, 2022). El análisis genético-poblacional llevado a cabo en nuestro estudio identificó por primera vez en el país y en Latinoamérica la frecuencia de variantes patogénicas, probablemente patogénicas y *oncdriver* de 206 genes relacionados con CCR. La comparación de las frecuencias alélicas con datos obtenidos de la base de datos gnomAD, permitió establecer diferencias estadísticamente significativas con la población mundial y con la Latinoamericana. Las mayores diferencias fueron obtenidas en poblaciones no Latinoamericanas, evidenciando que cerca del 46% de las variantes identificadas por nosotros, se presentan en mayor o menor frecuencia. Por el contrario, solo del 23% al 29% de las variantes de este estudio fueron diferentes a otras Latinoamericanas. Este hallazgo es de relevancia ya que reitera la necesidad de realizar estudios como el presente, en donde se realice caracterización genómica de poblaciones usualmente subrepresentadas en las bases de datos públicas.

Finalmente, llevamos a cabo un análisis de validación funcional de mutaciones patogénicas que fueron identificadas en genes no convencionalmente involucrados en los paneles genéticos de diagnóstico molecular, con el fin de aportar nuevo conocimiento de la potencial implicación en la etiología del CCR. En este contexto, mutaciones intrónicas en sitios canónicos de splicing fueron evaluadas mediante ensayo de minigenes. Este tipo de variantes son reconocidas por su potencial efecto sobre el splicing y la generación de transcritos aberrantes por mecanismos de omisión de exones, retención de intrones o formación de pseudogenes (Lim et al., 2011).

Nuestros hallazgos permitieron evidenciar que la variante *SMAD9*:c.781+2T>A produce la omisión de un exón, lo que resulta en la pérdida de 37 aminoácidos. Este efecto molecular afecta potencialmente la función de la proteína, que desempeña un papel crucial en la vía de señalización del TGFβ. Esta vía se ha relacionado con la regulación de procesos pro-oncogénicos, como la invasión, la transición epitelial-mesenquimal y promoción de la angiogénesis (Maslankova et al., 2022). Se han identificado SNPs en este gen en pacientes con CCR y se han asociado con un mayor riesgo, principalmente debido a su papel en la vía de señalización TGFβ (Chen et al., 2022; Djursby et al., 2020; Huyghe et al., 2019; Law et al., 2019; Yuan et al., 2021).

La variante c.702+1G>A en el gen *ERCC1* produce una pérdida del exón 6 (100 pb), que genera una mutación *frameshift*, con corrimiento en el marco de lectura y aparición de un codón de parada prematuro (p.Ser201Argfs*1). En este caso particular, hubo una preferencia por la isoforma con la omisión del exón tanto en WT como en MUT. A pesar de que se encontraron diferencias estadísticamente significativas entre el WT Y mutante en HCT-116 (P < 0.05), el transcrito predominante en las dos versiones fue el *exon skipping*, razón por la cual estimamos que no existe un claro efecto funcional derivado de la mutación *ERCC1* c.702+1G>A. En el contexto de la implicación biológica, no encontramos evidencia de literatura que soporte que la isoforma con el *exón skipping* sea activa, por lo que sugerimos que este resultado puede ser debido a un efecto inducido por el sistema *in vitro* utilizado. Este efecto ha sido previamente descrito en ensayos previos de minigenes, usando diferentes vectores de expresión, incluido el pSpliceExpress (Kishore et al., 2008).

La variante en *IL12RB1*:c.1791+2T>G tuvo el impacto más significativo, lo que resultó en la generación de dos transcritos aberrantes. Por un lado, provocó la omisión del exón 15 (76 pb) y, por otro lado, la retención del intrón 14 (98 pb). Ambos efectos causaron un cambio de marco de lectura y la generación de un codón de parada prematuro, lo que potencialmente afecta la integridad estructural de la proteína. *IL12RB1* es un receptor de interleucina que desempeña un papel en las vías de reparación del daño celular (Pietrzyk et al., 2015). Variantes en este gen en pacientes con CCR, y algunas proporcionan evidencia convincente de su papel en la predisposición a esta neoplasia (Chubb et al., 2016; Terradas et al., 2020).

Según los resultados que hemos obtenido, los ensayos con minigenes demostraron ser una herramienta robusta de validación de la consecuencia funcional de las mutaciones de splicing. Este enfoque proporciona un medio valioso para corroborar la patogenicidad de este tipo de variantes y la determinación del efecto molecular subyacente a este tipo de cambios genéticos (Bueno-Martínez et al., 2022; Cooper, 2005; Putscher et al., 2021; Rhine et al., 2019).

Comprender la predisposición genética al CCR es esencial para el diagnóstico temprano, la prevención y el tratamiento de los pacientes. Por esta razón, el uso de estrategias de NGS nos han permitido enriquecer el perfil genómico de pacientes colombianos de CCR, aportando al conocimiento de variantes y frecuencias poblacionales en nuestra población. Adicionalmente, la aplicación de filtros de clasificación usados en la práctica clínica, como ACMG/AMP, así como modelos de inteligencia artificial, nos permiten ampliar el número de variantes implicadas en la predisposición de CCR.

Si bien, las mejoras en las tecnologías de secuenciación y la aparición de plataformas bioinformáticas avanzadas de inteligencia artificial, como el modelo BoostDM, han ampliado nuestras herramientas para comprender los mecanismos por los cuales las variantes afectan a los genes o proteínas, incluyendo las variantes sinónimas que a menudo no son consideradas como patogénicas, es necesario ampliar la evidencia de la implicación molecular con análisis de validación funcional (Lin et al., 2023).

5. Conclusiones

- Nuestros hallazgos son pioneros en la determinación del perfil genómico de mutaciones germinales en pacientes con CCR no seleccionados de la población colombiana, utilizando un panel customizado ampliado, que incluyó genes establecidos en los paneles de diagnóstico y genes candidatos.
- Se evidenció la importancia de evaluar genes que suelen omitirse en el análisis molecular de rutina, resaltando la omisión de una proporción sustancial de variantes P/PP. Estos genes no son incorporados en los análisis genéticos, quizás por la falta de evidencia de su relación con la aparición de CCR, por lo que proyectos como el desarrollado generan nuevos elementos de potencial aplicación traslacional.
- Por primera vez se pudo establecer la frecuencia de variantes P/PP y *oncdriver* en una muestra de pacientes colombianos de CCR, usando un panel de 206 genes relacionados con CCR. Nuestros hallazgos demuestran mayores semejanzas entre nuestros pacientes y la población Latinoamericana que entre las frecuencias globales.
- El uso de modelos de inteligencia artificial amplía el enfoque y las estrategias de identificación de nuevas variantes con efectos patogénicos sobre genes de predisposición al CCR, ya que son entrenados con algoritmos diferentes a los usados en la práctica convencional de clasificación de variantes. Incluso, como fue demostrado en esta investigación, este nuevo enfoque resulta útil para evaluar la patogenicidad de las variantes sinónimas, que comúnmente no cumplen los criterios suficientes para una clasificación P/PP.
- Mediante el uso de minigenes, demostramos con éxito los efectos de tres variantes P/PP encontradas en nuestros pacientes y cómo afectaron el splicing. Este hallazgo fortalece la hipótesis de la potencial alteración a nivel de la proteína y proporciona criterios valiosos para una clasificación de patogenicidad más precisa.
- Los resultados obtenidos demostraron la utilidad de implementar un panel de genes ampliado para la identificación de variantes germinales en pacientes con CCR no seleccionado, los cuales son menos estudiados en relación con los casos familiares.

Este proyecto genera perspectivas desde el punto de vista de la medicina traslacional, ya que proporciona elementos para la inclusión de genes de interés en el análisis de casos de CCR no seleccionado. Así mismo, aporta información relacionada con variantes que han sido recientemente reconocidas por su implicación en el fenotipo, como es el caso de las variantes sinónimas. En este contexto, estimamos que la determinación de su patogenicidad mediante análisis *in silico* y de validación funcional es innovador. Con esta perspectiva, a la fecha hemos diseñado los protocolos para la validación funcional de variantes sinónimas derivadas de este trabajo, que representa una oportunidad para trabajos posteriores.

6. Referencias

- Ahmad, R., Singh, J. kumar, Wunnava, A., Al-obeed, O., Absulla, M., & Srivastava, S. K. (2021). Emerging trends in colorectal cancer: Dysregulated signaling pathways (Review). *International Journal of Molecular Medicine*, 47(3), 1–25. <https://doi.org/10.3892/ijmm.2021.4847>
- Alanazi, I. O., Shaik, J. P., Parine, N. R., Al Naeem, A., Azzam, N. A., Almadi, M. A., Aljebreen, A. M., Alharbi, O., Alanazi, M. S., & Khan, Z. (2021). NOTCH Single Nucleotide Polymorphisms in the Predisposition of Breast and Colorectal Cancers in Saudi Patients. *Pathology and Oncology Research*, 27. <https://doi.org/10.3389/pore.2021.616204>
- Alonso-Espinaco, V., Giráldez, M. D., Trujillo, C., van der Klift, H., Muñoz, J., Balaguer, F., Ocaña, T., Madrigal, I., Jones, A. M., Echeverry, M. M., Velez, A., Tomlinson, I., Milà, M., Wijnen, J., Carvajal-Carmona, L., Castells, A., & Castellví-Bel, S. (2011). Novel MLH1 duplication identified in Colombian families with Lynch syndrome. *Genetics in Medicine*, 13(2), 155–160. <https://doi.org/10.1097/GIM.0b013e318202e10b>
- Alvarez-Gomez, R. M., De la Fuente-Hernandez, M. A., Herrera-Montalvo, L., & Hidalgo-Miranda, A. (2021). Challenges of diagnostic genomics in Latin America. *Current Opinion in Genetics and Development*, 66, 101–109. <https://doi.org/10.1016/j.gde.2020.12.010>
- Bailey, S. F., Hinz, A., & Kassen, R. (2014). Adaptive synonymous mutations in an experimentally evolved *Pseudomonas fluorescens* population. *Nature Communications*, 5(1), 4076. <https://doi.org/10.1038/ncomms5076>
- Benson, A. B., Venook, A. P., Al-Hawary, M. M., Arain, M. A., Chen, Y.-J., Ciombor, K. K., Cohen, S., Cooper, H. S., Deming, D., Farkas, L., Garrido-Laguna, I., Grem, J. L., Gunn, A., Hecht, J. R., Hoffe, S., Hubbard, J., Hunt, S., Johung, K. L., Kirilcuk, N., ... Gurski, L. A. (2021). Colon Cancer, Version 2.2021, NCCN Clinical Practice Guidelines in Oncology. *Journal of the National Comprehensive Cancer Network*, 19(3), 329–359. <https://doi.org/10.6004/jnccn.2021.0012>
- Bessho, T. (1999). Nucleotide excision repair 3' endonuclease XPG stimulates the activity of base excision repair enzyme thymine glycol DNA glycosylase. *Nucleic Acids Research*, 27(4), 979–983. <https://doi.org/10.1093/nar/27.4.979>
- Billing, D., Horiguchi, M., Wu-Baer, F., Tagliabata, A., Leuzzi, G., Nanez, S. A., Jiang, W., Zha, S., Szabolcs, M., Lin, C.-S., Ciccia, A., & Baer, R. (2018). The BRCT Domains of the BRCA1 and BARD1 Tumor Suppressors Differentially Regulate Homology-Directed Repair and Stalled Fork Protection. *Molecular Cell*, 72(1), 127-139.e8. <https://doi.org/10.1016/j.molcel.2018.08.016>
- Blatter, R., Tschupp, B., Aretz, S., Bernstein, I., Colas, C., Evans, D. G., Genuardi, M., Hes, F. J., Hüneburg, R., Järvinen, H., Laloo, F., Moeslein, G., Renkonen-Sinisalo, L., Resta, N., Spier, I., Varvara, D., Vasen, H., Latchford, A. R., & Heinimann, K. (2020). Disease expression in juvenile polyposis syndrome: a retrospective survey on a cohort of 221 European patients and comparison with a literature-derived cohort of 473 SMAD4/BMP1A pathogenic variant carriers. *Genetics in Medicine*, 22(9), 1524–1532. <https://doi.org/10.1038/s41436-020-0826-1>
- Bolderson, E., Tomimatsu, N., Richard, D. J., Boucher, D., Kumar, R., Pandita, T. K., Burma, S., & Khanna, K. K. (2010). Phosphorylation of Exo1 modulates homologous recombination repair of DNA double-strand breaks. *Nucleic Acids Research*, 38(6), 1821–1831. <https://doi.org/10.1093/nar/gkp1164>
- Bork, P., Hofmann, K., Bucher, P., Neuwald, A. F., Altschul, S. F., & Koonin, E. V. (1997). A superfamily of conserved domains in DNA damage- responsive cell cycle checkpoint proteins. *The FASEB Journal*, 11(1), 68–76. <https://doi.org/10.1096/fasebj.11.1.9034168>
- Bueno-Martínez, E., Sanoguera-Miralles, L., Valenzuela-Palomo, A., Esteban-Sánchez, A., Lorca, V., Llinares-Burguet, I., Allen, J., García-Álvarez, A., Pérez-Segura, P., Durán, M., Easton, D. F., Devilee, P., Vreeswijk, M. P., de la Hoya, M., & Velasco-Sampedro, E. A. (2022). Minigene-based splicing analysis and ACMG/AMP-based tentative classification of 56 ATM variants. *The Journal of Pathology*, 258(1), 83–101. <https://doi.org/10.1002/path.5979>
- Carroll, B. L., Zahn, K. E., Hanley, J. P., Wallace, S. S., Dragon, J. A., & Doublié, S. (2021). Caught in motion: human NTHL1 undergoes interdomain rearrangement necessary for catalysis. *Nucleic Acids Research*, 49(22), 13165–13178. <https://doi.org/10.1093/nar/gkab1162>

- Chassaing, N., Davis, E. E., McKnight, K. L., Niederriter, A. R., Causse, A., David, V., Desmaison, A., Lamarre, S., Vincent-Delorme, C., Pasquier, L., Coubes, C., Lacombe, D., Rossi, M., Dufier, J.-L., Dollfus, H., Kaplan, J., Katsanis, N., Etchevers, H. C., Faguer, S., & Calvas, P. (2016). Targeted resequencing identifies PTCH1 as a major contributor to ocular developmental anomalies and extends the SOX2 regulatory network. *Genome Research*, *26*(4), 474–485. <https://doi.org/10.1101/gr.196048.115>
- Chen, L., Ye, L., & Hu, B. (2022). Hereditary Colorectal Cancer Syndromes: Molecular Genetics and Precision Medicine. *Biomedicine*, *10*(12). <https://doi.org/10.3390/biomedicine10123207>
- Chubb, D., Broderick, P., Dobbins, S. E., Frampton, M., Kinnersley, B., Penegar, S., Price, A., Ma, Y. P., Sherborne, A. L., Palles, C., Timofeeva, M. N., Bishop, D. T., Dunlop, M. G., Tomlinson, I., & Houlston, R. S. (2016). Rare disruptive mutations and their contribution to the heritable risk of colorectal cancer. *Nature Communications*, *7*(1), 11883. <https://doi.org/10.1038/ncomms11883>
- Cooper, T. A. (2005). Use of minigene systems to dissect alternative splicing elements. *Methods*, *37*(4), 331–340. <https://doi.org/10.1016/j.ymeth.2005.07.015>
- Cross, W., Kovac, M., Mustonen, V., Temko, D., Davis, H., Baker, A.-M., Biswas, S., Arnold, R., Chegwidan, L., Gatenbee, C., Anderson, A. R., Koelzer, V. H., Martinez, P., Jiang, X., Domingo, E., Woodcock, D. J., Feng, Y., Kovacova, M., Maughan, T., ... Tomlinson, I. P. M. (2018). The evolutionary landscape of colorectal tumorigenesis. *Nature Ecology & Evolution*, *2*(10), 1661–1672. <https://doi.org/10.1038/s41559-018-0642-z>
- Cuenta de alto costo (CAC) /HIGIA. (2022). <https://cuentadealtocosto.org/higia/cancer-morbimortalidad-demografico/>
- Dai, Q., Luo, H., Li, X. P., Huang, J., Zhou, T. J., & Yang, Z.-H. (2015). XRCC1 and ERCC1 polymorphisms are related to susceptibility and survival of colorectal cancer in the Chinese population. *Mutagenesis*, *30*(3), 441–449. <https://doi.org/10.1093/mutage/geu088>
- Das, L., Quintana, V. G., & Sweasy, J. B. (2020). NTHL1 in genomic integrity, aging and cancer. *DNA Repair*, *93*, 102920. <https://doi.org/10.1016/j.dnarep.2020.102920>
- Djursby, M., Madsen, M. B., Frederiksen, J. H., Berchtold, L. A., Therkildsen, C., Willemo, G. L., Hasselby, J. P., Wikman, F., Okkels, H., Skytte, A. B., Nilbert, M., Wadt, K., Gerdes, A. M., & van Overeem Hansen, T. (2020). New Pathogenic Germline Variants in Very Early Onset and Familial Colorectal Cancer Patients. *Frontiers in Genetics*, *11*(September), 1–16. <https://doi.org/10.3389/fgene.2020.566266>
- Dong, G., Guo, X., Fu, X., Wan, S., Zhou, F., Myers, R. E., Bao, G., Burkart, A., Yang, H., & Xing, J. (2012). Potentially functional genetic variants in KDR gene as prognostic markers in patients with resected colorectal cancer. *Cancer Science*, *103*(3), 561–568. <https://doi.org/10.1111/j.1349-7006.2011.02194.x>
- Esteban-Jurado, C., Franch-Expósito, S., Muñoz, J., Ocaña, T., Carballal, S., López-Cerón, M., Cuatrecasas, M., Vila-Casadesús, M., Lozano, J. J., Serra, E., Beltran, S., Brea-Fernández, A., Ruiz-Ponte, C., Castells, A., Bujanda, L., Garre, P., Caldés, T., Cubiella, J., Balaguer, F., & Castellví-Bel, S. (2016). The Fanconi anemia DNA damage repair pathway in the spotlight for germline predisposition to colorectal cancer. *European Journal of Human Genetics*, *24*(10), 1501–1505. <https://doi.org/10.1038/ejhg.2016.44>
- Ferlitsch, M., Reinhart, K., Pramhas, S., Wiener, C., Gal, O., Bannert, C., Hassler, M., Kozbial, K., Dunkler, D., Trauner, M., & Weiss, W. (2011). Sex-Specific Prevalence of Adenomas, Advanced Adenomas, and Colorectal Cancer in Individuals Undergoing Screening Colonoscopy. *JAMA*, *306*(12), 1352. <https://doi.org/10.1001/jama.2011.1362>
- Fujita, M., Liu, X., Iwasaki, Y., Terao, C., Mizukami, K., Kawakami, E., Takata, S., Inai, C., Aoi, T., Mizukoshi, M., Maejima, K., Hirata, M., Murakami, Y., Kamatani, Y., Kubo, M., Akagi, K., Matsuda, K., Nakagawa, H., & Momozawa, Y. (2022). Population-based Screening for Hereditary Colorectal Cancer Variants in Japan. *Clinical Gastroenterology and Hepatology*, *20*(9), 2132-2141.e9. <https://doi.org/10.1016/j.cgh.2020.12.007>
- Garg, S., Grenier, S., Misyura, M., Sukhai, M. A., Thomas, M., Kamel-Reid, S., & Stockley, T. (2020). Assessing the Diagnostic Yield of Targeted Next-Generation Sequencing for Melanoma and Gastrointestinal Tumors. *The*

Journal of Molecular Diagnostics, 22(4), 467–475. <https://doi.org/10.1016/j.jmoldx.2019.12.008>

- Gerloff, D. L., Woods, N. T., Farago, A. A., & Monteiro, A. N. A. (2012). BRCT domains: A little more than kin, and less than kind. *FEBS Letters*, 586(17), 2711–2716. <https://doi.org/10.1016/j.febslet.2012.05.005>
- Giacoletto, C. J., Benjamin, R., Deng, H.-W., Rotter, J. I., & Schiller, M. R. (2023). Most synonymous allelic variants in HIV tat are not silent. *Genomics*, 115(3), 110603. <https://doi.org/10.1016/j.ygeno.2023.110603>
- Globocan. (2020). *International Agency for research on cancer - WHO*. <https://gco.iarc.fr/today/home>
- Gong, R., He, Y., Liu, X. Y., Wang, H. Y., Sun, L. Y., Yang, X. H., Li, B., Cao, X. K., Ye, Z. L., Kong, L. H., Zhang, D. D., Li, Y. H., Xu, R. H., & Shao, J. Y. (2019). Mutation spectrum of germline cancer susceptibility genes among unselected Chinese colorectal cancer patients. *Cancer Management and Research*, 11, 3721–3739. <https://doi.org/10.2147/CMAR.S193985>
- Hamamoto, R., Koyama, T., Kouno, N., Yasuda, T., Yui, S., Sudo, K., Hirata, M., Sunami, K., Kubo, T., Takasawa, K., Takahashi, S., Machino, H., Kobayashi, K., Asada, K., Komatsu, M., Kaneko, S., Yatabe, Y., & Yamamoto, N. (2022). Introducing AI to the molecular tumor board: one direction toward the establishment of precision medicine using large-scale cancer clinical and biological information. *Experimental Hematology & Oncology*, 11(1), 82. <https://doi.org/10.1186/s40164-022-00333-7>
- Hampel, H., Bennett, R. L., Buchanan, A., Pearlman, R., & Wiesner, G. L. (2015). A practice guideline from the American College of Medical Genetics and Genomics and the National Society of Genetic Counselors: Referral indications for cancer predisposition assessment. *Genetics in Medicine*, 17(1), 70–87. <https://doi.org/10.1038/gim.2014.147>
- Howe, J. R., Dahdaleh, F. S., Carr, J. C., Wang, D., Sherman, S. K., & Howe, J. R. (2013). BMPR1A mutations in juvenile polyposis affect cellular localization. *Journal of Surgical Research*, 184(2), 739–745. <https://doi.org/10.1016/j.jss.2013.01.015>
- Huang, K., Luo, W., Fang, J., Yu, C., Liu, G., Yuan, X., Liu, Y., & Wu, W. (2023). Notch3 signaling promotes colorectal tumor growth by enhancing immunosuppressive cells infiltration in the microenvironment. *BMC Cancer*, 23(1), 55. <https://doi.org/10.1186/s12885-023-10526-w>
- Huyghe, J. R., Bien, S. A., Harrison, T. A., Kang, H. M., Chen, S., Schmit, S. L., Conti, D. V., Qu, C., Jeon, J., Edlund, C. K., Greenside, P., Wainberg, M., Schumacher, F. R., Smith, J. D., Levine, D. M., Nelson, S. C., Sinnott-Armstrong, N. A., Albanes, D., Alonso, M. H., ... Peters, U. (2019). Discovery of common and rare genetic risk variants for colorectal cancer. *Nature Genetics*, 51(1), 76–87. <https://doi.org/10.1038/s41588-018-0286-6>
- Ioannidis, N. M., Rothstein, J. H., Pejaver, V., Middha, S., McDonnell, S. K., Baheti, S., Musolf, A., Li, Q., Holzinger, E., Karyadi, D., Cannon-Albright, L. A., Teerlink, C. C., Stanford, J. L., Isaacs, W. B., Xu, J., Cooney, K. A., Lange, E. M., Schleutker, J., Carpten, J. D., ... Sieh, W. (2016). REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. *The American Journal of Human Genetics*, 99(4), 877–885. <https://doi.org/10.1016/j.ajhg.2016.08.016>
- Jiang, Q., & Greenberg, R. A. (2015). Deciphering the BRCA1 Tumor Suppressor Network. *Journal of Biological Chemistry*, 290(29), 17724–17732. <https://doi.org/10.1074/jbc.R115.667931>
- Jiang, T. J., Wang, F., Wang, Y. N., Hu, J. J., Ding, P. R., Lin, J. Z., Pan, Z. Z., Chen, G., Shao, J. Y., Xu, R. hua, Zhao, Q., & Wang, F. (2020). Germline mutational profile of Chinese patients under 70 years old with colorectal cancer. *Cancer Communications*, 40(11), 620–632. <https://doi.org/10.1002/cac2.12093>
- Jirka, G. W., Lefler, D. S., Russo, J., & Bashir, B. (2023). Colon adenocarcinoma and Birt–Hogg–Dubé syndrome in a young patient: case report and exploration of pathologic implications. *Cancer Biology & Therapy*, 24(1). <https://doi.org/10.1080/15384047.2023.2184153>
- Kaissarian, N. M., Meyer, D., & Kimchi-Sarfaty, C. (2022). Synonymous Variants: Necessary Nuance in Our Understanding of Cancer Drivers and Treatment Outcomes. *JNCI: Journal of the National Cancer Institute*, 114(8), 1072–1094. <https://doi.org/10.1093/jnci/djac090>

- Keijzers, G., Bakula, D., Petr, M., Madsen, N., Teklu, A., Mkrtychyan, G., Osborne, B., & Scheibye-Knudsen, M. (2018). Human Exonuclease 1 (EXO1) Regulatory Functions in DNA Replication with Putative Roles in Cancer. *International Journal of Molecular Sciences*, *20*(1), 74. <https://doi.org/10.3390/ijms20010074>
- Kishore, S., Khanna, A., & Stamm, S. (2008). Rapid generation of splicing reporters with pSpliceExpress. *Gene*, *427*(1–2), 104–110. <https://doi.org/10.1016/j.gene.2008.09.021>
- Lai, C.-Y., Hsieh, L.-L., Tang, R., Santella, R. M., Chang-Chieh, C. R., & Yeh, C.-C. (2016). Association between polymorphisms of APE1 and OGG1 and risk of colorectal cancer in Taiwan. *World Journal of Gastroenterology*, *22*(12), 3372–3380. <https://doi.org/10.3748/wjg.v22.i12.3372>
- Laitman, Y., Boker-Keinan, L., Berkenstadt, M., Liphstiz, I., Weissglas-Volkov, D., Ries-Levavi, L., Sarouk, I., Pras, E., & Friedman, E. (2016). The risk for developing cancer in Israeli ATM, BLM, and FANCC heterozygous mutation carriers. *Cancer Genetics*, *209*(3), 70–74. <https://doi.org/10.1016/j.cancergen.2015.12.006>
- Law, P. J., Timofeeva, M., Fernandez-Rozadilla, C., Broderick, P., Studd, J., Fernandez-Tajes, J., Farrington, S., Svinti, V., Palles, C., Orlando, G., Sud, A., Holroyd, A., Penegar, S., Theodoratou, E., Vaughan-Shaw, P., Campbell, H., Zgaga, L., Hayward, C., Campbell, A., ... Dunlop, M. G. (2019). Association analyses identify 31 new risk loci for colorectal cancer susceptibility. *Nature Communications*, *10*(1), 1–15. <https://doi.org/10.1038/s41467-019-09775-w>
- Lebeuf-Taylor, E., McCloskey, N., Bailey, S. F., Hinz, A., & Kassen, R. (2019). The distribution of fitness effects among synonymous mutations in a gene under directional selection. *ELife*, *8*. <https://doi.org/10.7554/eLife.45952>
- Leung, C. C. Y., & Glover, J. N. M. (2011). BRCT domains. *Cell Cycle*, *10*(15), 2461–2470. <https://doi.org/10.4161/cc.10.15.16312>
- Lewandowska, A., Rudzki, G., Lewandowski, T., Strykowska-Góra, A., & Rudzki, S. (2022). Title: Risk Factors for the Diagnosis of Colorectal Cancer. *Cancer Control*, *29*, 1–15. <https://doi.org/10.1177/10732748211056692>
- Li, Y.-K., Xu, Q., Sun, L.-P., Gong, Y.-H., Jing, J.-J., Xing, C.-Z., & Yuan, Y. (2020). Nucleotide excision repair pathway gene polymorphisms are associated with risk and prognosis of colorectal cancer. *World Journal of Gastroenterology*, *26*(3), 307–323. <https://doi.org/10.3748/wjg.v26.i3.307>
- Li, Y., Bollag, G., Clark, R., Stevens, J., Conroy, L., Fults, D., Ward, K., Friedman, E., Samowitz, W., Robertson, M., Bradley, P., McCormick, F., White, R., & Cawthon, R. (1992). Somatic mutations in the neurofibromatosis 1 gene in human tumors. *Cell*, *69*(2), 275–281. [https://doi.org/10.1016/0092-8674\(92\)90408-5](https://doi.org/10.1016/0092-8674(92)90408-5)
- Lim, K. H., Ferraris, L., Filloux, M. E., Raphael, B. J., & Fairbrother, W. G. (2011). Using positional distribution to identify splicing elements and predict pre-mRNA processing defects in human genes. *Proceedings of the National Academy of Sciences*, *108*(27), 11093–11098. <https://doi.org/10.1073/pnas.1101135108>
- Lin, B. C., Katneni, U., Jankowska, K. I., Meyer, D., & Kimchi-Sarfaty, C. (2023). In silico methods for predicting functional synonymous variants. *Genome Biology*, *24*(1), 126. <https://doi.org/10.1186/s13059-023-02966-1>
- Liu, Q., Liu, M., Liu, T., & Yu, Y. (2020). Familial juvenile polyposis syndrome with a de novo germline missense variant in BMPR1A gene: a case report. *BMC Medical Genetics*, *21*(1), 196. <https://doi.org/10.1186/s12881-020-01135-6>
- Lu, Y., Kweon, S.-S., Cai, Q., Tanikawa, C., Shu, X.-O., Jia, W.-H., Xiang, Y.-B., Huyghe, J. R., Harrison, T. A., Kim, J., Shin, A., Kim, D.-H., Matsuo, K., Jee, S. H., Guo, X., Wen, W., Shi, J., Li, B., Wang, N., ... Zheng, W. (2020). Identification of Novel Loci and New Risk Variant in Known Loci for Colorectal Cancer Risk in East Asians. *Cancer Epidemiology, Biomarkers & Prevention*, *29*(2), 477–486. <https://doi.org/10.1158/1055-9965.EPI-19-0755>
- Martinez-Romero, J., Bueno-Fortes, S., Martín-Merino, M., Ramirez de Molina, A., & De Las Rivas, J. (2018). Survival marker genes of colorectal cancer derived from consistent transcriptomic profiling. *BMC Genomics*, *19*(S8), 857. <https://doi.org/10.1186/s12864-018-5193-9>
- Maslankova, J., Vecurkovska, I., Rabajdova, M., Katuchova, J., Kicka, M., Gayova, M., & Katuch, V. (2022). Regulation of transforming growth factor- β signaling as a therapeutic approach to treating colorectal cancer. *World Journal of*

Gastroenterology, 28(33), 4744–4761. <https://doi.org/10.3748/wjg.v28.i33.4744>

- Massagué, J. (1996). TGF β Signaling: Receptors, Transducers, and Mad Proteins. *Cell*, 85(7), 947–950. [https://doi.org/10.1016/S0092-8674\(00\)81296-9](https://doi.org/10.1016/S0092-8674(00)81296-9)
- Moreno, V., Gemignani, F., Landi, S., Gioia-Patricola, L., Chabrier, A., Blanco, I., González, S., Guino, E., Capellà, G., & Canzian, F. (2006). Polymorphisms in Genes of Nucleotide and Base Excision Repair: Risk and Prognosis of Colorectal Cancer. *Clinical Cancer Research*, 12(7), 2101–2108. <https://doi.org/10.1158/1078-0432.CCR-05-1363>
- Muiños, F., Martínez-Jiménez, F., Pich, O., Gonzalez-Perez, A., & Lopez-Bigas, N. (2021). In silico saturation mutagenesis of cancer genes. *Nature*, 596(7872), 428–432. <https://doi.org/10.1038/s41586-021-03771-1>
- Nahorski, M. S., Lim, D. H. K., Martin, L., Gille, J. J. P., McKay, K., Rehal, P. K., Ploeger, H. M., van Steensel, M., Tomlinson, I. P., Latif, F., Menko, F. H., & Maher, E. R. (2010). Investigation of the Birt-Hogg-Dube tumour suppressor gene (FLCN) in familial and sporadic colorectal cancer. *Journal of Medical Genetics*, 47(6), 385–390. <https://doi.org/10.1136/jmg.2009.073304>
- National Comprehensive Cancer Network. (2023). *Clinical Practice Guidelines in Oncology. Genetic/Familial High-Risk Assessment: Colorectal*. https://www.nccn.org/professionals/physician_gls/pdf/genetics_colon.pdf
- Ngeow, J., Heald, B., Rybicki, L. A., Orloff, M. S., Chen, J. L., Liu, X., Yerian, L., Willis, J., Lehtonen, H. J., Lehtonen, R., Mester, J. L., Moline, J., Burke, C. A., Church, J., Aaltonen, L. A., & Eng, C. (2013). Prevalence of Germline PTEN, BMPR1A, SMAD4, STK11, and ENG Mutations in Patients With Moderate-Load Colorectal Polyps. *Gastroenterology*, 144(7), 1402-1409.e5. <https://doi.org/10.1053/j.gastro.2013.02.001>
- Ngeow, J., Yu, W., Yehia, L., Niazi, F., Chen, J., Tang, X., Heald, B., Lei, J., Romigh, T., Tucker-Kellogg, L., Lim, K. H., Song, H., & Eng, C. (2015). Exome Sequencing Reveals Germline SMAD9 Mutation That Reduces Phosphatase and Tensin Homolog Expression and Is Associated With Hamartomatous Polyposis and Gastrointestinal Ganglioneuromas. *Gastroenterology*, 149(4), 886-889.e5. <https://doi.org/10.1053/j.gastro.2015.06.027>
- Ni, M., Zhang, W., Qiu, J., Liu, F., Li, M., Zhang, Y., Liu, Q., & Bai, J. (2014). Association of ERCC1 and ERCC2 polymorphisms with colorectal cancer risk in a Chinese population. *Scientific Reports*, 4(1), 4112. <https://doi.org/10.1038/srep04112>
- Olkinuora, A. P., Peltomäki, P. T., Aaltonen, L. A., & Rajamäki, K. (2021). From APC to the genetics of hereditary and familial colon cancer syndromes. *Human Molecular Genetics*, 30(R2), R206–R224. <https://doi.org/10.1093/hmg/ddab208>
- Palacio-Rúa, K. A., Isaza-Jiménez, L. F., Ahumada-Rodríguez, E., Ceballos-García, H., & Muñetón-Peña, C. M. (2014). Genetic analysis in APC, KRAS, and TP53 in patients with stomach and colon cancer. *Revista de Gastroenterología de México*, 79(2), 79–89. <https://doi.org/10.1016/j.rgmx.2014.05.001>
- Palacio Rúa, K. A., Ceballos García, H., Ahumada Rodríguez, E., Isaza Jiménez, L. F., & Muñetón Peña, C. M. (2013). Molecular characterization of TP53 tumor suppressor gene in colorectal cancer. *Revista Colombiana de Gastroenterología*, 28(4), 294–300. <https://www.scopus.com/inward/record.uri?eid=2-s2.0-84893405008&partnerID=40&md5=f1544d1cf6ec55dfa1faff2c91c069e0>
- Passi, M., Gamble, L. A., Samaranayake, S. G., Schueler, S. A., Curtin, B. F., Fasaye, G.-A., Bowden, C., Gurram, S., Quezado, M., Miettinen, M., Koh, C., Heller, T., & Davis, J. L. (2023). Association of CDH1 Germline Variants and Colon Polyp Phenotypes in Patients With Hereditary Diffuse Gastric Cancer. *Gastro Hep Advances*, 2(2), 244–251. <https://doi.org/10.1016/j.gastha.2022.10.006>
- Patel, S. G., & Ahnen, D. J. (2012). Familial colon cancer syndromes: An update of a rapidly evolving field. *Current Gastroenterology Reports*, 14(5), 428–438. <https://doi.org/10.1007/s11894-012-0280-6>
- Pearlman, R., Frankel, W. L., Swanson, B., Zhao, W., Yilmaz, A., Miller, K., Bacher, J., Bigley, C., Nelsen, L., Goodfellow, P. J., Goldberg, R. M., Paskett, E., Shields, P. G., Freudenheim, J. L., Stanich, P. P., Lattimer, I., Arnold, M., Liyanarachchi, S., Kalady, M., ... Hampel, H. (2017). Prevalence and spectrum of germline cancer susceptibility

- gene mutations among patients with early-onset colorectal cancer. *JAMA Oncology*, 3(4), 464–471. <https://doi.org/10.1001/jamaoncol.2016.5194>
- Peña-Guerrero, J., Fernández-Rubio, C., García-Sosa, A. T., & Nguewa, P. A. (2023). BRCT Domains: Structure, Functions, and Implications in Disease—New Therapeutic Targets for Innovative Drug Discovery against Infections. *Pharmaceutics*, 15(7), 1839. <https://doi.org/10.3390/pharmaceutics15071839>
- Pietrzyk, L., Torres, A., Maciejewski, R., & Torres, K. (2015). Obesity and Obese-related Chronic Low-grade Inflammation in Promotion of Colorectal Cancer Development. *Asian Pacific Journal of Cancer Prevention*, 16(10), 4161–4168. <https://doi.org/10.7314/APJCP.2015.16.10.4161>
- Poliani, L., Greco, L., Barile, M., Buono, A. D., Bianchi, P., Basso, G., Giatti, V., Genuardi, M., Malesci, A., & Laghi, L. (2022). Canonical and uncanonical pathogenic germline variants in colorectal cancer patients by next-generation sequencing in a European referral center. *ESMO Open*, 7(6), 100607. <https://doi.org/10.1016/j.esmoop.2022.100607>
- Putscher, E., Hecker, M., Fitzner, B., Lorenz, P., & Zettl, U. K. (2021). Principles and practical considerations for the analysis of disease-associated alternative splicing events using the gateway cloning-based minigene vectors pdestsplice and pspliceexpress. *International Journal of Molecular Sciences*, 22(10). <https://doi.org/10.3390/ijms22105154>
- Raleigh, D. R., & Reiter, J. F. (2019). Misactivation of Hedgehog signaling causes inherited and sporadic cancers. *Journal of Clinical Investigation*, 129(2), 465–475. <https://doi.org/10.1172/JCI120850>
- Ramirez Reyes, J. M. J., Cuesta, R., & Pause, A. (2021). Folliculin: A Regulator of Transcription Through AMPK and mTOR Signaling Pathways. *Frontiers in Cell and Developmental Biology*, 9. <https://doi.org/10.3389/fcell.2021.667311>
- Rebuzzi, F., Ulivi, P., & Tedaldi, G. (2023). Genetic Predisposition to Colorectal Cancer: How Many and Which Genes to Test? *International Journal of Molecular Sciences*, 24(3). <https://doi.org/10.3390/ijms24032137>
- Reilly, N. M., Novara, L., Di Nicolantonio, F., & Bardelli, A. (2019). Exploiting DNA repair defects in colorectal cancer. *Molecular Oncology*, 13(4), 681–700. <https://doi.org/10.1002/1878-0261.12467>
- Rhine, C. L., Neil, C., Glidden, D. T., Cygan, K. J., Fredericks, A. M., Wang, J., Walton, N. A., & Fairbrother, W. G. (2019). Future directions for high-throughput splicing assays in precision medicine. *Human Mutation*, 40(9), 1225–1234. <https://doi.org/10.1002/humu.23866>
- Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W. W., Hegde, M., Lyon, E., Spector, E., Voelkerding, K., & Rehm, H. L. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genetics in Medicine*, 17(5), 405–424. <https://doi.org/10.1038/gim.2015.30>
- Roht, L., Tooming, M., Rekker, K., Roomere, H., Toome, K., Murumets, Ü., Šamarina, U., Õunap, K., & Kahre, T. (2022). The prevalence of germline pathogenic variants in Estonian colorectal cancer patients: results from routine clinical setting 2016–2021. *Frontiers in Genetics*, 13. <https://doi.org/10.3389/fgene.2022.1020543>
- Rubenstein, J. H., Tavakkoli, A., Koeppe, E., Ulintz, P., Inadomi, J. M., Morgenstern, H., Appelman, H., Scheiman, J. M., Schoenfeld, P., Metko, V., & Stoffel, E. M. (2020). Family History of Colorectal or Esophageal Cancer in Barrett’s Esophagus and Potentially Explanatory Genetic Variants. *Clinical and Translational Gastroenterology*, 11(4), e00151. <https://doi.org/10.14309/ctg.000000000000151>
- Sattler, E. C., Syunyaeva, Z., Reithmair, M., Dempke, W., & Steinlein, O. K. (2021). Colorectal cancer risk in families with Birt-Hogg-Dubé syndrome increased. *European Journal of Cancer*, 151, 168–174. <https://doi.org/10.1016/j.ejca.2021.04.013>
- Schmidt, L. S., & Linehan, W. M. (2018). FLCN : The causative gene for Birt-Hogg-Dubé syndrome. *Gene*, 640, 28–42. <https://doi.org/10.1016/j.gene.2017.09.044>

- Sebastian, A. M., & Peter, D. (2022). Artificial Intelligence in Cancer Research: Trends, Challenges and Future Directions. *Life*, *12*(12), 1991. <https://doi.org/10.3390/life12121991>
- Seminog, O. O., & Goldacre, M. J. (2013). Risk of benign tumours of nervous system, and of malignant neoplasms, in people with neurofibromatosis: population-based record-linkage study. *British Journal of Cancer*, *108*(1), 193–198. <https://doi.org/10.1038/bjc.2012.535>
- Stodolna, A., He, M., Vasipalli, M., Kingsbury, Z., Becq, J., Stockton, J. D., Dilworth, M. P., James, J., Sillo, T., Blakeway, D., Ward, S. T., Ismail, T., Ross, M. T., & Beggs, A. D. (2021). Clinical-grade whole-genome sequencing and 3' transcriptome analysis of colorectal cancer patients. *Genome Medicine*, *13*(1), 1–15. <https://doi.org/10.1186/s13073-021-00852-8>
- Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, *1*(0), 1–41. <https://doi.org/10.3322/caac.21660>
- Supek, F., Miñana, B., Valcárcel, J., Gabaldón, T., & Lehner, B. (2014). Synonymous Mutations Frequently Act as Driver Mutations in Human Cancers. *Cell*, *156*(6), 1324–1335. <https://doi.org/10.1016/j.cell.2014.01.051>
- Terradas, M., Capellá, G., & Valle, L. (2020). Dominantly Inherited Hereditary Nonpolyposis Colorectal Cancer Not Caused by MMR Genes. *Journal of Clinical Medicine*, *9*(6), 1954. <https://doi.org/10.3390/jcm9061954>
- Testa, U., Pelosi, E., & Castelli, G. (2018). Colorectal Cancer: Genetic Abnormalities, Tumor Progression, Tumor Heterogeneity, Clonal Evolution and Tumor-Initiating Cells. *Medical Sciences*, *6*(2), 31. <https://doi.org/10.3390/medsci6020031>
- Toledo, R. A., Garralda, E., Mitsi, M., Pons, T., Monsech, J., Vega, E., Otero, Á., Albarran, M. I., Baños, N., Durán, Y., Bonilla, V., Sarno, F., Camacho-Artacho, M., Sanchez-Perez, T., Perea, S., Álvarez, R., De Martino, A., Lietha, D., Blanco-Aparicio, C., ... Hidalgo, M. (2018). Exome Sequencing of Plasma DNA Portrays the Mutation Landscape of Colorectal Cancer and Discovers Mutated VEGFR2 Receptors as Modulators of Antiangiogenic Therapies. *Clinical Cancer Research*, *24*(15), 3550–3559. <https://doi.org/10.1158/1078-0432.CCR-18-0103>
- Uson, P. L. S., Riegert-Johnson, D., Boardman, L., Kisiel, J., Mountjoy, L., Patel, N., Lizaola-Mayo, B., Borad, M. J., Ahn, D., Sonbol, M. B., Jones, J., Leighton, J. A., Gurudu, S., Singh, H., Klint, M., Kunze, K. L., Golafshar, M. A., Esplin, E. D., Nussbaum, R. L., ... Jewel Samadder, N. (2022). Germline Cancer Susceptibility Gene Testing in Unselected Patients With Colorectal Adenocarcinoma: A Multicenter Prospective Study. *Clinical Gastroenterology and Hepatology*, *20*(3), e508–e528. <https://doi.org/10.1016/j.cgh.2021.04.013>
- Uson, P. S., Riegert-Johnson, D., Boardman, L., Borad, M., Ahn, D., Sonbol, M., Leighton, J., Gurudu, S., Kunze, K., Golafshar, M., Esplin, E., Nussbaum, R., Stewart, K., Bekaii-Saab, T., & Samadder, N. J. (2021). Germline cancer susceptibility gene testing in unselected patients with colorectal adenocarcinoma: a multi-center prospective study. *Molecular Genetics and Metabolism*, *132*(2021), S34–S35. [https://doi.org/10.1016/s1096-7192\(21\)00135-9](https://doi.org/10.1016/s1096-7192(21)00135-9)
- Vaccaro, C. A., López-Kostner, F., Adriana, D. V., Palmero, E. I., Rossi, B. M., Antelo, M., Solano, A., Carraro, D. M., Forones, N. M., Bohorquez, M., Lino-Silva, L. S., Buleje, J., Spirandelli, F., Abe-Sandes, K., Nascimento, I., Sullcahuaman, Y., Sarroca, C., Gonzalez, M. L., Herrando, A. I., ... Dominguez-Valentin, M. (2019). From colorectal cancer pattern to the characterization of individuals at risk: Picture for genetic research in Latin America. *International Journal of Cancer*, *145*(2), 318–326. <https://doi.org/10.1002/ijc.31920>
- van de Beek, I., Glykofridis, I. E., Wolthuis, R. M. F., Gille, H. J. J. P., Johannesma, P. C., Meijers-Heijboer, H. E. J., Moorselaar, R. J. A. van, & Houweling, A. C. (2020). No evidence for increased prevalence of colorectal carcinoma in 399 Dutch patients with Birt-Hogg-Dubé syndrome. *British Journal of Cancer*, *122*(4), 590–594. <https://doi.org/10.1038/s41416-019-0693-1>
- Waite, K. A., & Eng, C. (2003). From developmental disorder to heritable cancer: it's all in the BMP/TGF- β family. *Nature Reviews Genetics*, *4*(10), 763–773. <https://doi.org/10.1038/nrg1178>

- Wallace, S. S., Murphy, D. L., & Sweasy, J. B. (2012). Base excision repair and cancer. *Cancer Letters*, *327*(1–2), 73–89. <https://doi.org/10.1016/j.canlet.2011.12.038>
- Whitworth, J., Skytte, A.-B., Sunde, L., Lim, D. H., Arends, M. J., Happerfield, L., Frayling, I. M., van Minkelen, R., Woodward, E. R., Tischkowitz, M. D., & Maher, E. R. (2016). Multilocus Inherited Neoplasia Alleles Syndrome: A Case Series and Review. *JAMA Oncology*, *2*(3), 373. <https://doi.org/10.1001/jamaoncol.2015.4771>
- Williams, S. C., & Parsons, J. L. (2018). NTH1 Is a New Target for Ubiquitylation-Dependent Regulation by TRIM26 Required for the Cellular Response to Oxidative Stress. *Molecular and Cellular Biology*, *38*(12). <https://doi.org/10.1128/MCB.00616-17>
- Wu, D., Huang, H., Chen, T., Gai, X., Li, Q., Wang, C., Yao, J., Liu, Y., Cai, S., & Yu, X. (2023). The BRCA1/BARD1 complex recognizes pre-ribosomal RNA to facilitate homologous recombination. *Cell Discovery*, *9*(1), 99. <https://doi.org/10.1038/s41421-023-00590-8>
- Wu, Y., Mensink, R. G. J., Verlind, E., Sijmons, R. H., Buys, C. H. C. M., Hofstra, R. M. W., Berends, M. J. W., Kleibeuker, J. H., Post, J. G., Kempinga, C., van der Sluis, T., Hollema, H., & van der Zee, A. G. J. (2001). Germline mutations of EXO1 gene in patients with hereditary nonpolyposis colorectal cancer (HNPCC) and atypical HNPCC forms. *Gastroenterology*, *120*(7), 1580–1587. <https://doi.org/10.1053/gast.2001.25117>
- Xiu, M., Wang, Y., Li, B., Wang, X., Xiao, F., Chen, S., Zhang, L., Zhou, B., & Hua, F. (2021). The Role of Notch3 Signaling in Cancer Stemness and Chemoresistance: Molecular Mechanisms and Targeting Strategies. *Frontiers in Molecular Biosciences*, *8*. <https://doi.org/10.3389/fmolb.2021.694141>
- Xu, P., Sun, D., Gao, Y., Jiang, Y., Zhong, M., Zhao, G., Chen, J., Wang, Z., Liu, Q., Hong, J., Chen, H., Chen, Y. X., & Fang, J. Y. (2021). Germline mutations in a DNA repair pathway are associated with familial colorectal cancer. *JCI Insight*, *6*(18), 1–14. <https://doi.org/10.1172/jci.insight.148931>
- You, S., Zhou, J., Chen, S., Zhou, P., Lv, J., Han, X., & Sun, Y. (2010). PTCH1, a receptor of Hedgehog signaling pathway, is correlated with metastatic potential of colorectal cancer. *Uppsala Journal of Medical Sciences*, *115*(3), 169–175. <https://doi.org/10.3109/03009731003668316>
- Yuan, Y., Bao, J., Chen, Z., Villanueva, A. D., Wen, W., Wang, F., Zhao, D., Fu, X., Cai, Q., Long, J., Shu, X. O., Zheng, D., Moreno, V., Zheng, W., Lin, W., & Guo, X. (2021). Multi-omics analysis to identify susceptibility genes for colorectal cancer. *Human Molecular Genetics*, *30*(5), 321–330. <https://doi.org/10.1093/hmg/ddab021>
- Yurgelun, M. B., Kulke, M. H., Fuchs, C. S., Allen, B. A., Uno, H., Hornick, J. L., Ukaegbu, C. I., Brais, L. K., McNamara, P. G., Mayer, R. J., Schrag, D., Meyerhardt, J. A., Ng, K., Kidd, J., Singh, N., Hartman, A. R., Wenstrup, R. J., & Syngal, S. (2017). Cancer susceptibility gene mutations in individuals with colorectal cancer. *Journal of Clinical Oncology*, *35*(10), 1086–1095. <https://doi.org/10.1200/JCO.2016.71.0012>
- Zeng, Z., & Bromberg, Y. (2019). Predicting Functional Effects of Synonymous Variants: A Systematic Review and Perspectives. *Frontiers in Genetics*, *10*. <https://doi.org/10.3389/fgene.2019.00914>
- Zhang, J., Wang, X., de Voer, R. M., Hehir-Kwa, J. Y., Kamping, E. J., Weren, R. D. A., Nelen, M., Hoischen, A., Ligtenberg, M. J. L., Hoogerbrugge, N., Yang, X., Yang, Z., Fan, X., Wang, L., Liu, H., Wang, J., Kuiper, R. P., & van Kessel, A. G. (2017). A molecular inversion probe-based next-generation sequencing panel to detect germline mutations in Chinese early-onset colorectal cancer patients. *Oncotarget*, *8*(15), 24533–24547. <https://doi.org/10.18632/oncotarget.15593>
- Zhang, Y.-Q., Liang, Y.-K., Wu, Y., Chen, M., Chen, W.-L., Li, R.-H., Zeng, Y.-Z., Huang, W.-H., Wu, J.-D., Zeng, D., Gao, W.-L., Chen, C.-F., Lin, H.-Y., Yang, R.-Q., Zhu, J.-W., Liu, W.-L., Bai, J.-W., Wei, M., Wei, X.-L., & Zhang, G.-J. (2021). Notch3 inhibits cell proliferation and tumorigenesis and predicts better prognosis in breast cancer through transactivating PTEN. *Cell Death & Disease*, *12*(6), 502. <https://doi.org/10.1038/s41419-021-03735-3>
- Zhang, Y., He, B.-S., Pan, Y.-Q., Xu, Y.-Q., & Wang, S.-K. (2011). Association of OGG1 Ser326Cys polymorphism with colorectal cancer risk: a meta-analysis. *International Journal of Colorectal Disease*, *26*(12), 1525–1530.

<https://doi.org/10.1007/s00384-011-1258-9>

Zhao, M., Mishra, L., & Deng, C.-X. (2018). The role of TGF- β /SMAD4 signaling in cancer. *International Journal of Biological Sciences*, 14(2), 111–123. <https://doi.org/10.7150/ijbs.23230>

Zhou, X., Zhou, W., Wang, C., Wang, L., Jin, Y., Jia, Z., Liu, Z., & Zheng, B. (2021). A Comprehensive Analysis and Splicing Characterization of Naturally Occurring Synonymous Variants in the ATP7B Gene. *Frontiers in Genetics*, 11. <https://doi.org/10.3389/fgene.2020.592611>

CONTRIBUCIONES ESPECÍFICAS DEL ESTUDIANTE DENTRO DE LA INVESTIGACIÓN LLEVADA A CABO

La concepción del proyecto de investigación	()
El diseño del estudio	(X)
La adquisición de los datos a través de la experimentación	(X)
Análisis e interpretación de los datos	(X)
Elaboración del borrador del artículo	(X)
Revisión y aprobación definitiva de la versión que se presenta	(X)