



Incorporación de recursos computacionales para determinación de la reprogramación de vías metabólicas en cáncer colorrectal

por

Gastón Pini

Desarrollado en CICyTTP-CONICET,
para la obtención del título de Licenciado en Bioinformática,

de la

FACULTAD DE INGENIERÍA
UNIVERSIDAD NACIONAL DE ENTRE RÍOS

Directora de la tesina: Dra. Verónica Martínez Marignac.

Incorporación de recursos computacionales para determinación de la reprogramación de vías metabólicas en cáncer colorrectal

por

Gastón Pini

Resumen

El cáncer colorrectal (CCR) es el segundo cáncer más frecuente en nuestro país, representa el 12,6% del total de casos en ambos sexos, detrás del cáncer de mama (17,2%) y antes del de próstata (9,3%). Respecto a los factores de riesgo que predisponen al desarrollo de CCR, son la dieta y el estilo de vida; todos estos factores producen un efecto directo sobre el metabolismo y la homeostasis celular.

Gracias a los adelantos en Biología Molecular, la genómica y a las distintas “ómicas”, se permiten elucidar la expresión génica y secuencias en ADN y el ARN de más de un millón de variables biológicas. La Bioinformática es aquí donde entra en juego para poder manejar, organizar y analizar con herramientas informáticas y estadísticas, la información biológica compleja acerca de casos de CCR.

Se caracterizó la relevancia de la re-organización metabólica sufrida en tejidos colorrectales a partir del uso de bases de datos públicas de pacientes y controles, y plataformas de análisis de uso libre. Se hizo foco en vías de homeostasis metabólica de la glucosa y otros metabolitos intermediarios con la finalidad de obtener información valorable para futuros tratamientos con antitumorales.

Los datos obtenidos durante el trabajo, reflejaron una alteración en los niveles de expresión de genes relacionados a vías metabólicas de nutrientes, tales como lípidos y glicosaminoglicanos.

En los dos primeros capítulos, se detallan la información, las herramientas y los procedimientos con los que se cuenta y realiza el análisis de niveles de expresiones génicas. En los capítulos posteriores, se detallan y se discuten los resultados obtenidos a partir de estos.

Palabras clave: cáncer colorrectal, expresión génica, metabolismo, bases de datos públicas, herramientas de análisis genético.

Directora de tesina: Dra. Verónica Martínez Marignac.

Cargo: Investigadora Adjunta del CONICET, directora del IBIOGEM, CICYTTP (CONICET, Provincia de Entre Ríos y UADER).

Coordinador: Bioingeniero Carlos Pais.

*Alimenta tu voluntad,
y tus miedos se morirán de hambre.*

Agradecimientos

Quiero expresar mi gratitud, en primer lugar, a mi madre, a mi abuela Martha, y a Fito, por el constante apoyo y confianza en mí en todo momento, desde el inicio hasta la finalización de mis estudios académicos.

A la Facultad de Ingeniería, por haber hecho posible los estudios de la carrera de Licenciatura en Bioinformática.

A la directora Verónica Martínez Marignac, por la dirección y el asesoramiento en el aspecto biológico de los datos durante la realización de este proyecto.

Al Instituto CICYTTP, por concederme estadía en las instancias finales de la realización del proyecto.

A mi familia y mis incondicionales amigos, que han estado en las buenas y en las malas, siempre cuando más lo necesité.

Esta tesis de grado para la obtención del título de Licenciado en Bioinformática fue evaluada por el siguiente comité:

Dr. Rodrigo Gonzalo Parra. Investigador en European Molecular Biology Laboratory (EMBL), Heidelberg, Germany.

Dra. Rosana Ramírez. Investigadora Adjunta del CONICET – CCT – Santa Fe.

Índice

1. Introducción.....	7
1.1. Marco teórico.....	7
1.2. Antecedentes.....	7
1.3. Resultados asociados a metabolismo de CCR a través de la Bioinformática.....	8
1.4. Objetivos.....	9
1.4.1. General.....	9
1.4.2. Particulares.....	9
1.5. Herramientas bioinformáticas utilizadas para el estudio de CCR.....	9
1.5.1. Herramientas bioinformáticas de uso privativo.....	9
2. Materiales y métodos.....	10
2.1. Materiales.....	10
2.1.1. Equipos.....	10
2.1.2. ArrayExpress.....	10
2.1.3. Lenguaje de programación.....	11
2.1.3.1. Lenguaje R.....	11
2.1.3.2. Lenguaje Perl.....	11
2.1.4. DAVID.....	11
2.1.5. MeV4.....	12
2.1.6. Reactome.....	12
2.2. Métodos.....	13
2.2.1. Normalización y obtención de datos de estudio.....	13
2.2.2. Selección de muestras.....	13
2.2.3. Filtrado de sondas por error estándar.....	13
2.2.4. Filtrado por relevancia biológica - especie <i>Homo sapiens</i>	14
2.2.5. MeV v4.8.....	14
2.2.5.1. Sobreexpresión de sondas en poblaciones.....	14
2.2.5.2. Conversión de sondas a genes y análisis génico.....	14
2.2.6. Reactome: vías asociadas a genes.....	15
2.2.6.1. Vías metabólicas asociadas a sobreexpresiones.....	15
3. Resultados.....	16
3.1. Sondas seleccionadas por relevancia estadística y biológica.....	16
3.2. Sobreexpresiones mediante heatmaps.....	16
3.3. Vías activas en pacientes sanos y tumorales.....	21
3.4. Vías metabólicas activas por sobreexpresiones de genes.....	24
4. Discusión.....	30
4.1. Sobreexpresión de genes.....	30
4.2. Reactome: distinción de vías metabólicas y sus diferencias de sobreexpresión.....	31
4.2.1. Vías metabólicas sobreexpresadas en sondas vs genes.....	31
4.2.2. Conocimientos previos asociados al metabolismo de CCR.....	32
4.3. Uso en conjunto de herramientas de uso público.....	32
5. Conclusiones y trabajos futuros.....	33
Anexo: Material suplementario.....	34
Bibliografía.....	35

1. Introducción

1.1. Marco teórico

En el presente proyecto se plantea evidenciar la re-organización metabólica, a partir de diferencias en la expresión de genes involucrados con vías metabólicas, sufrida en tejido colorrectal a partir del uso de bases de datos públicas de pacientes y controles, y los softwares de análisis de uso libre DAVID y Mev4. Se hará foco en vías de homeostasis metabólica de la glucosa y otros metabolitos intermediarios, así como genes y vías blancos afectados por 5-fluorouracilo (5-FU).

En estudios previos de los proyectos PID UNER 6155 y 6176 [1, 2], se ha logrado la caracterización de las diferentes etapas del desarrollo de CCR en ratones utilizando técnicas de anatomía patológica clásicas, e inmunohistoquímica así como de daño al ADN por medio de radiaciones ionizantes, demostrando que poseen un efecto en vías de glutamina, glucógeno y carbohidratos [3]. Estos datos morfológicos y bioquímicos en modelo murino pueden ser correlacionados con la organización bioenergética de los colonocitos en pacientes y controles humanos. Sin embargo, actualmente existe una mayor cantidad de datos de expresión genica en bases de datos públicas que podría evidenciar posibles cambios metabólicos entre tejidos normales y tumorales, que permitan proponer nuevos blancos terapéuticos y conocimientos en el desarrollo tumoral y el papel de genes asociados a homeostasis celular [4]. Una forma de acercamiento a este intento de correlación, es a través de la implementación de recursos informáticos de análisis de datos de expresión génica y de niveles de expresión de vías metabólicas específicas, incluyendo la de carbohidratos y reparación del ADN. Todo esto resulta clave para establecer los mecanismos por el cual los cambios metabólicos actuarían y/o promoverían el desarrollo del CCR, así como posibles respuestas a anticancerígenos, y guiar en estudios futuros de mesada.

1.2. Antecedentes

El CCR es el segundo cáncer más frecuente en nuestro país, representando el 12,6% del total de casos en ambos sexos, detrás del cáncer de mama (17,2%) y antes del de próstata (9,3%). Según sexo, el CCR es segundo en mayor incidencia tanto en hombres (detrás del cáncer de próstata), como en mujeres (detrás del cáncer de mama).

Según los datos de incidencia de la Agencia Internacional de Investigación en Cáncer (IARC), en el 2008 han habido 11.000 nuevos casos de CCR (CCR), y más estimaciones indican que se han producido 13.558 casos en 2012, de los cuales 7.237 fueron hombres y 6.321 mujeres, con una tasa ajustada de 29,9 x 100.000 habitantes para hombres y de 19 x 100.000 habitantes para mujeres.

Los datos de mortalidad del Ministerio de Salud de la Nación (<https://www.argentina.gob.ar/salud/instituto-nacional-del-cancer/estadisticas/mortalidad>) muestran que en 2017, fallecieron 7.499 personas por esta causa, lo que corresponde al 27,1% del total de las muertes por tumores en la Argentina.

En el año 2018, han habido 15.692 nuevos casos de CCR, representando un 12,6% del total de tumores. De estos casos, 8.527 fueron hombres y 7.165 mujeres, con una tasa ajustada de 31 x 100.000 habitantes para hombres (<https://www.argentina.gob.ar/salud/instituto-nacional-del-cancer/estadisticas/incidencia>).

El 90% de los cánceres colorrectales se producen en personas mayores de 50 años de edad y tiene una incidencia levemente mayor en hombres que en mujeres. Alrededor del 75% de los casos de CCR son esporádicos, es decir que se desarrollan en personas que no presentan antecedentes personales ni familiares demostrados de la enfermedad y por lo tanto los factores dietéticos y

ambientales estarían implicados en su causa. El resto de los cánceres colorrectales (alrededor del 25%) se producen en personas con riesgo adicional debido a las siguientes situaciones: antecedente personal de adenomas únicos o múltiples, CCR o enfermedad inflamatoria intestinal, y antecedentes familiares de CCR o de adenomas únicos o múltiples [5].

La patología es detectada en los pacientes una vez los adenomas ya están transformados en adenocarcinomas, sus contrapartidas malignas.

Respecto a los factores de riesgo que predisponen al desarrollo de CCR, son la dieta y el estilo de vida [6]; estos factores producen un efecto directo sobre el metabolismo y la homeostasis celular. Por consiguiente, se sugiere que las células tumorales, en sus diferentes estadios de desarrollo, presentarán una reorganización metabólica distintiva [4]. Se conoce que las células cancerígenas poseen mayores requerimientos metabólicos que sus contrapartes normales [7]. Igualmente, diferentes cánceres o tumores responden metabólicamente en forma diferencial al estrés originado por tratamientos anticancerígenos y dicha respuesta será adaptativa para restablecer la homeostasis celular. Los cambios metabólicos en células cancerígenas en respuesta a terapias, pueden consistir en uno de los siguientes mecanismos: aumento de la glucólisis anaeróbica (efecto Warburg), de la glutaminólisis y de la fosforilación oxidativa, o la revitalización de la ruta de la pentosa fosfato [8].

Recientemente, se han reconsiderado los cambios metabólicos sufridos en distintos tipos de cánceres [9, 10, 11]. Las transformaciones metabólicas que sufren las células normales conforme se tornan cancerígenas fueron ignoradas por más de 50 años y actualmente se ha tomado conciencia que los cambios bioenergéticos resultan centrales para comprender los mecanismos que llevan al desarrollo y proliferación de células cancerosas [12]. Una teoría metabólica del cáncer versus a un origen genético del cáncer no fue retomada hasta los años 90' con el hallazgo del factor inducible por hipoxia (Hif1). Posteriormente, con el importante desarrollo de los conocimientos en genómica y las diferentes "ómicas", se evidenció que el factor genético del cáncer no es el principal desencadenante de la enfermedad, como se había propuesto anteriormente, ya que hasta en un 25% de los casos se deben a mutaciones heredadas y el resto presentan mutaciones de tipo esporádico y poco frecuentes [13].

1.3. Resultados asociados a metabolismo de CCR a través de la Bioinformática

Un estudio de CCR a nivel bioinformático realizado en una población de pacientes atendidos en la Provincia de Buenos Aires en los Hospitales Mendez y Martínez en los años 2003 y 2004, demuestra diferencias entre tejidos normales y tumorales asociados al inositol/fosfatidilinositol, apoptosis, lípidos y mitocondrias [14]. En este estudio, se compararon los tejidos cancerosos con la mucosa colónica normal. Realizaron una micromatriz de ADN-copia para evaluar la expresión de aproximadamente 19,200 genes, contenidos en la micromatriz SS-Humana 19Kv7 (University Health Network Microarray Center, Ontario, Canadá), derivada de la codificación y las secuencias de EST depositadas en GenBank. En 25 carcinomas colorrectales esporádicos, identificaron 584 genes conocidos expresados diferencialmente, siendo en su mayoría inhibidores apoptóticos (por ejemplo, BFAR, BIRC1, BIRC6).

La amplia variación geográfica en las tasas de incidencia del CCR y los datos estudiados, sugieren que los factores de riesgo del estilo de vida, como la dieta, la actividad física, la obesidad y la diabetes, desempeñan un papel fundamental en la etiología de la enfermedad [15]. El cáncer en consecuencia, sostiene un daño sobre el metabolismo de las mitocondrias, resultado tanto de una alteración nutricional como del impacto de los tóxicos sobre ellas. Estos factores afectan la generación de la energía metabólica. La hipoxia en el tumor hace que el sustrato energético utilice la glucosa de forma anaeróbica; en otros casos, se ha reportado que la glutamina y la vía de las mitocondrias, orgánulos celulares encargados de suministrar la mayor parte de la energía necesaria

para la actividad celular. Las mitocondrias como centrales energéticas de la célula y sintetizan ATP a expensas de los carburantes metabólicos, como la glucosa, ácidos grasos y aminoácidos.

La falta de oxígeno, asimismo, es una condición común de los tumores sólidos que determina la forma en la que se genera la energía metabólica. La hipoxia en el tumor no solo genera que el sustrato energético utilizado sea la glucosa de forma anaeróbica; sino que además en otros casos se ha reportado que la glutamina y la vía de pentosa fosfato [16] son preferenciales frente al resto de sustratos (ácidos grasos, cuerpos cetónicos, carbohidratos, lactato y aminoácidos como la glutamina) cuyo metabolismo es también dependiente del oxígeno.

1.4. Objetivos

1.4.1. General

Determinar el valor de la reprogramación de vías metabólicas en pacientes de CCR mediante recursos computacionales virtuales.

1.4.2. Particulares

- Hacer uso de bases de datos curadas y normalizadas de expresión génica.
- Acercamiento A: Determinar la re-organización metabólica por medio del análisis bioinformático de la expresión de sondas que representan genes en muestras de pacientes con estadios I al IV de CCR y en muestras controles.
- Acercamiento B: Evaluar la expresión de genes asociados al metabolismo del desarrollo tumoral y la homeostasis metabólica (carbohidratos, glucogenolisis y glutaminólisis, lípidos y ácidos grasos).
- Valorar las herramientas de uso público y gratuito para establecer líneas de estudios en cáncer colorrectal.

1.5. Herramientas bioinformáticas utilizadas para el estudio de CCR

1.5.1. Herramientas bioinformáticas de uso privativo

Existen herramientas que brindan información a partir de las grandes bases de datos de expresiones génicas, tales como Ingenuity Pathway Analysis, un software que toma datos referidos a las distintas "ómicas" para identificar nuevos blancos terapéuticos o potenciales biomarcadores en el contexto de vías biológicas, dando un acceso a los sistemas públicos de datos (<https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis/>).

Otro software de uso privado, es Invitrogen VectorNTI, el cual posee un amplio conjunto de herramientas de análisis de secuencias y diseño para administrar, ver, analizar, transformar, compartir y publicar diversos tipos de datos de biología molecular, en un entorno de análisis rico en gráficos (<https://www.thermofisher.com/ar/en/home/life-science/cloning/vector-nti-software.html>).

Como estos programas son de uso privado, a partir de licencias pagas, se hace imperante lograr el acceso libre a los datos y sus análisis para biólogos y científicos; en este marco, recurrir a herramientas con funciones semejantes, como las detalladas más arriba, son las intenciones de este trabajo, detalladas más adelante.

2. Materiales y métodos

2.1. Materiales

2.1.1. Equipos

El trabajo con la directora se realizó inicialmente a distancia a través de Internet, para luego seguir personalmente en la ciudad de Diamante. Se hizo uso de la computadora propia del estudiante y del laboratorio interdisciplinario de Biología y Genética Molecular (IBIOGEM), en el CICYTTP (CONICET, Provincia de Entre Ríos y UADER). Los softwares son de acceso público y libres, como asimismo las bases de datos. El uso de internet y de servidores fueron accesibles en el IBIOGEM. Se considera que todos los recursos informáticos y bibliografía se encuentran accesibles para realizar este proyecto en el CICYTTP. Se hizo uso de bibliografías de uso restringido, accedidos por medio del CICYTTP y su registro con la Secretaría de Ciencia, Tecnología e Innovación de la Nación.

El titular de este trabajo, también usó un computador portátil de 12 GB de RAM, con un procesador Intel(R) Core(TM) i7-7500U CPU @ 2.70GHz. Este portátil posee el sistema operativo Linux Ubuntu 18.04 (Bionic Beaver), una versión LTS lanzada el 26 de abril de 2018, con soporte de diez años. Esta versión de Ubuntu trae el Kernel 4.15 el cual se ha reescrito con algunos parches para optimizar el funcionamiento del sistema, y se colocó a Xorg como servidor principal.

Linux Ubuntu es el más sistema operativo más apropiado para la realización del presente proyecto, al consumir una menor cantidad de recursos inherentes del sistema operativo en segundo plano.

2.1.2. ArrayExpress

ArrayExpress (<https://www.ebi.ac.uk/arrayexpress/>) es una base de datos pública para datos genómicos funcionales de alto rendimiento. ArrayExpress consta de dos partes: el Repositorio, que es un archivo público de apoyo de información mínima sobre experimentos de microarreglos (MIAME) de datos de micromatrices, y el Almacén de datos, que es una base de datos de perfiles de expresión génica seleccionados del repositorio y re-anotados consistentemente. Los experimentos archivados pueden consultarse por atributos de experimentos, como palabras clave, especies, plataforma de matriz, autores, revistas o números de acceso. Actualmente, ArrayExpress contiene datos de más de 50.000 hibridaciones y más de 1.500.000 perfiles de expresión individuales.

El microarreglo seleccionado para este estudio es E-GEOD-50421 (Exon Array Data for Colorectal Tumor and Normal Samples), que contiene 49 muestras tanto de pacientes normales como de pacientes con los distintos niveles tumorales (Tabla 1). Su selección se basó en adenocarcinomas de colon como claves de búsqueda.

Como conocimiento adicional, se realizó un curso de entrenamiento provisto por el sitio de EMBL-EBI para tener un mejor entendimiento del manejo de microarreglos (<https://www.ebi.ac.uk/training/online/>).

Muestra	Etap	Edad	Sexo	Característica	Muestra	Etap	Edad	Sexo	Característica
GSM1218927 1	0	52	masculino	Adyacente normal	GSM1218951 1	I	74	masculino	Tumor
GSM1218926 1	0	74	masculino	Adyacente normal	GSM1218950 1	II	57	masculino	Tumor
GSM1218925 1	0	49	femenino	Adyacente normal	GSM1218949 1	IV	49	femenino	Tumor
GSM1218924 1	0	71	masculino	Adyacente normal	GSM1218948 1	II	52	masculino	Tumor
GSM1218923 1	0	57	masculino	Adyacente normal	GSM1218947 1	IV	63	femenino	Tumor
GSM1218922 1	0	63	femenino	Adyacente normal	GSM1218946 1	IV	76	femenino	Tumor
GSM1218921 1	0	76	femenino	Adyacente normal	GSM1218945 1	I	61	femenino	Tumor
GSM1218920 1	0	61	femenino	Adyacente normal	GSM1218944 1	IV	60	femenino	Tumor
GSM1218919 1	0	63	femenino	Adyacente normal	GSM1218943 1	IV	78	masculino	Tumor
GSM1218918 1	0	52	femenino	Adyacente normal	GSM1218942 1	I	70	masculino	Tumor
GSM1218917 1	0	28	femenino	Adyacente normal	GSM1218941 1	III	69	femenino	Tumor
GSM1218916 1	0	52	femenino	Adyacente normal	GSM1218940 1	II	56	masculino	Tumor
GSM1218915 1	0	66	masculino	Adyacente normal	GSM1218939 1	II	56	femenino	Tumor
GSM1218914 1	0	52	femenino	Adyacente normal	GSM1218938 1	IV	44	masculino	Tumor
GSM1218913 1	0	64	masculino	Adyacente normal	GSM1218937 1	II	84	masculino	Tumor
GSM1218912 1	0	97	femenino	Adyacente normal	GSM1218936 1	IIIB	61	femenino	Tumor
GSM1218911 1	0	69	femenino	Adyacente normal	GSM1218935 1	IIB	52	masculino	Tumor
GSM1218910 1	0	56	masculino	Adyacente normal	GSM1218934 1	IV	71	masculino	Tumor
GSM1218909 1	0	68	masculino	Adyacente normal	GSM1218933 1	IV	28	femenino	Tumor
GSM1218908 1	0	56	femenino	Adyacente normal	GSM1218932 1	III	66	masculino	Tumor
GSM1218907 1	0	56	femenino	Adyacente normal	GSM1218931 1	II	64	masculino	Tumor
GSM1218906 1	0	61	femenino	Adyacente normal	GSM1218930 1	II	56	femenino	Tumor
GSM1218905 1	0	44	masculino	Adyacente normal	GSM1218929 1	IV	61	femenino	Tumor
GSM1218904 1	0	84	masculino	Adyacente normal	GSM1218928 1	IV	49	femenino	Tumor
GSM1218903 1	0	61	femenino	Adyacente normal					

Tabla 1. Especificaciones de los pacientes a los que se les tomaron las muestras del microarreglo E-GEOD-50421.

2.1.3. Lenguaje de programación

Para el procesamiento y análisis de datos a utilizar, se hará uso de dos lenguajes de programación orientados a objetos: R y Perl.

2.1.3.1. Lenguaje R

El lenguaje R (<https://www.r-project.org/>) es un lenguaje enfocado al análisis estadístico. Es muy utilizado en la investigación científica, contribuyendo a la posibilidad de cargar diferentes bibliotecas y paquetes con funcionalidades de cálculo y gráfico.

Se hará uso de comandos en este lenguaje para descargar el microarreglo de estudio, por medio del paquete “ArrayExpress” de Bioconductor Release (3.11).

R version 4.0.2 (2020-06-22) – “Taking Off Again” (<https://cran.r-project.org/src/base/R-4/>).

2.1.3.2. Lenguaje Perl

Perl es un lenguaje imperativo basado en un estilo de bloques como el de C, que utiliza variables, expresiones, asignaciones, bloques de código delimitados por llaves, estructuras de control y subrutinas.

Información de Perl utilizada: perl 5, version 26, subversion 1 (v5.26.1) built for x86_64-linux-gnu-thread-multi (<http://www.cpan.org/src/>).

2.1.4. DAVID

DAVID (<https://david.ncifcrf.gov/>) consiste en una base de conocimientos biológicos integrados y herramientas analíticas destinadas a extraer sistemáticamente el significado biológico de grandes listas de genes y proteínas, capaz de convertir los identificadores de producto o gen de

entrada del usuario de un tipo a otro de una manera más completa y de alto rendimiento con una base de datos de mapeo ID-ID mejorada de forma única.

DAVID hace uso de Gene Ontology (GO), una herramienta bioinformática que proporciona anotaciones de lotes típicas y análisis de enriquecimiento para resaltar los términos más relevantes asociados con una lista de genes determinada. GO es una iniciativa de la Bioinformática que unifica la representación de los atributos de los genes y productos genéticos de todas las especies.

DAVID cuenta con el módulo Gene ID Conversion Tool, basado en evidencia de teoría de grafos para aglomerar los IDs de genes/proteínas específicos de especies en los repositorios de acceso públicos más populares, tales como NCBI, PIR y Uniprot/SwissProt.

2.1.5. MeV4

MultiExperiment Viewer (<http://mev.tm4.org/>) es un software enfocado en la integración y análisis de datos oncogenómicos clinicopatológicos, que permite la visualización de microarreglos procesados, y la identificación de genes y patrones de expresión de interés. La variedad de algoritmos de normalización y análisis de agrupamiento, permiten flexibilidad para crear vistas significativas de los datos de expresión.

2.1.6. Reactome

Reactome (<http://www.reactome.org/>) es una base de datos curada de vías y reacciones biológicas. Los datos anotados de Reactome describen reacciones posibles de todas las proteínas anotadas y moléculas pequeñas. Al superponer un conjunto de datos experimental en estas anotaciones, es posible realizar un análisis de representación excesiva de la ruta y su nivel asociación a la patología o tratamiento anotados.

Reactome devuelve para una lista de IDs, una serie de vías ordenadas por sus p-values, que representan la significancia estadística de una vía para el conjunto de datos dado. En Reactome, el método utilizado para calcular un p-value es la prueba binomial, y corregido con FDR, por el procedimiento Benjamini–Hochberg.

2.2. Métodos

2.2.1. Normalización y obtención de datos de estudio

La normalización de la expresión, desarrollada por las bases de datos públicas y centros generadores de datos de interés, se aplica en todos los casos que se desea evidenciar diferencias relativas en expresión génica. En el caso de tecnologías de inmunología, se utilizaron genes denominados de mantenimiento o constitutivos, los cuales tienen niveles de expresión constantes en todas las células y organismos. En los análisis de expresión a través de microarreglos, esta normalización resulta deficiente, es por eso, que se realizó una normalización mediante modelos lineales entre muestras por parte del paquete de Bioconductor en R llamado LIMMA (<https://bioconductor.org/packages/release/bioc/html/limma.html>).

Para la descarga de archivos con los valores de expresión a estudiar, se hizo uso de los paquetes “BiocManager” y “Arrayexpress” mediante los siguientes comandos:

```
> install.packages("BiocManager")
> BiocManager::install("ArrayExpress")
> library("ArrayExpress")
> GEOD50421 = getAE("GSE50421", type = "full")
> unzip("E-GEOD-50421.processed.1.zip")
> unzip("E-GEOD-50421.processed.2.zip")
> unzip("E-GEOD-50421.processed.3.zip")
```

2.2.2. Selección de muestras

Con la finalidad de mantener una homogeneidad metabólica, se seleccionaron 22 de origen masculino, con un rango etario de entre 44 y 84 años de edad, consistente con el aumento de incidencia a partir de los 50 años (Tabla 2).

Muestra	Edad	Etapas	Muestra	Edad	Etapas
GSM1218904	84	0	GSM1218951	74	I
GSM1218905	44	0	GSM1218935	52	II
GSM1218909	68	0	GSM1218931	64	II
GSM1218910	56	0	GSM1218937	84	II
GSM1218913	64	0	GSM1218940	56	II
GSM1218915	66	0	GSM1218948	52	II
GSM1218923	57	0	GSM1218950	57	II
GSM1218924	71	0	GSM1218932	66	III
GSM1218926	74	0	GSM1218934	71	IV
GSM1218927	52	0	GSM1218938	44	IV
			GSM1218943	78	IV

Tabla 2. Muestras de estudio.

2.2.3. Filtrado de sondas por error estándar

Cada muestra del microarreglo, posee 1.398.555 sondas analizadas. Este conjunto de sondas representa la expresión génica de cada paciente en su respectivo estadio tumoral. Una parte de estas expresiones, están relacionadas a vías metabólicas, las cuales son objetos de interés en este proyecto.

El error estándar (SE) es una medida de estadística que cuantifica el error de estimación de muestras, y biológicamente a mayor variancia, menor probabilidad que la expresión del gen

estuviese asociada a la categorización empleada (sano, tumoral I, II, III y IV). A nivel biológico, un $SE \leq 20\%$ es admisible, indicando que las sondas por debajo de este umbral, son representativas y homogéneamente descriptivas de cada estado bioquímico y genético.

Se generó un Perl script (material suplementario 1) para depurar las sondas que entre todas las muestras de la población sana, tuvieran un SE menor al 20%.

2.2.4. Filtrado por relevancia biológica - especie *Homo sapiens*

Con el módulo Gene ID Conversion Tool, se obtendrán las características de las sondas representativas. Como algunas sondas del microarreglo, son caracterizadas como IDs correspondientes a más de una especie, se depurará por medio de un Perl script (material suplementario 2), todas las sondas serán específicas a *Homo sapiens*.

Para posteriores estudios, para comparar con la población sana, se agruparán las muestras de un mismo nivel tumoral, utilizando sus valores medios como representativos de cada estadio, por medio de un Perl script (material suplementario 3).

2.2.5. MeV v4.8

Una vez depuradas las sondas pertenecientes a *Homo sapiens*, se convertirá el archivo con los IDs de sondas y valores de expresión a una estructura legible (material suplementario 4) para representar los valores individuales por medio matrices de colores llamadas heatmaps. Estos colores, definidos en base a $\log_2(A/B)$ de los valores numéricos entre ambas poblaciones, indican el nivel de expresión de cada sonda, ayudando a discernir visualmente las expresiones diferenciales.

2.2.5.1. Sobreexpresión de sondas en poblaciones

El fold change, es una medida que describe cuánto cambia en expresión una dada sonda o gen entre un grupo y otro. Se llama así porque se describe un aumento de X veces de un "aumento de X-fold". A medida que el valor del delta es mayor, se reduce la cantidad de sondas, caracterizando las diferencias génicas.

2.2.5.2. Conversión de sondas a genes y análisis génico

Por medio de Perl scripts (material suplementario 5), se convertirán los IDs de las sondas sobreexpresadas a los IDs de sus correspondientes genes. Para cada gen, se usarán los valores medios de la sumatoria de los valores de sus correspondientes sondas (material suplementario 6-7), para así dilucidar los genes sobreexpresados característicos, tanto en población sana, como en cada uno de los estadios tumorales (material suplementario 8-9).

Un detalle a aclarar, es que, dado que las sondas en el presente microarreglo, son partes de secuencias de determinados genes, y poseen variabilidad de expresión dentro de genes no representativos de la totalidad génica, hayan posibles errores de determinación de la expresión génica, o diferencias de especificidad por mutaciones en las secuencias, lo que afectará a la relevancia biológica de la información generada [17]. Esto se debe a que no se poseen las secuencias exactas de los genes sobreexpresados.

2.2.6. Reactome: vías asociadas a genes

Dentro del entorno de Reactome, se hará uso del módulo Analyze Data, ingresando los IDs de los genes sobreexpresados. A partir de los resultados generados, se obtendrán las vías activadas por los conjuntos de genes ingresados. Reactome también genera diagramas de las vías asociadas.

2.2.6.1. Vías metabólicas asociadas a sobreexpresiones

Reactome clasifica las vías en forma de diagramas de mejorados de alto nivel (EHL D), en los que se representan las vías muy complejas, como conjuntos de subvías. A partir de los resultados brindados por Reactome sobre los genes, se dilucidarán las vías (material suplementario 10), genes (material suplementario 11) y sondas asociadas (material suplementario 12-13-14) al metabolismo de nutrientes y metabolitos intermediarios, y sus impactos en la re-organización metabólica. El valor de p-value considerado como estadísticamente relevante, será de $p\text{-value} \leq 0,05$.

Acto seguido, se analizará el conjunto de sondas, sin haber convertido los IDs a genes, para determinar su relación con el metabolismo (material suplementario 15-16-17), y hacer una comparación metabólica entre ambos ensayos.

3. Resultados

3.1. Sondas seleccionadas por relevancia estadística y biológica

De las 1.398.555 sondas iniciales, se seleccionaron 1.201.578 con $SE \leq 20\%$ dentro del grupo control (material suplementario 18). Luego, se recurrió a DAVID para caracterizar estas sondas, obteniendo un total de 820.794. Las 820.794 sondas caracterizadas con DAVID, fueron filtradas por especie, dando 569.593 correspondientes a *Homo sapiens* y no conservadas en otras especies (Fig 1).

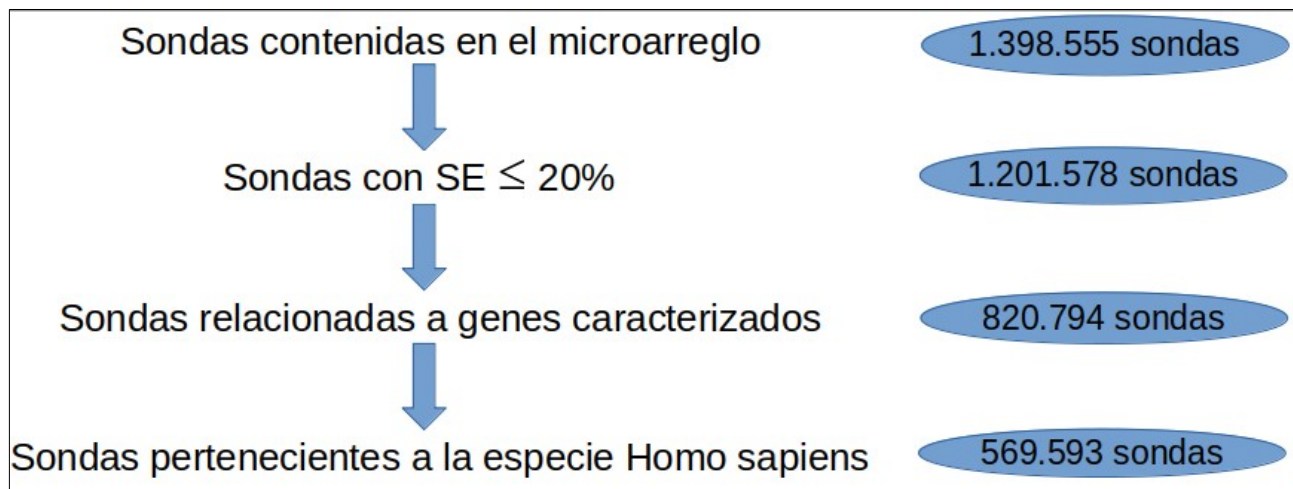


Fig 1. Esquema del procesamiento de sondas del proyecto E-GEOD-50421.

3.2. Sobreexpresiones mediante heatmaps

A partir de las 569.593 sondas de *Homo sapiens* (material suplementario 19), se pudo denotar una clara diferenciación de expresión entre poblaciones sanas y tumorales (Fig 2, Tabla 3), por medio del análisis de heatmaps con MeV (Fig 3). A diferencias de x3,5, no resultaron mayores valores de sobreexpresiones por parte de ninguna de las poblaciones, por lo que se presentan las comparaciones entre tejidos con deltas de x1,5, x2 y x2,5.

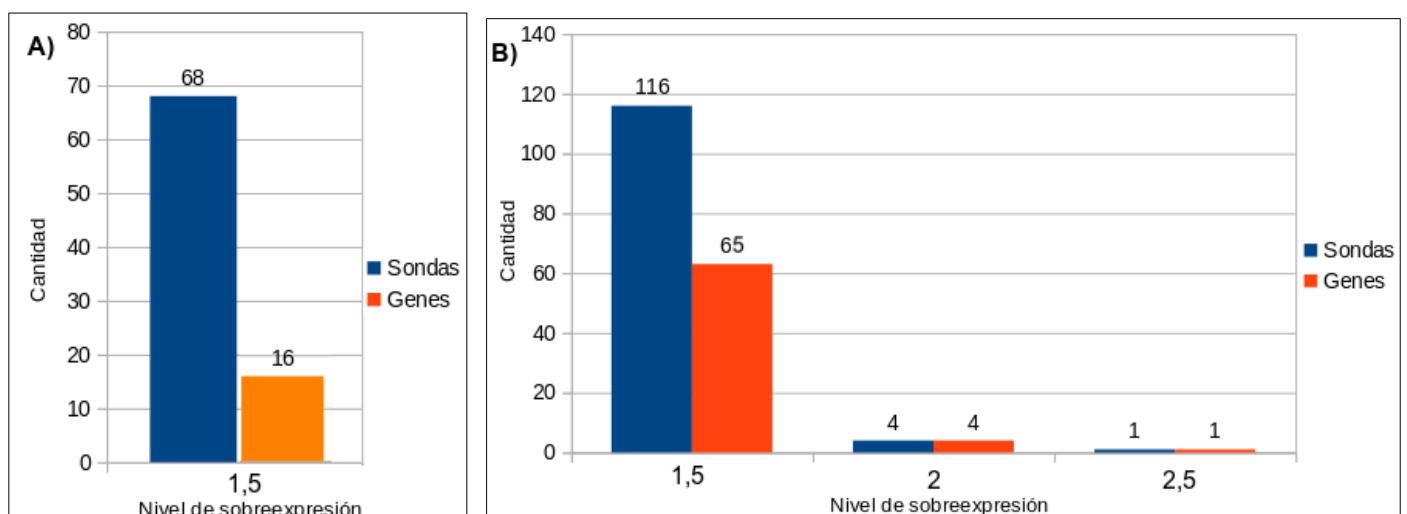


Fig 2. Cantidades de sondas y genes sobreexpresados en total de población con tumor. A) En población sana. B) En población tumoral. A medida que aumenta la tasa de cambio, se obtienen genes que representan específicamente cada población.

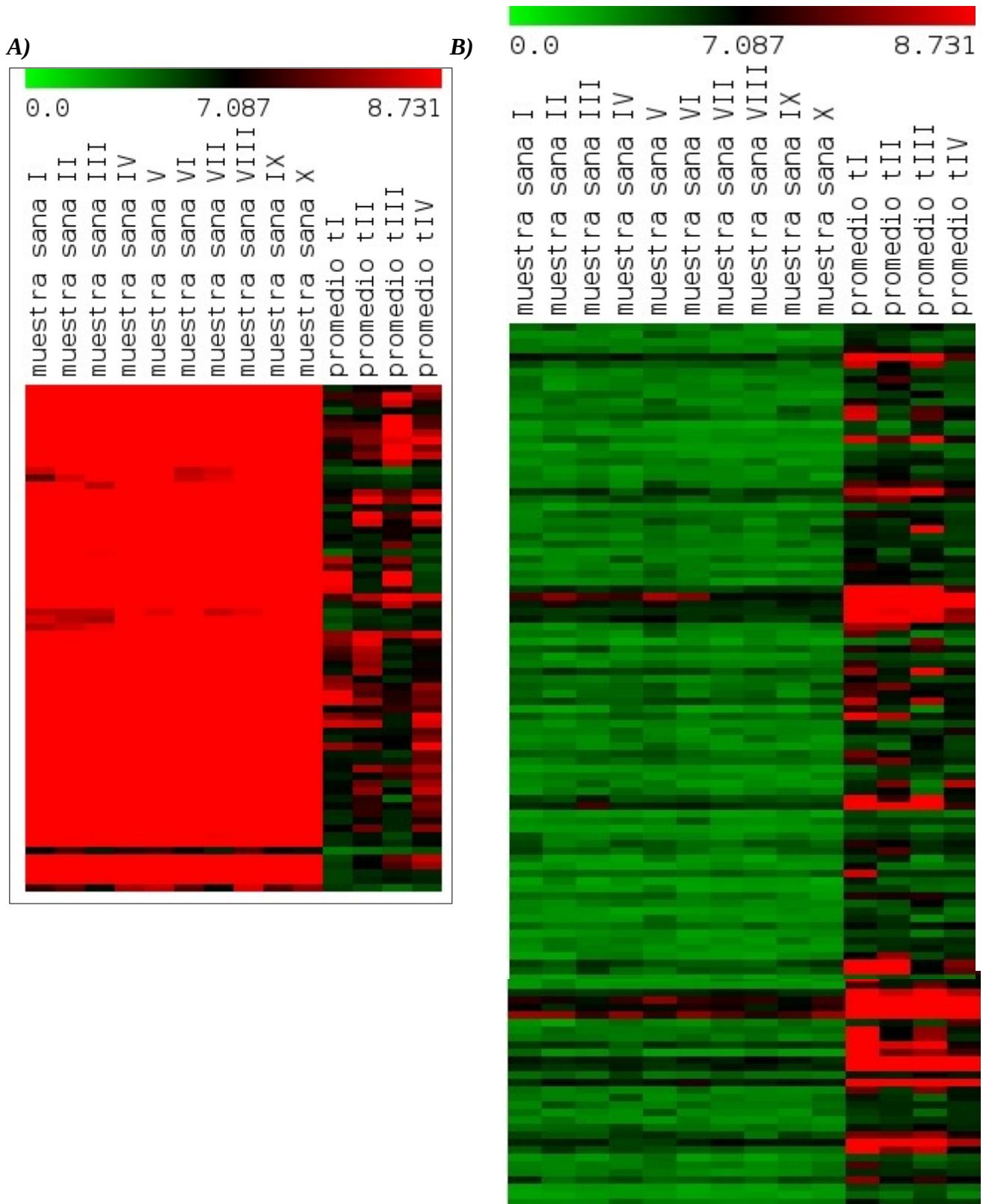


Fig 3. Heatmaps de sondas sobreexpresadas. A) Sobreexpresiones de hasta 1,5 veces en población sana. B) Sobreexpresiones de hasta 1,5 veces en población tumoral. Los gradientes de intensidad de expresión se representan en color verde para subexpresión, y en rojo para sobreexpresión, en base a la relación $\log_2(X/Y)$.

Sondas que representan genes sobreexpresados en población sana:

Los genes sobreexpresados en población sana (Fig 4A), codifican proteínas asociadas a la homeostasis celular general, transporte de moléculas y funcionamiento sanguíneo:

- Las anhidrasas carbónicas (genes CA1, CA2 y CA4), unas metaloenzimas que catalizan la conversión de CO₂ y agua a bicarbonato y protones. Las alteraciones de sus expresiones, son frecuentes en varios cánceres, incluyendo leucemia, adenocarcinomas y CCR [18].

- La beta-1,3-galactosiltransferasa (gen B3GALT1), involucrada en la biosíntesis de glucolípidos y glucoproteínas [19].

- La proteína reguladora de TFPI (gen ADTRP), que hidroliza ácidos grasos ésteres, y regula la expresión del inhibidor de la ruta del factor tisular [20, 21].

- Péptido YY (gen PYY), inhibidor de la motilidad gástrica, e incrementa la absorción de agua y electrolitos en el colon, factor importante en el mantenimiento de la homeostasis celular [22].

- Miembro 1 de neurexofilinas (genes NXPE), una familia de glucoproteínas que interactúan con las neurexinas, receptores premembrana transmembrana [23].

- El gen FER1L6 es un gen que tiene actividad ATPasa, asociada al movimiento transmembrana de sustancias [24].

- Proteína MS4A12 (gen MS4A12), que se sitúa en la superficie de la región apical de los colonocitos, siendo importante en la vía de EGFR [25].

- Dominio contenedor de inmunoglobulinas 2 (gen VSIG2) [26].

- Endotelina 3 (gen EDN3), miembro de la familia endotelina, encargadas del funcionamiento de venas sanguíneas, producción de hormonas y estimulación de crecimiento celular. [27].

Sondas que representan genes sobreexpresados una vez y media en población tumoral:

Dentro de los genes sobreexpresados una vez y media más en población tumoral (Fig 4B-1 y B-2), se destacan marcadores conocidos de CCR. Entre estos, están:

- El gen ANLN, asociado a retinoblastoma en cánceres [28]. La sobreexpresión de los genes AJUBA y GRIN2D, características de cánceres, incluido el CCR [29, 30].

- El gen NPSR1, cuya sobreexpresión promueve la activación de vías asociadas al CCR [31].

- Proteína NrCAM (gen NRCAM), un blanco de señalización de beta-cateninas en adenocarcinomas del colon, promoviendo el crecimiento y la motilidad [32].

- Los genes PSAT1, DIAPH3, KIF4A, ACSL1 y ACSL4 son marcadores de pronóstico en tejidos tumorales de CCR [33, 34, 35, 36, 37].

Sondas que representan genes sobreexpresados dos veces en población tumoral:

En este nivel de expresión (Fig 4C), se ven genes cuyas alteraciones de expresión tienen repercusiones en la homeostasis metabólica:

- El gen REG1A, asociado a AC19-9, conocido marcador de cáncer de páncreas y en CCR [38, 39, 40].

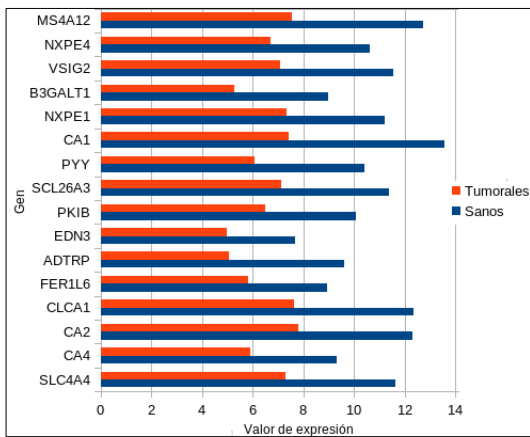
- El gen CPNE7 codifica la proteína copina-7, relacionada al tráfico de calcio intracelular y metabolismo lipídico celular.

- El producto del gen CEMIP, la proteína inductora de migración celular y unión a ácido hialurónico, está asociado a migración celular y a la vía Wnt, conocida por ser activada en CCR, interactuando con b-catenina [41, 42].

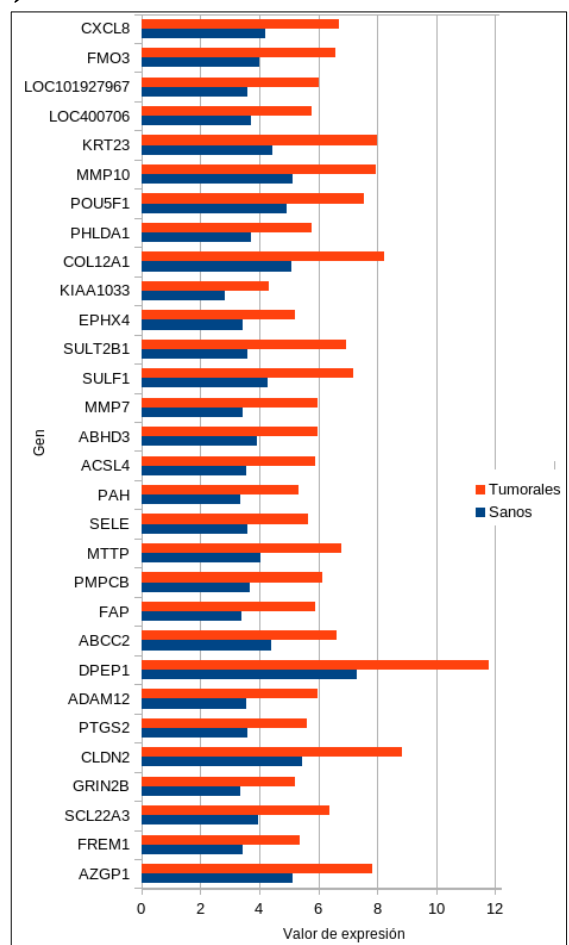
Sondas que representan genes sobreexpresados dos veces o más en población tumoral:

- ETV4 es un activador del factor de inducción de hipoxia (HIF), el cual promueve metástasis, y está asociado a LKB1, el sensor por excelencia del estrés metabólico, AMPK.

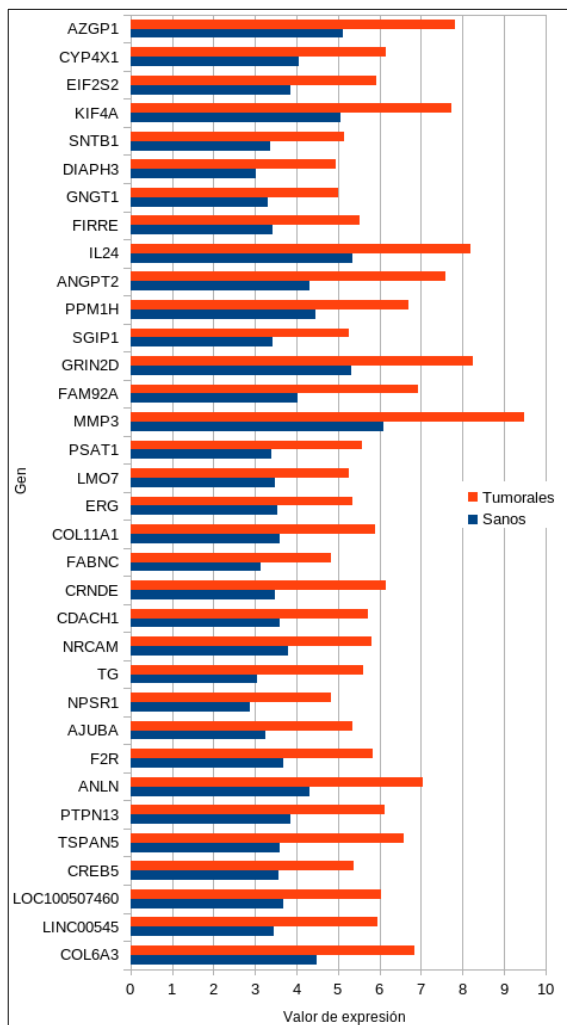
A)



B-2)



B-1)



C)

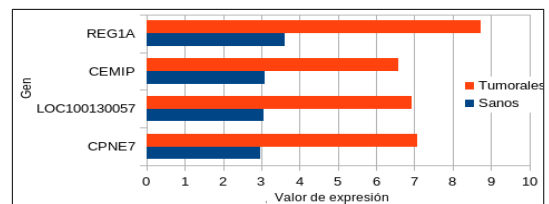


Fig 4. Comparación de sobreexpresión génica entre población sana (azul) y tumoral (rojo). A) Genes sobreexpresados hasta 1,5 veces superior en sanos. B-1, B-2) Genes sobreexpresados hasta 1,5 veces veces superior en sanos. C) Genes sobreexpresados hasta dos veces más en pacientes de CCR.

Sobreexpresión	Gen	Gen ID	Sobreexpresión	Gen	Gen ID
x1,5 en sanos	SLC4A4	8671	x1,5 en tumorales	GNGT1	2792
	CA4	762		DIAPH3	81624
	CA2	760		SNTB1	6641
	CLCA1	1179		KIF4A	24137
	FER1L6	654463		EIF2S2	8894
	ADTRP	84830		CYP4X1	260293
	EDN3	1908		AZGP1	563
	PKIB	5570		FREM1	158326
	SCL26A3	1811		SCL22A3	6581
	PYY	5697		GRIN2B	2904
	CA1	759		CLDN2	9075
	NXPE1	120400		PTGS2	5743
	B3GALT1	8708		ADAM12	8038
	VSIG2	23584		DPEP1	1800
	NXPE4	54827		ABCC2	1244
	MS4A12	54860		FAP	2191
x1,5 en tumorales	COL6A3	1293		PMPCB	9512
	LINC00545	440132		MTTP	4547
	LOC100507460	100507460		SELE	6401
	CREB5	9586		PAH	5053
	TSPAN5	10098		ACSL4	2182
	PTPN13	5783		ABHD3	171586
	ANLN	54443		MMP7	4316
	F2R	2149		SULF1	23213
	AJUBA	84962		SULT2B1	6820
	NPSR1	387129		EPHX4	253152
	TG	7038		KIAA1033	23325
	NRCAM	4897		COL12A1	1303
	CDACH1	1602		PHLDA1	22822
	CRNDE	643911		POU5F1	5460
	FABNC	2187		MMP10	4319
	COL11A1	1301		KRT23	25984
	ERG	2078		LOC400706	400706
	LMO7	4008		LOC101927967	101927967
	PSAT1	29968		FMO3	2328
	MMP3	4314		CXCL8	3576
	FAM92A	137392		401585	LOC401585
	GRIN2D	2906		259215	LY6G6F
	SGIP1	84251	x2 en tumorales	CPNE7	27132
	PPM1H	57460		LOC100130057	100130057
	ANGPT2	285		CEMIP	57214
	IL24	11009	x2,5 en tumorales	REG1A	5967
	FIRRE	286467		ETV4	2118

Tabla 3. Genes sobreexpresados en poblaciones sanas y tumorales.

3.3. Vías activas en pacientes sanos y tumorales

Las vías más representativas en cada conjunto de genes sobreexpresados (Tabla 4), según sus p-valores (material suplementario 20), indican una caracterización distintiva de vías en cada una de las poblaciones.

Vías activas en Sanos:

En la población sana, se presenta una sobreexpresión de genes relacionados a la activación de la fosfoinositol 3-quinasa (PI3K), asociada al sistema inmune mediante NTKR2 y MET. Los genes ERBB2 y ERBB4, pertenecientes a la familia HER2, participan en el desarrollo normal de tejidos, las alteraciones en sus expresiones son características de varios tipos de cánceres, por lo que su estudio mediante respuestas a radiación son relevantes a la hora de planificar un tratamiento.

Se presenta una normal activación de vías involucradas en el transporte de minerales y moléculas pequeñas, como CO₂, O₂, iones y aminoácidos, como SLC26A3 y SLC4A4, cuyas alteraciones generan diversas patologías.

Vías activas asociadas a sobreexpresiones una vez y media en Tumorales:

Como se observa en el apartado anterior, el tejido tumoral sobreexpresa genes asociados a adhesión celular, señalización de enzimas, y proliferación que promueven el crecimiento y la movilidad celular. Muchos corresponden a prognosis y diagnóstico tumoral en CCR como de diferentes tipos tumorales. Son genes que no estrictamente asociados al metabolismo, aunque estrechamente dependientes del metabolismo celular (Fig 5A).

-Los marcadores de cáncer NMDA [30] y NrCAM [32] activan a las vías NMDA y NRCAM. NMDA actúa activando al oncogen k-ras, el cual está fuertemente implicado en la progresión de CCR. El gen k-ras está asociado al gen APC, que se encuentra mutado en pacientes de Síndrome de Lynch. Este síndrome es un trastorno hereditario que aumenta el riesgo de tener cáncer de colon, cáncer de endometrio y otros tipos de cánceres.

-Las vías asociadas a adrenoreceptores se han encontrado activadas anormalmente en CCR y otros cánceres [43].

-Las monoaminos oxidasas son enzimas mitocondriales que catalizan la degradación de neurotransmisores y aminos dietéticas. Tumores de pacientes de cáncer de próstata y CCR, exhiben una mayor expresión tanto de MAOA como de MAOB [44, 45].

-Estudios varios, exhiben que muestras de CCR por medio de tinción inmunohistoquímica con anticuerpos monoclonales anti -DVL-3, también asociada a la vía Wnt, característica de estar alterada en varios cánceres [46].

Vías activas asociadas a sobreexpresiones dos veces en Tumorales:

Se detectó una sobreexpresión de genes que activan vías relacionadas al metabolismo de macromoléculas, como los carbohidratos, fosfolípidos, ácido hialurónico, glicosaminoglicanos, glicerofosfolípidos (Fig 5B).

-Los carbohidratos son los principales componentes de la dieta humana y el catabolismo de los monosacáridos, en particular la glucosa, derivados de ellos es una parte esencial del metabolismo energético humano. La glucosa se cataboliza a piruvato (glucólisis) y el piruvato sintetizado a partir de diversas fuentes se puede metabolizar para formar glucosa (gluconeogénesis). La glucosa también se puede polimerizar para formar glucógeno, y el glucógeno se puede descomponer en glucosa en respuesta al estrés o al hambre (glucogenólisis) [47].

La vía de la pentosa fosfato permite la síntesis de diversos monosacáridos a partir de glucosa, incluida la pentosa ribosa-5-fosfato, así como la generación de equivalentes reductores para procesos biosintéticos [47].

-El metabolismo de los glucosaminoglicanos se anota como parte del metabolismo de los carbohidratos. El ácido hialurónico es un glucosaminoglicano distribuido ampliamente en los tejidos conectivo, epitelial y neural, y más abundante en la matriz extracelular y la piel [48].

-Los glicerofosfolípidos son importantes componentes estructurales y funcionales de las membranas biológicas y constituyentes de las lipoproteínas séricas. Estos se forman utilizando ácidos grasos activados como donantes de acil-CoA. La mayoría de los glicerofosfolípidos se sintetizan en el retículo endoplásmico. Dado que la mayoría de estos se encuentran en todos los compartimentos de la membrana, hay una extensa red de transporte de glicerofosfolípidos de un compartimento de la membrana a otro a través de varios mecanismos que incluyen la difusión a través del citosol, la formación de complejos de transporte y la difusión a través de los sitios de contacto con la membrana (MCS) [49].

Vías activas asociadas a sobreexpresiones dos veces y media en Tumorales:

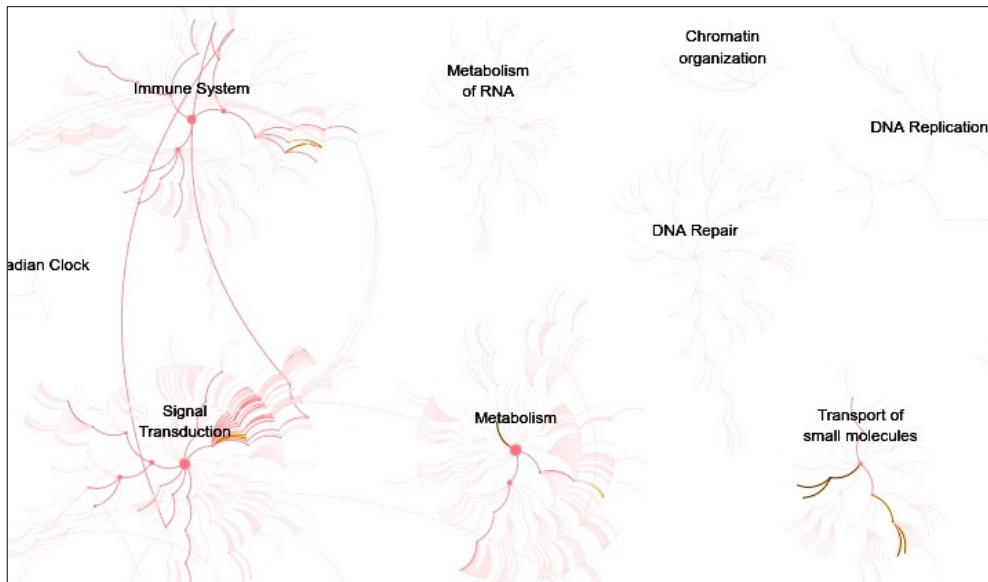
Se ven involucradas alteraciones en las vías asociadas al sistema inmune y transducción de señales (Fig 5C).

Se presenta una sobreactivación en las vías de señalización de interleucina-4 e interleucina-13, conocidos marcadores de CCR. También se ven involucradas sobreactivaciones de vías asociadas a la transducción de señales, como las asociadas a receptores y señalización de proteínas G (GPCR y G alfa), características de varios tipos de cánceres [50]. La sobreexpresión de la proteína ETV4, se asocia a la activación de vías GPCR, un marcador tumoral [51]. Otro marcador tumoral, entre ellos de CCR, es MAPK4 [52], cuya vía se encuentra activada una cantidad x2,5 veces a la normal.

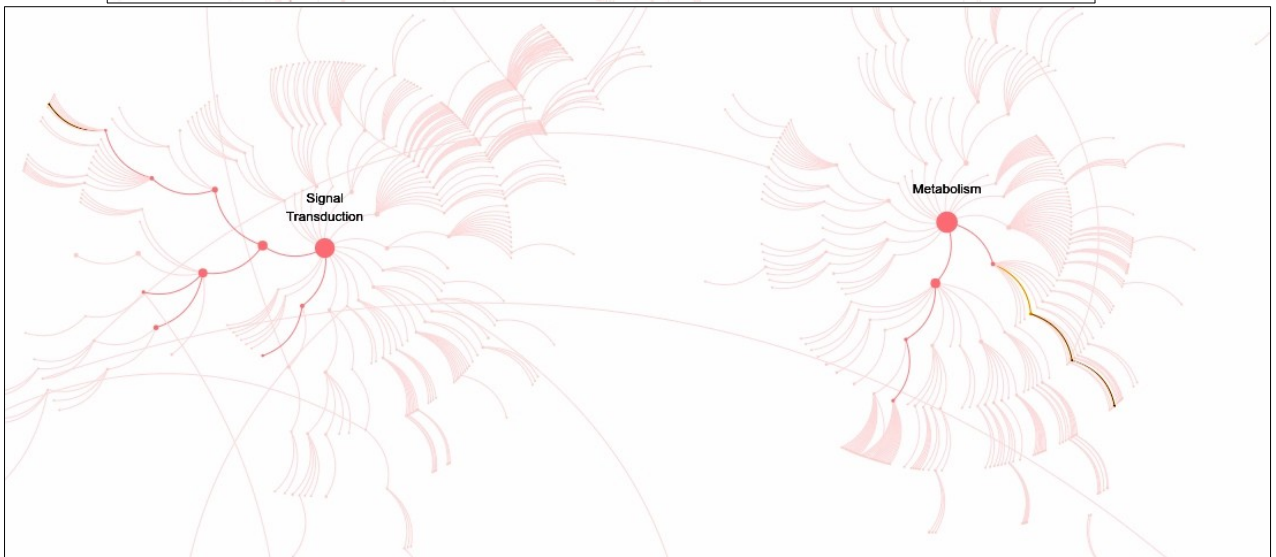
Sobreactivación	Nombre de vía	P-value	Sobreactivación	Nombre de vía	P-value
x1,5 en sanos	Hidratación reversible de dióxido de carbono	5,10E-07	x1,5 en tumorales	Regulación negativa por receptores de NMDA	0,022
	Absorción/liberación de O ₂ /CO ₂ por eritrocitos	9,60E-07		Regulación negativa por DVL	0,022
	Intercambio de O ₂ / CO ₂ en eritrocitos	2,26E-06		Fenilcetonuria	0,022
	Liberación/Absorción de O ₂ /CO ₂ en eritrocitos	2,26E-06		Formación de invadopodia	0,022
	SLC26A3 defectuoso	0,003		Activación de Ras mediante Ca ²⁺	0,022
	Activación de SLC4A4	0,003		FMO3 defectuosa	0,026
	Transporte de moléculas pequeñas	0,006		Moléculas sinápticas de adherencia	0,027
	Señalización de PDGFR	0,01		Adrenoceptores	0,031
	Transportadores de bicarbonato	0,013		Interacciones de NrCAM	0,039
	Unión a inhibidores de tirosina cinasa (TKI)	0,014		Desaminación oxidativa por MAOA y MAOB	0,055
	Intercambiadores de aniones	0,018	x2 en tumorales	Biosíntesis y exportación de hialuronano	2E-04
	Trastornos del transportador SLC	0,02		Metabolismo hialurónico	0,001
	Señalización SHC del receptor de interleucina	0,028		Metabolismo de glucosaminoglicanos	0,005
	Transporte de cargas inorgánicas y aminoácidos	0,03		Biosíntesis de glicerofosfolípidos	0,01
	Señalización por PDGF	0,044		Metabolismo de carbohidratos	0,014
	Biosíntesis del grupo sanguíneo de Lewis	0,049		Metabolismo de fosfolípidos	0,015
	Trastornos de los transportadores transmembrana	0,05		Metabolismo	0,042
	Receptores de unión a péptidos	0,057		Metabolismo de lípidos	0,053
	Biosíntesis de sistemas de grupos sanguíneos	0,06	x2,5 en tumorales	Unión receptores-ligandos	0,002
	Señalización por CBL	0,07		Contracción del músculo liso	0,01
	Señalización de NTRK2 por PI3K	0,087		Señalización de interleucina-4 e interleucina-13	0,028
	PI3K en señalización de ERBB4	0,09		Contracción del músculo	0,035
	PI3K en señalización de ERBB2	0,096		Receptores de unión a ligandos peptídicos	0,045
	Señalización de PI3K/AKT por MET	0,097		Señalización MAPK6/MAPK4	0,047
	PI3K activada por eritropoetina	0,1		Eventos de señalización G alfa (q)	0,049
x1,5 en tumorales	Degradación de colágeno	1,21E-06		Clase A1 de rodopsina	0,077
	Ensamblaje de fibras de colágeno	2,05E-05		Cascada de MAPK	0,083
	Degradación de la matriz extracelular	2,98E-04		Eventos de señalización G alfa (i)	0,09
	Formación de colágeno	4,23E-04		Unión de GPCR a ligandos	0,094
	Trimerización de colágeno	0,001		Transducción de señales	0,169
	Activación de las metaloproteinasas	0,003		Señalización de GPCR downstream	0,197
	Unión de quimiocina a receptores	0,007		Señalización mediante GPCR	0,236
	Receptores NMDA, unión y activación de glutamato	0,009		Señalización mediante interleucinas	0,249
	Biosíntesis de colágeno y enzimas modificadoras	0,012		Señalización de citocinas en el sistema inmune	0,317
	HEXA defectuoso causante de GM2G1	0,017		Sistema inmune	0,504

Tabla 4. Vías más representativas asociadas a genes sobreexpresados graficadas por Reactome.

A)



B)



C)

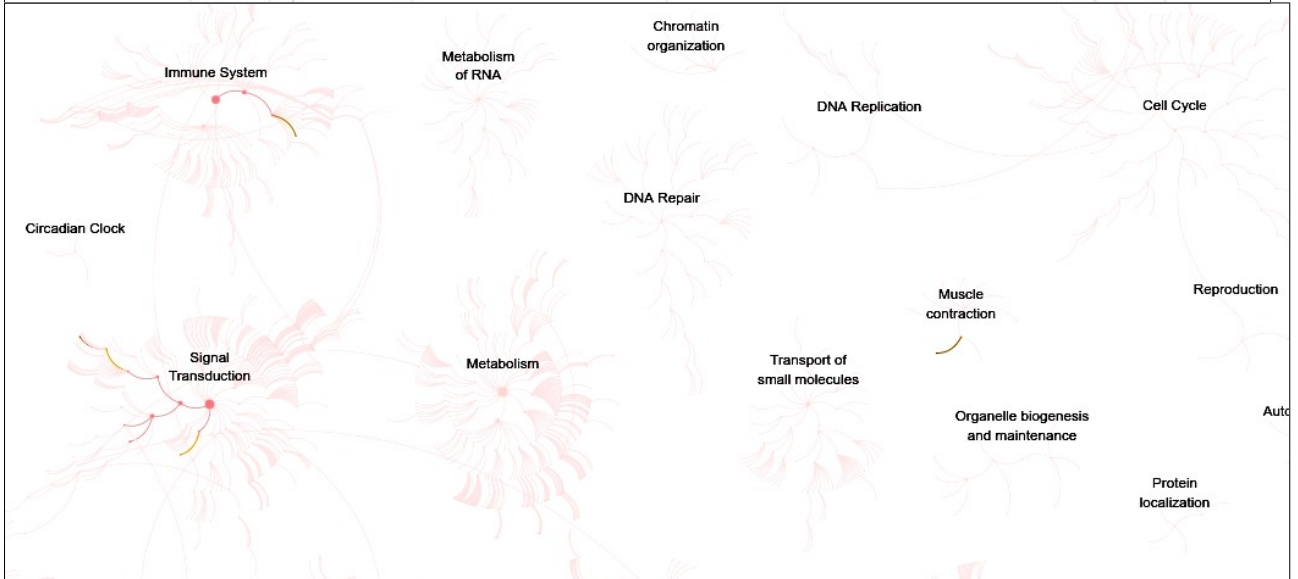


Fig 5. Vías clasificadas por jerarquías, asociadas a conjuntos de genes en muestras tumorales por niveles de sobreexpresiones. A) Vías activas hasta unas 1,5 veces. B) Vías activas hasta el doble de veces. C) Vías activas hasta unas 2,5 veces. La intensidad de color rojo, representa visualmente la significancia de la vía al conjunto dado.

3.4. Vías metabólicas activas por sobreexpresiones de genes

Los procesos metabólicos en las células humanas generan energía a través de la oxidación de moléculas consumidas en la dieta y median la síntesis de diversas moléculas esenciales no tomadas en la dieta así como la inactivación y eliminación de tóxicos generados endógenamente o presentes en el ambiente extracelular. Los procesos del metabolismo energético se pueden clasificar en siete grupos. Igualmente, para el presente trabajo, nos concentramos en dos grupos según se trate de moléculas derivadas de carbohidratos o de lípidos. Al mismo tiempo, todos estos procesos están estrechamente integrados [53].

Los intermedios en las reacciones de generación de energía son materiales de partida para la biosíntesis de aminoácidos y otros compuestos. Las enzimas oxidorreductasas de amplia especificidad pueden participar tanto en las reacciones de degradación como en las biosíntesis, y los procesos de señalización mediados por hormonas funcionan para coordinar la operación de generación de energía, y acoplarlas a otros procesos biosintéticos [53].

A partir de los datos obtenidos por medio de Reactome, se dilucidó que la reorganización metabólica en pacientes de CCR está asociada a alteración de genes sobreexpresados tanto una vez y media, como el doble de veces en tumorales por sobre personas sin esta enfermedad. Las subvías derivadas de la vía Metabolismo, para cada rango de sobreexpresión, se detallan en la Tabla 5.

Las alteraciones más influyentes sobre la homeostasis metabólica en la población tumoral, repercuten en el metabolismo tanto de carbohidratos como de lípidos (Fig 6A).

A niveles de sobreactivación de hasta x1,5 veces superiores, se han presenciado en Fig 6B, vías relacionadas al metabolismo y síntesis de metabolitos intermediarios en las vías de glúcidos, grasas y proteínas, entre ellos, hormonas, nucleótidos y aminoácidos (Tabla 6).

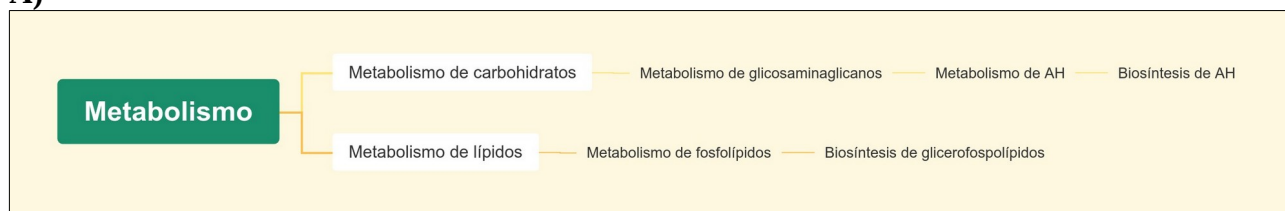
Sobreactivación	Nombre de vía	P-value
x1,5 en tumorales	Desaminación oxidativa por MAOA y MAOB	0,055
x2 en tumorales	Biosíntesis y exportación de HA	2E-04
	Metabolismo hialurónico	0,001
	Metabolismo de glicosaminoglicanos	0,005
	Biosíntesis de glicerofosfolípidos	0,010
	Metabolismo de carbohidratos	0,014
	Metabolismo de fosfolípidos	0,015
	Metabolismo	0,042
	Metabolismo de lípidos	0,054

Tabla 5. Vías metabólicas con una representación significativa en base a sus p-values.

En pacientes oncológicos ocurre una serie de alteraciones metabólicas complejas, caracterizadas por tener un gasto energético superior, relacionado a las reservas grasas y proteicas, más una utilización inadecuada de los glúcidos.

Los tumores sólidos obtienen su energía del metabolismo de la glucosa, lo que aumenta el consumo energético. Otra forma de neoglucogénesis es a partir de aminoácidos y glicerol, que se obtienen a través de la proteólisis muscular y la lipólisis. El aumento de la lipólisis, con mayor producción de ácidos grasos libres y glicerol, sirve de sustrato para la gluconeogénesis, la cual conlleva a un aumento del consumo energético.

A)



B)

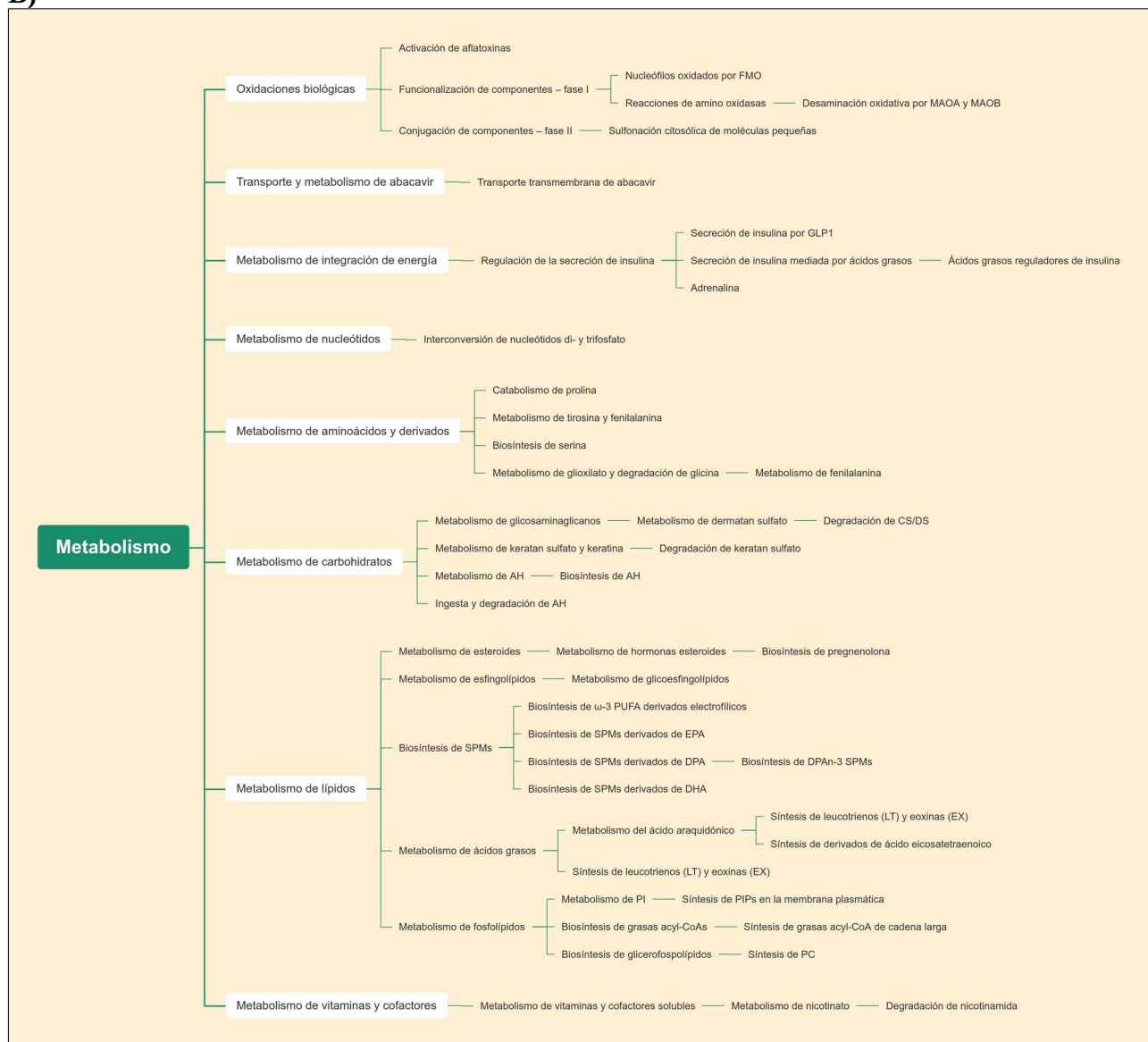


Fig 6. Vías metabólicas sobreactivadas en población tumoral. Vías obtenidas por Reactome, asociadas a metabolismo A) En una proporción de dos veces, siendo las más significativas. B) En una vez y media.

Nombre de vía	P-value
Nucleófilos oxidados por FMO	0,068
Catabolismo de prolina	0,068
Sulfonación citosólica de moléculas pequeñas	0,072
Reacciones de amino oxidasas	0,092
Síntesis de derivados de ácido eicosatetraenoico	0,092
Degradación de keratan sulfato	0,092
Metabolismo de fenilalanina	0,112
Ingesta y degradación de hialuronano	0,115
Degradación de CS/DS	0,119
Metabolismo de glicoesfingolípidos	0,139
Ácidos grasos reguladores de insulina	0,142
Transporte transmembrana de abacavir	0,153
Biosíntesis de pregnenolona	0,153
Oxidaciones biológicas	0,168
Metabolismo de tirosina y fenilalanina	0,175
Biosíntesis de serina	0,190
Biosíntesis de ω -3 PUFA derivados electrofílicos	0,197
Activación de aflatoxinas	0,200
Biosíntesis de DPAn-3 SPMs	0,207
Metabolismo de keratan sulfato y keratina	0,211
Biosíntesis de SPMs derivados de EPA	0,214
Biosíntesis de SPMs derivados de DPA	0,221
Adrenalina	0,224
Metabolismo del ácido araquidónico	0,233
Transporte y metabolismo de abacavir	0,235
Síntesis de prostaglandina y tromboxano	0,261
Metabolismo de dermatan sulfato	0,296

Nombre de vía	P-value
Síntesis de leucotrienos (LT) y eoxinas (EX)	0,305
Síntesis de grasas acyl-CoAs	0,311
Síntesis de PC	0,347
Regulación de la secreción de insulina	0,352
Degradación de nicotinamida	0,377
Metabolismo de esfingolípidos	0,382
Secreción de insulina por GLP1	0,386
Metabolismo de glioxilato y degradación de glicina	0,386
Biosíntesis de SPMs derivados de DHA	0,386
Biosíntesis de grasas acyl-CoA	0,440
Interconversión de nucleótido di- y trifosfato	0,442
Conjugación de componentes – fase II	0,447
Metabolismo de integración de energía	0,448
Funcionalización de componentes – fase I	0,466
Secreción de insulina mediada por ácidos grasos	0,466
Metabolismo de nicotinato	0,469
Metabolismo de hormonas esteroides	0,489
Biosíntesis de SPMs	0,549
Metabolismo de ácidos grasos	0,567
Síntesis de PIPs en la membrana plasmática	0,678
Metabolismo de nucleótidos	0,776
Metabolismo de aminoácidos y derivados	0,781
Metabolismo de PI	0,781
Metabolismo de vitaminas y cofactores solubles	0,815
Metabolismo	0,879
Metabolismo de vitaminas y cofactores	0,934
Metabolismo de esteroides	0,987

Tabla 6. Vías metabólicas sobreactivadas x1,5 veces en pacientes tumorales.

Los genes alterados en pacientes, asociados a las vías metabólicas representativas de cada intensidad de sobreactivación, son los codificadores de las proteínas FMO3, CEMIP y Copina-7 (Tabla 7).

Los niveles de FMO3 se aumentan en el padecimiento de CCR, estando asociado a metástasis.

Copina 7, está asociada al metabolismo lipídico y tráfico de metabolitos. El producto de gen CEMIP, está asociado al metabolismo de la glutamina y glúcidos, como el glucógeno.

Trabajos previos realizados en el IBIOGEM, demuestran que la molécula GSK3 fosforilada, que compite con CEMIP por b-cateninas, se halla igualmente sobreexpresada en el desarrollo tumoral. CEMIP, junto a GSK3 fosforilado, permiten la conversión de glucosa-6-fosfato por glucógeno (glutaminólisis).

A)

Gen ID	Nombre de gen	UniProt ID	Nombre de proteína
2328	FMO3	P31513	Monooxigenasa contenedora de flavina 3
27132	CPNE7	Q8WUJ3	Copina-7 (Copina VII)
57214	CEMIP	Q9UBL6	Proteína inductora de migración celular y unión a hialuronano (EC 3.2.1.35)

B)

Gen ID	Nombre de vía
FMO3	Desaminación oxidativa por MAOA y MAOB
CEMIP	Biosíntesis y exportación de AH
CEMIP	Metabolismo de AH
CEMIP	Metabolismo de glicosaminoglicanos
CPNE7	Biosíntesis de glicerofosfolípidos
CEMIP	Metabolismo de carbohidratos
CPNE7	Metabolismo de fosfolípidos
CPNE7, CEMIP	Metabolismo
CPNE7	Metabolismo de lípidos

Tabla 7. Genes asociados al metabolismo de carbohidratos y lípidos. A) FMO3, CEMIP y CPNE7 son genes que activan las vías metabólicas más preponderantes (B).

Las cantidades de vías sobreactivadas por los genes metabólicos, se detallan en la Tabla 8. Esto incluye vías que no metabólicas, entre ellas, vías asociadas al sistema inmune y al transporte de moléculas.

Gen	Cantidad de vías
FMO3	10
GNGT1	75
PTPN13	10
SELE	6
SULT2B1	17
PTGS2	38
PAH	7
GRIN2B	48
ACSL4	9
SCL22A3	7
COL11A1	26
DPEP1	23
PSAT1	3
ABHD3	5
ERG	3

Tabla 8. Genes asociados a metabolismo, junto a las cantidades de vías que alteran.

Para tener un mayor conocimiento, se obtuvieron a partir de los genes metabólicos, sus correspondientes sondas, dando un total de 31 sondas diferencialmente expresadas, asociadas a vías metabólicas representativas (Tabla 9A) y no representativas (Tabla 9B).

A)

Gen	ID de sondas	Cantidad
FMO3	2366966	1
CEMIP	3604209,3604179,3604182,3604157,3604212,3604188	6
CPNE7	3674233,3674239	2

B)

Gen	ID de sondas	Cantidad	Gen	ID de sondas	Cantidad
GNGT1	3012921	1	SCL22A3	2934557	1
SELE	2443503	1	COL11A1	2425832,2425833,2425845	3
SULT2B1	3837869	1	DPEP1	3674268,3674266,3674264	3
PTGS2	2448386	1	PSAT1	3175991	1
PAH	3468483,3468494,3468531	3	ABHD3	2852631	1
GRIN2B	3445418,3445429,3445340	3	ERG	3931788	1
ACSL4	4017838	1	PTPN13	2734730	1

Tabla 9. Sondas metabólicas sobrexpresadas en pacientes de CCR. A) Sondas correspondientes a los genes implicados en la alteración de vías metabólicas de carbohidratos y lípidos. **B)** Sondas correspondientes a genes asociados a vías de metabolitos varios.

A partir de valores de expresión de los genes sobreexpresiones en población tumoral, se dilucidaron los genes más representativos de cada estadio, notando que los estadios tumorales más acentuados génicamente, son los I y III, con unas cantidades de 31 y 19 genes mayormente expresados, respectivamente (Tabla 10). Las diferencias de expresiones, no son cuantitativamente elevadas entre los distintos estadios, la mayoría de los genes no superando una cantidad de 1,5 veces de sobreexpresión en un estadio por sobre otro (Fig 7).

Sobreexpresión en 1,5			
Estadio I	Estadio II	Estadio III	Estadio IV
LINC00545	COL6A3	LOC100507460	TSPAN5
CREB5	FANCB	PTPN13	CRNDE
F2R	COL11A1	ANLN	MMP7
AJUBA	ERG	NPSR1	LOC400706
TG	GNGT1	LMO7	
NRCAM	GRIN2B	PSAT1	
DACH1	ADAM12	MMP3	
FAM92A	FAP	GRIN2D	
SGIP1	ABHD3	KIF4A	
PPM1H	SULF1	EIF2S2	
ANGPT2	PHLDA1	CYP4X1	
IL24		AZGP1	
FIRRE		FREM1	
DIAPH3		SLC22A3	
SNTB1		MTTP	
CLDN2		SELE	
PTGS2		ACSL4	
DPEP1		SULT2B1	
ABCC2		MMP10	
LY6G6F			
PMPCB			
PAH			
EPHX4			
LOC401585			
KIAA1033			
COL12A1			
POU5F1			
KRT23			
LOC101927967			
FMO3			
CXCL8			

Sobreexpresión en 2			
Estadio I	Estadio II	Estadio III	Estadio IV
LOC100130057	CEMIP	CPNE7	---
		REG1A	

Sobreexpresión en 2,5			
Estadio I	Estadio II	Estadio III	Estadio IV
ETV4	---	---	---

Tabla 10. Genes sobreexpresados por estadio.

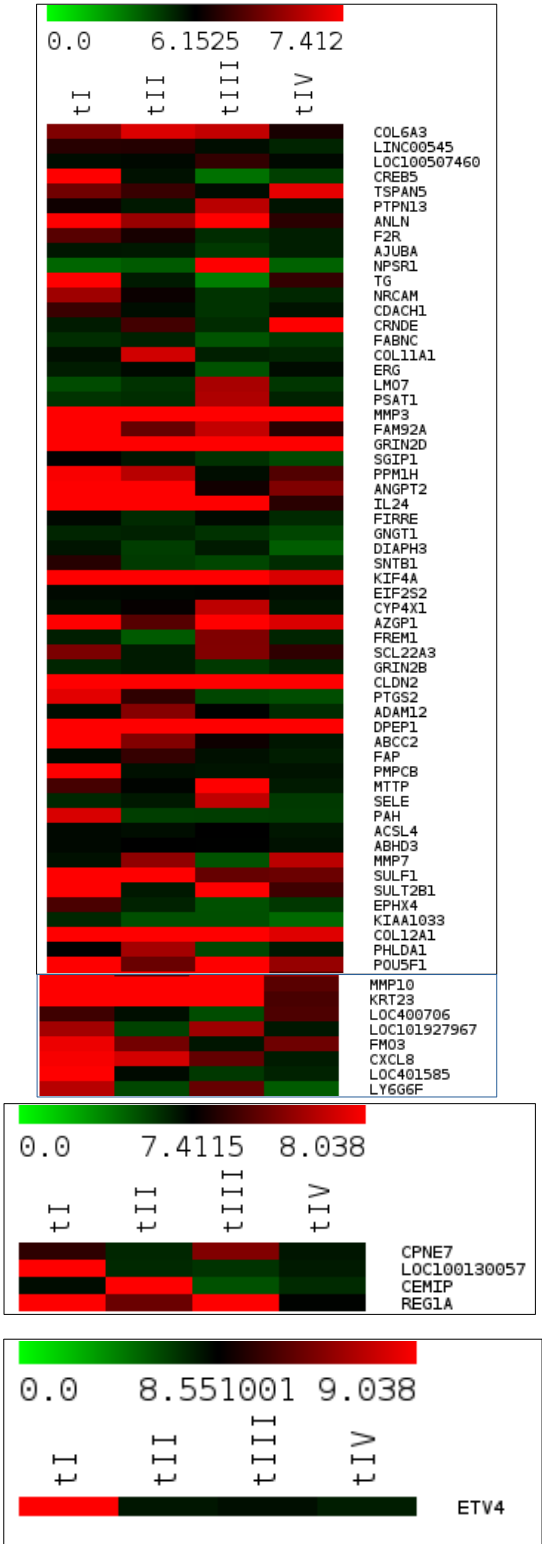


Fig 7. Representación de genes expresados.

En muestras sanas, hay una elevada expresión de genes relacionados al sistema inmune y el funcionamiento normal de los tejidos (Tabla 4), como son los casos de PI3K y HER2.

Cuando comienza el desarrollo de CCR, se producen alteraciones génicas que incluyen una reprogramación de la homeostasis metabólica, incluyendo un aumento de la glicólisis, glutaminólisis y síntesis de ácidos grasos (Tabla 7B).

Durante el estadio I, comienza a haber una disminución de genes relacionados al sistema inmune, y comienza a haber un aumento de marcadores oncogénicos, como NrCAM, AJUBA, y DIAPH3. Hay una alteración de genes relacionados al metabolismo de aminoácidos como la fenilalanina (gen PAH) [54], eicosanoides (PTGS2) [55], además de un mayor procesamiento de moléculas nitrogenadas (FMO3), como las proteínas provenientes de la dieta [56]. El aumento de expresión más acentuado, es del gen ETV4, coactivador de la señalización de Hif [57].

En el estadio II de CCR, se caracteriza por la expresión de marcadores de CCR, el continuo deterioro del sistema inmune, un aumento en la actividad de síntesis de colágeno (COL11A1) [58], fosfatidicolina (ABHD3) [59], y procesos de glutaminólisis asociadas a la proteína sobreexpresada, CEMIP. También, un aumento del uso del aminoácido glutamato, causada por la sobreexpresión del gen GRIN2B [60].

Durante el estadio III, se visualizó una mayor actividad metabólica ácidos grasos, entre ellos acil-CoA (ACSL4) [61] y fosfatidilserina (PSAT1) [62]. También hay un aumento en los tráfico de tirosina (PTPN13) [63], y calcio (CPNE7).

El más avanzado estadio IV, presenta un aumento en la expresión de tetraspanina 5 (TSPAN5), neoplasia colorrectal diferencialmente expresada (CRNDE) y metaloproteinasas de matriz 7 (MMP7), las tres asociadas al progreso de cáncer en etapas avanzadas y metástasis [64, 65, 66].

Vías metabólicas a partir de sondas

Como enfoque secundario, se realizó el análisis de metabolismo a partir de las sondas sobreexpresadas (Sección 3.2), por medio de Reactome.

Estos resultados (material suplementario 21), arrojaron un total de 18 sondas (material suplementario 22) asociadas a 43 vías metabólicas, sobreexpresadas en una cantidad de hasta x1,5 veces superior en población tumoral. De las 43 vías metabólicas, 41 fueron dilucidadas anteriormente por genes, las otras dos corresponden al metabolismo de porfirinas y degradación de grupos hemo.

La sonda corresponde al gen CEMIP, también obtenido en el apartado anterior, como preponderante en la sobreexpresión del metabolismo de glúcidos. Las vías metabólicas asociadas con valores relevantes de p-value, coinciden con las obtenidas anteriormente, correspondiendo al metabolismo de carbohidratos, AH, y glicosaminoglicanos en una cantidad de 2 veces superior en tumorales (Tabla 11).

probe_id	Nombre de vía	P-value
3604179	Biosíntesis y exporte de AH	0,000
3604179	Metabolismo de AH	0,001
3604179	Metabolismo de glicosaminoglicanos	0,004
3604179	Metabolismo de carbohidratos	0,012
3604179	Metabolismo	0,073

Tabla 11. Vías metabólicas significativas, asociadas a sondas sobreexpresadas. Las vías dilucidadas, coinciden con los resultados del análisis génico.

4. Discusión

4.1. Sobreexpresión de genes

El cáncer colorrectal es el cáncer originado en el colon o el recto, una vez que las células de esos órganos crecen descontroladamente. En lo que concierne a los factores incidentes, la vida sedentaria y el ambiente también influyen en su desarrollo.

El estudio de expresión génica en pacientes de CCR, reveló que esta enfermedad altera los niveles de expresión de genes asociados a adhesión celular y proliferación que promueven el crecimiento y la movilidad celular. Otros genes sobreexpresados, corresponden a marcadores de prognosis y diagnóstico tumoral en CCR como de diferentes tipos tumorales. Son genes estrechamente dependientes del metabolismo celular. Las diferenciaciones génicas por 2 veces en la población tumoral, fueron las relacionadas al metabolismo de carbohidratos y lípidos.

La expresión génica en muestras de tejidos normales, refleja un aumento de genes relacionados con el transporte de moléculas, interacciones y biosíntesis de macromoléculas (Sección 3.2), deficientes en CCR.

Durante el desarrollo de la enfermedad, las vías metabólicas son reprogramadas, aumentando la glicólisis, glutaminólisis y síntesis de ácidos grasos. La información sugiere que las células tumorales exhiben el “Efecto Warburg”. Por ejemplo, los tumores CCR primarios y metastásicos tienen un mayor consumo de glucosa que el tejido normal. Numerosos genes y proteínas importantes para la captación de glucosa y la glucólisis se regulan positivamente en el CCR, incluido el factor 1 α inducible por hipoxia (HIF-1 α), piruvato quinasa M2 (PKM2), y lactato deshidrogenasa A (LDHA, un monómero de LDH) [67]. De hecho, la regulación positiva de muchos de los genes obtenidos en los resultados, se han correlacionado con la agresividad del tumor y un mal pronóstico en el CCR.

La dependencia de la glutamina es otro fenotipo metabólico que apoya el metabolismo de Warburg y es común a muchos tipos diferentes de cáncer, incluido el CCR. El metabolismo lipídico desregulado se reconoce como otro fenotipo metabólico en CCR. La mayoría de las células normales utilizan preferentemente ácidos grasos circulantes derivados de la grasa de la dieta. Por el contrario, las células cancerosas no solo aumentan la captación de lípidos y lipoproteínas extracelulares, sino que también regulan al alza la biosíntesis de lípidos de novo y las vías de síntesis de colesterol, que producen metabolitos de lípidos cruciales para las membranas celulares. Los pacientes con CCR en estadio III y IV tienen niveles circulantes de ácidos grasos más altos que los pacientes con CCR en estadio I y II [67].

El noventa y tres por ciento de los tumores esporádicos de CCR exhiben una señalización hiperactiva de WNT, de los cuales el 80% es causado por la inactivación bialélica del gen supresor de tumores de la poliposis coli adenomatosa (APC) o mutaciones activadoras en la β -catenina (CTNNB1). La señalización de WNT es fundamental en el desarrollo embrionario y la homeostasis [67].

Las 3-quinasas de fosatidilinositol (PI3K) son una familia de enzimas de transducción de señales intracelulares críticas para la respuesta de las células normales a los factores de crecimiento, incluido el factor de crecimiento derivado de plaquetas (PDGF). Las mutaciones que conducen a la señalización hiperactiva de PI3K ocurren en la mayoría de los tumores de CCR.

Todo esto, resalta la importancia de poder realizar acciones terapéuticas lo antes posible del avance de la enfermedad, a partir de la obtención de información de diferenciación génica.

4.2. Reactome: distinción de vías metabólicas y sus diferencias de sobreexpresión

En los resultados obtenidos de la activación en vías celulares (Sección 3.3), la población padeciente de CCR presentan alteraciones en los niveles de expresión de vías relacionadas al metabolismo de nutrientes (Tabla 6), evidenciándose alteraciones en las vías de homeostasis metabólica como preponderantes en la discriminación de los distintos tejidos.

Las vías diferenciadas en hasta 1,5 veces en los casos de población sana, fueron las asociadas al correcto funcionamiento del sistema inmune, tal es el caso de PI3K, relacionada a la activación y fosforilación de ERBB2, ERBB4, NTKR2, NTRK3 y MET (Tabla 4). Cabe recalcar que se sobreexpresan, en menor medida, vías no metabólicas, como las de transporte de minerales y moléculas pequeñas, como CO₂ y O₂.

En las vías diferenciadas en 1,5 veces en población tumoral, se destacan las relacionadas a la adhesión celular, señalización y proliferación de tumores. También se presentan desordenes en el metabolismo de oxidaciones biológicas y cofactores (Fig 6B).

A diferencias de 2 veces, las vías metabólicas involucradas corresponden al metabolismo de carbohidratos y lípidos, entre ellos, AH, glicosaminoglicanos y glicerofosfolípidos (Fig 5A, Fig 6A, Tabla 5). Esto es biológicamente consistente, ya que se sabe que las células tumorales poseen un aumento de tasa metabólica (efecto Warburg), que solo visualizando los números, no hubiera podido notarse. Los resultados obtenidos además apoyan los estudios realizados en el IBIOGEM, CICYTTP al confirmar la asociación de la vía GSK3, WNT y beta-catenina en CCR [1].

En una cantidad de 2,5 veces, se sobreactivan las vías asociadas a la señalización de trombinas, e interleucina-4 y interleucina-13 (Tabla 4). Esto confirma el valor de la reorganización metabólica de las vías asociadas a glicosaminoglicanos y glicerofosfolípidos, además de reafirmar el aumento de consumo de energía de carbohidratos y lípidos. Como detalle adicional, se han detectado sobreactivada la vía conocida como marcador de pronóstico, GPCR.

4.2.1. Vías metabólicas sobreactivadas en sondas vs genes

Partiendo de las sondas, a niveles de sobreactivación de x1,5, se obtuvieron 43 vías metabólicas, respectivamente. De estas, 41 coincidieron con el conjunto de 63 vías obtenidas mediante el análisis de genes.

A niveles de x2, se detectaron alteraciones en 5 vías, ya incluidas dentro de las 8 vías metabólicas por genes, ambos métodos coincidiendo en que hay un aumento en el metabolismo de glúcidos.

Adicionalmente, el procesamiento con genes, pudo detectar un aumento el metabolismo de fosfolípidos. Mientras que, el enfoque con sondas, detectó un aumento en el metabolismo de porfirinas, los cuales suelen tener elevadas concentraciones en tejidos tumorales [68].

La diferencia en cantidad de resultados, se basa en el uso de los IDs de genes como punto de estudio, siendo reconocidos más bases de datos y plataformas genómicas. Esto también muestra la importancia de tener en cuenta ambos tipos de identificados para el estudio y análisis metabólico.

4.2.2. Conocimientos previos asociados al metabolismo de CCR

La desregulación de la ruta de señalización de PI3K/Akt en humanos, resulta en la activación del gen k-ras, una importante vía de la señal de transducción celular implicada en la proliferación y transformación maligna de tumores [69]. La activación de esta ruta en las células cancerígenas, da lugar a sucesos como la activación de mTOR, considerado como el principal regulador del crecimiento celular, el cual induce la pérdida de la función de las subunidades supresoras de tumores p85 y PTEN de PI3K, la mutación de la subunidad catalítica p110a de PI3K [70], y la amplificación de las isoenzimas de Akt, Akt1 y Akt2, las cuales inactivan y mutan proteínas reguladoras de mTOR asociadas a Akt, como las supresoras de tumores TSC1 y TSC2 [71]. En otras palabras, la inhibición de componentes de la vía PI3K puede hacer oposición a la resistencia y a la quimioterapia dirigidas en cáncer.

La AMPK es una serina/treonina quinasa y una vía de sensor de combustible celular sensible al aumento de la relación AMP/ATP, conectado a varios supresores de tumores humanos. AMPK es conocido por regular el metabolismo de lípidos, el colesterol y la glucosa en tejidos metabólicos [72]. Los tratamientos de antitumorales como metformina, un fármaco que activa el mecanismo metabólico por medio de AMPK, e interfiere con la mutación de PI3K y la deficiencia de p53 [46].

Trabajos de laboratorio han revelado que una de las principales vías controladas por el AMPK es la rampamicina mamífera (mTOR). mTOR es un integrador de entradas de nutrientes y factores de crecimiento, controlando el crecimiento celular en eucariotas, y está desregulado en la mayoría de los cánceres humanos. El complejo 1mTOR (mTORC1) controla la traslación de varios reguladores de factores de crecimiento celular, incluyendo la ciclina D1, el factor de inducción de hipoxia 1a (HIF-1 α), el cual promueve procesos como la angiogénesis.

En este trabajo, se observó la sobreexpresión en población tumoral de ETV4 que comparte una vía con HIF, el cual promueve metástasis, y está asociado a LKB1 [73], activador del sensor por excelencia del estrés metabólico, AMPK.

Los efectos de la metformina se explican principalmente por la activación de AMPK, que inhibe la síntesis de proteínas y la gluconeogénesis durante el estrés celular [71] y desarrollo tumoral, evitando una vía anabólica y de mTOR en tumores promoviendo la muerte celular tumoral.

El ribonucleótido 5-aminoimidazol-4-carboxamida (AICAR), es un intermediario en la generación de monofosfato de inosina. Este es un análogo del monofosfato de adenosina, capaz de estimular la actividad de la proteína quinasa dependiente de AMP. La AMPK es inducida por 5-FU, un fármaco de quimioterapia anticanceroso, mejora con AICAR, junto con la disminución de mTOR. Esto indica que la activación de AMPK por AICAR potencia la apoptosis celular, inducida por 5-FU en células tumorales. Hasta la actualidad, el 5-fluorouracilo (5-FU) sigue siendo un fármaco quimioterapéutico ampliamente utilizado en el tratamiento del carcinoma colorrectal [74].

Los resultados de trabajos previos de CCR, son concordantes con los obtenidos en el presente trabajo, pudiendo verse una alteración en la regulación de diversas vías, incluida la vía de PI3K, durante el desarrollo de la enfermedad. Se aporta información complementaria y de utilidad al analizar nuevos tratamientos con antitumorales.

4.3. Uso en conjunto de herramientas de uso público

El procedimiento de manipulación de datos mediante el uso de herramientas bioinformáticas fue factible, habiendo sido posible obtener información valorable, al igual que en los estudios con pacientes en la Provincia de Buenos Aires [14], donde a partir de un microarreglo, se evaluaron expresiones génicas utilizando herramientas informáticas, entre ellas por medio del software v3.0.

El uso conjunto de herramientas proyectan nuevas y variadas topologías de análisis con amplios potenciales de conocimientos, tal es el ejemplo de Ingenuity, que en conjunto con DAVID realiza cálculos estadísticos para análisis funcionales [75, 76].

5. Conclusiones y trabajos futuros

Gracias al esclarecimiento de los mecanismos que promueven el desarrollo del CCR en este trabajo, junto con los ensayos biológicos y bioquímicos utilizados, se pudo constatar que el enfoque bioinformático, tiene la posibilidad de brindar conocimientos sobre cómo avanzar en el estudio de los factores de riesgo, desarrollo de CCR, y evidenciar los blancos para el tratamiento con antitumorales a futuro. La combinación de herramientas bioinformáticas, genera un potencial provecho de trabajos in silico en la mesada, generando una búsqueda más focalizada y costo efectivo en nuevas muestras de pacientes de la región.

A continuación, se presentan algunos trabajos futuros que pueden desarrollarse como resultado de esta investigación y el empleo de las herramientas utilizadas en el presente, y solo nombramos por exceder el alcance de esta tesis. Además, se sugieren algunos desarrollos específicos para apoyar y mejorar el modelo y metodología propuestos. Entre los posibles trabajos futuros se destacan:

- Realizar estudios en una mayor cantidad de microarreglos, entre ellos, E-GEOD-63624, que con tiene datos de expresión para 52 tumores primarios estables a microsatélites (MSS) de pacientes con cáncer de colon proximal.

- Hacer foco en vías metabólicas ya conocidas, como gliceraldehído-3-fosfato deshidrogenasa (GAPDH), familia aldehído deshidrogenasa 3, miembro A1 (ALDH3A1), lactato deshidrogenasa A (LDHa) e isoenzima muscular de la piruvato quinasa (PKM2).

- El estudio de hipo e hiper metilación de islas CpG en supresores de tumores y otras características epigenéticas del cáncer. Tales eventos han sido identificados en varios cánceres, incluidos CCR y cáncer de colon [77].

- El uso complementario de paquetes de Bioconductor en lenguaje R, entre ellos, heatmaply, GEOquery, simpleaffy, ggplot2, Biobase y arrayQualityMetrics, con la finalidad de generar información de valor complementario a los del presente trabajo.

- El análisis de genes no caracterizados, visualizando la secuencia proteica de su producto, ya sea por simulación o resultados en laboratorio, para así saber su función y vías a las que se asociaría, por medio de herramientas bioinformáticas.

- La realización de análisis de expresión génica para conformar los resultados del proyecto en muestras colorrectales de la Argentina y la región, por medio de RT-PCR cuantitativas o semi-cuantitativas.

Anexo: Material suplementario

Disponible en: https://drive.google.com/drive/folders/1zDfb4J4rL_Jo4xA_tx_ncBlUf1SB63-f

MS-01: filtrado_por_SE.pl

MS-02: filtrado_por_Homo_sapiens.pl

MS-03: obtencion_promedios_tumorales.pl

MS-04: concatenacion_de_valores.pl

MS-05: sondas_sobreexpresadas.pl

MS-06: valores_expresion_genes_sobreexpresados.pl

MS-07: valores_expresion_genes_sobreexpresados_promedio.pl

MS-08: funciones_genes_sobreexpresados.pl

MS-09: DEG_tumorales.pl

MS-10: vias_metabolicas.pl

MS-11: genes_metabolicos.pl

MS-12: vias_asociadas_a_cada_gen.pl

MS-13: sondas_por_gen.pl

MS-14: sondas_metabolicas.pl

MS-15: sondas_metabolicas.pl

MS-16: vias_asociadas_a_cada_sonda.pl

MS-17: genes_por_sondas.pl

MS-18: valores_sondas_con_SE_menores_a_20.txt

MS-19: valores_expresion_sondas.txt

MS-20: Reactome_genes.zip

MS-21: Reactome_sondas.zip

MS-22: sondas asociadas a metabolismo.txt

Bibliografía

1. Ridel E, Ascolani J, Tognola F, Viale S, Romero N, Favant JL; Martinez Marignac VL. PID UNER 6155: Ensayo terapéutico preclínico de Cloroquina y 5-FU en el tratamiento de CCR. Ciencia, Docencia y Tecnología- Suplemento, 9 num.9 (2019): 399 – 413.
2. Martinez Marignac VL, Mondragon L, Oertlin G, Cervantes L, Cantero F, Favant JL. PID UNER 6176: Preclinical Study of Genuine ESSIA Formula: A Cancer Treatment Eight herbs' Tea Minimizes DNA Insult of X-rays. Clinical Cancer Investigation Journal, num.9 (2020). DOI: [10.4103/ccij.ccij.73.20](https://doi.org/10.4103/ccij.ccij.73.20).
3. Martinez Marignac VL; Mondragon L, Favant JL. Sources of ionizing radiation (IR) and their biological effects. An interdisciplinary view, from the physics to cell and molecular biology. Clinical Cancer Investigation Journal, 8 num.4 (2019): 129 – 138. DOI: [10.4103/ccij.ccij.108.18](https://doi.org/10.4103/ccij.ccij.108.18).
4. Seyfried T. Cancer as a metabolic disease: implications for novel therapeutics. Carcinogenesis. 2014 Mar; 35(3):515–527. DOI: [10.1093/carcin/bgt480](https://doi.org/10.1093/carcin/bgt480). PMID: [24343361](https://pubmed.ncbi.nlm.nih.gov/24343361/).
5. Castro-Mujica MC, Sullcahuamán-Allende Y, Barreda-Bolaños F, Taxa-Rojas L. Síndromes hereditarios de predisposición al cáncer colorrectal identificados en el Instituto Nacional de Enfermedades Neoplásicas (INEN), Lima, Perú. Rev Gastroenterol Peru. 2014;34(2):107-14.
6. Marchland et al. Association of sedentary lifestyle, obesity, smoking, alcohol use, and diabetes with the risk of colorectal cancer. Cancer Res. 1997 Nov 1;57(21):4787-94. PMID: [9354440](https://pubmed.ncbi.nlm.nih.gov/9354440/).
7. Warburg O. The metabolism of tumors in the body. J Gen Physiol. 1927 Mar 7; 8(6):519–530. DOI: [10.1085/jgp.8.6.519](https://doi.org/10.1085/jgp.8.6.519). PMID: [19872213](https://pubmed.ncbi.nlm.nih.gov/19872213/).
8. Seyfried T, Shelton L. Cancer as a metabolic disease. Nutr Metab (Lond). 2010 Jan 27. DOI: [10.1186/1743-7075-7-7](https://doi.org/10.1186/1743-7075-7-7). PMID: [20181022](https://pubmed.ncbi.nlm.nih.gov/20181022/).
9. Zachar Z, Marecek J, Maturo C. Non-redox-active lipoate derivatives disrupt cancer cell mitochondrial metabolism and are potent anticancer agents in vivo. J Mol Med (Berl). 2011 Nov;89(11):1137-48. DOI: [10.1007/s00109-011-0785-8](https://doi.org/10.1007/s00109-011-0785-8). PMID: [21769686](https://pubmed.ncbi.nlm.nih.gov/21769686/).
10. Jones N, Schulze A. Targeting cancer metabolism – aiming at a tumour's sweet-spot. Drug Discov Today. 2012 Mar;17(5-6):232-41. DOI: [10.1016/j.drudis.2011.12.017](https://doi.org/10.1016/j.drudis.2011.12.017). PMID: [22207221](https://pubmed.ncbi.nlm.nih.gov/22207221/).
11. Martinez VM, Smith S, Nader T, Bazile M, Aloyz R. Resistance to Dasatinib in primary chronic lymphocytic leukemia lymphocytes involves AMPK-mediated energetic reprogramming. Oncotarget. 2013 Dec; 4(12):2550–2566. DOI: [10.18632/oncotarget.1508](https://doi.org/10.18632/oncotarget.1508). PMID: [24334291](https://pubmed.ncbi.nlm.nih.gov/24334291/).
12. Durán R, 2019. Metabolismo del cáncer: comiendo con tu enemigo. Centro Andaluz de Biología Molecular y Medicina Regenerativa. DOI: [10.18567/sebbmdiv.RPC.2019.11.1](https://doi.org/10.18567/sebbmdiv.RPC.2019.11.1).
13. Alonso A, 2006. Mecanismos genéticos en la predisposición hereditaria al cáncer colorrectal. An. Sis San Navarra. 2006, vol.29, n.1, pp.59-76. ISSN 1137-6627.

14. Bianchini M. Comparative study of gene expression by cDNA microarray in human colorectal cancer tissues and normal mucosa. *Int J Oncol.* 2006 Jul;29(1):83-94. DOI: [10.3892/ijo.29.1.83](https://doi.org/10.3892/ijo.29.1.83). PMID: [16773188](https://pubmed.ncbi.nlm.nih.gov/16773188/).
15. Huxley RR. The impact of dietary and lifestyle risk factors on risk of colorectal cancer: A quantitative overview of the epidemiological evidence. *Int J Cancer.* 2009 Jul 1;125(1):171-80. DOI: [10.1002/ijc.24343](https://doi.org/10.1002/ijc.24343). PMID: [19350627](https://pubmed.ncbi.nlm.nih.gov/19350627/).
16. Richard SM, Martinez VM. Sensitization to oxaliplatin in HCT116 and HT29 cell lines by metformin and ribavirin and differences in response to mitochondrial glutaminase inhibition. *J Cancer Res Ther.* 2015 Apr-Jun;11(2):336-40. DOI: [10.4103/0973-1482.157317](https://doi.org/10.4103/0973-1482.157317). PMID: [26148596](https://pubmed.ncbi.nlm.nih.gov/26148596/).
17. Munkácsy G, 2018. Independent validation of induced overexpression efficiency across 242 experiments shows a success rate of 39%. *Scientific Reports* (2019) 9:343. DOI: [10.1038/s41598-018-36122-8](https://doi.org/10.1038/s41598-018-36122-8).
18. Mboge A.Y., Mahon B.P., McKenna I.H, Frost S.C. Carbonic Anhydrases: Role in pH Control and Cancer. *Metabolites.* 2018 Feb 28;8(1). pii: E19. DOI: [10.3390/metabo8010019](https://doi.org/10.3390/metabo8010019).
19. Schwientek T et al. Cloning of a Novel Member of the UDP-Galactose: b -N-Acetylglucosamine b 1,4-Galactosyltransferase Family, b 4Gal-T4, Involved in Glycosphingolipid Biosynthesis. *J. Biol. Chem.* 1998, 273:29331-29340. DOI: [10.1074/jbc.273.45.29331](https://doi.org/10.1074/jbc.273.45.29331).
20. Parsons W et al. AIG1 and ADTRP are Atypical Integral Membrane Hydrolases that Degrade Bioactive FAHFs. *Nat Chem Biol.* 2016 May; 12(5) 367–372. DOI: [10.1038/nchembio.2051](https://doi.org/10.1038/nchembio.2051).
21. Lupu C et al. Novel protein ADTRP regulates TFPI expression and function in human endothelial cells in normal conditions and in response to androgen. *Blood First Edition paper*, August 25, 2011; DOI [10.1182/blood-2011-05-355370](https://doi.org/10.1182/blood-2011-05-355370).
22. Álvarez M et al. Péptido YY: una nueva estrategia para el tratamiento de la obesidad. *Endocrinol Nutr.* 2009;56(6):317-30. DOI: [10.1016/S1575-0922\(09\)70186-X](https://doi.org/10.1016/S1575-0922(09)70186-X).
23. Missler M. Neurexophilins form a conserved family of neuropeptide-like glycoproteins. *The Journal of Neuroscience : the Official Journal of the Society for Neuroscience.* 1998 May 1;18(10):3630-3638. DOI: [10.1523/JNEUROSCI.18-10-03630.1998](https://doi.org/10.1523/JNEUROSCI.18-10-03630.1998). PMID: [9570794](https://pubmed.ncbi.nlm.nih.gov/9570794/).
24. Bonventre j et al, 2018. Fer1l6 is essential for the development of vertebrate muscle tissue in zebrafish. 2019 Feb 1;30(3):293-301. DOI: [10.1091/mbc.E18-06-0401](https://doi.org/10.1091/mbc.E18-06-0401).
25. Koslowski M. MS4A12 Is a Colon-Selective Store-Operated Calcium Channel Promoting Malignant Cell Processes. *Cancer Res.* 2008 May 1;68(9):3458-66. DOI: [10.1158/0008-5472.CAN-07-5768](https://doi.org/10.1158/0008-5472.CAN-07-5768). PMID: [18451174](https://pubmed.ncbi.nlm.nih.gov/18451174/).
26. GeneCards – VSIG2 Gene. <https://www.genecards.org/cgi-bin/carddisp.pl?gene=VSIG2&keywords=vsig2>.
27. MedlinePlus – EDN3 gene. <https://medlineplus.gov/genetics/gene/edn3/>.

28. Wang, Guanghui & Shen, Wei & Cui, Long & Chen, Wei & Hu, Xuguang & Fu, Jihong. (2016). Overexpression of Anillin (ANLN) is correlated with colorectal cancer progression and poor prognosis. *Cancer Biomarkers*. 16. 459-465. DOI: [10.3233/CBM-160585](https://doi.org/10.3233/CBM-160585).
29. Jia H, Song L, Cong Q, et al. The LIM protein AJUBA promotes colorectal cancer cell survival through suppression of JAK1/STAT1/IFIT2 network. *Oncogene*. 2017 May;36(19):2655-2666. DOI: [10.1038/onc.2016.418](https://doi.org/10.1038/onc.2016.418).
30. Ferguson H, Wragg J, Heath V, Bicknell R. Glutamate dependent NMDA receptor 2D is a novel angiogenic tumour endothelial marker in colorectal cancer. *Oncotarget*. 2016 Apr 12; 7(15): 20440–20454. DOI: [10.18632/oncotarget.7812](https://doi.org/10.18632/oncotarget.7812). PMID: [26943033](https://pubmed.ncbi.nlm.nih.gov/26943033/).
31. Pulkkinen V, Arola J. Neuropeptide S receptor 1 (NPSR1) activates cancer-related pathways and is widely expressed in neuroendocrine tumors. *Virchows Arch*. 2014; 465(2): 173–183. DOI: [10.1007/s00428-014-1602-x](https://doi.org/10.1007/s00428-014-1602-x). PMID: [24915894](https://pubmed.ncbi.nlm.nih.gov/24915894/).
32. Gorka B, Carnocka B. NrCAM, a neuronal system cell-adhesion molecule, is induced in papillary thyroid carcinomas. *British Journal of Cancer*. 2007; 97:531 – 538. DOI: [10.1038/sj.bjc.6603915](https://doi.org/10.1038/sj.bjc.6603915). PMID: [17667921](https://pubmed.ncbi.nlm.nih.gov/17667921/).
33. Fang Y, liang X. miR-424 targets AKT3 and PSAT1 and has a tumor-suppressive role in human colorectal cancer. *Cancer Manag Res*. 2018; 10: 6537–6547. DOI: [10.2147/CMAR.S185789](https://doi.org/10.2147/CMAR.S185789). PMID: [30555259](https://pubmed.ncbi.nlm.nih.gov/30555259/).
34. Foda AA. Role of MEK1 and DIAPH3 Expression in Colorectal Carcinoma. *Research in Oncology* 2018; Vol. 14,No. 2: 75-82. DOI: [10.21608/RESONCOL.2018.4042.1059](https://doi.org/10.21608/RESONCOL.2018.4042.1059).
35. Wu X, Deng F, Li Y, Daniels G. ACSL4 promotes prostate cancer growth, invasion and hormonal resistance. *Oncotarget*. 2015 Dec 29;6(42):44849-63. DOI: [10.18632/oncotarget.6438](https://doi.org/10.18632/oncotarget.6438). PMID: [26636648](https://pubmed.ncbi.nlm.nih.gov/26636648/).
36. Tang Y, Jiang YM, Lu GD. Fatty acid activation in carcinogenesis and cancer development: Essential roles of long-chain acyl-CoA synthetases.pdf. *Oncol Lett*. 2018 Aug; 16(2): 1390–1396. DOI: [10.3892/ol.2018.8843](https://doi.org/10.3892/ol.2018.8843). PMID: [30008815](https://pubmed.ncbi.nlm.nih.gov/30008815/).
37. Gao H, Chen X. Increased KIF4A expression is a potential prognostic factor in prostate cancer. *Oncol Lett*. 2018 May; 15(5): 7941–7947. DOI: [10.3892/ol.2018.8322](https://doi.org/10.3892/ol.2018.8322). PMID: [29725481](https://pubmed.ncbi.nlm.nih.gov/29725481/).
38. Nishio K, Kimura K. Preoperative predictors for early recurrence of resectable pancreatic cancer. *World J Surg Oncol*. 2017; 15:16. DOI: [10.1186/s12957-016-1078-z](https://doi.org/10.1186/s12957-016-1078-z). PMID: [28069033](https://pubmed.ncbi.nlm.nih.gov/28069033/).
39. Makawita et al. Validation of four candidate pancreatic cancer serological biomarkers that improve the performance of CA19.9. *BMC Cancer* 2013 13:404. DOI: [10.1186/1471-2407-13-404](https://doi.org/10.1186/1471-2407-13-404).
40. Astrosini C, Roefzaad C, Dai YY, Dieckgraefe BK, Jöns T, Kemmner W. REG1A expression is a prognostic marker in colorectal cancer and associated with peritoneal carcinomatosis. *Int J Cancer*. 2008;123(2):409-413. DOI: [10.1002/ijc.23466](https://doi.org/10.1002/ijc.23466).
41. Mojarad E, Zali M. The CpG island methylator phenotype (CIMP) in colorectal cancer. *Gastroenterol Hepatol Bed Bench*. 2013 Summer; 6(3): 120–128. PMID: [24834258](https://pubmed.ncbi.nlm.nih.gov/24834258/).

42. Duong HQ, Nemazanyy I. The Endosomal Protein CEMIP Links WNT Signaling to MEK1-ERK1/2 Activation in Selumetinib-Resistant Intestinal Organoids. *Cancer Res.* 2018 Aug 15;78(16):4533-4548. DOI: [10.1158/0008-5472.CAN-17-3149](https://doi.org/10.1158/0008-5472.CAN-17-3149). PMID: [29915160](https://pubmed.ncbi.nlm.nih.gov/29915160/).
43. Bruno G, Cencetti F. β 3-adrenoreceptor blockade reduces tumor growth and increases neuronal differentiation in neuroblastoma via SK2-S1P 2 modulation. *Oncogene* (2020) 39:368–384. DOI: [10.1038/s41388-019-0993-1](https://doi.org/10.1038/s41388-019-0993-1).
44. Wu J, Shih J, Chung L. Monoamine oxidase A mediates prostate tumorigenesis and cancer metastasis. *J Clin Invest.* 2014;124(7):2891-2908. DOI: [10.1172/JCI70982](https://doi.org/10.1172/JCI70982).
45. Yang Y, Chien M, Chen C. Monoamine Oxidase B Expression Correlates with a Poor Prognosis in Colorectal Cancer Patients and Is Significantly Associated with Epithelial-to-Mesenchymal Transition-Related Gene Signatures. *Int. J. Mol. Sci.* 2020, 21, 2813; DOI: [10.3390/ijms21082813](https://doi.org/10.3390/ijms21082813).
46. Gan X, Wang J, Xi Y, Li L. Nuclear Dvl, c-Jun, β -catenin, and TCF form a complex leading to stabilization of β -catenin–TCF interaction. *The Journal of Cell Biology*, Vol. 180, No. 6, March 24, 2008 1087–1100. DOI: [10.1083/jcb.200710050](https://doi.org/10.1083/jcb.200710050).
47. “Metabolism of carbohydrates”. Reactome, 2003-07-03, <https://reactome.org/content/detail/R-HSA-71387>.
48. “Hyaluronan metabolism”. Reactome, 2012-02-24, <https://reactome.org/content/detail/R-HSA-2142845>.
49. Williams, Mark. Glycerophospholipid Biosynthesis. Reactome - a curated knowledgebase of biological pathways (2011). DOI: [10.3180/REACT_121401.1](https://doi.org/10.3180/REACT_121401.1).
50. Aleena KS et al. GPCRs in Cancer: Protease-Activated Receptors. *Int J Mol Sci.* 2018 Jul; 19(7): 1886. DOI: [10.3390/ijms19071886](https://doi.org/10.3390/ijms19071886).
51. Lappano R, Maggiolini M. GPCRs and cancer. *Acta Pharmacol Sin.* 2012 Mar; 33(3): 351–362. DOI: [10.1038/aps.2011.183](https://doi.org/10.1038/aps.2011.183). PMID: [22266725](https://pubmed.ncbi.nlm.nih.gov/22266725/).
52. Wong W, Shen T, Dong B, Zhou W, Shi Q, Zhang Y. MAPK4 overexpression promotes tumor progression via noncanonical activation of AKT/mTOR signaling. *J Clin Invest.* 2019 Mar 1;129(3):1015-1029. DOI: [10.1172/JCI97712](https://doi.org/10.1172/JCI97712). PMID: [30688659](https://pubmed.ncbi.nlm.nih.gov/30688659/).
53. “Metabolism”. Reactome, 2011-07-07, <https://reactome.org/content/detail/R-HSA-1430728>.
54. Medline Plus – PAH gene – <https://medlineplus.gov/genetics/gene/pah/>.
55. Wilson M et al. Higher baseline expression of the PTGS2 gene and greater decreases in total colonic fatty acid content predict greater decreases in colonic prostaglandin-E2 concentrations after dietary supplementation with ω -3 fatty acids. November 01, 2018. DOI: [10.1016/j.plefa.2018.11.001](https://doi.org/10.1016/j.plefa.2018.11.001).
56. Medline Plus – FMO3 gene – <https://medlineplus.gov/genetics/gene/fmo3/>.

57. Wollenick K et al. Synthetic transactivation screening reveals ETV4 as broad coactivator of hypoxia-inducible factor signaling. *Nucleic Acids Research*, 2012, Vol. 40, No. 5. DOI: [10.1093/nar/gkr978](https://doi.org/10.1093/nar/gkr978).
58. GeneCards – COL11A1 Gene – <https://www.genecards.org/cgi-bin/carddisp.pl?gene=COL11A1>.
59. UniProt – ABHD3 – <https://www.uniprot.org/uniprot/Q91ZH7>.
60. MedlinePlus – GRIN2B gene – <https://medlineplus.gov/genetics/gene/grin2b/>.
61. GeneCards – ACSL4 Gene – <https://www.genecards.org/cgi-bin/carddisp.pl?gene=ACSL4>.
62. UniProt – PSAT1– <https://www.uniprot.org/uniprot/Q9Y617>.
63. GeneCards – PTPN13 Gene – <https://www.genecards.org/cgi-bin/carddisp.pl?gene=PTPN13>.
64. Hemler M. Tetraspanin proteins promote multiple cancer stages. *Nat Rev Cancer*. 2014 Jan;14(1):49-60. DOI: [10.1038/nrc3640](https://doi.org/10.1038/nrc3640).
65. Yu b et al. The Long Non-Coding RNA CRNDE Promotes Colorectal Carcinoma Progression by Competitively Binding miR-217 with TCF7L2 and Enhancing the Wnt/E-Catenin Signaling Pathway. *Cell Physiol Biochem* 2017;41:2489-2502. DOI: [10.1159/000475941](https://doi.org/10.1159/000475941).
66. Polistena A. MMP7 Expression in Colorectal Tumours of Different Stages. January-February 2014 vol. 28 no. 1 105-11. PMID: [24425843](https://pubmed.ncbi.nlm.nih.gov/24425843/).
67. Brown, R. E., Short, S. P., & Williams, C. S. (2018). Colorectal Cancer and Metabolism. *Current colorectal cancer reports*, 14(6), 226–241. DOI: [10.1007/s11888-018-0420-y](https://doi.org/10.1007/s11888-018-0420-y).
68. Cochón, Adriana Cristina. (1995). Porfiria y Cáncer: Antineoplásicos como agentes porfirinogénicos; influencia de tumores sobre el desencadenamiento de porfiria por hexaclorobenceno. Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires. http://digital.bl.fcen.uba.ar/Download/Tesis/Tesis_2779_Cochon.pdf.
69. Medrare A, Santos E. Ras in Cancer and Developmental Diseases. *Genes & Cancer* (2011) 2(3) 344-358. DOI: [10.1177/1947601911411084](https://doi.org/10.1177/1947601911411084).
70. Liu S, Ahmed AA. The structural basis of PI3K cancer mutations: from mechanism to therapy. *Cancer Res*. 2014 Feb 1; 74(3): 641–646. DOI: [10.1158/0008-5472.CAN-13-2319](https://doi.org/10.1158/0008-5472.CAN-13-2319). PMID: [24459181](https://pubmed.ncbi.nlm.nih.gov/24459181/).
71. Johnson SM et al. Novel Expression Patterns of PI3K-AKT-mTOR Signaling Pathway Components in Colorectal Cancer. *J Am Coll Surg*. 2010 May ; 210(5): 767–778. DOI: [10.1016/j.jamcollsurg.2009.12.008](https://doi.org/10.1016/j.jamcollsurg.2009.12.008).
72. Li S, Knapp S, Ahmed AA. Targeting AMPK for cancer prevention and treatment. *Cancer Res*. 2014 Feb 1; 74(3): 641–646. DOI: [10.18632/oncotarget.3629](https://doi.org/10.18632/oncotarget.3629). PMID: [25812084](https://pubmed.ncbi.nlm.nih.gov/25812084/).

73. Shackelford D, Shaw R. The LKB1-AMPK pathway: metabolism and growth control in tumor suppression. *Nat Rev Cancer*. 2009 Aug;9(8):563-75. DOI: [10.1038/nrc2676](https://doi.org/10.1038/nrc2676). PMID: [19629071](https://pubmed.ncbi.nlm.nih.gov/19629071/).
74. Yang J, Fang Y, Lou H. Use of Metformin Alone Is Not Associated with Survival Outcomes of Colorectal Cancer Cell but AMPK Activator AICAR Sensitizes Anticancer Effect of 5-Fluorouracil through AMPK Activation. *PLoS One*. 2014 May 21;9(5):e97781. DOI: [10.1371/journal.pone.0097781](https://doi.org/10.1371/journal.pone.0097781). PMID: [24849329](https://pubmed.ncbi.nlm.nih.gov/24849329/).
75. Bionaz M, Loor JJ. A novel dynamic impact approach (DIA) for functional analysis of time-course omics studies: validation using the bovine mammary transcriptome. *PLoS One*. 2012;7(3):e32455. DOI: [10.1371/journal.pone.0032455](https://doi.org/10.1371/journal.pone.0032455). PMID: [22438877](https://pubmed.ncbi.nlm.nih.gov/22438877/).
76. Liang & Li, 2016. Identification of key pathways and genes in colorectal cancer using bioinformatics analysis. *Med Oncol* (2016) 33:111. DOI: [10.1007/s12032-016-0829-6](https://doi.org/10.1007/s12032-016-0829-6).
77. Li W, Huang Y, Sargsyan, D. et al. Epigenetic alterations in TRAMP mice: epigenome DNA methylation profiling using MeDIP-seq. *Cell Biosci* (2018) 8:3. DOI: [10.1186/s13578-018-0201-y](https://doi.org/10.1186/s13578-018-0201-y).