

INTERACCIONES DE PEROXISOME PROLIFERATOR-ACTIVATED  
RECEPTOR GAMMA COACTIVATOR  $\gamma$  1 (PGC $\gamma$ 1) Y FACTORES  
MEDIOAMBIENTALES EN EL DESARROLLO DEL  
CÁNCER COLORRECTAL

Biomedicina y ciencias de la salud



Jéssica Alonso Molero

Departamento de Ciencias Biomédicas  
Área de Medicina Preventiva  
y Salud Pública

Universidad de León

2018





universidad  
de león

---

*Interacciones de Peroxisome  
Proliferator-Activated Receptor  $\gamma$   
Coactivator 1- $\alpha$  y factores ambientales  
en el desarrollo del cáncer colorrectal*

Programa de doctorado de Biomedicina  
y Ciencias de la Salud

---

**Jéssica Alonso Molero**

**Directores:**

Vicente Martín Sánchez  
Antonio José Molina de la Torre

León, Junio de 2018





universidad  
de león

Escuela de Doctorado

D./D<sup>a</sup> Jéssica Alonso Molero declara que la memoria de la tesis presentada bajo el título "Interacciones de Peroxisome Proliferator-Activated Receptor Gamma Coactivator 1 (PGC1) y factores medioambientales en el desarrollo del cáncer colorrectal" es, conforme al artículo 13.1 del R.D. 99/2011, de 28 de enero, un trabajo original de investigación, sin contribución significativa de otra persona que no aparezca reflejada en la misma, y citando adecuadamente la procedencia del contenido no original, conforme a la normativa vigente.

Asimismo, declaro que este trabajo no ha sido presentado y no lo será en el futuro como tesis doctoral, en ninguna universidad o institución de investigación, en España o en el extranjero.

Entiendo la política de tolerancia cero frente al plagio de la Universidad de León, la cual se reserva el derecho a retirar mi título de doctor y adoptar cuantas medidas procedan legalmente, en caso de incumplimiento de este compromiso.

León a 12 de junio de 2018

Fdo.: Jéssica Alonso Molero



## INFORME DEL TUTOR

El Dr. D. Vicente Martín Sánchez como Tutor de la Tesis Doctoral titulada "Interacciones de Peroxisome Proliferator-Activated Receptor Gamma Coactivator 1 (PGC1) y factores medioambientales en el desarrollo del cáncer colorrectal" realizada por D. Jéssica Alonso Molero en el programa de doctorado Biomedicina y ciencias de la salud, regulado por el R.D. 99/2011, de 28 de enero, informa favorablemente el depósito de la misma, dado que reúne las condiciones necesarias para su defensa.

Lo que firmo, en León a 12 de junio de 2018



*“No dejes que las mentes pequeñas te digan que tus sueños son demasiado grandes.”*

**Iván Alonso**



La presente Tesis Doctoral ha sido posible gracias al estudio MCC-Spain, el cuál ha recibido financiación de: “Acción Transversal del Cáncer”, aprobada por el Consejo de Ministros de España el 11 de octubre del 2007; por el Instituto de Salud Carlos III-FEDER (PI08/1770, PI08/0533, PI08/1359, PI09/00773-Cantabria, PI09/01286-León, PI09/01903-Valencia, PI09/02078-Huelva, PI09/01662-Granada, PI11/01403, PI11/01889-FEDER, PI11/00226, PI11/01810, PI11/02213, PI12/00488, PI12/00265, PI12/01270, PI12/00715, PI12/00150, PI14/01219, PI14/0613); por la Fundación Marqués de Valdecilla (API 10/09); por el ICGC International Cancer Genome Consortium CLL (The ICGC CLL-Genome Project is funded by Spanish Ministerio de Economía y Competitividad (MINECO) a través del Instituto de Salud Carlos III (ISCIII) y de la Red Temática de Investigación del Cáncer (RTICC) del ISCIII (RD12/0036/0036)); por la Junta de Castilla y León (LE22A10-2); por la Consejería de Salud de la Junta de Andalucía (2009-S0143); por la Conselleria de Sanitat of the Generalitat Valenciana (*AP*<sub>0</sub>61/10) y por la Recercaixa (2010ACUP 00310); por el Regional Government of the Basque Country; por la Consejería de Sanidad de la Región de Murcia; por la European Commission grants FOOD-CT-2006-036224-HIWATE; por la Asociación Española contra el cáncer (AECC); por el Catalan Government DURSI grant 2014SGR647; por la Fundación Caja de Ahorros de Asturias; y por la Universidad de Oviedo.

En el proceso de realización de esta tesis, he sido autora de dos artículos: “The RS4939827 polymorphism in the SMAD7 GENE and its association with Mediterranean diet in colorectal carcinogenesis.”, publicado en la revista BMC Medical Genetics (Q3 con un FI=2.198 en 2016) y “Alterations in PGC1 $\alpha$  expression levels are involved in colorectal cancer risk: a qualitative systematic review.”, publicado en BMC Cancer (Q2 con un FI=3.288 en 2016) recogidos en la sección de Anexos (Anexo I y Anexo II respectivamente). Además, he participado como coautora en la elaboración de otros cuatro artículos más: “The NRF2 transcription factor plays a dual role in colorectal cancer: A systematic review.”, publicado en PlosOne (Q1 con un FI=2.806 en 2016); “Reproductive risk factors in breast cancer and genetic hormonal pathways: a gene-environment interaction in the MCC-Spain project.”, publicado en BMC Cancer; “Validating a breast cancer score in Spanish women. The MCC-Spain study.” publicado en Scientific Reports (Q1 con un FI=4.259 en 2016) y “PGC-1 $\alpha$  as a biomarker of physical activity effect on colorectal cancer”, publicado en Cancer Prevention Research (Q2 con un FI=3.985 en 2016).

# Agradecimientos

---

Al finalizar cada una de las etapas de la vida, como es el desarrollo de una tesis doctoral, es costumbre hacer recuento de todos los momentos vividos en dicha época. Es en ese momento cuando nos damos cuenta de que, sin la participación de determinadas personas e instituciones, el camino sería imposible de recorrer. Por ese motivo, este espacio va dedicado a todos aquellos que, de un modo u otro, han estado presentes y han aportado algo en el desarrollo de este trabajo.

Primero, agradezco al Profesor Vicente Martín Sánchez y al Doctor Antonio José Molina de la Torre por permitirme realizar esta tesis bajo su dirección. Su confianza en mi trabajo ha permitido mi desarrollo como investigadora y mi formación como profesional en este campo. Además, debo agradecerles el haberme facilitado los medios suficientes para llevar a cabo las actividades propuestas. A este último agradecimiento debo sumar también a todos los nodos participantes en el proyecto MCC-Spain puesto que, sin sus datos, nada de esto hubiese sido posible.

También debo agradecer a mi nuevo equipo de investigación su acogida, su comprensión y su apoyo durante estos últimos años. Siempre han estado disponibles y nunca han dudado en compartir su experiencia y conocimiento conmigo. Además, debo añadir a este agradecimiento al casi Doctor Camilo Palazuelos Calderón y, a la que espero que lo sea pronto, María Fernández Ortiz, por su extremada paciencia y ayuda, por todas esas horas delante de la pizarra o de la pantalla del ordenador y por esos momentos de sangre, sudor y lágrimas, ellos mejor que nadie entienden y saben los motivos de mi agradecimiento.

A mi queridísimo equipo de investigación de León en general y en especial a Carmen González Donquiles, por ayudarme a ser una persona más fuerte y a crecer en diferentes aspectos de mi vida. También debo nombrar con privativo interés a Tania Fernández Villa por su paciencia y su capacidad para estar pendiente de todo y siempre con una sonrisa. Por supuesto, tampoco puede faltar Laura Vilorio Marqués quien, con sus acalorados, pero siempre fructuosos debates, me ha llevado a conocer en mayor profundidad el tema que en la presente tesis se debate.

Quiero recordar también a todos mis compañeros de doctorado, de carrera y de máster, compañeros de los diversos grupos de investigación por los que he pasado, del Laboratorio de Técnicas Instrumentales de la universidad de León y todas las posibles personas metidas en este mundo que es la investigación y que han hecho que, de un modo u otro, yo esté hoy donde estoy, consiguiendo este gran objetivo que es el depósito de mi tesis. Todos ellos han aportado un granito de arena en mis días y en mi conocimiento, me han ayudado a luchar por lo que quiero y han conseguido que dé lo mejor de mí. También debo agradecer con las mismas

palabras a otros colectivos de personas, más considerados como amigos, que no puedo nombrar uno a uno por su extenso número, pero que estoy segura de que sabrán quiénes son simplemente por llegar a leer este párrafo.

Por último, pero posiblemente el agradecimiento más importante, se lo tengo que dar a mi querida familia. A mis padres, Jose y Horten, porque con su apoyo y constancia, su continua preocupación por mi bienestar y sus largas horas escuchándome hablar de cosas que casi acabaron entendiendo, han hecho de mí la persona que hoy soy. A mi hermano Iván quien, a pesar de la distancia física, siempre ha estado ahí cuando lo he necesitado y quien, hoy en día, se ha convertido en un referente para mí y para la mitad de la población española. A mis abuelos, Pepe y Herminia y Gerardo y Josefina, por el orgullo que han mostrado siempre de su nieta. Y, por supuesto, a Jonatan Álvarez Quiñones, mi compañero de vida, esa persona que me sigue en todas mis locuras, que las apoya y las secunda, que nunca me abandona, que siempre me ha enseñado a levantarme sola cuando me he caído y sigue a mi lado por si las fuerzas vuelven a flaquear. Pero, sobre todo, esa persona que acepta mis logros como si fuesen propios, y es que en realidad lo son. Todo esto ha sido posible gracias a vosotros.

# Índice general

---

<b>1</b>	<b>Introducción</b>	<b>18</b>
1.1	El cáncer . . . . .	19
1.1.1	Metabolismo energético en el cáncer . . . . .	21
1.1.2	Las mitocondrias en el metabolismo energético . . . . .	22
1.2	El cáncer colorrectal . . . . .	24
1.2.1	Factores de riesgo del CCR . . . . .	26
1.3	Papel de la genética en el cáncer colorrectal . . . . .	27
1.4	Función de los SNPs en el cáncer colorrectal . . . . .	31
1.5	Métodos de análisis de los SNPs . . . . .	32
1.5.1	Análisis SNP a SNP . . . . .	32
1.5.2	Análisis de conjuntos de genes . . . . .	33
1.5.3	Bases de las técnicas de análisis de conjuntos de SNPs . . . . .	34
1.5.4	Técnicas de análisis de conjuntos de SNPs . . . . .	35
1.6	PGC1 $\alpha$ . . . . .	38
1.6.1	Papel del PGC1 $\alpha$ en la renovación del epitelio intestinal . . . . .	38
1.6.2	Características de PGC1 $\alpha$ , gen y proteína . . . . .	39
1.6.3	Funciones principales del PGC1 $\alpha$ . . . . .	41
1.6.4	Rutas metabólicas en las que participa PGC1 $\alpha$ . . . . .	42
<b>2</b>	<b>Objetivos</b>	<b>52</b>
<b>3</b>	<b>Mapas conceptuales de la función del PGC1<math>\alpha</math> en el CCR</b>	<b>54</b>
3.1	Introducción . . . . .	55
3.2	Material y métodos . . . . .	55
3.3	Resultados . . . . .	58
3.4	Discusión . . . . .	73
<b>4</b>	<b>Estudio de la asociación entre los SNPs del PGC1<math>\alpha</math> y el CCR</b>	<b>78</b>
4.1	Introducción . . . . .	79
4.2	Material y métodos . . . . .	80
4.3	Resultados . . . . .	82
4.4	Discusión . . . . .	87

<b>5</b>	<b>Comparación de las técnicas de análisis de conjuntos de genes</b>	<b>90</b>
5.1	Introducción . . . . .	91
5.2	Material y métodos . . . . .	91
5.3	Resultados . . . . .	95
5.4	Discusión . . . . .	100
<b>6</b>	<b>Análisis de rutas metabólicas en asociación con el CCR</b>	<b>104</b>
6.1	Introducción . . . . .	105
6.2	Material y métodos . . . . .	105
6.3	Resultados . . . . .	106
6.4	Discusión . . . . .	109
<b>7</b>	<b>Corolario</b>	<b>112</b>
<b>8</b>	<b>Conclusiones</b>	<b>117</b>
	<b>Bibliografía</b>	<b>120</b>



# Índice de figuras

---

Figura 1.	Fases del ciclo celular [3] . . . . .	19
Figura 2.	Fases del desarrollo del cáncer colorrectal por la inestabilidad cromosómica [63]. . . . .	29
Figura 3.	Estructura de la proteína PGC1 $\alpha$ . . . . .	40
Figura 4.	Ruta de señalización de la AMPK (KEGG). . . . .	43
Figura 5.	Ruta metabólica de señalización del glucagón (KEGG) . . . . .	44
Figura 6.	Ruta metabólica de regulación de la longevidad (KEGG) . . . . .	45
Figura 7.	Ruta de señalización de la insulina (KEGG) . . . . .	46
Figura 8.	Ruta de señalización de la adipocitoquina (KEGG) . . . . .	47
Figura 9.	Ruta metabólica de resistencia a la insulina (KEGG). . . . .	48
Figura 10.	Ruta metabólica de la Enfermedad de Huntington (KEGG) . . . . .	49
Figura 11.	Ruta metabólica de señalización de la Apelina (KEGG) . . . . .	50
Figura 12.	Resumen del proceso de selección de artículos. . . . .	59
Figura 13.	Evaluación de la homogeneidad entre los artículos incluidos en la revisión sistemática. . . . .	66
Figura 14.	Mapa conceptual basado en el trabajo de D'Errico et al. [30]. . . . .	68
Figura 15.	Mapa conceptual general. . . . .	69
Figura 16.	Esquema del flujo de trabajo para obtener los conjuntos de genes aleatorios. . . . .	93
Figura 17.	Capacidad de discriminación de cada técnica. . . . .	96
Figura 18.	Capacidad de discriminación de la selección de conjuntos de genes, considerando GSA-SNP(5). . . . .	97
Figura 19.	Diagrama de Venn: Conjuntos de genes significativos en la combinación de las técnicas. . . . .	98
Figura 20.	Influencia de la combinación de las técnicas en la capacidad de discriminación de los conjuntos de genes. . . . .	99
Figura 21.	Ratios de falsos positivos (A) y falsos negativos (B) para cada una de las técnicas. . . . .	100

# Índice de figuras

---

Tabla 1.	Características principales de las técnicas descritas. . . . .	37
Tabla 2.	Resumen de los conceptos clave de los estudios seleccionados . . . .	60
Tabla 3.	Materiales empleados en los artículos incluidos y sus características. . . . .	71
Tabla 4.	Caracterización según las variables de ajuste categóricas de los individuos participantes en el estudio. . . . .	82
Tabla 5.	Caracterización según las variables de ajuste cualitativas de los individuos participantes en el estudio. . . . .	83
Tabla 6.	Asociación entre el rs12505641 y el riesgo de padecer CCR. . . . .	84
Tabla 7.	Resultados del análisis de regresión logística para el SNP rs12505641 y el riesgo de desarrollar CCR estratificado por la localización del tumor. . . . .	85
Tabla 8.	Análisis de regresión logística del rs12505641 y el cáncer de recto estratificado por sexo y considerando diferentes factores de riesgo. . . .	86
Tabla 9.	Número de conjuntos de genes considerados significativos para cada técnica. . . . .	95
Tabla 10.	Comparación entre GSA-SNP considerando dos SNPs por gen (GSA-SNP(2)) y cinco SNPs por gen (GSA-SNP(5)). . . . .	97
Tabla 11.	Pathways en los que participa PGC1 $\alpha$ . . . . .	106
Tabla 12.	Resumen de los resultados de todos los métodos empleados para el consenso. . . . .	108

## Referencias de siglas y abreviaturas

---

$\Delta\Psi_m$	.....	Diferencia del potencial de membrana mitocondrial
<b>ACBP</b>	.....	Acyl CoEnzima-A binding protein
<b>ADIPOQ</b>	.....	Adiponectina
<b>ADNmt</b>	.....	ADN mitocondrial
<b>AFPG</b>	.....	Antecedentes familiares de primer grado
<b>AINES</b>	.....	Antinflamatorios no esteroideos
<b>Akt</b>	.....	AKT serine/threonine kinase
<b>AMPK</b>	.....	AMP-activated protein kinase
<b>APJ</b>	.....	Receptores de la Apelina
<b>ARNm</b>	.....	ARN mensajero
<b>ATP</b>	.....	Adenosín trifosfato
<b>BMP</b>	.....	Bone morphogenetic proteins
<b>CAT</b>	.....	Catalasa
<b>CaMK</b>	.....	Calcium/calmodulin dependent protein kinase II gamma
<b>CCR</b>	.....	Cáncer colorrectal
<b>CDC42P6</b>	....	Cell division cycle 42 pseudogene 6
<b>CLL</b>	.....	Leucemia crónica linfocítica
<b>CREB</b>	.....	cAMP response element binding
<b>CU</b>	.....	Colitis Ulcerosa
<b>ER</b>	.....	Elementos reguladores
<b>ERR<math>\alpha</math></b>	.....	Estrogen related receptor $\alpha$
<b>ES</b>	.....	Enrichment Score
<b>FAP</b>	.....	Poliposis adenomatosa familiar
<b>FDR</b>	.....	False discovery rate
<b>Fox01</b>	.....	Forkhead box O1

<b>FWER</b> .....	Family-wise error rate
<b>GLUT4</b> .....	Solute carrier family 2 member 4
<b>GSA</b> .....	Gene set analysis
<b>GSEA</b> .....	Gene set enrichment analysis
<b>G6Pasa</b> .....	Glucosa-6-Fosfatasa
<b>GWAS</b> .....	Genome-Wide Association Study
<b>HCF</b> .....	Host Cell Factor-1
<b>HWE</b> .....	Equilibrio de Hardy-Weimberg
<b>IC95</b> .....	Intervalo de Confianza del 95
<b>IMC</b> .....	Índice de masa corporal
<b>IRS1</b> .....	Insulin receptor substrate 1
<b>Kb</b> .....	Kilobases
<b>KEGG</b> .....	Kyoto encyclopedia of genes and genomes
<b>KSR1</b> .....	Kinase Supressor of RAS1
<b>MAF</b> .....	Frecuencia del menor alelo
<b>MAPK</b> .....	Mitogen-activated protein kinase
<b>Na-B</b> .....	Butirato de Sodio
<b>NES</b> .....	Enrichment Score normalizado
<b>NR</b> .....	Nuclear receptors
<b>NRF</b> .....	Nuclear Respiratory Factors
<b>OR</b> .....	Odds Ratio
<b>OXPHOS</b> ....	Fosforilación oxidativa
<b>pb</b> .....	Pares de bases
<b>PGC1<math>\alpha</math></b> .....	Peroxisome proliferator-activated receptor $\gamma$ coactivator 1 $\alpha$
<b>PGC1<math>\beta</math></b> .....	Peroxisome proliferator-activated receptor $\gamma$ coactivator 1 $\beta$

<b>PMC</b> .....	PubMed Central
<b>PRC</b> .....	PGC1 related coactivator
<b>ROS</b> .....	Especies reactivas de oxígeno
<b>rSNP</b> .....	SNP de región reguladora
<b>SD</b> .....	Desviación estándar
<b>SIRT1</b> .....	Sirtuin 1
<b>SNP</b> .....	Single nucleotide polymorphism
<b>SRC1</b> .....	SRC proto-oncogene, non-receptor tyrosine kinase
<b>SOD2</b> .....	Superóxido dismutasa
<b>Sp1</b> .....	Specificity protein 1
<b>SRT</b> .....	SNP ratio test
<b>TAF4</b> .....	TATA-box binding protein associated factor 4
<b>TGF-<math>\beta</math></b> .....	Transforming growth factor beta 1
<b>UC</b> .....	Colitis ulcerosa
<b>Upasa</b> .....	Uridin fosforilasa
<b>Wnt</b> .....	Wingless
<b>WOS</b> .....	Web of Science





CAPÍTULO 1

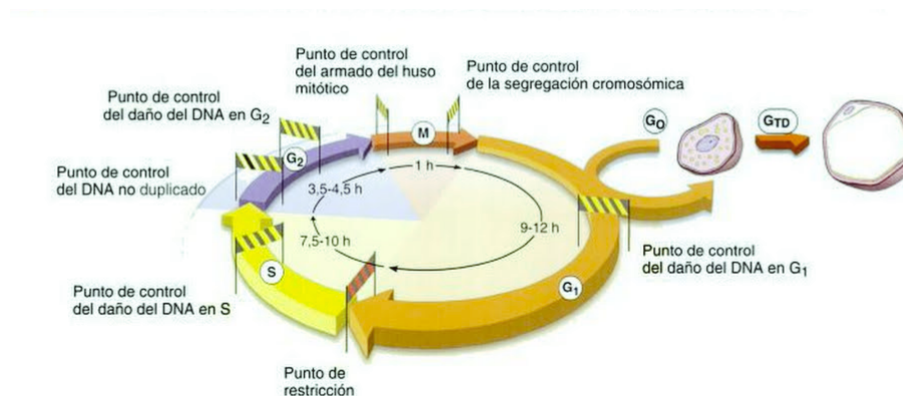
# Introducción

---

## 1.1. El cáncer

La palabra cáncer hace referencia a un conjunto de enfermedades que se caracterizan por una división celular descontrolada. Generalmente, el proceso de desarrollo de un tumor, conocido como carcinogénesis, se inicia a partir de una única célula o un grupo pequeño de células que, para llegar a ser cancerosas, han acumulado múltiples cambios genéticos en efectores que regulan el ciclo celular [1].

En condiciones normales, el ciclo celular está formado por dos fases principales, la *fase S* en la que se lleva a cabo la duplicación del ADN y la síntesis de proteínas y estructuras celulares y la *fase M* en la que se realiza tanto la mitosis como la meiosis. Estas dos fases están separadas por otras dos, conocidas como *fase G1* y *fase G2* respectivamente (Ver Figura 1). Cada cambio de fase está regulado por un punto de control denominado punto de restricción, en los que los efectores del ciclo celular, como CDK, RB o p53, entre otros, comprueban que la célula tiene las aptitudes necesarias para cambiar de fase. En caso contrario, producen su detención hasta que las células están preparadas para continuar el ciclo o, de no ser así, para provocar la apoptosis celular [2].



**Figura 1: Fases del ciclo celular [3].** El ciclo celular está formado por cuatro fases, la fase S, que dura aproximadamente entre 7 y 10h. La fase M que dura 1h y las fases intermedias entre ellas, denominadas como G1 (9-12h de duración) y G2 (3-4h de duración). En todas ellas, existen puntos de control que permiten comprobar el estado de las células, para permitir su división o, por el contrario, detenerla [3].

Sin embargo, cuando los efectores que controlan el ciclo celular están dañados, las células entran en un estado de división continua, dando lugar a células hijas que también acumulan alteraciones y que, salvo que sean detectadas por el sistema inmune, seguirán dividiéndose y acumulando mutaciones que cada vez las harán más agresivas. Para llegar a este punto, las células cancerígenas han tenido que ser capaces de violar tres aspectos del ciclo celular [1]:

## 1 . Introducción

---

1. Activar la división celular permanentemente, sin esperar a recibir las señales apropiadas. Para ello, se debe limitar la acción de los reguladores implicados en el control negativo de la progresión del ciclo celular. Esta inactivación conduce a la proliferación celular acelerada y no controlada [4, 5].
2. Inhibir el programa de apoptosis. La muerte celular se lleva a cabo cuando en la célula se detectan altos niveles de estrés o condiciones anormales de replicación del ADN, evitando así la acumulación de daños en sus genes. Sin embargo, las células cancerígenas limitan la actividad de los reguladores implicados en el acoplamiento entre la integridad del genoma y el ciclo celular, dando como resultado células que tienen alteraciones genéticas que se acumulan progresivamente durante la carcinogénesis [4, 5].
3. Sortear su “mecanismo de recuento de divisiones celulares”. Las células presentan un sistema de control de divisiones basado en las estructuras teloméricas del cromosoma, de tal modo que se pueden dividir un número limitado de veces. En cada proceso de división, los telómeros se acortan y, cuando son demasiado pequeños, las células entran en apoptosis. Para soslayar este control, las células cancerígenas deben activar una enzima denominada telomerasa, capaz de reemplazar las secuencias teloméricas perdidas durante la división celular, por lo que aumenta el número de divisiones celulares [6].

Cabe destacar que quebrantar estos tres pasos no es fácil. La maquinaria de regulación del ciclo celular presenta un alto grado de redundancia y solapamiento en sus funciones, por lo que la desregulación del ciclo celular en cualquiera de sus puntos requiere de la combinación de un gran número de alteraciones en los genes que codifican para proteínas encargadas de controlar el ciclo celular [4, 5]. Sin embargo, una vez quebrantados estos pasos, los tumores forman masas celulares desestructuradas que, incluso, tienen la capacidad para invadir el tejido normal adyacente [7, 8]. En base a esta invasión, los cánceres se pueden diferenciar en malignos o benignos, siendo los primeros los únicos capaces de mostrar esta característica.

Tras el inicio del desarrollo de la mayoría de los tumores, cada una de las enfermedades que forman el cáncer se caracterizará por la localización en la que se desarrolla y por otras anomalías genéticas acumuladas. Estas alteraciones producirán cambios en la expresión o en las funciones génicas que controlan a las células, proporcionando ventajas a las células cancerígenas, no sólo en su capacidad de crecimiento y de división, sino que también en su metabolismo energético, en su habilidad para comunicarse con las células vecinas o en su capacidad para

sortear al sistema inmune. En este sentido, cada tumor es personal, es decir, presenta una combinación única de cambios genéticos. Incluso en una misma masa tumoral diferentes conjuntos de células pueden mostrar diversidad en su genética [8].

### 1.1.1. Metabolismo energético en el cáncer

Existen evidencias de que los tumores sufren “olas” de reprogramación de la expresión génica, capaces de alterar su fenotipo metabólico durante la transformación maligna. Esta reprogramación genética puede dar lugar a diferentes tipos celulares en cuanto a su metabolismo energético, permitiendo diferenciar entre células tumorales que dependen totalmente del proceso de fosforilación oxidativa (OXPHOS) para la obtención de energía, células que consiguen ésta solamente por el proceso de glicólisis, o células que dependen parcialmente de ambas rutas [9]. Estas alteraciones son comunes en los tumores sólidos y se basan principalmente en la capacidad que presentan las células cancerígenas para suprimir las diferencias específicas de tejido en cuanto al fenotipo bioenergético de la mitocondria. De este modo, el metabolismo energético representa uno de los objetivos más prometedores para el tratamiento de estas enfermedades [9-11]. Esta hipótesis de trabajo fue planteada, por primera vez, por Warburg y Cori en 1924 [11, 12], quienes consiguieron demostrar que las células cancerígenas toman grandes cantidades de glucosa y excretan lactato [11].

En condiciones normales, en el citosol celular, las moléculas de glucosa son transformadas en piruvato mediante la glicólisis. Posteriormente, dicho piruvato viaja al interior de las mitocondrias para pasar a formar parte del ciclo de Krebs (o ciclo de los ácidos cítricos), donde se degrada hasta formar dióxido de carbono, obteniendo así 38 moléculas de energía en forma de ATP (Adenosín Trifosfato). Sin embargo, bajo condiciones anaeróbicas o de hipoxia, las células normales están limitadas en la ejecución de la glicólisis, por lo que deciden convertir la glucosa en lactato, el cual es secretado al espacio extracelular. Este es el proceso por el que se decantan las células cancerígenas, aún en presencia de grandes cantidades de oxígeno, según plantean Warburg y Cori [12]. Esta conversión a lactato muestra una serie de ventajas para las células cancerígenas, puesto que el lactato es un buen intermediario en numerosos procesos metabólicos, su rendimiento energético es similar al de la glicólisis y es capaz de estabilizar el factor inductor de la hipoxia y de aumentar la expresión del factor de crecimiento endotelial vascular [13, 14].

El fenómeno por el que las células se decantan por la formación de lactato fue denominado como “glicólisis aeróbica” por Warburg y se observa en una gran variedad de tumores, entre los que se incluye el Cáncer Colorrectal (CCR) [10, 15]. En la revisión realizada por Potter et al. en 2016 [11] sobre este fenómeno se

## 1 . Introducción

---

recogen diversos estudios que comprueban esta hipótesis, como el de DeBerardinis et al., en 2007 [16]; el de Fantin et al., en 2006 [17] o el de Schulz et al. en 2006 [18], quienes deducen que el principal factor metabólico en el crecimiento tumoral es la eficiencia en la conversión de la energía mitocondrial [11].

La hipótesis planteada por Warburg en 1927 defendía que este cambio en el metabolismo energético se llevaba a cabo porque las células tumorales presentaban una inactividad total en sus mitocondrias [10, 11]. Sin embargo, en los últimos 20 años, se ha concluido que esta hipótesis no es totalmente acertada. Las mitocondrias de estas células son funcionales y están activas en los tumores en desarrollo, aunque sí muestran diferentes alteraciones bioenergéticas en función del tipo de tumor. De hecho, es probable que muestren un comportamiento bimodal a lo largo del desarrollo del cáncer, exhibiendo un cambio dinámico entre el metabolismo oxidativo y la glicólisis. Aunque, de ser así, aún se desconocen tanto el método por el que se lleva a cabo como las causas exactas. Incluso se podría afirmar que esa flexibilidad metabólica podría estar vinculada, en cierto sentido, a las condiciones medioambientales o a la aparición de mutaciones genéticas asociadas al cáncer que, de algún modo, también tienen cierto impacto sobre el metabolismo [11, 19].

De hecho, en el 2011 Bhalla et al. [19] proponen que, además de la generación de ATP, existen muchos beneficios en el incremento del uso del metabolismo glucídico. La glucosa sirve como precursor de la biosíntesis de moléculas envueltas en el proceso de generación de biomasa, tales como los ácidos nucleicos o los lípidos. Por lo tanto, el incremento de estas dos moléculas juega un papel importante en el desarrollo de los tumores malignos. Y la habilidad de las células cancerígenas para coordinar el metabolismo glucídico es un aspecto muy importante en su fenotipo metabólico [19].

Concluyendo, la transformación celular es un proceso que requiere principalmente de alteraciones en el control del ciclo celular, pero que además depende de otros factores entre los que se puede destacar la reprogramación metabólica variable entre los diferentes tipos de tumor. Dicha variabilidad, junto con la capacidad de las células malignas de adaptarse a condiciones ambientales extremas, donde la cantidad de nutrientes y de oxígeno es limitada, pone de manifiesto la complejidad de esta enfermedad y la dificultad para el desarrollo de medidas de prevención o de terapias contra el cáncer [9, 11].

### 1.1.2. Las mitocondrias en el metabolismo energético

Las mitocondrias son los orgánulos encargados de la regulación del metabolismo energético celular, que se localizan en el citoplasma de las células eucariotas y se encargan de aportar la energía necesaria para la actividad celular [20]. El número de mitocondrias por célula varía según los tejidos en los que se encuentran [21, 22] y

está controlado por la homeostasis mitocondrial, encargada de conseguir el balance entre la biogénesis mitocondrial y la degradación de las mitocondrias [9].

La degradación mitocondrial o mitofagia es el proceso por el que se degradan y reciclan las mitocondrias dañadas, mientras que la biogénesis mitocondrial se trata del proceso contrario, por el que se forman las mitocondrias funcionales y sanas. De hecho, ambos procesos suelen estar coordinados por la proteína Parkin, consiguiendo así un proceso de renovación mitocondrial [23]. La biogénesis mitocondrial es un proceso coordinado que implica la replicación del ADNmt, su transcripción a ARNm (ARN mensajero) y su traducción a proteínas para su correcto funcionamiento, requiriendo de la síntesis de más de mil proteínas mitocondriales codificadas entre el genoma nuclear y el ADNmt (ADN mitocondrial). Las proteínas resultantes de la expresión de los genes nucleares se deben focalizar en cada una de las tres regiones de la membrana mitocondrial y en su matriz. Generalmente, este proceso se lleva a cabo por estímulos internos, como es el de Parkin para la renovación mitocondrial o en diferentes efectores del ciclo celular en los que se requieren grandes cantidades de energía. Sin embargo, cabe destacar que determinados estímulos externos también provocan su activación, entre los que se puede citar al ejercicio físico, la restricción calórica y la señalización hormonal [21, 24].

A pesar de que la función mitocondrial en la alteración del metabolismo energético no es del todo conocida, se han planteado hipótesis sobre que el cambio hacia la OXPHOS de las células cancerígenas puede ser debida a un desequilibrio en la coordinación entre las proteínas mitocondriales codificadas en el genoma nuclear y las codificadas en el genoma mitocondrial [9, 20]. Además, hace tiempo que se observan evidencias de que la disfunción mitocondrial está posiblemente relacionada con la oncogénesis ya que mutaciones en el ADNmt se relacionan con diferentes tipos de cáncer [25].

En general, se ha llegado a la conclusión de que las mutaciones del ADNmt pueden jugar un papel importante en las etapas tempranas del desarrollo del cáncer al aumentar la producción de especies reactivas de oxígeno (ROS), modificar las señales metabólicas o, incluso, al bloquear la apoptosis [26]. Por lo general, todas las mitocondrias de células cancerígenas son relativamente pequeñas, han perdido sus crestas y son deficientes en la subunidad  $\beta$  de la ATPasa F1F0. Además, su biogénesis se ve reducida por un descenso en la expresión del Peroxisome proliferator-activated receptor gamma coactivator 1-alpha (PGC1 $\alpha$ ) [9].

Este gen es el encargado de regular la función mitocondrial, lo que llevó a la especulación de que la reducción de su expresión en los tumores puede ser responsable del conocido “efecto Warburg” [19], nombre que recibe hoy la “glicólisis aeróbica” [19, 27]. Es más, hay estudios que pudieron observar una asociación

## 1 . Introducción

---

entre la reducción en la función mitocondrial y el incremento del crecimiento tumoral [28], aunque estudios más recientes muestran que la función mitocondrial es imprescindible para la transformación celular. Esto plantea la hipótesis de que las mitocondrias presentan un papel bimodal en el desarrollo tumoral controlado por el PGC1 $\alpha$ , quien puede tener un potencial papel carcinógeno [19, 28].

De hecho, aunque el papel del metabolismo energético en el desarrollo del cáncer no se consideró hasta el descubrimiento de Warburg, tras este punto de inflexión, cada vez más estudios sugieren que la mitocondria y el PGC1 $\alpha$  tienen un rol importante en diversos aspectos del desarrollo del tumor [29]. El trabajo llevado a cabo por Girnun, titulado “The diverse role of the PPAR $\gamma$  coactivator 1 family of transcriptional coactivators in cancer” [29], detalla los aspectos más importantes que vinculan al PGC1 $\alpha$  con el cáncer. Según se expone, el PGC1 $\alpha$  promueve la homeostasis energética al activar las rutas anabólicas, dando a las células la capacidad de obtener la biomasa necesaria para la división celular, favoreciendo un papel procancerígeno a este gen. Sin embargo, en el mismo trabajo se detalla como el gen es capaz de inducir la muerte celular al reducir la expresión de las enzimas antioxidantes, lo que promueve la acumulación de las ROS y, por lo tanto, estrés oxidativo y muerte celular, mostrando funciones anticancerígenas [29, 30]. Según se cree, el gen adquirirá este tipo de roles en función de las interacciones con factores ambientales y el nivel de expresión del gen [29]. Entender los mecanismos y patrones por los que el PGC1 $\alpha$  influye sobre el desarrollo del tumor, podrían dar lugar a un gran avance tanto en el conocimiento del cáncer como en su prevención y posibles terapias [31, 32].

En este aspecto, cabe destacar que las investigaciones sobre las alteraciones en la expresión de este gen lo han asociado principalmente con tumores de mama y hepatocarcinomas [33]. Sin embargo, a pesar de los estudios realizados, los mecanismos moleculares por los que actúa PGC1 $\alpha$  y que llevan al desarrollo de la enfermedad, aún no son conocidos, aunque se barajan diferentes hipótesis. Por otro lado, en otros tumores como el de colon y recto, la investigación aún es muy básica.

### 1.2. El cáncer colorrectal

La pared del colon se caracteriza por poseer, desde el lumen hacia el exterior, cuatro capas bien diferenciadas, denominadas capa mucosa, submucosa, muscular y serosa o adventicia [34]. Desde la capa mucosa hacia la capa muscular se extienden unas criptas cuyo epitelio, al igual que el de la superficie, está formado por células absortivas y caliciformes. Éstas últimas son más frecuentes en la parte baja de la cripta, mientras que las absortivas se encuentran repartidas por todo el epitelio y se caracterizan por su elevado contenido en mitocondrias [35], el cual es esencial



por el proceso de continua renovación que sufre este tejido. Este proceso es posible por la presencia de células madre pluripotentes en la región más baja de las criptas. Bajo condiciones fisiológicas, estas células madre entran en mitosis para dar lugar a células hijas capaces de diferenciarse, reducir su capacidad proliferativa y modificar su metabolismo de forma simultánea a su migración por el eje cripta-vellosidad. Una vez llegan a la zona alta de la vellosidad, en contacto con la luz intestinal, sufren la apoptosis, se descaman y son expulsadas al exterior junto con las heces [30, 36, 37].

Alteraciones en este proceso de renovación del epitelio intestinal constituyen una de las bases del desarrollo de enfermedades en el mismo, entre las que cabe destacar el CCR [30]. Esta enfermedad es un tipo de tumor maligno desarrollado en cualquiera de las regiones del intestino grueso y recto [38] y que supone el tercer tipo de cáncer más común y la cuarta causa de muerte por cáncer a nivel mundial, mostrando una incidencia de 1.36 millones de casos nuevos al año (según datos del Globocan 2012) y constituyendo uno de los problemas de salud pública más importantes a nivel global [39-41]. En España, es el tumor maligno de mayor incidencia considerando ambos sexos, con cifras que rondan los 35000 casos nuevos al año. Estos datos se pueden traducir en que 1 de cada 20 hombres y 1 de cada 30 mujeres estarán afectados por esta enfermedad antes de cumplir los 74 años en nuestro país [40].

Por otro lado, en España este tumor presenta una supervivencia a los 5 años que se encuentra por encima de la media europea, situándose en un 64 % frente a un 57 % respectivamente y siendo mayor en aquellos pacientes diagnosticados en un estadio temprano del desarrollo de la enfermedad [40]. Sin embargo, sólo un 39 % de los pacientes son diagnosticados a tiempo, cuando el tumor se encuentra en un estadio precoz [42], lo que demuestra la clara necesidad de sostener un cambio en el estilo de vida de la sociedad para prevenir la enfermedad y de implantar medidas de diagnóstico precoz para frenar el desarrollo de la misma.

Las vías de desarrollo de este tumor son principalmente de dos tipos, esporádicas y hereditarias. La mayor parte de los tumores de colon y/o recto diagnosticados pertenecen a los del primer grupo y suelen ser debidos a alteraciones genéticas somáticas o a cambios epigenéticos que sufren los colonocitos. Ambos procesos serán desarrollados posteriormente, en el apartado “Papel de la genética en el cáncer colorrectal”.

Por otro lado, el cáncer colorrectal parece ser uno de los más influenciados por factores epigenéticos [43], existiendo múltiples factores ambientales que pueden influir en su desarrollo.

### 1.2.1. Factores de riesgo del CCR

Las diferencias internacionales y los datos de migración dejan constancia de que el CCR es muy sensible a los cambios en el estilo de vida. De hecho, tanto los inmigrantes como su descendencia adquieren las ratios de incidencia del país donde se hospedan [43, 44]. Por este motivo, es especialmente interesante conocer los factores de riesgo que pueden influir en el desarrollo de este tumor. Entre los factores ambientales más comunes nos podemos encontrar los siguientes [45]:

- **Edad:** Es uno de los factores de riesgo más conocidos. Generalmente, la mayoría de los sujetos que padecen este tumor son mayores de 50 años, siendo esta la edad a partir de la cual el riesgo de incidencia crece de manera exponencial [45].
- **Sexo:** Los hombres presentan una mayor incidencia que las mujeres, exhibiendo un ratio de incidencia de 37.6 en hombres frente a un 24.2 en mujeres en los países desarrollados y un 20.4 frente al 14.6, respectivamente, en el mundo en general [40, 45, 46].
- **Antecedentes familiares:** Hace décadas que se conoce el hecho de que aquellos individuos que tienen familiares de primer grado que han padecido CCR presentan un riesgo aproximadamente duplicado de desarrollar este tumor. Este riesgo se incrementa si el paciente tiene más de un pariente en primer grado afectado por la enfermedad o si alguno de ellos fue diagnosticado antes de los 55 años. Por el contrario, si el parentesco es de segundo grado, el riesgo solo se incrementa un 25-50 % con respecto a la media [43, 47].
- **Tabaco:** Durante mucho tiempo no se ha podido observar una asociación entre el tabaco y el riesgo de desarrollar CCR [45] y, aún a día de hoy, sigue siendo un factor discutible. Sin embargo, la hipótesis más plausible parece dirigirse hacia que el riesgo en el consumo de tabaco depende de los factores genéticos del fumador [48, 49].
- **Alcohol:** El consumo de altas cantidades de este elemento se correlacionan con un incremento moderado del riesgo de padecer cáncer colorrectal (OR=1.34) [45]. Incluso se ha observado en un meta-análisis de ocho cohortes una relación dosis-dependiente entre el consumo de alcohol y el riesgo de desarrollar cáncer de colon o de recto, independientemente del tipo de alcohol consumido [49].

- **Índice de masa corporal (IMC):** Un alto IMC (Obesidad o superior;  $IMC \geq 30kg/m^2$ ) incrementa el riesgo de desarrollar cáncer de colon (OR=1.29) [45]. Sin embargo, esta relación no siempre es fácil de observar, puesto que este riesgo se puede corregir con una alta actividad física. Siendo estos resultados más consistentes en hombres que en mujeres [43].
- **Actividad física:** La práctica de actividad física se relaciona con un efecto protector sobre la probabilidad de desarrollar el tumor, siendo uno de los factores más consistentes en los resultados de la literatura epidemiológica, aunque hay cierta discordia sobre si influye por igual en colon y recto [50, 51]. En este sentido, el sedentarismo supone un factor de riesgo para este tumor, que presenta una OR=1.34 [45].
- **Dieta:** El papel que presenta el consumo de frutas y verduras en el desarrollo del CCR aún no está definido. Sin embargo, hay evidencia de que el uso de alimentos ricos en fibra, como fruta, verdura o legumbres reducen el riesgo de desarrollo de CCR [52]. Por otro lado, un elevado consumo de carnes rojas o procesadas y de grasas promueve la carcinogénesis (OR=1.29) [45, 53]. De este modo, los hábitos de la dieta occidental, entre los que se incluye el incremento de la ingesta de carnes rojas y grasas y el descenso de la ingesta de frutas y verduras, están asociados con el CCR [53]. Los mecanismos por los que la dieta influye en el desarrollo de la enfermedad aún no son conocidos [50, 54].
- **Consumo de fármacos:** El consumo de AINES (Antinflamatorios no esteroideos), como es el ácido acetil salicílico, muestran cierta controversia en su asociación con el CCR. A pesar de que tienen un efecto claramente preventivo, los datos del perfil riesgo/beneficio aún no son concluyentes [55].
- **Enfermedades:** Enfermedades inflamatorias del colon, como la colitis ulcerosa (UC) [56-58] o la enfermedad de Crohn [59], son fuertes factores de riesgo [60, 61]. Además, otras enfermedades como la Diabetes Mellitus, la hiperinsulinemia o la obesidad también suponen un factor de riesgo.

### 1.3. Papel de la genética en el cáncer colorrectal

La genética es uno de los principales factores de riesgo en el CCR. Las alteraciones genéticas que sufre el ADN pueden ser muy diversas, suponiendo desde cambios en un único nucleótido a otros que afectan a una gran región del genoma. Estas variaciones tienen consecuencias muy diferentes, entre las que se pueden

## 1 . Introducción

---

destacar la generación de proteínas anormales o la alteración en los patrones de expresión génica [62]. En el caso de las células cancerígenas, la acumulación de estas alteraciones les proporciona ventajas proliferativas y de autorrenovación sobre sus progenitoras [37].

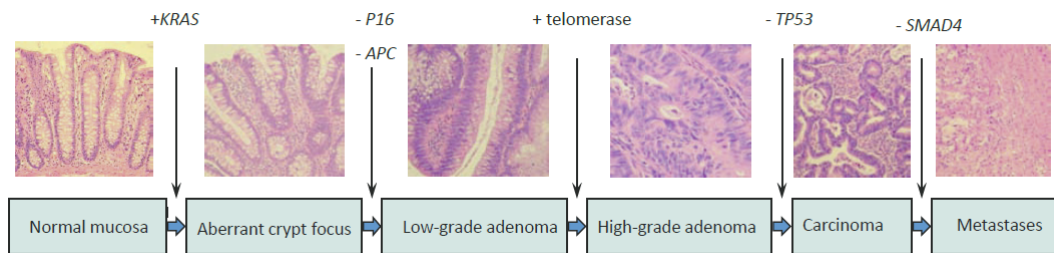
Hasta el momento, se conocen tres rutas genéticas asociadas al desarrollo del cáncer colorrectal bien descritas: la inestabilidad cromosómica, la inestabilidad de los microsátélites y los cambios en la metilación génica [63].

### ■ Ruta de la inestabilidad cromosómica:

Estas alteraciones están presentes en el 85% de los tumores de este tipo, por ese motivo están comunmente asociadas con el mecanismo genético de desarrollo del cáncer colorrectal. En este caso, el desarrollo del tumor se debe a la aparición de mutaciones somáticas y/o a cambios epigenéticos en los colonocitos que hacen que éstos no respondan debidamente a los mecanismos que regulan el proceso de renovación epitelial [1, 64]. Estas alteraciones genéticas actualmente están bien caracterizadas, lo que permite conocer la secuencia de los genes alterados en las diferentes etapas del desarrollo del adenoma-carcinoma (Ver Figura 2):

- Las alteraciones en el gen KRAS suceden en la primera etapa del desarrollo del CCR, en las células de la mucosa normal, provocando la aparición de criptas aberrantes [43, 63]. KRAS es un oncogen cuyas mutaciones, principalmente las localizadas en los codones 12 y 13, provocan una señal de división celular permanente [65].
- La inactivación del gen APC sucede también en las células de la mucosa normal, originando la misma consecuencia que las alteraciones en KRAS, la aparición de las criptas aberrantes. Posteriormente, en estas criptas, las alteraciones en APC generan señales de displasia, que dan lugar a un tumor de bajo grado, caracterizado por su pequeño tamaño y morfología tubular. El APC es un gen supresor de tumores del que se conocen mutaciones asociadas al cáncer desde hace más de 20 años [66]. Además, su inactivación supone la acumulación de la  $\beta$ -catenina, que llega al núcleo celular y activa la transcripción de sus genes diana, entre los que se encuentra el gen MYC [67], un proto-oncogen cuya sobreexpresión produce alteraciones entre las fases G1 a S del ciclo celular [68, 69].
- A estas etapas les sucede la activación de la telomerasa, una proteína encargada de la elongación de los telómeros, lo que confiere a las células la capacidad de división permanente [43, 63].

- Posteriormente la inactivación del gen supresor de tumores TP53, junto con la activación de la telomerasa, produce el cambio de carcinoma de bajo grado a carcinoma de alto grado, que se caracteriza por la pérdida de polaridad de sus células, las cuales además poseen un núcleo pleomórfico y presentan una gran capacidad mitótica. TP53 codifica para la proteína p53, encargada de funciones vitales para el mantenimiento celular, como son mantener la integridad celular o controlar múltiples rutas metabólicas celulares, entre las que se puede destacar la inhibición del proceso de proliferación celular, la sensibilización de las células ante la apoptosis o la autofagia, el control de los procesos inflamatorios, la represión de la metástasis y otras funciones capaces de controlar el desarrollo tumoral [43, 63, 70].
- Por último, la inactivación del gen SMAD4 puede dar a estos tumores la capacidad de iniciar una metástasis como última fase del desarrollo del tumor [43, 63].



**Figura 2: Fases del desarrollo del cáncer colorrectal por la inestabilidad cromosómica [63].** En la figura se pueden observar las fases del desarrollo del cáncer colorrectal y los diferentes genes implicados en cada una de ellas.

### ■ Ruta de la inestabilidad de los microsatélites:

Estas alteraciones están involucradas en el 15% de los tumores esporádicos de colon y recto y son debidas principalmente a alteraciones en la maquinaria de reparación del ADN. Entre los genes mutados en este caso están el MLH1 y el MSH2. MLH1 requiere de la formación de un complejo proteico con MLH3, mientras que MSH2 heterodimeriza con MSH6. La función de ambos complejos es la reparación de inserciones y deleciones en el ADN. Sin embargo, alteraciones en su secuencia impiden la formación de los complejos proteicos mencionados [71].

### ■ Ruta de metilación de las islas CpG:

Las consecuencias de las mutaciones genéticas de esta ruta se solapan en cierta medida con la anterior, pues se caracteriza por la metilación de los promotores de genes como el MLH1, resultando también en una inestabilidad de los microsatélites. Además, es común que sucedan mutaciones de este estilo en el gen BRAF, las cuales inician el desarrollo del tumor. Tras la metilación de estos genes, suelen proceder otras mutaciones en los oncogenes y genes supresores de tumores mencionados con anterioridad [72].

Cabe destacar que la genética no sólo influye en la aparición de tumores esporádicos. El CCR puede ser desarrollado también a partir de mutaciones germinales, las cuales dan lugar a los síndromes de cáncer colorrectal familiar, que suponen aproximadamente entre el 5 y el 10 % de los casos de esta enfermedad [52]. Entre estos síndromes se puede citar al FAP (Poliposis adenomatosa familiar) y el de Gardner, ambos caracterizados por mutaciones en el gen APC principalmente; o el síndrome de Lynch, que es el CCR hereditario más común y normalmente debido a mutaciones en los genes que codifican para la maquinaria de reparación del ADN, como es MLH1 o MSH2 [63].

Por último, los factores genéticos también influyen en el desarrollo de CCR a partir de determinadas enfermedades inflamatorias, como la enfermedad de Crohn o la colitis ulcerosa. En este caso, las alteraciones genómicas son similares a las del cáncer esporádico por inestabilidad cromosómica, pero en el que las alteraciones suceden al contrario, apareciendo en TP53 en las etapas más tempranas y en APC en las más tardías [63].

Los genes implicados en cada una de estas categorías han sido ampliamente estudiados, y sus funciones en el CCR son relativamente conocidas [43]. Sin embargo, el CCR representa un grupo de desórdenes muy heterogéneo tanto a nivel biológico como molecular [37], existiendo un gran número de tumores que no muestran ninguna de las variantes patogénicas conocidas [73]. Esto sugiere la posibilidad de que alteraciones en cada uno de los locus puedan presentar cierto efecto sobre el riesgo de desarrollar CCR, ampliando considerablemente el elenco de mutaciones que pueden causar cáncer. Estos polimorfismos que afectan a un único nucleótido del ADN son conocidos como SNPs (de sus siglas en inglés “*Single Nucleotide Polymorphism*”), y pueden actuar solos o en combinación con otros loci o factores ambientales [74].

## 1.4. Función de los SNPs en el cáncer colorrectal

Los SNPs suponen el tipo más común de alteraciones genéticas en el genoma humano, calculándose que pueden suceder, de media, en una de cada 300 pares de bases [75]. Este tipo de mutaciones juegan un papel importante en el desarrollo de múltiples enfermedades, entre las que se halla el CCR.

Los SNPs se localizan en regiones tanto codificantes como no codificantes del genoma, con diferentes consecuencias en cada caso. Cuando el SNP se localiza en una región codificante para proteína pueden darse dos situaciones. Por un lado, puede suceder que la funcionalidad de la misma se vea reducida, inhibida o, incluso, adquiera una funcionalidad diferente. Estos SNPs reciben el nombre de *SNPs no sinónimos* debido a que el aminoácido resultante de la alteración es diferente al original. Existen múltiples ejemplos de SNPs no sinónimos asociados a CCR, entre los que se pueden citar, como ejemplos, el rs4939287 del gen SMAD7 [76, 77] o el rs4779584 del gen GREM1-SCG5 [78]. Por otro lado, en caso de que la alteración resulte en el mismo aminoácido, el SNP recibe el nombre de *sinónimo*, de los cuales se considera que no tienen consecuencias visibles sobre la función de la proteína. Sin embargo, ya se ha observado al menos un caso en el que una mutación sinónima localizada sobre el gen MDR1 se asocia con un cambio en su especificidad por el sustrato [79]. Por último, cuando la alteración se localiza en una región no codificante, las consecuencias pueden ser múltiples, variando procesos de expresión génica como el empalme alternativo (también conocido como *splicing alternativo*), la unión de factores de transcripción o, incluso, puede llegar a producir la degradación del ARNm resultante de la transcripción [75].

La mayoría de estos polimorfismos que afectan a un único nucleótido se asocian con un riesgo entre leve y moderado con la enfermedad, por lo que reciben el nombre de polimorfismos de baja penetrancia. En los últimos años, este tipo de SNPs ha ganado cierta importancia en la investigación contra el cáncer, siendo la mayoría de ellos más importantes en el inicio que en la progresión de la enfermedad [80]. La identificación de alteraciones alélicas que modifican el riesgo de desarrollar el CCR puede permitir asentar las bases genéticas de dicha enfermedad y conocer el riesgo individual, pudiendo desarrollar estrategias de prevención y terapia más útiles.

Entre las herramientas disponibles para el análisis de los SNPs nos encontramos los GWAS (de su siglas en inglés *Genome-Wide Association Study*), herramientas que permiten el descubrimiento de múltiples SNPs de bajo riesgo. Estos SNPs, de forma aislada, sólo pueden explicar una parte de la contribución genética aún desconocida [81]. Sin embargo, parece plausible pensar que varias de estas alteraciones de baja penetrancia puedan actuar juntas en el desarrollo de la enfermedad, explicando en mayor medida la contribución genética a la enfermedad

y presentando, por lo tanto, una mayor interpretabilidad biológica [81, 82]. Con este objetivo se desarrolló el análisis de conjuntos de genes a partir de los datos de GWAS [83].

### 1.5. Métodos de análisis de los SNPs

La relación existente entre las alteraciones génicas y las enfermedades que éstas pueden generar no son fáciles de analizar ya que, en dichas asociaciones, suelen estar involucrados diversos factores, como la interacción entre diferentes genes o entre genes y factores ambientales [84]. Sin embargo, tras la aparición del primer GWAS en el 2005 [85], el análisis genético de las enfermedades complejas ha evolucionado mucho [86].

Los GWAS son herramientas diseñadas para entender estas asociaciones [82]. Explicados brevemente, los GWAS son bases de datos que contienen información de millones de SNPs repartidos de forma aleatoria y uniforme por todo el genoma humano y recogidos en una población concreta. Esta población está dividida en dos subpoblaciones, los casos (individuos que presentan la enfermedad analizada) y los controles (individuos libres de dicha enfermedad) [82, 84]. Su objetivo principal es evaluar la asociación existente entre la frecuencia del genotipo de un determinado SNP y una enfermedad [87].

Cabe señalar en la presente tesis que, en algunas ocasiones, estos GWAS no se realizan sobre el genoma completo, sino que se pueden realizar sobre un tipo de secuencias concretas, como puede ser el exoma, siendo ésta la información realmente disponible en el proyecto MCC-Spain empleada a lo largo del trabajo que nos ocupa.

Existen diversas técnicas para llevar a cabo el análisis de este tipo de información, entre los que se pueden destacar dos: El análisis SNP a SNP (*“Single SNP analysis”*) y el análisis de conjuntos de genes [82, 87], como ya se ha comentado previamente.

#### 1.5.1. Análisis SNP a SNP

Este análisis consiste en la valoración de la asociación de los SNPs con la enfermedad de manera individualizada [85] y ha sido uno de los análisis más empleados, siendo el primero en llevarse a cabo [85, 87, 88]. Sin embargo, la información que aporta, aunque relevante, sólo explica una pequeña proporción de todas las variaciones fenotípicas existentes [85].



Generalmente, este tipo de análisis se lleva a cabo mediante regresión logística, permitiendo así el análisis de interacción con factores ambientales. Las asociaciones significativas indican que la alteración génica analizada, o alguna próxima a la misma, puede presentar un papel importante en la susceptibilidad a la enfermedad [89]. Sin embargo, por lo general y salvo muy pocas excepciones (como la del primer GWAS publicado en 2005 por Klein et al. [90]), la mayoría de alelos de riesgo asociados con la enfermedad muestran Odds Ratios (OR) muy bajas ( $OR < 1.5$ ) [91]. A pesar de esto, en 2009, cuatro años después de la obtención del primer GWAS, ya se habían detectado más de 2000 polimorfismos asociados con diferentes enfermedades complejas [90].

Una de las principales limitaciones de este análisis es que, por definición, las enfermedades complejas están causadas por numerosas variantes génicas. Sin embargo, como ya se ha indicado, estos análisis sólo consideran un único SNP de cada vez y las interacciones entre ellos se omiten, reduciendo el tamaño del efecto [85]. Otro factor importante es que la proporción de variantes que explican un determinado fenotipo es realmente baja, reduciendo también la probabilidad de hallarlas. Por último, se ha observado que GWAS de enfermedades con una prevalencia elevada muestran variantes génicas menos significativas que aquellos basados en enfermedades de baja prevalencia [85].

Por estos motivos, los análisis llevados a cabo por estas herramientas han evolucionado hacia la valoración de conjuntos de SNPs [86]. Sin embargo, ambos tipos de análisis son complementarios, y el conocimiento aportado entre los dos enfoques puede ayudar a conocer mejor los factores genéticos e, incluso, los procesos biológicos que aumentan la susceptibilidad al desarrollo de enfermedades como el CCR [92].

### 1.5.2. Análisis de conjuntos de genes

Estos análisis, conocidos por las siglas GSA, de su nombre en inglés “*Gene set analysis*”, consisten en el estudio de un conjunto de genes relacionados en base a un criterio, ya que los genes no trabajan de forma aislada, sino que mantienen complejas redes de interacción [93]. Los criterios empleados pueden ser de diversa índole como, por ejemplo, participar en una misma ruta metabólica, que estén funcionalmente relacionados o que precisen de una regulación coordinada en su expresión [84, 86, 93].

Los primeros análisis llevados a cabo desde este enfoque fueron implementados con datos de expresión génica [94, 95]. Posteriormente, tras el trabajo de autores como Wang et al. [83], Chasman [96] o Holden et al. [97], estas herramientas de análisis se adaptaron a datos de SNPs [92]. En este sentido, han sido muchas las

## 1 . Introducción

---

herramientas desarrolladas para su realización, todas ellas con diferentes ajustes o hipótesis a probar [93, 98].

En este último aspecto, se pueden diferenciar dos tipos de técnicas en función de su hipótesis nula, las autocontenidas (*self-contained* en inglés) y las competitivas (*competitive*) [93, 98-100]. Las primeras comparan el test estadístico de los genes del conjunto (ruta metabólica, por ejemplo) con el test estadístico de los genes no asociados, mientras que la segunda compara las pruebas estadísticas de los genes del conjunto con la de los genes en el genoma [93]. Escoger entre un tipo u otro depende de cada estudio en particular, aunque hay autores que recomiendan el uso de las hipótesis competitivas por considerarlas más robustas [98].

Una de las principales ventajas de estas técnicas es que presentan un enfoque basado en el conocimiento biológico y sus resultados son más fáciles de interpretar desde este punto de vista.

### 1.5.3. Bases de las técnicas de análisis de conjuntos de SNPs

Las técnicas de GSA generalmente se basan en el método GSEA (de sus siglas en inglés *Gene set Enrichment analysis*). Éste fue uno de los primeros métodos generados, basado en datos de expresión génica [94]. Fundamentalmente, consiste en:

1. Calcular una puntuación de asociación SNP-enfermedad capaz de valorar las diferencias entre los casos y los controles de cada uno de los SNPs del Exome Array. Una de las pruebas más empleadas es el test de tendencia de Cochran Armitage.
2. Disponer todos los SNPs en una lista, ordenados por nivel de significación de su asociación, del más significativo al menos.
3. Calcular el ES (Enrichment Score) mediante una suma acumulada. Para ello, se toma la lista anterior y, de uno en uno, se considera si el SNP está en el conjunto de genes de interés o no. En caso afirmativo, el ES incrementará su valor, en caso contrario, lo reducirá. El valor de incremento cuando el SNP se encuentra en el conjunto de interés es equivalente al test estadístico empleado, mientras que la resta siempre es constante. El valor que adquiere el ES finalmente es el correspondiente a la máxima desviación desde cero de la suma acumulada.

4. El último paso de la técnica consiste en la normalización del ES (NES), de acuerdo al tamaño del conjunto de SNPs. La proporción de falsos positivos se controla mediante la realización de numerosas permutaciones de las etiquetas caso/control, con lo que se calcula el test de múltiples hipótesis (FDR = ‘False Discovery Rate’).

Si los SNPs del conjunto de datos se reparten al azar por toda la lista, se considera que el conjunto de genes probablemente no está asociado a la enfermedad, y su valor de ES será bajo. Sin embargo, si se acumulan en la parte alta o baja de la lista, la probabilidad de asociación se verá incrementada, alcanzando un valor de ES alto e, incluso, la significación estadística (Para más información [94]).

### 1.5.4. Técnicas de análisis de conjuntos de SNPs

Son muchas las herramientas disponibles para aplicar el GSA a datos de SNPs. En la revisión de Wang et al [93] se recogen las más empleadas, diferenciando aquellas que plantean una hipótesis nula competitiva o una autocontenida. A pesar de que todas ellas están basadas en el GSEA, comentado en la sección anterior, muestran diferencias básicas en su desarrollo, que van desde los algoritmos de selección de SNPs a las pruebas estadísticas realizadas para conocer la asociación SNP-enfermedad.

A continuación se exponen las técnicas que comprueban una hipótesis competitiva, por ser consideradas más robustas:

#### (i) GSEA-SNP:

Holden et al. [97] proponen esta técnica de análisis de datos de SNPs procedentes de dos poblaciones, los casos y los controles. Los conjuntos de SNPs son escogidos en base a los conjuntos de genes, seleccionando todos los SNPs disponibles de cada gen. El estadístico empleado es el MAX-test, que se define como el máximo test estadístico de los tres test del Cochran Armitage asociados a los diferentes mecanismos de herencia posibles que presenta cada SNP, entre los que se consideran el aditivo, el dominante y el recesivo. Otra característica a destacar es que los datos de entrada que requiere esta técnica es el genotipado completo y realiza corrección por múltiples hipótesis, empleando tanto FDR como por FWER (de sus siglas en inglés “*Family-wise error rate*”). Información más detallada sobre esta técnica se puede encontrar en [97] y en <https://www.nr.no/en/projects/software-genomics>.

## 1 . Introducción

---

### (II) **SNP ratio test (SRT):**

En la técnica propuesta por O’Dushlaine et al. [101] los SNPs seleccionados se basan en aquellos presentes en el conjunto de genes de interés, seleccionando sólo aquellos que tengan un valor estadístico superior a un umbral dado. El estadístico empleado en este caso es un test de tendencia. El tipo de datos de entrada es el genotipo completo y se puede ejecutar en la herramienta de análisis del PLINK, siendo así muy fácil de emplear. Un factor a tener en cuenta es que no hace ningún tipo de corrección por múltiples hipótesis. Más información disponible en [101] y en <https://sourceforge.net/projects/snpratiotest/>.

### (III) **i-GSEA4GWAS:**

Esta herramienta de análisis fue propuesta por Zhang et al. [87]. El tipo de datos de entrada admitidos son valores de p, no permitiendo el uso de los datos de genotipado “brutos”. La ventaja de este hecho es que permite seleccionar la prueba estadística más apropiada según el criterio del investigador que la aplique. Esta técnica calcula el logaritmo negativo del valor de p ( $-\log(\text{valordep})$ ), seleccionando el SNP más significativo de cada gen. En este caso se corrige por FDR. Detalles de esta técnica se pueden encontrar en [87] y en <http://gsea4gwas-v2.psych.ac.cn/>.

### (IV) **GenGen:**

Propuesta por Wang et al. [83], esta técnica calcula el  $\chi^2$  de tendencia y selecciona un SNP por gen, el que presente el mayor valor de dicho estadístico, sin considerar el hecho de que alcance un umbral de significación. Emplea el genotipado completo como tipo de datos de entrada y corrige por FDR y por FWER. Información más detallada se puede encontrar en <http://gengen.openbioinformatics.org/en/latest/> y en [83].

### (v) **GSA-SNP:**

El software desarrollado por Nam et al. [102] permite seleccionar un número de SNPs concreto por cada gen, el cual varía entre 1 y 5 SNPs, escogiendo siempre los polimorfismos más significativos por cada gen. Los autores señalan que la mejor opción es escoger los dos primeros. Este software además permite el desarrollo de diferentes test estadísticos con los que comprobar la asociación del conjunto de genes con la enfermedad. Entre sus opciones, se encuentra la del cálculo del ES mediante un test de tendencia del Cochran Armitage. Además, esta técnica corrige por FDR. Más información en <http://gsa.mudals.org> y en [102].

Todas estas técnicas emplean el GWAS completo como referencia para el cálculo de la significación ruta metabólica-enfermedad (característica esencial para ser consideradas técnicas competitivas). Sin embargo, la principal diferencia es que cada técnica selecciona un número diferente de SNPs. De este modo, en los extremos nos encontramos con el SRT, que sólo tiene en cuenta los genes que presenten al menos un SNP significativo, y el GSEA-SNP, que considera todos los SNPs de todos los genes, pasando por GenGen o iGSEA4GWAS que seleccionan un SNP por gen, el que presente un valor más significativo (aunque no lo sea) o por GSA-SNP que selecciona un número determinado de SNPs más significativos por gen. Las diferencias principales entre estas técnicas se pueden ver en la Tabla 1.

**Tabla 1: Características principales de las técnicas descritas.**

Método de análisis	Datos de entrada	Selección de SNPs	Prueba estadística	Corrección de múltiples hipótesis	Ref.
<b>SRT</b>	Genotipado.	Selecciona una lista de SNPs significativos a partir de un umbral.	Test de tendencia	No hace correcciones.	[101]
<b>GSEA-SNP</b>	Genotipado.	Selecciona todos los SNPs disponibles en la base de datos pertenecientes a los genes de la ruta de interés.	Max-test (Más información en [97])	FDR y FWER	[94, 97]
<b>i-GSEA4GWAS</b>	Valores de p.	Selecciona un SNP por gen, el que presenta el valor de significación más bajo.	Libertad de selección. (Test Cochran-Armitage).	FDR	[87, 103]
<b>GenGen</b>	Genotipado.	Selecciona un único SNP por gen. Este SNP es el que presenta el valor estadístico más alto de los asignados al gen.	Estadístico $x^2$ para casos-contróles.	FDR y FWER	[83]
<b>GSA-SNP</b>	Valores de p.	Selecciona los “k” mejores SNPs de cada gen. Los autores recomiendan usar los dos mejores.	Libertad de selección. (Test Cochran-Armitage).	FDR	[102]

Esta tabla muestra similitudes y diferencias entre las técnicas. Ref.= Referencias principales de cada técnica. FDR = False discovery rate; FWER = Family-wise error rate

### 1.6. PGC1 $\alpha$

Como ya se ha comentado anteriormente, las alteraciones en el proceso de renovación del epitelio intestinal constituyen una de las bases más importantes en el desarrollo del CCR [30, 37]. Este proceso es controlado principalmente por gradientes de proteínas tan importantes como Wnt (Wingless), BMP (Bone morphogenetic proteins) y TGF- $\beta$  (Transforming growth factor beta 1) [37]. Sin embargo, existen otros genes que también presentan funciones importantes en este proceso, aunque han sido menos estudiados. Uno de estos genes es el PGC1 $\alpha$ , cuya función ha sido muy estudiada en tejidos de alta demanda energética, pero no tanto en el intestino [30].

#### 1.6.1. Papel del PGC1 $\alpha$ en la renovación del epitelio intestinal

En condiciones fisiológicas normales y para preservar la integridad y la homeostasis celular, el intestino posee múltiples mecanismos de defensa, entre los que se destacan las enzimas antioxidantes CAT (catalasa) y SOD2 (Superóxido dismutasa). Estas enzimas, sin embargo, no se expresan por igual a lo largo del eje cripta-vellosidad del intestino, presentando una expresión mayor en las células madre de la parte inferior de las criptas, lo que ejerce una clara defensa frente a las ROS. Por el contrario, las células del epitelio intestinal que están en contacto con el lumen poseen una elevada concentración de ROS, debida probablemente a un desequilibrio entre el incremento de la respiración y la baja actividad de las enzimas antioxidantes [104].

Según proponen D'Errico et al. [30] en su trabajo, la expresión del gen PGC1 $\alpha$  es inversa a la de las enzimas antioxidantes, por lo que su expresión es más elevada cuanto más alejadas están las células del fondo de la cripta. Como ya se ha mencionado con anterioridad, el incremento de la expresión de este gen supone una activación de la biogénesis mitocondrial, lo que promueve procesos como la respiración mitocondrial o la actividad de la cadena de transporte de electrones. Como consecuencia, se produce un aumento de la concentración de ROS, capaz de activar la apoptosis celular, permitiendo así llevar a cabo la renovación del epitelio intestinal [30]. De este modo, alteraciones genéticas somáticas en este gen, o en las rutas metabólicas que regula, pueden actuar como factores de riesgo en el desarrollo de la enfermedad.

### 1.6.2. Características de PGC1 $\alpha$ , gen y proteína

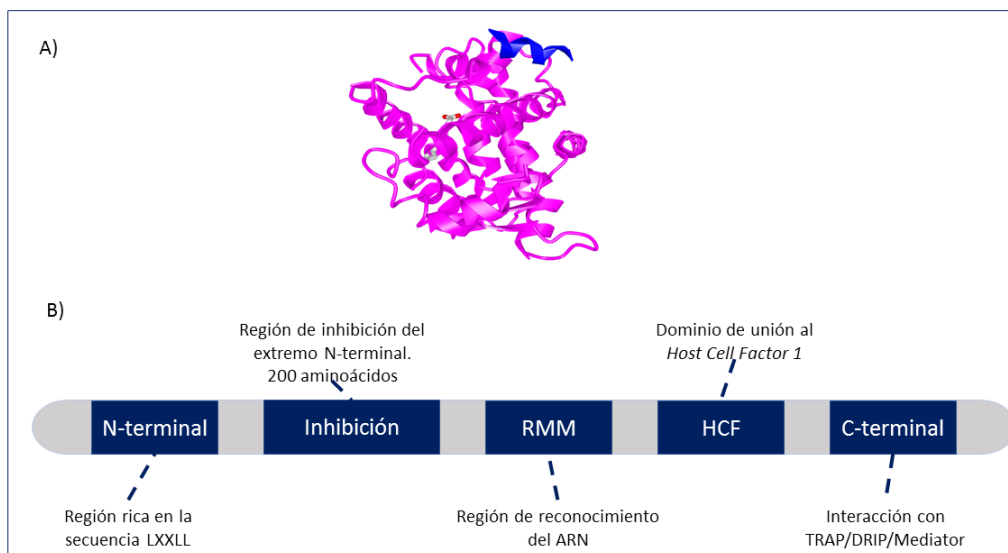
El Peroxisome proliferator-activated receptor gamma coactivator 1-alpha, comúnmente denominado PGC1 $\alpha$  (término empleado en la tesis), aunque también referido como LEM6, PGC1, PPARGC1a, PGC1v, PPARGC1, PGC-1alpha o PHC-1(alpha), es un gen del núcleo celular codificante para proteínas localizado en el brazo corto del cromosoma 4, concretamente en el 4p15.2. La localización exacta del gen va desde la posición 23792021 hasta la 24472975, por lo que su longitud es de 67 Kb aproximadamente. Está formado por 13 exones y codifica para una proteína de 798 aminoácidos [25, 105].

Este gen pertenece a la familia PGC1 $\gamma$  coactivator-1 (PGC-1). Esta familia fue identificada, en su origen, por su interacción funcional con el PGC1 $\gamma$  en el tejido adiposo marrón, un tejido rico en mitocondrias y cuya principal función es la termogénesis [106]. La familia PGC-1 está compuesta por tres miembros que se caracterizan por poseer un dominio de activación N-terminal, rico en la secuencia LXXLL encargada de interactuar con los factores nucleares. Adjunto a éste se encuentra una región de unos 200 aminoácidos, encargada de la inhibición de su actividad, seguido de una zona de reconocimiento de ARN denominado RMM que juega un papel importante en el empalme alternativo del ARNm, y continuado por un dominio de unión al HCF (*Host Cell Factor-1*). Por último, se encuentra el extremo Carboxilo terminal, encargado de la interacción con el complejo TRAP/DRIP/Mediator [29, 107] (Ver Figura 3).

El modo de actuación de esta proteína consiste en que el extremo amino-terminal interactúa con proteínas con actividad histon-acetil-transferasa, entre las que se incluyen las proteínas de unión a CREB(*cAMP response element binding*)/p300 y SRC-1 (*SRC proto-oncogene, non-receptor tyrosine kinase*). De este modo, la familia PGC-1 es capaz de remodelar las histonas de la cromatina consiguiendo un mayor acceso a los genes diana de la maquinaria transcripcional. Posteriormente, el complejo TRAP/DRIP/Mediator interactúa con el extremo carboxilo terminal, donde existen diferentes regiones que se encargan de llevar a cabo el empalme del pre-RNA con la transcripción [106].

## 1 . Introducción

---



**Figura 3: Estructura de la proteína PGC1 $\alpha$ .** La figura A muestra la estructura 3D de la proteína [108]. La figura B muestra un esquema de los dominios de la misma (Imagen de elaboración propia)

Esta familia de factores de transcripción es regulada a nivel de ARNm y, por lo tanto, de expresión de proteínas. Dicho proceso se lleva a cabo en respuesta a una gran variedad de rutas metabólicas de señalización involucradas en procesos como el crecimiento celular, la diferenciación o el metabolismo energético [107].

Los miembros pertenecientes a este grupo de genes son: i) El PGC1 $\alpha$  que es un coactivador transcripcional cuyo modo de actuación consiste en la interacción con el PPAR $\gamma$ , quien funciona como “comodín” para favorecer otras posibles interacciones posteriores con múltiples factores de transcripción. La proteína PGC1 $\alpha$  es capaz de regular la actividad del CREB y de los NRFs (*Nuclear respiratory factors*), entre otras, aportando un vínculo directo entre estímulos externos y la regulación de la biogénesis mitocondrial o la determinación del tipo de fibra muscular [25]; ii) El PGC1 $\beta$  (*Peroxisome proliferator-activated receptor gamma coactivator 1-beta*), quien comparte una relación muy estrecha con el primero en estructura y modo de actuación, compartiendo incluso factores de transcripción como diana; iii) el PRC (*PGC-1 related coactivator*), el cual no tiene una relación tan estrecha, pero es esencial en el desarrollo embrionario, donde sus funciones son muy similares a las de los anteriores [29, 107].

Los primeros estudios, llevados a cabo sobre células de mamíferos en cultivos *in vitro*, demostraron que tanto PGC1 $\alpha$  como PGC1 $\beta$  son los encargados de activar los programas de regulación génica que conducen al incremento de la



producción de energía celular. Posteriormente se observó que estos genes se expresan principalmente en tejidos con alta capacidad oxidativa, tales como el corazón, el músculo esquelético o el tejido adiposo marrón, donde tienen funciones críticas en la regulación de las funciones mitocondriales y del metabolismo energético celular [106]. Estos coactivadores dan respuesta a estímulos biológicos, colmando a las células de todo lo necesario para satisfacer las demandas energéticas de un entorno cambiante. Para ello activan procesos como la biogénesis mitocondrial, las tasas de respiración celular y la absorción y utilización del sustrato energético. Para ejercer este efecto solo requieren de la activación directa de los factores de transcripción NR (Receptores Nucleares) y no-NR involucrados en el control del metabolismo celular [106].

### 1.6.3. Funciones principales del PGC1 $\alpha$

Puesto que se han identificado múltiples patrones de actuación del PGC1 $\alpha$ , éste es considerado como el regulador pleiotrópico de las rutas metabólicas involucradas en el metabolismo energético celular tanto dentro como fuera de la mitocondria [106].

Entre los principales patrones de actuación de este gen, podemos observar que el PGC1 $\alpha$  es el encargado de la coactivación del NRF-1 (Nuclear respiratory factor-1) y NRF-2, con lo que consigue controlar el amplio proceso de la biogénesis mitocondrial. La familia de genes NRF son los encargados de regular la expresión del factor-A de transcripción mitocondrial, también conocido como Tfam, que es esencial para la replicación, manutención y transcripción del ADNmt. A su vez, los NRF también controlan la expresión de otros genes nucleares que codifican para las subunidades de la cadena respiratoria y otras proteínas necesarias para la función mitocondrial [106, 107].

Además, PGC1 $\alpha$  ejerce su función de coactivador sobre múltiples NR y no-NR a través de los cuales es capaz de controlar muchos aspectos del metabolismo energético mitocondrial. De este modo, PGC1 $\alpha$  ejerce control sobre rutas metabólicas mitocondriales como la oxidación mitocondrial de los ácidos grasos, la fosforilación oxidativa; y rutas metabólicas externas al mitocondrion como la gluconeogénesis, el transporte de glucosa, la lipogénesis o la condrogénesis [106]. Incluso hay evidencias de que es capaz de conferir especificidad a sus genes diana.

Por todas estas funciones, el PGC1 $\alpha$  es un gen altamente inducible en respuesta a diversas condiciones fisiológicas que demandan un incremento de la producción de energía mitocondrial. De este modo, la expresión del PGC1 $\alpha$  se ve estimulada ante la activación del músculo energético en un ejercicio físico, durante el ayuno o ante bajas temperaturas [106, 107]. La inducción de este gen está regulada por mecanismos de deacetilación o fosforilación, llevados a cabo por SIRT1 (Sirtuin-1)

## 1 . Introducción

---

y AMPK (AMP-activated protein kinase) respectivamente. Por su parte, las fosforilaciones además pueden mejorar también la síntesis de nuevas mitocondrias, al tiempo que promueven la destrucción de las células defectuosas y reducen el crecimiento celular [107].

### 1.6.4. Rutas metabólicas en las que participa PGC1 $\alpha$

En un inicio, la proteína codificada por este gen parecía actuar solo como coactivador del PPAR $\gamma$ . Sin embargo, cada vez hay más evidencias de su capacidad para interactuar con la mayoría de los miembros de la superfamilia de receptores nucleares y ligandos [29]. Tanto esta versatilidad como su ubicuidad permiten a PGC1 $\alpha$  ser el encargado de controlar la mayoría de programas biológicos en multitud de tejidos, siendo una de sus principales funciones, aunque no la única, el control del metabolismo energético [29].

Como ya se ha comentado con anterioridad (Sección: Las mitocondrias en el metabolismo energético), para llevar a cabo el control del metabolismo energético, el PGC1 $\alpha$  ejerce una función fundamental en la ruta de la biogénesis mitocondrial (hsa03029), que es la ruta encargada de generar mitocondrias funcionales. Esta ruta se puede activar tanto de forma endógena como exógena. Entre los activadores exógeno se hallan los factores ambientales capaces de causar agotamiento energético como la realización de ejercicio físico, los periodos de ayuno o la exposición al frío. Estos factores tienen como consecuencia una reducción en los niveles de ATP y un aumento de AMP, lo que activa a la AMPK, proteína que actúa como un sensor del estado energético celular. De este modo, se produce el inicio de la ruta de señalización de la AMPK.

La activación de la ruta de señalización de la AMPK (hsa04152; Figura 4) conlleva la inhibición de las rutas metabólicas de síntesis y de activación de las catabólicas para la obtención de ATP. Con este objetivo, entre otras funciones, el AMPK se encarga de fosforilar a PGC1 $\alpha$ , activando así a la biogénesis mitocondrial. Estos son sólo algunos de los sistemas de activación de la expresión del PGC1 $\alpha$  en condiciones fisiológicas, cuyo objetivo es activar la biogénesis mitocondrial y reponer los niveles de ATP consumidos.

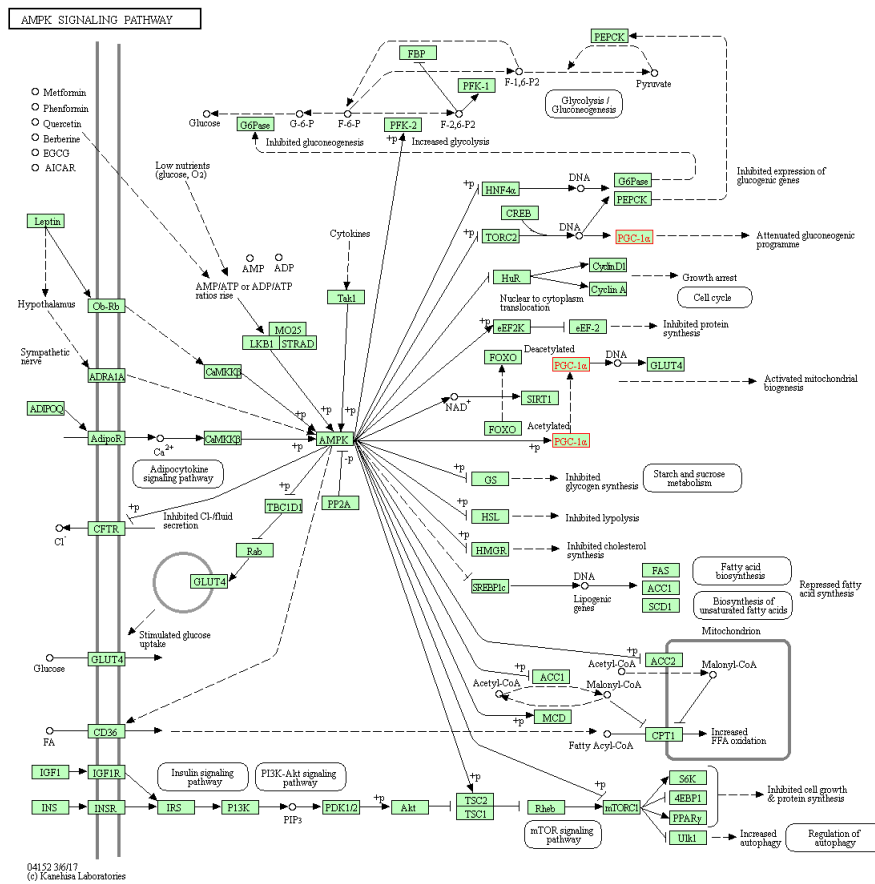


Figura 4: Ruta de señalización de la AMPK (KEGG)

Otras rutas metabólicas en las que el gen *PGC1α* tiene un papel fundamental, según la base de datos del KEGG (*Kyoto Encyclopedia of Genes and Genomes*; <http://www.genome.jp/kegg/> cuya última actualización consultada para esta tesis es del 1 de octubre de 2017) [105] son:

La ruta metabólica de señalización del glucagón (*hsa04922*), encargado de mantener la homeostasis de la glucosa mediante la regulación negativa de la actividad de la insulina para evitar la hipoglucemia. Con este propósito, activa la glucogenólisis o la gluconeogénesis e inhibe la glucólisis y la glucogénesis. Para llevar a cabo la gluconeogénesis, *PGC1α* se debe unir al *PPARα* y al *Fox01* formando dos complejos que desencadenan la activación de la *G6Pasa* (Glucosa-6-fosfatasa), encargada de la activación de dicho proceso (Figura 5).

# 1 . Introducción

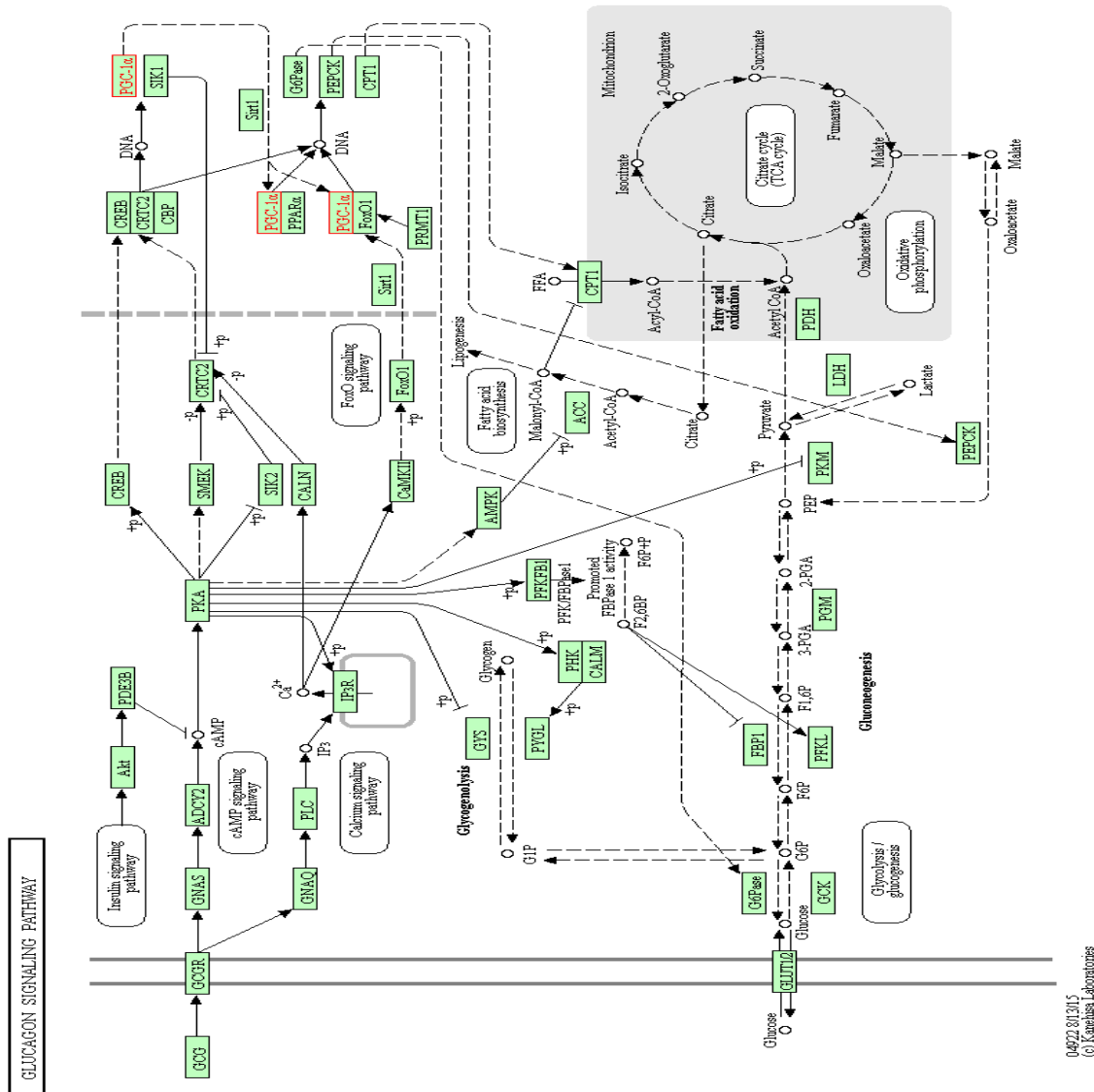


Figura 5: Ruta metabólica de señalización del glucagón (KEGG)

La ruta metabólica de regulación de la longevidad (hsa04211), dependiente de factores genéticos y ambientales, entre los que se puede destacar la restricción calórica como uno de los estilos de vida más reconocidos para alargar la longevidad. Como ya se ha visto con anterioridad, esta restricción conlleva un aumento de la expresión de PGC1 $\alpha$  por dos vías: i) por la activación de ADIPOQ (adiponectina) quien, de manera indirecta, incide sobre AMPK y ii) al fosforilar a CaMK (*calcium/calmodulin dependent protein kinase II gamma*), quien incide sobre PGC1 $\alpha$ . Ambas producen un aumento de la actividad de la biogénesis mitocondrial y, por lo tanto, de la longevidad (Figura 6).

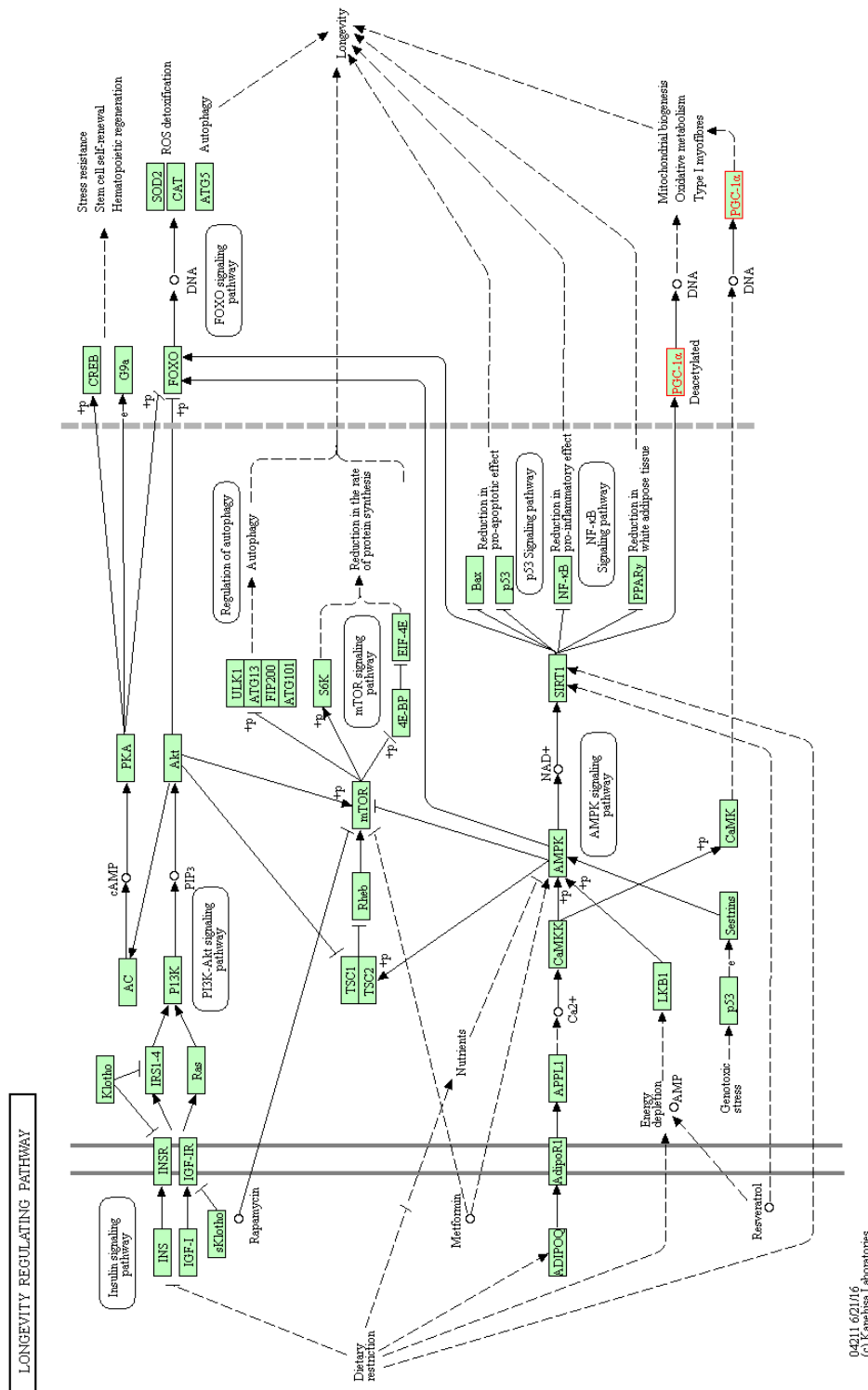


Figura 6: Ruta metabólica de regulación de la longevidad (KEGG)

## 1 . Introducción

La ruta de señalización de la insulina (hsa04910), en la que la unión de la insulina con su receptor desencadena la translocación de la proteína GLUT4 (*solute carrier family 2 member 4*) o la activación de la ruta metabólica RAS/MAPK (*mitogen-activated protein kinase*) independientemente del IRS1 (*Insulin receptor substrate 1*). Además, la insulina es capaz de activar a Akt (*AKT serine/threonine kinase*), encargado de la fosforilación de FoxO1 (*Forkhead box O1*), lo que impide la formación del complejo PGC1 $\alpha$ -FoxO1 y, por lo tanto, permite la glucólisis (Figura 7).

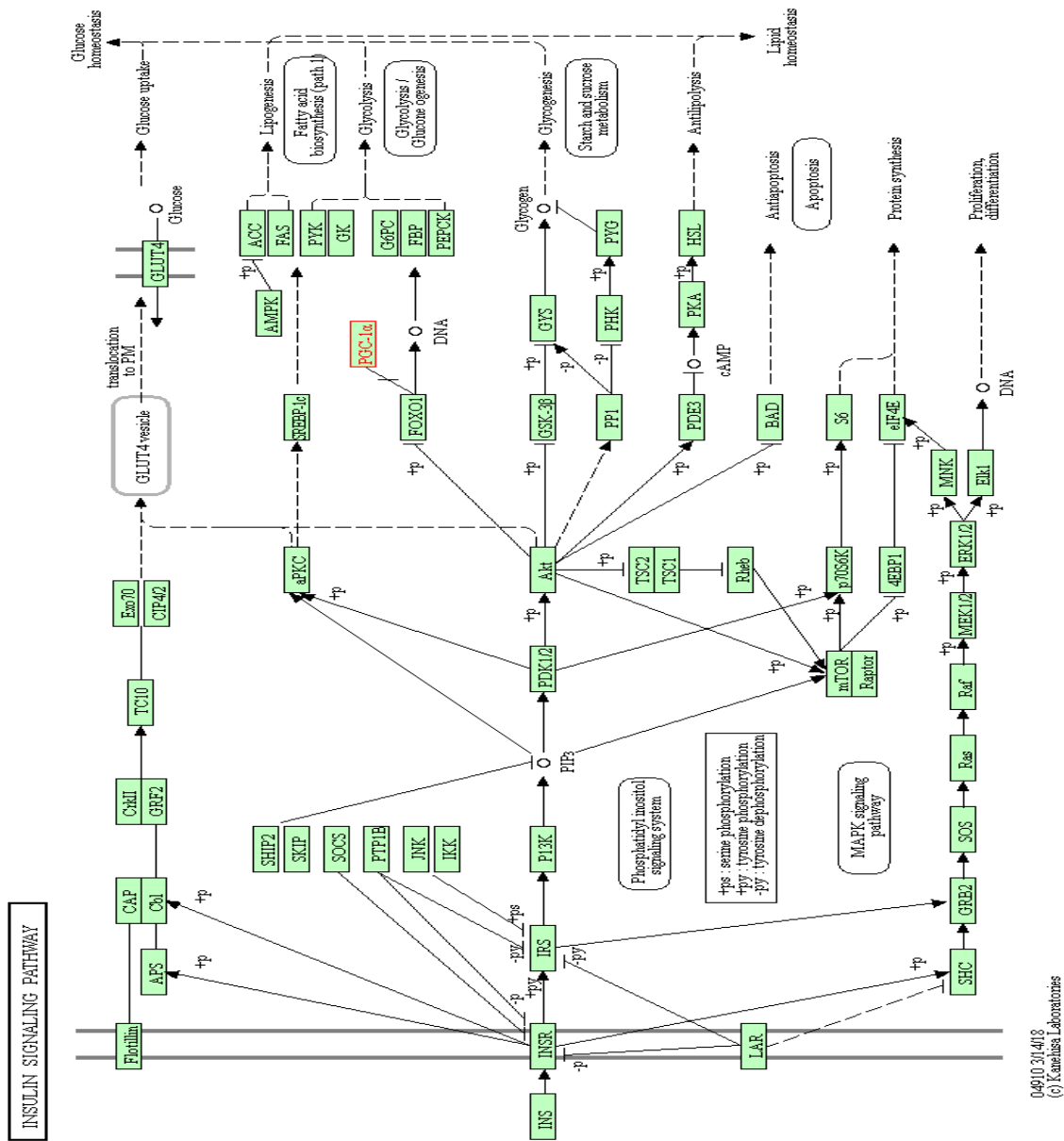


Figura 7: Ruta de señalización de la insulina (KEGG)

La ruta de señalización de la adipocitoquina (hsa04920), encargada de disminuir los niveles de glucosa y ácidos grasos en el plasma, en parte gracias a la activación de la ruta de la AMPK. Esta ruta está regulada por el incremento tanto del volumen como del número de adipocitos, que se correlaciona de forma positiva con la producción de leptinas y, negativa, con la ADIPOQ (Adiponectina). Esto permite a la leptina regular la ingesta de energía mediante la modulación de los niveles de PGC1 $\alpha$  y controlar así la  $\beta$ -oxidación (Figura 8).

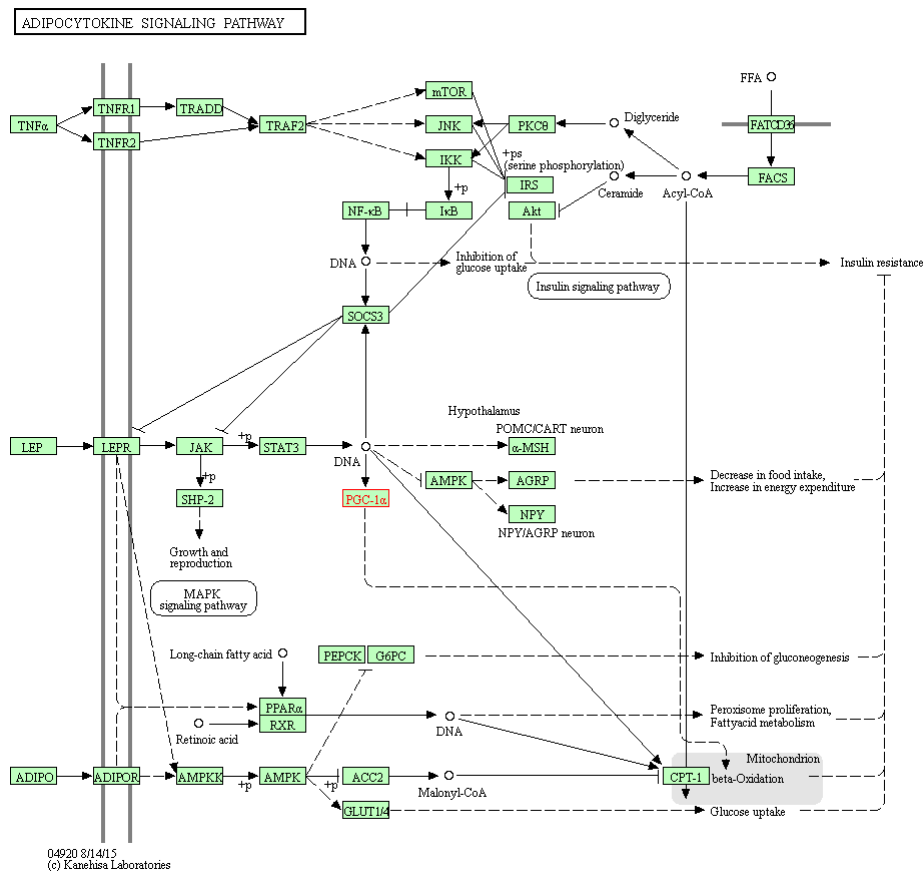


Figura 8: Ruta de señalización de la adipocitoquina (KEGG)

La ruta metabólica de resistencia a la insulina (hsa04931) es un problema en personas con algún tipo de desorden, como es la obesidad, la diabetes mellitus tipo 2, el hígado graso no alcohólico o enfermedades cardiovasculares. Entre los mecanismos que influyen en esta situación podemos citar la disfunción mitocondrial, el estrés oxidativo, la acumulación intracelular de derivados lipídicos o la inflamación, todos ellos procesos en los que PGC1 $\alpha$  tiene una función importante, como ya se ha visto anteriormente (Figura 9).

# 1. Introducción

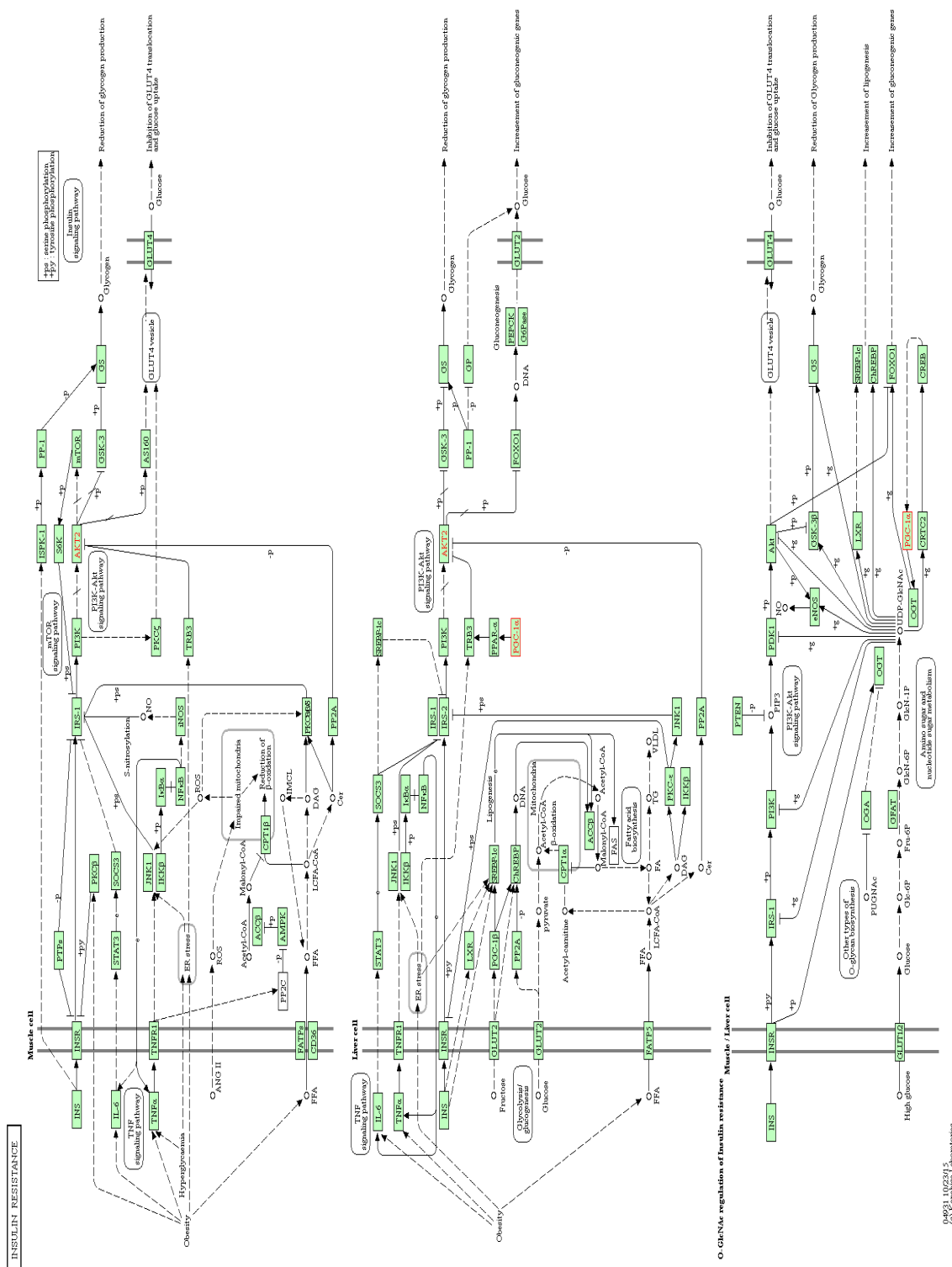


Figura 9: Ruta metabólica de resistencia a la insulina (KEGG)



La ruta metabólica de la Enfermedad de Huntington (hsa05016) es un desorden neurodegenerativo autosómico-dominante que afecta a las neuronas, provocando movimientos involuntarios, cambios en la personalidad y demencia. Entre otras consecuencias, la huntingtina mutada es capaz de formar complejo con CREB y TAF4 (*TATA-box binding protein associated factor 4*) para producir un incremento de PGC1 $\alpha$  y así desencadenar la activación de la termogénesis adaptativa y el sistema de defensa contra las ROS (Figura 10).

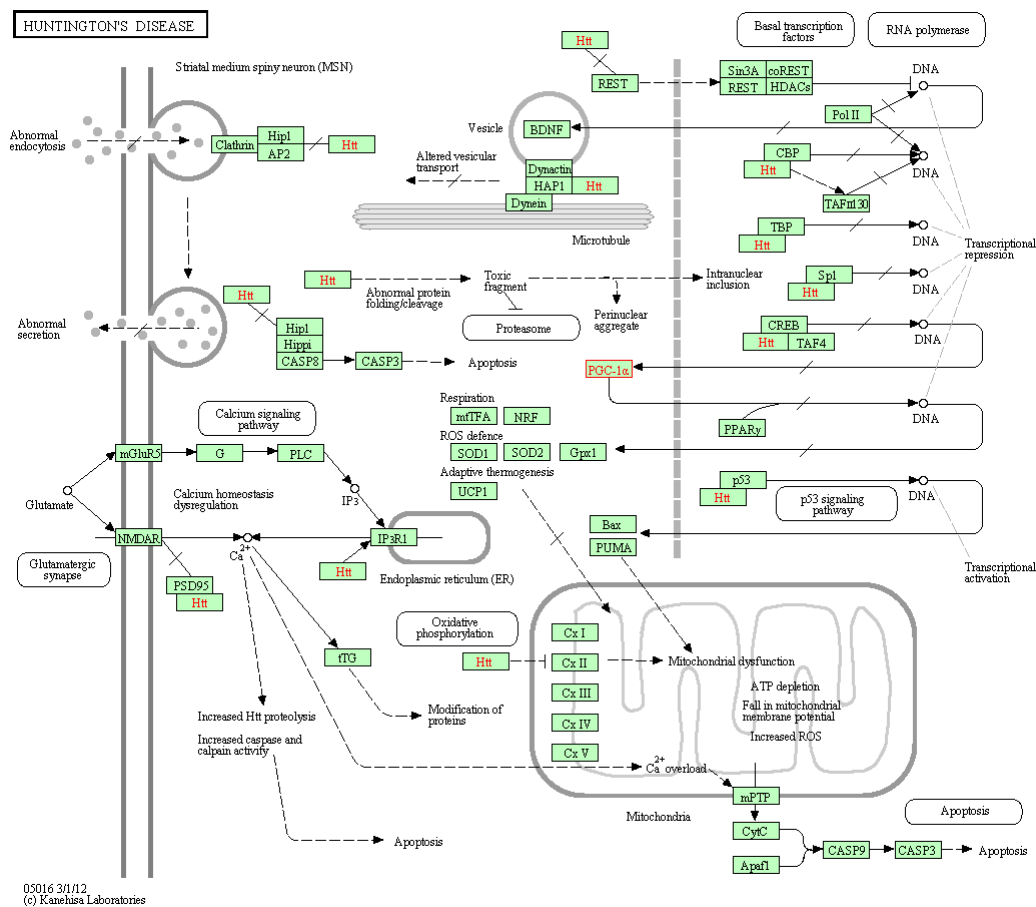


Figura 10: Ruta metabólica de la Enfermedad de Huntington (KEGG)

La ruta metabólica de señalización de la Apelina (hsa04530), que es un péptido endógeno de expresión muy ubicua y capaz de unirse a su receptor (APJ). Está implicada en diferentes procesos fisiológicos clave tales como la activación de la biogénesis mitocondrial, ya que la apelina unida a APJ es capaz de activar a AMPK, quien activa a PGC1 $\alpha$ . Además participa en la angiogénesis, proliferación celular y regulación del metabolismo energético. Además, está implicada en varias patologías que incluyen diabetes, obesidad, enfermedades cardiovasculares y cáncer (Figura 11).

# 1. Introducción

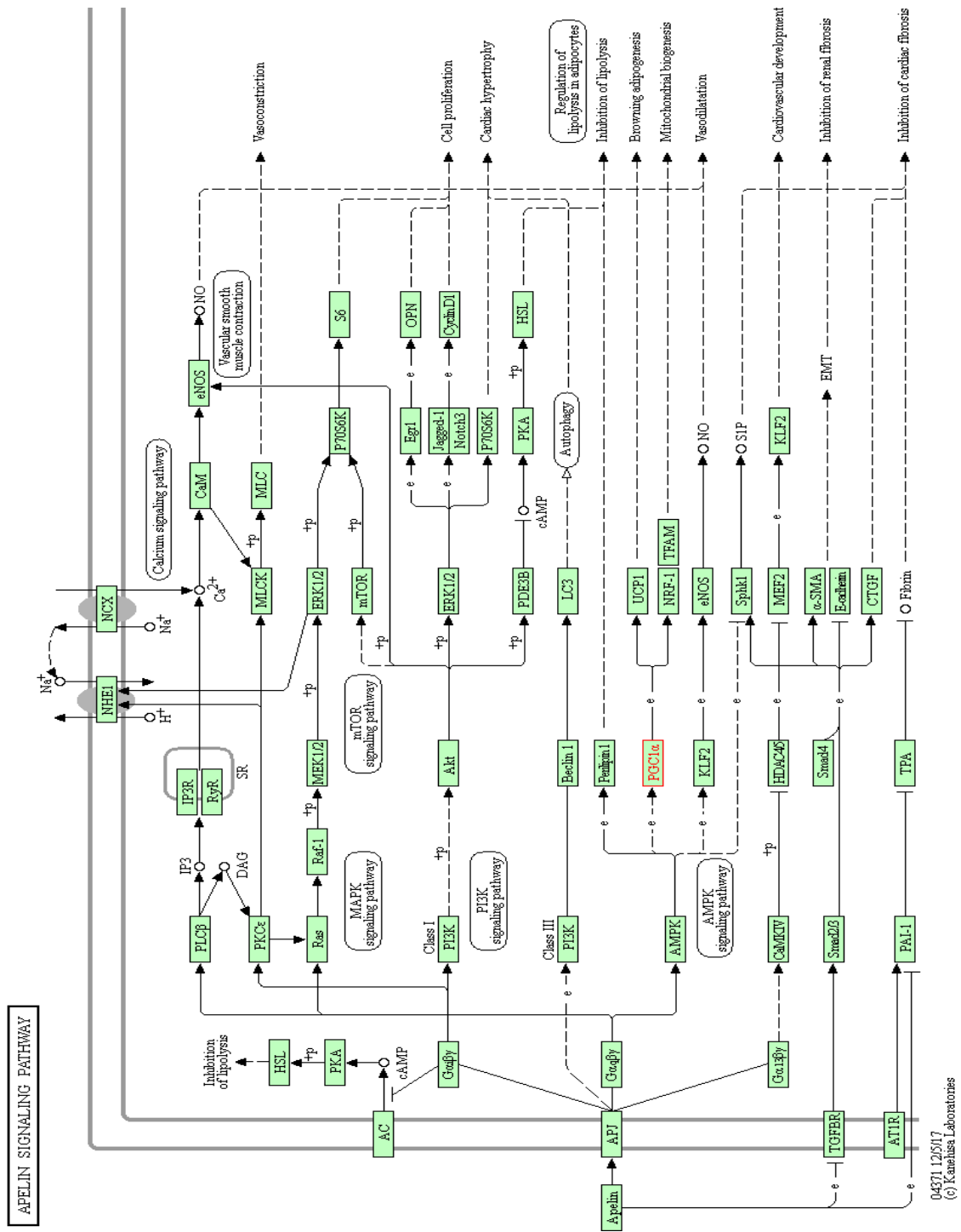


Figura 11: Ruta metabólica de señalización de la Apelina (KEGG)

Tras comprender la función del PGC1 $\alpha$  en la renovación del epitelio intestinal [30] y conocer su papel en las rutas metabólicas en las que participa, es importante destacar que alteraciones en los niveles de expresión de este gen pueden suponer un importante factor de riesgo para el desarrollo del CCR. Sin embargo, a pesar de que el CCR es un problema de salud pública a nivel mundial [46], el conocimiento con respecto a este tema aún está por explorar.

La presente tesis se centrará principalmente en conocer cómo se asocian los polimorfismos presentes en este gen con el riesgo de desarrollar CCR y cómo estas alteraciones, junto con otras presentes en genes que actúen en las mismas rutas metabólicas, pueden influir sobre la misma enfermedad.

CAPÍTULO 2

# Objetivos

---

### **Objetivo general:**

Evaluar la asociación del PGC1 $\alpha$  y de las rutas en las que éste participa con el riesgo de desarrollar cáncer colorrectal.

### **Objetivos específicos:**

1. Desarrollar dos mapas conceptuales que determinen la función del PGC1 $\alpha$  en el desarrollo del cáncer colorrectal en base al conocimiento actual.
2. Analizar los SNPs mapeados en el gen PGC1 $\alpha$  y su posible asociación con el cáncer colorrectal.
3. Analizar la influencia que presentan los factores ambientales en la asociación entre los SNPs del PGC1 $\alpha$  y el cáncer colorrectal.
4. Comparar las técnicas de análisis de conjuntos de genes, su capacidad de discriminación entre conjuntos de genes significativos y no significativos.
5. Evaluar las ventajas o desventajas que presenta contemplar como resultado principal el consenso de los resultados de varias técnicas.
6. Analizar la posible asociación entre el riesgo de desarrollar cáncer colorrectal y las diferentes rutas metabólicas en las que el PGC1 $\alpha$  presenta una función.

CAPÍTULO 3

# Mapas conceptuales de la función del PGC1 $\alpha$ en el CCR

---

## 3.1. Introducción

Como ya se ha visto en la introducción de la presente tesis, la disfunción mitocondrial parece estar involucrada en el desarrollo del CCR, aunque aún no se conoce mucho sobre el mecanismo de acción por el que las alteraciones mitocondriales pueden contribuir a la carcinogénesis [9, 109-111]. La biogénesis mitocondrial es un proceso celular fundamental para el mantenimiento de la función de dicho orgánulo y como mecanismo adaptativo para aportar la energía necesaria ante los cambios de requerimiento energético sufridos [112]. De este modo, factores tanto endógenos como exógenos, así como múltiples rutas metabólicas de señalización y diferentes patrones de expresión génica convergen en el proceso de la biogénesis mitocondrial para coordinar las necesidades energéticas de las células, del tejido o del organismo completo [9, 20, 112].

Muchos de estos factores son capaces de controlar la biogénesis mitocondrial modificando la acción del PGC1 $\alpha$  [9, 20]. Este gen se encarga de controlar la producción de proteínas mitocondriales [113] y es el coordinador principal del metabolismo energético y de la biogénesis mitocondrial, al regular la actividad de diferentes factores de transcripción [114]. De hecho, existen evidencias de que los niveles de expresión del PGC1 $\alpha$  están directamente relacionados con la actividad de la biogénesis mitocondrial [115].

En este capítulo, se pretende conocer la relación entre el CCR y el PGC1 $\alpha$ , ya que este es poco conocido actualmente. Sin embargo, son muchos los estudios que sugieren que el PGC1 $\alpha$  y genes relacionados con él, capaces de regular diversas rutas metabólicas, pueden jugar un papel importante en el desarrollo del CCR [113, 116]. Aunque no existe un mecanismo ampliamente aceptado que defina el modo de actuación del PGC1 $\alpha$  en la oncogénesis del colon o del recto, entender su modo de acción en el desarrollo del CCR, tal vez puedan ayudar a prevenir la enfermedad, reducir su riesgo e, incluso, desarrollar tratamientos terapéuticos contra este tumor.

## 3.2. Material y métodos

Para llevar a cabo la revisión bibliográfica de forma sistemática y totalmente reproducible, se realizaron los siguientes pasos:

### 1. Búsqueda Bibliográfica

La búsqueda bibliográfica se llevó a cabo de forma objetiva en fuentes online y de libre acceso. Las bases de datos consultadas fueron: Cochrane library, PubMed, Scopus, Web of Science (WOS), PsycINFO, Scielo, PLoS One y PubMedCentral (PMC). Tanto la búsqueda bibliográfica como la extracción

### 3 . Mapas conceptuales de la función del PGC1 $\alpha$ en el CCR

---

de datos y la evaluación de la calidad de los artículos seleccionados fueron realizados de acuerdo con la guía “Guidance on the Conduct of Narrative Synthesis in Systematic Reviews” [117] y la declaración PRISMA [118, 119].

Los términos empleados en la búsqueda fueron: “Colorectal”, “Colon”, “Rectum”, “Rectal”, “Cancer”, “Carcinoma”, “Tumor”, “Tumour”, “Neoplasm”, “Cancer Cells”, “Mitochondrial Biogenesis”, “Mitochondrial dysfunction”, “Mitochondria”, “Mitochondrion”, “Warburg effect”, “OXPHOS”, “Oxidative phosphorylation”, “Anaerobic glycolysis”, “PGC1A”, “PPARGC1A”, “Peroxisome Proliferator-activated Receptor gamma coactivator 1 alpha”. Todos ellos fueron combinados con operadores booleanos como “OR” y “AND”, respectiva y debidamente usados. La búsqueda fue restringida a artículos escritos en lengua inglesa.

#### 2. Proceso de selección y criterios de elegibilidad

Para la selección de los artículos, se consideraron diferentes criterios de elegibilidad, que fueron definidos considerando los consejos de las dos guías empleadas en la realización de la revisión sistemática [117, 119], los cuales fueron adecuados a las necesidades de la revisión.

La lista de artículos obtenida en el proceso de búsqueda fue examinada manualmente para identificar los artículos más relevantes. Para ello, el primer paso fue examinar los títulos y eliminar aquellos trabajos claramente irrelevantes. El segundo consistió en una lectura comprensiva de los resúmenes de cada artículo, eliminando los que no estuviesen directamente relacionados con el objetivo de la revisión. Posteriormente, se importó el conjunto de artículos seleccionados en un gestor de referencias bibliográficas (EndNote), lo que permitió eliminar los artículos duplicados. Finalmente, se leyó y estudió el texto completo del artículo, decidiendo la inclusión o exclusión del mismo en función de los siguientes criterios de elegibilidad:

El primer criterio empleado consistió en la exclusión por título. En este punto fueron eliminados todos aquellos artículos no relacionados con cáncer en general, con cáncer colorrectal o con las enfermedades inflamatorias del intestino (Colitis ulcerosa o enfermedad de Crohn). También se excluyeron artículos que tratasen genes diferentes al PGC1 $\alpha$ , al PGC1 $\beta$  o a genes cercanos a estos.

El segundo criterio empleado fue la exclusión por resumen, donde fueron excluidos todos aquellos artículos que no mantuviesen su objetivo en el



### 3 . Mapas conceptuales de la función del PGC1 $\alpha$ en el CCR

---

estudio del CCR o de las enfermedades inflamatorias que, en última instancia, dan lugar a la primera enfermedad. También se eliminaron aquellos que no trataran alguna de las isoformas del PGC1 $\alpha$ , PGC1 $\beta$  o cualquier otro gen/ruta metabólica relacionado con ellos.

Por último, se aplicó el criterio de exclusión por texto completo. En este aparato, y debido a la falta de material específico, los criterios de elegibilidad, aunque mucho más concretos, abarcan una amplia variedad de términos. i) Se incluyeron estudios de investigación básica, basados en líneas celulares o en animales. También se consideraron estudios en humanos, incluyendo estudios observacionales. ii) Se incluyeron artículos centrados en CCR directamente o en otras enfermedades siempre que, como último fin, se analizase el CCR. iii) No se consideraron factores demográficos debido al tipo de estudios seleccionados.

Tras la aplicación de los tres criterios anteriores, se llevó a cabo una replicación de la búsqueda por dos revisores independientes, quienes constituyeron sus propias listas de artículos de inclusión. Las tres listas obtenidas por los tres revisores totales se confrontaron, revisaron y discutieron para, en última instancia, acordar los artículos a considerar cumpliendo al máximo con los criterios de inclusión. Con esta lista, se llevó a cabo una síntesis preliminar, mediante el uso de tabulaciones y una representación visual de los datos que muestra las características clave de cada uno de los artículos.

#### 3. Relación entre artículos

Conocida la lista de artículos a incluir en la revisión, se comenzó por conocer la relación existente entre ellos. De este modo, se comprobó el nivel de homogeneidad que presentan, constatando la plausibilidad de su comparación. Para comprobar dicha homogeneidad se valoró, en términos generales, la presencia o ausencia de las características enumeradas a continuación en cada uno de los artículos:

- Cáncer colorrectal (Enfermedades inflamatorias previas al tumor = 0)
- Expresión del PGC1 $\alpha$  o PGC1 $\beta$ .
- Relación con otros genes.
- Presencia de especies reactivas de oxígeno.
- Biogénesis mitocondrial.
- Tratamiento por quimioterapia.

### 3 . Mapas conceptuales de la función del PGC1 $\alpha$ en el CCR

---

Todas las características fueron valoradas como 1 en caso de estar presentes en el artículo y como cero en caso de ausencia. Posteriormente se calculó la suma de todos los valores para obtener una gráfica que muestre la homogeneidad entre los artículos de la revisión.

#### 4. Cadena de ideas y mapa conceptual

Los resultados de la revisión se presentaron como cadenas de ideas. Este es un método que permite conceptualizar y explorar las conexiones entre los hallazgos de cada uno de los artículos. Usando estas cadenas, se procedió a la obtención de dos mapas conceptuales, es decir, dos figuras que unen todas las evidencias halladas en los diferentes trabajos incluidos en la revisión. De este modo, se construyó un modelo de conceptos clave que muestran las posibles relaciones existentes entre PGC1 $\alpha$  o PGC1 $\beta$  y el desarrollo del CCR.

#### 5. Comprobación de la validez de la síntesis

Por último, se intentó comprobar la validez de esta revisión mediante su comparación con otras revisiones sistemáticas similares.

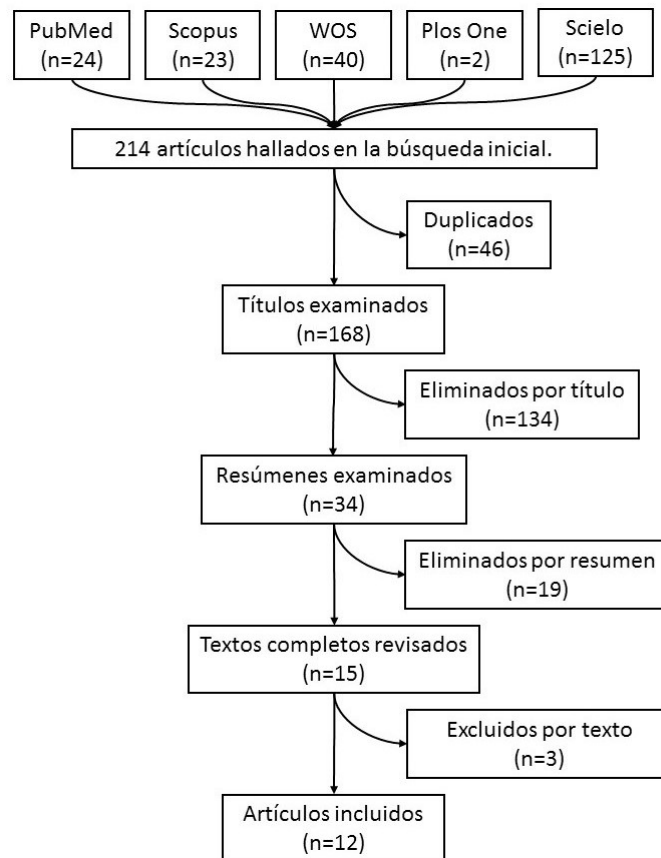
## 3.3. Resultados

### 1. Búsqueda bibliográfica

Mediante el uso de los términos de búsqueda descritos en la sección anterior (Criterios de elegibilidad), pero sin tener en cuenta aún los operadores booleanos, se identificó un total de 7688 manuscritos en la suma de todas las webs consultadas (PsycInfo y Cochrane no devolvieron ningún resultado). Al incluir los diferentes operadores booleanos, el número descendió a un total de 214 artículos, de los cuales se valoraron 34 resúmenes, ya que el resto fue eliminado por el primer criterio considerado (El título). De estos, sólo 15 fueron considerados para la revisión del texto completo y 12 fueron los elegidos para la elaboración de la revisión. La Figura 12 muestra el proceso de selección.

Los 18 artículos excluidos a partir de la selección por resumen, lo fueron por alguno de los siguientes motivos: i) El artículo menciona genes relacionados con el PGC1 $\alpha$  pero en ningún momento menciona al gen de interés en sí mismo; ii) El artículo menciona al gen pero no su relación con el cáncer; iii) El artículo menciona el gen y su relación con el cáncer o con enfermedades inflamatorias del intestino, pero no llegan a mencionar el CCR en concreto; iv) El artículo menciona tanto el gen como la enfermedad de interés, pero no una posible relación entre ellos.

### 3 . Mapas conceptuales de la función del PGC1 $\alpha$ en el CCR



**Figura 12: Resumen del proceso de selección de artículos.** Del total de 214 artículos hallados en la búsqueda bibliográfica, solo 12 cumplieron los criterios de selección y, por lo tanto, fueron considerados en la revisión.

#### 2. Síntesis preliminar

Se utilizó el método de tabulación para generar una tabla con las características principales de cada uno de los artículos seleccionados para la revisión. Estas características deben aportar información suficiente como para comprender perfectamente la posible relación entre el gen PGC1 $\alpha$  y la enfermedad de estudio, el CCR. El mensaje clave obtenido de todos los artículos permite comprender que alteraciones en la expresión del PGC1 $\alpha$  afectan al riesgo de desarrollar CCR a través de diversos mecanismos moleculares. Entre estos mecanismos se pueden incluir la biogénesis mitocondrial, los sistemas antioxidantes, las especies reactivas de oxígeno, la síntesis de novo de lípidos, la glicólisis o la alteración en la expresión de otros genes emparentados con el PGC1 $\alpha$ . Los conceptos principales de los artículos están recogidos en la Tabla 2.

Tabla 2: Resumen de los conceptos clave de los estudios seleccionados

Título	Primer autor	Año	Resultados principales	Conclusión
<b>PGC1<math>\alpha</math>/<math>\beta</math> upregulation is associated with improved oxidative phosphorylation in cells harboring nonsense mtDNA mutations [116]</b>	Sarika Srivastava	2007	<ol style="list-style-type: none"> <li>1) Tanto PGC1<math>\alpha</math> como PGC1<math>\beta</math> están sobreexpresados en V425.</li> <li>2) La sobreexpresión de estos genes estimula la respiración mitocondrial, al menos en osteosaroma.</li> <li>3) Además, la sobreexpresión de PGC1<math>\alpha</math> estimula la actividad del complejo IV.</li> <li>4) La sobreexpresión del PGC1<math>\alpha</math>/<math>\beta</math> también puede estimular la respiración en células deficientes en fosforilación oxidativa.</li> </ol>	En las células V425, los eventos de señalización dependientes de Ca <sup>2+</sup> + se activan durante periodos relativamente más largos, lo que puede activar genes nucleares (incluyendo PGC1 $\alpha$ / $\beta$ ) implicados en la invasión tumoral y la metástasis. Esta vía podría explorarse como un enfoque terapéutico para el tratamiento de enfermedades mitocondriales humanas.
<b>PGC-1<math>\alpha</math> Enhances Antiproliferative Activity of 5'-Deoxy-5-Fluorouridinein Cancer Cells through Induction of Uridine Phosphorylase [120]</b>	Xingxing Kong	2009	<ol style="list-style-type: none"> <li>1) PGC1<math>\alpha</math> induce la expresión de la UPasa tanto en células cancerígenas de mama como de colon.</li> <li>2) Esta inducción es mediada por ERR<math>\alpha</math>.</li> <li>3) La sobreexpresión de PGC1<math>\alpha</math> sensibiliza a las células cancerígenas frente a 5-DFUR.</li> </ol>	El PGC1 $\alpha$ parece ser el regulador de la transcripción de la UPasa, encargada de catalizar la transformación del 5-DFUR a 5FU, lo que inhibe la proliferación de células cancerígenas. En este caso, PGC1 $\alpha$ aumenta la sensibilidad de las células al tratamiento.

Continúa en la siguiente página

### 3 . Mapas conceptuales de la función del PGC1 $\alpha$ en el CCR

Título	Primer autor	Año	Resultados principales	Conclusión
Overexpression of PGC1-alpha enhances cell proliferation and tumorigenesis of HEK293 cells through the upregulation of Sp1 and Acyl-CoA binding protein[113]	Sung-Won Shin	2014	<ol style="list-style-type: none"> <li>1) PGC1<math>\alpha</math> es capaz de acelerar la proliferación de líneas celulares como HEK293 y CT-26.</li> <li>2) La inhibición de la expresión de PGC1-<math>\alpha</math> resulta en un descenso de la proliferación de células tumorales de CCR humanas.</li> <li>3) El PGC1<math>\alpha</math> promueve el potencial oncogénico de las células HEK293.</li> <li>4) La sobreexpresión de PGC1<math>\alpha</math> en células HEK293 disminuye su sensibilidad al estrés oxidativo.</li> </ol>	La sobreexpresión de PGC1 $\alpha$ incrementa la proliferación de diferentes líneas celulares de CCR (HEK293 y CT26). Además, su expresión se correlaciona con un aumento de la carcinogénesis y su inhibición resulta en un descenso de la proliferación celular.
PGC1 $\beta$ promotes enterocyte lifespan and tumorigenesis in the intestine [121]	Elena Bellafante	2014	<ol style="list-style-type: none"> <li>1) PGC1<math>\beta</math> se expresa de modo elevado en el epitelio intestinal y es capaz de modular la morfología del mismo.</li> <li>2) La sobreexpresión del PGC1<math>\beta</math> aumenta la defensa antioxidante de las células.</li> <li>3) La sobreexpresión intestinal de PGC1<math>\beta</math> promueve la carcinogénesis intestinal.</li> </ol>	PGC1 $\beta$ actúa como un regulador capaz de mantener el balance entre la actividad mitocondrial y la producción de especies reactivas de oxígeno.

Continúa en la siguiente página

### 3 . Mapas conceptuales de la función del PGC1 $\alpha$ en el CCR

Título	Primer autor	Año	Resultados principales	Conclusión
<b>AMPK Promotes Aberrant PGC1<math>\beta</math> Expression To Support Human Colon Tumor Cell Survival [122]</b>	Kurt W. Fisher	2015	<ol style="list-style-type: none"> <li>1) El PGC1<math>\beta</math> y el ERR<math>\alpha</math> son reguladores clave aguas abajo de K-Ras, KSR1, y AMPK1.</li> <li>2) Tanto la reducción de AMPK 1 como de K-Ras disminuye los niveles de proteína de PGC1<math>\beta</math>.</li> <li>3) PGC1<math>\beta</math> y ERR<math>\alpha</math> se sobreexpresan en el cáncer de colon y son necesarios para su supervivencia tanto in vivo como in vitro.</li> </ol> (El PGC1 $\alpha$ no es detectado en células HCT116.)	La expresión de PGC1 $\beta$ y ERR $\alpha$ tiene un papel importante en el desarrollo y promoción de las células cancerígenas y su supervivencia.,Ambos pueden suponer nuevas dianas terapéuticas.
<b>PGC1<math>\alpha</math> promotes tumor growth by inducing gene expression programs supporting lipogenesis [19]</b>	Kavita Bhalla	2011	<ol style="list-style-type: none"> <li>1) La pérdida de PGC1 protege contra la oncogénesis tanto en colon como en hígado.</li> <li>2) Su sobreexpresión promueve el crecimiento tumoral <i>in vivo</i>.</li> <li>3) PGC1<math>\alpha</math> es capaz de activar la síntesis de los ácidos grasos, promotores del crecimiento tumoral.</li> </ol>	PGC1 $\alpha$ muestra un papel como promotor de la carcinogénesis y del crecimiento tumoral. Además, es capaz de coordinar la inducción de genes que facilitan la conversión de glucosa a ácidos grasos. Por este motivo, PGC1 $\alpha$ es una potencial diana terapéutica.

Continúa en la siguiente página

### 3 . Mapas conceptuales de la función del PGC1 $\alpha$ en el CCR

Título	Primer autor	Año	Resultados principales	Conclusión
<b>Bax is necessary for PGC1<math>\alpha</math> pro-apoptotic effect in colorectal cancer cells [123]</b>	Ilenia D'Errico	2011	<ol style="list-style-type: none"> <li>1) PGC1<math>\alpha</math> induce la activación de Bax.</li> <li>2) PGC1<math>\alpha</math> incrementa la actividad mitocondrial.</li> <li>3) PGC1<math>\alpha</math> induce la apoptosis en presencia de Bax, pero no en su ausencia.</li> <li>4) PGC1<math>\alpha</math> inhibe el crecimiento tumoral en presencia de Bax.</li> </ol>	PGC1 $\alpha$ se encarga de inducir a Bax, de activarlo y de translocarlo a la mitocondria. Y, en presencia del mismo, el PGC1 $\alpha$ induce la acumulación de especies reactivas de oxígeno, conduciendo a las células cancerígenas a la apoptosis.
<b>Peroxisome proliferator-activated receptor-<math>\gamma</math> coactivator 1-<math>\alpha</math> (PGC1<math>\alpha</math>) is a metabolic regulator of intestinal epithelial cell fate [30]</b>	Ilenia D'Errico	2011	<ol style="list-style-type: none"> <li>1) Los niveles de expresión del PGC1<math>\alpha</math> son mayores en enterocitos diferenciados que en las células proliferativas del fondo de las criptas.</li> <li>2) PGC1<math>\alpha</math> induce la proliferación mitocondrial y la activación de células cancerígenas en el intestino humano.</li> <li>3) PGC1<math>\alpha</math> induce también la acumulación de especies reactivas de oxígeno y la apoptosis en tejidos específicos.</li> <li>4) PGC1<math>\alpha</math> estimula la biogénesis mitocondrial y la respiración in vivo.</li> </ol>	Los niveles de expresión del PGC1 $\alpha$ pueden decidir el destino de las células epiteliales del intestino al inducir modificaciones metabólicas que inducen la apoptosis. Además, su sobreexpresión puede estimular la biogénesis mitocondrial, las actividades metabólicas y la acumulación de especies reactivas de oxígeno (sin inducir los sistemas antioxidantes en el intestino).

Continúa en la siguiente página

### 3 . Mapas conceptuales de la función del PGC1 $\alpha$ en el CCR

Título	Primer autor	Año	Resultados principales	Conclusión
SIRT1/PGC1 $\alpha$ -Dependent Increase in Oxidative Phosphorylation Supports Chemotherapy Resistance of Colon Cancer [124]	Thomas T.Vellinga	2015	<p>1) La quimioterapia induce a SIRT1 para promover el metabolismo energético. Este gen controla la biogénesis mitocondrial al deacetilar a PGC1<math>\alpha</math>.</p> <p>2) La acción conjunta de SIRT1 y PGC1<math>\alpha</math> protege al CCR del tratamiento.</p>	La quimioterapia induce la fosforilación oxidativa en las células tumorales del colon y recto via SIRT1/PGC1 $\alpha$ , ayudando a las mismas a sobrevivir al tratamiento.
Validation of the Use of DNA Pools and Primer Extension in Association Studies of Sporadic Colorectal Cancer for Selection of Candidate SNPs [125]	Mette Gaustadnes	2006	<p>Los resultados fueron analizados usando el test de <math>\chi^2</math> con un <math>\alpha = 0.05</math>. En este análisis se encontraron cinco SNPs, aunque el análisis no fue reproducible. Sin embargo, el SNP (*604517)3'utr:96516 siempre fue estadísticamente significativo.</p>	Las diferencias observadas entre casos y controles en este artículo podrían ser estadísticamente significativas para una n de 600 casos y 600 controles.

Continúa en la siguiente página



### 3 . Mapas conceptuales de la función del PGC1 $\alpha$ en el CCR

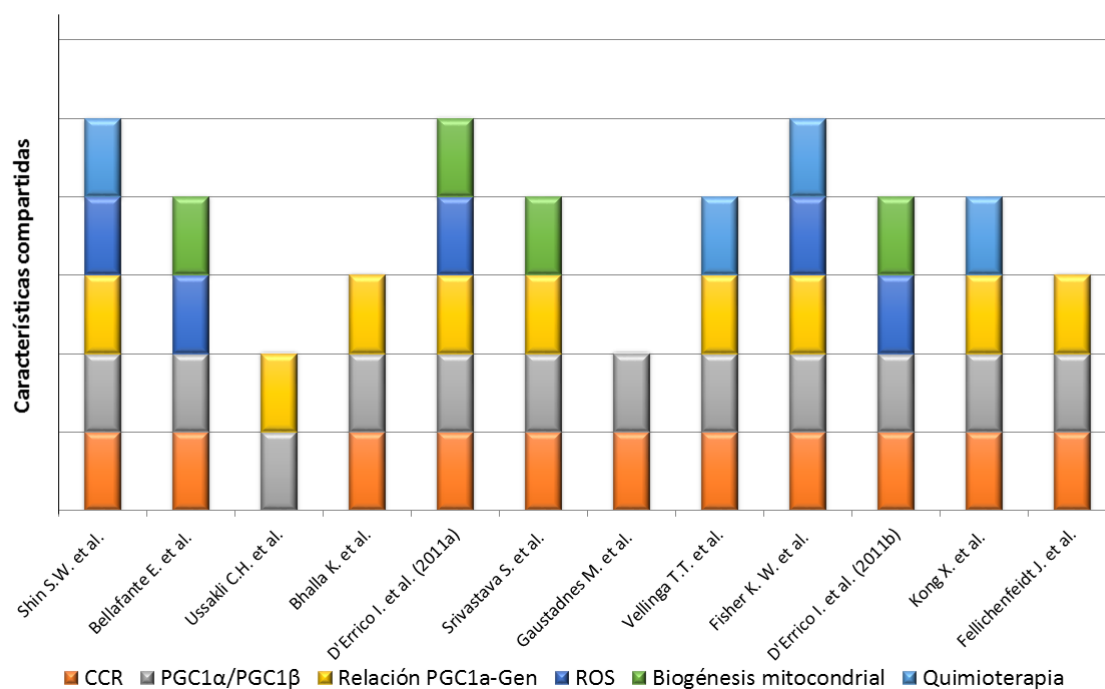
Título	Primer autor	Año	Resultados principales	Conclusión
<b>Peroxisome proliferator-activated receptors (PPARs) and associated transcription factors in colon cancer: reduced expression of PPAR<math>\gamma</math>-coactivator 1 (PGC-1) [126]</b>	Jonas Feilchenfeldt	2004	<ol style="list-style-type: none"> <li>1) La expresión del gen RXR<math>\alpha</math> en la mucosa tumoral es similar al de la mucosa normal.</li> <li>2) La expresión de PGC1<math>\alpha</math> en los tumores es significativamente menor en el tejido tumoral frente al normal.</li> </ol>	<p>Parece que PPAR<math>\beta/\delta</math> es capaz de reprimir a PPAR<math>\alpha</math> y PPAR<math>\gamma</math>, lo que conlleva una menor expresión del PGC1<math>\alpha</math>, y, por lo tanto, de su actividad supresora de tumores.</p>
<b>Mitochondria and Tumor Progression in Ulcerative Colitis [111]</b>	Cigdem Himmetoglu Ussakli	2013	<ol style="list-style-type: none"> <li>1) La comparación de la COX en las biopsias displásicas de los progresores y no progresores de la UC muestra que el desarrollo de tumores podría deberse a niveles anteriores de COX más bajos.</li> <li>2) Los niveles de COX se incrementan tras el desarrollo del tumor, siguiendo un patrón bimodal (El mismo que siguen las mitocondrias).</li> <li>3) PGC1<math>\alpha</math> puede ser el regulador de los cambios mitocondriales observados.</li> </ol>	<p>La pérdida de COX precede a la progresión tumoral en enfermos de UC. Esto puede representar una reducción del número de mitocondrias en la preneoplasia que, posteriormente, es restaurado en el cáncer. Todo ello parece estar regulado por el PGC1<math>\alpha</math>.</p>

### 3 . Mapas conceptuales de la función del PGC1 $\alpha$ en el CCR

La mayoría de los artículos recogidos en esta revisión están basados en investigación básica, usando como material de estudio diferentes líneas celulares o ratones, sin hallar presencia de otros tipos de estudios más avanzados, que recojan datos de ensayos clínicos o similar. Además, son muy pocos los estudios realizados hasta la fecha, lo que permite generar una idea de la necesidad de aumentar la investigación en este campo de estudio.

#### 3. Evaluación de la homogeneidad

Las diferencias entre los artículos incluidos en la revisión fueron analizadas en función de las características más importantes que debía presentar cada uno de ellos. Teniendo en cuenta la presencia o ausencia de estas características, se calculó una puntuación para cada uno de los trabajos incluidos en la revisión. Esta puntuación da una idea de la homogeneidad de los artículos y ha sido representada en la Figura 13.



**Figura 13: Evaluación de la homogeneidad entre los artículos incluidos en la revisión sistemática.** La presencia de la característica analizada se consideró como 1, mientras que su ausencia fue evaluada como cero. En el caso del CCR, la ausencia de esta característica indica la presencia de CU o de la enfermedad inflamatoria del intestino que, en última instancia, conducen al desarrollo de CCR.

Se evaluó la presencia o ausencia de un total de 6 características principales en cada uno de los artículos para comprobar su nivel de homogeneidad:

### 3 . Mapas conceptuales de la función del PGC1 $\alpha$ en el CCR

---

i) Estudio directo sobre el CCR, ii) Evaluación directa del gen PGC1 $\alpha$  o del PGC1 $\beta$  (Característica que debe estar presente en todos los artículos), iii) Análisis de la relación entre el PGC1 $\alpha$  y otro gen, iv) Evaluación de la presencia de especies reactivas de oxígeno en los enterocitos, v) Análisis sobre la biogénesis mitocondrial, vi) Evaluación de la relación entre el tratamiento contra el cáncer (quimioterapia) y el gen PGC1 $\alpha$ . De estas características, el 30.8 % de los artículos las presentaron todas; un 15.4 % presentabas 5 de las 6 características; otro 30.8 % recogía cuatro aspectos principales; el 7.6 % abarcaban la mitad de los puntos; y sólo un 15.4 % presentaba únicamente dos características. Estos porcentajes permitieron considerar que todos los artículos tratados son suficientemente homogéneos como para considerarlos en la revisión sistemática.

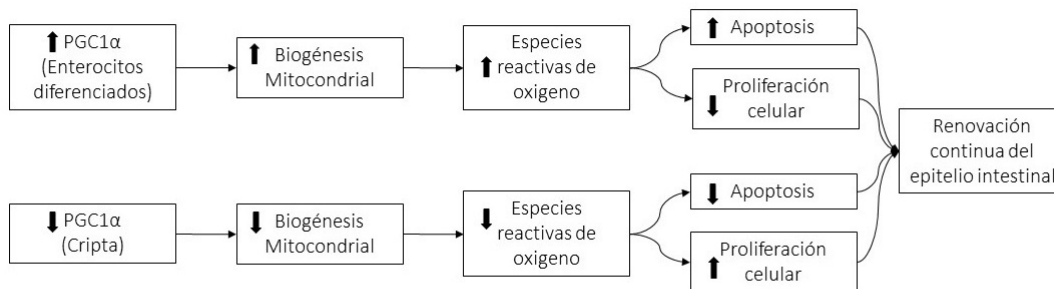
#### 4. Mapa conceptual

Conocida la información que aporta cada artículo al conocimiento actual sobre el papel del PGC1 $\alpha$  en el desarrollo del CCR, diseñamos diferentes cadenas de ideas con las que resumir los conceptos principales para, posteriormente, generar los dos mapas conceptuales que se muestran a continuación. El primero (Figura 14) está basado en el trabajo de D'Errico et al. [30] y muestra cómo varían los niveles de expresión del PGC1 $\alpha$  en el epitelio intestinal a lo largo del eje cripta-vellosidad en condiciones normales. Parece que, en la cercanía de la superficie intestinal, la expresión constitutiva del gen es elevada, aumentando la biogénesis mitocondrial. Esto conlleva un incremento de la acumulación de diferentes especies reactivas de oxígeno, lo que activa el proceso de apoptosis e inhibe la proliferación celular. Por otro lado, en el fondo de las criptas intestinales, la expresión del gen es menor, por lo que la biogénesis mitocondrial no está activa y no se acumulan especies reactivas de oxígeno, por lo que el proceso de proliferación celular se lleva a cabo adecuadamente. Así se completa el proceso de renovación del epitelio intestinal, en el que el PGC1 $\alpha$  parece tener un papel fundamental (Tal como se muestra en el apartado A de la Figura 14).

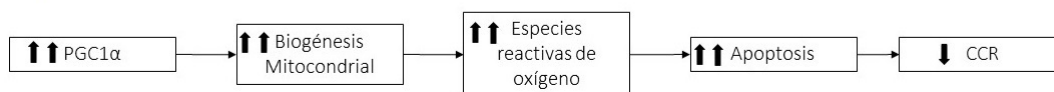
Por otro lado, estos autores también estudian la expresión anormal del gen en células cancerígenas del colon humano (Línea celular HT29). En este caso concluyen que la sobreexpresión del PGC1 $\alpha$  es capaz de activar diversos cambios metabólicos, entre los que se incluye la biogénesis mitocondrial. Se puede observar también un aumento del proceso de la apoptosis por la acumulación de una elevada cantidad de especies reactivas de oxígeno (Tal como se muestra en el apartado B de la Figura 14). La conclusión principal de este trabajo es que el PGC1 $\alpha$  parece comportarse como un regulador del destino de las células intestinales, capaz de protegerlas contra la carcinogénesis.

### 3 . Mapas conceptuales de la función del PGC1 $\alpha$ en el CCR

A) Condiciones normales. Expresión constitutiva del PGC1 $\alpha$  en el epitelio normal



B) Sobreexpresión del PGC1 $\alpha$  en células HT29 (Línea celular de cáncer de colon humano)



**Figura 14: Mapa conceptual basado en el trabajo de D’Errico et al. [30].**

A) Muestra las funciones del gen PGC1 $\alpha$  en condiciones fisiológicas normales en el epitelio intestinal. B) Muestra la posible función de este gen en células HT29 (Imagen de elaboración propia).

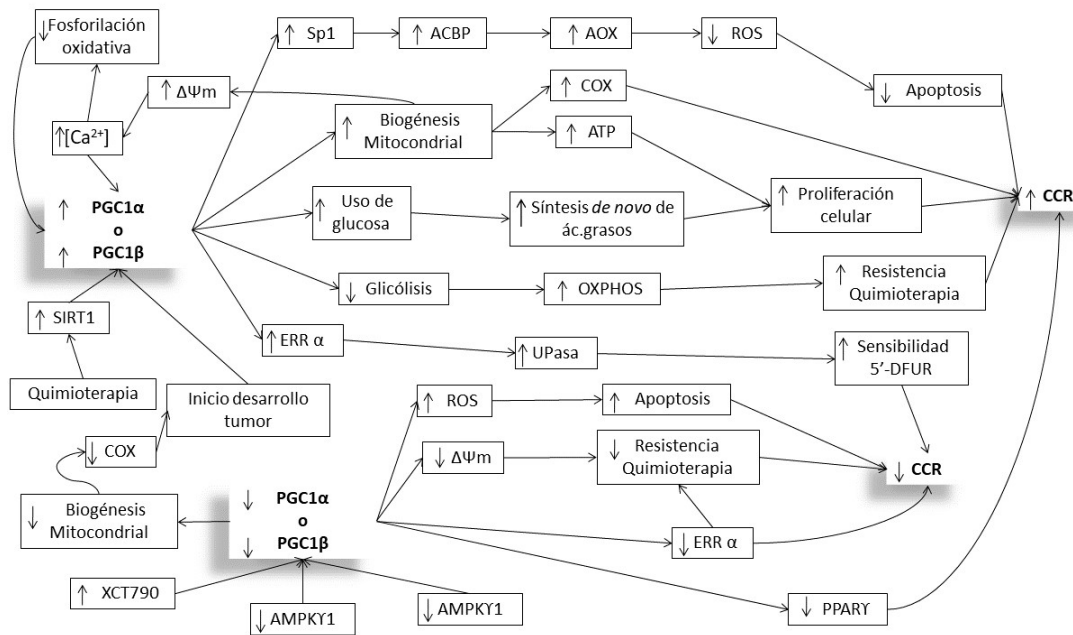
El segundo mapa conceptual muestra los posibles mecanismos de acción del PGC1 $\alpha$  en el desarrollo del CCR. Como se puede observar en la Figura 15, la mayoría de los artículos sugieren que la expresión del PGC1 $\alpha$  se ve incrementada durante el desarrollo y la progresión del cáncer, mientras que una expresión reducida del mismo parece reducir también el riesgo. No obstante, se observan discrepancias entre artículos.

Este mapa (Figura 15) sugiere la existencia de cuatro posibles mecanismos por los que la sobreexpresión del gen puede incrementar el riesgo de desarrollar CCR:

- Incrementando los sistemas antioxidantes, lo que reduce el número de especies reactivas de oxígeno y resulta en la inhibición de la apoptosis.
- Induciendo la biogénesis mitocondrial, lo que incrementa el crecimiento celular y, por lo tanto, la proliferación.
- Aumentando la captación de glucosa, lo que produce un incremento de la proliferación celular nuevamente.

### 3 . Mapas conceptuales de la función del PGC1 $\alpha$ en el CCR

d) Inhibiendo el proceso de glicólisis e incrementando la OXPHOS, lo que podría suponer un incremento de la resistencia celular a la quimioterapia.



**Figura 15: Mapa conceptual general.** Basado en el resto de artículos seleccionados en la revisión sistemática (Imagen de elaboración propia).

Además, el mapa muestra también algunos mecanismos que pueden aumentar la expresión del gen, como es la propia biogénesis mitocondrial (en una retroalimentación positiva) o algunos tratamientos de quimioterapia.

Por otro lado, hay estudios que también apoyan que el silenciamiento o los bajos niveles de expresión del gen pueden promover el desarrollo del tumor por dos mecanismos:

a) Impidiendo que la biogénesis mitocondrial funcione correctamente, lo que resulta en la retroalimentación positiva comentada con anterioridad.

### 3 . Mapas conceptuales de la función del PGC1 $\alpha$ en el CCR

---

- b) Reduciendo los niveles de PPAR- $\gamma$ . El mecanismo por el que esto conlleva un aumento del desarrollo del CCR aún no es conocido.

Por último, en el mapa conceptual general también se pueden observar algunas vías por las que el riesgo o el desarrollo del CCR o la expresión del PGC1 $\alpha$  se ven reducidos:

- a) La pérdida de PGC1 $\alpha$  incrementa los niveles de especies reactivas de oxígeno, lo que es capaz de activar la apoptosis.
- b) Los bajos niveles de expresión del gen conllevan una reducción del potencial de membrana mitocondrial, reduciendo así la resistencia a la quimioterapia.
- c) Una baja expresión del PGC1 $\alpha$  parece reducir los niveles de ERR $\alpha$  (estrogen-related receptor  $\alpha$ ), reduciendo a su vez el riesgo de padecer CCR. Por otro lado, también hay evidencias de que altos niveles de PGC1 $\alpha$  incrementan los niveles de ERR $\alpha$ , lo que da cierta sensibilidad a las células frente a la quimioterapia.
- d) La reducción de los niveles de PGC1 $\alpha$  puede ser debida a mecanismos de inhibición del KSR1 (Kinase suppressor of Ras 1), del AMPK- $\gamma$ 1 o por el tratamiento con XCT790.

Las diferencias entre los resultados de los artículos recogidos en la revisión podrían ser debidas a la desigualdad existente en el material empleado para el estudio (Las diferentes líneas celulares, el empleo de ratones, etc.). En la tabla Tabla 3 se resumen estas diferencias ente los artículos.

Tal como se puede ver en dicha tabla, las líneas celulares empleadas en la mayoría de los artículos son líneas tumorales de CCR humanas y epiteliales y el segundo material más empleado en esta investigación son los ratones.

### 3 . Mapas conceptuales de la función del PGC1 $\alpha$ en el CCR

**Tabla 3: Materiales empleados en los artículos incluidos y sus características.**

Título	Material de uso	Características
<b>PGC1<math>\alpha/\beta</math> upregulation is associated with improved oxidative phosphorylation in cells harboring nonsense mtDNA mutations [116]</b>	VACO425 (Linea celular)	i) Humana
	VACO429 (Linea celular)	i) Humana
<b>Peroxisome Proliferator-Activated Receptor Coactivator-1alpha Enhances Antiproliferative Activity of 5'-Deoxy-5-Fluorouridinein Cancer Cells through Induction of Uridine Phosphorylase [120]</b>	Colo320 (Linea celular)	i) Humana. ii) Indiferenciadas
	HCT116 (Linea celular)	i) Humana. ii) Epitelial
<b>Overexpression of PGC-1<math>\alpha</math> enhances cell proliferation and tumorigenesis of HEK293 cells through the upregulation of Sp1 and Acyl-CoA binding protein [113]</b>	HT-29 (Linea celular)	i) Humana ii) Epitelial
	SNU-C4 (Linea celular)	i) Humana
	CT-26 (Linea celular)	i) Ratón. ii) Epitelial
<b>PGC-1<math>\beta</math> promotes enterocyte lifespan and tumorigenesis in the intestine [121]</b>	Ratón iPGC1 $\beta$ inyectado con PGC1 $\beta$ Humano	
	iPGC1 $\beta$ ratones knockout	i) Ratón transgénico
	Ratones iPGC1 $\beta$ Apc Min/+	
<b>AMPK Promotes Aberrant PGC1<math>\beta</math> Expression To Support Human Colon Tumor Cell Survival [122]</b>	HCT116 (Linea celular)	i) Humana. ii) Epitelial
	Ratones inmunodeficientes (Células HCT116 injertadas en los ratones)	i) Ratón inoculado
<b>PGC1<math>\alpha</math> promotes tumor growth by inducing gene expression programs supporting lipogenesis [19]</b>	PGC1 $\alpha$ ratones knockout	
	PGC1 $\alpha$ +/+ ratones	i) Ratón transgénico
	SCID ratones (HT29)	i) Ratón inoculado
	HT29 (Linea celular)	i) Humana. ii) Epitelial
	Colo205 (Linea celular)	i) Humana. ii) Epitelial

Continúa en la siguiente página

### 3 . Mapas conceptuales de la función del PGC1 $\alpha$ en el CCR

Título	Material de uso	Características
Bax is necessary for PGC1 $\alpha$ pro-apoptotic effect in colorectal cancer cells [123]	HCT116 (Linea celular)	i) Humana ii) Epitelial
	Ratones nude (Inyectados subcutáneamente con las células)	i) Ratón inoculado
Peroxisome proliferator-activated receptor- $\gamma$ coactivator 1- $\alpha$ (PGC1 $\alpha$ ) is a metabolic regulator of intestinal epithelial cell fate [30]	HT29 (Linea celular)	i) Humana. ii) Epitelial
	HCT116 (Linea celular)	i) Humana. ii) Epitelial
	HT29p0 (Linea celular: pérdida total de ADN mitocondrial)	i) Humana. ii) Epitelial
	Ratones xenografiados con células HT29	i) Ratón inoculado
	iPGC1 $\alpha$ ratones transgénicos	i) Ratón transgénico
	Ratones PGC1 $\alpha$ +/+ y PGC1 $\alpha$ -/-	
SIRT1/PGC1 $\alpha$ -Dependent Increase in Oxidative Phosphorylation Supports Chemotherapy Resistance of Colon Cancer [124]	Cultivos en colonosferas shSIRT1 Cultivos en colonosferas shPGC1 $\alpha$	i) Cultivos en colonosferas (Humanos)
Validation of the Use of DNA Pools and Primer Extension in Association Studies of Sporadic Colorectal Cancer for Selection of Candidate SNPs [125]	Dos pools de DNA genómico (Pacientes con CCR esporádico + controles)	i) Humana
Mitochondria and Tumor Progression in Ulcerative Colitis [111]	Progresor de Colitis Ulcerosa (Humano) No progresor de Colitis Ulcerosa (Humano)	i) Humana
Peroxisome proliferator-activated receptors (PPARs) and associated transcription factors in colon cancer: reduced expression of PPAR $\gamma$ -coactivator 1 (PGC-1) [126]	Pacientes con CCR (Humana)	i) Humana



## 3.4. Discusión

El objetivo principal de esta revisión es resumir la evidencia actual del papel del PGC1 $\alpha$  en el inicio y el desarrollo del CCR. Para ello, se realizó un análisis de los diferentes estudios en los que se examina dicha relación. Una de las primeras conclusiones que se pudo obtener es que el número de trabajos que valoran esta cuestión es muy bajo y, por lo tanto, aún queda mucho que investigar en este campo.

Todos los trabajos observados en esta revisión tienen como objeto principal evaluar si un incremento o una disminución en los niveles de expresión del gen PGC1 $\alpha$  afectan al desarrollo del CCR. Sin embargo, cada trabajo fue realizado siguiendo un protocolo diferente, con diferentes materiales de análisis. De este modo, se constató que la mayoría de los estudios (61.5 %) se llevaron a cabo sobre líneas celulares [19, 30, 113, 116, 120, 122, 123]; de esos estudios, cuatro inyectaron dichas células subcutáneamente en los flancos de ratones [19, 30, 122, 123]. Además, dos estudios evaluaron los niveles de expresión del gen en ratones *in vivo* [19, 121]. Y, por último, el resto de estudios se llevaron a cabo sobre pacientes con CCR [111, 124-126], pudiendo destacarse que los resultados fueron similares a los de los estudios *in vitro*.

Cada uno de los artículos estudiados propone un posible mecanismo de acción molecular del PGC1 $\alpha$  en el desarrollo del CCR. Para poder comprender bien éstos, primero es necesario conocer cuál es el mecanismo de acción del gen en condiciones fisiológicas en el intestino. Este mecanismo fue propuesto por D'Errico et al. [30], quienes utilizan cultivos celulares y dos grupos de ratones transgénicos: unos iPGC1 $\alpha$  y otros iPGC1 $\alpha$  Apc<sup>Min/+</sup> para el análisis (Más información en la Tabla 3). En este trabajo se concluye que el PGC1 $\alpha$  es un regulador metabólico capaz de decidir el destino de las células intestinales y que tiene diferentes funciones según el metabolismo energético del tejido en el que se exprese. De este modo, en el fondo de las criptas intestinales, donde el metabolismo energético es elevado, la expresión reducida del gen PGC1 $\alpha$  permite el incremento de la producción de enzimas antioxidantes capaces de proteger a las células de las especies reactivas de oxígeno. Por el contrario, en la región alta de las vellosidades intestinales, con bajo metabolismo energético, la expresión del gen resulta en un incremento de la biogénesis mitocondrial, acumulación de especies reactivas de oxígeno e inducción de apoptosis [30]. De este modo, el PGC1 $\alpha$  parece ser el encargado de regular la renovación del epitelio intestinal, manteniendo un equilibrio entre la proliferación celular y el proceso de apoptosis en condiciones normales (Figura 14-A).

Bajo condiciones no fisiológicas, los mecanismos de actuación propuestos por los diferentes estudios son muy variados. Uno de los mecanismos apoya la idea de que la sobreexpresión del PGC1 $\alpha$  incrementa los niveles de expresión de Sp1

### 3 . Mapas conceptuales de la función del PGC1 $\alpha$ en el CCR

---

(Specificity protein 1). Sp1 es capaz de anclarse a su sitio de unión en ACBP (Acyl-CoA-binding protein), lo que conlleva un incremento de la expresión de este gen. Esta sobreexpresión incrementa la proliferación celular y reduce la sensibilidad al peróxido de hidrógeno, reduciendo así la apoptosis [113, 121].

El segundo mecanismo se basa en la idea de que el PGC1 $\alpha$  es el principal regulador de múltiples aspectos del metabolismo de la glucosa, ya que es capaz de regular las rutas metabólicas que convierten la glucosa en ácidos grasos. Este mecanismo defiende el hecho de que, a su vez, los ácidos grasos sintetizan promotores del crecimiento tumoral y rompen el balance entre la apoptosis y la proliferación celular, promoviendo el desarrollo del cáncer colorrectal [19].

Un posible tercer mecanismo de acción molecular se define mediante la sobreexpresión de PGC1 $\alpha$  o de PGC1 $\beta$ , que tiene dos consecuencias principales. Por un lado, incrementa la actividad de la transferencia de electrones de la cadena de respiración mitocondrial y, por el otro, aumenta la biogénesis mitocondrial [116, 121]. Ambos procesos son capaces de quebrantar el  $\Delta\Psi_m$  (la diferencia de potencial de la membrana mitocondrial) y la capacidad de taponamiento del calcio citosólico [Ca<sup>2+</sup>] produciendo, a su vez, dos consecuencias graves: 1) La OXPHOS disfuncional de las mitocondrias y, 2) la estimulación en cascada de la ruta de señalización de Ca<sup>2+</sup>, quien a su vez puede activar genes implicados en la invasión tumoral y en la metástasis. Asimismo, ambos efectos parecen ser capaces de aumentar la expresión del PGC1 $\alpha$ , lo que consigue un ciclo de retroalimentación positiva [116]. Además, la sobre activación de la biogénesis mitocondrial da a las células la energía necesaria para incrementar la longevidad y la división celular, promoviendo el crecimiento tumoral [116, 121].

Estos tres mecanismos han sido estudiados principalmente en líneas celulares y permiten concluir que tanto PGC1 $\alpha$  como PGC1 $\beta$  mantienen el equilibrio entre la actividad mitocondrial (producción de especies reactivas de oxígeno) y la protección citotóxica (activación de sistemas antioxidantes), manteniendo estable la proporción entre apoptosis y proliferación celular [19, 113, 116, 121]. Además, podemos concluir que estudios realizados sobre líneas celulares [19, 113, 116] y sobre ratones [121] obtienen los mismos resultados, dando soporte a la hipótesis de que una elevada expresión del PGC1 $\alpha$  incrementa el riesgo de padecer CCR.

Sin embargo, D'Errico et al. [123] aportan un mecanismo de actuación molecular que no es concordante con lo expuesto anteriormente. En este caso, la sobreexpresión de PGC1 $\alpha$  parece inducir la translocación de Bax a la mitocondria, quien tiene un efecto pro-apoptótico. Por lo tanto, en este caso, se está observando un efecto protector de la sobreexpresión del gen ante el desarrollo del tumor. Sin embargo, en ausencia de Bax, la sobreexpresión del PGC1 $\alpha$  no es capaz de oponerse al

### 3 . Mapas conceptuales de la función del PGC1 $\alpha$ en el CCR

---

crecimiento tumoral [123]. En este caso la línea celular empleada es inyectada subcutáneamente en ratones, y muestra la primera controversia entre resultados.

Por otro lado, la inhibición o la baja expresión del PGC1 $\alpha$  y PGC1 $\beta$  también parecen estar involucrados en el desarrollo del tumor. En este sentido, Fisher et al. [122] observan que PGC1 $\beta$  presenta una expresión anormal en las líneas celulares tumorales humanas. Esta expresión aberrante incrementa los niveles de ERR $\alpha$  y, por lo tanto, sus propiedades carcinogénicas. Apoyando esta hipótesis, estos autores advierten que en presencia de bajos niveles de KSR1 y AMPK, quienes se encuentran aguas arriba del promotor del PGC1 $\beta$ , se puede observar una disminución de la expresión de dicho gen y, en consecuencia, de ERR $\alpha$ . Lo cual disminuye la supervivencia de las células cancerígenas colorrectales [122].

Basándonos en estas evidencias, se puede deducir que PGC1 $\alpha$ /PGC1 $\beta$  se encuentra sobreexpresado en las líneas celulares de colon humano durante el desarrollo del cáncer, y que esta sobreexpresión es, además, requerida para la posterior proliferación y supervivencia celular y reducción de la apoptosis [19, 113, 116, 121, 122]. De hecho, solo uno de los trabajos que fueron revisados presenta un resultado diferente a lo anterior [123]. Es necesario destacar que, a pesar de que las líneas celulares empleadas en los trabajos son diferentes, todas ellas se caracterizan por ser de humano y epiteliales (Tabla 3), por lo que el origen de las diferencias entre los trabajos no puede ser explicado por este hecho.

Además de los estudios comentados hasta el momento, de la revisión bibliográfica se obtuvieron también otros dos trabajos cuyo objetivo principal se centra en el análisis de los efectos de la quimioterapia en la expresión del PGC1 $\alpha$  en cultivos celulares y colonosferas [120, 124]. Estos estudios sugieren otros tres posibles mecanismos de acción del PGC1 $\alpha$  en relación con el cáncer colorrectal:

1. La quimioterapia induce en las células tumorales colorrectales un aumento de la fosforilación oxidativa dependiente de SIRT1/PGC1 $\alpha$  que promueve la supervivencia tumoral durante el tratamiento. Explicado brevemente, la quimioterapia induce daños en el ADN que resultan en un incremento de la expresión de SIRT1, quien deacetila y, por lo tanto, activa a PGC1 $\alpha$ . Éste estimula la biogénesis y la respiración mitocondrial, aumentando la fosforilación oxidativa, lo que induce un cambio en el metabolismo energético de las células, consiguiendo proteger a las mismas de los diferentes daños citotóxicos que el tratamiento pudiera hacer [124].
2. El PGC1 $\alpha$  induce la expresión de la Uridin Fosforilasa (UPasa) a través de su unión con el ERR. De este modo, la sobreexpresión de PGC1 $\alpha$  actúa al

### 3 . Mapas conceptuales de la función del PGC1 $\alpha$ en el CCR

---

contrario de lo visto anteriormente, aumentando la sensibilidad de las células tumorales al tratamiento por 5-desoxi-5-fluorouridina, al inducir la apoptosis celular [120].

3. Por último, ERR $\alpha$  es capaz de regular la homeostasis del metabolismo energético al interactuar con PGC1 $\alpha$  o PGC1 $\beta$ . Por lo que la sobreexpresión de este gen disminuye la sensibilidad celular a la quimioterapia, mientras que su inhibición reduce las especies reactivas de oxígeno e induce el incremento de potencial de membrana mitocondrial ( $\Delta\Psi_m$ ), aumentando de la sensibilidad a la quimioterapia [122, 127].

Estos mecanismos apoyan la primera conclusión: el PGC1 $\alpha$  y PGC1 $\beta$  permiten a las células mantener el equilibrio entre: i) la actividad mitocondrial y la protección citotóxica ante la producción de especies reactivas de oxígeno y, ii) la apoptosis y la proliferación celular. Además, la mayoría de ellos sugieren que cualquier terapia contra el cáncer capaz de reducir la expresión de PGC1 $\alpha$  aumentará la sensibilidad de las células cancerígenas a la quimioterapia. Por lo que cada vez son más los autores que se decantan por el uso del PGC1 $\alpha$  como diana terapéutica [19, 120, 122, 124, 127]. Es más, Do et al. [128] muestran una estrategia para reducir los niveles de expresión del gen de interés y así mejorar la terapia contra el cáncer. Ellos demuestran, por primera vez, que la metformina (un agente reductor de la insulina) induce el miR-34, capaz de reducir SIRT1 en células cancerígenas con un p53 normal, pero no cuando el p53 está alterado. Sin embargo, a pesar de que este hecho se observa tanto en HCT116 (línea celular de colon) como en MCF-7 (de mama), sólo en esta segunda se ha podido demostrar que la reducción de SIRT1 conlleva una disminución en la expresión del PGC1 $\alpha$ , y consecuentemente una reducción de NRF2 y una mayor susceptibilidad al estrés oxidativo. Por otro lado, DeCensi et al. [129] sí pudieron constatar, mediante un meta-análisis, que la metformina está asociada con un descenso del riesgo de cáncer, incluyendo el CCR en pacientes con diabetes.

Por último, es necesario destacar que en esta revisión sistemática se analizaron también tres artículos basados en muestras humanas. Primero, Ussakli et al. [111] estudiaron 18 pacientes con colitis ulcerosa, de los cuales 9 presentaban un alto grado de displasia o cáncer y los otros 9 estaban libres de esta enfermedad. Estos autores concluyen que, mientras que el desarrollo de la displasia parece ser precedido por un descenso del número de mitocondrias, estos orgánulos son restaurados posteriormente en las células cancerígenas. Lo que sugiere que son necesarios para la posterior proliferación celular. Este patrón bimodal de las mitocondrias parece estar dirigido por la regulación transcripcional que el PGC1 $\alpha$

### 3 . Mapas conceptuales de la función del PGC1 $\alpha$ en el CCR

---

ejerce sobre la biogénesis mitocondrial [111], lo que es consistente con los resultados anteriormente expuestos y basados en líneas celulares [116, 121].

Los otros dos trabajos proponen una idea similar, observando de nuevo que la sobreexpresión del PGC1 $\alpha$  resulta en un mayor riesgo de padecer CCR. Por un lado, Gaustadnes et al. [125] analizan diferentes SNPs en un *pool* de muestras de ADN, encontrando que el polimorfismo rs96516 está asociado estadísticamente con el riesgo de desarrollar un CCR esporádico [125]. Por otro lado, Feilchenfeldt et al. [126] estudiaron los niveles de expresión de todas las isoformas del PPAR- $\gamma$  junto con sus cofactores, entre los que se incluye el PGC1 $\alpha$ . Para ello emplearon pacientes con diferentes estadios de cáncer de colon y encontraron que, a pesar de que los niveles de expresión de PPAR- $\gamma$  varían entre las diferentes etapas del cáncer, los de PGC1 $\alpha$  siempre se ven reducidos con respecto a las muestras normales [126].

Resumiendo, mientras que D'Errico et al. [30] mostraron que los ratones con PGC1 $\alpha$  silenciado son susceptibles a desarrollar un tumor colorrectal, la mayoría de los artículos incluidos en esta revisión muestran que una expresión elevada del gen PGC1 $\alpha$  en las células cancerígenas aumenta la progresión tumoral. Por lo tanto, la función del gen de interés en el riesgo de desarrollar cáncer colorrectal no está del todo clara, aunque sí parece evidente que posee un importante papel en el desarrollo de esta enfermedad. En los próximos capítulos de esta tesis, se pretende aumentar el conocimiento en este tema<sup>1</sup>.

---

<sup>1</sup>Este capítulo ha dado lugar a la publicación del artículo: “Alterations in PGC1 $\alpha$  expression levels are involved in colorectal cancer risk: a qualitative systematic review”, localizado en la sección de anexos: Anexo I

CAPÍTULO 4

# Estudio de la asociación entre los SNPs del PGC1 $\alpha$ y el CCR

---

### 4.1. Introducción

Como ya se ha comentado en la sección del Papel de la genética en el cáncer colorrectal, los SNPs son uno de los factores principales en el estudio de enfermedades complejas [130]. Éstos constituyen una de las variaciones genéticas más comunes y, en función de su localización en el genoma, pueden presentar consecuencias de diferentes dimensiones. Por ejemplo, SNPs en la región codificante del genoma pueden suponer cambios en la secuencia proteica, provocando una funcionalidad reducida o una disfuncionalidad completa de la misma [131]. Por otro lado, los SNPs que se presentan en la región no codificante, aunque no producen cambios en las proteínas, sí pueden provocar alteraciones en la expresión génica o en el empalme alternativo, entre otras [132].

Explicada de forma breve, la regulación de la expresión génica requiere de diversos elementos, entre los que se pueden destacar los factores de transcripción y los elementos reguladores (ER). Ambos necesitan unirse para formar complejos proteicos capaces de llevar a cabo la regulación. En el genoma, los ER pueden localizarse en tres regiones: i) A unos miles de pares de bases (pb) del extremo 5' del gen diana de regulación, en la zona conocida como promotor; ii) A cientos de kb del extremo 5' del gen o; iii) a cientos de Kb del extremo 3' del gen. En estas dos últimas situaciones, los ER funcionan como potenciadores (comúnmente conocidos como *enhancers*) o como silenciadores (comúnmente conocidos como *silencers*) [133]. Las mutaciones puntuales de un solo nucleótido en cualquiera de estos tres lugares se conocen como rSNP (SNP en regiones reguladoras), y suponen un 49 % del total de polimorfismos conocidos [130].

Para el análisis de los polimorfismo recogidos en un único gen, en este caso el PGC1 $\alpha$ , el método más empleado es el conocido como Análisis SNP a SNP o “Single SNP analysis” en inglés (Ver más en Análisis SNP a SNP)[89].

El objetivo principal de este capítulo es hacer un análisis de los SNPs del gen PGC1 $\alpha$  disponibles en la base de datos del Exome Array del proyecto MCC-Spain. Además, se pretende analizar las posibles interacciones de dichos SNPs con factores ambientales, para conocer la asociación entre estas mutaciones y el riesgo de desarrollar CCR<sup>1</sup>.

---

<sup>1</sup>Cabe destacar que el análisis SNP a SNP ha dado lugar a la publicación del artículo: “The RS4939827 polymorphism in the SMAD7 GENE and its association with Mediterranean diet in colorectal carcinogenesis”. Estos resultados se pueden consultar en la sección de anexos: Anexo II.

## 4.2. Material y métodos

### ■ Proyecto MCC-Spain

Para la obtención de los datos empleados en este capítulo, y a lo largo del desarrollo de la tesis, se empleó la base de datos generada en el proyecto MCC-Spain Study (<http://www.mccspain.org>). Este es un proyecto nacional dirigido por el CIBERESP (Spanish Consortium for Biomedical Research in Epidemiology and Public Health) basado en un estudio *multicaso – control* poblacional de los principales tumores de España. Su objetivo principal es analizar la posible interacción de factores genéticos y ambientales en el desarrollo de tumores de alta incidencia en nuestro país, para poder promover la investigación y la prevención de los mismos. Los tumores recogidos para el estudio son:

- Cáncer de próstata: con 1112 casos.
- Cáncer de mama: con 1738 casos.
- Cáncer colorrectal: con 2140 casos.
- Cáncer gastro-esofágico: con 459 casos.
- Leucemia crónica linfocítica (CLL): con 559 casos.

Todos los pacientes incluidos en el proyecto, junto con los individuos que aceptaron participar como controles, dieron su consentimiento por escrito en el momento de la inscripción, poseen edades comprendidas entre 20 y 85 años, han rellenado un cuestionario epidemiológico estándar, con el se que obtuvo constancia de la información demográfica, existencia de antecedentes familiares de primer grado, o hábitos de vida saludables de los individuos, y un cuestionario de frecuencia alimentaria. También se conoce su historia residencial (geocodificada) y ocupacional completa y se dispone de la información clínica de los casos. Además, se estimó el IMC, lo que permitió dividir a la población en cuatro grupos: bajo peso, normopeso, sobrepeso y obesidad. Por último, cada uno de los individuos ha consentido la extracción de una muestra de sangre con la que se ha obtenido el genotipado.

En concreto, para la realización de esta tesis, basada en CCR, se pudo contar con un total de 4093 individuos entre casos y controles con información genética eficiente (1349 casos y 2744 controles), excluyendo el resto de individuos.



##### ■ Selección de SNP

Para este análisis, se llevó a cabo una selección de todos los SNPs disponibles en la base de datos del MCC-Spain contenidos en  $\pm 500$ pb aguas arriba y abajo de la localización exacta del gen PGC1 $\alpha$ : 4p15.2 entre las bases 23793644 y 24474394 (GRCh37.p13) [134]. Dicha selección fue llevada a cabo mediante el software libre PLINK-1.09, un programa de línea de comandos desarrollado por Christopher Chang, con soporte desde NIH/NIDDK's Laboratory of Biological Modeling, la Escuela de Medicina Purcell Labat Mount Sinai y otros [135]. Los datos obtenidos por este software fueron pasados a STATA11 para ser analizados estadísticamente en combinación con el resto de las variables ambientales recogidas en el proyecto. Por último, se llevó a cabo una evaluación ad-hoc de los polimorfismos detectados, seleccionando aquellos SNPs adecuados para su estudio.

##### ■ Análisis estadísticos

Se llevó a cabo un análisis estadístico descriptivo en el que se detalla el número de casos y controles estratificado por cada variable de ajuste de la población general. Posteriormente, se ajustaron los datos para la población de análisis que posee el genotipado de los SNPs de interés, en la que se detallan las frecuencias para el sexo, antecedentes familiares de primer grado (AFP), actividad física y la adherencia a la dieta mediterránea; y la media y la desviación estándar (SD) para la edad y el IMC. A continuación se realizó una evaluación de la asociación entre genotipo y riesgo de desarrollo de CCR mediante una regresión logística no condicional. En ella se determinaron los OR ajustados y los Intervalos de Confianza al 95 % (IC95), empleando como referencia el genotipo homocigoto del alelo de mayor frecuencia.

Para el análisis de interacción SNP-ambiente se dispuso de la regresión lineal, interpretando principalmente el modelo aditivo. Si este modelo no resultaba significativo, se consideraban otros posibles modelos de herencia para comprobar significación estadística (dominante, codominante, recesivo y sobredominante). Por último, se llevó a cabo una regresión logística ajustada por edad, nivel educativo, antecedentes familiares, actividad física, consumo de tabaco, IMC, consumo de alcohol, frutas, verduras, lácteos y carnes rojas. Estratificando la población en función del sexo (hombres y mujeres) y la localización del tumor (recto, colon, colon derecho y colon izquierdo).

### 4.3. Resultados

#### ■ Análisis descriptivos

El número de individuos analizados en este capítulo fue seleccionado en función del tumor a analizar (CCR), de la disponibilidad de los datos genéticos que se requiere emplear ( $\pm 500$ pb aguas arriba y abajo del gen PGC1 $\alpha$ ) y de los factores de riesgo que se desea incluir. De este modo, se pudo incluir en el análisis un total de 4080 individuos, repartidos en 1336 casos (64.75 % hombres; 35.25 % mujeres) y 2744 controles (53.53 % hombres; 46.47 % mujeres). La edad media en este grupo de individuos se centra en 67 años ( $\pm 10.6$ ) en los casos y de 63.3 años ( $\pm 11.5$ ) en los controles (Tabla 4 y Tabla 5).

**Tabla 4: Caracterización según las variables de ajuste categóricas de los individuos participantes en el estudio.**

Factores de riesgo	Controles		Casos	
	N	%	N	%
<b>N</b>	2744	67.25	1336	32.75
<b>Sexo</b>				
<b>Hombres</b>	1469	62.94	865	37.06
<b>Mujeres</b>	1275	73.02	471	26.98
<b>AFIG</b>				
<b>No</b>	2376	90.79	1034	82.2
<b>Sí</b>	241	9.21	224	17.8
<b>Actividad Física</b>				
<b>Sedentario</b>	1033	37.65	620	46.41
<b>Baja</b>	421	15.34	156	11.68
<b>Intermedia</b>	342	12.46	128	9.58
<b>Activo</b>	948	34.55	432	32.34
<b>Adherencia al patrón de dieta mediterránea</b>				
<b>Bajo</b>	817	33.91	447	35.36
<b>Medio</b>	1055	43.79	492	31.8
<b>Alto</b>	537	22.29	224	29.43

Características de casos y controles de CCR para las variables categóricas. Los componentes empleados para obtener la adherencia al patrón de dieta mediterránea son el consumo de alcohol, vegetales, legumbres, carne, productos lácteos, fruta, cereales, pescado y aceite de oliva. Sólo se analizan los individuos con genotipado completo del gen PGC1 $\alpha$ .

#### 4 . Estudio de la asociación entre los SNPs del PGC1 $\alpha$ y el CCR

Para el resto de factores de riesgo estudiados, los datos se reparten del siguiente modo: El IMC presenta una media de 27.5 ( $\pm 4.6$ ) en los casos y 26.7 ( $\pm 4.4$ ) en los controles (Tabla 5). En ambos casos, la media observada permite considerar que la mayoría de la población pertenece a un IMC de sobrepeso. La población presenta un bajo porcentaje de AFPG, ya que el 82.2% en casos y el 90.8% en controles no poseen antecedentes familiares de este tipo. El índice de actividad física refleja una población principalmente sedentaria. El 46.41% de los casos y el 37.65% de los controles tienen un estilo de vida poco activo. Por último, en cuanto a la adherencia al patrón de dieta mediterráneo, un 61.2% de los casos y un 66.1% de los controles presentan una adherencia de media a alta (Tabla 4).

Si en lugar de considerar la adherencia al patrón de dieta mediterráneo, se tiene en cuenta el consumo medio de determinados alimentos, los resultados indican que la ingesta de vegetales es, de media, un 7% mayor en controles que en casos. Por el contrario, la ingesta de carnes rojas y procesadas es un 16% mayor en casos que en controles. Del mismo modo, el consumo de alcohol también es mayor en casos que en controles (aproximadamente un 28%). Por último, el consumo de frutas y productos lácteos no muestra diferencias entre ambos grupos (Estas variables se pueden ver detalladas en la Tabla 5).

**Tabla 5: Caracterización según las variables de ajuste cualitativas de los individuos participantes en el estudio.**

Factores de riesgo	Controles		Casos	
	Media	SD	Media	SD
Edad (Años)	67	11.46	63.3	10.59
IMC (Kg/m <sup>2</sup> )	26.74	4.4	27.49	4.63
Ingesta de vegetales (g/día)	176.18	119.27	162.83	103
Ingesta de carnes rojas (g/día)	62.35	39.93	74.59	51.2
Ingesta de productos lácteos (g/día)	337.31	182.83	330.75	177.56
Ingesta de fruta (g/día)	287.13	175.6	288.98	169.78
Ingesta de alcohol (g/día)	18.11	27.43	25.26	35.15

Características de casos y controles de CCR para las variables continuas. Sólo se analizan los individuos con genotipado completo del gen PGC1 $\alpha$ .

#### 4 . Estudio de la asociación entre los SNPs del PGC1 $\alpha$ y el CCR

---

- **SNPs presentes en PGC1 $\alpha$**

En la base de datos del proyecto MCC-Spain, en  $\pm 500$ pb aguas arriba y abajo de la localización exacta del gen PGC1 $\alpha$ , hay un total de 9 polimorfismos: rs12505641, rs885395, rs148144750, rs3736265, rs58772979, rs35437002, rs34514918, rs142669571 y rs12506649. En el análisis de asociación SNP-CCR de todos los SNPs presentes en el gen PGC1 $\alpha$  (disponibles en el Exome Array del proyecto MCC-Spain), sólo el rs12505641 se asocia con el riesgo de desarrollo de este tumor de forma mínima pero estadísticamente significativa, por ese motivo, es el único detallado a continuación.

- **Asociación rs12505641-CCR**

En la Tabla 6 se muestra el resultado de la evaluación de la asociación rs12505641 y CCR correspondiente al modelo aditivo tomando como referencia el alelo G (valor de  $p=0.03$ ). Por otro lado, se realizó el análisis de asociación de este polimorfismo con diferentes factores de riesgo, como son el IMC, el nivel de actividad física, AFPG o la adherencia al patrón de dieta mediterráneo. No obstante, en ninguno de los casos se pudo observar una asociación estadísticamente significativa (Tabla 6).

**Tabla 6: Asociación entre el rs12505641 y el riesgo de padecer CCR.**

rs12505641	CCR	IMC	Alcohol	AFPG	Actividad física	Adherencia
valor p	0.03	0.97	0.12	0.31	0.77	0.71

Se muestra el valor de significación de  $p$  para la asociación entre el polimorfismo y el CCR. Además se muestra la asociación entre el SNP y otros factores de riesgo, como son el IMC, la ingesta de alcohol, los AFPG, la actividad física y la adherencia al patrón de dieta Mediterránea.

- **Asociación entre rs12505641 y la localización del tumor**

Tras conocer la asociación estadísticamente significativa entre el SNP y el cáncer de colon, se realizó un análisis estratificado por localización del tumor. Éste permitió observar que, para los datos aquí empleados, el SNP rs12505641 se comporta como un factor de riesgo en la susceptibilidad al cáncer de recto para el genotipo homocigoto AA [OR=1.7; CI=1.06-2.75; valor de  $p=0.029$ ]. Por el contrario, no se observa ninguna relación significativa con el colon. Estos resultados se encuentran recogidos en la Tabla 7.

#### 4 . Estudio de la asociación entre los SNPs del PGC1 $\alpha$ y el CCR

**Tabla 7: Resultados del análisis de regresión logística para el SNP rs12505641 y el riesgo de desarrollar CCR estratificado por la localización del tumor.**

		Odds Ratio	Std. Err.	Valor de P	IC95 %	
<b>Recto</b>	<b>AG</b>	1.14	0.13	0.23	0.92	1.41
	<b>AA</b>	1.70	0.42	<b>0.03</b>	1.06	2.75
<b>Colon</b>	<b>AG</b>	0.98	0.09	0.85	0.82	1.18
	<b>AA</b>	1.29	0.29	0.26	0.83	1.99
<b>Colon derecho</b>	<b>AG</b>	1.22	0.15	0.10	0.97	1.55
	<b>AA</b>	1.21	0.38	0.54	0.66	2.23
<b>Colon izquierdo</b>	<b>AG</b>	0.80	0.10	0.06	0.63	1.01
	<b>AA</b>	1.34	0.36	0.27	0.79	2.28

Los Odds Ratios están ajustados por edad, nivel socioeconómico, y sexo.

#### ■ Asociación rs12505641-cáncer de recto en función del sexo

A la vista de los resultados del apartado anterior, se llevó a cabo un análisis de la interacción del SNP con el riesgo de desarrollo de cáncer de recto estratificando por sexo y teniendo en cuenta posibles factores de riesgo, como son la actividad física, el consumo de tabaco o de alcohol, el IMC o el consumo de alimentos como lácteos, vegetales, frutas o carnes rojas. Además, se ajustó por edad y nivel socioeconómico.

En estos resultados se puede advertir que el alelo A en los varones está asociado al riesgo de desarrollo de tumor, no observándose el mismo efecto en el caso de las mujeres. Además, en el caso de los hombres, se puede observar un efecto protector en los niveles de estudios superiores frente a inferiores, la baja realización de actividad física frente al sedentarismo y en el elevado consumo de vegetales (frente al primer tercil, o el bajo consumo de los mismos). Por otro lado, muestran efecto de riesgo factores como el elevado consumo de alcohol, tener un peso corporal por debajo de lo normal, o el consumo elevado de carnes rojas. Por su parte, las mujeres muestran resultados bastante similares a los hombres en cuanto a factores de riesgo o protección. En el primer grupo se pueden incluir la edad, el tabaco y el consumo abusivo de carnes rojas; mientras que en el grupo de factores de protección se halla el consumo elevado de vegetales (Tabla 8).

#### 4 . Estudio de la asociación entre los SNPs del PGC1 $\alpha$ y el CCR

**Tabla 8: Análisis de regresión logística del rs12505641 y el cáncer de recto estratificado por sexo y considerando diferentes factores de riesgo.**

	Recto									
	Hombres					Mujeres				
	Odds Ratio	Std. Err.	Valor de p	IC 95 %		Odds Ratio	Std. Err.	Valor de p	IC 95 %	
<b>rs12505641</b>										
AG	1.36	0.21	<b>0.04</b>	1.01	1.84	0.89	0.19	0.60	0.59	1.36
AA	2.45	0.80	<b>0.01</b>	1.29	4.65	0.69	0.44	0.56	0.20	2.39
<b>Edad</b>	1.01	0.01	0.46	0.99	1.02	1.03	0.01	<b>0.01</b>	1.01	1.05
<b>Nivel cultural</b>										
Educación primaria	0.73	0.14	0.10	0.50	1.06	1.04	0.26	0.87	0.64	1.68
Educación secundaria	0.49	0.11	<b>&lt;0.001</b>	0.32	0.76	0.54	0.18	0.06	0.29	1.02
Universidad	0.26	0.07	<b>&lt;0.001</b>	0.15	0.44	0.52	0.20	0.09	0.24	1.10
<b>Actividad física (METS)</b>										
Baja	0.47	0.13	<b>0.01</b>	0.28	0.81	0.60	0.18	0.09	0.33	1.09
Moderada	0.76	0.19	0.27	0.47	1.24	0.63	0.21	0.16	0.33	1.20
Alta	0.77	0.12	0.10	0.56	1.05	1.02	0.23	0.94	0.65	1.58
<b>Tabaco</b>										
Fumador	1.28	0.23	0.16	0.91	1.81	1.81	0.48	<b>0.03</b>	1.08	3.05
Exfumador	1.27	0.27	0.26	0.84	1.93	0.86	0.29	0.65	0.45	1.65
<b>Alcohol (Terciles)</b>										
2.º tercil	1.02	0.19	0.93	0.70	1.47	0.00	0.00	0.99	0.00	-
3.º tercil	1.53	0.27	<b>0.02</b>	1.08	2.17	0.00	0.00	0.99	0.00	-
<b>Índice de masa corporal (IMC)</b>										
Peso bajo	5.78	4.75	<b>0.03</b>	1.15	28.92	4.15	2.44	<b>0.02</b>	1.31	13.17
Sobrepeso	0.79	0.15	0.22	0.55	1.15	1.46	0.37	0.14	0.89	2.41
Obesidad	1.19	0.23	0.36	0.82	1.74	1.40	0.36	0.19	0.85	2.30
<b>Productos lácteos (Terciles)</b>										
2.º tercil	1.17	0.20	0.36	0.83	1.64	0.96	0.23	0.86	0.61	1.52
3.º tercil	1.03	0.18	0.88	0.72	1.46	0.83	0.20	0.43	0.52	1.33

Continúa en la siguiente página

#### 4 . Estudio de la asociación entre los SNPs del PGC1 $\alpha$ y el CCR

	Recto									
	Hombres					Mujeres				
	Odds Ratio	Std. Err.	Valor de p	IC 95 %		Odds Ratio	Std. Err.	Valor de p	IC 95 %	
<b>Consumo vegetales (Terciles)</b>										
2.º tercil	0.63	0.11	<b>0.01</b>	0.45	0.89	0.90	0.20	0.64	0.58	1.40
3.º tercil	0.58	0.11	<b>&lt;0.001</b>	0.41	0.84	0.43	0.11	<b>&lt;0.001</b>	0.26	0.71
<b>Consumo fruta (Terciles)</b>										
2.º tercil	1.10	0.20	0.59	0.78	1.56	1.42	0.35	0.16	0.87	2.32
3.º tercil	0.99	0.19	0.94	0.68	1.43	1.27	0.34	0.37	0.75	2.16
<b>Consumo carnes rojas (Terciles)</b>										
2.º tercil	1.33	0.26	0.15	0.91	1.95	1.32	0.32	0.25	0.82	2.13
3.º tercil	2.61	0.48	<b>&lt;0.001</b>	1.82	3.74	1.68	0.41	<b>0.04</b>	1.04	2.72

En los análisis se ha tomado como referencia el homocigoto GG y se ha ajustado por edad y nivel socioeconómico. Los puntos de corte para la actividad física fueron definidos como  $0 METs * hora/semana$  para el sedentarismo;  $> 0$  y  $< 8 METs * hora/semana$  para la baja actividad física;  $>= 8$  y  $< 16 METs * hora/semana$  para la actividad física moderada y  $>= 16 METs * hora/semana$  para considerar una elevada actividad física (De acuerdo con el American College of Sport Medicine; METS: Unidad de metabolismo). Las variables de ingesta de alcohol, consumo de productos lácteos, vegetales, fruta y carnes rojas se dividieron en terciles.

#### 4.4. Discusión

A pesar de la conocida importancia que tiene el genoma humano completo, aún se desconoce la función de la mayoría de los locus que lo conforman [136]. La necesidad de completar el conocimiento del genoma humano y de aumentar la información de todas las regiones del genoma es lo que lleva a realizar análisis como el que ocupa este capítulo de la presente tesis.

En el análisis SNP a SNP de los polimorfismos del gen PGC1 $\alpha$  recogidos en el Exome Array del proyecto MCC-Spain (comprendiendo  $\pm 500$ pb aguas arriba y abajo de la localización exacta del gen, 4p15.2), en el que se incluyeron un total de 1336 casos de CCR y 2744 controles, se comprobó que el genotipo AA del polimorfismo rs12505641 está asociado de forma estadísticamente significativa con el riesgo de desarrollar cáncer colorrectal, en concreto con el desarrollo de cáncer de recto.

El rs12505641 se caracteriza por ser un rSNP, es decir, un polimorfismo localizado en la región no codificante del ADN [130, 137]. Concretamente se localiza en el cromosoma 4, en la posición 23487269, y está rodeado por el gen CDC42P6

#### 4 . Estudio de la asociación entre los SNPs del PGC1 $\alpha$ y el CCR

---

(cell division cycle 42 pseudogene 6) (chr4:22726422-22728023) y el PGC1 $\alpha$  (chr4:23792021-24472771).

Ante la falta de información sobre este polimorfismo en la literatura actual, existen dos factores que sugieren que el rs12505641 podría tener influencia o actuar como un tag-SNP de otro SNP del gen PGC1 $\alpha$  :

- La distancia ente el rSNP y el promotor del PGC1 $\alpha$  es inferior a 305Kb, mientras que con el promotor del otro gen, la distancia está próxima a 761Kb [134].
- El CDC42P6 se corresponde con un pseudogen [138], es decir, una región del ADN que en un inicio tuvo una función pero que, debido a la acumulación de mutaciones en su secuencia, ya no cumple ningún cometido. Normalmente, los pseudogenes presentan un gen homólogo que lleva a cabo sus funciones biológicas o, de lo contrario, estas ya no son esenciales para la vida y se puede prescindir de ellas [139]. Por el contrario, el gen PGC1 $\alpha$  codifica para una proteína encargada de regular la expresión de genes involucrados en el metabolismo energético.

Por estos motivos, asumiremos que el rSNP detectado, u otro SNP que se encuentre en desequilibrio de ligamiento con el mismo, ejerce su efecto sobre la expresión del PGC $\alpha$ , aunque se requiere de análisis complementarios que verifiquen esto.

Según los resultados de este análisis, la presencia del alelo A en este SNP supone un factor de riesgo para el desarrollo del tumor de recto. Además, el análisis estratificado por sexo muestra que esta asociación es más plausible en los varones, donde el riesgo de desarrollo del tumor para el genotipo heterocigoto se ve incrementado en 1.36 veces (IC95 % 1.01-1.8) y para el genotipo homocigoto AA en 2.45 (IC95 % 1.29-4.65) con respecto al homocigoto GG.

Como ya se ha visto con anterioridad, el PGC1 $\alpha$  es el encargado de la regulación de múltiples procesos celulares, como el metabolismo energético, la apoptosis o la generación de especies reactivas de oxígeno. Alteraciones en estos procesos constituyen una de las principales causas para la formación de los diferentes tumores [140], y como vimos en el capítulo anterior (Capítulo 3: Revisión sistemática), modificaciones en la expresión de este gen conllevan trastornos en todos esos procesos [141]. De este modo, podemos plantear una hipótesis: La presencia del alelo A en el rSNP rs10505641 influye alterando la expresión normal del gen PGC1 $\alpha$ , pudiendo producir cambios en los diversos procesos celulares y, por lo tanto, el desarrollo del tumor de recto.



#### 4 . Estudio de la asociación entre los SNPs del PGC1 $\alpha$ y el CCR

---

Por otro lado, entre los factores ambientales analizados, se puede observar que el consumo de cantidades elevadas de alcohol supone un factor de riesgo para el desarrollo del tumor, mientras que la realización de actividad física supone un factor protector. Además, los alimentos ingeridos también parecen mostrar un gran efecto sobre esta asociación. Cuanto mayor es el consumo de vegetales, mayor es la protección contra el cáncer; mientras que, en el consumo de carnes rojas y procesadas, la relación es inversa, aumentando el riesgo.

Estos resultados permiten plantear la hipótesis de que el alelo A del rs12505641 ejerce alteraciones en la expresión del PGC1 $\alpha$  que, probablemente, sean inhibitorias. Por ese motivo, la realización de ejercicio físico ejerce un efecto protector al aumentar la expresión del gen, mientras que el consumo de alcohol refuerza la actividad del rSNP al reducir también la actividad del PGC1 $\alpha$ . Sin embargo, la ausencia de estudios previos sobre este polimorfismo dificulta en gran medida comprender por qué se observa este valor significativo de asociación con el cáncer de recto. Para poder completar esta información, sería necesario realizar análisis complementarios. Entre estos análisis se pueden destacar análisis de expresión génica mediante el uso de microarrays de ARN [142]; o análisis de proteínas mediante el uso de espectrometría de masas, técnica que permite identificar la proteína y analizar sus dominios de unión, pudiendo comprobar así su funcionalidad [143].

Además hay diferentes estudios que apoyan la posibilidad de desarrollar estas vías de análisis. Por ejemplo, según lo expuesto por Souza et al. [144] el incremento de la actividad física supone un aumento de la expresión del PGC1 $\alpha$ , frenando procesos como el envejecimiento o reduciendo la obesidad. Esto aumenta la sensibilidad a la insulina, ejerciendo una mayor protección frente al desarrollo del CCR. Por otro lado, el consumo de alcohol también puede afectar a la expresión del PGC1 $\alpha$ , incrementando el efecto del rSNP. Se ha visto que, al menos en hígado, la ingesta de esta bebida interrumpe las interacciones entre proteínas nucleares (PGC1 $\alpha$ ) y mitocondriales al modificar su translocación [145].

CAPÍTULO 5

# Comparación de las técnicas de análisis de conjuntos de genes

---

### 5.1. Introducción

Como ya se indicó anteriormente (en la introducción y en el capítulo 4), durante muchos años el marco de análisis de las enfermedades complejas, como el cáncer, estuvo centrado en el análisis SNP a SNP [85, 87, 88]. Sin embargo, este análisis tiene múltiples limitaciones, por lo que su uso cada vez es menor [85] y son otro tipo de análisis los que se llevan a cabo [87], como el análisis de conjuntos de genes comentado previamente en la introducción de la presente tesis.

Como se comenta en dicha sección, en un inicio, estos análisis se implementaron con datos de expresión génica [94, 95] y, en este sentido, estas técnicas han sido comparadas y comprobadas en múltiples ocasiones. Por ejemplo, Tarca et al. [100] identificaron las técnicas PLAGE, GLOBALTEST y PADOG como las mejores técnicas por generar un bajo número de falsos positivos [100], mientras que Maciejewski et al [99] llegaron a la conclusión de que Globaltest, GSA o GSEA son capaces de generar los resultados más biológicamente interpretables [99]. Además de este tipo de datos, cada vez son más las técnicas que permiten el uso de SNPs para estos análisis. Esta idea fue propuesta por primera vez por Wang et al. [83] y Chasman [96]. Sin embargo, estas herramientas han sido evaluadas en menor medida. Sólo se han llevado a cabo trabajos que detectan diferencias metodológicas entre ellas, sus posibles problemas o sus limitaciones [82, 93], pero no se ha comparado el uso de las técnicas o se ha llegado a la conclusión de qué técnica puede ser la mejor. Además, a pesar de que hay autores que recomiendan encarecidamente el uso de múltiples métodos de análisis para obtener un resultado fiable [93, 146], no se ha comprobado aún la ventaja real que presenta el uso del consenso de los resultados o si realmente compensa llevar a cabo múltiples técnicas para obtener un único resultado.

El objetivo de este capítulo es llevar a cabo una comparación de cinco técnicas de análisis de conjuntos de genes cuya hipótesis nula sea competitiva, y usar estos resultados en entender la necesidad o no de aplicar más de una técnica a los datos para obtener unos resultados más ajustados a la realidad.

### 5.2. Material y métodos

Los datos empleados en este capítulo han sido obtenidos del Exome Array disponible en el proyecto MCC-Spain. Como ya se ha comentado con anterioridad, estos datos corresponden a diferentes tumores. Sin embargo, para este análisis, se han eliminado las etiquetas de los casos, considerándolos a todos como un solo tipo de tumor. Además, la información genética fue filtrada en base al equilibrio de

## 5 . Comparación de las técnicas de análisis de conjuntos de genes

---

Hardy-Weimberg ( $HWE < 0,001$ ) y a la frecuencia del menor alelo ( $MAF < 0,05$ ). De este modo, disponemos de un total de 6857 sujetos de estudio con información genética completa (4034 casos y 2823 controles) y de 29374 SNPs.

### 1. Conjuntos de genes empleados

Para el desarrollo de estos análisis, los conjuntos de genes empleados fueron generados aleatoriamente, bajo el control del porcentaje de SNPs significativos que debe haber en cada uno de ellos. El proceso de generación de estos conjuntos se llevó a cabo en tres pasos consecutivos, a los que hemos denominado como: i) Clasificación de SNPs, ii) Barajado de SNPs y iii) Construcción de conjuntos de genes (Ver la Figura 16).

#### (I) Clasificación de los SNPs:

Se listaron todos los SNPs disponibles en la base de datos y se calculó su valor de  $p$  mediante el test de tendencia de Cochran-Armitage. Posteriormente se obtuvieron dos listas de SNPs, una con aquellos que poseen un valor de  $p$  significativo ( $p < 0,05$ ) y otra con los no significativos ( $p > 0,05$ ).

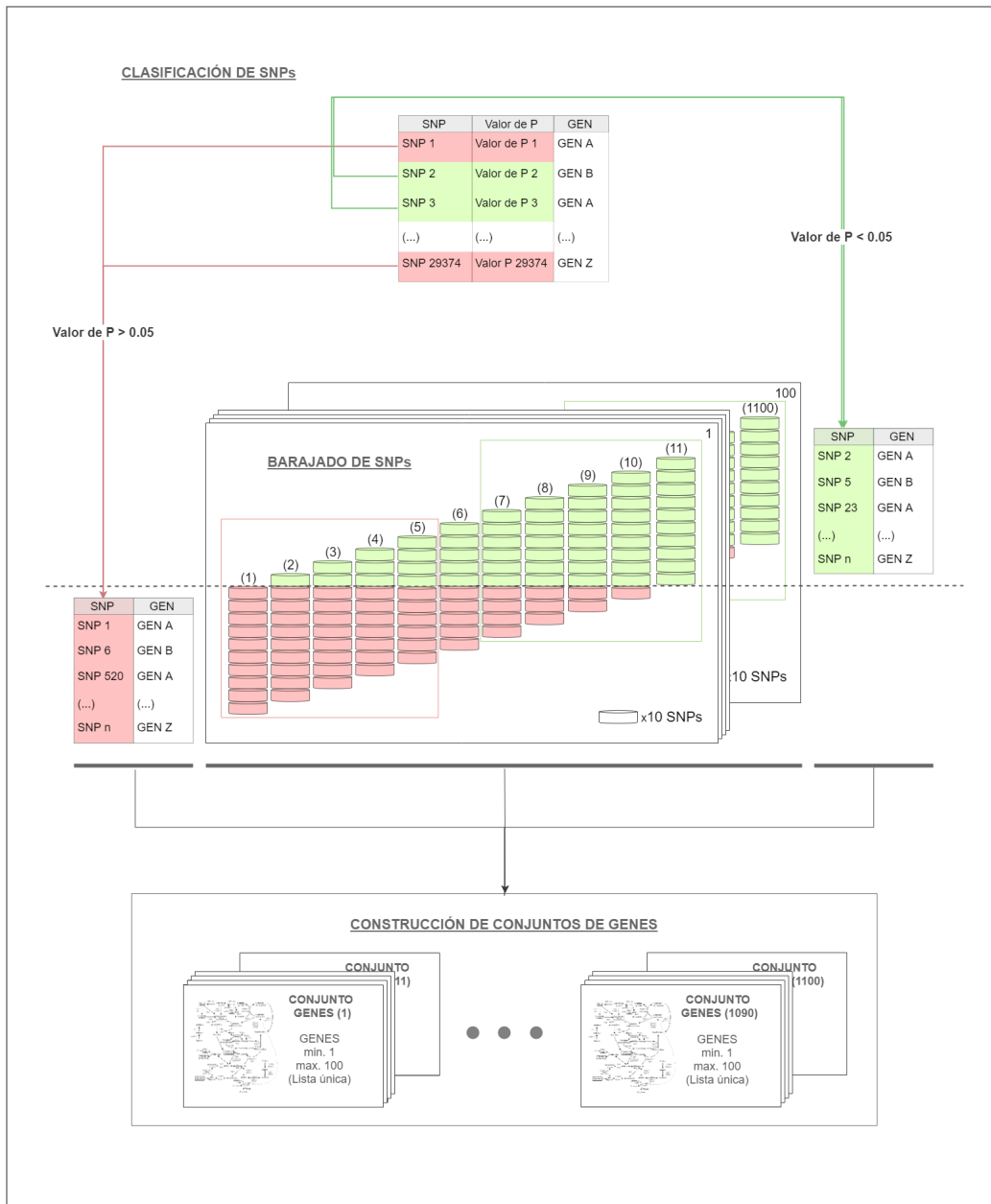
#### (II) Barajado de SNPs:

Conocida la clasificación de los SNPs, se llevó a cabo la mezcla de los mismos de forma aleatoria, pero teniendo en cuenta el porcentaje de polimorfismos que debían tomarse de cada una de las listas. De este modo, se obtuvieron 11 lotes de SNPs que van desde un 0% a un 100% de SNPs significativos aumentando de 10% en 10%. A su vez, esta selección aleatoria fue realizada 100 veces por cada uno de los lotes de SNPs, obteniendo un total de 1100 lotes diferentes.

#### (III) Construcción de conjuntos de genes:

De cada uno de los lotes se obtuvo un conjunto de genes comprobando a qué gen pertenece cada SNP. De este modo, cada conjunto de genes puede tener un máximo de 100 genes si cada SNP pertenece a un único gen y un mínimo de un único gen si los 100 SNP se localizasen en el mismo.

## 5 . Comparación de las técnicas de análisis de conjuntos de genes



**Figura 16: Esquema del flujo de trabajo para obtener los conjuntos de genes aleatorios.** Esta figura muestra los tres pasos realizados en el desarrollo de los conjuntos de genes: “Clasificación de SNPs” donde se dividen los SNPs en dos listas, significativos y no significativos; “Barajado de SNPs” en el que se mezclan ambas listas en diferentes proporciones; y “Construcción de conjuntos de genes” para obtener los conjuntos de genes.

## 5 . Comparación de las técnicas de análisis de conjuntos de genes

---

### 2. Técnicas empleadas

Las técnicas empleadas en este trabajo fueron cinco: GSEA-SNP, GenGen, SRT, iGSEA4GWAS y GSA-SNP. Su selección fue basada en los métodos expuestos en el trabajo de Wang et al [93] y que prueban una hipótesis competitiva. Estas técnicas están detalladas en la sección “Análisis de conjuntos de genes” de la introducción (Ver Tabla 1).

En cuanto a los ajustes empleados en cada técnica: se hicieron 1000 permutaciones en todas ellas y se seleccionó un valor de significación de  $p < 0,05$  en todos los análisis. El resto de ajustes considerados fueron los que presenta “por defecto” cada una de las técnicas, a excepción del GSA-SNP, donde es necesario seleccionar el método de análisis de ES. Además, en esta técnica se probó también el uso de los cinco mejores SNPs por gen.

### 3. Comparación de los resultados de las técnicas

Para contrastar los resultados de las diferentes técnicas se dispuso de dos diagramas de líneas, un diagrama de Venn y una comparación entre las ratios de falsos positivos y falsos negativos de cada técnica. El primer diagrama de líneas se empleó para valorar la capacidad que tiene cada una de las técnicas para discriminar los conjuntos de genes en función del porcentaje de SNPs significativos que estos presentan. El segundo diagrama de líneas se utilizó para conocer cómo influyen en dicha capacidad, la combinación de resultados de las técnicas. El diagrama de Venn permitió conocer el número de conjuntos de genes concretos en los que están de acuerdo cada una de las técnicas. Por último, la comparación entre los falsos positivos y falsos negativos de cada técnica fue representada por un diagrama de barras y se aplicó al conocimiento de qué técnicas generan un mayor número de falsos positivos o falsos negativos.

Ante la ausencia de un estándar de oro (*gold standard*) con el que calcular los ratios de falsos positivos y falsos negativos entre las técnicas de análisis, se acudió al método de Brown [147] para determinar la significación de los genes. Este método es una extensión del método de Fisher, cuya diferencia radica en que éste es capaz de combinar resultados de diferentes pruebas estadísticas potencialmente dependientes, siendo así una técnica apropiada para grandes conjuntos de datos y empleada en bases de datos biológicos [147]. Con el uso de este método, la definición de conjunto de genes significativo se relativizó al número de genes significativos que debería poseer cada conjunto para ser considerado como tal. Así se definieron diez puntos de corte calculados desde cero hasta el porcentaje máximo de genes significativos en nuestros datos por cada conjunto de genes definido (18.7%). Por cada uno de los puntos de corte

se compararon los resultados del estándar con los de cada una de las técnicas, considerando falsos positivos aquellos resultados positivos para la técnica, pero negativos para el estándar y falsos negativos a los contrarios.

### 5.3. Resultados

#### 1. Conjuntos de genes detectados como significativos

El número de conjuntos de genes detectados como significativos es diferente para cada una de las técnicas empleadas (Tabla 9). Mientras que el GSEA-SNP detecta un total de 721 conjuntos de genes significativos tras realizar la corrección por FDR, la técnica del iGSEA4GWAS detecta un total de 867. Por otro lado, el SRT, que no es capaz de corregir por FDR, detecta un total de 918 conjuntos de los 1100 analizados.

**Tabla 9: Número de conjuntos de genes considerados significativos para cada técnica.**

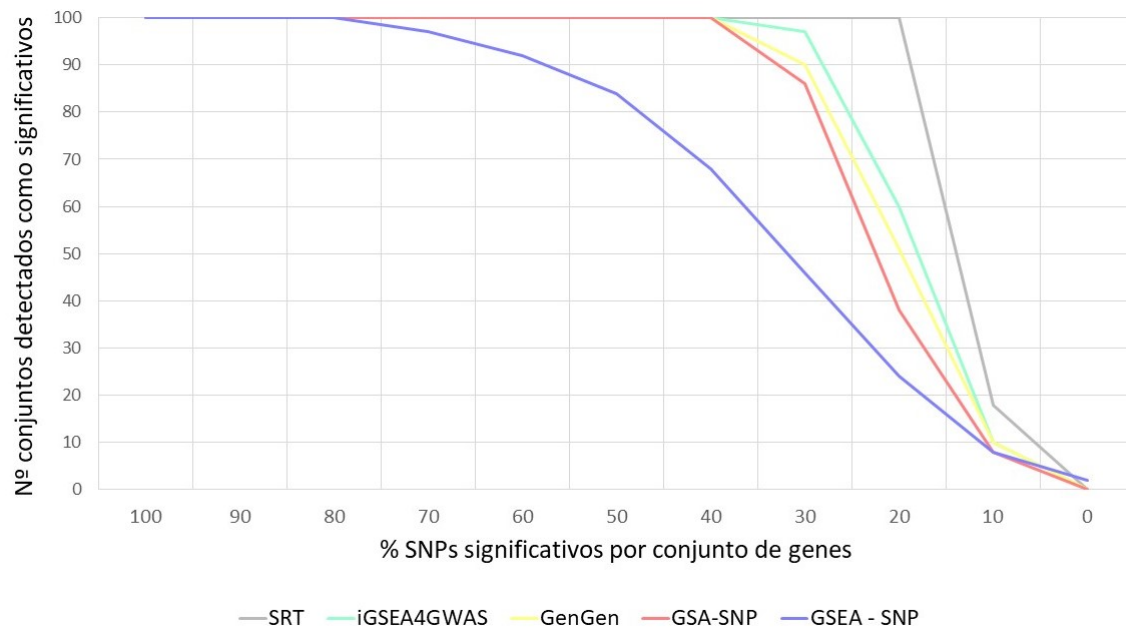
	Valor p < 0.05	FDR < 0.05
<b>GSEA-SNP</b>	791	721
<b>GSA-SNP</b>	859	832
<b>GenGen</b>	865	851
<b>iGSEA4GWAS</b>	870	867
<b>SRT</b>	918	-

El nivel de significación se constató para un valor de p o para un FDR inferior a 0.05.

#### 2. Capacidad discriminativa en la selección de conjuntos de genes

La Figura 17 muestra la capacidad que presenta cada una de las técnicas de descartar o admitir conjuntos de genes asociados con la enfermedad en función del porcentaje de SNPs significativos que presenta cada uno de ellos. Como se muestra en dicha figura, la técnica del SRT necesita un 80 % de SNPs no significativos en el conjunto para comenzar a descartarlos como no asociados al fenotipo de estudio. Por el contrario, el GSEA-SNP es la técnica que mejor se ajusta al porcentaje de SNPs significativos de cada conjunto.

## 5 . Comparación de las técnicas de análisis de conjuntos de genes



**Figura 17: Capacidad de discriminación de cada técnica.** Esta figura muestra el número de conjuntos de genes significativos (eje de ordenadas) que detecta cada una de las técnicas a nivel de FDR (excepto el SRT, valor de p) en función del porcentaje de SNPs significativos de cada uno de ellos (eje de abscisas). Los ajustes de cada una de las técnicas fueron los que presentan “por defecto”.

### 3. Incremento del número de SNPs que representa cada gen

Tras los resultados de la sección anterior, en el que se observó que la técnica que toma un mayor número de SNPs por gen es la que mejor se ajusta a la realidad de los datos, se hizo una comparación entre técnicas incrementando el número de SNPs que representan cada gen. Como sólo el GSA-SNP permite modificar esta opción, fue la única técnica en la que se valoró. La Tabla 10 muestra las diferencias entre la capacidad de discriminación que presenta el GSA-SNP cuando se seleccionan dos SNPs por gen (GSA-SNP(2)) y cuando se seleccionan cinco (GSA-SNP(5)).

El GSA-SNP(5) es capaz de detectar conjuntos de genes no significativos con un 60 % de SNPs significativos, mientras que el GSA-SNP(2) requiere, como ya se vio con anterioridad, de un total de 70 % de SNPs no significativos para poder detectar el conjunto como tal. La Figura 18 refleja la capacidad de adaptación a los datos que presentan estas técnicas, y permite su comparación con el resto, observando que GSA-SNP(5) se aproxima ahora más al GSEA-SNP.

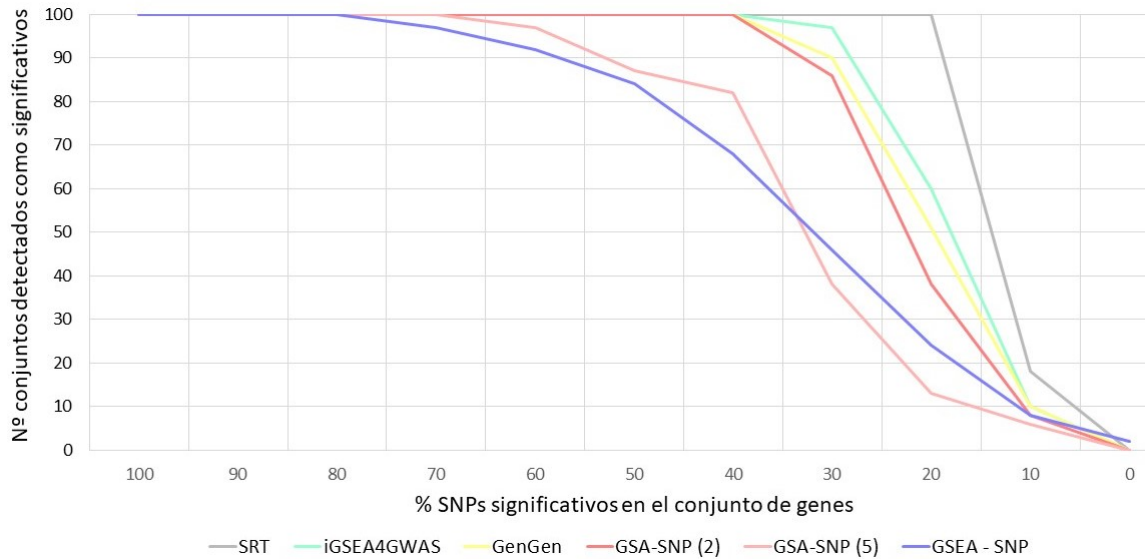


## 5 . Comparación de las técnicas de análisis de conjuntos de genes

**Tabla 10: Comparación entre GSA-SNP considerando dos SNPs por gen (GSA-SNP(2)) y cinco SNPs por gen (GSA-SNP(5)).**

	GSA-SNP (2)	GSA-SNP (5)
<b>100</b>	100	100
<b>90</b>	100	100
<b>80</b>	100	100
<b>70</b>	100	100
<b>60</b>	<b>100</b>	<b>97</b>
<b>50</b>	100	87
<b>40</b>	100	82
<b>30</b>	86	38
<b>20</b>	38	13
<b>10</b>	8	6
<b>0</b>	0	0

Esta tabla recoge el número de conjuntos de genes que detecta la técnica cuando se seleccionan los dos o cinco mejores SNPs por gen. En negrita se marca el porcentaje de SNPs significativos a partir del cuál ambas técnicas muestran diferencias.

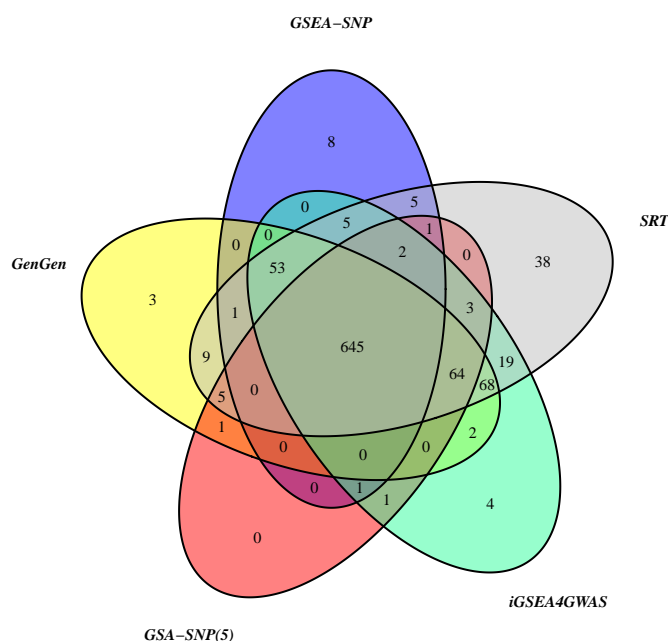


**Figura 18: Capacidad de discriminación de la selección de conjuntos de genes, considerando GSA-SNP(5).** Número de conjuntos de genes significativos (eje de ordenadas) que detecta cada una de las técnicas a nivel de FDR (excepto el SRT, valor de p) en función del porcentaje de SNPs significativos de cada conjunto (eje de abscisas). Los ajustes de cada una de las técnicas fueron los de “por defecto”, a excepción del GSA-SNP, en la que se añadió 5SNPs/gen (GSA-SNP(5)).

## 5 . Comparación de las técnicas de análisis de conjuntos de genes

### 4. Resultados de la combinación de técnicas

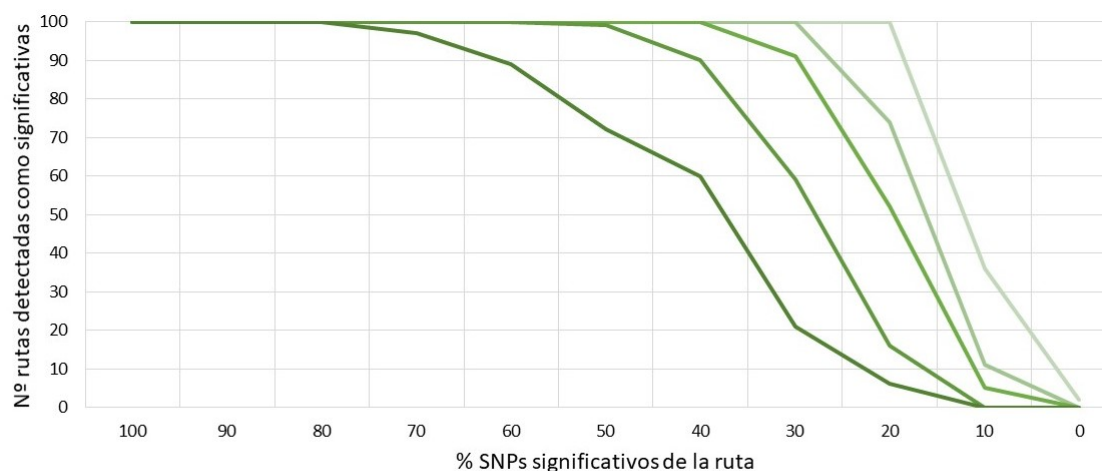
El diagrama de Venn (Figura 19) muestra el número de conjuntos de genes seleccionados en común por la combinación de técnicas, teniendo en cuenta el GSA-SNP(5). Un total de 645 conjuntos son seleccionados por todas las técnicas, mientras que el SRT detecta 38 que ninguna otra es capaz de detectar. Además todas las técnicas detectan 64 conjuntos que tan solo el GSEA-SNP descarta.



**Figura 19: Diagrama de Venn: Conjuntos de genes significativos en la combinación de las técnicas.** Representa el número de conjuntos de genes que convergen en los resultados de cada una de las técnicas. Se representan los conjuntos seleccionados por cada técnica a nivel de  $FDR < 0,05$ , a excepción del SRT que sólo considera valor de  $p < 0,05$ .

Apoyando estos resultados, representamos la capacidad de discriminación que muestra la combinación de los resultados de las técnicas (Figura 20). En este caso, una vez obtenidos los resultados de las cinco técnicas, se representan el número de conjuntos de genes seleccionados si se consideran como significativos todos los que hayan sido detectados por al menos una de las técnicas, por al menos dos, tres y así hasta considerar sólo como significativos aquellos que hayan sido detectados por las cinco técnicas. De este modo, se puede observar cómo, al aumentar el número de técnicas necesarias para determinar el resultado como significativo, aumenta también el ajuste al porcentaje real de SNPs significativos en la ruta metabólica.

## 5 . Comparación de las técnicas de análisis de conjuntos de genes



Rutas significativas detectadas considerando los resultados de:

— 1 Técnica    — 2 Técnicas    — 3 Técnicas    — 4 Técnicas    — 5 Técnicas

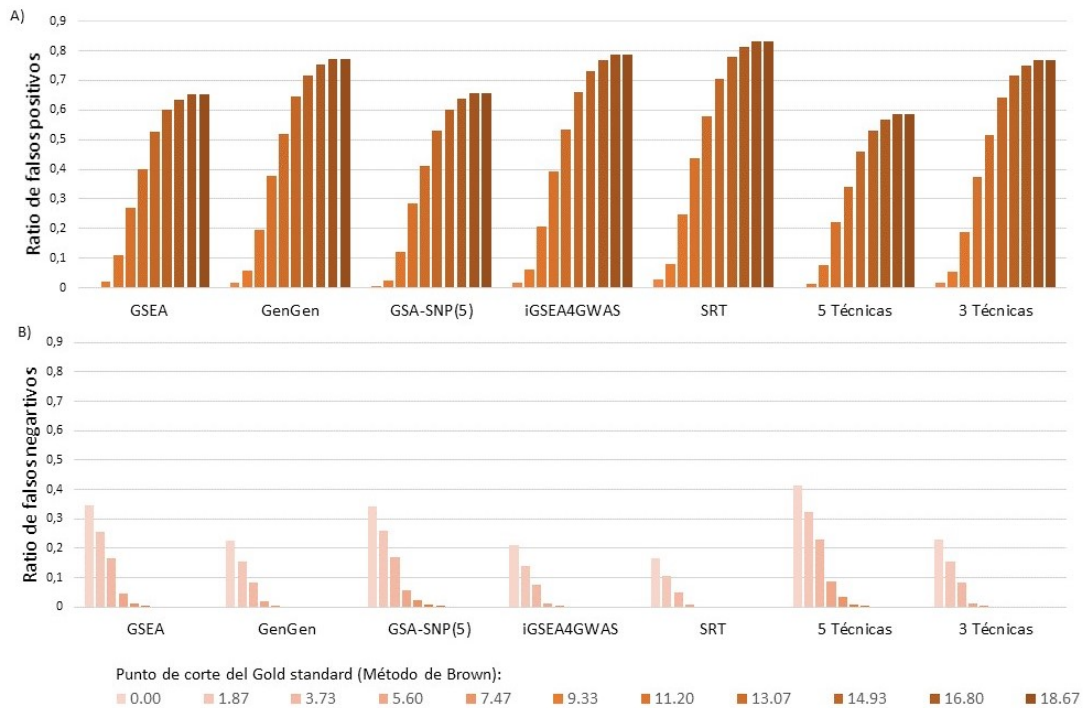
**Figura 20: Influencia de la combinación de las técnicas en la capacidad de discriminación de los conjuntos de genes.** (Considerando GSA-SNP (5)).

Representa el número de conjuntos de genes significativos (eje de ordenadas) que detecta la combinación de los resultados de las técnicas en función del porcentaje de SNPs significativos de cada conjunto (eje de abscisas).

### 5. Ratios de falsos positivos y falsos negativos

Las ratios de falsos positivos y falsos negativos que muestra cada una de las técnicas permiten generar una idea de su proximidad con la realidad (Figura 21). En ausencia de un estándar de oro determinado, se empleó el método de Brown para definir los puntos de corte para los que un conjunto de genes sería detectado como significativo. Obviamente, tal como muestra la Figura 21, cuando se considera significativo un conjunto que no posee ningún gen significativo, no pueden existir falsos positivos (Figura 21-A). Sin embargo, en esta situación, las técnicas sí muestran falsos negativos, siendo más abundantes en los métodos que seleccionan un mayor número de SNPs por gen, tales como el GSEA-SNP o el uso de los resultados consenso de 5 técnicas (Figura 21-B). Por el contrario, a medida que aumenta el número de genes significativos necesarios en el conjunto para considerar este como tal, la combinación de los resultados de las cinco técnicas es la que menos falsos positivos presenta, mientras que técnicas como el SRT o la combinación de los resultados de al menos tres técnicas (3-TÉCNICAS) muestran un gran número de falsos positivos.

## 5 . Comparación de las técnicas de análisis de conjuntos de genes



**Figura 21: Ratios de falsos positivos (A) y falsos negativos (B) para cada una de las técnicas.** Los puntos de corte son definidos como el porcentaje de genes necesario para considerar un conjunto como significativo. La definición de gen significativo ha sido obtenida en base al método de Brown [147].

### 5.4. Discusión

Los SNPs son un tipo de información válida para el análisis de conjuntos de genes, y estos análisis son muy útiles para dar una interpretación biológica a los resultados de los GWAS [148]. Wang et al. [83] fueron unos de los primeros autores en proponer el uso de SNPs para el análisis de conjuntos de genes y, junto con otros autores [82, 95, 98], fueron también unos de los primeros en hacer una revisión de las técnicas disponibles en este campo de análisis [93]. En esta revisión, los autores hablan de dos tipos de métodos en base al planteamiento de sus hipótesis nulas, diferenciando así entre las competitivas y las auto-contenidas. Además, expresan que, ante la falta de acuerdo entre el desarrollo de las técnicas y la ausencia de demostración de autoría por parte de una de ellas, la mejor opción está en el consenso del resultado de diferentes técnicas [93]. En base a esta observación, el objetivo de este capítulo es realizar una comparación de aquellas técnicas que plantean una hipótesis competitiva y que están recogidas en la revisión de Wang et al [93], conocer su capacidad de adaptación a los datos y valorar la necesidad de aplicar el consenso de los resultados entre técnicas.

## 5 . Comparación de las técnicas de análisis de conjuntos de genes

---

A este fin, y empleando el Exome Array disponible en el proyecto MCC-Spain con un total de 6857 individuos, se analizó un total de 1100 conjuntos de genes mediante la aplicación de cinco técnicas: GSEA-SNP, GSA-SNP, GenGen, iGSEA4GWAS y SRT. A partir de estos resultados, hemos sido capaces de observar grandes diferencias entre ellas. La más evidente es que el GSEA-SNP es la técnica con mayor capacidad de discriminación entre conjuntos de genes significativos y no significativos en base al porcentaje de SNPs significativos que estos muestran, definiendo como SNP significativo a aquellos que presentan una asociación significativa con el fenotipo de estudio. Mientras que, en el lado opuesto, el SRT muestra una capacidad de discriminación mucho menor. Esta disparidad podría ser debida a las diferencias en el desarrollo de las técnicas (Tabla 1), entre las que se puede destacar el hecho de que el GSEA-SNP recoge todos los SNPs disponibles de cada uno de los genes de la ruta metabólica [97], mientras que el SRT sólo representa en sus análisis aquellos genes que posean SNPs con un valor de  $p$  superior a un umbral dado [101]. La consecuencia es que el SRT selecciona un total de 197 conjuntos más que el GSEA-SNP.

La Figura 17 muestra la capacidad de discriminación de los conjuntos de genes no significativos de cada una de las técnicas. Teniendo en mente que los conjuntos fueron simulados considerando el porcentaje de polimorfismos significativos existentes en cada uno de ellos, es fácil observar que a medida que desciende este porcentaje, menor debería ser el número de conjuntos detectados como significativos por cada una de las técnicas. De nuevo, la técnica que mejor se adapta a esta situación es el GSEA-SNP, por lo que puede ser considerada como la mejor técnica entre las aquí comparadas.

Teniendo en cuenta lo anterior y puesto que la principal diferencia entre el GSEA-SNP y el resto es que ésta es la única técnica que selecciona un gran número de SNPs por cada uno de los genes del conjunto a analizar, parece plausible pensar que esta característica sea la que le permita mostrar una mejor adaptación a los datos. Por este motivo, decidimos incrementar el número de SNPs por gen seleccionados en la única herramienta que permite llevar a cabo esta acción, el GSA-SNP. El resultado de este incremento se observa en la Figura 18, donde se puede concluir que, efectivamente, el número de SNPs seleccionados por cada gen influye en los resultados, siendo mejor representar los genes por un mayor número de SNPs.

Sin embargo, ante esta afirmación, hay diversos autores con opiniones a favor y en contra. Por ejemplo, Nam et al. [102] o Kwon et al. [148] afirman que utilizar un menor número de SNPs permite reducir el total de falsos positivos detectados por la técnica sin incrementar la aparición de falsos negativos. A pesar de esto, también consideran que el uso de un mayor número de SNPs por gen ciertamente

## 5 . Comparación de las técnicas de análisis de conjuntos de genes

---

puede dar una distribución simétrica a los datos, disminuyendo la inflación de los genes considerados como significativos [102, 148]. Por su lado, Mooney et al. [146] concluyen que considerar todos los SNPs del gen puede añadir ruido al análisis, pero que descartar la gran mayoría de éstos puede suponer la pérdida de una gran cantidad de información fundamental. En base a estas reflexiones, se calculó la ratio de falsos positivos y falsos negativos que genera cada una de las técnicas (Figura 21). Los resultados observados en esta prueba ciertamente están de acuerdo con la afirmación de Kwon et al. [148], ya que las técnicas que seleccionan un mayor número de SNPs por gen muestran más falsos negativos, principalmente cuando se consideran como significativos aquellos conjuntos que poseen un bajo porcentaje de genes significativos (usando como estándar de oro el método de Brown [147]; Figura 21-B). Con estos resultados podemos determinar que GSEA-SNP es la técnica que más falsos negativos obtiene para puntos de corte muy bajos (Figura 21-B), pero a la vez, es la que menos falsos positivos genera a medida que se va aumentando dicho punto de corte (Figura 21-A). Por el contrario, el SRT a penas detecta falsos negativos, pero la proporción de falsos positivos que muestra es muy elevada.

Por lo tanto, estos resultados permiten concluir que, a pesar de que el GSEA-SNP es la técnica más conservadora en sus resultados, pues es la que muestra un mayor número de falsos negativos, también es el método que mejor se adapta al porcentaje de SNPs significativos de cada conjunto de genes. Mientras que, en el extremo opuesto se encuentra el SRT, el cuál presenta una muy baja capacidad de adaptación a esta situación y, además, exhibe un gran número de falsos positivos en sus resultados.

El segundo objetivo de este capítulo es valorar la necesidad de obtener un consenso entre el resultado de múltiples técnicas para lograr las mejores conclusiones. Esta postura ha sido defendida por Wang et al. [93] y por Mooney et al. [146] y ya había sido propuesta con anterioridad para las técnicas que emplean datos de expresión génica [149]. Sin embargo, aún no se ha comprobado realmente si el esfuerzo de emplear todas las técnicas compensa con los resultados ofrecidos.

En el diagrama de Venn (Figura 19) se pueden observar el consenso que muestran las diferentes técnicas. Cuando se tiene en cuenta los resultados de las cinco técnicas, el número total de conjuntos de genes considerados como significativos es de 645 de los 1100 totales generados. Del mismo modo, la capacidad de discriminación que muestran los resultados de las cinco técnicas en conjunto es mayor que la que exhiben los resultados de tener en cuenta al menos tres de las cinco técnicas, por ejemplo (Figura 20). Por último, al valorar la razón de falsos positivos y falsos negativos de los resultados en conjunto de las técnicas, podemos observar que el consenso de las cinco técnicas muestra una tendencia similar al GSEA-SNP, pero

## 5 . Comparación de las técnicas de análisis de conjuntos de genes

---

más definida (Figura 21). Es decir, cuando se considera como conjunto de genes significativo a aquel con un porcentaje muy bajo de genes significativos, los falsos negativos son muy elevados (Figura 21-B). Sin embargo, a medida que aumenta el punto de corte, la corrección de falsos positivos es mucho mayor (Figura 21-A). Por otro lado, la selección del consenso del resultado de tres de las cinco técnicas muestra menor número de falsos negativos, pero mayor número de falsos positivos por cada punto de corte (Figura 21).

Estos resultados están de acuerdo con la hipótesis de Wang et al. [93] y de Mooney et al. [146]. El consenso entre los resultados de diferentes técnicas permite obtener información más verídica que la aplicación de cada una de las técnicas individualmente, siendo más conservadores al incluir un mayor número de técnicas en el criterio de selección de conjuntos de genes significativos. Por otro lado, es aconsejable considerar siempre un número impar de técnicas en la selección de los mismos, para evitar tener un equilibrio entre los resultados.

CAPÍTULO 6

# **Análisis de rutas metabólicas en asociación con el CCR**

---



### 6.1. Introducción

Como ya se ha visto anteriormente en la presente tesis, entre los múltiples factores posiblemente asociados al desarrollo del CCR, se halla el gen *PGC1 $\alpha$*  [141]. Sin embargo, en el análisis SNP a SNP llevado a cabo en el capítulo 4 no se observó asociación directa entre los SNPs de este gen y el desarrollo del tumor de interés, hallándose sólo asociación entre el SNP intergénico rs12505641 y el cáncer de recto y en varones [76]. Esto puede ser debido a que el cáncer es una enfermedad compleja y este tipo de análisis presenta ciertas limitaciones en el estudio de dichas enfermedades [85]. Por ese motivo, y con el objetivo de cubrir dichas limitaciones, se diseñaron las técnicas expuestas en el capítulo 5, que se aplicarán ahora sobre los casos y los controles de CCR presentes en el Exome Array disponible en el proyecto MCC-Spain.

En base a los resultados de dicho capítulo, se aplicaron las cinco técnicas valoradas (GSEA-SNP, GSA-SNP, GenGen, iGSEA4GWAS y SRT) con el objetivo de considerar como resultado significativo aquél que mantenga consenso en al menos tres de ellas. No se consideraron las cinco técnicas en este caso debido a que, como ya se concluyó en el capítulo 5, la información obtenida sería muy conservadora y, puesto que el número de rutas analizadas es muy bajo, podría haber riesgo de pérdida de información.

El objetivo de este capítulo es conocer la posible asociación existente entre el CCR y las alteraciones existentes en las rutas metabólicas en las que participa el *PGC1 $\alpha$* .

### 6.2. Material y métodos

#### 1. Selección de SNPs

El Exome Array empleado en este estudio fue el recogido en el proyecto MCC-Spain. Para las técnicas empleadas en este análisis, no es necesaria una selección de SNPs previa ya que todas ellas requieren de toda la información disponible en el Exome Array completo. Sin embargo, para mejorar la calidad de los datos y evitar ruido en los análisis, los SNPs fueron filtrados por la frecuencia del alelo menor ( $MAF < 0,05$ ) y por el equilibrio de Hardy-Weimberg en controles ( $HWE < 0,001$ ). De este modo, el número total de SNPs disponibles es de 29473.

### 2. Rutas metabólicas para analizar

Las rutas metabólicas empleadas en el análisis fueron obtenidas de la base de datos del KEGG (<http://www.genome.jp/kegg/>) y seleccionadas en base a sus genes participantes. El requisito para considerar la ruta metabólica en el análisis es que el PGC1 $\alpha$  tenga una función importante en la misma. Las rutas metabólicas seleccionadas se muestran en la Tabla 11.

**Tabla 11: Pathways en los que participa PGC1 $\alpha$ .**

Identifer	Pathway
hsa04152	AMPK signaling pathway
hsa04211	Longevity regulating pathway
hsa04371	Apelin signaling pathway
hsa04910	Insulin signaling pathway
hsa04920	Adipocytokine signaling pathway
hsa04922	Glucagon signaling pathway
hsa04931	Insulin resistance
hsa05016	Huntington's disease
hsa03029	Mitochondrial Biogenesis

Rutas recogidas en la base de datos del KEGG en las que participa el gen PGC1 $\alpha$ .

### 3. Técnicas de análisis

En base a los resultados del capítulo anterior, se desarrolló el GSEA-SNP, GenGen, SRT, iGSEA4GWAS y GSA-SNP. En todas ellas se aplicaron 1000 permutaciones para el cálculo del nivel de significación del ES y se estimó un umbral para el valor de  $p$  menor a 0.05. Para el resto de variables de ajuste que cada técnica permite modificar se seleccionaron aquellas que vienen por defecto, a excepción del GSA-SNP donde se seleccionó el análisis por GSEA y ES y se consideraron los cinco SNPs más significativos por cada uno de los genes. Por último, para determinar si una ruta metabólica está asociada con la enfermedad, se consideró el consenso en los resultados de al menos tres de las cinco técnicas, admitiendo como significativas aquellas que muestren un  $FDR < 0,05$  (valor de  $p < 0,05$  para el SRT, puesto que no hace corrección por múltiples hipótesis).

## 6.3. Resultados

En el análisis de rutas metabólicas se contó con el total de 4093 sujetos, entre casos y controles de CCR, presentes en el estudio multicéntrico MCC-Spain. En las siguientes secciones se describen los resultados de cada una de las técnicas empleadas en el análisis y la Tabla 12 resume dichos resultados.

### 1. Presentación de resultados por cada técnica

#### ■ GSEA-SNP

Para esta técnica, las 9 rutas metabólicas analizadas adquieren un valor de  $ES$  positivo, aunque ninguna de ellas muestra un valor de FDR significativo. Sin embargo, cabe destacar que la ruta metabólica de la enfermedad de Huntington se sitúa en la parte más alta de la lista, con un  $ES = 0,42$  mientras que en la parte más baja de la misma se encuentra la ruta metabólica de resistencia a la insulina ( $ES = 0,3$ ). Dado que todas las rutas presentan un FDR cercano a 1, la asociación de las mismas con el fenotipo de enfermedad no puede ser demostrada.

#### ■ SRT

Los resultados devueltos por esta técnica no muestran significación a nivel de valor de  $p$  para ninguna de las rutas analizadas. Por lo tanto, esta técnica tampoco muestra asociación entre las rutas analizadas y el riesgo de desarrollar CCR. Cabe destacar que SRT no devuelve valores de FDR ni de  $ES$ .

#### ■ i-GSEA4GWAS V2

En el análisis realizado por la técnica de iGSEA4GWAs ninguna de las rutas metabólicas muestra una asociación significativa con la enfermedad de estudio a nivel de FDR. Esta técnica no devuelve valores de  $ES$ .

#### ■ GenGen

El estudio realizado mediante GenGen resulta en nueve valores de  $ES$  positivos. En este caso, la ruta de señalización de las Adipocitoquinas se coloca en el primer puesto de la lista con un  $ES = 0,65$ , mientras que la ruta metabólica de la Enfermedad de Huntington está en segundo lugar ( $ES = 0,59$ ). Por otro lado, en el fondo de la lista está la ruta de señalización del glucagón ( $ES = 0,42$ ). A pesar de esto, ninguna de las rutas metabólicas supera el nivel de significación del FDR, por lo que la asociación con la enfermedad no ha podido ser demostrada.

#### ■ GSA-SNP

Esta técnica es la única capaz de mostrar un valor de FDR significativo para la ruta de la enfermedad de Huntington, la cual presenta un  $ES = 0,38$  en este caso, mostrando una asociación significativa entre las alteraciones en esta ruta metabólica y el riesgo de presentar el fenotipo de la enfermedad.

## 6 . Análisis de rutas metabólicas en asociación con el CCR

**Tabla 12: Resumen de los resultados de todos los métodos empleados para el consenso.**

PATHWAYS		SRT	GSEA-SNP	i-GSEA4GWAS	GenGen	GSA-SNP
Enfermedad de Huntington	ES	-	0.42	-	0.59	0.38
	FDR	0.67*	1	1	0.41	<0.05
Ruta de señalización de la Apelina	ES	-	0.38	-	0.49	-
	FDR	0.19*	1	0.82	1	-
Biogénesis Mitocondrial	ES	-	0.4	-	0.44	-
	FDR	0.62*	1	1	1	-
Ruta de señalización de la AMPK	ES	-	0.34	-	0.45	-
	FDR	0.73*	1	0.81	0.95	-
Ruta de señalización del Glucagón	ES	-	0.34	-	0.42	-
	FDR	0.42*	1	1	1	-
Ruta de señalización de la Insulina	ES	-	0.33	-	0.5	-
	FDR	0.97*	1	0.91	0.94	-
Ruta de regulación de la longevidad	ES	-	0.32	-	0.51	-
	FDR	0.83*	1	1	0.99	-
Ruta de señalización de la Adipocytokina	ES	-	0.3	-	0.65	-
	FDR	0.85*	1	1	0.49	-
Ruta de resistencia a la Insulina	ES	-	0.3	-	0.56	-
	FDR	1*	0.97	1	1	-

ES = Enrichment Score, FDR = False Discovery Rate

Resultados de cada una de las técnicas empleadas en el análisis. El GSA-SNP sólo devuelve resultados de aquellos conjuntos de genes con resultados significativos.

\*Se representa el valor de FDR para todas a excepción del SRT, en el que se muestra el valor de p.

### 2. Representación de las rutas en los datos del Exome Array

La representación de las rutas metabólicas en la base de datos del proyecto MCC-Spain, según la herramienta del GSEA-SNP que es la única que aporta esta información, varía entre un 48 % para rutas como la de la señalización de la adipocitoquina, un 38 % para la ruta de la enfermedad de Huntington o un 13 % para rutas como la de la biogénesis mitocondrial, presentando ésta el porcentaje más bajo de genes en el Exome Array de todas las rutas analizadas.

### 3. Consenso de los resultados

Cuatro de las técnicas empleadas indican que no existen asociaciones estadísticamente significativas que permitan rechazar la hipótesis nula planteada. Solo una de las técnicas muestra una posible asociación entre el CCR y la ruta de la enfermedad de Huntington (GSA-SNP), tal como se puede observar en la Tabla 12. Debido a esto, y en base al consenso de los resultados de las técnicas, no podemos determinar que ninguna de las rutas metabólicas analizadas muestre una relación con el fenotipo del cáncer colorrectal, al menos para el Exome Array empleado en este análisis. Sin embargo, cabe destacar que las técnicas que devuelven valores de ES sitúan a la ruta de la enfermedad de Huntington en la parte alta de la lista.

## 6.4. Discusión

Las herramientas de análisis de conjuntos de genes son métodos que están en desarrollo y en continuo crecimiento en la actualidad [150], ya que son enfoques interesantes y prometedores en el análisis de los datos de un GWAS [82]. El objetivo principal de este capítulo fue encontrar la posible asociación existente entre la expresión del fenotipo de CCR y las rutas metabólicas en las que PGC1 $\alpha$  tiene una función según la base de datos del KEGG.

Como se ha visto a lo largo de la presente tesis, existen múltiples evidencias para considerar que rutas como la biogénesis mitocondrial o la ruta de señalización de la AMPK pueden estar asociadas con el riesgo de desarrollar CCR [141]. En esta línea, rutas que pueden activar a las anteriores u otras rutas también relacionadas con el metabolismo energético, como son la vía de la longevidad en el primer caso o las del glucagón o insulina en el segundo, también son susceptibles de presentar cierta asociación con el riesgo de padecer dicho tumor [141]. Por otro lado, en el caso de la ruta de señalización de la enfermedad de Huntington, a pesar de que no tiene una interpretación biológica tan pausable como las anteriores, su asociación con la enfermedad podría deberse a que comparte diversos genes con la cadena de fosforilación oxidativa, la cual está asociada con el metabolismo mitocondrial y el crecimiento tumoral, conocido también como efecto Warburg [11, 151]. Todas estas rutas muestran un aspecto común, y es que el PGC1 $\alpha$  presenta un papel importante en las mismas (Ver la sección “Rutas metabólicas en las que participa PGC1 $\alpha$ ” y el capítulo 3).

Los resultados observados de los análisis desarrollados sobre el Exome Array del proyecto MCC-Spain fueron coherentes en todas las técnicas empleadas. De hecho, de las cinco técnicas, cuatro no muestran ningún valor de significación para ninguna

## 6 . Análisis de rutas metabólicas en asociación con el CCR

---

de las rutas analizadas y sólo el GSA-SNP devuelve la ruta de la enfermedad de Huntington como el único conjunto de genes estadísticamente asociado con el riesgo de desarrollar CCR. Sin embargo, se debe resaltar que tanto el GSEA-SNP como el GenGen (las dos únicas herramientas que devuelven un valor de ES) muestran para la ruta metabólica de la enfermedad de Huntington un valor de ES elevado, aunque no sea estadísticamente significativo.

Autores que llevan a cabo estudios similares principalmente en datos de expresión génica se encuentran con resultados acordes a los aquí hallados. Cabe destacar que estos autores desarrollan sus análisis sobre bases de datos de rutas metabólicas completas, entre las que se considera el KEGG principalmente. De este modo, Lascorz et al. [151] fueron unos de los primeros autores en aplicar diferentes herramientas para el análisis de rutas incluidas en KEGG. En su artículo, los autores concluyen que, entre todas las rutas metabólicas recogidas en dicha base de datos, solo siete muestran asociación con el CCR, y una de ellas es la ruta de la enfermedad de Huntington, pero ninguna otra de las aquí estudiadas [151]. En este sentido, Quan et al. [152] usan una técnica sobre datos de SNPs en un GWAS basado también en el CCR. Nuevamente estos autores detectan la ruta de la enfermedad de Huntington en la base de datos analizada, sin mostrar más coincidencias con las otras ocho rutas [152]. Por otro lado, opuestos a estos resultados, Chen et al. [153] detectan como rutas metabólicas asociadas con el CCR el metabolismo de la nicotina y nicotinamida y la ruta de señalización del TGF $\beta$ , pero ninguna de las nueve rutas en las que participa el PGC1 $\alpha$  a pesar de emplear la base de datos del KEGG [153].

Por último, Li et al. [154] emplean dos técnicas diferentes para el mismo tipo de análisis, con las que intentan identificar rutas metabólicas descritas en la base de datos del KEGG posiblemente asociadas con el cáncer colorrectal [154]. En este estudio vuelve a presentarse cierta controversia en los resultados sobre la enfermedad de Huntington, ya que una confirma que la ruta de está asociada con el riesgo de presentar la enfermedad y la otra no [154].

Además de estos autores, hay otros que emplean otras bases de datos para el análisis. Kou et al. [155] trabajan con datos de expresión génica sobre las bases de datos KEGG y GO. De este modo, detectan la existencia de 3 rutas metabólicas sobre-reguladas y 10 rutas reprimidas asociadas con el CCR, encontrando de nuevo la ruta de la enfermedad de Huntington entre los rutas metabólicas detectadas en el segundo grupo [155]. Lu et al. [156] encuentran 17 posibles rutas metabólicas asociadas con la enfermedad en la base de datos GO, sin embargo, ninguna de ellas es coincidente con las nueve en las que se encuentra el gen PGC1 $\alpha$  [156]. Para finalizar, Hong et al. [157], quienes analizan los datos de expresión de diversas maneras, tampoco detectan ninguna de las rutas aquí analizadas [157].

Por lo tanto, los resultados basados en el análisis de la base de datos del KEGG y otras bases de datos son similares a los que aquí se presentan. Todos ellos [151-157] comprueban que ninguna de las rutas en las que el PGC1 $\alpha$  participa está asociada con el CCR, a excepción de la ruta de la enfermedad de Huntington, en la que hay discrepancias. En nuestros resultados, esta enfermedad muestra elevados valores para el ES en GSEA-SNP, GenGen y GSA-SNP, pero sólo esta última la considera significativa, por lo que, en el consenso de los resultados, no la podemos considerar como tal. Por su parte, los autores Hong et al. [157], Chen et al. [153] y Li et al. [154] contemplan unos resultados similares a los aquí hallados mientras que Quan et al. [152], Lascorz et al. [151] y Kou et al. [155] están a favor de que la ruta de Huntington muestra asociación con la enfermedad.

Es necesario destacar que, el hecho de que rutas metabólicas como la de la biogénesis mitocondrial (hsa03029), posiblemente asociada con el riesgo de padecer CCR según lo comentado en el capítulo 3 de la presente tesis, no muestre una asociación significativa en estos análisis puede ser debido a la baja representación que tienen sus genes en los datos del Exome Array empleado. Tal como muestran los resultados de la técnica del GSEA-SNP (ver dicha sección), mientras que la vía de la enfermedad de Huntington muestra una representación de un 38% de sus genes en los datos, la biogénesis mitocondrial sólo tiene un 13% y, sin embargo, el valor de ES que alcanza cada una tan sólo difiere en 0.02, siendo de 0.42 para la primera y de 0.4 para la segunda. Por este motivo cabe considerar que la falta de significación puede ser debida a falta de potencia estadística y, siguiendo la afirmación de Altman y Bland: “La ausencia de evidencia no significa que exista una evidencia de ausencia” [158], es recomendable continuar con la investigación en este campo.

La principal fortaleza de estas técnicas es que controlan el impacto de la red de interacción entre genes, facilitando la interpretación biológica de los resultados. Sin embargo, estas técnicas también presentan algunas limitaciones, como que una ruta metabólica estadísticamente significativa solo demuestra que los SNPs de los genes de dicha ruta no se distribuyen al azar en el Exome Array completo [157] o que cada ruta es analizada de modo independiente al resto, sin considerar genes comunes entre ellas, lo que puede generar cruces de información y superposición, dando como resultado un incremento en el número de falsos positivos [159]. Por otro lado, este análisis también presenta sus propias fortalezas y limitaciones en relación a nuestros datos. Como fortaleza principal está el uso de las cinco técnicas para lograr un consenso entre los resultados. Como limitación principal, es que muchas de las rutas metabólicas aquí analizadas tienen una baja representación en los datos del Exome Array empleado, lo que reduce el poder estadístico del estudio.

CAPÍTULO 7

# Corolario

---



El cáncer colorrectal es un importante problema de salud pública a nivel mundial y una de las principales causas de muerte debida a esta enfermedad [46]. Entre los factores de riesgo que pueden influir en el desarrollo de este tumor, se encuentran las alteraciones en las funciones normales del PGC1 $\alpha$ . Este gen es conocido como el regulador pleiotrópico de las rutas metabólicas involucradas en el metabolismo energético celular. Su expresión varía en función de los requerimientos energéticos de cada tejido, mostrando una expresión más elevada en tejidos de alta demanda energética o con una mayor capacidad oxidativa [106]. Para llevar a cabo su función, el PGC1 $\alpha$  se encarga de la activación de coactivadores como el NRF-1 y NRF-2, con los que se consigue controlar procesos como el de la biogénesis mitocondrial y otros aspectos del metabolismo energético [106, 107].

En el intestino y en condiciones normales, la función principal del PGC1 $\alpha$  es la de mantener el equilibrio entre la apoptosis y la proliferación celular (Figura 12-A del capítulo 3). Para ello, en la parte alta de las vellosidades intestinales, donde los enterocitos ya están totalmente diferenciados, la expresión del PGC1 $\alpha$  es elevada, produciendo un incremento de la concentración de especies reactivas de oxígeno y, por lo tanto, un incremento del suicidio celular. Por el contrario, en la parte baja de las criptas, donde las células aún no están diferenciadas, la expresión de este gen es muy baja, permitiendo que la proliferación celular se lleve a cabo. De este modo, se puede considerar que el PGC1 $\alpha$  es el encargado de la regulación de la renovación celular continua que padece el epitelio intestinal [30].

Alteraciones en el patrón de expresión de este gen pueden tener consecuencias graves para el correcto funcionamiento del organismo, pudiendo producir un desequilibrio entre el proceso de apoptosis y la proliferación celular [141]. Sin embargo, los factores que producen estas alteraciones en el patrón de expresión del gen o los mecanismos moleculares por los que las alteraciones incrementan o reducen el riesgo de padecer el cáncer, aún no han sido totalmente investigadas. En el capítulo 3 se recogen los pocos estudios realizados hasta la fecha sobre los posibles mecanismos de acción por los que el gen puede promover el desarrollo del tumor. Según la información obtenida hasta la actualidad, parecen existir cuatro mecanismos de acción principales por los que la sobreexpresión del gen conduce al desarrollo del tumor y otros dos por los que el silenciamiento o inhibición del PGC1 $\alpha$  obtienen la misma consecuencia (Figura 15 capítulo 3). Sin embargo, también se han observado resultados de artículos donde, tanto la represión como la sobreexpresión del gen ejercen un efecto protector frente al desarrollo tumoral, mostrando controversia entre los resultados de los análisis realizados. De estos datos se puede concluir que, a pesar de que la función del gen en el desarrollo del cáncer no está del todo determinada ni definida, sí parece evidente que alteraciones en los niveles de expresión del mismo presentan consecuencias en el desarrollo de esta enfermedad [141].

## 7 . Corolario

---

Con el fin de analizar las alteraciones en este gen, se llevó a cabo un análisis SNP a SNP de los polimorfismos disponibles en la base de datos del proyecto MCC-Spain mapeados  $\pm 500pb$  aguas arriba y abajo del gen  $PGC1\alpha$ , puesto que los SNPs suponen uno de los principales factores capaces de modificar los niveles de expresión de un gen [130]. En la discusión del capítulo 4 se observó que, de todos los SNPs presentes en este gen, tan sólo el rs12505641 muestra una asociación significativa con el riesgo de desarrollar CCR en general. Según los resultados de este análisis, el alelo A del rs12505641 supone un factor de riesgo para el desarrollo de tumor de recto en varones, principalmente. Este SNP se encuentra en una región intergénica entre el gen  $PGC1\alpha$  y el  $CDC42P6$  [137]. Sin embargo, debido a que i) el SNP está más próximo a la región promotora del  $PGC1\alpha$  que a la del  $CDC42P6$ ; y ii) a que el  $CDC42P6$  es un pseudogen que, en la actualidad, no desempeña ninguna función en el organismo [138]; se puede interpretar que este SNP ejerce su función sobre el gen de interés o actúa como tag-SNP de otro polimorfismo localizado en el mismo. Además, diferentes factores ambientales que tienen efectos conocidos sobre los niveles de expresión del  $PGC1\alpha$ , como es el consumo de alcohol, la realización de actividad física o diversos alimentos incluidos en la dieta, parecen aumentar o paliar el efecto del SNP sobre el gen [48]. Esto da soporte a la hipótesis de que el rs12505641 es capaz de ejercer su función sobre el promotor del  $PGC1\alpha$ , aunque se requiere de una mayor investigación sobre esta relación, ya que ésta no ha podido ser contrastada por la falta de literatura referida a este polimorfismo.

A pesar de que el análisis SNP a SNP fue uno de los primeros en llevarse a cabo, esta técnica muestra diversas limitaciones. Entre ellas, se puede destacar el hecho de que no considera más que un único polimorfismo cada vez, a pesar de que el desarrollo de las enfermedades complejas se caracteriza por la interacción entre diversas variantes genéticas. Esto puede provocar subestimaciones en el tamaño real del efecto del SNP [85]. Por este motivo, han surgido otros métodos de análisis de polimorfismos basados en conjuntos de genes [87], métodos que dan un nuevo sentido al análisis biológico de los polimorfismos, puesto que se realizan sobre conjuntos de SNP que se relacionan en base a un criterio, como puede ser pertenecer a la misma ruta metabólica. En este sentido, la interpretación biológica de sus resultados es más sencilla, ya que los genes trabajan manteniendo una compleja red de interacción [93].

La variedad de técnicas disponibles para desarrollar este objetivo es muy amplia y, sin embargo, han sido muy pocos los estudios que han realizado comparaciones entre ellas. Por ese motivo, se consideró necesario llevar a cabo una comparación entre cinco técnicas diferentes basadas en el trabajo de Wang et al. [93]. Este trabajo recoge las técnicas más empleadas en este tipo de análisis, diferenciando

dos grupos: técnicas competitivas y técnicas auto-contenidas. Puesto que, en base a la opinión de autores como Fridley et al. [98], las competitivas son más robustas, las técnicas comparadas en el capítulo 5 fueron GSEA-SNP, GSA-SNP, iGSEA4GWAS, GenGen y SRT.

Los resultados de esta comparación muestran que el GSEA-SNP es la técnica que parece adaptarse de forma más real a los datos de análisis, mientras que el SRT parece ser la menos cualificada en la adaptación a los mismos. Del mismo modo, al valorar los resultados falsos positivos que ofrece cada una de las herramientas, GSEA-SNP muestra la menor proporción de estos cuando el conjunto de genes analizado cuenta con entre un 2 y un 20% de genes significativos (definidos por el método de Brown [147]). Por el contrario, el SRT muestra el mayor número de falsos positivos en cada uno de los conjuntos. Por otro lado, autores como Lascorz et al. [151], Wang et al. [93] o Mooney et al. [146] defienden que es posible que la mejor definición de conjunto de genes significativo para un fenotipo la determine el consenso de más de una técnica de trabajo. En el capítulo 5, se observa que los resultados por consenso de las cinco técnicas ayudan a reducir aún más los falsos positivos, dejando ver también que estos resultados son bastante conservadores puesto que incrementan en cierta medida los falsos negativos en aquellos conjuntos de genes que presenten un porcentaje muy bajo de genes significativos (definidos por el método de Brown [147]). Esto permite concluir que, en función del objetivo del análisis, la información más conservadora la ofrece el consenso de los resultados de las cinco técnicas. Mientras que el consenso de al menos tres de ellas puede ser interesante en análisis donde ser conservador puede ser sinónimo de perder información relevante, como es en el análisis de un número bajo de rutas metabólicas.

Conocidos estos datos, se implementaron estas técnicas en el análisis de las nueve rutas metabólicas en las que el PGC1 $\alpha$  muestra una función principal (según la base de datos del KEGG). Puesto que poco se ha investigado sobre la asociación de estas rutas con el CCR y en base al análisis realizado en el capítulo 5, se decidió considerar como significativos aquellos conjuntos de genes que resultasen serlo en al menos tres de las cinco técnicas aplicadas. De este modo, en el capítulo 6 se pudo observar que ninguna de las rutas analizadas muestra una asociación significativa con el riesgo de padecer CCR. Estos resultados coinciden con los observados hasta la fecha en la bibliografía, a excepción de la ruta de la enfermedad de Huntington, la cual muestra controversia. Esto parece ser debido a que la ruta comparte múltiples genes con la cadena de la fosforilación oxidativa, asociada al metabolismo mitocondrial y al crecimiento tumoral [151]. Sin embargo, rutas metabólicas directamente asociadas con la mitocondria, como son la biogénesis mitocondrial o la ruta de señalización de la AMPK tampoco muestran asociación significativa con el desarrollo de la enfermedad.

## 7 . Corolario

---

Aunque cabe destacar que, en técnicas como la del GSEA-SNP, la ruta de la biogénesis mitocondrial en concreto muestra un elevado ES a pesar de tener una representación muy baja de sus genes en los datos empleados, indicando así que el hecho de que no alcance la significación estadística pueda ser debido a una falta de poder estadístico en los datos analizados.

CAPÍTULO 8

# Conclusiones

---

## 8 . Conclusiones

---

1. La función que presenta el PGC1 $\alpha$  en el riesgo de desarrollar CCR aún está por explorar. La investigación llevada a cabo hasta la fecha ha sido realizada *in vitro* y sobre animales, existiendo poca información en humanos.
2. Las alteraciones en los niveles de expresión del gen PGC1 $\alpha$  influyen sobre el riesgo de padecer CCR. La evidencia más común es que la sobreexpresión de este gen incrementa el riesgo de desarrollar el tumor, existiendo referencias a que bajos niveles de expresión también influyen en el riesgo de desarrollar esta enfermedad.
3. Entre los SNPs mapeados  $\pm 500pb$  aguas arriba y abajo del gen PGC1 $\alpha$ , disponibles en el Exome Array del proyecto MCC-Spain, sólo el SNP rs12505641 muestra asociación con el fenotipo de interés.
4. El rs12505641 se encuentra aguas arriba del promotor del PGC1 $\alpha$  y la presencia del alelo A en este locus parece tener consecuencias inhibitorias sobre la expresión del gen. Por ese motivo, al ajustar la asociación SNP-enfermedad por la realización de ejercicio físico, éste muestra un efecto protector, puesto que su función es aumentar la expresión del gen. Por el contrario, el consumo de alcohol refuerza la actividad del SNP al reducir también la actividad del PGC1 $\alpha$ .
5. El GSEA-SNP muestra una mayor capacidad de discriminación entre los conjuntos de genes significativos y no significativos comparado con el resto de las técnicas analizadas, siendo ésta la técnica que genera un menor número de falsos positivos. Sin embargo, es la técnica que genera un mayor número de falsos negativos.
6. El consenso entre los resultados de diferentes técnicas permite obtener un número menor de falsos positivos para todos los tipos de conjuntos de genes. Sin embargo, genera un mayor número de falsos negativos en conjuntos de genes con un bajo porcentaje de genes significativos. De este modo, la información más conservadora es aquella aportada por el consenso de los resultados entre técnicas.

7. De las nueve rutas metabólicas en las que participa el gen  $PGC1\alpha$  según la base de datos del KEGG, ninguna muestra una asociación significativa con el desarrollo del CCR en el consenso de resultados de las técnicas de análisis empleadas. Tan sólo la ruta de la enfermedad de Huntington muestra un resultado significativo para la técnica GSA-SNP, pero en el consenso de técnicas, es necesario descartarla. Estos resultados son similares a los disponibles en la bibliografía actual.

## Bibliografía

---



- 
- (1) Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011; 144 (5): 646-674.
  - (2) Dalton S, Coverdell PD. Linking the cell cycle to cell fate decisions. *Trends in Cell Biology*. 2015; 25 (10): 395-401.
  - (3) Ross MH, Pawlina W, Negrete JH. *Histología: texto y atlas color con biología celular y molecular*. Editorial Médica Panamericana; 2007 Disponible en: <https://books.google.es/books?hl=es%7B%5C%7Dlr=%7B%5C%7Ddid=NxYmIRZQi2oC%7B%5C%7Ddoi=fnd%7B%5C%7Dpg=PR1%7B%5C%7Ddq=ciclo+celular%7B%5C%7Dots=%7B%5C%7DyfUCFls5r%7B%5C%7Dsig=KdfJNyv5nQmjEWANJXwJz%7B%5C%7DacnBQ%7B%5C%7Dv=onepage%7B%5C%7Dq=celular%7B%5C%7Df=false> [Visitado el 02-04-2018].
  - (4) Lóbrich M, Jeggo PA. The impact of a negligent G2/M checkpoint on genomic instability and cancer induction. *Nature Reviews Cancer*. 2007; 7 (11): 861.
  - (5) Malumbres M, Barbacid M. Milestones in cell division: to cycle or not to cycle: a critical decision in cancer. *Nature Reviews Cancer*. 2001; 1 (3): 222.
  - (6) Shay JW. Role of telomeres and telomerase in aging and cancer. *Cancer Discovery*. 2016; 6 (6): 584-593.
  - (7) Ecsey J, Hunter D. *Textbook of cancer epidemiology*. Second Edition. Oxford; 2008 Disponible en: <https://books.google.es/books?hl=es&lr=&id=kuvMDqQuKEUC&oi=fnd&pg=PA308&dq=6.%09Hankinson+S,+Tamimi+R,+Hunter+D.+Breast+Cancer.+In:+Adami+H0,+Hunter+D,+Trichopoulos+D.+Textbook+of+Cancer+Epidemiology.+2nd+ed.+Oxford+University+Press,+2008,+pp:+403-445.&ot> [Visitado el 25-04-2017].
  - (8) National Cancer Institute. *What Is Cancer?* Disponible en: <https://www.cancer.gov/about-cancer/understanding/what-is-cancer> [Visitado el 20-03-2018].
  - (9) Smolková K, Plecítá-Hlavatá L, Bellance N, Benard G, Rossignol R, Ježek P. Waves of gene regulation suppress and then restore oxidative phosphorylation in cancer cells. *The International Journal of Biochemistry & Cell Biology*. 2011; 43 (7): 950-968.
  - (10) Weinberg RA. *The Biology of Cancer*. Second Edition. Garland Science; 2014 Disponible en: [https://books.google.es/books?id=MzMmAgAAQBAJ&printsec=frontcover&dq=cancer&hl=es&sa=X&redir\\_esc=y#v=onepage&q=cancer&f=true](https://books.google.es/books?id=MzMmAgAAQBAJ&printsec=frontcover&dq=cancer&hl=es&sa=X&redir_esc=y#v=onepage&q=cancer&f=true) [Visitado el 03-05-2017].
  - (11) Potter M, Newport E, Morten KJ. The Warburg effect: 80 years on. *Biochemical Society Transactions*. 2016; 44 (5): 1499-1505.
-

## BIBLIOGRAFÍA

---

- (12) Warburg O, Wind F, Negelein E. The metabolism of tumors in the body. *The Journal of General Physiology*. 1927; 8 (6): 519.
- (13) Goodwin ML, Gladden LB, Nijsten MW, Jones KB. Lactate and cancer: revisiting the Warburg effect in an era of lactate shuttling. *Frontiers in Nutrition*. 2015; 1 : 27.
- (14) Gladden L. Lactate metabolism: a new paradigm for the third millennium. *The Journal of Physiology*. 2004; 558 (1): 5-30.
- (15) Sakashita M, Aoyama N, Minami R, Maekawa S, Kuroda K, Shirasaka D et al. Glut1 expression in T1 and T2 stage colorectal carcinomas: its relationship to clinicopathological features. *European Journal of Cancer*. 2001; 37 (2): 204-209.
- (16) DeBerardinis RJ, Mancuso A, Daikhin E, Nissim I, Yudkoff M, Wehrli S et al. Beyond aerobic glycolysis: transformed cells can engage in glutamine metabolism that exceeds the requirement for protein and nucleotide synthesis. *Proceedings of the National Academy of Sciences*. 2007; 104 (49): 19345-19350.
- (17) Fantin VR, St-Pierre J, Leder P. Attenuation of LDH-A expression uncovers a link between glycolysis, mitochondrial physiology, and tumor maintenance. *Cancer Cell*. 2006; 9 (6): 425-434.
- (18) Schulz TJ, Thierbach R, Voigt A, Drewes G, Mietzner B, Steinberg P et al. Induction of oxidative metabolism by mitochondrial frataxin inhibits cancer growth Otto Warburg Revisited. *Journal of Biological Chemistry*. 2006; 281 (2): 977-981.
- (19) Bhalla K, Hwang BJ, Dewi RE, Ou L, Twaddel W, Fang Hb et al. PGC1 $\alpha$  promotes tumor growth by inducing gene expression programs supporting lipogenesis. *Cancer Research*. 2011; 71 (21): 6888-6898.
- (20) Mazzanti R, Giulivi C. Coordination of nuclear-and mitochondrial-DNA encoded proteins in cancer and normal colon tissues. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*. 2006; 1757 (5-6): 618-623.
- (21) Liang H, Ward WF. PGC-1 $\alpha$ : a key regulator of energy metabolism. *Advances in Physiology Education*. 2006; 30 (4): 145-151.
- (22) Jornayvaz FR, Shulman GI. Regulation of mitochondrial biogenesis. *Essays in Biochemistry*. 2010; 47 : 69-84.
- (23) Ashrafi G, Schwarz T. The pathways of mitophagy for quality control and clearance of mitochondria. *Cell Death and Differentiation*. 2013; 20 (1): 31.
- (24) Palikaras K, Tavernarakis N. Mitochondrial homeostasis: the interplay between mitophagy and mitochondrial biogenesis. *Experimental Gerontology*. 2014; 56 : 182-188.

- 
- (25) Passarge E, Wirth J. *Genética : texto y atlas*. Editorial Médica Panamericana; 2010 Disponible en: [https://books.google.es/books?id=bgQ%7B%5C\\_%7DxyJYkigC%7B%5C%7Dpg=PA150%7B%5C%7Ddq=ligamiento+gen%7B%5C'%7Be%7D%7Dtico%7B%5C%7Dhl=es%7B%5C%7Dsa=X%7B%5C%7Dredir%7B%5C\\_%7Ddesc=y%7B%5C%7Dv=onepage%7B%5C%7Ddq=ligamiento%20gen%7B%5C'%7Be%7D%7Dtico%7B%5C%7Df=true%20https://books.google.es/books?id=bgQ%7B%5C\\_%7DxyJYkigC%7B%5C%7Dpg=PA130%7B%5C%7Ddq=mitocondria%7B%5C%7Dhl=es%7B%5C%7Dsa=X%7B%5C%7Dredir%7B%5C\\_%7Ddesc=y%7B%5C%7Dv=onepage%7B%5C%7Ddq=mitocondri](https://books.google.es/books?id=bgQ%7B%5C_%7DxyJYkigC%7B%5C%7Dpg=PA150%7B%5C%7Ddq=ligamiento+gen%7B%5C'%7Be%7D%7Dtico%7B%5C%7Dhl=es%7B%5C%7Dsa=X%7B%5C%7Dredir%7B%5C_%7Ddesc=y%7B%5C%7Dv=onepage%7B%5C%7Ddq=ligamiento%20gen%7B%5C'%7Be%7D%7Dtico%7B%5C%7Df=true%20https://books.google.es/books?id=bgQ%7B%5C_%7DxyJYkigC%7B%5C%7Dpg=PA130%7B%5C%7Ddq=mitocondria%7B%5C%7Dhl=es%7B%5C%7Dsa=X%7B%5C%7Dredir%7B%5C_%7Ddesc=y%7B%5C%7Dv=onepage%7B%5C%7Ddq=mitocondri) [Visitado el 03-05-2017].
- (26) Schon EA, DiMauro S, Hirano M. Human mitochondrial DNA: roles of inherited and somatic mutations. *Nature Reviews Genetics*. 2012; 13 (12): 878.
- (27) Liberti MV, Locasale JW. The Warburg effect: how does it benefit cancer cells? *Trends in Biochemical Sciences*. 2016; 41 (3): 211-218.
- (28) Lauer C, Volkl A, Riedl S, Fahimi HD, Beier K. Impairment of peroxisomal biogenesis in human colon carcinoma. *Carcinogenesis*. 1999; 20 (6): 985-989.
- (29) Girnun GD. The diverse role of the PPAR $\gamma$  coactivator 1 family of transcriptional coactivators in cancer. *Seminars in Cell & Developmental Biology*. 2012; 23 (4): 381-388.
- (30) D'Errico I, Salvatore L, Murzilli S, Sasso GL, Latorre D, Martelli N et al. Peroxisome proliferator-activated receptor- $\gamma$  coactivator 1- $\alpha$  (PGC1 $\alpha$ ) is a metabolic regulator of intestinal epithelial cell fate. *Proceedings of the National Academy of Sciences*. 2011; 108 (16): 6603-6608.
- (31) Grabacka M, Pierzchalska M, Reiss K. Peroxisome proliferator activated receptor  $\alpha$  ligands as anticancer drugs targeting mitochondrial metabolism. *Current Pharmaceutical Biotechnology*. 2013; 14 (3): 342-356.
- (32) Wolf DA. Is reliance on mitochondrial respiration a “chink in the armor” of therapy-resistant cancer? *Cancer Cell*. 2014; 26 (6): 788-795.
- (33) Mastropasqua F, Girolimetti G, Shoshan M. PGC1 $\alpha$ : Friend or Foe in Cancer? *Genes*. 2018; 9 (1): 48.
- (34) Young B, Woodford P, O'Dowd G. *Wheater's Functional Histology E-Book: A Text and Colour Atlas*. Elsevier Health Sciences; 2013 Disponible en: [https://books.google.es/books?hl=es&lr=&id=HU1YAQAQBAJ&oi=fnd&pg=PP1&dq=Wheater%27s+Functional+Histology+E-Book:+A+Text+and+Colour+Atlas&ots=F7A1\\_\\_xNAy&sig=wSF1RXK3ZcUEa1cHv4LSD1qBaHc#v=onepage&q=Wheater's%20Functional%20Histology%20E-Book%3A%20A%20Text%20and%20Colour%20Atlas&f=false](https://books.google.es/books?hl=es&lr=&id=HU1YAQAQBAJ&oi=fnd&pg=PP1&dq=Wheater%27s+Functional+Histology+E-Book:+A+Text+and+Colour+Atlas&ots=F7A1__xNAy&sig=wSF1RXK3ZcUEa1cHv4LSD1qBaHc#v=onepage&q=Wheater's%20Functional%20Histology%20E-Book%3A%20A%20Text%20and%20Colour%20Atlas&f=false) [Visitado el 19-05-2017].
-

## BIBLIOGRAFÍA

---

- (35) Welsch U, Sobotta J. *Histología*. Editorial Médica Panamericana; 2008 Disponible en: <https://books.google.es/books?id=7zFxo6bmx10C&pg=PA380&dq=histolog%C3%ADa+del+intestino+grueso&hl=es&sa=X&ved=0ahUKEwjH2Jybsd7WAhVRY1AKHUEFC-wQ6AEIJjAA#v=onepage&q=histolog%C3%ADa%20del%20intestino%20grueso&f=false> [Visitado el 14-05-2017].
- (36) Eynard AR, Valentich MA, Rovasio RA. *Histología y embriología del ser humano: bases celulares y moleculares*. Editorial Médica Panamericana; 2008 Disponible en: <https://books.google.es/books?id=p1JSyuGai0oC&printsec=frontcover&dq=Histologi%CC%81a+y+embriologi%CC%81a+del+ser+humano+:+bases+celulares+y+moleculares&hl=es&sa=X&ved=0ahUKEwi8q-mcoLzbAhXLJ1AKHbaJAfsQ6AEIJzAA#v=onepage&q=Histologi%CC%81a%20y%20embriologi%CC%81a%20del%20ser%20humano%20%3A%20bases%20celulares%20y%20moleculares&f=false> [Visitado el 03-06-2017].
- (37) De Rosa M, Pace U, Rega D, Costabile V, Duraturo F, Izzo P et al. Genetics, diagnosis and management of colorectal cancer. *Oncology Reports*. 2015; 34 (3): 1087-1096.
- (38) Macarulla T, Élez E, Capdevila J, Tabernero J. *Cáncer de colon y recto*. Editorial AMAT; 2011 Disponible en: [https://books.google.es/books?id=QCNJlCR7dMMC&printsec=frontcover&dq=c%C3%A1ncer&hl=es&sa=X&redir\\_esc=y#v=onepage&q=c%C3%A1ncer&f=true](https://books.google.es/books?id=QCNJlCR7dMMC&printsec=frontcover&dq=c%C3%A1ncer&hl=es&sa=X&redir_esc=y#v=onepage&q=c%C3%A1ncer&f=true) [Visitado el 08-05-2017].
- (39) Piñol V, Andreu M, Castells A, Payá A, Bessa X, Jover R. Frequency of hereditary non-polyposis colorectal cancer and other colorectal cancer familial forms in Spain: a multicentre, prospective, nationwide study. *European Journal of Gastroenterology & Hepatology*. 2004; 16 (1): 39-45.
- (40) Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *International Journal of Cancer*. 2015; 136 (5): 359-386.
- (41) Center MM, Jemal A, Ward E. International trends in colorectal cancer incidence rates. *Cancer Epidemiology and Prevention Biomarkers*. 2009; 18 (6): 1688-1694.
- (42) Howlader NN, Krapcho M, Neyman N, Aminou R, Waldron W, Altekruse S et al. *SEER cancer statistics review (CSR), 1975-2014*. National Cancer Institute; 2017 Disponible en: [https://seer.cancer.gov/csr/1975%7B%5C\\_%7D2014/sections.html](https://seer.cancer.gov/csr/1975%7B%5C_%7D2014/sections.html) [Visitado el 08-05-2017].
- (43) Potter JD, Hunter D. *Textbook of Cancer Epidemiology*. Oxford; 2008 Disponible en: <https://books.google.es/books?id=rj9FDwAAQBAJ&>

- printsec=frontcover&dq=Textbook+of+Cancer+Epidemiology&hl=es&sa=X&ved=0ahUKEwiJ4-CqobzbAhUDbRQKHx6eDYIQ6AEIKjAA#v=onepage&q=Textbook%20of%20Cancer%20Epidemiology&f=false [Visitado el 20-03-2018].
- (44) McMichael AJ, Giles GG. Cancer in migrants to Australia: extending the descriptive epidemiological data. *Cancer Research*. 1988; 48 (3): 751-756.
- (45) Ibáñez-Sanz G, Díez-Villanueva A, Alonso MH, Rodríguez-Moranta F, Pérez-Gómez B, Bustamante M et al. Risk model for colorectal cancer in Spanish population using environmental and genetic factors: results from the MCC-Spain study. *Scientific Reports*. 2017; 7 : 43263.
- (46) Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *International Journal of Cancer*. 2010; 127 (12): 2893-2917.
- (47) Woolf CM. A genetic study of carcinoma of the large intestine. *American Journal of Human Genetics*. 1958; 10 (1): 42-7.
- (48) Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. *CA: A Cancer Journal for Clinicians*. 2015; 65 (2): 87-108.
- (49) Slattery ML, Curtin K, Anderson K, Ma KN, Ballard L, Edwards S et al. Associations between cigarette smoking, lifestyle factors, and microsatellite instability in colon tumors. *Journal of the National Cancer Institute*. 2000; 92 (22): 1831-1836.
- (50) Van Blarigan EL, Meyerhardt JA. Role of physical activity and diet after colorectal cancer diagnosis. *Journal of Clinical Oncology*. 2015; 33 (16): 1825.
- (51) Je Y, Jeon JY, Giovannucci EL, Meyerhardt JA. Association between physical activity and mortality in colorectal cancer: a meta-analysis of prospective cohort studies. *International Journal of Cancer*. 2013; 133 (8): 1905-1913.
- (52) Kancer W, Fonds O. *Food, nutrition, physical activity, and the prevention of cancer: a global perspective*. American Institute for Cancer Research; 2007; 1 : Disponible en: [http://www.aicr.org/assets/docs/pdf/reports/Second\\_Expert\\_Report.pdf](http://www.aicr.org/assets/docs/pdf/reports/Second_Expert_Report.pdf) [Visitado el 22-05-2018].
- (53) Hong WK, Kufe Bast RC, W. D, Hait W, Holland JF, Frei E et al. *Cancer Medicine*. American Association of Cancer Research; 2010; 8 : Disponible en: [https://books.google.es/books?id=R0FbhLsWHBEC&pg=PA925&dq=Holland-Frei+Cancer+Medicine,+6th+edition&hl=es&sa=X&ved=0ahUKEwj6wrH5\\_-HQAUI6xQKHbS9AwMQ6AEIHTAA#v=onepage&q=colorectal%20cancer&f=false](https://books.google.es/books?id=R0FbhLsWHBEC&pg=PA925&dq=Holland-Frei+Cancer+Medicine,+6th+edition&hl=es&sa=X&ved=0ahUKEwj6wrH5_-HQAUI6xQKHbS9AwMQ6AEIHTAA#v=onepage&q=colorectal%20cancer&f=false) [Visitado el 08-05-2017].

## BIBLIOGRAFÍA

---

- (54) Winkels RM, Heine-Bróring RC, Van Zutphen M, van Harten-Gerritsen S, Kok DE, Van Duijnhoven FJ et al. The COLON study: Colorectal cancer: Longitudinal, Observational study on Nutritional and lifestyle factors that may influence colorectal tumour recurrence, survival and quality of life. *BioMed Central Cancer*. 2014; 14 (1): 374.
- (55) Cuzick J, Otto F, Baron JA, Brown PH, Burn J, Greenwald P et al. Aspirin and non-steroidal anti-inflammatory drugs for cancer prevention: an international consensus statement. *The Lancet Oncology*. 2009; 10 (5): 501-507.
- (56) Persson PG, Bernell O, Leijonmarck CE, Farahmand BY, Hellers G, Ahlbom A. Survival and cause-specific mortality in inflammatory bowel disease: a population-based cohort study. *Gastroenterology*. 1996; 110 (5): 1339-45.
- (57) Mellemkjaer L, Olsen JH, Frisch M, Johansen C, Gridley G, McLaughlin JK. Cancer in patients with ulcerative colitis. *International Journal of Cancer*. 1995; 60 (3): 330-3.
- (58) Rogler G. Chronic ulcerative colitis and colorectal cancer. *Cancer Letters*. 2014; 345 (2): 235-241.
- (59) Gillen C, Walmsley R, Prior P, Andrews H, Allan R. Ulcerative colitis and Crohn's disease: a comparison of the colorectal cancer risk in extensive colitis. *Cancer Epidemiology Research Unit*. 1994; 35 (11): 1590-1592.
- (60) Riddell RH, Goldman H, Ransohoff DF, Appelman HD, Fenoglio CM, Haggitt RC et al. Dysplasia in inflammatory bowel disease: standardized classification with provisional clinical applications. *Human Pathology*. 1983; 14 (11): 931-68.
- (61) Scaringi S, Di Martino C, Zambonin D, Fazi M, Canonico G, Leo F et al. Colorectal cancer and Crohn's colitis: clinical implications from 313 surgical patients. *World Journal of Surgery*. 2013; 37 (4): 902-910.
- (62) Herceg Z, Hainaut P. Genetic and epigenetic alterations as biomarkers for cancer detection, diagnosis and prognosis. *Molecular Oncology*. 2007; 1 (1): 26-41.
- (63) Stewart BW, Wild CP. World Cancer Report 2014. *International Agency For Research On Cancer*. 2014; 422-431.
- (64) Markowitz SD, Bertagnolli MM. Molecular origins of cancer: Molecular basis of colorectal cancer. *New England Journal of Medicine*. 2009; 361 (25): 2449-2460.
- (65) The Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature*. 2012; 487 (7407): 330.

- 
- (66) Laurent-Puig P, Béroud C, Soussi T. APC gene: database of germline and somatic mutations in human tumors and cell lines. *Nucleic Acids Research*. 1998; 26 (1): 269-270.
- (67) Kwong LN, Dove WF. APC and its modifiers in colon cancer. *Advances in experimental medicine and biology*. 2009; 656 : 85-106.
- (68) Dang CV. MYC on the path to cancer. *Cell*. 2012; 149 (1): 22-35.
- (69) Guinney J, Dienstmann R, Wang X, de Reynies A, Schlicker A, Sonesson C et al. The consensus molecular subtypes of colorectal cancer. *Nature Medicine*. 2016; 21 (11): 1350-1356.
- (70) Olivier M, Hollstein M, Hainaut P. TP53 mutations in human cancers: origins, consequences, and clinical use. *Cold Spring Harbor Perspectives in Biology*. 2010; 2 (1): a001008.
- (71) Peltomáki P. Deficient DNA mismatch repair: a common etiologic factor for colon cancer. *Human Molecular Genetics*. 2001; 10 (7): 735-740.
- (72) Sinicrope FA, Sargent DJ. Molecular pathways: microsatellite instability in colorectal cancer: prognostic, predictive, and therapeutic implications. *Clinical Cancer Research*. 2012; 18 (6): 1506-1512.
- (73) Burt R, Bishop D, Lynch HT, Rozen P, Winawer S. Risk and surveillance of individuals with heritable factors for colorectal cancer. WHO Collaborating Centre for the Prevention of Colorectal Cancer. *Bulletin of the World Health Organization*. 1990; 68 (5): 655.
- (74) National Cancer Institute. *Genetics of colorectal cancer*. Disponible en: [https://www.cancer.gov/types/colorectal/hp/colorectal-genetics-pdq#section/\\_72](https://www.cancer.gov/types/colorectal/hp/colorectal-genetics-pdq#section/_72) [Visitado el 21-03-2018].
- (75) Shaw G. Polymorphism and single nucleotide polymorphisms (SNPs). *British Journal of Urology International*. 2013; 112 (5): 664-665.
- (76) Alonso-Molero J, González-Donquiles C, Palazuelos C, Fernández-Villa T, Ramos E, Pollán M et al. The rs4939827 polymorphism in the SMAD7 gene and its association with Mediterranean diet in colorectal carcinogenesis. *BioMed Central Medical Genetics*. 2017; 18 (1): 122.
- (77) Huang Y, Wu W, Nie M, Li C, Wang L. SMAD7 polymorphisms and colorectal cancer risk: a meta-analysis of case-control studies. *Oncotarget*. 2016; 7 (46): 75561.
- (78) Yang H, Gao Y, Feng T, Jin TB, Kang LL, Chen C. Meta-analysis of the rs4779584 polymorphism and colorectal cancer risk. *PLoS One*. 2014; 9 (2): e89736.
-

## BIBLIOGRAFÍA

---

- (79) Kimchi-Sarfaty C, Oh JM, Kim IW, Sauna ZE, Calcagno AM, Ambudkar SV et al. A “silent” polymorphism in the MDR1 gene changes substrate specificity. *Science*. 2007; 315 (5811): 525-528.
- (80) Hoskins JM, Ong PS, Keku TO, Galanko JA, Martin CF, Coleman CA et al. Association of eleven common, low-penetrance colorectal cancer susceptibility genetic variants at six risk loci with clinical outcome. *PLoS One*. 2012; 7 (7): e41954.
- (81) Picelli S, Zajac P, Zhou XL, Edler D, Lenander C, Dalén J et al. Common variants in human CRC genes as low-risk alleles. *European Journal of Cancer*. 2010; 46 (6): 1041-1048.
- (82) Wang L, Jia P, Wolfinger RD, Chen X, Zhao Z. Gene set analysis of genome-wide association studies: methodological issues and perspectives. *Genomics*. 2011; 98 (1): 1-8.
- (83) Wang K, Li M, Bucan M. Pathway-based approaches for analysis of genomewide association studies. *The American Journal of Human Genetics*. 2007; 81 (6): 1278-1283.
- (84) Tsafnat G, Jasch D, Misra A, Choong MK, Lin FPY, Coiera E. Gene-disease association with literature based enrichment. *Journal of Biomedical Informatics*. 2014; 49 : 221-226.
- (85) Stringer S, Wray NR, Kahn RS, Derks EM. Underestimated effect sizes in GWAS: fundamental limitations of single SNP analysis for dichotomous phenotypes. *PLoS One*. 2011; 6 (11): e27964.
- (86) Jin L, Zuo XY, Su WY, Zhao XL, Yuan MQ, Han LZ et al. Pathway-based analysis tools for complex diseases: a review. *Genomics, Proteomics & Bioinformatics*. 2014; 12 (5): 210-220.
- (87) Zhang K, Cui S, Chang S, Zhang L, Wang J. i-GSEA4GWAS: a web server for identification of pathways/gene sets associated with traits by applying an improved gene set enrichment analysis to genome-wide association study. *Nucleic Acids Research*. 2010; 38 (suppl.2): W90-W95.
- (88) Clarke GM, Anderson Ca, Pettersson FH, Cardon LR, Andrew P. Basic statistical analysis in genetic case-control studies Geraldine. *Nature Protocols*. 2011; 6 (2): 121-133.
- (89) Braun R, Buetow K. Pathways of distinction analysis: a new technique for multi-SNP analysis of GWAS data. *PLoS Genetics*. 2011; 7 (6): e1002101.
- (90) Klein RJ, Zeiss C, Chew EY, Tsai JY, Sackler RS, Haynes C et al. Complement factor H polymorphism in age-related macular degeneration. *Science*. 2005; 308 (5720): 385-389.



- 
- (91) Ku CS, Loy EY, Pawitan Y, Chia KS. The pursuit of genome-wide association studies: where are we now? *Journal of Human Genetics*. 2010; 55 (4): 195.
- (92) Weng L, Macciardi F, Subramanian A, Guffanti G, Potkin SG, Yu Z et al. SNP-based pathway enrichment analysis for genome-wide association studies. *BioMed Central Bioinformatics*. 2011; 12 (1): 99.
- (93) Wang K, Li M, Hakonarson H. Analysing biological pathways in genome-wide association studies. *Nature Reviews Genetics*. 2010; 11 (12): 843.
- (94) Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*. 2005; 102 (43): 15545-15550.
- (95) Tintle NL, Borchers B, Brown M, Bekmetjev A. Comparing gene set analysis methods on single-nucleotide polymorphism data from Genetic Analysis Workshop 16. *BioMed Central Proceedings*. 2009; 3 : S96.
- (96) Chasman DI. On the utility of gene set methods in genomewide association studies of quantitative traits. *Genetic Epidemiology*. 2008; 32 (7): 658-668.
- (97) Holden M, Deng S, Wojnowski L, Kulle B. GSEA-SNP: applying gene set enrichment analysis to SNP data from genome-wide association studies. *Bioinformatics*. 2008; 24 (23): 2784-2785.
- (98) Fridley BL, Biernacka JM. Gene set analysis of SNP data: benefits, challenges, and future directions. *European Journal of Human Genetics*. 2011; 19 (8): 837.
- (99) Maciejewski H. Gene set analysis methods: statistical models and methodological differences. *Briefings in Bioinformatics*. 2013; 15 (4): 504-518.
- (100) Tarca AL, Bhatti G, Romero R. A comparison of gene set analysis methods in terms of sensitivity, prioritization and specificity. *PLoS One*. 2013; 8 (11): e79217.
- (101) O'dushlaine C, Kenny E, Heron EA, Segurado R, Gill M, Morris DW et al. The SNP ratio test: pathway analysis of genome-wide association datasets. *Bioinformatics*. 2009; 25 (20): 2762-2763.
- (102) Nam D, Kim J, Kim SY, Kim S. GSA-SNP: a general approach for gene set analysis of polymorphisms. *Nucleic Acids Research*. 2010; 38 (suppl.2): W749-W754.
- (103) Zhang K, Chang S, Guo L, Wang J. I-GSEA4GWAS v2: a web server for functional analysis of SNPs in trait-associated pathways identified from genome-wide association study. *Protein & Cell*. 2015; 6 (3): 221-224.
-

## BIBLIOGRAFÍA

---

- (104) Turan A, Mahmood A. The profile of antioxidant systems and lipid peroxidation across the crypt-villus axis in rat intestine. *Digestive Diseases and Sciences*. 2007; 52 (8): 1840-1844.
- (105) Kanehisa Laboratories. *KEGG T01001: 10891 PPARG coactivator 1 alpha*. Disponible en: [http://www.genome.jp/dbget-bin/www%7B%5C\\_%7Dbget?hsa:10891](http://www.genome.jp/dbget-bin/www%7B%5C_%7Dbget?hsa:10891) [Visitado el 11-11-2016].
- (106) Finck BN, Kelly DP. PGC-1 coactivators: inducible regulators of energy metabolism in health and disease. *The Journal of Clinical Investigation*. 2006; 116 (3): 615-622.
- (107) Scarpulla RC, Vega RB, Kelly DP. Transcriptional integration of mitochondrial biogenesis. *Trends in Endocrinology & Metabolism*. 2012; 23 (9): 459-466.
- (108) National Center for Biotechnology Information Search database. *iCn3D: Web-based 3D Structure Viewer*. Disponible en: <https://www.ncbi.nlm.nih.gov/Structure/icn3d/full.html?showseq=1%7B%5C%7Dmmdbid=94968%7B%5C%7Dbuidx=1%7B%5C%7Dcomplexity=0> [Visitado el 09-04-2018].
- (109) Jose C, Bellance N, Rossignol R. Choosing between glycolysis and oxidative phosphorylation: a tumor's dilemma? *Biochimica et Biophysica Acta (BBA)-Bioenergetics*. 2011; 1807 (6): 552-561.
- (110) Bianchi G, Martella R, Ravera S, Marini C, Capitanio S, Orengo A et al. Fasting induces anti-Warburg effect that increases respiration but reduces ATP-synthesis to promote apoptosis in colon cancer models. *Oncotarget*. 2015; 6 (14): 11806.
- (111) Ussakli CH, Ebaee A, Binkley J, Brentnall TA, Emond MJ, Rabinovitch PS et al. Mitochondria and tumor progression in ulcerative colitis. *Journal of the National Cancer Institute*. 2013; 105 (16): 1239-1248.
- (112) Jones AW, Yao Z, Vicencio JM, Karkucinska-Wieckowska A, Szabadkai G. PGC-1 family coactivators and cell fate: Roles in cancer, neurodegeneration, cardiovascular disease and retrograde mitochondria-nucleus signalling. *Mitochondrion*. 2012; 12 (1): 86-99.
- (113) Shin SW, Yun SH, Park ES, Jeong JS, Kwak JY, Park JI. Overexpression of PGC-1 $\alpha$  enhances cell proliferation and tumorigenesis of HEK293 cells through the upregulation of Sp1 and Acyl-CoA binding protein. *International Journal of Oncology*. 2015; 46 (3): 1328-1342.
- (114) Seale P. Transcriptional regulatory circuits controlling brown fat development and activation. *Diabetes*. 2015; 64 (7): 2369-2375.

- 
- (115) López-Lluch G, Irusta PM, Navas P, de Cabo R. Mitochondrial biogenesis and healthy aging. *Experimental Gerontology*. 2008; 43 (9): 813-819.
- (116) Srivastava S, Barrett JN, Moraes CT. PGC-1 $\alpha/\beta$  upregulation is associated with improved oxidative phosphorylation in cells harboring nonsense mtDNA mutations. *Human Molecular Genetics*. 2007; 16 (8): 993-1005.
- (117) Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Rodgers M et al. Guidance on the conduct of narrative synthesis in systematic reviews. *A Product from the ESRC Methods Programme Version*. 2006; 1 : b92.
- (118) Urrútia G, Bonfill X. Declaración PRISMA: una propuesta para mejorar la publicación de revisiones sistemáticas y metaanálisis. *Medicina Clínica*. 2010; 135 (11): 507-511.
- (119) Moher D, Liberati A, Tetzlaff J, Altman D. The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine*. 2009; 6 (7): 264-9.
- (120) Kong X, Fan H, Liu X, Wang R, Liang J, Gupta N et al. Peroxisome proliferator-activated receptor  $\gamma$  coactivator-1 $\alpha$  enhances antiproliferative activity of 5-deoxy-5-fluorouridine in cancer cells through induction of uridine phosphorylase. *Molecular Pharmacology*. 2009; 76 (4): 854-860.
- (121) Bellafante E, Morgano A, Salvatore L, Murzilli S, Di Tullio G, DÓrazio A et al. PGC-1 $\beta$  promotes enterocyte lifespan and tumorigenesis in the intestine. *Proceedings of the National Academy of Sciences*. 2014; 111 (42): E4523-E4531.
- (122) Fisher KW, Das B, Kim HS, Clymer BK, Gehring D, Smith DR et al. AMPK promotes aberrant PGC1 $\beta$  expression to support human colon tumor cell survival. *Molecular and Cellular Biology*. 2015; 35 (22): 3866-3879.
- (123) D'Errico I, Lo Sasso G, Salvatore L, Murzilli S, Martelli N, Cristofaro M et al. Bax is necessary for PGC1 $\alpha$  pro-apoptotic effect in colorectal cancer cells. *Cell Cycle*. 2011; 10 (17): 2937-2945.
- (124) Vellinga TT, Borovski T, de Boer VC, Fatrai S, van Schelven S, Trumpi K et al. SIRT1/PGC1 $\alpha$  dependent increase in oxidative phosphorylation supports chemotherapy resistance of colon cancer. *Clinical Cancer Research*. 2015; 21 (12): 2870-2879.
- (125) Gaustadnes M, Ørntoft TF, Jensen JL, Topping N. Validation of the use of DNA pools and primer extension in association studies of sporadic colorectal cancer for selection of candidate SNPs. *Human Mutation*. 2006; 27 (2): 187-194.
- (126) Feilchenfeldt J, Bründler MA, Soravia C, Tótsch M, Meier CA. Peroxisome proliferator-activated receptors (PPARs) and associated transcription factors
-

## BIBLIOGRAFÍA

---

- in colon cancer: reduced expression of PPAR $\gamma$ -coactivator 1 (PGC-1). *Cancer Letters*. 2004; 203 (1): 25-33.
- (127) Wu F, Wang J, Wang Y, Kwok TT, Kong SK, Wong C. Estrogen-related receptor  $\alpha$  (ERR $\alpha$ ) inverse agonist XCT-790 induces cell death in chemotherapeutic resistant cancer cells. *Chemico-Biological Interactions*. 2009; 181 (2): 236-242.
- (128) Do MT, Kim HG, Choi JH, Jeong HG. Metformin induces microRNA-34a to downregulate the Sirt1/Pgc-1 $\alpha$ /Nrf2 pathway, leading to increased susceptibility of wild-type p53 cancer cells to oxidative stress and therapeutic agents. *Free Radical Biology and Medicine*. 2014; 74 : 21-34.
- (129) DeCensi A, Puntoni M, Goodwin P, Cazzaniga M, Gennari A, Bonanni B et al. Metformin and cancer risk in diabetic patients: a systematic review and meta-analysis. *Cancer Prevention Research*. 2010; 3 (11): 1451-1461.
- (130) Hernández-Romano J, Martínez-Barnetche J, Valverde-Garduño V. Polimorfismos reguladores y su participación en la patogenia de enfermedades complejas en la era posgenómica. *Salud Pública de México*. 2009; 51 : s455-s462.
- (131) Botstein D, Risch N. Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. *Nature Genetics*. 2003; 33 : 228.
- (132) Drake JA, Bird C, Nemesh J, Thomas DJ, Newton-Cheh C, Reymond A et al. Conserved noncoding sequences are selectively constrained and not mutation cold spots. *Nature Genetics*. 2006; 38 (2): 223.
- (133) Levine M, Tjian R. Transcription regulation and animal diversity. *Nature*. 2003; 424 (6945): 147.
- (134) National Center for Biotechnology Information Search database. *PPARGC1B PPARG coactivator 1 beta [Homo sapiens (human)]* - Gene - NCBI. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/133522%20http://www.ncbi.nlm.nih.gov/gene/133522> [Visitado el 04-05-2017].
- (135) Purcell S. PLINK (1.07) Documentation. 2010; 283.
- (136) The ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature*. 2012; 489 (7414): 57.
- (137) National Center for Biotechnology Information Search database. *Reference SNP (refSNP) Cluster Report: rs12505641*. Disponible en: [https://www.ncbi.nlm.nih.gov/projects/SNP/snp%7B%5C\\_%7Dref.cgi?rs=12505641](https://www.ncbi.nlm.nih.gov/projects/SNP/snp%7B%5C_%7Dref.cgi?rs=12505641) [Visitado el 11-01-2016].
- (138) National Center for Biotechnology Information Search database. *CDC42P6 cell division cycle 42 pseudogene 6 [Homo sapiens (human)]*. Disponible

- en: <https://www.ncbi.nlm.nih.gov/gene/?term=CDC42P6> [Visitado el 11-01-2016].
- (139) Gonçalves I, Duret L, Mouchiroud D. Nature and structure of human genes that generate retropseudogenes. *Genome Research*. 2000; 10 (5): 672-678.
- (140) Augenlicht LH, Heerdt BG. Mitochondria: integrators in tumorigenesis? *Nature Genetics*. 2001; 28 (2): 104.
- (141) Alonso-Molero J, González-Donquiles C, Fernández-Villa T, de Souza-Teixeira F, Vilorio-Marqués L, Molina AJ et al. Alterations in PGC1 $\alpha$  expression levels are involved in colorectal cancer risk: a qualitative systematic review. *BioMed Central Cancer*. 2017; 17 (1): 731.
- (142) Carvalho FI, Johns C, Gillespie ME. Gene expression analysis. *Biochemistry and Molecular Biology Education*. 2012; 40 (3): 181-190.
- (143) Auclair JR, Somasundaran M, Green KM, Evans JE, Schiffer CA, Ringe D et al. Mass spectrometry tools for analysis of intermolecular interactions. *Methods in Molecular Biology*. 2012; 896 (22): 387-398.
- (144) De Souza-Teixeira F, Alonso-Molero J, Ayán C, Vilorio-Marques L, Molina AJ, González-Donquiles C et al. PGC-1 $\alpha$  as a Biomarker of Physical Activity-Protective Effect on Colorectal Cancer. *Cancer Prevention Research*. 2018;
- (145) Lieber CS, Leo MA, Wang X, DeCarli LM. Alcohol alters hepatic FoxO1, p53, and mitochondrial SIRT5 deacetylation function. *Biochemical and Biophysical Research Communications*. 2008; 373 (2): 246-252.
- (146) Mooney MA, Nigg JT, McWeeney SK, Wilmot B. Functional and genomic context in pathway analysis of GWAS data. *Trends in Genetics*. 2014; 30 (9): 390-400.
- (147) Poole W, Gibbs DL, Shmulevich I, Bernard B, Knijnenburg TA. Combining dependent P-values with an empirical adaptation of Brown's method. *Bioinformatics*. 2016; 32 (17): i430-i436.
- (148) Kwon Js, Kim J, Nam D, Kim S. Performance comparison of two gene set analysis methods for genome-wide association study results: GSA-SNP vs i-GSEA4GWAS. *Genomics & Informatics*. 2012; 10 (2): 123-127.
- (149) Váremo L, Nielsen J, Nookaew I. Enriching the gene set analysis of genome-wide data by incorporating directionality of gene expression and combining statistical hypotheses and methods. *Nucleic Acids Research*. 2013; 41 (8): 4378-4391.
- (150) Sonachalam M, Shen J, Huang H, Wu X. Systems biology approach to identify gene network signatures for colorectal cancer. *Frontiers in Genetics*. 2012; 3 : 80.

## BIBLIOGRAFÍA

---

- (151) Lascorz J, Chen B, Hemminki K, Försti A. Consensus pathways implicated in prognosis of colorectal cancer identified through systematic enrichment analysis of gene expression profiling studies. *PLoS One*. 2011; 6 (4): e18867.
- (152) Quan B, Qi X, Yu Z, Jiang Y, Liao M, Wang G et al. Pathway analysis of genome-wide association study and transcriptome data highlights new biological pathways in colorectal cancer. *Molecular Genetics and Genomics*. 2015; 290 (2): 603-610.
- (153) Chen LS, Hutter CM, Potter JD, Liu Y, Prentice RL, Peters U et al. Insights into colon cancer etiology via a regularized approach to gene set analysis of GWAS data. *The American Journal of Human Genetics*. 2010; 86 (6): 860-871.
- (154) Li X, Shen L, Shang X, Liu W. Subpathway analysis based on signaling-pathway impact analysis of signaling pathway. *PLoS One*. 2015; 10 (7): e0132813.
- (155) Kou Y, Zhang S, Chen X, Hu S. Gene expression profile analysis of colorectal cancer to investigate potential mechanisms using bioinformatics. *OncoTargets and Therapy*. 2015; 8 : 745.
- (156) Lu X, Cao M, Han S, Yang Y, Zhou J. Colorectal cancer risk genes are functionally enriched in regulatory pathways. *Scientific Reports*. 2016; 6 : 25347.
- (157) Hong G, Zhang W, Li H, Shen X, Guo Z. Separate enrichment analysis of pathways for up- and downregulated genes. *Journal of the Royal Society Interface*. 2014; 11 (92): 20130950.
- (158) Altman DG, Bland JM. Statistics notes: Absence of evidence is not evidence of absence. *British Medical Journal*. 1995; 311 (7003): 485.
- (159) Khatri P, Sirota M, Butte AJ. Ten years of pathway analysis: current approaches and outstanding challenges. *PLoS Computational Biology*. 2012; 8 (2): e1002375.

## Anexos

---

# Anexo I

---




RESEARCH ARTICLE

Open Access



# Alterations in PGC1 $\alpha$ expression levels are involved in colorectal cancer risk: a qualitative systematic review

Jéssica Alonso-Molero<sup>1,2,5\*</sup> , Carmen González-Donquiles<sup>1</sup>, Tania Fernández-Villa<sup>1</sup>, Fernanda de Souza-Teixeira<sup>1,3</sup>, Laura Vilorio-Marqués<sup>1</sup>, Antonio J. Molina<sup>1</sup> and Vicente Martín<sup>1,4</sup>

## Abstract

**Background:** Colorectal cancer (CRC) is a major global public health problem and the second leading cause of cancer-related death. Mitochondrial dysfunction has long been suspected to be involved in this type of tumorigenesis, as supported by an accumulating body of research evidence. However, little is known about how mitochondrial alterations contribute to tumorigenesis. Mitochondrial biogenesis is a fundamental cellular process required to maintain functional mitochondria and as an adaptive mechanism in response to changing energy requirements. Mitochondrial biogenesis is regulated by peroxisome proliferator-activated receptor gamma coactivator 1- $\alpha$  (PPARGC1A or PGC1 $\alpha$ ). In this paper, we report a systematic review to summarize current evidence on the role of PGC1 $\alpha$  in the initiation and progression of CRC. The aim is to provide a basis for more comprehensive research.

**Methods:** The literature search, data extraction and quality assessment were performed according to the document Guidance on the Conduct of Narrative Synthesis in Systematic Reviews and the PRISMA declaration.

**Results:** The studies included in this review aimed to evaluate whether increased or decreased PGC1 $\alpha$  expression affects the development of CRC. Each article proposes a possible molecular mechanism of action and we create two concept maps.

**Conclusion:** Our systematic review indicates that altered expression of PGC1 $\alpha$  modifies CRC risk. Most studies showed that overexpression of this gene increases CRC risk, while some studies indicated that lower than normal expression levels could increase CRC risk. Thus, various authors propose PGC1 $\alpha$  as a good candidate molecular target for cancer therapy. Reducing expression of this gene could help to reduce risk or progression of CRC.

**Keywords:** PGC1 $\alpha$  or PPARGC1 $\alpha$ , Colorectal cancer (CRC), Signaling or metabolic pathways, Molecular mechanism

## Background

Colorectal cancer (CRC) is a major global public health problem and the second leading cause of cancer-related death. It is the third most commonly diagnosed cancer in men and the second in women [1]. CRC is the seventh and fourth most common cause of death and loss of life expectancy in Western Europe, respectively, and is associated with an elevated consumption of resources [2, 3].

Mitochondrial dysfunction has long been suspected to be involved in this type of tumorigenesis, as supported by an accumulating body of research evidence. However, little is known about how mitochondrial alterations contribute to tumorigenesis [4–7]. Mitochondrial biogenesis is a fundamental cellular process required to maintain functional mitochondria and as an adaptive mechanism in response to changing energy requirements [8]. Both endogenous and exogenous factors, as well as numerous signaling pathways and gene expression patterns, converge upon the mitochondrial biogenesis process to coordinate the energy needs of cells, tissues and the entire organism [4, 8, 9].

\* Correspondence: [jalonm00@estudiantes.unileon.es](mailto:jalonm00@estudiantes.unileon.es)

<sup>1</sup>Grupo de Investigación en Interacciones Gen-Ambiente y Salud, Universidad de León, León, Spain

<sup>2</sup>Universidad de Cantabria, Santander, Spain

Full list of author information is available at the end of the article



The master regulator of mitochondrial biogenesis is peroxisome proliferator-activated receptor gamma co-activator 1- $\alpha$  (PPARGC1A or PGC1 $\alpha$ ), because it controls production of mitochondrial proteins [10]. This gene is a transcriptional coactivator of the PGC-1 (peroxisome proliferator-activated receptor gamma co-activator 1) gene family, which has three known members, PGC1 $\alpha$ , PGC1 $\beta$  and PRC (PGC-1 related coactivator). PGC1 $\alpha$  and PGC1 $\beta$  are expressed in tissues with high energy demand, while PRC is expressed ubiquitously [10, 11]. While all three members of this family are potent regulators of mitochondrial function and biogenesis, PGC1 $\alpha$  is the most widely studied, and the other two are less well characterized [8, 11–13]. However, [11, 14].

PGC1 $\alpha$  acts as a master regulator of energy metabolism and mitochondrial biogenesis by integrating and coordinating the activity of other transcription factors, such as Nuclear respiratory factor 1, Nuclear factor 2, PPAR $\alpha$  (peroxisome proliferator-activated receptors  $\alpha$ ) and Mitochondrial transcription factor A [15]. Various endogenous and exogenous factors also regulate mitochondrial biogenesis through this gene [4, 9]. In addition, expression levels of PGC1 $\alpha$  appear to be directly related to mitochondrial biogenesis activity. As a multi-response factor, many agents and events regulate PGC1 $\alpha$  expression via multiple intracellular mediators [16].

Several mutations in nuclear and mitochondrial genes encoding for mitochondrial components have been reported to be associated with increased cancer risk [17], and mitochondrial loss is known to precede the development of dysplasia [6]. We believe that there is a clear relationship between CRC and PGC1 $\alpha$ ; however, the role of mitochondria and PGC1 $\alpha$  in CRC is poorly understood at present.

Several studies suggest that PGC1 $\alpha$  and related genes can regulate different pathways, such as mitochondrial biogenesis, antioxidant systems, reactive oxygen species, de novo lipid synthesis, and glycolysis, there by playing a role in risk for and development of CRC (See discussion) [10, 11]. Although there is no widely accepted mechanism to explain how PGC1 $\alpha$  is involved in human CRC, it is essential to understand this mechanism in order to reduce CRC risk, as well as the development of novel therapeutic tools to treat tumors and to support measures to reduce CRC risk.

In this paper, we report a systematic review to summarize current evidence on the role of PGC1 $\alpha$  in the initiation and progression of CRC. Since there is a limited evidence on this specific question, we used a broad, inclusive search strategy, with the aim of providing a basis for more comprehensive investigation.

## Methods

### Search strategy

The literature search, data extraction and quality assessment were performed according to the document Guidance on the Conduct of Narrative Synthesis in Systematic Reviews [18] and the PRISMA declaration [19]. This search was made between June and September of 2016.

We conducted a web-based search of 8 databases (Cochrane library, PubMed, Scopus, Web of Science, PsycINFO, Scielo, PLoS One and PubMedCentral (PMC)), using the following search terms: "Colorectal OR Colon OR Rectum OR Rectal" and "Cancer OR Carcinoma OR Tumor OR Tumour OR Neoplasm OR Cancer Cells" and "Mitochondrial Biogenesis OR Mitochondrial dysfunction OR Mitochondria OR Mitochondrion" and "Warburg effect" and "OXPHOS OR Oxidative phosphorylation OR Anaerobic glycolysis" and "PGC1A OR PPARGC1A OR Peroxisome Proliferator-activated Receptor gamma coactivator 1 alpha". The search was restricted to English language articles. We used this quite broad search strategy because of the paucity of research on this specific issue.

### Selection process

The list of articles obtained by this search was manually screened to identify relevant articles. We first read the titles and removed irrelevant articles, and then read the abstracts to eliminate those not directly related to the objective of this review. We imported the resulting set of articles into a reference management program (Endnote), which allowed us to detect duplicate articles. Finally, we read the full text and decided if the article should be included in this review according to the exclusion criteria described below.

### Study selection

We applied the following exclusion criteria during each stage of the selection process mentioned above:

1) Title – We eliminated articles that did not deal with cancer or inflammatory bowel diseases (e.g. Ulcerative Colitis or Crohn's disease, see below). We also excluded articles dealing with genes other than PGC1 $\alpha$  or PGC1 $\beta$  (there is evidence that both have a similar role in the organism).

2) Abstract – We eliminated articles that did not directly deal with CRC or inflammatory bowel diseases that ultimately develop into CRC, or did not deal with some isoform of PGC1 ( $\alpha$  or  $\beta$ ), or with related pathways.

3) Full-text – i) We included articles on basic research such as with cell lines or animals. For human studies, we included both basic research, and population-based observational studies. ii) We included articles on colorectal cancer, as well as those on other diseases, such as

ulcerative colitis, where these ultimately deal with CRC. We did not consider demographic factors during study selection because we considered a broad range of study types, including basic research.

4) The online search was replicated independently by two reviewers, who reviewed and filtered the titles, and created a draft list of titles. This list of articles to be included was agreed upon by three people (with the first author) based on the exclusion/inclusion criteria. The reviewers then independently reviewed the full text of the articles according to the inclusion criteria. Discrepancies were discussed and resolved in collaboration with the principal investigator.

#### **Preliminary synthesis**

We used tabulation and visual representations of data to reduce studies to their key characteristics, which could be important for understanding the objective of this review.

#### **Relationship between papers**

##### ***Evaluating heterogeneity***

In general terms, this technique focused on the characteristics of the various studies and their potential relationship with the findings. The following characteristics were assessed:

- Cells lines (human or animals = 1)
- Colorectal cancer (Inflammatory bowel diseases as previous disease = 1)
- PGC1 $\alpha$  or PGC1 $\beta$  expression
- Relationship with other genes
- reactive oxygen species
- Mitochondrial biogenesis
- Chemotherapy

These characteristics were scored as 1 if they were present and 0 if absent. We took the sum of these values to obtain a picture of the heterogeneity of the articles in this review. Using these values, we created a comparison graph to evaluate the shared features.

##### ***Idea webbing and concept mapping [18]***

Idea webbing is a method for conceptualizing and exploring connections among the findings reported by the studies included in the review (data not shown). Using idea webbing we obtained a concept map, a visual picture linking multiple pieces of evidence across several studies. The aim was to construct a model of key concepts related to PGC1 $\alpha$  and PGC1 $\beta$ , and to represent the relationships between these and the development of CRC.

#### ***Checking the synthesis with authors of primary studies***

We compared our results to those of other systematic reviews to support our ideas.

## **Results**

### **Literature search results**

Using the search terms described in the Methods section, but excluding the Boolean operators, we identified 7688 manuscripts from our search of PubMed, Scopus, Web Of Science (WOS), PsycInfo, Cochrane, Scielo and PLoSOne (no results were returned by PsycInfo or Cochrane). When the Boolean operators were included, 214 articles were returned (Fig. 1). We retained and analyzed 34 papers that met the eligibility criteria described in the Methods section. Of these 34 abstracts, 15 full-text articles were retrieved for detailed evaluation and 12 studies were included in the final analysis. Figure 1 illustrates the article screening and selection process.

Eighteen articles were excluded for the following reasons: i) Article mentions other genes related to PGC1 $\alpha$ , but does not deal with PGC1 $\alpha$  itself. ii) Article mentions PGC1 $\alpha$ , but does not deal with CRC. iii) Article mentions PGC1 $\alpha$  and its relationship with cancer, but does not deal with CRC. iv) Article mentions PGC1 $\alpha$  and its relationship with inflammatory bowel diseases, but does not deal with CRC. v) Article mentions PGC1 $\alpha$  and colorectal cancer but not define a relationship between them. A full summary of these data is shown in Additional file 1: Table S1 online.

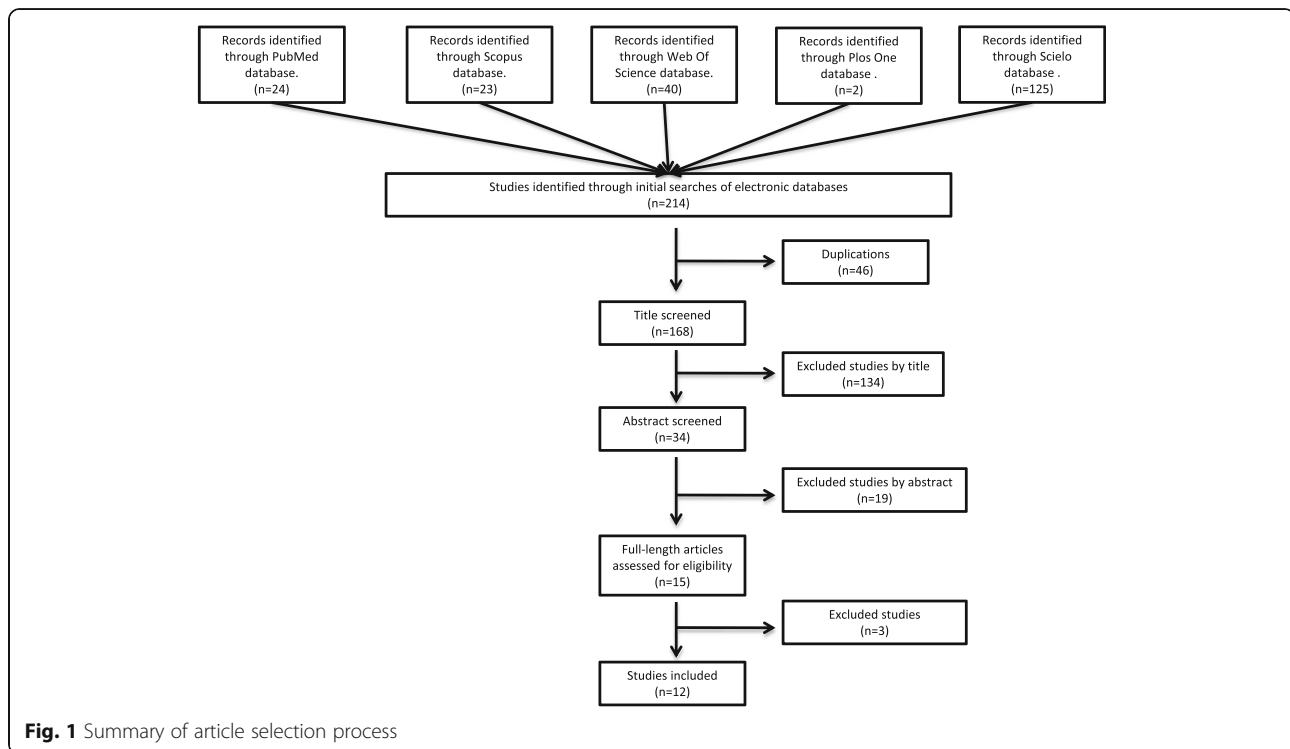
#### **Preliminary synthesis**

We chose tabulation to visually represent the data, with the aim of reducing studies to the key characteristics that could be important for understanding the relationship between PGC1 $\alpha$  and CRC. The key message observed in these articles was that patterns of altered PGC1 $\alpha$  expression affect risk and development of CRC via various molecular mechanisms, including mitochondrial biogenesis, antioxidant systems, reactive oxygen species, de novo lipid synthesis, glycolysis and alterations in the expression of other genes. Most of this information was obtained from basic research, such as cell culture, since there is little information about this field. These data are summarized in Table 1, and full details are provided in Additional file 1: Table S2 online. In addition, Table 2 summarizes the original material used in each of these articles.

#### **Relationships between papers**

##### ***Heterogeneity***

We analyzed differences in relevant characteristics between the selected articles studied using the graph shown in Fig. 2.



We assessed the qualitative outcomes of the 12 studies in terms of 6 features: study of CRC; assessment of PGC1 $\alpha$  or PGC1 $\beta$  (which should be present in all papers); relationship between PGC1 $\alpha$  and another gene; test for presence of reactive oxidative species; study of mitochondrial biogenesis; study of chemotherapy. Four papers (30.8%) had a maximum score of five points, and another 4 articles (30.8%) had a score of four points. One study (7.6%) had three points, two (15.4%) had two points, and another two (15.4%) had six points. These proportions show that the studies were sufficiently homogeneous to perform a systematic narrative review.

### Concept mapping

Based on the information in Table 1, we performed idea webbing for each article, from which we obtained two concept maps. The first concept map is based on D'Errico et al. [20], whose aim was to show that PGC1 $\alpha$  is highly expressed on the surface of the intestinal epithelium but is poorly expressed in the crypts, and is also reduced in intestinal tumors. D'Errico et al. analyzed the expression and function of PGC1 $\alpha$  along the crypt-to-villus axis under normal conditions, and observed that PGC1 $\alpha$  is poorly expressed in the proliferative compartment at the bottom of the crypts, but, conversely, is highly expressed at the villus tips, promoting mitochondria-mediated apoptosis via the accumulation of reactive oxygen species (Fig. 3a).

These authors also identified that overexpression of PGC1 $\alpha$  in human colon cancer cells (HT29) activates metabolic changes such as mitochondrial activation, which produce a proapoptotic effect via reactive species oxygen accumulation (Fig. 3b). Thus, they conclude that PGC1 $\alpha$  is a metabolic regulator of intestinal cell fate and protects against tumorigenesis.

The second concept map highlights patterns of altered gene expression in the development of colorectal cancer. Most papers included in this concept map suggest that PGC1 $\alpha$  expression is increased during the development of CRC, while reduced PGC1 $\alpha$  expression reduces risk and progression of this disease. Nonetheless, some studies have reported the opposite, that reduced PGC1 $\alpha$  expression levels can also increase cancer risk (Fig. 4).

Figure 4 shows various possible mechanisms of action of PGC1 $\alpha$  in relation to CRC development, and suggests four different mechanisms via which this gene can become overexpressed. First, high expression of PGC1 $\alpha$  seems to increase Antioxidant systems, which reduces the number of reactive oxygen species, resulting in inhibition of apoptosis. Second, overexpression of PGC1 $\alpha$  induces the mitochondrial biogenesis pathway, increasing cellular growth (proliferation). Third, overexpression of PGC1 $\alpha$  induces glucose uptake, increasing cell proliferation. Fourth, high levels of PGC1 $\alpha$  could reduce glycolysis and increase oxidative phosphorylation, which increases cells' resistance to chemotherapy. In the map, we can observe certain factors that trigger overexpression of PGC1 $\alpha$ , such as the

**Table 1** Summary of the key characteristics of the selected studies

Title	First Author	Year	Primary results	Conclusion
Overexpression of PGC1-alpha enhances cell proliferation and tumorigenesis of HEK293 cells through the upregulation of Sp1 and Acyl-CoA binding protein.	Sung-Won Shin	2014	1) PGC-1 $\alpha$ accelerates proliferation of HEK293 and CT-26 cells. 2) Knockdown of PGC1- $\alpha$ expression results in decreased cell proliferation of human colorectal cancer cells. 3) PGC-1 $\alpha$ promotes the oncogenic potential of HEK293 cells. 4) PGC-1 $\alpha$ overexpressing HEK293 cells have decreased sensitivity to oxidative stress.	PGC-1 $\alpha$ overexpression upregulates proliferation of HEK293 and CT26 cells. In addition, this expression correlates with enhanced tumorigenesis. Moreover, PGC-1 $\alpha$ siRNA transfection resulted in decreased cell proliferation. Further studies to clarify the molecular interactions are needed.
PGC-1 $\beta$ promotes enterocyte lifespan and tumorigenesis in the intestine	Elena Bellafante	2014	1) PGC-1 $\beta$ Is Highly Expressed in the Intestinal Epithelium and Modulates Intestinal Morphology. 2) Intestinal PGC-1 $\beta$ Overexpression Enhances Antioxidant Defense. 3) Intestinal PGC-1 $\beta$ Overexpression Promotes Intestinal Carcinogenesis.	PGC-1 $\beta$ seems to act as an adaptive self-point regulator, capable of providing a balance between mitochondrial activity and production of increased reactive oxygen species.
Mitochondria and Tumor Progression in Ulcerative Colitis	Cigdem Himmetoglu Ussakli	2013	1) Comparison of COX in No dysplastic Biopsies of UC Progressors and Nonprogressors show that tumor development could be due to previous lower COX levels. 2) Cox levels increase after the tumor development (Bimodal pattern). Mitochondria follow the same pattern. 3) PGC1 $\alpha$ may drive of the mitochondrial changes observed.	1) At the biomarker level, COX loss precedes tumor progression in UC. 2) At the biological level, the loss of COX represents a reduction in the number of mitochondria in preneoplasia, which is restored in cancers. It appears to be driven by PGC1 $\alpha$ .
PGC1 $\alpha$ promotes tumor growth by inducing gene expression programs supporting lipogenesis.	Kavita Bhalla	2011	1) Loss of PGC1 protects against both colon and liver tumorigenesis. 2) Overexpression PGC1 $\alpha$ promotes tumor growth in vivo. 3) PGC1 $\alpha$ mediated induction of fatty acid synthesis promotes tumor growth.	1) Novel role for PGC1 $\alpha$ in promoting carcinogenesis and tumor growth. 2) PGC1 $\alpha$ coordinates the induction of a gene expression program that facilitates the conversion of glucose to fatty acids. 3) PGC1 $\alpha$ is a potential therapeutic target for chemoprevention.
Bax is necessary for PGC1 $\alpha$ pro-apoptotic effect in colorectal cancer cells	Ilenia D'Errico	2011	1) PGC1 $\alpha$ induces Bax activation. 2) PGC1 $\alpha$ increases mitochondrial activity. 3) PGC1 $\alpha$ induces apoptosis in the presence of Bax, but not without Bax. 4) PGC1 $\alpha$ inhibits tumor growth in presence of Bax.	1) In the presence of Bax, the PGC1 $\alpha$ -induced accumulation of reactive oxygen species is one of the main apoptosis-driving factors in CRC cells. 2) PGC1 $\alpha$ is able to induce Bax activation and translocation to mitochondria, thus leading to apoptotic cascade.
PGC-1 $\alpha$ / $\beta$ upregulation is associated with improved oxidative phosphorylation in cells harboring nonsense mtDNA mutations	Sarika Srivastava	2007	1) PGC-1 $\alpha$ and PGC-1 $\beta$ are markedly upregulated in V425. 2) Overexpression of PGC-1 $\alpha$ and PGC-1 $\beta$ transcriptional coactivators stimulates mitochondrial respiration, at least in osteosarcoma cybrids. 3) Overexpression of PGC-1 $\alpha$ stimulates complex IV activity. 4) Overexpression of PGC-1 $\alpha$ / $\beta$ transcriptional coactivators can stimulate respiration in oxidative phosphorylation-deficient cells.	1) In V425 cells, the Ca <sup>2+</sup> -dependent signaling events are active for relatively longer periods, which in turn might activate the nuclear genes (including PGC-1 $\alpha$ / $\beta$ ) involved in tumor invasion and metastasis. 2) Overexpression of PGC-1 $\alpha$ / $\beta$ can stimulate respiration in oxidative phosphorylation deficient cells. 3) This pathway could be explored as a therapeutic approach for the treatment of human mitochondrial diseases.
Validation of the Use of DNA Pools and Primer Extension in Association Studies of Sporadic Colorectal Cancer for Selection of Candidate SNPs	Mette Gaustadnes	2006	Results were analyzed using the $\chi^2$ test with a level of significance $\alpha = 0.05$ . Five SNPs were found. The SNP analysis of the (*604517)3'utr96516 was not reproducible, but it was always statistically significant.	The results of this article allow us to conclude that the difference between cases and controls would be statistically significant for $n = 600$ cases and $n = 600$ controls.

**Table 1** Summary of the key characteristics of the selected studies (*Continued*)

Title	First Author	Year	Primary results	Conclusion
SIRT1/PGC1 $\alpha$ -Dependent Increase in Oxidative Phosphorylation Supports Chemotherapy Resistance of Colon Cancer	Thomas T.Vellinga	2015	1) Chemotherapy induces SIRT1 to promote oxidative energy metabolism. This gene controls mitochondrial biogenesis by deacetylation and activation of PGC1 $\alpha$ . 2) SIRT1 and PGC1 $\alpha$ protect colon cancer cells against chemotherapy.	Colorectal tumors shift their energy metabolism when challenged with chemotherapy. Chemotherapy induces oxidative phosphorylation in colon cancer cells via the SIRT1/PGC1 $\alpha$ axis to help them survive treatment.
AMPK Promotes Aberrant PGC1 $\beta$ Expression To Support Human Colon Tumor Cell Survival	Kurt W. Fisher	2015	PGC1 $\alpha$ is not detected in HCT116 cell line. 1) PGC1 $\beta$ and ERR $\alpha$ are key downstream effectors of K-Ras, KSR1, and AMPK1. 2) Both AMPK 1 and K-Ras depletion decreased the protein levels of PGC1 $\beta$ . 3) PGC1 $\beta$ and ERR $\alpha$ are overexpressed in colon cancer and are required for colon cancer survival both in vivo and in vitro.	The aberrant expression of PGC1 $\beta$ and ERR $\alpha$ that persists in additional tumors with oncogenic Ras alleles will reveal the importance of these transcriptional regulators in creating tumor cells and promoting their survival. This may represent a new therapeutic target.
Peroxisome proliferator-activated receptor- $\gamma$ coactivator 1- $\alpha$ (PGC1 $\alpha$ ) is a metabolic regulator of intestinal epithelial cell fate	Ilenia D'Errico	2011	1) Expression level of PGC1 $\alpha$ in the intestine is higher in differentiated enterocytes than in the proliferative compartment at the bottom of the crypts, where it has only a scattered expression. 2) PGC1 $\alpha$ Induces Mitochondrial Proliferation and Activation in Human Intestinal Cancer Cells. 3) PGC1 $\alpha$ induces tissue-specific accumulation of reactive oxygen species and apoptosis. 4) PGC1 $\alpha$ Stimulates Intestinal Mitochondrial Biogenesis and Respiration in vivo, and suppresses Colorectal Carcinogenesis	1) PGC1 $\alpha$ expression levels could influence intestinal epithelial cell fate by inducing mitochondrial-related metabolic modifications that induce apoptosis. 2) PGC1 $\alpha$ overexpression stimulates mitochondrial biogenesis, metabolic activities and accumulation of reactive oxygen species. 3) In tissues with high aerobic energy demand, PGC1 $\alpha$ preserves reactive oxygen species'homeostasis; In normal intestine, PGC1 $\alpha$ cannot induce reactive oxygen species scavenging systems.
Peroxisome Proliferator-Activated Receptor Coactivator-1alpha Enhances Antiproliferative Activity of 5'-Deoxy-5-Fluorouridine in Cancer Cells through Induction of Uridine Phosphorylase	Xingxing Kong	2009	1) PGC-1 Induces the Expression of UPase in Breast and Colon Cancer Cells. 2) PGC1 $\alpha$ -Dependent Induction of UPase Gene in Cancer Cells Is Mediated by ERR $\alpha$ . 3) Overexpression of PGC-1 Sensitizes Cancer Cells to 5 -DFUR.	1) PGC-1seems to be a regulator of UPase gene transcription, whose effect is mediated by ERR $\alpha$ . 2) PGC1 $\alpha$ has an effect on the absence of ERR $\alpha$ , suggesting the involvement of other regulatory factors. 3) In tumor cells, UPase catalyzes the transformation of 5 -DFUR to 5-FU, which inhibits their proliferation. In this way, PGC1 $\alpha$ enhances the cell's sensitivity to the treatment.
Peroxisome proliferator-activated receptors (PPARs) and associated transcription factors in colon cancer: reduced expression of PPAR $\gamma$ -coactivator 1 (PGC-1)	Jonas Feilchenfeldt	2004	1) RXR $\alpha$ expression in tumors is similar relative to normal mucosa. 2) PGC-1 expression in the tumors was significantly decreased relative to normal mucosa.	1) PPAR $\beta$ / $\delta$ may repress PPAR $\alpha$ and PPARY target gene expression. 2) Reduced coactivator levels of PGC-1 are compatible with reduced transcriptional activity of PPARY and hence reduced tumor suppressor activity. 3) Transcriptional activity of PPARY may not only be decreased by mutation and increased levels of the transcriptional repressor PPAR $\beta$ / $\delta$ but also by downregulation of coactivator PGC-1 of PPARY.

mitochondrial biogenesis pathway (via positive feedback), or chemotherapy.

In contrast, cancer development can also be promoted by silencing or downregulation of PGC1 $\alpha$ , also via two possible mechanisms. First, low levels of PGC1 $\alpha$  hinder the correct function of the mitochondrial biogenesis pathway, which promotes CRC growth. Second, low levels of PGC1 $\alpha$  reduce PPAR- $\gamma$ , which promotes CRC growth via an unknown process.

Figure 4 highlights some mechanisms that reduce CRC risk or development. In this case, reduction of PGC1 $\alpha$  levels is due to inhibition of KSR1 and AMPK- $\gamma$ 1 or XCT790 chemotherapy, via three possible mechanisms. First, lack of PGC1 $\alpha$  increases reactive oxygen species levels, which activates apoptosis. Second, reduced mitochondrial membrane potential ( $\Delta\Psi$ m) reduces resistance to chemotherapy. Third, low expression of PGC1 $\alpha$  seems to reduce ERR $\alpha$  expression, reducing

**Table 2** Cell lines and other materials, and characteristics of articles considered in this review

Title	Main Material	Characteristics
Overexpression of PGC-1 $\alpha$ enhances cell proliferation and tumorigenesis of HEK293 cells through the upregulation of Sp1 and Acyl-CoA binding protein	HT-29 (Cell line)	i) Human. ii) Epithelial
	SNU-C4 (Cell line)	i) Human
	CT-26 (Cell line)	i) Mouse. ii) Epithelial
PGC-1 $\beta$ promotes enterocyte lifespan and tumorigenesis in the intestine	iPGC1b mouse model with human PGC1b iPGC1b knockout mice iPGC1b Apc Min/+ Mice	i) Transgenic mouse
Mitochondria and Tumor Progression in Ulcerative Colitis	Ulcerative Colitis progressor (Human) Ulcerative Colitis non progressor (Human)	i) Human
PGC1 $\alpha$ promotes tumor growth by inducing gene expression programs supporting lipogenesis.	PGC1a knockout mice	i) Transgenic mouse
	PGC1a +/+ mice	
	SCID mice (HT29)	i) Inoculated mouse
	HT29 (Cell line)	i) Human. ii) Epithelial
Bax is necessary for PGC1 $\alpha$ pro-apoptotic effect in colorectal cancer cells	Colo205 (Cell line)	i) Human. ii) Epithelial
	HCT116 (Cell line) Nude mice (subcutaneously injected both cells)	i) Human ii) Epithelial i) Inoculated mouse
PGC-1 $\alpha$ / $\beta$ upregulation is associated with improved oxidative phosphorylation in cells harboring nonsense mtDNA mutations	VACO425 (Cell line)	i) Human
	VACO429 (Cell line)	i) Human
Validation of the Use of DNA Pools and Primer Extension in Association Studies of Sporadic Colorectal Cancer for Selection of Candidate SNPs	Two pools of genomic DNA (patients with sporadic CRC + controls)	i) Human
SIRT1/PGC1 $\alpha$ -Dependent Increase in Oxidative Phosphorylation Supports Chemotherapy Resistance of Colon Cancer	Colonosphere cultures shSIRT1 (Human colorectal tumor specimens)	i) Human colonosphere cultures
	Colonosphere cultures shPGC1 $\alpha$ (Human colorectal tumor specimens)	
AMPK Promotes Aberrant PGC1b Expression To Support Human Colon Tumor Cell Survival	HCT116 (Cell line)	i) Human. ii) Epithelial
	Inmunodeficient mice (HCT116 cells grafted into mice)	i) Inoculated mouse
Peroxisome proliferator-activated receptor- $\gamma$ coactivator 1- $\alpha$ (PGC1 $\alpha$ ) is a metabolic regulator of intestinal epithelial cell fate	HT29 (Cell line)	i) Human. ii) Epithelial
	HCT116 (Cell line)	i) Human. ii) Epithelial
	HT29p0 (Cell line: Completely lacks mitochondrial DNA)	i) Human. ii) Epithelial
	Xenograft mice using HT29 cells	i) Inoculated mouse
	iPGC1a transgenic mice	i) Transgenic mouse
	PGC1a +/+ mice and PGC1-/- mice	
Peroxisome Proliferator-Activated Receptor gamma Coactivator-1 Enhances Antiproliferative Activity of 5 -Deoxy-5-Fluorouridine in Cancer Cells through Induction of Uridine Phosphorylase	Colo320 (Cell line)	i) Human. ii) Undifferentiated
	HCT116 (Cell line)	i) Human. ii) Epithelial

**Table 2** Cell lines and other materials, and characteristics of articles considered in this review (Continued)

Title	Main Material	Characteristics
Peroxisome proliferator-activated receptors (PPARs) and associated transcription factors in colon cancer: reduced expression of PPARγ-coactivator 1 (PGC-1)	Colorectal cancers from patients (Human)	i) Human

CRC. In contrast, other mechanisms support the idea that PGC1α overexpression increases ERα expression, which may increase chemotherapy sensitivity.

These contrasting results could be due to the different starting material (e.g. cell lines, mice, etc.) in these studies. In Table 2 we summarize the material used in each article with the objective of trying to understand the cause of these differences. Note that we only indicate the main starting material, from which the authors obtained the principal results.

**Checking the synthesis with authors of primary studies**

We were unable to compare our results to those of other systematic reviews because no other reviews have dealt with this topic, as far as we are aware.

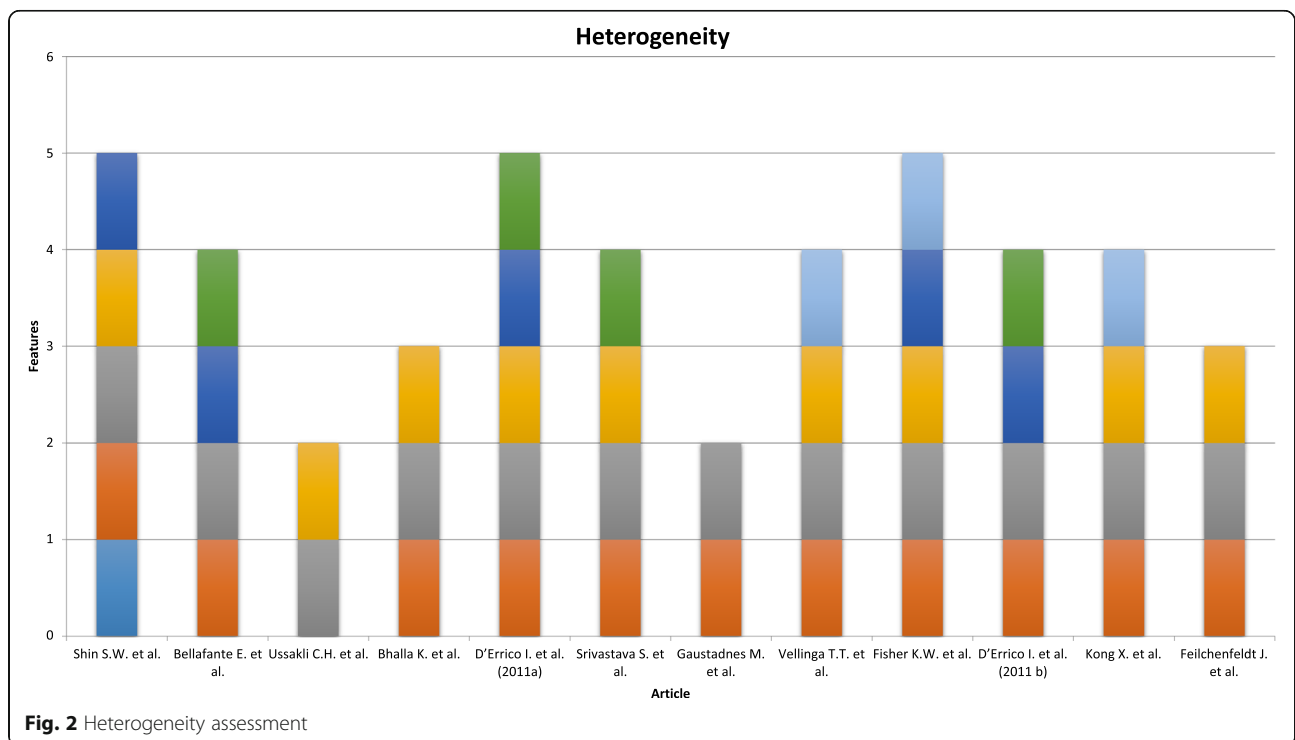
**Discussion**

The central question of this review is clear: summarize current evidence on the role of PGC1α in the initiation and progression of CRC. We tested the effect of different PGC1α expression levels during the development of colorectal cancer. There is little prior evidence on this

question, so only a small number of studies have been included in this review.

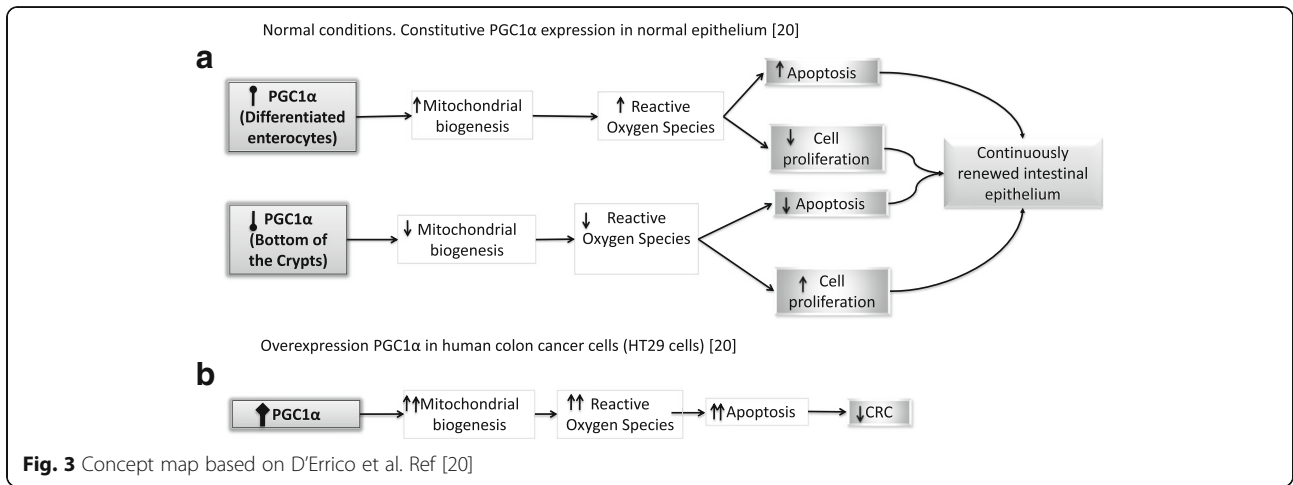
The studies included in this review aimed to evaluate whether an increase or decrease in PGC1α expression levels affects the development of CRC. Most studies (61.5%) were carried out using cell lines [10, 11, 20, 21, 23–25]: in four studies, cells were injected subcutaneously into the flanks of mice [20, 21, 23, 24], and two studies evaluated expression levels in mice in vivo [14, 21]. Other studies were carried out in samples of patients with disease [6, 13, 26, 27], with similar results to the in vitro studies.

Each article proposes a possible molecular mechanism of action. First, we try to understand the normal mechanism of PGC1α action proposed by Ilenia D’Errico et al. [20], which is derived from work on cell cultures and two groups of transgenic mice: iPGC1α transgenic mice and iPGC1α Apc<sup>Min/+</sup> mice (more information in Table 2). The authors conclude that PGC1α is a metabolic regulator of intestinal cell fate. In addition, PGC1α seems to have different functions in tissues with high aerobic energy metabolism, such as at the bottom of the crypts, where PGC1α expression increases production of



**Fig. 2** Heterogeneity assessment



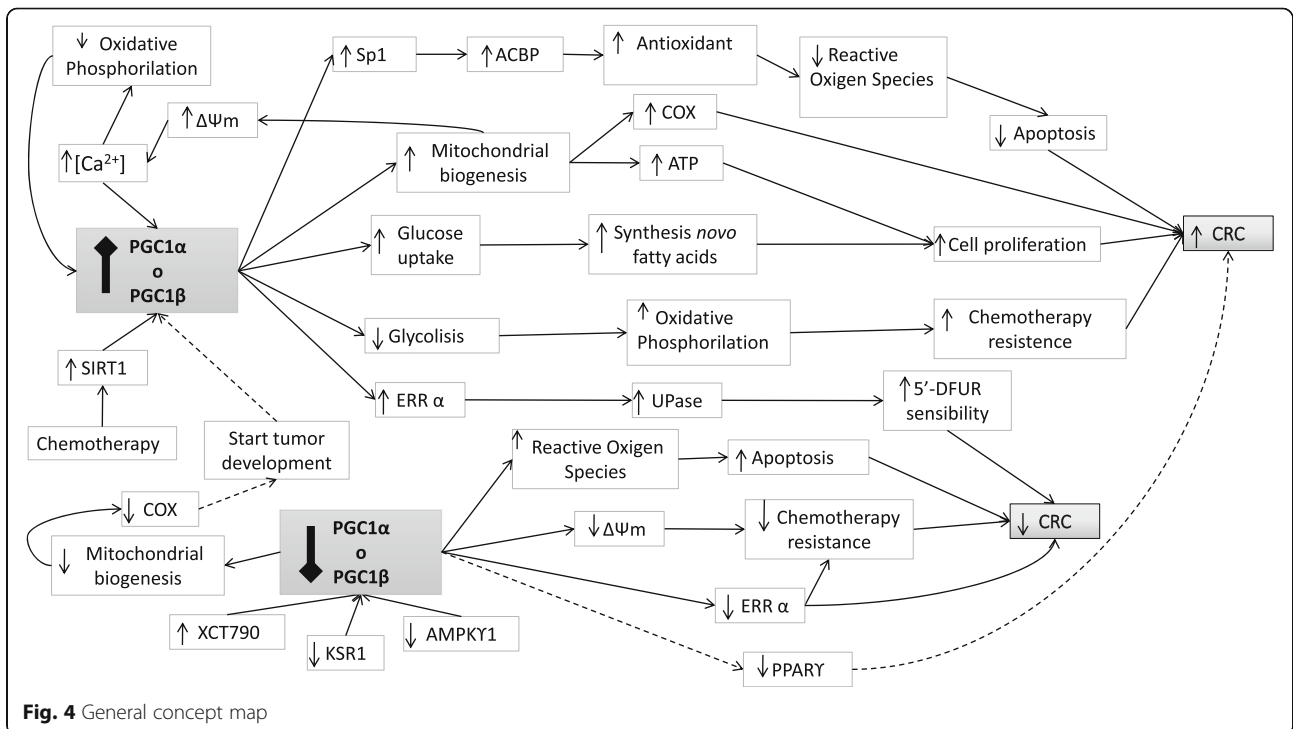


antioxidant enzymes that protect cells from reactive oxygen species. In contrast, at the top of the villi, where aerobic energy metabolism is low, PGC1α expression increases mitochondrial biogenesis, resulting in greater accumulation of reactive oxygen species than at the bottom of the crypts, which induces apoptosis [20]. In this way, there is a balance between cell proliferation and apoptosis under normal conditions.

Other mechanisms have been proposed on the basis of various studies of PGC1α overexpression under non-normal conditions, mainly in cell culture. One mechanism supports the idea that overexpression of PGC1α results in high levels of Sp1 (specificity protein 1), which is then thought to enhance expression of ACBP (Acyl-

CoA-binding protein) via ACBP's Sp1 binding site. ACBP upregulation increases cell proliferation and decreases sensitivity to H<sub>2</sub>O<sub>2</sub>-induced apoptosis [10, 14]. A second mechanism proposes that PGC1α is a key metabolic regulator of several aspects of glucose metabolism. PGC1α is thought to coordinate gene expression in metabolic pathways that convert glucose to fatty acids. In turn, fatty acid synthesis promotes tumor growth [21] and disrupts the balance between apoptosis and cell proliferation, promoting the development of colorectal cancer.

A third possible mechanism is based on overexpression of both PGC1α and PGC1β, which increases the electron transfer activity of the mitochondrial respiration



chain and augments mitochondrial biogenesis [11, 14]. Both processes partly disrupt  $\Delta\Psi_m$  and cytosolic calcium  $[Ca^{2+}]$  buffering ability, with two consequences: 1) Dysfunctional oxidative phosphorylation in mitochondria, and 2) stimulation of the  $Ca^{2+}$  signaling cascade, which in turn may activate genes involved in tumor invasion and metastasis. Both of these effects can upregulate PGC1 $\alpha$ , generating a positive feedback loop [11]. In addition, the increase in mitochondrial biogenesis seems to give the cells the necessary energy for increased longevity and cellular division, promoting tumor growth [11, 14].

From these three molecular mechanism we can conclude that PGC1 $\alpha$  and PGC1 $\beta$  allow cells to balance i) mitochondrial activity and cytotoxic protection in the production of reactive oxygen species, and ii) apoptosis and cell proliferation [10, 11, 14, 21]. In addition, we can observe that all studies in cell lines [10, 11, 21] and mice [14] support the idea that high expression of PGC1 $\alpha$  carries increased CRC risk. Note that while none of these articles work with the same cell line, the cells used are similar since the majority of cell lines are epithelial (Table 2). Shin et al. [10] use human colon cancer cells (HT29 and SNU-C4) and mouse colon cancer cells (CT-26). Srivastava et al. [11] use VACO425 and VACO429, another type of human colon cancer cell. Bhalla et al. [21] use other human colon adenocarcinoma cells, Colo205, as well as HT29, like Shin et al. Despite this, all of these studies arrived at the same conclusions.

However, Ilenia D'Errico [24] show that Bax is necessary for the pro-apoptotic effect of PGC1 $\alpha$  in colorectal cancer cells, which could be another molecular mechanism of action, although this is not consistent with other studies. In this study [24], PGC1 $\alpha$  overexpression seems to induce Bax translocation to mitochondria, and Bax-protein mediates the pro-apoptotic effect of PGC1 $\alpha$  [24]. In this case, the cell line used is HCT116 (human colon carcinoma, also epithelial cells) and they also use mice who were subcutaneously injected with these cells. Here there is some controversy: in absence of Bax, PGC1 $\alpha$  overexpression was not able to oppose tumor growth.

In addition to this research on overexpression of PGC1 $\alpha$ , other studies have focused on low expression, inhibition or silencing of PGC1 $\alpha$  and PGC1 $\beta$ . Using the same culture cells as Ilenia D'Errico [24], as well as other cells and in immunodeficient mice, Fisher et al. [23] showed that PGC1 $\beta$  is aberrantly expressed in human colon cell lines and tumors, and maintains ERR $\alpha$  levels, contributing to its tumorigenic properties. These authors also found that KSR1 (Kinase suppressor of Ras 1) and AMPK (AMP-activated protein kinase), which act upstream of the PGC1 $\beta$  promotor, were linked to the action of transcriptional regulators PGC1 $\beta$ /ERR $\alpha$  (estrogen-related receptor  $\alpha$ ). Reduced

KSR1 and AMPK expression also downregulated expression of PGC1 $\beta$  and ERR $\alpha$ , which inhibits the survival of colorectal cancer cells [23].

Based on all of this evidence, it seems that PGC1 $\alpha$ /PGC1 $\beta$  is overexpressed in human colon cell lines during tumor development, and this overexpression seems to be required for both cell proliferation and survival, and to reduce apoptosis [10, 11, 14, 21, 23]. Only one study obtained markedly different results despite using the same cell line as one of the above (HCT116), although this was based on Bax instead of PGC1 $\alpha$  [24]. Note that while all of these studies were performed using different cell lines with genetic changes, most of these are human epithelial cells (Table 2). For this reason, different results could not be attributed to the use of different cell lines. We consider that we cannot explain the origin of this differences.

Our systematic search returned two studies that analyzed the effects of chemotherapy on PGC1 $\alpha$  expression in cultured cells and colonosphere cultures (Table 2) [25, 26]. These studies suggest three other molecular mechanism of action of PGC1 $\alpha$ . 1) Chemotherapy of colorectal tumor cells induces a SIRT1/PGC1 $\alpha$ -dependent increase in oxidative phosphorylation that promotes tumor survival during treatment because chemotherapy induces a shift in tumor energy metabolism that protects tumor cells from cytotoxic damage [26]. Mechanistically, chemotherapy-induced DNA damage results in increased expression of SIRT1, which deacetylates and thereby activates PGC1 $\alpha$  as a transcriptional coactivator. PGC1 $\alpha$  acts in concert with several transcription factors to stimulate the expression of genes involved in mitochondrial biogenesis and respiration, resulting in increased oxidative phosphorylation and helping them survive treatment. In this case, the authors worked with human colonosphere cultures [26]. 2) PGC1 $\alpha$ -induced activation of Uridine phosphorylase (UPase) expression, which is mediated by an estrogen related receptor (ERR) binding site. Overexpression of PGC1 $\alpha$  via this mechanism sensitizes colon cancer cells to growth inhibition by 5-deoxy-5-fluorouridine, presumably by inducing apoptosis in tumor cells [25]. In this way, 3) ERR $\alpha$  is primarily thought to regulate energy homeostasis by interacting with PGC1 $\alpha$  or PGC1 $\beta$ . ERR $\alpha$  overexpression reduces sensitivity to chemotherapy, while inhibition of these genes reduces reactive oxygen species and  $\Delta\Psi_m$ , which increases sensitivity to chemotherapy [22, 23]. Here, Fisher et al. worked with HCT116 and other cell lines [23], but studied the effect of the chemotherapy on HepG2 [22].

These mechanisms support the conclusion mentioned above, that PGC1 $\alpha$  and PGC1 $\beta$  allow cells to balance i) mitochondrial activity and cytotoxic protection in the production of reactive oxygen species, and ii) apoptosis

and cell proliferation. In addition, they suggest that any therapy that reduces PGC1 $\alpha$  expression will increase cancer cells' sensitivity to chemotherapy. Thus, more and more authors support the idea of using PGC1 $\alpha$  as a potential target in cancer therapy [21–23, 25, 26]. In fact, Do et al. [28] reported a strategy to reduce PGC1 $\alpha$  levels to improve cancer therapy, though not in colon cancer cells. They showed for the first time that metformin, an insulin-lowering agent [29], induces miR-34a which reduces Sirt1 in wild-type p53 cancer cells, but does not occur in altered p53 cell lines. This fact was shown in HCT116 and MCF-7 cell lines among others. However, they only could show in an MCF-7 cell line (Breast cancer cells) that the reduction of Sirt1 involved reduced PGC1 $\alpha$  and, subsequently, reduced NRF2 and enhanced susceptibility to oxidative stress [28]. DeCensi et al. 2010 [29] showed using a meta-analysis that metformin is associated with decreased risk of cancer, including colon cancer, in diabetic patients compared with other treatments [29].

Our systematic search returned three articles based on human samples. First, Ussakli et al. [6] studied 9 non-dysplastic colon biopsies from Ulcerative Colitis (UC) patients with high-grade dysplasia or cancer and 9 dysplasia-free UC patients, and concluded that while the development of dysplasia is preceded by mitochondrial loss, mitochondria are restored in cancer cells, which suggests that they are needed for further proliferation. This bimodal pattern may be driven by transcriptional regulation of mitochondrial biogenesis by PGC1 $\alpha$  [6], which is consistent with the results of the cell culture studies described above [11, 14].

Two other papers propose the same idea, that overexpression of PGC1 $\alpha$  results in higher CRC risk, although these studies are not comparable with the one described above [6]. First, Gaustadnes et al. [27] screened a selection of SNPs in pooled DNA, and found that the rs96516 SNP in the PPARGC-1A (\*604517) 3'UTR was significantly associated with sporadic CRC risk, although this result was not reproducible in the same study [27]. Second, Feilchenfeldt et al. [13] studied expression levels of all isoforms of PPAR- $\gamma$  and transcriptional partners such as PGC1 $\alpha$  in patients with different stages of colon cancer, and found that expression levels of PPAR- $\gamma$  vary between isoforms and cancer stages, while those of PGC1 were reduced in all cancer samples, with respect to normal samples [13].

Summarizing, while D'Errico et al., 2011 [20] showed that PGC1 $\alpha^{-/-}$  mice are susceptible to intestinal tumorigenesis, several papers addressing the role of PGC-1 in tumor cell lines showed a role of PGC-1 in tumor

progression. Thus, the function of PGC1 $\alpha$  in colorectal cancer risk is not entirely clear, although it seems likely that it has a role in this disease.

## Conclusions

Colorectal cancer is a major global public health problem and the second leading cause of cancer-related death. Our systematic review indicates that altered expression of PGC1 $\alpha$  modifies CRC risk. Most studies showed that overexpression of this gene increases CRC risk [10, 11, 14, 21, 23, 25, 26], while some studies indicated that lower than normal expression levels could increase CRC risk [6, 13]. Thus, various authors suggest that PGC1 $\alpha$  is a good candidate as a molecular target for cancer therapy. Reducing expression of this gene could help to reduce risk or progression of CRC [22, 25, 26], at least in different cell lines and transgenic or nude mice. However, in our opinion, there are not enough data on the role of PGC1 $\alpha$  using human tumor samples to conclude a role of this gene in CRC. This question should be investigated further.

## Additional file

**Additional file 1: Table S1.** Reasons for exclusion of 22 articles according to abstract and full-text. Table that describes the reasons for exclusion of the selected articles in the first review. **Table S2.** Full details of the key characteristics of the selected studies. Table that describes the details of the selected articles. (PDF 245 kb)

## Abbreviations

ACBP: Acyl-CoA-binding protein; AMPK: AMP-activated protein kinase; CRC: Colorectal cancer; ER $\alpha$ : Estrogen-related receptor  $\alpha$ ; Fig.: Figure; KSR1: Kinase suppressor of Ras 1; PGC-1: Peroxisome proliferator-activated receptor gamma coactivator 1; PPARGC1A or PGC1 $\alpha$ : Peroxisome proliferator-activated receptor gamma coactivator 1- $\alpha$ ; Sp1: Specificity protein 1; UC: Ulcerative Colitis; UPase: Uridine phosphorylase;  $\Delta\Psi_m$ : Mitochondrial membrane potential

## Acknowledgements

We thank Silvia Gutierrez and Verónica Dávila their help with methodological and intellectual contributions to the review.

## Funding

No funding was received for this research.

## Availability of data and materials

All data generated or analyzed during this study are included in this published article and its supplementary information files.

## Authors' contributions

JAM: First author and corresponding author. Her functions: To search for a protocol to make the review, make the online review, check titles, check abstracts, check full text and write the review. She made substantial contributions to conception and design, acquisition of data, analysis and interpretation of data. In addition, she has been involved in drafting the manuscript and revising it critically for important intellectual content and she agreed to be accountable for all aspects of the work by ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. CGD: Second author. Expert in the subject of the review. Her functions: To make the online review, check titles, check abstracts and check full text. She made substantial contributions to acquisition of data, analysis and interpretation of data. TFV: Third Author. Not

an expert in the subject of the review. Her functions: To make the online review and check titles. She made substantial contributions to acquisition and analysis of data. FST: Fourth author. Expert in the subject of the review. Her functions: To check the review. She has been involved in revising the manuscript critically for important intellectual content. LVM: Fifth author. Expert in the subject of the review. Her functions: To check the review. She has been involved in revising the manuscript critically for important intellectual content. AJM: Sixth author. Not an expert in the subject of the review. His functions: To check the review. He has been involved in revising the manuscript critically for important intellectual content. VMS: Senior author. His functions: To approve the protocols to be followed in the study, perform manuscript correction and proof reading. He agreed to be accountable for all aspects of the work by ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. All of the authors have participated sufficiently in the work to take public responsibility for appropriate portions of the content. All of them have read and approved the final manuscript.

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

#### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

#### Author details

<sup>1</sup>Grupo de Investigación en Interacciones Gen-Ambiente y Salud, Universidad de León, León, Spain. <sup>2</sup>Universidad de Cantabria, Santander, Spain. <sup>3</sup>Superior Physical Education School, Federal University of Pelotas, Pelotas, Brazil. <sup>4</sup>CIBER Epidemiología y Salud Pública (CIBERESP), Madrid, Spain. <sup>5</sup>Departamento Medicina Preventiva y Salud Pública, Facultad Ciencias de la Salud, Campus Vegazana, s/n. León, C.P.: 24071 Castilla y León, Spain.

Received: 29 November 2016 Accepted: 30 October 2017

Published online: 09 November 2017

#### References

- Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer*. 2010; 127:12.
- Vega P, Valentín F, Cubiella J. Colorectal cancer diagnosis: Pitfalls and opportunities. *World J Gastrointest Oncol*. 2015;7:12.
- Abraha I, Giovannini G, Serraino D, Fusco M, Montedori A. Validity of breast, lung and colorectal cancer diagnoses in administrative databases: a systematic review protocol. *BMJ Open*. 2016;6:3.
- Smolková K, Plecítá-Hlavatá L, Bellance N, Benard G, Rossignol R, Ježek P. Waves of gene regulation suppress and then restore oxidative phosphorylation in cancer cells. *Int J Biochem Cell Biol*. 2011;43:7.
- Bianchi G, Martella R, Ravera S, Marini C, Capitanio S, Orengo A, et al. Fasting induces anti-Warburg effect that increases respiration but reduces ATP-synthesis to promote apoptosis in colon cancer models. *Oncotarget*. 2015;6:14.
- Ussakli CH, Ebaee A, Binkley J, Brentnall TA, Emond MJ, Rabinovitch PS, et al. Mitochondria and tumor progression in ulcerative colitis. *J Natl Cancer Inst*. 2013;105:16.
- Jose C, Bellance N, Rossignol R. Choosing between glycolysis and oxidative phosphorylation: a tumor's dilemma? *Biochim Biophys Acta Bioenerg*. 2011; 1807:6.
- Jones AWE, Yao Z, Vicencio JM, Karkucinska-Wieckowska A, Szabadkai G. PGC-1 family coactivators and cell fate: roles in cancer, neurodegeneration, cardiovascular disease and retrograde mitochondria-nucleus signalling. *Mitochondrion*. 2012;12:1.
- Mazzanti R, Giulivi C. Coordination of nuclear- and mitochondrial-DNA encoded proteins in cancer and normal colon tissues. *Biochim Biophys Acta Bioenerg*. 2006;1757:5–6.
- Shin SW, Yun SH, Park ES, Jeong JS, Kwak JY, Park JI. Overexpression of PGC-1 $\alpha$  enhances cell proliferation and tumorigenesis of HEK293 cells through the upregulation of Sp1 and acyl-CoA binding protein. *Int J Oncol*. 2015;46:3.
- Srivastava S, Barrett JN, Moraes CT. PGC-1 $\alpha$ / $\beta$  upregulation is associated with improved oxidative phosphorylation in cells harboring nonsense mtDNA mutations. *Hum Mol Genet*. 2007;16:8.
- Villena JA. New insights into PGC-1 coactivators: redefining their role in the regulation of mitochondrial function and beyond. *FEBS J*. 2015;282:4.
- Feilchenfeldt J, Bründler MA, Soravia C, Tötsch M, Meier CA. Peroxisome proliferator-activated receptors (PPARs) and associated transcription factors in colon cancer: reduced expression of PPAR $\gamma$ -coactivator 1 (PGC-1). *Cancer Lett*. 2004;203:1.
- Bellafante E, et al. PGC-1 $\beta$  promotes enterocyte lifespan and tumorigenesis in the intestine. *Proc Natl Acad Sci U S A*. 2014;111:42.
- Seale P. Transcriptional regulatory circuits controlling brown fat development and activation. *Diabetes*. 2015;64:7.
- López-Lluch G, Irueta PM, Navas P, De Cabo R. Mitochondrial biogenesis and healthy aging. *Exp Gerontol*. 2008;43:9.
- Van Gisbergen MW, Voets AM, Starman MHW, De Cood IFM, Yadakd R, Hoffmann RF, et al. How do changes in the mtDNA and mitochondrial dysfunction influence cancer and cancer therapy? Challenges, opportunities and models. *Mutat Res. Rev Mutat Res*. 2015;764:16–30.
- Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Rodgers M, et al. Guidance on the conduct of narrative synthesis in systematic reviews. *Biostat*. 1995;536(22):3.
- Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA GROUP. Preferred Reporting. Items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med*. 2009;6(7):264–9.
- D'Errico I, Salvatore L, Murzilli S, Lo Sasso G, Latorrec D, Martelli N, et al. Peroxisome proliferator-activated receptor-gamma coactivator 1-alpha (PGC1alpha) is a metabolic regulator of intestinal epithelial cell fate. *Proc Natl Acad Sci U S A*. 2011;108:16.
- Bhalla K, Hwang BJ, Dewi RE, Ou L, Twaddell W, Fang H, et al. PGC1 $\alpha$  promotes tumor growth by inducing gene expression programs supporting lipogenesis. *Cancer Res*. 2011;71:21.
- Wu F, Wang J, Wang Y, Kwok TT, Kong SK, Wong C. Estrogen-related receptor  $\alpha$  (ERR $\alpha$ ) inverse agonist XCT-790 induces cell death in chemotherapeutic resistant cancer cells. *Chem Biol Interact*. 2009;181:2.
- Fisher KW, Das B, Seok Kim H, Clymer BK, Gehring D, Smith DR, et al. AMPK Promotes Aberrant PGC1 $\beta$  Expression to Support Human Colon Tumor Cell Survival. 2015;35:22.
- D'Errico I, Lo Sasso G, Salvatore L, Murzilli S, Martelli N, Cristofaro M, et al. Bax is necessary for PGC1 $\alpha$  pro-apoptotic effect in colorectal cancer cells. *Cell Cycle*. 2011;10:17.
- Kong X, Fan H, Liu X, Wang R, Liang J, Gupta N, et al. Peroxisome proliferator-activated receptor gamma Coactivator-1 alpha enhances Antiproliferative activity of 5J-Deoxy-5-Fluorouridine in cancer cells through induction of uridine phosphorylase. *Mol Pharmacol*. 2009;76:4.
- Vellinga TT, Borowski T, de Boer VJ, Fatrai S, Van Schelven S, Trumpp K, et al. SIRT1/PGC1 $\alpha$ -dependent increase in oxidative phosphorylation supports chemotherapy resistance of colon cancer. *Clin Cancer Res*. 2015;21:12.
- Gaustadnes M, Ørntoft TF, Jensen JL, Torring N. Validation of the use of DNA pools and primer extension in association studies of sporadic colorectal cancer for selection of candidate SNPs. *Hum Mutat*. 2006;27:2.
- Do MT, Kim HG, Choi JH, Jeong HG. Metformin induces microRNA-34a to downregulate the Sirt1/Pgc-1 $\alpha$ /Nrf2 pathway, leading to increased susceptibility of wild-type p53 cancer cells to oxidative stress and therapeutic agents. *Free Radic Biol Med*. 2014;74:21–34.
- DeCensi A, Puntoni M, Goodwin P, Cazzaniga M, Gennari A, Bonanni B, Gandini S. Metformin and cancer risk in diabetic patients: a systematic review and meta-analysis. *Cancer Prev Res*. 2010;3:1451–61.

## **Anexo II**


---

RESEARCH ARTICLE

Open Access



# The RS4939827 polymorphism in the *SMAD7* GENE and its association with Mediterranean diet in colorectal carcinogenesis

Jéssica Alonso-Molero<sup>1</sup>, Carmen González-Donquiles<sup>2\*</sup> , Camilo Palazuelos<sup>3</sup>, Tania Fernández-Villa<sup>1</sup>, Elena Ramos<sup>1</sup>, Marina Pollán<sup>4,5</sup>, Nuria Aragonés<sup>4,5</sup>, Javier Llorca<sup>3,4</sup>, M. Henar Alonso<sup>4,6,7,8</sup>, Adonina Tardón<sup>4,9</sup>, Pilar Amiano<sup>4,10</sup>, José Juan Jiménez Moleon<sup>4,11</sup>, Rosana Peiró Pérez<sup>4,12</sup>, Rocío Capelo<sup>13</sup>, Antonio J. Molina<sup>1</sup>, Inés Gómez Acebo<sup>3,4</sup>, Marcela Guevara<sup>10</sup>, Beatriz Pérez Gómez<sup>4,5,14</sup>, Virginia Lope<sup>4,5</sup>, José María Huerta<sup>4,15</sup>, Gemma Castaño-Vinyals<sup>4,16,17,18</sup>, Manolis Kogevinas<sup>4,16,17,18,19</sup>, Victor Moreno<sup>4,6,7,8†</sup> and Vicente Martín<sup>4,20†</sup>

## Abstract

**Background:** The objective of our investigation is to study the relationship between the rs4939827 SNP in the *SMAD7* gene, Mediterranean diet pattern and the risk of colorectal cancer.

**Methods:** We examined 1087 cases of colorectal cancer and 2409 population controls with available DNA samples from the MCC-Spain study, 2008–2012. Descriptive statistical analyses, and multivariate logistic mixed models were performed. The potential synergistic effect of rs4939827 and the Mediterranean diet pattern was evaluated with logistic regression in different strata of adherence to the Mediterranean diet and the genotype.

**Results:** High adherence to Mediterranean diet was statistically significantly associated with colorectal cancer risk. A decreased risk for CRC cancer was observed for the CC compared to the TT genotype (OR = 0.65 and 95% CI = 0.51–0.81) of the rs4939827 SNP. Also, we could show an association between the Mediterranean diet pattern (protective factor) and rs4939827. Although the decreased risk for the CC genotype was slightly more pronounced in subjects with high adherence to Mediterranean diet, there was no statistically significant synergistic effect between genotype CC and adherence to the Mediterranean dietary pattern factors.

**Conclusion:** The *SMAD7* gene and specifically the allele C could be protective for colorectal cancer. An independent protective association was also observed between high adherence Mediterranean diet pattern and CRC risk. Findings from this study indicate that high adherence to Mediterranean diet pattern has a protective role for CRC cancer probably involving the Tumor Growth Factor- $\beta$  pathway in this cancer.

**Keywords:** Colorectal cancer, *SMAD7*, rs4939827, Mediterranean diet, Gene-environment

\* Correspondence: cgonzd00@estudiantes.unileon.es

†Equal contributors

<sup>2</sup>Centro de Investigación Biomédica en Red (CIBERESP) and Oviedo University; Departamento de Ciencias Biomédicas. Universidad de León. Campus de Vegazana, León, Spain

Full list of author information is available at the end of the article



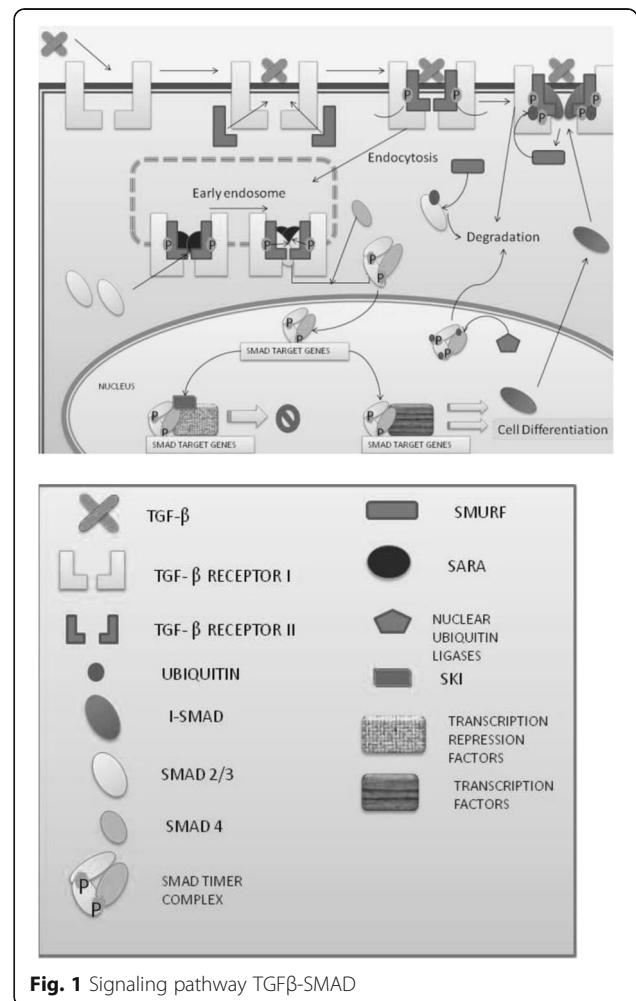
## Background

Colorectal cancer (CRC) has the third highest incidence in men, and the second highest in women worldwide [1]. This cancer has an increasing incidence in developing countries [2] probably due to the increasing prevalence of environmental factors that contribute to the development of CRC, for instance, dietary pattern and physical activity [3]. On the other hand, mortality levels are decreasing in countries with specialized care and better screening services [4]. Gene-Environment (GxE) interactions may play an essential role in increasing the susceptibility to developing colorectal cancer [3]. Among these factors, it is important to pay attention to the Mediterranean Diet Pattern adherence, since lower adherence to this pattern is more common as a direct consequence of the general westernization of the lifestyle's population [5–7]. Assessing the adherence to a Mediterranean dietary pattern is one of the particularly interesting approaches which may help to understand the relationship between diet and CRC [8]. However, the effect of this pattern over health depends on many individual aspects, involving genetic factors and polymorphisms. The lack of knowledge about the interactions between diet-polymorphism is a huge problem for the public health [9].

Earlier studies show a possible relationship between the diet pattern and the rs4939827 SNP in *SMAD7* (*SMAD family member 7*), which has been associated to CRC previously. The *SMAD7* gene acts as a Transforming Growth Factor Beta (TGF- $\beta$ ) family inhibitor by blocking the pathway signaling (Fig. 1) [10]. TGF- $\beta$  encodes for a cytokines family, which are multifunctional peptides that control some process like cell proliferation. In this way, when TGF- $\beta$  is inhibited by the interaction with *SMAD7*, cell proliferation is promoted and this could lead to the development of cancer. Another key gene in the undirected regulation of *SMAD7* expression is the gene *Ski* (*SKI proto-oncogene*), which acts by blocking the TGF- $\beta$  target genes, as *SMAD7* [11, 12].

The interaction between the Mediterranean diet and the *SMAD7* gene may be due to the action of other genes that are involved in the TGF- $\beta$  pathway, in particular *SMAD3* (*SMAD family member 3*), which is inhibited by the first one [13]. This pattern of diet is characterized by high fiber contents and before a fermentative process Na-butyrate is produced. This compound interacts with *SAMD3*, highlighting the signaling of this pathway [11]. In this way, we can check that the relationship between the dietary pattern and *SMAD7* expression is real.

The goal of this article is to investigate the relationship between the rs4939827 SNP, the Mediterranean Diet Pattern and the risk of CRC.



**Fig. 1** Signaling pathway TGF $\beta$ -SMAD

## Method

### Study population

MCC-Spain is a multicentric case-control study with population controls aiming to evaluate the influence of environmental exposures and their interaction with genetic factors in common tumors in Spain (prostate, breast, colorectal, gastroesophageal and chronic lymphocytic leukemia). Between September 2008 and December 2013, subjects between 20 and 85 years of age (matched by 5-year age interval) with histologically-confirmed newly-diagnosed colon or rectum cancer (ICD-10: C18, C19, C20, D01.0, D01.1, D01.2) were recruited in 23 Spanish hospitals from 12 Spanish provinces. Simultaneously, population-based controls frequency-matched to cases, by age, sex and region with the join distribution of the tumors included in MCC were randomly selected from primary care centers within hospitals' catchment areas. All participants signed an informed consent. Approval for the study was obtained from the ethical review boards of all recruiting centers. Additional information regarding the study design is provided elsewhere [14].

In our study, 3496 individuals were included from this MCC-Spain study. For this analysis, 1087 cases of colorectal cancer and 2409 controls were involved with available DNA samples for genetic analysis.

#### Lifestyle information

A computerized epidemiological questionnaire including self-reported socio-demographic and anthropometric data, family history of cancer, environmental exposures, use of selected drugs, reproductive history and current and past lifestyle behaviors (including leisure time physical activity and sedentary lifestyle) was administered by trained personnel in a face-to-face interview at enrolment. Waist and hip circumferences were measured by the interviewer [14]. Subjects were provided a previously validated semi-quantitative Spanish Food Frequency Questionnaire (FFQ) [15] which was modified to include regional products. The FFQ included 140 food items, and assessed usual dietary intake during the previous year. The FFQ included the specific cooking methods for meat and some pictures to establish how thoroughly-cooked participants prefer it. The FFQ was self-administered and returned by mail or filled out face to face (global response rate 88%). Frequency data was used to derive amount (g/day, g/1000 kcal) of each of the individual food types and summary variables. The food composition table has been a compiled table from the Centre for Higher Studies in Nutrition and Dietetics (CESNID) and other specific sources [16]. Cross-check questions on food groups intakes were used to adjust the frequency of foods eaten and reduce misreporting of food groups with large numbers of items [17, 18]. Data from this questionnaire was used to obtain a score in order to know the adherence to the Mediterranean diet by Sofi F [19]. Sofi F. et al. have computed the adherence to this type of diet taking into account 9 groups of foods (vegetables, legumes, meat, fish, integral cereals, fruit, dairy products, alcohol, olive oil) [19]. The range of this score goes from zero to eighteen points. From that score, we create subsets based on individuals with low (if the score is <9), medium (from 9 to 11 points) or high (if the score is higher than eleven points) adherence to this kind of diet [12]. Since the Mediterranean diet has a high fiber content and the *SMAD7* gene can interact with it, an analysis of this factor was considered essential.

For the variable of physical activity, the last 10 years were taken into account. This parameter also includes recreational physical activity. The assessed parameter was METs, a measured unit of metabolism. The responses obtained allowed us to create four subpopulations: sedentary, low physical activity, average physical activity and intense physical activity.

#### Samples processing

Peripheral blood (27 ml) was drawn from participants, which were aliquoted in whole blood, plasma, cellular fraction for DNA extraction, and serum and stored at  $-80^{\circ}\text{C}$ . Saliva was collected for subjects refusing to donate blood with the Oragene<sup>®</sup> DNA Kit and stored at room temperature until DNA extraction. We collected biological samples for DNA extraction for participants with interview, as well as toenail and hair samples. In 4 centers (Madrid, Cantabria, Asturias and Huelva) cases and controls also donated urine samples (60 ml) that were aliquoted and frozen at  $-80^{\circ}\text{C}$ . Fresh tumor biopsies or paraffin embedded samples are available in all participating hospitals. Standardized basic clinical and pathological information on the diagnosis and treatment of tumors was collected from hospital records by using a predefined format.

Genetic analyses were carried out within MCC-Spain and also through participation in international consortia. The InfiniumHumanExomeBeadChip from Illumina was used to genotype >200,000 coding markers plus 6000 additional custom variants on the pathways of interest [14].

#### SNP selection

After a literature search, different SNPs that were associated with CCR were selected. These SNPs were processed with the PLINK software in order to verify that they had been collected in the database of the MCC-Spain. This file contains the basic data needed for statistical analysis (cases and controls, SNPs, and the identifier of each individual). Finally, after an ad-hoc evaluation of the polymorphisms, rs4939827 was selected and we performed the statistical analyses described in the following section because only this SNP was considered relevant in its relation with colorectal cancer (see Additional file 1: Table S1).

#### Statistical methods.

First, a descriptive statistical analysis was performed to determine the characteristics of our study population. For age, mean and standard deviation were calculated. For the rest of variables, we calculate frequencies of cases and controls using STATA. Table 1 show variables used in analysis and how they were classified.

Then multivariate logistic mixed models, including the study region as a random effect term, were performed to evaluate the association between genotypes of selected polymorphisms and the risk of CRC. The same analysis was made in order to show the association of the component of the Mediterranean diet with the colorectal cancer. Adjusted odds ratios and confidence intervals at 95% were calculated by reference to the homozygous genotype of the most common allele (T allele). The



**Table 1** Descriptive analysis results. Characteristics of cases and controls

		CONTROLS		CASES		p-values
		N	%	N	%	
		2409		1087		–
Age	(Mean/SD)	63.13	11.5	66.64	10.4	–
Sex	Men	1300	54	708	65.1	<0.0001
	Women	1109	46	379	34.9	
First Degree Family History	No	2081	90.6	855	82.1	<0.0001
	Yes	216	9.4	186	17.9	
Cultural level	Less than primary school	409	17	288	26.5	<0.0001
	Primary school	803	33.3	458	42.1	
	Secondary school	698	29	221	20.3	
	University	499	20.7	120	11	
BMI	Normal weight	759	31.5	259	23.8	<0.0001
	Less than normal weight	23	1	13	1.2	
	Overweight	873	36.2	399	36.7	
	Obesity	754	31.3	416	38.3	
Total energy (Kcal/Day)	Low	795	33	296	27.2	<0.0001
	Medium	794	33	320	29.4	
	High	820	34	471	43.3	
METS	Sedentarism	855	35.5	468	43.1	<0.0001
	Low physical activity	385	16	134	12.3	
	Moderate physical activity	313	13	109	10	
	High physical activity	856	35.5	376	34.6	

N Number, BMI Body Mass Index, METS Metabolic Equivalent of Task. The cut point for category of total energy are <1586.3 Kcal/day, ≥ 1586.3 Kcal/day & < 2046.4 Kcal/day and ≤2046.4 for low, medium and high level respectively. For the physical activity, the cut points are 0 when there is no physical activity, < 8 METS\*hours/week for low physical activity, ≥ 8 and < 16 METS\*hours/week for moderate physical activity and >16 METS\*hours/week for high physical activity (according to American College of Sport Medicine)

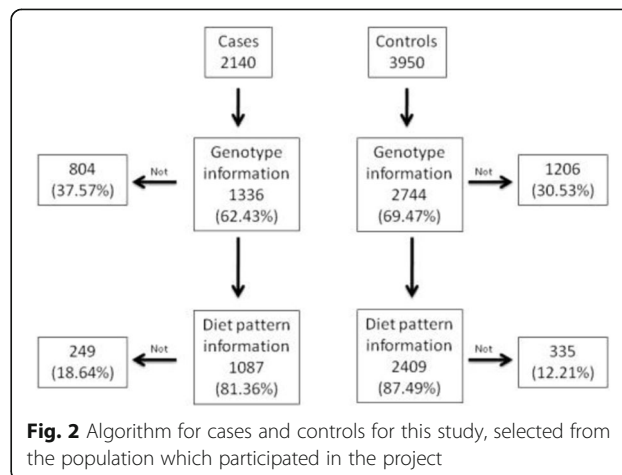
minimally adjusted odds ratio was calculated by adjusting by sex, age and educational background. The adjusted odds ratios were calculated taking into account the variables mentioned before and BMI, total energy intake, family history first degree, race and METS.

Finally, linear regression was used for associations between polymorphisms and the study variables. For this, the online software used was SNPstats, by the log-additive model mainly, although we also observed the results of other models (dominant, codominant, recessive and overdominant) [20]. By using STATA, the rs4939827 SNP association with the Mediterranean diet pattern was evaluated with logistic regression. For that, we stratified the variable of adherence to the Mediterranean diet and the genotype was also taken into account.

**Results**

The initial study population consists of 6090 individuals, distributed into 2140 cases and 3950 controls. Among them, 3496 had both genetic information and the dietary pattern available, the algorithm of missing data can be seen in Fig. 2. Table 1 shows the main

characteristics of those individuals which the analysis was performed with. Analyzing the missing data, it has been observed that only men have significant differences (in both case and control groups). The rest of variables do not show significant differences (Data not shown).



**Fig. 2** Algorithm for cases and controls for this study, selected from the population which participated in the project

### Descriptive analysis

In this study 3496 participants have been included, 1087 cases (65.1% males; 34.9% females) and 2407 controls (54% males; 46% females). The characteristics of the individuals of this study are described in Table 1. Data belong to a multi case-control study and matches were performed by frequencies in all types of tumors (colorectal, mama, prostate, leukemia, stomach), so it can be the reason for the mismatch between the sex distribution between affected cases and frequency-matched controls.

In the study population, the average age is 66.6 ( $\pm 10.4$ ) in cases and 63.1 ( $\pm 11.5$ ) in controls. Physical activity habits show a sedentary population, with 43.0% in cases and 35.5% in controls who do not practice physical

activity. BMI shows that only 23.8% in cases and 31.5% in controls have an adequate weight. Most of the population suffers from obesity or being overweight.

### Mediterranean diet pattern and components analysis

The components of the Mediterranean diet pattern by case-control status are shown in Table 2. We observe that there are significant differences in alcohol intake in the last 10 years between cases and controls, however we have not got any clear evidence that alcohol intake in the Mediterranean diet is associated with CRC. The opposite case can be found in vegetables and legumes, where a high intake of them may be protective against colorectal cancer. On the other hand, meat and dairy

**Table 2** Association between intake levels of food components in the Mediterranean diet and case (affection) status. The cut points are based on the score criteria for Sofi's Mediterranean diet pattern

		CONTROLS		CASES		OR minimally adjusted	CI (95%)	OR adjusted	CI (95%)
		N	%	N	%				
		2409		1087					
Alcohol	Moderate intake (12 g/day-24 g/day)	612	25.4	386	35.5	1	1	1	1
	High intake (>24 g/day)	1471	61.1	537	49.4	1.17	0.92-1.49	0.76	0.60-0.97
	Low intake (<12 g/day)	326	13.5	164	15.1	0.77	0.61-0.98	1.04	0.81-1.33
Vegetables	Low intake (<100 g/day)	563	23.4	306	28.2	1	1	1	1
	Moderate intake (100 g/day-250 g/day)	1378	57.2	624	57.4	0.82	0.68-0.98	0.85	0.70-1.03
	High/day intake (>250)	468	19.4	157	14.4	0.54	0.42-0.70	0.55	0.42-0.72
Legumes	Low intake (<70 g/day)	1561	64.8	670	61.6	1	1	1	1
	Moderate intake (70 g/day-140 g/day)	383	15.9	177	16.3	0.94	0.76-1.16	0.97	0.77-1.21
	High/day intake (>140 g/day)	465	19.3	240	22.1	0.75	0.60-0.92	0.77	0.61-0.96
Meat	Low intake (<80 g/day)	185	7.7	152	14.0	1	1	1	1
	Moderate intake (80 g/day-120 g/day)	442	18.3	236	21.7	1.4	1.15-1.70	1.29	1.05-1.60
	High/day intake (>120 g/day)	1782	74.0	699	64.3	1.98	1.54-2.55	1.74	1.32-2.30
Dairy products	Low intake (<180 g/day)	1541	64.0	693	63.8	1	1	1	1
	Moderate intake (180 g/day-270 g/day)	557	23.1	250	23.0	0.95	0.74-1.23	0.97	0.74-1.28
	High/day intake (>270 g/day)	311	12.9	144	13.2	1.02	0.82-1.29	1.01	0.79-1.30
Fruits	Low intake (<150 g/day)	0	0	0	0	1	1	1	1
	Moderate intake (150 g/day-300 g/day)	0	0	0	0	0	0	0	0
	High/day intake (>300 g/day)	2409	100	1087	100	omitted		omitted	
Cereals	Low intake (<130 g/day)	2327	96.6	1047	96.3	1	1	1	1
	Moderate intake (130 g/day-195 g/day)	71	2.9	32	2.9	0.97	0.62-1.51	1.08	0.68-1.74
	High/day intake (>195 g/day)	11	0.5	8	0.7	1.57	0.62-4.01	1.42	0.50-4.02
Fish	Low intake (<100 g/day)	169	7.0	91	8.4	1	1	1	1
	Moderate intake (100 g/day-250 g/day)	663	27.5	304	28.0	0.91	0.67-1.22	0.9	0.65-1.25
	High/day intake (>250 g/day)	1577	65.5	692	63.7	0.82	0.62-1.09	0.84	0.62-1.16
Olive oil	Low intake (<0.1 g/day)	115	10.6	245	10.2	1	1	1	1
	Moderate intake (0.1 g/day-0.99 g/day)	206	19.0	513	21.3	0.88	0.72-1.07	0.85	0.69-1.05
	High/day intake (>1 g/day)	766	70.5	1651	68.5	0.85	0.67-1.10	0.91	0.70-1.21

The minimally adjusted odds ratio (OR) was calculated adjusting by sex, age (as continuous variable) and educational level from a multivariate logistic mixed model. Area was used as random variable. The adjusted odds ratios were calculated taking into account the variables mentioned before and BMI, total energy intake, family history first degree, race and METS. (N: Number; CI: Confidence Interval). Cut-offs are based on score criteria of Sofi's Mediterranean Pattern [19]

products are risk factors when they are consumed in high quantities. Other components (integral cereals, fish and olive oil) do not show significant differences. Fruit cannot be analyzed because there are not enough subjects (Table 2).

Taking into account the adherence to the Mediterranean dietary pattern, 68.6% of cases and in 75.3% of controls have a high or medium adherence to the Mediterranean dietary pattern. All these variables show significant differences between cases and controls (Table 3).

#### Associations of rs4939827 with CRC and other factors

The rs4939827 SNP was statistically significantly associated with colorectal cancer according to several models. The most important result is referred to the Log-Additive model ( $p$ -value = 1.00E-04). The association of this polymorphism with CRC can also be observed in other models in a significant way: codominant model ( $p$ -value = 2e-04); dominant model ( $p$ -value = 6e-04); recessive model ( $p$ -value = 8e-04). The analysis of association of this polymorphism with risk factors is also shown in the summary table.

#### rs4939827 associations with risk factors of this study

After analyzing the association between rs4939827 and risk factors, statistically significant results can be observed in the case of the degree of adherence to the Mediterranean dietary pattern. Other factors assessed did not show a significant association with the SNP (data

**Table 3** Association between SNP rs4939827 and the Mediterranean dietary pattern adherence in relation with colorectal cancer taking into account the genotype

	Cases		Controls		Adjusted OR	
	n	%	n	%	OR	CI 95%
Low	341	31.4	594	24.7	1	–
TT	110	32.2	171	28.8	1	–
CT	169	49.6	302	50.8	0.83	0.59–1.17
CC	62	18.2	121	20.4	0.70	0.44–1.08
Medium	431	39.7	1010	41.9	0.85	0.70–1.03
TT	154	35.7	316	31.29	1	–
CT	211	49	478	47.2	0.87	0.65–1.15
CC	66	15.3	216	21.38	0.64	0.44–0.93
High	315	29.0	805	33.4	0.64	0.51–0.79
TT	123	39.1	242	30.1	1	–
CT	144	45.7	405	50.3	0.75	0.55–1.02
CC	48	15.2	158	19.6	0.63	0.42–0.96

CT vs TT: Adjusted Odd Ratio (OR) = 0.82 and 95% Confidence Interval (CI) = 0.69–0.98

CC vs TT: OR = 0.65 and 95% CI = 0.51–0.81

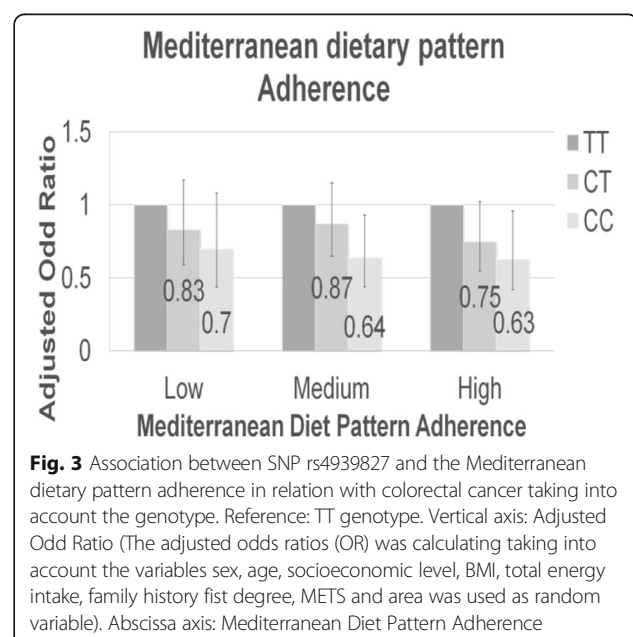
The adjusted odds ratios (OR) was calculating taking into account the variables sex, age, educational level, BMI, total energy intake, family history fist degree, METS. Area was used as random variable (N: Number; CI: Confidence Interval)

not shown). We conducted an analysis for rs4939827 with an interaction term to evaluate the independent main effect of the Mediterranean diet and main effect of the protective rs4939827 allele (see Additional file 2: Table S2). However, none of the results provide a significant  $p$ -value although the overdominant model is close to  $p$ -value = 0.05 ( $p$ -value = 0.07), determining that C allele it the dominant one.

The CC genotype (16.1% in cases and 20.5% in controls) seems to be protective compared to the TT (35.6% in cases and 30.3% in controls) genotype (ORa TTvsCT = 0.82 IC (95%) = 0.69–0.98; ORa TTvsCC = 0.65 IC (95%) = 0.51–0.81). When we check the association with adherence pattern to the Mediterranean diet (only statistically significant association), it can be observed that high adherence is a protective factor (Fig. 3 and Table 3). However, we could not see a synergistic effect between both genotype CC and adherence to the Mediterranean dietary pattern factors.

#### Discussion

SMAD7 is a target gene of the signaling pathway TGF $\beta$ -SMAD. For the TGF $\beta$  expression, it must be joined with its receptor type II. This binding stimulates the receptor type I, which binds to the above complex and, the complete assembly, is able to regulate the formation and activation (by phosphorylation) of the heterodimeric complex SMAD2/3. After the activation, this complex joins SMAD4 (SMAD family member 4) protein, which helps the complex translocation to the nucleus. In this organelle, the activation of the expression of target genes of TGF $\beta$  occurs, including the SMAD7 gene and other



genes responsible for proliferation inhibition and cell division. This happens when the SKI protein is not in the nucleus, which is capable of binding with the heterotrimeric complex and inhibit its function, inhibiting the expression of the *SMAD7* gene in this way (Fig. 1) [6, 7].

In this study of 1087 cases of CRC and 2409 controls, the rs4939827 SNP with the CC genotype was associated with a reduction of the risk for colorectal cancer. This may be due to *SMAD7* participation modulating the TGF $\beta$  pathway. SMAD7 protein is capable of binding with TGF $\beta$ , causing its ubiquitination and destruction. In this way, the action of TGF $\beta$  is inhibited and cell proliferation can occur. To clarify the role of *SMAD7* in cell growth in the CRC, Halder et al. [21], overexpressed this gene. This fact produces a higher cell growth than in normal conditions (21). At the same time, it has been observed that inhibition of *SMAD7* with a specific oligonucleotide reduces cell growth in CRC [22]. This data leads to the hypothesis that rs4939827 with genotype CC is a protective genotype because the correct expression of the *SMAD7* gene is inhibited, preventing cell proliferation and reducing the susceptibility to colorectal cancer.

This hypothesis is supported by B. Zhang et al. (2014) [23] who share the idea that the gene variant rs4939827 may be associated with an increase of CRC survivorship [23]. Similarly, Slattery M. et al. (2010) [24] also agree with our results as they show that the CC genotype is inversely associated with the risk of colon cancer. Thus, they conclude that individuals containing the homozygous recessive gene variant rs4939827 show a reduced risk of colon cancer by a 27% [24]. Thompson C.L. et al. (2009) [25] performed an analysis of the SNP taking into account the dominant model, stratifying the population according to gender. They noticed that women with the C allele were associated with a decreased risk of CRC. In contrast, they did not obtain significant results in men [25].

The CC genotype of this polymorphism was also associated with survival of other cancers. For example, according to Geng, T.T. et al. (2015) [26], the dominant model showed that the rs4939827 polymorphism was significantly associated with a decreased risk of esophageal cancer by 0.67 fold due to a possible inhibition of the *SMAD7* gene [26].

The T allele must also be taken into account. Although T allele does not show a significant risk in our results, many authors support the idea by Jung KJ et al. (2015) [27], which defended the hypothesis that the T allele of this SNP was a risk factor for colorectal cancer and rectal cancer [27]. Furthermore, according to analysis made by Baert-Desurmont S. et al. (2015) [28] an increased risk is observed depending on the number of T alleles present in genotype [28]. Ho J. W. et al. (2011) [29] analyzed the risk of TC heterozygous and TT homozygous

genotypes. They identified both as risk genotypes, but only the homozygous genotype showed a significant result [29]. Tenesa A. et al. (2008) [30], after taking into account the location of the tumor to perform the analysis, noted that the T allele is a risk factor mainly for rectal cancer and not for the colon cancer [30]. In our results, we cannot verify that information, since we did not identify the T allele as a risk allele, only the C allele could be identified as a protective allele. Tenesa A. et al. (2010) [31] also associated the T allele of the rs4939827 SNP with an increase of mortality [31]. Passarelli M.N. et al. (2011) [32] found that *SMAD7* variants that inhibit TGF $\beta$  completely may reduce its tumor suppressor activity (resulting in an increased risk), but can also reduce their ability to promote their metastatic promoter activity (resulting in a slower progression of the tumor and improved survival) [32]. Finally, Yao et al. (2015) [33] support that the T allele increases the risk of CRC in Caucasian population [33].

If we analyze the SNP association with diet, our results are statistically significant. However, few authors refer to the association of this factor with the rs4939827 polymorphism. As it is mentioned before, association with the diet is due to the TGF $\beta$  signaling pathway genes that interact with *SMAD7*. There is evidence that this association may be due to the ingestion of high amounts of fiber in the Mediterranean diet. Vegetables and fruit, for example, are highly fermentable fiber sources, while products such as wheat bran are low fermentable. The difference is the amount of short-chain fatty acid formed after fermentation of the fiber by the colon's bacteria in this organ, as described by Nguyen K.A. et al. (2006) [11]. So, food containing fermentable fiber (fruit and vegetables) will produce higher amount of fatty acids than that containing short chain poorly fermentable fiber. Therefore, there is a positive relationship between the consumption of large amounts of fiber and a lower risk of colon cancer. This may be due to the fact that the Na-butyrate (Na-B, the main product of the fermentation of fiber) is able to induce cell cycle arrest, cellular differentiation and, even, apoptosis. It seems that the Na-B is capable of inducing, selectively, Smad3 phosphorylation, so pathway activity of TGF- $\beta$  is enhanced [6, 7]. Despite this, we have not found a statistically significant interaction in our results. It may be due it has not got enough statistical power or the actual interaction is carried out with SMAD3. This produces that TGF- $\beta$  is not ever ubiquitinated and degraded, so cyclins jogging cell proliferation (p15, p21 and p27) are expressed [34]. Besides, one of *SMAD7* functions is to produce cell cycle arrest in G1 phase. However, because of this mutation, this activity cannot be carried out by this process. Na-B intake helps this action be covered, further minimizing the risk of uncontrolled cell proliferation.

One limitation in this work is that genotype x diet effect could be due to fiber intake rather than adherence to the Mediterranean diet, although both concepts are related.

## Conclusions

There is an important role of SMAD7 gene in colorectal cancer as the allele C is protective against this disease. Also, we can see a protective association between rs4939827 SNP and high adherence to the Mediterranean diet pattern and we suppose that high adherence to Mediterranean diet pattern probably has a protective role. So they may participate together in the TGF $\beta$  pathway in colorectal cancer.

## Additional files

**Additional file 1: Table S1.** Analysis of SNPs related to CRC. The minimally adjusted odds ratio (OR) was computed adjusting by sex, age, educational level. Area, places where cases and controls were recruited, was used as random variable. Association is evaluated for carrying 1 or 2 SNP minor frequency alleles relative to a reference of zero SNP minor frequency allele. (DOCX 22 kb)

**Additional file 2: Table S2.** Associations between polymorphisms and the Mediterranean Diet Pattern of Sofi considering different genetic models. Genetic models: codominant, dominant, recessive, overdominant, log-additive. (DOCX 16 kb)

## Abbreviations

CESNID: Centre for Higher Studies in Nutrition and Dietetics; CRC: Colorectal cancer; FFQ: Spanish Food Frequency Questionnaire; GxE: Gene-Environment; Na-B: Na-butyrate; ORa: Odd Ratio adjusted; *Ski*: *Ski proto-oncogene*; *SMAD3*: *SMAD family member 3*; *SMAD4*: *SMAD family member 4*; *SMAD7*: *SMAD family member 7*; SNP: Single Nucleotide Polymorphism; TGF- $\beta$ : Transforming Growth Factor Beta

## Acknowledgements

We thank all the subjects who participated in the study and all MCC-Spain collaborators.

## Funding

The study was partially funded by the "Accion Transversal del Cancer", approved on the Spanish Ministry Council on the 11th October 2007, by the Instituto de Salud Carlos III-FEDER (PI08/1770, PI08/0533, PI08/1359, PI09/00773-Cantabria, PI09/01286-León, PI09/01903-Valencia, PI09/02078-Huelva, PI09/01662-Granada, PI11/01403, PI11/01889-FEDER, PI11/00226, PI11/01810, PI11/02213, PI12/00488, PI12/00265, PI12/01270, PI12/00715, PI12/00150, PI14/01219, PI14/0613), by the Fundación Marqués de Valdecilla (API 10/09), by the ICGC International Cancer Genome Consortium CLL (The ICGC CLL-Genome Project is funded by Spanish Ministerio de Economía y Competitividad (MINECO) through the Instituto de Salud Carlos III (ISCIII) and Red Temática de Investigación del Cáncer (RTICC) del ISCIII (RD12/0036/0036)), by the Junta de Castilla y León (LE22A10-2), by the Consejería de Salud of the Junta de Andalucía (2009-S0143), by the Conselleria de Sanitat of the Generalitat Valenciana (AP\_061/10), by the Recercaixa (2010ACUP 00310), by the Regional Government of the Basque Country, by the Consejería de Sanidad de la Región de Murcia, by the European Commission grants FOOD-CT-2006-036224-HIWATE, by the Spanish Association Against Cancer (AECC) Scientific Foundation, by the Catalan Government DURSI grant 2014SGR647, by the Fundación Caja de Ahorros de Asturias and by the University of Oviedo.

## Availability of data and materials

**Samples:** Biological samples were stored at the biobanks supported by Instituto de Salud Carlos III- FEDER: Parc de Salut MAR Biobank (MARBIOBANC) (RD09/0076/00036), "Biobanco La Fe" (RD 09 0076/00021) and FISABIO

Biobank (RD09 0076/00058). Also at the Public Health Laboratory from Gipuzkoa, the Basque Biobank, the ICOBIOBANC (sponsored by the Catalan Institute of Oncology), the IUOPA Biobank from the University of Oviedo and the ISCIII Biobank.

**Genotyping:** SNP genotyping services were provided by the Spanish "Centro Nacional de Genotipado" (CEGEN-ISCIII).

## Authors' contributions

VMo and VMa have made substantial contributions to conception and design; JAM and CGD have made substantial contributions to conception and design, analysis and interpretation of data and have been involved in drafting the manuscript or revising it critically for important intellectual content. CP, TFV, ER, MP, NA, JL, MHA, AT, PA, JJJM, RPP, RC, AJM, IGA, MG, BPG, VL, JMH, GCV, MK acquired data and given final approval of the version to be published. All authors read and approved the final manuscript.

## Ethics approval and consent to participate

The protocol of MCC-Spain was approved by the Ethics committees of the participating institutions. All participants were informed about the study objectives and signed an informed consent. Confidentiality of data is secured removing personal identifiers in the datasets. The database was registered in the Spanish Agency for Data Protection, number 2102672171. Permission to use the study database will be granted to researchers outside the study group, after revision and approval of each request by the Steering Committee. More details on the organization of the project can be found online at <http://www.mccspain.org>.

## Consent for publication

Not applicable

## Competing interests

The authors declare that they have no competing interests.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Author details

<sup>1</sup>Grupo de Investigación en Interacciones Gen-Ambiente y Salud. Instituto de Biomedicina (BIOMED), Universidad de León, León, Spain. <sup>2</sup>Centro de Investigación Biomédica en Red (CIBERESP) and Oviedo University; Departamento de Ciencias Biomédicas. Universidad de León. Campus de Vegazana, León, Spain. <sup>3</sup>Universidad de Cantabria, Santander, Spain. <sup>4</sup>CIBER Epidemiología y Salud Pública (CIBERESP), Madrid, Spain. <sup>5</sup>Cancer and Environmental Epidemiology Unit, National Centre for Epidemiology, Carlos III Institute of Health, Madrid, Spain. <sup>6</sup>Cancer Prevention and Control Program, Catalan Institute of Oncology, Hospitalet de Llobregat, Barcelona, Spain. <sup>7</sup>Department of Clinical Sciences, Faculty of Medicine, University of Barcelona, Barcelona, Spain. <sup>8</sup>Colorectal Cancer Group, Bellvitge Biomedical Research Institute (IDIBELL), Hospitalet de Llobregat, Barcelona, Spain. <sup>9</sup>Oncology Institute IUOPA, Universidad de Oviedo, Oviedo, Asturias, Spain. <sup>10</sup>Instituto de Salud Pública de Navarra, Pamplona, Navarra, Spain. <sup>11</sup>Instituto de Investigación Biosanitaria de Granada (ibs.GRANADA), Hospitales Universitarios de Granada/Universidad de Granada, Granada, Spain. <sup>12</sup>Dirección General de Salud Pública, Fundación para el fomento de la investigación sanitaria y biomédica de la Comunidad Valenciana, FISABIO-Salud Pública, Barcelona, Spain. <sup>13</sup>Centro de Investigación en Salud y Medio Ambiente (CYSMA), Universidad de Huelva, Huelva, Spain. <sup>14</sup>Cancer Epidemiology Research Group, Oncology and Hematology Area, IIS Puerta De Hierro, Madrid, Spain. <sup>15</sup>Department of Epidemiology, Murcia Regional Health Council, IMIB-Arrixaca, Murcia, Spain. <sup>16</sup>ISGlobal, Centre for Research in Environmental Epidemiology (CREAL), Barcelona, Spain. <sup>17</sup>IMIM (Hospital del Mar Medical Research Institute), Barcelona, Spain. <sup>18</sup>Universitat Pompeu Fabra (UPF), Barcelona, Spain. <sup>19</sup>School of Public Health, Athens, Greece. <sup>20</sup>Grupo de Investigación en Interacciones Gen-Ambiente y Salud de la Universidad de León, León, Spain.

Received: 7 November 2016 Accepted: 18 October 2017

Published online: 30 October 2017

## References

- Ferlay J, Ervik M, Dikshit R, Eser S, Mathers C, AIE.GLOBOCAN. 2015. Available from: [http://globocan.iarc.fr/old/burden.asp?selection\\_pop=62968&Text-p=Europe&selection\\_cancer=5060&Text-c=Colorectal+cancer&pYear=8&type=0&window=1&submit=Execute](http://globocan.iarc.fr/old/burden.asp?selection_pop=62968&Text-p=Europe&selection_cancer=5060&Text-c=Colorectal+cancer&pYear=8&type=0&window=1&submit=Execute).
- Center MM, Jemal A, Ward E. International trends in colorectal cancer incidence rates. *Cancer Epidemiol Biomark Prev*. 2009;18(6):1688–94.
- Hu X, Feng F, Li X, Yuan P, Luan R, Yan J, et al. Gene polymorphisms related to insulin resistance and gene-environment interaction in colorectal cancer risk. *Ann Hum Biol*. 2015;42(6):560–8.
- Ait Ouakrim D, Pizot C, Boniol M, Malvezzi M, Boniol M, Negri E, et al. Trends in colorectal cancer mortality in Europe: retrospective analysis of the WHO mortality database. *BMJ*. BMJ Publishing Group; 2015 Oct 6;351:h4970.
- Pérez-Jiménez J, Elena Díaz-Rubio M, Saura-Calixto F. Contribution of macromolecular antioxidants to dietary antioxidant capacity: a study in the Spanish Mediterranean diet. *Plant Foods Hum Nutr*. 2015 Dec 20;70(4):365–70.
- Schwingshackl L, Hoffmann G. Adherence to Mediterranean diet and risk of cancer: an updated systematic review and meta-analysis of observational studies. *Cancer Med*. 2015;4(12):1933–47.
- Rotelli MT, Bocale D, De Fazio M, Ancona P, Scalera I, Memeo R, et al. IN-VITRO evidence for the protective properties of the main components of the Mediterranean diet against colorectal cancer: a systematic review. *Surg Oncol Elsevier Ltd*. 2015;24(3):145–52.
- Fung TT, Brown LS. Dietary patterns and the risk of colorectal cancer. *Curr Nutr Rep*. 2013;2(1):48–55.
- Frazier-Wood AC. Dietary patterns, genes, and health: challenges and obstacles to be overcome. *Curr Nutr Rep*. 2015;4(1):82–7.
- PB | Signaling by TGF-beta Receptor Complex. [cited 2016 May 22]. Available from: <http://www.reactome.org/PathwayBrowser/#/R-HSA-170834>
- Nguyen KA, Cao Y, Chen JR, Townsend Jr. CM, Ko TC. Dietary fiber enhances a tumor suppressor signaling pathway in the gut. *Ann Surg*. 2006;243(5):617–9.
- Kim YS, Milner J a. Dietary modulation of colon cancer risk. *J Nutr*. 2007;137:2576S–9S.
- Yan X, Liao H, Cheng M, Shi X, Lin X, Feng XH, et al. Smad7 protein interacts with receptor-regulated Smads (R-Smads) to inhibit transforming growth factor-B (TGF-B)/Smad signaling. *J Biol Chem*. 2016;291(1):382–92.
- Castaño-Vinyals G, Aragonés N, Pérez-Gómez B, Martín V, Llorca J, Moreno Y, et al. Population-based multicase-control study in common tumors in Spain (MCC-Spain): rationale and study design. *Gac Sanit. Elsevier*. 2015 Jul;29(4):308–15.
- García-Closas R, García-Closas M, Kogevinas M, Malats N, Silverman D, Serra C, et al. Food, nutrient and heterocyclic amine intake and the risk of bladder cancer. *Eur J Cancer*. 2007;43(11):1731–40.
- Cervera P, Farran A, Zamora-Ros R. Tablas de composición de alimentos CESNID = Taules de composició d'aliments del CESNID: Universitat de Barcelona; 2004. 247 p
- Nöthlings U, Hoffman K, Boeing H. Do cross-check questions improve food frequency questionnaire data? *IARC Sci Publ*. 2002:23–5.
- Calvert C, Cade J, Barrett JH, Woodhouse A. Using cross-check questions to address the problem of mis-reporting of specific food groups on food frequency questionnaires. UKWCS steering group. United Kingdom Women's cohort study steering group. *Eur J Clin Nutr*. 1997;51(10):708–12.
- Sofi F, Macchi C, Abbate R, Gensini GF, Casini A. Mediterranean diet and health status: an updated meta-analysis and a proposal for a literature-based adherence score. *Public Health Nutr*. 2014;17(12):2769–82.
- Sole X, Guino E, Valls J, Iniesta R, Moreno V. SNPStats: a web tool for the analysis of association studies. *Bioinformatics Oxford University Press*. 2006 Aug 1;22(15):1928–9.
- Halder SK, Beauchamp RD, Datta PK. Smad7 induces tumorigenicity by blocking TGF-β-induced growth inhibition and apoptosis. *Exp Cell Res*. 2005;307(1):231–46.
- Li Q, Zou C, Zou C, Han Z, Xiao H, Wei H, et al. MicroRNA-25 functions as a potential tumor suppressor in colon cancer by targeting Smad7. *Cancer Lett. Elsevier Ireland Ltd*. 2013;335(1):168–74.
- Zhang B, Jia W, Matsuo K, Shin A, Xiang Y, Matsuda K, et al. Genome-wide association study identifies a new SMAD7 risk variant associated with colorectal cancer risk in east Asians. *Int J Cancer*. 2014;135(4):948–55.
- Slattery ML, Herrick J, Curtin K, Samowitz W, Wolff RK, Caan BJ, et al. Increased risk of colon cancer associated with a genetic polymorphism of SMAD7. *Cancer Res*. 2010;70(4):1479–85.
- Thompson CL, Plummer SJ, Acheson LS, Tucker TC, Casey G, Li L. Association of common genetic variants in SMAD7 and risk of colon cancer. *Carcinogenesis*. 2009;30(6):982–6.
- Geng T-T, Xun X-J, Li S, Feng T, Wang L-P, Jin T-B, et al. Association of colorectal cancer susceptibility variants with esophageal cancer in a Chinese population. *World J Gastroenterol*. 2015;21(22):6898–904.
- Jung KJ, Won D, Jeon C, Kim S, Kim T. II, Jee SH, et al. a colorectal cancer prediction model using traditional and genetic risk scores in Koreans. *BMC Genet*. 2015;16(49):1–7.
- Baert-Desurmont S, Charbonnier F, Houivet E, Ippolito L, Mauillon J, Bougeard M, et al. Clinical relevance of 8q23, 15q13 and 18q21 SNP genotyping to evaluate colorectal cancer risk. *Eur J Hum Genet*. 2016;24(1):99–105.
- Ho JW, Choi S, Lee Y, Hui TC, Cherny SS, Garcia-Barceló M-M, et al. Replication study of SNP associations for colorectal cancer in Hong Kong Chinese. *Br J Cancer*. 2011;104(2):369–75.
- Tenesa A, Farrington SM, Prendergast JGD, Porteous ME, Walker M, Haq N, et al. Genome-wide association scan identifies a colorectal cancer susceptibility locus on 11q23 and replicates risk loci at 8q24 and 18q21. *Nat Genet*. 2008;40(5):631–7.
- Tenesa A, Theodoratou E, Din FVN, Farrington SM, Cetnarskyj R, Barnetson RA, et al. Ten common genetic variants associated with colorectal cancer risk are not associated with survival after diagnosis. *Clin Cancer Res*. 2010;16(14):3754–9.
- Passarelli MN, Coghill AE, Hutter CM, Zheng Y, Makar KW, Potter JD, et al. Common colorectal cancer risk variants in SMAD7 are associated with survival among pre-diagnostic non-steroidal anti-inflammatory drug users: a population-based study of post-menopausal women. *Genes Chromosom Cancer*. 2011;50(11):875–86.
- Yao K, Hua L, Wei L, Meng J, Correlation HJ. Between CASC8, SMAD7 polymorphisms and the susceptibility to colorectal cancer. *Medicine (Baltimore)*. 2015;94(46):e1884.
- Dancea HC, Shareef MM, Ahmed MM. Role of radiation-induced TGF-beta signaling in cancer therapy. *Mol Cell Pharmacol*. 2009;1(1):44–56.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

