



universidad
de león

**ASOCIACIÓN DEL CÁNCER COLORRECTAL
CON VARIANTES GENÉTICAS Y RUTAS
METABÓLICAS RELACIONADAS CON NRF**

*Association of colorectal cancer with NRF-related genetic
variants and metabolic pathways*

Programa de doctorado en Biomedicina y Ciencias de la Salud.

Carmen González Donquiles

DIRECTORES DE LA TESIS

Vicente Martín Sánchez

Antonio José Molina de la Torre

León, septiembre 2018



D./D^a **Carmen González Donquiles** declara que la memoria de la tesis presentada bajo el título **"ASOCIACIÓN DEL CÁNCER COLORRECTAL CON VARIANTES GENÉTICAS Y RUTAS METABÓLICAS RELACIONADAS CON NRF"** 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 2 de Octubre de 2018

Fdo.: Carmen González Donquiles



universidad
de león

Estudios de Doctorado

El Dr. D. Vicente Martín Sánchez como Director¹ de la Tesis Doctoral titulada “ASOCIACIÓN DEL CÁNCER COLORRECTAL CON VARIANTES GENÉTICAS Y RUTAS METABÓLICAS RELACIONADAS CON NRF” realizada por Dña. Carmen González Donquiles en el programa de doctorado de Biomedicina y Ciencias de la Salud, informa favorablemente el depósito de la misma, dado que reúne las condiciones necesarias para su defensa.

Lo que firmo, en León a 2 de Octubre de 2018

Fdo.: Vicente Martín Sánchez

¹ Si la Tesis está dirigida por más de un Director tienen que constar los datos de cada uno y han de firmar todos ellos.



El Dr. D. Antonio José Molina de la Torre como Director² de la Tesis Doctoral titulada *“ASOCIACIÓN DEL CÁNCER COLORRECTAL CON VARIANTES GENÉTICAS Y RUTAS METABÓLICAS RELACIONADAS CON NRF”* realizada por Dña. Carmen González Donquiles en el programa de doctorado de Biomedicina y Ciencias de la Salud, informa favorablemente el depósito de la misma, dado que reúne las condiciones necesarias para su defensa.

Lo que firmo, en León a 2 de Octubre de 2018

Fdo: Antonio José Molina de la Torre

² Si la Tesis está dirigida por más de un Director tienen que constar los datos de cada uno y han de firmar todos ellos.

“Los genes son como las letras, son imprescindibles para formar palabras, pero es el ambiente el que construye las frases y les da sentido”

Cornelia Bargmann, 2012

La presente Tesis Doctoral ha sido financiada para conseguir la Mención Internacional por el Centro de Investigación Biomédica en Red Epidemiología y Salud Pública, a través de una beca formativa disfrutada por la doctoranda. El trabajo ha sido desarrollado en el en el Área de Medicina Preventiva y Salud Pública de la Universidad de León, habiéndose realizado dos estancias, una nacional, en el Instituto Catalán de Oncología, y otra internacional de tres meses de duración en el Departamento de "Clinical Sciences and Community Health" de la "Università degli studio di Milano" (Milán, Italia). Para la realización de la presente tesis doctoral se han tenido en cuenta los datos recogidos en el proyecto MCC-Spain financiado por "Acción Transversal del Cáncer" aprobada por el Consejo de Ministros en 2007, por el instituto de Salud Carlos III-FEDER, Gobierno de España (PI08/1770, PI08/0553, 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) y por la Junta de Castilla y León (LE22A 10-2).

AGRADECIMIENTOS

A mis directores de Tesis, el Dr. Vicente Martín y Dr. Antonio José Molina de la Torre, por su generosidad al brindarme la oportunidad de recurrir a su capacidad y experiencia científica en un marco de confianza, afecto y amistad, fundamentales para la concreción de este trabajo.

A la Dra. Adonina Tardón García y al grupo CIBERESP CB06/02/1020, por su atención y disponibilidad a la hora de ofrecerme los medios necesarios para la elaboración de esta tesis.

Asímismo me complace agradecer la acogida, el apoyo y los medios recibidos en los distintos centros donde he desarrollado parte de mi Doctorado: el Instituto Catalán de Oncología y la Università degli estudio di Milano.

No me puedo olvidar del Área de Medicina Preventiva y Salud Pública:

A Laura V. por las consultas interminables de ciencia y de la vida; a Laura A., a Tania y a Lidia, por darme apoyo, cariño y tres florecillas; a Jessica, por nuestras investigaciones conjuntas y todo su apoyo; a Vero, por acercarme un poquito de sur.

A mis padres, porque con su sacrificio me dieron la posibilidad de llegar a esta instancia, por apoyarme de múltiples formas durante el desarrollo de esta tesis, por su constante estímulo, por enseñarme a enfrentar los obstáculos con optimismo, y que la perseverancia y el esfuerzo son el camino para lograr objetivos.

A la familia Campos Donquiles, por su apoyo constante e incondicional y su alegría ante mis logros, recordándome que una actitud positiva genera resultados brillantes.

A la familia Pérez Ruiz, por acogerme como a una hija y darme un hogar a 1000 km de mi tierra.

A Ramón, mi compañero preferido en este viaje, por su paciencia, ánimos y confianza en mí, porque no todos los días son fáciles cuando hay SNPs y genes de por medio.

A mi "seño Loli", por sentar los pilares básicos y ser el bordón que ha marcado y apoyado mis pasos, mi mochila cargada de ilusiones hacia mi particular "compostela".

A mis Chiquitrikis, mis amigas Emigrantes con "algo", ese algo que a pesar de la distancia siempre me ha hecho mirar hacia delante. A Alba y Julia, por sus empujones hacia la "victoria". Imposible rendirse teniendo un equipo como vosotras.

Agradecimientos también a esas otras personas que de una forma u otra han contribuido a que esta tesis sea posible.

A todos, muchas gracias.

ÍNDICE

| | |
|---|----|
| AGRADECIMIENTOS | 13 |
| ÍNDICE | 17 |
| ÍNDICE DE FIGURAS | 23 |
| ÍNDICE DE TABLAS | 29 |
| ABREVIATURAS..... | 33 |
| RESUMEN | 43 |
| ABSTRACT..... | 47 |
| CAPÍTULO I. INTRODUCCIÓN | 51 |
| CÁNCER COLORRECTAL | 53 |
| Determinantes del CCR | 54 |
| Factores modificables: | 54 |
| Factores no modificables: | 58 |
| USO DE DATOS CLÍNICO-EPIDEMIOLOGICOS ÓMICOS | 60 |
| BIOGÉNESIS MITOCONDRIAL Y CCR..... | 70 |
| Nuclear Respiratory Factor 1 (NRF1)..... | 75 |
| NRF1 y el cáncer..... | 78 |
| NRF1 y el CCR..... | 81 |
| Nuclear Respiratory Factor 2 (NRF2)..... | 81 |
| NRF2 y el cáncer..... | 84 |
| NRF2 y el CCR..... | 87 |
| CAPÍTULO II. OBJETIVOS..... | 89 |
| OBJETIVO GENERAL: | 91 |
| Objetivos específicos: | 91 |
| CAPÍTULO III. PLAN DE TRABAJO | 93 |
| PLAN DE TRABAJO | 95 |
| CAPÍTULO IV. MAPAS CONCEPTUALES DE NRF1 Y NRF2 EN EL CCR... | 97 |
| Metodología | 99 |

| | |
|---|-----|
| Resultados..... | 103 |
| Discusión | 114 |
| CAPÍTULO V. SNPS DE NRF1 Y NRF2 ASOCIADOS AL CCR..... | 121 |
| Metodología:..... | 123 |
| Identificación y extracción de SNPs..... | 125 |
| Identificación de posibles variables confusoras | 125 |
| Análisis estadístico. | 126 |
| Resultados:..... | 126 |
| Identificación y extracción de SNPs..... | 126 |
| Análisis estadístico | 128 |
| Discusión | 133 |
| Análisis descriptivo de la muestra | 133 |
| Selección de SNPs | 133 |
| Análisis SNP-CCR y SNP-posibles variables confusoras | 134 |
| CAPÍTULO VI. RUTAS METABÓLICAS ASOCIADAS AL CCR. | 137 |
| Metodología..... | 139 |
| Disponibilidad de datos genéticos. | 139 |
| Selección de pathways | 139 |
| " <i>SNP Ratio Test</i> " (SRT)..... | 140 |
| " <i>Fast - Gene Set Enrichment Score Analyses</i> " (ES) | 141 |
| " <i>Pathway of Distinction Analysis</i> " (PoDA)..... | 142 |
| Resultados..... | 143 |
| Selección de rutas: | 143 |
| " <i>SNP Ratio Test</i> "..... | 144 |
| " <i>Enrichment Score</i> " | 145 |
| " <i>Pathway of Distinction Analysis</i> "..... | 146 |
| Discusión | 149 |

| | |
|---|-----|
| CAPÍTULO VII. COROLARIO | 155 |
| CAPÍTULO VIII. CONCLUSIONES | 161 |
| CAPÍTULO IX. BIBLIOGRAFÍA | 165 |
| Artículos científicos relacionados con la presente tesis..... | 179 |
| CAPÍTULO IX. ANEXOS | 183 |

ÍNDICE DE FIGURAS

| | |
|--|-----------|
| Figura 1. Incidencia y mortalidad por tumores en España en el año 2015 .Basada en datos CancerToday. | 53 |
| Figura 2. Relación entre la obesidad y la aparición de CCR. | 58 |
| Figura 3. Edad al diagnóstico de CCR teniendo en cuenta el sexo. | 59 |
| Figura 4. Rutas por las que se desarrolla el CCR esporádico. | 60 |
| Figura 5. Clasificación de los SNPs en base a la función que desempeñan. | 61 |
| Figura 6. Algoritmo del proceso de ejecución de la metodología SRT. | 64 |
| Figura 7. Algoritmo del proceso de ejecución de la metodología "Fast Gene Set Enrichment Analysis". | 66 |
| Figura 8. Algoritmo del proceso de ejecución de la metodología "Pathway of Distinction Analysis". | 68 |
| Figura 9. Hipótesis de Warburg en células normales y tumorales. | 71 |
| Figura 10. Metabolismo de las células tumorales. | 72 |
| Figura 11. Génesis y desarrollo del cáncer debido a un aumento de ROS. | 73 |
| Figura 12. Avance y desarrollo del CCR. Basado en Cooper GM & Hausman RE (2007). | 74 |
| Figura 13. Estructura de NRF1. | 76 |
| Figura 14. Regulación de la expresión de NRF1. | 77 |
| Figura 15. Procesos desencadenados tras un aumento de la expresión de NRF1. | 79 |
| Figura 16. Procesos desencadenados tras la pérdida de función de NRF1. Basada en Permuth-Wey J. et al.(2011). | 80 |
| <i>Figura 17. Esquema de la proteína NRF2, señalizando los siete dominios (Neh1-Neh7).</i> | <i>82</i> |
| Figura 18. Relación entre NRF2 y el crecimiento y la progresión tumoral. | 84 |
| Figura 19. Génesis y desarrollo del cáncer tras la sobreexpresión de NRF2. | 85 |
| Figura 20. NRF2: de gen supresor de tumores a proto-oncogen. | 87 |
| Figura 21. Plan de trabajo. | 95 |
| Figura 22. Proceso de selección de artículos relacionados con NRF1 basado en la guía PRISMA. | 103 |
| Figura 23. Proceso de selección de artículos relacionados con NRF2 basado en la guía PRISMA. | 104 |
| Figura 24. Mapa conceptual 1. Expresión estándar de NRF2 y disminución del riesgo de CCR. | 112 |

| | |
|---|-----|
| Figura 25. Mapa conceptual 2. Sobreexpresión de NRF2 y aumento del riesgo de CCR..... | 113 |
| Figura 26. Mapa conceptual 3. Inhibición de NRF2 y aumento del riesgo de CCR. | 113 |
| Figura 27. Evaluación de la heterogeneidad de los artículos incluidos en la revisión. Cada color hace referencia a una característica señalada en la leyenda. | 115 |
| Figura 28. Metodología del capítulo VI y capítulo VII junto con los software utilizados..... | 124 |
| Figura 29. Posibles alelos del SNP rs12706898 y distribución de la frecuencia alélica en el mundo. Datos disponibles. en dbSNP y Ensembl..... | 127 |
| Figura 30. Histograma de frecuencias acumuladas referido a la ruta de las enfermedades inflamatorias del intestino. | 146 |
| Figura 31. Dot plots y diagramas de densidad para las distintas rutas con un resultado significativo obtenido mediante la metodología de análisis PoDA. | 148 |

ÍNDICE DE TABLAS

| | |
|--|-----|
| Tabla 1. Comparación de tres metodologías de análisis de pathways. | 68 |
| Tabla 2. Características de los artículos incluidos en la revisión sistemática cualitativa en los que NRF2 sigue un patrón de expresión estándar. | 105 |
| Tabla 3. Características de los artículos incluidos en la revisión sistemática cualitativa en los que NRF2 se sobreexpresa. | 109 |
| Tabla 4. Características de los artículos incluidos en la revisión sistemática cualitativa en los que NRF2 está inhibido. | 110 |
| Tabla 5. Características de los genes principales de estudio y SNPs pertenecientes a NRF1 o NRF2 recogidos en la base de datos MCC-Spain. | 127 |
| Tabla 6. Características de la población de estudio con el SNP rs20706898 genotipado. | 128 |
| Tabla 7. Distribución genotípica de la población de estudio. | 129 |
| Tabla 8. Asociación entre el CCR y los posibles genotipos del SNP rs12706898. | 131 |
| Tabla 9. Asociación entre el CCR y el SNP rs12706898 estratificada en función del sexo. | 132 |
| Tabla 10. Identificadores de los pathways evaluados en esta tesis. | 143 |
| Tabla 11. Resultados obtenidos tras analizar las rutas mediante "SNP Ratio Test". | 144 |
| Tabla 12. Resultados obtenidos tras analizar las rutas mediante "Fast Gene Set Enrichment Analysis". | 145 |
| Tabla 13. Resultados del análisis de rutas obtenidos mediante el método PoDA. ... | 147 |
| Tabla 14. Resultados obtenidos tras el análisis de las rutas mediante SRT, fGSEA y PoDA. | 149 |

ABREVIATURAS

| | |
|-------|---|
| 4HNE | Hydroxynonenal |
| 5-FU | 5-fluorouracilo |
| aa | Aminoácidos |
| ABC | ATP-Binding Cassette |
| AD1 | Dominio ácido 1 |
| AD2 | Dominio ácido 2 |
| ADN | Ácido desoxirribonucleico |
| ADNmt | Ácido desoxirribonucleico mitocondrial |
| AFPG | Antecedentes familiares de primer grado |
| AGP | Atmospheric pressure gas |
| AKRs | Aldoketoreductasas |
| AMPK | Proteín kinasa activada por AMP |
| AOM | Azoxymetano |
| APC | Adenomatous polyposis coli |
| ARE | Elementos de Respuesta Antioxidante |
| ARNm | ARN mensajero |
| ASK1 | Apoptosis signal-regulating kinase 1 |
| ATP | Adenosín trifosfato |
| Bcl-2 | B-cell lymphoma 2 |
| BRAF | B-Raf proto-oncogen |
| bZIP | Basic Leucine Zipper Domain |
| Ca | Calcio |
| CACC | Colitis asociada a cáncer de colon |

NRFs y cáncer colorrectal

| | |
|----------|--|
| cAMP | adenosín monofosfato cíclico |
| CAPE | Caffeic acid phenethyl ester |
| CCR | Cáncer Colorrectal |
| CIMP | Metilaciones en islas CpG |
| CIN | Inestabilidad cromosómica |
| C-MYC | MYC proto-oncogene |
| CNC | Can and collar |
| COX2 | Ciclooxigenasa-2 |
| cSNP | SNP codificantes |
| CU | Colitis ulcerosa |
| CXCR4 | C-X-C chemokine receptor type 4 |
| DBD | DNA binding domain |
| DHA | Ácido docosahexaenoico |
| DSS | Dextran sodium sulphate |
| ERK1/2 | Extracellular signal–regulated kinases 1/2 |
| ES | Enrichment Score |
| fGSEA | fast Gene Set Enrichment Score |
| GSEA | Gene Set Enrichment Score |
| GSH | Glutación |
| GSK3-B | Glycogen synthase kinase 3 beta |
| GWAS | Genome Whole Analysis Studies |
| Hb-alpha | Hemoglobina-alpha |
| HHB1 | N-terminal homology box-1 |

| | |
|------------|---|
| HHB2 | N-terminal homology box-2 |
| HIF-1alpha | Subunidad alfa del factor 1 inducible por hipoxia |
| HO-1 | Hemo-oxigenasa1 |
| Hrd1 | Synoviolin 1 |
| IER3 | Immediate early response 3 |
| IMC | Índice de masa corporal |
| JNK | c-Jun NH2-terminal kinases |
| Kb | Kilobases |
| Keap1 | Kelch-like ECH-associated protein 1 |
| KO | Knockout |
| KRAS | KRAS proto-oncogene |
| MAF | Frecuencia del menor alelo |
| MAFs | small Maf |
| MAPK | Mitogen-Activated Protein Kinases |
| MCC | Multicase-control |
| MEL | Melatonina |
| MSI | Inestabilidad por microsatélites |
| mTORC1 | Mammalian target of rapamycin complex 1 |
| NADPH | Nicotinamida adenina dinucleótido fosfato |
| Neh1 | Dominio Nrf2-ECH homology 1 |
| Neh1L | Dominio Nrf2-ECH homology 1-like |
| Neh2 | Dominio Nrf2-ECH homology 2 |
| Neh3 | Dominio Nrf2-ECH homology 3 |

NRFs y cáncer colorrectal

| | |
|------------|---|
| Neh3L | Dominio Nrf2-ECH homology 3-like |
| Neh4 | Dominio Nrf2-ECH homology 4 |
| Neh5 | Dominio Nrf2-ECH homology 5 |
| Neh6 | Dominio Nrf2-ECH homology 6 |
| Neh6L | Dominio Nrf2-ECH homology 6-like |
| Neh7 | Dominio Nrf2-ECH homology 7 |
| NES | Puntuación de enriquecimiento normalizada |
| NFE2L2 | Nuclear Factor (Erythroid-derived 2) like 2 |
| NFkB | Factor nuclear potenciador de las cadenas ligeras kappa de las células beta activadas |
| NHB1 | N-terminal homology box 1 |
| NHB2 | N-terminal homology box 2 |
| NO | Óxido nítrico |
| NRF1 | Nuclear Respiratory Factor 1 |
| NRF2 | Nuclear Respiratory Factor 2 |
| NRFs | Nuclear Respiratory Factors |
| NST | Glicodominio rico en Asn/Ser/Thr |
| NTD | Dominio N-terminal |
| O2 | Oxígeno molecular |
| OMS | Organización Mundial de la Salud |
| PGC1-alpha | Peroxisome proliferator-activated receptor-gamma coactivator 1 alpha |
| PI3K/AKT | Fosfoinositol-3-kinasa/AKT serine/threonine kinase 1 |
| PI3KCA | Fosfatidylinositol-4,5-bisfosfato 3-kinasa subunidad catalítica alfa |

| | |
|-------|--|
| PoDA | <i>"Pathway of Distinction Analysis"</i> |
| PS | Pterostilbeno |
| PTEN | Fosfatidilinositol-3,4,5-trisfosfato 3-fosfatasa |
| PYDDT | 2-(Pro-1-ynyl)-5-(5,6-dihidroxypenta-1,3-diynyl) thiophene |
| RE | Retículo Endoplasmático |
| ROS | Especies reactivas de Oxígeno |
| RS | Resveratrol |
| rSNP | SNP reguladores |
| RTK | Receptor Tirocín Kinasa |
| RXR | Receptor X retinoide |
| SAM | S-Adenosil Metionina |
| SM | Síndrome metabólico |
| SNP | Single nucleotide polymorphism |
| SR | Repetición de serinas |
| srSNP | SNP estructurales |
| SRT | <i>"SNP Ratio Test"</i> |
| Srx | Sulfirredoxin |
| TADs | Domino de transactivación |
| tBHQ | Terbutil Hidroquinona |
| WOS | Web of Science |
| WT | Wild Type |

"Todos tenemos algún valor añadido que nos puede ayudar a destacar en cualquier aspecto de la vida. Solo es cuestión de pararse a pensar, identificar las cosas que hacemos sin apenas dificultad y ponerlas muchas veces en práctica hasta afianzar nuestras fortalezas"
Gerver Torres.

RESUMEN

Introducción: El cáncer colorrectal es un tipo de cáncer multietápico y multifactorial influenciado por una gran cantidad de elementos ambientales y genéticos. La Biogénesis Mitocondrial, haciendo mención especial a los genes NRF1 y NRF2, podría estar relacionada con este tipo de cáncer ya que estos genes controlan procesos como la regulación de los ARE, la homeostasis y la activación de genes involucrados en la función mitocondrial, coordinando la expresión de genes nucleares y mitocondriales.

Objetivo: Analizar la relación entre los genes NRF1 y NRF2 y el CCR e identificar variantes genéticas y rutas metabólicas relacionadas con estos genes asociadas con el riesgo de CCR.

Metodología: Se llevó una revisión sistemática cualitativa para analizar el papel de NRF1 y NRF2 en el CCR. A partir de los datos del estudio MCC-Spain, un estudio multicaso-control, se realizó un análisis de SNP a SNP y un análisis de pathways en el que se emplearon tres técnicas diferentes: SNP Ratio Test, fast Gene Set Enrichment Score y Pathway of Distinction Analysis.

Resultados: De acuerdo a las evidencias científicas publicadas, se podría afirmar que existe relación entre la expresión/actividad alterada de NRF2 y el cáncer colorrectal, de tal manera que tanto una sobreexpresión como una inhibición del mismo se relaciona con este tipo de cáncer. El genotipo gT del SNP rs12706898 se detectó como un factor de riesgo para desarrollar cáncer de colon derecho en hombres. Mediante la técnica de PoDA, se detectaron 6 rutas metabólicas relacionadas con NRF1 y NRF2 cuya alteración podría estar relacionada con el cáncer colorrectal: la ruta de IBD, la ruta de las MAP Kinasas, la ruta de señalización NFK β , la ruta de señalización PI3K-AKT, la ruta de los transportadores ABC y la ruta de la biogénesis mitocondrial.

Conclusión: La presencia de SNPs en genes de rutas metabólicas relacionadas con NRF1 y NRF2 se asocia con una mayor probabilidad de sufrir CCR según los resultados obtenidos mediante la técnica PoDA.

ABSTRACT

Introduction: Colorectal cancer is a multi-stage, multi-factorial type of cancer influenced by a large number of environmental and genetic elements. Mitochondrial Biogenesis, specially the NRF1 and NRF2 genes, could be related to this type of cancer so they can control process such as ARE regulation, homeostasis and mitochondrial function-activators genes

Aim: To analyze the relationship among NRF1 and NRF2 and to identify genetic variants and genes-related pathways associated with the CRC risk.

Methodology: A qualitative systematic review was carried out to analyze NRF1 and NRF2 role in CRC. From MCC-Spain's data, a multicase-control study, an analysis of SNPs individually and a pathway analysis were carried out using three different techniques: SNP Ratio Test, fast Gene Set Enrichment Score and Pathway of Distinction Analysis.

Results: It could be say that there is a relationship between altered expression/activity of NRF2 and colorectal cancer, so both overexpression and inhibition are related to this type of cancer. The gT genotype of rs12706898 was detected as a risk factor for developing right colon cancer in men. Using the PoDA technique, 6 NRF1 and NRF2 –related metabolic pathways were detected whose disruption could be related to CRC: the IBD pathway, the MAP Kinases pathway, the NFK β signalling pathway, the PI3K-AKT signalling pathway, the ABC transporters pathway and the mitochondrial biogenesis pathway.

Conclusion: The presence of SNPs in NRF1 and NRF2-related metabolic pathways is associated with a higher probability of developing CRC according to the results obtained using the PoDA technique.

CAPÍTULO I. INTRODUCCIÓN

CÁNCER COLORRECTAL

El cáncer es un proceso de crecimiento y diseminación incontrolados de células que puede aparecer prácticamente en cualquier lugar del cuerpo, invadiendo el tejido circundante y pudiendo provocar metástasis (1). De forma general, todos los tumores tienen unas características comunes entre las que se encuentran la proliferación constitutiva, evasión de la acción de genes supresores de tumores, angiogénesis, inflamación, evasión de la muerte celular e inmortalidad. Recientemente, la reprogramación metabólica se ha descrito como la séptima característica adquirida por la célula tumoral ya que ésta se apropia de un metabolismo principalmente glicolítico (2).

Según la Organización Mundial de la Salud (OMS), el cáncer representa la tercera causa de muerte a nivel mundial y la segunda en los países desarrollados, tras las enfermedades cardiovasculares (3). El cáncer colorrectal (CCR) es la tercera neoplasia, en orden de frecuencia tumoral, en la población mundial. No obstante, en Europa se sitúa en segundo lugar tras el cáncer de pulmón y en España en primer lugar, con una incidencia estimada en 33000 casos al año. Cuando se diferencia entre sexos, el CCR ocupa el segundo lugar precedido por el cáncer de próstata en la población masculina y el cáncer de mama en la población femenina (Figura 1) (4,5).

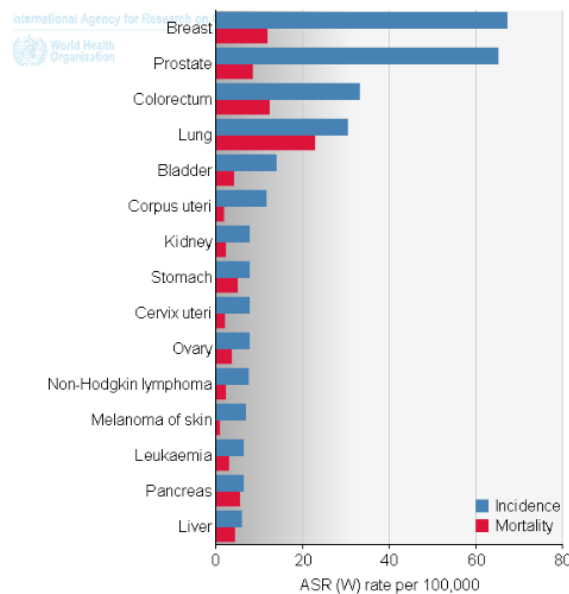


Figura 1. Incidencia y mortalidad por tumores en España en el año 2015. Basada en datos CancerToday.

En cuanto a la incidencia, se prevé un aumento que se multiplicará 1,2 veces desde el año 2020 al año 2030 a nivel mundial (6). Debido a la gran problemática que ello supone resulta esencial la aportación de nuevos conocimientos que permitan la prevención de esta enfermedad, un diagnóstico más precoz y la aplicación de tratamientos más eficaces.

El 40% de los diagnósticos de cáncer de colon se tratan de tumores localizados, alcanzando una supervivencia en 5 años del 90% mientras que si la enfermedad es regional (36% de los casos), el porcentaje de supervivencia a los 5 años se reduce hasta un 70%. Las cifras más bajas se alcanzan cuando los individuos diagnosticados presentan metástasis, variando la supervivencia desde un 20% hasta un 12%, a pesar de haber aumentado su esperanza de vida hasta un periodo de tiempo superior a los dos años (7)(8). En términos generales y sin tener en cuenta el estado del tumor en el momento del diagnóstico, EURO-CARE-5 estima la supervivencia media a los 5 años para el cáncer de colon en 57,1% y de recto en un 56,4% (9). A pesar de que más de la mitad de los individuos diagnosticados de CCR sobreviven, este tipo de cáncer en España es la segunda causa de muerte en relación a otros tipos de cáncer, siendo superado tan solo por el cáncer de pulmón.

DETERMINANTES DEL CCR

El CCR es un tipo de cáncer multietápico y multifactorial en el que los factores ambientales adquieren un rol esencial, interaccionando con factores genéticos (10). El 75-85% de los casos de CCR identificados en España ocurren de manera esporádica por interacción con la exposición ambiental, lo que lleva a pensar la gran influencia de estos sobre dicha enfermedad (11). Al hablar sobre los determinantes del CCR se pueden diferenciar entre aquellos modificables y no modificables.

Factores modificables:

La evidencia científica sugiere que los factores del estilo de vida, como por ejemplo la dieta o la actividad física, se asocian con el CCR más que con otros tipos de cáncer (8). A continuación, se describen en profundidad factores relacionados con el CCR informando sobre el riesgo y las posibles razones por las que influyen cada uno de ellos.

- **Exposición ocupacional:** El riesgo de CCR aumenta de dos a cuatro veces en individuos que trabajan expuestos a materiales industriales como gasolina, asbestos, ácido clorhídrico, humos, disolventes, pinturas, alquitrán, etc., de las fábricas o relacionados con la agricultura (12,13).

- **Consumo de cigarrillos:** El tabaco se ha mostrado como un factor de riesgo para el CCR, existiendo evidencia de un incremento del riesgo del 6% cuando el consumo es de hasta 5 paquetes/año, y del 26% cuando este ascendía a 30 paquetes/año, en comparación con los no fumadores, de acuerdo a los resultados obtenidos en un meta-análisis reciente. (14). El consumo de cigarrillos está diferencialmente asociado con ciertos tipos de CCR promoviendo el crecimiento de células tumorales residuales, a través de la resistencia a la quimioterapia o a través de la promoción de angiogénesis (15). Este hábito también podría inducir metilaciones aberrantes en el ADN y silenciar genes reguladores en la progresión del tumor además de inducir la inflamación, relacionada directamente con el desarrollo de CCR (16).

- **Consumo de alcohol:** La ingesta de alcohol se asocia con un incremento del riesgo de CCR habiéndose observado en un meta-análisis al respecto que el consumo de 2 bebidas alcohólicas conlleva un aumento del riesgo del 10% del riesgo en comparación con la no ingesta de alcohol (17). Uno de los riesgos de la ingesta de alcohol es que la mayoría de los individuos consumidores tienen dietas bajas en nutrientes esenciales, haciendo que el tejido se vuelva más susceptible a mutaciones que desencadenen el CCR. Además, el alcohol funciona como un disolvente promoviendo la penetración de moléculas carcinógenas a la mucosa del colon. (18).

- **Consumo de alimentos:** Dentro de la dieta existen distintos tipos de alimentos considerados factores de riesgo para el desarrollo de CCR. El principal grupo de alimentos considerado como factor de riesgo son las carnes rojas y procesadas incrementándose este hasta un 13% cuando el consumo es de 5 raciones/semana vs. 0 raciones/semana (14,19). La asociación positiva entre el consumo de carne y el CCR puede deberse a que a la hora de cocinarla a altas temperaturas se generan ciertos compuestos carcinógenos a partir de compuestos como las aminas heterocíclicas y los hidrocarburos aromáticos policíclicos (18,20). Además, la carne es una fuente abundante de grasas saturadas ,que aumentan la secreción de ácidos biliares promoviendo la inflamación y una posible

tumorogénesis, y sulfuros inorgánicos usados como conservantes ,que promueve mutaciones en el ADN de los colonocitos modificando su estado normal (21). Por último, el hierro contenido en las carnes rojas podría producir estrés oxidativo, proliferación descontrolada de los colonocitos y formación de compuestos N-nitrosos que son carcinógenos en el tracto intestinal. Sin embargo, se ha de mencionar que la carne de aves de corral se asocian con una reducción de la incidencia del CCR, aproximadamente del 20% (22,23).

Por otra parte, el consumo de grasas también intensifica el riesgo de desarrollo de este tipo de cáncer hasta un 70% debido a que favorece el desarrollo de flora bacteriana que degrada las sales biliares en compuestos potencialmente carcinógenos (18,24).

Además de grupos de alimentos que representan un riesgo, dentro de la dieta existen otros que debido a sus componentes nutricionales aportan cierta protección frente al CCR:

- Verduras y frutas. Se ha observado que el consumo de este grupo de alimentos correspondiente a 5 raciones a la semana comparado con el consumo nulo reduce el riesgo de CCR hasta un 14%, siendo el grupo de alimentos que mayor protección aporta (14). Las verduras y las frutas son una fuente rica en nutrientes que incluyen fibra, folato y vitaminas antioxidantes. La fibra podría disminuir el riesgo de CCR aumentando el volumen de heces, disminuyendo el tiempo de tránsito en el colon, diluyendo el contenido colónico e inhibiendo la formación de ácidos biliares secundarios debido al a disminución del pH por bacterias que se encargan de la fermentación de la fibra. Además, el folato promueve la metilación del ADN, la biosíntesis y la reparación y, por último, los nutrientes antioxidantes, incluyendo a los carotenos y a la vitamina A, C y E, podrían prevenir el CCR disminuyendo el estrés oxidativo (11,18,22,25).

- Pescado. El gran contenido en omega-3 y omega-6 del pescado es lo que lo hace beneficioso para prevenir el CCR, disminuyéndose el riesgo hasta un 12% para un consumo de una ración a la semana (26). El omega-3 protege frente al CCR inhibiendo a la ciclooxigenasa-2 (COX2), que suprime la biosíntesis de eicosanoides a partir de ácido araquidónico (27).

- Productos lácteos: estos alimentos podrían proteger frente a la neoplasia colorrectal reduciendo el riesgo hasta un 4%. Las causas de esta protección podrían ser debidas a su gran contenido en calcio ya que reacciona con

ácidos grasos libres que promueven la aparición del tumor en el lumen del colon, otros micronutrientes y constituyentes bioactivos. Los productos lácteos fermentados como el yogurt también se asocian de forma protectora con el CCR(28). El ácido láctico podría alterar la composición o función de la microbiota del hospedador, reduciendo la absorción de mutágenos de comidas cocinadas, inactivando los carcinógenos intestinales y reduciendo la inflamación (18,22).

- Cereales integrales. El consumo de 10g/día de cereales resultó en una reducción del riesgo de CCR próximo al 10% (29). Los cereales una fuente de sustancias anticancerígenas como la fibra (reduciendo la exposición a carcinógenos en el intestino debido a una mejora del tránsito y a la fermentación bacteriana), los antioxidantes (que protegen a las células del estrés oxidativo) y los fitoquímicos (22).

- Dieta: Además del análisis por grupos de alimentos, existen estudios en los que se ha valorado la dieta a partir de índices o "*scores*" para medir la adherencia de los individuos a un cierto patrón dietético, normalmente referido a un patrón dietético occidental o un patrón dietético mediterráneo (19). Se ha observado que una mayor adherencia al patrón de dieta mediterránea se traduce a una mayor protección frente al CCR, reduciéndose el riesgo hasta en un 11% aproximadamente (30).

- Obesidad: La obesidad tiene efectos tanto inmunológicos como endocrinos provocando la secreción de diferentes marcadores inflamatorios que regulan la inflamación sistemática y también contribuyen al proceso de carcinogénesis colorrectal, ascendiendo el riesgo a un 37% en hombres y un 7% en mujeres (31,32). Los factores nutricionales pueden influir en el desarrollo de CCR a través de sus efectos sobre la adiposidad y el trastorno metabólico. Un desequilibrio energético a largo plazo, representado por un exceso de consumo de energía, conduce al sobrepeso y a la obesidad, que son factores de riesgo que promueven la génesis del CCR (22). Al sufrir obesidad, aumenta la circulación de estrógenos y disminuye la sensibilidad a la insulina. Estos eventos junto con la inflamación asociada a la misma y la liberación de adipoquinas van a promover el inicio del ciclo celular, la proliferación celular y la evasión de la apoptosis. Todo esto lleva consigo la transición desde un epitelio normal hasta la aparición del adenoma, pasando por un estado hiperproliferativo en el epitelio y formación de criptas aberrantes, dando lugar al carcinoma colorrectal (Figura 2) (18,19). Además, un aumento de 2kg/m² supone un aumento del riesgo de CCR del 7% aunque a día de hoy no se conozca el mecanismo de manera exacta. Hay evidencias que sugieren una asociación entre el

IMC, la circunferencia de la cintura y el número, el tipo y la localización de pólipos de colon adenomatoso (31).

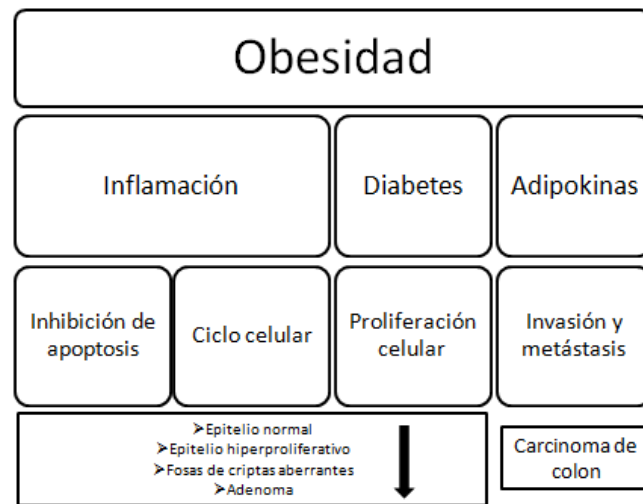


Figura 2. Relación entre la obesidad y la aparición de CCR.

- **Actividad física:** La actividad física aumenta la tasa de metabolismo del cuerpo, la motilidad intestinal, el metabolismo y la capacidad aeróbica, y reduce la presión sanguínea y la resistencia a la insulina. Por lo tanto, el riesgo de CCR desciende un 24%, variando este porcentaje en función de la frecuencia y la intensidad con que se realice esta actividad física (18,33). Además, un estudio reciente indica que aquellos individuos que realizan actividad física llevan asociado un estilo de vida saludable caracterizado por dieta saludable, un bajo consumo de alcohol, IMC correspondiente a normopeso y unas horas de sueño adecuadas (34).

Factores no modificables:

- **Edad:** La edad es sin duda el factor no modificable más fuerte para el desarrollo del CCR, con un aumento de la probabilidad de sufrir la enfermedad a partir de los 40 años, tal y como se puede observar en la Figura 3. Por lo tanto, el cáncer de colon se puede considerar una enfermedad adulta debido a que más del 90% de los casos tienen 50 años o más (35). Sin embargo, en la actualidad, el CCR parece estar aumentando continuamente entre las personas más jóvenes (18,36,37).

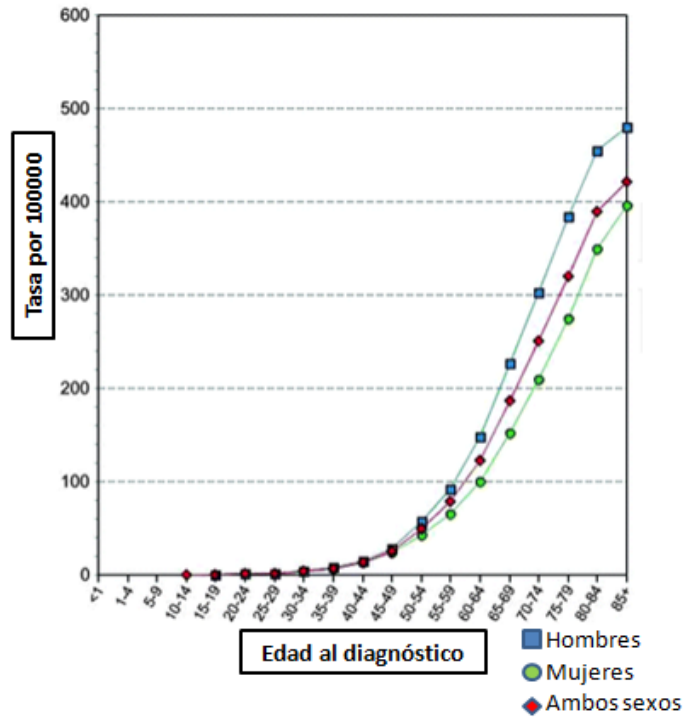


Figura 3. Edad al diagnóstico de CCR teniendo en cuenta el sexo.

- Sexo: Las tasas generales de incidencia y mortalidad por CCR mostraron tendencias crecientes a lo largo de los años, especialmente en las mujeres, y ambas siguen siendo sustancialmente más altas en los hombres. A pesar de las tasas más bajas para las mujeres, el número absoluto de casos para los dos sexos es más similar de lo que se podría esperar. Esto se debe al hecho de que el CCR es más común en los grupos de edad avanzada en los que las mujeres superan a los hombres (38).
- Nivel socioeconómico: Se ha demostrado que las personas con estatus socioeconómico más bajo tiene una alta incidencia de CCR, variando el riesgo desde un 42% en personas con menos de 12 años de estudios hasta un 12% en aquellos con titulación universitaria, si se comparan con sujetos post-graduados (39). Estos factores no influyen directamente sobre la aparición de cáncer de colon. Sin embargo, sí que pueden intervenir indirectamente condicionando otros factores como la dieta, exposiciones ocupacionales o el acceso al sistema sanitario (39,40).
- Historia personal de enfermedades inflamatorias del intestino: Los pacientes que han sufrido la enfermedad de Chron o la colitis ulcerosa aumentan la probabilidad de sufrir CCR entre 4-20 veces ya que tendrán una mayor tendencia a desarrollar displasia. Esto podría degenerar a CCR debido a la activación de factores

de transcripción que inducen la expresión de genes que estimulan la proliferación y supervivencia celular (18,41,42).

- Historia personal de pólipos adenomatosos de CCR: Los individuos con mayor tendencia a desarrollar CCR son aquellos que tienen mayor número de pólipos y aquellos cuyos pólipos son de gran tamaño (18).
- Historia familiar: este factor podría influir por diversos motivos entre los que se citan la herencia genética, los factores ambientales compartidos o una mezcla de ambos (18,43). Los familiares de primer grado de individuos con CCR tienen aproximadamente el doble de riesgo de desarrollarlo. Sin embargo, solo el 35 % de los casos de CCR pueden atribuirse a factores hereditarios, representando los síndromes familiares conocidos menos del 5% de todos los CCR. Por lo tanto, los factores genéticos también podrían estar involucrados en el CCR esporádico, entre los que se encuentran la vía de la inestabilidad cromosómica, la vía de inestabilidad por microsatélites (MSI) y la ruta de metilación de las islas CpG. (44). (Figura 4).

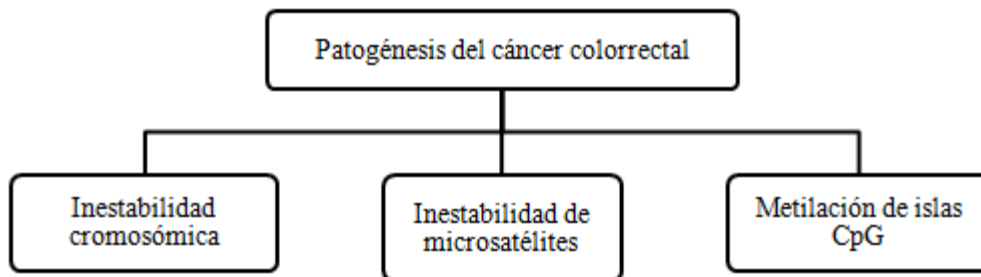


Figura 4. Rutas por las que se desarrolla el CCR esporádico.

USO DE DATOS CLÍNICO-EPIDEMIOLOGICOS ÓMICOS

En los últimos años, el uso de datos "ómicos" (datos a escala global) en Biomedicina se ha incrementado rápidamente, permitiendo estudiar las enfermedades desde un punto de vista biomolecular que anteriormente no estaba disponible. Para explorar en detalle la cantidad ingente de datos producidos por las tecnologías ómicas es indispensable la aplicación de técnicas nuevas de análisis y cálculo computacional que permitan extraer la información biológica disponible haciendo posible estudiar problemas médicos a través de la biología molecular, aplicando métodos estadísticos y computacionales (45).

Los datos ómicos pueden pertenecer a distintas áreas, abarcando datos pertenecientes a distintas categorías entre las que se podrían mencionar aquellas relacionadas con datos epidemiológicos o datos genéticos, entre los que se destacan los polimorfismos genéticos. Se denomina polimorfismo genético a la existencia de más de un alelo normal en un locus génico, con una frecuencia del alelo más raro superior al 1%. Cuando la variación se produce en tan solo una base, este polimorfismo se denomina *Polimorfismo de un Solo Nucleótido*, "Single Nucleotide Polymorphism" o "SNP". Estas variaciones son claves porque representan casi el 90% de todas las descritas en el genoma humano produciendo la variabilidad genética y fenotípica entre individuos.

Existen distintos tipos de SNPs: no funcionales, cuya presencia no promueve alteraciones de ningún tipo, y funcionales, que presentan consecuencias. A su vez, los SNPs funcionales se clasifican en rSNP o SNPs reguladores, es decir, aquellos que tienen implicaciones funcionales sobre la expresión génica, srSNP o SNPs estructurales en el caso de que alteren la traducción del ARN mensajero (ARNm), el corte y el empalme, la eficiencia para potenciar o inhibir el corte y empalme, la estabilidad de los ARNm y la función de las proteínas sin modificar su estructura, y cSNP o SNPs codificantes, que se encuentran en regiones codificantes y podrían alterar la formación y estructura proteica (Figura 5) (46).

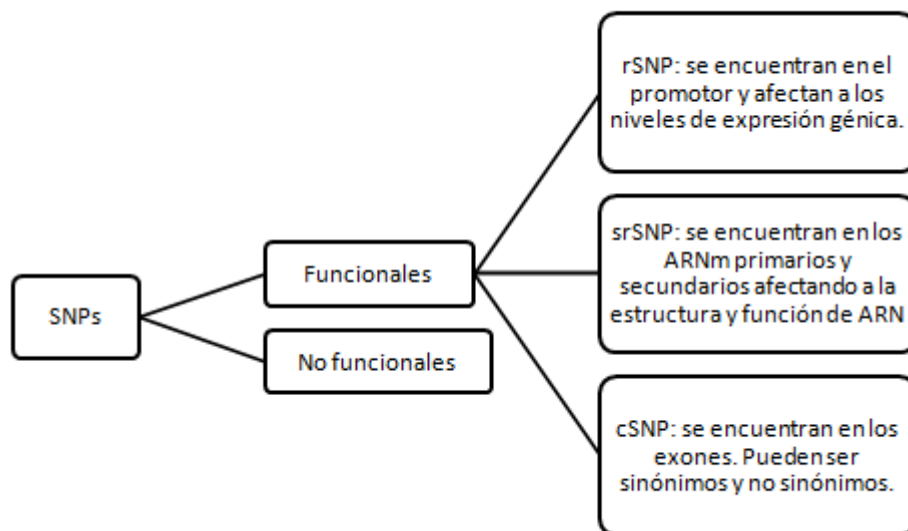


Figura 5. Clasificación de los SNPs en base a la función que desempeñan.

Aproximadamente el 89% de las variaciones asociadas a enfermedades residen en regiones intrónicas o intergénicas, haciendo difícil conocer las consecuencias moleculares directas de las variantes para la fisiología humana. Dado que las variantes genéticas comúnmente son dianas de elementos reguladores *cis*, principalmente potenciadores, en una manera específica en la enfermedad, tejido y célula, se ha sugerido que los principales contribuidores de predisposición a enfermedades comunes son variantes llamadas SNPs que modulan la unión a factores de transcripción. De hecho, *loci* asociados a enfermedades engloban comúnmente variantes que mapean sitios de unión a factores de transcripción y modulan la afinidad de esta unión con consecuencias directas sobre la expresión de genes diana (47).

Los análisis genómicos globales, "*Genome-Wide Association Study*" o GWAS buscan establecer una relación estadística entre variables genéticas poblacionales individuales y un fenotipo determinado (46). Mediante este tipo de análisis, se han identificado cerca de 50 *loci* relacionados con el riesgo de CCR basados en variantes del genoma nuclear. Sin embargo, esos *loci* tan solo explican una pequeña proporción de la heredabilidad del CCR (35). Hay otros estudios como el de Song N. *et al.*, (2016), Hong, Y *et al.*, (2016) y Zhan K. *et al.*, (2014) en los que se identifican *loci* genéticos asociados al CCR esporádico (48–50). Además, GECCO, un consorcio internacional, también cuenta con una gran cantidad de información genética en relación a este tipo de cáncer, habiendo identificado 11 *loci* nuevos relacionados con el CCR en el año 2012 (51).

Sin embargo, el análisis de todos los SNPs de un grupo de individuos sería prácticamente imposible, tanto por el tiempo como por la cuantía económica que supondría. Actualmente, se han detectado SNPs asociados al riesgo de CCR pero no son capaces de explicar por completo la aparición de esta enfermedad, por lo que se precisan nuevos métodos de identificación de SNPs que ayuden a predecir el riesgo de esta enfermedad (52).

De hecho, los genes en los sistemas biológicos interactúan unos con otros formando redes de genes que trabajan conjuntamente en funciones biológicas específicas. Estas relaciones pueden obtenerse a partir del análisis integrado del conocimiento previo almacenado en bases de datos biológicas (45,47). De esta

manera, un estudio de rutas aumentaría la eficiencia de los efectos y podría ayudar a revelar asociaciones entre el genoma y el riesgo de CCR con el objetivo de aproximarnos a una medicina de precisión cada vez más estratificada (53,54).

Para llevar a cabo el análisis de las rutas son diversos los métodos que se pueden encontrar en la literatura, diferenciándose en tres tipos:

1. Métodos que analizan la fracción de genes en una ruta que muestran cambios en la expresión de manera que tengan cierto efecto sobre el fenotipo.
2. Métodos que asignan puntuaciones a cada uno de los genes de la ruta puesto que suponen que pequeños efectos coordinados podrían tener grandes efectos sobre el fenotipo.
3. Métodos que tienen en cuenta la topología de la ruta y observan las interacciones entre los genes componentes del mismo.

En la presente tesis doctoral se examinan tres opciones de análisis de pathways que se describen a continuación: la primera opción, SNP Ratio Test (SRT), es un ejemplo del método explicado en el punto 1, mientras que las dos siguientes, Fast Gene Set Enrichment Score (fgSEA) y Pathway of Distinction Analysis (PoDA), son una representación de lo descrito en el segundo punto.

SNP Ratio Test (SRT)

Esta técnica lleva a cabo una comparación de la proporción de todos los SNPs significativos que forman parte de una ruta con respecto al total de SNPs de dicha ruta. La Figura 6 muestra un esquema explicativo de esta metodología, que tiene como base la distribución de los SNPs del GWAS en un ranking en función de la asociación de los SNPs con el fenotipo, para detectar cuantos SNPs de un pathway seleccionado superan un punto de corte establecido a partir de la selección de un p-valor en dicho ranking.

Se observa como a partir del GWAS original y conociendo el pathway que se desea evaluar, se lleva a cabo un test de asociación entre los SNPs del GWAS y la enfermedad de estudio con el objetivo de generar un ranking a través del cual se calcule un ratio (r_w), siendo el valor r_w el número de SNPs significativos de una ruta contenidos en el ranking de SNPs entre el número de SNPs total de la ruta (línea verde).

El siguiente paso consiste en permutar la etiqueta de casos y controles y, una vez que se genera un GWAS simulado, repetir los pasos anteriores hasta el cálculo del ratio (r'), representados sobre la línea azul del gráfico. En este caso, r' representa el número de SNPs significativos de una ruta contenidos en los rankings de SNPs simulados entre el número de SNPs total de la ruta.

Por último, se calcula un p-valor teniendo en cuenta los ratios obtenidos a partir de permutaciones (r') y el ratio obtenido a partir de datos reales (r_w). En caso de que este valor sea significativo, se podría concluir que el pathway de estudio es un buen predictor de la enfermedad (55). Este paso corresponde a las líneas rojas de la Figura 6, donde se observa como estas convergen en el cálculo de un p-valor empírico, donde se tiene en cuenta el valor S , siendo este el número de veces en el que el $r' \geq r_w$, y el valor N , correspondiente al número total de permutaciones.

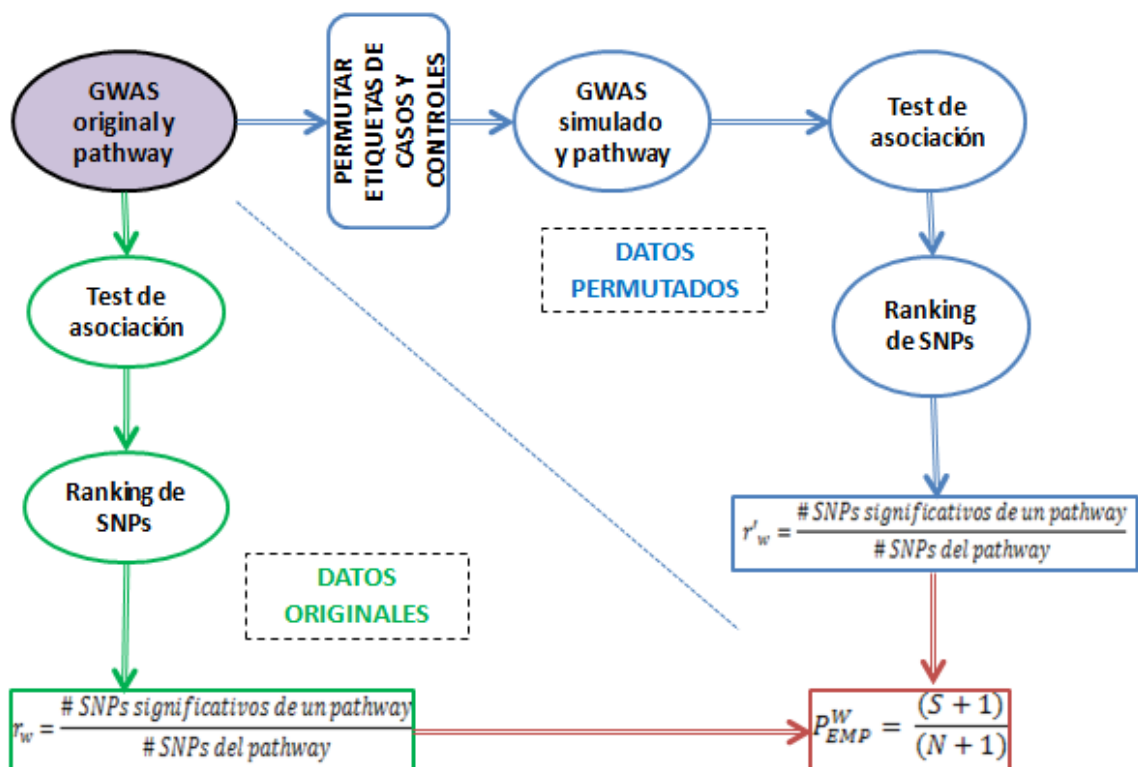


Figura 6. Algoritmo del proceso de ejecución de la metodología SRT.

Gene-Set Enrichment Analysis (GSEA)

El test de enriquecimiento (GSEA) tiene como objetivo principal determinar si los SNPs que conforman una ruta se encuentran agrupados al principio o al final de una lista de todos los SNPs del GWAS ordenados en función de su relación con el fenotipo. La Figura 7 muestra un esquema explicativo de esta metodología:

Partiendo del GWAS original, se calcula la puntuación de enriquecimiento normalizada (NES). Para ello, se realiza un test de asociación (test de χ^2) entre los SNPs del GWAS y el fenotipo para obtener un ranking de SNPs ordenados en función del nivel de significación. El paso siguiente consiste en realizar un "walking-down" por dicho ranking, sumando un valor relacionado con el valor χ^2 , el número de genes del ranking y el número de genes del pathway cuando el SNP del ranking pertenece a la ruta estudiada y restándolo si dicho SNP pertenece al ranking pero no a la ruta en cuestión. Estos valores se irán representando en un histograma de frecuencias acumuladas y la puntuación de enriquecimiento (ES) es la máxima desviación de cero. Finalmente, se normaliza este valor con el objetivo de tener una NES teniendo en cuenta el número de genes de la ruta metabólica. Estos pasos se describen en la línea verde de la Figura X.

A continuación, se realizan "n" permutaciones de las etiquetas del fenotipo conservando la proporción de casos y controles y se vuelven a recalcular "n" puntuaciones de enriquecimiento siguiendo el mismo procedimiento anterior. De esta manera, finalmente se obtienen "n" NES* (línea azul).

El siguiente paso, consiste en calcular un p-valor nominal relativo a esta distribución nula, teniendo en cuenta el número de veces en los que NES es mayor o igual a NES* y el número total de permutaciones. Finalmente, para limitar efecto de los múltiples análisis, se calcula, mediante False Discovery Rate, un valor q que se empleará como criterio de significación estadística para valorar la asociación entre la ruta estudiada y el fenotipo de interés (líneas rojas) (56).

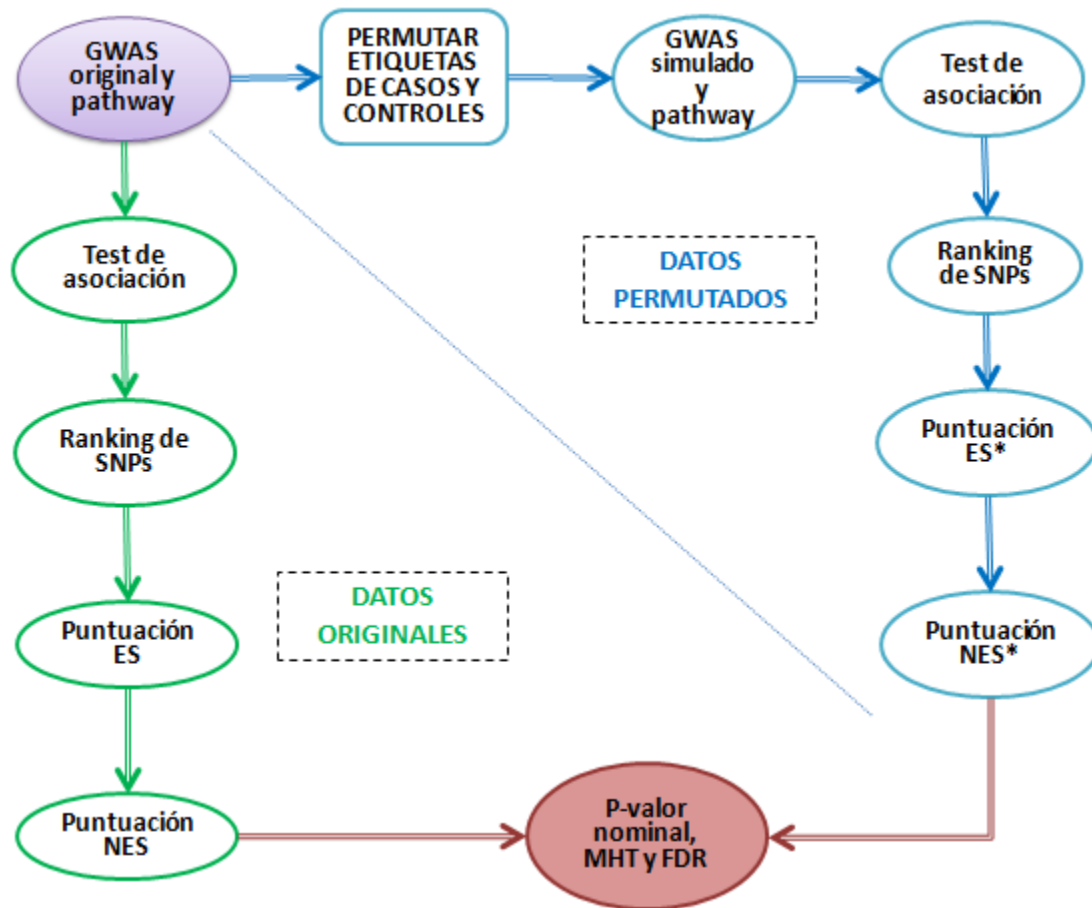


Figura 7. Algoritmo del proceso de ejecución de la metodología "Fast Gene Set Enrichment Analysis".

Pathway of Distinction Analysis (PoDA)

Esta técnica se basa en la hipótesis de que si una ruta está relacionada con el riesgo de padecer una determinada enfermedad, los casos tendrán mayor similitud con otros casos que con los controles para los SNPs asociados con esta ruta y viceversa. En la Figura 8, se observa un esquema explicativo de esta metodología:

El primer paso que se lleva a cabo consiste en generar un ranking con los SNPs pertenecientes a cada gen de tal manera que se selecciona un único SNP por gen (el más representativo) evitando que se produzca epistasis (57). A continuación, se calcula el "Minor Allele Frequency" (MAF) medio para casos y controles de los SNPs del GWAS y se calcula la distancia (D) de los SNPs del pathway al grupo casos y controles, teniendo en cuenta el genotipo de cada individuo para cada SNP (0=homocigoto dominante; 1: heterocigoto; 2; homocigoto recesivo). En la fórmula

que se observa en la Figura 8 para el cálculo de D , y_i corresponde al genotipo de un determinado SNP para un individuo, f_i se refiere al MAF medio de los controles para dicho SNP y g_i , al MAF medio de los casos para dicho SNP. De esta forma, se obtuvo un valor D para cada SNP.

A continuación, teniendo en cuenta los SNPs de cada individuo y el valor D de los mismos, se calcula el valor S , que se trata de un valor promedio y “normalizado” que cuantifica la cercanía de dicha muestra al grupo casos en relación con su distancia al grupo controles teniendo en cuenta la desviación estándar y el número de genes de la ruta. Resumiendo, el valor S , describe cómo es cada sujeto teniendo en cuenta los genotipos de los SNPs. El siguiente paso consiste en realizar el Test estadístico de Wilcoxon (W), para conocer si existen diferencias entre las distribuciones de los casos y los controles mediante la puntuación S de los mismos.

Los pasos explicados hasta ahora se repiten de nuevo tras realizar permutaciones de las etiquetas caso/control, de tal manera que se obtienen tantos valores W^* como permutaciones se hayan programado. Con estos datos, se calcula la puntuación DS_p , que describe si hay diferencias entre la distribución de los casos y los controles referido a una ruta en concreto tal y como se describe en la línea azul. Esta puntuación de cada pathway se obtienen teniendo en cuenta el valor W y la media y desviación estándar de W^* .

La técnica PoDA, además de las permutaciones de las etiquetas de fenotipo, realiza remuestreos en los que, teniendo en cuenta el número original de SNPs de pathway, se seleccionan al azar SNPs del GWAS para conformar nuevas rutas “dummies”. A partir de estas rutas, se repiten los cálculos descritos anteriormente para las rutas originales y, finalmente, se calcula un p-valor teniendo en cuenta los valores DS_p y DS_p^* , es decir, el valor obtenido a partir del pathway original junto con los valores obtenidos tras el remuestreo. El cálculo del p-valor se lleva a cabo teniendo en cuenta el número de veces en el que la puntuación DS_p es mayor o igual a DS_p^* entre el número total de permutaciones.

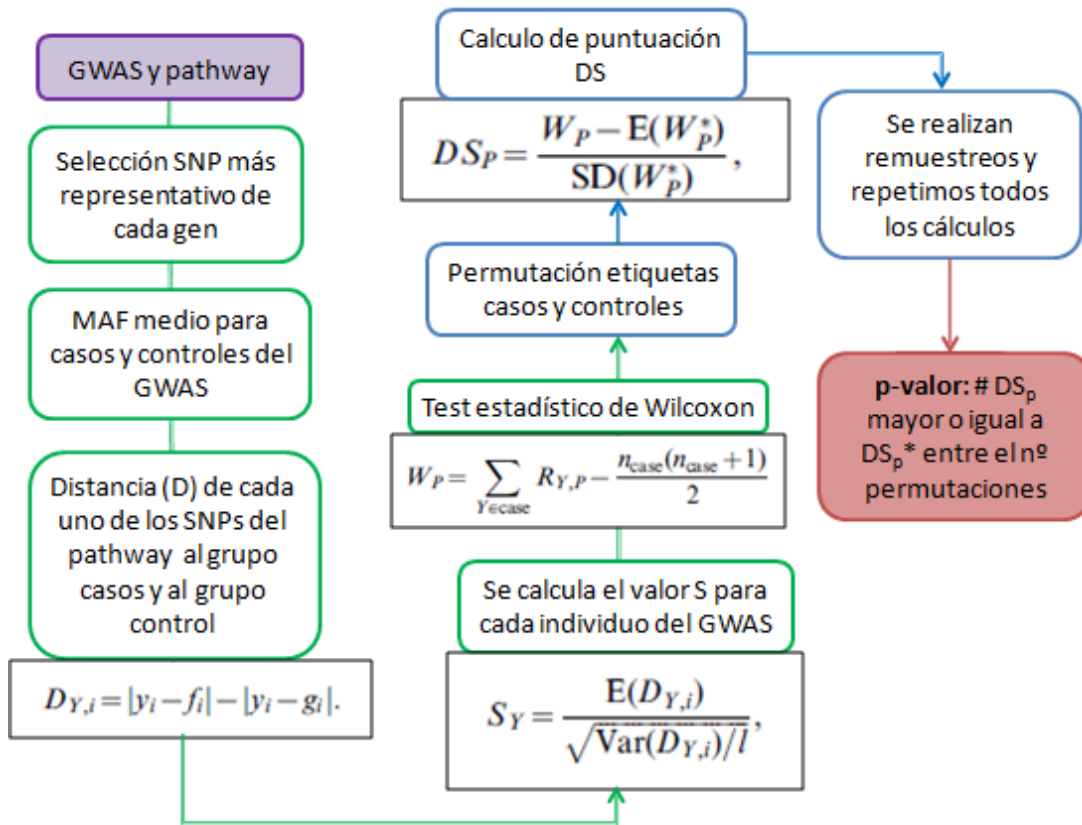


Figura 8. Algoritmo del proceso de ejecución de la metodología "Pathway of Distinction Analysis".

Comparativa de métodos de análisis de pathways

En la Tabla 1, se resumen las características principales de los métodos propuestos.

Tabla 1. Comparación de tres metodologías de análisis de pathways.

| Metodología | SRT | GSEA | PoDA |
|---|-----|------|------|
| Datos: SNPs | ✓ | ✓ | ✓ |
| Permutaciones etiquetas casos/controles | ✓ | ✓ | ✓ |
| Ranking SNPs del GWAS | ✓ | ✓ | ✗ |
| Valor para cada pathway ≠ p-valor | ✗ | ✓ | ✓ |
| Representación gráfica de resultados | ✗ | ✓ | ✓ |
| Distancias genéticas | ✗ | ✗ | ✓ |
| Desequilibrio de ligamiento | ✗ | ✗ | ✓ |
| Remuestros (pathways dummies) | ✗ | ✗ | ✓ |
| Valor para cada ID | ✗ | ✗ | ✓ |

Tanto el SRT como GSEA usa como base un ranking de los SNPs que forman parte del GWAS mientras que el PoDA se basa en distancias genéticas entre casos y controles. El ranking que se genera en el PoDA es referido a cada gen y se lleva a cabo para seleccionar el SNP más representativo de cada gen, siendo esta metodología la única que tienen en cuenta la posibilidad de existe equilibrio de ligamiento.

Las tres metodologías realizan permutaciones de las etiquetas de casos y controles manteniendo la proporción de los mismos para comprobar la veracidad de los datos obtenidos. Sin embargo, la metodología PoDA usa además remuestreos con el objetivo de tener en cuenta la posibilidad de que no sólo esté influyendo el hecho de ser caso o ser control sino también los SNPs componentes del pathway. Finalmente, para calcular el p-valor, las tres metodologías comparan los resultados obtenidos mediante permutaciones (y en el caso de PoDA también mediante remuestreos) con los datos originales.

El punto de corte que se establece en la metodología SRT para los SNPs del ranking considerados significativos es subjetivo y va a hacer que varíen los resultados obtenidos. De esta manera, si se selecciona un p-valor bajo, el número de SNPs seleccionados será menor pero estos se asociarán en mayor medida con el fenotipo evaluado. Al contrario, si el p-valor elegido es alto, el número de SNPs incluidos es mayor aunque su efecto sobre el fenotipo sea menor.

Por otra parte, el GSEA además del p-valor para cada pathway, da a cada una de ellas una puntuación de enriquecimiento que nos informa sobre el grado de asociación de esta con la enfermedad evaluada. Esta puntuación de enriquecimiento se va a normalizar previamente al cálculo del p-valor para que los resultados sean comparables. La metodología PoDA, además de dar un determinado valor a cada pathway como ocurre en el método GSEA, adjudica a cada individuo un valor que informa sobre la tendencia del mismo a ser caso o control (valor S) (57).

A pesar de sus diferencias, sería conveniente realizar las tres metodologías a la hora de realizar análisis de pathways asociadas a una enfermedad para comprobar si los datos son coincidentes. Además, mediante la aplicación de cada uno de los tres métodos se obtienen datos diferentes basados en hipótesis distintas que hacen que se

disponga de una mayor cantidad de información que facilite la interpretación de resultados.

BIOGÉNESIS MITOCONDRIAL Y CCR

Las mitocondrias son orgánulos semiautónomos que participan en el metabolismo energético, la producción de radicales libres y la apoptosis. Aparte del núcleo, la mitocondria es el único orgánulo celular que contiene su propio genoma y maquinaria genética. El ADNmt es una doble hebra de 16,6 kb, que se hereda de la madre, una molécula circular cerrada que codifica 37 genes (58,59). La biogénesis mitocondrial es un proceso esencial mediante el cual se obtienen nuevas mitocondrias, que requiere de la coordinación entre el genoma nuclear y mitocondrial (60).

Las mitocondrias, así como la mayoría de los procesos relacionados con estas, están estrechamente ligados con la génesis del cáncer (61). Por ello, resulta esencial el estudio de la biogénesis mitocondrial, así como conocer qué ocurre con estos orgánulos durante los procesos tumorales.

Hace 70 años, Otto Warburg hipotetizó que las células cancerígenas poseían un metabolismo alterado caracterizado por un incremento del consumo de glucosa y la glicolisis a pesar de un aporte de oxígeno adecuado para la respiración mitocondrial, un fenómeno conocido como "glicolisis aerobia" o "efecto Warburg" (Figura 9) (62). Warburg describió que esta tendencia hacia la glicolisis aerobia significaba una deficiencia en la respiración mitocondrial, representando una causa fundamental del cáncer. Esta observación se ha confirmado ahora en muchos tipos de células cancerígenas (53).

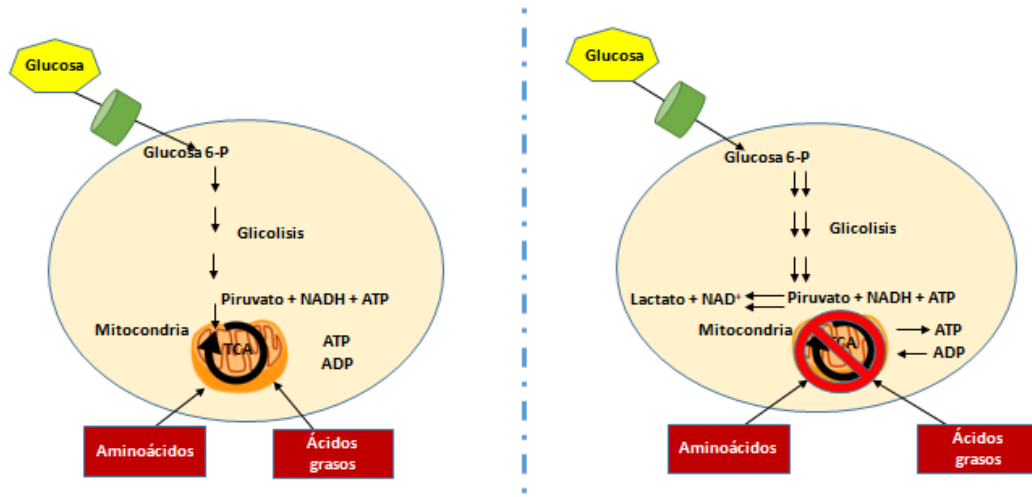


Figura 9. Hipótesis de Warburg en células normales y tumorales.

La provisión de energía metabólica por la fosforilación oxidativa, la ejecución de la muerte celular y la señalización intracelular por calcio (Ca) y especies reactivas de oxígeno (ROS) son las funciones principales de las mitocondrias. Las células que son incapaces de realizar la fosforilación oxidativa y llevan a cabo la glicolisis tienen un fenotipo resistente a la apoptosis. Este fenotipo celular confiere resistencia a la muerte celular y contribuye al crecimiento rápido de la masa tumoral (63). De esta manera, se puede decir que el crecimiento del tumor de una manera más rápida ocurre cuando la fosforilación oxidativa se inhibe y se da la glicolisis. Las mutaciones somáticas en el ADNmt podrían contribuir a defectos en la fosforilación oxidativa y esto podría reflejarse en la producción de ROS y en consecuencia en la progresión del crecimiento tumoral (61). Las células normales usan la fermentación de ácido láctico solo en condiciones anaerobias mientras que las células tumorales muestran un nivel elevado de producción de ácido láctico incluso ante condiciones normales de oxígeno (Figura 10) (62,64).

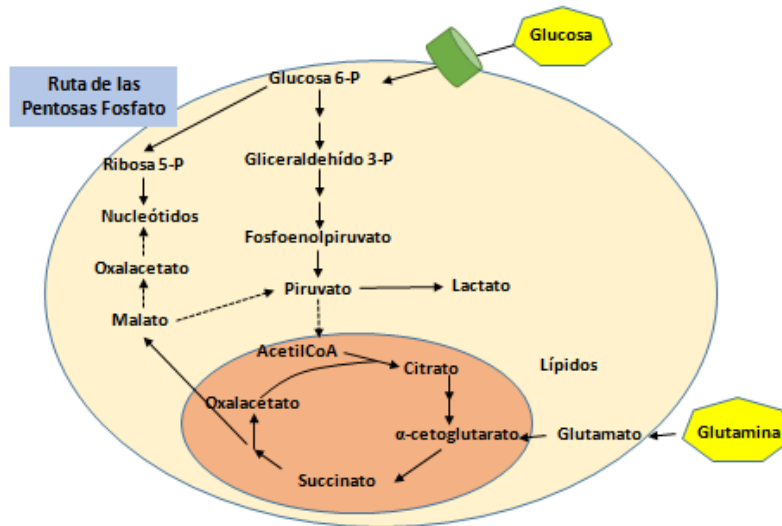


Figura 10. Metabolismo de las células tumorales.

Una ventaja para el crecimiento y la supervivencia podría ser adquirida a través de un segundo evento como alteraciones en el ADN causadas por ROS. Es posible que las mutaciones en el ADNmt causen un cambio funcional de la proteína afectada y de esta manera se disminuye la eficiencia de la cadena de transporte de electrones, que conduce a una bifurcación del flujo de electrones y a un aumento en la generación de ROS. Un aumento moderado de ROS se asocia a un aumento de la proliferación celular y la biogénesis mitocondrial. El incremento de ROS endógeno además da un estímulo constante para la proliferación celular y podría causar daños tanto en el ADN nuclear como en el mitocondrial, conduciendo al desarrollo del cáncer, la inestabilidad genética y la progresión de la enfermedad (Figura 11) (63,65).

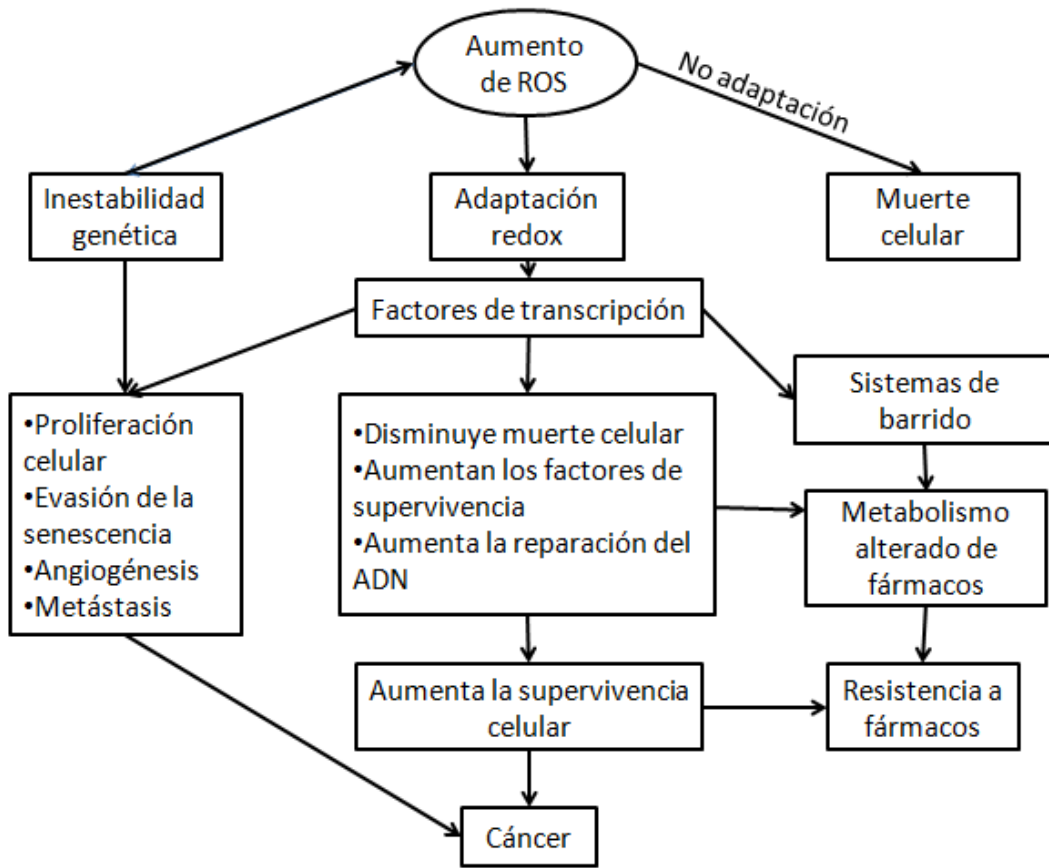


Figura 11. Génesis y desarrollo del cáncer debido a un aumento de ROS.

El CCR representa un sistema modelo ideal para estudiar el desarrollo y la progresión del tumor porque los colonocitos, tras adquirir mutaciones, siguen un proceso sistemático de proliferación, diferenciación y transformación a carcinoma (Figura 12) (66).

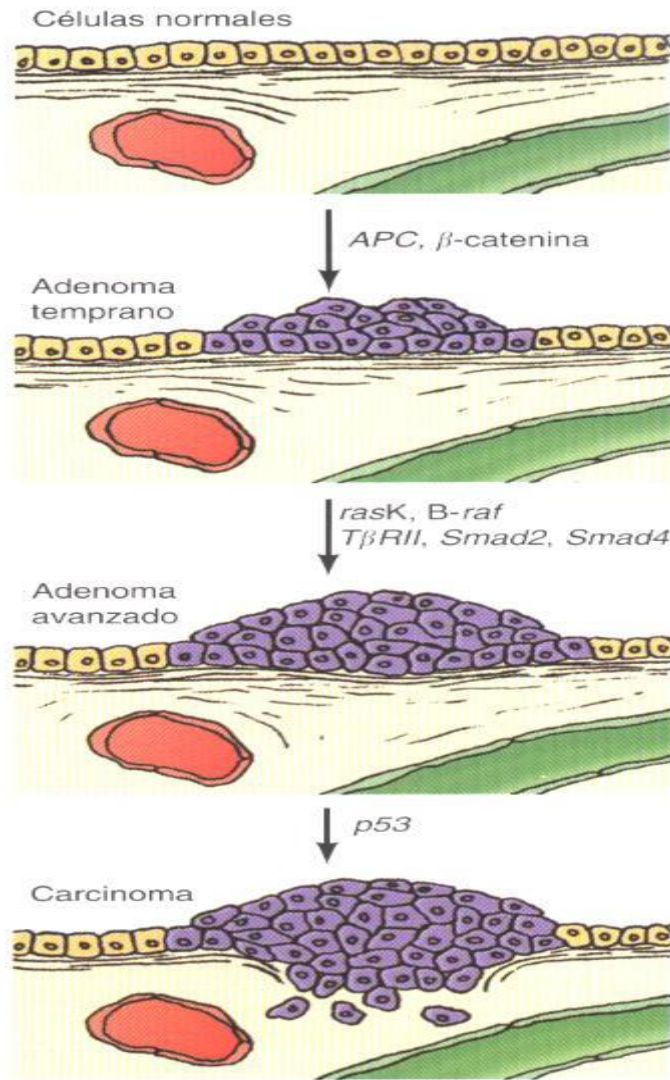


Figura 12. Avance y desarrollo del CCR. Basado en Cooper GM & Hausman RE (2007).

La progresión del CCR en humanos está íntimamente ligada a la alteración mitocondrial, a una mayor producción de óxido-radicales mitocondriales libres y al estrés oxidativo (67). El cáncer de colon comienza con una mutación en el ADN y se expande hasta colonizar una cripta. Luego, mediante fisión, se generan criptas nuevas que mantienen dicha mutación. De esta manera, la mutación continuará permaneciendo en las criptas hijas (68).

La progresión del tumor requiere mitocondrias activas tanto en las células tumorales como en las células adyacentes al tumor. En el tumor, un aumento del ATP glicolítico se asocia con una disminución de la actividad mitocondrial y con la proliferación celular, mientras que en el tejido adyacente la disfunción mitocondrial

se asocia con la apoptosis. Además, se ha observado que las mitocondrias de las células de cáncer de colon humano se encuentran en un estado de estrés oxidativo comparadas con las células del tejido adyacente (67). El estrés oxidativo, tal y como se mencionó anteriormente, también juega un papel fundamental en la carcinogénesis colorrectal. Los cambios cuantitativos y cualitativos del ADN podrán desarrollar un papel significativo en la carcinogénesis. Tanto un mayor como un menor número de copias del ADNmt se asocia con un incremento del riesgo de sufrir CCR, ya que se reduce la respuesta compensatoria al daño causado por estrés oxidativo y exposiciones ambientales (69). Además, estas alteraciones tienen como resultado la pérdida de potencial de transmembrana mitocondrial e induce una señalización retrógrada mitocondrial, dando lugar a una expresión alterada de los genes nucleares (70).

La biogénesis mitocondrial está regulada por los factores de transcripción "Cap 'n' collar" (CNC): Nuclear Respiratory Factor 1 (NRF1) y Nuclear Respiratory Factor 2 (NRF2) (71). En mamíferos, NRF1 y NRF2 se expresan de manera ubicua y representan los principales factores que regulan los genes que conducen a los *Elementos de Respuesta Antioxidante* (ARE) en tejidos no hematológicos cuando heterodimerizan con small Maf (*Mafs*). Ambos controlan la homeostasis y el desarrollo de rutas porque regulan la expresión basal e inducible de una batería de genes ARE, que codifican proteínas antioxidantes, enzimas detoxificantes, enzimas metabólicas y la subunidad proteosomal 26S (72,73). Además, los Nuclear Respiratory Factor (NRFs) activan varios genes nucleares involucrados en múltiples funciones mitocondriales y participan en la regulación de la expresión de genes que se codifican en el ADNmt, consiguiendo proporcionar una unión entre la expresión de genes nucleares y mitocondriales (74).

Nuclear Respiratory Factor 1 (NRF1)

El gen NRF1 en humanos se encuentra en el cromosoma 17q21.3 y ocupa 15 kb de ADN genómico. En tejidos adultos, los niveles más altos de los transcritos de NRF1 se detectan en corazón, músculo esquelético, grasa y cerebro mientras que a niveles más bajos se pueden detectar durante los procesos inflamatorios y en la fase proliferativa temprana en tejido de músculo esquelético en respuesta a un daño (75).

Estructura de NRF1

NRF1 cuenta con diversos dominios tal y como se muestra en la Figura 13. El dominio N-terminal (NTD) contiene una región de unión al retículo endoplasmático (conocido como "N-terminal homology box 1": NHB1) y una región de transmembrana ("N-terminal homology box 2": NHB2). Tanto NHB1 como NHB2 contienen una secuencia de anclaje con el retículo endoplasmático compuestas por un núcleo alfa-helical hidrofóbico que atraviesa la membrana, una región N-terminal básica que dicta la orientación de la inserción de una proteína al retículo endoplasmático y una región hidrofílica C-terminal que contiene un sitio de unión proteolítico. El dominio de transactivación (TADs) interaccionan con la maquinaria de transcripción basal para activar la transcripción. Los TADs de NRF1 están compuestos por los dominios ácidos 1 y 2 (AD1 y AD2). El dominio NST (asparagina / serina / treonina) está situado entre AD1 y AD2 y funciona como un glicodominio en el retículo endoplasmático. El dominio Neh6L (Nrf2-ECH de homología similar a 6) está situado entre el dominio SR (repetición de serina) y el DBD (dominio de unión a ADN), y contribuye a la regulación negativa de NRF1. La región de Neh1L (Nrf2-ECH homología similar a 1) contiene tanto dominios CNC como bZIP y funciones como DBD. Por último, NRF1 tiene un dominio C-terminal similar al dominio Neh3 de NRF2 (Neh3L) (Figura 13) (76).



Figura 13. Estructura de NRF1.

Regulación de NRF1

NRF1 se trata de una proteína integral de membrana direccionada hacia el retículo endoplasmático (RE) (72,76) y se procesa para producir diferentes isoformas (73). En la Figura 14 se puede observar que la regulación de NRF1 podría surgir a nivel transcripcional, traduccional, post-traduccional o por tratamiento químico o condición fisiológica. La regulación transcripcional ocurre cuando NRF1 se activa específicamente mediante "mammalian Target of Rapamycin Complex 1" (mTORC1) e induce la acción proteosomal como respuesta a factores de crecimiento o alimentación. La regulación traduccional se produce mediante splicing alternativo,

generándose distintas variantes de NRF1. La regulación post-traducciona se puede llevar a cabo de 6 maneras diferentes:

- Unión al retículo endoplásmico ("topogénesis"): Después de la orientación en el retículo endoplasmático, la secuencia N-terminal de NRF1 se ancla a la membrana. El siguiente paso sería la N-glicosilación.
- Glicosilación/Deglicosilación: La N-glicosilación reprime la actividad de NRF1 y su deglicosilación conduce a la activación de NRF1.
- Fosforilación: que afecta a la estabilidad de NRF1.
- Ubiquitinación y degradación proteosomal: la activación de NRF1 se inhibe totalmente con la inhibición máxima proteosomal usando altas dosis de inhibidores proteosómicos implicando al proteosoma como un factor esencial en la generación de péptidos NRF1.
- Proteolisis: se han descrito muchos sitios putativos de escisión proteolítica que permite activar los distintos fragmentos.
- Hidroxilación: existen secuencias en NRF1 que permiten su degradación mediante O²- dependientes de hidroxilación.

Por último, existe otro tipo de regulación por compuestos químicos como terbutil hidroquinona (t-BHQ) o condiciones fisiológicas como la hipoxia (76).

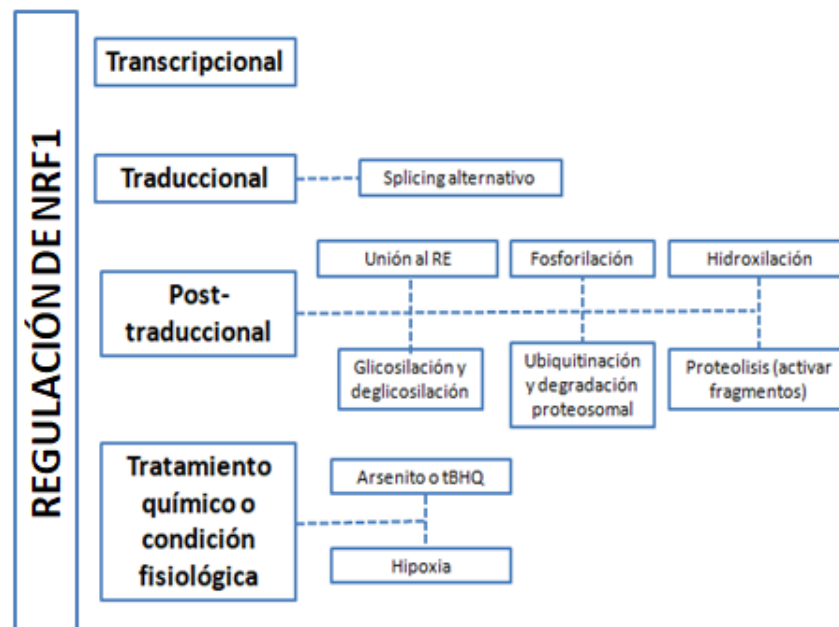


Figura 14. Regulación de la expresión de NRF1.

Función de NRF1

NRF1 es un factor de transcripción que se une al ADN y activa genes que están involucrados en la biogénesis mitocondrial y otras funciones celulares, tales como la traducción de proteínas, la reparación y la síntesis de ADN y la proliferación celular. NRF1 también aumenta la capacidad respiratoria mitocondrial e induce la expresión de un subconjunto de genes que controlan la actividad mitocondrial de una manera específica en cada tipo de célula (77). De esta manera, la expresión de NRF1 es probable que sea vital para la función correcta de la célula y aumenta la posibilidad de que la expresión aberrante de NRF1 promueva el desarrollo de enfermedades entre las que se podría citar al cáncer (78).

NRF1 y el cáncer

Para caracterizar la función de NRF1, se realizaron microarrays en los que se silenció a *NRF1* y se determinó que este gen regula rutas del metabolismo de lípidos, el metabolismo de aminoácidos, la degradación proteosomal, el ciclo de los ácidos tricarbónicos y la cadena respiratoria mitocondrial. Además, la deficiencia de este gen está ligado a una inestabilidad genética (76). La segregación cromosómica correcta durante la mitosis es esencial ya que errores en este proceso dan como resultado una aneuploidía, que podría jugar un papel en la tumorigénesis. De esta manera, la pérdida de NRF1 podría contribuir a la tumorigénesis promoviendo una segregación cromosómica anormal. En el estudio de Oh *et al.*, (2012) observaron que la frecuencia de fenotipos aberrantes mitóticos eran el doble en células con NRF1 "knockdown" comparadas con las células control. Estas anomalías incluyen células con cromosomas de revestimiento, puentes en la anafase y cromosomas no alineados. Por tanto, la falta de NRF1 produce una alteración de la mitosis normal (79) y una baja regulación de proteínas asociadas al centrómero (76) y promueve una tumorigénesis caracterizada por la acumulación de mutaciones de varios *loci*, así como la adquisición de inestabilidad genética (79).

Alteraciones en NRF1 se han relacionado con distintos tipos de cáncer, así como con otras enfermedades que aumentan el riesgo del mismo como, por ejemplo, la diabetes. De esta manera, se ha descrito que la falta de función de NRF1 en hepatocitos conduce a tumores en el hígado (80). Además, se conoce que cuando la expresión de este gen aumenta significativamente se asocia con la progresión del

grado de cáncer de mama (81). La sobreexpresión de NRF1 está relacionada con la proliferación celular descontrolada y con un cambio en la morfología de las mitocondrias (más alargadas y densas) característicos de procesos tumorales (82).

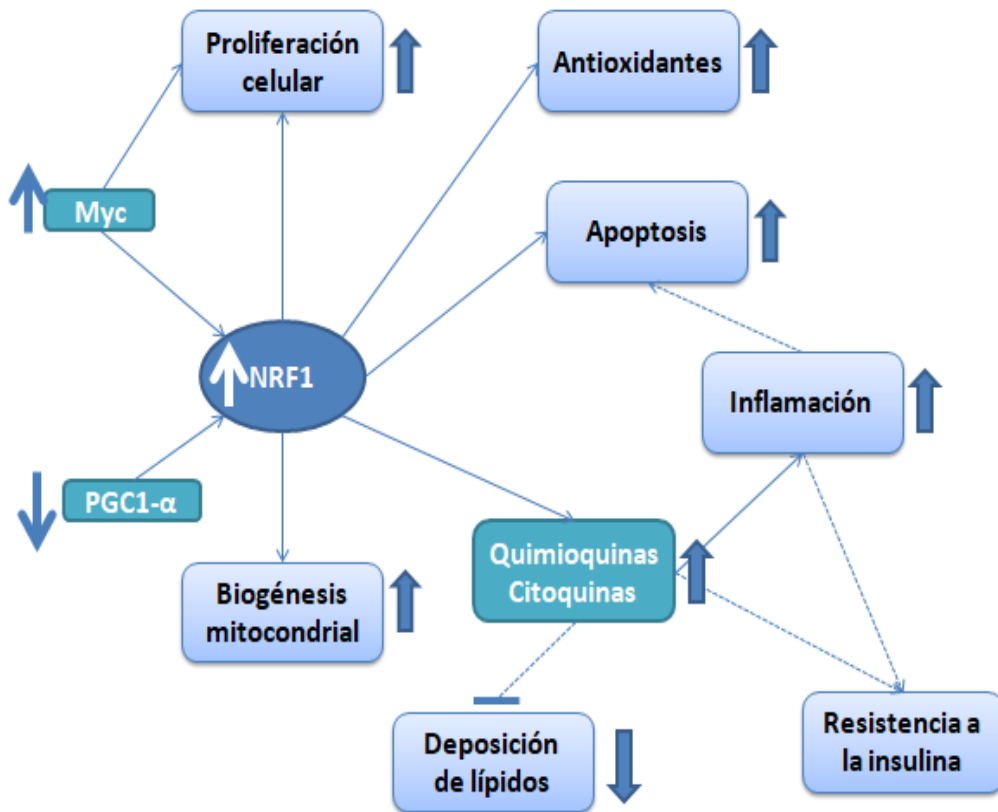


Figura 15. Procesos desencadenados tras un aumento de la expresión de NRF1.

La Figura 15 muestra las consecuencias directas de la sobreexpresión de NRF1, en la cual se observa como un aumento de la expresión de NRF1 está íntimamente relacionado con procesos típicos de un proceso tumoral como la inflamación, la resistencia a la insulina o la proliferación descontrolada (promovida en muchos casos a través de E2 mediante fosforilación).

De igual manera, la Figura 16 refleja los procesos desencadenados ante una pérdida de la función de NRF1, la cual se relaciona con la expresión alterada de genes detoxificantes y antioxidantes asociados con el estrés redox. Por otra parte, la pérdida de la función de NRF1 se relaciona con el estrés proteotóxico mediante la abolición de la expresión del proteosoma 26S, p97/VCP y co-factores. Además, ante

la expresión desregulada de genes que controlan el metabolismo de glúcidos y lípidos se genera estrés metabólico. Las consecuencias finales son estrés en el retículo endoplasmático y una serie de procesos desencadenados a partir de este entre los que se puede citar el ciclo celular, la inflamación y la carcinogénesis (83):

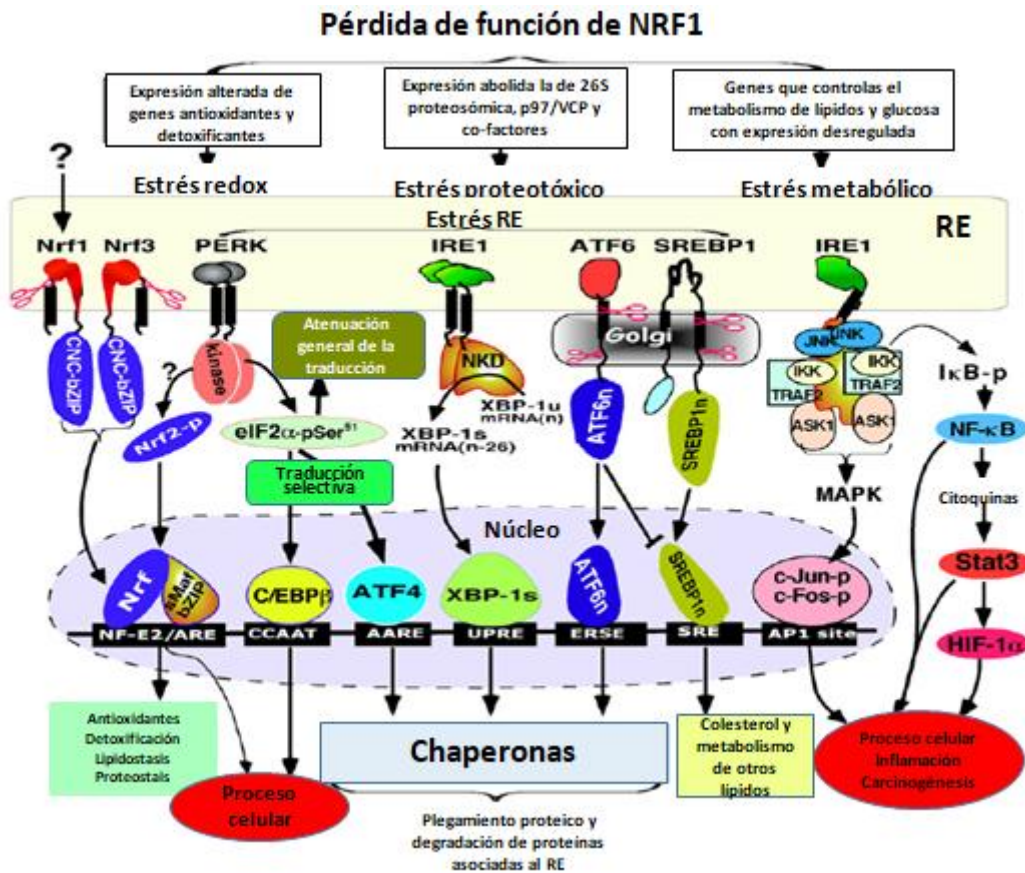


Figura 16. Procesos desencadenados tras la pérdida de función de NRF1. Basada en Permuth-Wey J. et al.(2011).

Por tanto, la presencia de SNPs (que podrían modificar tanto la activación como los niveles de expresión) en este gen podrían aumentar el riesgo de cáncer a través de cambios en las proteínas codificadas, una fosforilación oxidativa alterada y un exceso de producción de ROS (58).

NRF1 y el CCR

A día de hoy, tan solo existen tres estudios que relacionen el factor de transcripción NRF1 y el CCR (60,79,84). El primer artículo se publicó en el 2008 por Blesa *et al.*, y se puede observar como una sobreexpresión del factor de transcripción NRF1, promueve la sobreexpresión de TOMM34 en tumores colorrectales. Esto tiene sentido si observamos qué ocurre ante una sobreexpresión de NRF1: proliferación descontrolada, alteración del metabolismo, inflamación... Más tarde, el estudio llevado a cabo por Oh, 2012 informa de que la pérdida de función de NRF1 conduce a la inestabilidad genética promoviendo el aumento de la formación de micronúcleos y la aneuploidía promoviendo la aparición de cáncer de colon. El último, publicado en el año 2017 por Shukla *et al.*, consiste en un estudio experimental en el que inducen la expresión de NRF1 haciendo uso de un inhibidor de la aldolasa reductasa y observan que previene el CCR disminuyendo el daño mitocondrial gracias a la expresión de este gen junto con otros que participan en la ruta.

El hecho de que solo existan tres artículos sobre este tema en particular hace interesante analizar mutaciones puntuales en este gen recogidas en bases de datos que nos permitan afianzar conocimientos sobre la función que desempeña NRF1 en el CCR, pudiendo asemejarse a lo que ocurre en otros tipos de tumores.

Nuclear Respiratory Factor 2 (NRF2)

El gen *NRF2*, también conocido con el acrónimo "*Nuclear factor (erythroid-derived 2)-like 2*" (*NFE2L2*), en humanos se describió, clonó y caracterizó en 1994 por Moi *et al.*(85). El gen clonado está codificado por un transcrito de 2,2Kb, lo que predice una proteína de 589 aminoácidos (aa) con una masa molecular de 66,1KDa (86). NRF2 ha sido descrito como el principal factor de transcripción que se une a los elementos de respuesta antioxidante (ARE) (11).

Estructura de NRF2

NRF2 pertenece a la familia de CNC de factores de transcripción (87). Contiene 7 dominios altamente conservados que van desde Neh1 hasta Neh7. Neh1 contiene la

estructura CNC-bZIP que promueve la dimerización con sMaf y la unión al ADN. Además, Neh1 se une a las ligasas E2E3 para llevar a cabo la ubiquitinación formando un complejo que regula la estabilidad de NRF2 y su actividad. El dominio de Neh2 es responsable de la regulación negativa de la expresión de genes que conducen a la activación de los ARE mediante la unión a "Kelch-like ECH-associated protein 1" (Keap1). Neh3 promueve la transactivación de NRF2. El dominio Neh4 y Neh5 son dominios de la activación de la transcripción que cooperativamente se une a las proteínas de elementos de respuesta adenosin monofosfato cíclico (cAMP) y facilita la transactivación de NRF2. El dominio Neh6 es esencial para la degradación de NRF2 independiente a Keap1. El dominio Neh7 se asocia con el dominio de unión a "El receptor X retinoide alfa" (RXRalpha) e inhibe la expresión de genes mediados por ARE. NRF2 contiene motivos nucleares, tales como el motivo señal para la exportación nuclear y señales de localización nuclear en diferentes dominios, que hacen funciones importantes en la translocación nuclear y la localización de NRF2 (*Figura 17*) (88).

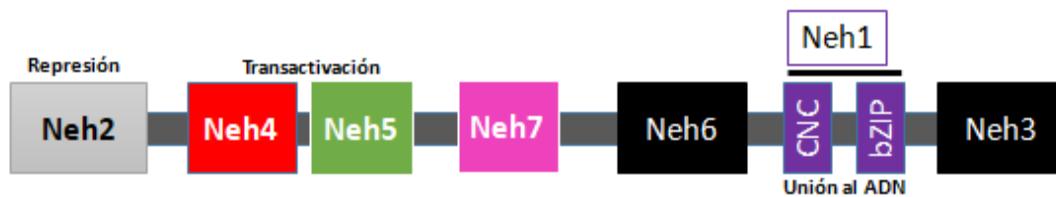


Figura 17. Esquema de la proteína NRF2, señalizando los siete dominios (Neh1-Neh7).

Regulación de NRF2

NRF2 se puede regular a nivel transcripcional y post-transcripcional. La regulación transcripcional se lleva a cabo a partir de oncogenes mediante la fosforilación de NRF2 por kinasas que se encuentran *aguas arriba*. La regulación post-transcripcional puede depender o no de Keap1. La regulación post-transcripcional dependiente de Keap1 se lleva a cabo mediante la ubiquitinación y degradación proteosomal de NRF2 o mediante la degradación autofágica de Keap1(86). Normalmente, se encuentra reprimido en el citosol unido a Keap1 y ante la presencia de inductores de ARE, el sistema se perturba haciendo que NRF2 se acumule en el núcleo (89). Allí, dimeriza con small Maf y se une a los ARE (87) Cuando la regulación post-transcripcional es independiente de Keap1, interviene "Glycogen synthase kinase 3 beta" (GSK-3beta) o Hrd1 (86).

Función de NRF2

NRF2 actúa como un regulador clave en la respuesta antioxidante, involucrado en la preservación de la estructura y la funcionalidad de las células normales sanas. Este factor de transcripción desempeña un papel fundamental en la regulación del estado redox celular, controlando tanto la expresión de antioxidantes y enzimas metabolizantes de drogas en la fase I y II como los transportadores de proteínas asociados a la resistencia a multifármacos. Las células tumorales muestran un incremento en la tasa de producción de ROS y toman ventaja sobreactivando la defensa antioxidante, en particular genes dependientes de NRF2 (90). Esta habilidad para adaptar y sobrevivir es dependiente de la expresión de un complejo comprendido por aproximadamente 500 genes, inducidos por NRF2 (86). La ruta de NRF2 es uno de los principales mecanismos de defensa contra el estrés oxidativo y contribuye a la protección contra varias patologías, entre las que se encuentran la carcinogénesis, toxicidad en el hígado, enfermedades respiratorias e inflamación (junto con el factor de transcripción "Factor nuclear potenciador de las cadenas ligeras kappa de las células β activadas" (NFkB) y otras citoquinas proinflamatorias *aguas abajo*) (91,92). NRF2 también actúa como un estimulante de proteínas anti-apoptóticas de la familia de " β -cell lymphoma 2" (Bcl-2) (91).

Además de genes antioxidantes y detoxificantes, NRF2 regula genes involucrados en el control metabólico de la célula y genes involucrados en la reparación y degradación de moléculas dañadas (47). El papel de este factor en la reprogramación metabólica es particularmente relevante. Este es el proceso a través del cual las células quiescentes cambian su metabolismo para entrar en un estado proliferativo en el cual necesitan tomar abundantes nutrientes para utilizar sus metabolitos en rutas anabólicas. NRF2 colabora con la ruta de señalización "*fosfoinositol-3-kinasa/AKT serine/threonine kinase 1*" (PI3K/AKT) para aumentar el metabolismo a través de la ruta de las pentosas fosfato y dirigir los metabolitos de la glucosa hacia la síntesis de nucleótidos *de novo*. NRF2 también induce la síntesis de glutatión (GSH) y nicotinamida adenina dinucleótido fosfato (NADPH) y regula genes involucrados en el metabolismo de lípidos. La disponibilidad de mayor número de metabolitos facilita la proliferación celular, en la cual puede tomar parte el crecimiento tumoral inducido por NRF2 (Figura 18) (87).

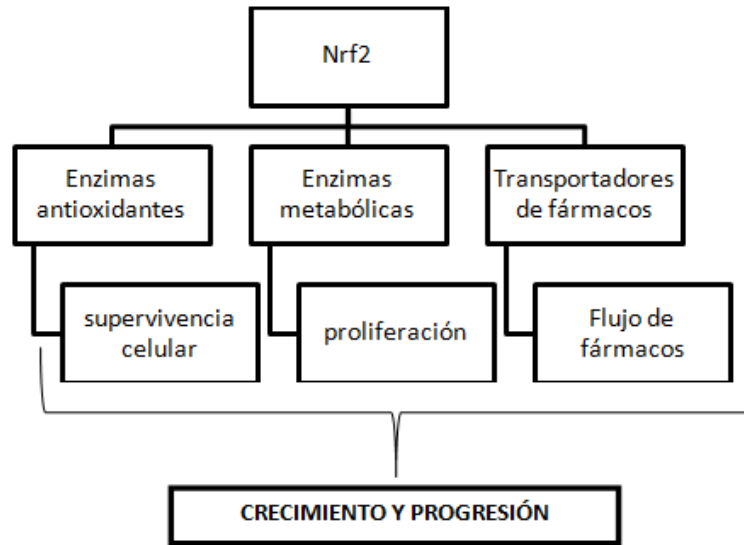


Figura 18. Relación entre NRF2 y el crecimiento y la progresión tumoral.

Estudios recientes sugieren que NRF2 tiene una función importante en el control de la proliferación y la homeostasis de células madre adultas y células madre hematopoyéticas e intestinales. La inhibición de NRF2 directamente altera la renovación de células madre humanas y la reprogramación celular, mientras que la activación de NRF2 promueve la diferenciación de este tipo de células(93). El proteosoma juega un papel clave en la proliferación de la las células madre auto renovadas, la diferenciación de las tres líneas germinales y la reprogramación celular, sugiriendo que la alta actividad proteosómica de las células madre de humano dependiente de NRF2 es una característica funcional importante (87). Uniendo todo esto, los datos sugieren una relación entre la ruta proteosoma-NRF2 y las marcas de las células madre humanas, pluripotencialidad y auto renovación (93).

NRF2 y el cáncer

NRF2 ha sido tradicionalmente considerado como un supresor de tumores (87). Sin embargo, dependiendo del tumor y el estadio, *NRF2* desempeña un papel favorecedor en la génesis y la progresión del tumor (94).

Bajo condiciones patológicas, la regulación de NRF2 es muy susceptible de ser alterada. Esto podría tener como resultado la pérdida de la respuesta ante la

presencia de agentes estresantes y la subsecuente vulnerabilidad de la célula a dichos agentes. La regulación de NRF2 podría desequilibrarse hacia la pérdida de la naturaleza inducible de NRF2 y la adquisición de un fenotipo activo constitutivamente. Esto confiere a la célula una ventaja en cuanto a la supervivencia ante condiciones adversas, la cual podría convertirse en un serio problema en el contexto de la patología del cáncer y su tratamiento (Figura 19). La activación constitutiva o aumentada de la ruta de NRF2, en la que este se comportaría como un proto-oncogen, podría ser decisiva para el destino celular durante la tumorigénesis y afectar a la respuesta a la quimioterapia (87).

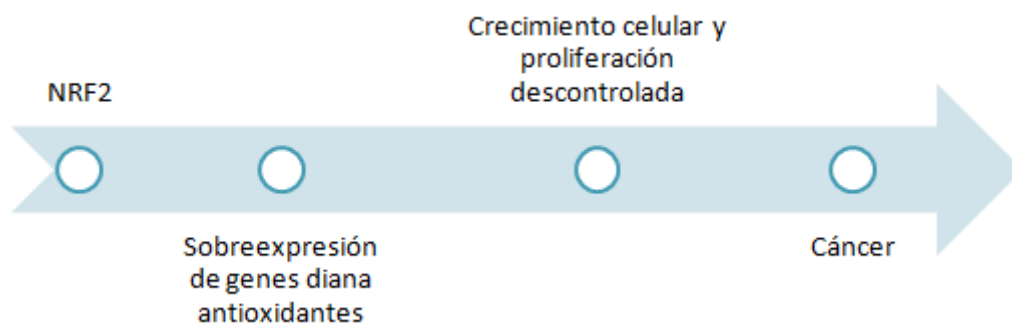


Figura 19. Génesis y desarrollo del cáncer tras la sobreexpresión de NRF2

Los cambios genéticos dan como resultado una expresión de los niveles proteicos de NRF2 anormalmente alta y la expresión de los genes de ARE producen un aumento de la resistencia al estrés oxidativo y una reprogramación metabólica característica de un ambiente tumoral (94).

La bibliografía científica muestra que las mutaciones somáticas en los genes que codifican para KEAP1 o NRF2 se encuentran solo en el 10% o 20% de ciertos tumores, mientras que patrones genéticos de NRF2 parecen estar más frecuentemente sobrerregulados en cáncer. Este hecho sugiere que otros mecanismos, además de la pérdida de la represión de NRF2 por KEAP1, podrían contribuir a la carcinogénesis (94).

Es posible que NRF2 pueda estar sobrerregulado en tumores a través de la activación de AKT y la inactivación permanente de GSK-3. De hecho, la participación de PI3K y su efector AKT *aguas abajo* en la activación de NRF2 se ha

evaluado y la capacidad de PI3K para reprogramar metabólicamente a las células tumorales a través de la activación de NRF2 también ha sido descrita.

"*Phosphatase and tensin homolog*" (PTEN), al contrario, participa en la represión de NRF2 (94). Cuando la función de PTEN se pierde, se ha demostrado que NRF2 participa en la carcinogénesis de una gran cantidad de tumores humanos, de tal manera que NRF2 ha sido reconocido como un factor de supervivencia para la proliferación descontrolada en células tumorogénicas (88,94).

Por tanto, la inducción de genes metabólicos por NRF2 presenta dos requerimientos: la ruta PI3K/AKT activa y una acumulación de NRF2 superior a la necesaria para la transcripción de genes diana antioxidantes. La sobreexpresión o la hiperactivación aberrante de NRF2 en células con una alta actividad de PI3K/AKT podría conducir a un fenotipo hiperproliferativo y, eventualmente, al crecimiento tumoral (87).

También existe relación entre NRF2 y la expresión de E-cadherina. La pérdida de la E-cadherina podría favorecer un aumento de la translocación nuclear de NRF2 y así conferir a las células mecanismos adicionales de supervivencias y proliferación que podrían contribuir a la progresión del tumor (87).

Hasta ahora, se ha hablado de sobreexpresión de NRF2 y su relación con el cáncer. Sin embargo, cuando la expresión de NRF2 es baja, se reprime la proliferación de células tumorales y se prolonga en el tiempo el crecimiento del tumor (95). De esta manera se puede decir que la actividad de NRF2 es deseable en estadios tempranos de la tumorogénesis, cuando el hospedador está intentando controlar la carcinogénesis induciendo la expresión de genes anti-inflamatorios y anti-oxidativos. De esta manera, se genera una citoprotección que tiene como resultado la supresión de la promoción y la progresión del tumor. Sin embargo, ocurre lo contrario en los estadios tardíos del cáncer, donde la actividad de NRF2 de forma permanente o constitutiva debido a la presencia de mutaciones podría convertir a las células en resistentes a la quimioterapia promoviendo una supervivencia de células tumorales. Los efectos de los inductores de NRF2 en estadios intermedios se desconocen y deben ser estudiados (Figura 20) (96).

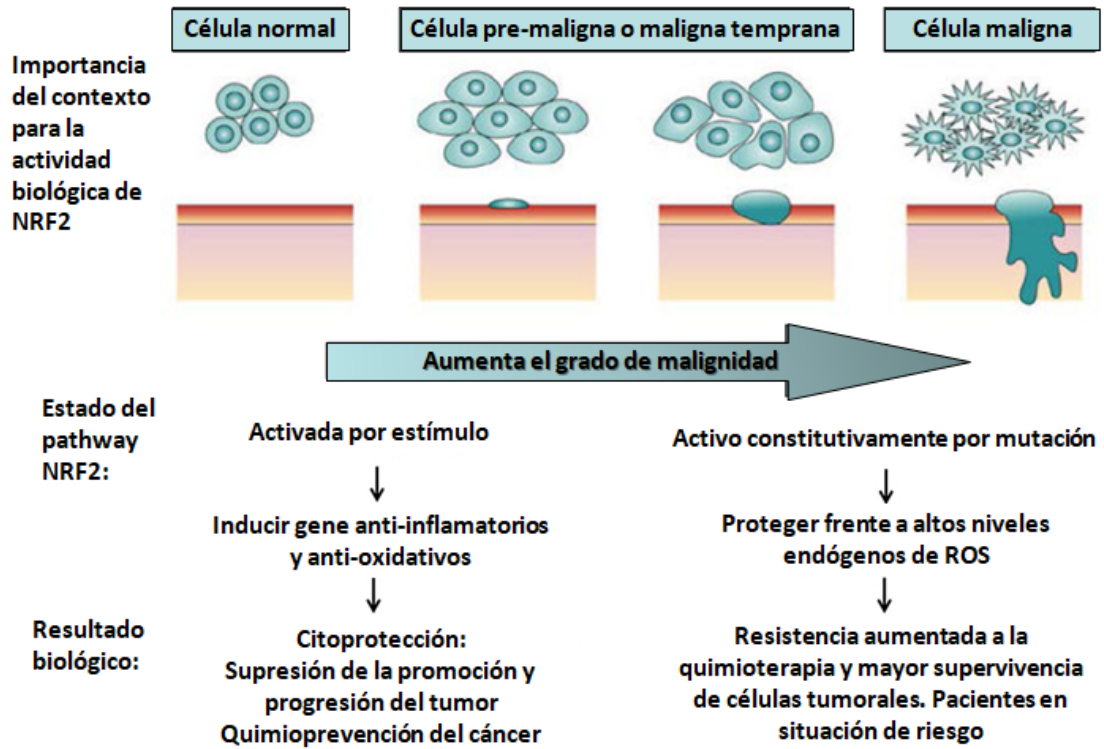


Figura 20. NRF2: de gen supresor de tumores a proto-oncogen.

A modo de resumen, NRF2 se puede activar por numerosos compuestos y está de manera frecuente desregulado en una gran cantidad de cánceres por mutaciones, regulación epigenética o por modificaciones post-traduccionales aberrantes. También puede ocurrir que están hiperactivadas las rutas de señalización oncogénicas que involucran a otros factores de transcripción. La sobreexpresión o hiperactivación de NRF2 puede participar en la tumorigénesis ayudando a las células a escapar de diversas formas de estrés oxidativo a través de la inducción de genes diana antioxidantes o promoviendo la proliferación, supervivencia y metástasis de las células tumorales (87).

NRF2 y el CCR

La inducción de genes diana de NRF2 ha sido relacionada con la prevención de patologías gastrointestinales. De esta manera, NRF2 se considera como un factor crítico para la citoprotección por su capacidad para sobreregular la actividad y la expresión de numerosas enzimas antioxidantes/detoxificantes a través de la

interacción con ARE (89). Sin embargo, las células tumorales pueden modificar la ruta de NRF2 para protegerse del estrés celular asociado con el aumento de la proliferación, el ambiente en hipoxia y la síntesis de proteínas desreguladas. Por tanto, la diferencia entre el tejido tumoral y el tejido normal en el perfil de la expresión de la ruta de NRF2 podría ofrecer una nueva dirección en el tratamiento del CCR (92).

En el intestino, NRF2 induce a los genes metabólicos y su acumulación promueve la proliferación celular que conduce a una hipertrofia epitelial. Considerando que el tejido epitelial del intestino está compuesto por células proliferativas, se asume que NRF2 podría efectivamente activar a genes metabólicos en presencia de señales proliferativas cuando sus niveles son superiores a los necesarios para que este factor de transcripción aporte citoprotección activando genes diana (97). Por otro lado, ante la ausencia de NRF2 se aumenta el daño en las células de las criptas del colon y se produce un aumento tanto en número como en tamaño de los pólipos (98). Como se puede observar, una alteración en la activación o expresión de NRF2 tiene consecuencias sobre la génesis o progresión de CCR. Además, en algunos casos lo que ocurre es que la activación de manera constitutiva de este gen en estadios tardíos del CCR conlleva a la reducción de la efectividad del tratamiento con agentes quimioterapéuticos como el 5-fluorouracilo (5-FU) (99).

La necesidad de conocer más la función de los NRFs en el CCR y el papel que desempeñan a partir de distintas vías metabólicas en la enfermedad, hace que se planteen diversas preguntas que corresponden a cada uno de los objetivos descritos en esta tesis

CAPÍTULO II. OBJETIVOS

Por todo lo descrito hasta ahora, se planteó la siguiente hipótesis: Modificaciones en la expresión o actividad de los genes Nuclear Respiratory Factors (NRF1 y NRF2) que participan en la Biogénesis Mitocondrial pueden influir en la aparición y desarrollo del CCR.

OBJETIVO GENERAL:

Analizar la relación entre los genes NRF1 y NRF2 y el CCR e identificar variantes genéticas y rutas metabólicas relacionadas con estos genes asociadas con el riesgo de CCR.

Objetivos específicos:

1. Caracterizar la relación entre los genes NRF1 y NRF2 y el CCR.
2. Evaluar la asociación de SNPs de los genes NRF1 y NRF2 con el CCR y factores ambientales.
3. Identificar rutas metabólicas relacionadas con NRF1 y NRF2 que se asocien con el riesgo de CCR.

CAPÍTULO III. PLAN DE TRABAJO

PLAN DE TRABAJO

De acuerdo a los objetivos planteados, se llevó a cabo un estudio que cuenta con tres partes bien diferenciadas (Figura 21).

La primera parte trató de identificar la relación existente entre los genes de estudio (NRF1 y NRF2) y el CCR a través de la elaboración de una revisión sistemática de los artículos publicados en este tema, lo que dio lugar al Capítulo III.

Una vez descrita esta relación, se llevó a cabo la identificación de SNPs pertenecientes a estos dos genes y recogidos en la base de datos del MCC-Spain. Tras un control de calidad, se procedió al análisis de los SNPs seleccionados de manera individual para conocer el efecto sobre el CCR y la posible interacción/asociación con diversas variables ambientales estudiadas, lo que se describió en el Capítulo IV.

La cuarta y última parte llevada a cabo en esta tesis consistió en evaluar conjuntos de genes con relación conocida o rutas metabólicas en las que participaran los genes de estudio directa o indirectamente haciendo uso de 3 metodologías diferentes: SNP Ratio Test, fast Gene Set Enrichment Analysis y Pathway of Distinction Analysis, lo que dio lugar al Capítulo V.

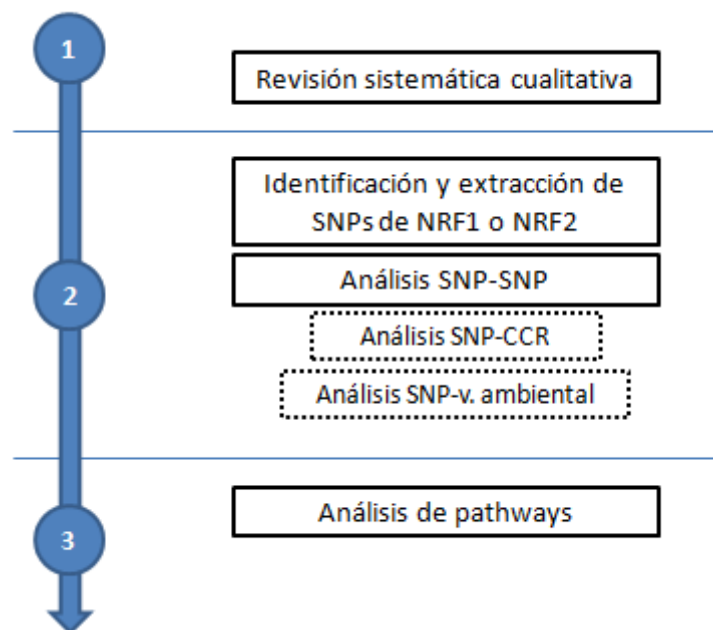


Figura 21. Plan de trabajo.

CAPÍTULO IV. MAPAS CONCEPTUALES DE NRF1 Y NRF2 EN EL CCR

METODOLOGÍA

Para la elaboración de la revisión sistemática, se ha tomado como referencia las consideraciones aportadas por PRISMA. Se tuvieron en cuenta el tutorial de Cochrane para realizar revisiones sistemáticas y la guía para llevar a cabo la síntesis narrativa en revisiones sistemáticas de Popay (100,101). A partir de los tres protocolos mencionados, se elaboró un tutorial propio para la realización de revisiones sistemáticas cualitativas que se adjunta en el Anexo II.

El tutorial planteado cuenta con los siguientes apartados que fueron los que se emplearon en la elaboración de la revisión sistemática cualitativa: el establecimiento de las bases de datos en las que se va a realizar la búsqueda, el diseño de la estrategia de búsqueda, la selección de estudios contrastando resultados con búsquedas realizadas por terceras personas, la síntesis preliminar, la exploración de la relación entre los papers incluidos y, por último, la reflexión crítica junto con la conclusión.

A continuación, se desarrollan cada uno de los puntos anteriormente mencionados para la presente revisión:

Establecer las bases de datos en las que se va a realizar la búsqueda.

Se realizaron búsquedas bibliográficas en distintas bases de datos online: Pubmed, Scopus, Web of Science, Scielo, Cochrane Library y PsycINFO. Las bases de datos que contenían información relevante de acuerdo a la temática tratada en esta tesis son Web of Science (WOS), Pubmed y Scopus.

Diseño de la estrategia de búsqueda

Para determinar los términos de búsqueda fue preciso tener en cuenta sinónimos, términos relacionados y variaciones ortográficas de los mismos. Además, fue importante hacer un buen uso de los operadores booleanos (AND, OR, NOT) para realizar búsquedas combinadas de los distintos términos de búsqueda de una manera óptima.

En esta revisión, los términos de búsqueda que se han utilizado en WOS y Pubmed son:

En el caso de NRF1:

1. '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 'NRF1 OR NFE1L1 OR Nuclear Respiratory Factor 1 OR Nuclear Factor Erythroid 1 Like 1'.
2. 'Colorectal OR Colon OR Rectum OR Rectal' AND 'Cancer OR Carcinoma OR Tumor OR Tumour OR Neoplasm OR Cancer Cells' AND 'Warburg effect' AND 'NRF1 OR NFE1L1 OR Nuclear Respiratory Factor 1 OR Nuclear Factor Erythroid 1 Like 1'.
3. 'Colorectal OR Colon OR Rectum OR Rectal' AND 'Cancer OR Carcinoma OR Tumor OR Tumour OR Neoplasm OR Cancer Cells' AND 'Oxidative phosphorylation OR OXPHOS OR Anaerobic glycolysis' AND 'NRF1 OR NFE1L1 OR Nuclear Respiratory Factor 1 OR Nuclear Factor Erythroid 1 Like 1'.

Cuando se usó la base de datos Scopus, para conseguir resultados, es necesario modificar los criterios de búsqueda:

1. 'CRC' AND 'NRF1 OR NFE1L1' AND 'Warburg'
2. 'CRC' AND 'NRF1 OR NFE1L1' AND 'OXPHOS'.

En el caso de NRF2:

1. '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 'NRF2 OR NFE2L2 OR Nuclear Respiratory Factor 2 OR Nuclear Factor Erythroid 2 Like 2'.
2. 'Colorectal OR Colon OR Rectum OR Rectal' AND 'Cancer OR Carcinoma OR Tumor OR Tumour OR Neoplasm OR Cancer Cells' AND 'Warburg effect' AND 'NRF2 OR NFE2L2 OR Nuclear Respiratory Factor 2 OR Nuclear Factor Erythroid 2 Like 2'.

3. 'Colorectal OR Colon OR Rectum OR Rectal' AND 'Cancer OR Carcinoma OR Tumor OR Tumour OR Neoplasm OR Cancer Cells' AND 'Oxidative phosphorylation OR OXPHOS OR Anaerobic glycolysis' AND 'NRF2 OR NFE2L2 OR Nuclear Respiratory Factor 2 OR Nuclear Factor Erythroid 2 Like 2'.

Cuando se usa la base de datos Scopus, para conseguir resultados, fue necesario modificar los criterios de búsqueda:

1. 'CRC' AND 'NRF2 OR NFE2L2' AND 'Warburg'
2. 'CRC' AND 'NRF2 OR NFE2L2' AND 'OXPHOS'.

Una vez que se realizó la búsqueda, se eliminaron las referencias bibliográficas duplicadas haciendo uso del gestor bibliográfico Mendeley (102).

Criterios de inclusión y exclusión

En primer lugar, se seleccionaron filtros de búsqueda para incluir tan sólo artículos en inglés publicados posteriormente a 2006. Además, se excluyeron las revisiones sistemáticas que ya existían sobre esta temática.

En segundo lugar, se examinaron los títulos y resúmenes para eliminar los artículos que no cumplieran los criterios de inclusión. Las características de exclusión principales fueron las siguientes: que en el título no presentara al menos una de las palabras clave de búsqueda y que no se estudiara la relación entre NRF1/NRF2 y el CCR.

Aquellos artículos que cumplieran las características de inclusión se leyeron en profundidad para posteriormente vincular y reunir aquellos que finalmente podían responder al objetivo de esta revisión sistemática cualitativa. Con ello se procedió a la obtención de datos relevantes de cada estudio con el objetivo de realizar una síntesis preliminar a partir de una tabulación de los mismos.

Explorar la relación entre artículos

Para explorar la relación entre los artículos, se han seleccionados dos técnicas diferentes. Por un lado, se ha evaluado la heterogeneidad mediante un gráfico de barras donde se pueden observar qué características establecidas contiene cada

publicación incluida en la tesis. Para ello, se eligen una serie de características comunes de los artículos seleccionados y se determina si están presentes en cada una de los artículos. En este caso, las características a tener en cuenta son las siguientes: estudiar procesos apoptóticos o inflamatorios, estudiar la función mitocondrial, considerar a los NRFs como quimiopreventivos, estudios *in vitro* y el estudio del CCR y NRF2 de manera directa y no a partir de otras enfermedades o genes relacionados.

Por otra parte, se realizaron mapas conceptuales que informaron de manera visual acerca de las relaciones existentes en esta revisión sistemática.

RESULTADOS

Tras realizar la selección de artículos, se realizó un algoritmo de selección para cada uno de los genes. La Figura 22, muestra el algoritmo de selección de artículos en los que se relaciona NRF1 con el CCR. Como se puede observar, tras la eliminación de duplicados, tan solo hay tres artículos que contemplen esta temática.

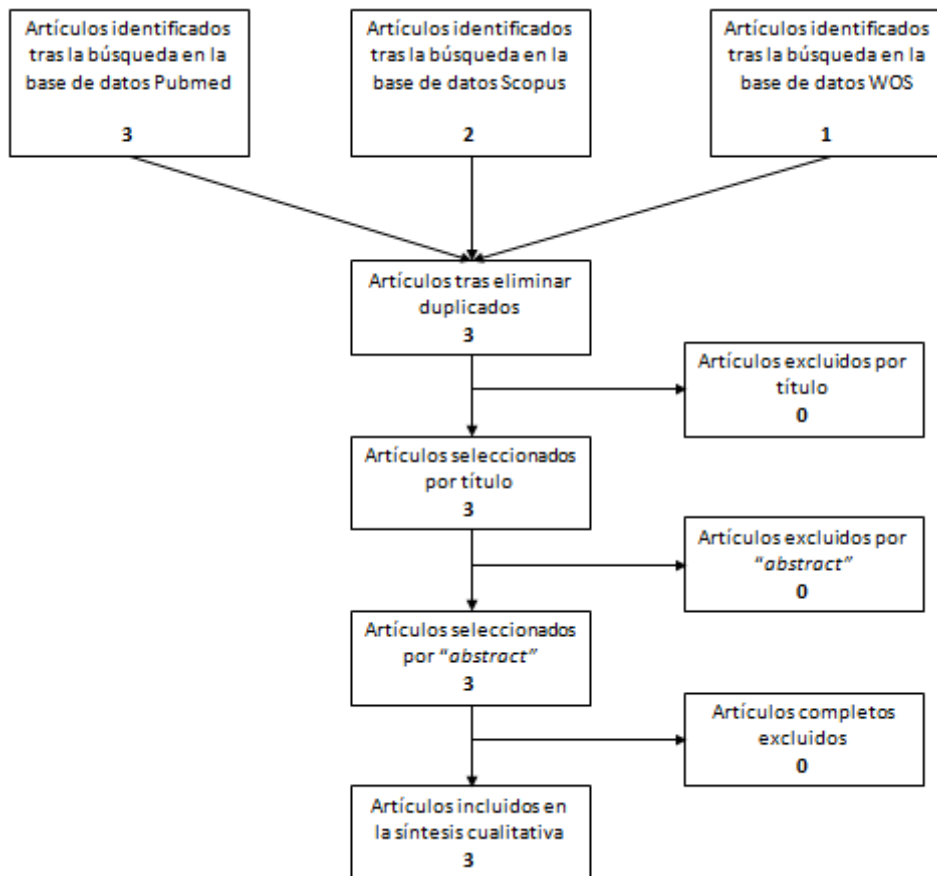


Figura 22. Proceso de selección de artículos relacionados con NRF1 basado en la guía PRISMA.

El segundo algoritmo de selección corresponde a la búsqueda relacionada con el gen NRF2 (Figura 23). Como se puede observar, tras la eliminación de duplicados, se cuenta con un total de 104 artículos. A continuación, tras la lectura del título y teniendo en cuenta las características de inclusión referenciadas, se excluyeron 22 artículos. De los 82 artículos restantes, se seleccionaron 60 tras la lectura del abstract. Finalmente, tras la lectura completa de los artículos, se excluyeron 20 y se procedió a la síntesis de la revisión sistemática cualitativa con los 40 artículos restantes, mostrándose los datos más relevantes de cada uno de ellos: Primer autor y

año de publicación, Organismo o tipo celular, Resultados primarios y conclusiones. Los 20 artículos que no fueron incluidos en la revisión se tabularon junto con el motivo de exclusión (Anexo III).

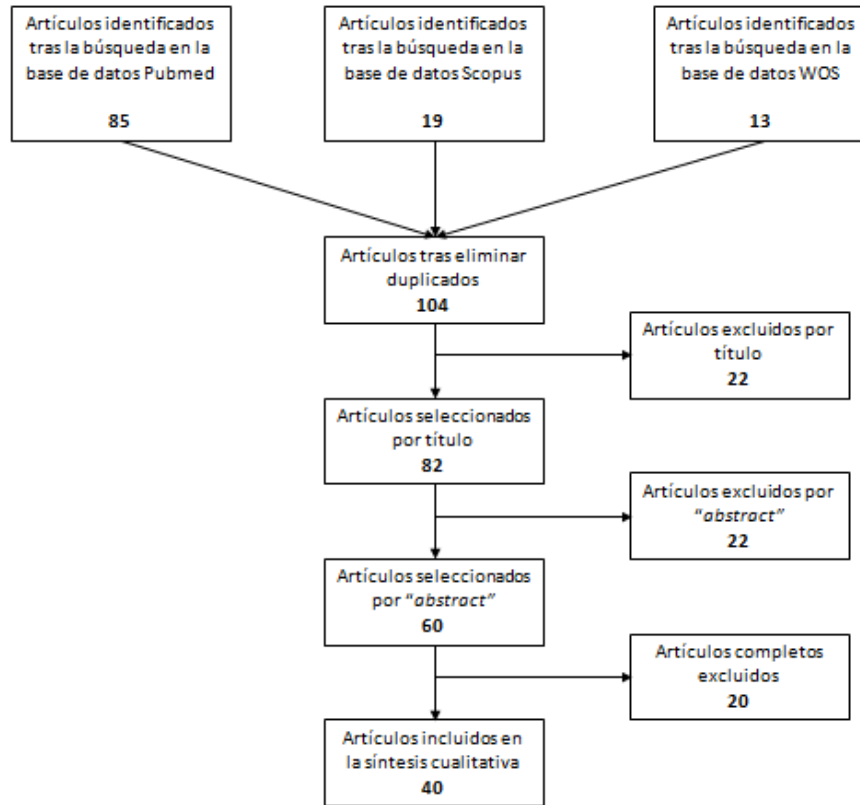


Figura 23. Proceso de selección de artículos relacionados con NRF2 basado en la guía PRISMA.

De acuerdo a los artículos seleccionados, se realizaron tres tablas, basadas en la actividad o expresión de NRF2. De esta manera, la Tabla 2 contenía los artículos en los que NRF2 tenía un patrón de expresión estándar, la Tabla 3 describía los artículos que evaluaban consecuencias de la sobreexpresión de este gen y, por último, en la Tabla 4 se identificaron aquellos artículos en los que NRF2 se encontraba inhibido total o parcialmente.

Tabla 2. Características de los artículos incluidos en la revisión sistemática cualitativa en los que NRF2 sigue un patrón de expresión estándar.

| Primer autor, año | Organismo / Tipo celular | Resultados primarios | Conclusión |
|-------------------|-----------------------------------|--|---|
| Jang, 2016 | Cultivo celular (HT29) | <ul style="list-style-type: none"> • La simvastatina induce la expresión de NRF2 y su traslocación nuclear. • La simvastatina induce la expresión de antioxidantes relacionados con NRF2. | <ul style="list-style-type: none"> • La simvastatina activa a NRF2 y promueve su traslocación al núcleo. • La expresión de antioxidantes relacionados con HO-1 se induce a través de la ruta ERK y PI3K/AKT en células HT29. • La simvastatina suprime la proliferación celular y aumenta la apoptosis en células de cáncer de colon. |
| Kruse, 2016 | Cultivo celular (NCM60) | <ul style="list-style-type: none"> • Las células tumorales NCM460 o Colo320 IMC-co-cultivadas son menos sensibles a la apoptosis debido a la actividad proteosómica inducida por NRF2. • La inmunostatina de tejidos "IBD" confirman la activación de NRF2 en áreas inflamadas del epitelio colónico y más expresión de proteínas proteosómicas. | <ul style="list-style-type: none"> • La adaptación de las células epiteliales del colon al estrés oxidativo causado por células inflamatorias es dependiente de NRF2. • Este proceso incluye un aumento de la actividad proteosómica y resistencia a la apoptosis (protegiendo contra el daño inducido por colitis y favoreciendo la carcinogénesis). |
| Lu, 2016 | Cultivo celular (HepG2 y NCM640) | <ul style="list-style-type: none"> • CPUY192018 aporta citoprotección contra el daño oxidativo causado por DSS en células NCM460. • El pre-tratamiento con CPUY192018 aumenta la tasa de supervivencia inhibiendo la apoptosis e impidiendo la parada del ciclo celular en la fase S inducida por DSS. • CPUY192018 induce un aumento de los niveles proteicos de NRF2 en células del colon NCM460 y su acumulación en el núcleo. | <ul style="list-style-type: none"> • La inhibición de la interacción Keap1-NRF2 es una estrategia prometedora para activar mecanismos de defensa y podría ser beneficioso para la colitis ulcerosa. |
| Pettersen, 2016 | Cultivo celular (SW620 and Caco2) | <ul style="list-style-type: none"> • DHA induce un aumento de ROS y causa la importación nuclear de NRF2. • Un aumento en los niveles de NRF2 se detectó en el núcleo de células CaCo2 y SW620. | <ul style="list-style-type: none"> • El desarrollo de herramientas de diagnósticas capaces de detectar la autofagia y reguladores de la misma son necesarias para verificar la pérdida de autofagia de las células cancerosas y comprobar sus sensibilidad al DHA. |

| | | | |
|-------------------|---|---|--|
| Xu, 2015 | Cultivo celular (SW620, HCT116 and SW480) | <ul style="list-style-type: none"> • PYDDT promueve la producción de ROS, aumenta la expresión de NRF2 y disminuye el riesgo de CCR debido a una reducción de la genotoxicidad y un aumento de la apoptosis. | <ul style="list-style-type: none"> • La activación de la ruta de señalización ROS-JNK provocada por PYDDT induce la apoptosis mitocondrial en células humanas de cáncer de colon. • PYDDT podría servir como agente antitumoral contra el cáncer de colon. |
| Ishaq, 2014 | Cultivo celular (Caco2, HCT116, SW480 and HT29) | <ul style="list-style-type: none"> • NRF2 está involucrado en la respuesta protectora de las células tumorales HT29 induciendo apoptosis a partir de la activación de las caspasas 3/7. | <ul style="list-style-type: none"> • NADPH oxidasa inducida por AGP genera ROS intracelular, que activa la ruta ASK1 en células de CCR. • Las células resistentes a AGP desarrollan un sistema eficiente antioxidante mediado por NRF2 en el cual Srx protege a las células del efecto tóxico de la terapia antitumoral. |
| Pandurangan, 2014 | Cultivo celular (HT29) | <ul style="list-style-type: none"> • Transactivación de NRF2. | <ul style="list-style-type: none"> • La luteolina controla el cáncer de colon de muchas formas a través de la ruta NRF2. |
| Stachel, 2014 | Cultivo celular (NCM460) y ratones C57BL/6 | <ul style="list-style-type: none"> • La sobreexpresión de IER3 inhibe la activación de NRF2. • La ausencia de IER3 causa la reducción de los niveles de ROS como resultado de un aumento de la actividad de NRF2. | <ul style="list-style-type: none"> • IER3 equilibra la interacción entre NF-kB y NRF2. • Un fallo podría contribuir al inicio de la oncogénesis iniciado por inflamación crónica, como en el caso de la colitis asociada al CCR. |
| Yang, 2014 | Cultivo celular (Caco2, HT29, HepG2 y HEK-293) | <ul style="list-style-type: none"> • Bajas concentraciones de digitoflavonoides son activadores potenciales de NRF2/ARE en cáncer de colon y estimulan la expresión de las proteínas de defensa antioxidantes, de NRF2 y su traslocación al núcleo. • Tras el tratamiento con estos, aumenta la fosforilación de AKT, ERK1/2 y AMPK p38. La inhibición de la fosforilación de AKT y ERK1/2 no influye sobre la activación de NRF2. La fosforilación de p38 es esencial para la activación de NRF2. • Los ratones tratados con digitoflavonoides: disminución en el n° y tamaño de las criptas. | <ul style="list-style-type: none"> • Los digitoflavonoides intestinales mejoran el poder antioxidante induciendo la expresión de las principales enzimas detoxificantes mediante p38 y NRF2. • Estos ejercen un efecto protector y podrían ser usados como agentes quimiopreventivos contra el CCR. |
| Bat-Chen, 2013 | Cultivo celular (Caco2 y HT29) | <ul style="list-style-type: none"> • La alicina induce la apoptosis en las células de cáncer de colon y estimula la acumulación nuclear de NRF2. | <ul style="list-style-type: none"> • La alicina promueve directamente efectos citotóxicos en células de cáncer de colon humanas induciendo la apoptosis a través de mecanismos asociados a NRF2. |

| | | | |
|-------------------|--|---|---|
| Jung, 2013 | Cultivo celular (HT29 and HCT116) | <ul style="list-style-type: none"> • Keap1 silenciado activa la señalización de NRF2 en líneas celulares de cáncer de colon. • El silenciamiento de Keap1 puede sobreregular la expresión de los AKRs y atenuar el daño oxidativo causado por estrés. | <ul style="list-style-type: none"> • Los AKRs son genes regulados por NRF2 en las células epiteliales del colon. • La inducción de AKR se asocia con la detoxificación 4HNE y es muy relevante para la carcinogénesis en colon. • Estos resultados apoyan el papel antitumoral de la ruta NRF2 en el tejido del colon. |
| Kim, 2013 | Cultivo celular (HCT116) | <ul style="list-style-type: none"> • Los residuos de catecol son esenciales para activar NRF2. | <ul style="list-style-type: none"> • CAPE reduce la inflamación del colon. |
| Xi, 2013 | Cultivo celular (HepG2 y HCT116) | <ul style="list-style-type: none"> • NRF2 tiene un efector inhibidor en el desarrollo de CCR. • NRF2 disminuye el desarrollo de displasias y se ha encontrado sobreexpresado en tejidos tumorales. | <ul style="list-style-type: none"> • La expresión de NRF2 en condiciones estándar reduce significativamente el desarrollo de cáncer de colon asociado a la inflamación. |
| Park, 2010 | Cultivo celular (HT29, A549, y HCT116) | <ul style="list-style-type: none"> • El rottlerin induce la traslocación nuclear de NRF2. | <ul style="list-style-type: none"> • Rottlerin induce la expresión génica de HO-1 en células HT29. • Este efecto se media por la generación de ROS, la activación de MAPK p38 y ERK y la traslocación nuclear de NRF2. |
| Wondrak, 2010 | Cultivo celular (TH29, HCT116 y FHC) | <ul style="list-style-type: none"> • El cinamaldehído sobre regula a NRF2 y a sus genes diana agua abajo. • NRF2 proporciona protección contra la genotoxicidad inducida por estrés oxidativo. | <ul style="list-style-type: none"> • La canela activa a NRF2 en las células del colon y proporciona beneficios quimiopreventivos contra la carcinogénesis colorrectal. |
| Li, 2009 | Cultivo celular (HCT116) | <ul style="list-style-type: none"> • La exposición de las células HTC116 a NO tiene como resultado la traslocación nuclear de NRF2 así como su sobre regulación transcripcional. | <ul style="list-style-type: none"> • En células HTC116, la activación de NRF2-Keap1 por NO confiere resistencia al NO por sí mismo. |
| Li, 2016 | Ratones C57BL/6 | <ul style="list-style-type: none"> • NRF2 se activa por luteolina. | <ul style="list-style-type: none"> • La luteolina protege contra el cáncer de colon a través de la ruta NRF2. |
| Pandurangan, 2015 | Ratones BALB/c | <ul style="list-style-type: none"> • El cacao activa la expresión de NRF2. | <ul style="list-style-type: none"> • El cacao es un agente quimiopreventivo en la carcinogénesis colorrectal. |
| Tan, 2015 | Ratones | <ul style="list-style-type: none"> • Los granos de cerveza promueven la activación de NRF2. | <ul style="list-style-type: none"> • Los granos de cerveza disminuyen el estrés oxidativo y la carcinogénesis del colon. |
| Pandurangan, 2014 | Ratones BALB/c | <ul style="list-style-type: none"> • NRF2 se activa por luteolina. | <ul style="list-style-type: none"> • La luteolina activa los genes diana de NRF2 y protege contra la progresión del CCR. |
| Trivedi, 2014 | Ratones Swiss Albino | <ul style="list-style-type: none"> • MEL aumenta la expresión de NRF2 y otros genes diana. | <ul style="list-style-type: none"> • La autofagia y la ruta de señalización de NRF2: nuevas visión hacia la biología del cáncer e intervenciones terapéuticas para el tratamiento de cáncer de colon asociado a colitis. |

NRFs y cáncer colorrectal

| | | | |
|---------------|--|---|---|
| Trivedi, 2013 | Ratones Swiss | <ul style="list-style-type: none"> • El ácido lipoico aumenta la expresión de NRF2 and HO-1 en el colon y protege frente al daño en el ADN. | <ul style="list-style-type: none"> • El ácido lipoico es un agente potente para aliviar la colitis ulcerosa y su daño sistémico asociado. |
| Chiou, 2011 | Ratones BALB/c | <ul style="list-style-type: none"> • PS y RS promueve la activación de NRF2 y reduce la tumorigénesis del colon. | <ul style="list-style-type: none"> • PS podría ser más potente que RS mediando el efecto anti-tumorigénico de la activación de NRF2. |
| Myers, 2014 | Individuos con enfermedad de Chron y CU y cultivo celular (NCM460 y SW620) | <ul style="list-style-type: none"> • Los niveles de NRF2 son menores en tejidos inflamados de pacientes con CU y enfermedad de Chron. • Los niveles proteicos eran el doble en pacientes con CU asociada al cáncer. | <ul style="list-style-type: none"> • Hb-a es un mediador del daño al ADN causado por ROS. El daño al ADN aumenta la probabilidad de CCR. • El uso de antioxidants o la estimulación de la ruta sensible a redox podrían disminuir la colitis asociada al CCR. |
| Chang, 2013 | Sujetos con un diagnóstico pre-operativo de CCR. | <ul style="list-style-type: none"> • Las proteínas de la ruta de NRF2 se expresan activamente en CCR. | <ul style="list-style-type: none"> • Las diferencias en el perfil de expresión de las proteínas de la ruta NRF2 entre tejido normal y tumoral podría ofrecer una nueva dirección del tratamiento para el CCR. |

Tabla 3. Características de los artículos incluidos en la revisión sistemática cualitativa en los que NRF2 se sobreexpresa.

| Primer autor, año | Organismo / Tipo celular | Resultados primarios | Conclusión |
|-------------------|--|---|--|
| Kang, 2016 | Cultivo celular (SNUC5) | <ul style="list-style-type: none"> • La expresión de NRF2 es mayor en células de colon resistentes a 5-FU. • La demetilación sobreexpone a NRF2 en este tipo de células. | <ul style="list-style-type: none"> • En las células cancerosas resistentes al 5-FU, NRF2 se encuentra sobreexpuesto por interacción con metiltransferasas del ADN y demetilinas de histonas a través de estrés oxidativo. |
| Zhao, 2015 | Cultivo celular (SNUC5 CCR y SNUC5R) | <ul style="list-style-type: none"> • La expresión de NRF2 es más alta en células de CCR que son resistentes al tratamiento con 5-FU. | <ul style="list-style-type: none"> • Las modificaciones epigenéticas de NRF2 podrían ayudar a bloquear la resistencia al 5-FU. |
| Kang, 2014 | Cultivo celular (SNUC5) | <ul style="list-style-type: none"> • NRF2 proporciona resistencia al tratamiento con 5-FU en células de CCR. | <ul style="list-style-type: none"> • NRF2 se propone como una diana terapéutica para prevenir la resistencia al 5-FU en CCR. |
| Sebens, 2011 | Cultivo celular (NCM460) | <ul style="list-style-type: none"> • NRF2 se activa cuando las células se exponen a macrófagos inflamatorios, aumentando la actividad proteosómica y la expresión de genes proteosómicos dependientes de NRF2. | <ul style="list-style-type: none"> • NRF2 tiene un alto potencial para inducir la expresión de genes proteosómicos y estimular la actividad proteosómica en células epiteliales no transformadas. • La respuesta oxidativa ocurre tras la exposición a células inflamatorias y se favorece la supervivencia de las células con fenotipo tumoral. |
| Arlt, 2009 | Cultivo celular (LoVo, Colo320 y NCM460) | <ul style="list-style-type: none"> • En el núcleo de las células de CCR hay una actividad alta de NRF2. • La sobreexpresión de NRF2 aumenta la actividad proteosómica. | <ul style="list-style-type: none"> • El aumento de la actividad de NRF2 en células de CCR promueve la actividad proteosómica y contribuye a la tumorigénesis. • Mientras que en estadios tempranos NRF2 es beneficioso, en estadios tardíos contribuye a la génesis del CCR. |
| Ji, 2014 | Pacientes de CCR con metástasis | <ul style="list-style-type: none"> • La expresión de NRF2 es más alta en tejido de CCR y se correlaciona positivamente con el estadio y el pronóstico clínico. | <ul style="list-style-type: none"> • La expresión de NRF2 aumenta el tejido de CCR. |
| Hu, 2013 | 76 pacientes con CCR | <ul style="list-style-type: none"> • NRF2 se expresa más en tejido tumoral que en tejido normal. • La sobreexpresión de NRF2 se relaciona con tumores más grandes en estadios avanzados y metástasis. • La metástasis direccional podría estar asociada con la activación de NRF2. | <ul style="list-style-type: none"> • CXCR4 y NRF2 juegan un papel muy importante en el desarrollo del CCR. |

Tabla 4. Características de los artículos incluidos en la revisión sistemática cualitativa en los que NRF2 está inhibido.

| Primer autor, año | Organismo / Tipo celular | Resultados primarios | Conclusión |
|-------------------|----------------------------------|---|---|
| Volonte, 2013 | Cultivo celular (HCT116) | <ul style="list-style-type: none"> • La caveolina inhibe la traslocación de NRF2 al núcleo y promueve la senescencia prematura de células de CCR. | <ul style="list-style-type: none"> • Se necesitan más estudios para aclarar el mecanismo por el que se regula el equilibrio caveolina-NRF2. • Las intervenciones terapéuticas objetivan alterar la señalización caveolin-1/NRF2 para ver si pueden mejorar la supresión de la progresión del tumor sin acelerar la angiogénesis y/o los fenotipos relacionados con la edad, o si pueden disminuir el proceso de angiogénesis sin promover el inicio del tumor y/o progresión. |
| Kim, 2011 | Cultivo celular (HCT116 y HT29) | <ul style="list-style-type: none"> • La angiogenesis inducida por hipoxia se puede bloquear inhibiendo NRF2, reduciendo el crecimiento tumoral y la angiogénesis. • La hipoxia no puede activar a HIF-1a en células tumorales de colon con NRF2 inhibido. • Los niveles de PGC1-a y NRF1 no se vieron afectados por la inhibición de NRF2. Sin embargo, la desestabilización de HIF-1a se asocia con una función mitocondrial débil en células de cáncer de colon. | <ul style="list-style-type: none"> • Un consumo reducido de O₂ por la mitocondria en células de cáncer de colon con NRF2 inhibido provoca un bloqueo en la señalización NIF-1a-VEGF y una disminución subsecuente en la angiogénesis y crecimiento tumoral. |
| Yokoo, 2016 | Experimentación animal (ratones) | <ul style="list-style-type: none"> • Ratones NRF2^{-/-} tratados con KBrO₃: hiperplasias atípicas, adenomas y adenocarcinomas en el intestino delgado. • Aumento en el n° de criptas aberrantes comparadas con su equivalente NRF2^{+/+}. | <ul style="list-style-type: none"> • La falta de NRF2 aumenta la susceptibilidad de la carcinogénesis en el intestino. • Estrés oxidativo inducido por la sobreexpresión de COX2 y progresión del ciclo celular. • Individuos con SNPs en el gen NRF2 aumenta el riesgo de CCR. |
| Cheung, 2012 | Ratones machos C57BL/6J | <ul style="list-style-type: none"> • Los ratones NRF2 KO: incremento del número y el tamaño de los pólipos y la proliferación celular. | <ul style="list-style-type: none"> • Ausencia de NRF2: hiperproliferación de las criptas y expansión en el tamaño y el número de pólipos en el colon. |
| Saw, 2011 | NRF2 KO y ratones WT | <ul style="list-style-type: none"> • Los ratones NRF2 KO: colitis severa tras un tratamiento con DSS. Tras añadir AOM, los ratones NRF2 KO tienen mayor tendencia a desarrollar CCR y displasia, y a padecer prolapso y sangrado rectal. • NRF2 se requiere para la protección contra el CCR asociado a la inflamación. • Ratones NRF2 KO: mayor susceptibilidad de desarrollar criptas aberrantes asociadas con la inflamación. | <ul style="list-style-type: none"> • Agentes quimiopreventivos conducidos por NRF2 bloquean la inflamación aberrante y reducen el estrés oxidativo, impidiendo el desarrollo de CCR. |
| Li, 2008 | NRF2 KO y ratones WT | <ul style="list-style-type: none"> • Ratones NRF2 KO: colitis ulcerosa más severa, pérdida de criptas y una infiltración de células inflamatorias, además del | <ul style="list-style-type: none"> • NRF2 protege contra la inflamación colónica y la carcinogénesis. |

| | | | |
|--------------|--|--|--|
| | | sangrado rectal. | |
| Osburn, 2007 | NRF2 KO y ratones WT | <ul style="list-style-type: none"> • Exposición con DSS en ratones NRF2 KO: formación de criptas aberrantes. • Tratamiento con AOM: tumorigénesis asociada a la colitis debido a la inflamación del colon. | <ul style="list-style-type: none"> • El CCR asociado a la colitis es un problema de salud grave. • El desarrollo de estrategias de quimiopreención contra el CCR asociado a colitis podría ser una forma prometedora para prevenirla. • El genotipo tiene gran influencia en la progresión de la tumorigénesis e inflamación mediada por NRF2. • La ruta de NRF2: estrategia quimiopreventiva muy útil. • Ratones NRF2 KO: aumento de la susceptibilidad a la colitis inducida por DSS debido a la disminución de la expresión de enzimas detoxificantes y antioxidantes y el aumento de la expresión de mediadores proinflamatorios. • NRF2 podría ser una diana importante para el desarrollo de estrategias de quimiopreención, particularmente para la prevención del CCR. |
| Khor, 2006 | Ratones NRF2 ^{-/-} y WT (C57BL/SV129) | <ul style="list-style-type: none"> • Ratones NRF2 KO: longitud media del colon menor. • Exposición con DSS: los ratones NRF2 KO mostraron alteraciones en las criptas con inflamación severa. | |

Al analizar los artículos seleccionados, se observaron tres tipos de resultados claramente diferenciados en función de la expresión evaluada de NRF2. De esta manera, cuando NRF2 se expresa siguiendo un patrón básico, se observa una reducción del riesgo de CCR puesto que se trata del principal mecanismo de defensa y actúa eliminando carcinógenos y reduciendo la inflamación, que es una característica principal del CCR. Al contrario de esta situación, cuando se altera la expresión de NRF2, tanto sobreexpresándose como inhibiéndose, se conoce que aumenta el riesgo de CCR como consecuencia del desencadenamiento de procesos relacionados con este tipo de cáncer. Estas tres situaciones quedan reflejadas en los mapas conceptuales, referidos a la expresión estándar de NRF2 (Figura 24), la sobreexpresión del mismo (Figura 25) o su inhibición (Figura 26).

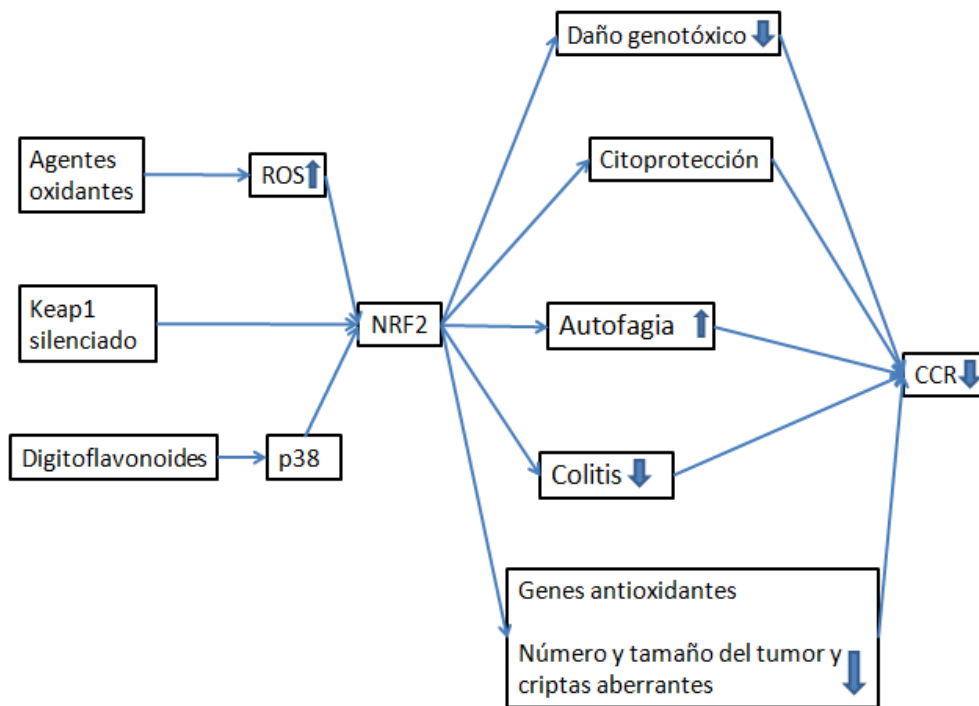


Figura 24. Mapa conceptual 1. Expresión estándar de NRF2 y disminución del riesgo de CCR.

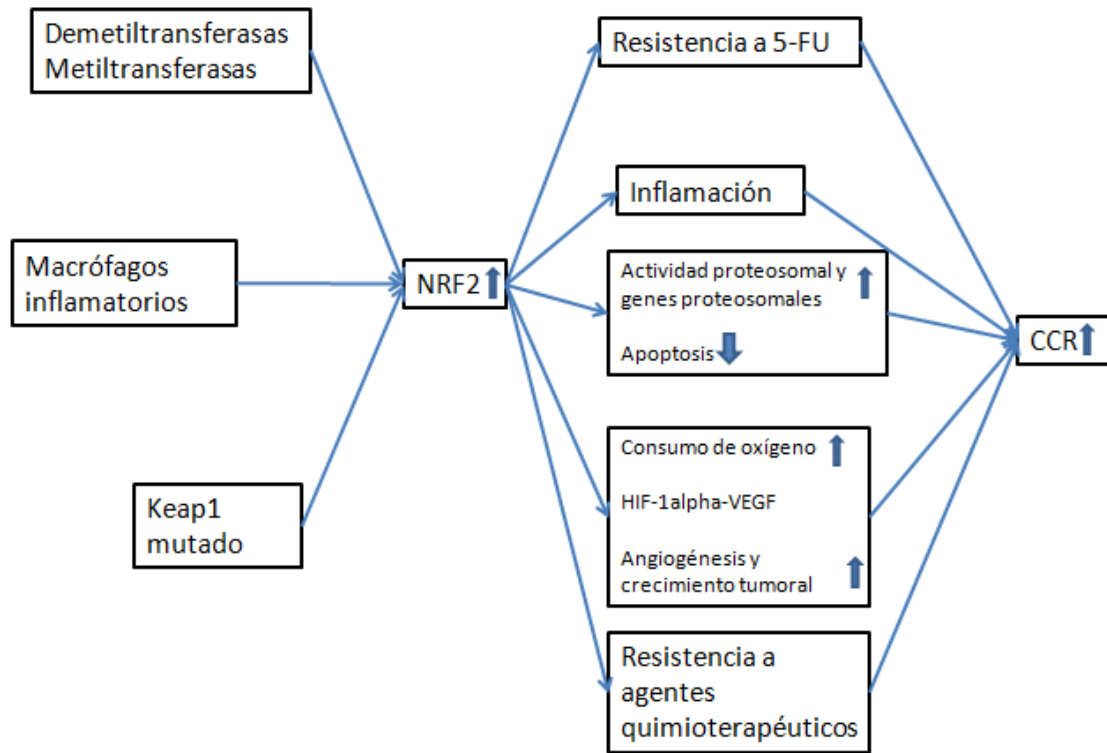


Figura 25. Mapa conceptual 2. Sobreexpresión de NRF2 y aumento del riesgo de CCR.

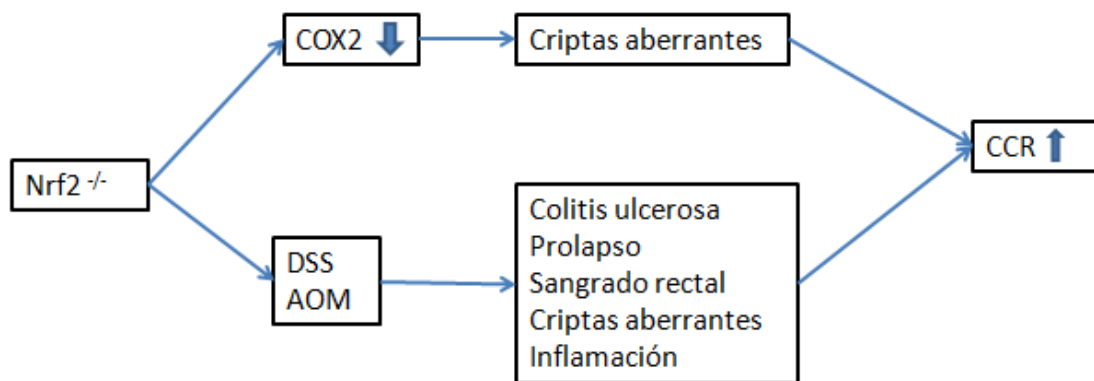


Figura 26. Mapa conceptual 3. Inhibición de NRF2 y aumento del riesgo de CCR.

Tras la evaluación de la heterogeneidad de los artículos, se observó que la mayoría de estudios se tratan de cultivos celulares o se realizan mediante experimentación animal, siendo limitados aquellos en los que se llevan a cabo análisis clínicos. En la Figura 27 se puntualizaron diferentes características mencionadas previamente en la metodología y si el artículo a evaluar las cumplía. Como se puede

observar todos evaluaban de alguna manera la expresión de NRF2. Además, todos los autores, salvo Lu *et al.*, (2016) y Myers *et al.*, (2014), relacionaban a NRF2 directamente con el CCR y no con otras enfermedades relacionadas con este como la colitis ulcerosa o enfermedad de Chron (Figura 27) (103,104).

DISCUSIÓN

La revisión tenía como objetivo relacionar NRF1 y NRF2 con el CCR. Sin embargo, debido a la falta de artículos que relacionen a NRF1 con el CCR, la revisión se limitó a conocer la relación entre el CCR y NRF2. La limitación de artículos relacionados con NRF1 es concordante con los descrito por Zhang *et al.*, (2015), que observa un número de artículos publicado sobre NRF1 muy inferior a los relacionados con NRF2 (72).

Tal y como se puede observar en el apartado de resultados, tras el análisis de las publicaciones incluidas en la revisión, se establecieron tres patrones distintos de comportamiento en función de la actividad y la expresión de NRF2.

El primer mapa conceptual encontrado muestra el efecto protector aportado ante un patrón de expresión estándar de NRF2 sobre el CCR seguida de una acumulación moderada de ROS debido a agentes externos específicos. Cuando NRF2 se activa, ejerce un efecto citoprotector disminuyendo el daño genotóxico y promoviendo la autofagia. A veces, la expresión de NRF2 conduce a reducir la proliferación celular y a promover la apoptosis de las células tumorales de colon. En todos los casos, los autores observar el mismo resultado final, principalmente una disminución del riesgo de CCR. Otro ejemplo de este comportamiento se describe en el estudio de Yang *et al.*, (2014), en el que los digitoflavonoides actúan sobre p38 e inducen la expresión de genes diana de NRF2 y antioxidantes (105). Esta investigación informa sobre una reducción tanto en el número y el tamaño de tumores como una disminución en el número de criptas aberrantes además de señalar el papel protector por parte de NRF2 en el CCR.

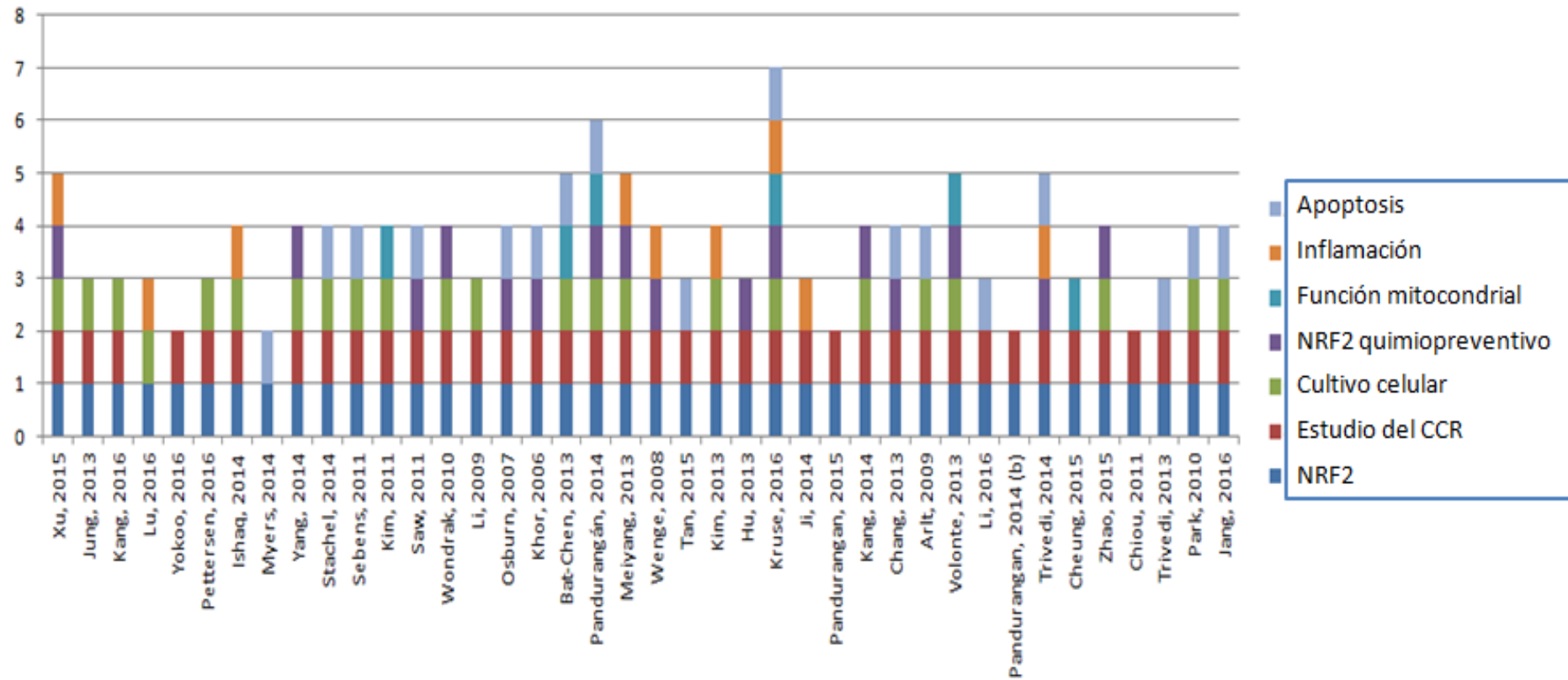


Figura 27. Evaluación de la heterogeneidad de los artículos incluidos en la revisión. Cada color hace referencia a una característica señalada en la leyenda.

Además, la expresión de NRF2 y sus genes diana, como la glutatión sintasa, juegan un papel crucial en mecanismos de citoprotección contra el CCR.

La expresión de NRF2, además de poder ser directamente inducida por agentes externos, también puede ocurrir indirectamente a partir de otros mecanismos, vía el represor Keap1. Por ejemplo, Jung y Kwak, 2013 silenciaron de manera temporal Keap1 y encontraron que la inducción de NRF2 promueve la expresión de miembros AKR, que están involucrados en el proceso de detoxificación, disminuyendo el riesgo de CCR. En otro estudio(106), Lu *et al.*, (2016) usaron un inhibidor de la interacción NRF2-Keap1 (CPUY1902018) y observaron que NRF2 actuaba como un factor protector ante la exposición a DSS, disminuyendo el riesgo de CCR (103).

En todos estos casos, la expresión estándar de NRF2 funciona como factor protector frente al CCR promoviendo citoprotección y activando genes diana antioxidantes, siendo estos datos coincidentes con los publicados en la revisión de Pandurangan y Esa, (2014) y Pandurangan *et al.*, (2014a, 2014b)(107–109).

Como se muestra en el segundo mapa conceptual, varios estudios han señalado que la sobreexpresión de NRF2 aumenta el riesgo de CCR. La sobreexpresión puede ocurrir por varias razones, incluyendo mutaciones constitutivas en el gen represor Keap1 o en el mismo NRF2. Stachel *et al.*, (2014) encontraron que los niveles excesivos de ROS inducían la sobreexpresión de NRF2, conduciendo a la inflamación del tejido colónico y promoviendo la tumorigénesis (110). En otro estudio, Sebens *et al.*, (2011) describieron que la sobreexpresión de NRF2 en tejidos expuestos a macrófagos inflamatorios y un aumento simultáneo en la actividad proteosomal de los genes conducían a una reducción de la apoptosis y a una proliferación descontrolada, lo que aumenta el riesgo de CCR (111).

Otros autores han observado resultados parecidos cuando NRF2 se sobreexpresa como consecuencia del silenciamiento, como un resultado de cambios epigenéticos en este gen. Otra manera de inducir la sobreexpresión de NRF2 es mediante el uso de compuestos químicos. Kim *et al.*, (2011), encontró que T-BHQ inducía la sobreexpresión de NRF2 y se producía un aumento del consumo de oxígeno, que se correlaciona con niveles más altos de HIF-1alpha y VEGF, que conducen al aumento de la angiogénesis y al crecimiento tumoral(112). Además, como se mencionó previamente, la sobreexpresión de NRF2 está muy unida al incremento del riesgo de

CCR. En este caso, el riesgo puede darse por medio de la inflamación del colon, una disminución de la apoptosis, o la promoción de la angiogénesis y la proliferación celular descontrolada.

Además, la sobreexpresión de NRF2 se ha relacionado con un aumento de la resistencia a agentes quimioterapéuticos como el 5-FU por lo que existe una respuesta mala al tratamiento basado en estos agentes (113).

El tercer mapa conceptual muestra que la expresión inhibida de NRF2 también tiene efectos sobre este tipo de cáncer, porque la función de NRF2 es proteger frente al daño y la sustancias carcinogénicas. Varios estudios señalan el papel de NRF2 en el CCR a través de su inhibición. Li *et al.*, (2008) describieron la relación de la inhibición de NRF2 con el aumento del CCR (114). Yokoo *et al.*, (2016) descubrieron que el silenciamiento de NRF2 tenía como resultado una disminución de la expresión de COX2, conduciendo a un aumento del número de criptas aberrantes, resultando la formación de adenomas, adenocarcinomas y, finalmente, CCR (115). Otros estudios han comparado los efectos del silenciamiento de NRF2 en ratones y exponiéndolos a tratamientos que inducen colitis (DSS) o carcinogénesis en el colon (AOM). Los ratones NRF2-knockout con colitis ulcerosa sufren prolapso, sangrado rectal, inflamación, y muestran un aumento del número de criptas aberrantes. Juntos, estos síntomas contribuyen al desarrollo del CCR. Estos hechos se corroboran de acuerdo a la revisión de Saw y Kong en 2011 y Cheung *et al.*, en 2014 (116,117).

Analizando el comportamiento de este gen, no en un momento en concreto, sino a través de varios estadios, Chang *et al.*, (2013) sugirió que la expresión diferencial de NRF2 en tumores versus tejido normal podría ser una nueva diana interesante en el CCR (92). Similarmente, Arlt *et al.*, (2009) observó que la expresión de NRF2 es beneficiosa durante estadios tempranos, pero podría contribuir a la carcinogénesis de colon en estadios tardíos (118). Esta idea es apoyada por autores como Menegon *et al.*, (2016) y Sporn y Liby (2012), que describieron a NRF2 como un oncogén y como un gen supresor de tumores (96,119). Son pocos los estudios que han analizado el papel de polimorfismos de un solo nucleótido en el gen y su promotor. Yokoo *et al.*, (2016) sugirió que algunos alelos de SNPs en el gen NRF2 podría aumentar el riesgo de CCR basado en previos estudios de estos polimorfismos y el riesgo de colitis ulcerosa (115).

Finalmente, otro factor importante relacionado con la expresión de NRF2 es el tipo de células usadas en los estudios incluidos en esta revisión. La mayoría de los artículos usan células madre de colon, que necesitan más energía y además requieren de una mayor expresión de NRF2 para obtener el ATP necesario para la proliferación, migración y diferenciación. La expresión alterada de NRF2 aumenta la inestabilidad genética, lo que conduce a aumentar el riesgo de CCR. *In vivo*, estas células migran a la parte apical de la cripta y se convierten en células epiteliales maduras, y la expresión alterada de NRF2 durante la migración y la diferenciación, conducirá a una acumulación de mutaciones que promueven la tumorigénesis debido a la ausencia de citoprotección. Además, la función de NRF2 es probablemente esencial tanto en las células madre del colon como las células epiteliales maduras (120).

Aunque se llevó a cabo una búsqueda de publicaciones en las principales bases de datos, empleando una lista exhaustiva de los términos relevantes relacionados con la pregunta de investigación, nuestro estudio tiene ciertas limitaciones, ya que tan solo se seleccionaron artículos publicados en inglés, lo que podría crear un sesgo de selección. Además, el uso de términos como fosforilación oxidativa, mitocondria o efecto Warburg, podría limitar la selección de artículos que relacionan NRF2 con el CCR, aunque las búsquedas llevadas a cabo sin tener en cuenta estos conceptos en varias bases de datos no mostraron ningún artículo adicional de interés.

Los resultados obtenidos han de ser interpretados con cautela, puesto que no están exentos de limitaciones. La principal limitación que presenta esta revisión bibliográfica es la ausencia de estudios clínicos en humanos, pues los estudios seleccionados e incluidos en ella se tratan de investigación básica y, en menor medida, de experimentación animal. Sin embargo, la fortaleza de esta revisión es el número de artículos incluidos en la misma (40 artículos) que permiten conocer de una manera clara el doble comportamiento de NRF2 en el CCR. Además, esta revisión cuenta con una metodología estandarizada que permite que el proceso sea fácilmente reproducible. Por otra parte, la elaboración de mapas conceptuales permite una interpretación de la información disponible hasta la fecha de una manera global.

Para concluir, está claro que el factor de transcripción de NRF2 tiene un papel muy complejo en la célula, y que su expresión puede ser fuertemente afectada por agentes externos. Cualquier alteración de su expresión estándar, ya sea sobreexpresión o

inhibición, puede promover la génesis y la progresión del CCR. Sin embargo, se precisan de más estudios para clarificar el papel de NRF2, particularmente para analizar el papel de varios SNPs que pudieran generar variabilidad en la expresión y la función de la proteína en humanos, así como estudiar rutas reguladas por este gen, y posibles interacciones con factores ambientales y comportamentales, aunque tales estudios requieren tamaños muestrales más grandes.

CAPÍTULO V. SNPS DE NRF1 Y NRF2 ASOCIADOS AL CCR.

METODOLOGÍA:

En el año 2008, el Centro de Investigación Biomédica en Red Epidemiología y Salud Pública (CIBERESP) del Instituto Carlos III (ISCI) inició el estudio MCC-Spain, un estudio multicaso-control poblacional para investigar la influencia de factores ambientales y su interacción con factores genéticos en tumores frecuentes o con características epidemiológicas peculiares en nuestro país, en las que los factores ambientales implicados no son suficientemente conocidos. MCC-Spain se recogió información entre septiembre de 2008 y diciembre de 2013 en 12 provincias españolas (Asturias, Barcelona, Cantabria, Girona, Granada, Gipuzkoa, Huelva, León, Madrid, Murcia, Navarra y Valencia). Los tumores recogidos son cáncer de mama, cáncer de próstata, cáncer de estómago, leucemia y CCR, siendo este último del que se han seleccionado los casos y los controles para llevar a cabo la presente tesis doctoral.

Este estudio implicó el reclutamiento de casos de los distintos tipos de cáncer que se compararon con un mismo grupo control. Todos los participantes debían tener entre 20 y 85 años de edad, haber residido en la zona de captación durante al menos 6 meses antes del reclutamiento y poder responder al cuestionario epidemiológico. La información sobre exposición se recogió mediante entrevistadoras previamente formadas a partir de un cuestionario estructurado, que incluyó información detallada sobre factores socio-demográficos, ocupación, antecedentes personales y familiares, hábitos de vida y dieta.

Cuando fue posible, se extrajeron 27 ml de sangre periférica de los participantes, que se alicuotaron en sangre total, plasma, fracción celular para la extracción de ADN y suero, y se almacenaron a -80 °C. Se recogió saliva de los sujetos de los que no se llevó a cabo la extracción de sangre con el kit de ADN Oragene® y se almacenó a temperatura ambiente hasta la extracción de ADN. Se recogieron muestras biológicas para la extracción de ADN para el 96% de los participantes entrevistados (76% sangre y 27% saliva, obteniendo en los casos en los que fue posible ambas muestras para un mismo individuo), así como muestras de uñas de los pies y cabellos de los participantes (79% y 84% respectivamente). En 4 centros (Madrid, Cantabria, Asturias y Huelva, que incluyen aproximadamente 1/3 de los participantes en el estudio) también se donaron muestras de orina (60 ml) alicuotadas y congeladas a -80°C. En todos los hospitales participantes se dispone de biopsias tumorales frescas o muestras incrustadas de

parafina. Los análisis genéticos se llevaron a cabo dentro de MCC-Spain y también a través de la participación en consorcios internacionales. El InfiniumHumanExomeBeadChip de Illumina se utilizó para genotipar más de 200000 marcadores de codificación de más de 6000 variantes personalizadas adicionales en las vías de interés para el proyecto.

Los controles fueron poblacionales, seleccionados de forma aleatoria a partir de los listados de personas que fueron atendidas en centros de atención primaria de las áreas sanitarias de referencia atendidas por los hospitales donde se reclutaron los casos. Los controles fueron entrevistados en su centro de atención primaria por las mismas entrevistadoras y en ese momento se recogían también las correspondientes muestras biológicas. Los controles fueron emparejados por frecuencia con los casos, por edad, sexo y región, asegurando que en cada región hubiera por lo menos un control del mismo sexo e intervalo de 5 años para cada caso. Para cada control, se seleccionó aleatoriamente un total de cinco participantes potenciales de edad, sexo y área de captación hospitalaria similares de las listas de médicos generalistas. Si el contacto con la primera persona de esta lista no era posible después de un mínimo de cinco intentos a diferentes horas del día, o si él/ella se negaba a participar, la siguiente persona de la lista fue contactada (121).

Las tasas de respuesta para los casos de CCR fueron del 68% y del 53% en los controles, variando por región. Para el 22% de los sujetos el contacto telefónico no fue posible debido a un número de teléfono erróneo o a la falta de respuesta.

Los pasos a seguir en este capítulo se recogen en la Figura 28, aunque el último paso adelanta la metodología referida a la siguiente sección. Como se puede observar, se especifican los distintos apartados así como el software usado para llevarlo a cabo.

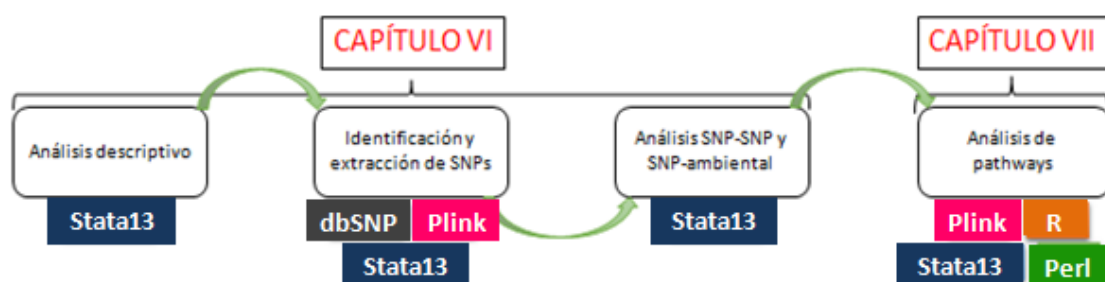


Figura 28. Metodología del capítulo VI y capítulo VII junto con los software utilizados.

Identificación y extracción de SNPs

Tanto para NRF1 como para NRF2, basándonos en la base de datos online Pubmed, se seleccionaron todas las variantes genéticas conocidas referentes a SNPs y se creó un archivo en los que cada línea correspondía a un SNP. A continuación, se cruzaron estos datos con los datos de MCC-Spain para detectar cuáles SNPs de NRF1 y NRF2 se encuentran recogidos en esta base de datos. Sobre estos SNPs se llevó a cabo un control de calidad ($MAF > 0,05$) y se eliminaron aquellos SNPs no genotipados (solo se incluye si el genotipado está realizado en el 90% de las muestras o más). Esta fase se realizó mediante el software Plink.

Una vez identificados los SNPs de NRF1 y NRF2 contenidos en MCC-Spain, se transformaron a un formato compatible con el software Stata13 para proceder a los análisis estadísticos.

Identificación de posibles variables confusoras

A partir de los cuestionarios del MCC-Spain realizados por los entrevistadores profesionales, se tomó información acerca de las variables edad, sexo, nivel educativo, IMC, antecedentes familiares de primer grado, actividad física, dieta y área por ser posibles variables confusoras. La edad se tomó como una variable continua. Las variables sexo y antecedentes familiares de primer grado se tratan de variables dicotómicas. El resto de variables mencionadas se categorizaron de la siguiente manera: nivel educativo (menos de estudios primarios, estudios primarios, estudios secundarios y estudios universitarios), IMC [bajo-peso ($< 18,5 \text{ kg/m}^2$), normopeso ($\geq 18,5 \text{ kg/m}^2 - 25 \text{ kg/m}^2$), sobrepeso ($\geq 25 \text{ kg/m}^2 - 30 \text{ kg/m}^2$) y obesidad ($\geq 30 \text{ kg/m}^2$)], actividad física [sedentario (0 MET*h/sem), bajo ($< 8 \text{ MET*h/sem}$), moderado (8-16 MET*h/sem), elevado ($> 16 \text{ MET*h/sem}$)] y dieta (adherencia media, baja o alta de acuerdo a la puntuación obtenida en el índice establecido por Trichopoulou (122)). La variable área hace referencia al lugar donde se situaba el hospital donde se realizó el reclutamiento.

Análisis estadístico.

Análisis descriptivo.

El primer tipo de análisis que se hizo corresponde a un análisis descriptivo para conocer las características principales de la población de estudio. Se realizó una clasificación diferenciando entre casos y controles calculándose la media y la desviación típica para las variables continuas y las frecuencias absolutas y relativas para las variables categóricas mencionadas en el apartado de "Identificación de posibles variables confusoras". Además, se realizó una tabla de frecuencias para conocer la distribución genotípica de los SNPs en la población de estudio.

Análisis de asociación

2.2.1. Asociación SNPs-CCR y SNPs-posibles variables confusoras

Con el objetivo de conocer si estos SNPs modifican el riesgo de CCR, se llevó a cabo un análisis mediante una regresión logística de modelos mixtos, usando como referencia siempre el genotipo más frecuente perteneciente a los controles.

De esta manera, se calculó la asociación mediante el cálculo de OR ajustados teniendo en cuenta las posibles variables confusoras. La variable área se tomó como variable aleatoria. También se realizaron análisis estratificando a la población en función del sexo y teniendo en cuenta la localización del tumor (recto, colon, colon derecho y colon izquierdo) para conocer las posibles diferencias entre hombres y mujeres, así como la localización de la aparición del tumor.

Finalmente, se realizaron diferentes análisis para identificar asociaciones entre el SNP y posibles variables confusoras. Para llevar a cabo el análisis se usó el software estadístico Stata 13.

RESULTADOS:

Identificación y extracción de SNPs

En la Tabla 5, se pueden observar los alias, los cromosomas y las regiones cromosómicas de los genes NRF1 y NRF2. Además, se listan los SNPs pertenecientes a cada uno de ellos contenidos en la base de datos de MCC (Tabla 5).

El SNP rs12706898 se trata de una variante intrónica de NRF1 localizada en la posición 501. Los dos alelos posibles son G/T, siendo el alelo G el alelo ancestral (no alelo de riesgo) y T, el alelo de riesgo (Figura 29) (123). Si se analiza la frecuencia alélica en el mundo y teniendo en cuenta los continentes, se observa la siguiente distribución: en Europa es donde se observa mayor frecuencia del alelo T (alelo de riesgo) (un 56% frente a un 44% del alelo G), mientras que el alelo G (alelo ancestral) presenta su mayor frecuencia en población africana (95%). Si se analizan estas frecuencias a nivel mundial, se observa que el alelo G es más común que el alelo T (T: 32%; G: 68%) (Figura 29) (124).

Tabla 5. Características de los genes principales de estudio y SNPs pertenecientes a NRF1 o NRF2 recogidos en la base de datos MCC-Spain.

| GENES | NRF1 | NRF2 | |
|-------------|------------------------|------------------------|-------------|
| ALIAS | ALPHA-PAL | NFE2L2-HEBP1 | |
| CROMOSOMA | 7 | 2 | |
| REGIÓN | NC_000007.13 | NC_000002.11 | |
| CROMOSÓMICA | (129251555..129396922) | (178095031..178129859) | |
| | rs12706898 | rs181294188 | rs60132461 |
| | rs149447357 | rs201560221 | rs35577826 |
| SNPs | | rs182276775 | rs201992337 |
| | | rs141363120 | rs200239262 |
| | | rs34154613 | rs35248500 |

Nota: Sólo el SNP rs12706898 perteneciente al gen NRF1 (negrita en la Tabla 5) se encuentra genotipado en más del 90% de las muestras y con un MAF > 0,05, por tanto, solo esta variante genética es óptima para llevar a cabo el análisis SNP-SNP.



Figura 29. Posibles alelos del SNP rs12706898 y distribución de la frecuencia alélica en el mundo. Datos disponibles. en dbSNP y Ensembl.

Análisis estadístico

Análisis descriptivo

Las características socio-demográficas y asociadas al estilo de vida de la población con la que se va a realizar los posteriores análisis quedan descritas mediante la Tabla 6.

Tabla 6. Características de la población de estudio con el SNP rs20706898 genotipado.

| Características | | Controles | | Casos | | p-tendencia |
|---|-----------------------------|---------------|------|---------------|------|-------------|
| | | 2741 | | 1336 | | |
| Edad | Media ± SD | 63,30 ± 11,46 | | 66,96 ± 10,59 | | |
| | | N | % | N | % | |
| Sexo | Hombre | 1468 | 53,6 | 865 | 64,7 | |
| | Mujer | 1273 | 46,4 | 471 | 35,3 | 0,001 |
| Nivel educativo | Menos de estudios primarios | 488 | 17,9 | 380 | 28,4 | |
| | Primaria | 912 | 33,3 | 552 | 41,3 | |
| | Secundaria | 786 | 28,7 | 261 | 19,5 | 0,001 |
| | Universidad | 555 | 20,3 | 143 | 10,7 | |
| IMC | Media ± SD | 26,74 ± 4,40 | | 27,74 ± 4,63 | | |
| | Bajo peso | 23 | 1,0 | 13 | 1,1 | |
| | Normopeso | 758 | 31,5 | 259 | 22,3 | |
| | Sobrepeso | 873 | 36,3 | 399 | 34,3 | 0,001 |
| | Obesidad | 752 | 31,3 | 492 | 42,3 | |
| AFPG | No | 2373 | 90,8 | 1034 | 82,2 | |
| | Si | 241 | 9,2 | 224 | 17,8 | 0,001 |
| Índice de Actividad Física (METS*h/sem) | Sedentarismo (0) | 1032 | 37,7 | 620 | 46,4 | |
| | Baja (<8) | 420 | 15,3 | 156 | 11,7 | |
| | Moderada (8-16) | 341 | 12,4 | 128 | 9,6 | |
| | Elevado (>16) | 948 | 34,6 | 432 | 32,3 | 0,001 |
| Adherencia a la dieta mediterránea | Baja | 817 | 34,0 | 447 | 38,4 | |
| | Media | 1052 | 43,7 | 492 | 42,3 | |
| | Alta | 537 | 22,3 | 224 | 19,3 | 0,004 |

En este estudio, participaron un total de 4077 individuos, repartidos en 1336 casos de CCR (865 hombres; 471 mujeres) y 2741 controles (1468 hombres; 1273 mujeres). La media de edad±DS en la población de estudio se centró en 66,96 años

$\pm 10,59$ en los casos y de 63,30 años $\pm 11,46$ en los controles. La Tabla 6 muestra la distribución en función de determinados factores asociados: El IMC presenta una media de 27,5 ($\pm 4,6$) en los casos y 26,7 ($\pm 4,4$) en los controles. Los antecedentes familiares de primer grado se caracterizan por su baja presencia, ya que el 82,2% en casos y el 90,8% en controles no poseen antecedentes familiares de primer grado. El índice de actividad física refleja una población principalmente sedentaria, el 46,4% de los casos y el 37,7% de los controles apenas realizan ningún tipo de actividad física. Por último, en cuanto a la adherencia al patrón de dieta mediterránea, un 61,6% de los casos y un 66,1% de los controles presentan patrón de adherencia medio o alto.

La distribución genotípica de la población de estudio se describió en la Tabla 7. Tras comprobar si sigue el equilibrio de Hardy-Weinberg se obtuvieron los siguientes resultados (HWE= 0,0068; $\chi^2=2,12$; p-valor=0,15).

Tabla 7. Distribución genotípica de la población de estudio.

| rs12706898 | Control (N/%) | Casos (N/%) | Total (N/%) |
|------------|---------------|--------------|---------------|
| TT | 893 32,58 | 396 29,64 | 1289 31,62 |
| gT | 1310 47,79 | 679 50,82 | 1989 48,79 |
| gg | 538 19,63 | 261 19,54 | 799 19,6 |
| Total | 2741 100 | 1336 100 | 4077 100 |

Análisis de asociación entre el SNP rs12706898 y el CCR y posibles variables confusoras.

El SNP rs12706898 no se asoció de una manera estadísticamente significativa con el CCR ni siquiera cuando se tuvo en cuenta la localización tumoral diferenciada. Al analizar los resultados obtenidos, en el caso del cáncer de colon derecho, el resultado fue próximo a alcanzar la significación estadística en aquellos individuos que presentaban el genotipo heterocigoto gT [OR=1,35; IC (95%) = (0,99-1,85)] tomando como referencia el genotipo TT. Además, se observó que los individuos con un genotipo homocigoto recesivo gg presentaban mayor riesgo de sufrir cáncer de colon izquierdo a pesar de no presentar un resultado estadísticamente significativo (Referencia: genotipo TT). En cuanto a la localización tumoral referida al recto, no se observaron modificaciones dependientes del genotipo, aportando en ambos casos OR próximos a 1. (Tabla 8)

Cuando se realizó una estratificación de la población por sexo a la vez que se tuvo en cuenta la localización del tumor, los hombres con el genotipo g/T presentaron riesgo de sufrir cáncer de colon derecho [OR=1,68; IC (95%) = (1,12-2,54)]. En el caso de las mujeres, a pesar de no alcanzar la significación estadística, el genotipo homocigoto recesivo provocó un aumento del riesgo de cáncer de colon derecho [OR=1,13; IC (95%) = (0,61-2,08)] mientras que el genotipo heterocigoto gT parecía no tener efecto alguno [OR=0,98; IC (95%) = (0,59-1,68)]. Si se analiza el CCR de manera global no se obtuvieron resultados estadísticamente significativos aunque se observó que, con respecto al genotipo TT, el genotipo gg y gT aportaban riesgo en la población masculina y protección en la población femenina. Este evento se repite cuando se evalúa el cáncer de colon izquierdo y el cáncer de recto (Tabla 9).

Tabla 8. Asociación entre el CCR y los posibles genotipos del SNP rs12706898.

| rs12706898 | CCR | | | | | Colon derecho | | | | | Colon izquierdo | | | | | Recto | | | | |
|------------|-------|-------|------|----------|------|---------------|-------|------|----------|------|-----------------|-------|------|----------|------|-------|-------|------|----------|------|
| | Co | Ca | OR | IC (95%) | | Co | Ca | OR | IC (95%) | | Co | Ca | OR | IC (95%) | | Co | Ca | OR | IC (95%) | |
| | (N/%) | (N/%) | | | | (N/%) | (N/%) | | | | (N/%) | (N/%) | | | | (N/%) | (N/%) | | | |
| TT | 893 | 396 | - | - | - | 893 | 112 | - | - | - | 893 | 129 | - | - | - | 893 | 152 | - | - | - |
| | 32,6 | 29,6 | | | | 32,6 | 29,1 | | | | 32,6 | 29,4 | | | | 32,6 | 30,6 | | | |
| gT | 1310 | 679 | 1,10 | 0,91 | 1,34 | 1310 | 208 | 1,35 | 0,99 | 1,85 | 1,310 | 216 | 1,07 | 0,79 | 1,44 | 1310 | 247 | 1,02 | 0,77 | 1,35 |
| | 47,8 | 50,8 | | | | 47,8 | 54,0 | | | | 47,8 | 49,2 | | | | 47,7 | 49,7 | | | |
| gg | 538 | 261 | 1,11 | 0,87 | 1,41 | 538 | 65 | 1,00 | 0,66 | 1,51 | 538 | 94 | 1,25 | 0,87 | 1,79 | 538 | 98 | 1,08 | 0,75 | 1,53 |
| | 19,6 | 19,5 | | | | 19,6 | 16,9 | | | | 19,6 | 21,4 | | | | 19,6 | 19,7 | | | |

Nota: Co= Controles; Ca= Casos

Tabla 9. Asociación entre el CCR y el SNP rs12706898 estratificada en función del sexo.

| rs12706898 | CCR | | | | Colon derecho | | | | Colon izquierdo | | | | Recto | | | | | | | |
|----------------|-------------|--------------|------|----------|---------------|-------------|-------------|-------------|-----------------|-------------|-------------|-------------|-------|------|------|-------------|-------------|------|------|------|
| | Co | Ca | OR | IC (95%) | Co | Ca | OR | IC (95%) | Co | Ca | OR | IC (95%) | Co | Ca | OR | IC (95%) | | | | |
| Hombres | | | | | | | | | | | | | | | | | | | | |
| TT | 494 33,7 | 248 28,67 | - | - | - | 494 33,7 | 60 26,2 | - | - | - | 494 33,7 | 83 28,6 | - | - | - | 494 33,7 | 103 30,8 | - | - | - |
| gT | 677 46,1 | 440 50,87 | 1,26 | 0,99 | 1,62 | 677 46,1 | 132 57,6 | 1,69 | 1,12 | 2,54 | 677 46,1 | 140 48,3 | 1,20 | 0,82 | 1,76 | 677 46,1 | 162 48,5 | 1,06 | 0,75 | 1,51 |
| Gg | 297 20,2 | 177 20,46 | 1,26 | 0,93 | 1,72 | 297 20,2 | 37 16,2 | 0,90 | 0,51 | 1,59 | 297 20,2 | 67 23,1 | 1,51 | 0,96 | 2,37 | 297 20,2 | 69 20,7 | 1,23 | 0,81 | 1,89 |
| Mujeres | | | | | | | | | | | | | | | | | | | | |
| TT | 399 31,3 | 148 31,42 | - | - | - | 399 31,3 | 25 33,3 | - | - | - | 399 31,3 | 46 30,9 | - | - | - | 399 31,3 | 49 30,1 | - | - | - |
| gT | 633 49,7 | 239 50,74 | 0,94 | 0,68 | 1,29 | 633 49,7 | 76 48,7 | 0,98 | 0,59 | 1,68 | 633 49,7 | 76 51,0 | 0,89 | 0,54 | 1,45 | 633 49,7 | 85 52,2 | 0,99 | 0,60 | 1,62 |
| Gg | 241 18,9 | 84 17,83 | 0,94 | 0,62 | 1,41 | 241 18,9 | 28 18,0 | 1,13 | 0,61 | 2,08 | 241 18,9 | 27 18,1 | 0,87 | 0,46 | 1,64 | 241 18,9 | 29 17,8 | 0,84 | 0,43 | 1,63 |

Nota: Co=Controles; Ca=Casos; Los datos que se muestran en negrita representan resultados estadísticamente significativos

DISCUSIÓN

Análisis descriptivo de la muestra

La muestra seleccionada para este análisis contó con 4077 individuos (1336 casos y 2741 controles) de 6090 individuos totales (2140 casos y 3950 controles) que cuenta el estudio MCC-Spain. Los individuos elegidos para llevar a cabo el análisis así como los individuos totales componentes del estudio MCC-Spain tenían una edad media de 63 años los controles y de 66 años los casos, no habiendo diferencias entre ambas poblaciones. Además, la proporción de hombres y mujeres en ambas poblaciones también es similar. Considerando estos datos, la muestra seleccionada para llevar a cabo los análisis podría ser representativa de la muestra total, pues el rango de edad y la cantidad de hombres y mujeres se mantiene, quedando representados el 69,5% de los controles y el 62,4% de los casos.

Selección de SNPs

Al observar los SNPs pertenecientes al gen NRF1 y NRF2, tan sólo un polimorfismo está genotipado y cumple las características de inclusión establecidas por el control de calidad. Esto podría deberse a que los genes NRF1 y NRF2 no estaban incluidos inicialmente como objeto de estudio prioritario en el proyecto de MCC-Spain. Además, se empleó el ExomaIllumina que limita identificar SNPs de este tipo.

Cuando se observan las frecuencias alélicas del SNP rs12706898 de la región europea y se compara con la de nuestra población, se puede observar que son bastante similares (alelo T en Europa = 56% vs alelo T en nuestra población= 56,4%; alelo G en Europa = 44% vs alelo G en nuestra población = 43,5%). Sin embargo, al comparar estas cifras con respecto a la población mundial, se observan grandes diferencias ya que el porcentaje referente al alelo T desciende hasta un 32%, siendo esta cifra más similar a la que se refiere al alelo G en nuestra población. Se podría decir que nuestra población tiene una distribución alélica similar a la de Europa pero inversa a la distribución mundial (124).

Análisis SNP-CCR y SNP-posibles variables confusoras

Tras realizar los análisis, tomando como referencia el genotipo TT, se puede observar que cuando se estudian las distintas localizaciones del tumor en el colon, el efecto de los posibles genotipos restantes del SNP (gT/gg) tiene un sentido protector en el caso de las mujeres mientras que en los hombres aporta cierto riesgo aunque no se alcancen resultados estadísticamente significativos, salvo en el caso de los hombres con cáncer de colon derecho y el genotipo heterocigoto del SNP de estudio. Al tratarse del genotipo heterocigoto, la explicación biológica de esta asociación es bastante complicada pudiendo resumirse de la siguiente manera: la selección natural no tiende a eliminar a ninguno de los dos alelos y a fijar un genotipo homocigoto, sino que, tras infinitas generaciones de selección, se alcanza un equilibrio a frecuencias intermedias en una situación en la que la eficacia biológica promedio de la población es máxima (125,126). Otra situación observada es la diferencia entre el colon derecho y colon izquierdo, lo que concuerda con lo descrito por Lee *et al.*, (2015). En este estudio apoyan la idea de que el cáncer colorrectal presenta diferencias en la incidencia, la patogénesis y las vías moleculares que afectan en función de la localización tumoral, corroborando que existen diferencias entre el cáncer de colon derecho y el cáncer de colon izquierdo (19).

Al analizar la posible relación entre el SNP rs12706898 con las posibles variables confusoras tampoco se observaron resultados estadísticamente significativos por lo que se podría llegar a pensar que estas no influyen en gran medida sobre el ajuste del análisis.

Fortalezas y limitaciones:

El estudio MCC-Spain cuenta con una base de datos donde se recogen las características de una población representativa de España, contando con datos clínicos, epidemiológicos y genéticos, por lo que se permite realizar estudios con una gran cantidad de información perteneciente a cada individuo. Sin embargo, no se planteó como objetivo primario la evaluación de los polimorfismos de los genes NRF1 y NRF2, por lo que el análisis se reduce al análisis de un SNP que cuenta con las características de inclusión. El análisis de este SNP en concreto se trata de una novedad puesto que tan solo existe un estudio que lo evalúe pero no lo relacionan con el cáncer (127). A pesar de tener la limitación del número de estudios acerca de este polimorfismo, nuestros

resultados abren el camino a futuras líneas de investigación en las que analicen de una manera experimental la mutación puntual en tejido tumoral de colon comparándola con tejido sano de colon, llegando a clarificar el papel que desempeña.

CAPÍTULO VI. RUTAS
METABÓLICAS ASOCIADAS AL
CCR.

METODOLOGÍA

Disponibilidad de datos genéticos.

Los datos usados para este análisis se encuentran disponibles en la base de datos del proyecto MCC-Spain descrita en los apartados anteriores, contando con información de 1336 casos de CCR y 2741 controles. Los datos genéticos que se evaluaron fueron obtenidos a partir del InfiniumHumanExomeBeadChip de Illumina (128). Para que los SNPs se incluyeran en el análisis se tuvo en cuenta que el genotipado fuera superior al 90% y el MAF mayor a 5%.

Selección de pathways

Se seleccionaron los pathways disponibles en las bases de datos KEGG y Reactome y se tuvieron en cuenta tanto las rutas en las que participa NRF1 y/o NRF2 como aquellas que regula o son reguladas por estos genes. Específicamente, se evaluaron las siguientes rutas: Inflammatory Bowel Disease (IBD) (129), Longevity regulating pathway (130), MAPK signaling pathway (131), NF-kappa B signaling pathway (132), PI3K-Akt signaling pathway (133), Pentose phosphate pathway (PPP) (134), ABC transporters (135), Mitochondrial Biogenesis pathway (136) and AMPK signaling pathway (137) (Tabla X). Teniendo en cuenta la existencia de evidencia científica previa sobre la asociación de la ruta de las enfermedades inflamatoria del intestino con el CCR (42,138,139) se consideró como un posible control positivo, mientras que debido a la inexistencia de estudios que relacionan CCR y la ruta relacionada con la longevidad, esta podría considerarse como probable control negativo.

El paso siguiente consistió en realizar dos mapeos. El primer mapeo se llevó cabo partiendo de los datos recogidos en las bases de datos KEGG y Reactome con el objetivo de conocer los genes componentes de cada uno de las rutas metabólicas. El segundo mapeo se realizó teniendo en cuenta la información disponible en la base de datos dbSNP de Pubmed para saber los SNPs pertenecientes a cada uno de los genes de los pathways.

A partir de esta información, para llevar a cabo el análisis de los pathways se han seleccionado 3 metodologías: SRT, fGSEA y PoDA (55,140,141).

"SNP Ratio Test" (SRT)

Con los datos del GWAS del MCC-Spain se creó 1 archivo donde se incluyó la información del GWAS [(identificador de cada individuo, fenotipo (caso/control) y características para cada SNP en el GWAS (presencia de 0, 1, o 2 alelos de riesgo)] en un formato de archivo binario apto para su uso en Plink.

A partir de estos archivos, mediante la aplicación del script "make_alt_pheno.pl" en Perl, se generó un conjunto de archivos con 1000 fenotipos alternativos, obtenidos mediante permutaciones aleatorias, almacenando dicha información en archivos con extensión ".txt". Dichos archivos se emplearon para calcular los p-valores para la asociación entre los SNPs del GWAS con los fenotipos generados en las distintas permutaciones, almacenándose dichos p-valores en los archivos correspondientes con extensión ".assoc" para cada uno de ellos. A su vez, también se calculó la asociación entre los SNPs contenidos en el GWAS original y el fenotipo. Estos pasos se llevaron a cabo haciendo uso del software Plink y teniendo en cuenta el modelo de herencia log-aditivo.

Mediante la ejecución del script parseassocfiles.pl. se creó un archivo que incluía el SNP id y los p-valores tanto de los datos originales como de los obtenidos con las 1000 permutaciones. Tras esto, se estableció un punto de corte (p-valor = 0,05) y se ejecutó el script "count_sig_snps.pl" para identificar los SNPs como significativo (por debajo del umbral) o no significativo (por encima del umbral), información que se almacenó en un archivo de salida con extensión "assoc.forSRT.p0.05.ratios". De esta manera, se conocen el número de SNPs que superan el punto de corte establecido en el ranking original y en los rankings obtenidos a partir de los datos simulados.

Por otro lado, a partir de los datos disponibles en KEGG y en Reactome, se generaron archivos de tipo txt en el que para cada pathway de estudio se incluían los SNPs conocidos en los genes de dichas rutas.

Mediante los comandos run_SRT_onSIMs.pl, cuando se trataba de los datos simulados, y run_SRT_on_ORIGINAL.pl., cuando se trataba de los datos originales, empleando los datos de los archivos de tipo "assoc.forSRT.p0.05.ratios" y del archivo txt que incluía la lista de SNP en cada pathway, se generaron archivos que contenían para cada pathway los siguientes datos dispuestos en 4 columnas:

En la primera columna el identificador del pathway, en la segunda columna el número de SNPs de dicho pathway con valores p significativos de acuerdo al punto de corte establecido, en la tercera columna, el número total de SNPs de dicho pathway, en la cuarta columna el cociente de la relación de SNPs significativos con respecto al total que se denominó r en el caso de los datos originales y r' en el caso de los datos simulados.

Aplicando el script "getSRTpvalue.pl", que permite comparar los valores r obtenidos para los datos originales con los valores r' obtenidos en las distintas permutaciones, se calculó para cada pathway un valor S que refería el número de veces en las que r' era mayor que r teniendo en cuenta el total de permutaciones realizadas. Tras ello, se calculó un valor p empírico, consistente en conocer el resultado del cociente $p_E = S+1/1001$.

Todos los script para llevar a cabo este análisis se encuentran disponibles online (142), y requirieron el empleo del software Plink y Perl.

"Fast - Gene Set Enrichment Score Analyses" (ES)

Tomando como referencia el archivo descrito en el apartado anterior que incluía la información del GWAS [(identificador de cada individuo, fenotipo (caso/control) y genotipo para cada SNP en el GWAS (presencia de 0, 1, o 2 alelos de riesgo)], se generaron 3 nuevos archivos. El primer archivo se trató de un ranking de los SNPs de dicho GWAS basado en la asociación de estos SNPs con la enfermedad que contenía dos columnas, el identificador del SNP (rs) y el valor de χ^2 obtenido mediante el software Stata13. El segundo archivo consistió en un conjunto de vectores que contenían información de los pathways a evaluar y los SNPs incluidos en cada uno de ellos. La información relativa a los SNPs se obtuvo a partir de la base de datos dbSNP de Pubmed. Por último, se creó un vector que describe el fenotipo de cada individuo (caso/control=TRUE/FALSE) teniendo en cuenta el orden seguido por los individuos en el primer archivo que se generó. Todos estos archivos se importaron al software R donde se llevaron a cabo los análisis.

Tras especificar que los pathways a tener en cuenta debían contener entre un tamaño mínimo de 15 SNPs y un máximo de 500 SNPs y que se permutaron 1000 veces

las etiquetas caso/control, mediante la aplicación del script "fgsea" en el software R se obtuvo un "dataframe" donde se tabuló la siguiente información: el nombre de la ruta evaluada, el identificador de las mismas, la fuente de donde provienen, el número de SNPs, la puntuación de enriquecimiento (ES), un valor normalizado de dicha puntuación (NES), un p-valor y un q-valor, que resultó del FDR. Para obtener estos datos, el script consta de una función llamada fgseaR encargada de realizar un test de asociación entre los SNPs del GWAS y los fenotipos permutados, generándose rankings de SNPs simulados. El siguiente paso consiste en calcular la ES, tanto de los datos originales como de los datos permutados, para después normalizarla y conseguir el valor de NES y NES*. Finalmente, el script compara el valor referente a los datos originales con los datos simulados para calcular un p-valor. Los resultados estadísticamente significativos se representaron gráficamente ejecutando la función "plotEnrichment" del script fgsea.

Los scripts para llevar a cabo este análisis se encuentran disponibles online en la biblioteca Bioconductor.org (143) y requirieron el empleo del software Stata13 y R.

"Pathway of Distinction Analysis" (PoDA)

A partir de los datos del GWAS del MCC-Spain se crearon 4 archivos: i) una matriz donde se representan los genotipos de cada SNP para cada individuo en función de la presencia del número de alelos de riesgo (0, 1 ó 2); ii) un vector booleano que indicó el fenotipo, es decir, si eran casos o controles (Casos= TRUE; Controles=FALSE) ordenado en función de la matriz previamente mencionada; iii) lista que mapeó los pathways a genes obteniendo la información necesaria de las bases de datos online KEGG y Reactome; iv) lista que mapea los genes a SNPs recolectando la información de la base de datos dbSNP de Pubmed. Los scripts para llevar a cabo este análisis se encuentran disponibles online (144) y requirieron el empleo del software R.

Una vez que se importaron estos archivos de entrada, se ejecutó en R el script "PoDA.R" para crear la función PoDA que permite calcular el MAF medio para casos y controles, el valor D, el valor S, el valor Wp y el valor DSp. En primer lugar, partiendo de los datos originales, se ejecutó el script "PoDA-example.R" para conseguir el valor W. El siguiente paso consistió en especificar el número de permutaciones de las etiquetas caso/control ("NsamplePerms "=1000) para ejecutar de nuevo el mismo script

y obtener 1000 valores W_p^* . Este paso conllevó la comparación de los valores W_p y W_p^* para generar el valor DSp. A continuación, se creó un objeto ("resamplepaths"=TRUE) que permitió generar aleatoriamente pathways "dummies" teniendo en cuenta la longitud del pathway original establecida en "pathGeneList" seleccionando al azar los SNPs componentes de la misma del GWAS. En nuestro caso, se estableció que se generaran 100 pathways "dummies" por cada pathway original. Finalmente, se ejecutó de nuevo el script "PoDA-example.R" que permitió obtener 100 DSp* para cada pathway. La última función de dicho script comparó los valores DSp y DSp*, calculando de esta manera un p-valor para cada pathway y un q-valor tras ejecutar el comando que realiza el test FDR.

Los resultados estadísticamente significativos se representaron gráficamente mediante la ejecución del script "plotSvals.R", que dio lugar a dos gráficos por cada pathway: un dotplot y un diagrama de densidad que permiten valorar las diferencias en la distribución de valores S entre casos y controles.

RESULTADOS

Selección de rutas:

En la Tabla 10 se pueden observar los identificadores de los pathways evaluados, así como el número de genes que compone cada pathway:

Tabla 10. Identificadores de los pathways evaluados en esta tesis.

| PATHWAY | ID | Genes |
|---|---------------|--------------|
| Ruta de las pentosas fosfato | hsa00030 | 30 |
| Ruta de los transportadores ABC | hsa02010 | 44 |
| Ruta de señalización MAPK | hsa04010 | 295 |
| Ruta de señalización NF-K β | hsa04064 | 95 |
| Ruta de señalización PI3K-Akt | hsa04151 | 354 |
| Ruta de señalización AMPK | hsa04152 | 120 |
| Ruta reguladora de la longevidad | hsa04213 | 62 |
| Ruta de las enfermedades inflamatorias de intestino | hsa05321 | 65 |
| Ruta de la biogénesis mitocondrial | R-HSA-1592230 | 96 |

El próximo paso consistió en seleccionar todos los genes de los pathways que se evaluaron, haciendo un total de 1098 genes. Al mismo tiempo, se tuvo en cuenta el GWAS que, tras un control de calidad, contenía un total de 6806 genes. Partiendo de los datos de MCC-Spain y teniendo en cuenta las rutas seleccionadas, una vez que se seleccionaron los genes, se recolectaron todos los SNPs de estos genes (121). Estos puntos son comunes a todas las metodologías de análisis de pathways que se van a emplear.

"SNP Ratio Test"

Mediante el método *"SNP Ratio Test"* se seleccionaron 1544 SNPs significativos del GWAS en función de la asociación con el CCR en base al punto de corte p-valor=0,05. Aplicando esta metodología para el análisis de rutas no se observaron resultados estadísticamente significativos, siendo el p-valor fue igual a 1 en todas las rutas evaluadas salvo en la ruta de señalización de las MAPK en la que el p-valor = 0,53 (Tabla 11).

Tabla 11. Resultados obtenidos tras analizar las rutas mediante "SNP Ratio Test".

| Rutas | ID | # r' > r | n° de permutaciones | p-valor |
|--|---------------|--------------------|----------------------------|----------------|
| Ruta de las pentosas fosfato | hsa00030 | 1000 | 1000 | 1 |
| Ruta de los transportadores ABC | hsa02010 | 1000 | 1000 | 1 |
| Ruta de señalización MAPK | hsa04010 | 533 | 1000 | 0,53 |
| Ruta de señalización NF-kappa-B | hsa04064 | 1000 | 1000 | 1 |
| Ruta de señalización PI3K-Akt | hsa04151 | 1000 | 1000 | 1 |
| Ruta de señalización AMPK | hsa04152 | 1000 | 1000 | 1 |
| Ruta de regulación de la longevidad | hsa04213 | 1000 | 1000 | 1 |
| Ruta de las enfermedades inflamatorias del intestino | hsa05213 | 1000 | 1000 | 1 |
| Ruta de la biogénesis mitocondrial | R-HSA-1592230 | 1000 | 1000 | 1 |

"Enrichment Score"

Si se analizan los resultados mostrados en la Tabla 12, se observa cómo tras realizar el test "False Discovery Rate" (FDR), la única ruta que presenta significación estadística es la ruta "Inflammatory Bowel Disease", con un p-valor =0,003, una q= 0,03.y una puntuación de enriquecimiento normalizada para este de 1,08 puntos. Este resultado se puede observar de una manera más visual mediante la representación en un histograma (Figura 30).

Tabla 12. Resultados obtenidos tras analizar las rutas mediante "Fast Gene Set Enrichment Analysis".

| Pathways | Ruta | Source | # of SNPs | ES | NE S | Nominal P-value | FDR |
|--|---------------|----------|-----------|------|------|-----------------|------|
| Ruta de las pentosas fosfato | hsa00030 | KEGG | 306 | 0,59 | 1,05 | 0,1 | 0,3 |
| Ruta de los transportadores ABC | hsa02010 | KEGG | 1731 | 0,54 | 0,96 | 0,97 | 0,97 |
| Ruta de señalización MAPK | hsa04010 | KEGG | 2303 | 0,56 | 1 | 0,42 | 0,54 |
| Ruta de señalización NF-kappa B | hsa04064 | KEGG | 706 | 0,57 | 1,02 | 0,18 | 0,34 |
| Ruta de señalización PI3K-Akt | hsa04151 | KEGG | 5167 | 0,54 | 0,98 | 0,9 | 0,97 |
| Ruta de regulación de la longevidad | hsa04152 | KEGG | 350 | 0,61 | 1,07 | 0,035 | 0,15 |
| Ruta de señalización AMPK | hsa04213 | KEGG | 1101 | 0,57 | 1,02 | 0,18 | 0,34 |
| Ruta de enfermedades inflamatorias del intestino | hsa05213 | KEGG | 545 | 0,61 | 1,08 | 0,003 | 0,03 |
| Ruta de la Biogénesis Mitocondrial | R-HSA-1592230 | Reactome | 605 | 0,57 | 1,02 | 0,19 | 0,34 |

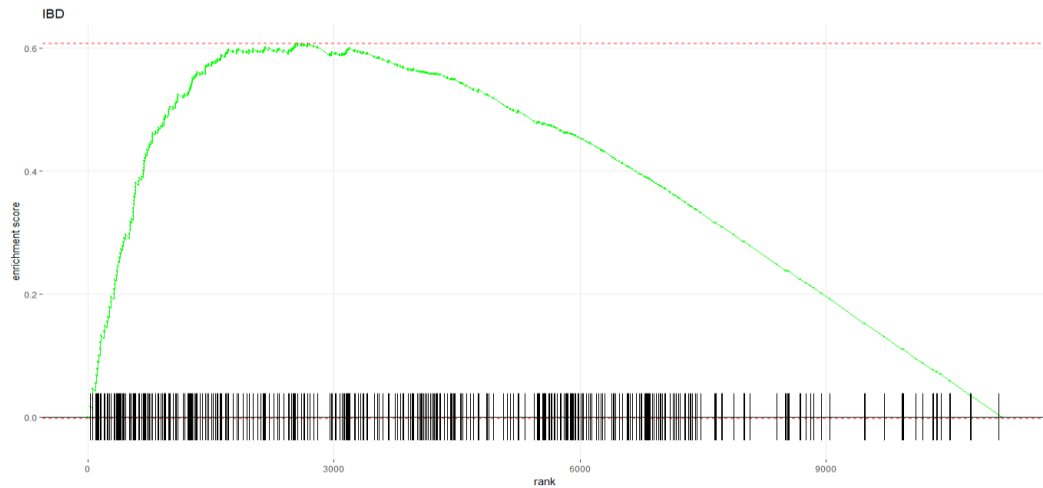


Figura 30. Histograma de frecuencias acumuladas referido a la ruta de las enfermedades inflamatorias del intestino.

"Pathway of Distinction Analysis"

En la Tabla 13 se representan los resultados obtenidos mediante PoDA. La ruta de IBD, la ruta de las MAP Kinasas, la ruta de señalización NFK β , la ruta de señalización PI3K-AKT, la ruta de los transportadores ABC transporter y la ruta de la biogénesis mitocondrial muestran asociación con el CCR, mientras que del resto de pathways no existen evidencias sobre la posible asociación.

Las gráficas referentes a las rutas que han mostrado resultados estadísticamente significativos muestran que el valor S de los casos es superior al de los controles. Además, si se hace referencia al diagrama de densidad, se observa que existen diferencias entre la curva de los controles (negra) y curva de los casos (roja). Observando la distribución de la población que ofrecen las gráficas se podría decir que existen diferencias entre casos y controles y se podrían establecer dos grupos diferenciados (Figura 31).

Tabla 13. Resultados del análisis de rutas obtenidos mediante el método PoDA.

| Ruta | Ruta | Fuente | Longitud | DSp | p(DSp) | p-Adj. (FDR) |
|--|---------------|----------|----------|-------|--------|------------------|
| Ruta de las pentosas fosfato | hsa00030 | KEGG | 30 | 1,21 | 0,27 | 0,34 |
| Ruta de los transportadores ABC | hsa02010 | KEGG | 45 | 10,58 | <0,001 | <0,001 |
| Ruta de señalización MAPK | hsa04010 | KEGG | 255 | 16,64 | <0,001 | <0,001 |
| Ruta de señalización NF-kappa-B | hsa04064 | KEGG | 95 | 7,76 | <0,001 | <0,001 |
| Ruta de señalización PI3K-Akt | hsa04151 | KEGG | 341 | 8,22 | <0,001 | <0,001 |
| Ruta de señalización AMPK | hsa04152 | KEGG | 114 | 0,71 | 0,87 | 0,87 |
| Ruta de regulación de la longevidad | hsa04213 | KEGG | 94 | 1,33 | 0,77 | 0,87 |
| Ruta de las enfermedades inflamatorias del intestino | hsa05213 | KEGG | 65 | 12,68 | <0,001 | <0,001 |
| Ruta de la biogénesis mitocondrial | R-HSA-1592230 | Reactome | 50 | 3,69 | <0,001 | <0,001 |

En la Tabla 14, se expone un resumen de los resultados obtenidos a partir del análisis de las rutas mediante distintas metodologías. Como se puede observar, la ruta de las enfermedades inflamatorias del intestino, seleccionada como control positivo, muestra asociación con el CCR cuando el análisis se lleva a cabo mediante fGSEA y PoDA. En cuanto al control negativo, el pathway relacionado con la longevidad, no se detecta asociación con el CCR independientemente del método seleccionado para el análisis.

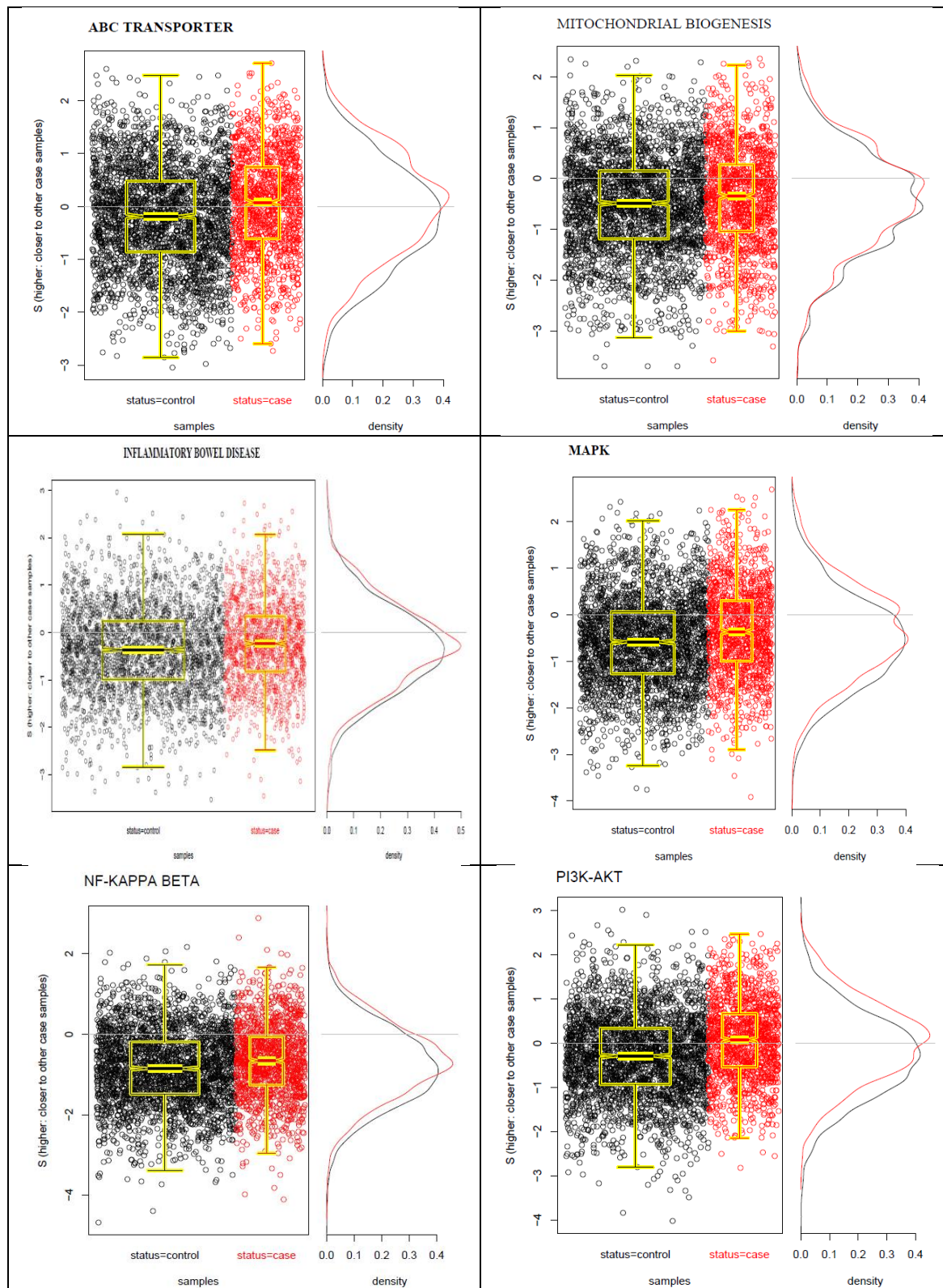


Figura 31. Dot plots y diagramas de densidad para las distintas rutas con un resultado significativo obtenido mediante la metodología de análisis PoDA.

Tabla 14. Resultados obtenidos tras el análisis de las rutas mediante SRT, fGSEA y PoDA.

| Ruta | SRT | fGSEA | PoDA |
|---|-------|-------|------------------|
| Ruta de regulación de la longevidad (hsa04213) | 1,000 | 0,150 | 0,870 |
| Ruta de las enfermedades inflamatorias del intestino (hsa05213) | 1,000 | 0,030 | <0,001 |
| Ruta de señalización MAPK (hsa04010) | 0,530 | 0,540 | <0,001 |
| Ruta de señalización NF-kappa-B (hsa04064) | 1,000 | 0,340 | <0,001 |
| Ruta de la biogénesis mitocondrial (R-HSA-1592230) | 1,000 | 0,340 | <0,001 |
| Ruta de señalización AMPK (hsa04152) | 1,000 | 0,340 | 0,870 |
| Ruta de señalización PI3K-Akt (hsa04151) | 1,000 | 0,970 | <0,001 |
| Ruta de las pentosas fosfato (hsa00030) | 1,000 | 0,300 | 0,340 |
| Ruta de los transportadores ABC (hsa02010) | 1,000 | 0,970 | <0,001 |

Los datos que se muestran para las técnicas fGSEA y PoDA son q-valores. Para la técnica SRT, se obtuvieron p-valores, que son los mostrados en esta tabla.

DISCUSIÓN

Tras analizar los resultados obtenidos, se observa que mediante PoDA se detectan 6 rutas significativas, mediante fGSEA solo se identifica una ruta significativa y mediante SRT no se encontró ninguna.

Al analizar los resultados obtenidos a partir de la metodología SRT de la Tabla 11, no se identifica ninguna ruta que sea significativa pues los ratios obtenidos a partir de las permutaciones siempre son mayores que el ratio real. Esto podría deberse a que la mayoría de los SNPs pertenecientes a las rutas estudiadas no están recogidos en el GWAS con el que se establece el ranking. En la metodología SRT, el p-valor elegido como punto de corte para seleccionar los SNPs significativos del ranking va a depender de la cantidad de variantes genéticas que se quieran tener en cuenta (55). De este modo, si se elige un p-valor estricto para definir los SNPs “asociados”, solamente los SNP altamente asociados que participan en la ruta serán seleccionados. El mejor resultado para una enfermedad de la cual se tiene poca información sería la elección de p-valores

menos estrictos, de tal manera que la ruta presenta un mayor número de variantes aunque éstas tengan menor efecto sobre la enfermedad (55).

Como fortaleza referida a esta metodología se puede señalar la facilidad para llevar a cabo el análisis. Sin embargo, existen ciertas limitaciones:

- Dificultad para detectar vías significativas cuando hay escasez de SNPs pertenecientes a las mismas. Además, en caso de que el número de SNPs significativos perteneciente a una ruta sea cero, la proporción siempre daría como resultado el valor cero.
- Si solo hay un único gen dentro de una ruta con SNPs significativos, el método SRT reduciría la señal de la ruta a la señal de un gen.
- El método SRT trata a todos los SNPs significativos por igual, lo cual puede provocar cierta dificultad para detectar rutas significativas.

Estas mismas rutas, cuando se analizaron mediante fGSEA, se observaron resultados diferentes. Mediante este análisis se identificó la ruta "*Inflammatory Bowel Diseases*" como buena predictiva del CCR ($p=0,03$), lo cual se esperaba de acuerdo a la gran cantidad de estudios que confirman la estrecha relación entre estas enfermedades y el CCR (138,145,146). En el resto de rutas no se han obtenido resultados estadísticamente significativos. Esta metodología presenta mejoras con respecto al SRT pues el nivel de significación para cada SNP es distinto y, además, el punto de referencia es la ruta que se va a evaluar. Subramanian *et al.*, (2015) señalaron que al no tener que establecer un punto de corte, tal y como ocurre en el SRT, se trata de un método más objetivo (147).

Como se puede observar, la técnica SRT no sería un buen método a la hora de evaluar la asociación de las rutas con la enfermedad debido a su poca sensibilidad. En el siguiente caso, mediante la técnica "*Fast Gene Set Enrichment Analysis*", aumenta la precisión pues cada ruta adquiere un valor estadístico relacionado con una puntuación de enriquecimiento (140). Hay que señalar que la ruta de las MAPK alcanza valores parecidos tanto por SRT como por FGSEA. Sin embargo, la única ruta que alcanza significación estadística es la ruta relacionada con las enfermedades inflamatorias del intestino, como se mencionó previamente. El análisis de enriquecimiento suple carencias presentes en el análisis mediante "*SNP Ratio Test*" ya que considera todos los

SNPs pertenecientes a la ruta y no solamente aquellos que superen un determinado valor de corte de significación establecido previamente. Sin embargo, es importante tener en cuenta que su principal limitación es que no tiene en cuenta la posibilidad de que haya desequilibrio de ligamiento porque el autor considera que las permutaciones van a disminuir la posibilidad de que esto ocurra (147).

El último método, PoDA, además de las permutaciones, realiza remuestreos para crear rutas "dummies" que sirvan para comprobar que los resultados obtenidos no se deben al azar (57). Mediante este test se detectaron 6 rutas estadísticamente significativas entre las que se encuentra la ruta de las enfermedades inflamatorias del intestino, por lo que coincide con el resultado obtenido por fGSEA. El resto de rutas significativas tiene un p-valor menor a $<0,001$ lo cual quiere decir que se tratan de rutas asociadas a la enfermedad de estudio. Otro dato a tener en cuenta en la metodología PoDA es la puntuación de distinción de rutas que cuantifica las diferencias de la distribución S en casos y controles (141). A medida que aumenta esta puntuación, aumenta la capacidad de predicción de la ruta. Por tanto, en este caso, la ruta de las MAPK es la que se encuentra más relacionada con el CCR, con una puntuación de distinción (DSp) de 16,64 seguida de las enfermedades inflamatorias del intestino (DSp=12,68).

De acuerdo con Azer *et al.*, (2013), la ruta de las enfermedades inflamatorias del intestino se asocia directamente con la génesis del CCR (148). Además, el estudio de Bondurant *et al.*, (2013) evaluó la asociación entre SNPs relacionados con las interleukinas (asociadas a procesos inflamatorios) y el CCR, mostrando asociación. En concordancia con Ran Kim *et al.*, (2014), el CCR asociado a las enfermedades inflamatorias del intestino incluían otros pathways entre los que se pueden citar la respuesta inmune mediante mediadores inflamatorios de la mucosa, estrés oxidativo y la microbiota intestinal (42). Todos estos hechos se corroboran con la revisión llevada a cabo por Sebastian *et al.*, (2014), en la que se analizaron varios estudios epidemiológicos concluyendo que hay una estrecha relación entre el CCR y esta ruta (145). Todos estos datos apoyan la idea de que este pathway fuera considerado como un control positivo para nuestro análisis, por lo que los resultados obtenidos están en concordancia con lo descrito (p-valor $<0,001$).

Por otra parte, aunque existen estudios que muestran relación entre una expresión/actividad adecuada de NRF2 y la longevidad (149), no se ha identificado hasta el momento ninguna relación entre esta ruta y el CCR, por lo que se podría considerar como un control negativo. Todo esto está en consonancia con los resultados estadísticamente no significativos obtenidos en el análisis (p -valor=0,870). La concordancia de estos resultados con evidencias anteriores nos da confianza para usar esta metodología para detectar pathways que se asocian con una enfermedad en concreto.

En relación a las otras rutas analizadas, es importante destacar que los resultados obtenidos tienen plausibilidad biológica y son consistentes con la evidencia científica. De esta manera, NRF2 es conocido como el principal mecanismo de defensa contra el estrés oxidativo o estrés electrofílico, por lo que resulta evidente que este gen participe en el transporte y metabolismo de fármacos (150). Este mecanismo podría llevarse a cabo mediante la ruta de los transportadores ABC (151). Esta ruta se asocia en gran medida con el CCR por lo que podría usarse como una herramienta predictiva en el futuro. La principal causa podría ser la acumulación de fármacos sin metabolizarse que podrían transformarse en compuestos carcinogénicos y promover el CCR (56). Relacionado con ello, Hu *et al.*, (2016) encontraron que los transportadores ABC sobreexpresados reducían el efecto de los fármacos antineoplásicos en el CCR (152).

Otro hecho importante es la activación de NRF2, que se puede producir por diversos mecanismos. Uno de estos mecanismos ocurre a través de la fosforilación mediante kinasas en sitios específicos a través de rutas de señalización como las MAPK, AMPK o PI3K, induciendo su acumulación en el núcleo y provocando la transactivación de genes que controlan la respuesta de elementos antioxidantes (ARE). De esta manera, NRF2 aporta protección y reduce el riesgo de CCR (153–156). El pathway de las MAPK parece ser un buen predictor del CCR. Gracias a la fosforilación promovida por este pathway, se activan algunos procesos que conducen a la génesis y desarrollo del CCR (157). Lo mismo ocurre con la ruta NF κ B que modula la activación y expresión de NRF2 (114) o con la ruta PI3K. Además, ambas rutas se relacionan y se encuentran a menudo alteradas en el CCR (158). Las tres rutas mencionadas pueden activar a NRF2 mediante la fosforilación, por lo que son esenciales para una expresión estándar. Sin embargo, se ha demostrado que la ruta de la AMPK, que también activa a

NRF2, se encuentra alterada en estadios tardíos, por lo que no siempre es un buen predictor en estadios tempranos del cáncer (159).

Además, NRF2 está involucrado en el metabolismo de la glucosa a través de la ruta de las pentosas fosfato (160). Esta ruta no muestra un resultado estadísticamente significativo con la aparición del CCR y podría deberse a que se trata de un metabolismo alternativo de carbohidratos, por lo que otros pathways podrían desempeñar esta función cuando esta ruta está alterada. Sin embargo, en discordancia con nuestros resultados, Kowalik *et al.*, (2017) describieron que existía una asociación entre este pathway y la tumorigénesis y la proliferación celular como consecuencia de una actividad sobrerregulada de NRF2 debido a la sobreexpresión de G6PD, si bien este resultado corresponde a estadios tardíos de la enfermedad por lo que podría no estar relacionada con la aparición sino con el desarrollo de la misma (161).

Por último, NRF2 también es uno de los genes responsables de la creación y formación de nuevas mitocondrias a través de la participación en la ruta de la biogénesis mitocondrial (162). La biogénesis mitocondrial se encarga de crear estas mitocondrias de una manera adecuada, tanto en morfología como en número (163). Dentro de la mitocondria ocurren una gran cantidad de procesos metabólicos por lo que es de suponer que cuando se modifican las mitocondrias, también ocurren modificaciones en los procesos que ocurren dentro de ellas, de acuerdo a lo descrito en el capítulo IV de esta tesis(164). Además, Cruz *et al.*, (2017) también supone que la reprogramación mitocondrial es un proceso que ocurre con cierta frecuencia en los estadios tempranos del CCR, asumiendo una relación entre la mitocondrial y el CCR (165).

Si se comparan las tres metodologías, PoDA es la que más capacidad posee para señalar diferencias estadísticas, por lo que podría tratarse del método más sensible. La metodología PoDA aumenta la eficacia de los métodos de enriquecimiento porque no es necesario que los SNPs tengan efectos independientes de alta penetrancia, si no que se trata de un método basado en distancias relativas de unos sujetos a otros (141,147). A diferencia de los tipos de análisis de enriquecimiento que buscan una sobreabundancia de marcadores significativos en un gen de interés, PoDA encuentra tanto rutas que contienen marcadores muy significativos como aquellas rutas que tienen marcadores sutiles pero coherentes (57). Por lo tanto, mediante el uso de esta metodología va a ser posible la identificación de rutas en las que la acción conjunta de diversas mutaciones

puntuales con poco efecto de manera individual confiera un determinado fenotipo. Otra ventaja con respecto a las técnicas de enriquecimiento es que el tamaño del conjunto de SNPs que se va a estudiar no está limitado (140). Además, PoDA aprovecha las relaciones biológicas que se conocen para detectar efecto de conjuntos de SNPs por lo que los resultados son más fiables e interpretables. Sin embargo, no todo son ventajas en cuanto a esta metodología pues el tiempo de ejecución es analíticamente costoso con respecto a las otras dos metodologías usadas en esta tesis, requiriendo mucha potencia de cálculo.

Fortalezas y limitaciones:

La principal fortaleza de este trabajo recae sobre la realización de un trabajo comparativo de tres metodologías de análisis de pathways con tamaños muestrales grandes y un gran número de SNPs, dando visibilidad a la gran sensibilidad que ha aportado la metodología PoDA, siendo pioneros en el uso de la misma, abriendo camino hacia una medicina cada vez más estratificada que conduzca a una individualización de la misma. Nuestros resultados proporcionan un fundamento para seguir investigando el papel de la genética en el desarrollo de CCR ya que se consiguen detectar mayor número de marcadores genéticos relacionados con la enfermedad que cuando se usa un análisis de variantes genéticas de manera individual.

Las limitaciones de este apartado hacen referencia a la necesidad de replicar el estudio en otras poblaciones haciendo uso de estudios multicéntricos internacionales para poder generalizar los resultados obtenidos. Además, una vez se han obtenido resultados es difícil corroborarlos a través de literatura científica pues hay pocos estudios que las hayan llevado a cabo.

CAPÍTULO VII. COROLARIO

Partiendo de la evidencia científica existente(83,87,166,167), se hipotetizó una posible relación entre los genes NRF1 y NRF2, y las rutas en las que participan con el riesgo de CCR. Por lo que como primer paso para intentar caracterizar el rol de estos genes en relación con este tipo de cáncer, se llevó a cabo una revisión sistemática, en la que se encontró que en el NRF1 esta relación ha sido poco estudiada existiendo solo tres estudios que relacionen el factor de transcripción NRF1 y el CCR (60,79,84) no pudiendo obtenerse una visión clara y profunda de dicha relación. Por el contrario, la relación entre NRF2 y el CCR, ha sido mucho más estudiada considerándose en la revisión 40 artículos que muestran 3 comportamientos dependientes de la expresión de este gen, observando que un patrón de expresión estándar se relacionó con una disminución del riesgo de enfermedad mientras que una alteración, ya sea referida a la sobreexpresión o inhibición del mismo, aumentaba el riesgo de CCR. Es importante señalar que la mayoría de estos resultados se obtuvieron a partir de estudios in vitro o mediante experimentación animal con ratones, siendo muy limitados los estudios en humanos. En referencia a los estudios con humanos, Yokoo et al., (2016) sugirió que algunas variantes en SNPs en el gen NRF2 podrían aumentar el riesgo de CCR (115).

Tomando esta posible influencia del efecto de la presencia de SNPs en estos genes, con el fin de analizar el efecto de las posibles alteraciones en estos genes se llevó a cabo un análisis de polimorfismos disponibles en la base de datos del estudio MCC-Spain. Al observar los SNPs pertenecientes al gen NRF1 y NRF2, disponibles en el array empleando en dicho estudio, sólo el SNP rs12706898 incluido en el gen NRF1 estaba genotipado y cumplía las características de inclusión establecidas por el control de calidad. Como se ha descrito en el capítulo V, tomando como referencia el genotipo TT, se puede observar que cuando se estudian las distintas localizaciones del tumor en el colon, el efecto de los posibles genotipos restantes del SNP (gT/gg) tiene un sentido protector en el caso de las mujeres mientras que en los hombres aporta cierto riesgo, alcanzándose el riesgo máximo en hombres con genotipo gT para el cáncer de colon derecho. El hecho de que existan diferencias entre el riesgo de cáncer de colon derecho y cáncer de colon izquierdo ante el mismo factor genético no es de extrañar, ya que de acuerdo a lo descrito por Lee et al., existen desigualdades en función de la localización del colon (19).

Teniendo en cuenta la falta de estudios que relacionen a NRF1 con el CCR, este resultado es pionero en identificar la asociación de un SNP de este gen con el cáncer de colon, si bien debe tomarse con cautela y requiere de ser replicado en futuros estudios.

Dado el limitado número de variantes genéticas de los genes de interés recogidas en la base de datos del estudio MCC-Spain y que además la mayoría de las variantes presentes no tienen una alta penetrancia, se planteó como siguiente paso, estudiar la relación de las rutas en las que participan los genes NRF1 y NRF2 con la aparición del CCR, mediante el análisis de conjuntos de genes o rutas metabólicas.

Puesto que no existe un gold standard entre las distintas metodologías de análisis de rutas, se seleccionaron 3 de ellas: SRT, fGSEA y PoDA (55,140,141), que presentan ciertas similitudes en el análisis pero también diferencias metodológicas, especialmente a la hora de seleccionar los SNPs. En este sentido, puesto que la técnica SRT solo selecciona los SNPs que cumplen con un nivel de significación estadística individualmente, mientras que PoDA selecciona todos los SNPs disponibles del GWAS para llevar a cabo el análisis, era de esperar que SRT fuera la técnica que detectará menos asociaciones al ser más estricta a la hora de incluir SNPs mientras que PoDA sería la técnica capaz de detectar la asociación del CCR con un mayor número de rutas relacionadas con NRF1 y NRF2, quedando fGSEA como una técnica con una capacidad intermedia para considerar significativas las asociaciones entre las rutas y el CCR. Los resultados de la presente tesis confirman esta hipótesis, habiéndose detectado la asociación del CCR con 6 rutas según PoDA, con 1 ruta según fGSEA y no detectando asociación con ninguna ruta mediante el SRT.

Cuando analizamos la plausibilidad biológica y la consistencia de las asociaciones detectadas, es reseñable que la asociación detectada, tanto por PoDA como por fGSEA, entre el CCR y la ruta de la Enfermedades Inflamatorias del Intestino es concordante con los datos obtenidos por Azer et al., (2013), que defienden que la ruta de las enfermedades inflamatorias del intestino se asocia directamente con la génesis del CCR (148). El resto de rutas metabólicas relacionadas con NRF2 asociadas con CCR detectadas mediante la técnica PoDA, es decir la ruta de la Biogénesis Mitocondrial, la ruta de las MAP Kinasas, la ruta de señalización NFK β , la ruta de señalización PI3K-AKT, la ruta de los transportadores ABC, también tienen plausibilidad biológica y son

consistentes con la evidencia científica, otorgando un papel más o menos relevante a NRF1 y NRF2.

De esta manera, se conoce que NRF2 junto con NRF1 son los encargados de coordinar la biogénesis mitocondrial, por lo que es de suponer que cuando se ve alterada esta ruta, existe una modificación de las mitocondrias y de los procesos que ocurren dentro de ellas, de acuerdo a lo descrito en el capítulo IV de esta tesis(164). Además, Cruz et al., (2017) también señalan que la reprogramación mitocondrial es un proceso que ocurre con cierta frecuencia en los estadios tempranos del CCR, asumiendo una relación entre la biogenesis mitocondrial y el CCR (165).

Por otro lado, la activación de NRF2 es un proceso importante que se produce a través de la fosforilación mediante kinasas en sitios específicos a través de rutas de señalización como la AMPK o PI3K, induciendo su acumulación en el núcleo y provocando la transactivación de genes que controlan los ARE. Por lo que alteraciones en estas rutas de señalización, pueden modificar la activación de NRF2 requerida para aportar protección y reducir el riesgo de CCR (153–156).

Finalmente indicar que NRF2 es conocido como un mecanismo de defensa esencial contra el estrés oxidativo (168), por lo que resulta evidente que este gen participe en el transporte y metabolismo de fármacos y otros compuestos vinculados a la ruta de los transportadores ABC (150), habiéndose sugerido que alteraciones en esta ruta, llevarían a una acumulación de compuestos sin metabolizarse que podrían transformarse en compuestos carcinogénicos y promover el CCR (56).

CAPÍTULO VIII. CONCLUSIONES

1. The systematic review showed that there are only 3 papers that have related NRF1 to CRC risk, so more studies are needed to explore this possible association.
2. The scientific evidence about NRF2 showed that an adequate basal function of NRF2 is essential to prevent oncogenic processes in colon, since this gene decreases genotoxic damage by providing cytoprotection, increases autophagy, reduces inflammation and activates a large number of antioxidant genes. All these processes controlled by NRF2 are associated with a reduction of the CRC risk.
3. According to the findings of the systematic review, NRF2 overexpression may promote inflammatory processes, increased proteosomal activity, and decreased apoptosis that may be associated with CRC risk. In addition, the overexpression of this gene could also increase the CRC development by promoting resistance to chemotherapeutic agents.
4. According to the available bibliography, NRF2 inhibition could be related to a higher susceptibility to suffer CRC caused by an increase in the number of aberrant crypts, as well as other processes related to this type of cancer such as ulcerative colitis, prolapse and rectal bleeding, among others.
5. The rs12706898 SNP located in the NRF1 gene did not show a statistically significant association with CRC risk considered globally, but, when tumor location and sex were taken into account, an association of the gT heterozygous genotype with the genesis of right colon cancer in men was observed.
6. In metabolic pathways analysis, the use of PoDA technique was shown to have the greatest capacity to detect association between the pathways studied and CRC risk, and the SRT as the least capacity, as expected according to the SNP selection criteria considered by each of them.
7. The inflammatory bowel diseases pathway was shown to be associated with CRC risk, both in the analysis by fgSEA and PoDA, coinciding with the previous bibliography using other analysis techniques.
8. According to the results obtained using the PoDA methodology, the MAP Kinases pathway, the NFK β signalling pathway, the PI3K-AKT signalling pathway, the ABC transporters pathway and the mitochondrial biogenesis pathway related to NRF1 and/or

NRFs y cáncer colorrectal

NRF2 genes showed association with the CRC risk. These results present biological plausibility although they need to be confirmed.

CAPÍTULO IX. BIBLIOGRAFÍA

1. OMS | Cáncer. *WHO* (2017). at <<http://www.who.int/topics/cancer/es/>>
2. Cicuéndez Ávila, R. A. Epidemiología del cáncer de colon y recto en hospitales públicos de la Comunidad Autónoma de Madrid. Análisis de supervivencia. (2015).
3. International Agency for Research of Cancer. Fact Sheets by Cancer. (2017). at <http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx?cancer=colorectal>
4. Galceran, J. *et al.* Cancer incidence in Spain, 2015. *Clin. Transl. Oncol.* **19**, 799–825 (2017).
5. Ferlay, J. *et al.* Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int. J. cancer* **136**, E359-86 (2015).
6. Globocan 2012 - Home. at <<http://globocan.iarc.fr/Default.aspx>>
7. Ji, L., Wei, Y., Jiang, T. & Wang, S. Correlation of Nrf2, NQO1, MRP1, cmyc and p53 in colorectal cancer and their relationships to clinicopathologic features and survival. *Int. J. Clin. Exp. Pathol.* **7**, 1124–31 (2014).
8. Van Blarigan, E. L. & Meyerhardt, J. A. Role of Physical Activity and Diet After Colorectal Cancer Diagnosis. *J. Clin. Oncol.* **33**, 1825–1834 (2015).
9. EUROCORE-5. at <<http://www.eurocare.it/Eurocare5/tabid/64/Default.aspx>>
10. Kantor, E. D. *et al.* Gene-environment interaction involving recently identified colorectal cancer susceptibility Loci. *Cancer Epidemiol. Biomarkers Prev.* **23**, 1824–33 (2014).
11. Rodríguez-Ramiro, I., Ramos, S., Bravo, L., Goya, L. & Martín, M. Á. Procyanidin B2 induces Nrf2 translocation and glutathione S-transferase P1 expression via ERKs and p38-MAPK pathways and protect human colonic cells against oxidative stress. *Eur. J. Nutr.* **51**, 881–892 (2012).
12. Dumas, S., Parent, M. E., Siemiatycki, J. & Brisson, J. Rectal cancer and occupational risk factors: a hypothesis-generating, exposure-based case-control study. *Int. J. cancer* **87**, 874–9 (2000).
13. Peters, R. K., Garabrant, D. H., Yu, M. C. & Mack, T. M. A case-control study of occupational and dietary factors in colorectal cancer in young men by subsite. *Cancer Res.* **49**, 5459–68 (1989).
14. Johnson, C. M. *et al.* Meta-analyses of colorectal cancer risk factors. *Cancer Causes Control* **24**, 1207–22 (2013).
15. Zhu, Y. *et al.* Influence of pre-diagnostic cigarette smoking on colorectal cancer survival: overall and by tumour molecular phenotype. *Br. J. Cancer* **110**, 1359–66 (2014).
16. Shiels, M. S. *et al.* Cigarette Smoking and Variations in Systemic Immune and Inflammation Markers. *JNCI J. Natl. Cancer Inst.* **106**, (2014).
17. Longnecker, M. P., Orza, M. J., Adams, M. E., Vioque, J. & Chalmers, T. C. A meta-analysis of alcoholic beverage consumption in relation to risk of colorectal

- cancer. *Cancer Causes Control* **1**, 59–68 (1990).
18. Rasool, S., Kadla, S. A., Rasool, V. & Ganai, B. A. A comparative overview of general risk factors associated with the incidence of colorectal cancer. *Tumor Biol.* **34**, 2469–2476 (2013).
 19. Lee, G. H. *et al.* Is right-sided colon cancer different to left-sided colorectal cancer? A systematic review. *Eur. J. Surg. Oncol.* **41**, 300–308 (2015).
 20. Carr, P. R., Walter, V., Brenner, H. & Hoffmeister, M. Meat subtypes and their association with colorectal cancer: Systematic review and meta-analysis. *Int. J. Cancer* **138**, 293–302 (2016).
 21. Santarelli, R., Pierre, F. & Corpet, D. Processed Meat and Colorectal Cancer: A Review of Epidemiologic and Experimental Evidence. *Nutr. Cancer* **60**, 131–144 (2008).
 22. Song, M., Garrett, W. S. & Chan, A. T. Nutrients, Foods, and Colorectal Cancer Prevention. *Gastroenterology* **148**, 1244–1260.e16 (2015).
 23. Spencer, E. A. *et al.* Meat , poultry and fish and risk of colorectal cancer : pooled analysis of data from the UK dietary cohort consortium. *Cancer Causes Control* **21**, 1417–1425 (2010).
 24. Kim, J. *et al.* Association between dietary fat intake and colorectal adenoma in korean adults. *Medicine (Baltimore)*. **96**, e5759 (2017).
 25. Kashino, I. *et al.* Vegetable consumption and colorectal cancer risk: an evaluation based on a systematic review and meta-analysis among the Japanese population. *Jpn. J. Clin. Oncol.* **45**, 973–979 (2015).
 26. Geelen, A. *et al.* Fish consumption, n-3 fatty acids, and colorectal cancer: A meta-analysis of prospective cohort studies. *Am. J. Epidemiol.* **166**, 1116–1125 (2007).
 27. Wu, S. *et al.* Fish Consumption and Colorectal Cancer Risk in Humans: A Systematic Review and Meta-analysis. *Am. J. Med.* **125**, 551–559.e5 (2012).
 28. Murphy, N. *et al.* Consumption of dairy products and colorectal cancer in the European Prospective Investigation into Cancer and Nutrition (EPIC). *PLoS One* **8**, e72715 (2013).
 29. Aune, D., Chan, D. S. M., Lau, R. & Vieira, R. Dietary fibre , whole grains , and risk of colorectal cancer : systematic review and dose-response meta-analysis of prospective studies. (2018). doi:10.1136/bmj.d6617
 30. Bamia, C. *et al.* Mediterranean diet and colorectal cancer risk: results from a European cohort. *Eur. J. Epidemiol.* **28**, 317–328 (2013).
 31. Tandon, K., Imam, M., Ismail, B. E. S. & Castro, F. Body mass index and colon cancer screening: the road ahead. *World J. Gastroenterol.* **21**, 1371–6 (2015).
 32. Dai, Z., Xu, Y. & Niu, L. Obesity and colorectal cancer risk: a meta-analysis of cohort studies. *World J. Gastroenterol.* **13**, 4199–206 (2007).

33. Wolin, K. Y., Yan, Y., Colditz, G. A. & Lee, I. M. Physical activity and colon cancer prevention: A meta-analysis. *Br. J. Cancer* **100**, 611–616 (2009).
34. Pérez Cueto, F. J. A. Colorectal cancer prevention with diet and physical activity: recent update. *Nutr. Hosp.* **26**, 1491
35. Wilkes, G. & Hartshorn, K. Clinical Update: Colon, Rectal, and Anal Cancers. *Semin. Oncol. Nurs.* **28**, e1–e22 (2012).
36. Campos, F. G. Colorectal cancer in young adults: A difficult challenge. *World J. Gastroenterol.* **23**, 5041–5044 (2017).
37. O’Connell, J. B. *et al.* Rates of colon and rectal cancers are increasing in young adults. *Am. Surg.* **69**, 866–72 (2003).
38. Gao, R.-N., Neutel, C. I. & Wai, E. Gender differences in colorectal cancer incidence, mortality, hospitalizations and surgical procedures in Canada. *J. Public Health (Bangkok)*. **30**, 194–201 (2008).
39. Doubeni, C. A. *et al.* Socioeconomic status and the risk of colorectal cancer: an analysis of more than a half million adults in the National Institutes of Health-AARP Diet and Health Study. *Cancer* **118**, 3636–44 (2012).
40. Le, H., Ziogas, A., Lipkin, S. M. & Zell, J. A. Effects of Socioeconomic Status and Treatment Disparities in Colorectal Cancer Survival. *Cancer Epidemiol. Biomarkers Prev.* **17**, 1950–1962 (2008).
41. Yao, J. *et al.* NF- κ B and Nrf2 signaling pathways contribute to wogonin-mediated inhibition of inflammation-associated colorectal carcinogenesis. *Cell Death Dis.* **5**, e1283 (2014).
42. Kim, E. R. & Chang, D. K. Colorectal cancer in inflammatory bowel disease: The risk, pathogenesis, prevention and diagnosis. *World J. Gastroenterol.* **20**, 9872 (2014).
43. Slattery, M. L. *et al.* Family history and colorectal cancer: predictors of risk. *Cancer Causes Control* **14**, 879–87 (2003).
44. Wark, P. A., Wu, K., van ’t Veer, P., Fuchs, C. F. & Giovannucci, E. L. Family history of colorectal cancer: a determinant of advanced adenoma stage or adenoma multiplicity? *Int. J. cancer* **125**, 413–20 (2009).
45. Aibar Santos, S. Bioinformática aplicada a datos genómicos para la caracterización de subtipos de cáncer: estudios integrativos en hemopatías malignas. (2015).
46. Gayardo Puyuelo, C. J. Susceptibilidad genética del cáncer colorrectal: influencia de polimorfismos genéticos en el desarrollo de cáncer colorrectal y lesiones preneoplásicas en familiares de primer grado de pacientes con cáncer colorrectal. (2016).
47. Kuosmanen, S. M. *et al.* The Effects of Sequence Variation on Genome-wide NRF2 Binding--New Target Genes and Regulatory SNPs. *Nucleic Acids Res.* **44**, 1760–1775 (2016).

48. Song, N., Shin, A., Park, J. W., Kim, J. & Oh, J. H. Common risk variants for colorectal cancer: an evaluation of associations with age at cancer onset. *Sci. Rep.* **7**, 40644 (2017).
49. Zhang, K., Civan, J., Mukherjee, S., Patel, F. & Yang, H. Genetic variations in colorectal cancer risk and clinical outcome. *World J. Gastroenterol.* **20**, 4167–77 (2014).
50. Hong, Y., Wu, G., Li, W., Liu, D. & He, K. A comprehensive meta-analysis of genetic associations between five key SNPs and colorectal cancer risk. *Oncotarget* **7**, 73945–73959 (2016).
51. Schmit, S. L. *et al.* Novel Common Genetic Susceptibility Loci for Colorectal Cancer. *J. Natl. Cancer Inst.* **111**, 1–12 (2018).
52. Jenkins, M. A. *et al.* Quantifying the utility of single nucleotide polymorphisms to guide colorectal cancer screening. *Futur. Oncol.* **12**, 503–513 (2016).
53. Li, Y. *et al.* Association of genes, pathways, and haplogroups of the mitochondrial genome with the risk of colorectal cancer: The multiethnic cohort. *PLoS One* **10**, e0136796 (2015).
54. de Leeuw, C. A., M. Neale, B., Heskes, T. & Posthuma, D. The statistical properties of gene-set analysis. *Nat. Rev. Genet.* **17**, 353–364 (2016).
55. O’Dushlaine, C. *et al.* The SNP ratio test: pathway analysis of genome-wide association datasets. *Bioinformatics* **25**, 2762–2763 (2009).
56. Lee, D., Lee, G. K., Yoon, K.-A. & Lee, J. S. Pathway-Based Analysis Using Genome-wide Association Data from a Korean Non-Small Cell Lung Cancer Study. *PLoS One* **8**, e65396 (2013).
57. Chen, Q.-R. *et al.* Multi-SNP Analysis of GWAS Data Identifies Pathways Associated with Nonalcoholic Fatty Liver Disease. *PLoS One* **8**, e65982 (2013).
58. Permut-Wey, J. *et al.* Inherited Variants in Mitochondrial Biogenesis Genes May Influence Epithelial Ovarian Cancer Risk. *Cancer Epidemiol. Biomarkers Prev.* **20**, 1131–1145 (2011).
59. Baar, K. *et al.* Skeletal muscle overexpression of nuclear respiratory factor 1 increases glucose transport capacity. *FASEB J.* **17**, 1666–1673 (2003).
60. Blesa, J. R. *et al.* NRF-1 is the major transcription factor regulating the expression of the human TOMM34 gene. *Biochem. Cell Biol.* **86**, 46–56 (2008).
61. Skonieczna, K., Malyarchuk, B. A. & Grzybowski, T. The landscape of mitochondrial DNA variation in human colorectal cancer on the background of phylogenetic knowledge. *Biochim. Biophys. Acta* **1825**, 153–9 (2012).
62. Weinberg, F. & Chandel, N. S. Mitochondrial Metabolism and Cancer. *Ann. N. Y. Acad. Sci.* **1177**, 66–73 (2009).
63. Sánchez-Aragó, M., Chamorro, M. & Cuezva, J. M. Selection of cancer cells with repressed mitochondria triggers colon cancer progression. *Carcinogenesis*

- 31**, 567–76 (2010).
64. Boland, M. L., Chourasia, A. H. & Macleod, K. F. Mitochondrial Dysfunction in Cancer. *Front. Oncol.* **3**, 1–28 (2013).
 65. Trachootham, D., Alexandre, J. & Huang, P. Targeting cancer cells by ROS-mediated mechanisms: a radical therapeutic approach? *Nat. Rev. Drug Discov.* **8**, 579–591 (2009).
 66. Mazzanti, R. & Giulivi, C. Coordination of nuclear- and mitochondrial-DNA encoded proteins in cancer and normal colon tissues. *Biochim. Biophys. Acta - Bioenerg.* **1757**, 618–623 (2006).
 67. Sanchez-Pino, M. J., Moreno, P. & Navarro, A. Mitochondrial dysfunction in human colorectal cancer progression. *Front. Biosci.* **12**, 1190–9 (2007).
 68. Greaves, L. C. *et al.* Mitochondrial DNA mutations are established in human colonic stem cells, and mutated clones expand by crypt fission. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 714–719 (2006).
 69. Thyagarajan, B., Wang, R., Barcelo, H., Koh, W. P. & Yuan, J. M. Mitochondrial copy number is associated with colorectal cancer risk. *Cancer Epidemiol. Biomarkers Prev.* **21**, 1574–1581 (2012).
 70. de Araujo, L. F. *et al.* Mitochondrial genome instability in colorectal adenoma and adenocarcinoma. *Tumor Biol.* **36**, 8869–8879 (2015).
 71. Wenz, T. Regulation of mitochondrial biogenesis and PGC-1 α under cellular stress. *Mitochondrion* **13**, 134–142 (2013).
 72. Zhang, Y. *et al.* The selective post-translational processing of transcription factor Nrf1 yields distinct isoforms that dictate its ability to differentially regulate gene expression. *Sci. Rep.* **5**, 12983 (2015).
 73. Chen, J. *et al.* Transcription factor Nrf1 is negatively regulated by its O-GlcNAcylation status. *FEBS Lett.* **589**, 2347–2358 (2015).
 74. Lee, E. J., Kang, Y. C., Park, W.-H., Jeong, J. H. & Pak, Y. K. Negative transcriptional regulation of mitochondrial transcription factor A (TFAM) by nuclear TFAM. *Biochem. Biophys. Res. Commun.* **450**, 166–171 (2014).
 75. Kim, H. M., Han, J. W. & Chan, J. Y. Nuclear Factor Erythroid-2 Like 1 (NFE2L1): Structure, function and regulation. *Gene* **584**, 17–25 (2016).
 76. Bugno, M., Daniel, M., Chepelev, N. L. & Willmore, W. G. Changing gears in Nrf1 research, from mechanisms of regulation to its role in disease and prevention. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1849**, 1260–1276 (2015).
 77. Zhao, L. *et al.* miR-504 mediated down-regulation of nuclear respiratory factor 1 leads to radio-resistance in nasopharyngeal carcinoma. *Oncotarget* **6**, 15995–16018 (2015).
 78. Biswas, M., Kwong, E. K., Park, E., Nagra, P. & Chan, J. Y. Glycogen synthase

- kinase 3 regulates expression of nuclear factor-erythroid-2 related transcription factor-1 (Nrf1) and inhibits pro-survival function of Nrf1. *Exp. Cell Res.* **319**, 1922–1931 (2013).
79. Oh, D. H., Rigas, D., Cho, A. & Chan, J. Y. Deficiency in the nuclear-related factor erythroid 2 transcription factor (Nrf1) leads to genetic instability. *FEBS J.* **279**, 4121–4130 (2012).
 80. Suliman, H. B., Carraway, M. S., Welty-Wolf, K. E., Whorton, A. R. & Piantadosi, C. A. Lipopolysaccharide Stimulates Mitochondrial Biogenesis via Activation of Nuclear Respiratory Factor-1. *J. Biol. Chem.* **278**, 41510–41518 (2003).
 81. Okoh, V. O. *et al.* Redox signalling to nuclear regulatory proteins by reactive oxygen species contributes to oestrogen-induced growth of breast cancer cells. *Br. J. Cancer* **112**, 1687–1702 (2015).
 82. van Tienen, F. H., Lindsey, P. J., van der Kallen, C. J. & Smeets, H. J. Prolonged Nrf1 overexpression triggers adipocyte inflammation and insulin resistance. *J. Cell. Biochem.* **111**, 1575–85 (2010).
 83. Zhang, Y. & Xiang, Y. Molecular and cellular basis for the unique functioning of Nrf1, an indispensable transcription factor for maintaining cell homeostasis and organ integrity. *Biochem. J.* **473**, 961–1000 (2016).
 84. Shukla, K., Sonowal, H., Saxena, A., Ramana, K. V. & Srivastava, S. K. Aldose reductase inhibitor, fidarestat regulates mitochondrial biogenesis via Nrf2/HO-1/AMPK pathway in colon cancer cells. *Cancer Lett.* **411**, 57–63 (2017).
 85. Moi, P., Chan, K., Asunis, I., Cao, A. & Kan, Y. W. Isolation of NF-E2-related factor 2 (Nrf2), a NF-E2-like basic leucine zipper transcriptional activator that binds to the tandem NF-E2/AP1 repeat of the beta-globin locus control region. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 9926–30 (1994).
 86. Furfaro, A. L. *et al.* The Nrf2/HO-1 Axis in Cancer Cell Growth and Chemoresistance. *Oxid. Med. Cell. Longev.* **2016**, 1–14 (2016).
 87. Gañán-Gómez, I., Wei, Y., Yang, H., Boyano-Adánez, M. C. & García-Manero, G. Oncogenic functions of the transcription factor Nrf2. *Free Radic. Biol. Med.* **65**, 750–764 (2013).
 88. Xiang, M., Namani, A., Wu, S. & Wang, X. Nrf2: bane or blessing in cancer? *J. Cancer Res. Clin. Oncol.* **140**, 1251–1259 (2014).
 89. Rodríguez-Ramiro, I., Ramos, S., Bravo, L., Goya, L. & Martín, M. Á. Procyanidin B2 induces Nrf2 translocation and glutathione S-transferase P1 expression via ERKs and p38-MAPK pathways and protect human colonic cells against oxidative stress. *Eur. J. Nutr.* **51**, 881–892 (2012).
 90. Köhler, U. A. *et al.* Activated Nrf2 impairs liver regeneration in mice by activation of genes involved in cell-cycle control and apoptosis. *Hepatology* **60**, 670–678 (2014).
 91. Gegotek, A. & Skrzydlewska, E. The role of transcription factor Nrf2 in skin

- cells metabolism. *Arch. Dermatol. Res.* **307**, 385–396 (2015).
92. Chang, L. C. *et al.* Immunohistochemical study of the Nrf2 pathway in colorectal cancer: Nrf2 expression is closely correlated to Keap1 in the tumor and Bach1 in the normal tissue. *Appl. Immunohistochem. Mol. Morphol.* **21**, 511–517 (2013).
 93. Jang, J., Wang, Y., Kim, H.-S., Lalli, M. A. & Kosik, K. S. Nrf2, a regulator of the proteasome, controls self-renewal and pluripotency in human embryonic stem cells. *Stem Cells* **32**, 2616–25 (2014).
 94. Rojo, A. I. *et al.* The PTEN/NRF2 Axis Promotes Human Carcinogenesis. *Antioxid. Redox Signal.* **21**, 2498–2514 (2014).
 95. Zhang, P. *et al.* Loss of Kelch-like ECH-associated protein 1 function in prostate cancer cells causes chemoresistance and radioresistance and promotes tumor growth. *Mol. Cancer Ther.* **9**, 336–346 (2010).
 96. Sporn, M. B. & Liby, K. T. NRF2 and cancer: the good, the bad and the importance of context. *Nat. Rev. Cancer* **12**, 564–571 (2012).
 97. Mitsuishi, Y. *et al.* Nrf2 Redirects Glucose and Glutamine into Anabolic Pathways in Metabolic Reprogramming. *Cancer Cell* **22**, 66–79 (2012).
 98. Cheung, K. L., Lee, J. H., Khor, T. O., Wu, T. & Li, G. X. Nrf2 Knockout Enhances Intestinal Tumorigenesis in Apc^{min} / β Mice Due to Attenuation of Anti-Oxidative Stress Pathway While Potentiates Inflammation. 1–8 (2012). doi:10.1002/mc.21950
 99. Kang, K. A., Piao, M. J., Ryu, Y. S., Kang, H. K. & Young, W. Interaction of DNA demethylase and histone methyltransferase upregulates Nrf2 in 5-fluorouracil-resistant colon cancer cells. *Oncotarget* **7**, 40594–40620 (2016).
 100. Centro Cochrane Iberoamericano traductores. Manual Cochrane de revisiones Sistemáticas de Intervenciones, versión 5.1.0 [actualizada en marzo de 2011]. (2012). at <<http://www.cochrane.es/?q=es/node/269>>
 101. Popay, J. *et al.* Guidance on the Conduct of Narrative Synthesis in Systematic Reviews. (2006).
 102. Mendeley | Gestor de referencias y red social académica gratuita | Elsevier. at <<https://www.elsevier.com/es-mx/solutions/mendeley>>
 103. Lu, M. C. *et al.* An inhibitor of the Keap1-Nrf2 protein-protein interaction protects NCM460 colonic cells and alleviates experimental colitis. *Sci. Rep.* **6**, 26585 (2016).
 104. Myers, J. N. *et al.* Implications of the colonic deposition of free hemoglobin- α chain: a previously unknown tissue by-product in inflammatory bowel disease. *Inflamm. Bowel Dis.* **20**, (2014).
 105. Yang, Y. *et al.* Chemoprevention of dietary digitoflavone on colitis-associated colon tumorigenesis through inducing Nrf2 signaling pathway and inhibition of inflammation. *Mol. Cancer* **13**, 48 (2014).

106. Jung, K. A. & Kwak, M. K. Enhanced 4-Hydroxynonenal Resistance in KEAP1 Silenced Human Colon Cancer Cells. *Oxid. Med. Cell. Longev.* **2013**, 1–16 (2013).
107. Pandurangan, A. K. & Esa, N. M. Dietary non-nutritive factors in targeting of regulatory molecules in colorectal cancer: An update. *Asian Pacific J. Cancer Prev.* **14**, 5543–5552 (2013).
108. Pandurangan, A. K., Kumar, S., Sadagopan, A., Dharmalingam, P. & Ganapasam, S. Luteolin, a bioflavonoid inhibits Azoxy methane-induced colorectal cancer through activation of Nrf2 signaling. **6516**, 13–20 (2014).
109. Pandurangan, A. K. & Esa, N. M. Luteolin, a bioflavonoid inhibits colorectal cancer through modulation of multiple signaling pathways: A review. *Asian Pacific J. Cancer Prev.* **15**, 5501–5508 (2014).
110. Stachel, I. *et al.* Modulation of nuclear factor E2-related factor-2 (Nrf2) activation by the stress response gene immediate early response-3 (IER3) in colonic epithelial cells: A novel mechanism of cellular adaptation to inflammatory stress. *J. Biol. Chem.* **289**, 1917–1929 (2014).
111. Sebens, S. *et al.* Inflammatory macrophages induce Nrf2 transcription factor-dependent proteasome activity in colonic NCM460 cells and thereby confer anti-apoptotic protection. *J. Biol. Chem.* **286**, 40911–40921 (2011).
112. Kim, T. H. *et al.* NRF2 blockade suppresses colon tumor angiogenesis by inhibiting hypoxia-induced activation of HIF-1 α . *Cancer Res.* **71**, 2260–2275 (2011).
113. Kang, K. A. *et al.* Epigenetic modification of Nrf2 in 5-fluorouracil-resistant colon cancer cells: involvement of TET-dependent DNA demethylation. *Cell Death Dis.* **5**, e1183 (2014).
114. Li, W. *et al.* Activation of Nrf2-antioxidant signaling attenuates NF-kappaB-inflammatory response and elicits apoptosis. *Biochem. Pharmacol.* **76**, 1485–1489 (2008).
115. Yokoo, Y. *et al.* Effects of Nrf2 silencing on oxidative stress-associated intestinal carcinogenesis in mice. *Cancer Med.* **5**, 1228–38 (2016).
116. Saw, C. L. L. & Kong, A. N. T. Nuclear factor-erythroid 2-related factor 2 as a chemopreventive target in colorectal cancer. *Expert Opin. Ther. Targets* **15**, 281–95 (2011).
117. Cheung, K. L. *et al.* Nrf2 knockout enhances intestinal tumorigenesis in Apcmin/+mice due to attenuation of anti-oxidative stress pathway while potentiates inflammation. *Mol. Carcinog.* **53**, 77–84 (2014).
118. Arlt, A. *et al.* Increased proteasome subunit protein expression and proteasome activity in colon cancer relate to an enhanced activation of nuclear factor E2-related factor 2 (Nrf2). *Oncogene* **28**, 3983–3996 (2009).
119. Menegon, S., Columbano, A. & Giordano, S. The Dual Roles of NRF2 in Cancer. *Trends Mol. Med.* **22**, 578–593 (2016).

120. D'Errico, I. *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.* **108**, 6603–8 (2011).
121. Castaño-Vinyals, G. *et al.* Population-based multicase-control study in common tumors in Spain (MCC-Spain): Rationale and study design. *Gac. Sanit.* **29**, 308–315 (2015).
122. Trichopoulou, A., Costacou, T., Bamia, C. & Trichopoulos, D. Adherence to a Mediterranean Diet and Survival in a Greek Population. *N. Engl. J. Med.* **348**, 2599–2608 (2003).
123. Reference SNP (refSNP) Cluster Report: rs12706898. at https://www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi?rs=12706898
124. rs12706898 (SNP) - Population genetics - Homo sapiens - GRCh37 Archive browser 92. at http://grch37.ensembl.org/Homo_sapiens/Variation/Population?r=7:129254497-129255497;v=rs12706898;vdb=variation;vf=7822059
125. Alonso, S. Poliformismo y sobredominancia en el genoma humano | Investigación y Ciencia | Investigación y Ciencia. at <https://www.investigacionyciencia.es/revistas/investigacion-y-ciencia/mundo-cuntico-y-relatividad-482/poliformismo-y-sobredominancia-en-el-genoma-humano-186>
126. Ziehe, M. & Gregorius, H.-R. The significance of over- and underdominance for the maintenance of genetic polymorphisms: I. Underdominance and stability. *J. Theor. Biol.* **117**, 493–504 (1985).
127. He, Z.-H. *et al.* PGC -related gene variants and elite endurance athletic status in a Chinese cohort: A functional study. *Scand. J. Med. Sci. Sports* **25**, 184–195 (2015).
128. Sequencing | Key methods and uses. at <https://www.illumina.com/techniques/sequencing.html>
129. Kenehisa Laboratories. KEGG PATHWAY: Inflammatory bowel disease (IBD) - Homo sapiens (human). at [http://www.kegg.jp/kegg-bin/highlight_pathway?scale=1.0&map=hsa05321&keyword=inflammatory bowel disease](http://www.kegg.jp/kegg-bin/highlight_pathway?scale=1.0&map=hsa05321&keyword=inflammatory%20bowel%20disease)
130. KEGG PATHWAY: Longevity regulating pathway - multiple species - Homo sapiens (human). at http://www.kegg.jp/kegg-bin/highlight_pathway?scale=1.0&map=hsa04213&keyword=LONGEVITY
131. KEGG PATHWAY: MAPK signaling pathway - Homo sapiens (human). at http://www.kegg.jp/kegg-bin/highlight_pathway?scale=1.0&map=hsa04010&keyword=MAPK
132. KEGG PATHWAY: NF-kappa B signaling pathway - Homo sapiens (human). at http://www.kegg.jp/kegg-bin/highlight_pathway?scale=1.0&map=hsa04064&keyword=nfkb

133. KEGG PATHWAY: PI3K-Akt signaling pathway - Homo sapiens (human). at http://www.kegg.jp/kegg-bin/highlight_pathway?scale=1.0&map=hsa04151&keyword=PI3K>
134. KEGG PATHWAY: Pentose phosphate pathway - Homo sapiens (human). at http://www.kegg.jp/kegg-bin/highlight_pathway?scale=1.0&map=hsa00030&keyword=PENTOSESPHOSPHATE>
135. KEGG PATHWAY: ABC transporters - Homo sapiens (human). at [http://www.kegg.jp/kegg-bin/highlight_pathway?scale=1.0&map=hsa02010&keyword=transporter drug](http://www.kegg.jp/kegg-bin/highlight_pathway?scale=1.0&map=hsa02010&keyword=transporter%20drug)>
136. May, B. *Mitochondrial biogenesis. Reactome - a curated knowledgebase Biol. pathways* **48**, (2014).
137. KEGG PATHWAY: AMPK signaling pathway - Homo sapiens (human). at http://www.kegg.jp/kegg-bin/highlight_pathway?scale=1.0&map=hsa04152&keyword=ampk>
138. Kruse, M.-L. L. *et al.* Colonic Lamina Propria Inflammatory Cells from Patients with IBD Induce the Nuclear Factor- κ B Related Factor-2 Thereby Leading to Greater Proteasome Activity and Apoptosis Protection in Human Colonocytes. *Inflamm. Bowel Dis.* **22**, 2593–2606 (2016).
139. Moura, F. A., de Andrade, K. Q., dos Santos, J. C. F., Araújo, O. R. P. & Goulart, M. O. F. Antioxidant therapy for treatment of inflammatory bowel disease: Does it work? *Redox Biol.* **6**, 617–39 (2015).
140. Sergushichev, A. An algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation. *bioRxiv* 060012 (2016). doi:10.1101/060012
141. Braun, R. & Buetow, K. Pathways of Distinction Analysis: A New Technique for Multi-SNP Analysis of GWAS Data. *PLOS Genet.* **7**, e1002101 (2011).
142. SNP ratio test download | SourceForge.net. at <https://sourceforge.net/projects/snpratiotest/>>
143. Bioconductor - fgsea. at <https://bioconductor.org/packages/release/bioc/html/fgsea.html>>
144. Braun, R. PoDA. at <http://braun.tx0.org/PoDA/>>
145. Sebastian, S. *et al.* Colorectal cancer in inflammatory bowel disease: Results of the 3rd ECCO pathogenesis scientific workshop (I). *J. Crohn's Colitis* **8**, 5–18 (2014).
146. Robles, A. I. *et al.* Whole-Exome Sequencing Analyses of Inflammatory Bowel Disease-Associated Colorectal Cancers. *Gastroenterology* **150**, 931–943 (2016).
147. Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci.* **102**, 15545–15550 (2005).

148. Azer, S. A. Overview of molecular pathways in inflammatory bowel disease associated with colorectal cancer development. *Eur. J. Gastroenterol. Hepatol.* **25**, 271–281 (2013).
149. Bruns, D. R. *et al.* Nrf2 Signaling and the Slowed Aging Phenotype: Evidence from Long-Lived Models. *Oxid. Med. Cell. Longev.* **2015**, 1–15 (2015).
150. Wang, X. *et al.* Nrf2 upregulates ATP binding cassette transporter expression and activity at the blood-brain and blood-spinal cord barriers. *J. Neurosci.* **34**, 8585–93 (2014).
151. Taylor, N. M. I. *et al.* Structure of the human multidrug transporter ABCG2. *Nature* **546**, 504–509 (2017).
152. Hu, T., Li, Z., Gao, C.-Y. Y. & Cho, C. H. Mechanisms of drug resistance in colon cancer and its therapeutic strategies. *World J. Gastroenterol.* **22**, 6876 (2016).
153. Mo, C. *et al.* The crosstalk between Nrf2 and AMPK signal pathways is important for the anti-inflammatory effect of berberine in LPS-stimulated macrophages and endotoxin-shocked mice. *Antioxid. Redox Signal.* **20**, 574–88 (2014).
154. Joo, M. S. *et al.* AMPK Facilitates Nuclear Accumulation of Nrf2 by Phosphorylating at Serine 550. *Mol. Cell. Biol.* **36**, 1931–1942 (2016).
155. Sun, Z., Huang, Z. & Zhang, D. D. Phosphorylation of Nrf2 at Multiple Sites by MAP Kinases Has a Limited Contribution in Modulating the Nrf2-Dependent Antioxidant Response. *PLoS One* **4**, e6588 (2009).
156. Nakaso, K. *et al.* PI3K is a key molecule in the Nrf2-mediated regulation of antioxidative proteins by hemin in human neuroblastoma cells. *FEBS Lett.* **546**, 181–4 (2003).
157. Huang, R. *et al.* Osteopontin Promotes Cell Migration and Invasion, and Inhibits Apoptosis and Autophagy in Colorectal Cancer by activating the p38 MAPK Signaling Pathway. *Cell. Physiol. Biochem.* **41**, 1851–1864 (2017).
158. Jana, A. *et al.* NFκB is essential for activin-induced colorectal cancer migration via upregulation of PI3K-MDM2 pathway. *Oncotarget* **8**, 37377–37393 (2017).
159. Zulato, E. *et al.* Prognostic significance of AMPK activation in advanced stage colorectal cancer treated with chemotherapy plus bevacizumab. *Br. J. Cancer* **111**, 25–32 (2014).
160. Heiss, E. H., Schachner, D., Zimmermann, K. & Dirsch, V. M. Glucose availability is a decisive factor for Nrf2-mediated gene expression. *Redox Biol.* **1**, 359–365 (2013).
161. Kowalik, M. A., Columbano, A. & Perra, A. Emerging Role of the Pentose Phosphate Pathway in Hepatocellular Carcinoma. *Front. Oncol.* **7**, 87 (2017).
162. Jornayvaz, F. R. & Shulman, G. G. I. Regulation of mitochondrial biogenesis. *Essays Biochem* **47**, 69–84 (2010).

163. Bragoszewski, P., Turek, M. & Chacinska, A. Control of mitochondrial biogenesis and function by the ubiquitin-proteasome system. *Open Biol.* **7**, 170007 (2017).
164. Gonzalez-Donquiles, C. *et al.* The NRF2 transcription factor plays a dual role in colorectal cancer: A systematic review. *PLoS One* **12**, e0177549 (2017).
165. Cruz, M. Dela *et al.* Metabolic reprogramming of the premalignant colonic mucosa is an early event in carcinogenesis. *Oncotarget* **8**, 20543–20557 (2017).
166. Jaramillo, M. & Zhang, D. The emerging role of the Nrf2–Keap1 signaling pathway in cancer. *Genes Dev.* **27**, 2179–2191 (2013).
167. Ralph, S. J., Rodríguez-Enríquez, S., Neuzil, J., Saavedra, E. & Moreno-Sánchez, R. The causes of cancer revisited: ‘Mitochondrial malignancy’ and ROS-induced oncogenic transformation - Why mitochondria are targets for cancer therapy. *Mol. Aspects Med.* **31**, 145–170 (2010).
168. Zhang, H., Davies, K. J. A. & Forman, H. J. Oxidative stress response and Nrf2 signaling in aging. *Free Radic. Biol. Med.* **88**, 314–336 (2015).

ARTÍCULOS CIENTÍFICOS
RELACIONADOS CON LA
PRESENTE TESIS

Los resultados de esta tesis han dado lugar a dos publicaciones científicas:

- Gonzalez-Donquiles, C., Alonso-Molero, J., Fernandez-Villa, T., Vilorio-Marqués, L., Molina, A. J., & Martín, V. (2017). The NRF2 transcription factor plays a dual role in colorectal cancer: a systematic review. *PloS one*, 12(5), e0177549. [Factor de impacto: 2.8; Q1 de la categoría Multidisciplinary Science]
- Gonzalez-Donquiles, C., Alonso-Molero, J., Vilorio-Marqués, L., Bustamante, M., Molina AJ., Fernández de Larrea, N. et al. Analysis of NRF2-related predictive pathways associated with colorectal cancer. Pendiente de publicación. Enviado a la revista "Scientific Reports".

CAPÍTULO IX. ANEXOS

Anexo I. Artículo 1: revisión sistemática

RESEARCH ARTICLE

The NRF2 transcription factor plays a dual role in colorectal cancer: A systematic review

C. Gonzalez-Donquiles^{1,2*}, J. Alonso-Molero², T. Fernandez-Villa², L. Vilorio-Marqués², A. J. Molina², V. Martín^{1,2}

1 CIBER Epidemiología y Salud Pública (CIBERESP), Madrid, Spain, **2** Gene-Environment and Health Research Group, University of Leon, León, Spain

* cgonzd00@estudiantes.unileon.es

Abstract

Background

Colorectal cancer is one of the most common cancers worldwide, and is influenced by the interplay of various factors, including a very strong genetic component. For instance, incorrect mitochondrial biogenesis is correlated with increased risk of developing colorectal cancer. Thus, it is important to understand the consequences of changes in both the expression and the correct function of the transcription factors that regulate mitochondrial biogenesis, namely NRF2.

Objectives

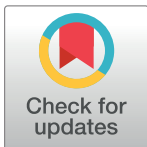
The main objective of this paper is to characterise the relationship between NRF2 and colorectal cancer by compiling data from an exhaustive literature search.

Methods

Information was obtained by defining specific search terms and searching in several databases. After a strict selection procedure, data were tabulated and the relationships between articles were assessed by measuring heterogeneity and by constructing conceptual maps.

Results and discussion

We found a general consensus in the literature that the presence of oxidizing agents as well as the inhibition of the NRF2 repressor Keap1 maintain NRF2 expression at basal levels. This predominantly exerts a cytoprotective effect on cells and decreases risk of colorectal cancer. However, if NRF2 is inhibited, protection against external agents disappears and risk of colorectal cancer increases. Interestingly, colorectal cancer risk is also increased when NRF2 becomes overexpressed. In this case, the increased risk arises from NRF2-induced inflammation and resistance to chemotherapy.



OPEN ACCESS

Citation: Gonzalez-Donquiles C, Alonso-Molero J, Fernandez-Villa T, Vilorio-Marqués L, Molina AJ, Martín V (2017) The NRF2 transcription factor plays a dual role in colorectal cancer: A systematic review. PLoS ONE 12(5): e0177549. <https://doi.org/10.1371/journal.pone.0177549>

Editor: Sergei Grivennikov, Fox Chase Cancer Center, UNITED STATES

Received: January 18, 2017

Accepted: April 28, 2017

Published: May 18, 2017

Copyright: © 2017 Gonzalez-Donquiles et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: The authors received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

Conclusion

The proper basal function of NRF2 and Keap1 are essential for preventing oncogenic processes in the colon. Consequently, any disruption to the expression of these genes can promote the genesis and progression of colon cancer.

Introduction

Colorectal cancer (CRC) is one of the most common cancers worldwide, and causes more than half a million deaths per year [1]. This type of cancer is a useful model for studying both carcinogenesis and tumour progression, as colonocytes follow a systematic process of proliferation, differentiation and transformation (adenoma to a carcinoma) [2]. CRC arises from stem cells in the colon [3], and its development is partly due to functional changes in the mitochondrial genome (mtDNA) [4].

Mitochondria play a crucial role in physiological processes, and their deregulation contributes to the genesis and progression of CRC. This is known as the Warburg effect, which consists of metabolic reprogramming that is characteristic of tumour processes. Thus, it is important that all genes involved in mitochondrial biogenesis are expressed according to a standard pattern. This process is highly coordinated and depends on the activity of several key proteins, including Nuclear Respiratory Factor 1 (NRF1) and Nuclear Respiratory Factor 2 (NRF2), which regulate the expression of respiratory chain proteins, and act as anti-oxidants [5].

NRF transcription factors control the functionality of the respiratory chain, the main pathway responsible for producing reactive oxygen species (ROS). ROS are extremely harmful for cellular micromolecules because they induce tumorigenesis through non-specific reactions with nucleic acids, proteins, and lipids [6].

NRF2 is considered to be a main defence mechanism and regulator of cell survival [7]. Under healthy conditions, NRF2 protects against tumorigenesis and cancer progression by attenuating genotoxic compounds that emerge both intrinsically and extrinsically. However, activation of the NRF2 defence response can promote the survival of both normal and cancer cells by creating an optimal environment for cell growth. Furthermore, NRF2 protects tumour cells from oxidative stress, chemotherapeutic agents, and radiotherapy, promoting tumour genesis and progression, and also metabolic reprogramming to anabolic pathways [8]. Fig 1 shows a simplified diagram of the mechanisms that can activate the transcription factor NRF2. NRF2 is sequestered by the inhibitor Keap1 and is transcriptionally inactive until Keap1 is dissociated by ROS. This dissociation allows NRF2 to form transcriptionally active complexes with other proteins, such as Mafs, and, after the transcription of cytoprotective and antioxidant genes, to remove ROS [9].

In this study, we provide a qualitative systematic review to summarize the role of the NRF2 transcription factor. Our goal was to assess the effects of differential expression of NRF2 on CRC.

Methods

Literature search strategy

Literature searches and quality assessment were performed according to 'Guidance on the Conduct of Narrative Synthesis in Systematic Reviews' (S1 File) [10]. The search results were

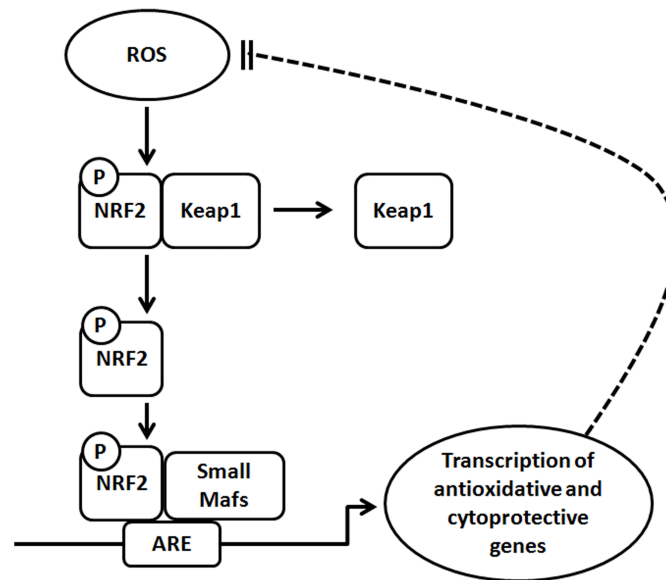


Fig 1. Simplified diagram of the activation of NRF2 transcription factor via ROS. When ROS is accumulated, NRF2 is released from Keap1 and heterodimerizes with Small Mafs. In this way, it binds to AREs and the transcription of antioxidant and cytoprotective genes occurs.

<https://doi.org/10.1371/journal.pone.0177549.g001>

independently assessed by two individuals to minimize bias (one was an expert on this topic and the other had no knowledge of it).

Design of search strategies and use of Boolean operators

For each search, synonyms, related terms and spelling variations were taken into account, and Boolean operators were used to relate the terms. Searches were performed on eight different databases (PubMed, Scopus, Web of Science (WOS), Cochrane library, PysCinfo, PLoS One, Scielo and PubMedCentral (PMC)) using the following 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 'NRF2 OR NFE2L2 OR Nuclear Respiratory Factor 2 OR Nuclear Factor Erythroid 2 Like 2'.

'Colorectal OR Colon OR Rectum OR Rectal' AND 'Cancer OR Carcinoma OR Tumor OR Tumour OR Neoplasm OR Cancer Cells' AND 'Warburg effect' AND 'NRF2 OR NFE2L2 OR Nuclear Respiratory Factor 2 OR Nuclear Factor Erythroid 2 Like 2'.

'Colorectal OR Colon OR Rectum OR Rectal' AND 'Cancer OR Carcinoma OR Tumor OR Tumour OR Neoplasm OR Cancer Cells' AND 'Oxidative phosphorylation OR OXPHOS OR Anaerobic glycolysis' AND 'NRF2 OR NFE2L2 OR Nuclear Respiratory Factor 2 OR Nuclear Factor Erythroid 2 Like 2'.

To retrieve optimal results when using the Scopus database, it was necessary to modify the search criteria. Thus, only the following more specific terms were used:

'CRC' AND 'NRF2 OR NFE2L2' AND 'Warburg'

'CRC' AND 'NRF2 OR NFE2L2' AND 'OXPHOS'.

We included only scientific reports written in English, and published after 2006, to obtain a current review of this issue. We collected articles using all combinations of search terms.

Selection process

The articles returned by the searches were manually screened to identify relevant articles to be included in this review: **1.** Search results were integrated using informatics programs. We used the EndNote reference manager to remove duplicate articles. **2.** Titles and abstracts were assessed to remove irrelevant articles. Only articles satisfying the following criteria were included: (i) the article considers CRC or other related diseases; (ii) the title of the article includes the word NRF2, or the name of another related gene; (iii) the article discusses the relationship between CRC and either mitochondrial biogenesis, OXPHOS or the Warburg effect. **3.** The full texts of potentially relevant articles were retrieved and analysed to verify that they met the inclusion criteria.

Fig 2 shows a flow chart summarising each step in the search process. Reports that did not meet the inclusion criteria were excluded from the analysis. S1 Table lists the articles excluded in the last step of the search process.

Assessment of results

Tabulation. We tabulated the following key data from each article: first author, publication year, organism/cell type, primary results and conclusion (S2 Table).

Heterogeneity. To assess the heterogeneity of the selected articles, we analysed the following aspects: expression of NRF2, relationship to CRC, use of cell culture, NRF2 as a chemotherapeutic agent, apoptosis, the inflammatory process and mitochondrial function. To more

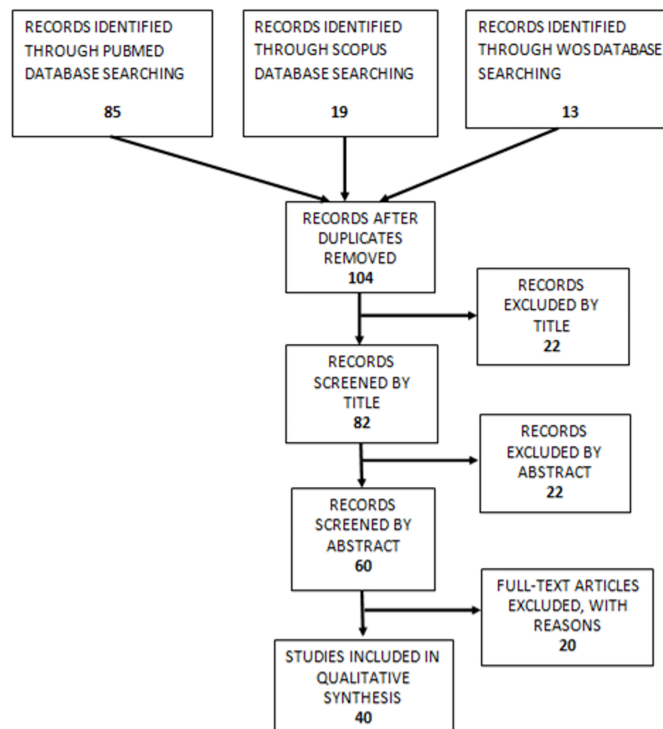


Fig 2. PRISMA 2009 flow diagram. This flow chart shows the process of selecting articles included in the qualitative systematic review.

<https://doi.org/10.1371/journal.pone.0177549.g002>

easily view the specific characteristics of each article, and to evaluate the homogeneity among them, we represented each article in a stacked bar chart (Fig 3).

Concept map. We constructed three maps to visualize the links between the selected articles. We aimed to relate three key concepts behind NRF2 expression and its role in the development of CRC so that we could draw conclusions and suggest ideas for future research.

Results

Article selection

The search described above identified relevant articles only in the Pubmed, Web of Science, and Scopus databases. Removing duplicated articles from these search results resulted in a total of 104 articles. We reviewed the titles of these articles and discarded those not related to our topic, reducing the number of articles to 82. We then reviewed the abstracts, further reducing the number of articles to 60. Finally, we reviewed the full text, and selected 40 articles for the review. This article selection process is summarized in Fig 2. Articles excluded in the last step are summarized in Supplementary Information 1.

Tabulation

Complete data from the 40 articles selected for this review are shown in Supplementary Information 2. All selected papers were published after 2006. The geographic settings of the studies vary, with nine from in the USA, nine from China, seven from Korea, four from Germany, four from Malaysia, three from India, two from Japan, and one each from the United Kingdom, Norway, and Australia. Most studies were based on human cancer cell lines. Only 13 studies used animal experimentation (mice), and four considered patients with ulcerative colitis or Crohn’s disease. The most important results from the selected articles are shown in Supplementary Information 2. Titles, author and publication date of the included articles are shown in Table 1.

Heterogeneity

The characteristics of each report are shown in Fig 3.

Building a concept map

To interconnect the ideas gathered from the articles included in this review, we constructed three conceptual maps focused on the expression of NRF2 and its effect on colon cancer. The

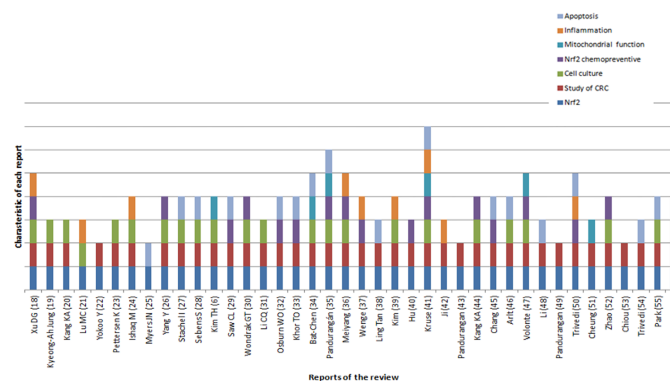


Fig 3. Heterogeneity assessment. This graph evaluates if certain features are present in the articles included in this review.

<https://doi.org/10.1371/journal.pone.0177549.g003>

Table 1. Summarized table about the included reports.

| First author | Primary results |
|-------------------------------|--|
| Kim et al., 2011 [6] | Hypoxia-induced angiogenesis can be blocked by inhibiting Nrf2. Inhibition of Nrf2 reduced tumour growth and decreased angiogenesis in mice xenografts, and was associated with a lower accumulation of HIF-1 α under hypoxic conditions. Hypoxia cannot activate HIF-1 α in Nrf2-inhibited colon cancer cells. PGC1- α and NRF1 levels were unaffected by inhibition of Nrf2. However, destabilization of HIF-1 α is associated with a weakened mitochondrial function in colon cancer cells. |
| Bat-Chen et al., 2010 [11] | Allicin induces apoptosis in colon cancer cells and stimulates Nrf2 nuclear accumulation. |
| Pandurangan et al., 2014 [12] | Nrf2 is activated by luteolin |
| Kim et al., 2013 [13] | Catechol residues are essential for activating Nrf2. |
| Pandurangan et al., 2015 [14] | Cocoa activates Nrf2 expression. |
| Li et al., 2016 [15] | Nrf2 is activated by luteolin. |
| Chiou et al., 2011 [16] | PS and RS promote Nrf2 activation and reduce colon tumorigenesis. |
| Trivedi et al., 2013 [17] | Lipoic acid (LA) increases Nrf2 and HO-1 expression in the colon. LA protects against DNA damage in the colon. |
| Pettersen et al., 2016 [18] | In SW620 cells, DHA induces an increase in ROS and causes the nuclear import of Nrf2. An increase in the level of Nrf2 was detected in the nucleus of CaCo2 and SW620 cells. |
| Myers et al., 2014 [19] | Nrf2 levels are lower in inflamed tissue of patients with UC and Crohn's disease. Nrf2 protein levels were twice as high in patients with diverticulitis than in patients with UC-associated cancer. |
| Xi et al., 2013 [20] | Nrf2 has an inhibitory effect on the development of the CRC. Nrf2 diminishes the development of dysplasia. Nrf2 is overexpressed in tumour tissues. |
| Park et al., 2010 [21] | Rottlerin induces Nrf2 nuclear translocation. |
| Jang et al., 2016 [22] | Simvastatin induces Nrf2 expression and its nuclear translocation. Simvastatin induces Nrf2-related antioxidant expression through the ERK and PI3K/Akt pathways. |
| Trivedi et al., 2016 [23] | MEL increases the expression of Nrf2 and other target genes. |
| Xu et al., 2015 [24] | PYDDT promotes ROS production, increases Nrf2 expression and decreases the CRC risk due to a reduction in genotoxicity and an increment of apoptosis. |
| Tan et al., 2015 [25] | Brewer's rice promotes Nrf2 activation. |
| Volonte et al., 2013 [26] | Calveolin inhibits Nrf2 translocation to the nucleus, thereby promoting premature senescence of CRC cells. |
| Kruse et al., 2016 [27] | IMC-co-cultured NCM460 or Colo320 cancer cells were less sensitive to TRAIL/etoposide-induced apoptosis due to Nrf2-induced proteasome activity. Immunostaining of IBD tissues confirmed Nrf2 activation within inflamed areas of the colonic epithelium, and greater proteasome protein expression. |
| Yang et al., 2014 [28] | Low concentrations of digitoflavonoids are potential Nrf2/Are activators in colon tumour, liver and kidney cells. Digitoflavonoids stimulate the expression of antioxidant defence proteins, the expression of Nrf2, and its translocation to the nucleus. After treatment with digitoflavonoids, AKT, ERK1/2 and p38 AMPK phosphorylation increases. Inhibition of AKT and ERK1/2 phosphorylation does not influence Nrf2 activation. p38 phosphorylation however is essential for Nrf2 activation. Mice treated with digitoflavonoids exhibited a decrease in the number and size of tumours and crypts as a result of Nrf2 (and its target genes) activation. |
| Pandurangan et al., 2014 [29] | Nrf2 transactivation and translocation. |
| Wondrak et al., 2010 [30] | Cinnamaldehyde upregulates Nrf2 and its downstream target genes. Nrf2 provides protection against oxidative stress-induced genotoxicity. |
| Li et al., 2009 [31] | Exposing HTC116 cells to NO results in the nuclear translocation of Nrf2 as well as its transcriptional upregulation. |

(Continued)

Table 1. (Continued)

| First author | Primary results |
|---------------------------|---|
| Jung et al., 2013 [32] | Silenced Keap1 activates Nrf2 signalling in colon cancer cell lines. Silencing Keap1 can upregulate the expression of AKRs and attenuate oxidative damage caused by stress. |
| Lu et al., 2016 [33] | CPUY192018 provides cytoprotection against oxidative damage caused by DSS in NCM460 cells. Pretreatment with CPUY192018 increases the survival rate by inhibiting apoptosis and by impeding the DSS-induced arrest of the S phase of the cell cycle. CPUY192018 induces an increase in Nrf2 protein levels in colonic NCM460 cells and its accumulation in the nucleus. |
| Stachel et al., 2014 [34] | IER3 overexpression inhibits Nrf2 activation. The absence of IER3 causes a reduction in ROS levels as a result of an increase in Nrf2 activity. |
| Sebens et al., 2011 [35] | Nrf2 is activated when cells are exposed to inflammatory macrophages. This increases proteosomal activity and the expression of proteosomal genes in a Nrf2-dependent manner. |
| Yokoo et al., 2016[36] | In Nrf2 ^{-/-} mice treated with KBrO ₃ , atypical hyperplasia, adenoma and adenocarcinoma were all observed in the upper small intestine. These same mice displayed an increase in the number of aberrant crypts compared with their Nrf2 ^{+/+} counterpart. |
| Kang et al., 2014 [37] | Nrf2 promotes resistance to 5-FU treatment in CRC cells. |
| Zhao et al., 2015 [38] | Nrf2 expression is higher in CRC cells that are resistant to 5-FU treatment. |
| Hu et al., 2013 [39] | Nrf2 is higher expressed in tumour tissue than in normal tissue. Nrf2 overexpression is related to larger tumours with advanced stage and metastasis. Directional metastases may be associated with activation of Nrf2. |
| Ji et al., 2014 [40] | Nrf2 expression is higher in CRC tissue, and is positively correlated with Duke's stage and clinical prognosis. |
| Kang et al., 2016 [41] | Nrf2 expression is higher in those colon cells resistant to 5-FU. Demethylation upregulates Nrf2 in 5-FU resistant SNUC5 cells. |
| Ishaq et al., 2014[42] | Nrf2 is involved in the protective response of HT29 cancer cells. It induces apoptosis through activating caspase 3/7. |
| Arlt et al., 2009 [43] | There is a high activity of Nrf2 in the nucleus of CRC cells. Nrf2 overexpression increases proteosomal activity. |
| Li et al., 2008 [44] | Compared with WT mice, Nrf2 KO mice suffer from more severe ulcerative colitis, loss of crypts, an infiltration of inflammatory cells, and rectal bleeding. |
| Khor et al., 2006 [45] | Compared with WT mice, Nrf2 KO mice had a smaller average colon length. After exposure to DSS, Nrf2 KO mice showed a loss in crypts with severe inflammation. |
| Osburn et al., 2007 [46] | DSS exposure in Nrf2-deficient mice led to the formation of multiple aberrant crypts. In addition, treatment with AOM promoted colitis-associated tumorigenesis (due to inflammation of the colon). |
| Saw et al., 2011 [47] | Nrf2 KO mice suffer more severe colitis than WT mice after treatment with DSS. After adding AOM, Nrf2 KO mice have a greater tendency to develop CRC and dysplasia, and display increased prolapse and rectal bleeding. Nrf2 is required for protection against inflammation-associated CRC. Nrf2 KO mice have a greater susceptibility of developing aberrant crypts associated with inflammation. |
| Cheung et al., 2014 [48] | Nrf2 KO mice have an increased number and size of polyps and cell proliferation. |
| Chang et al., 2013 [49] | Proteins of the Nrf2 pathway are actively expressed in CRC. |

<https://doi.org/10.1371/journal.pone.0177549.t001>

first conceptual map was constructed on the basis of standard NRF2 gene expression and how it is influenced by various mechanisms or external agents. The second conceptual map explains the various effects and mechanisms of NRF2 overexpression, and the third shows the effects of NRF2 silencing or inhibition. Analysing all three conceptual maps sheds light onto how NRF2 expression affects CRC. As shown in Fig 4, correct expression of NRF2 prevent

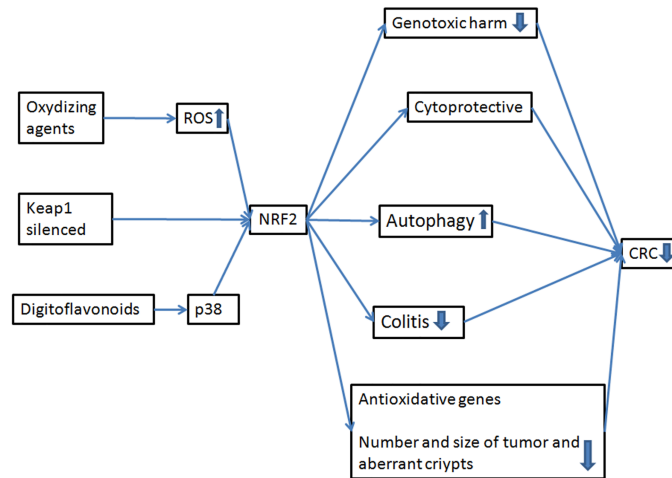


Fig 4. Standard expression of NRF2 and the relation with CRC. Oxidizing agents, the silencing of the Keap1 or the digitoflavonoids induce the expression of NRF2, decreasing the risk of CRC.

<https://doi.org/10.1371/journal.pone.0177549.g004>

CRC in humans, while altered expression of this gene promotes tumour genesis and progression (Figs 5 and 6).

Discussion

The articles selected show that research on the role of NRF2 in CRC can be classified into three broad groups: i) those that analyse standard expression and observe a cytoprotective effect on CRC [11–17], and those that analyse variations in expression, either ii) overexpression or iii) inhibition, both leading to increased CRC risk. Below, we discuss the findings in each of these three lines of research.

The first concept map (Fig 4) shows the protective effect of induced expression of the NRF2 gene on CRC following a moderate accumulation of ROS due to specific external agents [18–22]. When NRF2 activates, it exerts a cytoprotective effect by decreasing genotoxic damage and promoting autophagy [23]. Sometimes, NRF2 expression leads to reduced cell

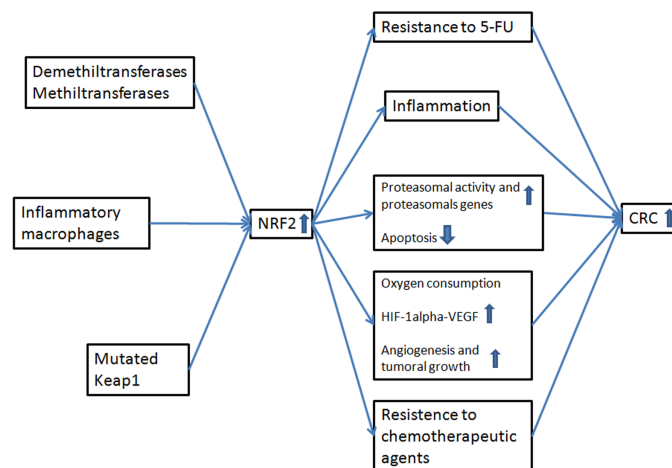


Fig 5. Overexpression of NRF2 and the relation with CRC. Constitutive activation of NRF2 promotes a series of events that lead to an increased risk of CRC.

<https://doi.org/10.1371/journal.pone.0177549.g005>

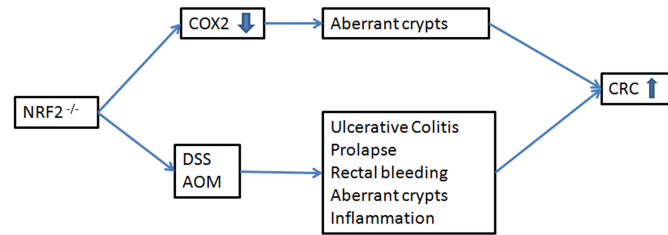


Fig 6. Inhibition of NRF2 and the relation with CRC. Inhibition or permanently silencing of NRF2 induce the formation of aberrant crypts and other typical processes related to the increased risk of CRC such as rectal bleeding or ulcerative colitis.

<https://doi.org/10.1371/journal.pone.0177549.g006>

proliferation and increased apoptosis of colon cancer cells [24–27]. In all cases, the authors observed the same final result, namely reduced risk of CRC. Another example of this behaviour was reported by Yang et al. (2014), in which digitoflavonoids were found to act upon p38 and to induce NRF2 target genes and antioxidants. They reported a reduction in both the number and size of tumours, as well as a decrease in the number of aberrant crypts, further pointing to a protective effect of NRF2 on CRC [28]. Therefore, expression of NRF2 and its target genes, such as glutathione synthase, likely play a crucial role in cytoprotective mechanisms against CRC [29–31].

In the studies mentioned up to this point, NRF2 expression is directly induced by the external agents. However, NRF2 expression can also be promoted indirectly by other mechanisms, via the repressor Keap1. For example, Jung and Kwak, 2013 silenced Keap1 and found that induction of NRF2 promoted the expression of AKR members, which are involved in the detoxification process, leading to reduced risk of CRC [32]. In another study, Lu et al., 2016 used an inhibitor of its interaction with NRF2, CPUY192018, and observed that activated NRF2 acted as a protective factor upon exposure to DSS, decreasing CRC risk [33].

In all of these cases, standard NRF2 expression levels function as a protective factor against CRC by providing cytoprotection and by activating antioxidant target genes. These results are corroborated by the review by Pandurangan and Esa, (2014) and Pandurangan et al (2014a, 2014b), who also support the idea that expression of NRF2 decreases the risk of CRC [12, 29, 50].

As shown in the second concept map (Fig 5), several studies have reported that NRF2 overexpression increases risk of CRC. This overexpression can occur for a number of reasons, including constitutive mutations in the Keap1 repressor gene or in the NRF2 gene itself. Stachel et al. (2014) found that excessive levels of ROS induced overexpression of NRF2, leading to inflammation of the colon tissue and promoting tumorigenesis [34]. In another study, Sebens et al. (2011) reported overexpression of NRF2 in tissues exposed to inflammatory macrophages, and a simultaneous increase in the activity of proteosomal genes, leading to reduced apoptosis and uncontrolled proliferation. This in turn results in greater risk of CRC [35]. Other authors have observed similar results when NRF2 is overexpressed as a consequence of silencing, or as a result of epigenetic changes in this gene [36–38, 51]. Another way to induce overexpression of NRF2 is using chemical compounds. Kim et al. (2011) found that T-BHQ-induced overexpression of NRF2 resulted in increased O₂ consumption, which is correlated with higher HIF-1 α and VEGF signalling, which in turn directly increases angiogenesis and tumour growth [6]. Therefore, as mentioned above, overexpression of the NRF2 gene is closely linked to increased risk of CRC [39, 40]. In this case, CRC risk can increase by promoting colonic inflammation, decreasing apoptosis, or promoting angiogenesis and uncontrolled cell proliferation.

Notably, NRF2 overexpression has also been linked to increased resistance to the chemotherapeutic agent 5-FU via the activity of demethylases and methyltransferases. In addition to the increase risk of CRC, this implies a poorer response to treatments based on this agent. [37,41].

The third concept map (Fig 6) shows that inhibited expression of NRF2 also has consequences for this type of cancer, since one of NRF2's functions is to protect against damaging and carcinogenic substances [42,43]. Various studies highlight the role of NRF2 on CRC via its inhibition. Li et al. (2008) related Nrf2 inhibition with increased risk of CRC [44]. Yokoo et al. (2016) found that silencing Nrf2 resulted in decreased expression of the COX2 gene, leading to an increase in the number of aberrant crypts, resulting in the formation of adenoma, adenocarcinoma and ultimately CRC [36]. Other studies have compared the effects of silencing the Nrf2 gene in mice and exposing them to treatments that induce ulcerative colitis (DSS) or colon carcinogenesis (AOM) [45–46]. Nrf2 knockout mice with ulcerative colitis suffer prolapse, rectal bleeding, inflammation, and as in the previous case, show an increase in the number of aberrant crypts. Together, these symptoms contribute to the development of CRC. This information is corroborated by reviews from Saw and Kong in 2011 [47] and Cheung et al. in 2014 [48].

Analysing the behaviour of this gene, not at a single time point, but throughout the various stages, Chang et al. (2013) suggested that the differential expression profile of NRF2 in tumour versus normal tissue could be an interesting new target for CRC treatment [49]. Similarly, Arlt et al. (2009) observed that NRF2 expression is beneficial during early stages, but can contribute to tumorigenesis in the colon at later stages [43]. These data are also supported by Menegon et al (2016) and Sporn M and Liby K (2012), who described NRF2 not only as an oncogene but also as a tumour suppressor gene [52,53]. In addition, few studies have analysed the role of Single Nucleotide Polymorphisms (SNPs) in this gene and its promotor, although Yokoo et al. (2016) suggested that some SNP alleles in the NRF2 gene could increase CRC risk based on previous studies of these polymorphisms and the risk of ulcerative colitis [36].

Finally, another important factor is the type of cell used in the studies included in this review. Most of the articles focus on colon stem cells, which need more energy and therefore require greater expression of NRF2 to obtain the necessary ATP for proliferation, migration and differentiation, such that altered expression of NRF2 increases genetic instability leading to increased CRC risk. *In vivo*, these cells migrate to the apical part of the crypt and become mature epithelial cells, and altered expression of NRF2 during migration and differentiation, will lead to an accumulation of mutations that promote tumorigenesis due to absence of cytoprotection. Thus, NRF2 function is probably essential in both the stem cells of the colon and in mature epithelial cells [54].

Although we searched several databases using an exhaustive list of relevant terms, our study has some limitations. We only selected articles published in English, which could result in selection bias. In addition, the use as search terms such as oxidative phosphorylation, mitochondria or Warburg effect, could limit the selection of articles that relate NRF2 to colorectal cancer, although searches performed without these concepts in several databases did not return any additional relevant articles.

In conclusion, it is clear that the NRF2 transcription factor has a very complex role in the cell, and that its expression can be strongly affected by external agents. Any disruption of its standard expression, be it overexpression or silencing, can promote the genesis and progression of CRC. Further studies are needed to clarify the role of NRF2, particularly to analyse the role of various SNPs that could generate variability in the expression and function of the protein in humans, as well as to study the pathways regulated by this gene, and possible interactions with environmental and behavioural factors, although such studies require large sample sizes.

Supporting information

S1 File. Guidance on the Conduct of Narrative Synthesis in Systematic Reviews.
(PDF)

S2 File. Check list.
(DOCX)

S1 Table. Excluded reports which not keep the inclusion criteria.
(DOCX)

S2 Table. Tabulation of included reports.
(DOCX)

Acknowledgments

I would like to recognize the advice and corrections given by Silvia Gutiérrez Cosio, Fernanda de Souza Teixeira and Verónica Davila Batista.

Author Contributions

Conceptualization: CGD.

Data curation: CGD AJM LVM.

Formal analysis: CGD JAM TFV.

Investigation: CGD JAM.

Methodology: CGD JAM TFV LVM AJM VM.

Project administration: VM.

Resources: CGD.

Software: CGD.

Supervision: AJM VM.

Validation: JAM TFV.

Visualization: CGD.

Writing – original draft: CGD AJM.

Writing – review & editing: CGD JAM TFV LVM AJM VM.

References

1. Siegel R, Naishadham D, & Jemal A. Cancer statistics, 2013. *CA Cancer J Clin.* 2013; 63(1), 11–30. <https://doi.org/10.3322/caac.21166> PMID: 23335087
2. 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):618–23.
3. Greaves LC, Preston SL, Tadrous PJ, Taylor RW, Barron MJ, Oukrif D, et al. Mitochondrial DNA mutations are established in human colonic stem cells, and mutated clones expand by crypt fission. *Proc Natl Acad Sci U S A.* 2006; 103(3):714–9. <https://doi.org/10.1073/pnas.0505903103> PMID: 16407113
4. Habano W, Sugai T, Yoshida T, Nakamura S ichi. Mitochondrial gene mutation, but not large-scale deletion, is a feature of colorectal carcinomas with mitochondrial microsatellite instability. *Int J Cancer.* 1999; 83(5):625–9. PMID: 10521798

5. Boland ML, Chourasia AH, Macleod KF. Mitochondrial Dysfunction in Cancer. *Front Oncol.* 2013; 3:1–28
6. Kim TH, Hur EG, Kang SJ, Kim JA, Thapa D, Mie Lee Y, et al. NRF2 blockade suppresses colon tumor angiogenesis by inhibiting hypoxia-induced activation of HIF-1 α . *Cancer Res.* 2011; 71(6):2260–75. <https://doi.org/10.1158/0008-5472.CAN-10-3007> PMID: 21278237
7. Jaramillo M, Zhang D. The emerging role of the NRF2–Keap1 signaling pathway in cancer. *Genes Dev.* 2013; 27:2179–91. <https://doi.org/10.1101/gad.225680.113> PMID: 24142871
8. Lau A, Villeneuve NF, Sun Z, Wong PK, Zhang DD. Dual roles of NRF2 in cancer. *Pharmacol Res.* 2008;(58)(5–6):262–270. <https://doi.org/10.1016/j.phrs.2008.09.003> PMID: 18838122
9. Rodríguez-Ramiro I, Ramos S, Bravo L, Goya L, Martín MÁ. Procyanidin B2 induces NRF2 translocation and glutathione S-transferase P1 expression via ERKs and p38-MAPK pathways and protect human colonic cells against oxidative stress. *Eur J Nutr.* 2012; 51: 881–92. <https://doi.org/10.1007/s00394-011-0269-1> PMID: 22042007
10. 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.
11. Bat-Chen W, Golan T, Peri I, Ludmer Z, Schwartz B. Allicin purified from fresh garlic cloves induces apoptosis in colon cancer cells via NRF2. *Nutr Cancer.* 2010; 62(7):947–57. <https://doi.org/10.1080/01635581.2010.509837> PMID: 20924970
12. Pandurangan AK, Esa NM. Luteolin, a bioflavonoid inhibits colorectal cancer through modulation of multiple signaling pathways: A review. *Asian Pacific J Cancer Prev.* 2014; 15(14):5501–8.
13. Kim H, Kim W, Yum S, Hong S, Oh J, Lee J, et al. Free Radical Biology and Medicine Caffeic acid phenethyl ester activation of NRF2 pathway is enhanced under oxidative state : Structural analysis and potential as a pathologically targeted therapeutic agent in treatment of colonic inflammation. *Free Radic Biol Med.* 2013; 65:552–62
14. Pandurangan AK, Saadatdoust Z, Hamzah H, Ismail A. Dietary cocoa protects against colitis-associated cancer by activating the NRF2/Keap1 pathway. *BioFactors.* 2015; 41(1), 1–14. <https://doi.org/10.1002/biof.1195> PMID: 25545372
15. Li Y, Shen L, Luo H. International Immunopharmacology Luteolin ameliorates dextran sulfate sodium-induced colitis in mice possibly through activation of the NRF2 signaling pathway. *Int Immunopharmacol.* 2016; 40:24–31.
16. Chiou YS, Tsai ML, Nagabhushanam K, Wang YJ, Wu CH, Ho CT, et al. Pterostilbene is more potent than resveratrol in preventing azoxymethane (AOM)-induced colon tumorigenesis via activation of the NF-E2-related factor 2 (NRF2)-mediated antioxidant signaling pathway. *J. Agric. Food Chem* 2011; 59(6), 2725–33. <https://doi.org/10.1021/jf2000103> PMID: 21355597
17. Trivedi PP, Jena GB. Role of α -lipoic acid in dextran sulfate sodium-induced ulcerative colitis in mice : Studies on inflammation, oxidative stress, DNA damage and fibrosis. *FOOD Chem Toxicol.* 2013; 59, 339–55. <https://doi.org/10.1016/j.fct.2013.06.019> PMID: 23793040
18. Pettersen K, Monsen VT, Hakvåg Pettersen CH, Overland HB, Pettersen G, Samdal H, et al. DHA-induced stress response in human colon cancer cells—Focus on oxidative stress and autophagy. *Free Radic Biol Med.* 2016; 90:158–72. <https://doi.org/10.1016/j.freeradbiomed.2015.11.018> PMID: 26585906
19. Myers JN, Schaffer MW, Korolkova OY, Williams AD, Gangula PR, M'Koma AE. Implications of the colonic deposition of free hemoglobin-alpha chain: a previously unknown tissue by-product in inflammatory bowel disease. 2014; 20(9), 1530–47.
20. Xi M, Sun Z, Sun H, Jiang J, Wang Y, Zhang M. 3-arylmethylene-2,3,6,7-tetrahydro-1H-pyrazino[2,1-a]isoquinolin-4(11bH)-ones as potent NRF2/ARE inducers in human cancer cells and AOM-DSS treated mice. *J. Med. Chem.* 2013; 56(20), 7925–38 <https://doi.org/10.1021/jm400944k> PMID: 24053646
21. Park EJ, Lim JH, Nam SI, Park JW, Kwon TK. Rottlerin induces heme oxygenase-1 (HO-1) up-regulation through reactive oxygen species (ROS) dependent and PKC δ -independent pathway in human colon cancer HT29 cells. *Biochimie.* 2010; 92(1), 110–15. <https://doi.org/10.1016/j.biochi.2009.10.001> PMID: 19833168
22. Jang HJ, Hong EM, Kim M, Kim JH, Jang J, Park SW, et al. Simvastatin induces heme oxygenase-1 via NF-E2-related factor 2 (NRF2) activation through ERK and PI3K/Akt pathway in colon cancer. *Oncotarget.* 2016; 7(29).
23. Trivedi PP, Jena GB, Tikoo KB, Kumar V. Melatonin Modulated Autophagy and NRF2 Signaling Pathways in Mice With Colitis-Associated Colon Carcinogenesis. *Mol Carcinog.* 2016 Mar; 55(3):255–67. <https://doi.org/10.1002/mc.22274> PMID: 25598500

24. Xu DG, Lv W, Dai CY, Zhu FF, Xu GH, Ma ZJ, et al. 2-(Pro-1-ynyl)-5-(5,6-dihydroxypenta-1,3-diyndyl) thiophene induces apoptosis through reactive oxygen species-mediated JNK activation in human colon cancer SW620 cells. *Anat Rec*. 2015; 298(2):376–85.
25. Tan BL, Norhaizan ME, Huynh K, Yeap SK, Hazilawati H, Roselina K. Brewers' rice modulates oxidative stress in azoxymethane-mediated colon carcinogenesis in rats. 2015; 21(29):8826–35.
26. Volonte D, Liu Z, Musille PM, Stoppani E, Wakabayashi N. Inhibition of nuclear factor-erythroid 2-related factor (NRF2) by caveolin-1 promotes stress-induced premature senescence. *Mol Bio Cell*. 2013; 24(12), 1852–62.
27. Kruse ML, Friedrich M, Arlt A, Röcken C, Egberts JH, Sebens, et al. Colonic Lamina Propria Inflammatory Cells from Patients with IBD Induce the Nuclear Factor-E2 Related Factor-2 Thereby Leading to Greater Proteasome Activity and Apoptosis Protection in Human Colonocytes. *Inflamm Bowel Dis*. 2016; 22(11), 2593–2606. <https://doi.org/10.1097/MIB.0000000000000925> PMID: 27661668
28. Yang Y, Cai X, Yang J, Sun X, Hu C, Yan Z, et al. Chemoprevention of dietary digitoflavone on colitis-associated colon tumorigenesis through inducing NRF2 signaling pathway and inhibition of inflammation. *Mol Cancer*. 2014; 13:48. <https://doi.org/10.1186/1476-4598-13-48> PMID: 24602443
29. Pandurangan AK, Kumar S, Sadagopan A, Dharmalingam P, Ganapasam S. Luteolin, a bioflavonoid inhibits Azoxymethane-induced colorectal cancer through activation of NRF2 signaling. *Toxicol. Mech. Meth*. 2014; 6516(1):13–20.
30. Wondrak GT, Villeneuve NF, Lamore SD, Bause AS, Jiang T, Zhang DD. The cinnamon-derived dietary factor cinnamic aldehyde activates the NRF2-dependent antioxidant response in human epithelial colon cells. *Molecules*. 2010; 15(5), 3338–3355. <https://doi.org/10.3390/molecules15053338> PMID: 20657484
31. Li CQ, Kim MY, Godoy LC, Thiantanawat A, Trudel LJ, Wogan GN. Nitric oxide activation of Keap1/NRF2 signaling in human colon carcinoma cells. *Proc Natl Acad Sci U S A*. 2009; 106(34):14547–51. <https://doi.org/10.1073/pnas.0907539106> PMID: 19706542
32. Jung K-A, Kwak M-K. Enhanced 4-Hydroxynonenal Resistance in KEAP1 Silenced Human Colon Cancer Cells. *Oxid Med Cell Longev*. 2013; 2013:1–16.
33. Lu M-C, Ji J-A, Jiang Y-L, Chen Z-Y, Yuan Z-W, You Q-D, et al. An inhibitor of the Keap1-NRF2 protein-protein interaction protects NCM460 colonic cells and alleviates experimental colitis. *Sci Rep. Nature Publishing Group*; 2016; 6:26585.
34. Stachel I, Geismann C, Aden K, Deisinger F, Rosenstiel P, Schreiber S, et al. Modulation of nuclear factor E2-related factor-2 (NRF2) activation by the stress response gene immediate early response-3 (IER3) in colonic epithelial cells: A novel mechanism of cellular adaption to inflammatory stress. *J Biol Chem*. 2014; 289(4):1917–29. <https://doi.org/10.1074/jbc.M113.490920> PMID: 24311782
35. Sebens S, Bauer I, Geismann C, Grage-Griebenow E, Ehlers S, Kruse ML, et al. Inflammatory macrophages induce NRF2 transcription factor-dependent proteasome activity in colonic NCM460 cells and thereby confer anti-apoptotic protection. *J Biol Chem*. 2011; 286(47):40911–21. <https://doi.org/10.1074/jbc.M111.274902> PMID: 21990354
36. Yokoo Y, Kijima A, Ishii Y, Takasu S, Tsuchiya T, Umemura T. Effects of NRF2 silencing on oxidative stress-associated intestinal carcinogenesis in mice. *Cancer Med*. 2016;
37. Kang KA, Piao MJ, Kim KC, Kang HK, Chang WY, Park IC, et al. Epigenetic modification of NRF2 in 5-fluorouracil-resistant colon cancer cells: involvement of TET-dependent DNA demethylation. *Cell Death Dis*. 2014; 5:e1183. <https://doi.org/10.1038/cddis.2014.149> PMID: 24743738
38. Zhao XQ, Zhang YF, Xia YF, Zhou ZM, Cao YQ. Promoter demethylation of nuclear factor-erythroid 2-related factor 2 gene in drug-resistant colon cancer cells. *Oncol Lett*. 2015; 10(3), 1287–92. <https://doi.org/10.3892/ol.2015.3468> PMID: 26622665
39. Hu T, Yao Y, Yu S, Guo H, Han L, Wang W, et al. Clinicopathologic significance of CXCR4 and NRF2 in colorectal cancer. *Int J of Biomed Res*. 2013; 27(4):283–90.
40. Ji L, Wei Y, Jiang T, Wang S. Correlation of NRF2, NQO1, MRP1, cmyc and p53 in colorectal cancer and their relationships to clinicopathologic features and survival. *Int J Clin Exp Pathol*. 2014; 7(3):1124–31. PMID: 24695690
41. Kang KA, Piao MJ, Ryu YS, Kang HK, Young W. Interaction of DNA demethylase and histone methyltransferase upregulates NRF2 in 5-fluorouracil-resistant colon cancer cells. *Oncotarget*. 2016
42. Ishaq M., Evans M. D., & Ostrikov K. K.. Atmospheric pressure gas plasma-induced colorectal cancer cell death is mediated by Nox2-ASK1 apoptosis pathways and oxidative stress is mitigated by Srx-NRF2 anti-oxidant system. *Biochim Biophys Acta*. 2014; 1843(12), 2827–2837. <https://doi.org/10.1016/j.bbamcr.2014.08.011> PMID: 25173814
43. Arlt A, Bauer I, Schafmayer C, Tepel J, Mu SS, Brosch M, et al. Increased proteasome subunit protein expression and proteasome activity in colon cancer relate to an enhanced activation of nuclear factor

- E2-related factor 2 (NRF2). *Oncogene*. 2009; 28(45), 3983–96. <https://doi.org/10.1038/onc.2009.264> PMID: 19734940
44. Li W, Khor TO, Xu C, Shen G, Jeong WS, Yu S, et al. Activation of Nrf2-antioxidant signaling attenuates NFκB-inflammatory response and elicits apoptosis. *Biochem. Pharmacol.* 2008; 76(11), 1485–89. <https://doi.org/10.1016/j.bcp.2008.07.017> PMID: 18694732
 45. Khor TO, Huang MT, Kwon KH, Chan JY, Reddy BS, Kong AN. NRF2-deficient mice have an increased susceptibility to dextran sulfate sodium-induced colitis. *Cancer Res.* 2006; 66(24):11580–4. <https://doi.org/10.1158/0008-5472.CAN-06-3562> PMID: 17178849
 46. Osburn WO, Karim B, Dolan PM, Liu G, Yamamoto M, Huso DL, et al. Increased colonic inflammatory injury and formation of aberrant crypt foci in NRF2-deficient mice upon dextran sulfate treatment. *Int J Cancer.* 2007; 121(9):1883–91. <https://doi.org/10.1002/ijc.22943> PMID: 17631644
 47. Saw CLL, Kong A-NT. Nuclear factor-erythroid 2-related factor 2 as a chemopreventive target in colorectal cancer. *Expert Opin Ther Targets.* 2011; 15(3), 281–95. <https://doi.org/10.1517/14728222.2011.553602> PMID: 21261563
 48. Cheung KL, Lee JH, Khor TO, Wu TY, Li GX, Chan J., et al. NRF2 knockout enhances intestinal tumorigenesis in *Apcmin*+/+ mice due to attenuation of anti-oxidative stress pathway while potentiates inflammation. *MOL CARCINOGEN.* 2014; 53(1), 77–84.
 49. Chang LC, Fan CW, Tseng WK, Chen JR, Chein HP, Hwang CC, et al. Immunohistochemical study of the NRF2 pathway in colorectal cancer: NRF2 expression is closely correlated to Keap1 in the tumor and Bach1 in the normal tissue. *Appl Immunohistochem Mol Morphol.* 2013. 21(6), 511–517. <https://doi.org/10.1097/PAI.0b013e318282ac20> PMID: 23455180
 50. Pandurangan AK, Esa NM. Dietary non-nutritive factors in targeting of regulatory molecules in colorectal cancer: An update. *Asian Pacific J Cancer Prev.* 2013; 14(10):5543–52.
 51. Yao J, Zhao L, Zhao Q, Zhao Y, Sun Y, Zhang Y, et al. NF-κB and NRF2 signaling pathways contribute to wogonin-mediated inhibition of inflammation-associated colorectal carcinogenesis. *Cell Death Dis.* 2014; 5:1283.
 52. Menegon S, Columbano A, Giordano S. The Dual Roles of NRF2 in Cancer. *Trends Mol Med.* 2016; 22: 578–93. <https://doi.org/10.1016/j.molmed.2016.05.002> PMID: 27263465
 53. Sporn MB, Liby KT. NRF2 and cancer: the good, the bad and the importance of context. *Nat Rev Cancer.* 2012; 12: 564–71. <https://doi.org/10.1038/nrc3278> PMID: 22810811
 54. D'Errico I, Salvatore L, Murzilli S, Sasso GLO, Latorre D, Martelli N, et al. Peroxisome proliferator-activated receptor-γ coactivator 1-α (PGC1α) is a metabolic regulator of intestinal epithelial cell fate. *Proc. Natl. Acad. Sci. USA.* 2011; 108(16), 6603–08. <https://doi.org/10.1073/pnas.1016354108> PMID: 21467224

Anexo II. Protocolo para la elaboración de revisión sistemática.

PROTOCOLO PARA REALIZAR REVISIONES SISTEMÁTICAS CUALITATIVAS

Introducción:

La introducción debe contener los siguientes puntos:

- Teorías, conceptos e ideas clave = CCR, desarrollo, tipos, etc.
- Fuentes clave = quién descubrió el CCR (opcional)
- Cuestiones y problemas = biogénesis mitocondrial
- Origen y definición del tema = BM y gen
- Controversias sobre el tema= BM y gen y CCR. Idea básica ¿efecto Warburg?
- Importancia de la revisión
- Objetivos = "Evaluar los efectos de GEN en la BM para el CCR" (Sub-objetivos opcional)

Los objetivos se deberían establecer previos a la redacción de la discusión y no se deben variar, pero no deben convertirse en un corsé que impida la exploración profunda de los temas de interés. En caso de cambiarlos, se debería explicar por qué, cuál es la motivación del replanteamiento y si son apropiadas las estrategias de búsqueda para el nuevo objetivo.

Métodos:

Criterios de elegibilidad

Tipos de participantes:

- Presencia/Ausencia de la enfermedad = CCR
- Características importantes de la población = Posible desarrollo de otras enfermedades previas.
- Factores demográficos relevantes (Edad, sexo, raza, nivel socioeconómico). =
Cuando no hay seguridad de que haya diferencias importantes de los efectos entre varios subgrupos de personas, puede ser mejor incluir todo lo relacionado con los subgrupos y luego buscar en los análisis si hay diferencias importantes y verosímiles del efecto.
- Ámbito de interés = ¿Trabajo con células colorrectales?

Tipos de intervenciones:

- Tipo de intervención experimental = cultivos, animales, células humanas, etc.
- Variaciones en la intervención = Forma de detección niveles gen, análisis de expresión o de otros tipos, duración del estudio, análisis de polimorfismos, etc.
- ¿Hay alguna técnica inapropiada para el estudio del gen?
- ¿Cómo se manejan los ensayos que incluyen solo parte de los datos de interés?
¿y los que incluyen los datos de interés combinados con otros factores? = Por

ejemplo, estudios que analizan la BM y el CCR sin GEN. O que analicen GEN y diabetes y CCR.

Tipos de desenlace:

Enumerar los desenlaces más relevantes. Se deben incluir todos para destacar posibles lagunas de conocimiento en la investigación primaria a falta de resultados triviales. Priorizar los desenlaces: principales o primarios y secundarios.

Búsqueda de estudios

Búsqueda amplia, objetiva y reproducible de una alta gama de fuentes.

Consultar diferentes bases de datos online.

Se recomienda que para todas las revisiones se consulten CENTRAL y MEDLINE, como mínimo, junto con EMBASE si está disponible.

Diseñar estrategias de búsqueda

Para realizar una búsqueda en las diferentes bases de datos hay que tener en cuenta los sinónimos de nuestras palabras clave, así como términos relacionados y variaciones ortográficas. Además, es importante usar operadores booleanos que permitan relacionar los términos de búsqueda.

El manejo de las referencias y documentación del proceso de búsqueda.

Un proceso típico para seleccionar los estudios para su inclusión en una revisión es el siguiente (el proceso se debe detallar en el protocolo de la revisión).

1. Integrar los resultados de la búsqueda mediante programas informáticos de gestión de referencias bibliográficas, y eliminar los registros duplicados de un mismo informe.
2. **Examinar los títulos y resúmenes** para eliminar los informes claramente irrelevantes (en este estadio los revisores deberían, en general, incluir más que excluir).
3. Recuperar el texto completo de los informes potencialmente relevantes.
4. Vincular y reunir los informes múltiples de un mismo estudio.
5. **Examinar el texto completo de los informes** para verificar el grado de cumplimiento que tienen los estudios de los criterios de elegibilidad.
6. Tomar una decisión definitiva sobre la inclusión del estudio y proceder a la obtención de los datos.

La lista de estudios excluidos debe ser tan breve como sea posible, dando un motivo breve y conciso sobre la exclusión.

Síntesis preliminar

Descripción textual

Se usa para resumir los papers y empezar a extraer la información de una manera sistemática.

Tabulación

Permite comparar los artículos seleccionados para la revisión sistemática. Se ha de tener en cuenta el autor y el año de publicación, el tipo de intervención, los resultados principales y la conclusión final.

Explorar relaciones entre papers.

Explorar la heterogeneidad

Se seleccionan características principales que se valorarán identificándolas en cada uno de los artículos:

Cadena de ideas y mapa conceptual

Este tipo de gráfico sirve para visualizar de una manera rápida el contenido general de todos los artículos contenidos en la revisión, es decir, se pueden observar las relaciones estudiadas.

Reflexión crítica y conclusión

En este apartado se explicará en más detalle los mapas conceptuales agrupando y teniendo en cuenta los distintos resultados aportados por los artículos que van a conformar la revisión sistemática. Finalmente, de acuerdo a lo descrito, se redactará una conclusión final.

Anexo III. Tabulación de artículos excluidos en la revisión sistemática

Tabla S 1. Artículos excluidos de la revisión sistemática.

| Título | Motivo de exclusión |
|---|----------------------------|
| Anti-inflammatory/Anti-oxidative stress activities and differential regulation of Nrf2-mediated genes by non-polar fractions of tea <i>Chrysanthemum zawadskii</i> and licorice <i>Glycyrrhiza uralensis</i> . | No discute sobre CCR |
| Coffee constituents as modulators of Nrf2 nuclear translocation and ARE (EpRE)-dependent gene expression. | No discute sobre CCR |
| Dietary squalene supplementation improves DSS-induced acute colitis by downregulating p38 MAPK and NFκB signaling pathways. | No relaciona NRF2 y CCR |
| Induction of Nrf2-mediated cellular defenses and alteration of phase I activities as mechanisms of chemoprotective effects of coffee in the liver. | No discute sobre CCR |
| Mechanistic insight into beta-carotene-mediated protection against ulcerative colitis-associated local and systemic damage in mice. | No analiza NRF2 |
| Methylation of the KEAP1 gene promoter region in human colorectal cancer | No analiza NRF2 |
| Nadroparin sodium activates Nrf2/HO-1 pathway in acetic acid-induced colitis in rats. | No relaciona NRF2 y CCR |
| NRF-1 is the major transcription factor regulating the expression of the human TOMM34 gene | No analiza NRF2 |
| Oxidative stress in health and disease: the therapeutic potential of Nrf2 activation. | No discute sobre CCR |
| Proteasome inhibitors MG-132 and bortezomib induce AKR1C1, AKR1C3, AKR1B1, and AKR1B10 in human colon cancer cell lines SW-480 and HT-29. | No relaciona NRF2 y CCR |
| Red meat and colorectal cancer: Nrf2-dependent antioxidant response contributes to the resistance of preneoplastic colon cells to fecal water of hemoglobin- and beef-fed rats. | No relaciona NRF2 y CCR |
| Regulation of human carbonyl reductase 3 (CBR3; SDR21C2) expression by Nrf2 in cultured cancer cells. | No relaciona NRF2 y CCR |
| Somatic mutations of the KEAP1 gene in common solid cancers. | No relaciona NRF2 y CCR |
| Synergy between sulforaphane and selenium in protection against oxidative damage in colonic CCD841 cells. | No analiza NRF2 |
| The proteasome inhibitor lactacystin enhances GSH synthesis capacity by increased expression of antioxidant components in an Nrf2-independent, but p38 MAPK-dependent manner in rat colorectal carcinoma cells. | No relaciona NRF2 y CCR |
| The secretome of colon cancer stem cells contains drug-metabolizing enzymes. | No analiza NRF2 |
| The sensitivity of cancer cells to pheophorbide a-based photodynamic therapy is enhanced by Nrf2 silencing. | No relaciona NRF2 y CCR |
| The use of nitric oxide-donating nonsteroidal anti-inflammatory drugs in the chemoprevention of colorectal neoplasia. | No relaciona NRF2 y CCR |
| Tomato powder impedes the development of azoxymethane-induced colorectal cancer in rats through suppression of COX-2 expression via NF-κB and regulating Nrf2/HO-1 pathway. | No analiza NRF2 |
| GRP78 knockdown enhances apoptosis via the down-regulation of oxidative stress and Akt pathway after epirubicin treatment in colon cancer DLD-1 cells. | No relaciona NRF2 y CCR |

Anexo IV. Artículo 2: NRF2 y PoDA. (pendiente de publicación)

Analysis of NRF2-related predictive pathways associated with colorectal cancer

Carmen Gonzalez-Donquiles^{1,2,3}, Jessica Alonso-Molero^{3,4}, Laura Vilorio-Marqués³, Mariona Bustamante^{5,6,7}, Antonio J. Molina³, Nerea Fernández de Larrea^{1,8}, Javier Llorca^{1,4}, Pilar Amiano^{1,9}, M^a Eva Ardanaz Aicua^{1,10}, Guillermo Fernandez-Tardon^{1,2,11}, Benito Mirón-Pozo¹², Rosana Peiró^{1,13}, Juan Alguacil^{1,14}, Ana Espinosa^{1,5,6}, Tomás González Elosua¹⁵, Beatriz Pérez-Gómez^{1,8}, Camilo Palazuelos⁴, Luisa C. de la Peña-Negro^{16,17}, Tania Fernández-Villa³, Marina Pollán^{1,8}, Manolis Kogevinas^{1,5,6,18}, Vicente Martín^{1,3}.

¹ Centro de Investigación Biomédica en Red Epidemiología y Salud Pública (CIBERESP), Madrid, Spain.

² Universidad de Oviedo, Oviedo, Spain

³ Grupo de Investigación Interacción Gen-Ambiente y Salud. Instituto de Biomedicina (IBIOMED). Universidad de León. León, Spain.

⁴ Instituto de Investigación Marqués de Valdecilla (IDIVAL). Universidad de Cantabria, Santander, Spain.

⁵ ISGlobal, Barcelona, Spain.

⁶ Pompeu Fabra University (UPF), Barcelona, Spain.

⁷ Genomics and Disease Group, Bioinformatics and Genomics Program, Centre for Genomic Regulation (CRG), Barcelona, Spain.

⁸ Cancer Epidemiology Unit, National Centre for Epidemiology, Instituto de Salud Carlos III. Madrid, Spain.

⁹ Public Health Division of Guipuzkoa, BioDonostia Research Institute, San Sebastian, Spain.

¹⁰ Navarra Public Health Institute, Pamplona, Spain; IdisNA, Navarra Institute for Health Research, Pamplona, Spain.

¹¹ Instituto Universitario de Oncología del Principado de Asturias (IUOPA), Oviedo, Spain.

¹² Unidad de Gestión Clínica de Cirugía General y del Aparato Digestivo, Complejo Hospitalario Universitario de Granada, Granada, Spain.

¹³ Centre for Research in Public Health, Valencia, Spain.

¹⁴ Centro de Investigación en Recursos Naturales, Salud, y Medio Ambiente (RENSMA), Universidad de Huelva, Huelva, Spain.

¹⁵ Servicio de Cirugía General, Complejo Asistencial Universitario de León (CAULE), León, Spain.

¹⁶ Gastroenterology Service, Bellvitge University Hospital (HUB), Barcelona, Spain.

¹⁷ Cancer Prevention and Control Program, Catalan Institute of Oncology (ICO), ONCOBELL Program, Bellvitge Biomedical Research Institute (IDIBELL), Hospital de Llobregat, Barcelona, Spain.

¹⁸ IMIM (Hospital del Mar Medical Research Institute), Barcelona, Spain.

ABSTRACT

Colorectal cancer (CRC) is the third most commonly diagnosed malignancy and the fourth leading cause of cancer-related deaths in the world. To better understand pathogenesis of CRC, we performed the pathway of distinction analysis (PoDA) on a genome-wide association study dataset of 1,336 CRC cases and 2,744 controls matched by age and sex, who were recruited in different Spanish Hospitals. The aim of this study was to search pathways whose variations were related to the development of CRC. Selected pathways were the following: Longevity regulating pathway, Inflammatory Bowel Disease, MAPK signaling pathway, NF-Kappa-B signaling pathway, Mitochondrial Biogenesis pathways, AMPK signaling pathway, PI3K-AKT signaling pathway, Pentose phosphate pathway and Transporter ABC pathway. The results provide evidence by which genetic factor (beyond the effects of SNPs) could promote the genesis and development of CRC. It would be advisable to undertake this same type of analysis with different study populations in order to corroborate the present results and to establish if they can be extrapolated.

Keywords: PoDA, Pathway analysis, NRF2, colorectal cancer

INTRODUCTION

Colorectal cancer (CRC) has worldwide the third highest incidence in men, and the second highest in women [1]. This type of cancer evidences an increasing prevalence in developing and developed countries [2]. The gene-environmental interaction (GxE) is important to understand the etiology of this disease, therefore, exposition to a determined environmental factor increases the disease risk only of those patients with a specific genetic profile [3].

Over the past few years many genome-wide association studies (GWAS) have been performed using SNP analyses in order to detect the punctual mutations associated with an increased CRC risk [4–7]. Nonetheless, this methodology presents some limitations such as the enormous quantity of SNPs in need to be analyzed or else the possibility of working with an elevated alpha error. Furthermore, if the effect of a specific SNP on the disease is low, then it could not be detected. Yet another limitation could entail the lack of analytical reproducibility [8]. For the above cited reasons, pathway analysis has been proposed.

Pathways of Distinction Analysis (PoDA) is a current and useful methodology used in pathway analysis which supplements the shortcomings of other methodologies. The running hypothesis of this method is based on the fact that if SNPs of a pathway play roles in the occurrence of a disease, then there will be a greater similarity within-classes than across-classes in the genotypes of those SNPs [9]. If the selected pathways are good CRC predictors, it will allow for the formation of two well-differentiated groups (cases and controls) based on the genotypes of the studied SNPs. On the other hand, should the opposite occur, then the two before mentioned groups cannot be clearly established and, therefore, these pathways will not comprise good CRC predictors [9].

In this study, the Nuclear Respiratory Factor 2 (NRF2) locus was selected as the reference gene owing to studies that support the idea that the disrupted activity or expression of this gene could be related to CRC. Consequently for CRC risk to decrease, it is necessary that the associated pathways including a role of this gene should not manifest any disruptions [10].

Once the NRF2-CRC relationship has been established, the main objective of this study has been to test if the selected NRF2-related pathways could embody good CRC predictors.

MATERIAL AND METHODS

1. Subject study

The selection of subjects included in this study was made using the PLINK software. Subjects belonged to the MCC-Spain database which is a multi-center case-control study with population controls aiming to evaluate in Spain the influence of environmental exposures and their interactions with genetic factors regarding common tumors (prostate, breast, colorectal, gastroesophageal and chronic lymphocytic leukemia). Between September of 2008 and December of 2013, subjects aged between 20-85 years with histologically-confirmed newly-diagnosed colon or rectum cancers (ICD-10: C18, C19, C20, D01.0, D01.1, D01.2) were recruited throughout 18 Spanish hospitals located in 11 Spanish provinces. Simultaneously, population-based controls with a frequency-matched to cases, by age, sex and region, with the joint distribution of the tumors included in the MCC-Spain were randomly selected from primary care centers within of the hospitals' catchment areas. All participants signed an informed consent. Approval for the study was obtained from the ethical review boards of all of the recruiting centers (Ethical Committee of Clinical Research of Barcelona, Cantabria, Girona, Gipuzkoa, Huelva, León, Principado de Asturias, Madrid, Navarra and Valencia). Additional information regarding the study design is provided elsewhere [11].

The study was carried out according to Spanish laws on biomedical research. All procedures were performed with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments.

2. Sample processing

Peripheral blood was drawn from participants, which was aliquoted in whole blood, plasma, the cellular fraction for DNA extraction and serum, all stored at -80 °C. Saliva was collected from subjects refusing to donate blood with the Oragene® DNA Kit and stored at room temperature until DNA extraction. We collected biological samples for

DNA extraction from the interviewed participants, 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 °C. Fresh tumor biopsies or paraffin embedded samples were available in all participating hospitals. Standardized basic clinical and pathological information on the diagnosis and treatment of tumors was collected from the hospital records by using a predefined format.

Genetic analyses were carried out within of the MCC-Spain and also through the participation of international consortia. The InfiniumHumanExomeBeadChip of Illumina was used to genotype >200,000 coding markers plus 6,000 additional custom variants related to the pathways of interest [11].

3. Pathways of Distinction Analysis (PoDA)

The analyzed number of cases and controls were respectively 1,336 and 2,744. The PoDA analysis was run systematically in the pathways subsequently mentioned in section 5. This method compares the difference between cases and controls based on a genetic distance due to the observed variations within and between groups. A distance score is computed with respect to each pathway and each sample measuring the distance to the remaining case samples and control samples. The distinction score quantifies the score distribution difference between cases and controls for each pathway. Significance was assessed by resampling dummy pathways with a same length followed by computing the fraction of distinction scores greater than those observed [9]. A summary of the steps with the equations and the objective in each of them is shown in table 1.

Table 1. Summary of the PoDA method's mathematical model

| Steps | Equation | Aim |
|---|--|--|
| χ^2 | $\chi^2 = \frac{(f_o - f_e)^2}{f_e}$ <ul style="list-style-type: none"> f_o and f_e: observed and expected frequencies | To pick the most representative SNP belonging to each gene |
| relative distance (D) | $D_{Y,i} = y_i - f_i - y_i - g_i $ <ul style="list-style-type: none"> f_i and g_i: MAFs of the SNP i in the samples F (controls) and G (cases) y_i: [0,1,2] is the Y genotype | Knowing the relative distance of a SNP to the cases and the controls |
| S score | $S_Y = \frac{Mean(D_{Y,i})}{\sqrt{Var(D_{Y,i})/l}}$ <ul style="list-style-type: none"> l: all loci considered | Quantifying the case and control closeness distributions |
| Wilcoxon rank sum statistic | $W = \sum_{Y \in case} R_{Y,p} - \frac{n_{case}(n_{case} + 1)}{2}$ <ul style="list-style-type: none"> $R_{Y,p}$: rank of $S_{Y,p}$ amongst all samples Y | Quantifying differences in the S distribution for cases and controls |
| Case and control labels permutations and repetitions of all steps | | |

| | | |
|-------------------------------|---|--|
| Standardized pathway DS score | $DS_p = \frac{W_p - E(W_p^*)}{SD(W_p^*)}$ <ul style="list-style-type: none"> • W_p^*: set of W_p obtained for pathway P across of the permutations | To evaluate the true significance of W |
| P-value** | $P_{value} = \frac{\# DS\ score_{theoretical} > DS\ score_{observed}}{n\ permutations}$ | To determine statistical significance of a pathway |

**Note that it is convenient to adjust the P-values using the False Discovery Rate (FDR) in order to identify false-positive results.

4. Splot.

The PoDA method contains a script which allows to easily observe the S-value distribution for cases and controls by means of a dotplot graphic and a density diagram. In the instance that both case and control distributions are overlapped, the assessed pathway does not entail a good disease predictor. However, if case and control distributions are different, it follows that the pathway will assist to predict CRC.

5. Pathway list

After reviewing the bibliography and selecting the pathways to assess, all of the corresponding pathway genes were selected. To carry this out, the KEGG and Reactome databases were consulted. Pathways regulated by NRF2, as well as pathways which regulate NRF2 together with pathways in which NRF2 participates were taken into account. Specifically, the following pathways were assessed: Inflammatory Bowel Disease (IBD) [12], Longevity regulating pathway [13], MAPK signaling pathway [14], NF-kappa B signaling pathway [15], PI3K-Akt signaling pathway [16], Pentose phosphate pathway (PPP) [17], ABC transporters [18], Mitochondrial Biogenesis pathway [19] and AMPK signaling pathway [20] (Table 1). The inflammatory bowel diseases pathway could be proposed as a positive control because its association with CRC is clear and verified by scientific evidences [21–24]. On the other hand, the longevity regulating pathway could be suggested as a negative control due to the lack of studies that relate CRC with this pathway.

The IDs of the assessed pathways are the following:

[Table 2. Assessed pathway IDs.](#)

| PATHWAY | ID |
|---------|----|
|---------|----|

| | |
|---|---------------|
| Inflammatory Bowel Disease (IBD) | hsa05321 |
| Longevity regulating pathway | hsa04213 |
| MAPK signaling pathway | hsa04010 |
| NF-kappa B signaling pathway | hsa04064 |
| PI3K-Akt signaling pathway | hsa04151 |
| Pentose phosphate pathway | hsa00030 |
| ABC transporter | hsa02010 |
| Mitochondrial Biogenesis pathway | R-HSA-1592230 |
| AMPK signaling pathway | hsa04152 |

Next, all pathway genes were selected amounting to a total of 1,089 genes analyzed. The next step consisted in selecting all of the SNPs. At the same time, we took into account a GWAS which, after a quality control, contained a total of 6,806 genes.

For each pathway, once all the genes had been selected, all of the SNPs contained in the MCC-Spain database were selected [25].

RESULTS

1. Study subjects

Our study included a total of 6,090 individuals derived from the MCC-Spain database consisting of 2,140 CRC cases and 3,950 controls. Out of this total number, only subjects with genotyped SNPs were chosen. After undertaking a descriptive analysis, 1,336 cases and 2,744 controls were counted. Both groups comprised more men than women. Individuals within of the case group were on average older than those belonging to the control group and this last group had a higher educational level. The study encompassed cases and controls residing in 10 Spanish provinces, although Barcelona and León were the two cities that provided most cases to the study (Table 3).

Table 3. Subject characteristics

| | | Cases | Controls |
|------------------|--|---------------------|---------------------|
| | | N/% | N/% |
| | | 2,140 (35.1) | 3,950 (64.9) |
| Genotyped | | 1,336 (32.8) | 2,744 (67.2) |

| | | | |
|--------------------------|---------------------------------|-------------|--------------|
| | | | |
| Sex | Men | 865 (64.8) | 1,469 (53.5) |
| | Women | 471 (35.2) | 1,275 (46.5) |
| | | | |
| Age (Mean/SD) | | 67.0 ± 11.2 | 63.3 ± 11.5 |
| | | | |
| Educational level | Less than Primary School | 380 (28.4) | 488 (17.8) |
| | Primary School | 552 (41.3) | 913 (33.3) |
| | Secondary School | 261 (19.5) | 787 (28.7) |
| | University Studies | 143 (10.7) | 556 (20.2) |
| | | | |
| Region | Asturias | 60 (4.5) | 167 (6.1) |
| | Barcelona | 371 (27.8) | 560 (20.4) |
| | Cantabria | 143 (10.7) | 323 (11.8) |
| | Granada | 52 (3.9) | 156 (5.7) |
| | Guipúzcoa | 88 (6.6) | 314 (11.4) |
| | Huelva | 33 (2.5) | 42 (1.5) |
| | León | 280 (21.0) | 284 (10.4) |
| | Madrid | 185 (13.9) | 650 (23.7) |
| | Navarra | 79 (5.9) | 157 (5.7) |
| | Valencia | 45 (3.4) | 91 (3.3) |

2. PoDA

The IBD pathway, MAP Kinase pathway, NFK β signaling pathway, PI3K-AKT signaling pathway, ABC transporter pathway and the mitochondrial biogenesis pathway exemplify association with CRC. The remaining three pathways assessed did not evidence any significant results. The respective p-values corresponding to the tested pathways are displayed in Table 4.

Table 4. P-values and adjusted p-values of the assessed pathways in relation to CRC.

| Pathway | Source | Length | DSp | p(DSp) | p-Adj. (FDR) |
|----------------------------------|---------------|---------------|------------|---------------|---------------------|
| Inflammatory Bowel Disease (IBD) | KEGG | 65 | 12.68 | <0.001 | <0.001 |
| Longevity regulating pathway | KEGG | 94 | 1.33 | 0.770 | 0.870 |
| MAPK signaling pathway | KEGG | 255 | 16.64 | <0.001 | <0.001 |
| NFK β pathway | KEGG | 95 | 7.76 | <0.001 | <0.001 |
| PI3K-AKT pathway | KEGG | 341 | 8.22 | <0.001 | <0.001 |
| Pentose Phosphate Pathway | KEGG | 30 | 1.21 | 0.270 | 0.340 |
| ABC Transporter pathway | KEGG | 45 | 10.58 | <0.001 | <0.001 |
| Mitochondrial Biogenesis | Reactome | 50 | 3.69 | <0.001 | <0.001 |

| pathway | | | | | |
|------------------------|------|-----|------|-------|-------|
| AMPK signaling pathway | KEGG | 114 | 0.71 | 0.870 | 0.870 |

3. Splot

The scatter plots of the distance scores plus overlaid boxplots regarding the representative pathways are provided in Figures 1-6 of the Supplementary Material. Higher S values indicate that the sample was at a closer distance to other cases than it was to other control samples. Although distribution in cases and controls partially overlapped in all the pathways, higher differences between the groups were observed in MAPK signaling pathway and PI3k-AKT signaling pathway.

DISCUSSION

Out of the nine NRF2-related pathways assessed, six were associated with CRC with adjusted p-values lower than 0.001 for the Inflammatory Bowel Disease pathway, MAPK signaling pathway, NF-Kappa B signaling pathway, PI3K-Akt signaling pathway, ABC transporter and the Mitochondrial Biogenesis pathway, suggesting an interesting predictive ability for these pathways.

According to Azer *et al.* (2013), the Inflammatory Bowel Disease pathway is directly associated with CRC genesis [24]. Furthermore, the study of Bondurant *et al.* (2013) assessed the association between SNPs related to interleukins (associated to inflammatory processes) and colorectal cancer, showing association. In accordance with Ran Kim *et al.* (2014), IBD-associated CRC involved other pathways including the following: immune response by mucosal inflammatory mediators, oxidative stress, and intestinal microbiota[22]. All this events were corroborated in the review of Sebastian *et al.* (2014), where analyzed several epidemiological studies to conclude that there is a close relationship between CRC and IBDs [23]. All this data, suggest that this pathway could be consider as positive control for our analysis, and the obtained results (p-value<0.001) are in concordance with this proposal.

On the other hand, although studies exist also showing a relationship between an adequate NRF2 activity/expression and longevity [26] no link has been found between this pathway and colorectal cancer, so we propose to consider this pathway as a

negative which agree with the statistically non-significant association of this pathway with CRC in our analysis (p-value=0.870). The concordance of these results with previous evidence, provide us confidence in using this methodology to detect pathways that are associated with a particular disease.

In relation to the other assessed pathways, it is important to note that the results found have biological plausibility and are consistent with the evidence. In this way, NRF2 is known as the main cellular defense mechanism against oxidative or electrophilic stress, so it is clear that this gene participates in drug transport and metabolization. [27]. This mechanism could be performed by the ABC transporter pathway [28]. This path is highly associated to CRC so it could be used as a predictive tool. The main cause may be that accumulated drugs without a metabolization possibility could be converted into carcinogenic compounds and promote CRC[29]. This is supported by Hu *et al.* (2016). These authors found some overregulated ABC transporters that reduced the effect of antineoplastic drugs in CRC [30].

Another important event is the activation of NRF2, which can be produced by a number of mechanisms. One of these mechanisms is via of the specific-site phosphorylation by kinases. The phosphorylation could either be produced by the MAPK pathway, the AMPK pathway or else the PI3K pathway, inducing NRF2 accumulation in the nucleus and targeting gene transactivation controlling the antioxidant response elements (ARE). In this way, NRF2 provides protection and reduces CRC risk [31–34]. The MAPK pathway seems to be a good CRC predictor. Owing to the phosphorylation promoted by this pathway, some processes are activated that trigger the genesis and development of CRC [35]. This pathway is activated by a great number of different stimuli so it can easily be mutated in CRC. The same occurs with the NF κ B pathway (that modulates the activation and expression of NRF2 [36]) or with the PI3K pathway. Furthermore, both pathways are related and often disrupted in CRC [37]. All three pathways can activate NRF2 by phosphorylation, hence being essential for an adequate expression. However, the AMPK pathway which also activates NRF2, has been shown to be mainly disrupted in later stages, therefore not always embodying a good predictor at the disease genesis stage [38].

Furthermore, this gene is involved in glucose metabolism by means of the pentose phosphate pathway [39]. The Pentose phosphate pathway did not reveal a statistical

significant association with CRC prevention. This could be due to the fact that this pathway is an alternative carbohydrate metabolism pathway, so other pathways exist which could support the same function when the pentose phosphate pathway is disrupted. This is the reason why this pathway could not be used as a CRC predictor in our population. However, Kowalik *et al.* (2017) support the idea that an association exists between the pentose phosphate pathway in relation to tumorigenesis and cell proliferation due to an unregulated NRF2 activity as a consequence of G6PD overexpression at later disease stages [40].

Lastly, NRF2 is also responsible of the creation and formation of new mitochondria owing to its participation in the mitochondrial biogenesis pathway [41]. The last pathway assessed in this study was the Mitochondrial Biogenesis route responsible of creating new mitochondria with an adequate morphology and in an appropriate number [42]. Inside of mitochondria, a lot of metabolic processes occur, so it can be assumed that if mitochondria are modified, therefore also their processes. A systematic revision exists in which Gonzalez-Donquiles *et al.* (2017) support the relationship between CRC and mitochondrial biogenesis via NRF2 [10]. Furthermore, Cruz *et al.* (2017) assume that mitochondrial reprogramming frequently occurs at early stages of CRC, so it is clear that a relationship between both exists [43].

One of the major constraints of previous studies relating pathways to colorectal cancer is that they have been conducted either on an experimental basis or else through literature reviews, pathway analysis using bioinformatics techniques is a novel approach. Therefore, it would be advisable to undertake this same type of analysis with different study populations in order to corroborate the present results and to establish if they can be extrapolated.

CONCLUSIONS

CRC is included in a complex disease group. While some genes associated to this disease have been identified, pathway analysis is a more complex method which allows to make reliable and coherent biological interpretations. Genetic variants in the following NRF2 related pathways allowed the discrimination of CRC cases and controls: IBD pathway, NFK β pathway, MAPK signaling pathway, PI3K-AKT signaling pathway, ABC Transporter pathway and Mitochondrial Biogenesis pathway.

Additional information:**Competing Interests**

The authors declare no competing interests.

Ethics approval and consent to participate

The MCC-Spain protocol was approved by the Ethics committees of the participating institutions. All participants were informed about the objectives of the study and signed an informed consent. Confidentiality of the data was secured by removing any personal identifiers within of 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 regarding the organization of the project can be found online at <http://www.mccspain.org>.

Consent for publication

Not applicable.

Availability of data and material

Samples: Biological samples were stored at the biobanks supported by the 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 of Gipuzkoa, the Basque Biobank, the ICOBIOBANC (sponsored by the Catalan Institute of Oncology), the IUOPA Biobank of the University of Oviedo and the ISCIII Biobank.

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

Funding: The study was partially funded by the “Acción Transversal del Cancer”, approved by the Spanish Ministry Council on the 11th of 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 the 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 grant 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.

Authors' contributions

JL, PA, EAA, RP, JA, BPG, MP, MK and VM contributed substantially to the acquisition of data. AJM and VM designed the study. CGD performed the experiments. CGD, JAL, LVM, TFV and GFT analyzed the data. CGD and AJM wrote the manuscript. All authors (CGD, JAM, LVM, MB, AJM, NFL, JL, PA, EAA, GFT, BMP, RP, JA, AE, TGE, BPG, CP, LPN, TFV, MP, MK and VM) read and approved the final manuscript.

Acknowledgements

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

REFERENCES

1. IARC - INTERNATIONAL AGENCY FOR RESEARCH ON CANCER [Internet]. [cited 2017 Jun 12]. Available 2017 Jun 12, from <https://www.iarc.fr/>
2. Bishehsari F, Mahdavinia M, Vacca M, Malekzadeh R, Mariani-Costantini R. Epidemiological transition of colorectal cancer in developing countries: environmental factors, molecular pathways, and opportunities for prevention. *World J Gastroenterol* [Internet]. Baishideng Publishing Group Inc; 2014 [cited 2017 Jun 14]; 20: 6055–72. doi: 10.3748/wjg.v20.i20.6055.
3. Lin X, Lee S, Christiani DC, Lin X. Test for interactions between a genetic marker set and environment in generalized linear models. *Biostatistics* [Internet]. 2013 [cited 2017 Jul 27]; 14: 667–81. doi: 10.1093/biostatistics/kxt006.
4. Mates IN, Jinga V, Csiki IE, Mates D, Dinu D, Constantin A, Jinga M. Single nucleotide polymorphisms in colorectal cancer: associations with tumor site and TNM stage. *J Gastrointest Liver Dis* [Internet]. 2012 [cited 2017 Jun 14]; 21: 45–52. Available from <http://www.ncbi.nlm.nih.gov/pubmed/22457859>
5. Li F-X, Yang X-X, Hu N-Y, Du H-Y, Ma Q, Li M. Single-nucleotide polymorphism associations for colorectal cancer in southern chinese population. *Chin J Cancer Res* [Internet]. Beijing Institute for Cancer Research; 2012 [cited 2017 Jun 14]; 24: 29–35. doi: 10.1007/s11670-012-0029-7.
6. Tuupanen S, Turunen M, Lehtonen R, Hallikas O, Vanharanta S, Kivioja T, Björklund M, Wei G, Yan J, Niittymäki I, Mecklin J-P, Järvinen H, Ristimäki A, et al. The common colorectal cancer predisposition SNP rs6983267 at chromosome 8q24 confers potential to enhanced Wnt signaling. *Nat Genet* [Internet]. Nature Publishing Group; 2009 [cited 2017 Jun 14]; 41: 885–90. doi: 10.1038/ng.406.
7. Tanskanen T, van den Berg L, Välimäki N, Aavikko M, Ness-Jensen E, Hveem K, Wettergren Y, Bexé Lindskog E, Tönisson N, Metspalu A, Silander K, Orlando G, Law PJ, et al. Genome-wide association study and meta-analysis in Northern European populations replicate multiple colorectal cancer risk loci. *Int J cancer* [Internet]. 2018 [cited 2017 Dec 18]; 142: 540–6. doi: 10.1002/ijc.31076.
8. F M Van der Velden WJ, Feuth T, C Stevens WB, Donnelly JP, A Blijlevens

- NM. Issues in genetic association studies: limitations of statistical analysis and biological plausibility. *Bone Marrow Transplant* [Internet]. 2010 [cited 2017 Jun 14]; 46: 906–7. doi: 10.1038/bmt.2010.211.
9. Braun R, Buetow K, Narasimhan B, Chu G, Kavanagh J. Pathways of Distinction Analysis: A New Technique for Multi-SNP Analysis of GWAS Data. Schork NJ, editor. *PLoS Genet* [Internet]. Public Library of Science; 2011 [cited 2017 Jun 14]; 7: e1002101. doi: 10.1371/journal.pgen.1002101.
 10. Gonzalez-Donquiles C, Alonso-Molero J, Fernandez-Villa T, Vilorio-Marques L, Molina AJ, Martin V. The NRF2 transcription factor plays a dual role in colorectal cancer: A systematic review. Grivennikov S, editor. *PLoS One* [Internet]. Public Library of Science; 2017 [cited 2017 Jun 14]; 12: e0177549. doi: 10.1371/journal.pone.0177549.
 11. Castaño-Vinyals G, Aragoes N, Perez-Gomez B, Martin V, Llorca J, Moreno V, Alzibar JM, Ardanaz E, de Sanjose S, Jimenez-Moleon JJ, Tardon A, Alguacil J, Peiro R, et al. Population-based multicase-control study in common tumors in Spain (MCC-Spain): rationale and study design. *Gac Sanit* [Internet]. 2015 [cited 2017 Jun 14]; 29: 308–15. doi: 10.1016/j.gaceta.2014.12.003.
 12. Kenehisa Laboratories. KEGG PATHWAY: Inflammatory bowel disease (IBD) - Homo sapiens (human) [Internet]. [cited 2017 Jun 14]. Available 2017 Jun 14, from [http://www.kegg.jp/kegg-bin/highlight_pathway?scale=1.0&map=hsa05321&keyword=inflammatory bowel disease](http://www.kegg.jp/kegg-bin/highlight_pathway?scale=1.0&map=hsa05321&keyword=inflammatory%20bowel%20disease)
 13. KEGG PATHWAY: Longevity regulating pathway - multiple species - Homo sapiens (human) [Internet]. [cited 2017 Jun 14]. Available 2017 Jun 14, from http://www.kegg.jp/kegg-bin/highlight_pathway?scale=1.0&map=hsa04213&keyword=LONGEVITY
 14. KEGG PATHWAY: MAPK signaling pathway - Homo sapiens (human) [Internet]. [cited 2017 Jun 14]. Available 2017 Jun 14, from http://www.kegg.jp/kegg-bin/highlight_pathway?scale=1.0&map=hsa04010&keyword=MAPK

15. KEGG PATHWAY: NF-kappa B signaling pathway - Homo sapiens (human) [Internet]. [cited 2017 Jun 14]. Available 2017 Jun 14, from http://www.kegg.jp/kegg-bin/highlight_pathway?scale=1.0&map=hsa04064&keyword=nfkb
16. KEGG PATHWAY: PI3K-Akt signaling pathway - Homo sapiens (human) [Internet]. [cited 2017 Jun 14]. Available 2017 Jun 14, from http://www.kegg.jp/kegg-bin/highlight_pathway?scale=1.0&map=hsa04151&keyword=PI3K
17. KEGG PATHWAY: Pentose phosphate pathway - Homo sapiens (human) [Internet]. [cited 2017 Jun 14]. Available 2017 Jun 14, from http://www.kegg.jp/kegg-bin/highlight_pathway?scale=1.0&map=hsa00030&keyword=PENTOSESPHOSPHATE
18. KEGG PATHWAY: ABC transporters - Homo sapiens (human) [Internet]. [cited 2017 Jun 14]. Available 2017 Jun 14, from [http://www.kegg.jp/kegg-bin/highlight_pathway?scale=1.0&map=hsa02010&keyword=transporter drug](http://www.kegg.jp/kegg-bin/highlight_pathway?scale=1.0&map=hsa02010&keyword=transporter%20drug)
19. May B. Mitochondrial biogenesis [Internet]. Reactome - a curated knowledgebase of biological pathways. 2014 [cited 2017 May 31]. doi: 10.3180/REACT_200788.1.
20. KEGG PATHWAY: AMPK signaling pathway - Homo sapiens (human) [Internet]. [cited 2017 Jun 14]. Available 2017 Jun 14, from http://www.kegg.jp/kegg-bin/highlight_pathway?scale=1.0&map=hsa04152&keyword=ampk
21. Triantafyllidis JK, Nasioulas G, Kosmidis PA. Colorectal cancer and inflammatory bowel disease: epidemiology, risk factors, mechanisms of carcinogenesis and prevention strategies. *Anticancer Res* [Internet]. 2009 [cited 2018 Feb 16]; 29: 2727–37. Available from <http://www.ncbi.nlm.nih.gov/pubmed/19596953>
22. Kim ER, Chang DK. Colorectal cancer in inflammatory bowel disease: The risk, pathogenesis, prevention and diagnosis. *World J Gastroenterol* [Internet].

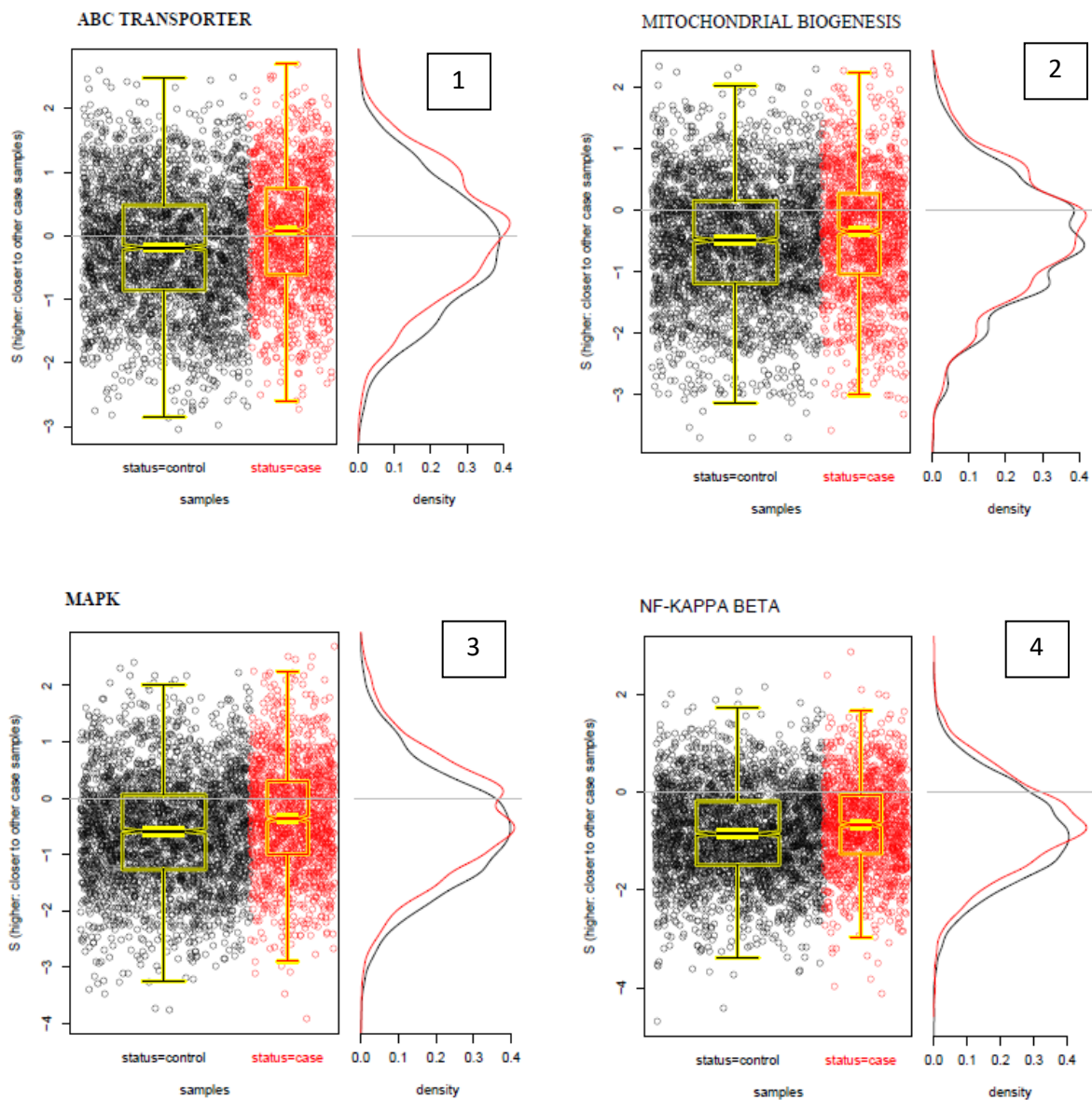
- Baishideng Publishing Group Inc.; 2014 [cited 2018 Feb 16]; 20: 9872. doi: 10.3748/wjg.v20.i29.9872.
23. Sebastian S, Hernández V, Myrelid P, Kariv R, Tsianos E, Toruner M, Marti-Gallostra M, Spinelli A, van der Meulen-de Jong AE, Yuksel ES, Gasche C, Ardizzone S, Danese S. Colorectal cancer in inflammatory bowel disease: Results of the 3rd ECCO pathogenesis scientific workshop (I). *J Crohn's Colitis* [Internet]. Oxford University Press; 2014 [cited 2018 Feb 16]; 8: 5–18. doi: 10.1016/j.crohns.2013.04.008.
 24. Azer SA. Overview of molecular pathways in inflammatory bowel disease associated with colorectal cancer development. *Eur J Gastroenterol Hepatol* [Internet]. 2013 [cited 2017 Jun 14]; 25: 271–81. doi: 10.1097/MEG.0b013e32835b5803.
 25. Castaño-Vinyals G, Aragonés N, Pérez-Gómez B, Martín V, Llorca J, Moreno V, Altzibar JM, Ardanaz E, de Sanjosé S, Jiménez-Moleón JJ, Tardón A, Alguacil J, Peiró R, et al. Population-based multicase-control study in common tumors in Spain (MCC-Spain): rationale and study design. *Gac Sanit* [Internet]. 2015 [cited 2018 Jan 18]; 29: 308–15. doi: 10.1016/j.gaceta.2014.12.003.
 26. Bruns DR, Drake JC, Biela LM, Peelor FF, Miller BF, Hamilton KL. Nrf2 Signaling and the Slowed Aging Phenotype: Evidence from Long-Lived Models. *Oxid Med Cell Longev* [Internet]. 2015 [cited 2017 Jun 14]; 2015: 1–15. doi: 10.1155/2015/732596.
 27. Wang X, Campos CR, Peart JC, Smith LK, Boni JL, Cannon RE, Miller DS. Nrf2 upregulates ATP binding cassette transporter expression and activity at the blood-brain and blood-spinal cord barriers. *J Neurosci* [Internet]. Society for Neuroscience; 2014 [cited 2017 Jun 14]; 34: 8585–93. doi: 10.1523/JNEUROSCI.2935-13.2014.
 28. Taylor NMI, Manolaridis I, Jackson SM, Kowal J, Stahlberg H, Locher KP. Structure of the human multidrug transporter ABCG2. *Nature* [Internet]. 2017 [cited 2017 Jun 14]; . doi: 10.1038/nature22345.
 29. Lee D, Lee GK, Yoon K-A, Lee JS. Pathway-Based Analysis Using Genome-

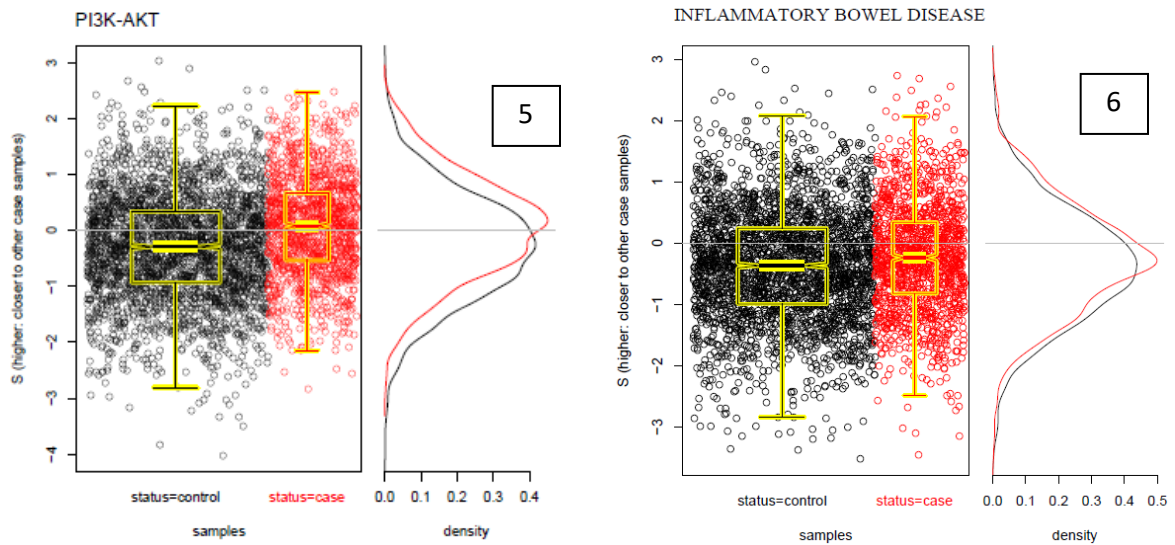
- wide Association Data from a Korean Non-Small Cell Lung Cancer Study. Cho WCS, editor. PLoS One [Internet]. Public Library of Science; 2013 [cited 2018 Feb 19]; 8: e65396. doi: 10.1371/journal.pone.0065396.
30. Hu T, Li Z, Gao C-Y, Cho CH. Mechanisms of drug resistance in colon cancer and its therapeutic strategies. *World J Gastroenterol* [Internet]. Baishideng Publishing Group Inc; 2016 [cited 2017 Sep 4]; 22: 6876–89. doi: 10.3748/wjg.v22.i30.6876.
 31. Mo C, Wang L, Zhang J, Numazawa S, Tang H, Tang X, Han X, Li J, Yang M, Wang Z, Wei D, Xiao H. The crosstalk between Nrf2 and AMPK signal pathways is important for the anti-inflammatory effect of berberine in LPS-stimulated macrophages and endotoxin-shocked mice. *Antioxid Redox Signal* [Internet]. Mary Ann Liebert, Inc.; 2014 [cited 2017 Jun 14]; 20: 574–88. doi: 10.1089/ars.2012.5116.
 32. Joo MS, Kim WD, Lee KY, Kim JH, Koo JH, Kim SG. AMPK Facilitates Nuclear Accumulation of Nrf2 by Phosphorylating at Serine 550. *Mol Cell Biol* [Internet]. 2016 [cited 2017 Jun 14]; 36: 1931–42. doi: 10.1128/MCB.00118-16.
 33. Sun Z, Huang Z, Zhang DD. Phosphorylation of Nrf2 at Multiple Sites by MAP Kinases Has a Limited Contribution in Modulating the Nrf2-Dependent Antioxidant Response. Morty RE, editor. *PLoS One* [Internet]. 2009 [cited 2017 Jun 14]; 4: e6588. doi: 10.1371/journal.pone.0006588.
 34. Nakaso K, Yano H, Fukuhara Y, Takeshima T, Wada-Isoe K, Nakashima K. PI3K is a key molecule in the Nrf2-mediated regulation of antioxidative proteins by hemin in human neuroblastoma cells. *FEBS Lett* [Internet]. 2003 [cited 2017 Jun 14]; 546: 181–4. Available from <http://www.ncbi.nlm.nih.gov/pubmed/12832036>
 35. Huang R, Quan Y, Chen J, Wang T, Xu M, Ye M, Yuan H, Zhang C, Liu X, Min Z. Osteopontin Promotes Cell Migration and Invasion, and Inhibits Apoptosis and Autophagy in Colorectal Cancer by activating the p38 MAPK Signaling Pathway. *Cell Physiol Biochem* [Internet]. 2017 [cited 2017 Jun 14]; 41: 1851–64. doi: 10.1159/000471933.

36. Li W, Khor TO, Xu C, Shen G, Jeong W-S, Yu S, Kong A-N. Activation of Nrf2-antioxidant signaling attenuates NF-kappaB-inflammatory response and elicits apoptosis. *Biochem Pharmacol* [Internet]. 2008 [cited 2017 Jun 14]; 76: 1485–9. doi: 10.1016/j.bcp.2008.07.017.
37. Jana A, Krett NL, Guzman G, Khalid A, Ozden O, Staudacher JJ, Bauer J, Hyun Baik S, Carroll T, Yazici C, Jung B. NFkB is essential for activin-induced colorectal cancer migration via upregulation of PI3K-MDM2 pathway. *Oncotarget* [Internet]. 2017 [cited 2017 Jun 14]; 8: 37377–93. doi: 10.18632/oncotarget.16343.
38. Zulato E, Bergamo F, De Paoli A, Griguolo G, Esposito G, De Salvo GL, Mescoli C, Ruge M, Nardin M, Di Grazia L, Lonardi S, Indraccolo S, Zagonel V. Prognostic significance of AMPK activation in advanced stage colorectal cancer treated with chemotherapy plus bevacizumab. *Br J Cancer* [Internet]. 2014 [cited 2017 Jun 14]; 111: 25–32. doi: 10.1038/bjc.2014.274.
39. Heiss EH, Schachner D, Zimmermann K, Dirsch VM. Glucose availability is a decisive factor for Nrf2-mediated gene expression. *Redox Biol* [Internet]. 2013 [cited 2017 Jun 14]; 1: 359–65. doi: 10.1016/j.redox.2013.06.001.
40. Kowalik MA, Columbano A, Perra A. Emerging Role of the Pentose Phosphate Pathway in Hepatocellular Carcinoma. *Front Oncol* [Internet]. Frontiers Media SA; 2017 [cited 2017 Sep 4]; 7: 87. doi: 10.3389/fonc.2017.00087.
41. Jornayvaz F, Shulman G. Regulation of mitochondrial biogenesis. *Essays Biochem* [Internet]. 2010; 47: 69–84. doi: 10.1042/bse0470069.Regulation.
42. Bragoszewski P, Turek M, Chacinska A. Control of mitochondrial biogenesis and function by the ubiquitin-proteasome system. *Open Biol* [Internet]. 2017 [cited 2017 Jun 14]; 7: 170007. doi: 10.1098/rsob.170007.
43. Cruz M Dela, Ledbetter S, Chowdhury S, Tiwari AK, Momi N, Wali RK, Bliss C, Huang C, Lichtenstein D, Bhattacharya S, Varma-Wilson A, Backman V, Roy HK. Metabolic reprogramming of the premalignant colonic mucosa is an early event in carcinogenesis. *Oncotarget* [Internet]. Impact Journals, LLC; 2017 [cited 2017 Sep 4]; 8: 20543–57. doi: 10.18632/oncotarget.16129.

Analysis of NRF2-related predictive pathways associated with colorectal cancer

Carmen Gonzalez-Donquiles, Jessica Alonso-Molero, Laura Vilorio-Marqués, Mariona Bustamante, Antonio J. Molina, Nerea Fernández de Larrea, Javier Llorca, Pilar Amiano, M^a Eva Ardanaz Aicua, Guillermo Fernandez-Tardon, Benito Mirón-Pozo, Rosana Peiró, Juan Alguacil, Ana Espinosa, Tomás González Elosua, Beatriz Pérez-Gómez, Camilo Palazuelos, Luisa C. de la Peña-Negro, Tania Fernández-Villa, Marina Pollán, Manolis Kogevinas, Vicente Martín.





Figures 1. Scatter plots of the distance scores plus overlaid boxplots regarding to ABC Transporter pathway

Figures 2. Scatter plots of the distance scores plus overlaid boxplots regarding to Mitochondrial Biogenesis pathway

Figures 3. Scatter plots of the distance scores plus overlaid boxplots regarding to MAPK signaling pathway

Figures 4. Scatter plots of the distance scores plus overlaid boxplots regarding to NFkB signaling pathway

Figures 5. Scatter plots of the distance scores plus overlaid boxplots regarding to PI3K-AKT signaling pathway

Figures 6. Scatter plots of the distance scores plus overlaid boxplots regarding to Inflammatory Bowel Disease pathway.

