

Grado en Biología

Memoria del Trabajo de Fin de Grado

Estudio proteómico de biopsias de cáncer de colon mediante array de proteínas.

Estudo proteómico de biopsias de cancro de colon mediante array de proteínas.

Proteomic study of colon cancer biopsies by proteinarray.

Antonio Morano Torres

Septiembre, 2018

Directora Académica: Valentina Calamia

Codirectora: María Esther Rodríguez Belmonte

TRABAJO DE FIN DE GRADO

Dña María Esther Rodríguez Belmonte y Dña Valentina Calamia autorizan la presentación del trabajo de Fin de Grado "Estudio proteómico de biopsias de cáncer de colon mediante array de proteínas." Presentado por Antonio Morano Torres para su defensa ante el tribunal calificador.

En A Coruña, a 11 de Septiembre de 2018

A handwritten signature in blue ink, consisting of several loops and a long horizontal stroke at the end.

Fdo: María Esther Rodríguez Belmonte

A handwritten signature in blue ink, appearing as a stylized 'V' followed by a horizontal line.

Fdo: Valentina Calamia

ÍNDICE

1. Resumen	1
2. Introducción.....	3
2.1. Cancer.....	3
2.1.1. Cáncer de Colon (CCR)	3
2.2. Proteómica	6
2.2.1. Proteómica dirigida.....	7
2.2.1.1. Arrays de proteínas	7
2.3. Biomarcadores	9
2.3.1. Biomarcadores en CCR.....	11
3. Objetivos	12
4. Material y métodos	13
4.1. Procesamiento de las muestras	13
4.2. Extracción de proteínas.....	14
4.3. Cuantificación de proteínas	15
4.4. Luminex y xMAP.....	17
4.5. Análisis estadístico	19
5. Resultados y discusión.....	20
5.1. THBS1	20
5.2. AHSG	22
5.3. RBP4.....	23
5.4. Estudio de correlación	25
5.5. Discusión general.....	25
6. Conclusiones.....	27
7. Bibliografía	28

1.a Resumen

El cáncer de colon es el tercer cáncer con mayor incidencia a nivel global y el segundo con mayor índice de mortalidad en España. La supervivencia a este cáncer varía entre un 90% y un 5-15% según el estadio en el que se detecte. A pesar de la importancia de un diagnóstico precoz, en la actualidad no se encuentran herramientas específicas para la detección temprana de esta enfermedad debido a su heterogeneidad y al elevado coste que esto supone.

Una buena estrategia, fácilmente implementable en la práctica clínica, consistiría en el estudio de los niveles de biomarcadores específicos de cada estadio de cáncer de colon, para la monitorización precisa de la enfermedad y la evaluación de la efectividad de nuevas estrategias terapéuticas.

En este trabajo se pretende investigar una cohorte de 165 muestras de biopsias de cáncer de colon mediante microarrays analítico de anticuerpos en formato sándwich frente a proteínas previamente descritas como potenciales biomarcadores. Para la realización de estos arrays se utilizará la tecnología xMAP de Luminex, basada en arrays multiplex en formato esfera.

Palabras Clave: Cáncer de colon, proteómica, array multiplex de anticuerpos, biomarcadores.

1.b Resumo

O cancro de colon é o terceiro cancro con maior incidencia en todo o mundo e o segundo con maior mortalidade en España. A supervivencia a este cancro varía entre o 90% e o 5-15% dependendo do estadio en que se detecte. A pesar da importancia do diagnóstico precoz, actualmente non existen ferramentas específicas para a detección precoz da enfermidade, debido a súa heteroxeneidade e os elevados custos implicados.

Unha boa estratexia facilmente implementable na práctica clínica, consistiría no estudo dos niveles de biomarcadores específicos de cada estadios do cancro de colon, para a monitorización precisa da enfermidade e a evacuación da eficacia de novas estratexias terapéuticas.

Neste estudo pretendese investigar un conxunto de 165 mostras de biopsias de cancro de colon por microarrays analíticos de anticorpos formato sándwich frente a proteínas que foron relatadas en estudos anteriores con potencial biomarcador do cancro de colon. Para a realización destes arrays empregouse a tecnoloxía Xmap de Luminex, baseada en arrays multiplex de formato esfera.

Palabras Clave: Cáncer de colon, proteómica, array multiplex de anticorpos, biomarcadores.

1.c Abstract

Colon cancer is the third cancer with the highest incidence worldwide and the second with the highest mortality rate in Spain. Survival to this cancer differs between 90% and 5-15% depending on the stage in which it is detected. Despite the importance of early diagnosis, no specific tools are currently available for the early detection of this disease due to its heterogeneity and the high cost involved.

A good strategy, easily implementable in clinical practice, would consist in the study of the level of specific biomarkers proteins in different stages of colon cancer, which would ideally allow the precise monitoring of the disease and the evaluation of the effectiveness of new therapeutic strategies.

In this work the goal was to investigate a cohort of 165 colon cancer tissue samples by analytical microarrays of antibodies in sandwich format against proteins that have been reported in previous studies with biomarker potential of colon cancer. For the development of these arrays the xMAP technology from Luminex, based on multiplex sphere format arrays, will be used.

Key words: Colon cancer, proteomics, multiplex antibody array, biomarkers.

2. Introducción

2.1 Cancer

El cáncer es una enfermedad genética, desencadenada por una única célula debido a una serie de mutaciones que se producen en el ADN en diferentes etapas. La acumulación de estas mutaciones provoca la síntesis de proteínas que interfieren en la regulación del ciclo celular, ocasionando un crecimiento y proliferación descontroladas de la célula mutada y en consecuencia de las células hijas, originándose así un tumor (Hejmadi, 2010). Según la patogenicidad y capacidad de invasión de las células cancerígenas, se distinguen cuatro estadios diferentes de cáncer, correspondiéndose el estadio 0 a una hiperplasia celular benigna. Los estadios del I al III corresponden a un tumor cada vez mayor que puede además invadir tejidos cercanos. Por último, el estadio IV corresponde a un cáncer con metástasis en órganos lejanos al tumor original (Figura 1) (Greene et al, 2002; Holland-Frei et al, 2003).

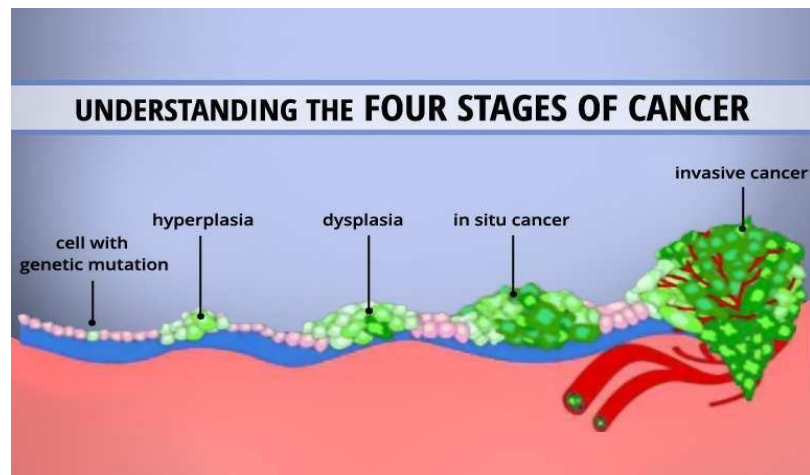


Figura 1: Representación gráfica de los distintos estadios de la progresión tumoral (<https://thetruthaboutcancer.com/understanding-four-stages-cancer/>).

Esta enfermedad afecta aproximadamente a uno de cada tres individuos (Hejmadi, 2010) y es globalmente la segunda causa de muertes según la World Health Organization (WHO). Las principales causas de la elevada prevalencia y mortalidad de esta enfermedad son, entre otras, el envejecimiento de la población y la adopción de malos hábitos de vida como el sedentarismo, el tabaquismo, el alcoholismo o las dietas altas en grasas animales y bajas en consumo de frutas y verduras (Siegel et al, 2018).

2.1.1 Cáncer de Colon (CCR)

Entre todos los tipos de cáncer que se diagnostican anualmente, el cáncer colorrectal (CCR) es el tercero con mayor incidencia si se consideran ambos sexos o únicamente los hombres, y alcanza el segundo puesto cuando se

consideran solo las mujeres, según los datos disponibles de la WHO (Figura 2). En el caso concreto de España encontramos datos más recientes, datados marzo 2018, de la Asociación Española Contra el Cáncer (AECC), donde el CCR ocupa el segundo puesto entre los cánceres con mayor mortalidad entre los diagnosticados tanto en hombres como en mujeres.

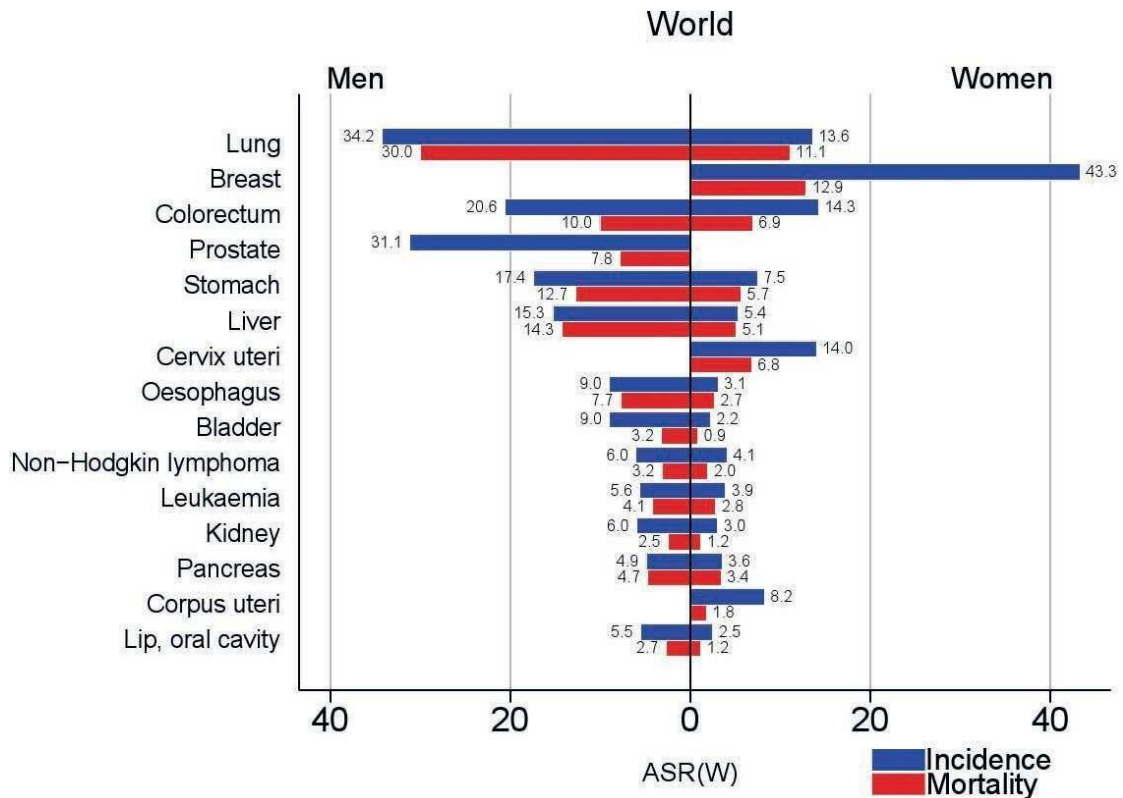


Figura 2: Grafico de incidencia y mortalidad de los diferentes tipos de cáncer (Stewart, 2014).

El cáncer de colon se define como un carcinoma, normalmente un adenocarcinoma, que presenta una progresión escalonada, evolucionando de una lesión benigna precursora (adenoma) a una lesión invasiva (adenocarcinoma) con capacidad metastásica (Figura 3) (Stewart et al, 2014). Esta primera lesión precursora se produce a nivel de las criptas del epitelio en la pared del intestino grueso, donde una célula clonogénica (célula madre) debido a una acumulación de desajustes genéticos y epigenéticos transforma el epitelio normal en un epitelio displásico (Jimenez et al, 2010). Este epitelio displásico atraviesa una etapa prolongada donde las células presentan una capacidad de crecimiento autónomo, pero no presentan capacidad invasiva ni metastásica: a esto se le conoce como adenoma. Si esta lesión progresa adquiriendo capacidad invasiva y metastásica se habla de adenocarcinoma. Afortunadamente, tan solo el 10% de los adenomas evolucionan a adenocarcinoma, tardando este proceso entorno a los 10-20 años desde la formación de la lesión primaria (Luo et al, 2013).

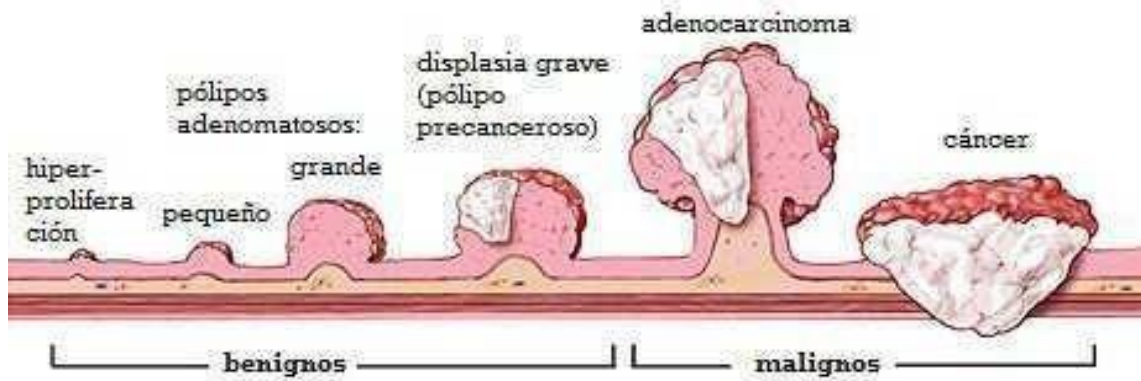


Figura 3: Evolución de intestino sano a intestino cancerígeno (<http://varimed.ugr.es/index.php?op=viewconcepto&idconcepto=577>).

Las formas de expansión primarias del CCR son la invasión de la pared del intestino desde la cripta, metástasis a nodos linfáticos cercanos y en etapas más tardías metástasis en pulmones e hígado (Stewart et al, 2014). El conocimiento de estas etapas de expansión medidas en términos de invasión local, invasión de nodos linfáticos y órganos distales es fundamental para el diagnóstico y pronóstico de estos tumores (Jimenez et al, 2010). Una clara evidencia de esto es que la supervivencia a cinco años desde la detección de dicho cáncer varía de un 90% cuando se diagnostica en grado 1, a una supervivencia del 60% cuando es diagnosticada en el estadio 3 y llegando a estar entorno al 5-15% cuando se detecta en el estadio 4 (Esteve et al, 2007).

Actualmente la detección del CCR se basa en test que incluyen sangre oculta en muestras fecales, sigmoidoscopia flexible, enema de bario o colonoscopia. Sin embargo, ninguno de estos métodos es perfecto, debido a sus baja sensibilidad y especificidad, por lo que se hace necesario encontrar nuevos métodos de diagnóstico que sean más eficaces y menos invasivos (Luo et al, 2013). En este aspecto toma un papel relevante la búsqueda de biomarcadores específicos de CCR. Las técnicas proteómicas serían las más adecuadas respecto a otras ciencias "ómicas", como la genómica o la transcriptómica, para la búsqueda de biomarcadores útiles para el diagnóstico precoz de esta enfermedad contribuyendo además a la comprensión de los mecanismos patogénicos responsables de la transformación neoplásica (Ortega, 2010). En este sentido cabe destacar que el 98% de las dianas terapéuticas actuales están constituidas por proteínas (Mustafa et al, 2011). Esto se debe a que las proteínas juegan un papel importante en las enfermedades, así como en la respuesta y recuperación, y ayudan a comprender las bases moleculares de dichas patologías. Además, la mayoría de las proteínas secretadas se expresan durante fases específicas del crecimiento, por tipos particulares de células o durante respuestas específicas, y es por esto que son las más utilizadas comúnmente como biomarcadores (Barderas et al, 2010; Mustafa et al, 2011; Hernandez-Sanchez et al, 2011). Hasta el momento se han descrito más de sesenta proteínas expresadas de manera diferencial entre una mucosa

intestinal normal y una cancerígena, entre las que destacan proteínas como anexinas, citoqueratinas o quinasas (Barderas et al, 2010). En un futuro próximo, podríamos incluso llegar a la caracterización de los perfiles proteicos específicos de cada tipo tumoral mejorando así el manejo clínico de los pacientes, lo cual permitiría intervenciones terapéuticas más personalizadas según la agresividad de cada tumor y las características clínicas de cada paciente (Sanchez-Carballo, 2007).

2.2 Proteómica

El término proteómica hace referencia a la ciencia y tecnología asociada con el estudio de la expresión de todas o la mayoría de las proteínas de un sistema vivo. Un estudio proteómico engloba no solo la identificación de proteínas, sino también el estudio de su papel en procesos fisiológicos y patológicos (Nambiar et al, 2010). A nivel biomédico, las técnicas proteómicas están encontrando una amplia aplicación en la determinación de las bases moleculares asociadas a enfermedades (proteómica funcional), las bases de acciones de fármacos, así como su toxicidad (farmacoproteómica), y el desarrollo de nuevos marcadores específicos para cada enfermedad (proteómica clínica).

El estudio del proteoma de una célula cancerosa se puede abordar usando dos estrategias proteómicas diferentes. La primera supone un conocimiento previo del proteoma a ser investigado y se conoce como proteómica dirigida; en el segundo caso el proteoma se desconoce por lo que se utilizan las técnicas de proteómica al azar (no dirigidas). Las diferencias entre estas dos aproximaciones se reflejan en la tabla 1:

Tabla 1: Características principales de las técnicas de proteómica dirigida y no dirigida.

Característica/Tipo	No dirigida	Dirigida
Proteínas medidas	Desconocidas	Conocidas
Objetivo	Identificación	Cuantificación
Ventajas	Descubrimiento	Validación
Desventajas	Reproducibilidad, cuantificación, automatización	Requiere reactivos y protocolos optimizados
Tecnologías	Geles bidimensionales, Cromatografía líquida, espectrometría de masas.	Arrays de tejido, Western blotting, arrays de proteínas y anticuerpos.

Dentro de las técnicas proteómicas, las inmunoanalíticas poseen una alta sensibilidad y selectividad, permitiendo la cuantificación de posibles biomarcadores incluso a bajas concentraciones, garantizando así la detección de patologías en fases tempranas (Olives et al, 2010).

2.2.1 Proteómica dirigida

La proteómica proporciona un conjunto de herramientas muy poderosas para el estudio a gran escala de la función de las proteínas, de su estructura, localización, modificaciones postraduccionales e interacciones. Entre todas las técnicas utilizadas solo se entrará en detalle en las técnicas basadas en arrays de proteínas por ser las relevantes para este trabajo.

2.2.1.1 Arrays de proteínas

El término microarray, también conocidos como chips o matrices de proteínas, hace referencia a la miniaturización de miles de pruebas en una misma placa (Sutandy et al, 2013). En esencia se trata de un inmunoensayo en miniatura, en los que actualmente se llega a fijar hasta 10^5 sondas en superficies de 5 a 1,6 cm^2 (Herraez, 2010). Consta de una superficie sólida, normalmente vidrio, donde miles de proteínas diferentes se inmovilizan en un espacio reducido formando una densa matriz (Hu et al, 2011). Estas proteínas serán posteriormente interrogadas con otras moléculas (Lourido, 2015).

Actualmente esta técnica constituye una tecnología de alto rendimiento capaz de estudiar de forma simultánea un alto número de proteínas sin necesidad de una separación previa de la muestra (Lourido, 2015). Los arrays se han impuesto a otras técnicas más tradicionales como la espectrometría de masas o los ensayos de inmunoabsorción ligados a enzimas (ELISA) debido a su capacidad de detectar proteínas con una alta especificidad y sensibilidad usando para ello menos de 1 μL de muestra. Esto supone una gran ventaja al utilizar mucha menos cantidad de muestra y poder además analizar cientos de analitos a la vez (Sutandy et al, 2013).

Para la generación y la detección de la señal en estos ensayos se pueden utilizar diferentes técnicas: fluorescencia, quimioluminiscencia o métodos sin marcaje como el quantum dots o el plasmón de resonancia. De todos ellos, los métodos basados en fluorescencia son los más utilizados, puesto que la señal ofrece una elevada sensibilidad y amplio rango dinámico de detección, permitiendo incluso detectar y cuantificar la señal de cada uno de los puntos del array de forma casi simultánea (Herraez, 2010).

Los arrays pueden ser en esfera, también conocidos como arrays en suspensión, o planos según el soporte sobre el que se realicen. Además, se pueden clasificar en dos tipos atendiendo al compuesto fijado: arrays de antígenos y arrays de anticuerpos.

Los arrays de antígenos permiten realizar tanto ensayos analíticos para ver la abundancia de la proteína de interés en la muestra, como ensayos funcionales para la detección de interacciones, modificaciones postraduccionales o actividad enzimática (Hu et al., 2011). Los antígenos son inmovilizados en el array, y la interacción antígeno-anticuerpo se visualiza mediante el empleo de un anticuerpo secundario conjugado con una etiqueta que permita amplificar la señal de fluorescencia.

Los arrays de anticuerpos son arrays analíticos cuya principal aplicación es la detección de proteínas diferencialmente expresadas y su abundancia, basándose para ello en la fijación de anticuerpos de captura a un soporte sólido para detectar dichas proteínas. Los arrays de anticuerpos se pueden presentar en varios formatos: arrays directos o arrays tipo “sándwich”.

En los arrays directos las proteínas son capturadas por un único anticuerpo inmovilizado en el soporte y la señal de la proteína se detecta a través del marcaje de la muestra. Este marcaje puede ser de dos tipos, directo o indirecto (Figura 4). En el marcaje directo, las proteínas se marcan normalmente con un fluorocromo tipo cianina. En el marcaje indirecto, las proteínas se marcan a través de moléculas intermedias como la biotina o la digoxigenina que son posteriormente detectadas por compuestos fluorescentes (Lourido, 2015). La principal ventaja de los arrays directos es que son ensayos competitivos, lo cual conlleva a una mejora de la linealidad de la respuesta y del rango en comparación a los que no lo son, sin embargo, podría producirse una distorsión en la interacción analito-sonda limitando así la sensibilidad y especificidad en relación con los no competitivos (Sanchez-Carballo, 2007).

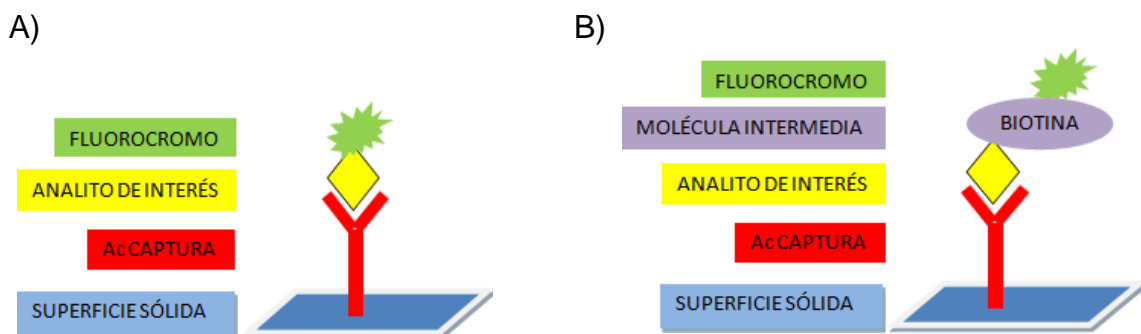


Figura 4: A) Microarray plano con marcaje directo a través de fluorocromo. B) Microarray plano con marcaje indirecto a través de molécula intermedia.

En los arrays en formato sándwich se emplean dos tipos de anticuerpos diferentes para detectar una proteína específica. Un anticuerpo, denominado anticuerpo de captura, inmoviliza la proteína problema en una fase sólida; el otro anticuerpo, conocido como anticuerpo reportero o de detección, genera una señal para el sistema de detección, dando una medida de la abundancia de proteína inmovilizada (Sutandy et al, 2013). Este tipo de arrays solo permite ensayos no competitivos, lo que provoca que se obtengan curvas sigmoideas en la relación analito-sonda y se necesita realizar curvas de calibración con estándares de concentración conocida (Sanchez-Carballo, 2007).

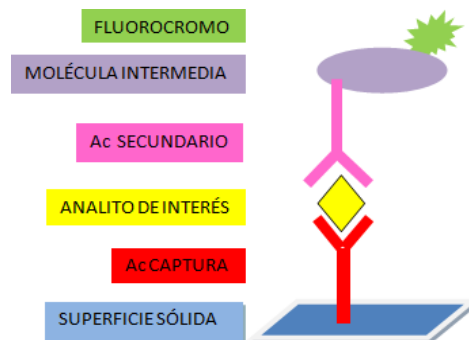


Figura 5: Microarray plano en formato sándwich.

Una de las últimas y más eficaces tecnologías creadas en este campo ha sido desarrollada por la compañía Luminex bajo el nombre de xMAP (*unknown multianalyte profiling*). Se trata de una técnica muy novedosa que permite realizar ensayos de múltiples analitos de forma simultánea. Luminex ha desarrollado también la tecnología MagPlex, complementaria a xMAP, que consiste en el uso de anticuerpos acoplados a microesferas magnéticas para la detección de proteínas diana. La gran ventaja de esta tecnología reside en el hecho de que las esferas sean magnéticas, ya que esto permite la captura de proteínas solubles con una alta reproducibilidad gracias a un aumento en la automatización del ensayo (Angeloni et al, 2014; Reslova et al, 2017).

2.3 Biomarcadores

Los biomarcadores son un grupo diverso de moléculas como secuencias de ADN, ARNs, proteínas y metabolitos (Zamanian-Azodi et al, 2015). Sirven como indicadores de fases biológicas específicas y pueden ser utilizados tanto para el diagnóstico (preferiblemente en una etapa temprana) y la progresión de una enfermedad, permitiendo identificar los estadios de la misma, como también para optimizar las terapias, pudiendo predecir y monitorizar la respuesta a distintos tratamientos y seleccionar así el más adecuado (Barderas et al, 2010; Wit et al, 2013). Debido a las diferentes utilidades que pueden abarcar los biomarcadores se clasifican en biomarcadores de riesgo, que son aquellos que

informan sobre la predisposición a desarrollar una patología, biomarcadores de diagnóstico, que son aquellas proteínas o metabolitos que incrementan sus niveles con la proliferación de la enfermedad y biomarcadores pronóstico, que son los que ayudan a monitorizar el desarrollo y efecto de los fármacos en la enfermedad (Olives et al, 2010).

El biomarcador ideal debe cumplir unos requisitos muy estrictos: ha de ser específico de un tejido o una enfermedad concreta para evitar la ocurrencia de falsos positivos (Wit et al, 2013). Además, los biomarcadores suelen aparecer a muy bajas concentraciones, por lo que es muy importante medirlos con una alta sensibilidad. También han de poseer un elevado poder predictivo (>80%), para ser realmente útiles al predecir cambios y monitorizar las diferentes fases de una enfermedad. Además, para que un biomarcador sea empleado en la práctica clínica debe ser fácilmente detectable en muestras de uso rutinario en un hospital como suero, plasma u orina (Casal y Vivanco, 2014).

El objetivo final de un biomarcador es el de introducir un indicador específico y sensible para cierto tipo de enfermedad (Zamanian-Azodi et al, 2015). Por lo tanto, para que llegue a ser utilizado de forma rutinaria en la práctica clínica ha de superar un largo proceso que va desde su descubrimiento hasta su entrada en el mercado (figura 6). Los experimentos de proteómica al azar para el descubrimiento de nuevas proteínas con potencial biomarcador deben ir seguidos del diseño de estrategias dirigidas, con el fin de poder verificar su utilidad biomarcadora en un número más amplio de muestras. Una vez seleccionados los posibles biomarcadores, se inicia una fase de desarrollo en la que se evalúa la sensibilidad y especificidad de cada uno de ellos. Si el candidato cumple todos los requisitos necesarios exigidos durante la fase clínica, será lanzado al mercado (Ruiz-Romero y Blanco, 2010).



Figura 6: Fases de desarrollo de un biomarcador.

Sin embargo, la realidad es que para la mayoría de las enfermedades no existe un único y específico biomarcador que sirva para determinarla, sino que es un conjunto de proteínas las que la determinan. Por esto en los últimos años han ganado protagonismo los paneles de biomarcadores sobre los biomarcadores únicos (Zamanian–Azodi et al, 2015). Prueba de esto son paneles de biomarcadores que se encuentran actualmente en uso para la detección de

diferentes cánceres, como el suministrado por la FDA (Food and Drug Administration) americana bajo el nombre MammaPrint, que es un test para la detección de cáncer de mama basado en la detección de la expresión de biomarcadores relacionados con esta enfermedad, o el CupPrint, un dispositivo para la detección de cáncer de origen primario desconocido (Ortega, 2010).

2.3.1 Biomarcadores en CCR

Los biomarcadores de cáncer se conocen como marcadores tumorales (MT), aunque no todos son específicos de procesos cancerígenos. Para asegurarse de que estos MT realmente correspondan a un proceso tumoral se recurre a tres verificaciones. La primera consta en determinar la concentración sérica: a mayor concentración en suero, mayor probabilidad de un proceso tumoral. En segundo lugar, hay que tener en cuenta que en los riñones e hígado es donde ocurre la mayor degradación de estos MT, por lo que los procesos cancerígenos suelen conllevar insuficiencias renales y hepáticas. Por último, se determina su concentración a intervalos de 2-3 semanas (intervalos superiores a la vida media de los MT), por lo que de mantenerse constante o incrementarse su concentración sería un claro signo de cáncer (Ocaña y Aceituno, 2014).

En cuanto al CCR encontramos un alto número de investigaciones referentes a posibles biomarcadores en la literatura. Revisiones como la de Luo *et al* 2013 hacen referencia a proteínas como el antígeno carcinoembrionario (CEA), el antígeno carbohidratado 19-9 (CA19-9), la proteína 1 de unión final (EB1), y la ribonucleoproteína nuclear heterogénea A1 (hnRNP A1). La recopilación realizada por Zamanian-Azodi *et al* en 2015 recoge antígenos específicos de cancer de colon (CCSA) como CCSA-2, CCSA-3, CCSA-4 o CCSA-5. Encontramos estudios sobre otros posibles biomarcadores como la trombospondina 1 (THBS1) (Miyanga et al, 2002), la fetuina A (AHSG) (Swallow et al, 2004) o la proteína 4 de unión al retinol (RBP4) (Fei et al, 2017). De todos ellos actualmente tan solo el CEA es utilizado como biomarcador clínico de CCR para monitorizar la respuesta sistémica a distintas terapias (Wit et al, 2013). Esto se debe a las dificultades que existen en cuanto a la validación de estos biomarcadores. La heterogeneidad de esta enfermedad (Jimenez et al, 2010), así como el escaso número de muestras utilizadas en los trabajos y los altos costes de los análisis necesarios (Barderas et al, 2010) suponen los principales problemas para la validación clínica de estos biomarcadores.

3. Objetivos

Este trabajo tiene como objetivo principal la validación proteómica de posibles biomarcadores de CCR. Dicho proyecto se realizará a través de la utilización de un array en formato esfera empleando la tecnología xMAP de Luminex, lo que permitirá llevar a cabo el análisis múltiple de un panel de proteínas en biopsias de pacientes de diferentes estadios.

Para alcanzar el objetivo principal se plantean los siguientes objetivos secundarios:

- Extracción de proteínas a partir de muestras de tejido humano.
- Cuantificación de los extractos proteicos.
- Desarrollo de un ensayo múltiple para la detección simultánea de tres posibles biomarcadores de CCR: THBS1, AHSG y RBP4.

4. Material y métodos

4.1 Procesamiento de las muestras

Para el desarrollo de este proyecto se utilizó un set de muestras de tejido procesadas en la unidad de anatomía patológica del Hospital Universitario de A Coruña entre los años 2007-2008, y criopreservadas en el Biobanco de A Coruña en bloques de OCT (temperatura de corte óptima). El OCT es un compuesto formado por una disolución de glicoles solubles y resinas en agua destilada (H₂O dd), y se utiliza para preservar muestras entre -10 y -80°C. Se procesaron en total 165 biopsias de CCR procedentes de pacientes con distintas características (ver anexo 1).

Debido a las diferencias en el tamaño entre las diferentes muestras, el primer paso que se realizó fue una medida del área superficial de los tejidos incluidos en OCT con el fin de conseguir la mayor homogeneidad posible antes de la extracción. Estas medidas se utilizaron para establecer el número de cortes a realizar en cada muestra, ya que cada corte se hizo de un tamaño de 20µm. Finalmente se realizaron de dos a diez cortes según el tamaño de cada muestra (Figura 7). Estos cortes fueron realizados por el personal del biobanco y se guardaron en tubos de 2ml a -20°C, para su posterior procesamiento.

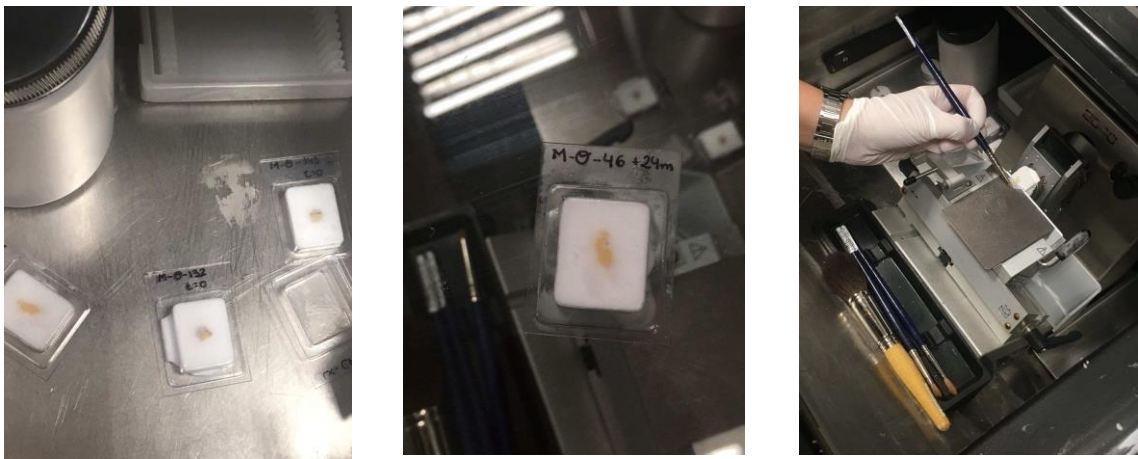


Figura 7: Imágenes tomadas de los bloques de OCT y del criostato empleado para la realización de los cortes (Reichert-Jung 2800; Leica Instruments).

En la figura 8 se muestra el flujo de trabajo que se siguió durante el desarrollo de este experimento:

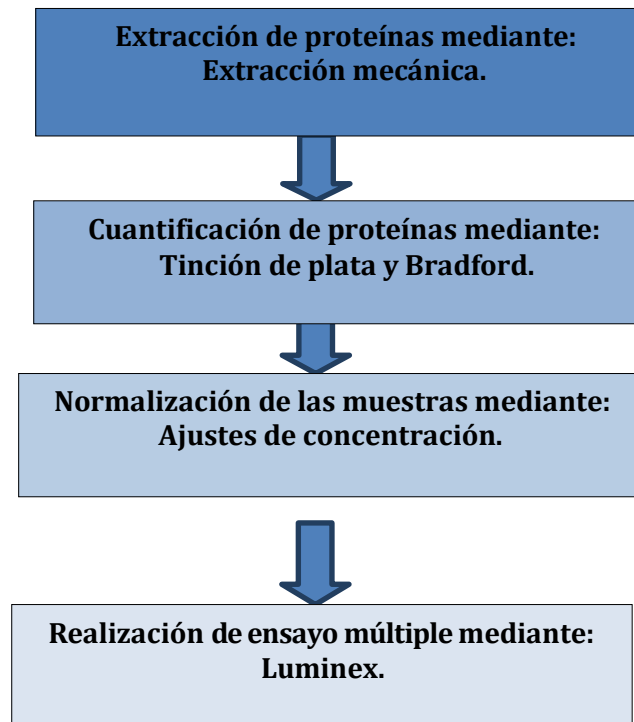


Figura 8: Esquema del flujo de trabajo realizado para la validación de biomarcadores de CCR.

4.2 Extracción de proteínas

La extracción de las proteínas se realizó siguiendo un protocolo de extracción mecánica, manteniendo las muestras en hielo siempre que fuera posible. Dicho protocolo consta de los siguientes pasos:

1. Se añadió 1ml de PBS filtrado a los tubos que contenían los cortes para eliminar el OCT. Se centrifugó 5 minutos a 12000 g a 4°C y se retiró el sobrenadante.
2. Se añadió un volumen determinado según el tamaño de cada muestra de tampón de lisis comercial compatible con Luminex (Invitrogen, FNN0071), que se completó con el cocktail de inhibidor de proteasas (10uL/mL) (Sigma, P8340).
3. Se añadieron dos bolitas de zirconio de 5mm de diámetro a cada tubo para facilitar la disgregación del tejido. Estos tubos se colocaron en unos soportes refrigerados y se agitaron en un homogeneizador (Mixer Mill MM 200, Retsch) a una frecuencia de 1500 hz durante 15 minutos.

4. Se retiraron los tubos de los soportes, se agitaron en un vortex a máxima velocidad durante 30 minutos y se dejaron durante 2 minutos en un baño de ultrasonidos.
5. Por último, se centrifugaron 5 minutos a 12000 g a 4°C. Tras este paso no debería quedar ningún pellet en los tubos y las proteínas deberían estar completamente solubilizadas, por lo que se pasó el sobrenadante a unos tubos nuevos de 1,5mL que se conservaron a -20°C.

4.3 Cuantificación de proteínas

La cuantificación de proteínas puede realizarse de manera tanto relativa como absoluta. En este proyecto se realizaron ambos métodos de cuantificación. El primero consistió en la tinción de plata de un gel de poliacrilamida y su posterior análisis densitométrico. El segundo consistió en la realización de un ensayo de Bradford que permite extrapolar la concentración absoluta de proteína a partir de una recta patrón realizada con una proteína estandar (comúnmente BSA).

- Tinción de plata de gel de acrilamida

La cuantificación de proteínas mediante la tinción de plata se llevó a cabo en cuatro pasos diferentes, el primero consistió en realizar un gel de acrilamida para seguidamente correr las muestras en una cubeta de electroforesis. El tercer paso consistió en la tinción y revelado de dicho gel con nitrato de plata. Por último, se adquirieron las imágenes de los geles con un digitalizador de imágenes y posteriormente se analizaron utilizando el programa Image Quant 5.2.

Los geles de poliacrilamida permiten separar proteínas de acuerdo a su tamaño molecular al migrar a través de una matriz porosa que presenta dos fases, el separador y el concentrador (Figura 9). Los geles monodimensionales permiten la separación de proteínas entre 10 y 300 Kilo Dalton (kDa) debido a la solubilización de las mismas en una red de dodecil sulfato sódico (SDS).

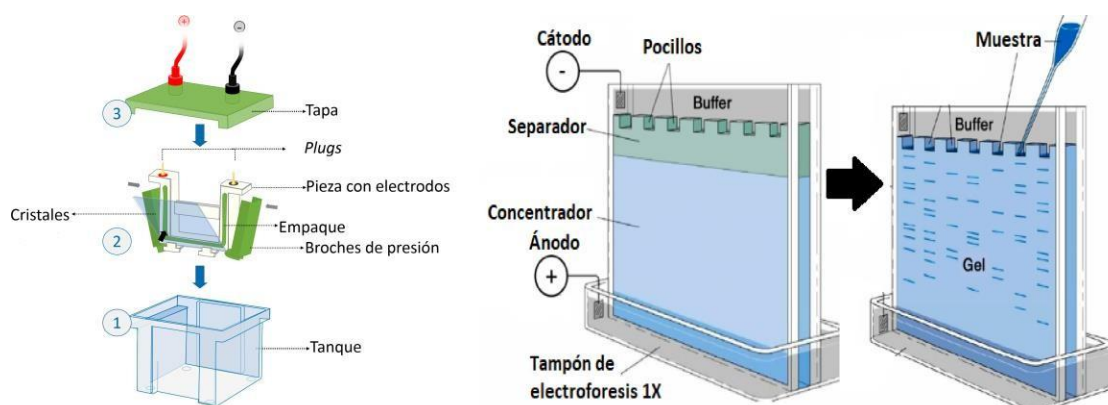


Figura 9: Gel de acrilamida y montaje en una cubeta de electroforesis.

Para la cuantificación de proteínas en gel se utilizó la tinción con nitrato de plata (Figura 10). Esta tinción se debe a la formación de complejos entre las proteínas y los iones de plata, que forman coloides de plata en la superficie de los geles, formando bandas visibles para su cuantificación con una sensibilidad entorno a los 0.5-1 ng (Shevchenko, 1996).

El último paso a realizar sería la densitometría, ayudándonos para ello de dos programas informáticos: Photoshop e Image Quant 5.2. Este último es un programa informático que mide la densidad de píxeles en las bandas y las transforma en datos numéricos, que en nuestro caso corresponderían a datos de concentración relativa de proteína en cada carril (muestra).



Figura 10: Tinción de plata. De izquierda a derecha encontramos: marker, control y muestras.

- Método de Bradford

La técnica de Bradford es un método colorimétrico para la cuantificación de proteínas que consiste en la tinción de proteínas con un colorante, el Coomassie Brilliant Blue G-250, que existe en tres formas iónicas diferentes: una catiónica roja, una neutra verde y una aniónica azul. De este modo bajo condiciones ácidas predomina la forma protonada roja, que al unirse a las proteínas se desprotona, tornándose azul y llegando a una absorbancia máxima de 595nm. Esta segunda forma será medida en los ensayos, sirviéndose para ello de una recta patrón creada a partir de concentraciones conocidas de albumina sérica bovina (BSA) (Bio-rad).

La cuantificación de proteínas mediante Bradford se llevó a cabo en placas de 96 pocillos, añadiendo una recta por triplicado con diferentes concentraciones de BSA. Cada muestra también se midió por triplicado para intentar disminuir posibles errores de pipeteo y conseguir unos coeficientes de variación adecuados. Se siguió el protocolo establecido por la casa comercial Bio-rad y se midió la absorbancia de las placas a 570nm en un espectrofotómetro (Nano Quant Infinite M200, Tecan).

4.4 Luminex y xMAP

La tecnología xMAP, desarrollada por la compañía Luminex permite la realización de ensayos de múltiples analitos de manera simultánea, utilizando

para ello un único volumen de muestra. Esta novedosa tecnología se basa en el concepto de array líquido (en suspensión) y utiliza como soporte para los ensayos esferas microscópicas de poliestireno. Luminex produce diferentes tipos de microesferas, pero para este trabajo se utilizaron microesferas magnéticas compatibles con el equipo MagPix, siempre de luminex. Estas microesferas presentan un diámetro de 6.5 micras y están impregnadas por toda la superficie con metal que contiene partículas magnéticas. Esto supone una gran ventaja frente a otras microesferas al optimizar los tiempos de lavado y al minimizar la pérdida de esferas, aumentando así la reproducibilidad de los ensayos. Distribuidos por toda la superficie de las esferas encontramos hasta 100 millones de grupos carboxilo que permiten la unión entre los anticuerpos y las esferas formando enlaces carbodimide entre las aminas primarias de los anticuerpos y estos grupos carboxilo. Internamente las microesferas presentan una combinación de hasta tres fluorocromos, produciendo en total hasta 500 colores diferentes en un único set de microesferas, teniendo cada población de esferas su perfil de emisión característico.

En este ensayo se realizó un array de tipo sándwich por lo que además del anticuerpo de captura, unido a las esferas, se utilizó un segundo anticuerpo de detección (Figura 11), marcado con biotina y cuya señal es amplificada por el sistema estreptavidina/ficoeretrina (SAPE). Este sistema puede ser excitado a una longitud de onda determinada reportando una fluorescencia que permite identificar una muestra problema capturada en la superficie de las microesferas.



Figura 11: Proceso secuencial de array tipo sándwich en microesferas magnéticas Luminex.

De esta manera el sistema de detección MagPix, acoplado a la tecnología xMAP se basa en la interacción de dos láseres diferentes (Figura 12). El primero es un láser de clasificación rojo con una longitud de onda de 635nm, que excita los fluorocromos internos de las microesferas consiguiendo así la identificación de cada microesfera gracias a su señal de espectro. El segundo láser reportero verde actúa a una longitud de onda de 525-532nm que excita la biotina del anticuerpo de detección por lo que, si el analito no está capturado en la superficie de la microesfera no devolverá ninguna señal, pero en caso de estar presente dará un perfil de emisión que podrá ser medido como índice de fluorescencia media (MFI). Estas medidas se llevan a cabo mediante el programa xPONENT que recoge docenas de lecturas para cada esfera produciendo distintos resultados para cada analito presente en la muestra.

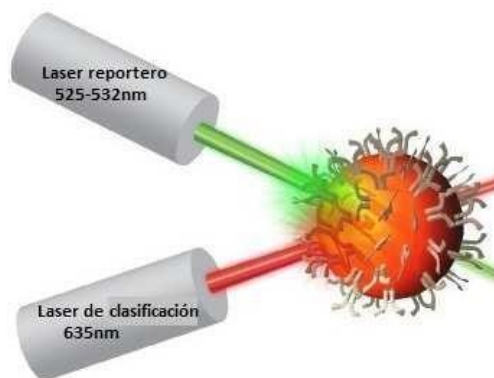


Figura 12: Esquema de funcionamiento de los láseres del MagPix.

Para realizar los inmunoensayos en formato sándwich con Luminex se necesitan: las muestras, una curva estándar a partir de concentraciones conocidas de la proteína recombinante y un control negativo formado por las microesferas magnéticas solas (bare beads, región 21). Para comenzar preparamos una dilución 1/50 con las tres regiones de esferas (13,18,26) acopladas a los tres anticuerpos que se desean analizar (THBS1, AHSG, RBP4) en tampón de bloqueo (PBS, 0,1% BSA, 0,02% Tween20, 0,05% NaN₃, pH=7,4). Se sonica durante dos minutos para evitar que se unan las esferas entre ellas. Tras ello, en una placa de 96 pocillos añadimos 5uL de esta disolución a cada pocillo (10³ esferas de cada región por pocillo), añadimos los volúmenes de proteína necesarios para cada punto de la curva estándar obteniendo un volumen final de 45uL en cada pocillo. A su vez se añaden diferentes diluciones de las muestras en pocillos individuales y se deja incubar durante dos horas y media en agitación y oscuridad a temperatura ambiente para que las proteínas (tanto de la curva como de las muestras) se unan a sus respectivos anticuerpos. Tras ello se lava 3 veces con 100uL de PBS-Tween 0,05% usando para ello un soporte magnético que facilitará el lavado al quedar las esferas magnéticas adheridas al fondo de la placa. Una vez realizados los lavados añadimos el anticuerpo de detección biotinilado (1ug/mL) hasta un volumen final de 45uL por pocillo e incubamos una hora de nuevo en agitación, oscuridad y temperatura ambiente para conseguir que el anticuerpo de detección se una a la muestra. Transcurrida la hora se lava de nuevo 3 veces de la misma manera, se añaden 50uL de SAPE a una dilución 1/1000 en PBS-Tween 0,05% e incubamos de nuevo 20 minutos bajo las mismas condiciones que las anteriores veces consiguiendo así la amplificación de la señal del anticuerpo biotinilado. Por último, lavamos otras 3 veces de la misma forma, añadimos 100uL de PBS-Tween 0,05% a cada pocillo y leemos la placa en el lector de fluorescencia MagPix.

4.5 Análisis estadístico

Para los análisis estadísticos de los datos se utilizaron una serie de test tanto paramétricos como no paramétricos empleando el programa informático GraphPad Prism 5.

- Test paramétricos

Dentro de los test paramétricos se utilizó el test de Shapiro Wilk como test de medida de la normalidad. Se realizó también el test de Bartlett para comprobar la homocedasticidad (varianzas iguales). El test de análisis de varianzas ANOVA se utilizó para comprobar si existían diferencias de las medias entre grupos y el test post-hoc de Tukey para comprobar entre que grupos existían estas diferencias en caso de resultar significativo el test ANOVA.

- Test no paramétricos

Dentro de los test no paramétricos se utilizó el test de Kruskal Wallis como homólogo al test ANOVA y como post hoc de este estadístico se realizó el test de Dunn. A mayores también se realizaron test de Mann Whitney U para comprobar las diferencias entre dos grupos individualmente en caso de ser necesario.

También se realizó un análisis de correlación de Pearson que nos indica la relación lineal que hay entre dos variables (dos proteínas).

La significación utilizada para todos los estadísticos realizados fue del 95%, por lo que se utilizó un $\alpha=0,05$.

5. Resultados y discusión

Tras la extracción y cuantificación proteica de las 165 muestras de CCR se comprobó que presentaban un rango muy amplio de concentración de proteínas, por lo que se decidió analizar solamente las muestras que presentaban una concentración proteica más parecida. En total se testaron 24 muestras (6 muestras de cada estadio) mediante ensayo multiplex de 3 posibles biomarcadores de CCR: THBS1, AHSG y RBP4.

La trombospondina (THBS1) es una glucoproteína trimérica multifuncional que participa en la angiogénesis y estimula el factor de crecimiento transformante beta 1 (TGF- β 1) fomentando la tumorigenesis (Miyanga, 2002). La Fetuina A (AHSG) es una glucoproteína formada por dos cadenas polipeptídicas implicada en la actividad del TGF- β 1, bloqueando su unión a receptores de membrana e inhibiéndolo, lo que sugiere su posible implicación en la progresión tumoral (Swallow, 2004). La proteína 4 de unión al retinol (RBP4) es una adipoquina que promueve la resistencia a la insulina y está presente en casos de obesidad. Además, se ha descrito que RBP4 puede estar implicada en el inicio de la transformación carcinogénica (Abola, 2015).

5.1 THBS1

Para la THBS1 se realizó una curva estándar según se detalla en la Figura 13. Se utilizaron nueve puntos diferentes con concentraciones que iban desde 0,3125 ng/mL a 80 ng/mL de proteína recombinante, alcanzando una R^2 de 0,9964.

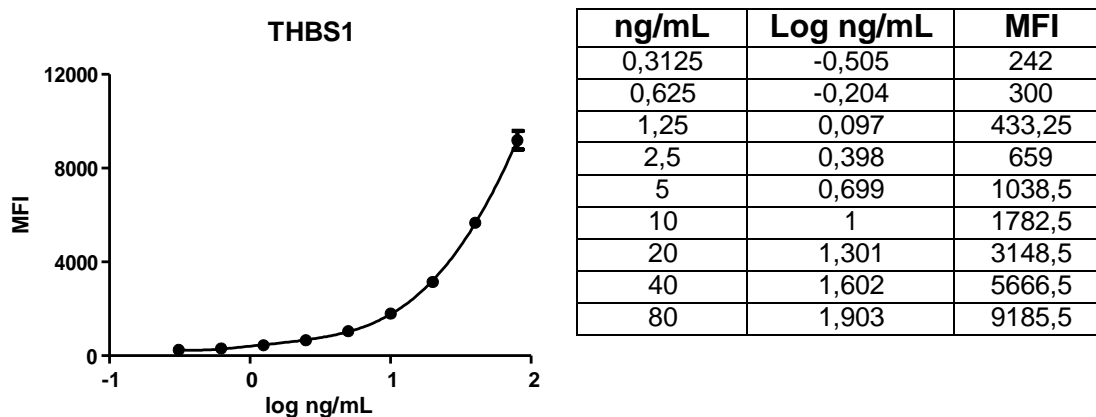


Figura 13: Curva estándar de THBS1 empleada en el ensayo Luminex.

La curva estándar se utilizó para extrapolar los datos de concentración de THBS1 en las 24 muestras testadas (Tabla 2). Dichos resultados se agruparon por estadios y se estudiaron las diferencias entre ellos mediante un test estadístico de análisis de varianza (ANOVA). Para ello se guardaron los residuos y se comprobó que siguiesen una distribución normal con un test de

Shapiro Wilk ($p=0,9419$) y que tuviesen varianzas iguales con un test de Bartlett ($p=0,3174$). El ANOVA realizado fue significativo ($p=0,00001$) lo que nos indicó que había diferencias entre grupos. Dados estos resultados se realizó como post-hoc el test de comparación múltiple de Tukey que a una significación del 95% nos dió los resultados que se indican en la tabla 3.

Tabla 2: Media de las concentraciones de THBS1 obtenidas para cada muestra

Estadio I	Estadio II	Estadio III	Estadio IV
713,299 ng/mL	3761,464 ng/mL	7321,507 ng/mL	1393,3 ng/mL
484,39 ng/mL	7445,788 ng/mL	14426,35 ng/mL	5597,844 ng/mL
1073,487 ng/mL	4182,348 ng/mL	15084,73 ng/mL	3328,484 ng/mL
1548,349 ng/mL	3386,666 ng/mL	7307,536 ng/mL	2346,086 ng/mL
1635,479 ng/mL	3355,147 ng/mL	8045,158 ng/mL	1180,919 ng/mL
693,886 ng/mL	2528,882 ng/mL	4979,118 ng/mL	770,649 ng/mL

Tabla 3: Test de comparación múltiple de Tukey entre grupos para THBS1.

Comparación múltiple	$\zeta P < 0,05?$	Resumen
I vs II	No	ns
I vs III	Si	***
I vs IV	No	ns
II vs III	Si	**
II vs IV	No	ns
III vs IV	Si	***

Los datos de las medias de concentración fueron además representados gráficamente junto con sus barras de error indicando los resultados significativos obtenidos por el test de Tukey (Figura 14).

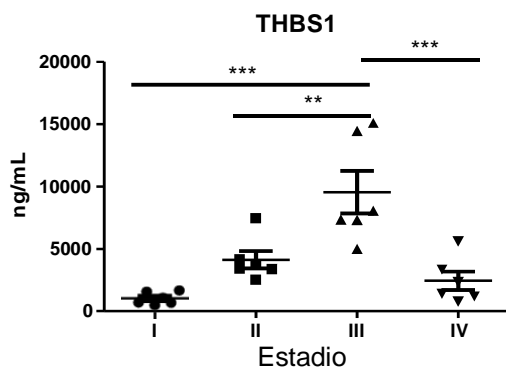


Figura 14: Gráfico de las concentraciones de THBS1 para los distintos estadios. ** $p \leq 0,01$; *** $p \leq 0,001$ según Tukey.

En esta gráfica se aprecia claramente que la concentración de THBS1 en el estadio III es significativamente aumentada en comparación al resto de grupos lo que nos permite proponer esta proteína como un posible biomarcador de CCR específica de este estadio.

5.2 AHSG

AHSG se analizó de la misma manera que THBS1. Primero se realizó una curva estándar con nueve puntos diferentes con concentraciones que iban de 1,17ng/mL a 300ng/mL y que presentó una $R^2= 0,9976$ (Figura 15).

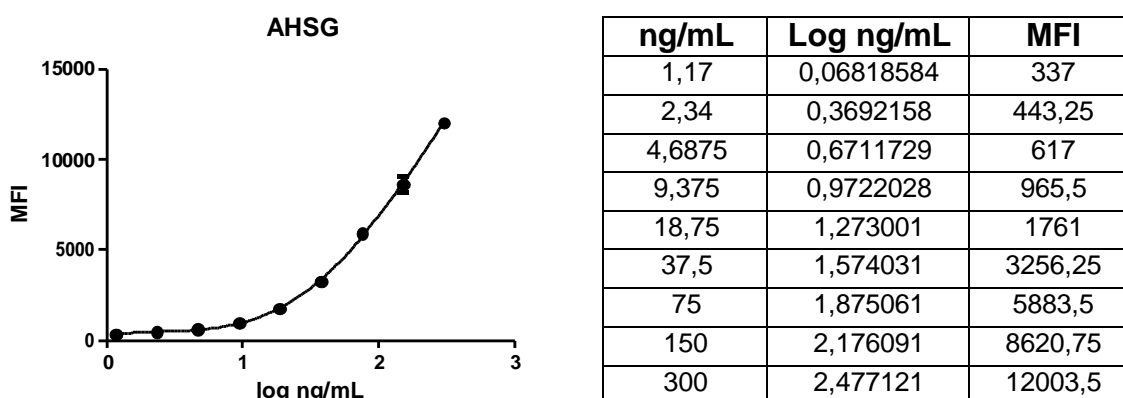


Figura 15: Curva estándar de AHSG empleada en el ensayo Luminex.

La curva estándar se utilizó para extrapolar los datos de concentración de AHSG en las 24 muestras testadas (Tabla 4) y se analizaron las diferencias entre ellas. En este caso el análisis de la normalidad de los residuos nos proporcionó una $p=0,02421$, por lo que no se pudo asumir la normalidad de los datos y se utilizaron en consecuencia test no paramétricos. Se realizó por tanto el estadístico Kruskal Wallis ($p=0,0437$) lo que nos indicó que sí había diferencias entre grupos. Dados estos resultados se realizó como post-hoc el test de comparación múltiple de Dunn que no reportó diferencias entre grupos, por lo que se realizaron comparaciones de dos en dos con el test de Mann Whitney U. Este estadístico reportó una diferencia entre los grupos 2 y 4 ($p=0,0411$), diferencia que se indica en el gráfico de la Figura 16.

Tabla 4: Media de las concentraciones de AHSG obtenidas para cada muestra.

Estadio I	Estadio II	Estadio III	Estadio IV
983,794 ng/mL	3068,728 ng/mL	5114,245 ng/mL	895,199 ng/mL
1348,416 ng/mL	2556,079 ng/mL	3440,816 ng/mL	4487,002 ng/mL
2908,428 ng/mL	2588,36 ng/mL	5274,458 ng/mL	2654,138 ng/mL
3363,939 ng/mL	4578,744 ng/mL	4880,143 ng/mL	1646,285 ng/mL
2685,594 ng/mL	7952,005 ng/mL	5346,408 ng/mL	916,066 ng/mL
1741,87 ng/mL	6951,175 ng/mL	805,442 ng/mL	493,8 ng/mL

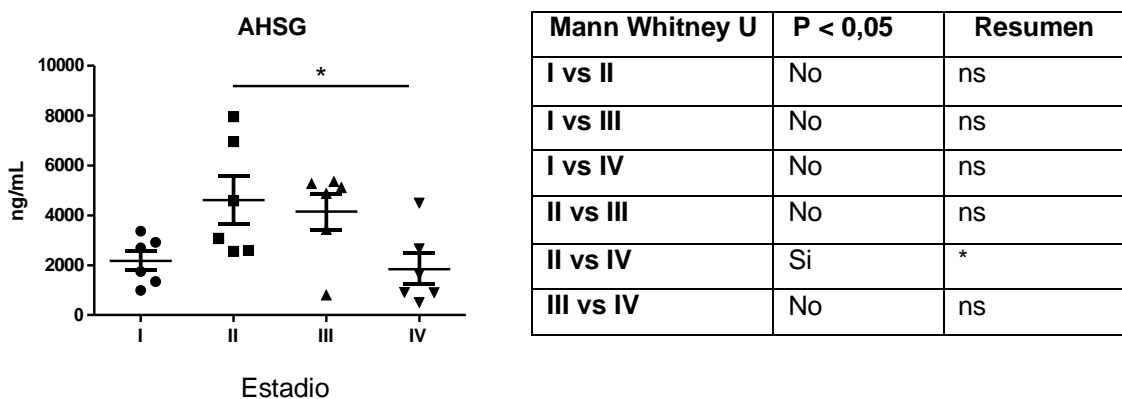


Figura 16: Gráfico de las concentraciones de AHSG junto con las diferencias según el test de Mann Whitney U.

La decisión de realizar un test de Man Whitney se debió a la ausencia de diferencias estadísticas al realizar el test de Dunn, estas ausencias podrían deberse al bajo número de muestras utilizadas para el test y a que el propio test de Kruskal Wallis nos reportó una significación muy baja, apenas 0,0063 por debajo de $p = 0,05$, lo cual pudo derivar en esta pérdida de significación al realizar el test post-hoc.

5.3 RBP4

La última proteína analizada fue la RBP4. De la misma manera que para las dos anteriores se realizó una curva estándar para extrapolar los datos de las muestras. Dicha curva se realizó con nueve puntos que iban de una concentración de 0,781 ng/mL a 100ng/mL de proteína recombinante y que presentaron una $R^2 = 1$ (Figura 17).

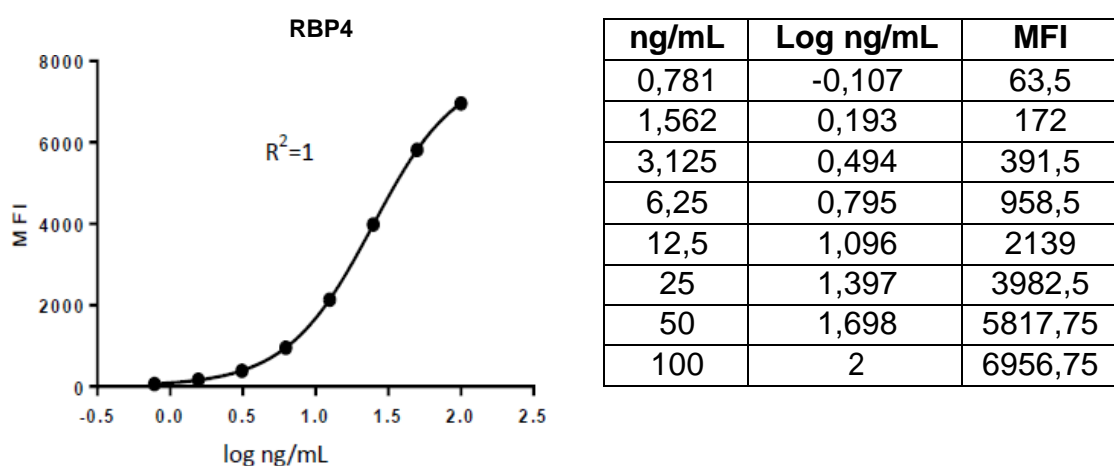


Figura 17: Curva estándar de la RBP4 empleada en el ensayo Luminex.

Los datos obtenidos de las muestras a partir de esta curva (Tabla 5), se analizaron mediante un ANOVA ($p = 0,0029$) presentando de nuevo diferencias

entre los estadios de CCR, tras haber comprobado de nuevo la normalidad de los residuos ($p= 0,3176$) y la homocedasticidad con un test de Bartlett ($p=0,5367$). Se realizó el análisis post hoc con el test de Tukey (Tabla 6). En la Figura 18 se representan gráficamente las medias de las concentraciones de RBP4 junto con las barras de error y las diferencias obtenidas en el test de Tukey.

Tabla 5: Media de las concentraciones de RBP4 obtenidas para cada muestra.

Estadio I	Estadio II	Estadio III	Estadio IV
13,85 ng/mL	72,93 ng/mL	52,55 ng/mL	12,76 ng/mL
26,91 ng/mL	36,8 ng/mL	63,31 ng/mL	45,84 ng/mL
44,93 ng/mL	44,31 ng/mL	65,91 ng/mL	34,76 ng/mL
41,18 ng/mL	54,39 ng/mL	43,63 ng/mL	29,56 ng/mL
27,69 ng/mL	80,48 ng/mL	73,41 ng/mL	16,75 ng/mL
33,05 ng/mL	81,08 ng/mL	17,48 ng/mL	15,19 ng/mL

Tabla 6: Test de comparación múltiple de Tukey para la RBP4 entre grupos.

Comparación múltiple	¿P < 0,05?	Resumen
I vs II	Yes	*
I vs III	No	ns
I vs IV	No	ns
II vs III	No	ns
II vs IV	Yes	**
III vs IV	Yes	*

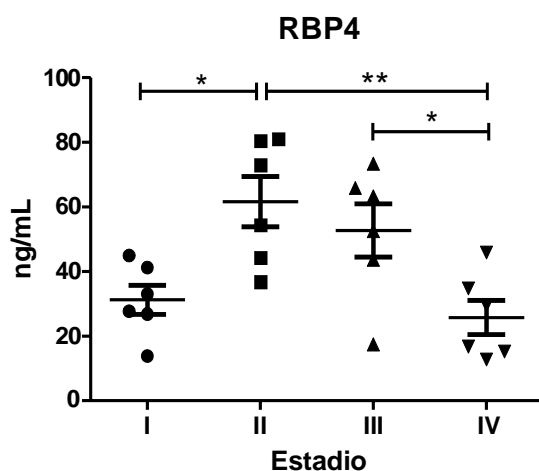


Figura 18: Gráfico de las MFI de RBP4 para los distintos estadios. * $p\leq 0,05$; ** $p\leq 0,01$ según Tukey.

En el caso de la RBP4, a pesar de tener diferencias estadísticas del grupo II con otros grupos, los resultados de concentración de proteínas fueron muy bajos en comparación con las otras proteínas analizadas (Figura 17), lo cual

nos indica una escasa implicación de esta proteína como biomarcador tisular de CCR. A pesar de estos resultados, RBP4 podría estar implicada igualmente en el CCR, ya que es una proteína que se expresa mayoritariamente en el hígado, el principal órgano afectado en la metástasis de CCR. Además, RBP4 participa activamente en la acumulación de grasa abdominal, que se ha reportado como un factor de riesgo en el desarrollo del CCR, por lo que, a pesar de no encontrarse en el tejido, esta proteína podría tener una implicación indirecta en este tipo de cáncer (Fei et al, 2017).

5.4 Estudio de correlación

Por último, se realizó una comparación global de las tres proteínas conjuntamente. Para ello se representaron gráficamente las concentraciones medias de cada proteína en cada estadio (Figura 19).

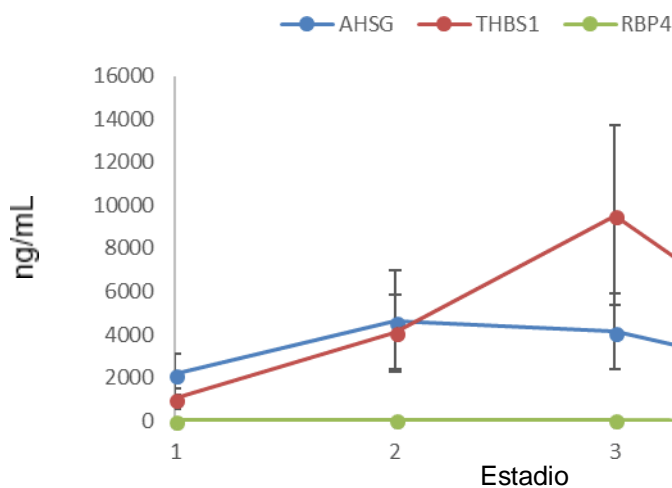


Figura 19: Representación de las concentraciones de AHSG (Azul), THBS (Rojo) y RBP4 (verde).

Dado que según la literatura AHSG y THBS1 tienen un efecto antagónico en cuanto al $TGF\beta 1$, que se refleja también en esta gráfica, decidimos realizar un test de correlación de Pearson. El resultado de este estadístico nos aportó una $p=0,6807$, por lo que podemos descartar que THBS1 y AHSG estén relacionadas entre sí.

5.5 Discusión general

Para el estudio de biomarcadores, las muestras de tejido presentan una ventaja frente a las muestras séricas debido a que la mayoría de las proteínas que se extraigan estarán directamente implicadas en los procesos patológicos que ocurren en el tejido estudiado. En cambio, no se dispone fácilmente de muestras de tejido sobre todo si queremos compararlas con controles sanos de los cuales no se realizan biopsias de colon de forma rutinaria. Por otra parte, las muestras de suero representan una fuente más accesible en ámbito

hospitalario para la realización de este tipo de estudios pese a reducirse la especificidad de las proteínas detectadas (Alvarez-Chaver et al, 2018).

El hecho de haber analizado en los análisis Luminex solo 24 de las 165 muestras procesadas representa, junto a la falta de controles sanos, una limitación a la hora de conseguir resultados significativos.

Por todo esto una manera adecuada de continuar en el futuro sería aumentar la n para obtener una mayor significación, testar un mayor número de biomarcadores mediante ensayos multiplex y corroborar los datos obtenidos en este proyecto en muestras de suero procedente de pacientes con CCR y controles sanos.

6.a. Conclusiones

Tras la realización del estudio y ante los resultados obtenidos, se han establecido una serie de conclusiones respecto a los objetivos que nos habíamos planteado:

1. Se ha optimizado un protocolo de extracción de proteínas de CCR a partir de biopsias de tejido humano.
2. Se ha determinado la concentración de proteínas en muestras de tejido de CCR mediante tinción de plata y método de Bradford.
3. Se ha identificado THBS1 como un posible biomarcador del estadio III de CCR.

6.b. Conclusión

Despois de realizar o estudo e ante os resultados obtidos, estableceronse unha serie de conclusións respecto dos obxetivos que nos plantexáramos:

1. Foi posible optimizar un protocolo de extracción de proteínas do CCR a partir de biopsias de tecido humano.
2. Foi posible determinar as concentracións de proteínas nas muestras de tecido do CCR mediante tinción de prata e método de Bradford.
3. Identificouse un posible biomarcador do estadio III do CCR na THBS1.

6.c. Conclusions

After carrying out the study and considering the results achieved we have established several conclusions concerning our goals:

1. A reliable protocol for CCR proteins from human biopsies was optimized.
2. Protein concentrations measurements in CCR tissue samples were performed by silver staining and Bradford method.
3. THBS1 was identified as potential biomarker of stage III of CCR.

7. Bibliografía

1. Abola, M., Thompson, C., Chen, Z., Chak, A., Berger, N., Kirwan, J., Li, L. (2015). Serum level of retinol-binding protein 4 (RBP4) and risk of colon adenoma. *Endocrin related cancer*. 22(2) 1-7.
2. Alvarez-Chaver, P., De Chiara, L., Soledad, V. (2018). Proteomic profiling for colorectal cancer biomarker Discovery. *Springer Nature*. 1765: 241-269.
3. Angeloni, S., Cordes, R., Dunbar, S., García, C., Gibson, G., Martin, C., Stone, V. (2014). *xMAP® Cookbook. A collection of methods and protocols for developing multiplex assays with xMAP Technology*. 2ª ed. Luminex Corporation.
4. Barderas, R., Babel, I., Casal, J.I. (2010). Colorectal cancer proteomics, molecular characterization and biomarker Discovery. *Proteomics clinical applications*. 4(2), 159-178.
5. Centro Nacional de Epidemiología. Área de Epidemiología Ambiental y Cáncer [página en Internet]. Madrid: Instituto de Salud Carlos III. Disponible en: <http://www.isciii.es/ISCIII/es/contenidos/fd-servicios-cientifico-tecnicos/fd-vigilancias-alertas/fd-epidemiologia-ambiental-y-cancer/Mortal2011.pdf>
6. Casal, JI., Vivanco, F. (2014). Proteómica clínica. Biomarcadores y dianas terapéuticas. Corrales, F. Calvete, JJ. (Eds.). En: *Manual de proteómica*, pp. 583-601.
7. Domenech, A., Pérez-Gómez, B., Aragonés, N., Pollán, M., López-Abente, G. (2009). *La situación del cáncer en España, 1975-2006*. Madrid, Centro nacional de epidemiología ISCIII.
8. Esteva, M., Ramos, M., Cabeza, E., Llobera, J., Ruiz, A., Pita, S., Segura, J.M., Córtes, J.M., Gonzalez-Lujan, L. (2007). Factors influencing delay in the diagnosis of colorectal cancer: a study protocol. *BMC cancer*. 7:86
9. Fei, W., Chen, L., Chen, J., Shi, Q., Zhang, L., Liu, S., Li, L., Zheng, L., Hu, X. (2017). RBP4 and THBS2 are serum biomarkers for diagnosis of colorectal cancer. *Oncotarget*. 8(54), 92254-92264.
10. Greene, F., Page, D., Fleming, I. Fritz, A., Blach, C., Haller, D., Morrow, M. (2002). *AJCC Cancer staging manual*. Chicago, Springer.
11. Hejmadi, M. (2010). *Introduction to cancer biology*. Dinamarca, Bookboon.
12. Herraiz, A. (2010). Análisis de biomarcadores empleando biochips y matrices de biomoléculas. *Anales de la real academia de farmacia*, in Monografía XXX: Biomarcadores: Analítica, diagnóstico y terapéutica.

13. Hernández Sánchez, M. L., y Gil García, C. (2011). La proteómica un reto constante en biomedicina. Encuentros multidisciplinares. 13(38), 2-8.
14. Holland, J., Frei, E., Kufe, D., Pollock, R., Weichselbaum, R., Bast, R., Gansler, T. (2003). Cancer Medicine 6th edition. Hamilton, BC Decker.
15. Hu, S., Xie, Z., Qian, J., Blackshaw, S., & Zhu, H. (2011). Functional Protein Microarray Technology. Wiley Interdisciplinary Reviews. Systems Biology and Medicine. 3(3), 255–268.
16. Jiménez, C., Knol, J., Meijer, G., Fijneman, R. (2010). Proteomics of colorrectal cáncer: Overview of Discovery studies and identificación of commonly identified cáncer-associated proteins and candidate CRC serum markers. Journal of proteomics. 73(10), 1873-1895.
17. Lourido, L. (2015). Búsqueda de biomarcadores de artrosis mediante técnicas proteómicas. Tesis doctoral.
18. Luo, Y., Wang, L., Wang, J. (2013). Developing proteomics-based biomarkers for colorrectal neoplasms for clinical practice: Opportunities and challenges. Proteomic clinical applications. 7, 30-41.
19. Miyanga, K., Kato, Y., Nakamura, T., Matsumura, M., Amaya, H., Horiuchi, T., Chiba, Y., Tanak, K. (2002). Expression and role of thrombospondin-1 in colorrectal cáncer. Anticancer Research. 22: 3941-3948.
20. Mustafa, S.A., Hoheisel, J., Alhamdani, M. (2011). Secretome profiling with antibody microarrays. Molecular Biosystems. 7, 1795-1801.
21. Nambiar, P., Gupta, R., Misra, V. (2010). An “Omics” based survey of human colon cáncer. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis. 693 (1-2), 3-18.
22. Ocaña, E., Aceituno, M^aI. (2014). Utilidad clínica de los marcadores tumorales. Revista médica de Jaén. Recuperado de: <https://www.juntadeandalucia.es/servicioandaluzdesalud/chjaen/files/pdf/1414585159.pdf>
23. Olives, A., Del Castillo, B., Martin, M.A. (2010). Técnicas analíticas luminiscentes y de separación aplicadas a la identificación y cuantificación de biomarcadores. Anales de la real academia de farmacia, in Monografía XXX: Biomarcadores: Analítica, diagnóstico y terapéutica.
24. Ortega, F. (2010). Estrategias Analíticas En La Investigación de Nuevos Biomarcadores. Anales de la real academia de farmacia, in Monografía XXX: Biomarcadores: Analítica, diagnóstico y terapéutica.
25. Pando, RV., Lanz-Mendoza, H. (2009). The significance of proteomics in public health. Salud pública México. 51, 386-394.
26. Quick start Bradford protein assay, instruction manual. Recuperado de: <http://www.bio-rad.com/webroot/web/pdf/lsr/literature/4110065A.pdf>

27. Reslova, N., Michna, V., Kasny, M., Mikel, P., Kralik, P. (2017). xMAP technology: applications in detection of pathogens. *Frontiers in microbiology*. 8: 55.
28. Reymond Sutandy, F., Qian, J., Chen, C.-S., & Zhu, H. (2013). Overview of Protein Microarrays. *Current Protocols in Protein Science*. 0 27, Unit–27.1.
29. Ruiz-Romero, C., Blanco, FJ. (2010). Proteomics role in the search for improved diagnosis, prognosis and treatment of osteoarthritis. *Osteoarthritis and cartilage*. 18: 500-509.
30. Sanchez-Carballo, M. (2007). Aplicación de arrays de anticuerpos en el estudio del cáncer vesical. *Actas urológicas españolas*. 31(9), 1082-1088.
31. Sandra K, Moshir M, D'hondt F, Tuytten R, Verleysen K, Kas K, et al. Highly efficient peptide separations in proteomics (2009). Part 2: bi- and multidimensional liquid-based separation techniques. *Journal of Chromatography and Analytics in Technology Biomedicine of Life Sciences*. 877(11-12), 1019-39.
32. Shevchenko, A., Wilm, M., Vorn, O., Mann, M. (1996). Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. *Analytical chemistry*. 68(5), 850-858.
33. Siegel, RL., Miller, KD., Jemal, A. (2018). Cancer statistics, 2018. *CA cáncer J Clinicians*. 68(1), 7-30.
34. Stewart, B., Wild, C. (2014). *World cáncer report 2014*. Lyon, International agency for research on cancer.
35. Sutandy, R., Qian, J., Chen, C., Zhu, H. (2013) Overview of protein microarrays. *National institutes of health*. 27(1), 1-21.
36. Swallow, C., Partridge, E., Macmillan, J., Tajirian, T., DiGuglielmo, G., Hay, K., Szweras, M., Jahnen-Dechent, W., Wrana, J., Redston, M., Gallinger, S., Dennis, J. (2004). α 2HS-glycoprotein, an antagoinst of transforming growth factor β in vivo, inhibits intestinal tumor progression. *Cancer Research* 64: 6402-6409.
37. Wit, M., Fijneman, R., Verheul, H., Meijer, G., Jimenez, C. (2013). Proteomics in colorrectal cáncer translational research: Biomarkers Discovery for clinical applications. *Clinical Biochemistry*. 46(6), 466-479.
38. Zamanian–Azodi, M., Rezaei–Tavirani, M., Hasanzadeh, H., RahmatiRad, S., Dalilan, S., Gilanchi, S., Manzour, H. (2015) Introducing biomarker panel in esophageal, gastric, and colon cancers; a proteomic approach. *Gastroenterol Hepatol Bed Bench*. 8(1), 6-18

8. Anexo

Listado de muestras agrupadas por estadio. En las tablas se indican código de referencia, año de muestreo, edad del paciente y género (H= hombre; M= mujer).

Estadio I:

Código	Año	Edad	Género
B07-01747	2007	63	H
B07-02707	2007	68	H
B07-04331	2007	53	M
B07-04818	2007	72	H
B07-06324	2007	78	M
B07-06453	2007	73	M
B07-06903	2007	70	M
B07-07136	2007	76	H
B07-07921	2007	76	H
B07-09017	2007	69	M
B07-09573	2007	54	M
B07-10897	2007	70	M
B07-11002	2007	70	M
B07-12037	2007	66	H
B07-12376	2007	75	H
B07-12906	2007	64	H
B07-13380	2007	75	M
B07-16350	2007	53	M
B07-16629	2007	67	M
B07-16707	2007	49	H
B07-16985	2007	77	M
B07-17570	2007	84	H
B07-18850	2007	80	H
B07-21572	2007	64	M

Código	Año	Edad	Género
B08-00038	2008	59	H
B08-03954	2008	59	H
B08-04383	2008	72	H
B08-04718	2008	79	M
B08-05553	2008	67	H
B08-05946	2008	47	H
B08-06659	2008	80	H
B08-09095	2008	64	H
B08-09332	2008	71	H
B08-09806	2008	68	M
B08-14073	2008	78	M

Estadio II:

Código	Año	Edad	Género
B08-01044	2008	72	H
B08-01145	2008	73	H
B08-01146	2008	80	M
B08-02075	2008	48	H
B08-03042	2008	58	H
B08-03606	2008	70	H
B08-03942	2008	64	H
B08-04051	2008	71	H
B08-04503	2008	70	M
B08-04513	2008	60	M
B08-04632	2008	71	H
B08-05084	2008	74	H
B08-05915	2008	67	H
B08-06541	2008	74	H
B08-07032	2008	88	M
B08-08454	2008	68	H
B08-08458	2008	82	M
B08-09581	2008	71	H
B08-10212	2008	75	H
B08-11178	2008	80	M
B08-13164	2008	64	H
B08-13442	2008	72	M
B08-13812	2008	68	M
B08-14066	2008	72	H

Código	Año	Edad	Género
B07-01396	2007	54	H
B07-01640	2007	52	M
B07-01873	2007	84	H
B07-02473	2007	64	M
B07-02474	2007	63	H
B07-02608	2007	73	H
B07-02956	2007	75	H
B07-03286	2007	76	H
B07-03650	2007	76	M
B07-07283	2007	77	M
B07-08056	2007	34	H
B07-08487	2007	66	M
B07-08590	2007	78	M
B07-09450	2007	45	H
B07-09467	2007	73	H
B07-10022	2007	79	H
B07-10377	2007	75	H
B07-10575	2007	52	M
B07-10780	2007	69	H
B07-11546	2007	62	H
B07-11776	2007	85	H
B07-12174	2007	69	H
B07-12565	2007	69	H
B07-12664	2007	84	M
B07-12676	2007	69	H
B07-13360	2007	81	H
B07-13480	2007	72	H
B07-13682	2007	43	M
B07-14192	2007	70	M
B07-14271	2007	73	M
B07-15355	2007	70	H
B07-15547	2007	74	M
B07-16851	2007	79	M
B07-17494	2007	70	H
B07-17654	2007	85	H
B07-18837	2007	84	H
B07-18965	2007	71	H
B07-21142	2007	71	M
B07-21798	2007	88	H
B07-23034	2007	74	M
B07-23818	2007	70	H
B07-23904	2007	79	H
B07-24300	2007	67	H
B07-24432	2007	53	H
B07-25859	2007	65	H

Estadio III:

Código	Año	Edad	Género
B07-01290	2007	42	M
B07-02123	2007	79	H
B07-05942	2007	76	H
B07-08362	2007	75	H
B07-09346	2007	59	H
B07-09707	2007	70	H
B07-09795	2007	74	M
B07-10463	2007	75	H
B07-14386	2007	65	H
B07-14571	2007	59	H
B07-16237	2007	66	H
B07-16500	2007	66	M
B07-17516	2007	81	M
B07-17911	2007	53	H
B07-17928	2007	81	H
B07-19388	2007	37	H
B07-19803	2007	60	H
B07-19837	2007	77	M
B07-22774	2007	61	H
B07-23019	2007	77	H
B07-23703	2007	74	M

Código	Año	Edad	Género
B08-00142	2008	63	M
B08-00284	2008	69	M
B08-01263	2008	65	H
B08-01854	2008	77	H
B08-02395	2008	79	H
B08-02524	2008	65	H
B08-02534	2008	67	M
B08-03185	2008	54	H
B08-04184	2008	57	M
B08-04964	2008	61	M
B08-06270	2008	72	M
B08-07401	2008	70	H
B08-07625	2008	78	H
B08-08869	2008	77	M
B08-08873	2008	72	H
B08-09100	2008	65	H
B08-09359	2008	62	M
B08-09613	2008	90	H
B08-10094	2008	84	M
B08-11435	2008	82	M
B08-11802	2008	70	M
B08-12955	2008	75	H
B08-13695	2008	73	H

Estadio IV:

Código	Año	Edad	Género
B07-0328	2007	43	M
B07-04125	2007	59	M
B07-04693	2007	79	M
B07-0827	2007	73	M
B07-08493	2007	62	H
B07-08585	2007	80	H
B07-08897	2007	66	H
B07-12285	2007	79	H
B07-13851	2007	83	H
B07-14168	2007	90	M
B07-18848	2007	54	H
B07-21013	2007	73	M
B07-21469	2007	59	H
B07-25380	2007	72	M
B07-26017	2007	70	H

Código	Año	Edad	Género
B08-00920	2008	56	M
B08-03044	2008	81	H
B08-04049	2008	48	H
B08-05192	2008	77	H
B08-06053	2008	71	M
B08-09462	2008	64	M
B08-10958	2008	85	H